International Energy Agency

Real Time Simulation of HVAC Systems for Building Optimization, Fault Detection and Diagnosis. Technical Papers of IEA Annex 25

Energy Conservation in Buildings and Community Systems Programme. Annex XXV. Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnosis

Real Time Simulation of HVAC Systems for Building Optimization, Fault Detection and Diagnosis

Technical Papers of IEA Annex 25

November 1996

editor

Juhani Hyvärinen

VTT Building Technology
P. O. Box 1804 (Lämpömiehenkuja 3)
FIN-02044 VTT (Espoo)

FINLAND

This report is part of the work of the IEA Energy Conservation in Buildings and Community Systems Programme

Annex 25 Real Time Simulation of HVAC Systems for Building Optimization, Fault Detection and Diagnosis

ABSTRACT

This technical report describes the technical and scientific work carried out in Annex 25, and the applications developed during the annex. Annex 25, Real Time Simulation for Building Optimisation, Fault Detection and Diagnosis, was a part of the work of the IEA Energy Conservation in Buildings and Community Systems Programme.

This report includes papers on Building Optimisation and Fault Diagnosis (BOFD) system applications, BOFD-method applications and some papers on BOFD-tools. Method applications deal with fault detection and diagnosis of a single component or of a subprocess whereas system applications include fault analysis of HVAC-system in consideration and the steps leading to development of BOFD-methods for the components and processes of that system.

Annex 25 reports consist of two volumes:
    Volume I: Building opimization and fault diagnosis source book
    Volume II: Technical papers of IEA Annex 25

This report, volume II, consists of three parts: system applications, method applications, and tools. Each part gives examples of applications of the concepts that are presented in volume I. The parts and sections in this volume are written as separate technical papers, reviewed and only published together without any further edition. Some of the papers are reprints of conferences and other publications, and the purpose of reprinting them also here is that the major part of the research reported in them was carried out as contribution to Annex 25.

PREFACE

International Energy Agency

The International Energy Agency (IEA) was established in 1974 within the framework of the Organisation for Economic Co-operation and Development (OECD) to implement an International Energy Programme. A basic aim of the IEA is to foster co-operation among twenty-one IEA Participating Countries to increase energy security through energy conservation, development of alternative energy sources and energy research development and demonstration (RD&D) This is achieved in part through a programme of collaborative RD&D consisting of forty-two Implementing Agreements, containing a total of over eighty separate energy RD&D projects.

Energy Conservation in Buildings and Community Systems

The IEA sponsors research and development in a number of areas related to energy. In one of these areas, energy conservation in buildings, the IEA is sponsoring various exercises to predict more accurately the energy use of buildings, including comparison of existing computer programs, building monitoring, comparison of calculation methods, as well as air quality and studies of occupancy. Seventeen countries have elected to participate in this area and have designated contracting parties to the Implementing Agreement covering collaborative research in this area. The designation by governments of a number of private organisations, as well as universities and government laboratories, as contracting parties, has provided a broader range of expertise to tackle the projects in the different technology areas than would have been the case if participation was restricted to governments. The importance of associating industry with government sponsored energy research and development is recognised in the IEA, and every effort is made to encourage this trend.

The Executive Committee

Overall control of the programme is maintained by an Executive Committee, which not only monitors existing projects but identifies new areas where collaborative effort may be beneficial. The Executive Committee ensures that all projects fit into a pre-determined strategy, without unnecessary overlap or duplication but with effective liaison and communication. the Executive Committee has initiated the following projects to date (completed projects are identified by *):

I Load Energy Determination of Buildings * II Ekistics and Advanced Community Energy Systems * III Energy Conservation in Residential Buildings * IV Glasgow Commercial Building Monitoring * V Air Infiltration and Ventilation Centre VI Energy Systems and Designs of Communities * VII Local Government Energy Planning * VIII Inhabitant Behaviour with Regard to Ventilation * IX Minimum Ventilation Rates * X Building HVAC Systems Simulation * XI Energy Auditing * XII Windows and Fenestration * XIII Energy Management in Hospitals * XIV Condensation * XV Energy Efficiency in Schools * XVI BEMS - 1: Energy Management Procedures * XVII BEMS - 2: Evaluation and Emulation Techniques * XVIII Demand Controlled Ventilating Systems * XIX Low Slope Roof Systems XX Air Flow Patterns within Buildings * XXI Environmental Performance XXII Energy Efficient Communities XXIII Multizone Air Flow Modelling * XXIV Heat, Air and Moisture Transport XXV Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnosis XXVI Energy Efficient Ventilation of Large Enclosures XXVII Evaluation and Demonstration of Domestic ventilation Systems XXVIII Low Energy Cooling Systems XXIX Daylight in Buildings XXX Bringing Simulation to Application

ACKNOWLEDGEMENTS

Dr.Ing F. SCHMIDT, IKE University Stuttgart Wissensverarbeitung & Numerik und Technische Gebäudeausrüstung

JAPAN

Prof Emeritus Nobuo NAKAHARA, Nagoya University
Prof Harunori YOSHIDA, Kyoto Univ. / Fac. Engineering, Div. of Global Environment Engineering
Prof Hisahiro ITOH, Mie University / Dept of Architecture, Faculty of Engineering
Mr Tatsuhiro IWAMI, Shinryo Reinetsu Co. Ltd.
Mr Kazuyuki KAMIMURA, Yamatake Honeywell Co. Ltd.
Mr Katashi MATSUNAWA, Nikken Sekkei Ltd.
Mr Tomoaki MIYAZAKI, Ohbayashi Corporation
Mr Masaharu NAKAMURA, Yokogawa Johnson Controls Corp.
Dr Hideharu NIWA, Nikken Sekkei Ltd.
Prof. Kazunobu SAGARA, Mie University / Dept. of Architecture, Faculty of Engineering
Mr Hiroaki TSUTSUI, Yamatake-Honeywell Co., Ltd
Mr Hideki YUZAWA, Nikken Sekkei Ltd.
Dr Minjie ZHENG, Sanko Kucho Ltd.
Dr Yi JIANG, Tsinghua University / Dept. of Thermal engineering
Dr Yingxin ZHU, Thsinghua University / Dept. of Thermal engineering
Dr Tianzhen HONG, Tsinghua University / Dept. of Thermal engineering

SWEDEN

Mr Per BLOMBERG, KTH (Royal Institute of Techn.) Div. Bldg Services Engineering
Dr Per ISAKSSON, KTH (Royal Institute of Techn.) Div. Bldg Services Engineering
Prof T. G. MALMSTRÖM, KTH (Royal Institute of Techn.) Div. Bldg Services Engineering
Dr L. Göran OLSSON, KTH (Royal Institute of Techn.) Div. Bldg Services Engineering
Mr Svein RUUD, Swedish National Testing and Research Institute Energy Tech., HVAC and Sanitat

SWITZERLAND

Dr Jürg TOEDTLI, Landis & Gyr (Europe) AG
Dr Andrew S. GLASS, Ingenieurschule Luzern
Dr Peter GRUBER, Landis & Gyr Technology Innovation
Mr Henri van KUIJK, SEC Sulzer Energieconsulting SA
Dr Daniel MAURER, Ulrich Amman AG, Langenthal
Mr. Roger RAEBER, ETHZ, Energy systems laboratory
Mr Horst F. SCHIEL, H. F. Schiel Consulting Service, Winterthur

Dr Peter SPRECHER, Landis&Gyr (Europe) AG

THE NETHERLANDS

Mr Henk PEITSMAN, TNO Building and Construction Research, Dept. of Indoor Environment, Building Physics and Systems
Dr Lue L., SOETHOUT, TNO Building and Construction Research, Dept. of Indoor Environment, Building Physics and Systems

U.S.A

Dr George E. KELLY, NIST
Prof James BRAUN, Purdue University Ray W. Herrick Laboratories
Dr John M. HOUSE, NIST
Mr Les NORFORD, MIT
Dr Cheol PARK, NIST
Mr Todd ROSSI, Purdue University Ray W. Herrick Laboratories
Mr John SEEM, Johnson Controls Inc. Controls Research Dept
Dr Rex C. STRATTON, Pacific Northwest Laboratories BATTELLE Battelle Boulevard
Won-Yong LEE, Korea Instit. of Energy Research Taejean Yuseong Gu

UNITED-KINGDOM

Dr Arthur DEXTER, University of Oxford, Department of Engineering Science
Dr Philip HAVES, Loughborough University of Tech., Dept. of Civil and Building Engineering
Dr Mourad BENOUARETS, University of Oxford, Department of Engineering Science
Mr Richard FARGUS, Building Research Establishment
Mr Simon HEPWORTH, University of Oxford, Department of Engineering Science
Mr Newton MARUYAMA, University of Oxford, Department of Engineering Science
Mr Darius NGO, University of Oxford, Department of Engineering Science
Mr Tim SALSBURY, Loughborough University of Tech., Dept. of Civil and Building Engineering
Dr Jon WRIGHT, Loughborough University of Tech., Department of Civil and Building Engineering

LIST OF CONTENTS

7

Qualitative approaches to FDD

Generic BOFD tools

Late paper

FOREWORD

This report is a collection of technical papers describing the technical work carried out by participants in Annex 25. The papers are written as separate papers, reviewed internally in the annex and only published together without any further edition. Some of the papers are reprints of conferences and other publications. The purpose of reprinting them also here is that the major part of the research reported in them was carried out as contribution to Annex 25.

The background information and the main results of the Annex 25 collaboration is published as a separate volume and it is called the Building Optimisation and Fault Diagnosis Source book. This technical report is organised following the outline of the Source Book. The work carried out and the results achieved in the annex can roughly be divided in three parts: BOFD systems applications for HVAC systems, BOFD method applications, and tools for developing BOFD methods and systems. This same division can be found in the outline of the Source Book chapters 3 - 5 and the same is followed here in order to assist the reader to associate the technical papers of this report to sections in the Source Book. Some of the technical papers could be associated to more than one section of the Source Book . In those cases the place of a paper is chosen by the operating agent of the annex according to the advice from the authors of that paper.

A Neural Network Prototype for Fault Detection and Diagnosis of Heating Systems

Xiaoming Li, Jean-Christophe Visier and Hossein Vaezi-Nejad [*]

CSTB (French Scientific and Technical Building Centre)
Department ENEA
BP 02, F-77421 Marne La Vallee, Cedex 2, France

ABSTRACT

An Artificial Neural Network (ANN) prototype for Fault Detection and Diagnosis (FDD) in complex heating systems is presented in this paper. The six operating modes with faults used to develop this prototype stemmed from a detailed investigation in co-operation with heating systems maintenance experts, and are among the most important operating faults for this type of system. The prototype has been developed by using the daily values obtained by a pre-processing procedure of the simulation data of one reference heating system, and then generalized to four heating systems not used during the training phase.

This paper demonstrates the feasibility of using ANN for detecting and diagnosing faults in heating systems provided that training data representative of the behaviour of the systems with and without faults are available.

Key-words:

boiler, control, energy conservation, gas fired, maintenance, office building, space heating , valve.

[*]   X. Li, J.C. Visier and H. Vaezi-Nejad are PhD and Research Engineers in the

HVAC Department (ENEA), CSTB (French Scientific and Technical Building Center), Marne-la-Vallée, France.

ABSTRACT

An Artificial Neural Network (ANN) prototype for Fault Detection and Diagnosis (FDD) in complex heating systems is presented in this paper. The six operating modes with faults used to develop this prototype stemmed from a detailed investigation in co-operation with heating systems maintenance experts, and are among the most important operating faults for this type of system. The prototype has been developed by using the daily values obtained by a pre-processing procedure of the simulation data of one reference heating system, and then generalized to four heating systems not used during the training phase.

This paper demonstrates the feasibility of using ANN for detecting and diagnosing faults in heating systems provided that training data representative of the behaviour of the systems with and without faults are available.

INTRODUCTION

As technical systems of buildings develop, their operations get more difficult for an average operator to understand. Buildings get more intelligent but the users do not. Understanding the relationship between cause and effect is more difficult than in the past because of complex relationships in the building's processes. So it is necessary to develop the tools which support the operator in decision-making for building optimisation as well as recovery from a faulty state. The tools should focus on the underlying defects and give instructions on corrective action to be taken in a simple and understandable way.

Various methods can be used to resolve a failure in a process. The development of defects can be monitored with special condition monitoring instrumentation to obtain information on the need for maintenance. These monitoring systems are usually separated from building automation systems and need specific instrumentation of their own. For instance, vibration analysis systems which are used in industrial processes can be used for condition monitoring. Also various maintenance programs can be used to prevent serious defects and to schedule maintenance for maximum convenience. In maintenance programs the process is inspected and maintained at constant time intervals independent of the true condition of the process. The main disadvantages of all these fault detection methods are: that they are expensive because of special instrumentation; or that they cannot be operated in real-time application.

The development of Energy Management and Control Systems (EMCS) makes possible to measure a great number of variables on Heating, Ventilating, and Air-Conditioning (HVAC)-systems of middle or large size. Today, it is known that these systems are widely used for automating HVAC-system operation, for its remote control, and for detecting operating faults with great magnitudes. But The makers of EMCS do not propose the tools to aid the operator in diagnosing the defects that cause the faulty process operation. A·small number of users of EMCS have developed this type of tool; however, because of the lack of time and means, they can only develop some very limited tools of this type, and this prevents them from diffusing easily these tools.

The International Energy Agency (IEA) established, in the framework of Annex 25, a research program whith the main goal to develop methodological procedures, within a defined concept, for the real-time and automatic performance optimisation, diagnosis and fault detection of HVAC processes. The ultimate goal of the project will be Building Optimisation and Fault Detection (BOFD) prototypes that are implementable in EMCS (Hyvärinen, 1993).

The work reported in this paper is a part of the work undertaken by French partners which concerns the application of Artificial Neural Networks (ANNs) to develop FDD methodologies in HVAC-systems.

The paper is composed of: data base generation; neural network structures development; retained network generalization; discussion and conclusion about this work.

DATA BASE GENERATION

A data base representing the different operating modes of a system with and without fault is necessary for developing FDD knowledge-based methods. This data base can be obtained either by measurement on a real system or a testing bench or by computer simulation. For this project, taking advantage of mathematical models already established to simulate the running modes of a heating system without faults , computer simulation was used to obtain the data base.

The work achieved consists of: determining a reference heating system as a support for typical fault assessment and fault simulation, modelling the components with the most important faults and simulating the behaviour of the whole system with these faults, selecting simulation data to develop FDD methodologies, and choosing the pertinent parameters which can most easily discriminate the different chosen faults.

Reference System

The system used for computer simulations is a simplified version of the reference heating system (see Figure 1) described by Visier and Li (Hyvärinen, 1993) in the framework of IEA-Annex 25. Note that this type of system is very common in Europe.
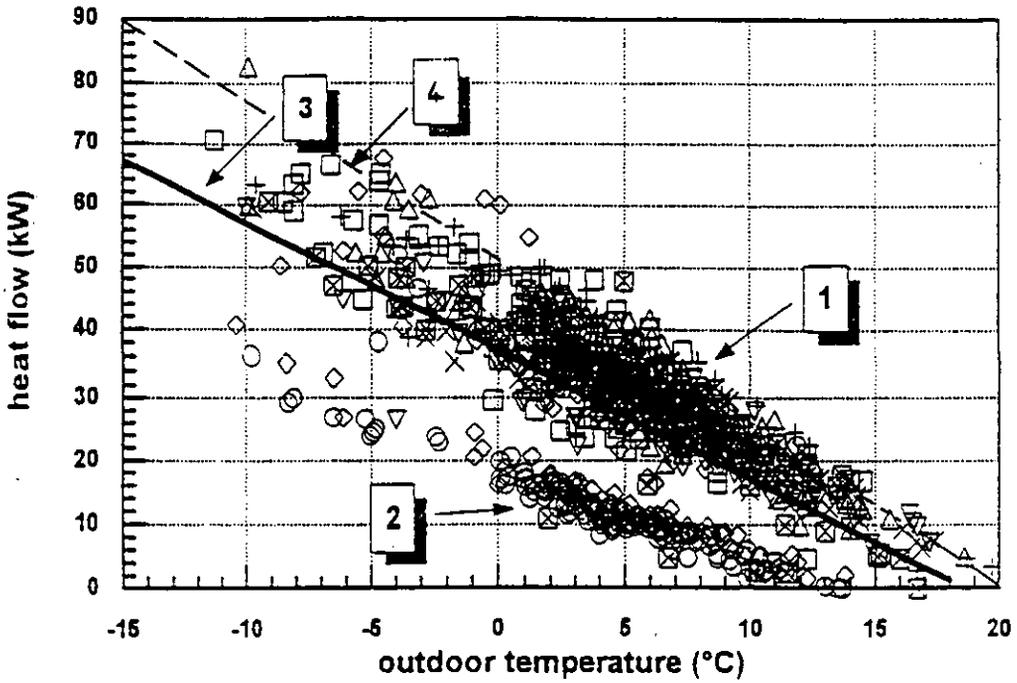
fig. 3: hourly means of heat consumption (Mo-Su)

1 - 04.00 a.m. - 10.00 p.m.
2 - 10.00 p.m. - 04.00 a.m.
3 - 00.00 a.m. - 12.00 p.m. (regression)
4 - theor.
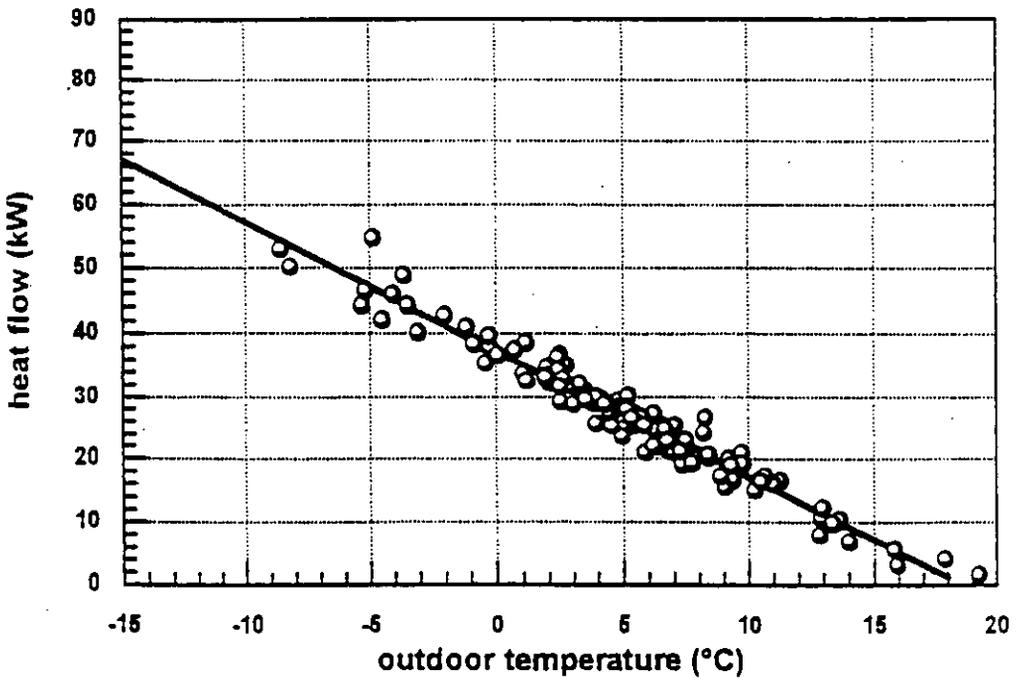
fig. 4: daily means of heat consumption (Mo-Su)
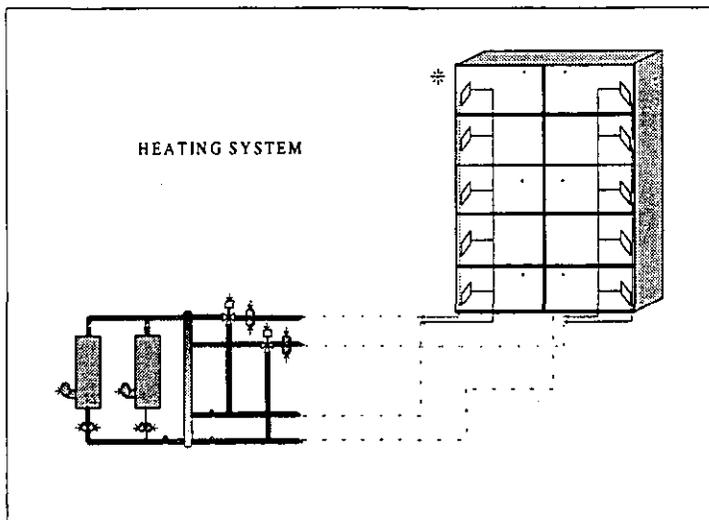
Figure 1 - Reference heating system

The main characteristics of the system are as follows:

- The building is representative of a middle sized office building or of a school (2000m^2 -21500 ft^2, 5 floors).

- The heat generation plant includes two classical gas fired boilers. The secondary loop (distribution loop) has two circuits : north and south circuits.

- The modelling of the office building and of the radiators in this building was simplified to two zones (north and south zones).

- Each heating circuit (northern/southern) has its own control system and can be controlled independently. One supposes a typical weekly occupation, meaning days of occupancy from 8 a.m. to 6 p.m. and two days of unoccupancy during all the heating season (a typical office building or school occupancy). The Operational Mode Controller (OMC) of each circuit determines in which mode the circuit is controlled. The OMC distinguishes four operation modes:

1° Stop heating. During an unoccupied period the heating of the building is stopped partially or totally. *The beginning of stop heating and the beginning of the unoccupied period are coincident.*

2° Reduce heating. During an unoccupied period, when the indoor temperature of building goes down to its low limit (antifreeze for example), the OMC generates a signal which effects the starting of the reduced heating and the maintenance of the indoor temperature at this low limit.

3° Boost heating. After a period of stop or reduce heating, the boost heating beginning time is calculated mainly according to the outdoor temperature, the low limit of indoor temperature and the building inertia in order to optimize the energy consumption and to ensure the user's comfort at the beginning of occupancy period. During a boost heating period, the two boilers are working with full power in order to minimize the boost heating duration. During this period the two three-way-valves (north and south) are open totally. At the beginning of the boost period the opening of these valves can be slowed down to prevent the primary return water temperature from being too low. The OMC transfers control to the regulator (normal heating) at the beginning of the occupancy period or as soon as the measured indoor temperature passes beyond its set point before the occupancy period.

4° Normal heating. During an occupancy period, the OMC keeps the indoor temperature at its set point by means of the heating curve (supply water temperature control) and the thermostatic valve in each room.

Modelling-Simulation of the Operating Faults

The simulation of the building and its heating system was performed using a commercial simulation software and a library (models) of HVAC components. This library includes a dynamic representation of the different components (building, boilers, pipes, valves, radiators, controllers...). The building model has been simplified to a two zones'model. It is derived from the second order model initially developed by Laret (Laret, 1980). A description of the component models and their validation is given by Caccavelli (Caccavelli et al. 1991).

The simulation of the plant includes a detailed modelling of the control system: OMC, heating curve and thermostatic valve. The models of the control systems were developed in order to represent standard and actually existing components instead of optimal components. The control of the temperatures of the supply water to the heated zones includes an outside temperature reset (heating curve) which is mandatory in French regulations. It does not include a sun radiation reset which is not often implemented on such plants. The modelled optimum start controller which is integrated in the modelled OMC has the same limits as existing optimum start controllers (Visier et al., 1994) and does not evaluate perfectly the duration of the boost heating.

Simulation of the most important operating faults was necessary for the construction of the data set. These operating faults came from the results of the questionnaires answered by 46 maintenance experts of heating systems in France (Li and Visier, 1995). Design faults were not simulated at the present time. However it will be necessary to verify the robustness of the developed FD method in the presence of design faults.

The seven operating modes (six categories of fault mode and one reference normal mode), that is seven classes in terms of pattern recognition (Dubuisson, 1990), under consideration in our simulation are modelled as follows:

Class n°0 (referred as "**normal** class"): a reference fault free mode. This normal mode is not optimal but is representative of what can be considered as a normal mode in a real building. In particular the sub optimal behaviour of existing controllers which are described above are not considered as faults.

Class n°1 (referred as "**combustion** class"): bad combustion of the burner of the boiler. There are many types of bad combustion : air excess rate too high, air excess rate insufficient, large variation of calorific value. The fault of air excess rate too high was chosen to simulate this type of fault.

Class n°2 (referred as "**heat exchanger** class"): dirtiness (flue gas side) and scale formation (water circulating side) of the heat exchanger of boiler. This fault is simulated by changing the heat transfer coefficient between water side and flue gas side.

Class n°3 consisting of two sub classes: the first is related to the faults of heating curve too high (class n°31), the second to the faults of heating curve too low (class n°32 referred as "**heating curve** class"). The latter means the supply hot water temperature of the distribution circuit is too low to satisfy the need of heat power to maintain an indoor air temperature at its set point. At the first stage only the class n°32 was simulated for all the two zones (north and south) of the building. This fault was simulated by giving the supply water temperature a value which is 10°C lower than its design value.

Class n°4 is seperated into two classes: the class n°41 i.e. the boost heating starts too early (referred as "**early boost** class"); and the class n°42 i.e. the boost heating starts too late (referred as "**late boost** class"). To simulate these classes, the maximum boost heating duration in the OMC model was increased (simulation of early boost) or decreased (simulation of late boost). The maximum boost heating duration is defined according to a basic outdoor temperature (-7°C for example) and a low limit indoor temperature (14°C for example). Either for a real heating system or with the help of a numerical simulator, this value is obtained by proceeding by trial and error and is generally different according to different buildings.

Class n°5 (referred as "**leaky valve** class"): three-way-valve (for the control of the secondary circuits) non-watertight in closed position (leakage from the direct way to bypass) or limiter stroke badly positioned. This fault was simulated by an opening of 10% of the north circuit three-way-valve which should be closed in a normal operating mode.

All these classes have been simulated in five different heating systems for purposes of constructing a data base. This data base will be used not only to develop the neural network structures but also to test the generalization capacity of the developed network structures.

The five heating systems are briefly described here in below:

- System 0: a reference heating system with the following characteristics:

  - normal room temperature set point :19°C

  - radiators size adapted to a supply temperature of 73°C for an outdoor temperature of -5°C

  - low inertia building

  - reduced room temperature set point :14°C

- System 1: with the same characteristics of the reference one but with normal room temperature set point of 21°C

-System 2: with the same characteristics of the reference one but with oversized radiators. The radiators size adapted to a supply temperature of 63°C for an outdoor temperature of -5°C

- System 3: with the same characteristics of the reference one but with a high inertia building

- System 4: with the same characteristics of the reference one but with reduced room temperature set point equal 16°C:

Choice of the Simulation period for Training Network Structures

Both the choice of the simulation duration and that of the parameters selection for the purpose of training are based only on the reference system (that is System 0).

The simulation duration of each fault for each heating system, extends a whole heating season of one year (September-April), representing 196 days (28 weeks x 7 days/week).

We have used **a data base comprising 70 days** (24h per day) (14 weeks from Monday to Friday) for all the seven classes of the reference system in order to develop the neural network structures (Li, 1996). This data base is divided into two parts: first (35 days) for training the networks; second (35 days) for testing the networks. Both training and testing data sets include fall, winter and spring data.

Data Pre-Processing and Parameters Selection

Well-chosen pertinent parameters mean well-chosen network inputs. One facet of neural networks is that a statistical understanding of the relationships between the independent and the dependent parameters is not needed. However, as with any modelling method, improved performance for a network can be expected when well-chosen independent parameters are used.

For this type of system, a daily diagnosis seems sufficient. This is why the useful information of the data base is filtered to daily values. The filtering procedure is defined below. It consists in averaging values of different variables on specific time periods. The time periods are selected to represent a specific operating mode of the system. Consequently, each simulated class comprises **70 patterns corresponding to 70 simulation days**. The whole of the seven classes comprises **490 patterns** (70 patterns/class x 7 classes) which will be used for training and testing the neural networks. The following parameters have been chosen to construct the network inputs (Li, 1996):

(1) Tfgb (mean flue gas Temperature of boiler during the boost heating period) which helps to discriminate the classes n°1 and n°2 from the other classes;

(2) Trocc (mean room temperature from 10 a.m. to 6 p.m.) which helps to discriminate the class n°32 from the other classes. This variable is representative of the room Temperature during occupied period;

(3) Tr6(room Temperature at 6 a.m. i.e. 2 hours before beginning of occupation period) which helps to discriminate the class n°41 from the other classes,

(4) Tr8(room Temperature at 8 a.m. i.e. beginning of occupation period) which helps to discriminate the class n°42 from the other classes;

(5) Twunoc(water supply Temperature at the beginning of the unoccupied period i.e. from 6 p.m. to 11 p.m.) which helps to discriminate the class n°5 from the other classes;

(6) Tout (mean daily outdoor Temperature), note that this type of heating system is principally controlled according to the outdoor temperature, and it is of interest to determine how this parameter influences the neural networks to be trained;

(7) LUP (Length of Unoccupied Period). The method developed is adapted to non permanently occupied buildings with intermittent heating such as office buildings and schools. In such buildings the room temperature obtained at the beginning of the occupation period or the decrease of room temperature during unoccupied periods is linked to its length. So LUP is used to take this effect into account.

Note that the seven parameters mentionned above are normalized in order that these parameters are numerically comparable. The normalization strategy is as follows:

$$T_{normalized} = \frac{T_{measured} - T_{ref2}}{T_{ref1} - T_{ref2}}$$

The choice of T_{ref1} and T_{ref2} was based on a physical analysis of the heating systems and took into account their genericity in relation to different heating systems and buildings (Systems 0, 1, 2, 3, 4, 5 for our study).

DEVELOPING NEURAL NETWORK STRUCTURES

Introduction

The central idea is as follows: use only the reference heating system (System 0) to develop the best neural network structure, and then to test its generalization capacity by applying it to the other systems (Systems 1 -4) not used during the training stage.

So, a data base of 490 patterns for the seven classes of the System 0, that is 70 patterns for each class (see the section *Data Pre-Processing and Parameters Selection*), has been used to develop several structures of neural networks. We have finally retained two types of structure.

The first is named the **Multiple Artificial Neural Networks** (MANN) structure which is composed of two or more neural networks to discriminate seven classes.

The second is named the **Single Artificial Neural Networks** (SANN) structure which has only one neural network for discriminating seven classes.

These two structures are both multilayered feed-forward networks with a tan-sigmoid transfer function. They were trained using a commercial software (Demuth and Beale, 1992) with an improved back-propagation algorithm, i.e. Levenberg-Marquardt optimization. Both structures are two-layer networks, i.e. one hidden layer and one output layer.

The **number of input neurons** (or input terminals) is equal to one bias neuron (for the threshold) plus the number of components which code the training patterns and the testing patterns in one specific representation space. For example, the patterns coded by two components correspond to three input neurons.

The procedure to determine the **number of neurons in a hidden layer** consists first in training different networks with increasing number of neurons. The Sum-Squared Error (SSE) between the target outputs and the actual outputs of all the training patterns is then computed. This error decreases in principle when the number of neurons increases. A compromise between the number of neurons and the SSE should be found. The solution consists in adding a neuron only if it leads to a large variation of the SSE. Because false classification of 1 pattern leads to an increase of the SSE equal to about 4, we accept to add a new neuron only if it leads to a decrease of SSE equal to or larger than to 4 (Li, 1996).

The **number of neurons in the output layer** must be large enough to code the different classes to be discriminated during the training phase. For our study, we have used as many neurons as the number of the classes to be discriminated during the training phase.

We present first the Multiple Artificial Neural Networks Structure.

Multiple Artificial Neural Networks Structure

The Multiple Artificial Neural Networks (MANN) structure is composed of two neural networks to discriminate seven classes:

- MANN1 (discriminating the heating curve from the other classes),

- MANN2 (discriminating the classes other than the heating curve class - classes n°0, n°1, n°2, n°41, n°42, n°5).

MANN1 (Multiple Artificial Neural Network 1)

MANN1 was trained to discriminate the heating curve class from the other classes. It is well known that the tuning of the heating curve and the detection of heating curve faults are generally difficult for maintenance teams since first, it is not easy to define a reference operating mode indicating a well tuned heating curve; second, the heating curve faults are frequently mixed either with another control system fault or with the normal operating mode. As to our study, the idea is that the heating curve too low - must be detected (if it exists), and then readjusted before diagnosing the six other classes (the class n°0 and the classes n°1, n°2, n°41, n°42, n°5).

Structure

Three parameters - Trocc, Tout, and LUP - were used for coding the 70 training patterns and the 245 testing patterns, i.e. these training patterns and the testing patterns are represented in a three-dimensional representation space. So MANN1 has *four input neurons*.

Using only two classes (normal and heating curve class) to train MANN1, the output patterns of MANN1 are coded by two groups corresponding to the two training classes, i.e. *two neurons in the output layer*. The desired outputs of normal class are [+1 -1]; the desired outputs of heating curve class are [-1 +1].

After many tries, two *neurons* have been retained *in the hidden layer*. This architecture gives the "best" results for training and more importantly for testing. Note that if different neuron numbers give nearly the same testing result, the lowest was chosen as the final structure.

The structure of MANN1 is schematically presented by Figure 2. The training and testing results are presented below.
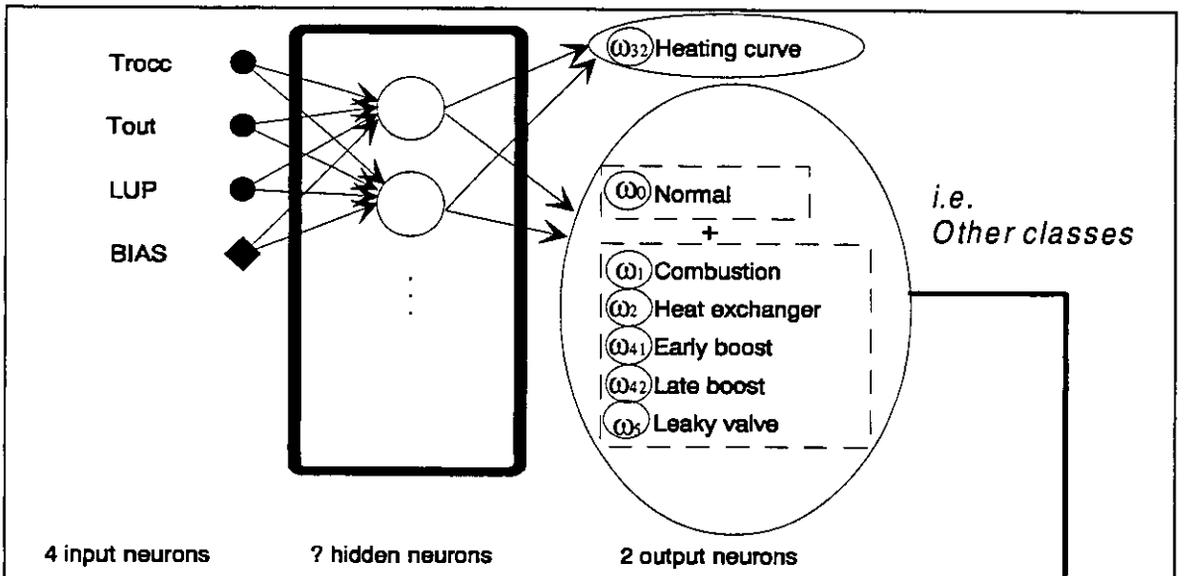
Figure 2 - Structure of MANN1

Troc
Tout
LUP
BIAS

ω_{32} Heating curve

ω_0 Normal
+
ω_1 Combustion
ω_2 Heat exchanger
ω_{41} Early boost
ω_{42} Late boost
ω_5 Leaky valve

i.e.
Other classes

4 input neurons   ? hidden neurons   2 output neurons

Tfgb
Tr6
Tr8
Twunoc
Tout
BIAS

ω_{32} Heating curve

ω_0 Normal
ω_1 Combustion
ω_2 Exchanger
ω_{41} Early boost
ω_{42} Late boost
ω_5 Leaky valve

(6 input neurons)   (? hidden neurons)   (6 output neurons)

Figure 3 - Structure of MANN2

Training and testing

Only two classes (normal and heating curve class), i.e. 70 training patterns (35 days x 2 classes/day), were used to train the network ANN1. But all the seven classes (n°0, n°1, n°2, n°41, n°42, n°5 and n°32), i.e. 245 testing patterns (seven classes/day x 35 days), have been used for testing it. Three possibilities are posible for each pattern. It can be classified in class 32 (heating curve) in other classes(without further distinction) or not classified. The testing result is shown in the Table 1.

To understand this table two points shall be considered:

- Because MANN1 has only two output classifications (i.e. normal and heating curve), the outputs of the other classes (i.e. normal class, combustion class, heat exchanger class, early boost class, late boost class, leaky valve class) are represented here only by the outputs of normal class.

- A not-classified class or pattern means that the network can not classify clearly the actual outputs of one pattern into one specific class and that the classification hesitates among at least two classes. For example, the outputs [+1 +1] can not be classified.

Among the four (i.e. 11% = 4/35) incorrectly classified patterns of the heating curve class (n°32), there are three corresponding to the days with high solar gains (Li, 1996).

actual class	classified in		
	heating curve	other classes	not classified
heating curve (32)	**89 %**	11 %	0
normal (0)	11 %	**89 %**	0
combustion (1)	9 %	**91 %**	0
heat exchanger (2)	9 %	**91 %**	0
early boost (41)	0	**100 %**	0
late boost (42)	26 %	**71 %**	3 %
leaky valve (5)	0	**100 %**	0

Table 1 - Classification rates during testing - MANN1

MANN2 (Artificial Neural Network 2)

Because the MANN1 is used to discriminate the heating curve class, the MANN2 (Figure 3) will be constructed to discriminate the classes other than heating curve. Thus heating curve fault data were not used to train this structure.

Structure

In this case the room temperature during occupancy period (Trocc) and the length of unoccupied period (LUP), which help to discriminate heating curve class from the others, are not presented in the input layer of MANN2 structure. Thus only five parameters have been used for coding the training patterns and the testing patterns and all these patterns are represented in a five-dimensional representation space. As a result, MANN2 has *six input neurons*.

The output patterns of the MANN2 are coded by six groups corresponding to six training classes, i.e. *six neurons in the output layer*.

Training and testing

Different numbers of hidden neurons have been tried during the training phase and an architecture with 4 hidden neurons was finally retained as the best choice.

Both the training set and the testing set are composed of 6 classes (n°0, n°1, n°2, n°41, n°42, n°5), i.e. 210 patterns (35 days x 6 classes/day). The testing results are presented in Table 2.

actual class	classified in						
	normal	combustion	heat exchange	early boost	late boo	leaky valve	not classified
normal	**88%**	0	0	3%	6%	0	3%
combustion	0	**97%**	0	0	0	0	3%
heat exchanger	0	6%	**94%**	0	0	0	0
early boost	9%	0	0	**91%**	0	0	0
late boost	0	0	0	0	**100%**	0	0
leaky valve	0	0	0	0	0	**91%**	9%

Table 2 - Classification rates during testing - MANN2

Comments on the Multiple-ANN Structure

About MANN1

The MANN1 can not detect the heating curve too low (fault 32) for the days with a strong solar influence (see Table 1). During these days the effect of the fault is hidden by the solar gains. The fact that the fault 32 will not be detected on such days does not seem to be very important for following two reasons:

- building occupants will not notice the fault due to the solar gains, because there will be no effect on occupant comfort;

- secondly, no false alarm will be sent to maintenance people.

So we can say that the non detection of the fault 32 on very sunny days does not reflect negatively on MANN1.

MANN1 classifies boost heating too late (fault 42) as heating curve too low (fault 32) in 26% of the cases (Table 1). Why ? When the boost heating is late the room temperature at the beginning of occupation (Tr8) is too low. Consequently the room temperature during occupation (Trocc) is sometimes too low. As it is an input of MANN1, this variable leads in some cases to a wrong diagnosis by MANN1. So if the boost heating is late the whole detection and diagnosis system consisting of MANN1 and MANN2 will in 26% of the cases detect a fault but make a wrong diagnosis. However, this situation can be sensibly improved by using a Single-ANN structure (presented later).

For the other classes the false detection rate of MANN1 are not higher than 11% and the non detection rate is zero. However it is not obvious to determine if such a false detection rate is acceptable for a maintenance team. If MANN1 is used in a real building a compromise will be defined by maintenance team between the time delay necessary to detect the fault and the false alarm rate. If a false alarm rate of 11% is not acceptable, the solution to reduce it will be to trigger an alarm only if the fault is detected on two or three successive days. This will slow down the detection process but reduce drastically the false alarm rate.

About MANN2

The training results of MANN2 leads to incorrect or non-classification for less than 3% (3% = 1/35 pattern) whatever the class to be detected. Only 3 training patterns in total are not correctly classified. Note in particular that 1 pattern of the early boost class has been classified as normal class; this point will be explained below with testing results analysis. The architecture selected does not enable to make a 100% valid diagnosis on the training data, but the results are considered to be good.

Testing MANN2 shows incorrect or non-classification rate comprises between 0 and 12 % depending on the type of class.

Table 2 shows that most of the bad or non classifications concern the confusion between early boost, normal and late boost. This is probably due to the sub-optimal behavior of the "Optimum Start Controller" (OMC) used in the numerical simulation software. Because of the sub-optimality of this OMC, it happens some days that even if the system is properly tuned the boost heating is too late or too early . In such cases a pattern corresponding to a well tuned system will be classified by MANN2 as a late or early boost. If the system is tuned in such a way that most of the days the boost heating will be too early, it will happen some days that the duration of the boost will be correct. The corresponding patterns will be classified by MANN2 in the normal class.

We will find that the confusion between normal class, early boost class and late boost class can be reduced with the Single-ANN structure.

Single Artificial Neural Network Structure

The idea is that only one neural network should be used for discriminating all the seven classes. So, the Single Artificial Neural Network (SANN) structure has to be trained with all the classes (seven in total) to be discriminated.

Structure

To discriminate seven classes, six parameters (Tfgb, Trocc, Tr6, Tr8, Twunoc, Tout) have been used for coding all the training and testing patterns. So the networks SANN has seven input neurons. LUP is not included in the parameters because tests have proven that with this network structure it is not usefull to include it.

With seven training classes, the output patterns of SANN are coded by seven neurons in the output layer.

The structure of SANN networks is schematically presented by Figure 4.
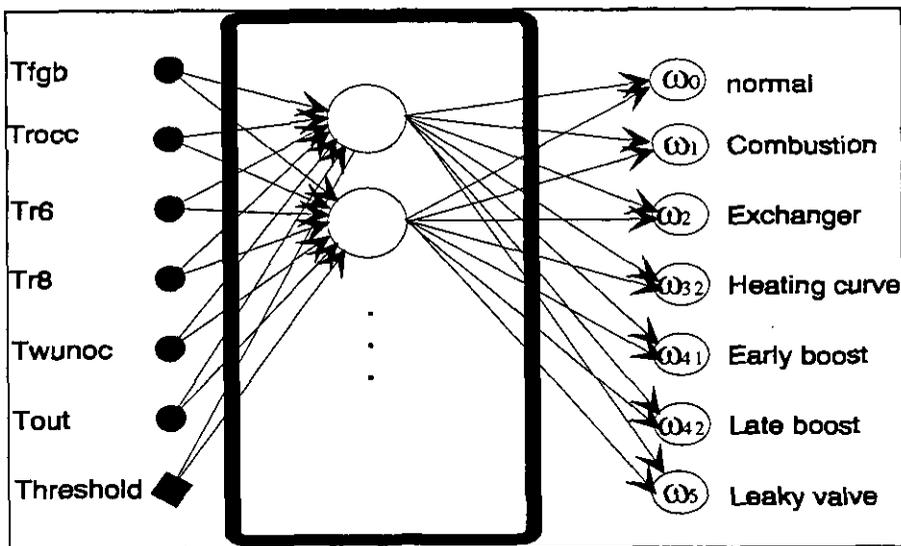
Figure 4: Structure of SANN

Training and testing

The data base of the reference system described-above (490 patterns) has been separated into two parts:

- the first 245 patterns (35 patterns/class x seven classes) are used for training,

- the second 245 patterns are used for testing.

This structure has been trained and tested with different numbers of hidden neurons. Finally one SANN with only 5 hidden neurons has been retained.

The testing result is presented in Table 3.

Tested class	Classified in							
	Normal	Combustion	Exchanger	Heating curve	Early boost	Late boost	Leaky valve	Non-classification
Normal	91%	0	0	0	6%	0	0	3%
Combustion	3%	91%	0	0	0	3%	0	3%
Exchanger	0	0	100%	0	0	0	0	0
heating curve	0	0	0	100%	0	0	0	0
Early boost	3%	0	0	0	91%	0	0	6%
Late boost	0	0	0	0	0	100%	0	0
Leaky valve	6%	0	0	0	0	0	94%	0

Table 3: Classification rates (35 patterns per class) - SANN

Comments on the Single-ANN's Structure

The correct classification rates varies from 91% to 100%, so is very satisfactory.

The detection of the heating curve class and of the late boost class is much easier with one SANN (100% for the 2 classes, see Table 3) than with one MANN (respectively 11% and 26% of bad-classification, see Table 1). And the detection of the normal class is also easier with one SANN.

Why does a SANN perform better than a MANN ? Is it always true ? Mathematically it is difficult to answer these questions. However, we can probably say that a net-structure composed of one single network learns more easily a global knowledge than that composed of two multiple networks.

For all these reasons, the SANN structure has been retained for further study. This study will consist, at the first stage, testing the generalisation capacity of the developed SANN; it will be applied to heating systems not used during the training phase. This approach was chosen because it is most desirable to train a network structure using only one heating system and then to test the capability of the network to generalize on a series of heating systems.

GENERALIZING THE DEVELOPED NETWORK STRUCTURE

We have, up to now, developed one network structure (SANN) by using only 70 simulation days of the reference heating system. The testing result using the reference heating system (used also for training) is very satisfactory (see Table 3).

However we wonder whether this SANN structure can always generate good classification in the case if it is applied to other heating systems not used for training and when the diagnosis period is longer (one whole heating season for example).

To confirm this, the trained SANN structure has been tested by applying it to detect the faults of the 5 simulated heating systems mentioned-above (that is the System 0 and the Systems 1 - 4).

For each system the data corresponding to the faulty and non-faulty states were used to test the SANN structure. 840 patterns (120 patterns /class x 7 classes) of each system have been used for this generalization. One shall note that among these 840 patterns 245 issued form system 0 in total were already used for training the SANN structure. 120 patterns correspond to 120 days from Monday to Friday, which mean 24 weeks as well. This period covers almost the whole heating season of one year. Consequently, there are in the testing data base certainly the days during which the heating systems (especially the boilers) stay in no- or partial running state, so that the SANN structure has more difficulties to detect faults during these days.

The generalisation result are analyzed in two different approaches:

- daily diagnosis which means one diagnosis per day;

- weekly diagnosis which means one diagnosis per week.

Daily Diagnosis Approach

The classification results obtained according to the daily diagnosis approach are presented below. These results are represented by the classification rates which show how different patterns are classified by the SANN structure. Three tables (Tables 4, 5, 6) are presented here-in-below among five: the first is the reference system, the second (system 1) gives the worst classification results and the third (system 4) generates the best classification results for systems which were not used during the training phase.

Tested class	Classified in							
	Normal	Combustion	Exchanger	Heating curve	Early boost	Late boost	Leaky valve	Non-classification
Normal	**78%**	0	0	3%	2%	0	6%	11%
Combustion	1%	**88%**	0	0	0	1%	6%	4%
Exchanger	0	1%	**96%**	0	0	0	1%	2%
heating curve	2%	0	0	**92%**	1%	0	4%	1%
Early boost	2%	0	0	1%	**77%**	0	10%	9%
Late boost	2%	0	0	0	2%	**85%**	1%	10%
Leaky valve	2%	0	0	0	0	0	**97%**	1%

Table 4: Classification Probabilities (120 patterns per class) - System 0

Tested class	Classified in							
	Normal	Combustion	Exchange	Heating curve	Early boost	Late boost	Leaky valve	Non-classification
Normal	**66%**	0	0	0	5%	0	4%	25%
Combustion	16%	**73%**	0	0	0	1%	7%	3%
Exchanger	0	1%	**94%**	0	0	0	0	5%
heating curve	13%	0	0	**58%**	5%	0	3%	21%
Early boost	11%	0	0	0	**31%**	0	11%	47%
Late boost	2%	0	1%	0	0	**92%**	0	6%
Leaky valve	0	0	0	0	0	0	**97%**	3%

Table 5: Classification Probabilities (120 patterns per class) - System 1

Tested class	Classified in							
	Normal	Combustion	Exchange	Heating curve	Early boost	Late boost	Leaky valve	Non-classification
Normal	**75%**	0	0	2%	2%	0	7%	14%
Combustion	0	**86%**	0	0	0	0	7%	7%
Exchanger	0	1%	**94%**	0	0	0	1%	4%
heating curve	2%	0	0	**91%**	1%	0	4%	2%
Early boost	1%	0	0	1%	**80%**	0	10%	8%
Late boost	3%	0	1%	0	1%	**80%**	3%	12%
Leaky valve	5%	0	0	0	0	0	**90%**	5%

Table 6: Classification Probabilities (120 patterns per class) - System 4

In order to facilitate the comparison of the SANN's generalization capacity to different heating systems, the results will be compressed into two global probabilities as defined below:

- RP (Recognition Probability) : probability of good classification for all the seven tested classes. The RP of each system is calculated from the good classification probabilities of 120 patterns of each system. For example, the RP of System 0 presented in Table 7 was calculated from Table 4, that is:

$$RP = \frac{78 + 88 + 96 + 92 + 77 + 85 + 97}{7} = 88$$

- FAP (False Alarm Probability) : probability of detecting a non-existing fault. The FAP of each system is calculated from the bad classification probabilities of 120 patterns of each system. For example, the FAP of System 0 presented in Table 7 was calculated from Table 4, that is:

$$FAP = \frac{(3+2+6)+(1+6)+(1+1)+(1+4)+(1+10)+(2+1)+(0)}{7} = 6$$

Table 7 presents the global probabilities calculated from the classification rates of 5 heating systems (systems 0, 1, 2, 3, 4).

	System 0	System 1	System 2	System 3	System 4	MEAN
RP	88%	73%	80%	85%	85%	82%
FAP	6%	5%	8%	7%	6%	6%

Table 7: Classification Probabilities Based on Daily Diagnosis

Weekly Diagnosis Approach

If we want to do a weekly diagnosis, i.e. one diagnosis per week, the classification probabilities (RP and FAP) are probably different. For this approach we have studied three different diagnoses, that is:

1. One class will be identified if two patterns of this class appeared in one week (one week means 5 days from Monday to Friday). If three patterns or days out of the week have a diagnosis of class X and two patterns have a diagnosis of class Y, This week is considered to be class X.

2. One class will be identified if three patterns of this class appeared in one week (one week means 5 days from Monday to Friday).

3. One class will be identified if four patterns of this class appeared in one week (one week means 5 days from Monday to Friday).

The performances of 4 different diagnosis, i.e. 1 daily diagnosis and 3 weekly diagnoses, are shown (Table 8) by means of two probabilities (RP and FAP), each comprising 5 systems (System 0, System 1, System 2, System 3, System 4).

Diagnosis Approach	Daily	Weekly		
		2 days	3 days	4 days
Recognition Probability	82%	87%	86%	77%
False Alarm Probability	6%	3%	3%	1%

Table 8: Classification Probabilities Based on Daily and Weekly Diagnosis

This comparison shows that:

- based on the principle that one class will be detected if two patterns of this class appeared in one week, the weekly diagnosis produces very satisfactory result (RP=87% & FAP=3%, see Table 8) for all the five heating systems in which four were not used during training phase.

- and one can further reduce the false alarm probability to 1% if it is acceptable to decrease the recognition probability to 77%.

DISCUSSION

Neural Network Structures

The choice of the pertinent parameters to do fault diagnosis (i.e. the choice of representation space for class discrimination) is particularly important, because the choice of a representation space has a great influence on the quality of the discrimination result. In one representation space, the classes can be very badly separated, thus increasing the difficulty of discrimination. A better choice of the representation space gives very distinct classes, so there is not any problem for discrimination. Unfortunately, there isn't any general rule for choosing the representation space; this choice is only based on the knowledge of the problem to be solved, and there isn't any algorithmic method that can be applied to formalize it.

The architecture of a back propagation network is not completely constrained by the problem to be solved. The number of inputs to the network is constrained by the problem and must be well chosen, and the number of neurons in the output layer is constrained by the number of outputs required by the problem. However, the number of layers between network inputs and the output layer and the size of the layers are up to the designer. This work might be very tedious but is crucial. It demands a powerful calculation tool. In our study only the networks with one hidden layer were trained and tested. A criteria was defined to optimize the number of neurons in the hidden layer for each network structure.

Representative Data Base Generation

Another difficulty is linked to the training procedure. To train the neural networks, data representative of the behavior of the system without fault and with each of the faults are needed. The goal of the research is to get a FDD tool which can be implemented in different building-heating systems. It seems impossible to have a training phase for each building-heating system. So we have to find a generic neural network structure which can be used for a series of building-heating systems without re-training. The hypothesis was the possibility to adapt the neural network structure to one particular building-heating system simply by modifying the parameters used to normalize the inputs of the network structure. This is why we used at first one system to develop one network structure (SANN) and secondly four other systems to test its genericity.

We must know that using only the simulation data is not sufficient to have one reliable FDD tool implementable in a real EMCS system. An effort must be done for obtaining a data base from real heating systems-buildings in order to test the performance of the developed network structure (SANN).

If satisfactory results can be obtained by applying the FDD prototype to real heating systems data, it will be possible to prepare a test in a real building and by implementing the FDD prototype in a real EMC-System.

Retained Network Structure (SANN) and Its Generalization

Two types of neural network structures have been presented in this paper: Multiple-ANN and Single-ANN. The testing of these structures have shown that the Single-ANN demonstrates a better performance, in particular to detect the normal class, the heating curve class and the late boost class. The fact that a Single-ANN generates a better result than a Multiple-ANN is probably because the former learns more easily the global knowledge than the latter. So the SANN structure as a FDD prototype has been retained for further study.

The generalization of the SANN structure has been performed with five simulation heating systems. The results were analyzed in two different approaches. The first is based on a daily diagnosis; the second on a weekly diagnosis.

The daily diagnosess approach generates acceptable generalization result. The mean Recognition Probability (RP) of all the 5 systems with seven classes is 82%, with a False Alarm Probability (FAP) of 6% (*) (see Table 7), although a whole heating season simulation data have been used. FAP of 6% means approximately seven false alarms (here is 7 days) for a duration of six heating months. Note that it is more difficult to generalize the SANN structure to a heating system for which the indoor temperature setpoint during occupancy period has increased, especially for detecting the early boost fault and the heating curve fault of this system.

The weekly diagnoses give a satisfactory generalization result, provided that we identify one class when it is detected two/three times per week. In this case the RP is 87% and the FAP is 3%. And the FAP can be reduced to 1% if one accepts reducing the RP to 77% (see Table 8). Note that the false alarms concentrate in particular on the days with very high outdoor temperature (effect of solar gains).

CONCLUSION

The work presented in this paper concerns the development of a Fault Detection and Diagnosis prototype for heating systems by using artificial neural networks. This prototype is expected to diagnose faults in a series of heating systems located in non-permanently occupied buildings such as schools, offices buildings, etc.

Six different types of fault were selected for this work: two related to the gas fired boilers, one to the control valve and three to the control system. The ultimate goal of this work is the development of a FDD tool which can be implemented in real EMC-Systems and enable the automatic detection and diagnosis of these types of faults.

At the first stage of our study, a data base representing the behavior of five buildings and of their heating systems with and without faults was generated by numerical simulation. Then, by using neural networks, a FDD prototype has been proposed and generalized.

This prototype presents a good generalization capacity, although it was trained only by one simulated heating system and a data base of five simulated heating systems during one year heating season (more than six months) was applied for testing this genericity.

The study demonstrates the feasibility of using neural networks for fault detection and diagnosis in such heating systems. Nevertheless the application of this methodology implies training the neural networks with a reliable data base representative of the behavior of heating systems in faulty and non faulty cases. Such a data base is often difficult to obtain.

Also this study shows the possibility that one FDD prototype can be applied to different heating systems without re-training it. The only adaptation is the modification of certain normalisation parameters.

However, the FDD prototype has been up to now studied only on simulation data. Further work will be its validation, which can be performed in two different ways:

- off-line validation by data bases obtained from the measurements of real heating systems-building;

- on-line testing by implementing the FDD prototype in a real EMC-System.

ACKNOWLEDGEMENTS

REFERENCES

D. Caccavelli, E. Hutter, S. Nibel, and P. Picard 1991, Gas Fired heating plant management, *Conference proceedings of Building Simulation '91*, IBPSA, pp. 169-174.

H. Demuth, M. Beale, 1992, *MATLAB Neural Network Toolbox: User's guide.* Natick Massachussetts: The Math Work Inc.

B. Dubuisson, 1990. *Diagnostic et reconnaissance des formes.* Paris, Editions Hermès.

J. Hyvärinen and al.. 1993. *Building Optimization and Fault Diagnosis (BOFD) system concept* (collective document - IEA Annex 25). VTT, Finland.

L. Laret 1980 *Contribution au développement de modèles mathématiques du comportement thermique transitoire de structures d'habitation*, Thèse de doctorat en sciences appliquées, Université de Liège.

X. Li, and J.C. Visier. 1995. *Méthode automatique de détection des défauts et dérives des installations de chauffage - Les défauts de fonctionnement des installations de chauffage collectif.* Cachier de CSTB, Livraison 365 - Décembre 1995, France.

X. Li, H. Vaezi-Nejad and J.C. Visier. *Development of a Fault Diagnosis Méthod for Heating Systems Using Neural Networks.* ASHRAE Trans. Winter meeting February 1996, Atlanta, USA.

J.C Visier, H. Vaezi-Nejad and M. Jandon 1994. Results of the application of a quality assessment procedure to different building energy management systems. 3rd IEEE conference on control applications, Vol 3, pp. 1867-1872.

FAULT DETECTION AND DIAGNOSIS IN DISTRICT HEATING SYSTEM

Yi Jiang, D. Eng.

ABSTRACT

The importance of fault detection in district heating (DH) system has been introduced first. Common faults in DH systems are then listed according to a survey. Faults in DH are then divided into three groups according to its feather. Newly developed procedures, S region contraction method and Fault Direction Space method are then introduced for treatment of different types of fault. The key to rise the sensitivity and reliability of these two methods are discussed in detail. Real procedures used in engineering practices is briefly summarised finally.

INTRODUCTION

During the last ten years, district heating systems (DH) have been widely developed in China. Large and middle size of DH systems are built almost in every big and middle cities in the North and Northeast of China with the capacities of 50 MW to 800 MW for each and heated building area of 1 to 10 million square meters. Figures 1 is one of DH networks in Shenyiang, as an example, with the capacity of 400 MW and about 10 million square meters building heated. Figure 2 is the scheme of a substation in this network. There are more than 100 substations like this connected with the water network. As the system is getting larger, the monitor and control of the network and substations is getting important. Distributed computer system is used to look after the network as well as substations. Figure 2 also shows the control and measuring points of the computer controller in a substation. A centre computer in the control centre links with all the controllers through a computer network. This kind of computer system can be found in most of DH systems in North Europe. It starts to be popular in China. There are around 20 DH systems installed with this kind of computer system. Therefore what are the tasks of the computer system in DH? It is learnt from a survey as the follows:

- collect real time operation data, such as temperature, pressure and flowrate, from each substation;
- count the heat, electric and water consumption of each substation as well as each part of the network;

Yi JIANG is Professor, Department of Thermal Energy Engineering, Tsinghua University, Beijing 100084, P R China.
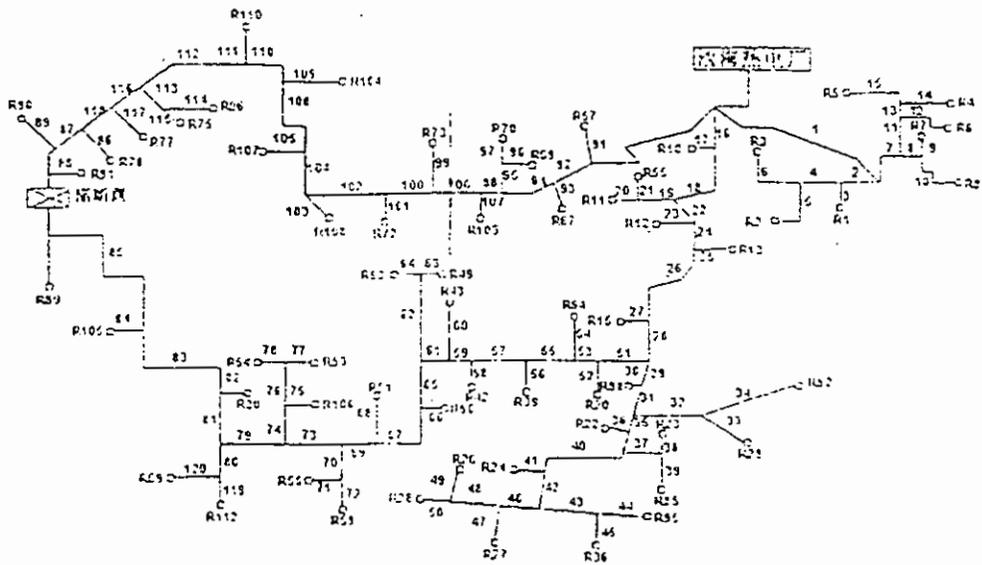
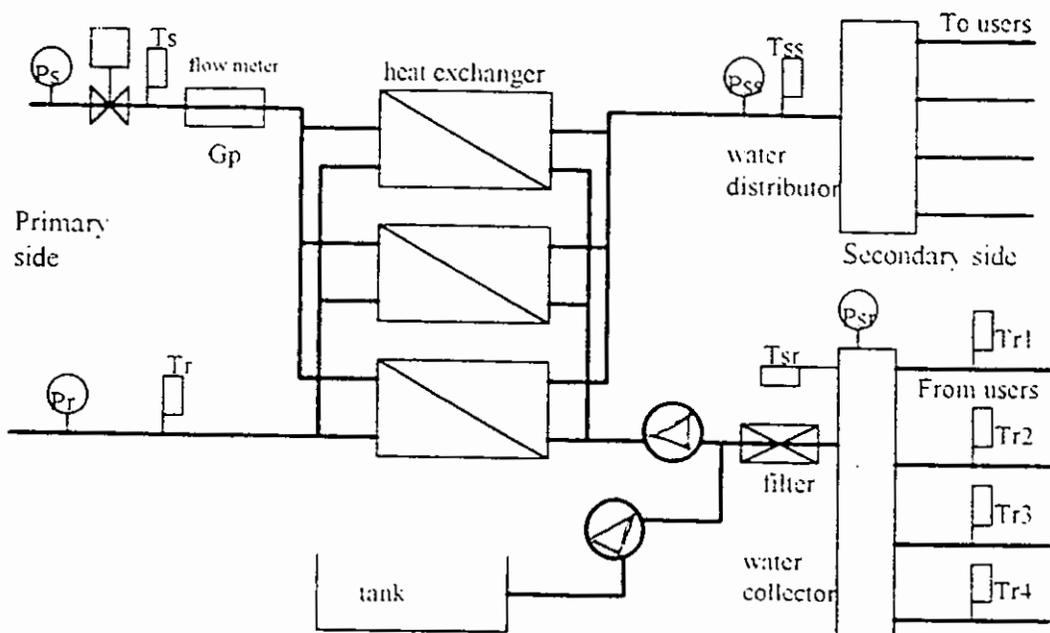Figure 1 Primary network of Shenyiang DH system

Figure 2 Scheme of a substation and the measurments

36

- regulate the water flow rate distribution by controlling of the speed of circulation pumps or the open ratio of motor-drive valves to match the variation in the requirements of substations.
- alarm, if some unusual measured data have been found such as a pressure exceeds the upper limit value or the flow rate in a branch is getting smaller than the low limitation.

One of the most important tasks of the computer system for DH is to tell operators what is happened in the DH network. So that operators can make some relative regulation if some operation state changes; or treat faults on time to avoid accident when a fault is found; or plan the maintenance schedules according to the equipment real situation learnt from the measured data. However, as a large DH system, the number of measuring points can be as many as few thousands, it is a hard work to study these huge data from more than thousands measuring points and to discover fault as well as abnormal state. Furthermore, most of fault and abnormal state cannot be found out directly from measured data in the way of simlpe limit checking. What appears in measured data when something fail is the change in relationship among measurements. For instance, if there is a leak between the primary side and the secondary side of a plate type heat exchanger, the temperature and pressure at the both sides may change more or less but none of the measurement could be come to the up or low limit. It can hardly be discovered without detailed analysis at the beginning when the leak is not so serious. Therefrom, *on-line fault detection facilities* is highly recommended so that:
- operators can be liberated from looking around the thousands measurements. Data can be checked automatically and fault and abnormal state can then be discovred as the result;
- the in-direct fault can also be discovered using the powerful computing and logic analysing abilities of computers.
- reports can be made by computers to give guidances to the repairman for treatment of the fault discovered. Maintenance schedule can also be made according to the equipment situation the computer detected. High level maintenance can be achieved by aiding of the computer system.

This is one of the goals for distributed computer monitor and control systems of DH. It is far more than the up and down limits checking of the measured data. The centre point is:

Can and How we get more information from the measured data and make clear judgement on the system?

As part of the IEA Annex 25 "Fault Detection and Diagnosis in HVAC" (FDD) project, theoretical research and engineering practices has been carried out in Tsinghua University. This paper is the summary of this research. Faults need to be detected in DH system have been listed first. The problems and difficulties in real FDD procedure for DH system has discussed then. The basic philosophy of the DH FDD in our research is introduced. The S region contraction approach and the Fault Direction Space (FDS) method is presented in detail. Engineering applications in some Chinese DH systems have been brief introduced finally.

FAULTS IN DISTRICT HEATING NETWORKS

According to a survey from some DH systems, faults in DH can be classified as follows:

Sensors fault: as there are few thousands sensors of temperature, pressure and flow-meter in a DH network, any procedure and analysis cannot relied on the hope that all the data from

sensors are correct. To produce a correct result from the measurement, the first thing is to pick out the wrong measurements from the data and find out fault sensors. There are three types of fault sensors in generally:

- dead sensor, that gives a zero or a constant output;
- high frequency noise sensor of which the output changes in a greater range than permissible one and varys with a high frequency;
- slowly floating sensor that keeps going up or down slowly as the time being or along with the environment temperature changes.

Limits check combining with a filter such as Kaman's filter can take the first two types of fault away. However the third one is not easy to be discovered by the measured history data itself. For instance if the output from a pressure sensor in the network increased 0.1% per day, no filter nor limit checking can get it. This can only be detected systematically by analysing the pressure distribution of the whole DH network.

Fault Component in substation: circulating pump at a substation stops; motor-drive valve cannot open/close; water filter blocked as well as some types of fault in heat exchanger at a substation, which can be:

- leakage between the primary side and the secondary side;
- block in the primary side or the secondary side;
- poor performance of heat transfer due to scaling;
- unbalanced water dispensing among heat exchangers in a substation, water flow rate in one heat exchanger is much more than in another;

Among these types of fault, pump stopping can be discovered by checking of the pressure difference inlet and outlet of the pump or check the signal from the electric panel to know if the contact for pump is still closed.

Fault in motor-drive valve can be detected in two steps: comparing the measured open-radio with the required one to know if the electric and electronic part of the valve is correct; turn the valve open/close to a greater or less extent then to see if the pressure measured on the down stream of the valve does change.

Partly blocked water filter can also be checked out by the pressure different at the two sides of the filter if there are pressure sensors. If sensors are not available on these positions, it has to be judged from the temperature, pressure and flow rate measured within the substation by some systematically analysis method.

The most difficult detected fault in a substation is the fault of heat exchanger. As there is not a direct signal to indicate its working state, a systematically analysis procedure is needed to estimate the heat exchanger performance and detect if there is a fault.

Fault in the primary network: The primary network is defined as the network from the thermal plant to the substation. Four types of fault may be happened in a primary network of a DH system:

- high heat loss causing by poor insulated pipe;
- high pressure drop causing by a partly blocked branch;
- leakage in the network;

- short circuit between a supply branch and nearby return branch due to a mis-opened or damaged short circuit valve.

The poor insulation of pipe can be discovered directly by checking of the supply water temperature difference between two nearby substations. Thermal resistance from the trunk to outside, R, can be calculated as

$$R = \frac{L \int (T_{w,m} - T_{outside})d\tau}{\int \Delta TG d\tau}$$

where the G is the flow rate through the trunk which can be obtained by summing up the measured flow rate of each substation in the down stream. $T_{w,m}$ is the average supply water temperature. As the temperature may change with time, integration is needed to avoid the dynamic effect so to obtain the steady state thermal resistance from the water to outside, R. This value should be higher than a given limit value, limit checking on the R value can then find the part of poor insulated trunk.

However other faults listed above cannot be discovered as simple as the "poor insulation" one. Leak at a pipe of the trunk cause some of pressures go down, while "short circle" is very possible get the similar effect. To discover these fault, the pressure and flow rate distribution over the network has to be studied in detail and systemic analysis has to be carried out to find the type of fault and its location.

Fault in secondary networks: The secondary network means the network distributing hot water from the substation to every buildings it serviced. There is normally a water distributor in the substation. Few of branches from the distributor to buildings and back from buildings to a water collector in the substation as shown in Figure 2. Thereby water temperature of every branch returning from buildings can be measured at the substation. Faults at the branch level in secondary networks normally are:

- leak in one of branch of the secondary network;
- unbalanced water distribution due to mis-opened one of the branches or block/close another;
- short circuit in one branch between the supply pipe and the return pipe due to mis-open a short circuit valve.

These types of fault cannot be detected directly from the measured pressure and temperature unless the extent of fault is very serious. Systematically analysis is needed for these types of fault.

THE LIMITATION IN ENGINEERIG PRACTICS

From the listed possible fault above, it is understood that some types of fault cannot be detected simply by limit checking or some kinds of rule basis analysis locally. It has to be analysed systematically. To carry on this analysis, the network structure and the flow resistance of each branch of the network should be know so to predicate the system performance and comparing with the measured data. As a network of DH system, the topologic structure is determined according to the design. However the flow resistance performance
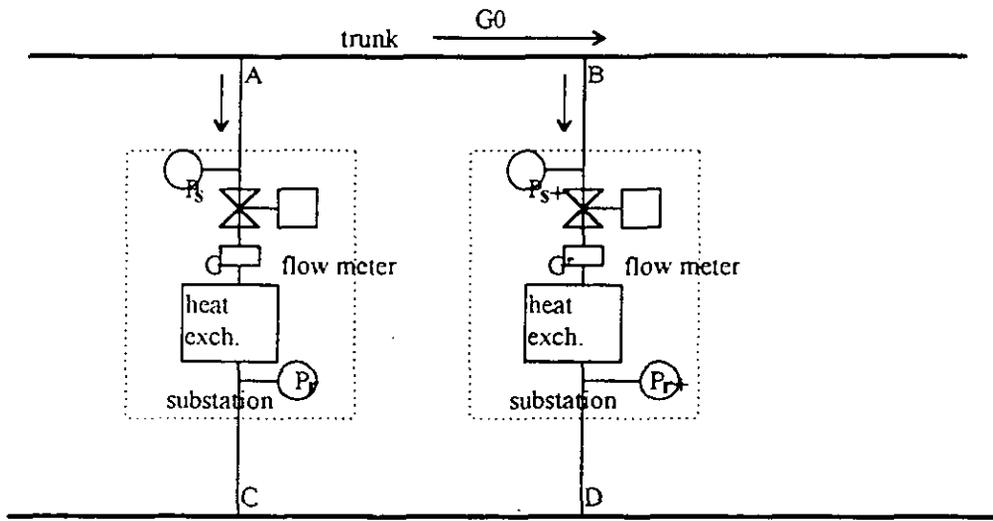
Figure 3 Part of the primary network in DH system

of each branch of the network cannot be provided from the design. Defining the flow resistance coefficient of a branch. S, is

$$\Delta P = SG^2 \tag{1}$$

where the ΔP is the pressure drop from one end of the branch to another. The S value can be assumed as flow rate independent. However as there are many pipe bends, three-way tubes and manual valves of which the flow resistance performance depends on the installation and local initial regulation, the S value for each branch cannot be pre-determinate, it has to be on-line estimated from the measured data.

Figure 3 is a part of the primary network of a DH system. If install pressure sensors at each junctions of the network as the points A, B, C, ... in that drawing, and install flow meters on each branch, the S value of every branch is very easy to be estimated from equation (1). However, in most of DH systems, pipes are lied underground directly. Sensors can only be installed inside of substations. Then we cannot put any pressure sensor on the junctions as the points A, B, C, ..., nor put flow meters on the trunk outside of the substation as indicated in Figure 3. The measuring points can only be located within substations. In some real system, the distance from the junction to the substation where the pressure sensor is allowed to be located can be as long as few hundred meters, and the pressure drop from the junction to the pressure measuring point can be dozens kP. Due to this practical limitation, most of S values of the network cannot be measured directly.

Beside of the location problem, the number of sensors is also limited. As the cost of sensors especially the one for pressure and the one for flow rate is quit high, to keep the total investment at an accepted level, the number of installed sensors should be as less as possible. The question is rising then in order to know the system performance and detect the faults above, what is the minimum number of sensors needed?

BASIC PRINCIPLE AND PROCEDURE OF FAULT DETECTION

Now, from the discussion above, there is not an universal method to detect all kinds of fault. More than one approaches are needed for detecting faults with different features. Three groups can then be divided according to the performance of faults: the fault that can be detected according to the measured data from one or two sensors individually; the fault that has to be detected using the measured data within the substation where the fault is; and the fault that has to take the operation data over the whole network to analysis systematically.

Dead or high frequency noise sensors belong to the first group. They can be picked up by limit checking and a data filter on the measured data. Damaged valve, blocked water filter, pump stopping as well as poor insulation of pipe can also belong to this group. They can be detected according to an "if-then" rule base knowledge. The only thing needs to be designed carefully is the threshold of each rule. This is *individual level of fault detection*.

Fault in a group of heat exchangers, such as leak between its two sides or the poor heat transfer performance due to scaling, and fault in the secondary network under a substation are in the second group, *substation level fault detection*, which takes the measured data within a substation as the start point. While the fault in the primary network, such as a leak or block of a branch in the primary network, has to be analysed according to the pressure and flow rate distribution over the whole network. This belongs to the third group, the *centre level fault detection*.

Centre level Fault detection

Let us go back to Figure 3 to study the primary network. The part through a substation is drawn as a branch in the network. Two pressure sensors are located at the inlet and outlet pipes of the primary loop in the substation for measuring the inlet and outlet pressures. Flow meter is also located in the substation for measuring of the flow rate in the primary loop of this substation. The squat in the drawing represents a heat exchanger or a secondary local network. An electric motor drive valve is located before the heat exchanger which can regulate the flow rate according to the heat load of the substation. The question is how we can detect if one of the pressure sensor or flow meter gives a wrong report or if there is a block or leak in one of the branches of the primary network? If we know the values of resistance coefficient of each branch, S, the pressure and flow rate can be calculated according to the S values of each branches and the pressures at inlet and outlet points of main circulation pumps. Comparing the calculated pressure and flowrate with the measured one, faults can be detected and classified from the differences:

If only one or two values are different between measured and calculated data, it is very possible to be the sensor failure. If there is a part of the network that most of the data in this part are different between measuring and calculation, there must be a block or a leak in one of the branches in/nearby this part. [1] explains how to detect the block/leak pipe from these differences. However in most cases, the reliable values of the resistance coefficients, S, of each branch of the network are not available as discussed above. The S value is also hardly determined directly from equation (1) as the limitation of the location for pressure sensor. What we know is the relationship between the measured data. If the trunk does not change, this relationship should be fixed regardless what the operation state is. From this idea, a new procedure has been developed which just based on this relationship. Study the history data to know the relationship then discover the range of S value for each branch and detect fault in the primary network in the same time.

According to equation (1) and figure 3, relationship between two pressure measurement points can be described as

$$P_{s+} + G_+^2 S_{s+} - G_{0+}^2 S_{0+} - G^2 S_s - P_s = 0 \qquad (2)$$

where the P and G are measured pressure and flow rate. The flow rate at the trunk pipe, G0, can be determined by adding all the measured flow rate at each down stream substation branches as

41

$$G_{0+} = \sum_{i \in down} G_i$$

. If the permitted errors of sensor are ΔP and ΔG, the relationship between the S's should be

$$(G_+ + \Delta G)^2 S_{s+} - (G_{0+} - \Delta G)^2 S_{0+} - (G - \Delta G)^2 S_s > P_s - P_{s+} - 2\Delta P$$
$$(G_+ - \Delta G)^2 S_{s+} - (G_{0+} + \Delta G)^2 S_{0+} - (G + \Delta G)^2 S_s < P_s - P_{s+} + 2\Delta P \quad (3)$$

This gives the up and down limitation of the S's values of the branches in this piece of network. Similar equations can also be made for the returning part of the network. In some part of the network where the trunk splits into two branches, four different branches may be involved in one equation as the relationship of S values. In this way, every S of supply pipe and return pipe for each substation will appear in two groups of limitation equations and every S for trunk pipe will appear in a group of limitation equations. These limitation equations defines a possible region of S values, Ω. As time being, the flow state will change as the motor drive valves are regulated or/and operation state of circulation pumps change,

the measured flow rate and pressure changes correspondingly. A new group of limitation equations can be obtained and a new region of S values can also be determinated as Ωnew. However as the network is not changed, the S values of each branch should keep as the same as previous. The real region for S values, Ωr, should then be

$$\Omega_r = \Omega \bullet \Omega_{new} \quad (4)$$

The Ωr defines a narrower region of S values than Ω and Ωnew. Repeat the procedure of equation (4) for each time, the S region, Ω , can be contracted step by step and a narrower but more accuracy region of S values can be obtained finally.
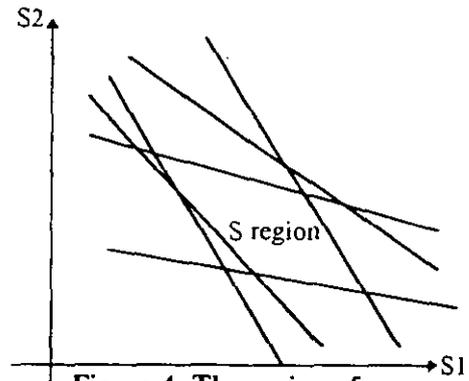
Figure 4 The region of

Figure 4 is a simplified two dimensional example of this procedure. In a real process, the S region will become a polyhedron in a multidimensional space. Once a sensor fault or a branch block/leak, there will be no common region between the Ωnew, the S region determinated from the newly measured data, and the Ω, the previous S region obtained from the history measured data. The Ωr from equation (4) will be a zero set. This indicates that there is a fault somewhere in the network. Further analysis can then find out the type and the location of the fault. This is the basic idea of the S region contraction method. Detailed procedure is described as follow:

0. Set-up three or four dimensional squares for each relationship equations as the initial region of Ω. The length of the square should be designed according to the un-determinated level of the S value. For instance if the estimated S value is S0 with the un-determinated level of ΔS, the bound of the square in this S direction should be from S0-ΔS to S0+ΔS. There should be two squares for each substation, or says two Ω's, one for the supply part and other for return part.

1. Produce a planes according to equation (2) basing on the measured data for every region.

2. Check if the plane cross over the S region, Ω;

3. If so, produce another S region Ωnew according to Equation (3) then obtain the modified Ωr according to the Equation (4). The S region has been contracted in this way.

4. If NOT in step 2, produce another S region Ωnew according to Equation (3) and check if the Ωnew and the previous Ω have a common region;

5. If so, it means the reason that the plane is outside of the previous Ω is because the errors in measurement are in the permissible range. In this situation the Ω is not going to be changed.

6. If NOT in step 4, it means there must be something wrong. Calculate the distance between the centre point of the Ω and the measured plane, Δfs and Δfr of each substation can be obtained as

$$
\begin{aligned}
\Delta f_s &= G_+^2 \overline{S_{s+}} - G_{0+}^2 \overline{S_{0+}} - G^2 \overline{S_s} + P_{s+} - P_s \\
\Delta f_r &= G_+^2 \overline{S_{r+}} - G_{0+}^2 \overline{S_{r0+}} - G^2 \overline{S_r} - P_{r+} + P_r
\end{aligned}
\tag{5}
$$

The type of fault can then be classified from the Δf 's of the substation and nearby. According to simulations and analysis, Table I lists how the Δf 's change corresponding to each type of fault. From this result it is understood that most types of fault can be classified from the changes of Δf in the substation and nearby substations. There is a few of situations when the symptoms are the same for different types of fault:

	Type of fault	Δfs	Δfr	Δfs, down	Δfr, down	Δfs, every up	Δfr, every up
1	sup. branch block	↓		↑			
2	sup., trunk block	↓					
3	ret. branch, block		↓		↑		
4	ret., trunk block		↓				
5	P sup., ↑	↓		↑			
6	P sup., ↓	↑	↓				
7	P ret., ↑		↑		↓		
8	P ret., ↓		↓		↑		
9	G ↑	↓	↓	↑	↑	↓	↓
10	G ↓	↑	↑	↓	↓	↑	↑
11	user's branch leak		↓		↑		↓
12	sup. branch leak	↑	↓			↑	
13	ret. branch leak		↓		↑		↓
14	sup. trunk leak	↑				↑	↑
15	ret. trunk leak		↑			↑	↑
16	Short circle between sup. and ret.	↑	↑			↑	↑

Table 1 How Δf's change corresponding with the types of fault

where Δfs,down/Δfr, down: Δfs/Δfr of the down stream substation

Δfs, every up/Δfr, every up: Δfs/Δfr of the substations in every up stream

sup. branch: from the junction between supply trunk and substation branch to the location where the pressure sensor for supply branch is located.

ret. branch: from the junction between return trunk and substation branch to the location where the pressure sensor for return branch is located.

user's branch: from the point of the supply pressure measurement to the point of return pressure measurement

P↑, P↓, G↑, G↓: the measured data is too high/low due to sensor fault

- Block in supply branch and supply pressure sensor producing too high output appear in the same symptom: (Δf s) goes down and the (Δf s,down) goes up. In this case, computer can close the motor-drive valve of this branch and let the flow rate at this branch be zero. Check the Δf again. If the supply branch is blocked, the Δf 's should go to normal. While if it is the sensor fault, Δf 's should still be the same trend.

- Block in return branch and return pressure sensor producing too low output appear in the same symptom: (Δf r) goes down and the (Δf r,down) goes up. In this case, computer can close the motor-drive valve of this branch and let the flow rate at this branch be zero. Check the Δf again. If the return branch is blocked, the Δf 's should go to normal. While if it is the sensor fault, Δf 's should still be the same direction.

- Leak at return branch and leak at user's branch appear in same symptom. This cannot be judged unless install another pressure sensor at the point between the motor-drive valve and the user's branch. Close the valve then measure the pressure different between this additional sensor and the return pressure sensor, the location of the leak can then be detected.

The key of the whole procedure is the establishment and modification of the S region, Ω. To record every region Ω for the relationship of each supply part and return part of substation, the co-ordinate of every peak point of the Ω has to be recorded comparing with the knowledge of how the peak points make up a edge and what edges make up a plane. Geometry calculation is needed to cut off some old peak points and edges and produce some new ones so to contract the old Ω to a new one. A special geometry calculation procedure has been developed for this narrowing down process. Further analysis by simulations shows that to make the process successful is relied on the correct estimation of the value for normal errors of each measurement. Improper values set to sensors as the normal errors may cause the Ω too large that cannot indicate a fault if the values of permissible errors are set too large or cause the Ω deviate to the region of which the real S values are outside if the errors values are set too small. After a fault sensor has been discovered, its output should be replaced by an estimated value. The estimation can be done by a calculation basing on the S coefficients of the network and other measured data. S values at the centre of the Ω can be used for this calculation.

Simulation and on-line test shows that at beginning, the Ω reduced very fast providing the operation state changes in a large range. After a period, the Ω stops reducing as the Ωnew is almost equal or larger than the old Ω. The length of this period depends on how the system has been regulated. In the beginning of a heating season, valves have been regulated frequently in a large range as the heat load distribution has been changed during the un-heating season. This is the best chance to estimate the Ω. Volume of each Ω can also be calculated to indicate how the estimation process getting on and how well of the Ω is. If the volume of Ω keeps to be at a certain high level and cannot go down any more, it may be the reason that the range which the state of the system changes within is too small. In this situation, some of valves can be closed a little while and then opened back again to make the system state change so to get better estimation of the region Ω.

Substation level Fault detection

This is to detect fault inside of the substation and in the connected secondary network. From the previous discussion, typical faults involved are:

- fault sensor that produces too high or too low output due to slowly float. The flow meter and pressure sensors at primary side has been checked out at the centre level, so only the pressure sensor and temperature sensors at the secondary side and the return water temperature sensor at the primary side need to be checked.
- faults in heat exchanger. Such as leak between the two sides, block in one side or poor heat transfer performance due to scaling. If there are more than one set of heat exchanger, another possible fault is the unbalanced water distribution among the heat exchangers.
- fault in the secondary network as listed before.

All the fault listed above cannot be easily detected by a rule base reasoning procedure. It can be hardly judged if the change from an output of sensor is because the real state has been changed or is because a fault happens. One of the approach is to estimate a model for heat exchanger/secondary network basing on the previous measured data first. Using this model to predicate the performance of the heat exchanger/secondary network and comparing the result with measured data. Fault can then be indicated and classified from the differences. This approach works in some cases however it involves too many stages. Computer has to switch from the "learning process" to the "predicating and fault detecting process". The first phase basis on the idea that the sensors and the system are all right. Computer just learns from this period of operation and take the performance in this period as the standard. In the second phase, the task is to find the difference from the first phase in the system performance. When we should switch from the first phase to the second one? If where is a difference, should we point out that there is a fault or modify the model to make it compatible with the newly measured data? The logic is not very clear in some sense.

To make the process simple and to avoid this confuse, a new procedure "Fault Direction Space" (FDS) has been developed and used in DH system since last two years. The basic idea of the FDS is to merge the three steps (model estimation, performance prediction and match difference to fault) into a global exercise to make it simple and uniform. Character Parameters (CP) are estimated according to the structure of the object to be detected. The CP is a composition of measurements in one hand. It is also a composition of structure properties of the object on the other hand. As it is the composition of measurement, its value can be calculated from the measured data during operation. As it is the composition of the structure properties, it should not change during a normal fault less operation. Observe the CP value during the operation. When great changes of the CP have been examined, a fault appears. Put the change, ΔCP, to a pre-designed FDS space, the location of ΔCP in the FDS space indicates the type of fault. Following is how the FDS method used in fault detection at the substation level.

FDS for heat exchanger. Figure 2 is a heat exchanger in substation and the measuring points around it. The heat and flow mechanical balance equations can then be written as

$$Q = UF \frac{T_{PS} - T_{PR} - T_{SS} + T_{SR}}{Ln(\frac{T_{PS} - T_{SS}}{T_{PR} - T_{SR}})} = UF\Delta T_{Li}$$

$$Q = G_p c_p (T_{PS} - T_{PR}) = G_S c_p (T_{SS} - T_{SR})$$

$$G_p = \sqrt{(P_{PS} - P_{PR})/S_p} \qquad G_s = \sqrt{(P_{SS} - P_{SR})/S_s}$$

where the F and U is the area and heat transfer coefficient respectively. From this balance equation, two CPs can be estimated as:

$$CP_1 = S_p / S_i = \frac{(T_{PS} - T_{PR})\sqrt{P_{PS} - P_{PR}}}{(T_{SS} - T_{SR})\sqrt{P_{SS} - P_{SR}}}$$

$$CP_2 = c_p / UFS_i = \frac{\Delta T_{Ln}}{(T_{SS} - T_{SR})\sqrt{P_{SS} - P_{SR}}}$$

Because the resistance coefficient of the heat exchanger at two sides, Sr ans Ss, and the heat transfer coefficient, U, as well as the heat transfer area, F, should not change within a normal working state, the two Character Parameters (CPs) should also keep in constant. In this way the CP becomes an indicator of fault of the heat exchanger. In a fault-free state, the CPs change only within a small range regardless how the operation state changes. If a fault

Figure 5 FDS for heat exchanger

happens, the CP will change greatly. Then the type of fault can be judged from the shape how the CP's change. Using the change of the two CP's, ΔCP1 and ΔCP2 as the co-ordinator, Figure 5 shows how the CP change for different types of fault. It is clearly that the direction to which the ΔCP's deviate from the zero point can indicate the type of fault. The space consisting of ΔCP1 and ΔCP2 is called Fault Direction Space (FDS).

As the CP comes from the steady state balance equation, the dynamic influence in the measured data should be filtered out before to composite it. As the thermal mass of the heat exchanger, outlet temperature cannot follow the change of the inlet temperature on time. Move average needs to smooth the measured temperature data before it is used in CP. Figure 6 shows how the two CPs change during a fault less process. This is measured from a real DH system. It shows that the CP changes around a constant value within a small range although the moving average treatment on the temperature has been done. This is because the moving average treatment cannot filter all the dynamic influence out. There is still something remainder. The no-linear properties of the heat exchanger, such as the U value and S value change along with the change in flow rate, can also make the CP change. In addition, the normal error in sensor's output is another reason for the change of the CP. Considering these reasons, a normal region should be defined in the FDS. When the ΔCP is inside of this region, it is considered as a fault less state. If the ΔCP is out of this region, warning can then be made to point out the type of fault according to the direction of the ΔCP.

Further study shows that the normal region is operation state dependent. As the CP1 and CP2 are the quotient by ΔT and ΔP, same measurement errors and non-linear effect can cause different level in CP if the ΔT and ΔP is different as at different operation state. Therefore, to obtain an uniform normal region, normalised ΔCP, the ΔCP* has to be used to replace ΔCP:
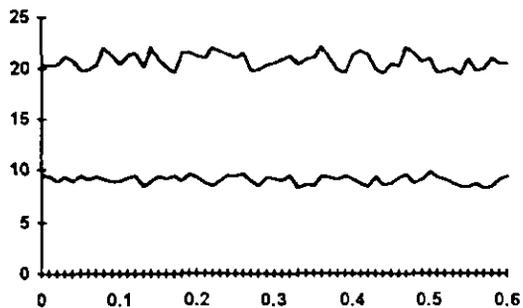
Figure 6 Measured CP1 and CP2

$$\Delta CP^* = \Delta CP / DEV(\Delta CP)$$

46

where DEV(ΔCP) is the deviation of ΔCP causing by the sensor's error at normal level. In this way the normal region in the FDS as well as the direction of each type of fault can be independent with operation state.

Detailed study has been done to know the sensitivity, reliability and distinguishable ability of this procedure. The sensitivity study is to know how serious the fault is when it is able to be discovered. It is learnt that the sensitivity is strongly dependent on the normal errors of sensors. High precision sensor used will result in high sensitivity. However after the precision of sensor comes to a certainly level, the non-linear properties becomes the wall against rising the sensitivity. Reliability study is to know how often a fault report is made when it is a fault less state? This is very dependent on the normalisation of ΔCP which then related on the estimation of the range of normal errors of sensors output. Distinguishable ability study is to know if this procedure can point out the type of fault correctly. Fuzzy match can then be used to find out the type of fault when the direction of ΔCP is between two types of fault directions. These results are present in [2].

Fault in secondary network: From the similar analysis, another FDS can be constructed to detect faults in secondary networks. As drawing in Figure 2, temperature sensors are located at each branches of the collector to measure the return water temperature from each user's branch. CP's for detection of fault in each user branch can be structured as

$$CP_0 = \frac{T_{ss} + T_{r0} - 2T_{outdoor}}{(T_{ss} - T_{r0})\sqrt{P_{ss} - P_{sr}}}$$

$$CP_i = \frac{T_{ss} + T_{ri} - 2T_{outdoor}}{(T_{ss} - T_{ri})\sqrt{P_{ss} - P_{sr}}}$$

$i = 1, 2, 3, \ldots$
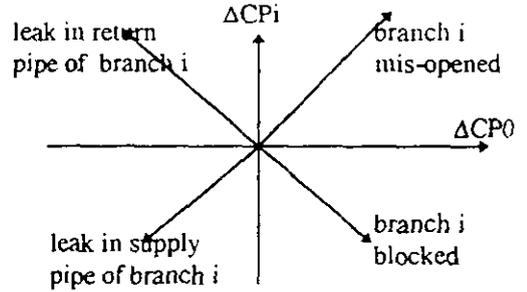
Figure 7 FDS of secondary network

Figure 7 is the FDS constructed with ΔCP0 and ΔCpi. The largest mod of CPi can be determinated from all the branches within a substation. If the point of this ΔCPi combining with the ΔCP0 is still inside of the Normal region , it can be thought as being a fault less state. If this point is out of the Normal region, the type of fault can then be distinguished by the direction of the point deviated from the zero point. The possible faults it is able to detected are:

- the branch i blocked;
- leak in the supply pipe of i'th branch;
- leak in the return pipe of the i'th branch;
- the water flow rate in i'th branch is too large.

Fault sensors in substation: Slow floated pressure sensor and flow meter at the primary side of the heat exchanger has been detected at the centre detection level. However, the pressure sensor and temperature sensor at the secondary side and the temperature sensor for return water outlet from the heat exchanger at the primary side should also be checked so to make sure the change in CPs used above is not caused by fault sensors. Temperature sensors can be checked by the relationships among them as

Tps > Tss;   supply temp. should be higher than the return temp. of the primary side
Tss > Tsr;   heated water should be warmer than the water inlet of the exch. at secondary side

47

Tpr > Tsr;   return water at primary side should be warmer than that at secondary side
max(Tsr,i) > Tsr > min(Tsr, i); inlet water at secondary side should be the mixture of the water
from each branch backing to the collector.

If the relationship above cannot be satisfied with measured data, it means one of the temperature sensors is wrong. Further analysis can point out the fault sensor [3].

Another CP can be added to indicate the fault pressure sensor at the secondary side of the heat exchanger as

$$CP_3 = \frac{(T_{ps} - T_{pr})G_p}{(T_{ss} - T_{sr})\sqrt{P_{ss} - P_{sr}}}$$

Combining with CP1 and CP2 in the heat exchanger section, these three CP's can work for both the heat exchanger and the sensors around. If there is a leak between the two sides of the exchanger, in addition of the change in CP1 and CP2 (go down), the CP3 will go up. In all the other situations the CP3 will be constant unless the pressure sensor fault. If it is pressure sensor fault causing the ΔP too small, CP1, CP2 and CP3 will go up. If the ΔP is too large due to pressure sensor fault, CP1, CP2 and CP3 will go down.

ENGINEERING PRACTIC

The fault detection procedures discussed above has been used in a few of DH systems in China as part of the distributed computer control and monitoring system. The computer system consists of a control centre installed with few computers and distributed controllers at every substation. Computer network links controllers to the control centre. The fault detection is carried out in the control centre. Each local controller send the measured data to the control centre every minute. A filter is used to remove the high frequency noise of sensor's output and discover some types of fault sensor if the range of the noise is too high. "if -- then" rules are then used to detect if there is a circulation pump in a substation being stop or is a damaged motor-drive valve or other individual fault. Fault detection for primary network is then carried on for every three or four minutes. Fault pressure sensor and flow meters, leaked or blocked branches of primary network as well as poor insulated pipe are then detected by the procedure described in this paper. The possible regions of S value for each branch in the primary network are estimated step by step in the same time by this procedure. These estimated S values are also used for regulation of the network, in addition of detecting fault, to distribute flow rate according to the varied load distribution. Fault detection of substation is then taken at every three or four minutes. This is for checking the situation of heat exchangers and every secondary network under each substation. Problems in heat exchangers, in secondary networks and in sensors at the secondary side are discovered in this exercise. If something detected, warning is displayed on the screen of the control centre. Fault report is also printed out responding to the operator's command. Detailed information about the fault is printed in the report such as the type of fault, the location of the fault, the measured data before/after the fault, the reason to consider there is the fault and the guidance for treatment. Fault information is also recorded into a special database so that it can be checked and analysed later. During the last two years operation, it seems work well. This system makes operators much easy to discover and repair fault on time and enable the DH system at high reliable working state.

CONCLUSION

As the computer industry has being developed very fast, computer control and monitoring system has been widely used in the HVAC field as well as urban distribution networks. However the master computer rests at most of time in most of the applications because there is not enough work to do and the computer is getting more and more powerful. Therefore one task for researchers in this field is how can we let the computer system do more jobs for us and what more benefits we can get from the powerful computer system? One thing worth to try is the fault detection and diagnosis of the being controlled system. The experiment on DH system described in this paper is one of examples.

What we have learnt from this engineering practice? First, it makes a great difference by adopting this fault detection and diagnosis facility. The computer system is no longer a tool for data collection and system regulation. It feels much clever. It can analysis the system performance and find out faults much fast and correct than the operator. Records on the fault detection result is also the great help for maintaining and improving the DH system. Second, in order to put the FDD technology into practices, a standard uniform procedure is very important. We cannot keep to write software or different knowledge bases for every DH systems. Common software and common knowledge base must be developed in general. The difference in system structure should only be described in a configuration file. This is part of the reason for us to develop the S region contract procedure and the FDS method. The former one can used in any networks providing the topologic structure of the network is written in a special file. The later can be widely used in any component and sub-system in the HVAC field if the explanation of CP is estabilished and defined in a configuration file. As the start point of FDD, the measured data is the base for any FDD procedure. Therefore the way for treatment of the measurements is one of the keys for making the FDD be successful. Any sensor does has a error at a certain level. For instant the error in a temperature sensor can be 0.1 °C or 0.2 °C. This must be taken into count. Ingenious treatment of the permissible sensor error can achieve much reliable result. This is the third conclusion.

References

[1] Y. Jiang etc., Leakage and block detection in water network of district heating system. IEA Annex 25 Tokyo meeting, April, 1993, also Proceedings of Chinese national conference on HVAC and Refrigeration, 1994
[2] Y. Jiang, etc., Fault Direction Space method for HVAC on-line fault detection and diagnosis. IEA Annex 25 Boston meeting, April, 1994, also to be published in ASHRAE Transactions, 1995
[3] C. Yang, etc. Sensor fault detection of HVAC system. Proceedings of Pan Pacific Symposium on building and urban environmental conditioning, Nagoya, 1995

SENSOR FAULT DEFECTION OF HVAC SYSTEM
—SYSTEM CONSTRAINTS AND VOTING—

Cheng H. Yang, D. Eng   **Yi Jiang, D. Eng.**

ABSTRACT

In an HVAC system many sensors are used for the purpose of control and measurement. Sensor fault may cause wrong measuring and incorrect control operation. In order to increase the reliability of HVAC system, sensor fault must be detected. In this paper we present a method for detection of sensor fault in HVAC system.

1. INTRODUCTION

In order to make HVAC system work effectively, sensor fault should be detected while the system is running.

The process in an air processing equipment is a physical process, and physical laws should be obeyed. These physical laws play numerical constraints on the process. For example, the air temperature before a cooling coil should be lower than that after the it, and the heat lost by the water on one side of a heat exchanger should be equal to that obtained by the water on the other side of the heat exchanger. This idea was first proposed in reference [1].

An HVAC system consists of many air processing equipment. The processed air flows from one equipment to another. The air state at the outlet of one equipment is the same as that at the inlet of the next equipment. All the constraints both in an equipment and between two equipment should be satisfied in a real system. If the measurement data show any violation of the constraints, sensor fault must occurs.

The constraints are divided into two types: equality constraints and inequality constraints. From equality constraints, whether any of the sensor outputs contained in the equation is wrong can be determined. Inequality constraints can determine fault occurred in the sensor output contained in the inequality, if it is violated. However, the inequality constraints are not violated, it is not necessarily no fault occurs.

Usually, one sensor output is contained in several constraints, and different sensor outputs contained in different constraints. Different sensor fault causes different constraints violated, Therefore, the sensor fault may be determined through checking the violated constraints.

This paper presents a method to construct and infer the system constraints and determine sensor fault by voting. A simple example is given to illustrate the method, and the result is analysed at the end of the paper.

Cheng Han Yang is postal doctoral research fellow, Department of Thermal Engineering, Tsinghua University, and Yi Jiang is professor, Tsinghua University.

2. SYSTEM CONSTRAINTS

In a system equipment are used for air processing, and each of these equipment performs special function. It influences the temperature, humidity or other parameters of the air flow through the equipment. In other words, the air flow through each equipment will be changed in temperature, humidity and so on. For each equipment, its function is determinate. So that the change of air parameter through the equipment is also determined. For example, when air flows through a heater, its temperature will be increased, while the air flow through a cooling coil, its temperature will decreased and the humidity may also be decreased or maintained as before.

Because the function of each equipment in the system is definite and when air flows through an equipment, the change of air parameter is also definite. So, there exist some constraints between the state of the air at different locations of the system. If there is any violation of the constraints in the data measured from sensors, there must be at least one sensor failed. Each constraint links two sensors, and if it is violated the two sensors are possibly failed. To determine which of the two sensors is failed, additional information is needed. Therefor it is necessary to expand all equality and inequality constraints to get more information.

3. MEASUREMENT DATA COMPARING

We have mentioned that in the system physical constraints should be obeyed, but in real application the sensor output values are not accurate, and errors are contained in the measurement data. If an error of sensor output value is great enough that is beyond the allowed range of that sensor, the sensor is considered failed. If the error is within the allowed range, it is considered work properly, but the output value is not equal to the real value the sensor measured. So that, the constraints should be obeyed under the meaning of threshold. That is to say, if a constraint shows that two sensor output values should be the save, and the measurement data of these two values are close enough, the constraint is considered satisfied. For an inequality constraint $t_a > t_b$, if the measurement data shows that $t_a - t_b >$ the threshold, the inequality is considered satisfied.

The threshold is determined by the greatest allowed error of the sensors. If a constraint reflect the relationship of temperature sensor output value, the threshold is determined by the measuring error of the temperature sensor, and if it reflect the relationship of two humidity sensor output value, the threshold is determined by the measuring error of humidity sensor.

4. VIOLATION NUMBER

When a sensor fails, there may be more than one constraints be violated. For each constraint, at least two sensors are related. From violation of a single constraint, it can not be determined which of the two sensors is fault, and both of them are possibly failed. At this time, both of the two sensors are marked with one violation. For each sensor, the violation number is the sum of violated constraints that related to the sensor. The sensor with greatest violation number is usually a really failed one.

5. VIOLATION DIRECTION

If an equality or inequality constraint is violated, the sensors related to it are marked possibly failed with certain direction. For example, if an inequality $t_a > t_b$ is not satisfied, that is the measurement data shows $t_a < t_b$, there are two possibilities: (1) the measurement value t_a is smaller than its real value and (2) the measurement value t_b is greater than its real value. So both of the two sensors are marked with one violation, but with different direction, t_b is possibly greater while t_a is possibly smaller. The violation direction is the direction of the error of a sensor output. There are two possible directions one for positive error, while the other for negative error. Correspondingly, the violation number of a sensor is recorded by two variables each for a possible direction.

6. CONFIRM NUMBER

Contrary to the violate number, there is a confirm number for a sensor, if the sensor is related to a equality constraint. For example, if in the system there is a constraint of $t_b = t_c$, and the measurement data t_b and t_c are close enough, it is considered that this equality is satisfied and both of the two sensors work properly. So that both t_b and t_c are confirmed once. If a sensor appears in more than one equality and is confirmed more than once, its confirm number is great than 1. The confirm number of a sensor is the number of satisfied equalities that related to the sensor.

7. CONSTRAINT EXPANSION

A sensor measurement appeared only in the constraints determined by the equipment before and after the sensor, so that it appears only in a few of constraints. Some times when a sensor fails, there is only one constraint be violated, in this case, there is not enough information to determine which of the two sensors failed. In order to determine the failure of a sensor, at least two violation with same direction are needed. There for it is necessary to expand equalities and inequalities to list out all the implied constraints directly. From the above constraint $t_o > t_a$ and $t_a > t_b$ we can get another constraint that is $t_o > t_b$, and expand all the equality and inequality constraints like this. In this way sensor fault can be detected by voting. For instance, if the measurement $t_b > t_a$ and $t_b > t_o$, it can be determined that t_b is failed, because two constraints are violated by t_b, while one violated by t_o and one by t_a. The violation of t_b is 2, while that of t_o is 1, and that of t_a is also 1. The sensor with greatest violation number is t_b, and it must be a real failure and violation numbers of t_o and t_a are cause by t_b, and so are to be excluded.

Expanding of system constraints is one of the most important steps, from this step, all implied constraints can be derived from the constraints derived from the function of each air processing equipment, and list out all the implied relationship directly. In an unexpanded constraint system, a sensor output value usually appears in a few constraints and is usually restricted by only one value in each direction, and if the relation between the sensor output and the restricting value do not satisfy the constraint, it is impossible to determine which of the two values is wrong. In an expanded constraint system, all implied constraints are listed out directly, the greatest value and the lowest values are restricted by equalities, and each of the sensor output value is restricted at least by two values in each direction. Therefor if any of the sensor output value is far away from its real value, there must be at least two constraints be violated by that value and it is easy to be detected by checking the violation number. It is easy for computer to check if there is any of the constraints be violated and count the violation of each sensor in each direction.

8. FAULT DETECTION AND FALSE ALARM EXCLUSION

In the step of fault detection, it is assumed the violations, if there is any, are caused fewest wrong values, or in other words, to make the all constraints satisfied, as few as possible values should be corrected. Each equality and inequality constraint is related to two sensors, If it is violated, two sensors linked to it may be marked possibly fail, but actually the violation is usually caused by only one sensor fault and the other is a false alarm. Once the real failed sensor be detected, the false alarms should be to be excluded.

To detect a sensor fault, violation number of the sensor is checked. The violation number of a sensor is the sum of constraint which are violated by the sensor measurement data, so that it reflect the possibility of the sensor failure. The sensor with greatest violation number of a certain direction is most possibly the failed one. It is determined fail and the violation of the constraints is caused by the fault of the failed sensor and the violation number of other related sensor is caused by the failed sensor and should be excluded.

In our example t_b is detected higher and two constraints violated by it is $t_o > t_b$ and $t_a > t_b$. these two constraints lead the two other sensors t_o and t_a. When t_b is detected fail, the alarm of

t_o and t_a are considered false alarm. So that the violation number of both t_o and t_a are reduced by 1, to exclude the false alarm caused by t_b.

The violation number for each sensor is recorded by two variables, one for higher value the other for lower value. In the above example, higher violation number for t_b is 2, while the lower number is 0, the lower number for t_o and t_a is 1, while the higher number is 0. When t_b is detected higher, so that the lower number of t_o and t_a is to be reduced by 1. So that false alarms of t_o and t_a are excluded. Sometimes, sensor fault can not be detected just by inequalities, because of only one inequality constraint is violated by a pair of sensor output. In this case, equality is needed to determine which sensor is failed or not.

9. EXAMPLE

In a system of one handle unit shown in Fig. 1, it consists of a cooling coil, a heater and a steam humidifier. The constraints can be listed out from each step of air processing.

$max(t_o, t_r) \geq t_m \geq min(t_o, t_r)$

$max(d_o, d_r) \geq d_m \geq min(d_o, d_r)$

$t_m > t_1$ (if the cooling coil work)

$t_m = t_1$ (if the cooling coil not work)

$d_m > d_1$ (if the cooling coil work)

$d_m = d_1$ (if the cooling coil not work)

$t_a > t_1$ (if the cooling coil work)

$t_b > t_a$ (if the cooling coil work)

$t_2 > t_1$ (if heater works)

$t_2 = t_1$ (if heater not work)

$d_2 = d_1$

$t_c > t_2$ (if heater works)

$t_c > t_d$ (if heater works)

$d_i > d_2$ (if humidifier works)

$d_i = d_2$ (if it not work)

$t_i = t_2$

If any of the constraint in the system is violated, there must be a sensor failed. For example, if the measurement data shows that $d_i < d_2$, there must be one sensor failed in d_i and d_2. But from this violation, it cannot be determined which of the two sensors failed.

Expand all constraints derived from each single equipment, additional constraints can be derived. For example, from $d_1 = d_2$ and $d_i \geq d_2$, a new constraint $d_i \geq d_1$ can be derived. If the above measurement data shows $d_i < d_2$, $d_i > d_1$ and $d_1 < d_2$, it can be determined that d_2 is failed, because there are two constraints about d_2 are violated with same direction, while for each of d_i and d_1, there is only one constraint violated. When the failure of d_2 is determined, the violation of constraints about d_1 and d_i are considered false alarm and should be excluded.

10. ANALYSES

A HVAC system is a dynamic system, and the state of the air processed is changed continuously. Sometimes a measurement value is restricted in wide range while in some other time, it is restricted in narrow range. When a sensor output is restricted in a narrow range, the error of the sensor output value can be easily detected. In the system of above example and in the season of summer, at noon the outside temperature t_o is much higher than the temperature in the room t_r. At this time, the range for temperature of mixed air is wide and a small error of the sensor output t_m is difficult to detect. But at night the temperature outside is close to that in the room, namely to is close enough to t_r, at this time the range for t_m is restricted in a narrow range

and a small error of t_m is easy to be detected. But if the state of system does not change. some errors may not be detected.

For a system of more than one subsystems the outside temperature t_o should be the same. and t_a,and t_c for all subsystems should be close enough if they use same heat and cool sources. and the temperature in the room t_r should be close to each other. These are strong constraints for a large system and that makes each sensor output value appears in at least two constraints. that is necessary for voting method.

11. APPLICATION

This method has been used in a system for rubber plant with six subsystems. The result shows that is effective for detection of sensor faults.

12. CONCLUSION

Detection of faulted sensor of a HVAC system is an important work and detection of slowly floating error of a sensor output is one of the most difficult tasks. Because the slowly floating error does not influence the dynamic characteristic of the system. the fault cannot be detected from the dynamic characteristic of the sensor output itself. It can only be detected by the relationships of the air state at different locations of the system.

This paper presents a new method for detecting of the slowly floating error of a sensor output by taking use of the relationships of air state at different location of the system and comparing the output values of all sensors.

Simulation and application shows that this method is effective especially when used in a large system with many subsystem.

REFERENCE

[1] Yi Jiang, Chenghan Yang. "Detection of Faulted Sensor in HVAC System". IEA Annex 25 Internal circle. 1994
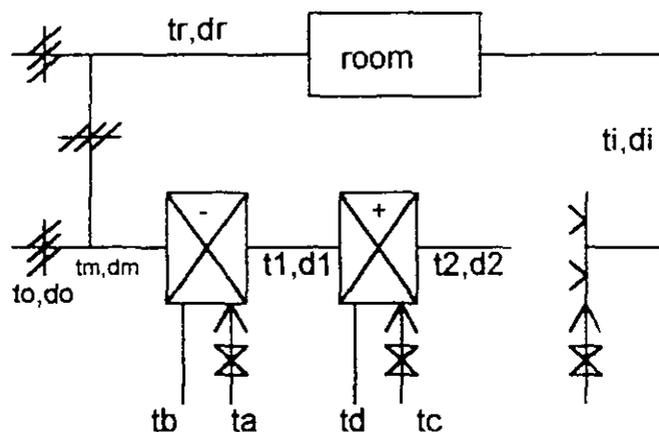
Fig. 1 An Example of AC Processing System

Fault Detection and Optimization for a heating system

Knabe, G.; Kremonke, A.; Wagner, H.-J.
Dresden University of Technology
Thermodynamics and Technical Installation of Buildings
Germany

Abstract

A sensitivity analysis on heating systems is carried out. It reveals the dominating impact of an inlet temperature fault. The heat consumption analysis proofs that there is a linear dependence between daily means of heat consumption and daily means of outdoor temperature. An operating failure can be detected by comparing the heat consumption with an outdoor dependent threshold. Room temperatures are mainly determined by occupants behaviour. Each room has a different temperature evolution. Furthermore the room temperature is dependent on the control system in use. Parameters to optimize are heating curve and night set back recovery.

1. System description and Objective

A single pipe and a two pipe system (fig. 1) have been investigated by employees of the TU Dresden for two years. Both systems are installed to run a 5-storey residential building in Dresden. The effectivness of refurbishing strategies and a comparison of heat allocators are of major interest. A direct link connects the remote power station with the substations of the residential buildings which deliver warm water to heat 20 flats.
Central outdoor dependent inlet temperature of the substation is accomplished by entraining return water. Heat demand of the radiators is controlled locally in 10 flats by thermostatic valves (TRV) and single room controllers (ERS) respectively. A on-off control is used to adjust heating-power as part of the single room controller.
During the investigation measurements are taken for monitoring systems behaviour as well for fault detection [1], [2], [3].

2. Sensitivity analysis for faults in heating systems under stationary conditions

The following equation shows the simplified heat transfer scheme for the system heating-room-building:

$$\dot{Q} = \dot{m}c(\vartheta_V - \vartheta_R) = kA\,\Delta\vartheta m = C(\vartheta_i - \vartheta_a) + \dot{Q}_S \tag{1}$$

| cooling of heat medium | heat flow via radiators | heat balance of building |

where

$$k = k_N \left(\frac{\Delta \vartheta_m}{\Delta \vartheta_{m,N}} \right)^m \left(\frac{\dot{m}}{\dot{m}_N} \right)^n$$

$$\Delta \vartheta_m = \frac{\vartheta_V - \vartheta_R}{\ln \dfrac{\vartheta_V - \vartheta_i}{\vartheta_R - \vartheta_i}} \, ,$$

where

A	-	heating surface
C	-	specific heat loss of the room (insulation and air exchange)
$\dot{Q}$	-	heat output of the radiator
$\dot{Q}_S$	-	gains (solar heat gains, inner heat gains)
k	-	U-value
$\dot{m}c$	-	heat capacity flow
ϑ	-	temperature

indicees:

N	-	nominal value
R	-	outlet
V	-	inlet
a	-	outdoor
i	-	room

When doing fault detection the following influencing quantities are considered:

- inlet temperature ϑ_V
- heating surface A
- mass flow $\dot{m}$
- insulation and coefficient for ventilation C.

When performing the sensitivity analysis these quantities are related to reference values, i.e. their values in "faultless" and undisturbed conditions. Reference values will be marked by '*'.
The following characteristics describing the faults are introduced:

inlet temperature fault $\qquad C_1 = \dfrac{\vartheta_V - \vartheta_a}{\vartheta_V^* - \vartheta_a}$ resp. $C_1' = \vartheta_V - \vartheta_V^*$ $\qquad\qquad$ (2)

heating surface fault $\qquad C_2 = \dfrac{A}{A^*}$ $\qquad\qquad\qquad\qquad\qquad$ (3)

mass flow fault $\qquad C_3 = \dfrac{\dot{m}}{\dot{m}^*}$ $\qquad\qquad\qquad\qquad\qquad$ (4)

specific heat loss
$$C_4 = \frac{C}{C^*} \tag{5}$$

The influence of the disturbance variable is expressed by the ratio of free heat gains and heat losses

$$C_5 = \frac{\dot{Q}_S}{C^*(\vartheta_i^* - \vartheta_a)} \tag{6}$$

To determine the outdoor temperature dependent reference value ϑ_V^* the heating balance (1) is setup for the faultless and undisturbed ($\dot{Q}_S = 0$) case:

$$\dot{Q}^* = \dot{m}^* c(\vartheta_V^* - \vartheta_R^*) = k_N A^* \cdot \left(\frac{\Delta \vartheta_m^*}{\Delta \vartheta_{m,N}}\right)^m \left(\frac{\dot{m}^*}{\dot{m}_N}\right)^n \Delta \vartheta_m^* = C^*(\vartheta_i^* - \vartheta_a)$$

(7)

For the heating systems considered here $\dot{m}^* = \dot{m}_N$. Their design is expected to supply $\dot{m}$, $\Delta \vartheta_{m,N}$, ϑ_i^*, and for $\vartheta_a = \vartheta_{a,N}$, $\vartheta_V^*(\vartheta_{a,N}) = \vartheta_{V,N}$ and $\vartheta_R^*(\vartheta_{a,N}) = \vartheta_{R,N}$. Given these date it is possible to determine ϑ_V^* from eqs. (7) for any outdoor temperature $\vartheta_a < \vartheta_i^*$.
Equations (1), (2), (3), (4), (5), (6) and (7) lead to

$$\frac{\dot{Q}}{\dot{Q}^*} = C_3 \frac{C_1(\vartheta_V^* - \vartheta_a) - (\vartheta_R - \vartheta_a)}{\vartheta_V^* - \vartheta_R^*} \approx C_2 \left(\frac{\Delta \vartheta_m}{\Delta \vartheta_m^*}\right)^{1+m} \cdot C_3^n = C_4 \frac{\vartheta_i - \vartheta_a}{\vartheta_i^* - \vartheta_a} + C_5 \tag{8}$$

The deviation of room temperature in response to the fault extent is esteemed a value to determine fault-sensitivity. The room temperature is dependent on the fault and the outdoor temperature. The most essential deviation is caused by an inlet temperature fault (see figure 2).

The example below illustrates this statement.

Design: $\vartheta_{V,N} = 90\ °C$; $\vartheta_{R,N} = 70\ °C$; $\vartheta_i^* = 20\ °C$; $\vartheta_{a,N} = -15\ °C$; $m = 0,33$; $n = 0$

With equation (7) and (8) the room temperature and the relation $\dfrac{\dot{Q}}{\dot{Q}^*}$ can be determined by iteration (see figure 2) for any outdoor temperature.

For an outdoor temperature of 0 °C depending on the fault mentioned the room temperature becomes ($C_i = 1$ exept for those in consideration and $C_5 = 0$):

1. inlet temperature $C_1^J = 10\ K$: $\vartheta_i = 23,5\ °C$
2. heating surface $C_2 = 1,2$: $\vartheta_i = 21,8\ °C$
3. mass flow $C_3 = 1,2$: $\vartheta_i = 20,8\ °C$
4. insulation, vent $C_4 = 0,8$: $\vartheta_i = 22,8\ °C$

With the occurence of the disturbance C_5 the temperature becomes

$$C_5 = -0,8 : \quad \vartheta_i = 24,8 \ °C$$

3. Fault detection based on the buildings heat-consumption-plot

In the heating system considered here errors may occur in the substations management and locally in the tenants control systems. In the substation the following errors are likely to appear:

- pump breakdown
- breakdown of BEMS-components
- defective valve (leakages, contamination, blocking....)

Pump breakdown can be easily detected by temperature and pressure measurements. To determine other failures a model or a graphic scheme is needed.
Figure 3 shows the hourly heat consumption in dependence of the outdoor temperature.
The system is run by an outdoor dependent inlet temperature. While still following the outdoor temperature the inlet temperature is reduced by 20 K during night hours.
Two regression curves can be perceived for daytime and night setback respectively. The deviation of single values are considerable when hourly means are taken. At lower outdoor temperatures fewer measurements could be taken. When scanning daily mean values of heat consumption a deviation decisively smaller can be expected (refer to figure 4). Therefore the monitoring of daily heat consumption is useful for fault detection. A diagnostic routine is to be launched when measurements are recorded violating a given error threshold.

4. Obstacles to fault analysis based on recordings of room temperatures

The heating of rooms aims for comfortable indoor air conditions during the occupancy time. The energy consumption has to be a minimum.
A fault analysis drawn from measurements of the room temperature evolution is difficult as one has to pay attention to the occupants behaviour. Room temperature measurements were carried out in 100 rooms of a single pipe and a two pipe system respectively. Each room has its characteristic room temperature plot related to its occupants behaviour. It depends on the installed local controller too. Figure 5 shows the mean room temperature evolution of the single pipe system covering all tenants. Furthermore the related daily mean value for living room, childrens room, bathroom, bedroom and kitchen regarding to ERS and TRV is exposed. There are only small variations between working days (Mo-Fr) and weekends (Sa,Su). While in the living rooms and bathrooms the highest room temperature were recorded the room temperatures of bedrooms are the lowest.
The daily mean room temperatures related to the ERS - system are higher compared with those of the TRV system except of the bathroom. The biggest deviations concerning the local controllers influence on room temperatures can be detected when monitoring kitchen room temperatures. As programming functions are available with the ERS system kitchen room temperatures vary over a wider range. For the ERS system the maximum of the variation range

over a day is 2 K when considering living rooms. This value is smalller for TRV controlled systems (all examples are related to the single pipe system). In fact each tenant proofs a different occupants behaviour. Figure 6 shows the daily mean evolution over a year of the room temperature for a tenant with a high scale energy consumption (A) and a low scale energy consumption (B) respectively. The curve related to tenant B indicates only one occupancy period *- that is in the evening. Tenant A`s plot reveals two peaks of room temperatures- that is in the* morning and in the evening. In the bedroom temperatures for both tenants a noticable cooling down process can be percieved.

Figure 7 shows the frequency distribution of room temperatures related to means. This figure indicates a normal distribution for the room temperature. Data of living room temperature for the tenants A and B in the time period of 7 a.m. to 9 a.m. were used. This plot is based on hourly means of room temperatures covering all days of the heating period. Means and daily evolutions of room temperatures are exposed in figure 8. Measurements taken during the occupancy time within the heating period were utilized. Furthermore the maximum and minimum of hourly means of the room temperatures having occured within the related time range are shown. The plot is completed by thresholds based on probalility ranges of 80 % and 90 %.

A fault detection based on room temperatures can be performed in the following way: Over a certain time period room temperatures have to be compared with a plot given in figure 8. If occured values does not fit within the boundaries a local malfunction may be expected.

With more rooms indicating a similiar trend a central error is taking place and an error diagnosis has to be launched.

An optimization of the systems operation has to consider the occupants behaviour.

The determination of categories of tenants is a necessity before forming optimization strategies and doing fault detection.

A suggestion for such a category-related-investigation could be:

Category   I: ϑ_i = 18 °C   (ϑ_i < 19,5 °C)

Category II: ϑ_i = 21 °C   (19,5 ≤ ϑ_i ≤ 22,5 °C)

Category III: ϑ_i = 24 °C   (22,5 °C < ϑ_i)

Analysis to this continue.

5.   Optimization of the heating curve

In the substation the inlet temperature is controlled dependent on the outdoor temperature. The determination of the inlet temperature is based on a linear equation.

When judging the current heating curve one has to pay attention to occupants behaviour and local control units. When discussing the ERS-system one has to monitor the proportion between heating-up and cool-down time. A proportion of 1:1 is esteemed an appropriate value. Deviations from that relation may be caused by power to high or low respectively.

Figure 9 exposes a heating curve too high powering the single pipe system. An occupants behaviour related to category III (refer to figure 10,11) instead of category II corresponds with the desired proportion. Instability may often be perceived at TRV-controlled radiators following

a power too high (refer to figure 12). As catagory III is demanded for the Kitchen's room temperature, radiator surface matching fourfold air change per hour was installed. This high air change scarcely occurs. Due to low valve authority instability is more likeley to be recorded at two-pipe-systems. At TRV controlled radiators the outlet temperature evoluation signifies whether a power supply too high or low is applied. Figure 13 shows that because off high supply there is a eminent flow restriction bringing outlet temperature nearer to the room temperature. The heating curve has to match the room with highest demand on the inlet temperature. For a single pipe system with an upper distribution scheme it would be the uppermost radiator.

6. Optimization of heating up time

Figure 14 shows the inlet and outlet temperature evolution as well as the heat flow over a one week period in the single pipe system.

The plot allows the conclusion that heating up time does not match with tenants needs. Apparently during heating up time occupants demand is less than heating supply. That inevitably leads to a fast rise of the outlet temperature.

The provided outlet temperature limiter initiates a lowering of the inlet temperature. Therefore the transfered heat flow is reduced (close to zero at 5.45 a.m.). An unnecessarily high heating up peak is recorded. The time for night set back recovery is chosen too early.
In the mean time investigations to optimize the heating up time are launched.

7. Summary

This report introduces a useful error detection for buildings checking the experimentally determined heat consumption curve.
An error analysis based on room temperatures has to regard the wide range of occupant behaviour.
The statistic values of room temperature covering occupant behaviour depend on the local controller used. Following issues are of major interest:

- adjusting the heating curve
- optimized heating up time.

Bibliography

[1] Knabe, G.; Kremonke, A.; Wagner,H.-J.: Building optimization and fault defection for single pipe and two pipe heating systems. IEA Annex 25 Stuttgart meeting September 1994.

[2] Knabe, G.; Kremonke, A.; Wagner,H.-J.: Fault detection for a heating system. IEA Annex 25 Oxford meeting 1995.

[3] Visier, J.C.: Data Base on the typical Troubles and Faults in HVA C Processes. Heating Systems. IEA Annex 25, CSTB Liege 1992.
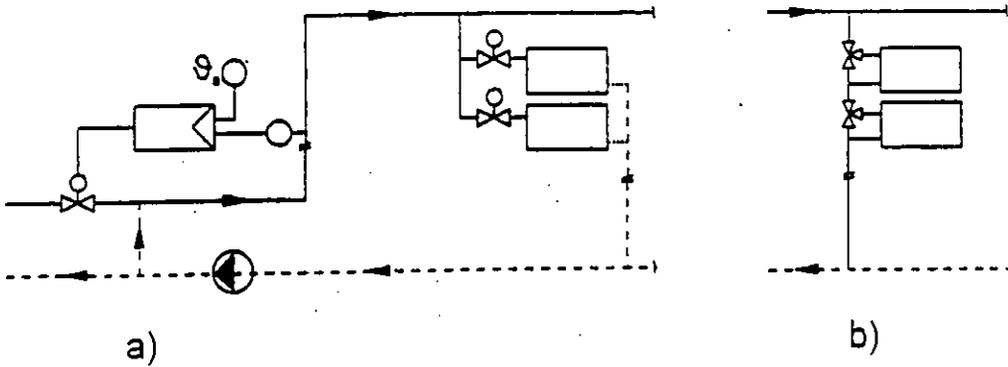
fig. 1:   scheme heating system
   a)   two pipe system
   b)   single pipe system
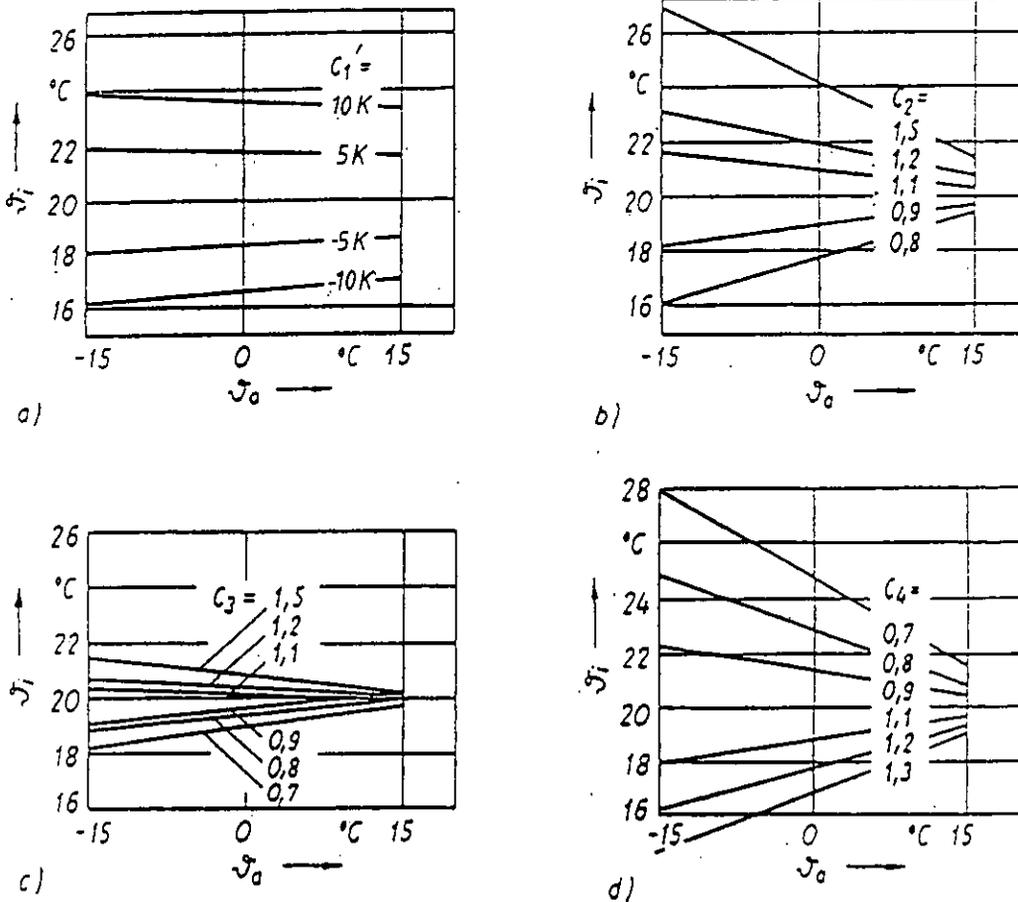
fig. 2:   sensitivity analysis for a heating system
          (90/70/20/-15; m = 0,33; n = 0; C_5 = 0)

a) inlet temperature fault C'_1 (C_2 = C_3 = C_4 = 1)
b) surface of radiator C_2 (C'_1 = 0, C_3 = C_4 = 1)
c) mass flow C_3 (C'_1 = 0, C_2 = C_4 = 1)
d) heat loss of the building C_4 (C'_1 = 0, C_2 = C_3 = 1)

64
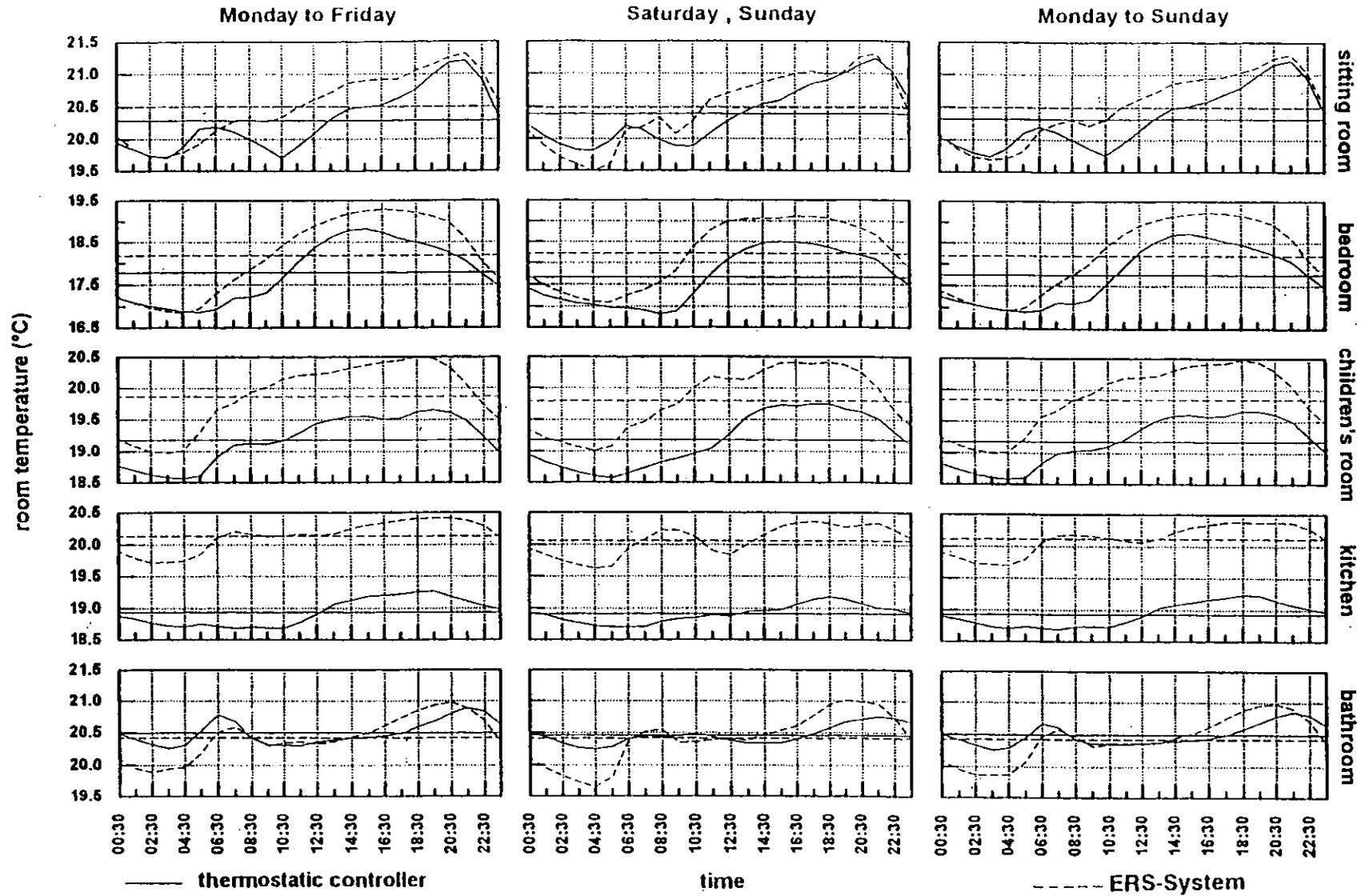
fig. 5: daily mean evoluation of room temperature (averages over 10 flats and over one heating period)

living room
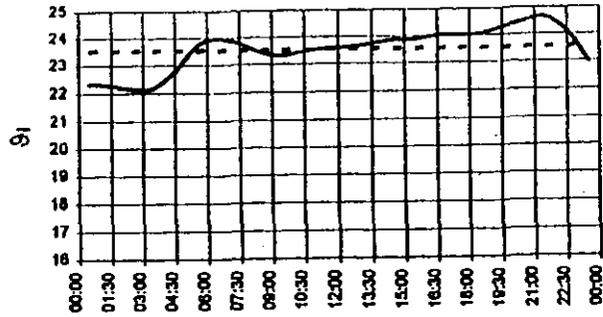
bathroom
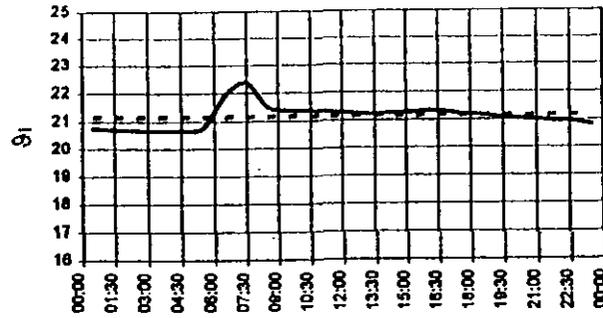
children's room

kitchen

bedroom

a)
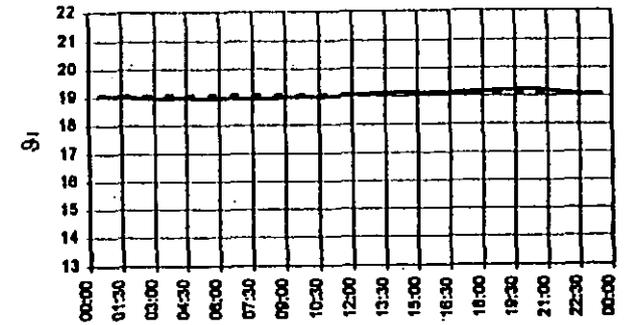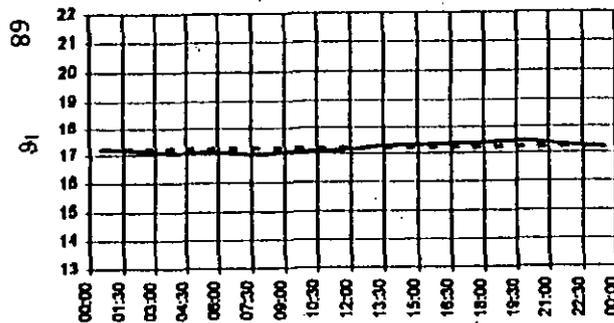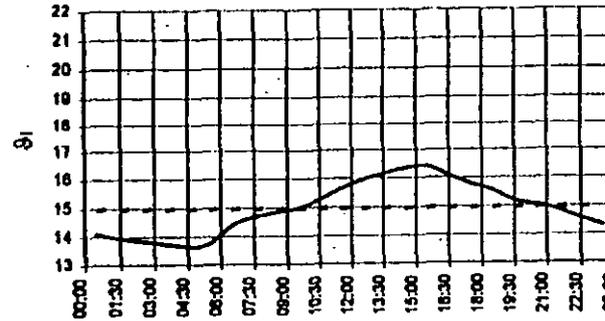
fig. 6: daily mean evoluation of the room temperature over a one year period
for two selected tenants
       a) tenant characterized by large scale heat consumption
       b) tenant characterized by small scale heat consumption
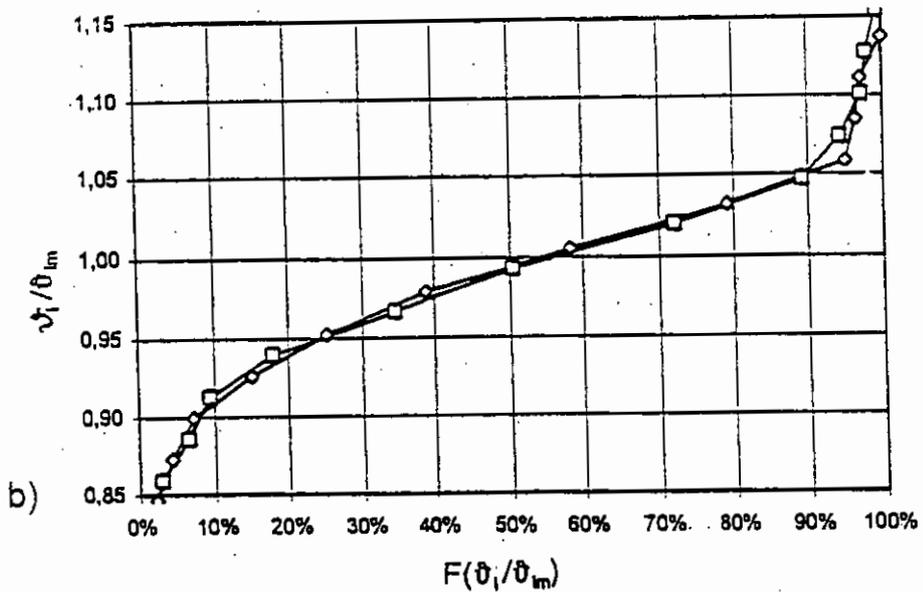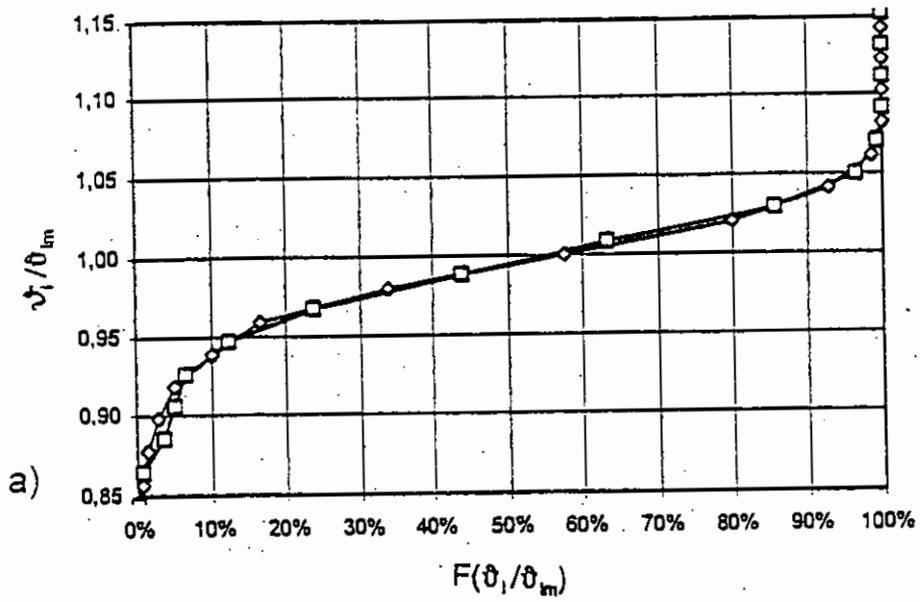
fig. 7: frequency distribution of room temperatures (ϑ_i is related to the mean of the considered time range : 7 a.m. - 8 a.m. and 8 a.m. - 9 a.m.)
   a) tenant characterized by large scale heat consumption
   b) tenant characterized by small scale heat consumption
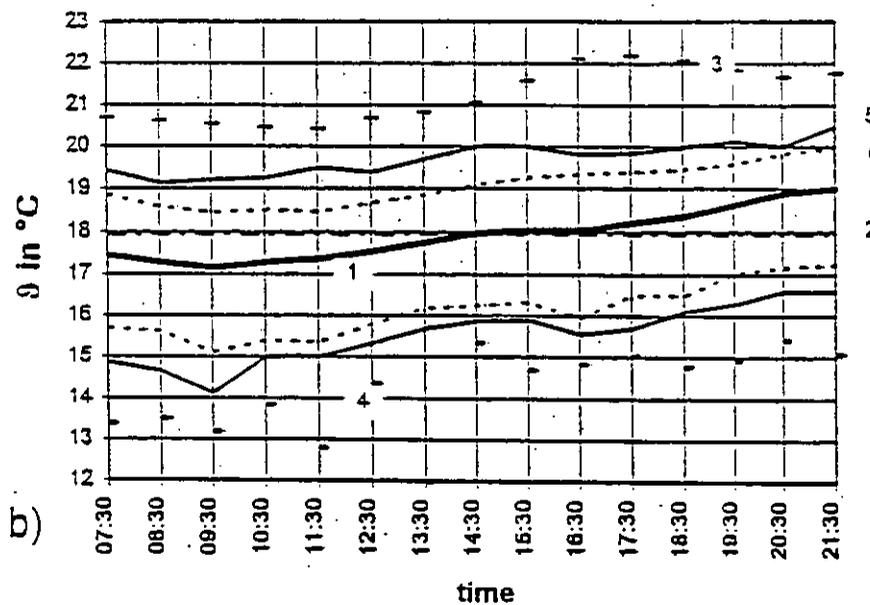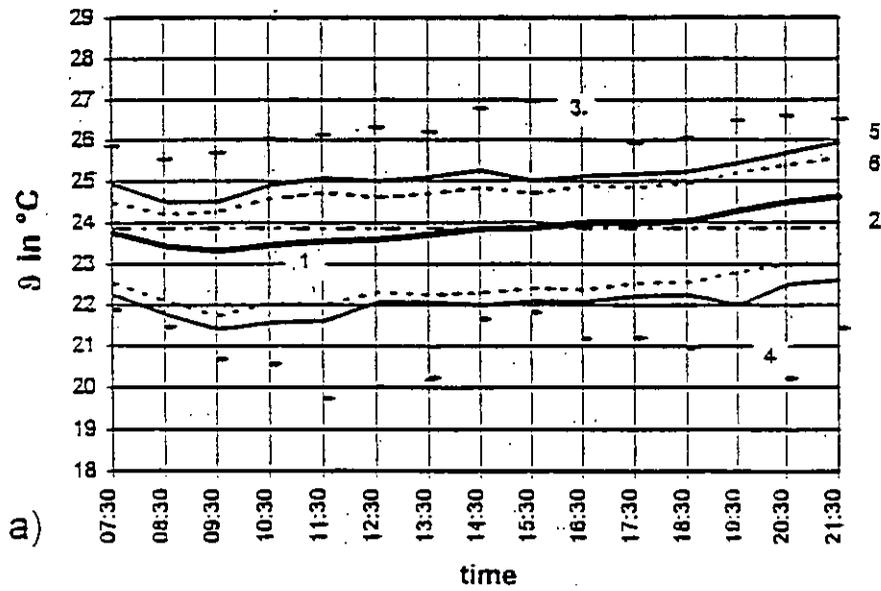
fig. 8: daily mean evoluation of the room temperature during occupancy period
for living rooms
a) tenant characterized by large scale heat consumption
b) tenant characterized by small scale heat consumption

1 - mean daily evolution
2 - mean
3 - maximum
4 - minimum
5 - probability range = 90%
6 - probability range = 80%
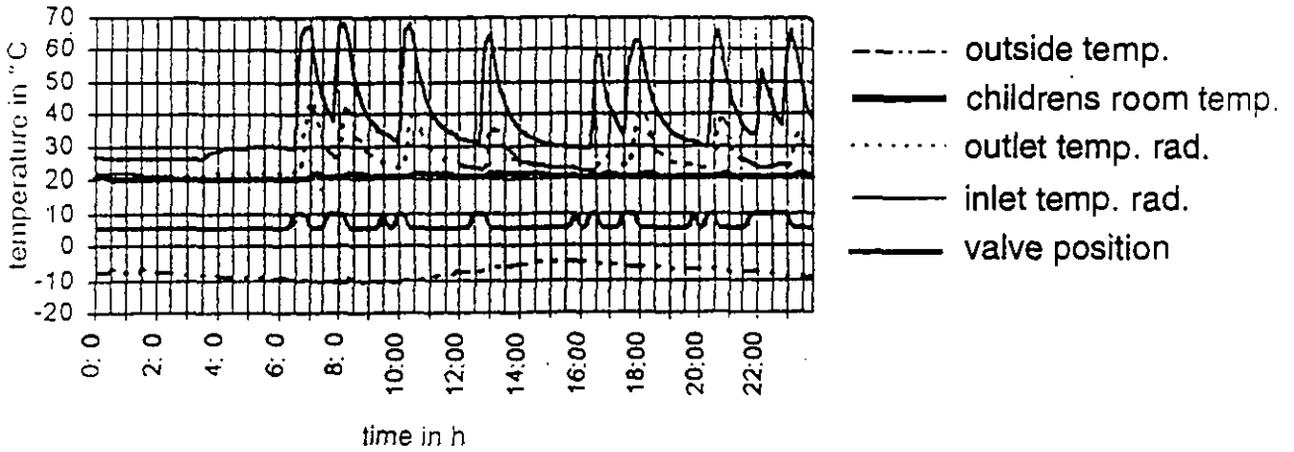
fig. 9: Temperature behaviour for the childrens room 4
(low outside temperature; single pipe heating
system; ERS)
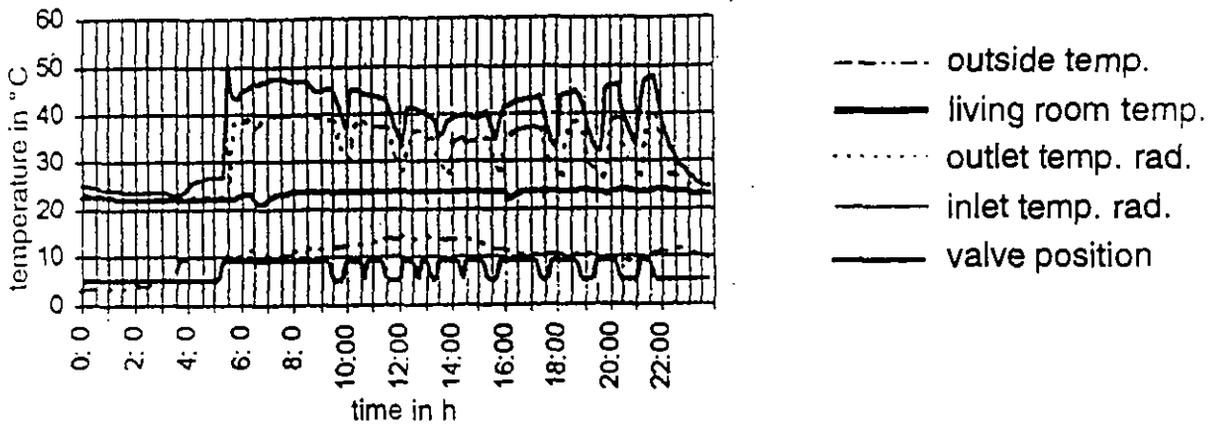
fig. 10: Temperature behaviour for the living room 10
(high outside temperature; single pipe heating
system; ERS)
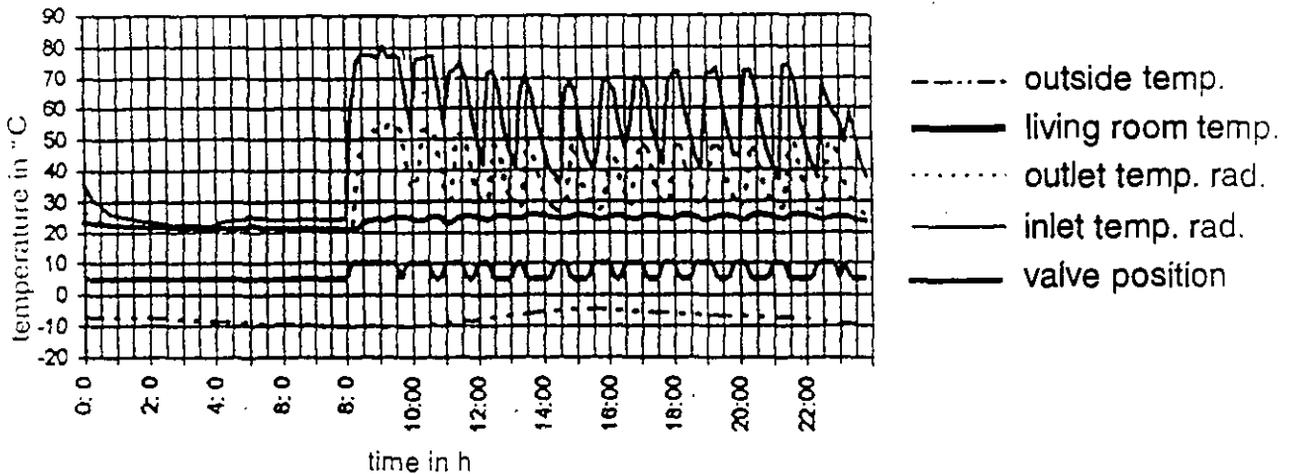
fig.11: Temperature behaviour for the living room 20
(low outside temperature; two pipe heating
system; ERS)
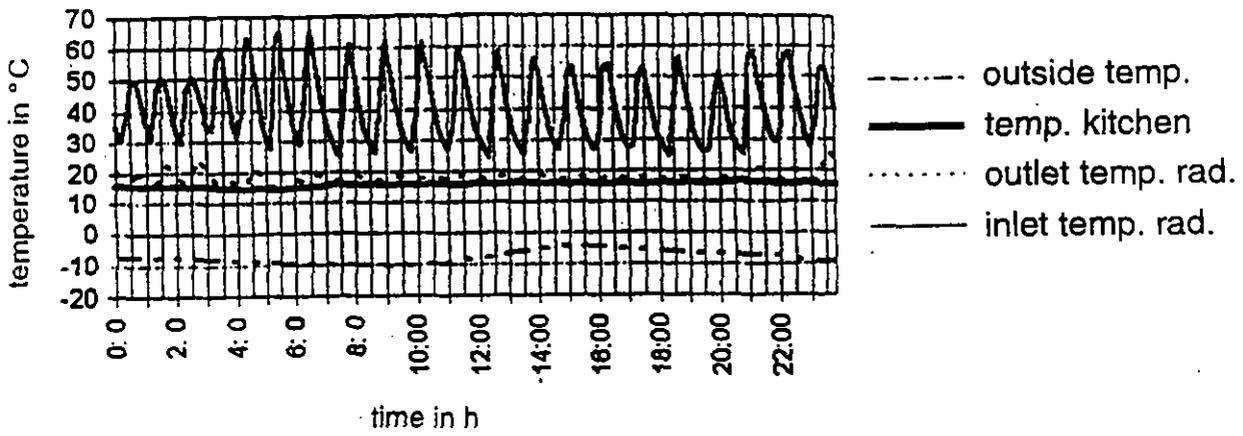
fig. 12: Temperature behaviour for the kitchen 1
(low outside temperature; two pipe heating
system; TRV)
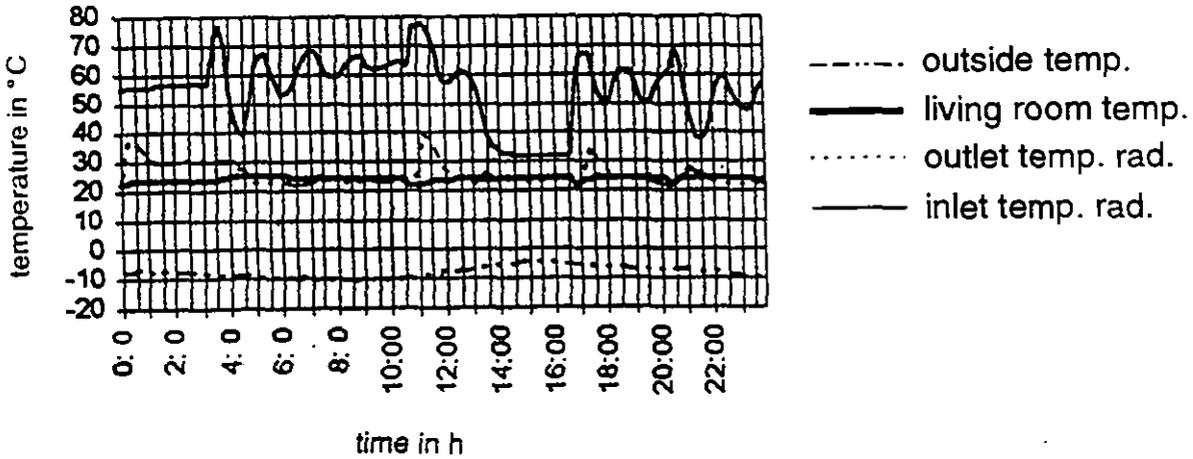
fig.13: Temperature behaviour for the living room 9
(low outside temperature; single pipe heating
system; TRV)
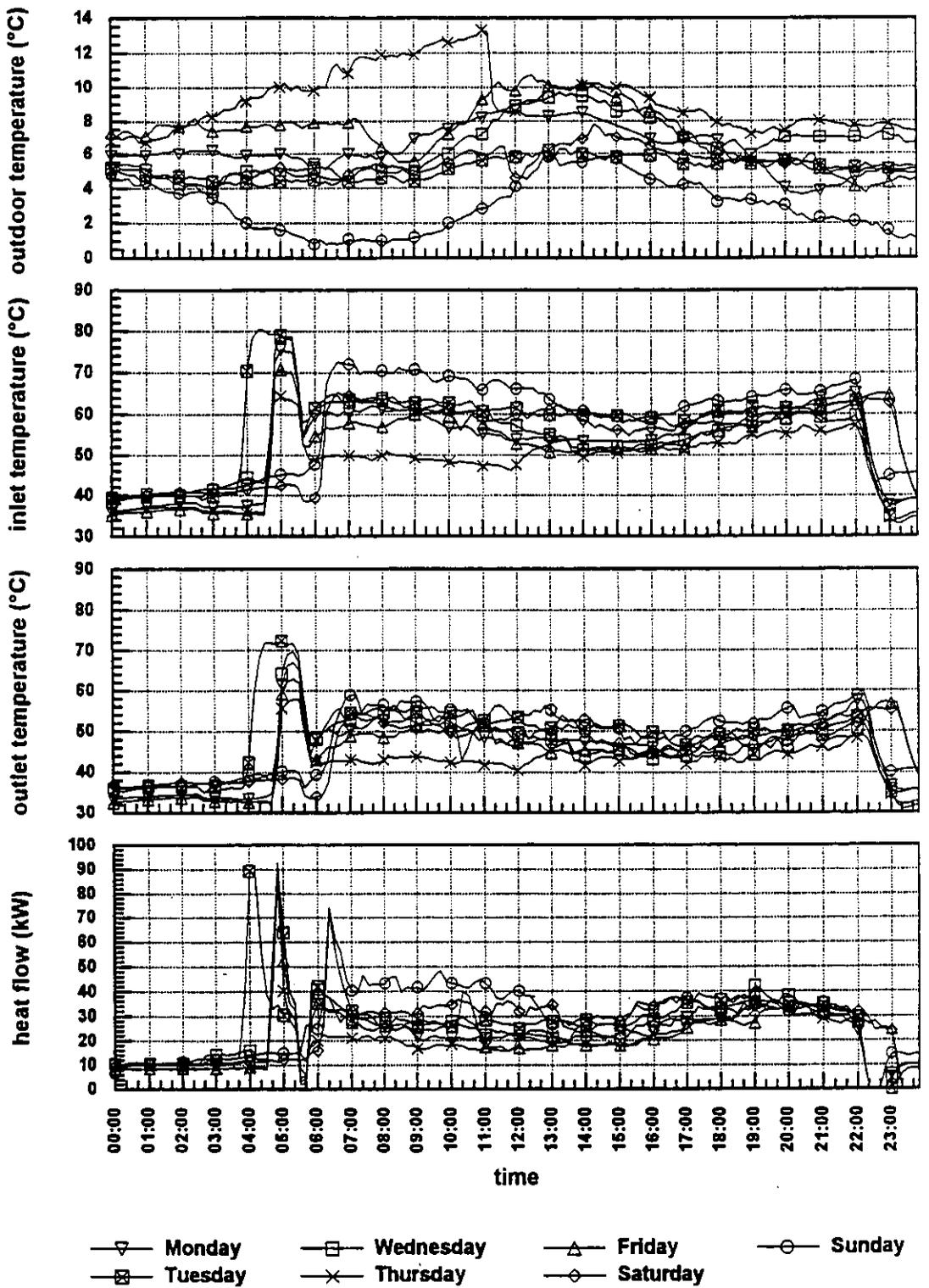
fig. 14: outdoor-, outlet-, inlet temperature and heat consumption
during 10.01. - 16.01.94

PERFORMANCE MONITORING, FAULT DETECTION AND DIAGNOSIS OF RECIPROCATING CHILLER

Meli Stylianou and Darius Nikanpour

EDRL-CANMET, 1615 Lionel-Boulet Blvd, Varennes, Quebec, Canada

ABSTRACT

This paper presents a methodology which uses a combination of techniques: thermodynamic modelling, pattern recognition and expert knowledge to determine the "health" of a reciprocating chiller and to diagnose selected faults.

The system is composed of three modules. The first one deals with the detection of faults that are more discernible when the chiller is off such as sensor drift. The second module detects faults during start-up and deals with those related to refrigerant flow characteristics, which are generally more apparent during the transient period. Finally, the third module detects deterioration in performance followed by diagnosis when the unit is operating in a steady-state condition.

The approach has been experimentally tested on one laboratory unit and results presented. It is emphasized that further data is required to establish the repeatability of the emerging patterns and validate the applicability of the approach to the reciprocating chillers in general.

INTRODUCTION

Vapour compression refrigeration systems constitute the largest portion of commercial and industrial refrigeration capacity, accounting for an important portion of energy consumption in these sectors. In large office buildings for example, it is estimated that 10% to 25% of the total electricity consumption can be attributed to cooling systems alone (Huang, Akbari, Reiner and Ritschard, 1991). Moreover, these percentages can be significantly higher if a cooling system is operating at low performance levels due to the presence of faults (Herzog and LaVine, 1992).

Fault detection systems for commercial chillers currently focus primarily on preventing mechanical failures, which is generally achieved through the use of switches that cut out the unit when temperatures and/or pressures exceed preset limits. Generally, these systems do not provide direct information as to the "health" of the chiller prior to its shut-down, resulting in unexpected periods of unavailability and to extended periods during which the unit operates under abnormal conditions. However, as chillers become better instrumented, advanced low-cost fault detection and diagnosis modules become increasingly more attractive.

Fault detection and diagnosis systems (FDDS), developed for application to vapour compression units, generally employ one of the following two procedures:
1. Estimation and
2. Pattern Recognition.

Estimation-based procedures are used in systems which employ physical or statistical models to estimate relevant process variables. Figure 1(a) shows the structure of this type of FDDS. In general, process system models are used to estimate the values of the process variable. These

values are subsequently compared with measured process variables. The resulting differences (or so called innovations) are then supplied to a separate module where the diagnosis is made.

(a) Fault detection and diagnosis through estimation

(b) Fault detection and diagnosis through pattern recognition
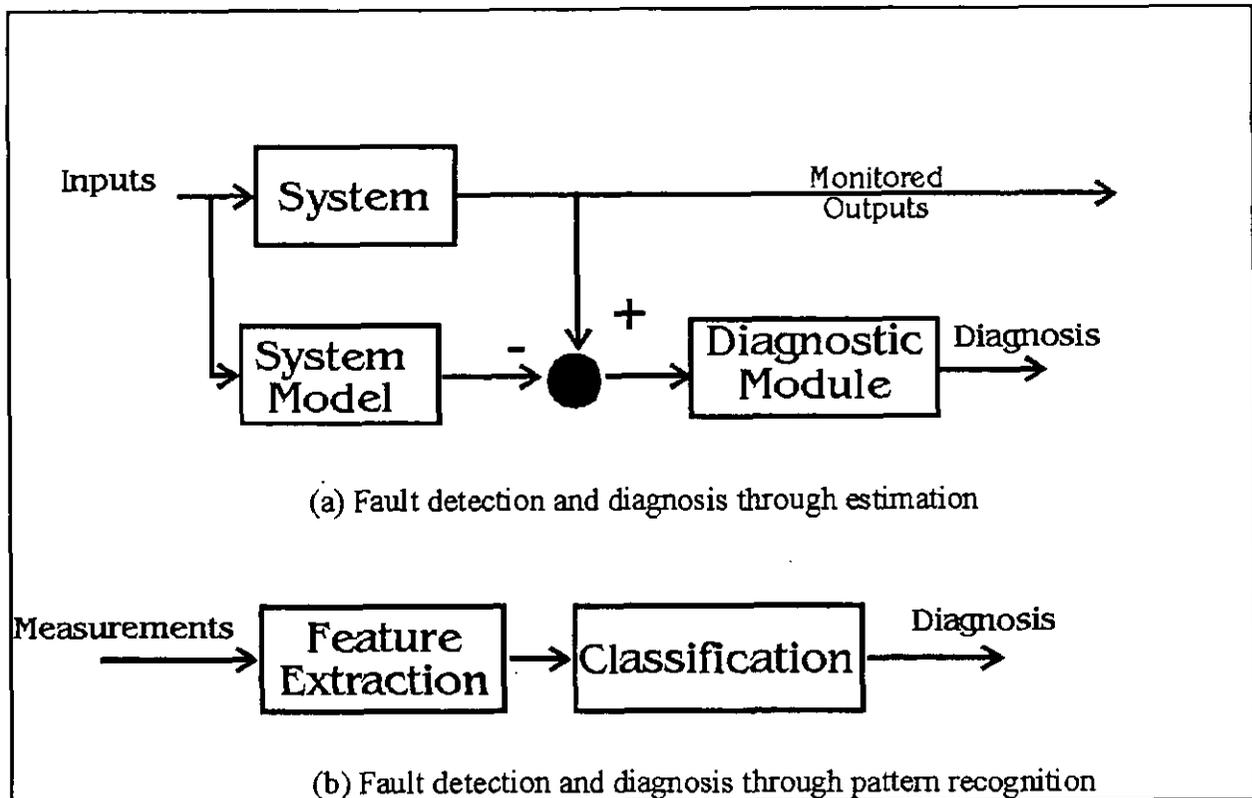
Figure 1   Procedures for fault detection and diagnosis

Figure 1(b) shows the structure of pattern recognition-based FDDS. In this type of scheme, features particular to the process in question are extracted from measured data. These features are then matched to pre-determined classes (normal operation, fault 1, fault 2, ... etc.) which have been generated a priori from training data.

Researchers have used FDDS employing one of the above two procedures and in some cases estimation and pattern recognition procedures have been used together (Shoureshi and Wagner, 1992; Sami, Zhou and Tulej 1993; Grimmelius, Klein Woud and Been, 1994). For example, Grimmelius, Klein Woud and Been (1994) developed a system that used a non-linear regression model of a chiller to estimate the process variables. These estimations were subsequently used to generate innovations which were classified by a fuzzy pattern classification routine. The system was developed for a water chiller equipped with a thermal expansion valve and was tested on common faults of this type of machine.

Grimmelius, Klein Woud and Been (1994) identified three areas for further development:

1.   Generic chiller models requiring no training data;
2.   Fault detection capability during transient conditions; and
3.   Methodologies identifying a wider range of faulty behaviour.

The purpose of the work presented here is the development of a methodology for a performance monitoring, fault detection, and diagnosis system. This methodology is developed for reciprocating chillers, and integrates techniques derived from physical modelling, artificial intelligence and pattern recognition. This paper focuses on the detection and diagnosis of faults for a reciprocating chiller during transient and steady-state conditions. An innovative approach is the utilisation of the transient characteristics during start-up to provide useful information on the "health" of the laboratory test unit.

TEST UNIT
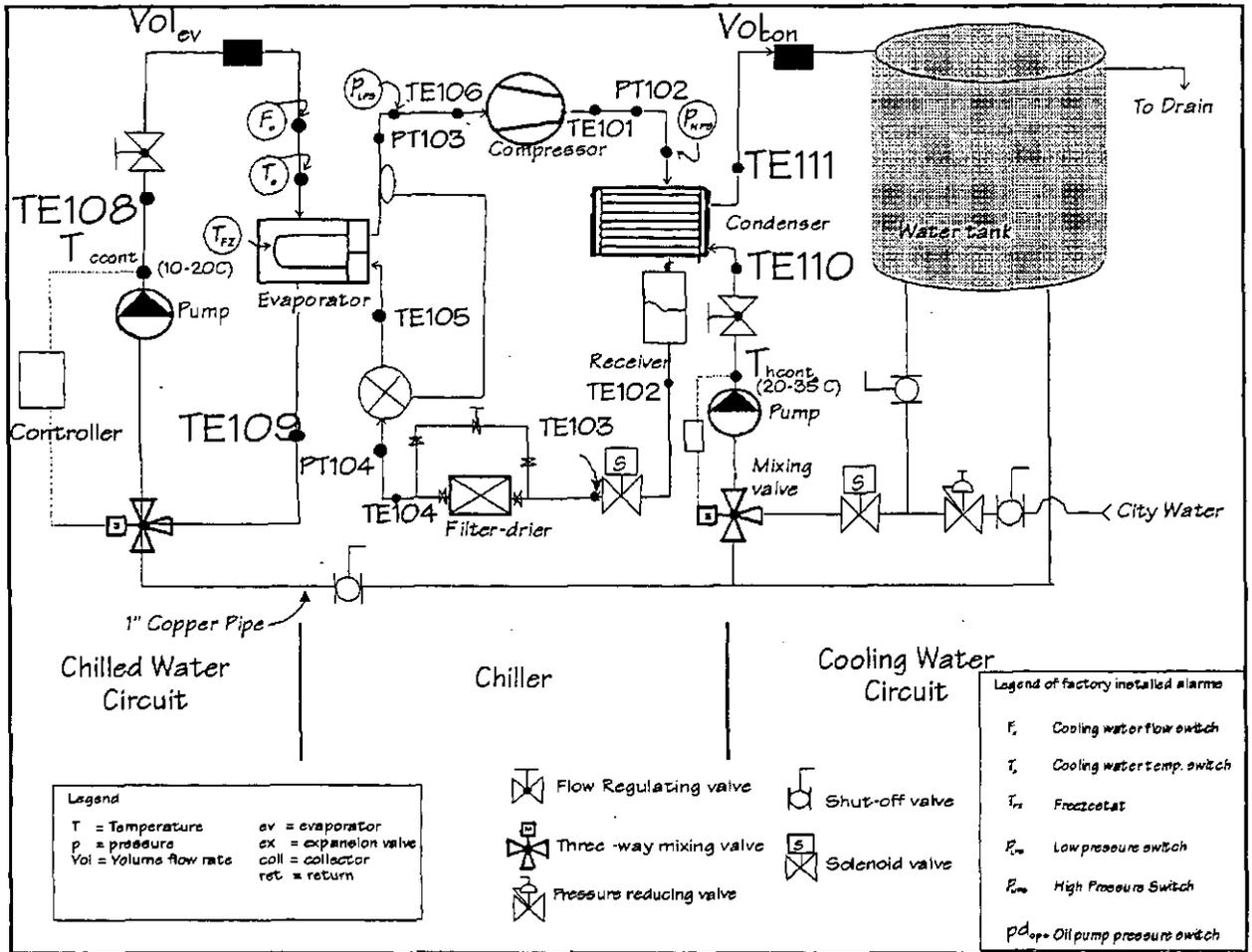
Figure 2   Test Unit

Experiments and FDDS development was carried out on a refrigeration test unit, shown in Figure 2. It is based on a commercially available reciprocating chiller using refrigerant R22 with a cooling capacity of 17.6 kW (5 RT).

The chiller components include a two cylinder semi-hermetic compressor, a cleanable shell and tube type condenser with water circulating through the tubes, and a direct expansion shell and tube evaporator. The chiller is equipped with a thermal expansion valve.

The test conditions for the chiller are achieved using a single storage tank and two three-way mixing valves serving two PID controllers. In order to have a controlled set point temperature at the inlet to the evaporator and condenser, the three-way valve supplying the evaporator mixes warm water from the storage tank with return water, while the valve serving the condenser mixes city water with storage tank water respectively.

Instrumentation and Data Acquisition System
The instrumentation of the testbench is composed of eleven platinum Resistance Temperature Detectors (RTDs) (tolerance: ±0.25+0.0042T°), four pressure transducers and two flow meters as given in Table 1.

With the exception of the RTD which measures the crankcase oil temperature (TE107), dry surface-mounted RTDs are used. This type of installation was chosen not only to avoid problems with refrigerant leaks, but also to duplicate the most likely way RTDs would be installed in the field. Pressures are measured using optical pressure transducers (accuracy: ±6.5 kPa (±1 psi)), mounted in the manner usually employed for pressure gauges, and as close to the desired point as conditions would allow. Flow rates are measured using positive displacement flow meters (accuracy: ±0.077 l/s (±1.2 gpm)).

Data acquisition is carried out using a microcomputer-based system which enables the user to establish sampling frequencies of up to 1Hz. The output data files can be stored on the PC serving the test unit and can also be transferred to other platforms with specialized software applications for further analysis.

Table 1 Measured variables

TE101	Discharge temperature, °C (°F)
TE102	High pressure (HP) liquid line temperature, °C (°F)
TE103	HP liquid line temp. (before filter dryer), °C (°F)
TE104	HP liquid line temp. (after filter dryer), °C (°F)
TE105	Low pressure liquid line temperature, °C (°F)
TE106	Suction line temperature, °C (°F)
TE107	Crankcase oil temperature, °C (°F)
TE108	Evaporator entering water temperature, °C (°F)
TE109	Evaporator leaving water temperature, °C (°F)
TE110	Condenser entering water temperature, °C (°F)
TE111	Condenser leaving water temperature, °C (°F)
PT101	Crankcase oil pressure, kPa (psi)
PT102	Discharge pressure, kPa (psi)
PT103	Suction pressure, kPa (psi)
PT104	High pressure liquid line pressure, kPa (psi)
FT101	Condenser water flow rate, l/s (gpm)
FT102	Evaporator water flow rate, l/s (gpm)

EXPERIMENTAL METHODOLOGY
The unit was used to map the normal range of operating conditions for commercial chillers. This was accomplished by changing the setpoints for the entering water temperatures at the evaporator and the condenser, simulating different cooling and chilled water return conditions. Cooling water temperature was varied between 22°C (71.6°F) and 34°C (93.2°F), while the chilled water entering temperature was varied between 10°C (50°F) and 15°C (59°F). The flow was set at 1.0 l/s (15.9 gpm) for both the condenser and evaporator for all experiments.

Five faults were used to develop and test of the Chiller Fault Detection and Diagnosis System, namely: refrigerant floodback at start-up, refrigerant leak, refrigerant line flow restriction and condenser and evaporator fouling. The fault associated with the refrigerant floodback was the only fault not deliberately introduced, as it was inherent in the chiller system due to a slightly undersized evaporator unit. The remaining four faults were introduced as shown in Table 2. The experiments were performed under ambient conditions, and temperature, pressure and flow variables were monitored every 10 seconds.

Separate files were created for transient and steady-state behaviour, and were subsequently used to develop and test the CFDDS.

Table 2: Fault Types

Fault Type	Fault Introduction Mechanism
Refrigerant leak	Removal of refrigerant
Refrigerant line flow restriction (Plugged filter-drier, obstructions in the piping etc.)	Throttling of line after condenser
Condenser water side flow resistance (Pump fault, fouling, etc.)	Reduction of water flow
Evaporator water side flow resistance (Pump fault, fouling, etc.)	Reduction of water flow

CHILLER PERFORMANCE MONITORING, FAULT DETECTION AND DIAGNOSIS SYSTEM

The system is subdivided into three distinct modules as shown in Figure 3:

The *Off-Cycle Module* is activated as soon as the chiller is turned "Off" for night
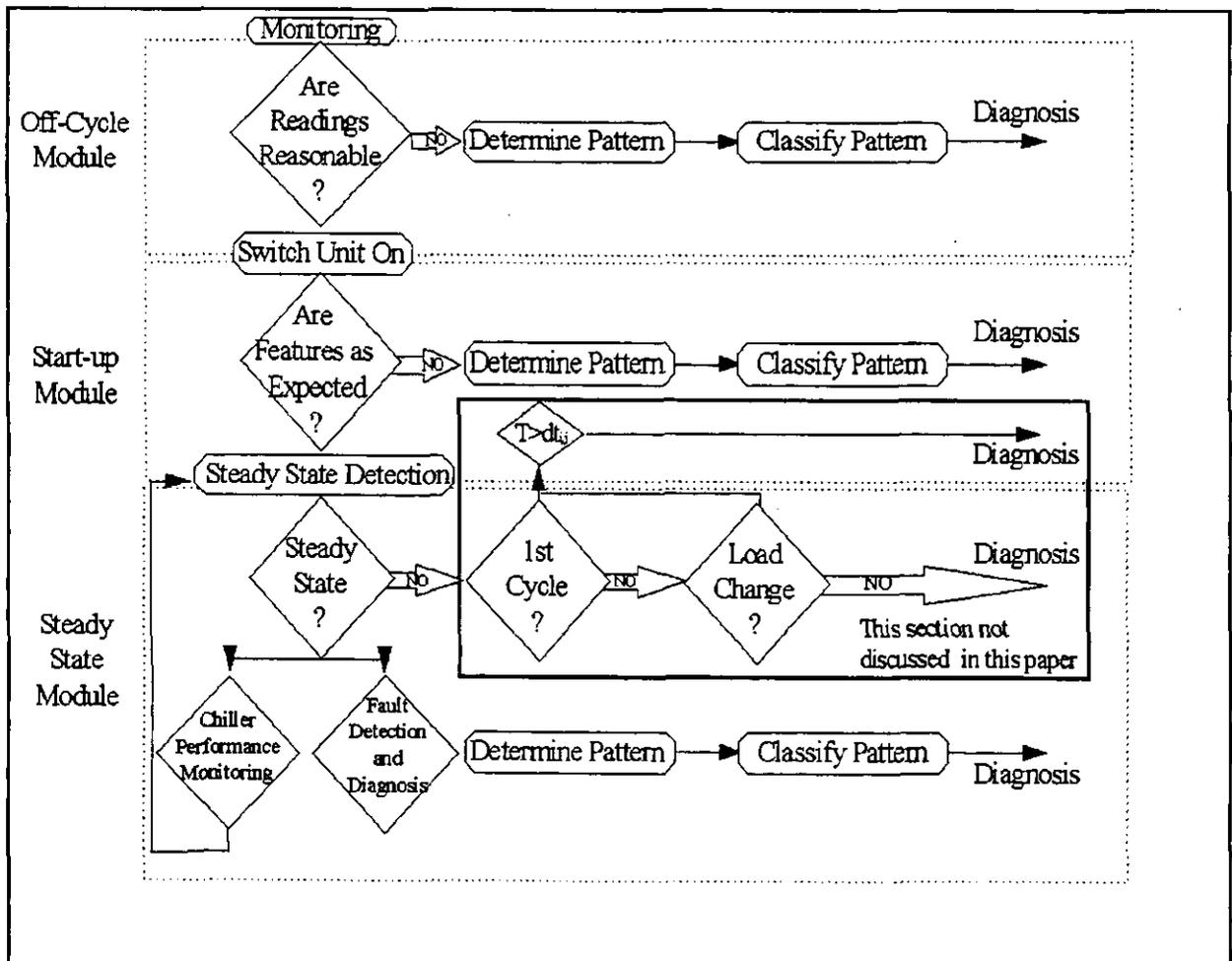
Figure 3 Flowchart for a Chiller Performance Monitoring, Fault Detection and Diagnosis System

time. It checks the performance of selected sensors and alerts the operator if one or more mechanical problems are detected before the next start-up.

The *Start-Up Module* is activated as soon as the chiller is turned "On" in the morning and remains active for about 15 minutes of operation.
It detects faults associated with refrigerant flow control that are easier to detect before the system reaches steady state.

The *Steady-State Module* is activated once the chiller reaches steady state. It remains activated while the machine's operating status is not changed. It ensures that the unit is operating within acceptable energy performance limits and performs fault detection and diagnosis primarily for faults that can not be compensated by the thermal expansion valve.
A detailed description of the individual modules follows.

Off-Cycle Module
Temperature sensor failure can take one of two forms: sensor bias and sensor drift. Sensor bias can be manifested as either a hard failure, where the sensor fails completely, or a soft failure, where the sensor still operates but produces incorrect readings. In the case of sensor drift, the bias exhibited by the sensor continuously changes in a given direction with time and can be seen as a special case of sensor bias.

In aerospace applications, where the reliability of the sensors is essential, hardware redundancy ensures the continuity of measurement in the case of a faulty sensor. Due to the high costs associated with multiplicity of sensors required for hardware redundancy, software redundancy is an area of active research (Patton, Frank and Clark, 1989). This method uses a number of approaches to model the expected behaviour of a system so that any measurement errors will be detected by deviations (or innovations) between state variables determined using the expected and measured data. In commercial chillers, it is possible to have both software and hardware redundancy with no sensor duplication and minimal computational requirements. This is possible primarily due to the mode of operation of the chiller. In a typical office building, commercial chillers are turned "Off" between 10:00 pm and 7:00 am. This provides a period during which the temperature of the unit approaches that of the mechanical room. The only part of the unit that does not reach room temperature is the crankcase oil reservoir, which is equipped with a heater designed to maintain the oil temperature between 8°C (14.4°F) and 14°C (25.2°F) above ambient conditions. The Off-Cycle Module uses this cooling period in two ways: for the verification of the dynamic performance of certain sensors and of the steady-state measuring capability of the sensor.

Immediately after the chiller is turned off, the temperature decay may, in some cases, be approximated by a first order model. The variables (Equation 1) may be established during commissioning. As these variables rely on the unit's time constant (τ) and the ambient conditions, the model will be valid provided the environmental conditions in the mechanical room are not changed significantly.

$$T^{\circ} = A + Be^{t/\tau} \tag{1}$$

In addition to dynamic testing, sensors are tested for bias. This is accomplished by using the sensor measuring the crankcase oil temperature (TE107) as an independent measurement. The Off-Cycle Module takes advantage of the built-in control of the oil crankcase heater to ensure the reliability of TE107. The module checks that the sensors have reached steady-state by monitoring the respective rates of change of temperature with time. The method used is based on functional variation, whereby the variation of the measured variable over a fixed time interval is compared to a hand-tuned value. This temperature variation, for a fixed time interval t, may be represented by:

$$V_n(t) = \frac{1}{N} \sum_{k=n-N+1}^{n} \left| T_k^* - T_{k-1}^* \right|$$

(2)

Once steady-state is achieved, the difference between individual sensors and TE107 are computed and compared to values established during commissioning. The procedure used, generally known as hypothesis testing, determines whether the temperature differs from the established one using the t-distribution:

$$t_{1-\frac{\alpha}{2}}(n-1) > \frac{\left| \Delta T - \eta_0 \right|}{\sqrt{s^2/n}}$$

(3)

If the test is successful, the samples are added to the population from which η_0 is estimated. If, however, the test fails then the sensor in question is declared faulty.

An application of the first order model is shown in Figure 4, which indicates the predicted and measured decay of the temperature read by sensor TE101. The model predicts the decay of
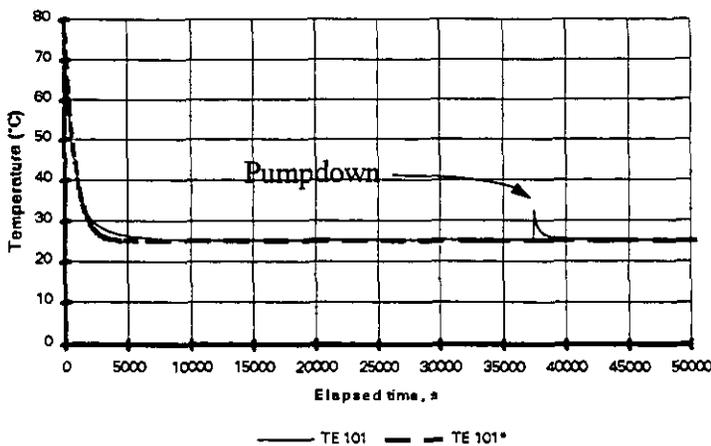
Figure 4 First order model approximation for the decay of TE101

TE101 with an error of less than 3% between 78°C (172.4°F) and 35°C (95°F), the usual operating range of the sensor. A larger error (< 7%) is encountered at lower temperatures, from 30°C (86°F) to 26°C (78.8°F).

The spike on the measured value of TE101, shown in Figure 4 is representative of the pump-down transient. The Off-Cycle Module measures the frequency at which this transient occurs and compares it to a pre-determined value. Any increase in the number of pump-downs is indicative of a deterioration of either the compressor valve seal or that of the solenoid valve, which isolates the high pressure from the low pressure parts of the unit (Figure 2).

Start-Up Module
The Start-Up Module is activated following the night (or weekend) "Off" period and monitors four of the 12 variables installed on the test unit: the discharge temperature (TE101) the crankcase oil temperature (TE107), and the refrigerant temperature entering (TE105) and leaving
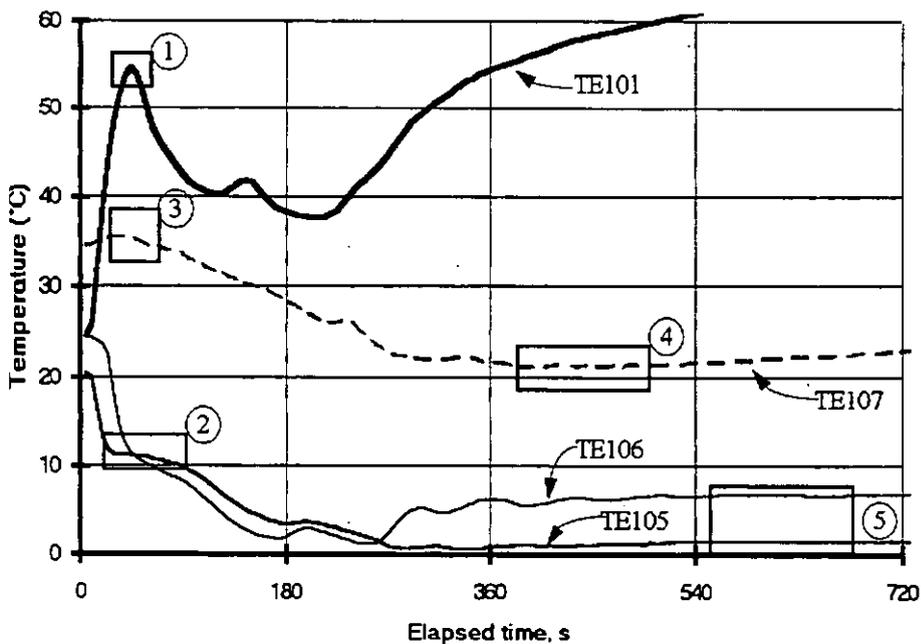
Figure 5 Features with diagnostic significance

(TE106) the evaporator.

Figure 5 shows a "baseline" transient response of the test unit during start-up. The features are indicated by boxes in the plots, with each feature being defined by its location in terms of time after start-up (x-axis) and the magnitude of the variable (y-axis). These features are particular to the unit and, as shown in figure 6, are repetitive. The room temperatures under which the tests were carried out varied by about 6°C. This variation did not have a detectable impact on the start-up transient profiles.
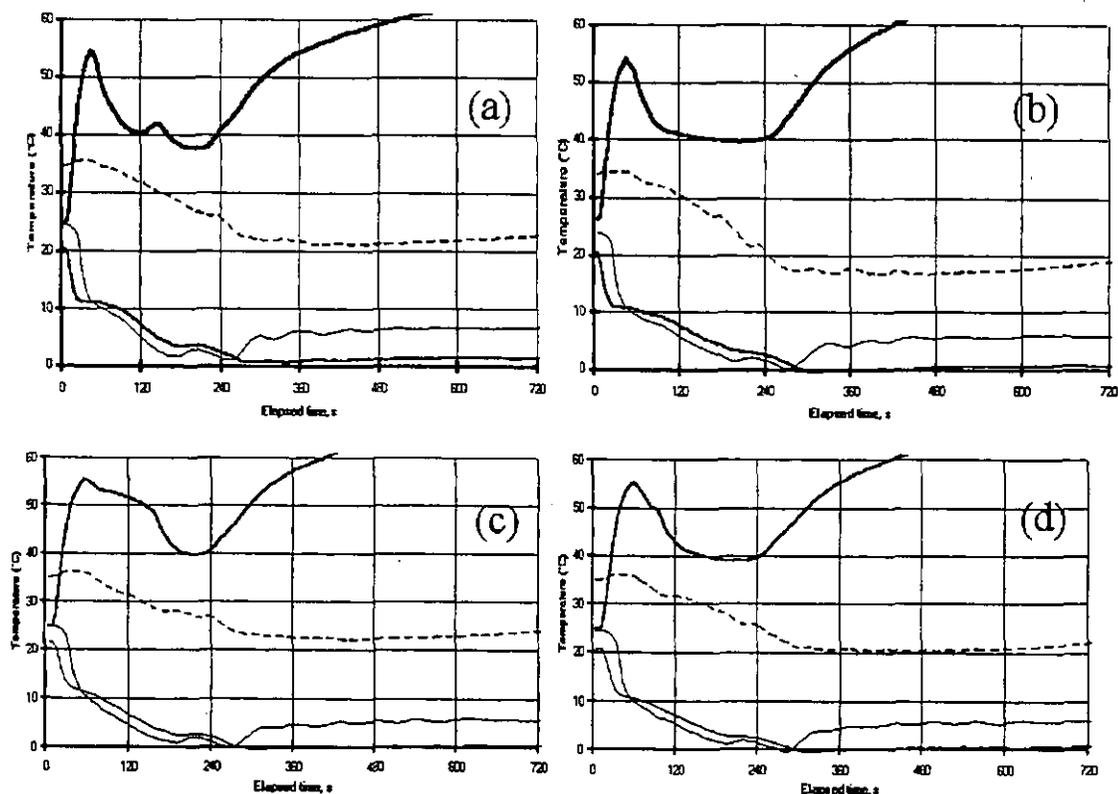
Figure 6 Start-up transients for the test unit

A typical list of faults and start-up features are given in Table 3. It is noted that the list of faults is not exhaustive and that the effefect of simultaneous faults on the transient behaviour of the unit needs to be examined.

Table 3 List of start-up transient features and their possible faults

Feature*		Possible faults associated with shifts in magnitude or timing of features
No.*	Feature Description	
1	Peak of TE101	Liquid refrigerant floodback, refrigerant loss, obstruction in refrigerant circuit
2	Inflection of TE105	Obstruction in refrigerant circuit, refrigerant loss
3	Slope of TE107	Liquid refrigerant floodback, thermal expansion valve faults
4	Minimum of TE107	Liquid refrigerant floodback, thermal expansion valve faults
5	ΔT (TE105-TE106)	Refrigerant leak, obstruction in refrigerant circuit

* Refer to Figure 5.

Either the presence or a shift in features can be used as a diagnostic indicator. This shift is influenced by the ambient conditions and, in case of outdoor packaged system applications, needs to be normalized. However, as chillers are indoor units, they do not experience a wide range of ambient conditions, thereby ensuring that the Start-Up module encounters nearly similar conditions every time it is turned on.

In case of refrigerant liquid floodback to the compressor, the features associated with this condition include the decrease in the discharge (TE101) and oil (TE107) temperatures, and the minimum of the oil temperature. The slope of TE107 (Figure 5) indicates that liquid refrigerant continues to enter the compressor until the minimum is reached, about 6 minutes later. This minimum coincides with achieving the 7°C of the superheat as shown by the difference between the low pressure liquid line temperature (TE105) and the suction temperature (TE106). This time period is therefore associated with the thermal expansion valve response time to arrive at the operating superheat level of the unit.

The variation between the start-up transients under "normal" or baseline conditions and for progressive refrigerant loss is shown in Figure 7. It is apparent that start-up transients of the "normal" chiller are markedly different from those of the faulty chiller in terms of mainly peak temperatures (TE101), and this feature is progressively modified according to the severity of the fault. Figure 8 displays the temperature transients for a progressively more severe flow obstruction in the refrigerant line. As shown, there is a clear upward shift of the TE101 peak as
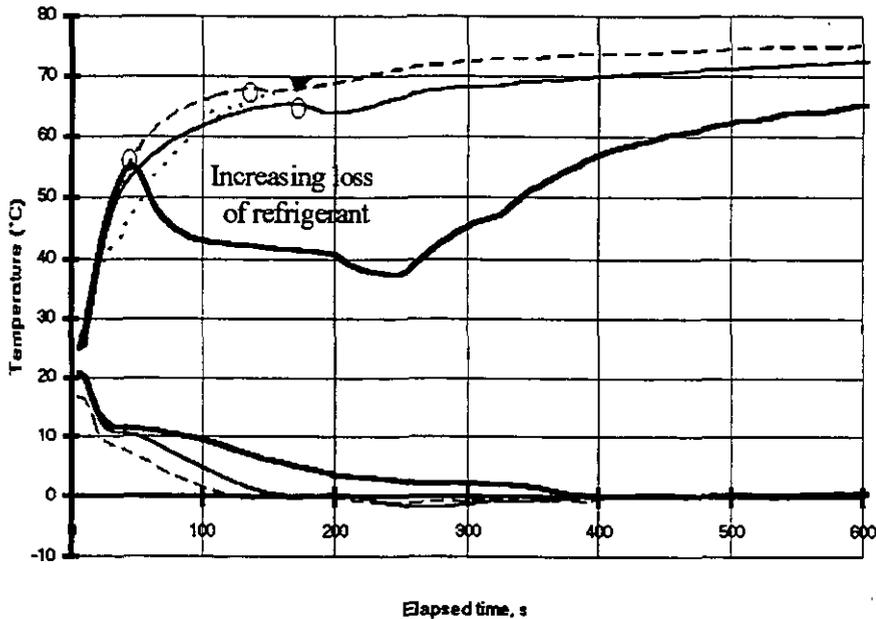
Figure 7 Comparison of start-up transients between baseline and unit with progressively lower refrigerant charge

84

well as a displacement in time while the inflection of TE105 appears to occur at a lower temperature than the baseline case. Increased obstruction produces a further shift to the
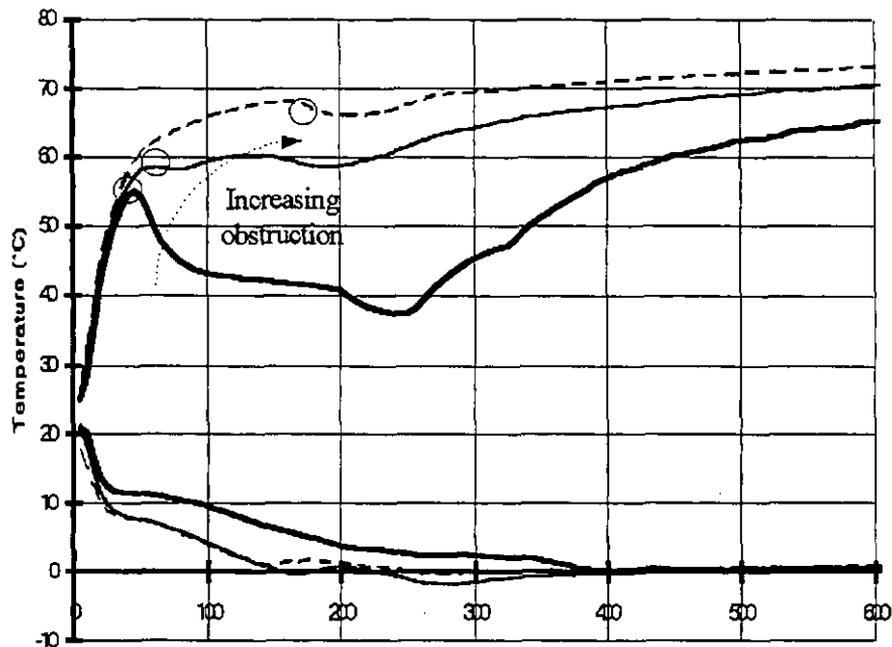
Figure 8 Comparison of start-up transients between baseline and unit with progressively bigger obstruction in liquid refrigerant line

magnitude and timing of the peak. The inflection of TE105 was not affected by the increase in pressure drop.

The features exhibited by the increased obstruction are nearly similar to those exhibited by refrigerant loss as seen in figure 7. Although this process allows us to detect a faulty condition, it may also permit the classification of each fault, but this needs to be further investigated.

Steady-State Module

Steady-state condition is determined using the same detection method used in the Off-Cycle module (Equation 2). Once steady-state is reached, the module performs two functions: the first one deals with the units' performance and the second one seeks to detect and diagnose a predetermined number of faults. This approach was devised so that the Chiller FDDS can detect deterioration in performance even in cases where no fault models have been defined, or in the event of multiple simultaneous faults. The performance level is established using a simple thermodynamic model of the chiller based on the manufacturers' data, while the fault detection and diagnosis procedure uses multiple linear regression models to generate innovations. The pattern of these innovations is subsequently classified in predetermined fault classes according to rules derived from literature and experiments.

Performance Level Monitoring

A simple thermodynamic model for the chiller is used to determine the actual performance of the unit, as described in detail by Gordon and Ng (1994). The following expressions for the condenser and evaporator temperatures are derived from heat balance considerations:

$$T_{cond} = T_{cond}^{in} - \frac{Q_{evap}\left(1+\frac{1}{COP}\right)[1-e^{(-NTU_{cond})}]}{(\dot{m}C)_{cond}} \tag{4}$$

$$T_{evap} = T_{evap}^{out} - \frac{Q_{evap}[e^{(-NTU_{evap})}-1]}{(\dot{m}C)_{evap}} \tag{5}$$

In addition to the approximation with respect to condenser and evaporator temperatures, the model makes a number of assumptions that are specific to commercial chillers. In particular, the

model assumes that the zero net change in entropy of the chiller fluid can be expressed as:

$$\Delta S = 0 = \frac{Q_{cond} - q_{cond}^{loss}}{T_{cond}} - \frac{Q_{evap} + q_{evap}^{loss}}{T_{evap}} \tag{6}$$

where q^{loss} refers to the losses due to heat leaks, fluid friction, throttling and superheating. All heat transfer processes are approximated as isothermal since the deviations from isothermal behaviour is small relative to the absolute temperatures T_{cond} and T_{evap}.

An expression for the COP is obtained by combining equation (6) with the expressions for T_{cond} and T_{evap}:

$$\frac{1}{COP} = -1 + \frac{T_{cond}^{in}+T_{cond}'}{T_{evap}^{out}-T_{evap}'} + \frac{q_{cond}^{loss}+\dfrac{q_{evap}^{loss}\left(T_{cond}^{in}+T_{cond}'\right)}{T_{evap}^{out}-T_{evap}'}}{Q_{evap}} \tag{7}$$

Where:

$$T_{evap}' = \frac{Q_{evap}}{(\dot{m}C)_{evap}[e^{(-NTU)_{evap}}-1]} \tag{7a}$$

$$T_{cond}' = \frac{Q_{evap}[1+(1/COP)]}{(\dot{m}C)_{cond}[1-e^{(-NTU)_{cond}}]} \tag{7b}$$

Based on chiller performance data, the only term on the right-hand side of Equation 7 that depends on cooling rate and which contributes significantly is the term proportional to $1/Q_{evap}$. Therefore, for typical commercial chiller operating ranges, Equation 7 can be approximated

86

linearly by:

$$\frac{1}{COP} = -1 + \frac{T_{cond}^{in}}{T_{evap}^{out}} + \frac{q_{cond}^{loss} + (q_{evap}^{loss} T_{cond}^{in} / T_{evap}^{out})}{Q_{evap}}$$

(8)

In order to use the particular characteristics of commercial chillers, additional assumptions are imposed regarding the linearity of heat transfer and the isentropic nature of throttling and superheating (Cerepnalkovsky, 1991). Under these assumptions q^{loss} is approximated by:

$$q_{cond}^{loss} = -A_0 + A_3 T_{cond}$$

(9)

and

$$q_{evap}^{loss} = -A_2 + A_4 T_{evap}$$

(10)

Using equations (8)-(10) and setting $A_1 = A_3 + A_4$, one obtains:

$$\frac{1}{COP} = -1 + \frac{T_{cond}^{in}}{T_{evap}^{out}} + \frac{-A_0 + A_1 T_{cond}^{in} + A_2 (T_{cond}^{in} / T_{evap}^{out})}{Q_{evap}}$$

(11)

A_0, A_1, and A_2 are constants that characterize the irreversibilities of a particular chiller and are derived from data routinely provided by chiller manufacturers using linear regression. The required data consists of tables representing the performance of the chillers (Q_{evap}, P) for a number of entering and leaving water temperatures at the condenser and evaporator respectively (T_{cond}^{in}, T_{evap}^{out}). To estimate A_0, A_1, and A_2, use is made of the bilinear nature of equation 11, which effectively is a function of two variables, namely T_{cond}^{in} and $T_{cond}^{in} / T_{evap}^{out}$. This is as follows: the first step in estimating the constants is to plot:

$$\left(\frac{1}{COP} + 1 - \frac{T_{cond}^{in}}{T_{evap}^{out}} \right) Q_{evap} \quad against \quad \frac{T_{cond}^{in}}{T_{evap}^{out}}$$

(12)

The resulting lines, one for each of the given T_{cond}^{in}, have a slope that is equivalent to the value of A_2.
Once A_2 is estimated a second plot of:

$$\left(\frac{1}{COP} + 1 - \frac{T_{cond}^{in}}{T_{evap}^{out}} \right) Q_{evap} + A_2 \frac{T_{cond}^{in}}{T_{evap}^{out}} \quad against \quad T_{cond}^{in}$$

(13)

produces a straight line with A_0, as the y-axis intercept and A_1, as the slope. A detailed description of the method for determining A_0, A_1, and A_2 is described in Gordon and Ng (1994).

The above method was applied to develop the model for the test unit. Figure 9 displays the calculated COP values for various load conditions for the test unit.
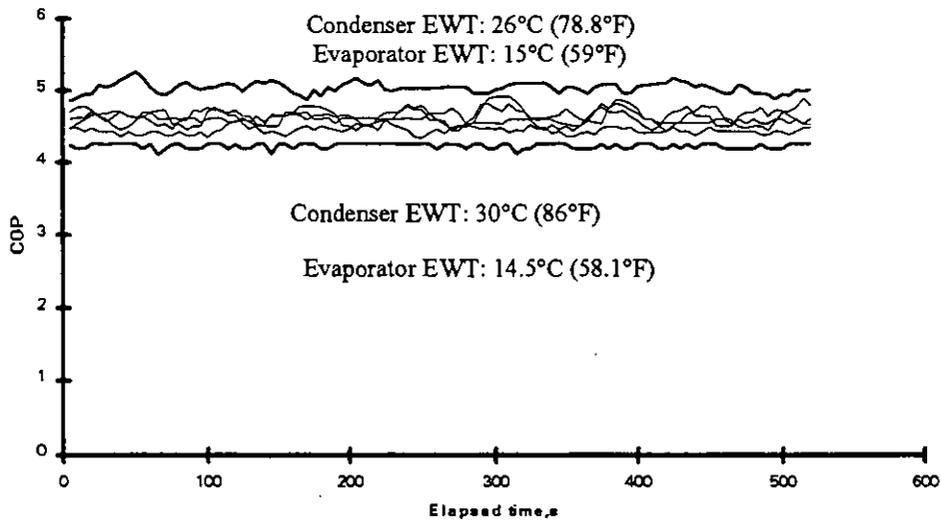
Figure 9 Calculated COPs for test unit under different operating conditions
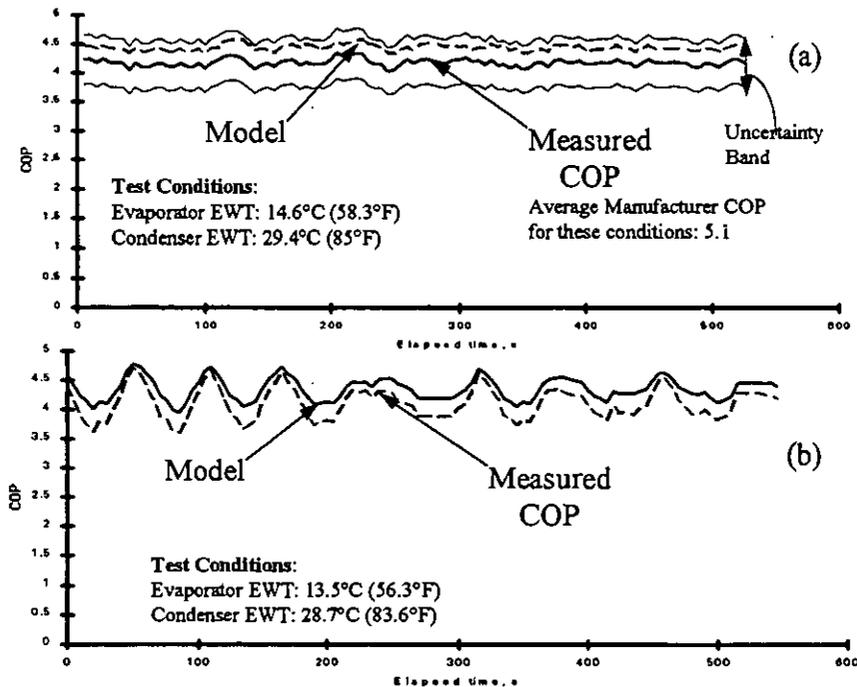
Figure 10 Comparison of measured and predicted COP for test unit

The calculated COP values compare well with the measured COP and, as shown in Figure 10(a), the difference between the model-based (calculated) COP and the measured COP is less than 6%, falling within the measurement uncertainty band. Figure 10(b) shows the model COP following closely the measured COP during a test where the throttle valve was induced to hunting.

Fault Detection and Diagnosis

Symptom matrices for a number of different faults encountered in vapour compression machines have been developed (Grimmelius, Klein Woud and Been (1994) and Rossi and Braun (1995)). The matrix used in the present work was modified using data obtained from experiments and used for the detection and diagnosis of faults in the second function of the Steady-State Module. Table 4 displays the modified matrix used for the detection and diagnosis of faults. The "+" represents an increase in the value of the variable over the normal operation of the chiller while "-" indicates a decrease.

Table 4 Fault patterns used in the diagnostic module

Fault	TE101	TE102	PT102	TE105	TE106	PT103	ΔTcond	ΔTev
Restriction in Refrigerant Line	+	-	-	-	+	-	-	+
Refrigerant Leak	+	-	-	-	+	-	-	+
Restriction in Cooling Water	+	+	+	-	-	-	+	-
Restriction in Chilled water	+	-	-	-	-	-	-	+

Linear regression models are used to generate estimates of the pressure and temperature variables. The estimated variables are then compared to their measured values, the innovations thus detected are then matched, using a rule-base to the patterns shown in Table 4, thereby diagnosing the fault. As indicated in equation 14, the temperature and pressure variables represented by vector y are a function of the two independent temperature variables, namely the entering water temperatures to the condenser (TE110) and the evaporator (TE108).

$$y_i = \beta_0 + \beta_1 TE108 + \beta_2 TE110 \tag{14}$$

Variable	Adjusted R^2	Std. Error of Estimate
TE101	0.99917	0.11280
TE102	0.99987	0.03416
TE103	0.99988	0.03334
TE105	0.99810	0.63781
TE106	0.99480	0.10531
TE107	0.96945	0.29026
TE109	0.99985	0.01872
TE111	0.99996	0.01969
PT101	0.99811	0.14052
PT102	0.99892	0.51470
PT103	0.99879	0.11641
PT104	0.99878	0.53479

Table 5 Results of Regression Analysis
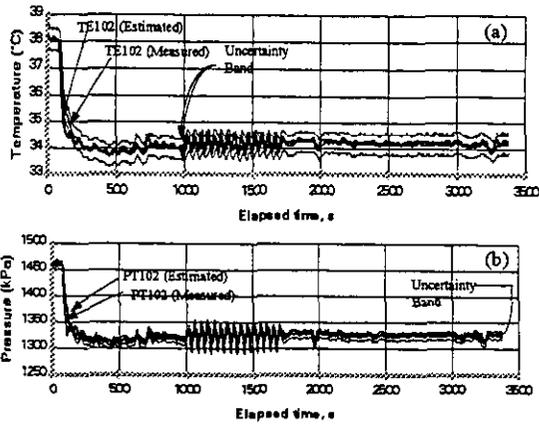
Figure 11 Comparison of measured and predicted liquid line temperature (TE102) and discharge pressure (PT102)
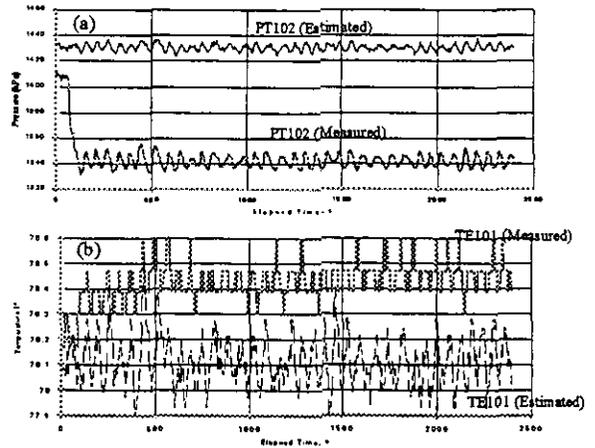
Figure 12 Differentiation of predicted and measured values during and after the introduction of a restriction in the liquid refrigerant line

Figure 11(a) and 11(b) display the measured and predicted values for TE102 and PT102 based on the above model.

The measured pressure (PT104) upstream of the thermal expansion valve and its predicted value using the regression model during the introduction of a restriction in the liquid line is presented in Figure 12(a). As can be seen, the pressure after the restriction is decreased by about 85 kPa. As expected, the effect of this restriction is propagated at other points of measurement around the cycle. One such point is the temperature at the discharge of the compressor (TE101), as shown in Figure 12(b). The average values for the temperature and pressure values were used to compensate for temperature and pressure fluctuations. The difference between the measured, averaged temperature (TE101) and the predicted, averaged value, together with the differences of the other measured parameters, are used to create a fault diagnostic pattern that can be matched to those described in Table 4.

The experimental results for the case of a flow restriction of 85 kPa, and for that of 140 kPa are shown in Table 6, and compared to their predicted values ("U", signifies differences that fall within the Uncertainty band, with "U+" and "U-" signifying tendencies to the upper and lower limits of the uncertainty band respectively).

It is apparent that the pattern of the differences between the measured and predicted values for the first experiment is less discernible than that of the second experiment. This is not only the result of the increased severity of the obstruction introduced, but is also influenced by the compensating action of the Thermal eXpansion Valve (TXV) .

Table 6 Patterns generated using the linear regression models for operation of the chiller with two different obstructions in the refrigerant liquid line

Variable	~85 kPa Pressure Drop		Pattern	~140 kPa Pressure Drop		Pattern
	Average, Measured Value	Average, Predicted Value		Average, Measured Value	Average, Predicted Value	
TE101 °C	78.65	78.14	+	79.38	78.09	+
TE102 °C	35.44	35.54	U−	35.41	35.52	U−
TE103 °C	34.08	34.16	U−	34.04	34.14	U−
TE105 °C	0.57	0.87	−	0.45	0.9	−
TE106 °C	6.59	6.59	U	6.87	6.62	U+
TE107 °C	41.71	41.51	U+	42.62	41.48	+
TE109 °C	9.48	9.43	U+	9.52	9.46	U+
TE111 °C	32.78	32.79	U−	32.72	32.76	U−
PT101 kPa	616.71	621.90	−	613.28	622.31	−
PT102 kPa	1366.65	1372.81	−	1370.37	1372.09	U−
PT103 kPa	486.29	491..85	−	484.60	492.26	−
PT104 kPa	1343.95	1431.42	−	1290.44	1430.70	−

The TXV being the control element in the chiller unit, maintains the refrigerant entering the compressor at a pre-set superheat level, thereby preventing damage to the compressor by liquid refrigerant. The superheat is maintained throughout the normal load conditions of the unit by modulating the TXV. In addition to changing load conditions, the TXV is actuated by disturbances caused by faults such as obstructions in the refrigerant circuit and the loss of refrigerant. In these cases, if the valve orifice area were maintained constant, an elevated superheat would have been produced. Since the TXV can modulate the flow, it reacts to compensate for the increase in superheat by allowing more refrigerant to flow through to the evaporator, thereby maintaining the pre-set level of superheat.

This feedback action of the TXV tends to compensate for the disturbances caused by the faults modifying the diagnostic patterns, thereby limiting the sensitivity of the rule-based pattern recognition approach as presented in Table 3.

Discussion

The paper presents a system for detecting and diagnosing faults occurring in reciprocating chillers. The nature of the chiller fault detection and diagnosis system requires that the signals carrying the sensor data be filtered to eliminate disruptions in the fault detection and diagnosis process. The present configuration of the system includes a low pass filter which effectively acts as a three point moving average smoothing function. Similar filtering will be required to minimize false alarms from electrical noise or other interference in the quality of the signal.

This system employs three modules each dedicated to a specific operating condition of the chiller. The off-line module assumes that the sensors are calibrated and properly installed. The module uses data collected during the periods when the chiller is "off" for extended periods (overnight, weekends) to estimate parameters that are used to detect temperature sensor faults. Initial estimates of the parameters may be determined during commissioning and fine-tuned through additional collection of data following the occupation of the building.

The second module, dedicated to monitoring the initial start-up of the chiller, uses features that may also be determined during commissioning to detect and diagnose faults since different chillers may display different start-up transients (Tassou and Al Nizari, 1993). The initial transient profiles have been shown to contain useful information on the "health" of the test unit. This information is obtained through the observation of the variations of the transients' particular features and classified based on a predetermined rule-base. Further work is necessary to ascertain if the variations in the features exhibited in the laboratory test unit are generally applicable to this type of chiller.

The methodology followed in the start-up module allows a rapid determination of the chiller condition which is independent of the behaviour of the thermal expansion valve. In particular for the test unit, it only takes about one minute for the discharge temperature to reach the first peak, therefore, it is not influenced by the response of the thermal expansion valve, whose response is in about five minutes. Consequently the start-up transients, on which the fault detection and diagnosis are based, is not influenced markedly by the final load conditions.

The detection and diagnosis of faults in this manner accomplishes the detection of faults before they become serious ones or cause mechanical failure in the machine. However, this method requires fast sampling, of the order of 5 seconds, and its performance depends on the accuracy and precision of the sensors employed.

The Steady-State module requires eight temperature sensors and two pressure transducers to perform fault detection and diagnosis as described in the present paper. Although this may represent a significant cost, further work may reduce this number through the use, for example, of differential temperature measurements.

The data collected from the condenser and evaporator entering water temperatures are used by

the Steady-State module to develop a bilinear linear regression to estimate selected variables for fault detection and diagnosis during steady-state operation. This necessitates the collection of training data which are not contaminated by faulty conditions. Data representing the operating envelope of the unit would not be available upon installation and will therefore require several months of operation before the data base is complete enough to develop the model. During this period of training, the "health" of the unit may be determined by the Start-up module in combination with the performance monitoring function of the Steady-State module. This function employs a linear model (Gordon and Ng, 1994) for reciprocating chillers. The model was applied to the test unit and was found to produce acceptable results. Because the physical basis for the model is clearly formulated and not based on empirical fits, one can state the condition under which the linear approximation will be inadequate. This deviation from linearity occurs when the temperature differences across each chiller component are large, leading to increase in magnitude of the contributions of the linear approximations used to arrive at equation 11. It has also been shown (Gordon and Ng, 1994) that for the normal operating range of water chillers the magnitude of these approximations does not affect the validity of the model. The model was tested for 30 chillers varying in size from 30 kW to 1300kW and the difference between the calculated and measured COP was within the measurement uncertainty band.

The results presented in this paper pertain to tests performed on one test unit under laboratory conditions. Furthermore as the extent of the tests is limited, further work is required to verify the applicability of the approach to the more general population of reciprocating chillers, both new and retrofit.

CONCLUSIONS
The present paper presents a global approach to the performance monitoring, fault detection and diagnosis of commercial chillers. The methodology presented integrates elements from artificial intelligence, pattern recognition and physical modelling do determine the "health" of a reciprocating chiller and to diagnose selected faults.

The approach taken pre-supposes that all operating and non-operating phases of the chiller can contribute in establishing fault detection conditions. The three modules of the diagnostic system include the *off-cycle, (initial/morning) start-up* and *steady-state modules* which cover the spectrum of the chillers' operating states. This allows the detection and diagnosis of faults at operating conditions providing the maximum level of information. In the envisaged application of the methodology, faults that influence refrigerant flow such as refrigerant flow obstructions or refrigerant leak will be detected and diagnosed by the start-up module prior to the actuation of the thermal expansion valve. This is an approach that has not been previously used for the detection and diagnosis of chiller faults, and which has been shown in the present paper to provide useful information on the "health" of a laboratory chiller test unit. Faults currently not in the knowledge base, as well as simultaneous faults affecting the energy performance of the chiller, will be detected by the steady-state module. In addition, the steady-state module can detect and diagnose faults associated with load conditions as well as providing fault detection capability in cases of more severe faults which may not be detected by the start-up module due to unpredicted changes in ambient conditions.

An important aspect of the methodology is the need for training data. In the case of the off-cycle, the start-up modules and the performance monitoring portion of the steady-state module, the data can be gathered during commissioning. However, the diagnostic portion of the steady-state module requires a larger amount of training data which would not be available during commissioning. On-line measurement of variables to establish the normal steady-state operational envelope needs to be carried out to allow build-up of linear regression models. The start-up module and the model developed from manufacturers' data will be used to build-up the training database by filtering out faulty conditions. This requirement will need to be satisfied with a minimum influence on the performance of fault detection and diagnosis and to the operation of the chiller.

Generally, steady state and transient fault patterns are modified by a number of factors. These depend on the severity of the fault and, for the steady-state case, the compensatory actions of the thermal expansion valve.

As there were a limited number of experimental tests performed, there is need for further experimental data to verify the repeatability of the emerging patterns with a high statistical confidence level.

Furthermore this methodology needs additional treatment to improve the definition of the thresholds for classification of emerging patterns and to establish the range of applicability of these patterns. Further study is required to produce an exhaustive list of faults and their associated transient features and steady-state patterns.

Nomenclature

A	Constant, °C
A_0	Constant, kW
A_1	Constant, kW/K
A_2	Constant, kW
B	Constant, °C
C	Specific heat, kW/kgK
COP	Coefficient of performance
EWT	Entering water temperature
LWT	Leaving water temperature
m	Mass flow rate, kg/s
n	Sample size
N	Sample size
NTU	Number of transfer units
P	Power, kW
Q	Heat transfer rate, kW
q^{loss}	Heat losses, kW

s	Standard deviation
S	Entropy, kW/K
t	Time, s; t-distribution
$T°$	Temperature, °C
T	Temperature, K
TXV	Thermal xpansion valve
α	Confidence level
β	Regression constant
η_0	Mean
τ	Time constant, s

Subscripts

cond	Condenser
evap	Evaporator

Superscripts

in	Entering flow
out	Leaving flow
*	Estimated

References

Cerepnalkovsky, I., 1991. *Modern Refrigerating Machines*. Amsterdam, Elsevier.

Gimmelius, H.T., J. Klein Woud and G. Been 1995. On-line failure diagnosis for compression refrigeration plants. *Rev. Int. Froid* 18 (1): 31- 41.

Gordon, J.M. and Kim Choon Ng 1994. Thermodynamic modelling of reciprocation chillers. *J. Appl. Phys.* 75 (6): 2769-2774.

Gordon, J.M. and Kim Choon Ng 1995. Predictive and diagnostic aspects of a universal thermodynamic model for chillers. *Int. J. Heat Transfer 38 (5): 807-818.*

Herzog, P., L. LaVine, 1992. Identification and quantification of the impact of improper operation of midsize Minnesota office buildings on energy use: A seven building case study. *Proc. ACEEE 1992 Summer Study on Energy Efficiency in Buildings*, pp.3.121-3.129

Huang,J., H. Akbari, L. Reiner and R. Ritschard 1991. *481 Prototypical commercial buildings for 20 urban market areas*. GRI Report GRI-90/0326.

Patton, R., P. Frank and R. Clark 1989. *Fault diagnosis in dynamic systems*. Cambridge: Prentice Hall.

Rossi, T. and J. Braun, 1995. Thermodynamic impact for detecting refrigerant leaks in vapor compression equipment. Accepted for 1995 American Control Conference, IEEE Control Systems Society. Seattle Washington, USA

Sami, S.M., Y. Zhou and P.J. Tulej 1993. Development of a diagnostic expert system for heat pumps. *Proc. 2nd Annual Conf. Heat Pumps in Cold Climates.* Moncton, New Brunswick, Aug. 16-17, pp.475-489.

Shoureshi, R. and J. Wagner 1992. Failure detection diagnostics for thermofluid systems. *Journal of Dynamic Systems, Measurement, and Control* 114: 699-706.

Tassou, S.A. and H.O. Al -Nizari 1993. Effect of refrigerant flow control on the thermodynamic performances of reciprocating chillers. *Applied Energy* 45:101-116

MAN MACHINE INTERFACE FOR A FAILURE DIAGNOSIS EXPERT SYSTEM (EnDEx)

Daniel E. Maurer
Swiss Federal Institute of Technology
Energy Systems Laboratory
CH-8092 Zurich
Switzerland

Abstract

A system for on-line process visualisation and failure diagnosis is explained in terms of a prototype realisation and its underlying concepts. The first section sets the scope: The tasks are failure diagnosis and on-line process visualisation of a heat pump plant. The goals comprise openness of the system, userfriendliness and a clear knowledge methodology. Section 2 addresses the issues of the chosen system architecture. In section 3 the man machine interface for a heat pump application is explained by means of the different screens. The sections 4 and 5 discuss the expert system diagnostics and the knowledge engineering methodology.

1 INTRODUCTION

1.1 BASIC SYSTEM ARCHITECTURE

EnDEx (Energy System Diagnosis Expert System) was designed as a generic framework for supervisory tasks in the domain of HVAC equipment and plants (see Table 1). Data from the plant is periodically collected from the plant DDC or PLC and pasted in a blackboard database (see Figure 1).
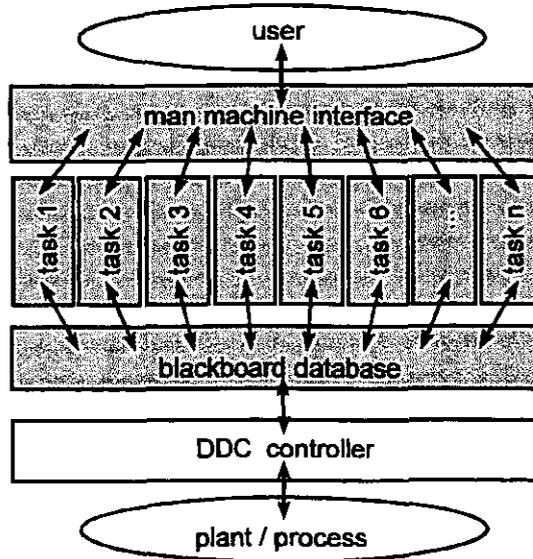
Fig. 1: Basic architecture of EnDEx (in grey) as framework for different tasks.

From here the different tasks can get the data they need or can exchange information with other tasks. The processed information is displayed in a uniform man machine interface.

1.2 TASKS AND SCOPE OF THE DESCRIBED SYSTEM

plant subsystem / task type	on-line process visualisation	failure diagnosis	fault detection	....
boiler	task 01	task 02	task 03	.....
heat pump	task 11	task 12	task 13	.....
heat distribution	task 21	task 22	task 23	.....
ventilation	.....	.....	.....	.....
....	.....	.....	.....	.....

Tab. 1: The current EnDEx prototype incorporates the tasks 11 and 12

- Task types: The present state of EnDEx incorporates the task types of *on-line process visualisation* and *failure diagnosis*. Extended detection mechanisms as well as long term wear models are considered for later implementation.

- Subsystems: The chosen task types were carried out for the *heat pump* subsystem. Other subsystems could be incorporated.

- Development stage: EnDEx was developed to the stage of a demonstration prototype. Testing and validation was carried out at the Scheco Ltd. laboratory plant in Winterthur.

1.3 GOALS

- For EnDEx as a generic framework: Being open for the various plant subsystems and the different tasks requires the use of *standard* hardware and software components and communication facilities.

- For EnDEx as man machine interface: To be operable by unskilled service staff a maximal *user-friendliness* has to be achieved.

- For the diagnostic part in EnDEx: To ensure a good maintainability of the diagnostic knowledge a homogenous structure should be realised with the help of a suitable *knowledge engineering methodology*.

2 SYSTEM DESIGN

2.1 THE PHYSICAL COMPONENTS

- Different analog sensors (temperature, pressure, etc.) and binary switches (high pressure, flow watcher, etc.) send their signals to the DDC-controller.

- The DDC or PLC plant controller acquires the signals, converts and buffers it as digital values. Besides the standard tasks of the controller (e.g. start-/stop-sequence), a simple data processing may take place. The controller is equipped with a serial interface.

- An IBM-compatible PC serves as *hardware* platform for the EnDEx-System.

- A high-speed modem with phone- or ISDN-connection access makes the link to the outer world. The entire local PC screen can be transmitted to a remote terminal.

sensors/switches   controller   PC   modem
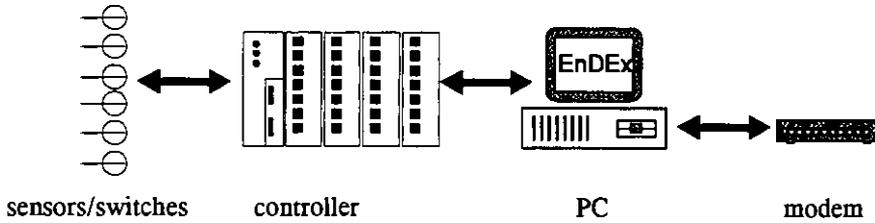
Fig. 2: The hardware components of the overall system

2.2 FUNCTIONS OF THE DIAGNOSTIC TASK

- *Failure detection:* Failures are detected directly from the security switches or from threshold supervision of analog data within the controller.

- *Event management in the plant controller:* When a failure event has occurred the controller generates an alarm message. An alarm script handles data buffering and event data recording.

- *Communication of the event data:* The buffered values from the last alarm event are uploaded from the DDC to the PC.

- *Diagnosis by EnDEx:* Determination of the process state, plausibility check of input data and deduction of the cause.

- *Diagnostic output of EnDEx:* It includes failure location, maintenance domain, explanation of the reasoning path, suggestions for correction and input data are displayed. For details see section 3.7 to 3.8.

2.3 FUNCTIONS OF THE ON-LINE VISUALISATION TASK

The current process state is displayed in EnDEx by means of:

- gauges of some important process states,
- sensor values in PI-charts or in the process data table,
- calculated process values.

For details see section 3.5 and 3.6.

2.4 PROGRAM MODULES AND SOFTWARE PACKAGES

As EnDEx is a prototype standard software was chosen for the different program modules (see fig. 3) .

- Since up to now no time critical operations had to be performed no special real time *operating system* has to be considered. To facilitate the development of user friendly and widely known man machine interfaces Microsoft Windows™ was chosen.

24.02.1996

```
┌─────────────────────────────────────────────┐
│  man machine interface    ToolBook           │
│  ┌────┬────┬────┬────┐         ┌──────────┐  │
│  │exp │onl │    │    │         │          │  │
│  │ert │ine │    │    │         │          │  │
│  │sys │pro │    │    │         │          │  │
│  │tem │ces │ :  │ :  │         │ device   │  │
│  │Nex │s   │    │    │         │ driver   │  │
│  │pert│vis │    │    │         │ DynaComm │  │
│  └────┴────┴────┴────┘         └──────────┘  │
│  blackboard  Excel                           │
│  operating system  Windows                   │
│  remote control      Carbon Copy             │
└─────────────────────────────────────────────┘
```
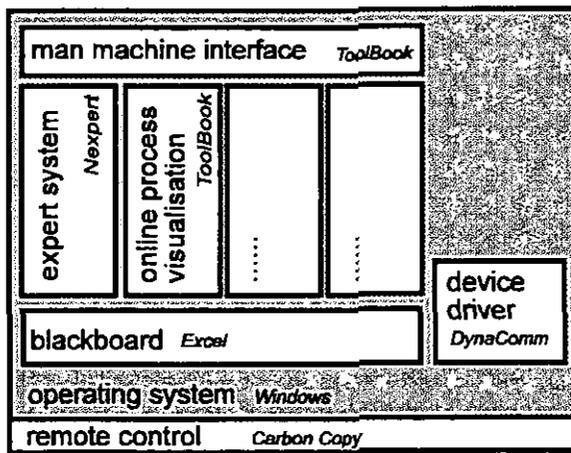
Fig. 3: The program modules are programmed in different standard software packages.

- In particular the multi media authoring software ToolBook™ from Asymmetrix, Inc. speeds the development of uniform *screens*. The object oriented approach (with inheritance features) allows the design of generic screens that can be used by the specific screens.
- The *black board* as central database is implemented using Microsoft Excel™
- The generic *device driver* to the DDC equipment of the plant was realised with the communication software DynaComm™ from FutureSoft, Inc. A fairly mature scripting language allows quick access to the DDC. To move from the prototype to a commercial system, a custom built device driver (written in C or Pascal) would replace the current configuration increasing the communication rate. The physical link is done using a RS 232 serial interface.
- The expert system knowledge base of the diagnostic task is stored and processed within Neuron Data's NexpertObject™. A combined rule-based and object-oriented knowledge representation and processing is provided. The graphical display of the rule network eases development and debugging.
- Remote access and control can be achieved with a standard remote control package such as Microcom's Carbon Copy™. No explicit communication has to be programmed, since this software transmits the entire graphic screen to the guest station. A high-speed modem with integrated compression and fault correction capabilities is required.

The different program modules may communicate with each other either via the blackboard or directly via the DDE-facilities (dynamic data exchange) provided by MS-Windows™.

3 SCREENS

3.1 INTRODUCTION

User-friendliness in a supervisory system can only be achieved when certain aspects are taken into account:

- All the different controls and commands should be uniformly handled: on different screens the same buttons should be placed at the same location; for the same commands always the same command name should be used.

- No user entries should be expected from the user: only buttons, checkboxes, choicelists and menus are used to control the application.

- Always up to date status information should be displayed. Old information has to be cleared immediately.

In the following the man machine interface is explained with the German screens of EnDEx. For the better understanding most of the screen texts are translated (= German Screen Text).

3.2 SEQUENCE OF SCREENS
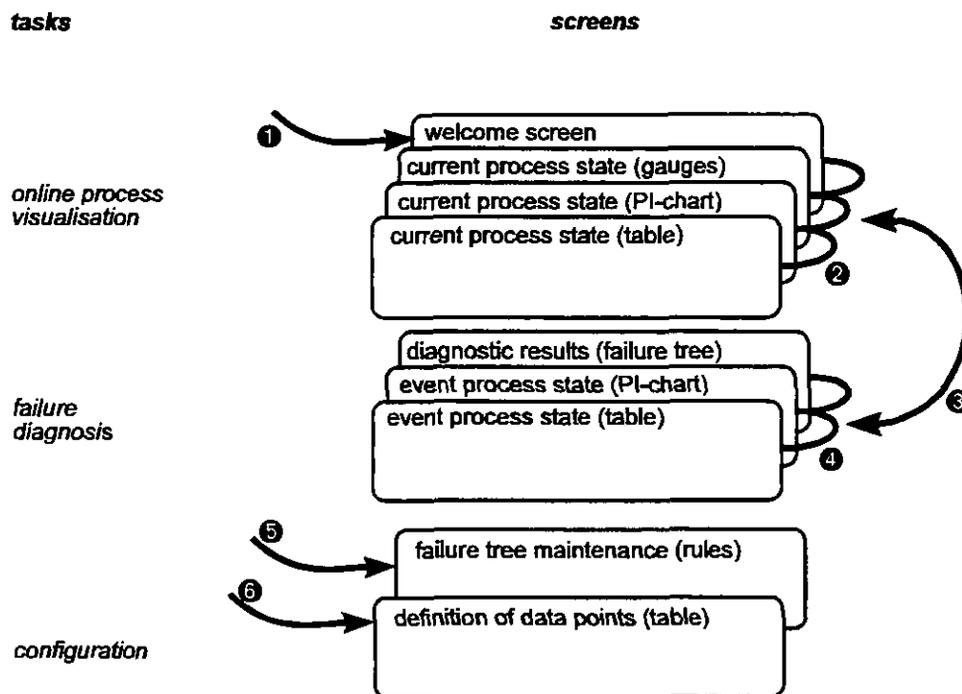
Fig. 4: Sequence of screens

The users enter EnDEx ❶ (Figure 4) in a welcome screen. From here they can choose between the different screens ❷ of the on-line process visualisation task or switch to the failure diagnosis task ❸. Within the latter they can get the different screens in the same way ❹. The configuration task of EnDEx has to be accessed from Excel ❺ or Nexpert ❻ directly.

3.3 GENERAL SCREEN LAYOUT (BACKGROUND)

Besides following the Windows conventions for screen design, a background screen layout (see figure 5) for all screens was designed in order to give the user always the same command and display elements at the same location.

The *title section* explains the topic that is shown in detail further down in the display section. On the right hand side in the *controls section*, all the controls are located to get to the other screens, other tasks or to get further help on the current topic. All the controls and many of the variable screen elements (e.g. icons, PI-charts) of the *display section* are featured with status text that is indicated in the status bar section whenever the mouse is located above that control. The *status bar* also shows internal states of the program such as "loading", "data acquisition", etc. A minimum of optional commands for the mature user are provided in the *menu bar*.
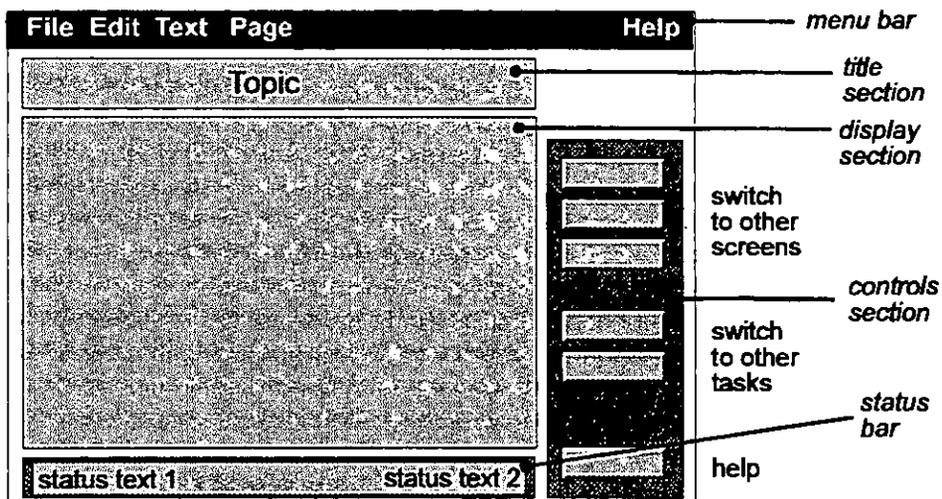
Fig. 5: Screen elements and sections

3.4 WELCOME SCREEN (=TITELBLATT)

Clicking on the welcome screen's big logo, some information pops up about the systems functions and handling. Further information is displayed on pressing the "introduction" (=Einführung..) button located at the lower right corner. By clicking the button "to gauges" (=Zu Anzeigen) in the control section, the user gets to the screen "current process state (gauges)"

Fig. 6: Welcome screen

3.5 ON-LINE: CURRENT PROCESS STATE (GAUGES) (=ANZEIGEN)

Practitioners like gauges of the main process states of a plant. The position of scales can obviously easier be remembered than an absolute numerical value.

Fig. 7: Gauges show some important values of the process

24.02.1996

Up to twelve gauges can be configured, each with labels for name and unit of the measurement as well as for the displayed and the critical range (red zone). In the top right gauge in Fig. 7 the compressor pressure (=p_HD) currently shows 13 bar with a critical range beginning at 18 bar. The needle's position is refreshed every 10 seconds. In ToolBook a gauge can be defined as an instance/object of a custom programmed class. Drawing, positioning and updating is thus performed at class level.

By clicking the button "to PI-chart" (=Zu Anlage) in the control section, the user gets to the screen "current process state (PI-chart)"

3.6 ON-LINE: PI-CHARTS AND PROCESS DATA TABLE (= ANLAGE)

Fig. 8: Process chart and measurement values

To overview the total process behaviour of a plant a process and instrumentation chart is useful. A complex plant can be divided into several subsystems. The screen should not be overloaded with labels for the different measurements. More than 20 labels are not recommended on one screen when using a standard VGA screen and a minimal font size of 8 pt. In order not to cover the process chart with the various labels, the user can show or hide it by selecting the corresponding checkbox at the lower right of the controls section. Labels are provided for name (= Name), value (= Istwert) and unit (= Einheiten) of every measurement. Special labels for the components description (= Kompos) can also be popped up.

Double-clicking on one of the minimised charts in the bottom of the boosts it up and places it in the centre of the display section.

In fig. 9 the two blocks left are just measured (= gemessen) process values, whereas the right block shows some calculated (= berechnet) values. The table view of the process values can additionally show the threshold values of each measurement.

Fig. 9: Process data table

3.7 DIAGNOSIS: FAILURE TREE (= FEHLERBAUM)

Fig. 10: Diagnosis and failure tree

The failure tree screen is explained based on a high pressure alarm case.

The display section of the screen " failure tree" is divided into four subsections:

24.02.1996

107

1. *Failure Diagnosis* (= Störungsdiagnose)

   Displayed are

   - name and location of the plant (= Anlage),
   - time of occurrence (= Zeit),
   - priority of the alarm (= Priorität) and
   - number of diagnosed causes (= Anzahl Diagnosen).

2. *Failure tree* (= Fehlerbaum für Alarm)

   The failure tree is described in section 4.3. Each node of the tree corresponds to a hierarchical failure section.

   - Within the title bar of that subsection the name of the alarm in abbreviated form; here high pressure (= HD) is indicated.
   - The failure tree picture dynamically reconstructs the tree as processed and evaluated within Nexpert-Object. The first level of diagnosis is reserved for a plausibility check of the alarm. The following levels represent the hierarchical failure sections (see 4.3). Each node of the failure trees changes its colour according to its diagnostic state (bright = true; dark = false; very bright = not evaluated).

3. *Explanation of the node* (= Erklärung des Knotens)

   On clicking one of the buttons in the failure tree the corresponding failure section *explains* itself in this scroll window:

   - The first item covers the *type of the node* (starting node, intermediate or "else"-node and final node) and the *name of the failure section*. In our case we have an intermediate node (= Fehlerbaum-Mittelknoten) called "heating circuit mixing valve fully open" (=offen).
   - The next item indicates the *name of the test* (= Test) of that failure section or end diagnosis. The test "mixing valve in position mixing" (=HKr_Mischventil Beimischung) was carried out.
   - The following item gives a verbal *explanation of that test* (= Testerklärung): "Heating water temperature before condenser {=49°C} is higher than return temperature {=37 °C}" (= T_Heizwasser vor Kd {=49 °} ist höher als T_Rücklauf)". In order to support the understanding of the user all values of the relevant process states are included in brackets.
   - The last item covers the *explanation of the node* (=Knotenerklärung): The effect of the faulty section on the neighbouring sections or the overall system: "An open mixing valve rises the condenser inlet temperature and thus leads to higher pressure" (= Weil das Mischventil offen ist, wird die Temp vor Kondensator noch erhöht).

4. *Correcting actions* (= Abhilfemassnahmen)

   - Diagnosed cause (= Diagnose): "Heat transfer of the condenser water side restricted" (= Kd wasserseitig behindert)

- Maintenance domain / specialist (= Fachmann): "Heating or refrigerating technician" (= Kälte- / Heiztechniker).

- Correcting action (= Abhilfe): "Eliminate fouling sources on the heating water side; clean condenser"(= wasserseitige Verschmutzungsherde eliminieren, Kondensator reinigen).

In ambiguous cases several possible causes with correcting actions are displayed one after another. This gives the user the possibility to continue the failure isolation himself. In our case another diagnosed cause was "heating circuit return temperature high" (= HKr_T_Rücklauf_Heizgr_zu hoch). For further information the user has to scroll the window downwards.

The described case and node is corresponding to node 2.5.1 in the failure tree of table 3.

No.	Node	Test	
Start	high pressure	high pressure ALARM	
1	false alarm	complementary	K
2	condenser side: pressure to high	$p\_HP > p\_HP_{max}$	
2.1	condenser on water side inhibited	$A^*k\_Con < A^*k\_Con_{normal}$	H
2.2	condenser pressure elevated by non-condensible. gas	$\Delta T\_Con\_subc > \Delta T\_Con\_subc_{norm}$	
2.2.1	wrong refrigerant	$\Delta T\_Ev\_superh < \Delta T\_Ev\_superh_{norm}$	K
2.2.2	air in condenser	complementary	K
2.3	congestion of refrigerant in the condenser	$L\_Coll > L\_Coll_{max}$	K
2.3.1	flow in pipes inhibited	$\Delta T\_Ev\_superh > \Delta T\_Ev\_superh_{norm}$ (U normal: LP-alarm)	K
2.3.2	overflow of refrigerant	complementary (U especially after service)	K
2.4	heating cycle: flow rate to low	$\dot{m}\_Con < \dot{m}\_Con_{set}$	
2.4.1	pump output low	$\Delta p\_ConP < \Delta p\_ConP_{set}$	
2.4.1.1	pump broken down	$\Delta p \cong 0$	H
2.4.1.2	pump defective / cavitated	else, especially when old	H
2.4.1.3	air in pump	else, especially after service	H
2.4.2	flow resistance in pipes to high	$\Delta p\_ConP > \Delta p\_ConP_{nominal}$	
2.4.2.1	throttle badly adjusted	complementary (especially after service)	H
2.4.2.2	condenser clogged	$A^*k > A^*k_{set}$ (see 2.1)	H
2.5	heating cycle: return temp to high	$T\_bCon\_H > T\_ret_{max}$	
2.5.1	mixing valve still open	$Pos\_V \neq 0$ or. $T\_bCon\_H > T\_ret$	
2.5.1.1	badly controlled	command controller $\neq 0$	R
2.5.1.2	mixing valve blocked	command controller $= 0$	H
2.5.1.3	stop node to 2.5.1.n	complementary	
2.5.2	T_ret from storage tank to high	$T\_ret\_St > T\_ret_{max}$	
2.5.2.1	Thermostat storage tank , others.	visually	R
2.5.3	T_ret from heating to high	$T\_ret\_H > T\_ret_{max}$	R
2.5.4	stop node to 2.5.n	complementary	
2.6	stop node to to 2.n	complementary	

Tab. 2: Sample failure tree for high pressure alarm with tests.

3.8 DIAGNOSIS: PI-CHARTS AND PROCESS DATA TABLE

The screens are identical to the ones of the on-line task. But the displayed values are corresponding to the state of the event.

3.9 CONFIGURATION: FAILURE TREE MAINTENANCE (RULES)

The maintenance of the failure trees mainly takes place within Nexpert Object.

a) The order of failure sections and the contents of the tests are represented as backward-chaining rules and forward-chaining contexts. Any changes are performed in the rule editor.

b) In special cases it is necessary to change the type of a node. This operation can be carried out by simply changing from one class membership to another.

c) The test thresholds are not stored in Nexpert; they have to be set in a table (see next sub-section).

3.10 CONFIGURATION: DEFINITION OF DATA POINTS

The blackboard of EnDEx is realised in MS-Excel.

- The configuration spreadsheet contains several tables. The main definition table holds name, value and unit of each data point.

- For each failure tree a separate table is set up containing the threshold values for the various tests.

4 HVAC DIAGNOSTICS OF ENDEX

4.1 PLANT CHARACTERISTICS AND MAINTENANCE

HVAC plants are all quite similar in their general set-up, but very different in detail. It's stated that each plant is a "prototype" made up by standard components from different vendors. The components are linked by pipes or air ducts incorporating many different primitive elements such as throttles, fire dampers, valves etc. The status of these devices is normally not monitored.

The plants are equipped with several *security devices* that switch it off in emergency situations. They generate alarm messages that can serve for failure detection. *Sensors* are normally only used for control purposes, energy measurements and for security functions.

Expert interviews have shown that in practice only a limited number of alarms occur, but with many different causes. Failures leading to a breakdown are in most cases single faults.

The *maintenance actions* are undertaken by specialists from the *domains* of

- heating,
- refrigeration,
- air conditioning,
- control systems and
- the electrical part.

In future, each service group maintaining DDC controlled plants will be equipped with portable PCs to access the plant status or the information from a central service station by means of a standard modem link.

4.2 PRECONDITIONS

Preconditions and effects on the reasoning of EnDEx:

1. Since no physical interaction is possible in the case of remote access the *order* of questions in the diagnostic process needs not be taken into account.

2. For safety reasons no active tests have up to now been allowed by the customers. A remote diagnostic system has to come out with a *passive data* acquisition. However, the optional input of human observations should be possible.

3. A simple failure detection is implemented in the plant *(security switches)* or in the DDC-controller (detection of no-start, of exceeded comfort intervals etc.).

4. Most parts of the HVAC-plant can be treated quasi-*stationary*. Dynamic models are possible but not necessary.

5. The potential failures or at least failure groups can be foreseen. No failure synthesis has to be performed.

4.3 EXPERT SYSTEM REASONING

A special approach is made to the deduction of the causes. It is based on the fault tree analysis known in safety technology. For its diagnostic use, however, the reasoning direction and the resolution mechanisms have to be changed.
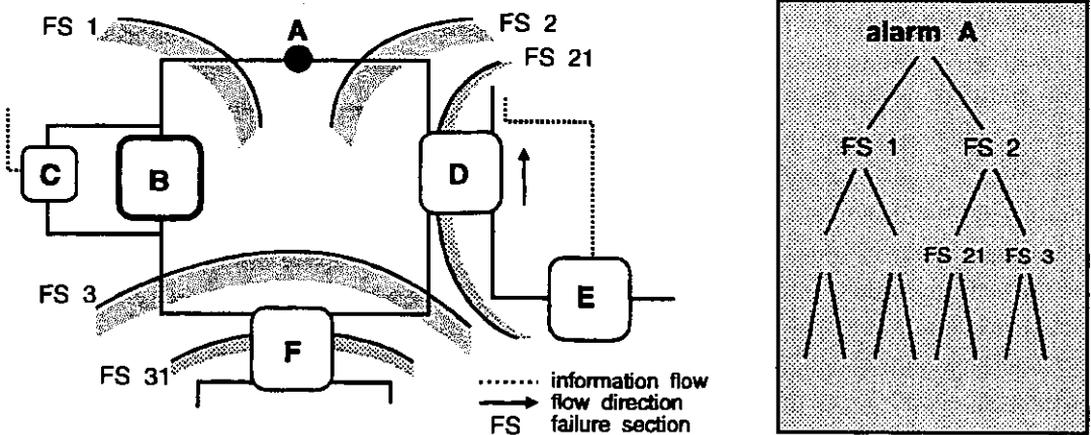
Fig. 11: From plant structure and component function to the failure tree.

The reasoning process follows a failure section isolation scheme. Starting from the place of the failure detection (e.g. alarm from the high pressure switch) the search goes on to the neighbour failure sections (FS) (see Figure 11). This neighbourhood is defined by information, mass flow and energy transfer. According to the experts experience, certain FS's can be dropped if not relevant, or directly sub-divided if no test for that FS can be found.

Each suspicious FS has to prove itself by a *test*, based on an expected failure or function behaviour compared to the actual state. A test may be performed on relevant and robust process features that are derived from direct sensor measurements, controller states or characteristic values calculated with analytical or empirical models.

From one established FS, another set of sub-FS is considered and tested until an end hypothesis can be found or until no sub-FS can be established. The last mentioned condition is important in order to draw conclusions even if no end hypothesis can be found. This leads to a more "graceful degradation" than the one of other systems.

4.4 DISCUSSION

The proposed expert system method comprises both *heuristic* and *model - based* diagnosis. The structure as well as the tests of the failure tree are strongly derived from a mental or described model of the plant structure and failure behaviour. But the form is still free allowing heuristic knowledge to be embedded, which is recommended to take advantage of the a priori knowledge of experts e.g. in order to exclude theoretical cases that never occur in practice.

5 KNOWLEDGE ENGINEERING METHODOLOGY

This section describes the knowledge engineering methodology used to generate the failure tree. It is derived from practical experience and from [1] to [5]. The full methodology is explained in [6].

5.1 ABOUT KNOWLEDGE

According to Table 3 the diagnostic task incorporates reasoning over various accessible knowledge resources with different orientation and type.

Resource	Type	Orientation	Contents
expert knowledge	informal	failure-/ function oriented	- a priori knowledge about weak links of the plant
expert strategy	informal	diagnosis oriented	- diagnostic strategy - shortcuts of the diagnostic process - simplifications
model - based knowledge	formal	function oriented	- plant scheme - component functions / processes - control strategy
case based knowledge	formal	diagnosis oriented	- data form diagnosed cases - failure statistics
state data	formal	none	- data of the actual operating point including temperatures, pressures, energy etc.

Tab. 3: Different knowledge resources can be accessed within the diagnostic expert system.

A major challenge of a diagnostic system is the integration of these different resources. In contrast to other fields, the real component behaviour is poorly known. As the components are quite cheap and the competition on the market is tough, no manufacturer is ready to publish more component behaviour data than necessary for the layout at a nominal operating point.

This is one of the main reasons why simply a classic model - based approach cannot be applied.

For a remote diagnosis the data of the plant state have to be gained from sensors and controller-states only.

5.2 EXPERT INTERVIEWS AND KNOWLEDGE STRUCTURING

A key issue is the elicitation and representation of the knowledge. As a result of that work, the diagnostic knowledge should be compiled into the form of explained *failure trees*.

A clear but open methodology helps the novice knowledge engineer to get a complete diagnostic tree structure out of the plant scheme, out of qualitative knowledge of the component functions and out of expert knowledge (see Figure 12):
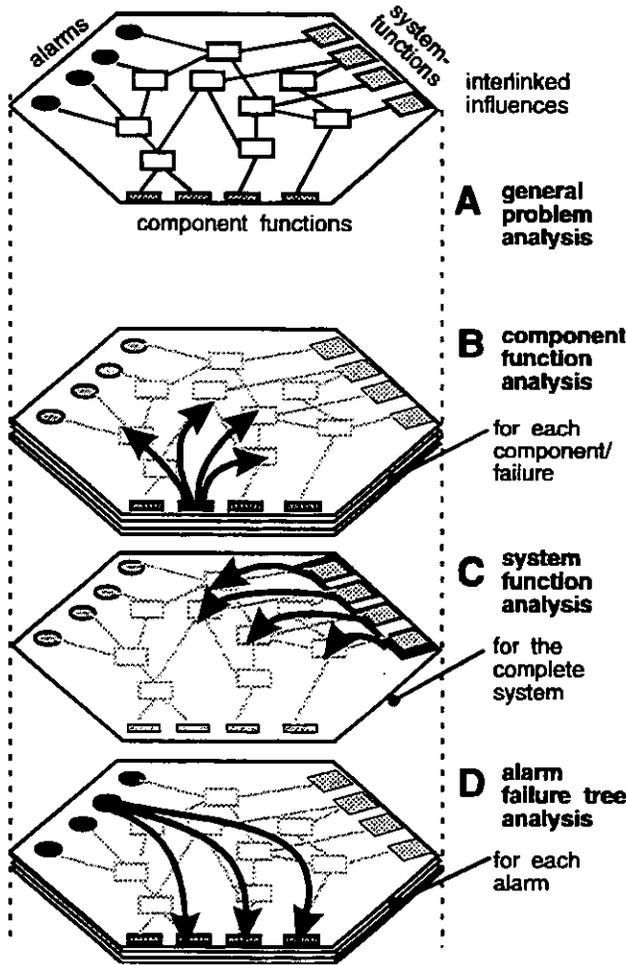

Fig. 12: Multiple approach to the failure analysis.

A. **General problem analysis** (together with a maintenance engineer): Overview of the functions and control strategies of the plant and identification of the weak points within its subsystems.

B. **Analysis of the component behaviour** concerning functions, failure modes and its local consequences (by the knowledge engineer).

C. **Analysis of the system tasks** (knowledge engineer with expert): Building up a functional hierarchy for the main tasks of plant and subsystems.

D. **Failure tree analysis** (together with a domain expert): With the acquired knowledge the failure tree can be set up. Tests have to be found for each decision. Optimisation of the tree.

E. **Input of the failure tree** into a diagnostic program, testing of the system (knowledge engineer).

Discussion:

In contrast to the approach of some existing diagnostic tools the above described failure tree does not correspond to a hierarchy of functions but rather to a topological structure according to the components in the plant.

6 SUMMARY AND DISCUSSION

With the demonstration prototype of EnDEx the following issues could be shown:

- EnDEx as an open framework is able to embed several different tasks to run simultaneously.
- A uniform graphical man machine interface facilitates the handling of the different tasks.
- Common data can easily be exchanged between the different applications by means of a shared blackboard. This is accessed by the Windows dynamic data exchange facilities (DDE).
- For cost-effective prototype development, a PC-based solution with standard software can be chosen. Limitations may arise in case of real-time-tasks. Remote access to the plant can be performed with standard modem links and remote access or remote control software.
- The on-line process visualisation task of EnDEx comprises the indication of the process state in PI-charts, data tables and in the form of gauges.
- The diagnostic task of EnDEx comprises the various indication of the process state at the time of the event. The data is processed in the expert system. The results are displayed in the graphical form of an failure tree. On selecting a tree node more information about the meaning of that specific failure section, about the tests and the actions to be taken are given.
- In order to develop EnDEx from prototype to a commercial product, the following functions would have to be realised: Continuous data acquisition; alarm history management; replay and re-evaluation of former failure situations.

The second part of the paper has shown the use of a knowledge engineering methodology:

- The failure tree analysis can be carried out from an analysis of the component behaviour (cause-effect-relations) and an analysis of the system tasks (hierarchy of the system functions). This method incorporates model - based as well as heuristic approaches.
- The model - based part helps the knowledge engineer to achieve a completeness of possible causes whereas the heuristic style of rule implementation offers a great flexibility. Several different knowledge resource types can be embedded. The use of fault descriptions instead of detailed component behaviour models drastically reduces the development time over the one with a model - based approach.

24.02.1996

- The drawback of the methodology compared to fully model - based approaches is the amount of paperwork prior to programming. Also after a topological change of the plant, a mental re-engineering phase has to be carried out.

7 REFERENCES

[1] P. W. Brothers; Knowledge Engineering for HVAC Expert Systems, ASHRAE Transactions 1988, V.94, Pt.1

[2] B. Chandrasekarean, Sanjai Mittal; Deep Versus Compiled Knowledge Approaches to Diagnostic Problem-solving; from: Int. Journal on Man-Machine Studies (1983)19, p.425-436

[3] Charles H. Culp; Expert Systems in Preventive Maintenance and Diagnostics; ASHRAE Journal, August 1989

[4] Dan Diaper (Ed.); Knowledge Elicitation: Principles, Techniques and Applications; Ellis Horwood/Wiley, 1989

[5] Anna Hart; Knowledge Acquisition for Expert Systems, McGraw-Hill, 1986

[6] Daniel E. Maurer; Expertensystem für die Störungsdiagnose an wärmetechnischen Anlagen von Gebäuden, Dissertation 9872 /Forschungsbericht Nr.14, Labor für Energiesysteme, ETH Zürich, 1992 (in German)

[7] Daniel E. Maurer, Peter Neuenschwander; An Expert System Failure Diagnosis Methodology for HVAC-Plants, Tooldiag '93, Int. Conference on Fault Diagnosis, Toulouse, France, 1993

ACKNOWLEDGEMENTS

C:\DANIEL\IEA\ENDEX06.DOC

24.02.1996

BOFD SYSTEMS IN TOTAL AC SYSTEM AND THERMAL STORAGE SUBSYSTEM

N. NAKAHARA, Professor

Nagoya University

Nagoya, Japan

Abstract

The HVAC system is one of the most complicated systems whose subsystems interact each other, so that its performance can not be optimized except for being considered from the total system point of view. Another important idea is that faults take place anywhere in the production and operation stage, so that life cycle commissioning, e.g. the life cycle BOFD, is requested in the future. The present paper discusses the method of total system BOFD from the life cycle point of view. In case of the water thermal storage system, most of the design and control faults appears as the abnormal temperature profiles in the tank. An bottom-up process for the thermal storage faults utilizing the symptoms in the temperature profiles and other process variables is also discussed at the latter part of the paper.

1 BOFD IN TOTAL SYSTEM

1.1 TOTAL HVAC SYSTEM

The flow diagram showing four kinds of subsystems in a HVAC total system, the heat generating plant, air and water transporting, air handling unit and building model is shown in **Figure 1,** to which the energy source and BEMS may be added.
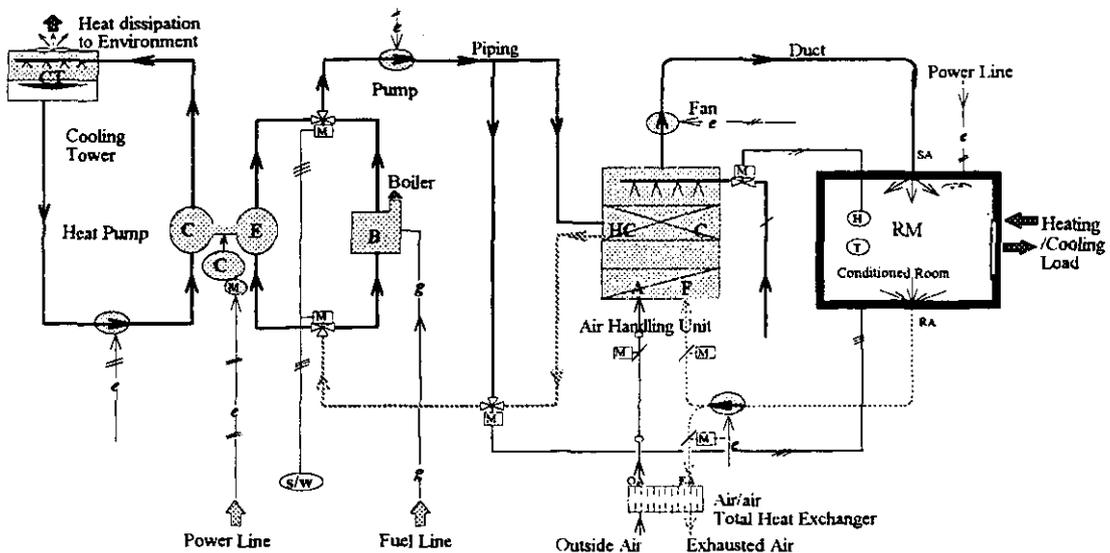
Figure 1 Flow Diagram of HVAC Total System

It is clearly seen from this figure how additional energy consumption is needed to drive the system and then reduces to thermal energy to warm up our environment. Building Optimization, or BO, is to minimize energy consumption without sacrificing, or rather to say optimizing, human environment.

1.2 THE BASIC STRUCTURE OF PRIMARY ENERGY CONSUMPTION

Each item of the Building Services Engineering, that is, air-conditioning, sanitary engineering, lighting, power supply or transport systems, has the load which results in the 'secondary energy consumption' evaluated by the final style of energy, such as electricity, and then in the 'primary energy consumption' which is evaluated as the oil consumption.

The **figure 2** [1] shows how an initial load causes a certain amount of energy consumption in the course of energy transfer, energy transport and pumping energy from a low temperature to a high temperature. The philosophy of energy conservation technology is clearly shown in the diagram, that is,
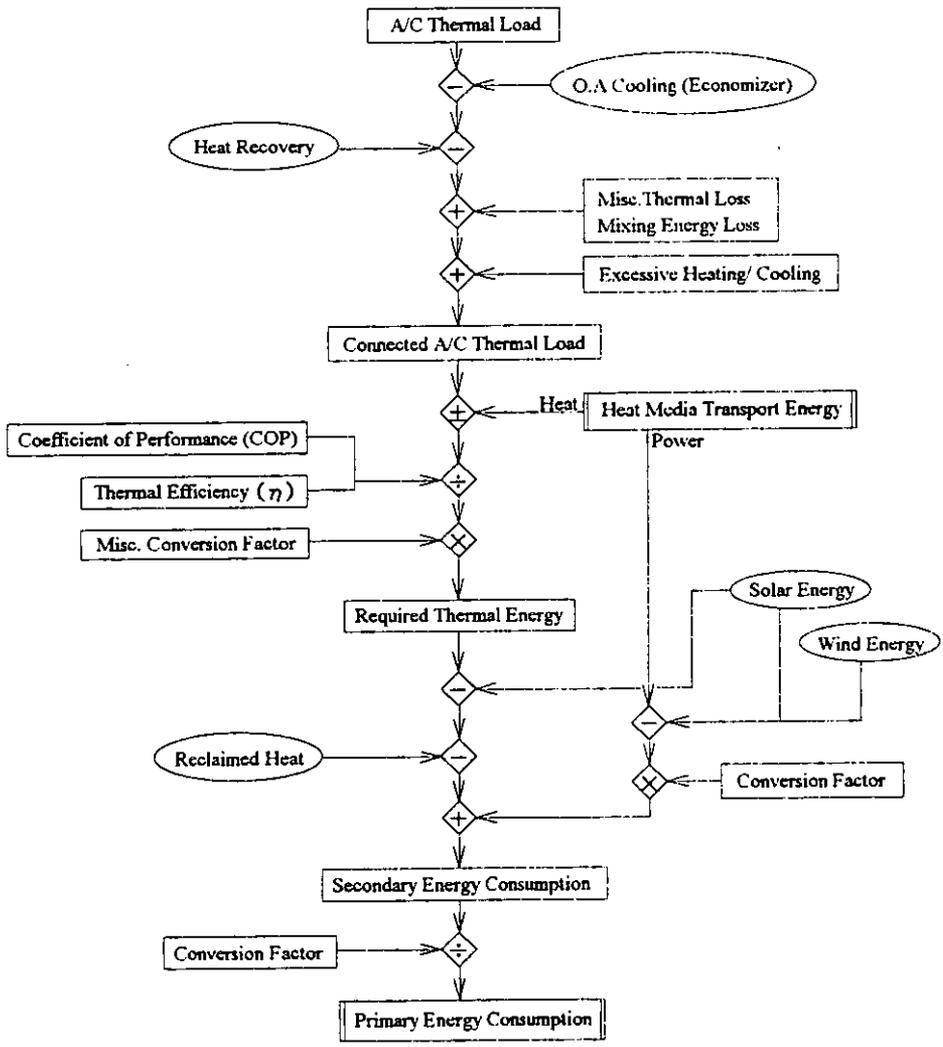
Figure 2 Basic Structure of Orimary Energy Consumption

1) to reduce thermal loss

2) to reduce mixing energy loss[2]

3) to recover heat and/or power

4) to make use of natural energy such as solar energy, wind energy and temperature difference from room temperature such as well water, surface water and atmosphere.

5) to reduce air and/or water transport energy

6) to maximize the efficiency of energy conversion

7) to maximize the COP of heat pump

8) to minimize harmful effluent gas on the global environment

1.3 THE ROLE OF OPTOMAL DESIGN AND BOFD

The Building Optimization can only be performed after the optimal design based on minimum energy consumption with maximum environmental satisfaction is realized, and after an optimal control for each subsystem based on a practical performance function with constraints and FDD for system components is accomplished as well.

Any original design necessarily includes faults or inoptimality. Also, no system has ever proved to operate satisfactorily without any adjustment of design factors as well as various parameters in control algorithms. All the components deteriorate, reduce their performance and offset their characteristic values from their initial values.

Even if no deterioration occurs, control parameters, which should guarantee satisfactory environment and life cycle of actuators, must be optimally tuned in the course of load change. Thus, the BO of total HVAC system is a summed up result of the BOFD of each subsystem.

1.4 THE WAY TO SUCCESSFUL BOFD OF THE TOTAL SYSTEM

The way of BOFD in the total HVAC system will be established as follows.

1.4.1 Around the time of Completion

1) Review the design documents and understand the philosophy of design, the targets of energy and environmental performance that the designer guarantees to the orderer.

2) Review the commissioning documents and understand how well or how poorly the system satisfies the design goal at the time of completion. If there are no commissioning documents, make use of reports of system adjustment by constructors and completion documents of components manufacturers.

3) Review the operational manual which must have been prepared by designers for fundamental design concept and by control manufacturers for detailed specifications. If there are no operational manuals prepared by designers, it should be requested to submit them, because they must have been obliged to submit it.

4) Make sure if the full training course for operators on BEMS/BOFD operation at the manufacturers training site is

prepared before delivery.

Commissioning using emulator/tester at the factory site is desirable. And understand the methods and the objects of control algorithms and check if the algorithms are still appropriate at the time of completion. Status of occupancy may have changed during the course of construction.

1.4.2 During one to two year after completion

Two years of experience are necessary to identify the HVAC system performance and to accumulate data for learning process for BEMS/BOFD. Two years will be an important period from the viewpoint of guarantee of the system in general. The following are the must items.

1) Make sure if each function of BEMS works satisfactorily.

2) Make notice of how control functions and algorithms work well and realize the initial targets of the controls.

3) Watch the movement of actuators if there are no hunting movements at a certain time or season.

4) Compare the energy performance with the reference values and the goal of design.

5) Analyze the environmental claims from occupants. Check the conceived causes of the claim from the following.

a) different load condition from designed value

b) insufficient adjustment of system components during construction.

c) malfunctioning of system components

d) set point error

e) insufficient capacity at the design stage

f) improper zoning design

g) excessive requirements of occupants over the design value

h) fundamental problems included in the conventional design, such as neglected radiative effect in comfort design/control.

6) Analyze the optimality of optimal control based on the performance function and constraints.

7) Enrich the knowledge database from operators' experiences.

8) Implement BOFD algorithms in combination with all items above described.

Operators should not hesitate to ask the designers, constructors and manufacturers of the system, if there are any doubts in the system design, adjustment data and component performances.

1.4.3 After one year

BOFD software described in this manual and the technical reports can be applied, if the system has these functions. In case of design and commissioning faults Chapter 6 of the Source Book should be referred.

For each subsystems as classified in Fig. 1, corresponding sections and technical reports will be helpful. However, it should be noticed that these methods are not sufficiently verified, so that operators should not fully depend on the original

algorithms, but consult with the vendors or consulting engineers who are responsible for the BOFD system.

Top-down approach for the total system is also described in the next chapter, BOFD in the thermal storage system. Watching thermal storage characteristic values may often lead to BO of total system, because the storage tank is the buffer of the faulty operation phenomena as well as the heat. The next chapter should be reffered.

1.4.4 Introduction of the new BOFD system to the existing system

BOFD system is under development for now. However, it is a very expecting technology for the future to realize both energy conservation and to guarantee environmental satisfaction during life cycle of the building. Cost/benefit analysis is advised to carry into effect in case of new introduction of BOFD system, referring the Section 5.5 of the Source Book.

2 BOFD SYSTEMS IN THERMAL STORAGE SYSTEM (TES)

2.1 FAULT TREE ANALYSIS WITH TYPICAL FAULTS AND CAUSE-EFFECT RELATIONSHIPS

The reference system and its typical faults has been shown in the Section 3.4 of the Source Book. More detailed explanation was reported in references[3]. Simply expressed system diagram is shown in Figure 3.

Cause-Effect relationships were listed in the reference[4] in which symptoms for implementation and relations among BOFD subsystems are also included. Several kinds of fault simulations will strengthen the qualitative knowledge by experts to include into knowledge database and further discover interactions among faulty phenomena. These knowledge lead a fault tree to diagnose the system and localize the point of target as the candidate of the cause of faults. In order to attain a successful result, try and error method will be applied and fault simulations during diagnosis may be necessary.
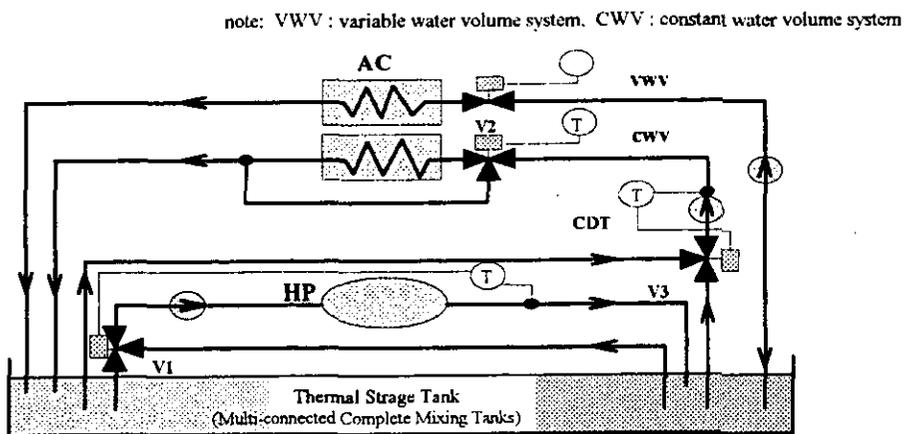

Figure 3 Simplified Reference System Diagram of TES

2.2 STORAGE EFFICIENCY AND IMPORTANT FAULTS

The storage efficiency[5] as defined by

$$\eta_s = \frac{Q_s}{\Delta\theta \cdot V_0}$$

plays an important role in BOFD. The Q_s is the heat to be stored, or the heat actually available, $\Delta\theta$ is the weighted mean temperature difference through the coils and V_0 is the volume of the tank water. The efficiency is largely effected by design factors such as temperature difference across the cooling coil, allowable temperature rise, or down in case of warm storage, system control methods and air-conditioning load profile. The estimation table of the storage efficiency for the multi connected complete mixing tanks based upon factorial effects by three levels of significant factors is shown in **Table 1**[5][6][7], which has been revised from the original table inthe bibliography[5].

Table 1 Significant Factors and Factorial Effects of Storage Efficiency

FACTOR (I)		LEVEL			FACTOR. EFFECT(FF)			INTERACTION (IR)				
(significant only)		(j)			levels			comb.	level of X	level of Y		
I	factor name	1	2	3	1	2	3	X×Y		1	2	'3
B	Min/Max Load ratio	0.8	0.5	0.2	2.85	0.94	-3.79	B×D	1	-5.92	0.18	5.73
									2	-0.30	-0.49	0.79
									3	6.22	0.31	-6.51
C	Yes/No of CDT for CWV	yes (2)	yes (1)	No	9.47	1.76	-11.23	C×D	1	-5.03	-0.32	5.35
									2	-1.06	-0.23	1.29
									3	6.09	0.54	-6.64
D	CWV Load Ratio to Total Load	0.2	0.5	0.8	3.63	0.20	-3.83	C×F	1	1.93	-1.29	-0.64
									2	2.49	0.35	-2.84
									3	-4.42	0.94	3.48
F	Limit Temp. Differential Ratio	0.4	0.3	0.2	7.19	0.76	-7.95	F×G	1	-1.01	-3.05	4.06
									2	0.96	1.03	-1.99
									3	0.05	2.01	-2.07
G	No. of Tanks	40	20	10	6.20	-0.21	-5.98	Storage	96.41	+ ΣFF		
H	Operate hours	0-24	18-12	22-8	5.50	-7.29	1.79	Effi-ciency		+ ΣIR	(%)	

note
B: ratio of Minimum load to Maximum load in a day at the peak load C: existance of Constant water Delivery Temperature control,
refer Figure 3 D: ratio load ratio of VWV system load to CWV system load F: ratio of allowable temperature rise of delivery tank to
design temperature difference through the coil at peak load G: number of tanks H: schedule of thermal storage operation

When a highly efficient system falls into faulty state due to malfunctioning of these significant factors for storage efficiency, it is easily conceived that the system may not sufficiently afford the air-conditioning load, even in case of

partial load, which must result in some kind of abnormal temperature profiles in the tank. This leads to applicability of pattern recognition method for BOFD with temperature profile analyses[8].

2.3 KNOWLEDGE DATA THROUGH FAULT SIMULATION

Thermal storage system simulation necessarily includes HVAC total system, because thermal storage tank is not used solely by itself, but must be combined with secondary HVAC system as well as primary heating/cooling plant. Therefore, the temperature transition profiles represent all the thermal incidents which are being realized in the whole system. In the same way all the faulty operation in the system must result in the abnormal profiles.

<div style="text-align:center">
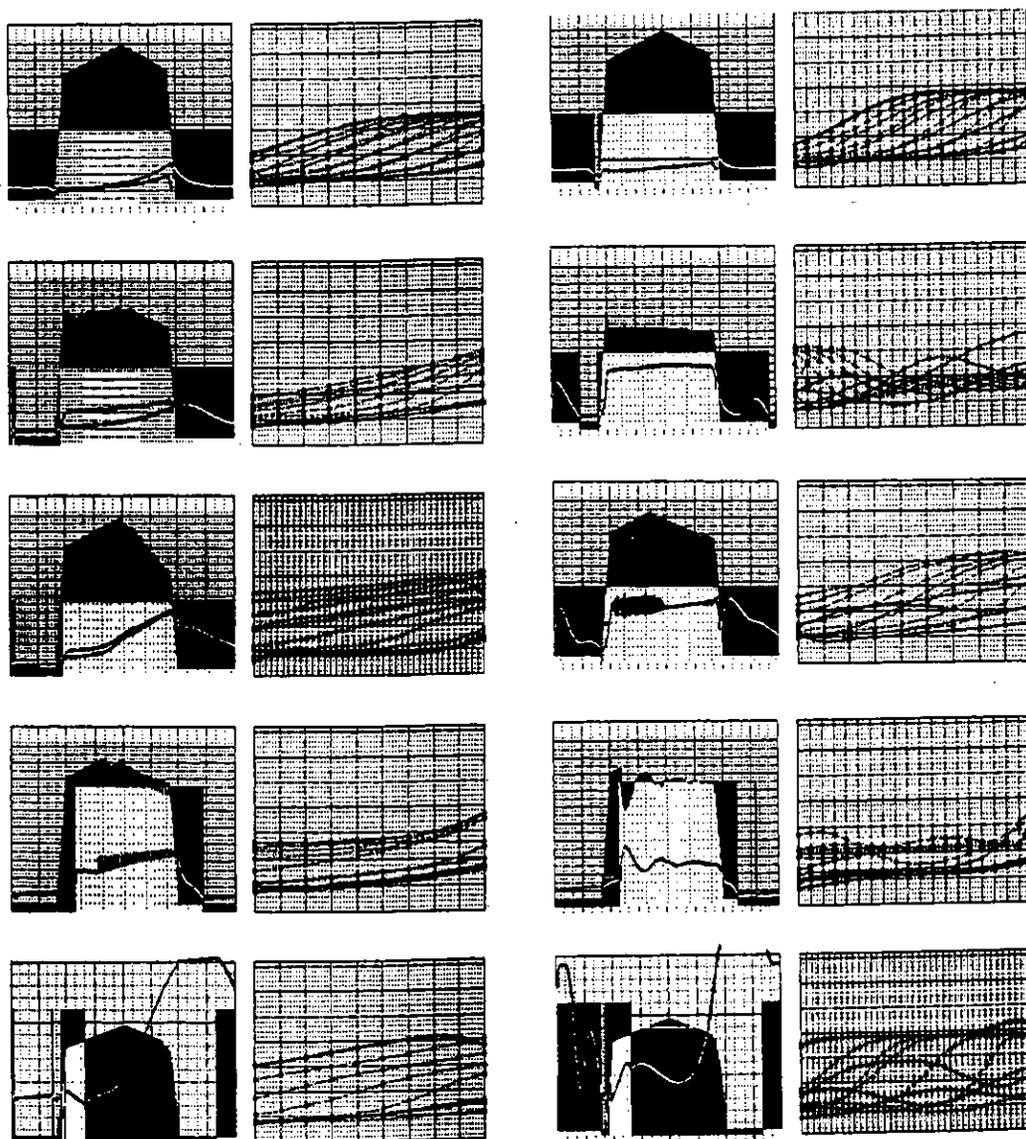
(a) Fault in V1   (b) Fault in V2 & V3

note: horizontal axis is time for the load profiles and the number of tank for the temperature profiles

Figure 4 Examples of fault simulation results, load profiles (left) and temperature profiles (right)

</div>

In the same manner as practiced in establishing estimation tables of storage efficiency using design of experiment [5][6], fault simulations with some typical faulty conditions[7] were performed and produce a lot of new knowledge for BOFD concerning to the fault symptoms and relationships between faults and abnormal temperature transition profiles[4][8]. Figure 4 shows examples of fault simulation results[4].

2.4 APPLICATION OF PHYSICAL MODEL TO INSULATION AND WATER PROOF DAMAGE

The application of physical model to the on-line fault detection caused by insulation damage and/or water proof damage was applied[9]. Physical model of the multi-connected complete mixing tanks was developed and two kinds of parameters, one representing the heat transference coefficient of the tank wall and the other representing the effective volume ratio, was introduced.

Normal values of parameters should be identified beforehand during normal operation. Fault simulation studies showed the method successfully applicable, while in the actual system it was clarified that lots of noises, such as measuring errors, sensor errors, data transmission errors and some manual errors in reading and copying should be eliminated beforehand.

2.5 TEMPERATURE PROFILES AND PATTERN RECOGNITION OF FAULTS

The temperature profile means the visual graph of the temperature distribution of the tank or tanks in the course of time. There are two types of expressions for the two dimensional temperature profile. Type A, shown as Figure 5, is the graph in which the horizontal axis is the time of a day and the parameter is the number of tanks. The other type B, as shown in Figure 4, has the horizontal axis as either the position of the tank or the number of tanks corresponding to whether the tank type is thermal stratified tank or the multi-connected complete mixing tanks[5][6]. The cool storage is supposed for explanation purpose in the following.
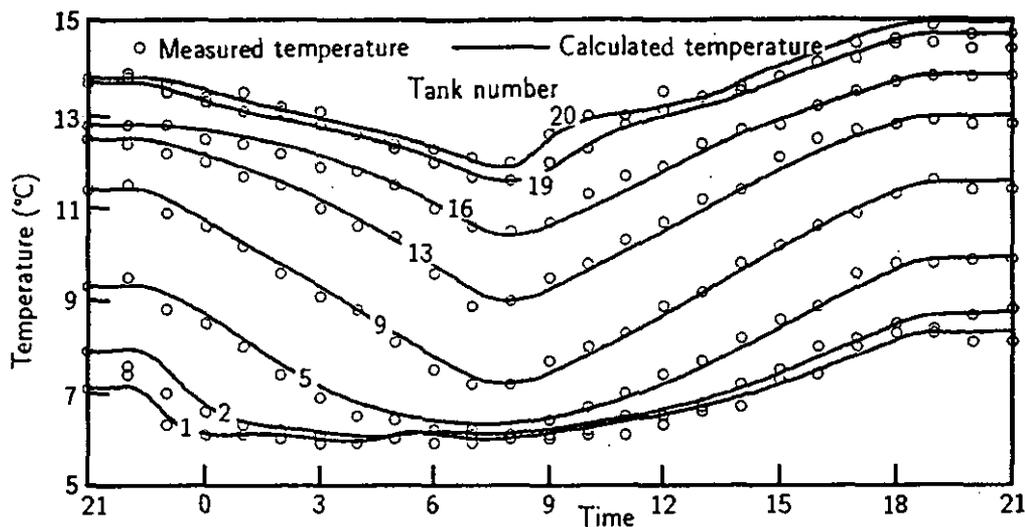
Figure 5 Temperature Profile Type A

The type A is used in BOFD with physical model for the insulation and/or water proof damage.[9] The type B is appropriate to recognize the macroscopic performance of the thermal storage tank and use for defining storage efficiency[5][6][7]. A critical value is the limit temperature rise at the coolest side of the tank which directly affects the cooling power of the coil in AHU. The characteristic parameters for BOFD in this profile are the phase, cycle and amplitude of the curbs. Different kinds of faults will result in different kinds of patterns with corresponding characteristic parameter values[8].

In case of nearly peak cooling demand the maximum temperature rise in the coolest part is an easiest and best parameter of symptoms to detect the existence of any faults. In the off-peak cooling demand the value does not always become abnormal, so that the pattern recognition method is useful for both fault detection and diagnosis. These characteristic values as described above analyzes directly in time domain, as shown in two dimensions, Figure3, as well as in three dimensions as shown in **Figure 6**, or in Fourier domain as shown in **Figure 7**[8]. Actually, mapping data using some kind of parameters are useful for FDD as shown in **Figure 8** just as used in the Fault Vector Direction Method[10].
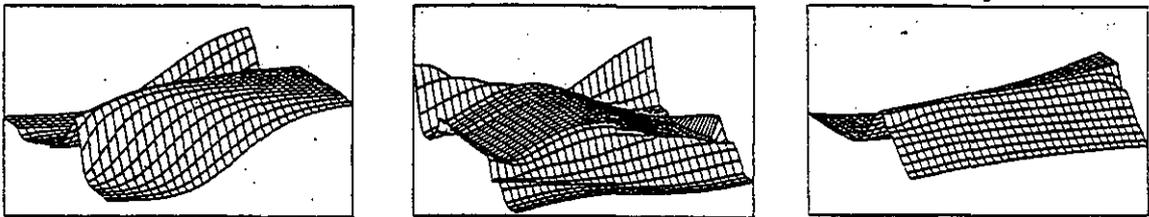
(a) three dimension graphs for a normal case   (b) three dimension graphs for a faulty case

note for axes: x:number of tank, y:time, z:temperature

Figure 6 Three Dimensional Temperature Profiles for the same case as in Figure 4
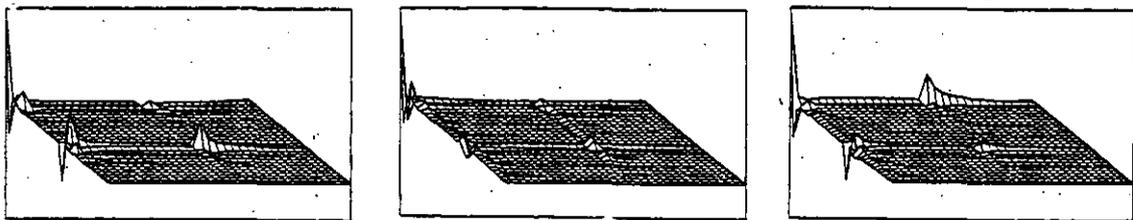
(a) Fourier domain graph for a normal case   (b) Fourier domain graph for a faulty case

note for axes: x: cosine components, y: sine components, z: Fourier value

Figure 7 Temperature Profile as expressed in Fourier Domainfor the same case as in Figure 4
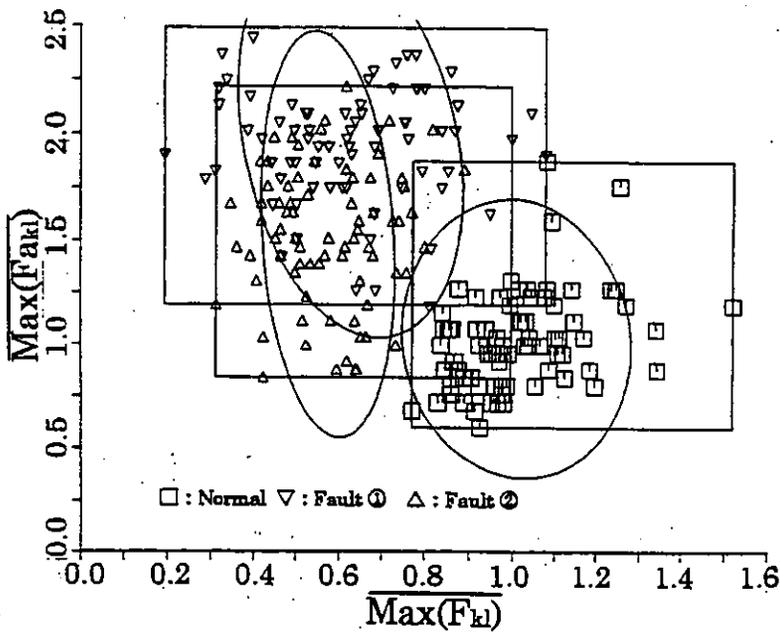
Figure 8 Example of Data Mapping on the Plane

2.6 BEMS/BOFD SYSTEM STRUCTURE

The BOFD system is to be combined with the BEMS, either existing or not, in practical applications. The **figure 9** shows an example structure of the BEMS/BOFD system[11].

Knowledge database consisting of various reference values, experts knowledge and updating database with statistically analyzed on line.

The main frame of FDD system consists of the software establishing FTA, cause-effect relationships, inspection manual and video display of condition charts. The FD software as described in Chapter 4 of the Source Book may be included either in the mainframe or outside the mainframe.

On-line data from the field are collected by the BEMS which includes data logging and analyzing software, alarming software, predetermined and regulatory and control software, optimization control software, maintenance software and, sometimes, the FD software.
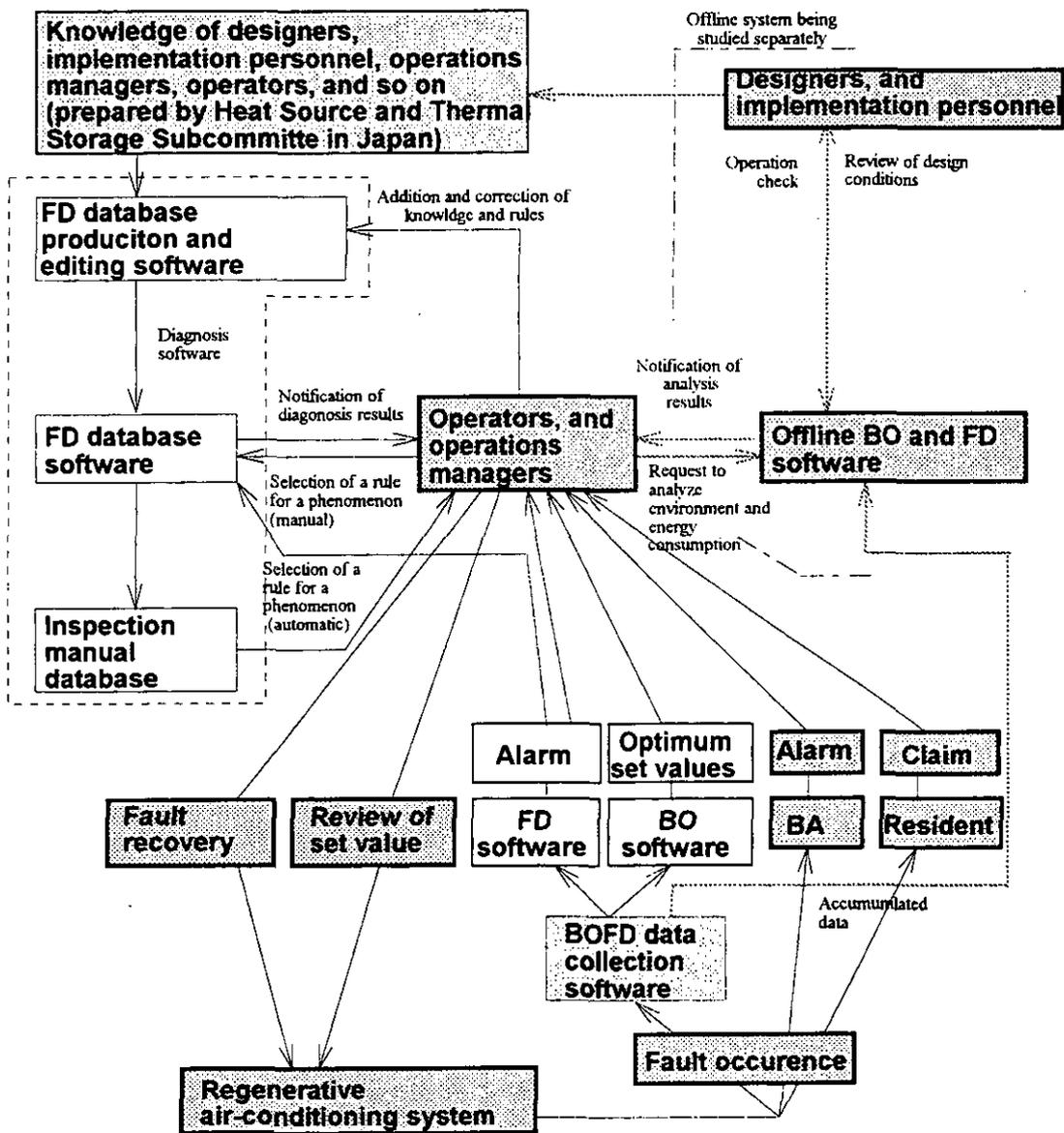
Figure 9   BEMS/BOFD Structure

References

[1]Nobuo Nakahara et al.: Planning of Building Services Engineering System, Shin-Kentikugaku Taikei, Vol.27, Shokokusya, 1982

[2]Nobuo Nakahara, Hisahiro Ito:PREDICTION OF MIXING ENERGY LOSS IN A SIMULTANEOUSLY HEATED AND COOLED ROOM:PART 1-PART 3, ASHRAE TRANSACTIONS, 99-1, January, 1993

[3]IEA Annex25: Building Optimization and Fault Diagnosis system concept, Oct.1993

[4]Nobuo Nakahara: Fault Diagnosis Techniques-Discussion on Mapping Data to Faults-, IEA Annex25, MIT meeting, Apr.1994

[5]N.Nakahara, K.Sagara: WATER THERMAL STORAGE TANK: PART 1-BASIC DESIGN CONCEPT AND STORAGE ESTIMATION FOR MULTI-CONNECTED COMPLETE MIXING TANKS, ASHRAE TRANSAC-

TIONS, 94-2, June 1988

[6]K.Sagara, N.Nakahara, M.Tsujimoto: WATER THERMAL STORAGE TANK: PART 2-MIXING MODEL AND STORAGE ESTIMATION FOR TEMPERATURE- STRATIFIED TANKS, ASHRAE TRANSACTIONS, 94-2, June 1988

[7]NOBUO NAKAHARA: FAULT ANALYSIS OF THE CONTROL VALVE MALFUNCTION IN HVAC USING THERMAL STORAGE SYSTEM SIMULATION, IEA Annex25, Tokyo meeting, Apr.1993

[8]N.Nakahara, M.Zheng, : STUDY ON FAULT DETECTION AND DIAGNOSIS OF THERMAL STORAGE CONTROL SYSTEMS WITH PATTERN RECOGNITION, Method Application Volume, IEA Annex25 Final Report, 1996

[9]K.Sagara, N.Nakahara: FAULT DETECTION IN THERMAL STORAGE TANK USING PHYSICAL MODEL, IEA Annex25, Method Application Volume, IEA Annex25 Final Report, 1996

[10]J.Li: Detailed procedure of fault detection method with some numerical examples, Annex25, Tokyo meeting, Apr.1993

[11]K.Kamimura, T.Miyazaki: Fault Diagnosis System and Demonstration Plan for a Hospital Thermal Storage System, IEA Annex25, MIT meeting, Apr.1994

The FTA System For Application To HVAC Systems.

Hideki Yuzawa,   Nikken Sekkei Limited

Masami Suzuki,   Shinryo Corporation

Hiroshi Takei,   Obayashi Corporation.

Harunori Yoshida,   Kyoto University

Abstract

This paper briefly describes the fault diagnosis method as developed by the authors for application to HVAC systems. The proposed diagnosis method is featured by the fact that it aims at improving the technical ability of the personnel in charge of maintenance/operation of HVAC systems. This method is intended to help the maintenance personnel detect the cause of improperly high room temperature by the aid of the " fault tree " formulated on the basis of the experts' knowledge. (The proposed diagnosis method is hereafter referred to as " the FTA system ".) Here, FTA means Fault Tree Analysis. For verifying the validity of " the FTA system " , some simulation studies were conducted by applying the FTA system to a number of faulty HVAC system models.

1. Introduction

As different types of HVAC systems are in use, different kind of faults naturally occur in them. Generally, it is not easy for an average HVAC system maintenance crew to detect faults in HVAC systems, and even if they are notified of a fault by the building's occupants, it is often difficult for them to diagnose the precise cause of such fault efficiently ; and it is in most cases more difficult for them to remedy such fault. Thus, all that most maintenance crews can do when they are notified that the temperature in certain rooms is excessively high is to reset the temperature for the problematic rooms to a lower degree. Various factors including not limited to equipment deterioration, control system malfunctions and internal heat generation exceeding the HVAC system capacity can cause excessively high room temperatures, Hence, it is important for the maintenance personnel to detect the real cause of the trouble out of numerous possible causes and eliminate the real cause quickly. To make this possible, the building maintenance crew who keep vigilant watch of the system should be able to make an appropriate diagnosis of the trouble cause quickly. The fault diagnosis method proposed by authors (i.e., "the FTA system") has two

main features : one is that various on-going measured values concerning the HVAC system are displayed on the control room screen in a graphic form easily comprehensible for the maintenance personnel ; and another is that the system is designed to enable the maintenance personnel to diagnose the cause of the problem by the use of the " fault tree " formed on the basis of the experts' knowledge.

2. The FTA system comprising a part of BOFDD

The hierarchy of BOFDD total structure and the FTA system is shown in Fig.1. [1]
Here, BOFDD means Building Optimization and Fault Detection and Diagnosis.
As you follow the chart vertically from below to upward, you will be able to see that the optimizing process of the objective system increases. As you follow the chart horizontally from right to left, you will be able to see the process through which the faults (or the anti-optimizing factors) are eliminated. The position of this system in BOFDD is shown in Fig.1 as " Fault Diagnosis". Once any fault has been detected , this system enables the maintenance crew to diagnose the cause of the fault utilizing the experts' knowledge and at the same time the system indicates to the crew a proper method   by which to eliminate the cause of the fault.
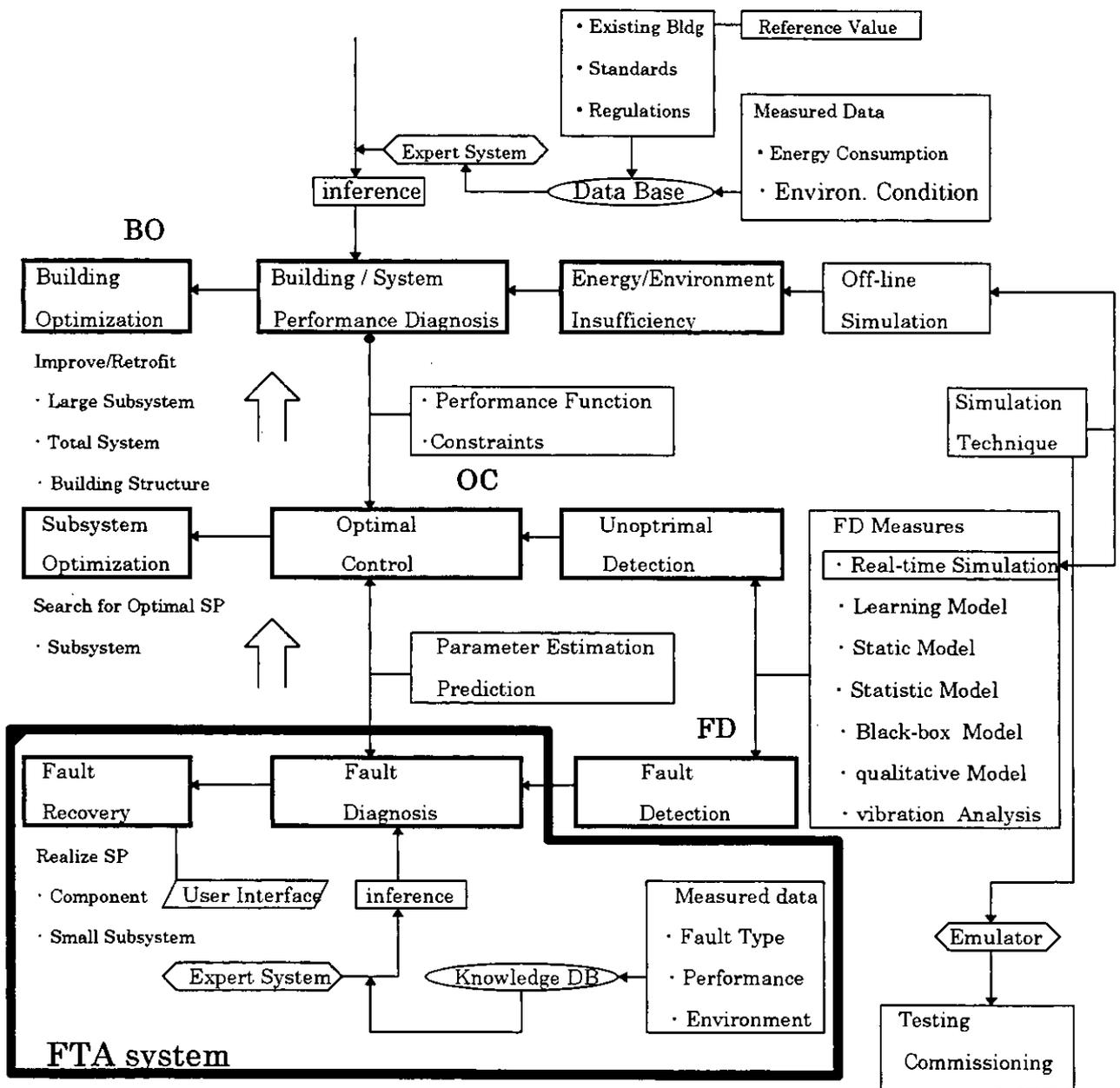
Fig.1 Hierarchy of BOFDD [1]

3. Description of the reference system

The reference system which was used in developing the FTA system is described in this section. If an object system is similar to this reference system, the FTA system can be used to diagnose a fault.

(1) Construction of the reference system

Construction of a reference system is shown in Fig.2.

This reference system is a VAV system without terminal reheating, and no preheating is included.
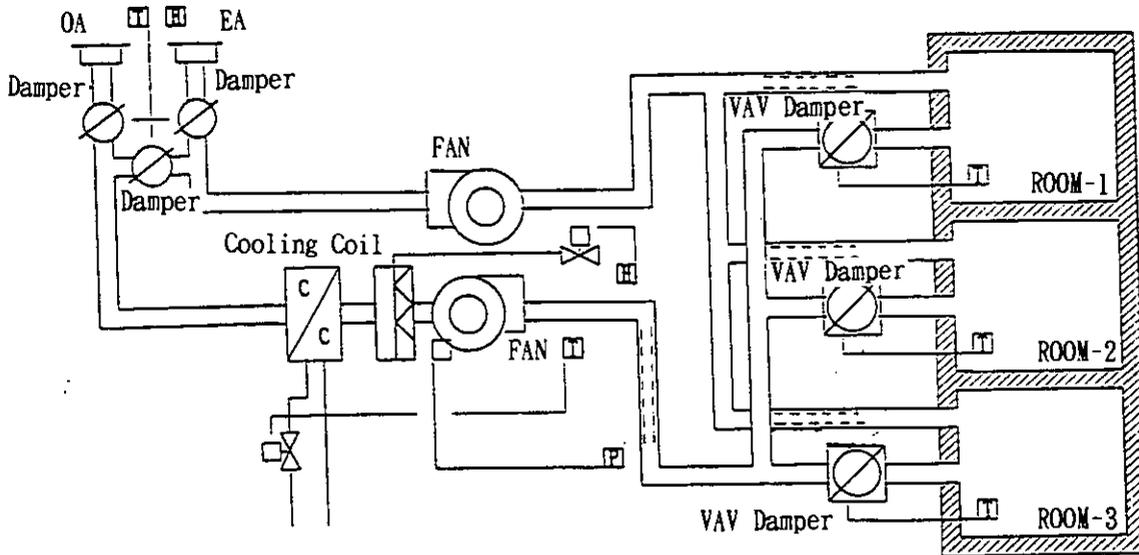
Fig.2　Construction of the reference system

(2) Control method

The HVAC system control method compatible with the FTA system is described below　and the design parameters of the model system are as shown in Fig.3 and　Table.1.

a) Control for the VAV damper opening

The VAV opening is regulated by P control. The room air temperature is set at 26° C, with a proportional gain of 0.5. The VAV damper opening changes within a range of 40-100% when the room air temperature is within a　$25\text{-}27\,^{\circ}$ C range.

b) Supply air temperature control method

The two way valve opening for the cooling coil is regulated by PID control to keep supply air temperature within 14.6-18.7 $^{\circ}$ C range.

c) Supply air volume control method

The number of revolutions (r.p.s) of the supply fan is controlled by the measured pressure inside duct. The pressure is set at 249 Pa at a 2/3 point of the total duct length from the entrance of the duct.
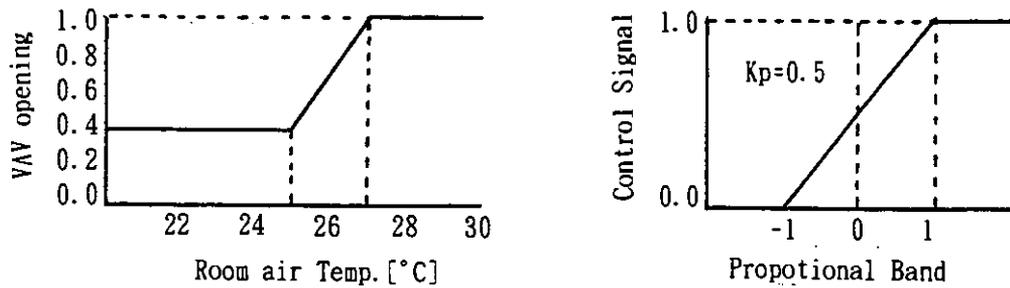
Fig.3 Control algorithm of VAV damper

Table.1 Design Parameters of Controllers

CONTROLER	Kp	Ti	Td
VAV	0.5	-	-
SUPPLY AIR TEMPERATURE	0.01	100(s)	0.001 (s)
FAN SPEED	0.20	200(s)	0.0005(s)

(3) Inputs to the FTA system

The information that are used for diagnosis is shown in Table.2.

Table.2 indicates the information that is usually completed within the local controllers and is not reported to the central supervisory machine. Such items of information include a VAV opening signal. The FTA system can be used more advantageously and more conveniently, if the amount of information required for diagnosis decreases. This, therefore, poses a problem which must be solved in future.

Table.2 Information by FTA system inputted

```
-MESUREMENT
        ·Room air temperature
        ·Supply air temperature
        ·Supply air fan speed
        ·Supply air fan static pressure
        ·Inlet   water temperature of cooling coil
        ·Outlet water temperature of cooling coil
-CONTROL SIGNAL
        ·Load reset signal of VAV
        ·VAV opening signal
-SET UP VALUE
        ·Supply air temperature
        ·Supply air fan static pressure
        ·Inlet water temperature of cooling coil
-ALGORITHM
        ·Control algorithm of VAV
```

133

4. Process of making the fault tree

The process of making the fault tree which is utilized when in diagnosing the cause of " excessively high room temperature" is described in this section. An HVAC system consists of various sub-units such as rooms, AHUs, VAV box units, piping systems, and duct systems. Various factors in each sub-unit may trigger faulty condition. Thus, the authors made a fault tree by arranging the relations between the causes and the condition of faults in these sub-units.

Firstly, a fault arising from excessively high room temperature is assumed, and the probable causes of such a fault can be enumerated below.

EVENT : " Room air temperature is excessively high "

PROBABLE CAUSES:

  a. Inadequate VAV supply air temperature

  b. Excessive internal heat generation

  c. Lack of VAV supply air volume

  d. AHU start is too late

  e. Inadequate position of air diffuser

  f. Wall insulation not meeting design requirements

  g. Windows larger than those specified

  h. Keeping windows or doors open

  I. Inadequate use of shade

Some of these causes may be confirmed on site while, the other require further checking in detail. Second, the causes of nine probable causes listed above can be considered. As an example, two causes are selected. " Inadequate position of air diffuser " and " Inadequate VAV supply air temperature " will be considered the examples of the former case and the latter case respectively. Next, the probable causes of " Inadequate VAV supply air temperature " can be enumerated as;

EVENT : " Inadequate VAV supply air temperature "

PROBABLE CAUSES :

  a. Inadequate supply air temperature from AHU

  b. Poor insulation of duct

Next, the probable causes of " Inadequate supply air temperature from AHU" can be enumerated as;

EVENT : " Inadequate supply air temperature from AHU "

PROBABLE CAUSES :

  a. Heat load exceeds a design capacity of cooling coil

  b. Failure of supply air temperature controller .... etc.

The process described above was repeated to make up fault trees formed by the events (i.e., the faults) and the probable causes thereof. A fault tree applicable to a cause for an excessively high/low room temperature shown in Fig.4. [3]

room air temperature
deviation from set point

A
excessive internal heat generation

B
inadequate VAV supply air temp

inadequate supply air temp from AHU

poor insulation of duct

B C
air temp reset control defect

defect of a reset controller

VAV trouble

inadequate mixed air temp in AHU

E
coil inlet water temp deviation

F G
failure in AHU valve control

degradation of AHU valve
deterioration of valve actuator

insufficient size of coil

scale on coil fin

failure in AHU valve control

outdoor air cooling control miss

failure in OA/EA/RA damper control

degradation of OA/EA/RA damper
deterioration of damper actuator

outdoor air temp is high/low

VAV damper regulation miss

air leakage from duct

temp deviation of from set point

poor insulation of piping

water leakage from piping

false wiring connection

noise

deterioration of controller

inadequate parameter setting

C
lack/excess of VAV supply air volume

failure in VAV damper control

D
lack/excess of static pressure

degradation of VAV damper
deterioration of VAV damper actuator

VAV damper regulation miss

VAV unit is too small

air leaking from duct

failure in control algorithm
for air volume

false wiring connection

noise

deterioration of controller

inadequate parameter setting

failure in fan control

insufficient size of fan

pulley establishment defective

deterioration of fan

static pressure sensor defective

failure in control algorithm
for static pressure

false wiring connection

noise

deterioration of controller

inadequate parameter setting

AHU start is too late

inadequate position of air diffuser

less wall insulation than
design value

large window than design value

window/door is open

inadequate use of shade

example

fault can be judged by "condition operating chart"

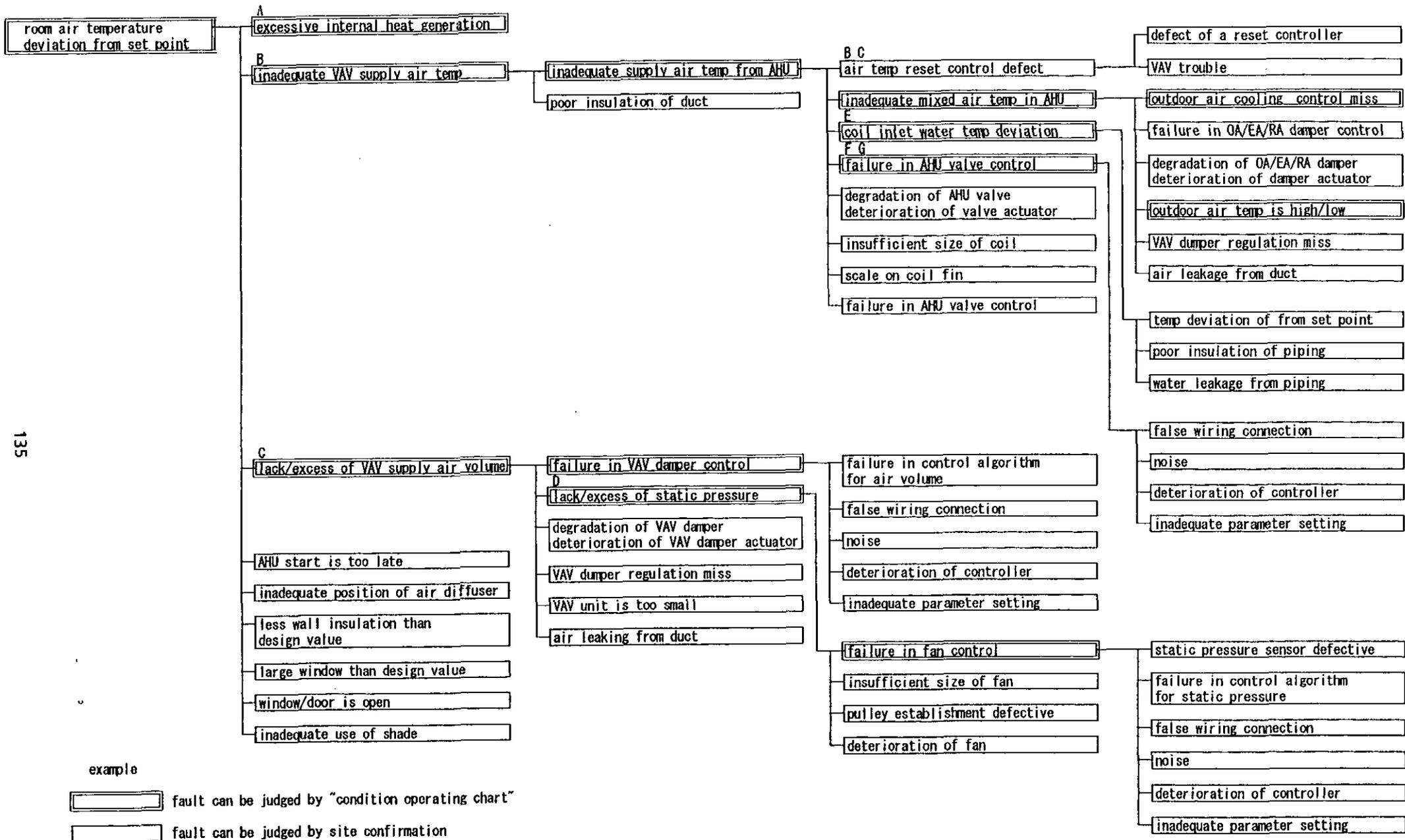fault can be judged by site confirmation

135

Fig.4   Fault tree for the fault that the room air temperature is excessive high/low

5. The diagnosing method by using the fault tree

The flow diagram of the diagnosis using "the FTA system" is shown in Fig.5.[2] This system asks the maintenance personnel questions and presets the diagnosis result. When diagnoses are made by the use of the fault trees, the method for judging whether or not the measured values are compatible with the rules of the diagnosis becomes a highly important factor. Hence, proper judgment standards are required for all cases. For instance, whether or not the room temperature is really too high can be readily ascertained by comparing it with the preset temperature. However, if the proposed diagnosis method is to be really trusted by the maintenance personnel, the process of such diagnosis must be clearly shown and understood by them. To make this possible, the diagnosis method includes the use of a screen on which the judgment standard as well as the measured values is displayed. What is displayed on the screen is referred to as "operating condition charts." However, there are also the cases when the data and information necessary for preparation of such "operating condition charts" are not available. Therefore, firstly, an attempt is made to detect probable causes of a fault by the use of "operating condition charts," and after such probable causes have been narrowed down, real causes are ascertained by on-site visual inspection as necessary. In this chapter, the essence of "the operating condition charts" as well as the utilization method and the judgment standard for some of such charts are described. The judgment standard given here has been obtained through analyses of the data of a number of simulated cases (i.e., the data nine cases including one normal fault-free case).

①Installation of FTA system

Knowledge of designers. contractors. maintenance personnel

⑥Update the knowledge

③Input operation data to FTA system

Software for making FT

④FT diagnosis. Show the result

Measurements

Maintenance data

DATABASE Determine the causes of faults

Select rules

Maintenance personnel

Condition operating charts

FTA system

②Detecting faults

alarm   claim   alarm

BA Occupant FD system
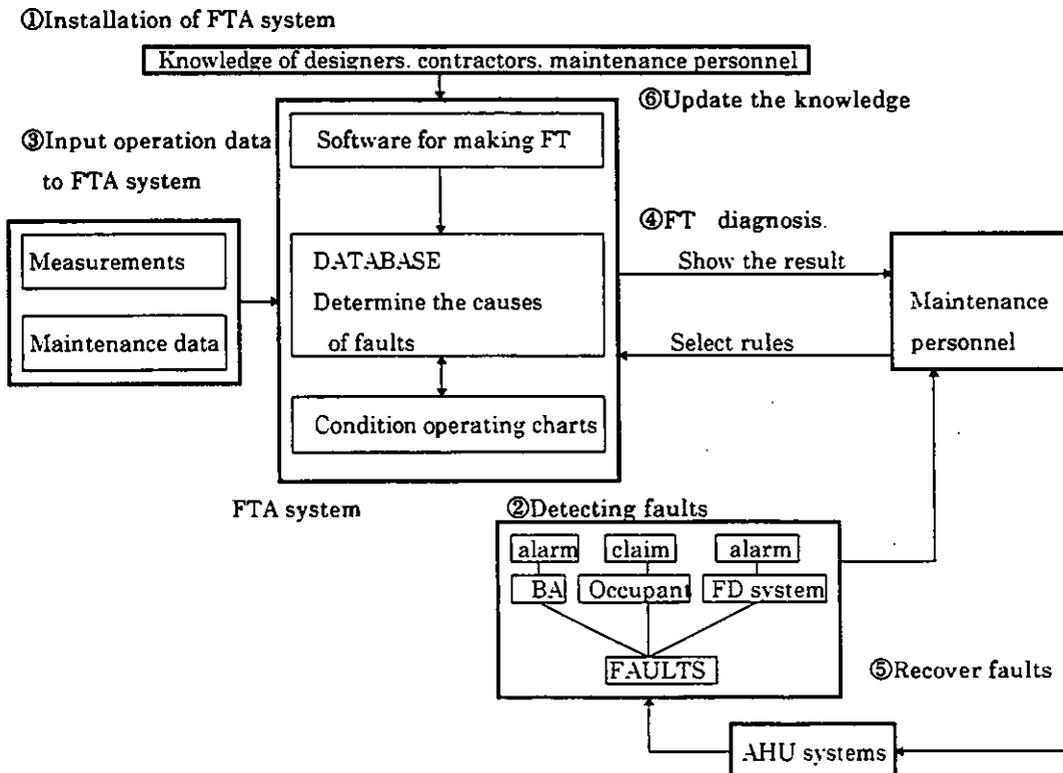
FAULTS

⑤Recover faults

AHU systems

Fig.5 Flow diagram of the diagnosis with FTA system [2]

(1) The room sensible heat load variations by the use of "the operating condition chart A"

1) Composition of the "operating condition chart A"

This " the operating condition chart " consists two of sub-charts. The lower sub-chart shows changes in room sensible heat load and the design load. This example shows that sensible heat load exceeds the design load value during 11:40-11:50 and 13:50-14:00.
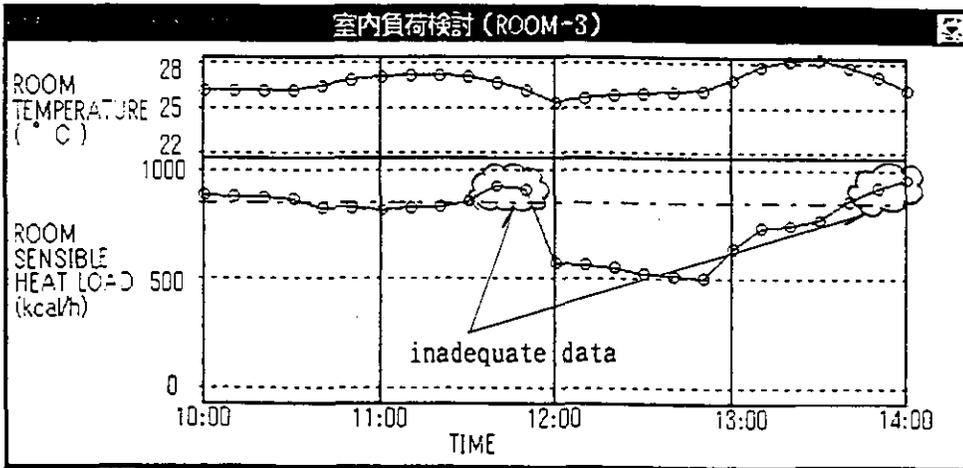

Fig.6 The operating condition chart A

2) Faults indicated by this chart

" excessive internal heat generation "

" insufficient wall insulation not meeting the design requirement "

" windows larger than those specified "

" window / door is open "

" inadequate use of shade "

3) How to interpret the "operating condition chart A"

① Judgment as to whether or not the sensible heat load in any room is in excess of the specified value is made by checking time-sequential changes in the air temperature and sensible heat load in that room.

② By checking the time-sequential data, judgment is made when the room sensible heat load exceeded the design load (in other words, when a fault in this respect took place).

4) Data necessary for forming the "operating condition chart A"

① Room sensible heat load (computed value)

$$Q = C \times (T1 - T2) \times V$$

where,   C : a constant

T1 : room air temperature (= return air temperature)

T2 : supply air temperature

V : supply air volume presumable from VAV opening signal

② Design room sensible heat load

(2) Supply air temperature control by using the "operating condition chart B"

1) Composition of the "operating condition chart B"

The chart given below indicates the changes in the supply air temperature in reference to the preset temperature and the control signals, The charts also indicates the time period during which the measured temperature exceeded the preset temperature (13:20-13:40)
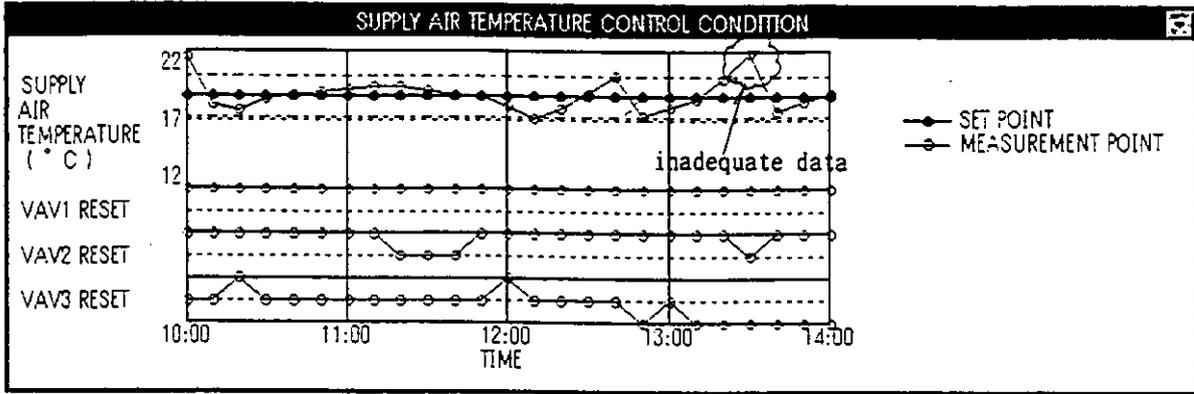

Fig.7　The operating condition chart B

2) Faults indicated by this chart

" VAV supply air temperature is inadequate " and/or

" supply air temperature is set inadequately "

3) How to interpret the "operating condition chart B"

① Judgment as to whether or not the supply air control is satisfactory is made by referring to the preset supply air temperatures and the time-sequential data on actual supply air temperatures. Under normal conditions, the changes in the actual supply air temperatures follow those in the preset ones with some time lags but almost in a similar way. In this case study, a temperature difference of ±2℃ and a time lag of 30 minutes or less were considered to be allowable.

② Judgment as to when a fault occurred in the supply air temperature control is made by reference to the time-sequential air temperature data.

③ Judgment as to whether or not the supply air control is satisfactory is made by referring to the mode of reset control signal and the time-sequential data on actual set point of the supply air temperature. If the system is working satisfactorily, the supply air temperature changes when the load setting signal is given according to the preset supply air temperature program. In the system now being discussed, the supply air should change within a range of 14.6℃～18.7℃.

④ The time sequential temperature data is frequently examined so that, if any fault occurs with respect of the reset control, the time of such fault occurrence can be correctly determined.

⑤ Symbols in the charts: The thin line-and-circle symbols and the thick line and symbols represent the actual supply air temperatures and the preset supply air temperatures respectively.

4) Data necessary for forming the "operating condition chart-B"

① supply air temperature

② preset supply air temperatures

③ preset VAV signaling points as estimated from the room temperature and the VAV control algorithm.

(3) VAV control by using the "Operating condition chart C"

1) Composition of the "Operating condition chart C"

The "Operating condition chart-C " consists of four sub-charts. The sub-chart on the left side shows the algorithm related to VAV damper opening with respect to the room temperature as seen in Fig.8. The three sub-charts to the left show in a top-to-bottom, the room air temperature variation, VAV degrees of opening and VAV damper opening deviation. The chart given here indicates that the VAV damper for Room-3 was in some trouble during 12:30~14:00.
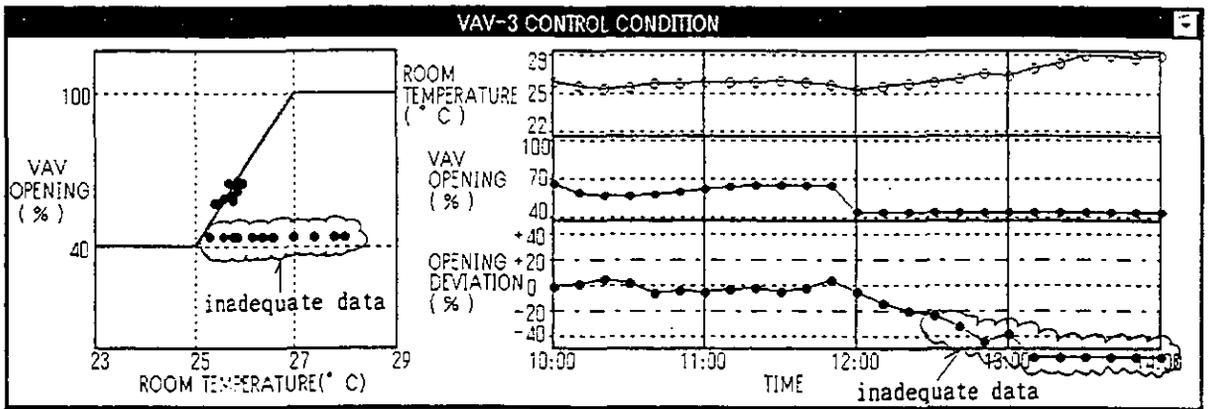
Fig.8   The operating condition chart C

2) Fault indicated by this chart

" lack / excess of VAV supply air volume "

3) How to interpret the "operating condition chart C"

① The relation between the room air temperature and VAV opening signal shown in the chart indicates that the VAV was opened to an incorrect position. Under normal conditions, VAV opening control signals are controlled according to the control algorithm. (In this system, the degrees of VAV opening are supposed to change within a range of 40%~100% in proportion to the room air temperature that changes within a 25℃~27℃ range.)

② From the time-sequential data which indicates a certain deviation from the algorithm, the time when some trouble occurred to the VAV control system can be determined. In this specific case, judgment was based on the condition that deviation exceeding ±20% from the preset value would regards as a fault.

139

4) Data necessary for forming the "operating condition chart C"

① room air temperature (return air temperature)

② VAV opening signal

③ conditions under which the VAV control algorithm was established

(4) Control of supply air fan speed by using "Operating condition chart D"

1) Composition of the "operating condition chart D"

The captioned chart is composed of four sub-charts as seen in Fig.9. The sub-chart in the left shows the fan speed pressure relation of the supply fan while those in the right show changes in the pressure, the fan speed, and the fan speed. deviation of the supply fan.

This sample chart indicates that the supply fan pressure deviated from the design pressure during 10:00~14:00
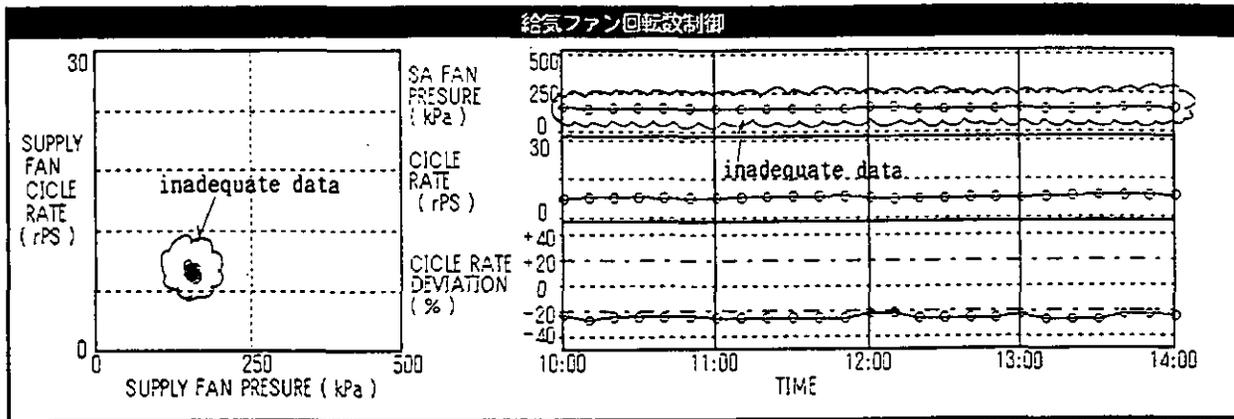
Fig.9   The operating condition chart D

2)   The cause of a fault indicated by the chart

" The static pressure was too low or too high "

3) How to interpret the "operating condition chart D"

① From the relation between the fan speed of the supply fan and the static air pressure developed by it, it may be judged that the supply air fan was at fault. If the conditions had been normal, the supply air fan would have developed generally stabilized pressure irrespective of the fan's fan speed In this sample case, the static pressure developed by the supply air fan was taken as 249 Pa with an allowance of ± 20%.

② The time at which any fault occurred at the supply air fan can also be determined.

4) Data necessary for forming the "operating condition chart D"

① The fan speed of the supply air fan

② The static pressure developed by the supply air fan

6. Example of FTA system application

In order to verify the effectiveness of the FTA system, the system was applied to some HVAC system models developed by the use of HVACSIM+ in such a way as to contain various faults. In this paper, a model in which the controller regulating the r.p.s. of the supply air fan is at fault will be selected for a case study.

Cause of a fault : Inadequate range setting of the pressure sensor
(The value was amplified ten times.)

A process of diagnosis by using the FTA system is summarized below.

(1) Investigating the air temperature in each room. (Fig.10-a,b,c)

Although the room air temperature in each room is within the predetermined range (25-27° C), the air temperature in Room-3 is on a little high side; so, analysis will be conducted for this room.
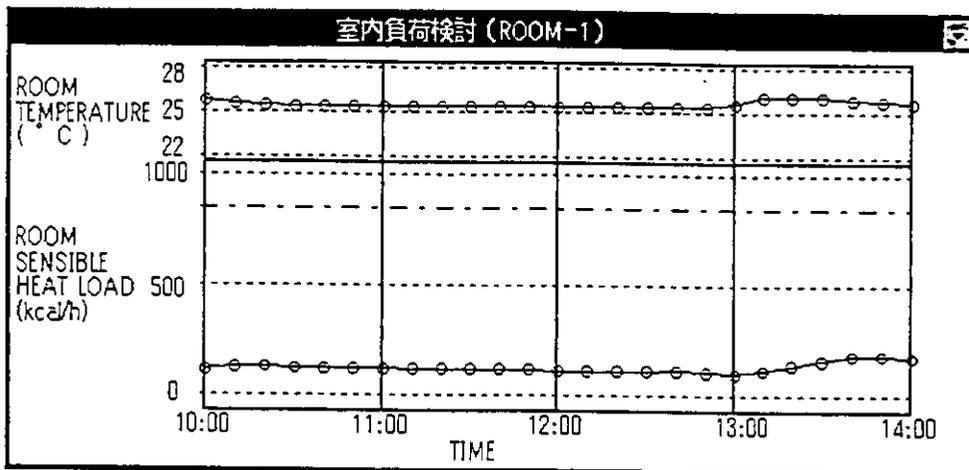
Fig.10-a Room temperature and internal heat generation in ROOM-1
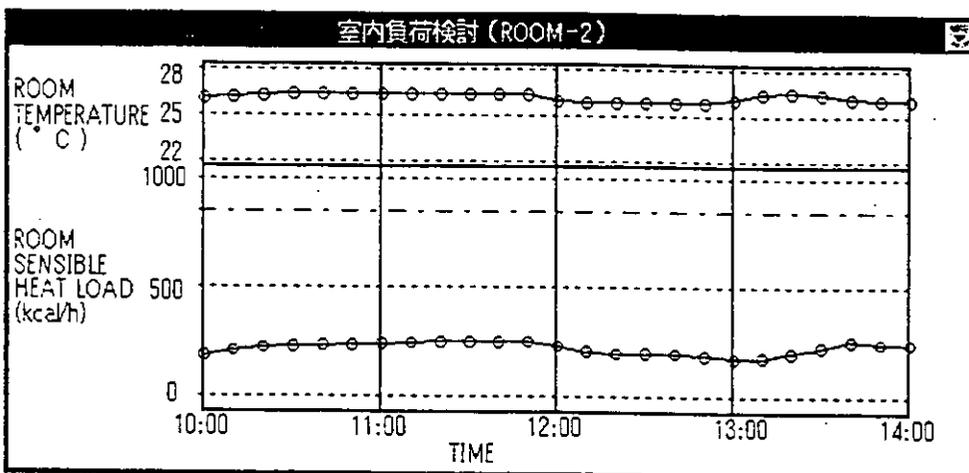
Fig.10-b Room temperature and internal heat generation in ROOM-2
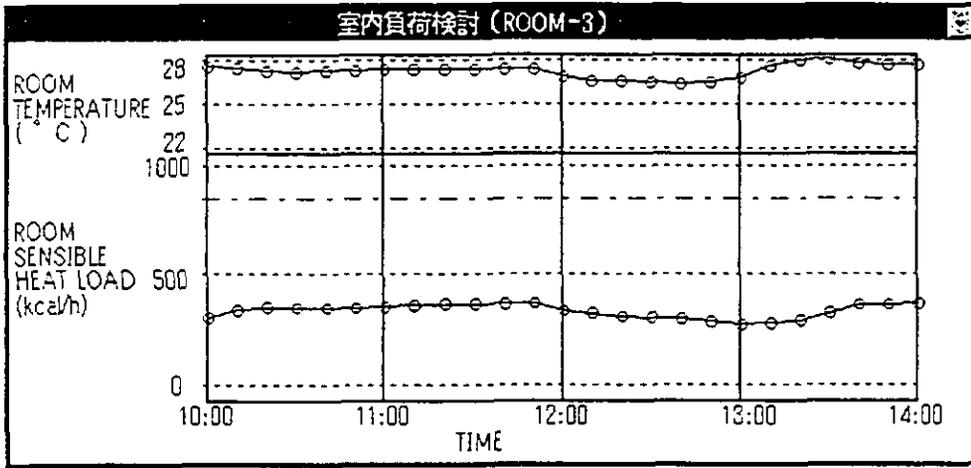
Fig.10-c Room temperature and internal heat generation in ROOM-3

(2) Confirming the internal heat generation. (Fig.10-a,b,c)

In each room, the internal heat generation is lower than the design value; hence, the internal heat generation is not the cause of the fault.

(3) Confirming the supply air temperature from VAV. (Fig.11)

· The supply air temperature deviates from the preset range during a certain period of time (i.e., 13:00~13:30 hours).

· Supply air temperature variation is within the preset temperature range.(The temperature is so controlled as to be in the neighborhood of 14.7 ℃ which is the lower limit of the preset temperature range.)

· Even when the supply air temperature is 14.7 ℃ which is the lowest within the preset range, the air temperature Room-3 is still too high.

→ Hence, the cause of the fault is not the supply air temperature.
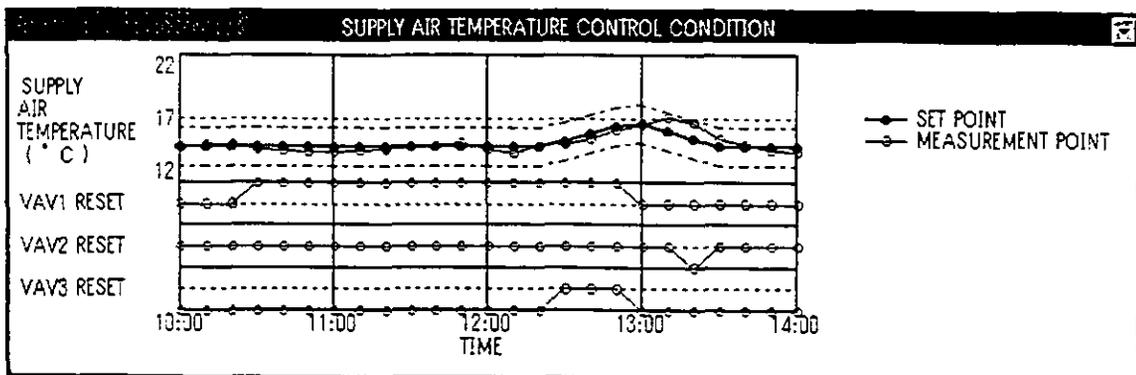
Fig.11 Supply   air temperature

(4) Confirming the opening of damper of VAV-3. (Fig.12)

Control values are distributed on the control line.

In case that room air temperature is high, the VAV damper opening is maintained by 100%.

→ Cause of the fault isn't the VAV damper opening.
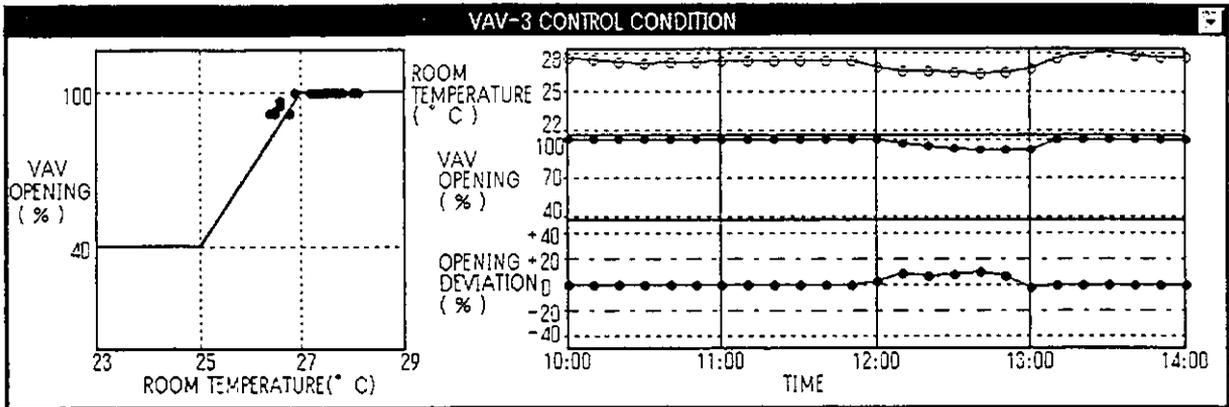
Fig. 12 The VAV damper opening of VAV-3

(5) Confirming the supply air volume. (Fig.13)

The fan speed of the supply fan is about half of the design value at all times.

→Hence, it is considered highly probable that something is wrong with the system which controls the revolutions of the supply fan. If the fan speed of the supply fan is controlled satisfactory, the fan speed is closed to the design value (249Pa).
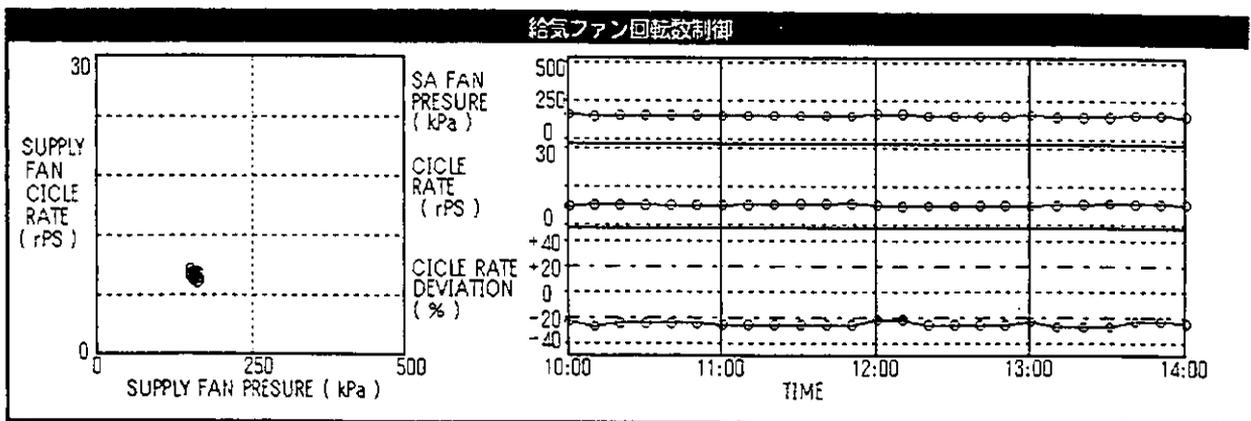
Fig. 13 Supply air volume

143

(6) Advice which the FTA system gives to the operator

> ・ It is highly probable that the malfunction of the system which controls the fan speed of the supply air fan is responsible for the faults.
> ・ The controlling parameters related with the supply air temperature need to be readjusted.

Subsequent to the above diagnoses, the FTA system only can give above advice to the operator. After the operator receives that advice, he requests the building automation maintenance company to check the system which controls the fan speed of the supply air fan. The maintenance crew of the building automation maintenance company will find out that the cause of fault is inadequate range setting of the pressure sensor.

7. Conclusive Remarks

The merits and demerits of the FTA system may be summarized at follows.

MERITS

(1) All significant information concerning the system operation is displayed graphically in the form of "the operating condition charts"; hence, the maintenance crew can easily follow the minute-to-minute state of the system operation correctly.

(2) Even if the input data leaves out some items of information, it is still possible for the maintenance crew to conduct diagnosis properly.

(3) Since the FTA system expresses the information in ordinary writing, even the third party persons can understand the essence of the information without difficulty.

DEMERITS

(1) Diagnoses are impossible if any faults occur due to causes not covered by the FTA system input program.

(2) The users need to have the information necessary for making judgment as well as the proper judgment criteria; so, the means to provide the users with such information (e.q., one as provided by the sensor) and criteria must be considered.

(3) The ability of the FTA system is limited to the diagnoses of air temperature-related faults. If the FTA system is to serve the practical purposes more usefully, it should be capable of diagnosing energy-related problems, too. Future development of the system to cover this area is considered necessary.

Reference

[1] Nakahara N.:Fault Diagnosis Techniques, 4/1994, IEA Annex 25, MIT

[2] Yoshida H.,Yuzawa H.,Suzuki M., Takei H.,:The development of FTA system on AHU systems.,4/1994,IEA Annex 25,MIT

[3]   Yoshida H., Yuzawa H.,Suzuki M., Takei H.,:The   application of FTA system.,4/1995,IEA Annex 25,Montreal

Impact of Operational Faults in a VAV System on Energy Consumption, Thermal Comfort and Monitored Process Variables

Madjid Madjidi

University of Stuttgart
IKE-HLK Pfaffenwaldring 35
D-70550 Stuttgart, Germany

Abstract

This paper presents the results of a simulation study on typical operational faults in a VAV system of a commercial building. Basic equations for the main simulation models are described. The static models for coils and fans are based on characteristic curves. They are also recommended for real time applications in an automatic fault detection system. The simulation results demonstrate the applicability of system simulation as a generic tool for constructing fault detection and diagnosis systems. Simulation aids choosing the suitable process variables to be monitored in an automatic fault detection system. By computing the expected magnitude of changes in energy consumption and thermal comfort the importance of faults can be evaluated and considered in fault diagnosis.

1 Introduction

Fault detection and diagnosis (FDD) of HVAC systems is the main research field within IEA project Annex 25 /14/. Operational faults of a HVAC system can be found automatically, if the model behaviour of a system operating under regular conditions deviates from the behaviour of the real monitored system. The lower part of Figure 1 illustrates the basic idea adopted from /22/. Hereby the process of recognizing fault effects is called *fault detection* whereas the process of searching symptom's causes is called *fault diagnosis*.

Automatical fault detection and diagnosis in HVAC systems is a young research field. Only few expert knowledge yet exists according to the importance of operational faults and their influence on total systems.

Independently by the methods and algorithms to be implemented in a FDD system, when constructing such a system, HVAC system related information are required /25/.
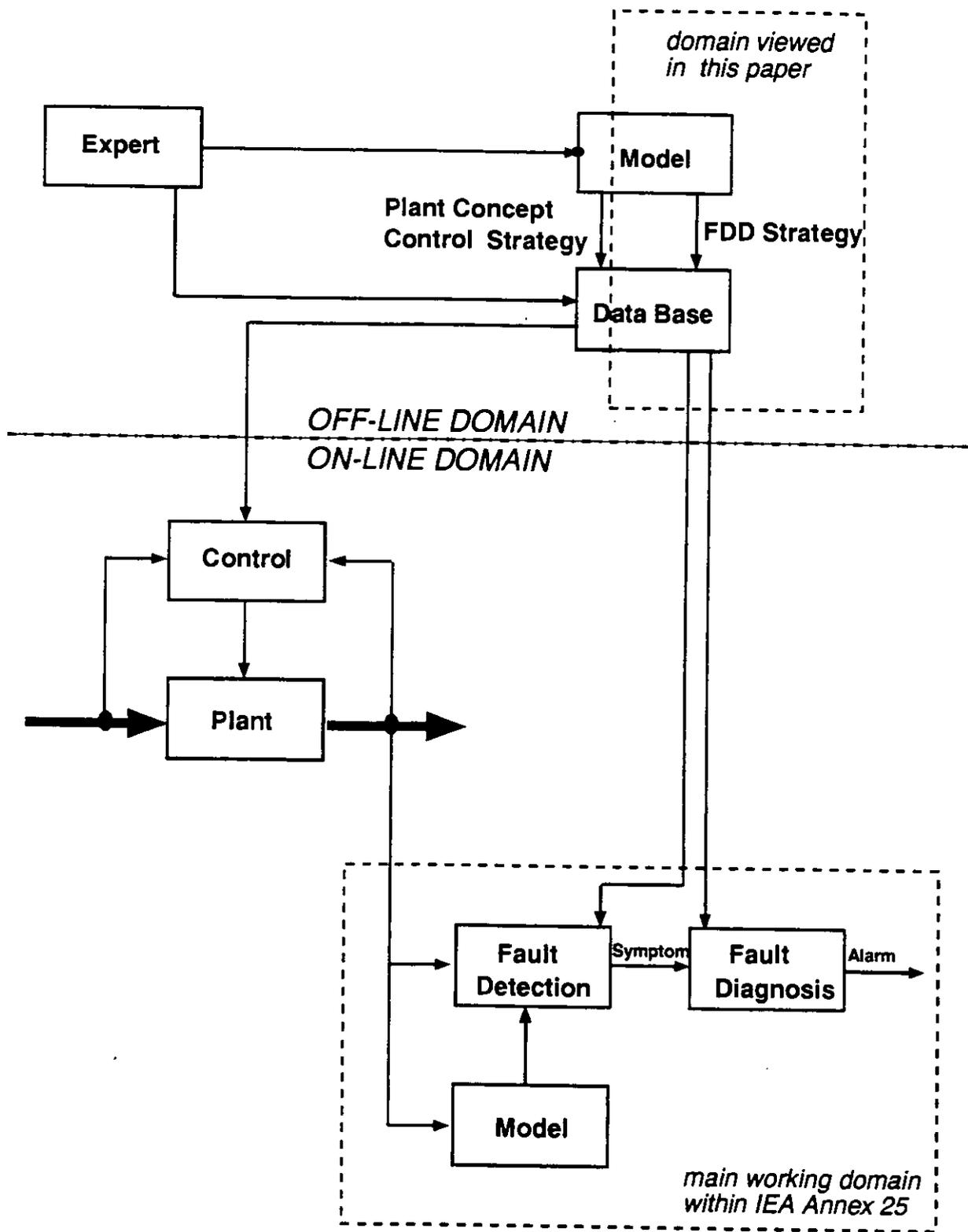
Figure 1: Design, control and FDD of HVAC systems

These information can be named as FDD project data and they can be classified as follows:

- **system topology** (components, interconnections, instrumentation)

- **main process variables** (those monitored variables which might provide robust symptoms)

- **possible faults and their effects** (impact of faults on monitored variables, security, environment, thermal comfort, energy consumption, operational costs, service and maintenance costs)

- **tresholds** (to avoid false alarms the grade of the deviation should be considered and an alarm should be released only , if the deviation exceeds a certain limit)

- **setpoints** (controller setpoints and how they are determined and the possible operational modes of a system or subsystem)

The sources for these information can be design data, human logic, practical experience, lab experiments and system simulation (see upper part of Figure 1).

Tresholds can be found by statistical methods and by expert knowledge. They may depend on accuracy of process variable measurements (sensors and converters), control tolerances of local loop controllers, individual tolerances (concerning comfort, energy waste, environmental damage, service cost), selection of reference model (type: physical, neuronal networks, fuzzy logic, characteristic curves, ...) and model accuracy (numerical accuracy) /1/. But also the classification of faults may help to eliminate problems of treshold adjustment /21/.

The selection of the main process variables to be monitored can be derived from the knowledge on possible faults and their effects. The aim of the presented study is to demonstrate the applicability of system simulation as a generic FDD tool, especially for the selection of suitable process variables to be monitored in an automatic fault detection system. Furthermore, by computing the expected magnitude of changes in energy consumption and thermal comfort the importance of faults can be evaluated and considered in fault diagnosis.

In the following example faults are introduced just by *intuitivly tuning* some of the characteristic values of the simulation models. The effect of model accuracy on the results is not analyzed. Such a study would require large experimental investigations according to fault models (impact of degradation, fouling, clogging, leackage, ... on characteristic values of simulation models). The presented study gives for the first time a global impression about the possible impact of faults on energy consumption, thermal comfort and usually monitored process variables.

2 Simulation Example

The example presented here is chosen because the Variable Air Volume (VAV) system is similar to the Annex 25 reference air handling unit /13/. Design data and real measured data form the base of the developped simulation model /17/. The complete weather data of a german test reference year are used to represent realistic climatic conditions (TRY05 /2/) for the execution of one-year-period simulations.

2.1 Building

The building is the administration and education center of a computer company near Stuttgart. It has a total area of 74 000 m^2. Figure 2 shows the shape and the blueprint of the building and the selected training rooms. The training rooms are oriented to the east. Each room has a ground area of 40 m^2 and a height of 3 m The observed HVAC system supplies only 6 of these nearly identical rooms. All internal walls have the same structure. Table 1 includes the list of building structure materials for these rooms. The windows are double glazed. They cover 70 % of the external walls and they have a global heat transfer coefficient of 2.9 W/mK.

2.2 HVAC System

Figure 3 shows the scheme of the HVAC system which supplies the teaching rooms. It consists of a single-duct pressure independent VAV system with a central air handling unit and local terminals and an additional hydronic heating system equipped with radiators in each conditioned zone. The additional hydronic heating system allows to have lower supply air temperatures and airflow rates in winter than observed in traditional VAV-concepts. The VAV-system consists of mixing box, air filter, finned tube preheating and cooling coils, centrifugal supply and extract fans with revolving speed control and finned tube reheating coils and VAV-boxes for each zone. In Table 2 nominal characteristic values for the coils and fans are listed.

2.3 Control Strategy

The complete system is controlled by a Direct Digital Control (DDC) unit and local controllers. Figure 4 represents the implemented control strategy. Also the daylight is controlled. The DDC-unit includes the supervisory control strategy for the optimal set points of temperatures, massflow rates and pressures to be realized with local PI-controllers. Figure 5 gives the setpoint temperatures for zone air and supply air as functions of the outdoor temperature. Variable air volume is realized only in the

blueprint of
a section

training rooms

training rooms

isometrical view

Figure 2: Building

149
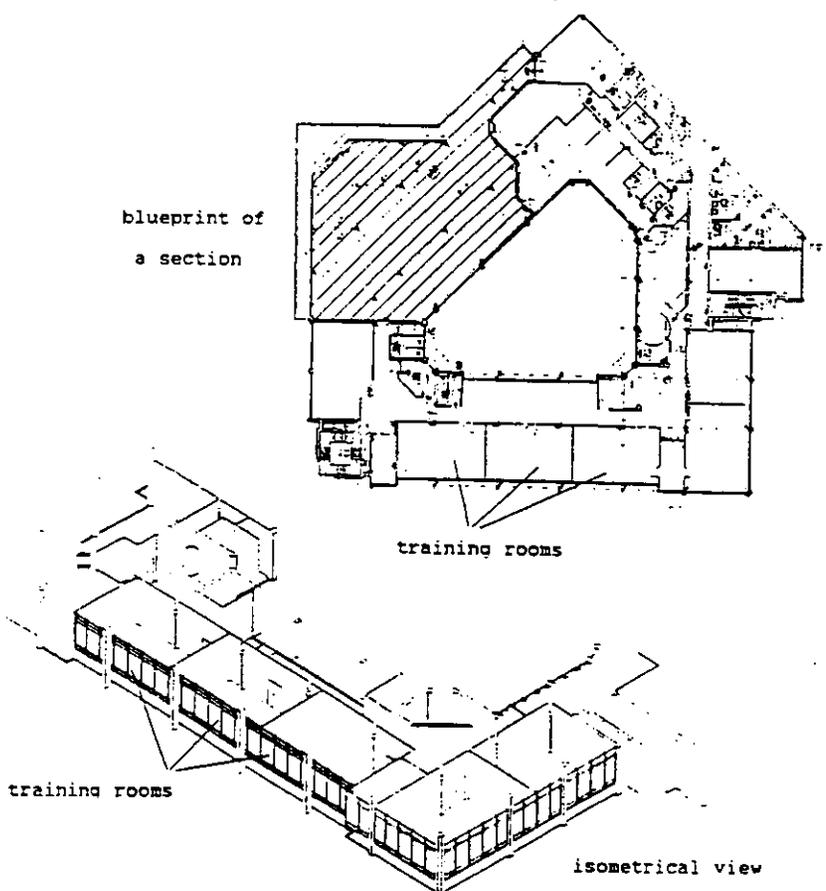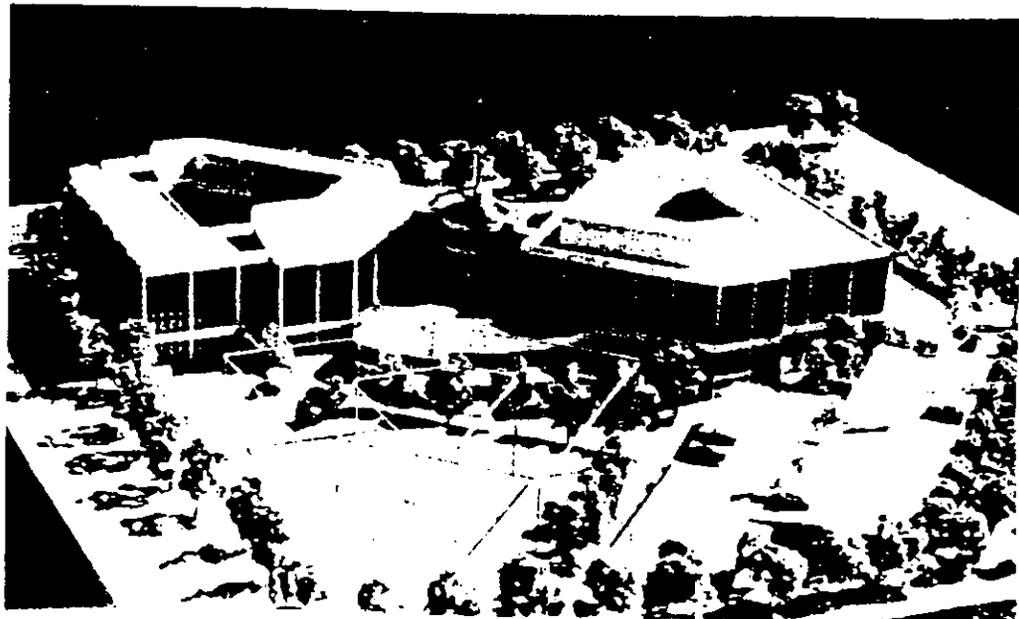
Table 1: Building structure materials

section	composition	thickness m	conductivity W / (mK)	specific heat kJ / (kg K)	density kg / m³
external wall	plaster	0.005	0.87	1	1400
	concrete	0.3	2.035	0.92	2400
	insulation	0.09	0.045	0.84	75
internal wall	insulation	0.1	0.045	0.84	75
floor	carpet	0.007	0.06	1.3	100
	pavement	0.075	1.4	0.92	2000
	concrete	0.12	2.035	0.92	2100
ceiling	concrete	0.12	2.035	0.92	2100
	pavement	0.075	1.4	0.92	2000
	carpet	0.007	0.06	1.3	100

Table 2: Nominal values of HVAC components

coils	preheat	cooling	reheat
power in kW	97	49	3
air flow in m³/h	7000	7000	1000
water flow in m³/h	4.15	9.01	0.1
air inlet temp. in °C	-15	26	20
air outlet temp. in °C	25	16	30
water inlet temp. in °C	80	7	80
water outlet temp. in °C	60	12	60
air pressure drop in Pa	40	120	10

supply and return fan	
air flow in m³/h	7000
pressure increase in Pa	600
revolving speed in 1 / min	2000

radiators	
power in kW	1
supply in °C	80
return in °C	60
zone air in °C	20
exponent	1.3

component	filter	preheat	cooling	reheat
air flow resistance (see equation 9)	24.28	7.47	22.41	91.82
exponents (see equation 9)	2	2	2	2
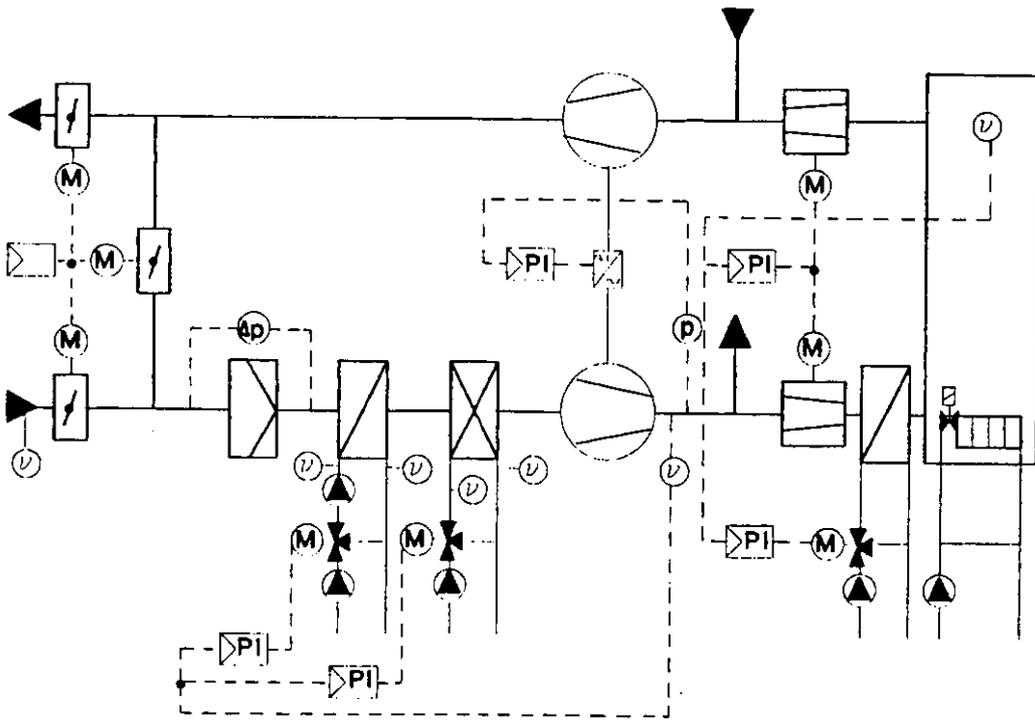
Figure 3: HVAC system

	outdoor temperature		
	< 10°C	> 10°C	
mode	winter	summer	
radiator	on	off	
preheating coil	on	off	
cooling coil	off	on	
		room air temperature	
		> setpoint	< setpoint
VAV damper	minimum air flowrate	VAV	minimum air flowrate
fan	400 Pa	400 Pa	400 Pa
	room air temperature		
	> setpoint	< setpoint	
reheating coil	off	on	off	off

Figure 4: Control strategy

summer mode. In winter the zone temperatures are controlled by variable zone supply temperatures modulated by the reheting coils. The minimum air flow rate for each zone is about 70 % of the maximum value. Both fans are fitted with variable frequeny control for varying the airflow rate. A single controller is used to control both fans. It is set to hold a static pressure of 400 Pa at the measured point indicated in Figure 3. Further assumtions are

- operation of the plant between 9 am and 8 pm;

- 100 % outside air during operation period in summer mode, additional daily one-hour-preheating periods in winter mode with 100 % return air;

- solar beam radiation used to control the daylight: if beam radiation is greater than zero, the total solar gains through windows is reduced 70 %.

3 Simulation Models

The simulation program TRNSYS 13.1 /28/ is used to realize the coupling between a detailed building model and the plant model. The chosen simulation time step is 36 seconds. The VAV-system is modelled in a modular way. Each component is represented by a stand-alone program called TYPE. In the following subsections the models for building and plant components are described.

3.1 Building

The TRNSYS subroutine TYPE 56 based on the response factor method is applied. The walls are modelled according to the transfer function relationships from wall surface to wall surface /28/. The long-wave radiation exchange between the surfaces within a zone and the convective heat flux from inside surfaces to zone air is considered according to a star network approach /23/. The computation of radiation transmittance through windows takes into account the incident angle of solar beams and glazing type /27/.

3.2 Centrifugal Fans

The performance of the fans is modelled by a relationship between the fan pressure increase Δp , air volume rate $\dot{V}$ and revolving fan speed n /7/. No geometrical data are needed, contrary to the dimensionless representation given in the IEA Annex 10 and 17 specifications /3,29/.
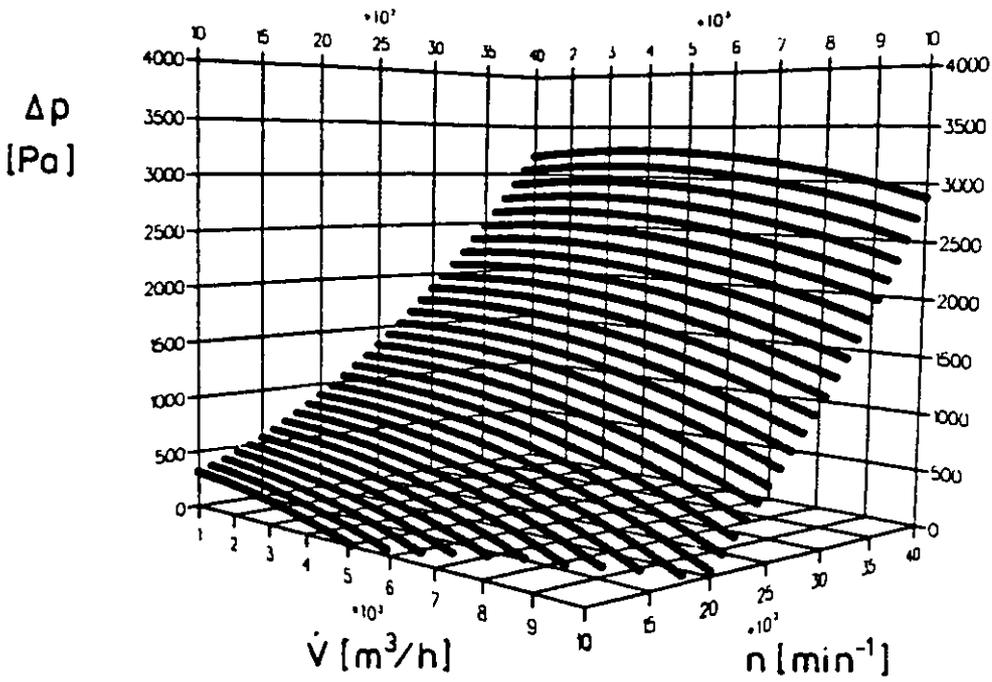
characteristic fan curves (catalogue)

Figure 6: Static model of fans

Even if manufacturers often use the fan laws to generate performance curves at different rotation speeds, just a single square polynomial can describe all curves in a more compact but still accurate way:

$$\Delta p = c_1 \cdot n^2 + c_2 \cdot n + c_3 \cdot n \cdot \dot{V} + c_4 \cdot \dot{V} + c_5 \cdot \dot{V}^2 + c_6 \ . \tag{1}$$

The basic advantage of the approach is that the parameters of the model are only the six polynomial coefficients of this equation. In this example manufacturers' measurements (catalogue data) are used for the estimation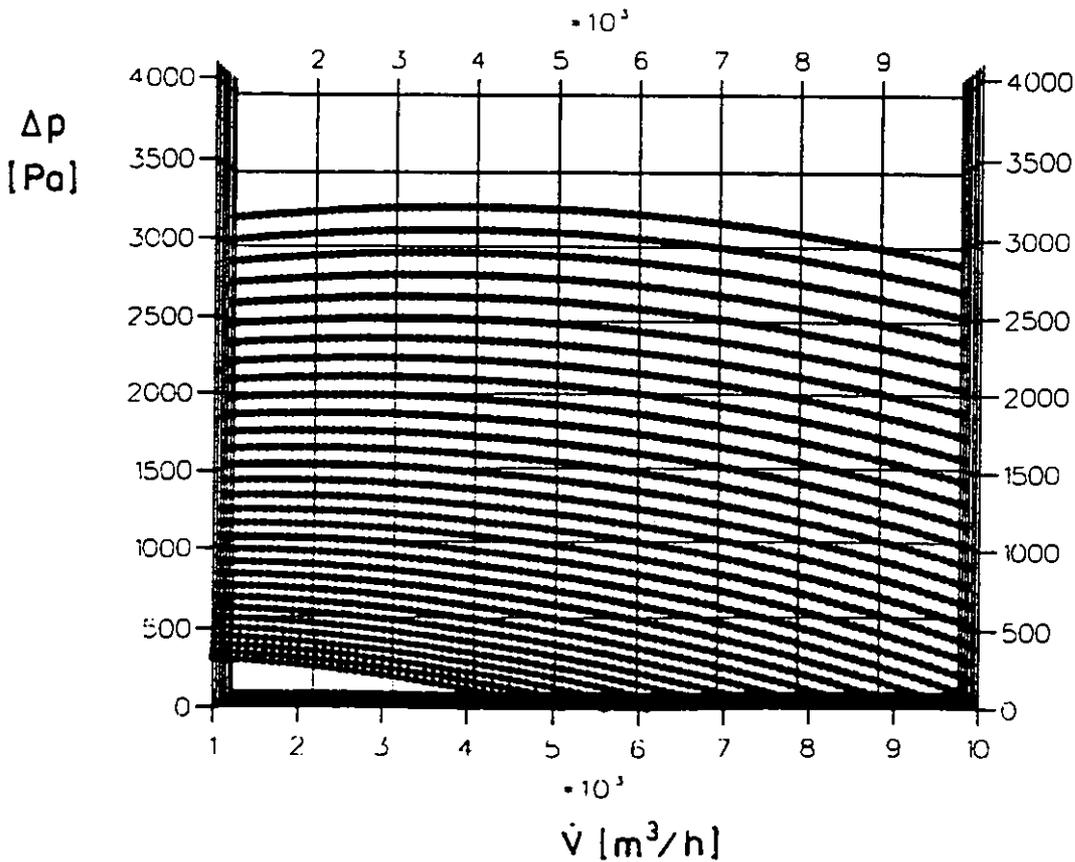 of the supply and exhaust fan model. The lower part of Figure 6 displays the curves measured by the fan manufacturer. The upper part of Figure 6 describes the same characteristic curves as a 3D-space-surface. It is achieved with the following polynomial coefficients :

$$c_1 = +0.00019 \ , c_2 = -0.01862 \ , c_3 = +0.00003 \ ,$$
$$c_4 = -0.04369 \ , c_5 = -0.00001 \ , c_6 = +171.7724 \ .$$

It is assumed that all electrical energy is converted to an increase of air enthalpy and kinetic energy. A polytropical process is assumed to calculate the increase of air temperature. The fan total efficiency η_t is assumed to be constant (0.7) in the feasible operational range (partial load varies in this example only between 75 % and 100 % total air volume rate and about 500 to 1000 Pa pressure increase).

The outlet temperature can be calculated from the energy balance:

$$\vartheta_2 = \vartheta_1 + \frac{\Delta p}{\varrho_L \cdot c_{p,L} \cdot \eta_t} \tag{2}$$

3.3   Coils

Within the IEA projects Annex 10 and 17 a detailed physical coil model was developped /10/ and applied for system simulation applications. Most of the parameters of this model are geometrical data of a coil, like fin thickness, fin spacing and so on. In this model the global thermal resistance is the sum of the waterside thermal resistance, the coil material resistance and the airside thermal resistance. Empiric relationships are given for the computation of the thermal resistances on each flow side depending on the flow velocity . Especially relationships given for the thermal resistance on the airside should be used with care. They often relate to various coil constructions, especially with different settings of water tubes (e.g. shift in columns and rows) /8,9/. Furthermore, geometrical data are often not available or manufacturers hide these even as a secret.

A more simplified way to roughly characterize the heat transfer in dry coils can be done with the socalled coil effectiveness Φ /5/. It is defined as the fraction of the

total heat transfer to the product of the maximum possible temperature difference and the minimum flow capacity.

For an air to water heat exchanger the total heat transfer is expressed by the following equation

$$\dot{Q}_{Tot} = (\dot{m} \cdot c_p)_{min} \cdot \Phi \cdot (\vartheta_{Air,1} - \vartheta_{Wat,1}) \tag{3}$$

with

$\vartheta_{Air,1}$: inlet air temperature
$\vartheta_{Wat,1}$: inlet water temperature

and the minimum thermal flow capacity

$$(\dot{m} \cdot c_p)_{min} = MINIMUM \left[(\dot{m}_{Air} \cdot c_{p,Air}), (\dot{m}_{Wat} \cdot c_{p,Wat}) \right] \tag{4}$$

If measured data are available (catalogue, BEMS) the model of a specific coil can be easily calibrated. The result of a model calibration for dry air conditions could be a relationship between the effectiveness and the air and water flowrates:

$$\Phi = \Phi (\dot{m}_{Air} , \dot{m}_{Wat}) \tag{5}$$

If such a relationship is available, the outlet air and water temperature can be computed with the energy balance:

$$\dot{Q}_{Tot} = \dot{m}_{Wat} \cdot c_{p,Wat} \cdot (\vartheta_{Wat,2} - \vartheta_{Wat,1}) \tag{6}$$

$$- \dot{Q}_{Tot} = \dot{m}_{Air} \cdot c_{p,Air} \cdot (\vartheta_{Air,2} - \vartheta_{Air,1}) \tag{7}$$

Again a square polynomial approach is chosen /18/:

$$\Phi = c_1 \cdot \dot{m}_{Air}^2 + c_2 \cdot \dot{m}_{Air} + c_3 \cdot \dot{m}_{Air} \cdot \dot{m}_{Wat} + c_4 \cdot \dot{m}_{Wat} + c_5 \cdot \dot{m}_{Wat}^2 + c_6 \tag{8}$$

The parameters of the model are then only the six polynomial coefficients of this equation. Measurements under relatively dry conditions have been used for the calibration of the coil models. The latent duty is neglected.
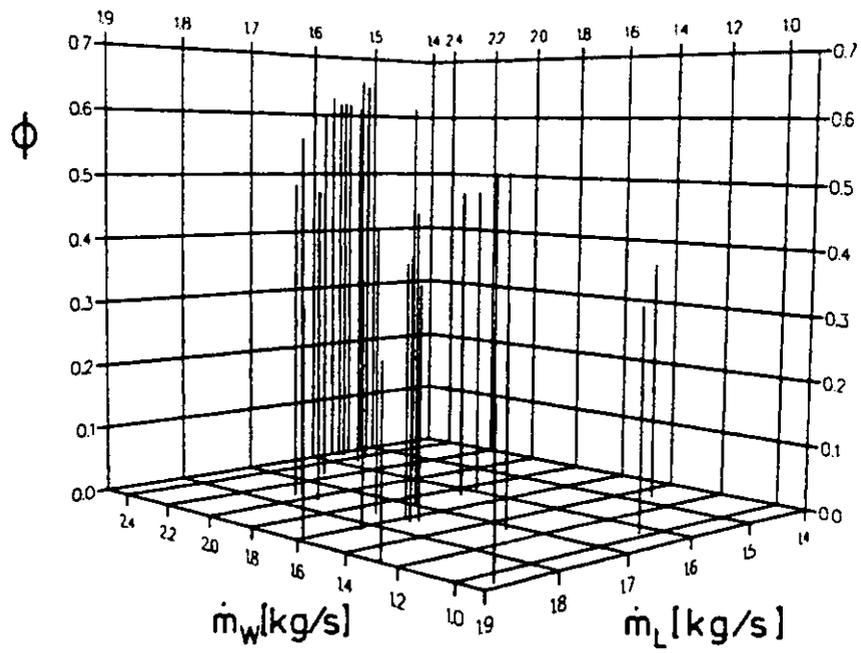
cooling coil effectiveness (measured)
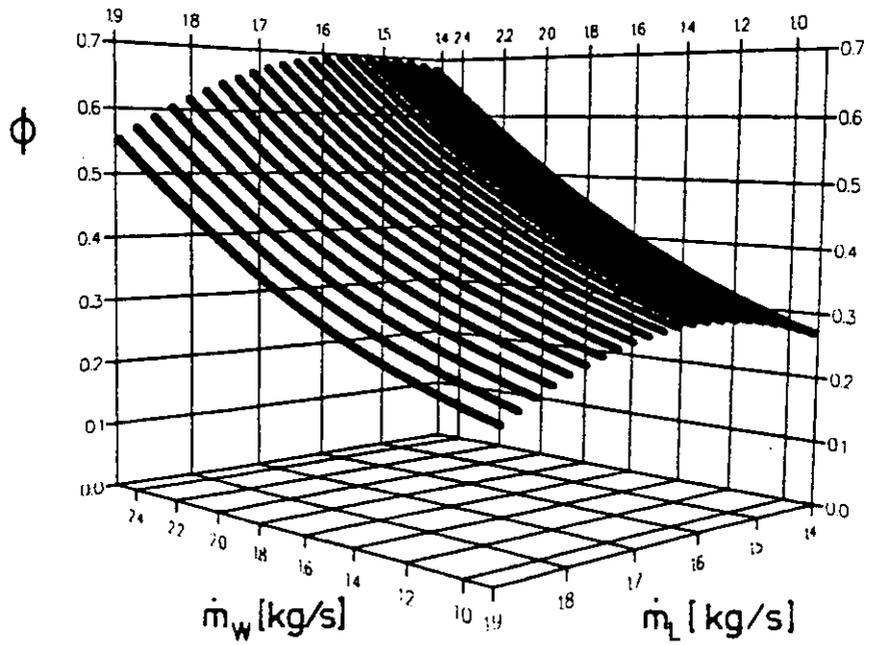
Figure 7: Static model of cooling coil

The upper part of Figure 7 displays measured values for the cooling coil in a 3-dimensional plot. The lower part of Figure 7 displays the calibrated characteristic curves achieved with the following coefficients :

$$c_1 = +0.06945 , c_2 = +0.14889 , c_3 = -0.08750 ,$$
$$c_4 = +3.45875 , c_5 = -1.06667 , c_6 = -2.56158 .$$

It might be surprising to see a maximum in the variation of the effectiveness with the air flow rate since the effectiveness approaches unity asymptotically. But it should be noticed that the outlet air temperature is measured after the supply fan. Therefore, the curves consider the temperature increase caused by the fan. The value for the cooling coil air outlet temperature can be estimated by computing backwards using equation 2.

The pressure drop is modelled by using the following equation

$$\Delta p_{coil} = R_{coil} \cdot \dot{m}_{air}^n . \tag{9}$$

The air flow resistances are determined by using the manufacturers values of pressure drop and air flowrate under nominal conditions (see Table 2).

3.4 Air Filter

Any characteristic curve of an air filter /6/ can be transfered to the following equation

$$\Delta p_{filter} = R_{filter} \cdot \dot{m}_{air}^n . \tag{10}$$

3.5 Ducts

For calculating heat transmission through the duct walls and insulation an overall heat transfer coefficient $U \cdot A$ is used /19/. The approach assumes steady-state conditions:

$$\dot{m} \cdot c_{p,L,f} \cdot (\vartheta_{air,1} - \vartheta_{air,2}) = U \cdot A \cdot \frac{\vartheta_{air,1} - \vartheta_{air,2}}{\ln \frac{\vartheta_{air,1} - \vartheta_{env}}{\vartheta_{air,2} - \vartheta_{env}}} . \tag{11}$$

3.6 PI-Controllers

PI-controllers are modelled explicitly and as to be digital:

$$Y = Y_0 + K_p \cdot [X_S - X_M] + K_I \sum_{t=0}^{t=t} \cdot [X_S - X_M] \cdot \Delta t . \qquad (12)$$

with

X_M :   input signal (measured)
X_S :   set point
Y :   output signal
Y_0 :   constant value
K_P :   proportional coefficient
K_I :   integral coefficient
Δt: measurement time interval

Valves and dampers are modelled to provide massflow rates which are directly proportional to controller output signals.

3.7   Air Pressure Control

As mentioned in 2.3 in this example the fan controller is set to hold a certain static pressure p_{static} at the measured point indicated in Figure 3. The mixing box is closed during plant operation. The pressure drop between fan and pressure sensor is negligible due to relatively small distance. Therefore, the necessary pressure increase Δp provided by the supply fan is caluculated as a sum of pressure drops in each component and the demanded static pressure :

$$\Delta p = \Delta p_{filter} + \Delta p_{heating\ coil} + \Delta p_{cooling\ coil} + p_{static} \qquad (13)$$

3.8   Radiators

The exponential relationship between heat emission and temperature difference of radiator and zone is implemented. The characteristic curve of a radiator is related to nominal values (see Table 2) on a radiator test bench /4/:

$$\dot{Q} = \dot{Q}_{nom} \cdot (\frac{\Delta \vartheta_{ln}}{\Delta \vartheta_{ln,nom}})^n . \qquad (14)$$

Hereby the logarithmic temperature difference is computed with

$$\Delta \vartheta_{ln} = \frac{\vartheta_{supply} - \vartheta_{return}}{ln\left(\frac{\vartheta_{supply} - \vartheta_{zone}}{\vartheta_{return} - \vartheta_{zone}}\right)} . \qquad (15)$$

4 Simulation Cases

The simulation results for the reference case and for all fault cases described in the following section are one-year periods time series and yearly energy consumtions. As an example Figure 8 and 9 show simulated process variables for a three-day midseason period in the reference case (no faults).

The yearly energy consumption in the reference case is for electricity (fans and pumps) 30.1 $kWh/m2$, for cool (cooling coil) 19.1 $kWh/m2$ and for heat (heating coils and radiators) 69.1 $kwh/m2$.

Considering surveys on typical faults in VAV systems /12, 24, 30/ the following faults are investigated in the presented study:

- a clogged air filter,

- a defective supply air temperature sensor,

- a defective preheat coil water pump,

- a fouled cooling coil with a reduced heat transfer performance.

4.1 Clogged Filter

The Figures 11 and 12 show the revolving fan speeds in the reference case and in the case that the air filter is clogged. Usually, the fault is detected by the pressure drop increase monitored at the air filter. So, it is also assumed that the pressure instrument at the air filter which is indicated in Figure 3 is defect. The black bars result from the on/off switching of the plant during workdays. The gaps in both plots indicate the weekends.

The full filter is simulated by increasing the flow resistance R. The filter pressure drop at design air flow rate is now 400 Pa instead of 130 Pa in the reference case.

Figure 14 shows the influence of this fault on main process variables and yearly energy consumption. The fault has no influence on zone temperatures. The additional pressure drop is compensated by higher revolving fan speeds. This leads to an increase of air temperature at the supply fan. But the small temperature increase can be easily compensated by the supply air temperature controller. On few hours in summer the cooling coil valve position is therefore up to 19.1 % higher than it is in the reference case. But the influence on total energy consumption for cooling remains negligible. Only the yearly electricity consumption of the air handling unit (chiller excluded) is 46.9 % higher than the consumption in the reference case.
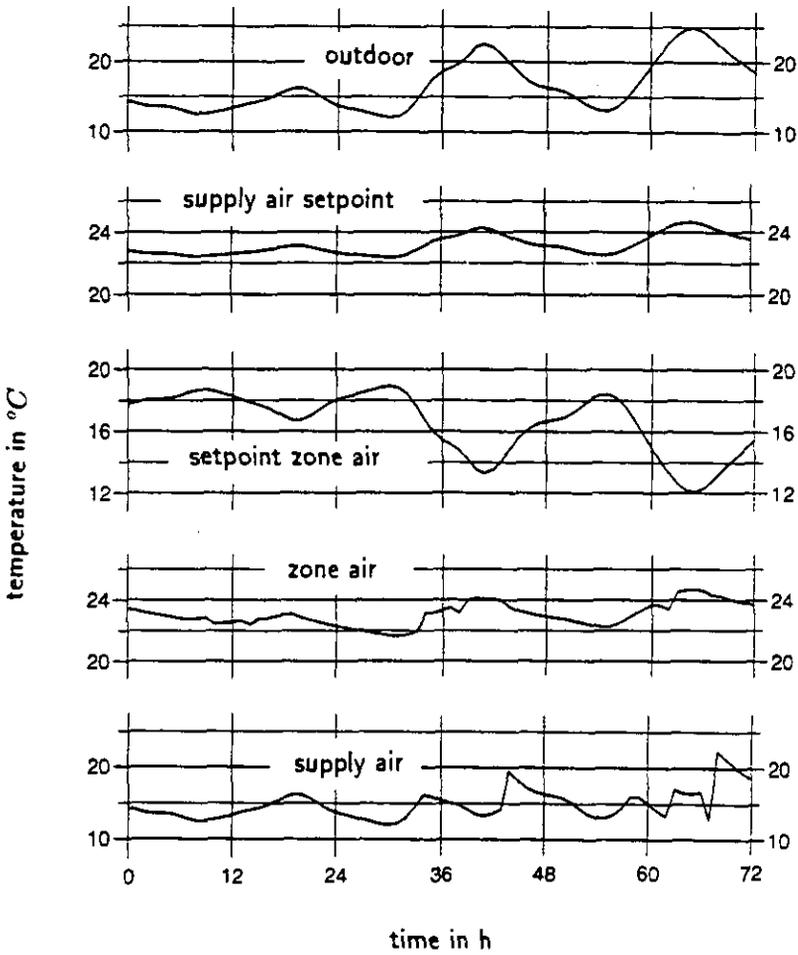
Figure 8: Simulated temperatures for a three-day midseason period
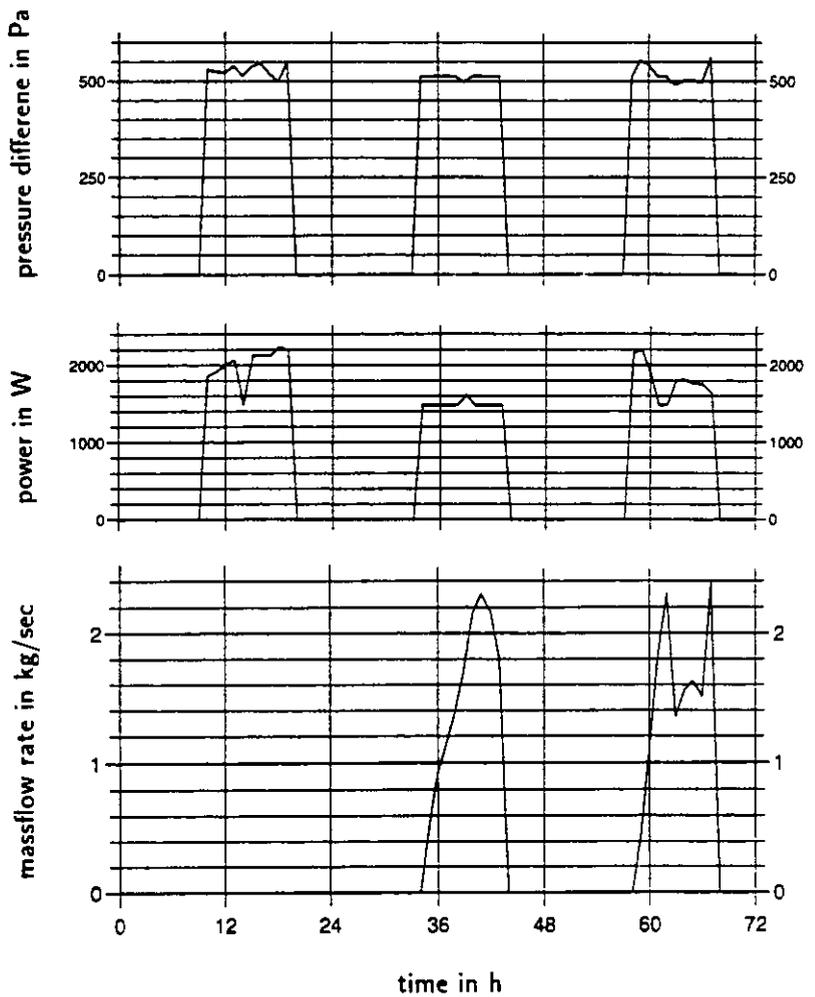
Figure 9: Simulated process variables for a three-day midseason period (top: pressure increase provided by supply fan, mid: supply fan electrical power, bottom: cooling coil water massflow rates)
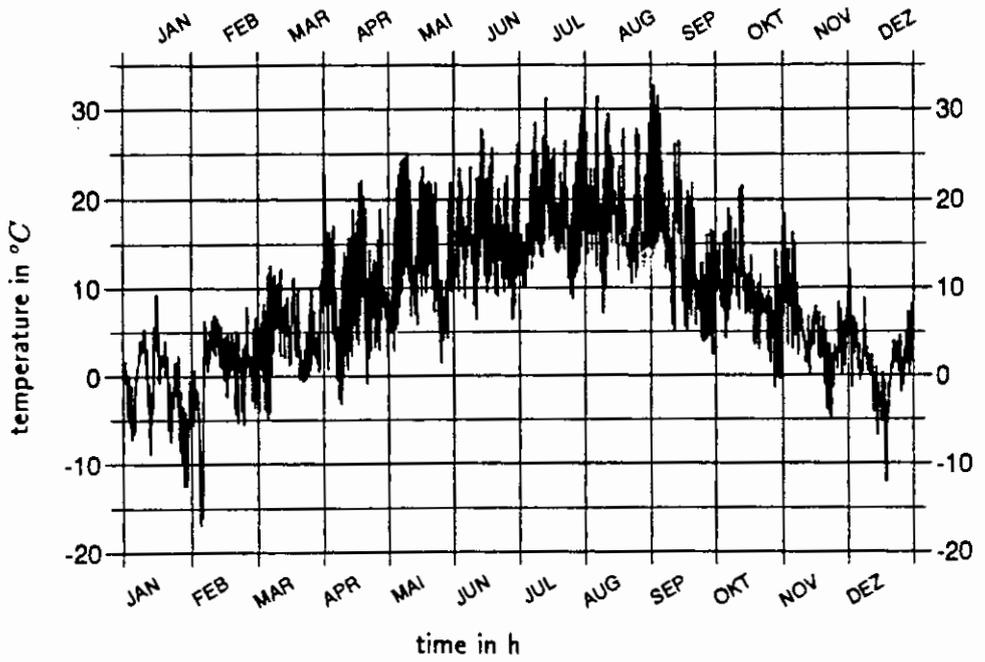
Figure 10: One-year-period outdoor temperatures
(from [TRY05])
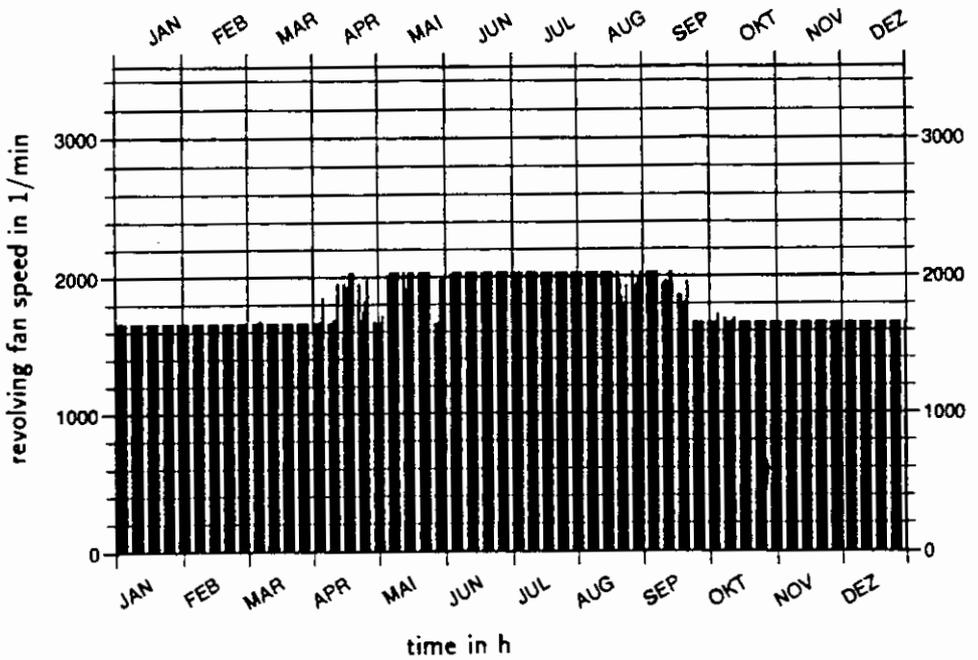
Figure 11: Revolving fan speeds (reference case)
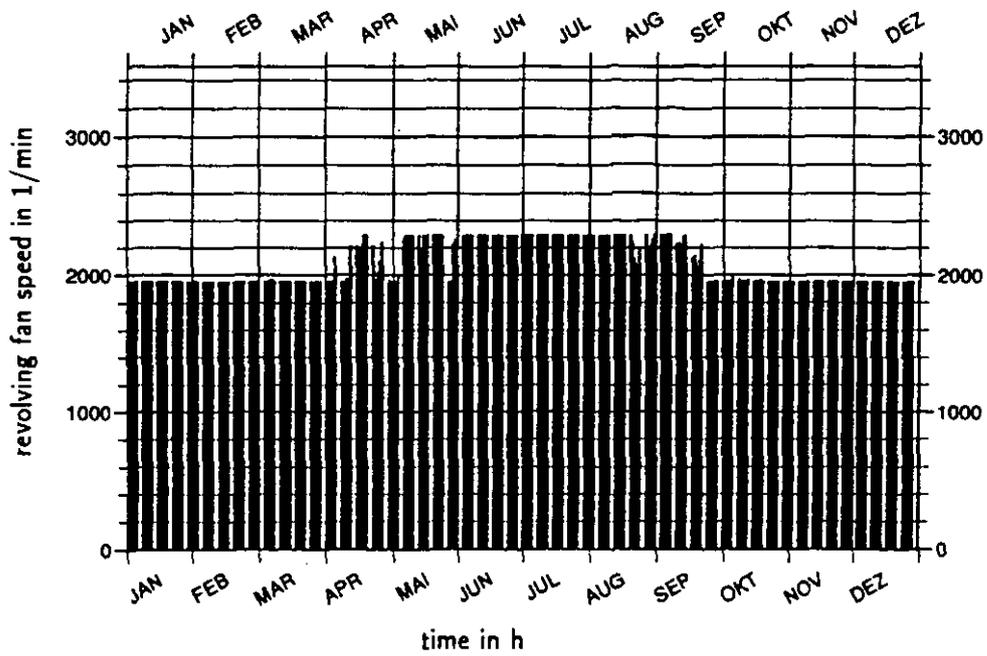
Figure 12: Revolving fan speeds (full filter)
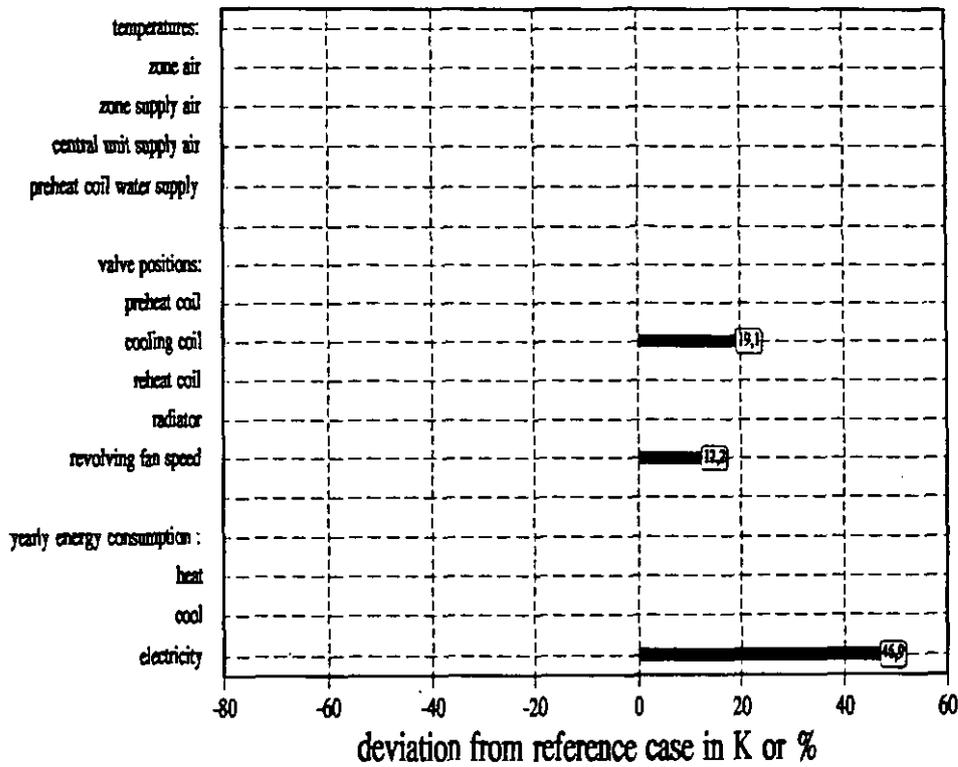
Figure 13: Influence of a full filter on main process variables and yearly energy consumption

164

When using even a static fan model (as presented in section 3.2) the revolving fan speed seems to be a robust process variable for the automatic detection of such a fault or similar faults. Of course, the value of the actual air flow rate must be also available. This might be computed from the position of the dampers in the VAV-terminals.

4.2 Defective Sensor

It is now assumed that during the whole year the measured value of the supply air temperature is 2 K lower than the real value.

The Figures 14 and 15 show the cooling coil water massflow rates in the reference case and in the case that the supply air temperature sensor fails. The Figures 16 and 17 show the reheat coil water massflow rates in both cases. Figure 18 summerizes the influence on main process variables and yearly energy consumption. This fault is compensated in summer by higher air flow rates and it is compensated in winter by lower reheat massflow rates. Therefore, nearly no influence on thermal comfort occurs.

The fault has a stronger impact on thermal comfort, if the reheat coils are smaller, or if no radiators are in the zones.

The influence on total energy consumption for cooling is strong. The yearly consumption is 35 % lower than the consumption in the reference case. But this does not mean that the original setpoint is not optimal! The lower energy consumption is achieved by lower thermal comfort (higher zone temperatures in summer). The yearly heat consumption is 2 % higher than the consumption in the reference case. Because sometimes the higher supply air temperatures are not needed on warm winter days. On such days the reheat valves are closed in both cases.

4.3 Defective Pump

In this case it is assumed that the warm water pump in the preheat coil is defect. Due to bad hydraulic balance the water massflow rate in this circuit is assumed to be not zero but reduced 75 %.

The Figures 19 and 20 show the preheat coil supply water temperatures in the reference case an in the case when the pump is defect. Figure 21 summerizes the influence of the defect on main process variables and yearly energy consumption. Again the fault has nearly no influence on thermal comfort.

The yearly heat consumption is 13.9 % lower than the consumption in the reference case. But again the lower energy consumption is achieved by lower thermal comfort (this time lower zone temperatures in winter). Due to full load limits the reheating coils and the radiators cannot fully compensate the fault.
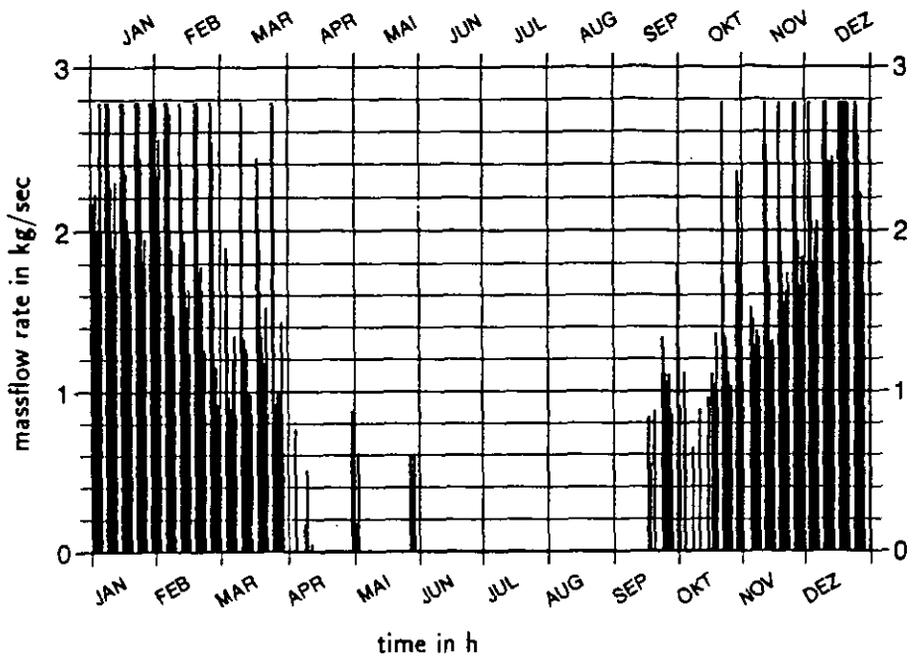
Figure 16: Reheat coil water massflow rates (reference case)
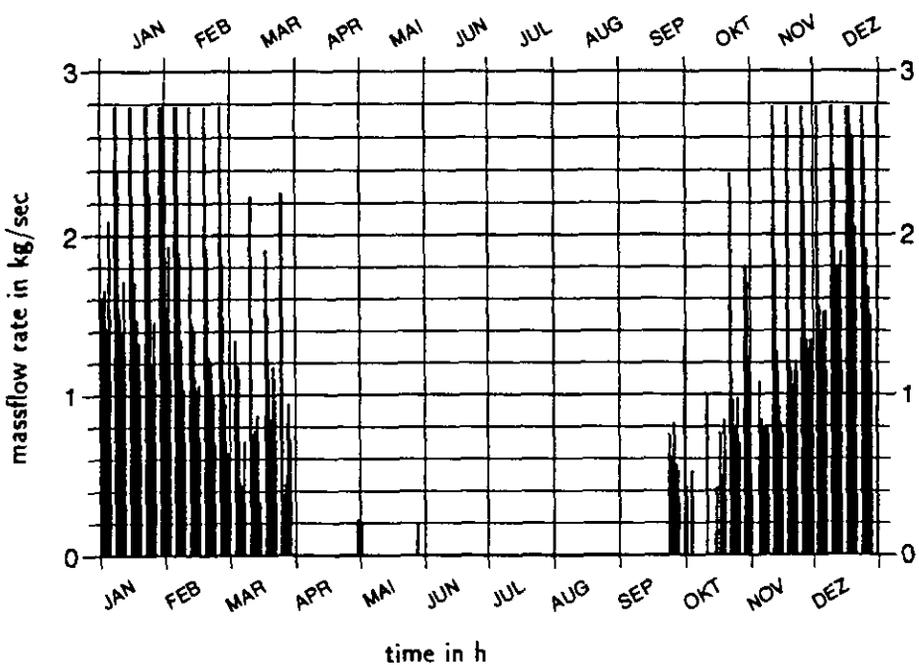
Figure 17: Reheat coil water massflow rates (defect sensor)
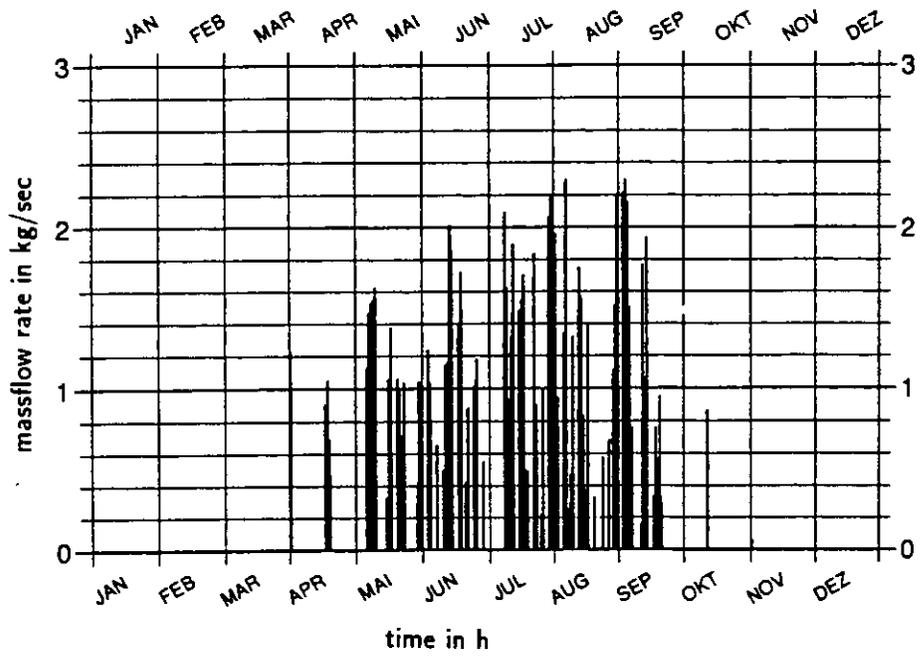
Figure 14: Cooling coil water massflow rates (reference case)
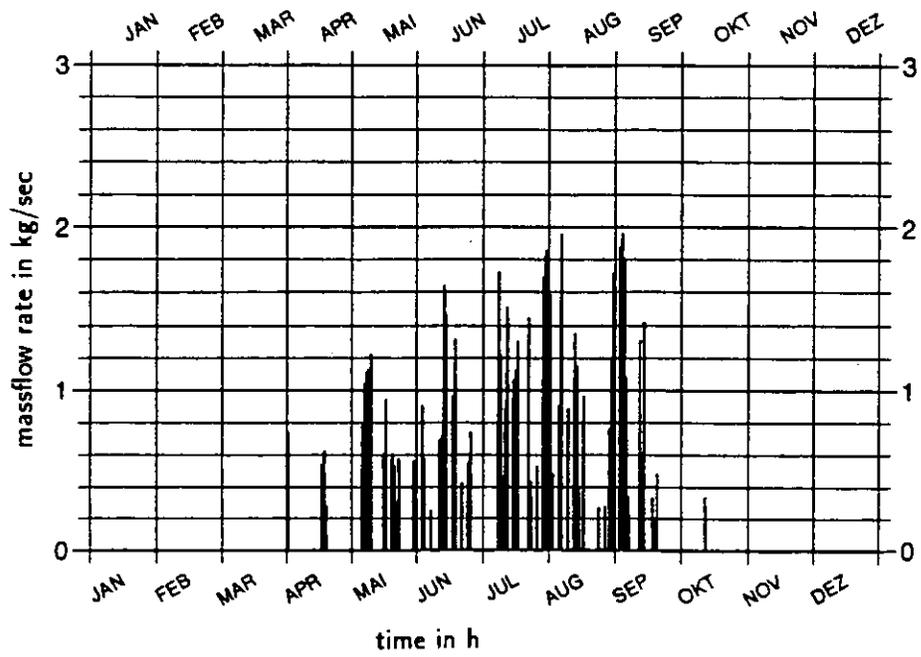
Figure 15: Cooling coil water massflow rates (defect sensor)
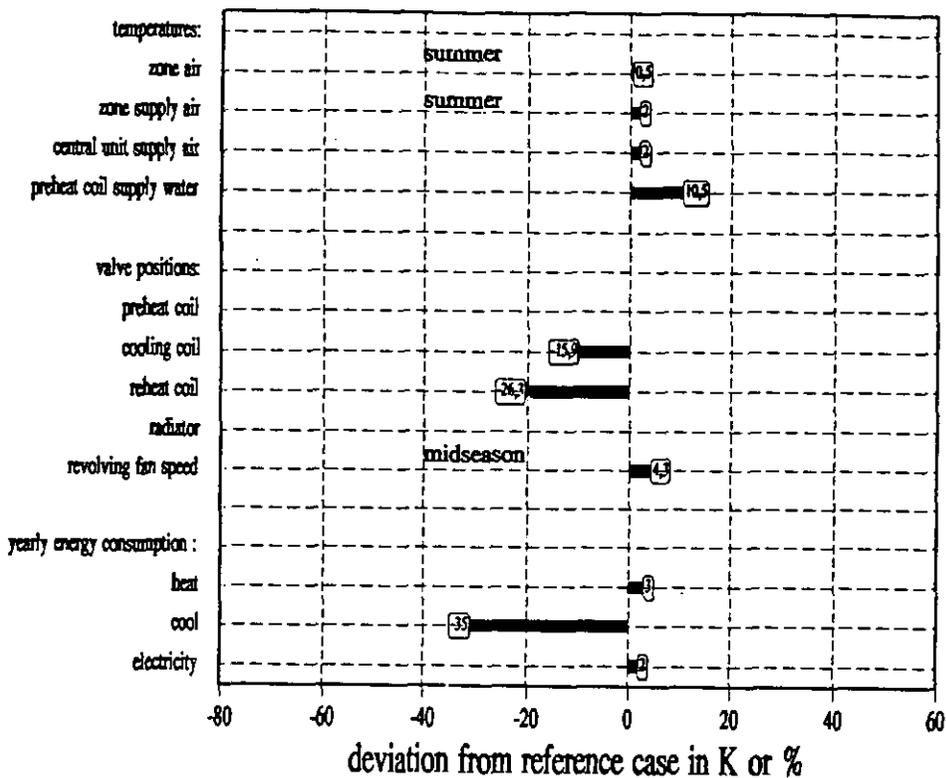
Figure 18: Influence of a defect sensor on main process variables and yearly energy consumption
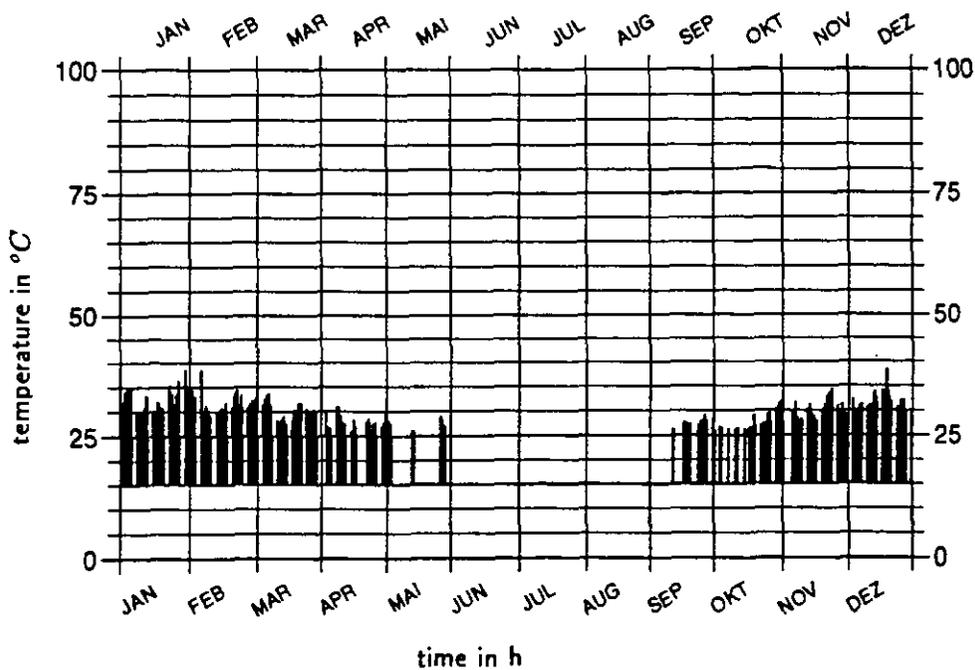
Figure 19: Preheat coil water temperatures (reference case)
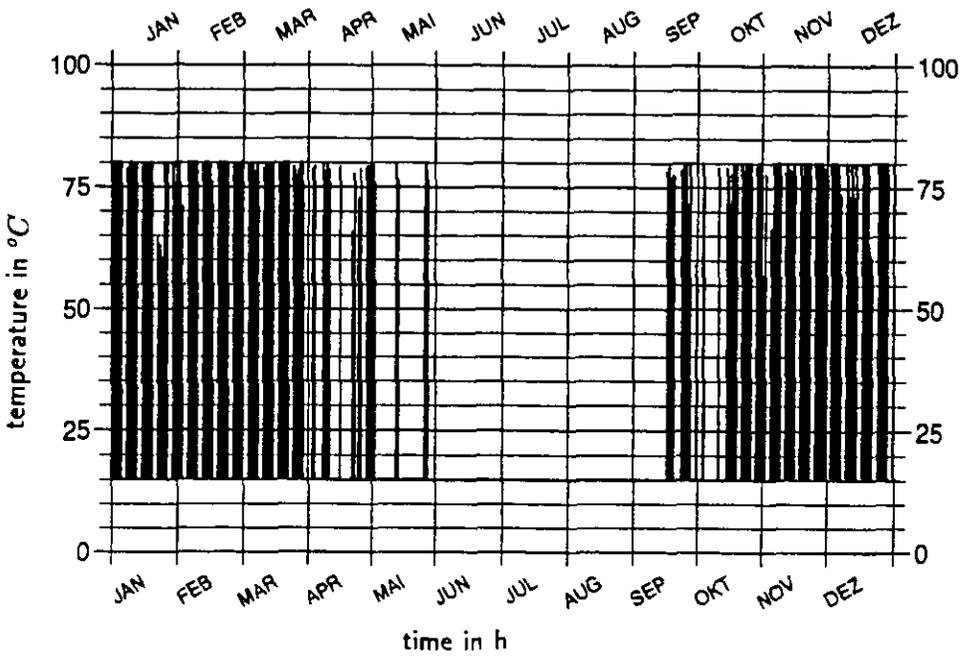
Figure 20: Preheat coil water temperatures (defect pump)
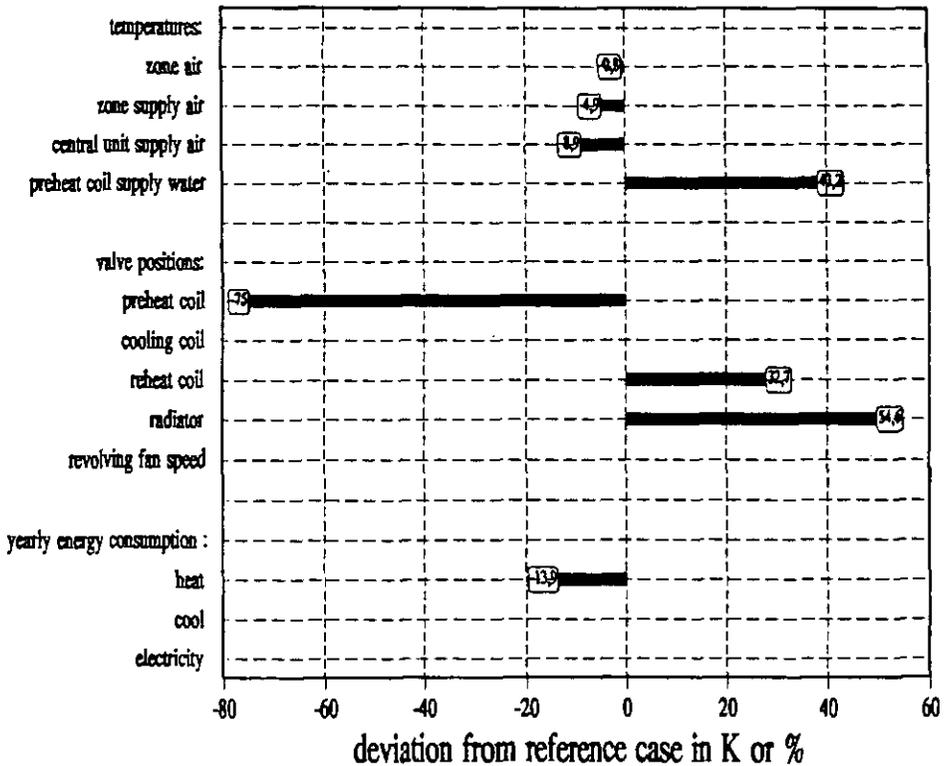
Figure 21: Influence of a defect pump on main process variables and yearly energy consumption

169

4.4 Fouled Coil

It is assumed that due to water side fouling the cooling coil effectiveness is simply 50 % reduced in all partial load conditions:

$$\Phi_{fouledcoil} = 0.5 \cdot \Phi_{reference} \left(\dot{m}_{Air} , \dot{m}_{Wat} \right) \tag{16}$$

The constant factor reduction in effectiveness might be a too simple assumption. But there is still no experimental investigation available on the impact of fouling on the thermal performance of a coil. Fouling should increase the thermal resistance between the two fluids and so far it reduces the UA-value and the effectiveness of a coil. Probably it influences the coil performance stronger at low flow rates.

The simulation results show that the reduced coil performance is compensated by higher chilled water massflow rates. Figures 22 shows the cooling coil water massflow rates in the case of the fouled cooling coil. The Figures 23 and 24 show the air temperatures in a zone in the reference case and in the case of the fouled coil. In opposite to the other cases thermal comfort is strongly violated on summer days. The reason is that the fault compensation is limited by the maximum coil waterflow rate. This leads to higher air supply temperatures (up to 5 K). The cooling loads in the zones must be therefore exported with higher airflow rates.

Figure 25 shows the influence of the fault on main process variables and yearly energy consumption.

The water massflow rate might be a good process variable for the automatic detection of such a fault. A more practical value is the corresponding valve position.

By using even a static coil model based on calibrated characteristic curves (as presented in section 3.3) such a fault could be easily detected.

In the presented example also the exhaust water temperatures of the heating and cooling coil are monitored. The measured values are used to calibrate the coil models. So, here also the coil exhaust water temperature might be a good process variable for the detection of such a fault. But often this value is not monitored.

5 Conclusion

The simulation results demonstrate the applicability of system simulation as a generic tool when constructing fault detection and diagnosis systems. Simulation aids choosing the best process variables to be monitored for the detection of certain faults. As a result of the presented simulation study **fault/symptom relationships** as shown in Figure 26 can be derived. Figure 26 shows the relationship between the investigated
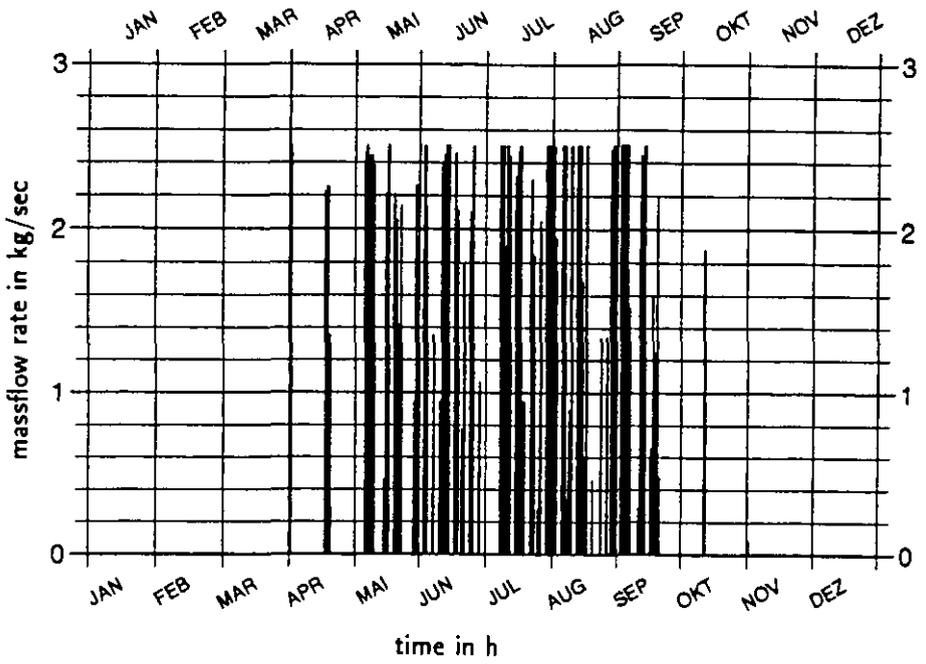
Figure 22: Cooling coil water massflow rates (fouled coil)
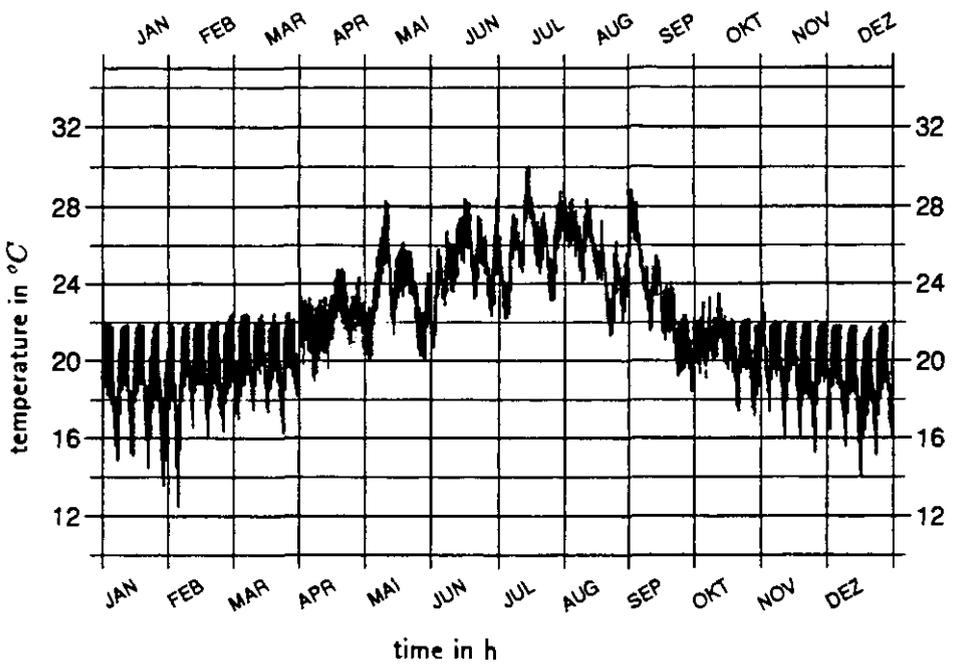
Figure 23: Zone air temperatures (reference case)
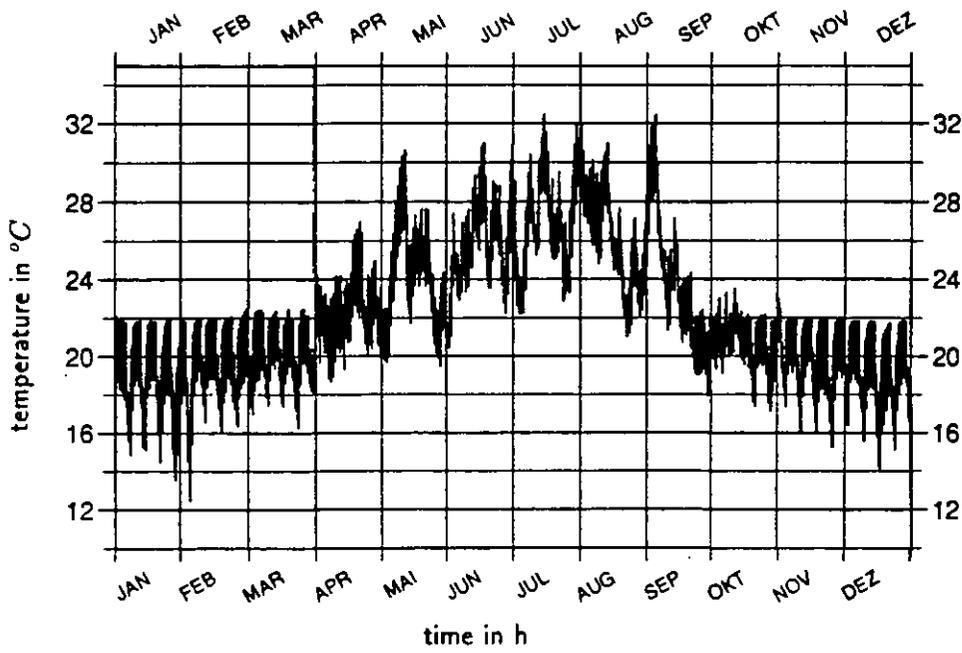
Figure 24: Zone air temperatures (fouled coil)
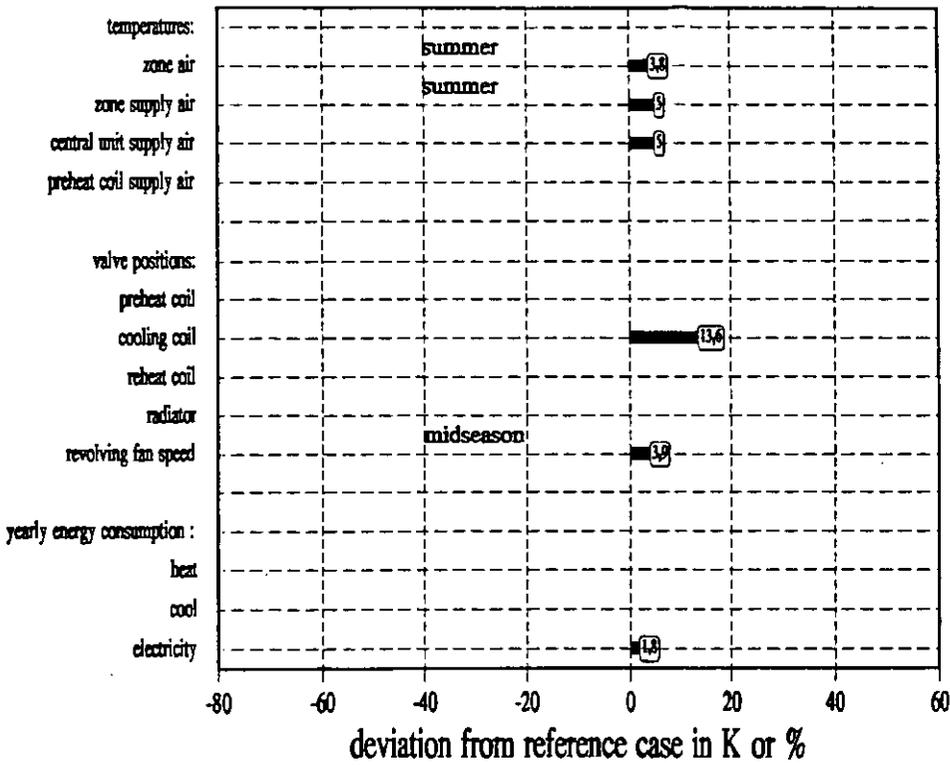
Figure 25: Influence of a fouled coil on main process
variables and yearly energy consumption

172
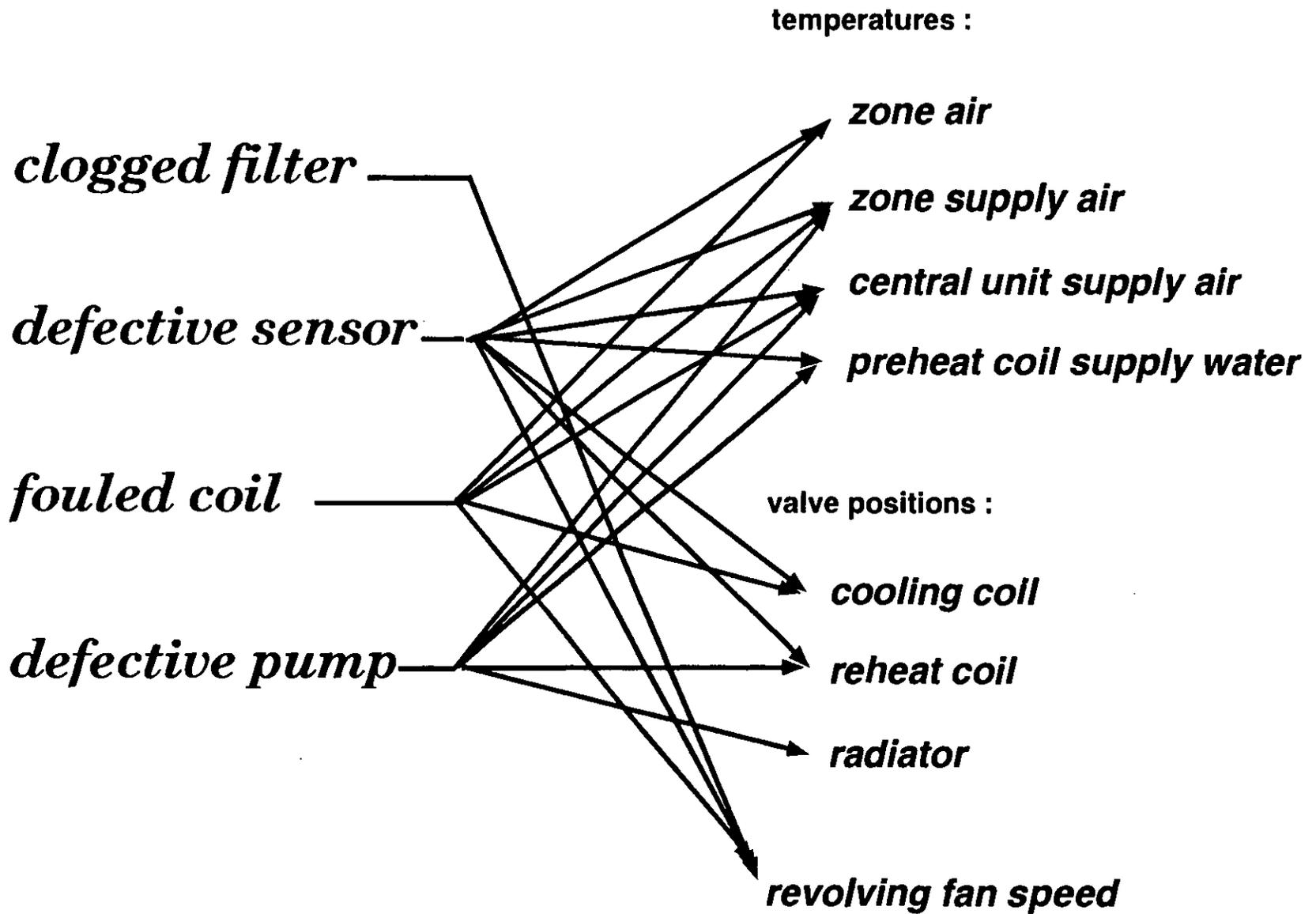
Figure 26: Relationship between faults and process variables

faults in the presented example and typical process variables which are usually already monitored for control and energy management purposes.

The impact of operational faults on yearly energy consumption and thermal comfort is shown. But the computed values cannot be generalized. They strongly depend on concept, sizing and control strategy of an HVAC system. Due to small sizing the presented VAV system runs only on few weeks of the year with variable air flow rates. In summer most of the time maximum available airflow rates are needed to export the cooling loads from the zones. Also the faults according to the heating mode (defective sensor, defective preheat pump) have a stronger impact on thermal comfort than it is shown here, if the system is equipped with smaller reheat coils, or if no radiators are available in the zones. For more systematic classifications further investigations are required. They will be executed within IEA Annex 25.

6 References

1. Bauer, M.; Madjidi M.: *Treshold Examples in Commissioning HVAC Systems*, IEA Annex 25 discussion paper AN25/FRG/250993/1.0, University of Stuttgart, Zurich, September 1993.

2. Blümel, K.; Hollan, E.; Kähler, M.; Peter, R.; Jahn, A.: *Entwicklung von Testreferenzjahren (TRY) für Klimaregionen der Bundesrepublik Deutschland.* final report BMFT-FB-T 86-051, July 1986.

3. Corrado, V.; Mazza, A.: *Parameter Selection of an Axial Fan.* IEA Annex 17 working report, Politecnico di Torino, November 1991.

4. *DIN 4704: Prüfung von Raumheizkörpern.* Deutsches Institut für Normung, Berlin.

5. *VDI 2076: Leistungsnachweis fürr Wärmetauscher mit zwei Massenströmen.* Verein Deutscher Ingenieure, Düsseldorf, 1969.

6. *DIN 24185: Prüfung von Luftfiltern.* Deutsches Institut für Normung, Berlin.

7. *DIN 24163 Ventilatoren - Leistungsmessung, Normprüfstände.* Deutsches Institut für Normung, Berlin, 1985.

8. Elmahdy, A.H., Biggs, R.C. : *Finned Tube Heat Exchanger: Correlation of Dry Surface Heat Transfer Data.* ASHRAE Transactions Vol 85 Part 2, pp. 262-268, Atlanta 1979.

9. Fornasieri, E., Mattarolo L.: *Airside Heat Transfer and Pressure Loss in Finned Tube Heat Exchangers: State of Art.* University of Padova, proceedings of the 3rd conference on finned tube heat exchangers, Stuttgart, 1991.

10. Holmes, M.J. : *Heating and Cooling Coils.* IEA Annex 10 specification AN10-880212-01, Ove ARUP Research and Development London, 1988.

11. Hyvarinen, J.:*Static Models and Characteristic Curves.* IEA-Annex 25 working document, VTT, Espoo, August 1992.

12. Kelly, G. E.: *Common Faults in Air Conditioning Systems*. IEA Annex 25 working document, NIST, Delft, October 1991.

13. Kelly, G. E.: *Description of a Reference Air Handling System*. working document, AN25/USA/11092/2.0, NIST, Liege, October 1992.

14. Kohonen, R., Hyvärinen, J.: *A System Concept for Real-Time Simulation of HVAC Systems for Building Optimization, Fault Detection and Diagnosis (BOFD)*. CIB-Conference, VTT, Rotterdam, October 1992.

15. Lebrun, J.: *IEA-ANNEX 17: Building Energy Management Systems, Evaluation and Emulation Techniques*. synthesis report, University of Liege, March 1993.

16. Madjidi, M.: *Simulation of Faults in a VAV-System - Preliminary Example*. IEA Annex 25 working report AN25/FRG/250193/1.0, University of Stuttgart, Tokyo, 1993.

17. Madjidi, M.: *Design Data and Preliminary Measurements in a VAV system for model calibration and fault detection purposes*. IEA Annex 25 working report AN25/FRG/150993/1.0, University of Stuttgart, Zurich, 1993.

18. Madjidi, M.: *First Trial on the Calibration of a Coil Model for Applications in a VAV Fault Detection System*. working paper, AN25/FRG/150993/1.0, University of Stuttgart, Zurich, 1993.

19. Malmström, T.G.; Olsson, L.G. :*Ducts*. IEA-Annex 10 Componet Specification, Royal Institute of Technology, Stockholm, 1987.

20. Nakahara, N.: A guide to sensors for BEMS, IEA Annex 16 building and energy management systems : user guidance, final report, ISBN 0946075611, Nagoya University, May 1991.

21. Pakanen, J.: *Classification of faults as a method to eliminate problems of treshold adjustment*. IEA Annex 25 working paper, VTT, Espoo, July 1993.

22. Rossi, T., Braun, J.: *Classification of Fault Detection and Diagnosis Methods*. IEA Annex 25 working document, University of Purdue, Liege, 1992.

23. Seem, J.E.: *Modelling of Heat in Buildings*. Universität Wisconsin, PhD thesis, Madison, 1987.

24. Schiel, H.F.: *Extract of the Summary Concerning the Questionnaire Modules of a Ventilation and Air Conditioning Plants*. IEA Annex 25 working document, Liege, October 1992.

25. Schöckle, M.; Hinkelmann, M.: *Development of a Fault Detection Tool Applying Associative and Rule Based Methods*. IEA Annex 25 working document, University of Stuttgart, Boston, March 1994.

26. Schmidt, F.; Hinkelmann, M.; Madjidi, M; Bach, H.: *Optimization of HVAC Systems by Integration of Planning, Simulation and Operation*. proceedings of Clima 2000 conference, University of Stuttgart, London, October 1993.

27. Schuler, M.: *Extension of Type 56 - Calculation of the Radiation Transmittance Depending on Incident Angle and Glazing Type*. TRNSYS News, Vol. 4 Nr. 1, University of Wisconsin, Madison, November 1991.

28. TRNSYS 14.1: *A Transient System Simulation Program*. Solar Energy Laboratory, University of Wisconsin, Madison, USA 1994.

29. Wright, J.A.; Hanby, V.I. :*Fans*. IEA-Annex 10 Component Specification, University of Liverpool, University of Technology Loughborough, 1988.

30. Yoshida, H.; Yamada, A.; Yuzawa, H.; Iwami, T.: Suzuki, H.: *List of Typical Faults of AHU Systems*. IEA Annex 25 working document, Yamatakee-Honeywell, Liege, October 1992.

12. Kelly, G. E.: *Common Faults in Air Conditioning Systems.* IEA Annex 25 working document, NIST, Delft, October 1991.

13. Kelly, G. E.: *Description of a Reference Air Handling System.* working document, AN25/USA/11092/2.0, NIST, Liege, October 1992.

14. Kohonen, R., Hyvärinen, J.: *A System Concept for Real-Time Simulation of HVAC Systems for Building Optimization, Fault Detection and Diagnosis (BOFD).* CIB-Conference, VTT, Rotterdam, October 1992.

15. Lebrun, J.: *IEA-ANNEX 17: Building Energy Management Systems, Evaluation and Emulation Techniques.* synthesis report, University of Liege, March 1993.

16. Madjidi, M.: *Simulation of Faults in a VAV-System - Preliminary Example.* IEA Annex 25 working report AN25/FRG/250193/1.0, University of Stuttgart, Tokyo, 1993.

17. Madjidi, M.: *Design Data and Preliminary Measurements in a VAV system for model calibration and fault detection purposes.* IEA Annex 25 working report AN25/FRG/150993/1.0, University of Stuttgart, Zurich, 1993.

18. Madjidi, M.: *First Trial on the Calibration of a Coil Model for Applications in a VAV Fault Detection System.* working paper, AN25/FRG/150993/1.0, University of Stuttgart, Zurich, 1993.

19. Malmström, T.G.; Olsson, L.G. :*Ducts.* IEA-Annex 10 Componet Specification, Royal Institute of Technology, Stockholm, 1987.

20. Nakahara, N.: A guide to sensors for BEMS, IEA Annex 16 building and energy management systems : user guidance, final report, ISBN 0946075611, Nagoya University, May 1991.

21. Pakanen, J.: *Classification of faults as a method to eliminate problems of treshold adjustment.* IEA Annex 25 working paper, VTT, Espoo, July 1993.

22. Rossi, T., Braun, J.: *Classification of Fault Detection and Diagnosis Methods.* IEA Annex 25 working document, University of Purdue, Liege, 1992.

23. Seem, J.E.: *Modelling of Heat in Buildings.* Universität Wisconsin, PhD thesis, Madison, 1987.

24. Schiel, H.F.: *Extract of the Summary Concerning the Questionnaire Modules of a Ventilation and Air Conditioning Plants.* IEA Annex 25 working document, Liege, October 1992.

25. Schöckle, M.; Hinkelmann, M.: *Development of a Fault Detection Tool Applying Associative and Rule Based Methods.* IEA Annex 25 working document, University of Stuttgart, Boston, March 1994.

26. Schmidt, F.; Hinkelmann, M.; Madjidi, M; Bach, H.: *Optimization of HVAC Systems by Integration of Planning, Simulation and Operation.* proceedings of Clima 2000 conference, University of Stuttgart, London, October 1993.

27. Schuler, M.: *Extension of Type 56 - Calculation of the Radiation Transmittance Depending on Incident Angle and Glazing Type.* TRNSYS News, Vol. 4 Nr. 1, University of Wisconsin, Madison, November 1991.

28. TRNSYS 14.1: *A Transient System Simulation Program.* Solar Energy Laboratory, University of Wisconsin, Madison, USA 1994.

29. Wright, J.A.; Hanby, V.I. :*Fans.* IEA-Annex 10 Component Specification, University of Liverpool, University of Technology Loughborough, 1988.

30. Yoshida, H.; Yamada, A.; Yuzawa, H.; Iwami, T.: Suzuki, H.: *List of Typical Faults of AHU Systems.* IEA Annex 25 working document, Yamatakee-Honeywell, Liege, October 1992.

ARX MODELS AND
REAL-TIME MODEL-BASED DIAGNOSIS

V.E. Bakker, H.C. Peitsman
TNO Building and Construction Research
Delft, The Netherlands

Abstract

The approach used in this research is based on a combination of two technologies:
- Model-based reasoning
- ARX modelling
Model-based reasoning is a technique capable of finding possible diagnoses based on behaviour descriptions and interconnections of the separate components contained in the whole system. It can best be understood as the interaction of observation and prediction. On one side there is the actual device, on the other side the model, that can make predictions about the behaviour of the device. A significant difference between the actual observations and the model predictions indicates the monitored device has a malfunction.

The behaviours are defined using ARX models. An ARX model of the complete system is used to detect a degrading performance of the system. After the system model has indicated a possible malfunction, the model-based reasoning process will try to calculate a diagnosis using the ARX models of the separate components.

The fault detection process and the model-based reasoning process are captured in a programmable real-time environment, that controls the selection of the models and the used model-based reasoning algorithm.

The joint evaluations exercise was based on a VAV system of NIST.

1 SYSTEM IDENTIFICATION

In the system identification approach described in this report, two steps are distinguished:

- system level

- component level

The system level is used to detect faults in the system, based on a degrading performance of the system. After a fault has been detected, the fault has to be diagnosed on component level. The use of a system model for detection and component models for diagnosis has the advantage that system models are easier to calculate, and hence faster in use and more suitable for real-time applications. Continuous propagation of the component model is slower and might lead to deviations in the estimated values. This is further elaborated in the real-time chapter.

A black-box model (ARX-type) is used to predict the output of the system based on the inputs of the system. The predicted output is compared with the 'measured' output to detect faults which occur in the system.

Three data sets were used; first one without malfunctions, second one with one known and two unknown malfunctions and one with four unknown malfunctions. The healthy data set (no faults) is used for training the models and the other two to detect the fault.
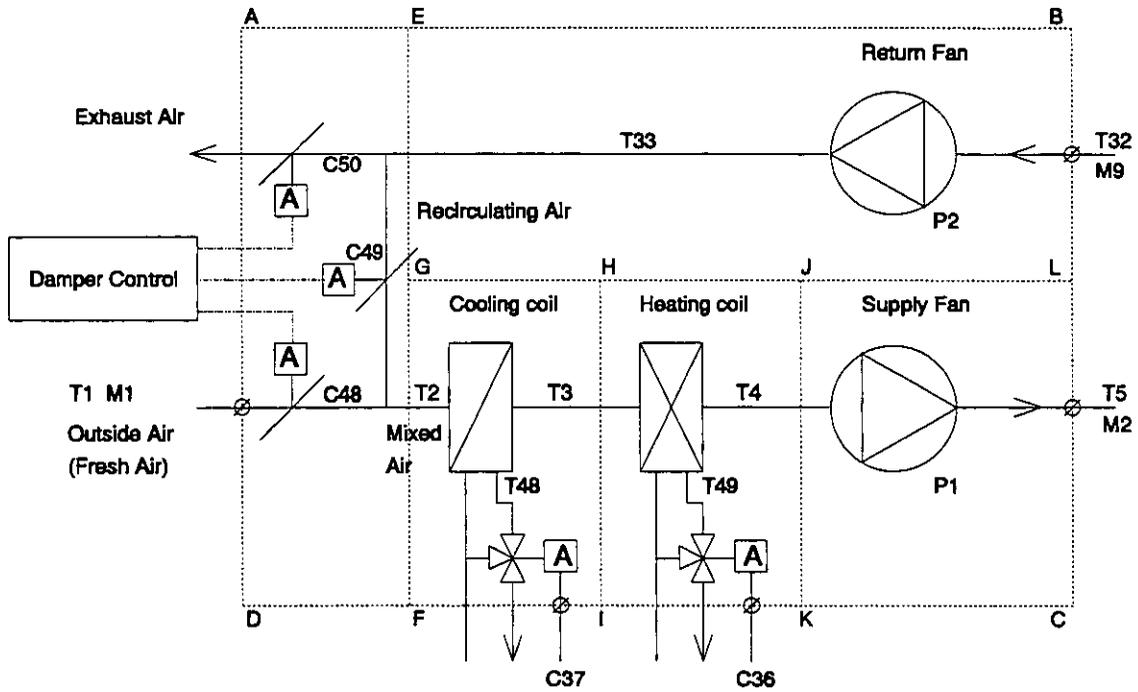
Figure 1 A simple PI-diagram of the AHU of the VAV system of NIST

* C=Ctrl, M=Flow, P=Power, T=Temp.

The next models can be defined:

ABCD	= System model AHU of the VAV system.
AEDF	= Component model Mixing box.
GHFI	= Component model Cooling coil.
HJIK	= Component model Heating coil.
JLKC	= Component model Supply fan.
EBGL	= Component model Return fan.

1.1 SYSTEM MODEL
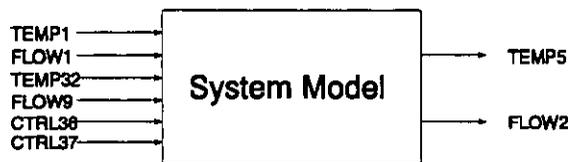
Figure 2 The system model ABCD of the VAV system

178

Inputs:

temp1 flow1 temp32 flow9 ctrl36 ctrl37

ctrl36*ctrl36 ctrl37*ctrl37 flow9*flow9 temp32*temp32 temp1*flow9 temp32*flow1

The structure of the system model is an

  ARX(5,5) with 12 inputs.
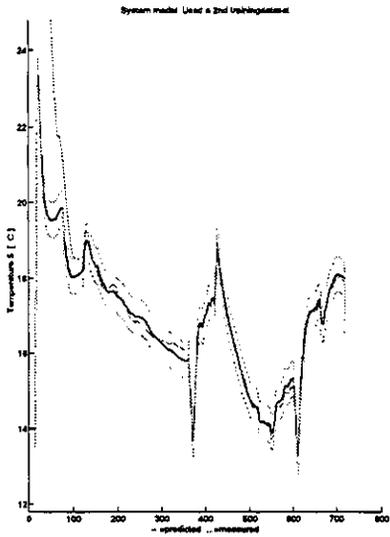
Figure 3a Results ARX model on system level TEMP5
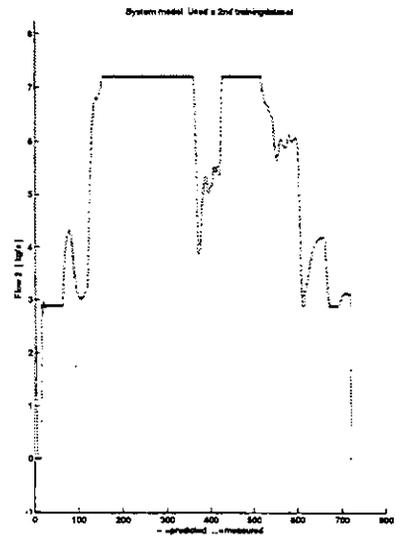
Figure 3b Results ARX model on system level FLOW2
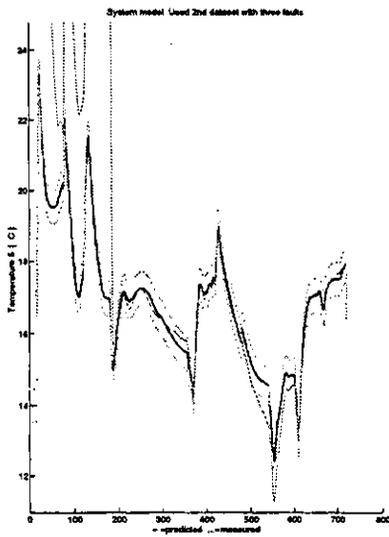
Figure 4a System level TEMP5 with dataset 3 faults.
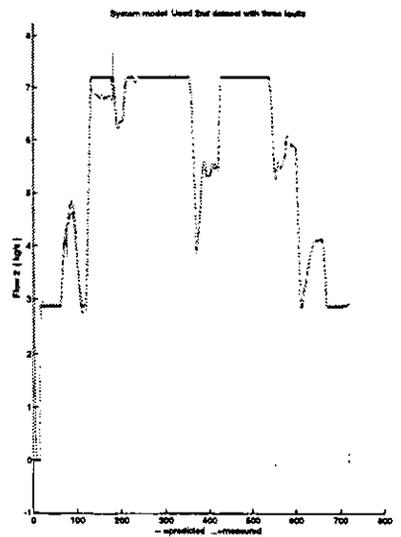
Figure 4b System level FLOW2 with dataset 3 faults.

179

The training results illustrated in Figure 3 show that the system models of TEMP5 and FLOW2 fit well. The margins depicted in the figures are exactly 3 times the standard deviation of the residuals of the training results. This margin is generally used in system identification to determine if a fault occurs. In Figure 4 the results for the dataset holding three faults are presented.

1.2 COMPONENT MODEL

For the component model five components were distinguished.

<u>Return fan model</u>

Figure 5 The Return Fan model EBGL of the VAV system

Inputs:

power2 temp32

The structure of this component model is an

ARX(1,1) with 2 inputs.

<u>Mixing box model</u>

Figure 6 The Mixing Box model AEDF of the VAV system

Inputs:

ctrl48 ctrl49 ctrl50 temp1 temp33

temp1*temp1

The structure of this component model is an

ARX(4,4) with 6 inputs.

<u>Cooling coil model</u>
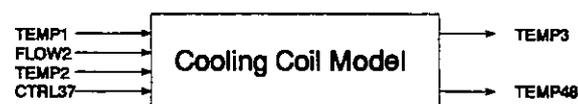
Figure 7 The Cooling Coil model GHFI of the VAV system

Inputs:

ctrl37 temp1 temp2 flow2

ctrl37*temp1 ctrl37*temp2 ctrl37*flow2 temp2*flow2 ctrl37*ctrl37 temp1*temp1 flow2*flow2

The structure of this component model is an

ARX(4,4) with 11 inputs.

Heating coil model

Figure 8 The Heating Coil model HJIK of the VAV system

Inputs:

temp1 flow1 temp3 flow9 ctrl36

temp3*flow1 temp3*flow9 temp1*temp1 flow1*flow1 temp3*temp3 flow9*flow9 ctrl36*ctrl36

The structure of this component model is an

ARX(2,2) with 12 inputs.

Supply fan model

Figure 9 The Supply Fan model JLKC of the VAV system

Inputs:

temp4 power1 flow9 flow1

The structure of this component model is an

ARX(1,1) with 4 inputs.

In Figure 10 some component model results for the dataset with three faults are depicted. The component models of the heating coil and supply fan are not presented because no faults were found in these components.
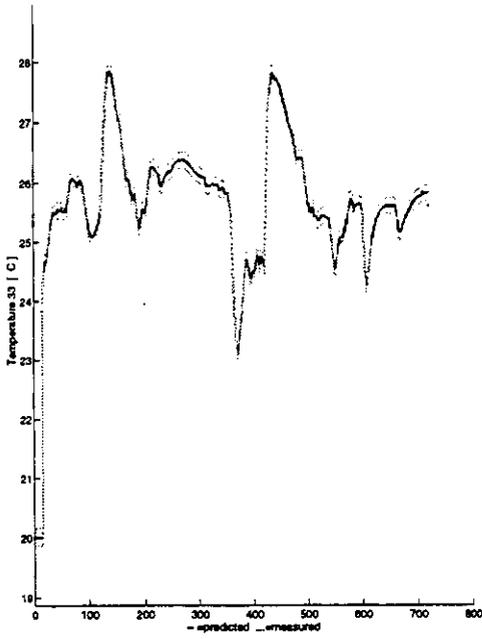
Figure 10a Return fan results, TEMP33
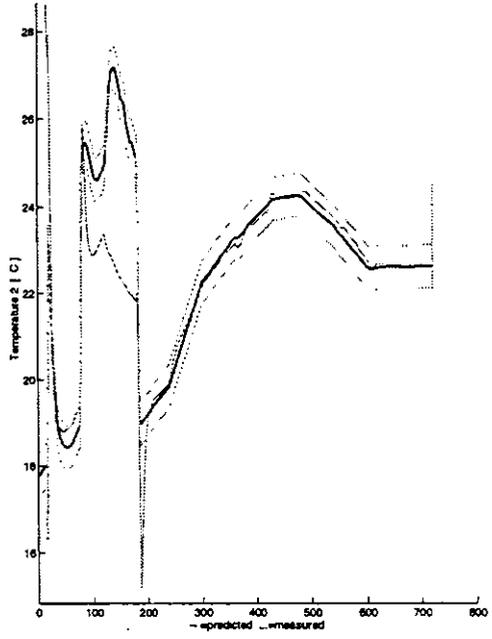
Figure 10b Results Mixing Box, TEMP2
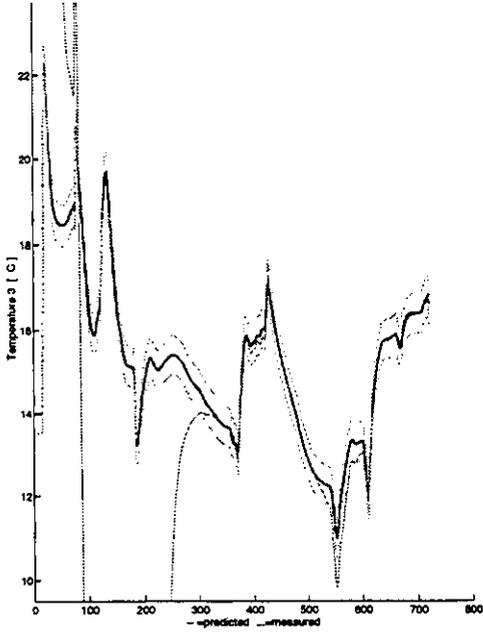
Figure 10c Cooling Coil Results, TEMP3
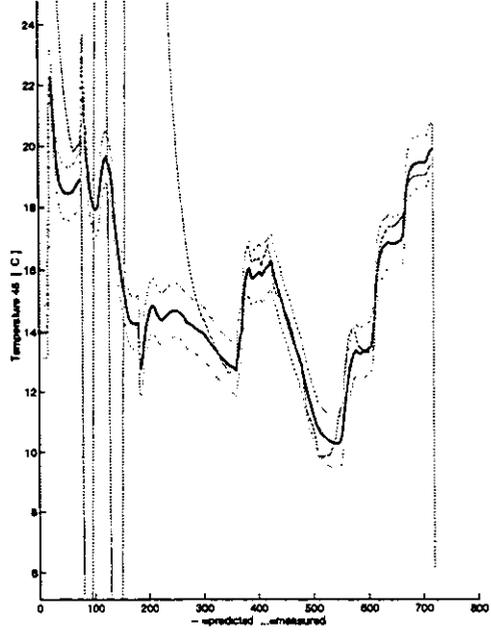
Figure 10d Cooling Coil Results, TEMP48

182

2 DIAGNOSING THE FAULTS

2.1 DATA SET WITH THREE FAULTS

Detection of the first fault

System level :

The system model detects a fault after 60 minutes. The next faults are detected:

- the predicted TEMP5 is much higher than the observations (figure 3a),

- the predicted FLOW2 is lower than the observations (figure 3b).

After the fault detection on component level, the Model Based Diagnosis (MBD) program switch to component level.

Component level :

The next step is to diagnose the fault in the component:

- *Return fan*

    No fault is detected in the return fan

- *Mixing box*

    Fault is detected in Mixing box; Symptom: the predicted TEMP2 is much lower than the observations (figure 10b).

- *Cooling Coil*

    Fault is detected in Cooling Coil; Symptom: the predicted TEMP3 is much lower than the observations (figure 10c).

- *Heating Coil*

    No fault detected in Heating Coil

- *Supply Fan*

    No fault detected in Supply Fan

Diagnosis

Fault in Mixing Box or Cooling Coil

The cooling coil is situated after the mixing box, therefore the fault in the cooling coil can be caused by the mixing box.

Detection of the second fault

System level :

The system model detects a second fault after 500 minutes. The next fault is detected:

- the predicted TEMP5 is lower than the observations (figure 4a),

Component level :

The next step is to diagnose the fault in the component:

- *Return fan*

    No fault is detected in the return fan

- *Mixing box*

    No fault is detected in Mixing Box

- *Cooling Coil*

    Fault is detected in Cooling Coil; Symptom: the predicted TEMP3 is lower than the observations (figure 10c) and the predicted TEMP48 is lower than the observations (figure 10d), but between the allowed margins

- *Heating Coil*

    No fault detected in Heating Coil

- *Supply Fan*

    No fault detected in Supply Fan

Diagnosis

Fault in Cooling Coil

2.2 DATASET WITH FOUR FAULTS

Detection of the first fault

System level :

The system model detects a fault after the startup period of the VAV system. The next fault is detected:

- the predicted TEMP5 is much higher than the observations,

Component level :

The next step is to diagnose the fault in the component:

- *Return fan*

    No fault is detected in the return fan

- *Mixing box*

    No fault is detected in Mixing Box

- *Cooling Coil*

    Fault is detected in Cooling Coil; Symptom: the predicted TEMP3 is higher than the observations (figure 10c)

- *Heating Coil*

    No fault detected in Heating Coil

- *Supply Fan*

    No fault detected in Supply Fan

184

Diagnosis

Fault in Cooling Coil

Note: possible false alarm, because the fault is in the startup period of the VAV system

Detection of the second fault

System level :

The system model detects a fault after 500 minutes. The next fault is detected:

-       the predicted TEMP5 is much lower than the observations (figure 4a)

Component level :

The next step is to diagnose the fault in the component:

-       *Return fan*

        No fault is detected in the return fan

-       *Mixing box*

        No fault is detected in Mixing Box

-       *Cooling Coil*

        Fault is detected in Cooling Coil; Symptom: the predicted TEMP3 and TEMP48 are much lower than

        the observations (figure 10c, 10d)

-       *Heating Coil*

        No fault detected in Heating Coil

-       *Supply Fan*

        No fault detected in Supply Fan

Diagnosis

Fault in Cooling Coil

3 MODEL-BASED REASONING

Model-based diagnostic methods are used for localisation of defective components in (technical) systems. The kernel of model-based diagnosis is the model, which describes the functionality of the concerned apparatus.

The principle of model-based reasoning for diagnosis can best be understood as a combination of observation and prediction. In the one hand we have an actual device, in the other hand we have a model of that device that can make predictions about its intended behaviour. Observations indicate what the device is actually doing, predictions indicate what it's supposed to do. Any difference between these two indicates a problem in the most broad definition. The difference between the system and its model is called a discrepancy.

Model-based reasoning has several advantages over the use of rule-bases:

One advantage is formulated by Dvorak [Dvorak, 1991]:

"Unlike many diagnostic methods, model-based diagnosis does not rely on a set of symptom-fault patterns. Such patterns are often incomplete, since it is difficult for an expert to anticipate all possible faults and predict their symptoms, especially the symptoms of interacting faults."

Another advantage of MBD is that even with a new system, for which no repair experience exists, MBD can be used; it only needs the model, which is always obtainable. In this context it can also be stated that the creation/setup-time for an MBD is relatively short (if only the model needs to be created).

Different reasoning methods have been developed for model-based diagnosis. An important method for model-based reasoning is General Diagnostic Engine (GDE), defined by De Kleer & Williams [Kleer & Williams, 1987]. In the GDE approach, diagnoses (candidates with De Kleer & Williams) are deducted from conflicts. A conflict is a collection of components whose correct function leads to inconsistencies. All existing conflicts are generated by a process called constraint propagation. Constraint propagation performs calculations on the model in as many different ways as possible.
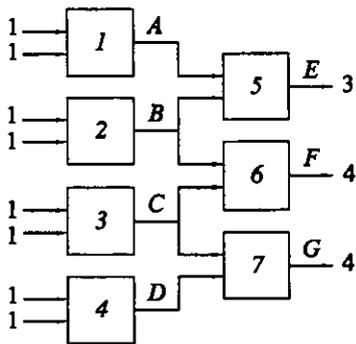
Figure 11 A collection of adders

This can be illustrated by Figure 11. Output A of component 1 can be calculated assuming component 1 works (result $A=2$), but it can also be calculated assuming components 2 and 5 are functional (result $A=1$). The result of each propagation step is combined with the corresponding assumptions made and recorded in a dependency record. If different values are calculated for one and the same output, the assumptions made for the calculations are inconsistent and result in a conflict. The collection of conflicts for Figure 11 is given in Table 1.

dependency records	conflicts
[A=2, {1}] [A=1, {2,5},{3,5,6},{4,5,6,7}]	<1,2,5>, <1,3,5,6>, <1,4,5,6,7>
[B=2, {2},{3,6},{4,6,7}] [B=1, {1,5}]	<1,2,5>, <1,3,5,6>, <1,4,5,6,7>
[C=2, {3},{2,6},{4,7}] [C=3, {1,5,6}]	<1,3,5,6>, <1,4,5,6,7> [1]
[D=2, {4},{3,7},{2,6,7}] [D=1, {1,5,6,7}]	<1,4,5,6,7>

Table 1: Dependency records and conflicts using GDE on Figure 11.

When conflicts are transformed to diagnoses, only minimal conflicts and minimal diagnoses are processed, which means superfluous elements in the sets are omitted. In this example, <1,2,5> is a minimal conflict, while <1,2,5,6> contains one more element <6>.
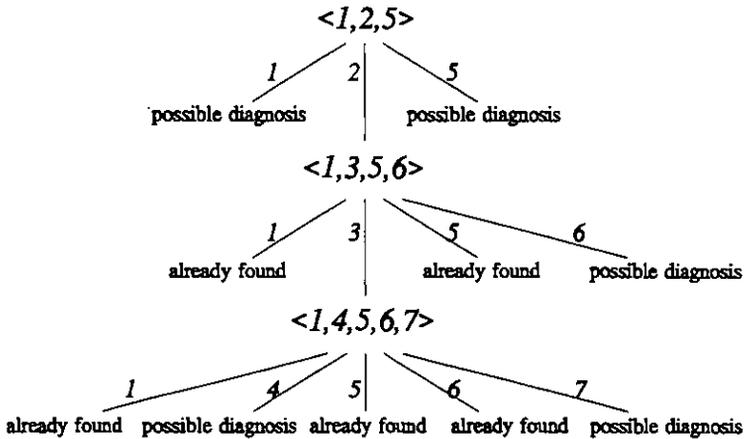
Figure 12 Generation of diagnoses from conflicts using a Hitting Set Tree

A proper diagnose has to clarify each discrepancy; it has to contain at least one component assumption of every conflict. The transformation of conflicts to diagnoses is illustrated by Figure 12. This collection of diagnoses/conflicts is called the *Hitting Set Tree* ([Reiter 1987]).

The hitting set tree, of which the nodes represent the generated conflicts, can be used to search for diagnoses that 'cover' every conflict. A diagnose that 'covers' every conflict contains at least one component of each conflict. A diagnose consists of the labels of a path from a 'diagnose-leave' of the tree to the root. Paths starting at a repeating leaves form a not minimal diagnose.

---

[1] Assumptions that are already contained in a conflict are not used for further calculations. <1,2,5,6> is not regarded a new conflict because <1,2,5> already is. This is part of the *"assumption-based truth maintenance system"* that GDE uses: "unreliable environments are not used for further calculations".

In the Diagnose application, the GDE algorithm is used. No special arguments underline the choice of GDE for the Diagnose application; however, it can be stated that the separation of conflict calculation and diagnosis generation in GDE is very suitable for time-efficient FDD, for as long as the system is functioning correctly, only conflict calculation has to be performed.

Once the conflicts have been calculated, generating diagnoses is a quite straight-forward process. A diagnose is an assumption that explains every encountered conflict, i.e. a conflict "lightswitch is on" and "the lightbulb does not glow" can be explained by the diagnose "LIGHTSWITCH is DEFECT". In practice, only minimal diagnoses are generated; if there exists a diagnose "LIGHTSWITCH is DEFECT" the possible diagnose "LIGHTSWITCH is DEFECT and LIGHTBULB is DEFECT" is ignored.

The generation of diagnoses from conflicts comes down to a breadth-first search method. The generation process is already illustrated in Figure 12. The nodes of the tree represent conflicts, the leaves represent possible diagnoses corresponding to the elements of the conflicts. An 'Already found' on the lowest level of the tree indicates a correct diagnose that covers each conflict.

3.1 MODEL REQUIREMENTS

The number one question when creating a model for MBD is 'What is a suitable model for MBD?'. A useful definition of the desired properties of an MBD model is given in [Soest, 1993]:
'An adequate model is one that enables a troubleshooter advisor to advise efficiently on the cheapest effective repair'

Unfortunately, as good as the properties of the MBD tool are defined, as difficult it is to transform the final MBD demands to model properties. Still, five requirements are clearly stated by [Soest, 1993]:

1) the states should be uniquely identifiable;
2) as few components as possible;
3) as few connections as possible;
4) total cost of probing points as low as possible;
5) behavioral descriptions as simple as possible.

These requirements still leave some questions unanswered, such as 'Which parts have to be represented in the model' and 'which influences between the parts have to be represented'. The specific information that is needed to make optimal choices on this account of the modelling process is system-dependent, and therefore dependent on the insight of the model designer. Using the term 'desirable properties' instead of 'requirements' allows for more freedom in the modelling process.

3.2 HIERARCHICAL MODELLING

Performing MBD on a system can be quite time-consuming, depending the number of components in a model. As there is an exponential relation between the number of components and the number of calculations, the time needed for constraint propagation increases fast with the complexity of the model. In [Bakker,1989] the calculation complexity for different approaches is calculated to be $O(2^n)$, where n is the number of components in the model. A trade-off between the detail of the diagnose (the number of components) and the reactiveness of the diagnose system is obvious.

This trade-off can be bypassed with *model reduction*[2]. By using different models for the same system depending the state of diagnosis, the most effective model, in other words the model adding the most information to the diagnosis, is used.

The structure of the interdependent models can be best defined as a tree; on the highest level only one model is used for fault detection. When a fault is detected the diagnosis tool descends one level in the model tree, to the model containing the different components.

As soon as a fault is globally localised, the model tree node containing the fault can be expanded. The tree structure can be derived from the system structure. An example for a Hi-Fi system is given in Figure 13.

Hi-Fi Set

| | |
Radio   Amplifier   Speakers

Antenna Power Supply Pre-Amplifier Power-Amplifier Cone Connections Cables
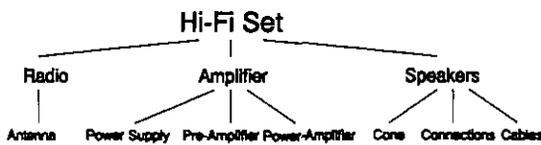
Figure 13 Example System Structure of Hi-fi Set

How the efficiency of the diagnosis process can be influenced by choosing the right subsequent abstraction levels will be illustrated by an example of a very simple model with few in- and outputs. In this example the functionality of the system is of minor interest, and therefore not defined.

3.3 LIMITS TO HIERARCHICAL MODELLING

Unfortunately, the resolution of the computed diagnoses can not be refined infinitely. The effect of defining more subsystems depends on two model properties:

- the number of available observations
- the component interconnection structure (connectivity of the model)

---

[2]   *Model Reduction - The process of discarding certain modes of motion while retaining others in the model used by an active control system, in order that the control system can compute control commands with sufficient rapidity.*

This can be made obvious by the following examples (Figure 14).
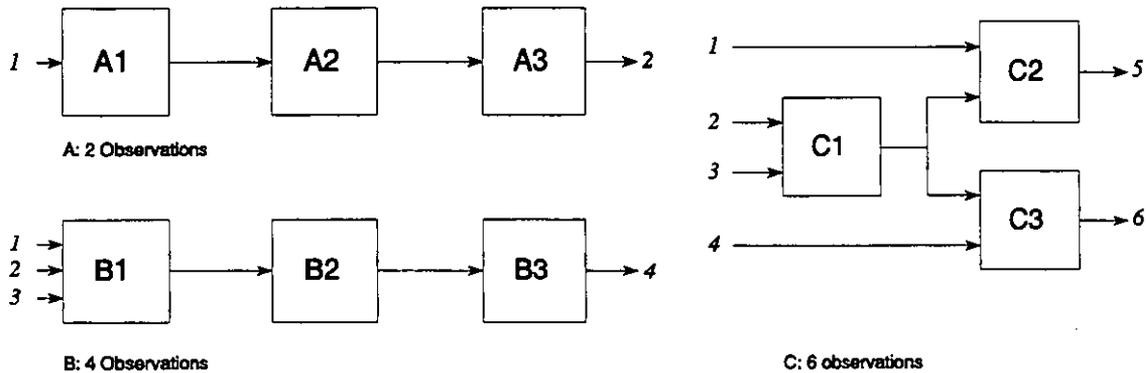
Figure 14 Observations and Connectivity

All three examples consist of three components. In model A, two observations (1 input, 1 output) are available. No observations are available to discriminate diagnose [A1] from [A2] or [A3]. In model B, four observations (3 inputs, 1 output) are available, but no observations to allow discrimination of diagnoses. The third model, model C, has six observations (4 inputs, 2 outputs). The structure of the model combined with enough observations allows for some diagnosis.

The example showed the choice of components/subsystems in the model definition should be based on a combination of the number of observations and analysis of the model structure. The intuitive restrictions that can certify the efficiency of hierarchical descriptions are discussed by [Mozetič, 1990]. In practice, it would be wise to determine the number of sensors depending the desired diagnostic performance.

The MBD process can reason about sensor faults if sensors are admitted to the model. Introducing the sensors as VAV components in leads to a model as presented in Figure 15.
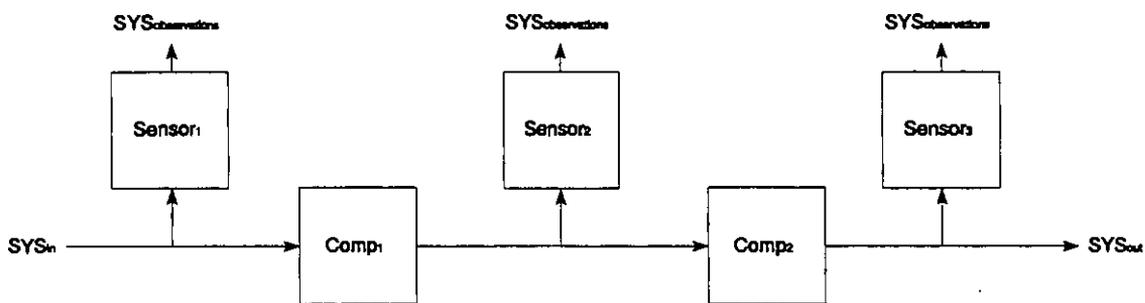
Figure 15 Component model including sensors

This model automatically implies the sensors as part of every detected conflict. Combining conflicts will eventually indicate defective sensors; i.e. if in the model presented in Figure 15 the conflicts $<Sensor_1,Comp_1,Sensor_2>$ and $<Sensor_2,Comp_2,Sensor_3>$ are detected, the intersection diagnose of the conflicts 'Sensor$_2$ Defective' is more probable than any other diagnosis.

In order to model the dependency between the model calculations and the sensors or measured values, two-port sensors are used in two configurations:

Input configuration is used to model the dependency of the model calculations on the observations; the input values of the model are always measured by sensors, so a deviation of the calculated value can also be caused by a faulty model input value, e.g. a defective sensor.

Output configuration is used to model the dependency of the observations at the outputs of the system on the sensors; any information regarding the output values of the system is collected by sensors, so a deviation of the calculated value can also be caused by a faulty output observation sensors.

3.4 MODEL STRUCTURE OF THE VAV SYSTEM

In order to decrease the introductory threshold and keep realisation costs low, all models are based on sensors that are most common in HVAC systems. The *structure* of the models will apply to VAV systems in general, the fitted ARX models are system-dependent.

The system model is used to verify the performance of the whole system. The verification of the performance is based on predictable system outputs that can be labelled representative for the system performance in general. For the VAV system as presented in Figure 1, two outputs can be defined:

- T5 Outlet air temperature after the supply fan
- M2 Supply air mass flow rate to the zones

For the component model, five separate ARX models have to be created; one for each of the identified components Mixing box, Cooling coil, Heating coil, Supply fan and Return fan. Besides the ARX models for the VAV subsystems, the component model contains a collection of sensor components. To enable model-based reasoning, the behaviours of and the connections between the separate components have to be specified. The outputs of some ARX models are inputs for other ARX models.

Note: MBD demands the connections between the components to be ideal. In the VAV model, the connections are not ideal. In assuming they are, the possibility of a broken or bad performing connection as independent diagnosis is excluded.

Explicitly introducing connection models would in this case only make sense if enough sensors are available to let the MBD system distinguish between connections and components; a luxury that does not apply to the VAV system. Since increasing the number of sensors is not desirable, connections are not explicitly accounted for in the model; they are part of the connected component. In example, in the VAV system this means that a diagnosis <Supply Fan Defect> on account of M1 (the flow) could mean that the *fan* itself is defect, as well as the *flow connection* to the fan. In general, component borders are defined by location of the sensors, as is depicted in Figure 16.
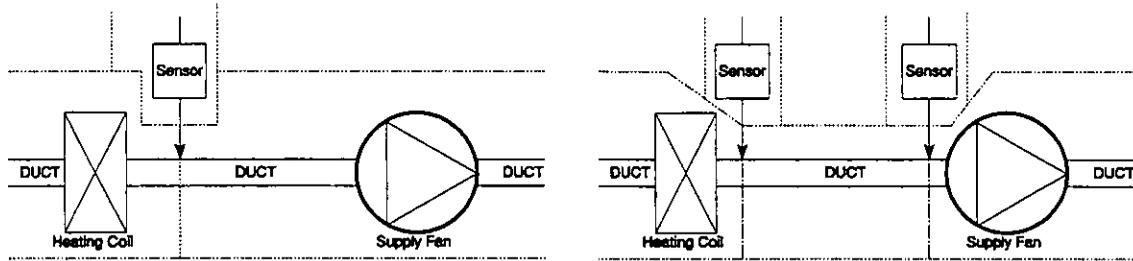
Figure 16 Connections implicitly and explicitly modelled

In Figure 17 the combination and interconnection of the separate ARX models is illustrated. The dependency of the components is clearly visible. Note that there is no complete depencency relation between any components. The effect of the partial dependencies and partial direct inputs for components in model-based reasoning is still researched.
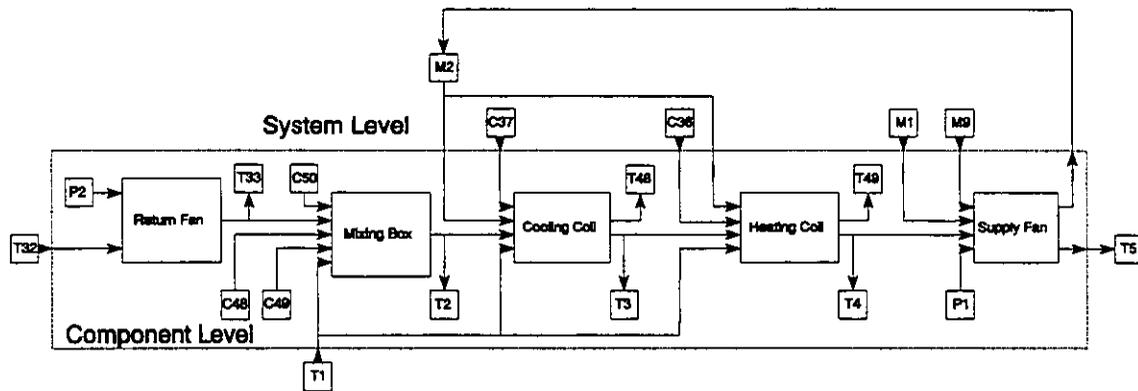
Figure 17 Interconnection of the component models

The representation of this model useable for the diagnosis tool is given in Appendix I. The effect of the use of estimated values as input for following models on the accuracy of the component model is not yet studied. Partly this is bypassed by using models on system level where possible.

4 REAL-TIME FAULT DETECTION AND DIAGNOSIS

To identify the explicit tasks of a real-time MBD system the existing situation can be analyzed. In general HVAC systems are monitored by BEMS and interpreted by system operators. The functions that a human system operator performs in real-time system monitoring/diagnosis are shown in Figure 18.
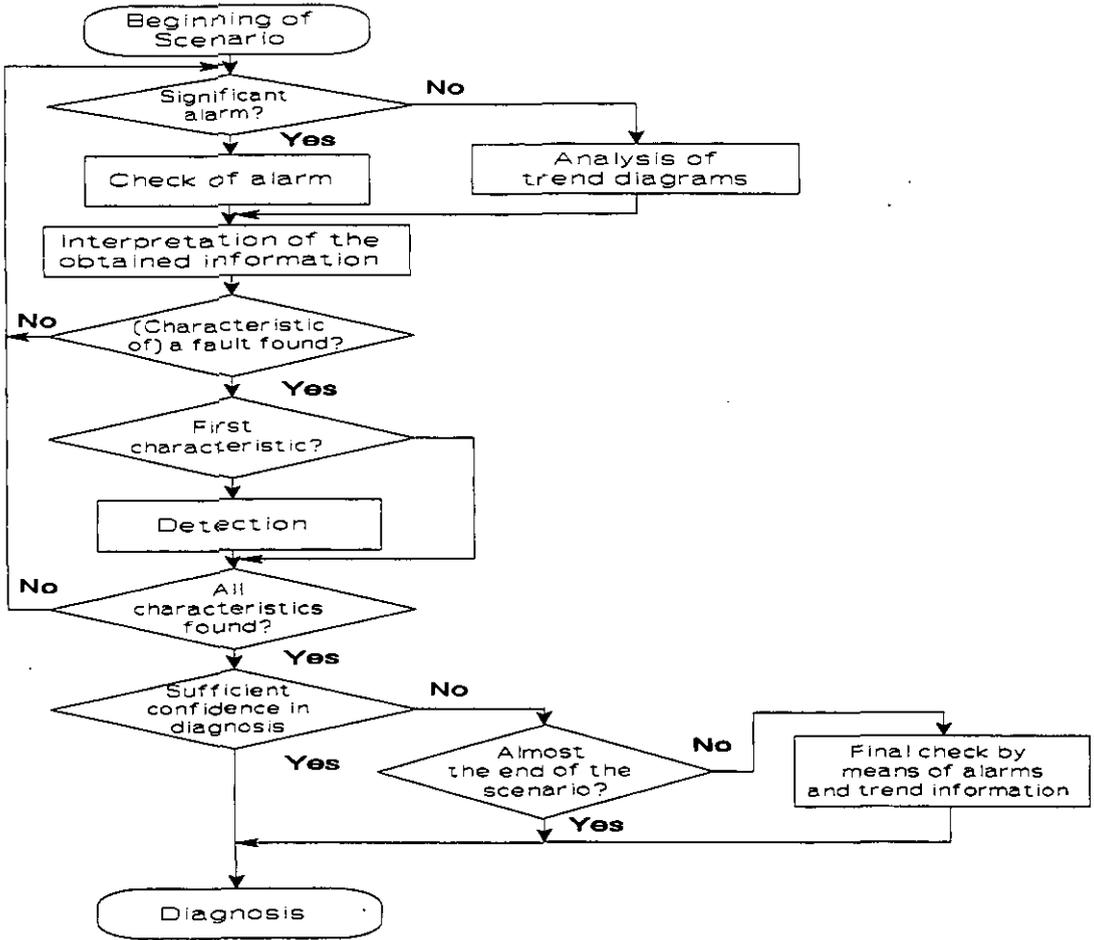
Figure 18 Operator Strategy

The actions of the system operator can be simplified into five subsequent steps:

1) Monitoring the system

2) Detecting discrepancies / conflicts

3) Creating Diagnoses

4) Advising Testvector / probing for discrimination

5) Repairing system

Basically, these are the functions the real-time MBD system must perform. Since diagnosis systems are not yet capable of mechanical probing and repairing, the operator's actions can be transformed to three system tasks:

1)   the output of the system is compared with the expected behaviour

2)   if there is a difference between actual and expected output, the difference must be explained as either an allowed deviation or an indication of malfunction

3)   if a malfunction is suspected, it must be traced

If the systems operator is replaced by a MBD system, a situation like in Figure 19 arises.
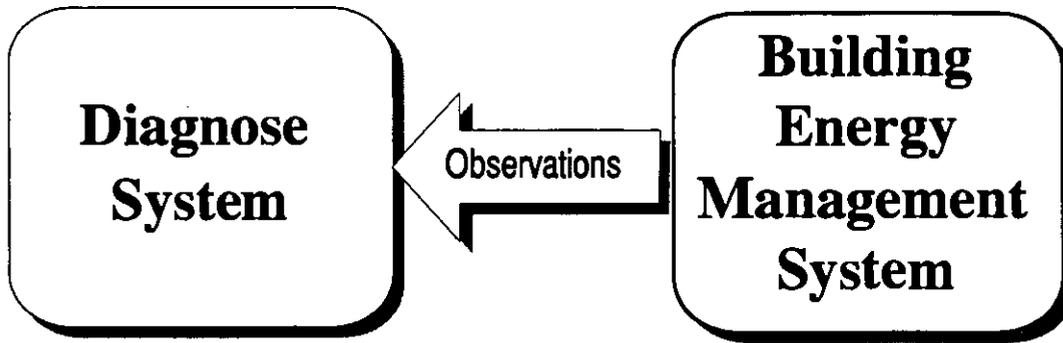
Figure 19 Monitoring the BEMS

In this structure, the MBD system is a separate system that communicates with the Building Energy Management System, which controls the HVAC system. The underlying structures of the BEMS and the MBD system will be discussed in the next chapter.

The need for flexible problem-solving strategies can be explained on two levels:

 - the system has to be flexible enough to choose the best problem-solving strategy

 - the system has to be flexible enough to cope with a variety of given strategies

The second approach is preferred, for the following three arguments:

1)   since no information is available about how real-time model-based reasoning should be accomplished on HVAC systems, it is desirable to create a tool for testing different diagnostic strategies

2)   it is most difficult to make a machine choose between different strategies if the performance demands are unknown

3)   for HVAC systems of the same type, strategies will most likely be identical, so once a optimal strategy is found it can be universally applied to similar systems

As a realisation of a semi-flexible way to let a system act like a system operator, a simple diagnostic programming language holding the different system operators actions as well as typical model-based reasoning actions is implemented.

This language should i.e. be able to perform the following actions:

- load different reasoning models depending the diagnostic situation
- load observation data if available
- perform different types of conflict calculation and diagnosis construction
- check the reasoning results
- simple imperative loop and jump constructions to let the system run automatically
- perform simple calculations for loop counting etc.
- give output on system conditions
- trigger an alarm situation

An example of the implemented language and program is given in Appendix II. A program (=real-time diagnosis strategy) was designed that accomplished the following actions:

1   Load the system model

2   Load observations and calculate conflicts until a conflict is found

3   Load the component model

4   If no conflicts are found return to system model (Step 1)

5   If conflicts are found calculate diagnoses

6   If possible diagnoses hold sensor/component combinations that can checked using the corroboration models load the corresponding corroboration model

7   check for conflicts

8   If corroboration model gives no conflicts, the corresponding sensor is indeed functioning incorrectly

Based on the functional implementation, the conclusion is justified that model-based diagnosis can be performed real-time. The implementation described in this chapter gives a possible approach. Specific tests will provide information about the actual performance of the real-time MBD system. Test results will be discussed in the next chapter.

4.1 TEST WITH CORRECT DATA SET 1 BASED ON THE VAV SYSTEM

Running the MBD program with the correct dataset should give no problems. However, one conflict was detected. This conflict occured at the start-up period of the system; the ARX models were not able to fit the start-up pulse. The conflict that was detected at system level, caused the diagnosis tool to switch to the component model to start the diagnosis procedure! Propagation of the component model did not reveal a conflict, so the system concluded it was a false alarm. In fact, the reliability of the ARX models in the startup phase is very low, since they need a certain time to reach a 'stable' condition. The figures of the training estimations illustrate this.

4.2 TEST OF DATA SET 2 WITH THREE FAULTS

<u>Fault Detection</u>

For this data set, fault detection starts after the ARX models have reached a 'stable' condition. A somewhat critical situation occured, since the first fault is close to, or almost in the unstable area of the ARX models. Still, the ARX model on system level was able to detect each of the three faults.

<u>Fault Diagnosis</u>

Using the component model to enable automatic diagnosis did not result in the correct diagnoses. For each fault the supply fan was diagnosed to be defective. It is suspected that this is caused by the narrow margin of the Supply fan model and the fact that it is the last

component in the row of connected components[3].

4.3 TEST OF DATA SET 3 WITH FOUR FAULTS

The dataset with four faults provided the same results as the dataset with three faults.

4.4 TEST OF SENSOR STUCK-AT FAULT

An other possible and common fault was manually introduced as a stuck-at sensor fault. A stuck-at fault is a situation in which a measured value becomes static, e.g. stuck at a certain value, which can be explained as a defective sensor or a broken sensor connection. For the test, the TEMP5 sensor was set at 0 °C from a certain observation on.

---

[3] Due to propagation of estimated values it is possible that a well-fitted component exceeds the margins while the component actually causing the deviation does not.

As the results illustrate, the Diagnose tool almost immediately identified the broken sensor.

```
  SYS_8:    0.000000       SYS_8:    0.349229   Marg:  0.420000
  SYS_7:    0.000000       SYS_7:   14.254783   Marg:  0.640600

Conflicts Found
Now loading Component Model...

SUPFAN 6:    0.000000     SUPFAN 6:   14.824574   Marg:  1.500000
SUPFAN 5:    0.000000     SUPFAN 5:    0.340577   Marg:  0.500000
MIXBOX 6:   17.860600     MIXBOX 6:   17.530922   Marg:  2.100000
RETFAN 3:   20.000000     RETFAN 3:   20.110648   Marg:  0.390990
FLOW9 2:     0.000000     FLOW9 2:     0.000000   Marg:  0.390990
Diagnosis 1 : [ (TEMP5, DEFECT) ]
    A-priori    : 0.010453
    A-posteriori: 0.019005
Diagnosis 2 : [ (FLOW1, DEFECT) ]
    A-priori    : 0.010453
    A-posteriori: 0.019005
Diagnosis 3 : [ (FLOW9, DEFECT) ]
    A-priori    : 0.010453
    A-posteriori: 0.019005
Diagnosis 4 : [ (TEMP4, DEFECT) ]
    A-priori    : 0.010453
    A-posteriori: 0.019005
Diagnosis 5 : [ (POWER1, DEFECT) ]
    A-priori    : 0.010453
    A-posteriori: 0.019005
Diagnosis 6 : [ (SUPFAN, DEFECT) ]
    A-priori    : 0.004951
    A-posteriori: 0.009002
```

In this example, SYS_8 corresponds with FLOW2 and SYS_7 with TEMP5. The left values are the observations, the right values the estimations. The allowed margins are also available on-screen. Obviously SYS_7 (=TEMP5) is observed to be zero, while the ARX estimation is still about 14 °C. Propagation of the component model results in 6 possible diagnoses, including TEMP5. Of course, the MBD system is unable to distinguish between the different sensors, because for that purpose, observations regarding the sensors would be needed.

5 REQUIRED RESEARCH

The research described in this report is only the beginning. A lot of questions are still unanswered. A few points of further research are:
- the mixed use of direct observations and propagated values in the component model
- the requirements of the ARX models to enable proper diagnosing (i.e. the margins)
- the effect of switching between the system level and component level models
- the consequenses of the use of historical values in the ARX estimations for MBD

REFERENCES

Bakker, V.E., 'A real-time fault detection and diagnosis application for HVAC systems', Thesis University of Twente, August 1995

Bakker, V.E., Peitsman, H.C., 'Application of black-box models to HVAC systems for fault detections', to be published in IEA-ANNEX 25, ASHRAE, USA, February 1996

Davis, R., Hamscher, W., 'Model-based reasoning: Troubleshooting', in *Exploring Artificial Intelligence; survey talks from the National conferences on Artificial Intelligence*, Morgan Kaufman Publishers, San Mateo, 1987, pp 297-346

Dexter, A.L., Benouarets, M., 'Generic modelling of HVAC plant for fault diagnosis', Paper Department of Engineering Science, University of Oxford, 1995

Dvorak, D., Kuipers, B., 'Process monitoring and diagnosis', *IEEE Expert June*, 1991, pp 67-74

Kleer, J. de, Williams, B.C., 'Diagnosing Multiple Faults', *Artificial Intelligence 32*, 1987

Kohonen, R., Hyvärinen, J., 'Real-time simulation of HVAC systems for building optimisation, fault detection and diagnosis', IEA ANNEX 25 Workplan, Technical research centre of Finland, Laboratory of heating and ventilation, Espoo, Finland

Ljung, L., 'System Identification - Theory for the User', Prentice Hall, Englewood Cliffs, New York, 1987

Moor, J.H. de, Peitsman, H.C., 'Joint Evaluation Exercise Air-Conditioning Group', IEA ANNEX 25, April 1995

Mozetiĉ, I., 'Reduction of diagnostic complexity through model abstractions', in *Working Notes First Intl. Workshop on Principles of Diagnosis*, Stanford University 1990, pp 102-111

Raaphorst, J.P.F., Peitsman, H.C., 'The use of ARX-models of the TNO reference system in the Model-Based Diagnostic engine', TNO Building and Construction Research, Delft, 1994

Soest, D.C. van, 'Modelling for model-based diagnosis', Thesis University of Twente, Enschede, 1993

APPENDIX I

In this appendix the component model, as used by the MBD tool is presented.

VAVCOMPONENT

COMPONENTS:

TEMP1   real_sensor
FLOW1   real_sensor
TEMP32 real_sensor
FLOW9   real_sensor
CTRL36 real_sensor
CTRL37 real_sensor
TEMP5   real_sensor
FLOW2   real_sensor
POWER2 real_sensor
TEMP33 real_sensor
CTRL48 real_sensor
CTRL49 real_sensor
CTRL50 real_sensor
TEMP2   real_sensor
TEMP3   real_sensor
TEMP48 real_sensor
TEMP4   real_sensor
TEMP49 real_sensor
POWER1 real_sensor

RETFAN matlab_2in_1out

DESCRIPTION: forward1
ARX_INPUTS: in1 in2
 ARX_TIME_ORDERS: [1 [1 1] [0 0]]
  ARX_COEFFICIENTS:
    -4.1194995e-001 7.9163022e-002 5.9358247e-001
ARX_MARGIN: 0.390990

MIXBOX matlab_5in_1out

DESCRIPTION: forward1
ARX_INPUTS: in1 in2 in3 in4 in5 in4*in4
 ARX_TIME_ORDERS: [4 [4 4 4 4 4 4] [0 0 0 0 0 0]]
  ARX_COEFFICIENTS:
    -6.481282832371505e-001 -6.582179894622993e-001 -7.777401291724149e-002 4.106377540038953e-001
     7.928607489118034e+000 0.000000000000000e+000 0.000000000000000e+000 0.000000000000000e+000
     0.000000000000000e+000 -1.670719803668529e+001 0.000000000000000e+000 5.069951836068407e+000
     0.000000000000000e+000 0.000000000000000e+000 3.609596103597613e+000 0.000000000000000e+000
    -4.481230593102304e+001 1.194832994638849e+002 -9.043515017315432e+001 1.579869470563974e+001
    -3.290166424167588e-001 6.508828773980396e-001 -2.211060477951576e-001 -1.000621262907474e+000
     1.097141766681237e+000 -2.905944801950731e+000 2.206250699577930e+000 -3.976353955008034e-001
  ARX_MARGIN: 0.500863

COOLCO matlab_4in_2out_a

DESCRIPTION: forward1
ARX_INPUTS: in1 in2 in3 in4 in1*in2 in1*in3 in1*in4 in3*in4 in1*in1 in2*in2 in4*in4
 ARX_TIME_ORDERS: [4 [4 4 4 4 4 4 4 4 4 4 4] [0 0 0 0 0 0 0 0 0 0 0]]
  ARX_COEFFICIENTS:
    -2.012807661788014e+000 9.611874477139300e-001 2.377856465345267e-001 -1.847711113946610e-001
    -3.983342813160675e+001 4.657855621108710e+001 1.615091705392232e+001 -2.286511153550188e+001
     2.099169802766818e+000 -2.068266293727453e+000 -7.060682302942644e+000 6.757115020178849e-001
     9.687399106564606e-002 -5.397092031803175e-002 -1.336250061516967e-001 9.166985227262406e-002
     2.501191343948820e-001 -1.636036177572186e-001 -5.844632459061060e-001 4.971508933975154e-001
    -4.983950890124521e-001 9.193447624360170e-001 -6.738185685063136e-002 -2.975285069535521e-001
     2.165698624291771e+000 -2.997458018892557e+000 -4.761767513596907e-001 1.248539498220706e+000
     1.208203283005456e+000 -1.705310413584127e+000 -1.296512012957139e-001 6.296078300905024e-001
     2.903635447608961e-003 -1.524787473569807e-002 4.252186789205873e-002 -2.997972308134883e-002
    -6.540872655192198e+000 1.120207256277638e+001 -1.423271438596947e+000 -3.226465925037212e+000
    -6.015412869400266e-002 6.094353476681819e-002 1.795990682953589e-002 -1.878020142337192e-002
    -2.543048246158963e+000 2.864231682368595e-002 -2.367673999824511e-003 -1.041857064749989e-003
  ARX_MARGIN: 0.471489

DESCRIPTION: forward2
ARX_INPUTS: in1 in2 in3 in4 in1*in2 in1*in3 in1*in4 in3*in4 in1*in1 in2*in2 in4*in4
 ARX_TIME_ORDERS: [4 [4 4 4 4 4 4 4 4 4 4 4] [0 0 0 0 0 0 0 0 0 0 0]]
  ARX_COEFFICIENTS:
    -1.396872472386783e+000 9.791815461869906e-002 2.981405424388132e-001 4.257484420335442e-003
     2.857544682843398e+002 -2.897614266881635e+002 -2.493881199522118e+002 2.534855352869522e+002
    -5.967897996961657e+000 4.390889978493386e+000 9.873085036296974e+000 -8.295506153858902e+000
     5.217269899272074e-002 1.839285567123947e-002 -5.893671933895991e-002 -9.382280870646010e-003
    -4.566776565043955e-001 5.419179902151765e-001 -6.816840167412777e-002 -3.004788907126029e-002
     2.959063185746617e+000 -4.000676406592807e+001 -2.381563370802887e+001 1.061054763128685e+001
    -4.204008513118066e+001 5.231485768135096e+001 1.127052654169966e+001 -2.151284690448416e+001
     1.641553494359437e-001 7.187888225620663e-002 -9.844472749198958e-002 -1.371423268674522e-001

199

```
        3.941929602181036e-002  -3.206605088345849e-002  -5.554795208844733e-003  -7.361094810113747e-004
        3.931382535152229e+000  -2.204145466433774e+000  -5.457266147481993e+000   3.825352599361262e+000
        2.092677066596587e-001  -1.775928150166493e-001  -2.849399440442625e-001   2.532443179006902e-001
       -2.849981558895286e-002   3.578270237319903e-003   1.311939097071185e-002   1.124105918859059e-002
   ARX_MARGIN: 0.868900
```

HEATCO matlab_4in_2out_b

```
DESCRIPTION: forward1
ARX_INPUTS: in1 in2 in3 in4
            in1*in2 in1*in3 in1*in4 in3*in4
            in1*in1 in2*in2 in3*in3 in4*in4
 ARX_TIME_ORDERS: [ 2 [ 2 2 2 2 2 2 2 2 2 2 2 2 ] [0 0 0 0 0 0 0 0 0 0 0 0 ] ]
   ARX_COEFFICIENTS:
      -1.211250124912917e+000   4.375242844740948e-001   0.000000000000000e+000   0.000000000000000e+000
      -6.707774351941087e-001   6.667280198320056e-001   6.057454734326838e-001  -3.736884213369088e-001
       2.665383117643850e-001  -2.595741280866179e-001  -1.475541354963033e-001   1.415044932925165e+001
       0.000000000000000e+000   0.000000000000000e+000   0.000000000000000e+000   0.000000000000000e+000
      -1.913460868876778e-002   1.886881663846546e-002   0.000000000000000e+000   0.000000000000000e+000
       1.770772761954355e-002  -1.762632204163374e-002  -5.584806043116148e-003   5.399609571142501e-003
       1.008137477923758e-002  -1.044208007554232e-002
ARX_MARGIN: 0.023148
```

```
DESCRIPTION: forward2
ARX_INPUTS: in1 in2 in3 in4
            in1*in2 in1*in3 in1*in4 in3*in4
            in1*in1 in2*in2 in3*in3 in4*in4
 ARX_TIME_ORDERS: [ 2 [ 2 2 2 2 2 2 2 2 2 2 2 2 ] [0 0 0 0 0 0 0 0 0 0 0 0 ] ]
   ARX_COEFFICIENTS:
      -3.510550680349328e-001   3.177282293840597e-001   0.000000000000000e+000   0.000000000000000e+000
       4.204685163172011e+000  -4.304845051785863e+000   4.580860973643669e+000  -3.489968908859943e+000
       3.134552831771393e+000  -3.059181559406307e+000  -5.906940879319706e+000  -5.630101639722969e+001
       0.000000000000000e+000   0.000000000000000e+000   0.000000000000000e+000   0.000000000000000e+000
      -1.686190808456064e-001   1.657427145451551e-001   0.000000000000000e+000   0.000000000000000e+000
      -1.114748450953000e-001   1.137269647717678e-001  -1.341159634481772e-001   1.304200965040484e-001
      -8.014921488079371e-002   7.716143520191962e-002
ARX_MARGIN: 0.087871
```

SUPFAN matlab_4in_2out_c

```
DESCRIPTION: forward1
ARX_INPUTS: in1 in2 in3 in4
 ARX_TIME_ORDERS: [ 1 [ 1 1 1 1 ] [0 0 0 0 ] ]
   ARX_COEFFICIENTS:
       2.381369923515381e-001   3.463933217979033e-003   1.234503451901340e-002
1.259769859321848e+000
       7.815893954051693e-002
ARX_MARGIN: 0.1
```

```
DESCRIPTION: forward2
ARX_INPUTS: in1 in2 in3 in4
 ARX_TIME_ORDERS: [ 1 [ 1 1 1 1 ] [0 0 0 0 ] ]
   ARX_COEFFICIENTS:
      -5.871170150442758e-001   4.199759928387796e-001  -1.936053085033960e-002   1.472518466377957e-001
      -7.288116708521500e-003
ARX_MARGIN: 0.2
```

```
CONNECTIONS:

POWER2  2 RETFAN 1
TEMP32  2 RETFAN 2
RETFAN  3 TEMP33 1
CTRL48  2 MIXBOX 1
CTRL49  2 MIXBOX 2
CTRL50  2 MIXBOX 3
TEMP1   2 MIXBOX 4
RETFAN  3 MIXBOX 5
MIXBOX  6 TEMP2  1
CTRL37  2 COOLCO 1
TEMP1   2 COOLCO 2
MIXBOX  6 COOLCO 3
FLOW2   2 COOLCO 4
COOLCO  5 TEMP3  1
COOLCO  6 TEMP48 1
CTRL36  2 HEATCO 1
TEMP1   2 HEATCO 2
COOLCO  5 HEATCO 3
FLOW2   2 HEATCO 4
HEATCO  5 TEMP4  1
HEATCO  6 TEMP49 1
HEATCO  5 SUPFAN 1
POWER1  2 SUPFAN 2
FLOW9   2 SUPFAN 3
FLOW1   2 SUPFAN 4
SUPFAN  5 FLOW2  1
SUPFAN  6 TEMP5  1
```

APPENDIX II

The 'diagnosis language' program that is used to accomplish real-time fault detection and diagnosis:

```
VAR count
VAR compo

:lop
rem First start with the system model
Cls
Load_Model sys

count=0
rem *** A loop to read historical data
:lopje
Load_Observations
count=count+1
If count<6 Goto :lopje

:loop
rem *** This is where the conflict detection loop starts
Cls
Load_Observations
count=count+1
Calculate_prop2
If No_Conflicts Goto :loop

rem *** A conflict is found
Pos 5,5
Echo Conflicts Found
Beep
pos 5,6
echo Now loading Component Model..
rem *** if historical observations are available,
rem *** these are automatically read when a new model is loaded
Load_Model comp
Calculate_prop2
If No_Conflicts=0 Goto :ohno

rem *** seems like a false alarm
rem *** stay in component mode for a few runs to be shure
rem *** If after 5 runs still no conflicts are found
rem *** the system assumes it was a false alarm due to inaccuracy

compo=0
:checkloop
Load_Observations
compo=compo+1
Calculate_prop2
If compo=5 Goto :lop
If No_Conflicts Goto :checkloop

:ohno
rem *** Now we're in serious trouble..
rem *** a conflict is found in the component model
Calculate_Hit_Best
Show_Diagnoses
Keypress
```

Qualitative Model-Based
Fault Detection in Air-Handling Units

A.S. Glass, P. Gruber, M. Roos, and J. Tödtli

The feasibility of a qualitative approach for detecting faults in an air-conditioning system is considered. The system considered is a multi-zone variable air volume air-handling unit, and the faults investigated include types which result in deterioration of operation, as distinct from actual failure. The operating modes of the sequential controller for the central air-handling plant can be matched to a corresponding qualitative classification of steady-state temperatures. Observed mismatches indicate the presence of faults. Trials of the method in an air-conditioning test laboratory are reported.

Introduction

This article extends and consolidates the investigations reported by Fornera et al. [1] and Glass [2], in which a *qualitative approach* to detecting a class of faults in a variable air volume air-handling system was developed. The qualitative approach identifies certain so-called "landmark" states of the controller regulating a central air-handling unit and relates them to corresponding "landmarks" of the relevant air temperatures and flows.

Using this approach, a wide class of operating faults can be analyzed. In particular, we deal with faults which result in deterioration of operation as distinct from actual failure—a situation known to occur frequently in practice. Conflicts between qualitative values of the controller states as actually observed and those expected—as derived from the measured qualitative temperature states—provide evidence of such faults.

This article describes a technique for analyzing the symptoms expected to result from given faults, and by applying this approach to known faults, a rule-table can be compiled to aid *diagnosis*.
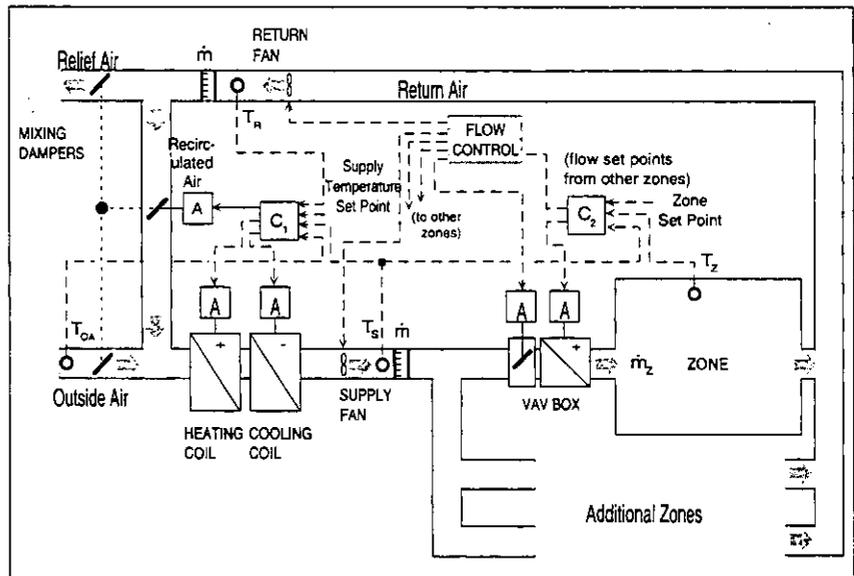
Fig. 1. The simplified reference air-handling system.

The qualitative detection method described presupposes the temperatures and control settings of the central air handling unit (CAHU) to be quasi-stationary. Consequently a class of prototype steady-state detectors has been developed and implemented in a fault detection system in the simulation model. Furthermore, it has been tested on a laboratory heating, ventilation, and air-conditioning plant (HVAC plant). A theoretical analysis of the threshold problem for various designs of a steady-state detector (SSD) is discussed by Tödtli and Gruber [3].

Description of the Test System

Simulation tests have been carried out using a SIMULINK model of a simplified variable air-volume system (VAV system) based on the Annex 25 reference air-handling system described by Kelly [4] (SIMULINK is a computer simulation environment, part of the MATLAB software package supplied by The MathWorks Inc.). The system comprises a controlled central air-handling unit supplying air to three independently controlled zones with differing loads. Both the reference system and the simplified version are described in more detail in Fornera et al. [1,5]. We summarize the components and control strategy of the simplified system below.

The Component System

The overall system under test is illustrated in Fig. 1. It consists of a central air-handling unit, depicted in Fig. 2, supplying air to a number of separately controlled zones. The CAHU comprises

a bypass mixer and a heating coil followed by a cooling coil in a single air duct. The bypass dampers are controlled to provide a mixture of outside air and recirculated air; the amount of outside air may vary between 20% and 100% of the maximum airflow through the CAHU. The air for each of the three zones is processed through a VAV box containing a damper and a reheating coil. The airflow to each zone as regulated by the damper may vary between 40% and 100% of the specified maximum. In point of fact, the maximum flow from the CAHU is not sufficient to provide maximum flows through all three zones simultaneously, so limiting effects can be expected to occur.

The system illustrated differs from standard North American practice in not including dehumidification of the cooling coil. However, the qualitative fault detection technique described below can also be applied to some such systems.

The Control Strategy of the System

Three controllers share the task of regulating the fans, dampers, heating coils, and cooling coils so as to attain the required temperatures and airflows in each of the three zones. Referring to Fig. 1, Controller C_1 regulates the CAHU and attempts to maintain the supply air temperature T_S at its set point by operating the preheating coil, dampers, and cooling coil in sequence. Controller C_2, which regulates a single zone, attempts to maintain the zone temperature T_Z at its set point by operating in sequence the damper and reheating coil in the VAV box. The third controller is an idealized "flow control" governing the airflow through the CAHU and the various zones. It ensures that the flows $\dot{m}_{z_1}$, $\dot{m}_{z_2}$, and $\dot{m}_{z_3}$ meet the requirements set by the zone controller C_2, provided the resulting total airflow $\dot{m}$ through the CAHU does not exceed its specified maximum. If the CAHU airflow maximum is reached, the total flow is shared among the three zones in proportion to their respective requirements.

The Central Air-Handling Unit

The fault detection methods described below are applied specifically to the CAHU. As mentioned above, its task is to supply air at a controlled, fixed temperature T_S by operating the preheating coil, dampers, and cooling coil in sequence. The corresponding outputs of Controller C_1 are denoted U_H, U_D, and U_C, respectively. The controller also includes an economy control feature, in which the control action on the bypass is reversed whenever the outside air temperature T_{OA} exceeds the return air temperature.

As mentioned, this system does not provide for *dehumidification*. If, however, the temperature control in such a system operates as described above, and the humidity control only operates while the cooling coil is in operation, then the fault detection method described below can be applied.

The overall system, including the zones, affect the CAHU indirectly. The VAV boxes, acting in response to any loads in the zones, determine overall air flow $\dot{m}$, while the return air temperature T_R results from mixing the air extracted from the zones. Thus, controller C_1 must respond to three quantities over which it has no control, T_{OA}, T_R, and $\dot{m}$, which can be formally regarded as disturbances in this control loop.

The Qualitative Model-Based Fault Detection Method

The detection method described below involves reducing measured controller outputs to *qualitative* values and at the same time using temperature measurements to predict expected qualitative controller outputs in steady state. This procedure necessitates reliable *steady-state detection*. Although we classify faults in terms of observed discrepancies using rule tables, such rules are derived *analytically* from the models as distinct from empirical data. Thus our fault detection method incorporates a *model-based* approach.

The Fault Detection and Diagnosis Strategy

The current fault-detection strategy relies on analyzing the steady-state behavior of the system, including controls. Our objective has been to identify *qualitative* modeling methods that can lead to successful fault detection and diagnosis procedures (FDD procedures).

Qualitative models are investigated because, even if conventional quantitative mathematical models are available for the system components, it is frequently impracticable if not impossible to obtain all of the relevant physical parameters of the system. Thus FDD methods based on qualitative models are particularly *robust*. Obviously, there is a tradeoff in that the same methods are often less *sensitive* and therefore may not be able to detect faults in all operating states of the system. Furthermore, they may not be able to *discriminate* between different types of fault. One of the goals of this work was to assess the circumstances under which it is possible to detect given faults in a CAHU using qualitative information.

The strategy adopted corresponds to the general scheme described in Fornera et al. [1]. The overall structure of such a qualitative model-based fault detector is shown in Fig. 3. This conforms to the structure of the so-called general diagnostic engine (de Kleer & Williams [6], Dexter & Glass [7]). The fault detector models we consider are also related to the generic FDD scheme proposed by Rossi and Braun [8].

From the central air-handling unit the measured values of the temperatures T_{OA}, T_R, T_S and the control variables U_H, U_D, U_C are obtained and fed into the first stage of the analog pre-processor, which serves to test whether the system is in steady state.
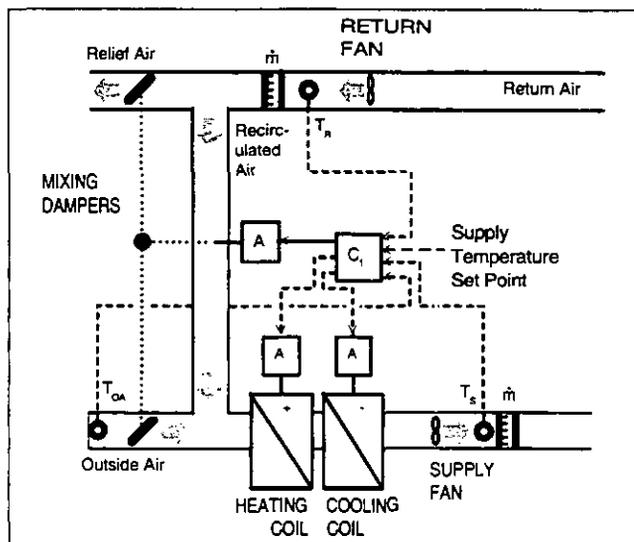
Fig. 2. The central air-handling unit.
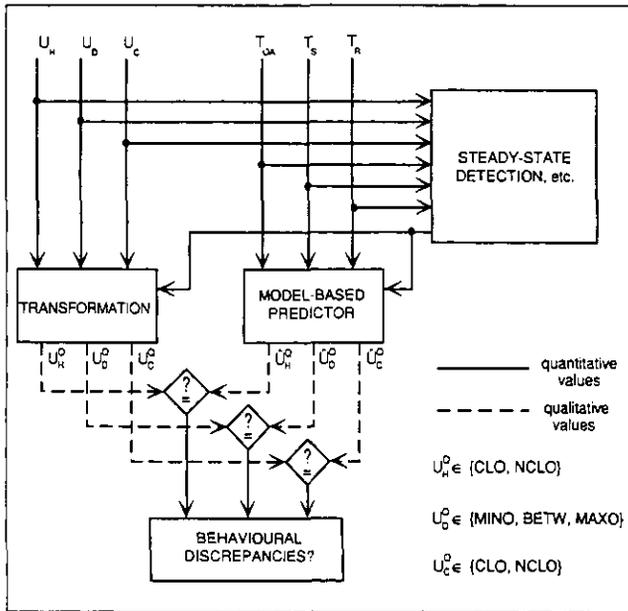
Fig. 3. The qualitative model-based fault-detection strategy.

Referring to the left half of Fig. 3, the controller outputs are converted to qualitative values U_H^Q, U_D^Q, and U_C^Q. In the right half of the figure, the temperature data is input to a model-based predictor which outputs $\hat{U}_H^Q$, $\hat{U}_D^Q$, and $\hat{U}_C^Q$, the expected *qualitative* controller states under steady-state conditions. The qualitative values of the outputs of the transformation and predictor blocks are chosen from the following sets:

$$\begin{cases} U_H^Q \in \{CLO, NCLO\}, \\ U_D^Q \in \{MINO, BETW, MAXO\}, \\ U_C^Q \in \{CLO, NCLO\}, \end{cases} \tag{1}$$

$$\begin{cases} \hat{U}_H^Q \in \{CLO, NCLO\}, \\ \hat{U}_D^Q \in \{MINO, BETW, MAXO\}, \\ \hat{U}_C^Q \in \{CLO, NCLO\}, \end{cases} \tag{2}$$

where the linguistic mnemonics "CLO," "NCLO," "MINO," "MAXO," and "BETW" stand for "closed," "not closed," "maximally open," "minimally open," and "between," respectively.

Faults are detected on the basis of discrepancies between the *measured* qualitative controller outputs and the corresponding *model-based predictions* based on temperature measurements.

In point of fact, the method involves *quantitative* pre-processing of temperature data, but particular stationary states are identified in which the *qualitative* settings of the control signals ("minimum," "maximum," or "between") betray the presence of faults. This differs from the type of purely qualitative formalism in which, for instance, *all* physical quantities, such as temperatures, etc., are strictly described in terms of intervals.

Transition States of the CAHU Controller

The strategy hinges on analyzing the *transition states* of the sequential controller regulating the CAHU. One transition occurs when the controller switches from operating the heating coil to operating the bypass dampers; a similar transition occurs when the controller switches from the bypass dampers to the cooling coil. A third transition occurs when the controller switches into economy mode and reverses the direction in which the dampers are operated. These transitions correspond to the idea of *landmarks* used in some approaches to qualitative physics (Kuipers, [9,10,11]; Dexter & Glass, [7]). Landmarks are physical values of special significance. For example, freezing and boiling temperatures can serve as landmarks because phase transitions occur.

The values taken by qualitative variables can either be on one of the landmarks themselves or in an interval between two landmarks. In our case, neglecting the issue of economy control for the time being, we identify five qualitative values of the CAHU controller states, depicted in Table 1.

An equivalent classification occurs when the system goes into *economy mode*, the only difference being that χ_{max} and χ_{min} should be interchanged in the above table.

The two landmark controller transition states can be simply related to corresponding critical temperatures. Given a particular flow $\dot{m}$, supply air temperature T_S and return air temperature T_R, critical values of the outside air temperatures, $T_{OA}^{(C1)}$ and $T_{OA}^{(C2)}$, can be calculated corresponding to controller states (4) and (2)

Table 1. Qualitative States of the CAHU Sequential Controller

Qualitative Controller State	Controller outputs to ...		
	Heating Coil	Bypass Dampers	Cooling Coil
1. The controller output sets the outside-air dampers at minimum, the cooling coil off, and the *heating coil on*.	$U_H > 0$	$U_D = \chi_{min}$	$U_C = 0$
2. Landmark state: *transition* between heating coil operation and bypass damper operation.	$U_H = 0$	$U_D = \chi_{min}$	$U_C = 0$
3. The controller output sets the bypass *dampers* somewhere *between* their extreme values.	$U_H = 0$	$\chi_{min} < U_D, U_D < \chi_{max}$	$U_C = 0$
4. Landmark state: *transition* between bypass damper operation and cooling coil operation.	$U_H = 0$	$U_D = \chi_{max}$	$U_C = 0$
5. The controller output sets the outside-air dampers at maximum, the heating coil off, and the *cooling coil on*.	$U_H = 0$	$U_D = \chi_{max}$	$U_C > 0$
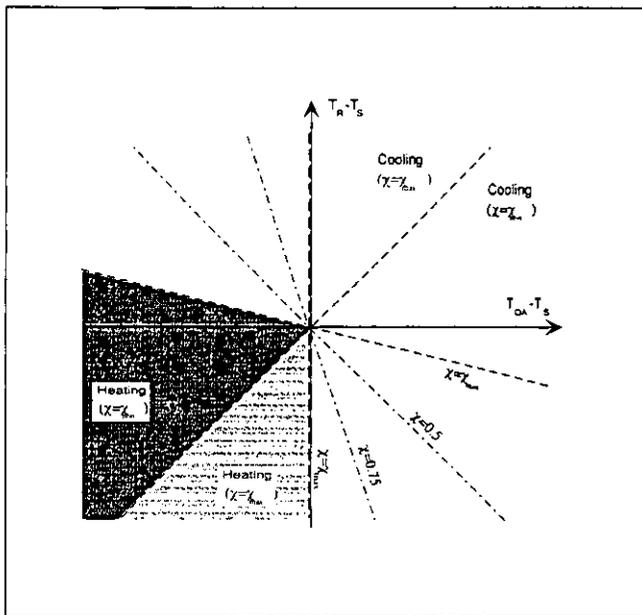
Fig. 4. Graphical representation of CAHU controller operating regimes in terms of steady-state temperature conditions.

respectively. Outside air temperatures are then classified qualitatively, according to whether they are on or somewhere between the critical temperatures.

Steady-State Behavior

To illustrate the relevant steady-state conditions which may occur we consider a plot of $T_R - T_S$ vs. $T_{OA} - T_S$ as depicted in Fig. 4.

In order to optimize energy consumption, the controller must attempt to mix the outside air and the return air so as to ensure that the temperature T_M of the mixed air is as close as possible to the supply air set-point temperature. If the fraction of outside air in the mix is given by χ, the temperature of the mixed air is given by

$$T_M = \chi T_{OA} + (1 - \chi) T_R , \qquad (3)$$

where χ is constrained between a specified minimum χ_{min} and maximum χ_{max}.

$$0 < \chi_{min} \leq \chi \leq \chi_{max} \leq 1 \qquad (4)$$

In our case, $\chi_{min} = 0.2$ and $\chi_{max} = 1.0$.

Combinations of T_{OA} and T_R yielding a given mixed air temperature correspond to straight lines in this representation. Combinations in which $T_M = T_S$ are straight lines through the *origin*. However, in view of the constraints imposed on χ, only those lines with slopes between

$$\frac{\chi_{max}}{1 - \chi_{max}} \quad \text{and} \quad \frac{\chi_{min}}{1 - \chi_{min}} \qquad (5)$$

correspond to situations in which the supply temperature can be achieved by operating just the dampers. Otherwise, depending

on whether or not $T_{OA} \leq T_R$ and $T_S < T_R$, the optimal damper setting will correspond to χ_{min} or χ_{max}.

The various regimes are depicted in the diagram. The line $T_{OA}=T_R$ determines when the economy control feature should switch χ from χ_{min} to χ_{max} or vice versa. The unshaded regions are those in which it is only necessary to operate the dampers to achieve the desired effect. The shaded regions in the upper right sectors require cooling, with $\chi = \chi_{max}$, if $T_{OA} \leq T_R$, or $\chi = \chi_{min}$ otherwise. Similarly, the shaded regions in the lower left sectors require heating, with $\chi = \chi_{min}$ if $T_{OA} \leq T_R$, or $\chi = \chi_{max}$ otherwise. We summarize the six operating regimes in Table 2 (the numbering of the operating regimes differs from that used in Fornera et al. (1993) & Glass (1993)).

As previously noted, the above analysis also applies to air-handling units which incorporate *dehumidification* in the cooling mode, *provided* humidity control is subservient to temperature control and the dehumidifier can only be operated while the cooling coil is active.

Temperature States Corresponding to Controller Transitions

To ensure that a fault detection strategy using controller landmark values functions correctly, the temperature states corresponding to the landmark values must be modeled accurately. Indeed, this aspect of the analysis involves *quantitative* modeling. However, the models used are sufficiently simple that the method as a whole may be considered qualitative.

For the sake of simplicity we shall assume that $T_{OA} \leq T_R$ and consider only the first three operating regimes from Table 2. The first transition we consider occurs when the controller switches between cooling and operating the bypass dampers. The second occurs when the controller switches between operating the dampers and heating.

Under steady state conditions the first transition occurs when the mixed air temperature exactly matches the supply air temperature for damper settings with maximum outside air. The second transition point occurs in the corresponding situation when the dampers are set for minimum outside air.

$$T_S = \chi_{max} T_{OA} + (1 - \chi_{max}) T_R \qquad (6)$$

or

$$T_S = \chi_{min} T_{OA} + (1 - \chi_{min}) T_R . \qquad (7)$$

The corresponding critical outside air temperatures are:

$$T_{OA}^{(C1)} = \frac{T_S - (1 - \chi_{max}) T_R}{\chi_{max}} ; \qquad (8)$$

or

$$T_{OA}^{(C2)} = \frac{T_S - (1 - \chi_{min}) T_R}{\chi_{min}} . \qquad (9)$$

For general values of χ_{min} and χ_{max} these temperatures are time-varying quantities depending on the current return air temperature; the supply temperature itself may reasonably be re-
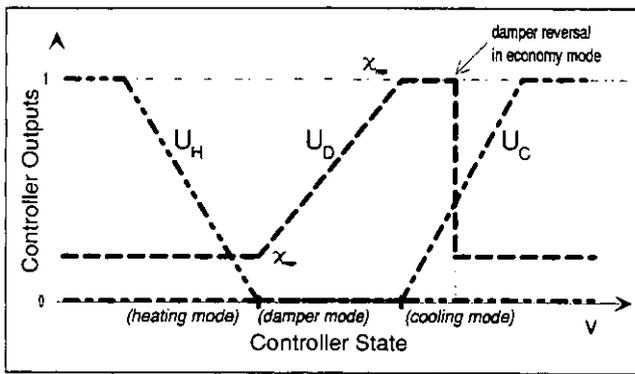
Fig. 5. Controller outputs in terms of controller state.

placed by the corresponding set point. However, in our examples, since $\chi_{max} = 1.0$, the first critical temperature reduces to

$$T_{OA}^{(C1)} = T_S , \qquad (10)$$

which is the situation discussed in Examples 1 and 3 of the fault detectors described in Fornera et al. [1,5].

We summarize the qualitative temperature states in Table 3. To take the effect of economy mode operation into account, we note that the direction of damper operation reverses whenever $T_{OA} = T_R$. We label the extreme damper settings by the symbolic values χ_1 and χ_2, in which χ_1 is the damper setting at the controller transition to cooler operation and χ_2 is the damper setting at the transition to heater operation. When $T_{OA} \leq T_R$ (the normal situation), they correspond to χ_{min} and χ_{max} as given above

$$\chi_1 = \chi_{max}, \qquad \chi_2 = \chi_{min}. \qquad (11)$$

If $T_{OA} > T_R$, however (economy mode), the direction of damper operation is reversed, and

$$\chi_1 = \chi_{min}, \qquad \chi_2 = \chi_{max}. \qquad (12)$$

Detecting Faults

In order to predict the *observable qualitative discrepancies* which allow faults to be detected, it is necessary to analyze relevant fault models in more detail.

For single faults, rules can be derived as to what qualitative discrepancies (between temperature states and controller states) will be observed under what circumstances. Whenever a fault is detected, it can be matched to a table of such rules to generate a list of possible *candidate faults* compatible with the observed discrepancy, thereby furnishing a partial *fault diagnosis*.

We consider the behavior of a CAHU in which the operating ranges of the three components—bypass, heating coil, cooling coil—be constrained to less than their full range (bypass 20% to 100% of flow), heating and cooling coils (0% to 100% of heat transfer capacity). This simulates a typical class of faults in which dampers or valves suffer partial mechanical blockage. In addition, the analysis described below can be extended to sensor offsets, etc.

The objective is to identify such faults by observing only whether the control signals are maximum, minimum, or simply somewhere in between. In particular, we consider what happens when such faults occur *singly*.

To analyze graphically what discrepancies are to be expected when faults occur, it is useful to develop an alternative representation to Fig. 3. Fig. 3 subdivides a two dimensional representation of the three relevant temperatures, T_{OA}, T_R, and T_S, into regions corresponding to heating, damper operation, and cooling. In the alternative representation we plot *one* of the temperatures directly against the controller state.

In Fig. 5 the three CAHU controller outputs are depicted graphically in terms of a single PI controller state v. In the following figures v is used to furnish a one-dimensional description of the (quantitative) controller state.

To obtain a one-dimensional representation of temperature, it is convenient to assume T_R, T_S, and $\dot{m}$ to be *constant* and to plot the controller variable v against T_{OA}, as depicted in Fig. 6. The characteristic curve shown in this diagram is based on the assumption that the temperatures, the flow, and the controller outputs are in *steady state*.
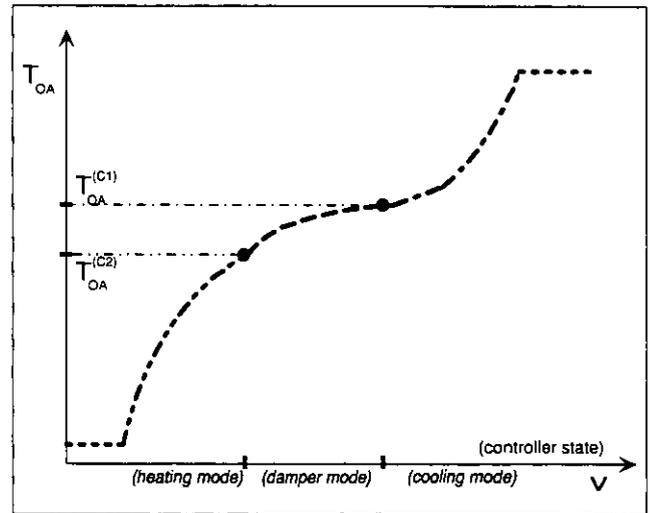
Fig. 6. Characteristic curve of T_{OA} vs. controller state.

The characteristic curves shown are indeed qualitative in nature. Neither the heating nor the cooling curve can be calculated without having a reasonably sophisticated *quantitative* model of the thermal effects taking place. By the same token, calculating the damper curves requires knowing the pneumatic response of the air flow to particular angular settings of the dampers and fan pressures. In general, this characteristic curve is *monotonically increasing* (between maximum heating and maximum cooling), but in the absence of an adequate quantitative model its slope is unknown.

In contrast, the transition points, illustrated by two dots L_1 and L_2 in the diagram, calculate. Strictly speaking, the critical temperature corresponding to the minimal damper setting will only be known accurately if the CAHU is indeed able to maintain its minimum outside air specification exactly.

It may be noted that a kink appears in the cooling mode curve. This corresponds to the transition to the economy mode: the increased slope results from the increased efficiency of cooling

	Economy mode?	Qualitative temperature state	Qualitative controller state	[Q.C.S. index]
(a)	$T_{OA} \le T_R$	TOA comparatively low	Dampers set for minimal outside air and controller operates heating coil	[1]
(b)	$T_{OA} \le T_R$	TM = TS within operating range of the bypass dampers	Heating and cooling switched off and controller operates dampers in normal mode.	[3]
(c)	$T_{OA} \le T_R$	TOA comparatively high	Dampers set for maximal outside air and controller operates cooling coil	[5]
(d)	$T_{OA} > T_R$	TOA comparatively low (but higher than TR)	Dampers set for maximal outside air and controller operates heating coil	[1]
(e)	$T_{OA} > T_R$	TM = TS within operating range of the bypass dampers	Heating and cooling switched off and controller operates dampers in *economy mode*	[3]
(f)	$T_{OA} > T_R$	TOA comparatively high.	Dampers set for minimal outside air and controller operates cooling coil.	[5]

when the quantity of warm outside air is reduced. This suggests introducing a *third* controller landmark corresponding to economy control switching. The corresponding temperature is $T_{OA} = T_R$, as shown in Fig. 7.

However, for the sake of simplicity, we omit further discussion of the economy mode switching for the time being.

A benefit of this qualitative approach is that it can be applied without significant modification to related central air-handling units. If the temperature control strategy of a CAHU is as described above, and, in addition, *dehumidification* may occur during cooling, this merely has the effect of reducing the efficiency of cooling (in terms of energy expended to achieve a given temperature drop), and therefore reduces the slope of the right-hand curve segment in Fig. 6, but does not alter the positions of the "landmarks" L_1 and L_2. If, on the other hand, humidity control is autonomous, then the position of the transition L_1 (if not both) will shift, and the above analysis will no longer apply without modification.

In another variant of the above system, the bypass mixers can be replaced by a *heat recovery wheel*. The control strategy of such a system is analogous; assuming $T_{OA} \le T_R$, when the cooling coil is operating, both the heat recovery wheel and the heating coil are switched off. When the heating coil is operating, the cooling coil is switched off and the heat recovery wheel is operating at maximum capacity.

When a known fault is present, it can be modeled qualitatively in terms of the above characteristic curves. An example is illustrated in Fig. 8, in which the heating coil valve cannot close completely, but otherwise operates normally. For cold outside air temperatures, when substantial heating may be required, the characteristic curve is normal. However, for controller states to the right of the threshold at which valve blockage occurs, the same minimal amount of heat continues to be delivered, and so the characteristic curve remains flat (at the temperature corresponding to that threshold heating setting) until the transition to damper operation is reached. Thereafter, remaining characteristic curve segments are the same *with respect to the modified transition point* L_2'.
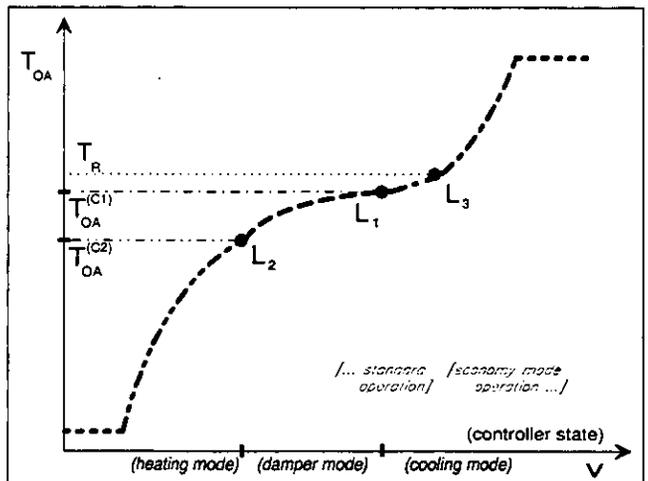
Fig. 7. Characteristic curve of T_{OA} vs. controller state with three landmarks when $T_S < T_R$.
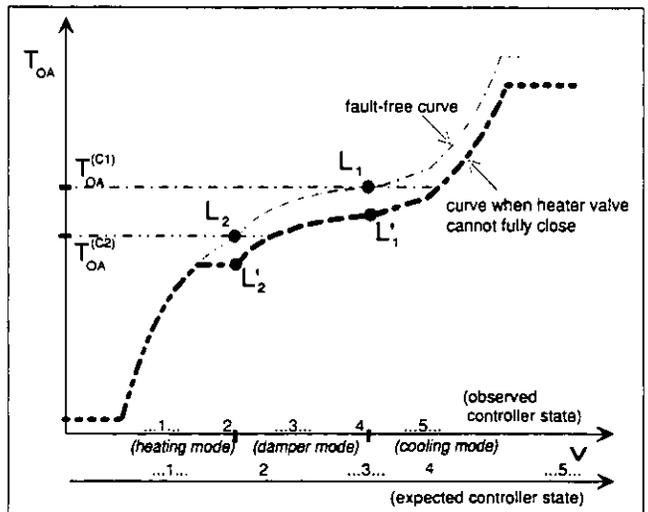
Fig. 8. Characteristic curve of TOA vs. controller state when a fault is present.

To illustrate the qualitative discrepancies which may occur, two parallel axes for the controller state are included: one for the expected behavior, and one for the observed behavior. The observed behavior corresponds to that illustrated in Fig. 6. The five observed qualitative

Table 3. Qualitative Temperature States	
Qualitative Temperature State in Terms of T_{OA}	Qualitative Temperature State in Terms of T_S
1. $T_{OA} < T_{OA}^{(C2)}$	$T_S > \chi_2 T_{OA} + (1 - \chi_2) T_R$
2. $T_{OA} = T_{OA}^{(C2)}$	$T_S = \chi_2 T_{OA} + (1 - \chi_2) T_R$
3. $T_{OA}^{(C2)} < T_{OA} < T_{OA}^{(C1)}$	$\chi_1 T_{OA} + (1 - \chi_1) T_R < T_S;$ $\qquad T_S < \chi_2 T_{OA} + (1 - \chi_2) T_R$
4. $T_{OA} = T_{OA}^{(C1)}$	$T_S = \chi_1 T_{OA} + (1 - \chi_1) T_R$
5. $T_{OA} > T_{OA}^{(C1)}$	$T_S < \chi_1 T_{OA} + (1 - \chi_1) T_R$

states of the controller are indicated on the upper v-axis in the diagram. States 2 and 4 correspond to the transitions (landmark values) themselves.

The *predicted* qualitative controller states are obtained by determining the controller values on the *modified curve* which correspond to the two transition temperatures $T_{OA}^{(C1)}$ and $T_{OA}^{(C2)}$. The corresponding values have been projected onto the lower v-axis in the diagram. Comparing the two axes shows that two temperature ranges exist for which qualitative discrepancies will be apparent.

The detectability of the fault in this example is summarized in Table 4. It can be seen that for both temperatures at the landmark transition points, as well as in intervals below these values, the fault can be detected.

Similar behavior occurs for other faults. When $T_{OA} = T_{OA}^{(C1)}$ (and $T_{OA} \leq T_R$) we expect the intake of outside air to be maximal and both the heating and cooling coils to be switched *off*. If *either* the heating coil fails to switch off completely or if the bypass mixer takes in less than the specified maximum proportion of outside air, the controller will switch on the cooling coil to compensate. Conversely, if the cooler fails to switch off, the controller will compensate by operating the dampers and heaters in sequence.

The situation when more than one fault occurs simultaneously is more complex, but can be analyzed by repeated application of the technique described graphically in Fig. 8. When $T_{OA} = T_{OA}^{(C1)}$ (and $T_{OA} \leq T_R$) and both the dampers fail to open fully and the heating coil fails to switch off completely, compensatory cooling will result. However, if both heating and cooling coils cannot switch off completely, the net effect could be either one of cooling or heating, and so the symptoms cannot be predicted without further, quantitative information. The same is true for the cooling-coil fault in combination with the bypass fault, as well as for all three faults in combination. When $T_{OA} = T_{OA}^{(C2)}$ the effect of combination of the cooling coil not switching off completely and the bypass failing to close down to χ_{min} will lead to compensatory heating. The symptom resulting from any other combination of this group of faults is unpredictable.

Table 4. Detectability of Heater Valve That Cannot Close Completely			
Steady-state configuration		Fault-free behavior	Symptom when heater cannot switch off completely
1(a)	$T_{OA} \ll T_{OA}^{(C2)}$	$U_H \gg 0; U_D = \chi_2; U_C = 0$	
1(b)	$T_{OA} < T_{OA}^{(C2)}$	$U_H > 0; U_D = \chi_2; U_C = 0$	$U_H = 0$ or $U_D > \chi_2$ or $U_C > 0$
2.	$T_{OA} = T_{OA}^{(C2)}$	$U_H = 0; U_D = \chi_2; U_C = 0$	$U_D > \chi_2$ or $U_C > 0$
3(a)	$T_{OA}^{(C2)} < T_{OA}; T_{OA} \ll T_{OA}^{(C1)}$	$U_H = 0; \chi_2 < U_D \ll \chi_1; U_C = 0$	
3(b)	$T_{OA}^{(C2)} \ll T_{OA}; T_{OA} < T_{OA}^{(C1)}$	$U_H = 0; \chi_2 \ll U_D < \chi_1; U_C = 0$	$U_D = \chi_1$ or $U_C > 0$
4.	$T_{OA} = T_{OA}^{(C1)}$	$U_H = 0; U_D = \chi_1; U_C = 0$	$U_C > 0$
5(a)	$T_{OA} > T_{OA}^{(C1)}$	$U_H = 0; U_D = \chi_1; U_C > 0$	
5(b)	$T_{OA} \gg T_{OA}^{(C1)}$	$U_H = 0; U_D = \chi_1; U_C \gg 0$	

Steady-State Detection

A prototype steady-state detector has been implemented in SIMULINK. It has been both tested separately and incorporated in the FDD module of the overall simulation model.

The separate tests have been carried out with generated data comprising a mixture of step disturbances and both damped and undamped sinusoidal oscillations. In addition, initial tests have been carried out on measured data from the new Grafenau Building at Landis & Gyr's headquarters in Zug, and, more recently, on data generated from the Landis & Gyr HVAC testing laboratory.

The prototype detector computes a geometrically weighted running variance with respect to a geometrically weighted running average. An alternative would be to compute the variance of data within a sliding time interval, but the former scheme has been implemented because it has the advantage that the computations are recursive, requiring a minimum of memory. Moreover, it is sensitive in reacting promptly whenever the current data depart from their steady-state values.

Denoting the sequence of data by

$$\{x_0, x_1, x_2, ..., x_n\} \tag{13}$$

the geometrically weighted average is defined

$$\overline{X}_n(\alpha) = \frac{\sum_{k=0}^{n} \alpha^{n-k} x_k}{\sum_{k=0}^{n} \alpha^{n-k}} \tag{14}$$

where α is the (constant) geometric weighting factor $(0 < \alpha < 1)$. The geometrically weighted variance can be formally defined with a different weighting factor β $(0 < \beta < 1)$:

$$S_n^2(\alpha,\beta) = \frac{\sum_{k=0}^{n} \beta^{n-k}\left(x_k - \overline{X}_n(\alpha)\right)^2}{\sum_{k=0}^{n} \beta^{n-k}} . \tag{15}$$

The above can be computed recursively in terms of the running moments $X_n^{(m)}(\alpha)$:

$$X_n^{(m)}(\alpha) = \sum_{k=0}^{n} \alpha^{n-k}\left(x_k\right)^m ; \tag{16}$$

$$X_{n+1}^{(m)}(\alpha) = \left(x_{n+1}\right)^m + \alpha X_n^{(m)}(\alpha) ; \tag{17}$$

$$S_n^2(\alpha,\beta) = \frac{X_n^{(2)}(\beta)}{X_n^{(0)}(\beta)} - 2\frac{X_n^{(1)}(\beta)}{X_n^{(0)}(\beta)}\frac{X_n^{(1)}(\alpha)}{X_n^{(0)}(\alpha)} + \left(\frac{X_n^{(1)}(\alpha)}{X_n^{(0)}(\alpha)}\right)^2 . \tag{18}$$

The variable x_n is deemed to be in steady state whenever the weighted *deviation* falls below a pre-determined threshold ε_{ss} [or, equivalently, the variance falls below ε_{ss}^2]

$$Sn(\alpha,\beta) \le \varepsilon_{ss} . \tag{19}$$

The effect of using *differing* parameters α and β can be summarized in the equation:

$$S_n^2(\alpha,\beta) = S_n^2(\beta,\beta) + \left(\overline{X}_n(\alpha) - \overline{X}_n(\beta)\right)^2 . \tag{20}$$

In the simulation tests carried out, the geometrically weighted variance has been used with a *single* parameter $\alpha = \beta$. At the moment, the merits of using differing weighting parameters are not apparent.

The (single) parameter α can be related to an effective time "window" of length τ_{ss} by means of the weighted average

$$\tau_{ss} = \frac{\sum_{k=0}^{\infty} \alpha^k (k \Delta t)}{\sum_{k=0}^{\infty} \alpha^k} . \tag{21}$$

where Δt is the time increment between measurements. Thus

$$\alpha = \frac{\tau_{ss}}{\tau_{ss} + \Delta t} , \tag{22}$$

Tuning the SSD has involved adjusting both the threshold ε_{ss} and the time window τ_{ss}. The latter is set to match the typical relaxation times of the test system, and the former must be set according to how much "noise" is to be tolerated when the system is deemed to have reached a steady state. In the simulations and laboratory tests, a hysteresis-type threshold was used to avoid rapid fluctuations of the steady-state signal.

The questions of threshold settings and alternative steady-state detector designs are discussed in fuller detail elsewhere (Tödtli & Gruber [3]).

Simulation Tests

As reported previously (Glass, [2]), simulated tests were carried out using the simulation program developed in SIMULINK and incorporating *discrete-time* modules.

Tests of the steady-state detector itself, after initial tuning had been carried out, yielded satisfactory results. When incorporated into the simulation model of the reference air-handling unit, the steady states were flagged approximately where expected.

Running the full system including diagnostic modules tuned to the two steady-state configurations described above produced the expected effects. After the system had started up (with constant outdoor air temperatures and temperature set points), it reached steady state after about 10 minutes, and if the operating range of one of the components had been constrained, symptoms were correctly flagged in accordance with Table 4.

The simulations also reveal the necessity of proper tuning with regard to thresholds. For instance, whenever the steady-state configuration of temperatures ought to result in the heating being off and the dampers in minimum position, the controller outputs in reality *fluctuate* around this transition point, and the heating may sometimes be briefly switched on a small amount or the dampers opened slightly. If the fault detector is not to be triggered
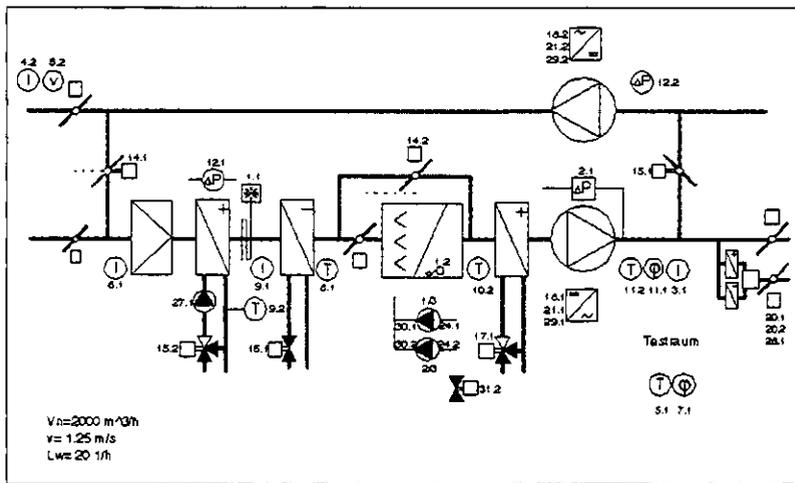
Fig. 9. Schematic diagram of HVAC laboratory air-handling unit.

triggered with false alarms, the thresholds must be set to tolerate such "normal" departures from theoretical values.

Experimental Tests

Experimental tests were conducted on a laboratory air-handling unit with two goals: first, to test the performance of the steady-state detection method of the last section on data deriving from realistic conditions, and second, to test the detection method described previously.

The plant used was the air-handling unit in the Landis & Gyr HVAC laboratory, illustrated schematically in Fig. 9. It differs from the Annex 25 reference plant in supplying a *single zone* with controlled constant air volume. In addition, the AHU is equipped to control humidity. The laboratory is equipped to modify the effective load in the zone as well as the outside air temperature as may be required to achieve desired experimental conditions: the former by simulating both heat gains and the thermal behavior of the walls, and the latter via an additional air-handling unit external to the one being tested.

. The laboratory AHU in Fig. 9 was operated so as to match the behavior of the CAHU illustrated in Fig. 2. The supply-air temperature prior to the zone was controlled using only the heating coil, the cooling coil, and the bypass mixer; dehumidification was switched off. Experiments were carried out using both naturally occurring outside-air temperatures as well as temperatures raised by an approximately constant amount. Plant data was captured via an especially developed interface and recorded in a SIMULINK software environment, to allow post-processing analysis with SSD and FDD modules. Nevertheless, the special interface would allow steady-state detection and fault detection to be carried out on-line in *real time*.

Data were recorded for runs lasting between one and five days. Faults of the type described graphically in Fig. 8 were simulated during part of some of the runs by re-programming the operating ranges of subsidiary controllers serving the individual component in question.

A typical run is illustrated in Fig. 10. The upper graph illustrates the outdoor, return, and supply-air temperatures T_{OA}, T_R, and T_S for a duration of approximately 118 hours. The supply temperature set point T_S^{SP} was deliberately kept low (initially 16°C, later 12°C) in order to provide conditions in which the outdoor temperature was close to its upper critical value $\left(T_{OA} \approx T_{OA}^{(CI)}\right)$. At approximately 98 hours T_{OA} was boosted by means of the external air-handling unit in order to avoid freezing temperatures requiring operation of the frost-protection heater, and at the same time a higher T_S^{SP} was chosen (18°C). The changes are evident in the diagram. During this last phase a fault was introduced at approximately 102 hours: the cooling coil was prevented from operating below 15% of its maximum power. A small drop in T_S is evident in the diagram at that point, but the main sequential controller ensured that T_S was brought back to its set point of 18°C.

The lower graph in Fig. 10 illustrates the results of filtering the three temperatures with steady-state detectors. The top curve is the binary (logical 0 or 1) output of the SSD acting on T_{OA}, the second curve the same for T_R, and the third the same for T_S. The bottom curve is the output showing when *all three* temperatures are deemed to be *simultaneously* in steady state. All three steady-state detectors were characterized by time constants τ_{ss} of 15 minutes and thresholds ε_{ss} of 0.1°C.

Fig. 11 illustrates the same data during the first 20 hours in more detail. The top graph displays T_{OA}, T_R, and T_S, as well as the supply-air set-point temperature T_S^{SP}. The second graph shows the corresponding controller outputs. The upper curve is
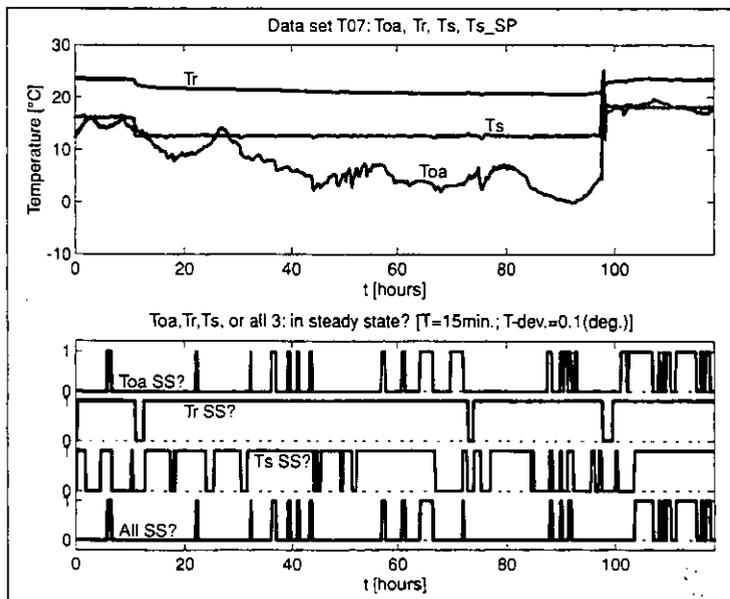
Fig. 10. Laboratory test: temperatures and steady-state detection during approximately 120 hours' operation. Top graph: Temperatures T_{OA}, T_R, and T_S. Bottom graph: Steady-state status of T_{OA}, T_R, T_S, and all three simultaneously
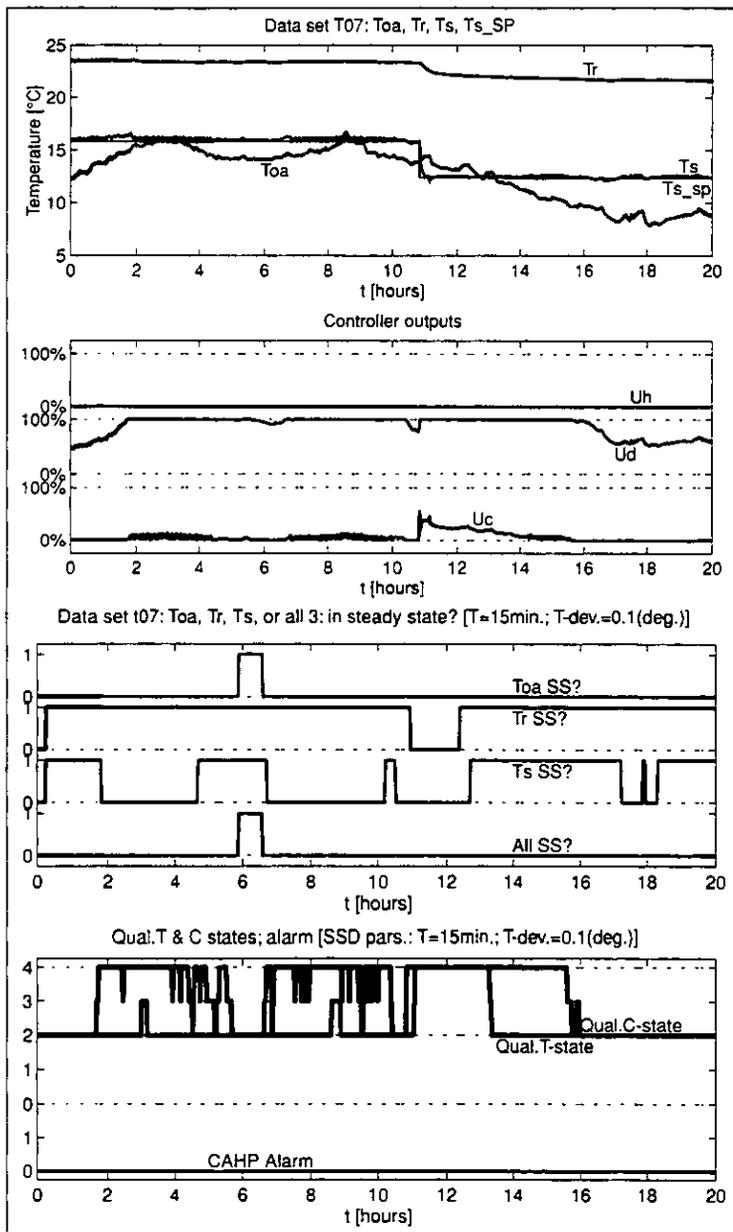
Fig. 11. Laboratory test: temperatures, controller outputs, SSD & FDD analysis when no faults present in system. Top graph: Temperatures T_{OA}, T_R, T_S, and T_S^{SP}. Second graph: Controller outputs U_H, U_D, and U_C. Third graph: Steady-state status of T_{OA}, T_R, T_S, and all three simultaneously. Bottom graph: Qualitative temperature and controller states; CAHU fault-detection alarm status.

U_H, the output to the heating coil (constantly at its minimum), the middle curve is U_D, the output to the bypass dampers (mostly at maximum), and the lower curve is U_C, the output to the cooling coil. For the sake of clarity, U_H and U_D have been offset by 2.0 and 1.0 respectively in order to separate the curves.

The third graph depicts the results of steady-state testing of the three temperatures (cf. lower graph of Fig. 10), and the bottom graph depicts the results of qualitative fault-detection analysis. The upper half of the graph shows the qualitative values of the measured temperature states and the corresponding con-

troller outputs. In terms of the classification schemes in Table 1 and Table 3, Controller State #1 and Temperature State #1 have been assigned a value of 2 and Controller State #5 and Temperature State #5 a value of 6; the other states have been assigned corresponding integer values in between. The bottom curve is the binary (logical 0 or 1) output of the fault-detection module (remaining constant at zero, in this case).

After about 16 hours both qualitative states are at level 4 (i.e., State #3), meaning that the system is regulating the supply temperature by using the dampers only. It may be noted that between hours 2 and 10 there is an almost constant discrepancy between the temperature state and the controller state: the controller is operating the cooling coil (minimally) even though the outside-air temperature is below the supply-air set point. However, no fault is signaled, because there is only one short period when all three temperatures and the airflow $\dot{m}$ are in steady state, namely from approximately 6 to 6.5 hours. During this period the qualitative states of both the temperatures and the controller are in State #4, so no discrepancy is registered.

Fig. 12 illustrates the same data as Fig. 10, but for the last approximately 19 hours' of operation of the laboratory test run. The information incorporated in the four graphs corresponds to that in Fig. 11. However, a significant difference between this part of the laboratory test and that illustrated in Fig. 11 is that from approximately 101.5 hours on, the cooling coil was constrained to operate at least 15% of its maximum power.

Although $T_{OA} \geq T_S^{SP}$ until about 113.5 hours it is evident that prior to approximately 101.5 hours $U_C > 0$, whereas afterwards, up to 102 hours, $U_D < \chi_{max}$ and then, after 102 hours, $U_H > 0$ for the remainder of the run. During much of this period, the system is deemed to be in steady state, and, since the qualitative states of the controller ("observations") and the qualitative temperature states (classified in terms of the expected controller behavior: "predictions") are incompatible, a fault is flagged.

The nature of the fault can be at least partially diagnosed from the particular combination of incompatible qualitative states observed. It may also be noted that in Fig. 12 the fault alarm and the three-temperature steady-state status do not completely coincide. This is because the FDD module also includes a SSD for the flow $\dot{m}$ (as measured by pressure differences). All four quantities, TOA, TR, TS, and $\dot{m}$, must be deemed to be in steady state for the fault detector to output an alarm.

Generalization and Applicability

Classifying the examples from Fornera et al. [1] in terms of transition points of the sequential controller lends this class of faults generic character, which ought to facilitate its generalization to other systems. However a feature of the faults considered is that the transitions occur at the ends of the operating range of

the *bypass dampers*, in which the relevant *quantitative* models are comparatively simple and reliable. In other words, the effects of adjusting the dampers is highly predictable. If one attempts to apply the same idea to the VAV boxes, for instance, one comes up against the obstacle that the loads in the zones are not predictable in practice, which makes determination of the conditions for transition more difficult.

On the other hand, the ideas applied to the CAHU in the reference AHU should, in principle, be applicable to other types of CAHU, provided a sequential control strategy is used which involves similarly predictable transitions.

The concept described ought to be implementable in practice. Simulations with the FDD module incorporating operating on *time-discrete* signals appear to be sufficiently fast to allow them to be used on-line with operating HVAC systems. The memory requirement is in any case trivial.

A more serious question is the applicability of methods which require steady states to be achieved which may not occur in the course of normal operation. This would mean that the test would have to be used as part of a program of active testing (HVAC building tests during night-time, for instance) as opposed to a pure monitoring function.
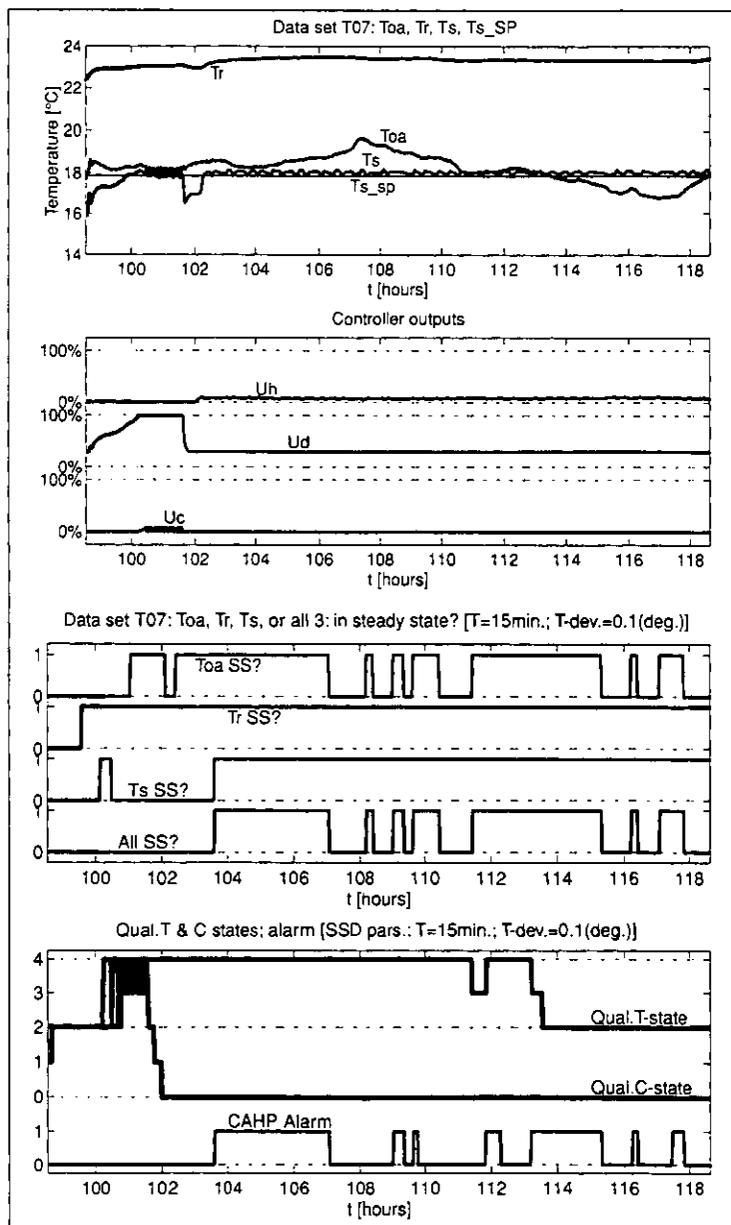
Fig. 12. *Laboratory test: temperatures, controller outputs, SSD and FDD analysis when cooling coil cannot switch off completely. Top graph: Temperatures T_{OA}, T_R, T_S, and T_S^{SP}. Second graph: Controller outputs U_H, U_D, and U_C. Third graph: Steady-state status of T_{OA}, T_R, T_S, and all three simultaneously. Bottom graph: Qualitative temperature and controller states; CAHU fault-detection alarm status.*

Conclusions and Outlook

The fault detection strategy considered here detects faults in a central air-handling unit by analyzing certain steady states of the plant (or a subsystem of the plant) in terms of *qualitative* criteria. In the overall FDD strategy, however, at least some quantitative pre-processing of data is required. The behavior of the fault detection system implemented in the simulation program was as expected. Symptoms arising from single faults can be detected. If multiple faults are assumed, the expected qualitative symptoms are, in general, not uniquely determined.

Initial results of the HVAC laboratory tests verified the fault-detection method experimentally. Steady-state conditions of the system as a whole—and the outside-air temperature in particular—were observed for a sufficient length of time during the diurnal cycle to allow faults to be detected qualitatively. Yet to be investigated is the question as to just how big a fault needs to be before it can be detected by the above qualitative method.

Future work will also be concerned with investigating the applicability of the above methods to other HVAC systems. Attention will also be given to linking knowledge derived from the comparative qualitative analysis of two (or more) different steady states. Finally, the possibility of applying qualitative methods to the analysis of transient behavior (cf. work by Koch [12]) or to other non-steady-state situations needs to be investigated.

Acknowledgments

References

[1] L. Fornera, A.S. Glass, P. Gruber, and J. Tödtli, "Fault Detection Based on Logical Programming Applied to the VAV Reference System: First Trials," *I.E.A. Annex 25* Working Paper AN25/CH/050493/1.0, Tokyo, April 14-16, 1993.

[2] A.S. Glass, "Rule-Based Qualitative Fault Detection Applied to the VAV Reference System: Further Results," *I.E.A. Annex 25* Working Paper AN25/CH/200993/1.0, Zürich, Sept. 27-29, 1993.

[3] J. Tödtli and P. Gruber, "Stochastic Model for Threshold Evaluation in Steady State and Fault Detectors," *I.E.A. Annex 25* Working Paper AN25/CH/070494/1.0, Boston, April 11-13, 1994.

[4] G. Kelly, "Air Conditioning Unit," §5.3 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen, eds., Technical Research Centre of Finland (VTT), Laboratory of Heating & Ventilation, P.O. Box 206, SF-02151 Espoo, Finland, 1993.

[5] L. Fornera, A.S. Glass, P. Gruber, and J. Tödtli, "Qualitative Fault Detection Based on Logical Programming Applied to a VAV Air Handling Unit," *Second IFAC Workshop on Computer Software Structures Integrating AI/KBS Systems in Process Control*, Lund, Aug. 10-12, 1994.

[6] J. de Kleer and B. Williams, "Diagnosing Multiple Faults," *Artificial Intelligence*, 32(1), 97-130, 1987.

[7] A.L. Dexter and A.S. Glass, "The Use of Qualitative Models in Fault Detection and Diagnosis," §4.2 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen, eds., Technical Research Centre of Finland (VTT), Laboratory of Heating & Ventilation, P.O. Box 206, SF-02151 Espoo, Finland, 1993.

[8] T. Rossi and J. Braun, "Classification of Fault Detection and Diagnostic Methods," §2 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen, eds., Technical Research Centre of Finland (VTT), Laboratory of Heating & Ventilation, P.O. Box 206, SF-02151 Espoo, Finland, 1993.

[9] B. Kuipers, "Commonsense Reasoning About Causality: Deriving Behavior from Structure," *Artificial Intelligence*, 24(1-3), 169-203, 1984.

[10] B. Kuipers, "The Limits of Qualitative Simulation," *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, William Kaufman, Los Altos, USA, 1985.

[11] B. Kuipers, "Qualitative Simulation," *Artificial Intelligence*, 29(3), 289-388, 1986.

[12] G. Koch, "Knowledge-Based Coordination of Qualitative Online Diagnostic Tests," *IFAC Symposium on Intelligent Components and Instruments for Control Applications*, Malaga, Spain, May 20-22, 1992.

Andrew Glass received his honors B.Sc. in mathematics and physics from the University of British Columbia and his M.A. and Ph.D. in physics from Princeton University, completing his studies in 1971. After a postdoctoral fellowship in theoretical physics at the Swiss Federal Institute of Technology, he joined the Mathematics Department of the University of British Columbia in 1973. Since 1979 he has been a research scientist in the Mathematics and Systems Theory Group of the Corporate R&D Department of Landis & Gyr, Zug, Switzerland. Since Corporate R&D was dissolved in 1993, he has been active in the Control Systems Group of the newly formed Landis & Gyr Technology Innovation Corp. He is also a lecturer at the ATIS Engineering College in Lucerne. His professional interests encompass several fields in engineering, physics, and mathematics, in particular the analysis and modeling of heating ventilation and air-conditioning systems and theoretical and computational analysis of physical systems, including diffractive optics, and digital coding theory. His recent activities have focused on qualitative reasoning methods applied to fault detection and diagnosis in building automation systems.

Peter Gruber received his bachelor's, master's, and doctorate degrees in electrical engineering (Dipl. El.-Ing. ETH, Nachdiplom and Dr. sc. techn.) in 1971, 1974, and 1978, respectively, from the Swiss Federal Institute of Technology, Zurich. From 1978 to 1980, he was at Queen's University, Kingston, Canada, where he held a postdoctoral position in the Electrical Engineering Department and worked simultaneously for the Canadian Institute of Guided Ground Transport. Since 1981, he has been with the Mathematics and Systems Theory Group in the Corporate R&D Department of Landis and Gyr, Zug, Switzerland. Since 1993, he has been with the Control Systems Group of the Landis & Gyr Technology Innovation. He has also given a number of lecture courses in control and signal processing. His current interests are in signal processing applied to electricity metering, including multi-rate digital filtering, sigma-delta conversion and estimation, and control of uncertain and nonlinear systems, especially heating, ventilation, and air conditioning systems.

Markus Roos received his bachelor's and doctorate degrees in theoretical physics (Dipl. Phys. and Dr. sc. nat.) in 1986 and 1990, respectively, from the Swiss Federal Institute of Technology, Zurich. From 1990 to 1994, he worked in the Computational Physics Group in the Corporate R&D Department of Landis and Gyr, Zug, Switzerland. Since 1994 he has been working as senior scientist in the Institute for Structural Engineering at the Swiss Federal Institute of Technology. His special interests are in modeling and simulation of devices and systems (using various methods including finite element techniques), Bayesian belief networks, and dynamic systems.

Jurg Todtli received his bachelor's, master's, and doctorate degrees in electrical engineering (Dipl. El.-Ing. ETH, Nachdiplom and Dr. sc. techn.) in 1969, 1970, and 1977, respectively, from the Swiss Federal Institute of Technology, Zurich. In 1978, he held a postdoctoral position at the University of Calgary and, after that, at the Swiss Federal Institute of Technology. From 1984 to 1992, he worked as Head of the Mathematics and Systems Theory Group in the Corporate R&D Department of Landis & Gyr, Zug, Switzerland. Since 1993, he has been Head of the Application and Control Group in the Commercial Buildings Segment of Landis & Gyr. He is also a lecturer at the Swiss Federal Institute of Technology and a member of the board of directors of the Swiss Society for Automatic Control (SGA). His current interests are in control, monitoring, and optimization of energy systems in buildings (including man-machine interaction aspects), uncertain dynamic systems (estimation theory in particular), stochastic control, and fuzzy logic.

0967-0661(95)00213-8

QUALITATIVE FAULT DETECTION BASED ON LOGICAL PROGRAMMING APPLIED TO A VARIABLE AIR VOLUME AIR-HANDLING UNIT

L. Fornera*, A.S. Glass**, P. Gruber*** and J. Tödtli***

*via Case Conti 3, CH-6616 Losone, Switzerland
**Zugerstrasse 60, CH-6318 Walchwil, Switzerland
***Landis & Gyr, CH-6301 Zug, Switzerland

Abstract. A qualitative approach for detecting faults or sub-optimal operation in a class of heating, ventilation and air-conditioning systems is presented. In particular, qualitative models are used to describe the steady-state operation of a controlled central air-handling plant. Included among the variants of the underlying fault-detection approach considered is one using a model-based predictor in a logical programming environment. The predictions of the qualitative steady-state models are compared with numerical simulations of the dynamical system behaviour.

Key Words. heating ventilation and air-conditioning (HVAC), central air-handling unit, fault detection and diagnosis, qualitative modelling, logical programming.

1. INTRODUCTION

This paper considers the feasibility of a qualitative approach for diagnosing a class of faults in a variable air-volume air-handling system. This paper reports the results of some preliminary investigations making use of logical programming methods (Fornera, et al., 1993, 1994).

The system considered is a simplified version of the I.E.A. Annex 25 [1] reference air-handling system described by Kelly (1993). The paper deals with some illustrative faults of the central air-handling plant which result in deterioration of operation, as distinct from actual failure — indeed, some of the faults chosen reflect situations known to occur frequently in practice.

The investigations described here comprise simulation of the air-handling system, with and without faults, an analysis of the faults considered in terms of qualitative models, and testing the models' predictions within a PROLOG logical programming environment.

In what follows, Section 2 describes the system being modelled, as well as the simulation methods employed. In Section 3 the relevant features of the steady-state behaviour of the CAHP system are discussed. Section 4 deals with the detectors themselves, including some relevant examples. Finally, conclusions and the outlook for future work are presented in Section 5. The results of the quantitative simulations are included in the appendices.

---

[1] I.E.A. is an abbreviation for the International Energy Agency. Annex 25 is part of the I.E.A. programme, "Energy Conservation in Buildings and Community Systems"; and is specifically concerned with "Real-time simulation of HVAC systems for building optimisation, fault detection and diagnosis".

2. DESCRIPTION OF THE AIR-HANDLING SYSTEM

2.1 The reference system

These investigations have concentrated on the central air-handling plant ("CAHP") of a simplified system based on the I.E.A. Annex 25 reference system described by Kelly (1993). The reference system – with a slightly different notation from that used by Kelly – is depicted in Figure 1.

The central unit comprises a bypass mixer and a heating coil followed by a cooling coil in a single air channel. The bypass dampers are controlled to provide a mixture of outside air and recirculated air; the amount of outside air may vary between 20% of the maximum airflow through the CAHP and 100%. The air for each of the three zones is processed through a variable air volume ("VAV") box containing a damper and a re-heating coil. The airflow to each zone as regulated by the damper may vary between 40% and 100% of the specified maximum. In point of fact, the maximum flow from the CAHP is not sufficient to provide maximum flows through all three zones simultaneously, so limiting effects could occur.

Five controllers share the task of regulating the fans, dampers, heating coils and cooling coils so as to attain the required temperatures and airflows in each of the three zones. Referring to Figure 1, Controller C_1, which mainly regulates the CAHP, attempts to maintain the supply air temperature T_S at its set point by operating the preheating coil, dampers and cooling coil in sequence. Controller C_2, which regulates a single zone, attempts to maintain the zone temperature T_Z at its set point by operating in sequence the damper and re-heating coil in the VAV box. Controllers C_3 and C_4 regulate the airflow through the CAHP: C_3 attempts to maintain the pressure P_S in the main supply air duct at its set point, and C_4 controls the return fan airflow rate so as to ensure that the difference $\dot{m}_R - \dot{m}_S$ between the return and supply airflow rates is a fixed, positive
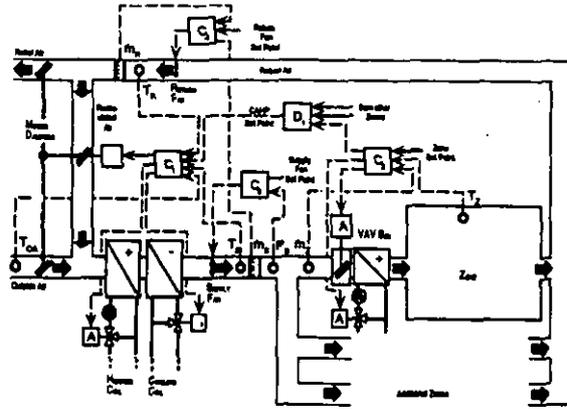
amount.

Controller C_2 actually consists of two cascaded controllers: the main zone controller attempts to maintain the zone temperature T_Z at its set point by acting in sequence on the reheater coil and a secondary controller, to which it supplies the zone air-flow set point. The secondary controller attempts to maintain the zone air flow at that set-point value.

The discriminator D_1 receives information from the various zone controllers C_2 and determines the highest supply air temperature set-point in a fixed range (13.9°C–18.3°C) compatible with ensuring that the demands of all the zones can be met.

2.2 The simplified system

The system that has been modelled thus far differs from the reference system in a number of points. First and foremost, for reasons of tractability, neither the component models nor the control strategies have yet been implemented in full detail. Besides that, there are some minor differences in the details of the control strategies used.

The modified system is illustrated in Figure 2. It differs from the reference system with regard to the

Table 1. Notation used

Quantity	Symbol [Units]	Corresp. Symbol in Kelly (1993)	
Outdoor air temperature	T_{OA} [°C]	T1	
Supply air temperature	T_S [°C]	T5	
Return air temperature	T_R [°C]	T2	
Zone temperature	T_Z [°C]	T6	
Supply airflow rate	$\dot{m}_S$ [kg/s]	F1	(vol. flow rate)
Return airflow rate	$\dot{m}_R$ [kg/s]	F2	(vol. flow rate)
Airflow rate to zone	$\dot{m}_Z$ [kg/s]	VP	(vol. flow rate)
Pressure in supply duct	P_S [kg/s]	SP	
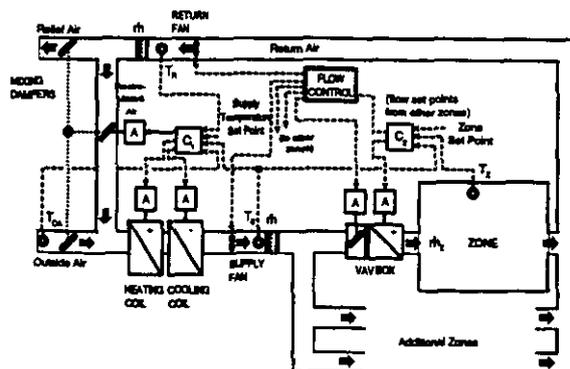
Fig. 2. The Simplified System

following points:

1. All heating coils and cooling coils omit the actual *physical heat exchange* between the air and the hot water or coolant. The effect on the air of a prescribed heat transfer rate is modelled.

2. No reference to pressure is made in this model. Simulation models for pressure-dependent airflows have been developed, but have not yet been integrated in the present system.

3. The difference between supply and return airflows is assumed to be 0 ($\dot{m} = \dot{m}_R = \dot{m}_S$). Air losses through open windows, leakage, etc., are ignored. The total airflow through the CAHP is assumed to be the sum of the flows through the various zones.

$$\dot{m} = \dot{m}_{Z_1} + \dot{m}_{Z_2} + \dot{m}_{Z_3}. \quad (1)$$

4. Controllers C_3 and C_4 have been replaced by an idealized "*flow control*" which ensures that the flows $\dot{m}_{Z_1}$, $\dot{m}_{Z_2}$ and $\dot{m}_{Z_3}$ meet the requirements set by the zone controller, subject to the limitation that the resulting total airflow $\dot{m}$ through the CAHP does not exceed a specified maximum. Whenever the total demand exceeds this limit, the maximum available flow is shared among the three zones in proportion to the requirements set by their respective controllers.

5. Discriminator D_1 has been omitted. At the moment, the correct supply temperature set point is assumed to be fixed. It is planned to add the discriminator to future versions of the simulated system.

6. Controller C_2, the zone controller, does not include the zone airflow among its inputs. It is assumed that the zone airflow control operates *ideally* so as to maintain the zone airflow at its prescribed value $\dot{m}_{Z_1}$ as described in Point 4 above. In its current implementation, however, it *includes an economy mode* feature not prescribed in the reference system, and therefore takes the supply air temperature into account.

7. Controller C_1, the CAHP controller, functions similarly to its reference system counterpart except for the fact that the supply air temperature set point is presumed constant (as mentioned in 5, above) and except for the differences in control strategy mentioned below.

8. When Controller C_1 operates the dampers in the bypass mixer, the minimal proportion of outside air is 20% of the *current total flow*, whereas the reference system sets a minimum of 20% of the maximum flow. The reference system specification will be implemented in future versions.

9. In this implementation of the economy control, reversing the direction of operating the bypass dampers occurs when the outside air temperature T_{OA} matches the *return air* temperature T_R rather than the fixed value of 22.2°C as proposed in the reference system specifications.

2.3 Simulation

For the preliminary investigations, the layout shown in Figure 2 was simulated on an Apple® Macintosh® IIfx computer using the MATLAB™/SIMULINK™ applications software package. The simulation program was used to investigate the quantitative dynamic behaviour of the system under the various fault conditions discussed in the following sections.

In addition, a MATLAB™ program was developed to investigate the steady-state behaviour of the system. Such investigations are extremely useful in assessing whether the specified design conditions can be met. In this case, pending a working simulation model of the discriminator D_1, such analyses were helpful in deciding what supply temperature set-points would ensure correct operation of the system in given situations.

3. THE STEADY-STATE BEHAVIOUR OF THE SYSTEM

Consider the operating regimes of the CAHP controller C_1 under steady-state conditions. In order to optimize energy consumption, the controller attempts to mix the outside air and the return air so as to ensure that the temperature T_M of the mixed air is as close as possible to the supply air set-point temperature. If the fraction of outside air in the mix is given by χ, the temperature of the mixed air is given by

$$T_M = \chi T_{OA} + (1 - \chi)T_R, \quad (2)$$

where χ is constrained between a specified minimum χ_{min} and maximum χ_{max}.

$$0 < \chi_{min} \leq \chi \leq \chi_{max} \leq 1. \quad (3)$$

In this case, $\chi_{min} = 0.2$ and $\chi_{max} = 1.0$.

Since it is the temperature *difference* between T_M and T_S that is crucial, the various situations that may occur can be visualized in terms of a graphical representation with axes $T_{OA} - T_S$ and $T_R - T_S$ as depicted in Figure 3.
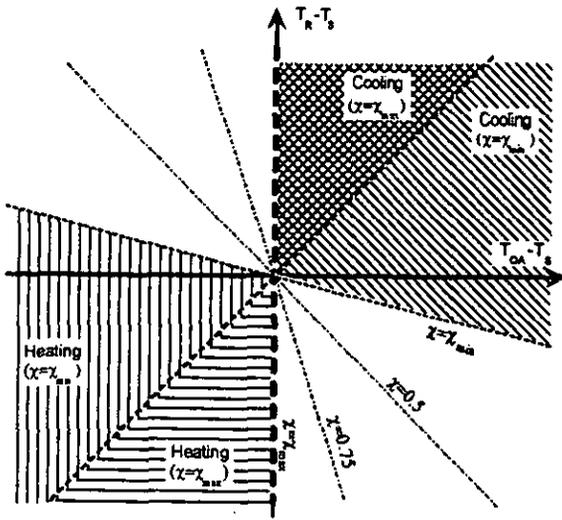
Fig. 3. Graphical representation of CAHP controller operating regimes in terms of steady-state temperature conditions.

Combinations of T_{OA} and T_R yielding a given mixed air temperature correspond to straight lines in this representation. Combinations in which $T_M = T_S$ are straight lines through the *origin*. However, in view of the constraints imposed on χ, only those lines with slopes between

$$\frac{\chi_{max}}{1 - \chi_{max}} \quad \text{and} \quad \frac{\chi_{min}}{1 - \chi_{min}} \qquad (4)$$

correspond to situations in which the supply temperature can be achieved by operating just the dampers. Otherwise, depending on whether or not $T_{OA} \le T_R$ and $T_S < T_R$, the optimal damper setting will correspond to χ_{min} or χ_{max}.

The various regimes are depicted in the diagram. The line $T_{OA} = T_R$ determines when χ should be switched from χ_{min} to χ_{max} or vice versa. The white regions are those in which it is only necessary to operate the dampers to achieve the desired effect. The shaded regions in the upper right sectors require cooling, with $\chi = \chi_{max}$, if $T_{OA} \le T_R$, or $\chi = \chi_{min}$ otherwise. Similarly, the shaded regions in the lower left sectors require heating, with $\chi = \chi_{min}$, if $T_{OA} \le T_R$, or $\chi = \chi_{max}$ otherwise. The six operating regimes can be summarized as follows:

1. $T_{OA} \le T_R$ and T_{OA} comparatively low: dampers set for minimal outside air and controller operates heating coil.

2. $T_{OA} \le T_R$ and $T_M = T_S$ can be achieved within the operating range of the bypass dampers: heating and cooling switched off and controller operates dampers.

3. $T_{OA} \le T_R$ and T_{OA} comparatively high: dampers set for maximal outside air and controller operates cooling coil.

4. $T_{OA} > T_R$ and T_{OA} comparatively high: dampers set for minimal outside air and controller operates cooling coil.

5. $T_{OA} > T_R$ and $T_M = T_S$ can be achieved within the operating range of the bypass dampers: heating and cooling switched off and controller operates dampers.

6. $T_{OA} > T_R$ and T_{OA} comparatively low: dampers set for maximal outside air and controller operates heating coil.

However, Cases 5 and 6 above can normally be ruled out. Since VAV boxes are not equipped with cooling coils, the only cooling effects in the zones are thermal losses when either the outside air is cold or the walls are cold. Thus, in steady-state situations, it would only be expected that $T_R < T_Z$ would occur if the outside temperature is sufficiently *cold* - i.e. that $T_{OA} < T_R$. In fact, barring transient effects (taking the complete system into account), no temperature states may be expected to be found anywhere in the lower right quadrant of Figure 3, so that the main cases of interest are Cases 1 to 4, which correspond to the four operating regimes prescribed for Controller C_1 in the reference system (Kelly, 1993).

4. DESIGN OF QUALITATIVE FAULT DETECTORS OF THE CENTRAL AIR-HANDLING PLANT

4.1 General structure of the fault detector

The task of a fault detector is to recognize those changes of transient or steady-state behaviour of the air-handling system which arise from faults. The types described here operate on the central air-handling plant and are fault *detectors* in the strict sense: they only detect the *presence* of faults, without attempting to diagnose possible causes. They are particularly simple illustrative examples and are not capable of *localizing* or *diagnosing* faults. Three qualitative model-based fault detectors are considered.

The overall structure of such a qualitative model-based fault detector is shown in Figure 4. This conforms to the structure of the so-called general diagnostic engine ("GDE") (de Kleer and Williams, 1987; Dexter and Glass, 1993). The fault detector models being considered are also related to the generic fault detection and diagnosis ("FDD") scheme proposed by Rossi and Braun (1993).
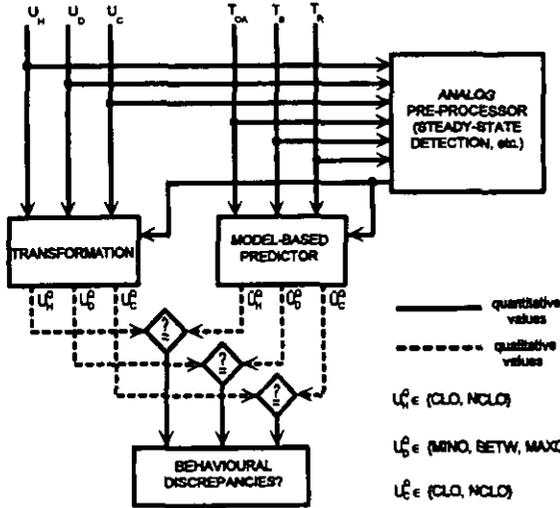
Fig. 4. A qualitative model-based fault detector.

From the central air-handling plant the observed values of the temperatures T_{OA}, T_R, T_S and the control variables U_H, U_D, U_C are obtained and fed into the detector. There they are used in the steady-state detector which initializes both the transformation of the quantitative control values U_H and U_D, U_C to qualitative values U_H^Q, U_D^Q and U_C^Q and, at the same time, the generation of the corresponding predicted values $\hat{U}_H^Q$, $\hat{U}_D^Q$ and $\hat{U}_C^Q$.

The qualitative values of the outputs of the transformation and predictor blocks are chosen from the following sets:

$$\begin{cases} U_H^Q \in \{CLO, NCLO\}, \\ U_D^Q \in \{MINO, BETW, MAXO\}, \\ U_C^Q \in \{CLO, NCLO\}. \end{cases} \quad (5)$$

$$\begin{cases} \hat{U}_H^Q \in \{CLO, NCLO\}, \\ \hat{U}_D^Q \in \{MINO, BETW, MAXO\}, \\ \hat{U}_C^Q \in \{CLO, NCLO\}. \end{cases} \quad (6)$$

This corresponds to the idea of defining qualitative values in terms of so-called "landmarks" (Kuipers, 1984, 1985, 1986; Dexter and Glass, 1993), the landmarks in this case being provided by the extreme settings of the various actuators in the CAHP.

The predictor block is based on a qualitative prediction model of the plant to be supervised. The steady-state behaviour of the central air-handling plant is used as prediction model with the temperatures T_{OA}, T_R and T_S as inputs and the predicted control variables $\hat{U}_H^Q$, $\hat{U}_D^Q$ and $\hat{U}_C^Q$ as the outputs. The outputs of the transformation block and the predictor are then compared and checked for behavioural discrepancies, as can be seen in Figure 4.

Fig. 5. Graphical representation of predicted qualitative actuator settings in terms of steady-state temperature conditions.

The differences among the three detectors presented here concern the *predictor*.

4.2 Fault Detector 1

The first type of fault detector uses the graphical representation of Figure 3 as the prediction model. For each temperature triple T_{OA}, T_R, T_S a point in the plane of Figure 3 can be determined whose location fully determines the qualitative outputs $\hat{U}_H^Q$, $\hat{U}_D^Q$ and $\hat{U}_C^Q$ of the predictor. Figure 5 shows how the plane of the graphical representation of Figure 3 can be partitioned into regions according to the predicted actuator settings. Each region is labelled with the qualitative predicted actuator settings that correspond to the correct behaviour of the plant. The predicted actuator settings are also shown on some of the lines separating the regions. On the line $T_R = T_{OA}$, however, the predicted values of $\hat{U}_D^Q$ are ambiguous.

Three examples illustrate how Fault Detector 1 works. The corresponding steady-state temperatures are shown in Figure 5.

Example 1

A steady state with the following values is detected:

$$T_{OA} = T_S, \quad T_{OA} < T_R, \quad \text{fault: none} \quad (7)$$

It follows that (as indeed the simulations confirm)

$$U_H = 0, \qquad U_D = 1, \qquad U_C = 0 \quad (8)$$

The transformation block generates the following qualitative control values:

$$U_H^Q = CLO, \quad U_D^Q = MAXO, \quad U_C^Q = CLO \quad (9)$$

The steady-state temperatures which are fed into the predictor lead to the points labelled "Ex.1" in Figure 5. This produces the qualitative prediction values:

$$\hat{U}_H^Q = CLO, \quad \hat{U}_D^Q = MAXO, \quad \hat{U}_C^Q = CLO \quad (10)$$

No discrepancy is found, so no fault is detected, which is indeed correct.

Example 2

A steady-state with the following values is detected:

$$T_{OA} < T_S, \quad T_{OA} < T_R, \quad \textit{fault: cooling valve}$$
$$\textit{cannot close if } U_C = 0$$
$$(11)$$

For example:

$$T_{OA} = -15°C; \quad T_S = 18°C; \quad T_R = 20°C \quad (12)$$

Simulation yields the result that

$$U_H = 0.593, \quad U_D = 0.2 \text{ (min.)}, \quad U_C = 0 \quad (13)$$

The transformation block now generates

$$U_H^Q = NCLO, \quad U_D^Q = MINO, \quad U_C^Q = CLO \quad (14)$$

The steady-state temperatures which are fed into the predictor lead to the points labelled "Ex.2" in Figure 5. The prediction is

$$\hat{U}_H^Q = NCLO, \quad \hat{U}_D^Q = MINO, \quad \hat{U}_C^Q = CLO \quad (15)$$

No discrepancy is found, so no fault is detected, although one is in fact present. Thus, in this particular operating steady-state, the fault described cannot be detected. The air is unnecessarily heated up and cooled down, which results in wasted energy. In the simulation of the above example *without* the fault, the steady-state heating valve setting turned out to be $U_H = 0.514$.

Example 3

The same steady-state is detected as in Example 1, but this time the fault of Example 2 occurs, namely the cooling valve cannot close completely if $U_C = 0$.

$$T_{OA} = T_S, \quad T_{OA} < T_R \quad \textit{fault: cooling valve}$$
$$\textit{cannot close if } U_C = 0$$
$$(16)$$

---

[2] Note that the order of the controller outputs in this program is (U_D, U_H, U_C), whereas throughout the remainder of this paper the outputs have been shown in the order corresponding to the controller sequence, namely (U_H, U_D, U_C).

Simulation yields the values

$$U_H = 0, \quad U_D = 0.916, \quad U_C = 0 \quad (17)$$

The transformation block generates the following qualitative control values:

$$U_H^Q = CLO, \quad U_D^Q = BETW, \quad U_C^Q = CLO \quad (18)$$

The location of this steady-state in the diagrams of Figure 5 is the same as in Example 1. Consequently, the predictor produces the following qualitative

```
airhandler(AirhandleState,Ud,Uh,Uc,D_as,D_rs,D_ra) :-
    sum(D_rm,D_ms,D_rs),
    sum(D_am,D_ms,D_as),
    sum(D_mh,D_hs,D_ms),
    damper(DampState,Ud,D_am,D_rm),
    heater(HeatState,Uh,D_mh),
    cooler(CoolState,Uc,D_hs),
    controller(ContrState,Ud,Uh,Uc,D_ra),
    sum(D_ra,D_as,D_rs),
    airhandlcond(AirhandleState,DampState,HeatState,
        CoolState,ContrState).

%airhandlcond(AirhandleState,DampState,HeatState,
        CoolState,ContrState). airhandlcond(ok,ok,ok,ok,ok).

%damper (DampState,Ud,D_am,D_rm)
damper(ok,maxo,zero,Temp) :- anyTemp(Temp).
damper(ok,betw,neg,pos).
damper(ok,mino,neg,pos).
damper(ok,betw,zero,zero).
damper(ok,mino,zero,zero).
damper(ok,betw,pos,neg).
damper(ok,mino,pos,neg).

%heater(HeatState,Uh,D_mh).
heater(ok,clo,zero).
heater(ok,nclo,neg).

%cooler(CoolState,Uc,D_hs).
cooler(ok,clo,zero).
cooler(ok,nclo,pos).

%controller(ContrState,Ud,Uh,Uc,D_ra)
controller(ok,mino,nclo,clo,pos).
controller(ok,DampPos,nclo,clo,zero) :- anyDampPos(DampPos).
controller(ok,maxo,nclo,clo,neg).
controller(ok,maxo,clo,nclo,pos).
controller(ok,DampPos,clo,nclo,zero) :- anyDampPos(DampPos).
controller(ok,mino, clo,nclo,neg).
controller(ok,DampPos,clo,clo,Temp) :- anyDampPos(DampPos),
    anyTemp(Temp).

%rest
sum(X,X,X) :- anyTemp(X).
sum(X,zero,X) :- anyTemp(X).
sum(zero,X,X) :- anyTemp(X).
sum(neg,pos,Temp) :- anyTemp(Temp).
sum(pos,neg,Temp) :- anyTemp(Temp).

anyTemp(neg).
anyTemp(zero).
anyTemp(pos).

anyDampPos(mino).
anyDampPos(betw).
anyDampPos(maxo).
```

Fig. 6. PROLOG program[2] for Fault Detector 2.

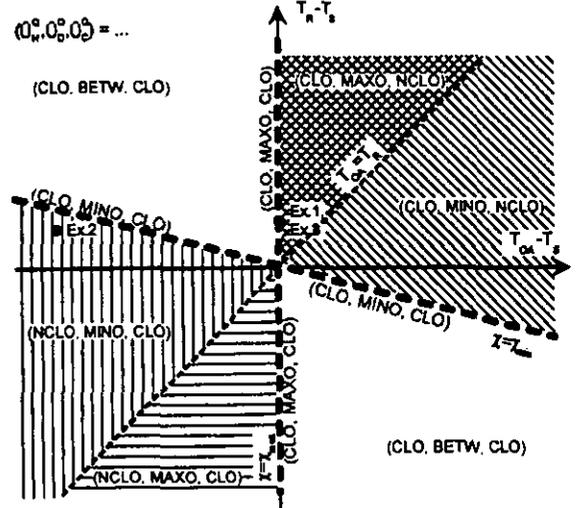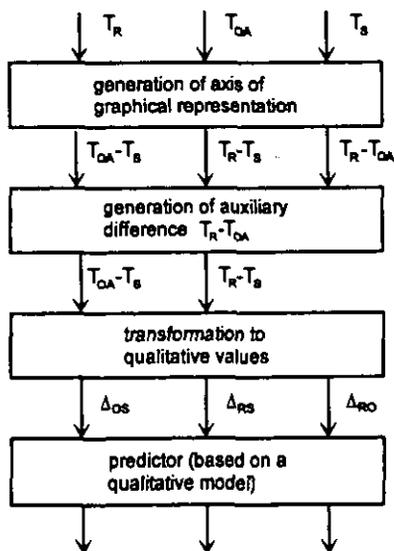Fig. 7   Structure of the predictor for Fault Detector 2

Fig. 8. Graphical   representation   of   predicted qualitative   actuator   settings   in   terms   of *qualitative* steady-state temperature conditions (classified in terms of the signs of Δ_{OS}, Δ_{RS} and Δ_{RO}).

predicted control values:

$$\hat{U}_H^Q = CLO, \qquad \hat{U}_D^Q = MAXO, \quad \hat{U}_C^Q = CLO \qquad (19)$$

In this case the first predicted value differs from U_D^Q. That means a fault has now been detected.

If in Figure 3 the separation line representing the minimal damper position is missing or not known exactly, a reduced version of Fault Detector 1 can be derived, which is simpler. The reduced version is also able to detect the fault of Examples 2 and 3. This remark applies especially to the situation occurring in the reference system, in which the minimal outside air flow is absolute (20% of the maximum flow) rather than proportional. In this case, the slope of the line corresponding to χ_{min} will depend on the instantaneous total airflow through the CAHP.

4.3 Fault detector 2

The design of Fault Detector 2 starts by describing the qualitative behaviour of the *components* of the central air-handling plant: damper, cooler, heater and controller. Then the interaction of these components among themselves is described by defining the *structure* of the system incorporating them. The resulting qualitative prediction model is represented by logical clauses.

The form of the clauses is such that it can be executed as a PROLOG program, which is shown in Figure 6. It is structured as follows: in the first part of the code the structure of the air-handling unit is programmed – i.e. the way in which the components are linked. The second part deals with the components: damper, heater, cooler and controller. Finally, there are some auxiliary statements needed for the logical clauses.

The overall structure of the predictor is shown in Figure 7. It makes use of the auxiliary temperature differences

$$\Delta_{OS} = T_{OA} - T_S; \qquad\qquad \Delta_{RS} = T_R - T_S;$$

$$\Delta_{RO} = T_R - T_{OA}. \qquad\qquad\qquad (20)$$

The qualitative predictor consists of the qualitative prediction model and the PROLOG interpreter.

In the qualitative predictor block a prediction-type question such as

$$?- airhandler\left(ok, \hat{U}_H^Q, \hat{U}_D^Q, \hat{U}_C^Q, 0, pos, pos\right) \quad (21)$$

is posed to the PROLOG program. The predicted controller states compatible with fault-free operation are shown in Figure 8, based the graphical representation of Figure 3. It differs from Figure 5 in that the plane is subdivided into regions according to whether the *auxiliary* temperature differences Δ_{OS}, Δ_{RS} and Δ_{RO} are postive, negative or zero. As a result, three different qualitative controller states may be predicted for steady-state temperatures in the upper left or lower right quadrants.

The operation of Fault Detector 2 is illustrated using the same three examples as for Fault Detector 1.

Example 1'

The steady-state is as in Example 1. The following qualitative controller values are observed:

$$U_H^Q = CLO, \qquad U_D^Q = MAXO, \quad U_C^Q = CLO \quad (22)$$

From the temperature measurements the qualitative values of the three temperature differences are calculated as:

$$\Delta_{OS} = 0, \qquad \Delta_{RS} = pos, \qquad \Delta_{RO} = pos \qquad (23)$$

With the assumption that the behaviour of the air handler is OK, the following predictor-type question can be asked:

$$? = airhandler\left(\underbrace{\underset{|}{ok,}}_{assumption} \underbrace{\hat{U}_H^Q, \hat{U}_D^Q, \hat{U}_C^Q,}_{predicted~result} \underbrace{0, pos, pos}_{measurements} \right) \qquad (24)$$

The result of the PROLOG program is (see appendix)

$$\hat{U}_H^Q = CLO, \qquad \hat{U}_D^Q = MAXO, \qquad \hat{U}_C^Q = CLO \qquad (25)$$

So again there are no discrepancies. Therefore no fault is detected, which is correct.

Example 2'

The steady-state is as in Example 2. The observed qualitative controller values are the same:

$$U_H^Q = NCLO, \quad U_D^Q = MINO, \quad U_C^Q = CLO \qquad (26)$$

The qualitative temperature differences are:

$$\Delta_{OS} = neg, \qquad \Delta_{RS} = pos, \qquad \Delta_{RO} = pos \qquad (27)$$

The question posed now is

$$? = airhandler\left(\underbrace{\underset{|}{ok,}}_{assumption} \underbrace{\hat{U}_H^Q, \hat{U}_D^Q, \hat{U}_C^Q,}_{predicted~result} \underbrace{neg, pos, pos}_{measurements} \right)$$
$$(28)$$

The result of the PROLOG program is (see appendix)

$$\hat{U}_H^Q = NCLO, \quad \hat{U}_D^Q = MINO, \quad \hat{U}_C^Q = CLO \qquad (27)$$

Again the fault remains undetected, because no discrepancies are observed.

Example 3'

The steady-state is as in Example 3. The same question is asked with the same input arguments generating the same predicted qualitative control values. The check for the behavioural discrepancy leads to the same result as in Example 2: the fault detector has detected that there is a fault.

Modelling the prediction model in this way has been motivated by work of Bratko *et al.* (1989) and Burkhard (1992). But in contrast to the latter's work, the plant has been modelled here in a way that is independent of the particular sizing of the plant being monitored. This is important, since the adaptation of the fault detector (or FDD-system) to the plant should be as simple as possible and, furthermore, be based on easily available information.

4.4 Fault detector 3

Fault detector 3 is the same as Fault Detector 2 but the predictor part is in a compiled form, i.e. the predictor is in a form which, for a given input (Δ_{OS}, Δ_{RS}, Δ_{RO}), outputs the corresponding possible controller states ($\hat{U}_H^Q$, $\hat{U}_D^Q$, $\hat{U}_C^Q$). Such a form is, for example, the decision table shown in Table 2.

This table has been derived by posing for each possible input to the predictor a predictor-type question to the PROLOG program. For two input cases this leads to situations with three possible predictions (see Table 2).

Table 2. Predictor part of Fault Detector 3, realized as a decision table. Signs of Δ_{RO} in brackets are irrelevant to the prediction. Predictions in curly brackets are sets of alternative possibilities.

Δ_{OS}	Δ_{RS}	Δ_{RO}	Prediction No.	$\hat{U}_H^Q$	$\hat{U}_D^Q$	$\hat{U}_C^Q$
pos	pos	pos	1	CLO	MAXO	NCLO
pos	pos	0	1	CLO	{MINO, BETW, MAXO}	NCLO
pos	pos	neg	1	CLO	MINO	NCLO
pos	0	(neg)	1	CLO	MINO	NCLO
pos	neg	(neg)	1	CLO	MINO	NCLO
			2	CLO	MINO	CLO
			3	CLO	BETW	CLO
0	neg	(neg)	1	CLO	MAXO	CLO
neg	neg	neg	1	NCLO	MAXO	CLO
neg	neg	0	1	NCLO	{MINO, BETW, MAXO}	CLO
neg	neg	pos	1	NCLO	MINO	CLO
neg	0	(pos)	1	NCLO	MINO	CLO
neg	pos	(pos)	1	NCLO	MINO	CLO
			2	CLO	MINO	CLO
			3	CLO	BETW	CLO
0	pos	(pos)	1	CLO	MAXO	CLO
0	0	(0)	1	CLO	{MINO, BETW, MAXO}	CLO

Examples:

If the same examples as for Fault Detector 2 are considered, the same results are obtained.

4.5 Comparison of the three fault detectors

1) As mentioned before, all three fault detectors have the structure of a GDE and all distinguish the same qualitative values for the control variables U_H^Q, U_D^Q, U_C^Q. Their differences lies in the predictors.

2) Fault Detectors 2 and 3 are equivalent in their overall behaviour. This means that they give the same output in the same situation. Fault detector 1 however is not equivalent to Fault Detectors 2 and 3. It *generates different answers in some situations*. This can be seen by comparing the graphical representation of the predictor of Fault Detector 2 or 3 with the graphical representation of the predictor of Fault Detector 1. The comparison shows that the predictor of Fault Detector 2 and 3 *corresponds to the reduced* version of Fault Detector 1. But although Fault Detector 1 is not equivalent to Fault Detector 2 and 3, *there is no contradiction in their* behaviour. This verifies the derivations of the fault detectors, at least to some extent.

3) The difference between Fault Detectors 2 and 3 lies in the predictor. The predictor of Fault Detector 2 has to run a PROLOG program in each detected steady-state of the central air-handling plant. In contrast to this, the predictor of Fault Detector 3 has only to look up some values in a decision table. The predictor of Fault Detector 2 is model-based, and the model is a deep model in the sense that it is composed of component models and a description of how the components are linked together. In contrast, the predictor of Fault Detector 3 does not explicitly contain a deep model, even if it is derived originally from a deep model of the plant.

5. CONCLUSIONS AND OUTLOOK

As is evident from the examples discussed, it is possible to define a fault detector for a central air-handling plant based solely on *qualitative* observable features. The results of the predictions using this qualitative model are in agreement with the quantitative simulations carried out.

As the examples show, qualitative tests require only a minimal knowledge of the system parameters. On the other hand, as expected (Dexter and Glass, 1993), the qualitative tests investigated here were not always able to discern faults that quantitative methods might

have identified. Moreover, a potential limitation of the method tried is that the particular steady-states of the system that lend themselves to such diagnosis may occur very infrequently in *normal operation*. One way of overcoming such a problem would be to devise active tests – say during the time when a building is unoccupied – to set up the conditions that allow for such fault detection.

This work is being continued by implementing a steady-state detector in the simulation program being developed, and linking the diagnostic tests with actual simulations (Glass *et al.*, 1994, 1995). It is also planned to implement the types of *active* diagnostic tests mentioned in the previous paragraph. As the research work progresses, the methods developed for the CAHP will be extended to the full reference air-handling system and to the actual localization and diagnosis of faults [e.g. using an assumption-based truth maintenance system (Gelle, 1993)].

The fault detectors considered here detect faults by analysing one steady-state of the plant (or a subsystem of the plant). Qualitative fault detectors which also detect faults by comparing two (or more) different steady-states together or by analysing transients will be considered in future work. Knowing that a function is monotonic, for example, can be exploited in many cases [cf. work by Koch (1992)].

6. ACKNOWLEDGEMENT

7. REFERENCES

Bratko, I., Mozetic, I., and Lavrac, N. (1989). *KARDIO: A study in Deep and Qualitative Knowledge for Expert Systems*, M.I.T. Press, Cambridge, U.S.A.

Burkhard, E. (1992). "Wissensbasierte Modellierung einer Heizungs-Lüftungs-Klima-Anlage" Diploma thesis, Lehrstuhl für Logik und Informatik, ETH Zürich, September, 1992.

Dexter, A.L. and Glass A.S (1993). "The use of qualitative models in fault detection and diagnosis", §4.2 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of

Finland (VTT), Laboratory of Heating and Ventilation, P.O. Box 206 (Lämpömiehenkuja 3), FIN-02151 Espoo, Finland (ISBN952-9601-16-6).

de Kleer, J. and Williams, B.(1987). "Diagnosing Multiple Faults", *Artificial Intelligence*, 32(1), 97-130.

Fornera, L., Glass, A.S., Gruber, P. and Tödtli, J. (1993). "Fault Detection Based on Logical Programming Applied to the VAV Reference System: First Trials", *I.E.A. Annex 25* Working Paper AN25/CH/050493/1.0, Tokyo meeting, 14-16.4.1993.

Fornera, L., Glass, A.S., Gruber, P. and Tödtli, J. (1994), "Qualitative Fault Detection Based on Logical Programming Applied to a VAV Air Handling Unit", *2nd IFAC Workshop on Computer Software Structures Integrating AI/KBS Systems in Process Control*, Lund, 10-12.8.1994.

Gelle, E. (1993). "Diagnostic system for a heating-cooling unit", diploma thesis, Laboratoire d'Intelligence Artificielle, École Polytechnique Fédérale de Lausanne, March, 1993.

Glass, A.S., Gruber, P. , Roos,"M. and Tödtli, J. (1994). "Preliminary Evaluation of a Qualitative Model-Based Fault Detector for a Central Air-Handling Unit", *1994 IEEE Conference on Control Applications*, Glasgow, August 24-26, 1994.

Glass, A.S., Gruber, P. , Roos,"M. and Tödtli, J. (1995). "Qualitative Model-Based Fault Detection in Air-Handling Units", *IEEE Control Systems Magazine*, 15(4), 11-22, August, 1995.

Kelly, G.(1993) "Air conditioning unit", §5.3 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland (VTT), Laboratory of Heating and Ventilation, P.O. Box 206 (Lämpömiehenkuja 3), FIN-02151 Espoo, Finland (ISBN952-9601-16-6).

Koch, G.(1992). "Knowledge-Based Coordination of Qualitative Online Diagnostic Tests", *IFAC Symposium on Intelligent Components and Instruments for Control Applications*, May 20-22, 1992, Malaga, Spain.

Kuipers, B (1984). "Commonsense reasoning about causality: Deriving behaviour from structure", *Artificial Intelligence*, 24(1-3), 169-203.

Kuipers, B. (1985). "The Limits of Qualitative Simulation", *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, William Kaufman, Los Altos, USA.

Kuipers, B (1986). "Qualitative Simulation", *Artificial Intelligence*, 29(3), 289-388.

Rossi, T. and Braun, J., "Classification of Fault Detection and Diagnostic Methods" (1993). §2 in

Building Optimisation and Fault Diagnosis System Concept, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland (VTT), Laboratory of Heating and Ventilation, P.O. Box 206 (Lämpömiehenkuja 3), FIN-02151 Espoo, Finland (ISBN952-9601-16-6).

APPENDIX I

PROLOG Program Pertaining to the Decision Table of Fault Detector 3 (cf. Table 2)

```
:- airhandler(ok, Ud, Uh, Uc, pos, pos, pos)
3 times        Ud=maxo,Uh=clo,Uc=nclo

:- airhandler(AirhandleState, Ud, Uh, Uc, pos, pos, zero)
3 times        AirhandleState=ok,Ud=maxo,Uh=clo,Uc=nclo
3 times        AirhandleState=ok,Ud=betw,Uh=clo,Uc=nclo
3 times        AirhandleState=ok,Ud=mino,Uh=clo,Uc=nclo

:- airhandler(ok, Ud, Uh, Uc, pos, pos, neg)
3 times        Ud=mino,Uh=clo,Uc=nclo

:- airhandler(ok, Ud, Uh, Uc, pos, zero, neg)
1 time         Ud=mino,Uh=clo,Uc=nclo

:- airhandler(ok, Ud, Uh, Uc, pos, neg, neg)
3 times        Ud=betw,Uh=clo,Uc=clo
3 times        Ud=mino,Uh=clo,Uc=clo
1 time         Ud=mino,Uh=clo,Uc=nclo

:- airhandler(ok, Ud, Uh, Uc, zero, neg, neg)
9 times        Ud=maxo,Uh=clo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, neg, neg, neg)
3 times        Ud=maxo,Uh=nclo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, neg, neg, zero)
3 times        Ud=maxo,Uh=nclo,Uc=clo
3 times        Ud=betw,Uh=nclo,Uc=clo
3 times        Ud=mino,Uh=nclo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, neg, neg, pos)
3 times        Ud=mino,Uh=nclo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, neg, zero, pos)
1 time         Ud=mino,Uh=nclo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, neg, pos, pos)
3 times        Ud=betw,Uh=clo,Uc=clo
3 times        Ud=mino,Uh=clo,Uc=clo
1 time         Ud=mino,Uh=nclo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, zero, pos, pos)
9 times        Ud=maxo,Uh=clo,Uc=clo

:- airhandler(ok, Ud, Uh, Uc, zero, zero, zero)
infinite times  Ud=maxo,Uh=clo,Uc=clo
infinite times  Ud=betw,Uh=clo,Uc=clo
infinite times  Ud=mino,Uh=clo,Uc=clo
```
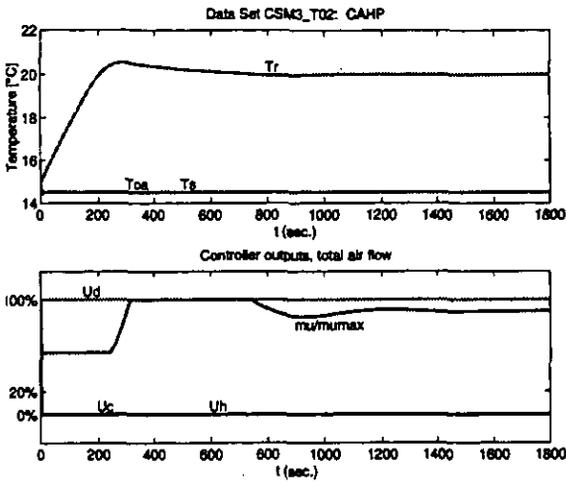
Fig. 9. Simulation of the CAHP in Example 1, showing T_{OA}, T_S & T_R (upper plot) and U_H, U_D, U_C and relative air flow (lower plot).

Fig. 10. Simulation of one of the zones in Example 1, showing T_Z (upper plot) and the VAV controller outputs U_H & U_C (lower plot).

APPENDIX II

Simulation of the simplified reference Air-Handling System

The dynamic behaviour of the simplified reference system was simulated numerically using the SIMULINK™ package using the physical parameters specified by Kelly (1993). In each simulation the outside air temperature T_{OA} was constant and the initial temperatures of the zones was 15°C. The system was allowed to run for a simulated time of 30 min, with the zone controllers acting to try and achieve a set-point temperature of 20°C and the CAHP controller acting to try and achieve the supply-air temperatures noted below.

In each case the simulated system required about 10-15 min. to reach steady-state in both the central air-

Fig. 11. Simulation of the CAHP in Example 2, showing T_{OA}, T_S & T_R (upper plot) and U_H, U_D, U_C and relative air flow (lower plot).

handling plant and the zones. In all cases, zone temperatures acceptably close to 20°C were achieved within approximately 10 min and the zone controller outputs attained an acceptable equilibrium within 15 to 20 min. The supply-air set-point temperatures were, as a rule, attained much faster.

Example 1

Relevant inputs:

Outside air temperature	T_{OA}	14.5°C
Supply air set point temperature	T_S^{SP}	14.5°C
Zone set point temperature	$T_{Z_t}^{SP}$	20°C

Faults: *none.*

Values of relevant quantities attaining steady-state.

Return air temperature	T_R	20.001°C
Supply air temperature	T_S	14.500°C
Controller outputs	U_H	0.000
	U_D	1.000
	U_C	0.000

Example 2

Relevant inputs:

Outside air temperature	T_{OA}	-15°C
Supply air set point temperature	T_S^{SP}	18°C
Zone set point temperature	$T_{Z_t}^{SP}$	20°C

Fault: *cooling valve cannot close completely - cooling coil operates at a minimum of 2.5% of its 40600 kW capacity.*

Values of relevant quantities attaining steady-state:

Return air temperature	T_R	c. 20°C
Supply air temperature	T_S	c. 18°C
Controller outputs	U_H	0.593
	U_D	0.200
	U_C	0.000

Example 3

Relevant inputs:

Outside air temperature	T_{OA}	14.5°C
Supply air set point temperature	T_S^{SP}	14.5°C
Zone set point temperature	$T_{Z_x}^{SP}$	20°C

Fault: *cooling valve cannot close completely* - cooling coil operates at a minimum of 2.5% of its 40600 kW capacity.

Values of relevant quantities attaining steady-state.

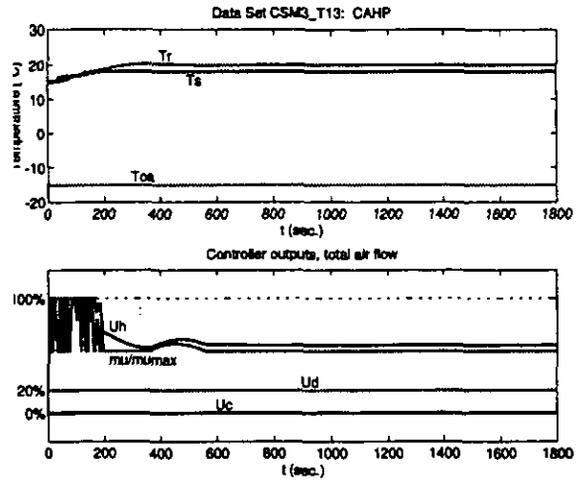|		
Return air temperature	T_R	20.001°C
Supply air temperature	T_S	c. 14.5°C
Controller outputs	U_H	0.000
	U_D	0.916
	U_C	0.000
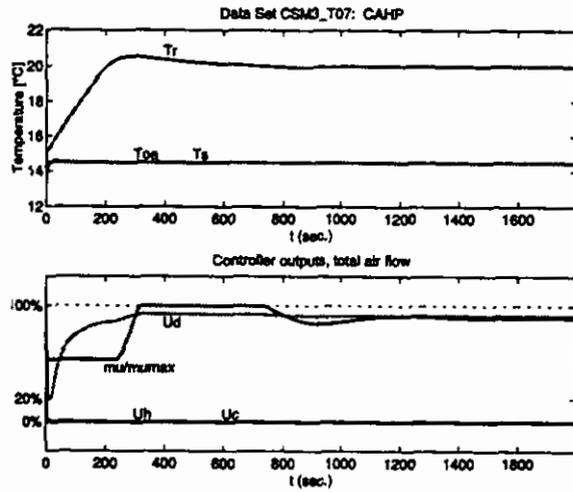
Fig. 12. Simulation of the CAHP in Example 3, showing T_{OA}, T_S & T_R (upper plot) and U_H, U_D, U_C and relative air flow (lower plot).
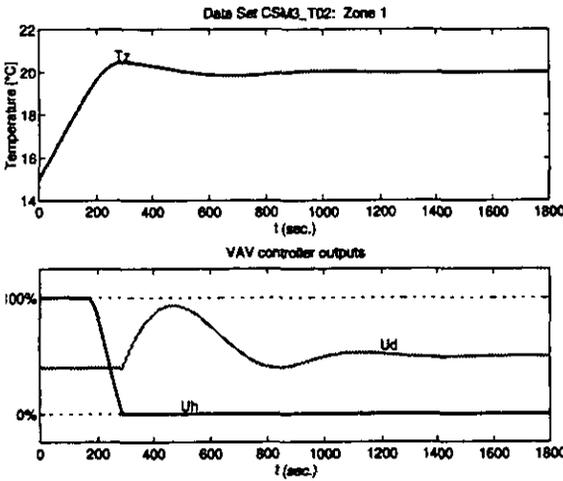
Testing Qualitative Model-Based Fault Detection for Air-Handling Units Using Operational Building Data

A.S. Glass[*] & J. Tödtli

Landis & Gyr, CH-6301 Zug, Switzerland

Abstract

The feasibility of a *qualitative approach* for detecting faults in an air-conditioning system is field-tested on a fully operational central air-handling plant. The particular interest in qualitative methods derives from the fact that many relevant parameters of an HVAC system and building are very difficult to measure and, in consequence, usually remain undetermined in practice. In the approach investigated, plant controller outputs are *predicted* on the basis of temperature data and reduced to qualitative values which are then compared with the corresponding *observed* qualitative values.

The building considered has at its heart a central air-handling unit with more-or-less constant air volume, supplying air to several zones with independently controlled ceiling-mounted air coolers and hot-water radiators. The methods developed for the Annex 25 reference air-handling system were applied to this plant, which differs from the Annex 25 standard in using a heat-recovery wheel in lieu of an air bypass mixer. The steady-state detection parameters were tuned to the known building data and the overall method was applied off-line to operational data captured from the building energy management system.

It was demonstrated that the required degree of stationarity is reached sufficiently often in practice for the method to be useful. Moreover, an unplanned fault was detected and it showed up sufficiently often to demonstrate the practical usefulness of the qualitative approach. Finally, the successful application of a method developed for one type of central air-handling unit to a related, but different plant type illustrates the generic nature of this qualitative fault-detection procedure.

1 INTRODUCTION

This paper extends and consolidates the investigations initially reported by Fornera *et al* [1] and Glass [2], in which a *qualitative model-based* approach to detecting a class of faults in a variable air-volume air-handling system was developed. As reported by the above authors [1,3,4,2,5] the potential of the method was successfully demonstrated on numerically simulated models of the Annex 25 standard variable air-volume air-handling system (Kelly [6]) in a MATLAB/SIMULINK[1] simulation environment. This was followed by successful trials in a heating, ventilation and air-conditioning laboratory (abbreviated ."HVAC laboratory") in which the method succeeded in detecting faults under controlled conditions (Glass, *et al* [5,7,8]). This paper reports on successful tests carried out on building energy-management data from a fully operational commercial building in Switzerland.

The qualitative approach identifies certain so-called "landmark" states of the controller regulating a central air-handling unit and relates them to corresponding "landmarks" of the relevant air temperatures and flows. Using this approach, a wide class of operating faults can be analysed. In particular, we also deal with faults which result in deterioration of operation as distinct from actual failure — a situation known to occur frequently in practice. Conflicts between qualitative values of the controller states as actually observed and those expected — as derived from the measured qualitative temperature states — provide evidence of such faults. Furthermore, the procedure relies solely on measurements required by the control system; no extra sensors are needed to detect these faults.

The qualitative detection method described pre-supposes the temperatures and control settings of the central air handling unit (abbreviated ."CAHU") to be quasi-stationary. Consequently' a prototype steady-state detector (abbreviated ."SSD") was developed and tested in both simulation models and in the HVAC laboratory (Glass [2,5]). An important question concerns appropriate *tuning* of the *time-frame* and *threshold* parameters so that the steady-state detectors are matched optimally to the fault detectors for which they are needed. A theoretical analysis of the threshold problem for various designs of a steady-state detector was provided by Tödtli and Gruber [9] and a practical procedure for tuning the SSD parameters to a given air-handling plant (abbreviated "AHU") has been successfully developed by Gruber [10,11]. An overview of steady-state detection is included in the IEA Annex 25 Final Report [12].

Figure 1 *The Simplified Annex 25 Reference Air-Handling System*
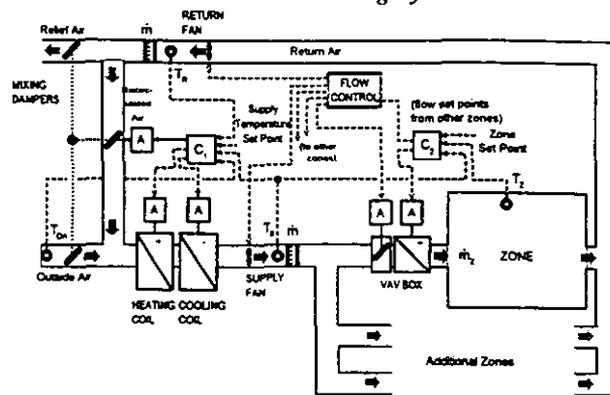
This paper includes a description of the type of systems under test: both the CAHU for which the fault-detection method was developed and the CAHU actually encountered in the test-site. This is followed by an outline of adapting the fault-detection procedure to the modified type of plant, a brief section on relevant fault models, and a section on tuning the steady-state detectors. Finally selected data from the test-site is analysed.

2. DESCRIPTION OF THE TEST SYSTEMS

The fault-detection tests described in this paper were carried out on the central air-handling unit used in a commercial building in central Switzerland. It differs in a number of respects from the Annex 25 reference air-handling system, which is more closely representative of *North American* as distinct from European HVAC practice. However, the central air-handling unit at the test-site is *generically* sufficiently similar to the reference system to allow effective testing of the qualitative fault-detection method developed for the latter. We begin by reviewing the features of the Annex 25 reference air-handling system and then describe the differences found in the plant at the test site.

2.1 THE ANNEX 25 REFERENCE AIR-HANDLING SYSTEM

At the beginning of the investigations, simulation tests were carried out using a SIMULINK model of a *simplified* variable air-volume system (abbreviated "VAV system") based on the Annex 25 reference air-handling system described by Kelly [6]. This system comprises a controlled central air-handling unit (CAHU) supplying air to three independently controlled zones with differing loads. Both the reference system and the simplified version are described in more detail in Fornera *et al* [1,3,4]. We summarize the components and control strategy of the simplified system below.

2.1.1 The components of the simplified Annex 25 reference system

The overall simplified reference system is illustrated in Figure 1. It consists of a central air-handling unit (CAHU), depicted in Figure 2, supplying air to a number of separately controlled zones. The CAHU comprises a bypass mixer and a heating coil followed by a cooling coil in a single air duct. The bypass dampers are controlled to provide a mixture of outside air and recirculated air; the amount of outside air may vary between 20% and 100% of the maximum airflow through the CAHU. The air for each of the three zones is processed through a VAV box containing a damper and a re-heating coil. The airflow to each zone as regulated by the damper may vary between 40% and 100% of the specified maximum. In point of fact, the maximum flow from the CAHU is not sufficient to provide maximum flows through all three zones simultaneously, so limiting effects can be expected to occur.

The system illustrated differs from standard North American practice in *not* including dehumidification of the cooling coil. However, the *qualitative* fault detection technique described below can also be applied to some such systems.

2.1.2 The control strategy of the simplified Annex 25 reference system

Three controllers share the task of regulating the fans, dampers, heating coils and cooling coils so as to attain the required temperatures and airflows in each of the three zones. Referring to Figure 1, Controller C_1 regulates the CAHU and attempts to maintain the supply air temperature T_S at its set point by operating the preheating coil, dampers and cooling coil in sequence. Controller C_2, which regulates a single zone, attempts to maintain the zone temperature T_Z at its set point by operating in sequence the damper and re-heating coil in the VAV box. The third controller is an idealized "flow control" governing the airflow through the CAHU and the various zones. It ensures that the flows $\dot{m}_{Z_1}$, $\dot{m}_{Z_2}$ and $\dot{m}_{Z_3}$ meet the requirements set by the zone controller C_2, provided the resulting total airflow $\dot{m}$ through the CAHU does

Figure 2   *The Central Air-Handling Unit of the Simplified Annex 25 Air-Handling System*

not exceed its specified maximum. If the CAHU airflow maximum is reached, the total flow is shared among the three zones in proportion to their respective requirements.
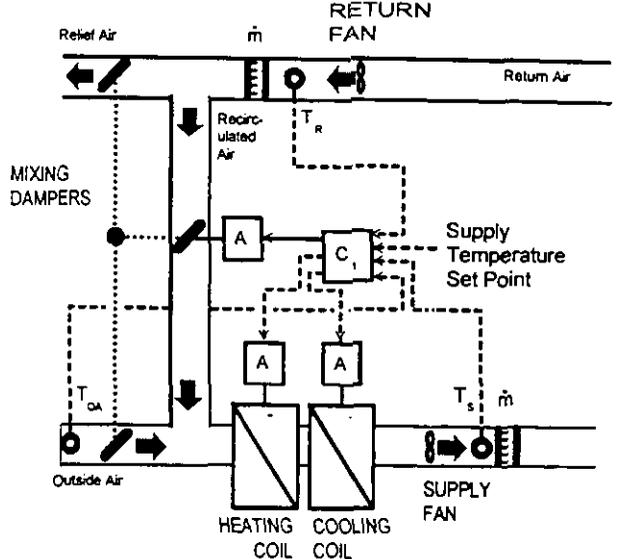
2.1.3 The central air-handling unit of the simplified Annex 25 reference system

The fault detection methods described below were developed especially for the CAHU. As mentioned above, its task is to supply air at a controlled, fixed temperature T_S by operating the preheating coil, dampers and cooling coil in sequence. The corresponding outputs of Controller C_1 are denoted U_R, U_D and U_C, respectively. The controller also includes an economy control feature, in which the control action on the dampers is reversed whenever the outside air temperature T_{OA} exceeds the return air temperature.

As mentioned, this system does not provide for *dehumidification*. If, however, the temperature control in such a system operates as described above, and the humidity control only operates while the cooling coil is in operation, then the fault detection method described below can be applied.

The overall system, including the zones, affect the CAHU indirectly. The VAV boxes, acting in response to any loads in the zones, determine overall air flow $\dot{m}$, while the return air temperature T_R results from mixing the air extracted from the zones. Thus, controller C_1 must respond to three quantities over which it has no direct control, T_{OA}, T_R and $\dot{m}$, which can be formally regarded as disturbances in this control loop.

2.2 THE TEST-SITE AIR-HANDLING SYSTEM

The overall plant differs from the Annex 25 reference air-handling system in a number of respects.

- In conformity with modern European practice, the central air-handling unit uses a *heat-recovery wheel* (abbreviated "HRW") instead of a bypass mixer.

- The central air-handling unit operates with more-or-less constant air volume.

- The air-handling plant serves to *ventilate* and *cool* the zones only. Heating is accomplished by radiators.

- The CAHU control strategy incorporates a dead zone in the control sequence between the HRW mode and the cooling mode. One set-point is

Figure 3   *The Air-Handling System at the Test Site*
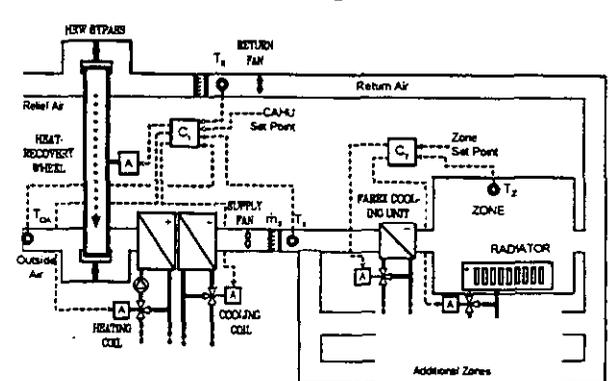
used for the supply-air temperature when the outside air must be warmed, and another set point, 4°C higher, is used when the outside air must be cooled.

* The switching to economy mode incorporates a hysteresis effect. The system switches to economy mode when the outside-air temperature exceeds the return air-temperature by 2°C and is switched off if the outside-air temperature drops below the return-air temperature. Furthermore, when the system switches to economy mode, the HRW is not subject to continuous control, but is switched on at full capacity.

2.2.1 The components of the test-site system

Figure 4 *The Central Air-Handling Unit at the Test Site*

The overall system under test is illustrated in Figure 3. It consists of a central air-handling unit (CAHU), depicted in Figure 2, supplying air to ventilate, and possibly cool a number of separately controlled zones. The CAHU comprises a heat recovery wheel and a heating coil followed by a cooling coil in a single air duct. The use of the heat recovery wheel ensures that a maximum of fresh air is supplied at all times. The air for each zones is passed through a ceiling unit which can locally cool the air to the zone as may be required. Heating is supplied by hot-water radiators in each of the zones.

In response to Swiss building regulations, which demand economy of energy use, the system is *not* equipped to dehumidify the air.

2.2.2 The control strategy of the test-site system

Two main controllers share the task of ultimately supplying air at a regulated temperature and volume to the zones. Controller C_1 regulates the CAHU and attempts to maintain the supply air temperature T_S at its set point by operating the preheating coil, heat-recovery wheel and cooling coil in sequence. Controller C_2, which regulates a single zone, attempts to maintain the zone temperature T_Z at its set point by operating the radiators or ceiling cooling coils in sequence. The HRW is equipped with bypass dampers to reduce the pressure differential in the system when it is not needed. The fans operate at two fixed levels and have the task of supplying a constant volume of air from the CAHU.

The zone temperature set point is largely prescribed by the central building energy management system (abbreviated "BEMS") in response to external conditions. This set point may vary from about 21°C, during most of the year, up to a maximum of 26°C during hot summer weather. In addition, the set points of the zone controllers can be varied manually by ±2°C from the prescribed norm, to accommodate individual needs. The nominal set point actually consists of two, slightly differing set points differing by a constant amount: a heating set point and a cooling set point. Thus there is a dead zone in the sequential control between the temperature levels requiring heating and cooling.

2.2.3 The central air-handling unit of the test-site system

The test-site CAHU differs from that of the Annex 25 reference system both in its use of a HRW and in some features of the control strategy. As mentioned above, its task is to supply air at a controlled temperature T_S by operating the preheating coil, heat-recovery wheel and cooling coil in sequence. The corresponding outputs of Controller C_1 are denoted U_H, U_W and U_C, respectively. The controller also includes an economy control feature, in which the HRW assists the cooling coil whenever T_{OA} exceeds the return air temperature.

As previously mentioned, the control strategy of the CAHU differs from that in the reference system in two main respects. The first one is that there is an effective dead zone in the control sequence. Specifically, the system uses two set-point temperatures for day-time operation: the lower set point is 16°C and the higher one is 20°C. The heating coil and HRW are operated in sequence to keep T_S from falling below 16°C, and the cooling coil is operated to keep T_S from climbing above 20°C. Whenever T_S is *between* the two set-points, *all three* components, the heating coil, the HRW and the cooling coil, remain switched off.

The second difference concerns the nature of the economy control. In the investigations preceding these building tests, the control concept simply had the dampers reverse the direction of operation whenever $T_{OA} > T_R$. As already mentioned, the test-site HRW simply switches on *at full capacity* whenever economy operation is required. To avoid situations in which the HRW might be all too frequently switched on and off, hysteresis switching is programmed in the controller. Economy mode is switched on as soon as $T_{OA} > T_R + 2°C$; it is switched off again as soon as $T_{OA} < T_R$.

The overall system, including the zones, affect the CAHU indirectly. The radiators and ceiling coolers, acting in response to loads in the zones, influence the return air temperature T_R extracted from the zones. Thus, controller C_1 must respond to two quantities over which it has no direct control, T_{OA}, and T_R, which can be formally regarded as disturbances in this control loop.

3 THE QUALITATIVE MODEL-BASED FAULT DETECTION METHOD

The detection method described below involves reducing measured controller outputs to *qualitative* values and at the same time using temperature measurements to predict expected qualitative controller outputs in steady state. This procedure necessitates reliable *steady-state detection*. Although we classify faults in terms of observed discrepancies using rule tables, such rules are derived *analytically* from the models as distinct from empirical data. Thus our fault detection method incorporates a *model-based* approach.

3.1 THE FAULT DETECTION AND DIAGNOSIS STRATEGY

The current fault-detection strategy relies on analysing the steady-state behaviour of the system, including controls. Our objective has been to identify *qualitative* modelling methods that can lead to successful fault detection and diagnosis procedures (abbreviated "FDD procedures").

Qualitative models are investigated because, even if conventional quantitative mathematical models are available for the system components, it is frequently impracticable if not impossible to obtain the values of all the relevant physical parameters of the system. Thus FDD methods based on qualitative models are particularly *robust*. Obviously, there is a trade-off in that the same methods are not able to detect all types of faults, or cannot detect certain faults in all operating states of the system. Furthermore they may not be able to *discriminate* between different types of fault. One of the goals of this work was to assess the circumstances under which it is possible to detect given faults in a CAHU using qualitative information.

The strategy adopted corresponds to the general scheme described in Fornera *et al* [1]. The overall structure of such a qualitative model-based fault detector is shown in Figure 5. This conforms to the structure of the so-called general diagnostic engine (de Kleer & Williams [13], Dexter & Glass [14]). The fault detector models we consider are also related to the generic FDD scheme proposed by Rossi and Braun [15,16].
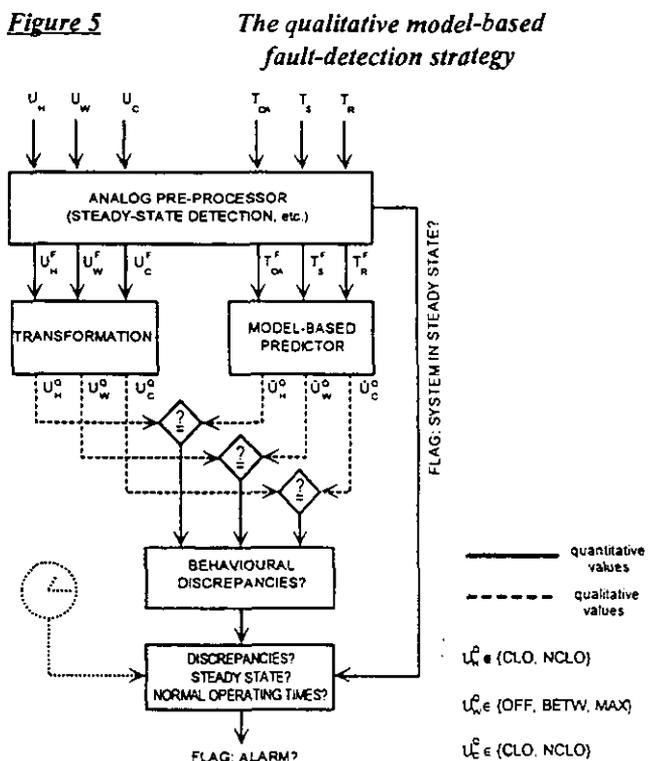
Figure 5   *The qualitative model-based fault-detection strategy*

From the central air-handling unit the measured values of the temperatures T_{OA}, T_R, T_S and the control variables[2] U_H, U_W, U_C are obtained and fed into the first stage of the analog pre-processor, which serves to:

- test whether the system is in "steady state", and

---

[2]   The outputs U_H, U_D, U_C in the case of the reference system with bypass mixer.

- optionally filter the data prior to further processing (in the analysis of building data described in this paper, moving weighted averages were used which arose naturally from the steady-state testing).

Referring to the left half of Figure 5, the filtered controller outputs U_H^F, U_W^F and U_C^F are converted to qualitative values U_H^Q, U_W^Q and U_C^Q. In the right half of the figure, the filtered temperature data T_{OA}^F, T_R^F, T_S^F is input to a model-based predictor which outputs $\hat{U}_H^Q$, $\hat{U}_W^Q$ and $\hat{U}_C^Q$, the expected (or predicted) *qualitative* controller states under steady-state conditions. The qualitative values of the outputs of the transformation and predictor blocks are chosen from the following sets:

$$\begin{cases} U_H^Q \in \{CLO, NCLO\} \\ U_W^Q \in \{OFF, BETW, MAX\} \\ U_C^Q \in \{CLO, NCLO\} \end{cases} \tag{1}$$

$$\begin{cases} \hat{U}_H^Q \in \{CLO, NCLO\} \\ \hat{U}_W^Q \in \{OFF, BETW, MAX\} \\ \hat{U}_C^Q \in \{CLO, NCLO\} \end{cases} \tag{2}$$

where the linguistic mnemonics "CLO", "NCLO", "OFF", "MAX" and "BETW" stand for "closed", "not closed", "off", "maximum" and "between", respectively.

This is a modification of the situation in the reference system, in which the qualitative value of the controller output to the dampers was assumed to take the values:

$$U_D^Q \in \{MINO, BETW, MAXO\} \tag{3}$$

where the linguistic mnemonics "MAXO", "MINO" and "BETW" stand for "maximally open", "minimally open" and "between", respectively.

Faults are detected on the basis of discrepancies between the *measured* qualitative controller outputs and the corresponding *model-based predictions* based on temperature measurements. The system signals an alarm whenever such discrepancies are detected *and* the system is deemed to be in steady state. Alarms may also be subject to further criteria, such as requiring the current time to be during hours of normal daytime operation, etc.

In point of fact, the method involves *quantitative* pre-processing of temperature data, but particular stationary states are identified in which the *qualitative* settings of the control signals ("closed" or "not closed") betray the presence of faults. This differs from the type of purely qualitative formalism in which, for instance, *all* physical quantities, such as temperatures, etc. are strictly described in terms of intervals.

3.2 TRANSITION STATES OF THE CAHU CONTROLLER

The strategy hinges on analysing the *transition states* of the sequential controller regulating the CAHU. One transition occurs when the controller switches from operating the heating coil to operating the HRW; a similar transition occurs when the controller switches from the HRW to the cooling coil. A third transition occurs when the controller switches into economy mode and reverses the direction in which the dampers are operated. These transitions correspond to the idea of *landmarks* used in some approaches to qualitative physics (Kuipers, [17,18,19]; Dexter & Glass, [14]). Landmarks are physical values of special significance. For example, freezing and boiling temperatures can serve as landmarks because phase transitions occur.

Figure 6   *Controller outputs in terms of controller state*
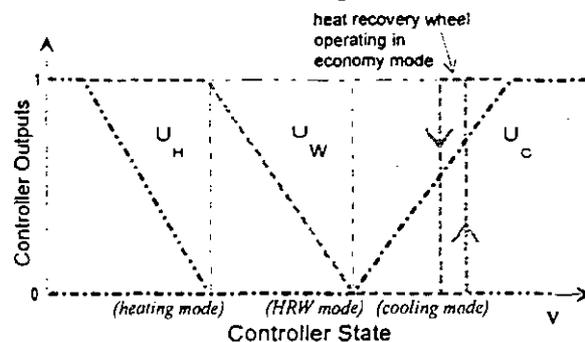
The values taken by qualitative variables can either be on one of the landmarks themselves or in an interval between two landmarks. In our case, neglecting the issue of economy control for the time being, we identify qualitative values of the CAHU controller states, depicted in Table 1. Qualitative Controller States 0 to 4

correspond to those used earlier in this series of investigations. In the context of the test site, however, it is necessary include a sixth state (Controller State No.5) to take economy-mode operation into account. In this instance it proves inappropriate to define a transition state between normal cooling and economy-mode cooling, because the HRW is switched *immediately* from off to maximum; the controller output is in fact discontinuous.

The two landmark controller transition states can be simply related to corresponding critical temperatures. Given a particular supply air temperature T_S and return air temperature T_R, critical values of the outside air temperatures, $T_{OA}^{(C1)}$ and $T_{OA}^{(C2)}$, can be calculated corresponding to Controller States Nos. 3 and 1 respectively. Outside air temperatures are then classified qualitatively, according to whether they are on or somewhere between the critical temperatures.

3.3 STEADY-STATE BEHAVIOUR

Assuming the system to be operating in steady-state conditions, it is possible to predict the outputs of the sequential controller in terms of the stationary temperatures T_{OA}, T_S and T_R. The controller sequence is illustrated in Figure 6. For moderate outside-air temperatures T_{OA} below the lower supply-air set point, the system optimizes energy consumption by operating the heat-recover wheel at the appropriate rate of rotation. For sufficiently low outside-air temperatures T_{OA} the HRW, running at its maximum capacity, must be assisted by the heating coil. For outside-air temperatures T_{OA} above the higher supply-air set point, the system operates the cooling coil.

To predict which stationary controller states correspond to which temperature states, it is helpful to consider the plot of $T_R - T_S$ versus $T_{OA} - T_S$ depicted in Figure 7. Under normal operating conditions $T_R > T_S$, so that a sequence of

Table 1   Qualitative states of the CAHU sequential controller

Qualitative Controller State	Controller outputs to...		
	Heating Coil	Heat-Recovery Wheel	Cooling Coil
0. The controller output sets the heat-recovery wheel at maximum, the cooling coil off and the *heating coil on*.	$U_H > 0$	$U_W = U_W^{(max)}$	$U_C = 0$
1. Landmark state: *transition* between heating coil operation and heat-recovery operation.	$U_H = 0$	$U_W = U_W^{(max)}$	$U_C = 0$
2. The controller output sets the heat-recovery wheel at less than maximum.	$U_H = 0$	$0 < U_W < U_W^{(max)}$	$U_C = 0$
3. Landmark state: *transition* between heat-recovery operation and cooling coil operation.	$U_H = 0$	$U_W = 0$	$U_C = 0$
4. The controller output sets the heat-recovery wheel off, the heating coil off and the *cooling coil on*.	$U_H = 0$	$U_W = 0$	$U_C > 0$
5. The controller output sets the heating coil off, the *cooling coil on* and the *heat-recovery wheel at maximum*.	$U_H = 0$	$U_W = U_W^{(max)}$	$U_C > 0$

temperature states leading to the controller outputs shown in Figure 6 will normally appear as a trajectory in the *upper half plane* in Figure 7. The transition to cooling occurs whenever $T_{OA} > T_S$, and the corresponding boundary in the diagram is the vertical axis. Thus the transition from cooling to HRW operation corresponds to passing from the first to the second quadrant in the diagram. The transition from HRW operation to heating is more complex to model for this type of plant than for a CAHU using bypass dampers, but to a good approximation the transition occurs across a straight line through the origin with negative slope as shown in Figure 7.

In the reference system, economy control is exercised whenever $T_{OA} > T_R$, which corresponds to the 45° line through the origin in the diagram. The hysteresis which is a feature of the test-site plant is represented by a second, parallel line to the right. The figure only shows the hysteresis boundary in the first quadrant, since the test-site control strategy implicitly assumes $T_R > T_S$.

The corresponding sequence that pertains to the lower half plane in the diagram has the transition between heating and HRW operation across the axis $T_{OA} = T_S$, but further discussion is omitted in this article, because

conditions at the test site in which $T_R < T_S$ during normal operation would necessarily mean that either was much hotter than it was supposed to be or the zones were much colder than they were supposed to be.

It might also be asked how the *dead zone* in the controller sequence has been represented in the diagram (Figure 7). In fact, it is not necessary to make any modification, since switching between HRW operation and cooling is governed (under steady-state conditions) by $T_{OA} - T_S$. As an initially cool outside-air temperature T_{OA} slowly rises, the operating state approaches the ($T_R - T_S$)-axis from the left. When it reaches the axis, the controller sequence enters the dead zone and the HRW is switched off. As long as neither the HRW nor the cooling coil are needed, the supply-air temperature T_S should be more-or-less equal to T_{OA}[3].

3.4 TEMPERATURE STATES CORRESPONDING TO CONTROLLER TRANSITIONS

For purposes of processing building data, it is useful to compare the outside-air temperature with critical values derived from the relevant boundaries in Figure 7. For the transition between cooling and HRW operation, the critical value $T_{OA} = T_S$ can be used, so that

$$T_{OA}^{(C1)} = T_S, \qquad (4)$$

The second, lower transition between cooling and HRW operation occurs on the boundary for which

$$\left(T_R - T_S\right) = -\eta\left(T_{OA} - T_S\right), \qquad (5)$$

where η is a positive parameter depending on the thermal transfer capacity of the HRW at maximum speed, and possibly on T_{OA}, T_S and T_R as well. Thus, in terms of this model,

$$T_{OA}^{(C2)} = -\frac{1}{\eta}T_R + \left(1 + \frac{1}{\eta}\right)T_S, \qquad (6)$$

Figure 7   Graphical representation of CAHU controller operating regimes in terms of steady-state temperature conditions (including the mnemonics for the corresponding qualitative controller states)
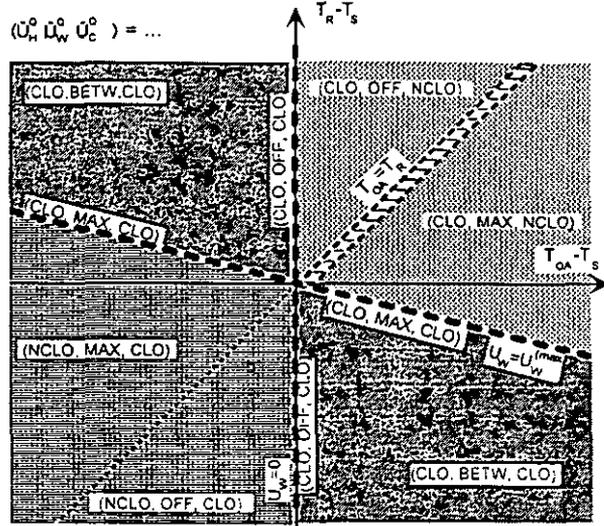
Figure 8   Graphical representation of *restricted* prediction of qualitative controller outputs in terms of *signs* of $T_{OA} - T_S$ and $T_R - T_S$

The analogous formula applying to a bypass mixer depends on χ_{min}, the minimum proportion of outside air included in the recirculated mixture.

Finally, when taking economy control, into account, the relevant critical temperature is

$$T_{OA}^{(ec)} = T_R + \varepsilon_{on/off}^{(ec)}, \qquad (7)$$

where $\varepsilon_{on/off}^{(ec)}$ is a switching threshold incorporating possible hysteresis effects. In the case of the test-site:

---

[3] There will be some slight discrepancy, since T_S does not follow rapid fluctuations in T_{OA} exactly. The response of T_S to changes T_{OA} is dampened by the heat capacities of the HRW, heating and cooling coils as well as the ducting.

$$\varepsilon_{on}^{(ec)} = 2K \; ; \qquad \varepsilon_{off}^{(ec)} = 0K \; . \qquad (8)$$

The qualitative temperature states are summarized in Table 3, where they are matched to the corresponding predicted controller states. In the case of restricted prediction, States 0, 1, and 2 must be treated as a single group.

For purposes of testing, it is difficult to calibrate the actual factor η for the HRW. Furthermore, the calibration would have to be carried out under conditions known to be *fault-free*, something which cannot be guaranteed in an unknown building under test. In this situation, it is necessary to manage without precise knowledge of $T_{OA}^{(C2)}$, which corresponds to the situation described in Fornera, et al [1,3,4], described as the *restricted* qualitative fault-detection scheme.

As illustrated in Figure 8, the predicted controller states in the second and fourth quadrants are necessarily *ambiguous*. Since in practice we typically observe temperature states in the first and second quadrants, this ambiguity becomes equivalent to being unable to discriminate in terms of the lower critical temperature $T_{OA}^{(C2)}$. We can only identify transitions in terms of $T_{OA}^{(C1)}$ and $T_{OA}^{(ec)}$. Either the system is expected to be cooling, or it is expected to exercise a *combination* of heat recovery and heating. In terms of Table 3, Qualitative Temperature States 0, 1 and 2 become indistinguishable.

Table 3   Qualitative temperature states

Qualitative Temperature State in terms of T_{OA}	Qualitative Controller State
0. $T_{OA} < T_{OA}^{(C2)}$	$U_H > 0$; $U_w = U_w^{(max)}$; $U_C = 0$
1. $T_{OA} = T_{OA}^{(C2)}$	$U_H = 0$; $U_w = U_w^{(max)}$; $U_C = 0$
2. $T_{OA}^{(C2)} < T_{OA} < T_{OA}^{(C1)}$	$U_H = 0$; $0 < U_w < U_w^{(max)}$; $U_C = 0$
3. $T_{OA} = T_{OA}^{(C1)}$	$U_H = 0$; $U_w = 0$; $U_C = 0$
4. $T_{OA}^{(C1)} < T_{OA} < T_{OA}^{(ec)} + \varepsilon_{on/off}$	$U_H = 0$; $U_w = 0$; $U_C > 0$
5. $T_{OA} \geq T_{OA}^{(ec)} + \varepsilon_{on/off}$	$U_H = 0$; $U_w = U_w^{(max)}$; $U_C > 0$

4 FAULT MODELS

Once a fault has been detected following the strategy outlined in the preceding sections, there remains the question of analysing the observed discrepancies with a view to *diagnosing* the causes. To be able to do this, it is useful to predict the *observed qualitative discrepancies* which might be expected from known faults. Consequently, it is necessary to analyse relevant fault models in more detail.

Examples of faults in a CAHU include the possible blockage of heating coil (or cooling coil) valves, so that the heating coil (or cooling coil) cannot fully shut off, or is otherwise constrained to operate within less than its full range. Another example is a temperature sensor with an offset.

Once the characteristic symptoms of particular faults are known, rules can be derived as to what qualitative discrepancies (between temperature states and controller states) will be observed under what circumstances. Whenever a fault is detected, the observed symptoms can be matched to a table of such rules to generate a list of possible *candidate faults* compatible with the observed discrepancy.

To analyse graphically what discrepancies are to be expected when faults occur, it is useful to develop an alternative representation to Figure 7. Figure 7 subdivides a two dimensional representation of the three relevant temperatures, T_{OA}, T_R, and T_S, into regions corresponding to heating, HRW operation and cooling. In the alternative representation just *one* of the temperatures is plotted directly against the controller state, symbolized by the variable v as used in Figure 6. In the following figures v is used to furnish a one-dimensional description of the (quantitative) controller state.

To obtain a one-dimensional representation of temperature, it is convenient to imagine T_R, T_S and

Figure 9   *Characteristic curve of T_{OA} vs. controller state*
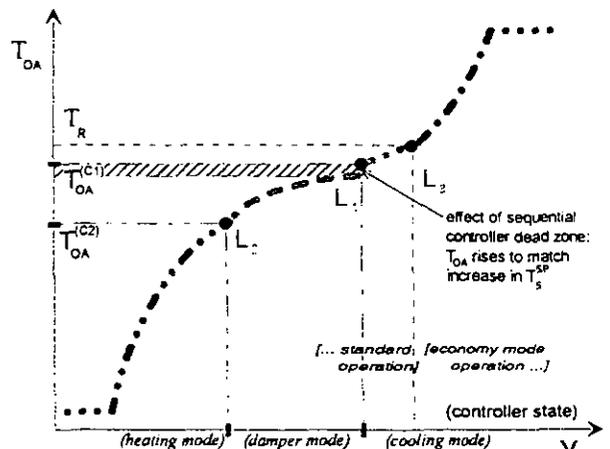
Figure 10 *Characteristic curve of* T_{OA} *vs. controller state when the cooling coil is prevented from fully closing*
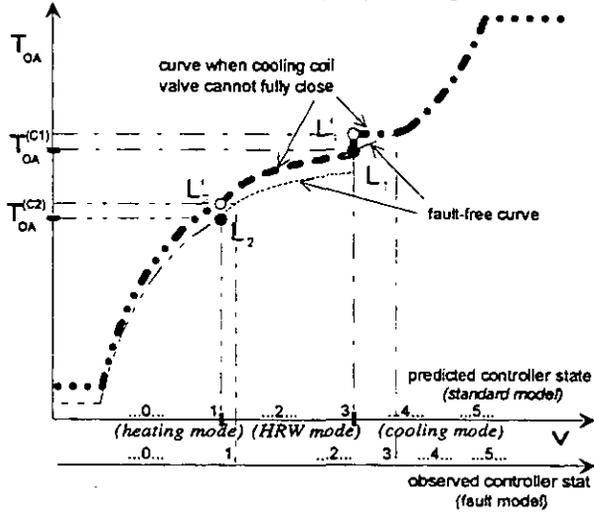
Figure 11 *Characteristic curve of* T_{OA} *vs. controller state when the measured value of the outside-air temperature is too high*
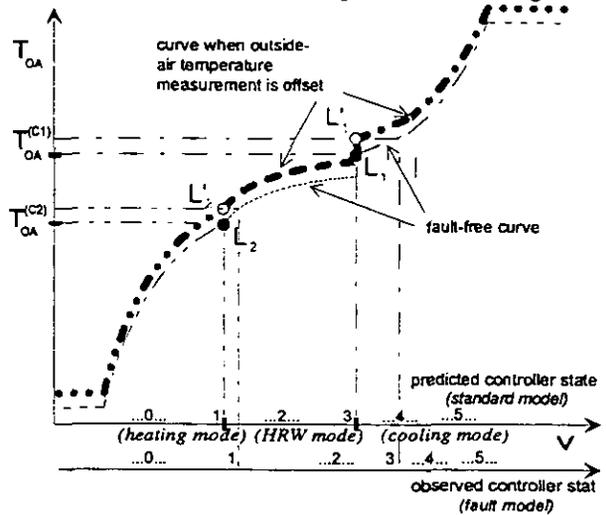
$\dot{m}$ to be *constant* and to plot the controller variable v against T_{OA}, as depicted in Figure 9. The characteristic curve shown in this diagram is based on the assumption that the temperatures, the flow and the controller outputs are in *steady state*. Furthermore, for the sake of simplicity, the hysterisis of the economy-mode switching is neglected.

The effect of the dead zone in the sequential controller program is to introduce a vertical offset at the point where the supply-air set point is allowed to rise to its upper value.

The characteristic curves shown are indeed qualitative in nature. Neither the heating nor the cooling curve can be calculated without having reasonably sophisticated *quantitative* model of the thermal effects taking place. By the same token, calculating the heat-recovery wheel curves requires knowing its thermal transfer properties for particular rotation speeds and temperature differences between outside air and return air. In general, this characteristic curve is *monotonically increasing* (between maximum heating and maximum cooling) but, in the absence of an adequate quantitative model, its slope is unknown.

In contrast, the transition points, illustrated by two dots L_1 and L_2 in the diagram, can be obtained. L_1 in particular is simply determined as $T_{OA}^{(C1)} = T_S$. L_2 is not as easy to determine since it depends on the performance of the heat-recovery wheel, but it could be calibrated during the commissioning of the plant.

The kink which appears in the cooling mode curve corresponds to the transition to the economy mode: the increased slope results from the increased efficiency of cooling when the quantity of warm outside air is reduced. This suggests introducing a *third* controller landmark L_3 corresponding to economy control switching. The corresponding temperature is $T_{OA} = T_R$, as shown in Figure 9.

When a known fault is present, it can be modelled qualitatively in terms of the above characteristic curves. An example is illustrated in Figure 10, in which the cooling coil valve cannot close completely, but otherwise operates normally. For warm outside air temperatures, when substantial cooling may be required, the characteristic curve is normal. However, for controller states to the left of the threshold at which valve blockage occurs, the same minimal amount of cooling continues to be delivered, and so the characteristic curve remains flat (at the temperature corresponding to that threshold cooling setting) until the transition to HRW operation is reached. Thereafter, remaining characteristic curve segments are the same *with respect to the modified transition point* L_1'.

To illustrate the qualitative discrepancies which may occur, two parallel axes for the controller state are included: one for the predicted behaviour, and one for the observed behaviour. The *predicted* behaviour corresponds to that illustrated in Figure 9. The six predicted qualitative states of the controller, numbered from 0 to 5, are indicated on the upper v-axis the diagram. States 1 and 3 correspond to the transitions (landmark values), State 5 corresponds to the economy mode.

Table 4 Observable qualitative discrepancies for some selected faults

	Steady-state configuration	Fault free behaviour	Symptom	Heater cannot switch off	Cooler cannot switch off	Meas-ured T_{OA} too high	Meas-ured T_{OA} too low
0	$T_{OA} < T_{OA}^{(C2)}$	$U_H > 0$;	$U_H = 0$	●			●
		$U_W = U_W^{(max)}$;	$U_W < U_W^{(max)}$	●			●
		$U_C = 0$	$U_C > 0$	●			●
1	$T_{OA} = T_{OA}^{(C2)}$	$U_H = 0$;	$U_H > 0$		●	●	
		$U_W = U_W^{(max)}$;	$U_W < U_W^{(max)}$	●			●
		$U_C = 0$	$U_C > 0$	●			●
2	$T_{OA}^{(C2)} < T_{OA}$; $T_{OA} < T_{OA}^{(C1)}$	$U_H = 0$	$U_H > 0$		●	●	
		$0 < U_W < U_W^{(max)}$	$U_W = U_W^{(max)}$		●	●	
		$U_C = 0$	$U_W = 0$	●			●
			$U_C > 0$	●			●
3	$T_{OA} = T_{OA}^{(C1)}$	$U_H = 0$	$U_H > 0$		●	●	
		$U_W = 0$;	$U_W > 0$		●	●	
		$U_C = 0$	$U_C > 0$	●			●
4	$T_{OA} > T_{OA}^{(C1)}$	$U_H = 0$	$U_H > 0$		●	●	
		$U_W = 0$;	$U_W > 0$		●	●	
		$U_C > 0$	$U_C = 0$		●	●	

The _observed_ qualitative controller states are obtained by extrapolating the _temperature_ values of the _modified landmarks_ L_1' and L_2' to the fault-free curve and determining the corresponding controller values. The corresponding values have been projected onto the lower v-axis in the diagram. Comparing the two axes shows that two temperature ranges exist for which qualitative discrepancies will be apparent.

A second example is shown in Figure 11. In this case, the measured T_{OA} exceeds the true outside-air temperature by a constant offset.

Symptoms for these two and two other, related faults are summarized in Table 4. It can be seen that for both temperatures at the landmark transition points, as well as in intervals below these values, such faults are detectable. Moreover, the presence of the dead zone at the transition between HRW operation and cooling has a distinctly beneficial side-effect. The transition state (No. 3), instead of being a brief transition, is typically maintained for significant lengths of time, thereby affording adequate opportunity to reveal any unexpected difference between T_{OA} and T_S.

Table 4 also illustrates a limitation of diagnosis based on qualitative fault detection. It can be seen that the qualitative "signature" of a heating coil valve that cannot fully close is _identical_ to that of a measured outside-air temperature that is too low. By the same token, a cooling coil valve that cannot fully close manifests identical qualitative discrepancies to those which occur when the measured value of the outside-air temperature is too high. This ambiguity must be resolved by carrying out further diagnostic analysis on the system, in order to isolate and identify the cause of the fault detected.

Table 4 does _not_ include the analysis of faults around L_3. Possible fault symptoms associated with economy control are considerably more subtle. The controller behaviour around L_1 and L_2 depends only on the measured value of T_S, together with the (implicit) _physical dynamics_ of the system. In this case, T_{OA} and T_R effectively serve to provide an _independent_ check that the system is working properly. In contrast, the transition point L_3 makes use of all three temperatures, so that the same type of check lacks the required independence. An offset

in the measured value of T_{OA}, for instance, would *not* reveal an obvious fault around L_3 because the controller is using the *same* measured value as the fault detection system.

5 STEADY-STATE DETECTION

5.1 GENERAL PRINCIPLE

The prototype steady-state detector described by Glass, *et al* [2,5,7,8], was successfully applied to the data from the test site. The essential theoretical issues are described elsewhere [12] and an alternative approach has been used by, for example, Dexter & Benouarets [20]. Basically, the steady-state detector used computes geometrically weighted averages and variances, which has two advantages:

- the computations can be carried out recursively

- the attribute "quasi-stationary" can reasonably be applied to the last measured time point; this SDD is especially sensitive to *departures* from quasi-stationarity.

To briefly review its definition, if one denotes the sequence of data by

$$\{x_0, x_1, x_2, \ldots, x_n\},\tag{9}$$

the geometrically weighted average is defined by

$$\overline{X}_n(\alpha) = \frac{\sum_{k=0}^{n} \alpha^{n-k} x_k}{\sum_{k=0}^{n} \alpha^{n-k}},\tag{10}$$

where α is the geometric weighting parameter, constrained to the range $0 < \alpha < 1$. The corresponding variance is defined by

$$S_n^2(\alpha) = \frac{\sum_{k=0}^{n} \alpha^{n-k} \left(x_k - \overline{X}_n(\alpha)\right)^2}{\sum_{k=0}^{n} \alpha^{n-k}}.\tag{11}$$

The variable x_n is deemed to be in steady state whenever the weighted *deviation* falls below a pre-determined threshold ε_{SS} [or, equivalently, the variance falls below ε_{SS}^2]

$$S_n(\alpha) \leq \varepsilon_{SS}.\tag{12}$$

The parameter α can be related to the characteristic time τ_{SS} of the response of the SSD to a step change in the input sequence. The effective time "window" is about three times this parameter.

$$\alpha = \frac{\tau_{SS}}{\tau_{SS} + \Delta T},\tag{13}$$

where ΔT is the time increment between measurements.

5.2 TUNING THE STEADY-STATE DETECTOR

Tuning the SSD requires choosing appropriate values of both the threshold ε_{SS} and the time parameter τ_{SS} in relation to the dynamics of the test object. The tuning procedure described was developed by Gruber [11], and the reader is referred to that paper for the details. Basically it proves possible to work with a limited set of plant specifications, *provided* the system is known to be properly tuned. The correct controller settings provide the necessary *implicit* information concerning the plant and building dynamics.

The rates of thermal transfer of the heating coil, heat-recovery wheel and cooling coils are needed together with the air flow. All of these parameters should be available from plant design and installation specifications. In

addition, the controller settings are required, which can be extracted from the BEMS. To apply the ideas to the test-site plant, the following equations from [11] are relevant[4]:

5.2.1 Time window for the steady-state detector

Characteristic time responses of the controlled heating coil, heat-recovery wheel and cooling coil are computed separately. Taking the *heating coil* first, its settling time $\tau_{settle,SP}^H$ in response to a change in the supply-air set point T_S^{SP} is estimated. The relevant expression is

$$\tau_{settle,SP}^H = 3\frac{\tau_n^H}{2\xi^2 K_H K_r^H},$$ (14)

where

$$K_H = \frac{\dot{Q}_H}{c\dot{m}},$$ (15)

τ_n^H and K_r^H are the PI controller parameters, $\dot{Q}_H$ is the maximum rate of heat transfer of the heating coil, $\dot{m}$ is the mass flow of air through the CAHU. In this context, ξ is a dimensionless parameter relating derived from the type of response the controlled system makes to a step change of its set point, and in a correctly tuned system it must be chosen in the range

$$0.4 \leq \xi \leq 0.7.$$ (16)

Values of ξ closer to the upper value are to be preferred.

In the case of the plant under test, the specifications are $\dot{m} = 5.23\,\text{kg/s}$, $\dot{Q}_H = 51.4\,\text{kW}$, $\tau_n^H = 120\text{s}$, and $K_r^H = 0.10\text{K}^{-1}$.

In the case of the *heat-recovery wheel*, the formulas derived for bypass dampers are applied directly. In the test-site CAHU the PI controller parameters are $\tau_n^W = 180\text{s}$ and $K_r^W = 0.13\text{K}^{-1}$. The nominal rate of heat transfer is $\dot{Q}_W = 141\,\text{kW}$ but this figure is considerably influenced by the temperature difference $T_R - T_{OA}$ (see Gruber [11]), so that for purposes of estimation we use $\dot{Q}_W^{(eff)} = 50\,\text{kW}$. The settling time $\tau_{settle,SP}^W$, which is defined a 3 times the characteristic decay time of the response to a step change of set point, is given by

$$\tau_{settle,SP}^W = 3\frac{\tau_n^W}{2\xi^2 K_W^{(eff)} K_r^W},$$ (17)

where

$$K_W^{(eff)} = \frac{\dot{Q}_W^{(eff)}}{c\dot{m}}.$$ (18)

In the *cooling coil* at the test-site, the rate of heat transfer is $\dot{Q}_C = 100\,\text{kW}$, and the PI controller parameters are $\tau_n^C = 120\text{s}$ and $K_r^C = 0.06\text{K}^{-1}$. The settling time $\tau_{settle,SP}^C$ is given by a formula slightly different from those above:

$$\tau_{settle,SP}^C = 3\frac{\tau_n^C\left(1 + K_C K_r^C\right)}{2\xi^2 K_C K_r^C},$$ (19)

where

$$K_C = \frac{\dot{Q}_C}{c\dot{m}}.$$ (20)

---

4   Some of the notation has been changed slightly from that in the original paper.

Since the steady-state detector itself has an exponential-decay response to step changes in the measured signal, τ_{SS} should be compared directly with exponential response times rather than the settling times themselves. Consequently, if the steady-state detector is to be matched as closely as possible to the response times of the plant, the following value should be used

$$\tau_{SS} = \max\left\{\frac{\tau^H_{settle,SP}}{3}, \frac{\tau^W_{settle,SP}}{3}, \frac{\tau^C_{settle,SP}}{3}\right\} = \max\left\{\frac{\tau^H_n}{2\xi^2 K_H K^H_r}, \frac{\tau^W_n}{2\xi^2 K_w K^W_r}, \frac{\tau^C_n(1+K_C K^C_r)}{2\xi^2 K_C K^C_r}\right\}, \tag{21}$$

In view of the observed behaviour of the tuned CAHU, it was deemed reasonable to take ξ close to its upper limit

$$\xi = \frac{1}{\sqrt{2}}. \tag{22}$$

Applying this value of ξ and the plant specification data to the above formulas, the resulting estimate of τ_{SS} is

$$\tau_{SS} = \max\{360s, 454s, 660s\} \approx 12\,min., \tag{23}$$

5.2.2 Threshold values for the steady-state detector

The threshold values are set in response to two considerations. One is that the minimum threshold $\epsilon^{T_s}_{SS}$ pertaining to the supply-air temperature T_S must take residual high-frequency noise in the signal into account, whose standard deviation is denoted here by σ_{T_s}. In the case of the test data, a minimum standard deviation of T_S under stable operating conditions was found to be approximately $\sigma_{T_s} = 0.1K$. Secondly, in view of the fact that the response of the supply-air temperature to step changes in its set point is approximately exponential, some further allowance has to be made for the residual effect of such a step change after the settling time $\tau_{settle,SP}$, which is approximately 5% of the original step. The residual amount is denoted by δ_{T_s}, and the resulting modification to the threshold $\epsilon^{T_s}_{SS}$ is given by

$$\epsilon^{T_s}_{SS} = \sqrt{\sigma^2_{T_s} + \delta^2_{T_s}}. \tag{24}$$

Given that supply-air set point changes of up to 5K might be expected, the resulting residual disturbance is $\delta_{T_s} = 0.25K$, and consequently $\epsilon^{T_s}_{SS}$ is estimated at

$$\epsilon^{T_s}_{SS} = 0.27K, \tag{25}$$

To apply the same idea to the outside-air temperature, the standard deviation $\sigma_{T_{OA}}$ which would lead to the observed minimum standard deviation of the supply-air temperature is estimated. The relevant formula in the context of the test-site CAHU is

$$\sigma_{T_{OA}} = 0.82\sqrt{\frac{\tau_{settle,SP}}{\Delta T}}\sigma_{T_s}, \tag{26}$$

where ΔT is the sampling time of the signal. Thus a value of $\sigma_{T_{OA}} = 0.49K$ is estimated. Given that the maximum sudden changes observed in the outside-air temperature were of the order of about 2K within the space of half an hour, the relevant residual response to a step change was estimated at $\delta_{T_{OA}} = 0.1K$. The resulting threshold is given by

$$\epsilon^{T_{OA}}_{SS} = \sqrt{\sigma^2_{T_{OA}} + \delta^2_{T_{OA}}}, \tag{27}$$

resulting in an estimate $\epsilon^{T_{OA}}_{SS} = 0.50K$.

For practical purposes, the following, rounded values were used in processing the test data:

$$\epsilon^{T_s}_{SS} = 0.3K, \qquad\qquad \epsilon^{T_{OA}}_{SS} = 0.5K, \tag{28}$$

6 THRESHOLDS FOR CLASSIFICATION OF QUALITATIVE STATES

When classifying temperatures and controller outputs in terms of qualitative states, the question of *threshold tolerances* must be addressed. For the testing of building data described below, a fairly pragmatic approach has been used, using values derived from the tuning of the steady-state detector.

Referring to the classification concepts shown in Figures 7 & 8, further relevant details pertinent to the procedure actually implemented are illustrated in Figure 12. The operating points actually encountered were all in the upper half-plane of the diagram, in which

$$T_R > T_S,\tag{29}$$

as may be reasonably expected under steady-state operating conditions. The qualitative states labelled "0", "1", "2", "3" & "4" correspond the classification in Table 3. State 0 denotes heating coil operation, State 2 heat-recovery wheel operation, State 4 cooling coil operation, and State 5 cooling coil

Figure 12   *Graphical representation of prediction scheme as applied to building data*

operation in the economy mode, in which the heat-recovery wheel provides additional cooling. State 1, on the diagonal boundary in the second quadrant, corresponds to the transition between heating and heat recovery, while State 3, on the vertical boundary, corresponds to the transition between heat recovery and cooling.

The transition to economy mode is *not* represented by a separate qualitative state. In this instance, there is a hysteresis effect between States 4 and 5, as illustrated by the region of overlap. Whenever the outside-air temperature T_{OA} is *rising*, State 4 persists to the right of the main diagonal boundary in the first quadrant until it reaches the second, parallel boundary, before switching to State 5. Conversely, whenever T_{OA} is *falling*, the transition from State 5 to State 4 is deferred until the main diagonal boundary is reached.

6.1 BOUNDARY TOLERANCES FOR THE CLASSIFICATION OF TEMPERATURE STATES

6.1.1 Boundary tolerances for the general classification scheme

At the vertical and diagonal boundaries corresponding to transitions in the CAHU controller sequence, a tolerance ε_T^Q has been introduced. The temperatures are deemed to be in State 3, whenever

$$\left|T_{OA} - T_{OA}^{(C1)}\right| = \left|T_{OA} - T_S\right| < \varepsilon_T^Q.\tag{30}$$

The temperatures are deemed to be in State 1, whenever

$$\left|T_{OA} - T_{OA}^{(C2)}\right| < \varepsilon_T^Q.\tag{31}$$

In the investigations carried out, the steady-state threshold for the supply-air temperature (derived in the preceding section) was used directly for the tolerances in both (28) and (29):

$$\varepsilon_T^Q = \varepsilon_{SS}^{T_S} = 0.3K.\tag{32}$$

6.1.2 Boundary tolerances for the restricted classification scheme

In the plant investigated, it was not possible to estimate the slope the heating/heat-recovery boundary with sufficient reliability. Consequently, the restricted classification scheme pertains as shown by the shaded areas in Figure 2. In an orthodox application of the restricted qualitative classification scheme (29) is replaced by

$$\left|T_R - T_S\right| < \varepsilon_T^Q.\tag{33}$$

In practice, however, temperatures were still subjected to a formal classification on the basis of Table 3. Discrepancies between predicted and observed values were simply ignored if both the qualitative states involved were chosen from among State 0, 1 & 2.

6.2 BOUNDARY TOLERANCES FOR THE CONTROLLER OUTPUTS

Corresponding tolerances for the controller transitions can be derived. The factors K_H, $K_w^{(eff)}$ & K_C specified above in (14), (17) & (19), respectively, describe the steady-state response of the supply-air temperature T_S to changes in the controller output to a particular CAHU component. For instance, if the cooling-coil is in operation,

$$\Delta T_S = K_C \Delta U_C = \frac{\dot{Q}_C}{c\dot{m}} \Delta U_C .$$
(34)

Thus, substituting ε_T^Q for ΔT_S and ε_U^Q for ΔU_C and solving, the appropriate tolerances for the controller are obtained:

$$\varepsilon_{U_H}^Q = \frac{\varepsilon_T^Q}{K_H} ; \qquad \varepsilon_{U_w}^Q = \frac{\varepsilon_T^Q}{K_w^{(eff)}} ; \qquad \varepsilon_{U_C}^Q = \frac{\varepsilon_T^Q}{K_C} .$$
(35)

For the building data analysed in the following section, the corresponding numerical values are

$$\varepsilon_{U_H}^Q = 0.030 ; \qquad \varepsilon_{U_w}^Q = 0.010 ; \qquad \varepsilon_{U_C}^Q = 0.015 .$$
(36)

7 TESTS ON BUILDING ENERGY-MANAGEMENT SYSTEM DATA

For testing purposes, a medium-sized commercial building occupied by the Landis & Gyr Corporation was selected in Central Switzerland. Working in collaboration with the staff responsible for operating the building energy management system, arrangements were made to monitor relevant building data on a regular basis over the course of several months.

For purposes of *qualitative fault detection*, special consideration had to be given to the *frequency* with which the data were collected, since the data must be pre-processed to identify *quasi-stationary states* of the BEMS. Normally a large selection of data points was registered every 10 minutes. For qualitative testing, however, the relevant temperatures and CAHU controller outputs were recorded *every minute* during the month of May, 1995. The plant investigated was the central air-handling unit serving the Farex ceiling coolers as described above and illustrated in Figures 3 & 4.
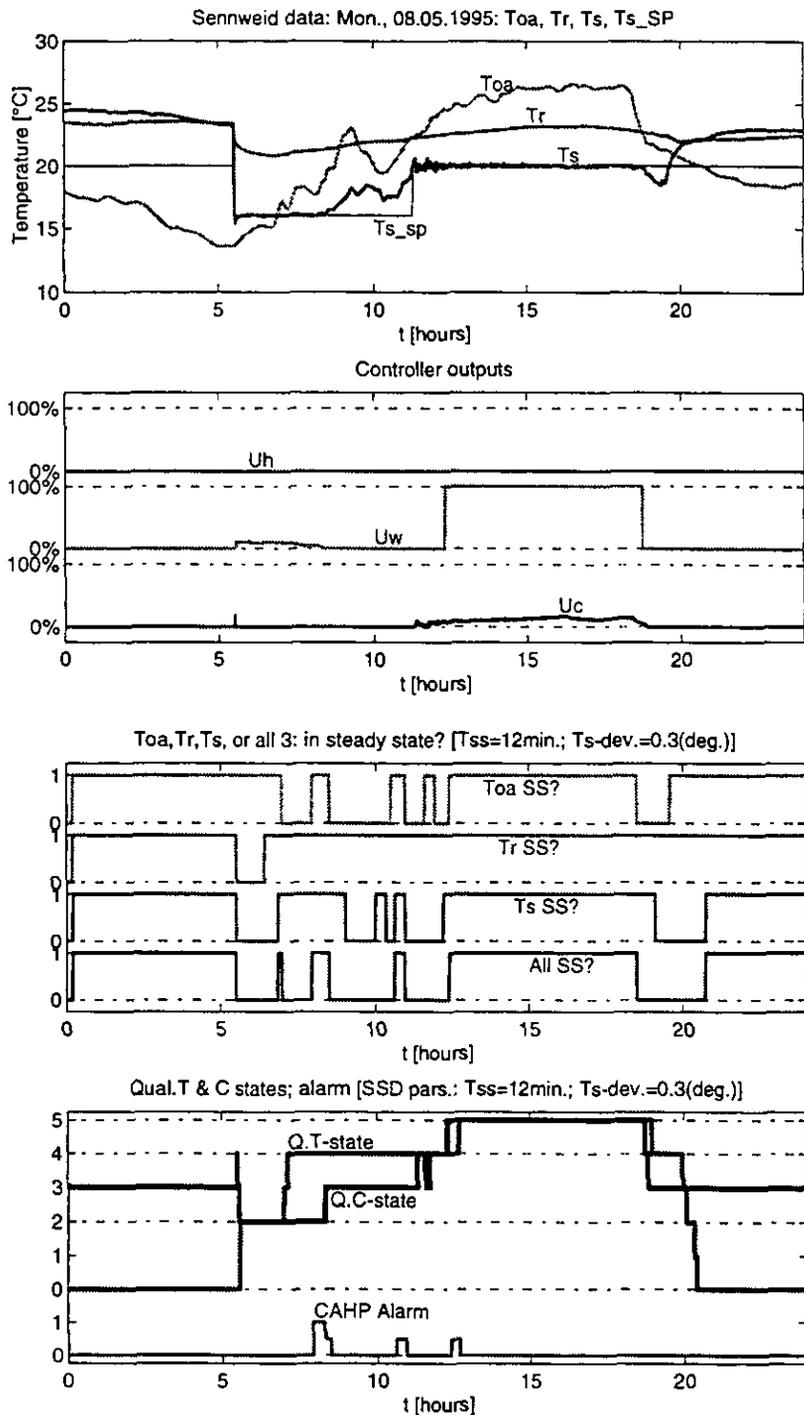
The testing programme had two goals:

- to determine approximately how often *quasi-stationary* states of the system could be expected to occur, and

- to apply the qualitative fault-detection method to the building data, but *without knowing in advance* whether faults might be expected to be present.

As described in the previous section, having reviewed the plant specifications, the steady-state time constant τ_{SS} was set at 12 min. The residual high frequency fluctuations in the supply-air temperature T_S were estimated using the building data. Giving due consideration to expected step changes in T_S and T_{OA}, thresholds of 0.3K and 0.5K respectively were used. The value $\tau_{SS} = 12$ min. was applied to T_S, T_{OA}, and T_R. The same threshold was used for both T_S and T_R.

A typical run is illustrated in Figure 13. The upper graph illustrates the outdoor, return and supply-air temperatures T_{OA}, T_R and T_S[5] for a duration of 24 hours on May 8, 1995. The system is operated between the hours of 5:30 and 19:30, during which the supply temperature is maintained between the lower and upper set-point values of 16°C and 20°C respectively. Outside these times the supply-air temperature was not controlled. The second graph illustrates the controller outputs for the same period. It may be noted that the May 8 data shows the system operating in economy-mode between about 12:00 and 19:00h.

---

[5] For technical reasons T_{OA}, T_R, T_S and T_S^{SP} are shown in the graphs as "Toa", "Tr", "Ts" and "Ts_sp" respectively.

Figure 13   *Building data test: temperatures, controller outputs, SSD & qualitative FD analysis on day when no fault evident*

Sennweid data: Mon., 08.05.1995: Toa, Tr, Ts, Ts_SP

Controller outputs

Toa,Tr,Ts, or all 3: in steady state? [Tss=12min.; Ts-dev.=0.3(deg.)]

Qual.T & C states; alarm [SSD pars.: Tss=12min.; Ts-dev.=0.3(deg.)]

Top Graph: Temperatures T_{OA}, T_R, T_S & T_S^{SP}.

2nd Graph: Controller outputs U_H, U_W & U_C.

3rd Graph: Steady-state status of T_{OA}, T_R, T_S & all 3 *simultaneously*.

Bottom Graph: Qualitative temperature & controller states; CAHU fault-detection alarm status (only activated between 5:30 and 19:30 each day).

The third graph in Figure 13 illustrates the results of filtering the three temperatures with steady-state detectors. The top curve is the binary (logical 0 or 1) output of the SSD acting on T_{OA}, the second curve the same for T_R, and the third the same for T_S. The bottom curve is the output showing when *all three* temperatures are deemed to be *simultaneously* quasi-stationary.

The bottom graph compares the qualitative temperature states with the qualitative controller states. In spite of the fact that the *restricted fault-detection* scheme was being applied, both controller states and, in particular, the temperature states have been classified in terms of the *full* discrete scale, ranging from 0 to 5. In order to classify the lower temperature states, a formal value of $\eta = 0.25$ has been used in the expression for $T_{OA}^{(C2)}$ (Equation (6)). This value is in fact based on the specifications supplied for the heat-recovery wheel. However, in the absence of careful calibration, it may not be supposed sufficiently accurate for reliable fault detection around the relevant operating points.

The bottom curve is an alarm state, taking the values 0, 0.5 or 1. An alarm was flagged, if simultaneously:

- the system was in normal operating mode (between the hours of 5:30 and 19:30),

- all three temperatures were quasi-stationary, and

- a discrepancy between the qualitative temperature state and the qualitative controller state was observed[5].

A full alarm was flagged if the qualitative temperature state differed from the qualitative controller state by more than one

---

5   A discrepancy in the sense of the *restricted* scheme requires that one of the states be at least Level 3, since States 0, 1 & 2 are lumped together (compare Figures 7 and 8).

Figure 14 *Building data test: temperatures, controller outputs, SSD & qualitative FD analysis on day when fault evident*
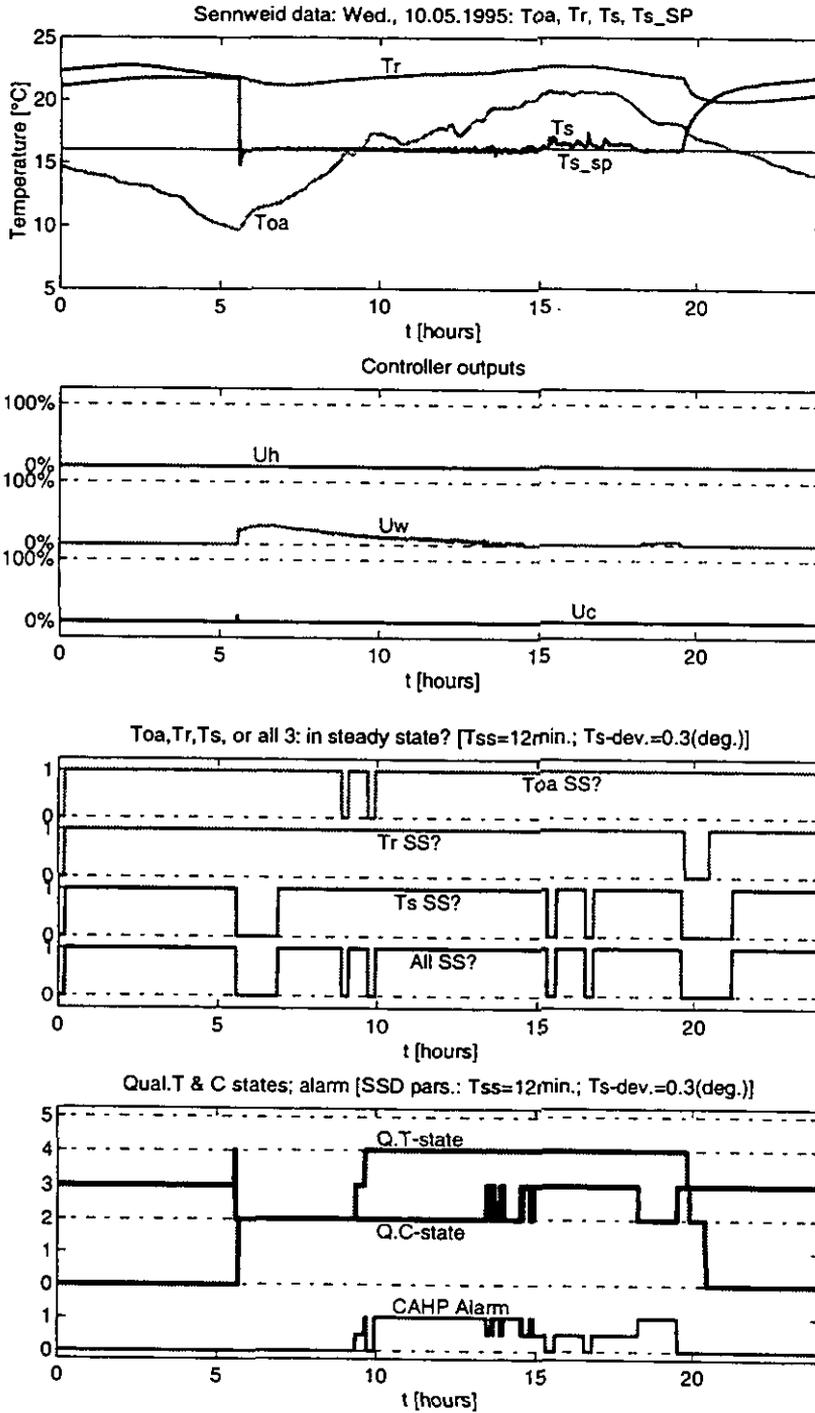
Top Graph: Temperatures T_{OA}, T_R, T_S & T_S^{SP}.

2nd Graph: Controller outputs U_H, U_W & U_C.

3rd Graph: Steady-state status of T_{OA}, T_R, T_S & all 3 *simultaneously*.

Bottom Graph: Qualitative temperature & controller states; CAHU fault-detection alarm status (only activated between 5:30 and 19:30 each day).

level, otherwise a 50% alarm was flagged. The reason for making the distinction, is that an observed discrepancy involving *neighbouring* qualitative states could conceivably result from quantities close to the adjacent borderline. It is desirable to ensure that observed discrepancies result from *quantitative* differences exceeding a reasonable mimimum tolerance, and the simplest means of ensuring this is to require that qualitative states differ by more than one level.

Looking at the results of the steady-state analysis, it may be noted that for about 50% of the operating period 5:30 to 19:30 all three temperatures are simultaneously quasi-stationary. This proportion is certainly adequate to ensure that steady-state fault detection procedures have a fair opportunity to discern possible discrepancies. However, on this particular day, although a persistent discrepancy between the qualitative temperature and controller states can be observed between about 8:00 and 11:00, the associated alarm was very brief, because the system was not sufficiently stationary. In fact during this time, the controller had entered its dead zone and the supply-air temperature drifted between its lower set point and its upper set point.

Figure 14 illustrates comparable data recorded on May 10, during which the outside-air temperature did not go high enough to cause the system to operate in economy mode. During the main operating period 5:30 to 19:30 the system was found to be quasi-stationary approximately 80% of the time. As can be seen in the top graph, the outside-air temperature remains above the supply-air temperature from about 9:00 to 19:30. In spite of that the heat-

recovery wheel continues to operate as late as about 15:00. As may be expected, the fault detection analysis shows a persistent discrepancy between the qualitative temperature and controller states. In this case, a persistent alarm results.

The observed symptom is compatible with both a cooling coil valve that cannot switch off and an outside-air temperature sensor that yields measurements above the true value. In this case, other evidence supports the diagnosis that the outside-air temperature sensor yields measured values that are too high. In the May 8 data, for instance, the dead zone between the lower and upper supply-air set points shows a difference of about 3K between the outside-air and supply-air temperatures, although they ought to be more-or-less equal.

8 GENERALIZATION AND APPLICABILITY

Classifying the examples from Fornera et al [1,3,4] in terms of transition points of the sequential controller lends this class of faults generic character, which facilitates its generalization to other systems. It has been successfully generalized to a related system in which the bypass dampers have been replaced by a heat recovery wheel.

It should be noted, however, that a feature of the system considered is that the transitions occur at the ends of the operating range of the *bypass dampers*, in which the relevant *quantitative* models are comparatively simple and reliable. In other words, the effects of adjusting the dampers is highly predictable. If one attempts to apply the same idea to the VAV boxes, for instance, one comes up against the obstacle that the loads in the zones are not predictable in practice, which makes determination of the conditions for transition more difficult.

Even within the central air-handling unit, the controller transition at the lower end of the range can only be utilized for fault detection if the system has been calibrated under fault-free conditions. The minimum prescribed proportion of outside air leaving the bypass dampers must be maintained as specified. Likewise, the performance of a heat recovery-wheel operating at maximum capacity must be known in order to utilize the lower critical temperature at which the heating coil may be expected to commence operation.

A more serious question is the applicability of methods which require steady states to be achieved which may not occur in the course of normal operation. In the plant considered, quasi-stationary states were in fact observed to occur satisfactorily often during normal operation. If this had not been the case, there would still be possibility of active testing (HVAC building tests during night-time, for instance) as opposed to simply monitoring the data.

9 CONCLUSIONS AND OUTLOOK

The fault detection strategy considered here detects faults in a central air-handling unit by analysing certain steady states of the plant (or a subsystem of the plant) in terms of *qualitative* criteria. In the overall FDD strategy, however, at least some quantitative pre-processing of data is required. In particular, the building energy-management system data must be filtered to detect quasi-stationary states. Symptoms arising from single faults can be detected under suitable temperature and controller conditions. If multiple faults are assumed, the expected qualitative symptoms are, in general, not uniquely determined. Diagnosis tables of single faults can be compiled from qualitative fault models. However, typical fault symptoms, even when observed in different temperature states, do not lead to unique diagnoses.

The steady-state detectors were tuned using the practical criteria for time-frame and threshold parameters developed by Gruber [10,11]. Analysis of data from six consecutive working days in May, 1995, in which the SSD thresholds were chosen to match a residual "noise" level of 0.3K in the supply-air temperature and 0.5K in the outside-air temperature, revealed that quasi-stationary states were maintained on average approximately 65% of the time during operational hours. A more stringent threshold of 0.1K for the supply air and 0.3K for outside air still resulted in an acceptance quota of 20% on average. Thus attaining steady state in a reasonably tuned plant does not pose a practical barrier to the application of this method.

The data collected during May proved to be suitable for testing in that the outside-air temperature was frequently in a range around the expected transition between heat-recovery-wheel operation and cooling. Moreover, the sequential controller has a dead zone in the supply-temperature set point at this transition, thereby increasing the chances of detecting faults around this operating point. Previous evidence had indicated that the outside-air temperature measurement is often a few degrees in excess of the correct value. Examination of the data indicated that during extended periods when the outside-air temperature exceeded the supply-air temperature the system,

instead of cooling, was either operating the heat-recovery wheel or was in transition between heat-recovery and cooling. Using the more conservative steady-state detection threshold, the qualitative fault detector flagged a fault on average 30% of operational time during the six working days tested.

These results adequately demonstrate that the qualitative fault-detection method can be practically and usefully applied to monitoring central plant performance in real buildings.

Possible future work includes adapting the above fault-detection method to monitor building energy-management system data on-line. This requires programming the method to operating within the BEMS software. In addition qualitative techniques described might possibly be applied to additional HVAC systems. Finally, the possibility of applying qualitative methods to the analysis of transient behaviour (cf. work by Koch [21]) or to other non-steady-state situations needs to be investigated.

Acknowledgements

References

[1] L. Fornera, A.S. Glass, P. Gruber and J. Tödtli, "Fault Detection Based on Logical Programming Applied to the VAV Reference System: First Trials", *I.E.A. Annex 25* Working Paper AN25/CH/050493/1.0, Tokyo, 14-16.4.1993.

[2] A.S. Glass, "Rule-Based Qualitative Fault Detection Applied to the VAV Reference System: Further Results", *I.E.A. Annex 25* Working Paper AN25/CH/210993/1.0, Zürich, 27-29.9.1993.

[3] L. Fornera, A.S. Glass, P. Gruber and J. Tödtli, "Qualitative Fault Detection Based on Logical Programming Applied to a VAV Air Handling Unit", *2nd IFAC Workshop on Computer Software Structures Integrating AI/KBS Systems in Process Control*, Lund, 10-12.8.1994

[4] L. Fornera, A.S. Glass, P. Gruber and J. Tödtli, "Qualitative Fault Detection Based on Logical Programming Applied to a Variable Air Volume Air-Handling", *IFAC Control Engineering Practice*, 4(1), 105-116, 1996

[5] A.S. Glass, "Detectability of faults using a qualitative method applied to a central air-handling unit", *I.E.A. Annex 25* Working Paper AN25/CH/310394/1.1, Boston, 11-13.4.1994.

[6] G. Kelly, "Air conditioning unit", §5.3 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland (VTT), Laboratory of Heating & Ventilation, P.O. Box 216, SF-02151 Espoo, Finland, 1993.

[7] A.S. Glass, P. Gruber, M. Roos & J. Tödtli, "Preliminary Evaluation of a Qualitative Model-Based Fault Detector for a Central Air-Handling Unit", Proceedings of the 3rd IEEE Conference on Control Applications, Strathclyde University, Glasgow, August 24–26, 1994, pp.1873-1882, ISBN 0-7803-1872-2, IEEE Catalogue No. 94CH3421-7.

[8] A.S. Glass, P. Gruber, M. Roos & J. Tödtli, "Qualitative Model-Based Fault Detection in Air-Handling Units", IEEE Control Systems Magazine, 15(4), 11-21, August, 1995.

[9] J. Tödtli & P. Gruber, "Stochastic model for threshold evaluation in steady state and fault detectors", Working Paper, *I.E.A. Annex 25* Working Paper AN25/CH/070494/1.0, Boston, 11-13.4.1994.

[10] P. Gruber, "Determination of the Tuning Parameters of the Steady State Detector for a Central Air Handling Unit (CAHU)", I.E.A. Annex 25 Draft Working Paper AN25/CH/230994/1.0, Stuttgart meeting, 26-28.9.1994.

[11] P. Gruber, "Determination of the Tuning Parameters of the Steady State Detector for a Central Air Handling Unit (CAHU)", I.E.A. Annex 25 Technical Report AN25/CH/230994/3.0, I.E.A. Annex 25 Final Report.

[12] A. Glass & P. Gruber, Section on "Steady-state detection", *I.E.A. Annex 25 Final Report* Source Book, 1996.

[13] J. de Kleer and B. Williams, "Diagnosing Multiple Faults", *Artificial* Intelligence, **32**(1), 97-130, 1987.

[14] A.L. Dexter and A.S. Glass, "The use of qualitative models in fault detection and diagnosis", §4.2 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland (VTT), Laboratory.of Heating & Ventilation, P.O. Box 216, SF-02151 Espoo, Finland, 1993.

[15] T. Rossi and J. Braun, "Classification of Fault Detection and Diagnostic Methods", §2 in *Building Optimisation and Fault Diagnosis System Concept*, J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland (VTT), Laboratory.of Heating & Ventilation, P.O. Box 216, SF-02151 Espoo, Finland, 1993.

[16] T. Rossi and J. Braun, §5.1, Introduction to Chapter 5, "Fault Detection & Diagnosis Methods", *I.E.A. Annex 25 Final Report* Source Book, 1996.

[17] B. Kuipers, "Commonsense reasoning about causality: Deriving behaviour from structure", *Artificial Intelligence*, **24**(1-3), 169-213, 1984

[18] B. Kuipers, "The Limits of Qualitative Simulation", *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, William Kaufman, Los Altos, USA, 1985.

[19] B. Kuipers, "Qualitative Simulation", *Artificial Intelligence*, **29**(3), 289-388, 1986.

[20] A.L. Dexter. & M. Benouarets, "Transient detection", *I.E.A. Annex Final Report*, 1996

[21] G. Koch, "Knowledge-Based Coordination of Qualitative Online Diagnostic Tests", *IFAC Symposium on Intelligent Components and Instruments for Control Applications*, Malaga, Spain, 21-21.5.1992.

A Generic Approach to Identifying Faults in HVAC Plants

Arthur L. Dexter, Ph.D., C. Eng.
Member ASHRAE

Mourad Benouarets, Ph.D.

ABSTRACT

The paper describes a semi-qualitative model-based method of fault diagnosis that is suitable for generic applications over a range of different sizes and designs of heating, ventilating, and air-conditioning (HVAC) plant items, such as terminal boxes and heating coils. The method requires no training data from the actual plant and is suitable for real-time implementation in packaged digital controllers or in the outstations of energy management and control systems. The scheme uses reference models describing faultfree and faulty operation that are generated from data produced by simulating a number of plants of the same type as the plant under test. The method of diagnosis takes account of the ambiguity introduced by using such generic reference models that can arise if the symptoms of correct and faulty operation, or of different faults, are similar at certain operating points. The results presented demonstrate that the scheme can successfully detect and identify faults in the cooling coil subsystem of an air-handling unit.

INTRODUCTION

It is difficult to obtain adequate representations of the complex, ill-defined, and often highly nonlinear behavior of a faulty plant for use in fault-diagnosis schemes (Frank 1992). Qualitative mathematical models called fuzzy models have been proposed to take account of the uncertainties and imprecision associated with describing the behavior of such systems (Sugeno and Yasukawa 1993). A fuzzy model consists of a set of IF-THEN rules that describe the essential features of the behavior of a system. A particular model is defined by the fuzzy sets that are used to describe its inputs and outputs or by a fuzzy relational matrix that indicates the extent to which each rule correctly describes the behavior of the system around a particular operating point (Yager and Filev 1994). The rules may be based on

expert knowledge or may be generated from measured data taken from the plant.

In practice, it is unlikely that data that are representative of faulty behavior can be collected from the actual plant, since it would normally be unacceptable to introduce faults in an occupied building and the physical defects that cause some faults are difficult to reproduce (for example, water-side fouling). Generic fuzzy models, which describe the underlying behavior of a class of plants of a similar design, must be used where it is impossible to obtain detailed information about or measurement data from a specific plant. Either domain knowledge or data produced by computer simulation of plants of similar design, with and without the faults, can be used to create models of this type (Benouarets et al. 1995). The development of suitable generic models involves a number of issues, such as the appropriate choice and normalization of the input and output variables, the size of the class to be represented, and the modeling accuracy demanded by the particular application. Models that are too generic must be avoided since they may not allow the fault-diagnosis scheme to distinguish between correct and faulty behavior of the plant (Dexter and Benouarets 1995).

FUZZY FAULT DIAGNOSIS

The fuzzy-model-based fault-diagnosis scheme proposed here identifies faults by comparing a partial model that describes the current behavior of the system with a set of generic reference models. One of the reference models describes the correct operation of the system and each of the other models describes the behavior of the system in the presence of a particular fault. The partial model is identified on-line from the measured data using a numerically simple fuzzy identification scheme that requires minimal processing power (Xu and Lu 1987). The identification scheme, which does not involve optimization and uses fixed reference sets, estimates values for the elements of a fuzzy relational matrix, each element of which specifies the credibility with which an associated rule correctly describes the behavior of

Arthur L. Dexter is a university lecturer and Mourad Benouarets is a research assistant in the Department of Engineering Science at the University of Oxford, Oxford, U.K.

the system. The values are in the range of 0 to 1, where a value of 0 indicates no confidence in the rule and a value of 1 indicates complete confidence in the rule. An overview of the scheme is shown in Figure 1. The data-processing unit consists of a moving-average filter (which removes any high-frequency noise) and a transient detector (which determines whether the system is sufficiently close to steady state if static fuzzy reference models are to be used) (Maruyama et al. 1995).
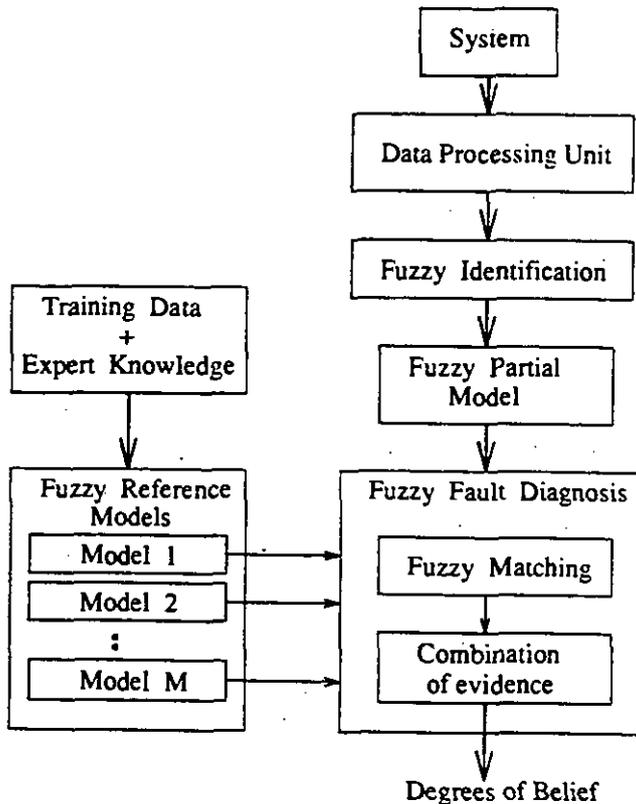

Figure 1   The fault diagnosis scheme.

The term *partial model* is used since the model only describes the behavior of the system at the current operating point and will have a large number of rules whose credibilities are zero. Since the reference models describe the behavior of the system at all possible operating points, they will have many more rules with nonzero credibility values. The rules of the partial fuzzy model are compared to the rules of the fuzzy reference models, using a fuzzy matching scheme (Dexter and Benouarets 1994). A degree of similarity, S, of the fuzzy models is calculated by considering the models themselves as fuzzy sets with discrete membership functions defined by the credibilities of the rules. For example, the degree of similarity between the partial model (R_P) and one of the reference models (R_i) is given by

$$S(R_P, R_i) = \frac{\sum_{n=1}^{N} min\{r_{R_P}(n), r_{R_i}(n)\}}{\sum_{n=1}^{N} r_{R_P}(n)}, \quad (1)$$

where $r_{R_P}(n)$ and $r_{R_i}(n)$ are the credibilities of the nth rule in the partial model and the ith reference model, respectively, and N is the number of rules in the models.

In practice, the symptoms of the behavior described by the partial model may be similar to those described by more than one of the reference models. The same approach can be used to calculate the degree of similarity of the partial model and any two or more of the reference models. For example, the degree of similarity of the partial model and two of the reference models (R_i) and (R_j) is given by

$$S(R_P, R_i, R_j) = \frac{\sum_{n=1}^{N} min[r_{R_P}(n), r_{R_i}(n), r_{R_j}(n)]}{\sum_{n=1}^{N} r_{R_P}(n)}. \quad (2)$$

The results of the diagnosis will be at least partially ambiguous whenever the generic reference models of the faultfree and faulty operation, or of different faults, exhibit similar symptoms. The ambiguity must be taken into account in the diagnosis if it is not feasible to install additional sensors to allow unique symptoms to be observed at all operating points. A measure of the level of ambiguity is obtained by calculating the similarity of the partial model and a particular reference model and all of the other reference models (Dexter 1995). For example, in the case of M reference models, the ambiguity associated with the nth rule of the partial model and the reference model R_i and the nth rule of all of the other reference models is given by

$$\lambda_{R_i}(n) = min[r_{R_P}(n), r_{R_i}(n), max_{l=1, l \neq i}^{M}\{r_{R_l}(n)\}]. \quad (3)$$

Similarly, the ambiguity associated with the nth rule of two reference models, R_i and R_j and the nth rule of all the other reference models is given by

$$\lambda_{R_i, R_j}(n)$$
$$= min[r_{R_P}(n), r_{R_i}(n), r_{R_j}(n), max_{l=1, l \neq ii \neq j}^{M}\{r_{R_l}(n)\}]. \quad (4)$$

The degrees of similarity and the levels of ambiguity are used to generate the strength of the evidence (called a normalized basic assignment) that the system is in any one of the set of possible operating states. For example, the strength of the evidence $m(\{R_i\})$ that the system is in the state described by the reference model R_i is given by

$$m(\{R_i\}) = S(R_P, R_i) - \Lambda_{R_i} \quad (5)$$

where the total ambiguity, Λ_{R_i}, associated with the reference model R_i is given by

$$\Lambda_{R_i} = \frac{\sum_{n=1}^{N} \lambda_{R_i}(n)}{\sum_{n=1}^{N} r_{R_P}(n)}. \quad (6)$$

In the same way, the degree to which the partial model is only similar to both R_i and R_j is used to estimate the strength of
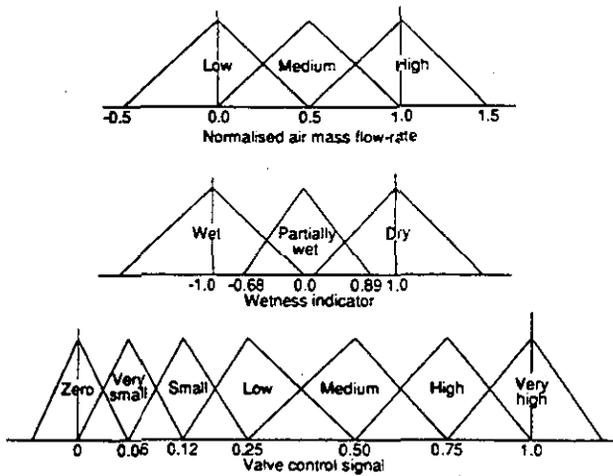
AT-96-3-4

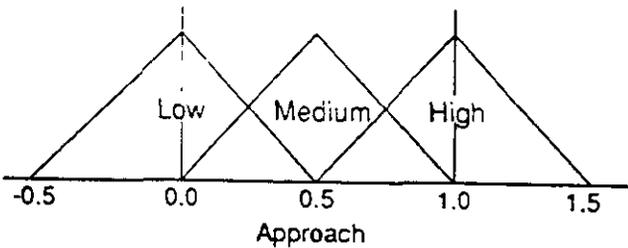**Figure 2** *Fuzzy partitions of the input space.*

Figure 3 *Fuzzy partitions of the output space.*

evidence that the system is in either the state associated with model R_i or the state associated with model R_j. Thus,

$$m(\{R_i, R_j\}) = S(R_P, R_i, R_j) - \Lambda_{R_i, R_j} \qquad (7)$$

where the total ambiguity, Λ_{R_i, R_j}, associated with the reference models R_i and R_j is given by

$$\Lambda_{R_i, R_j} = \frac{\sum_{n=1}^{N} \lambda_{R_i, R_j}(n)}{\sum_{n=1}^{N} r_{R_P}(n)}. \qquad (8)$$

The method of calculating the strengths of evidence ensures that $0 \le m(\{\}) \le 1$ and $\Sigma m(\{\}) \le 1$. A value of zero indicates that there is no evidence and a value of one indicates that there is complete evidence to support the conclusion. New values for the normalized basic assignments are computed each time the partial fuzzy model is identified.

If there is a sufficiently long interval between the current and last identification of the partial model, the normalized basic assignments will be evaluated from data collected at different operating conditions, and the Dempster rule of combination (Klir and Folger 1988) can be used to combine new evidence

(m_{new}) with evidence collected previously (m_{old}). For example, if only two reference models, R_1 and R_2, are used, then

$$m_{old, new}(\{R_1\}) = \frac{\begin{aligned}&m_{old}(\{R_1\})m_{new}(\{R_1\}) \\ &+ m_{old}(\{R_1, R_2\})m_{new}(\{R_1\}) \\ &+ m_{new}(\{R_1, R_2\})m_{old}(\{R_1\})\end{aligned}}{\begin{aligned}&1 - [m_{old}(\{R_1\})m_{new}(\{R_2\}) \\ &+ m_{old}(\{R_2\})m_{new}(\{R_1\})]\end{aligned}} \qquad (9)$$

$$m_{old, new}(\{R_2\}) = \frac{\begin{aligned}&m_{old}(\{R_2\})m_{new}(\{R_2\}) \\ &+ m_{old}(\{R_1, R_2\})m_{new}(\{R_2\}) \\ &+ m_{new}(\{R_1, R_2\})m_{old}(\{R_2\})\end{aligned}}{\begin{aligned}&1 - [m_{old}(\{R_1\})m_{new}(\{R_2\}) \\ &+ m_{old}(\{R_2\})m_{new}(\{R_1\})]\end{aligned}} \qquad (10)$$

and

$$m_{old, new}(\{R_1, R_2\}) = \frac{m_{old}(\{R_1, R_2\})m_{new}(\{R_1, R_2\})}{\begin{aligned}&1 - [m_{old}(\{R_1\})m_{new}(\{R_2\}) \\ &+ m_{old}(\{R_2\})m_{new}(\{R_1\})]\end{aligned}} \qquad (11)$$

are the strength of the evidence supporting the conclusion that the system could be in either of the states described by reference models R_1 and R_2.

The degree of belief, *Bel*, that the system is in a particular state is then calculated from the combined evidence. Thus,

$$Bel(\{R_1\}) = m_{old, new}(\{R_1\}), \qquad (12)$$

$$Bel(\{R_2\}) = m_{old, new}(\{R_2\}), \qquad (13)$$

and

$$Bel(\{R_1, R_2\}) = m_{old, new}(\{R_1\}) + m_{old, new}(\{R_2\}) + m_{old, new}(\{R_1, R_2\}). \qquad (14)$$

The degrees of belief are updated every time a new partial model is identified and new evidence is obtained. The most recent values of belief are either displayed, so that the plant operator can make the final diagnosis, or a fault is identified and an alarm set when the associated value of belief reaches a user-defined threshold.

The method of diagnosis can be easily extended to deal with simultaneous faults by including additional references models that define the behavior of the system in the presence of combinations of two or more faults at the same time. However, the ambiguity levels are likely to be much higher, and the computational complexity of combining evidence will increase significantly if a large number of reference models is used.
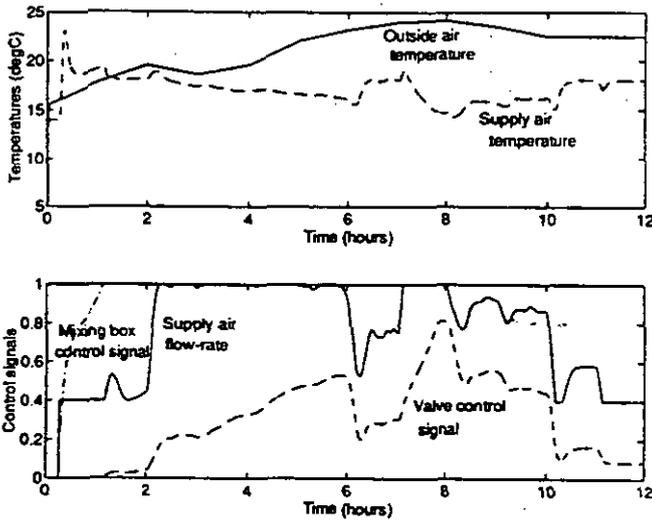
Figure 4 Test data with no faults.

APPLICATION OF THE SCHEME

Experiments have been performed using a detailed nonlinear dynamic simulation of a three-zone variable-air-volume (VAV) air-conditioning system (Haves 1994). The sizing of the plant is based on a detailed design for an office building in London, England.

Two types of degradation faults are introduced in the cooling coil subsystem: water-side fouling resulting from a buildup of scale inside the tubes of the coil and valve leakage caused by wear to the plug or seat of the valve controlling the flow of water through the coil. Both are faults that are difficult to detect since (a) they do not cause sudden changes in the observed behavior but result in a gradual deterioration of plant performance over a long period, (b) the behavior of the system in the presence of these faults is similar to that for correct operation at many operating points, and (c) there is always some doubt about the amount of fouling and leakage that should be regarded as a fault.

The generic reference models are identified off-line using input-output training data obtained from computer simulations of three different designs of cooling coil subsystems. Table 1 shows the main design parameters for the three designs and for the cooling coil subsystem in the VAV system used for the tests. Designs 1 and 2 have similar duties and face areas but operate at different chilled-water supply temperatures. Design 3 operates at an airflow rate that is roughly double that of designs 1 and 2. The reference models represent the behavior of the cooling coil when it is operating correctly {C} and with the fouling {F} and leakage {L} faults. The models defining the symptoms of faulty behavior are generated using data collected from computer simulations of the plant with 1 mm of fouling and 3% leakage, respectively.

Each model is a qualitative description of the steady-state relationship between the inputs and output. All models have the same structure, and design parameters are used to normalize the input and output variables. The inputs are the valve control signal, u_c; the normalized air mass flow rate,

TABLE 1 Design Data for the Cooling Coil Subsystems

Design Parameters	Design			
	1	2	3	Test Plant
Coil duty (Kw)	74.3	63.2	146.6	63.5
Airflow Kg/s	4.3	4.3	8.6	6.0
Supply temp. (°C)	7	10	7	7
Water flow (Kg/s)	6.3	7.4	10.0	6.3
No. of rows	6	6	7	6
Height of coil (m)	1.44	1.5	2.0	1.44
Width of coil (m)	1.36	1.36	1.50	1.36
No. of circuits	45	50	75	45
Valve capacity (m³/h)	31.0	36.4	49.2	31.0
Coil resistance (0.001Kg·m)	0.30	0.23	0.12	0.522

$(\dot{m} - \dot{m}_{min})/(\dot{m}_{max} - \dot{m}_{min})$; and a variable, $W = (T_{ao} - T_{dew})/\Delta$, that indicates whether the coil is operating dry or wet. T_{dew} is the dew-point temperature, and Δ is a user-selected parameter whose value (+5°C) reflects the uncertainties associated with this simple method of estimating the wetness of the coil. The fuzzy sets describing each of the input variables are shown in Figure 2. The output of the models is the air-side approach, $\alpha = (T_{ai} - T_{ao})/(T_{ai} - T_{cws})$, where T_{ai} and T_{ao} are the temperatures of the air entering and leaving the cooling coil and T_{cws} is the design value of the chilled-water supply temperature. The fuzzy sets used to describe the air-side approach are shown in Figure 3.

Each fuzzy model includes 189 possible rules of the following form:

IF the coil is WET
    (or PARTIALLY WET or DRY)
AND the valve control signal is ZERO
    (or VERY SMALL or SMALL or LOW or MEDIUM
    or HIGH or VERY HIGH)
AND the normalized supply airflow rate is LOW
    (or MEDIUM or HIGH)
THEN the air-side approach is LOW
    (or MEDIUM or HIGH).

The training data are obtained by simulating the open-loop behavior of each design of coil subsystem while varying the boundary conditions so as to follow staircase waveforms covering the entire operating space. To ensure that the training data fit the structure of the model, the magnitude of the steps is chosen so that the values of the normalized input variables coincide with the values at the apex of the fuzzy reference sets, where this is possible. The steady-state value of the temperature of the air leaving the coil is recorded for all combinations of values of the boundary conditions.

Five sets of test data are collected every 15 seconds throughout the occupancy period on a typical summer day, during which the system is subjected to a wide range of operating conditions

4

and is operating under closed-loop control. Data sets are collected when:

- the system is operating correctly,
- there is a 0.2-mm buildup of scale inside the tubes of the coil,
- there is a 1-mm buildup of scale inside the tubes of the coil,
- there is 1% leakage through the cooling coil valve, and
- there is 3% leakage through the cooling coil valve.

The measurements used for the diagnosis are restricted to those normally available from sensors connected to the energy management and control system (EMCS). Since the measurement of mixed-air temperature can be unreliable and a sensor is often not installed, the outside air temperature measurement is used as a proxy for the temperature of the air entering the coil whenever the fresh air dampers are fully open. The damper control signal can be used to determine when the dampers are fully open if it is assumed that the mixing box is operating correctly. A crude estimate of the temperature of the air leaving the coil is obtained from the supply-air temperature measurement by assuming a constant temperature difference across the supply fan. Figure 4 shows the raw test data obtained when the cooling coil subsystem is operating correctly. The raw test data after 0.2 mm and 1 mm of scale has built up on the inside of the tubes and with 1% and 3% leakage through the valve of the cooling coil are shown in Figures 5 through 8. The raw data are filtered (using a moving-average filter with a rectangular window of 2.5 minutes) to remove short-term changes caused by unmeasured disturbances or measurement noise and then preprocessed to remove measurements obtained when the fresh air dampers are not fully open. A transient detector is then used to identify those sections of the data when the system is considered to be close enough to steady state for the data to be used to

identify a static model (Maruyama et al. 1995). The system is assumed to be sufficiently close to steady state whenever the average activity of all of the signals (within a 10-minute window) is less than a threshold value that is based on the maximum acceptable prediction error of the static model (in this case, 5%). The average values of the measured signals over the previous 10 minutes are then used to identify the fuzzy partial model.

Table 2 shows the maximum degrees of belief obtained during each of the five tests. The variations of the degrees of belief when there has been 0.2-mm and 1-mm buildups of scale inside the tubes of the cooling coil are shown in Figures 9 and 10, respectively.
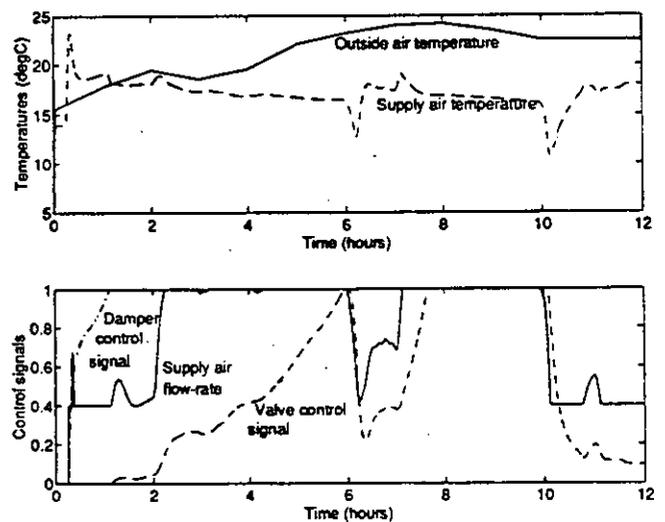

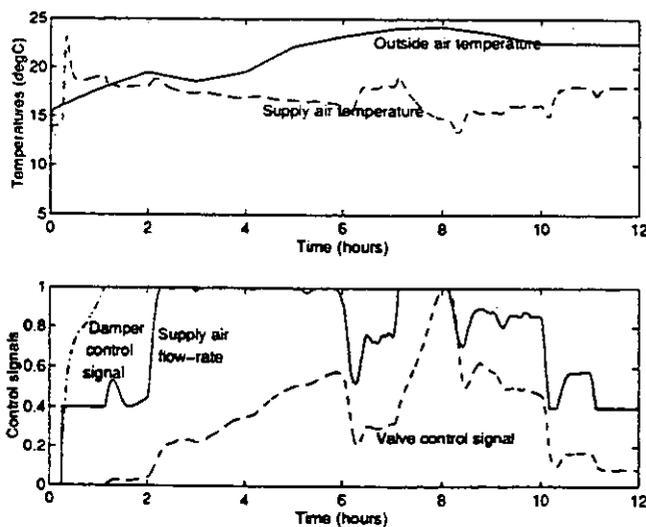**Figure 6** *Test data with 1 mm buildup of scale inside the tubes of the cooling coil.*

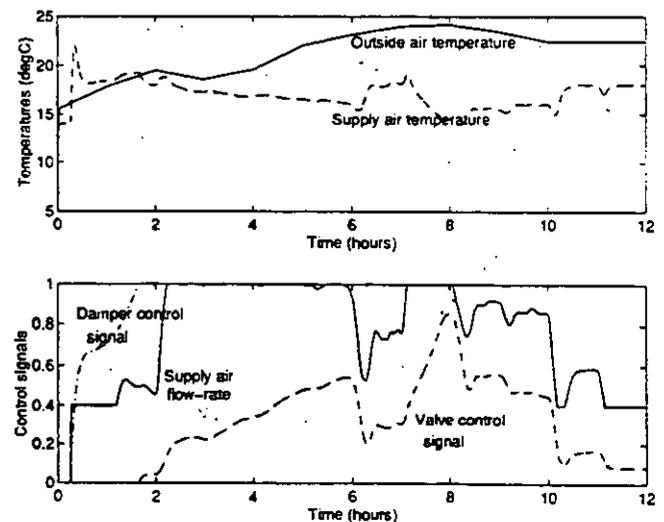**Figure 5** *Test data with 0.2 mm buildup of scale inside the tubes of the cooling coil.*

Figure 7 *Test data with 1% leakage through the cooling coil valve.*

TABLE 2   Maximum Degrees of Belief (%)

Test Condition	{C} Correct	{F} Fouled	{L} Leaky	{C, F} Correct or Fouled	{C, L} Correct or Leaky	{F, L} Fouled or Leaky
Correct operation	5.3	4.8	1.3	0.0	43.4	8.0
0.2 mm buildup of scale	0.3	8.8	4.0	0.0	34.0	22.1
1 mm buildup of scale	0.0	81.7	0.0	0.0	0.0	99.8
1% leakage	5.8	5.8	2.0	0.0	50.9	9.8
3% leakage	3.3	0.1	61.0	0.0	83.2	61.5

Figure 8   *Test data with 3% leakage through the cooling coil valve.*
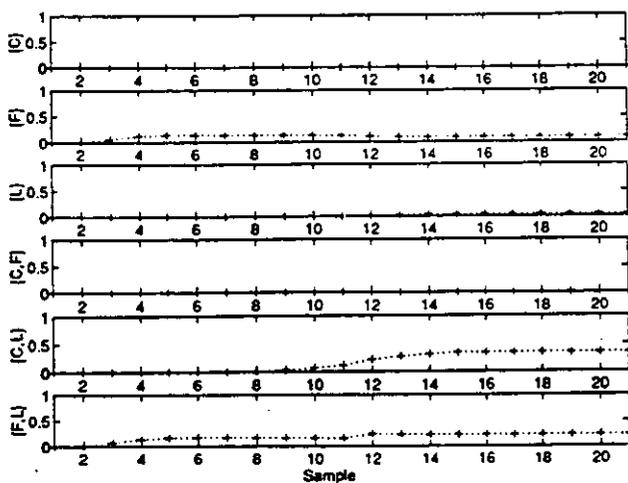
Figure 9   *Variations of the degrees of belief (0.2mm fouling).*

DISCUSSION OF RESULTS

The results presented in Table 2 clearly demonstrate the ambiguous nature of the diagnosis made in all five tests. When the system is operating correctly, the unambiguous belief in correct operation is small since most of the usable test data are obtained at high and medium cooling loads, when the symptoms of correct and leaky operation are similar. However, the diagnosis does generate a reasonably high belief in either correct or leaky operation {C, L} and a low belief in all of the other operating states.

The correct diagnosis is made when faults are introduced that have the same magnitude as those used in the generation of the reference models. As can be seen in Figure 10, the diagnosis generates a high, unambiguous belief in fouling when there is a 1-mm buildup of scale, and there is a significant belief in leakage when there is 3% leakage through the valve. Since symptoms are similar at certain operating points and there is some mismatch between the behavior of the actual plant and that described by the generic models, ambiguous results are also produced. However, it should be noted that, for example, when there is a 1-mm buildup of scale, the belief in fouled or leaky operation {F, L} of 99.8% includes the unambiguous 81.7% belief in fouling (the strength of evidence that the system is fouled is 0.817, that it is leaky is 0.000, and that it is either fouled or leaky is 0.181).
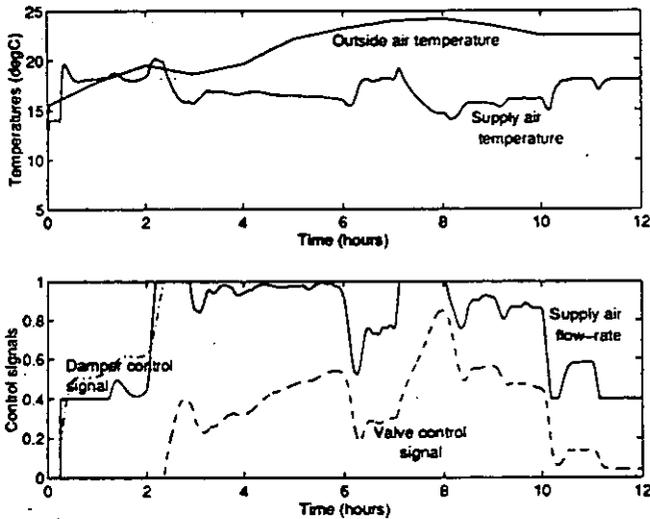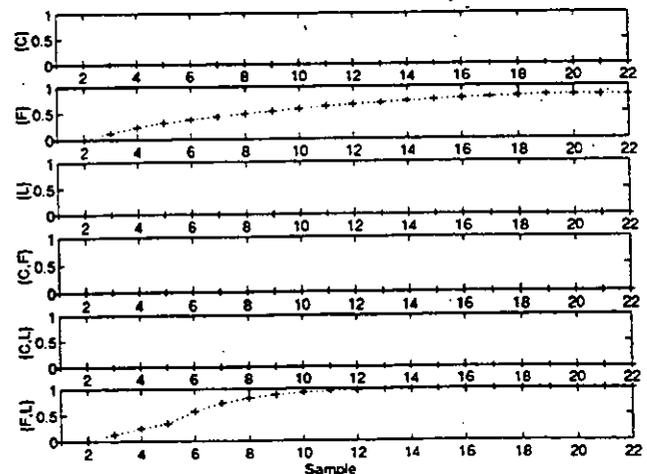

Figure 10   *Variations of the degrees of belief (1 mm fouling).*

Even for small amounts of fouling or leaking, the problem is still identified correctly to some degree, although the diagnosis is less convincing when the size of the fault is smaller than that used in the reference models. When there is a 0.2-mm buildup of scale, the mismatch between the behavior of the actual plant and that described by the generic reference models is clearly greater than the difference between fouled and correct, or fouled and leaky, operation. As can be seen in Figure 9, some evidence of both correct or leaky operation {*C, L*} and fouled or leaky operation {*F, L*} is generated, resulting in a significant level of belief in both these states. When there is 1% leakage, the unambiguous belief in leaky operation is less than the unambiguous belief in correct or fouled operation. The only significant belief is in correct or leaky operation {*C, L*} since there are few usable test data at the lower cooling loads, where the symptoms of correct and leaky operation are different.

CONCLUSIONS

The feasibility of fault diagnosis based on semi qualitative generic reference models has been demonstrated. Results have been presented that show that a fuzzy-model-based fault diagnosis scheme, which would require no training on the actual plant, can successfully identify faults in a simulated HVAC subsystem if the size of the faults is sufficiently large. By using fuzzy matching, the scheme is able to account for the ambiguity that arises from fault-free and faulty operation, or different faults having similar symptoms at some operating points, and reduce the occurrence of false alarms. The proposed fuzzy method of fault diagnosis is computationally simple enough for it to be used on more complex subsystems such as air-handling units, it is insensitive to measurement uncertainties, and it is suitable for real-time implementation in EMCS outstations or packaged digital controllers.

ACKNOWLEDGMENTS

REFERENCES

Benouarets, M., A.L. Dexter, R.S. Fargus, P. Haves, T.I. Salsbury, and J.A. Wright. 1995. Model-based approaches to fault detection and diagnosis in air-conditioning systems. *Proc. 4th Int. Conf. on System Simulation in Buildings*, Liege, Belgium, pp. 529-547.

Dexter, A.L. 1995. Fuzzy model-based fault diagnosis using fuzzy matching. *Proc. IEE, Pt. D, Control Theory and Applications* 142 (6): 545-550.

Dexter, A.L., and M. Benouarets. 1994. Fuzzy model-based fault diagnosis using fuzzy matching. Oxford University Report No. OUEL 2032/94. Oxford, U.K.: Oxford University.

Dexter, A.L., and M. Benouarets. 1995. Generic modeling of HVAC plant for fault diagnosis. Proc. 4th IBPSA Int. Conference, Madison, Wisc., pp. 339-345.

Frank, P.M. 1992. Principles of model-based fault detection. *Proc. IFAC/IFIP/IMACS Int. Symp. on AI in Real-Time Control*, Delft, the Netherlands, pp. 363-370.

Haves, P. 1994. Component based modeling of VAV systems. *Proc. 4th Int. Conf. on System Simulation in Buildings*, Liege, Belgium, pp. 31-53.

Klir, G.J., and T.A. Folger. 1988. *Fuzzy sets, uncertainty and information.* Englewood Cliffs, N.J.: Prentice Hall.

Maruyama, N., M. Benouarets, and A.L. Dexter. 1995. Detecting faults in nonlinear dynamic systems using static neuro-fuzzy models. IEE Colloquium on Qualitative and Quantitative Modeling Methods for Fault Diagnosis, York, U.K., Digest No. 1995/079, 8/1-8/10.

Sugeno, M., and T. Yasukawa. 1993. A fuzzy-logic approach to qualitative modeling. *IEEE Transactions of Fuzzy Systems* 1(1): 7-31.

Xu, C.-W., and Y.-Z. Lu. 1987. Fuzzy model identification and self-learning for dynamic systems. *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-17, no. 4, pp. 683-689.

Yager, R.R., and D.P. Filev. 1994. *Essentials of fuzzy modeling and control.* New York: John Wiley and Sons, Inc.

ON-LINE DIAGNOSTIC TESTS APPLIED TO FAULT DETECTION AND ISOLATION OF AN AIR-HANDLING UNIT

Jouko Pakanen
Technical Research Centre of Finland (VTT)
Building Technology
Oulu, Finland

Abstract

On-line diagnostic testing is one choice, when practical fault detection and isolation methods are considered for automated processes. Performing a test means exciting a process by means of prescribed input signals, supervising the responses and comparing results with a process model. An on-line diagnostic test is repeated similarly every time, in similar process conditions, making modeling an uncomplicated task. Fault detection is a direct consequence of the comparison, but fault isolation is based on elementary constraints, decomposed from the process model. A rough description of a fault can be achieved by heuristic reasoning, which enables application of the method in practice. A more specified fault description is accomplished by learning from old solutions. The reasoner accumulates information making decisions of the classifier gradually more precise through acquired experience. The method is best for successive installations, in which knowledge can be cumulated. On-line diagnostic tests are generic in character, but in this paper they are configured for an air handling unit of an office building.

1 INTRODUCTION

If technical processes and systems of buildings are subjected to effective and robust fault detection and isolation (FDI) methods, several benefits will be obtained. These include energy and water savings, reduced maintenance costs, lower safety and health risks and increased quality of living. Today, these benefits are only partly attained and more efforts in research & development of diagnostic methods and systems are needed.

Technical prerequisites for constructing diagnostic systems already exist, but there are several obstacles. One is the poor robustness of FDI methods. Besides robustness other aspects of HVAC environment also need to be considered when the objective is practical implementation of FDI. The method should be generic, i.e., easy to apply to different processes, without disturbing occupants and ordinary operation of the process. The method should also be capable of being integrated and embedded in building automation systems. In addition to the above, the FDI must meet the requirements of the occupants, and the service and maintenance personnel of the building. The last aspect contains several issues concerning a user interface, configuration of new applications, updating diagnostic data, and supporting software tools for diagnostic operations [1], [2], [3]. These kinds of basic issues are seldom presented in FDI publications, although they definitely have an influence on FDI method development.

Air handling of a building is a demanding process for fault detection and isolation. Temperature control of the supply air may consist of several cascaded control circuits. The non-linear character of subprocess make many of the known FDI methods difficult to apply. Pragmatic aspects, like minimum extra instrumentation, also set limits on available methods and tools. This paper presents a FDI method for an air handling unit (AHU) of an office building, as one possible way of tackling these problems. The method is called the on-line diagnostic test (ODT). The subject has been earlier handled by Koch [4], but he focuses on the knowledge-level system description. This paper presents a detailed procedure for fault detection and isolation using an ODT.

On-line diagnostic tests were examined in a real air handling unit. Field tests concentrated largely on the preheating process of the AHU, but the results are equally generalized to other subprocesses. Those tests are not presented here. But this paper presents conclusions and configuration of the diagnostic test methodology, which followed the experiments.

2 CONCEPT OF THE DIAGNOSTIC TEST

An on-line diagnostic test (ODT) is a series of control and monitoring actions applied to a process to try to reveal possible faults of the process. Performing an on-line diagnostic test means exciting an automated process

by means of proper input signals, disturbances or loads, causing dynamic changes in the output of the process. In this paper only input signal excitation is considered. If abnormal responses are generated, the process is faulty (Figure 1). An ODT is focused on one process at a time. When the entire process consists of several subprocesses, faults are better isolated by testing each subprocess separately.
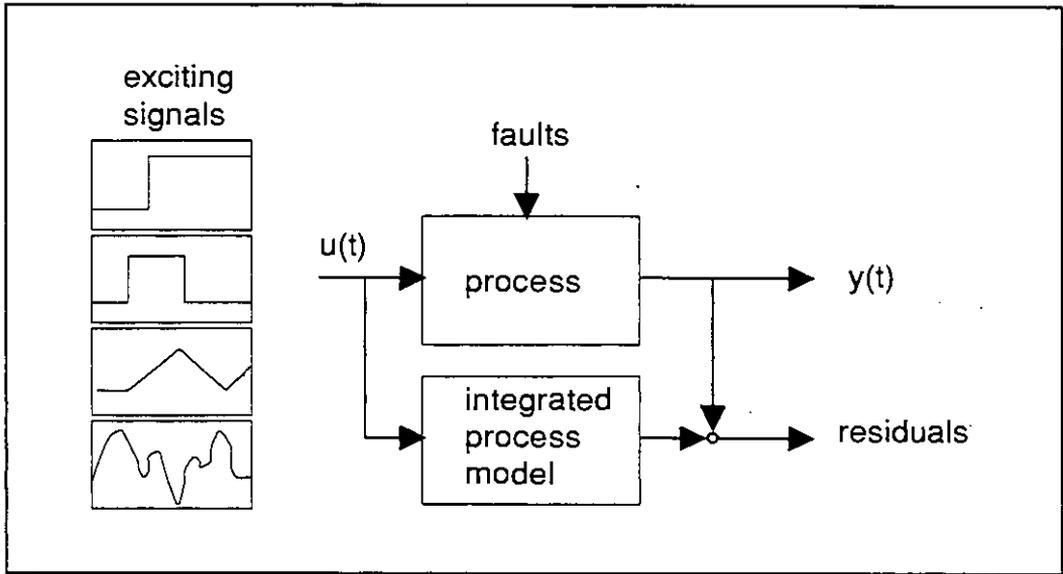

Figure 1. Principle of fault diagnosis using on-line diagnostic tests.

The ability of the diagnostic test to detect and locate faults is based on a comparison between the behaviour of faulty and normal processes. An ODT is repeated similarly every time, in similar process conditions. The approach is simple and robust because process models are necessary only for responses to specific input signals, making modeling an uncomplicated task. On the other hand, one must utilize all the information generated by the test in order to ensure fault detectability and isolability of the ODT. Thus, on-line diagnostic tests share mainly model-based diagnosis methods but may also utilize other FDI approaches.

One diagnostic test comprises of three different parts: identification, fault detection and fault isolation (Figure 2). Data gathered during the identification period represents characteristic operation of the process in normal status. Later, when the process may be faulty, this data is compared with data collected during the fault detection period. Fault isolation is a result of reasoning based on comparison of the data of current and earlier periods. One test may refer to several possible faults. By combining the results of earlier tests and the tests of other subprocesses, redundant fault alternatives are excluded.

Diagnostic tests are performed on-line, during a normal operating sequence of the process and controlled by an automation system. However, the execution time of the test are selected so that disturbances, load or environmental conditions are similar to those of earlier tests. Achievement of the right process conditions is taken care of by an automation system, which has good control over the process. An ODT can be scheduled to start at a prescribed time or it can be initiated by an alarm or by a user.
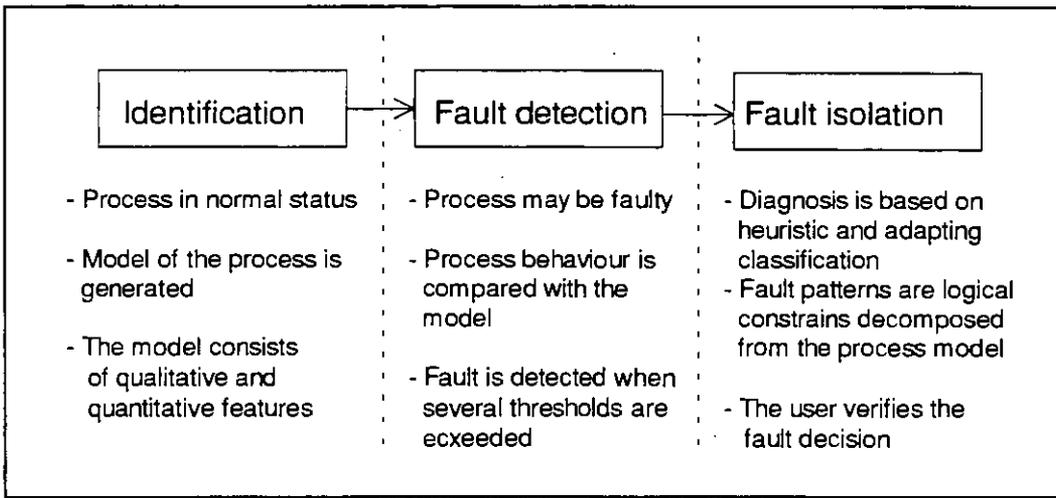
Figure 2. Decomposition of on-line diagnostic tests into identification, fault detection and fault isolation.

3 DESCRIPTION OF THE AIR HANDLING UNIT

Methodology of the ODT was examined in a real air handling process. Figure 3 represents a simplified schematic of the AHU. The system consists of a heat recovery unit, mixing dampers, and preheating, humidifying, cooling and heating processes. The heat recovery unit, dampers, heating and cooling need continuous control signals $(z_r, z_d, z_p, z_b, z_c)$, but the humidifier is controlled by an on-off signal (z_m). The dampers in the figure are connected to a single control signal. An alternative would be a separate control signal for each damper. Supply and return fans can be driven at two different speeds, controlled by signal z_f. In addition, there are temperature measurements of the outdoor air (u_a), mixed air (u_i), supply air (u_s), return air (u_r), leaving water of the preheating coil (u_p) and the humidity of the return air (u_m). Usually, the set point temperature of the zone is maintained using a cascade control algorithm, but during the tests each subprocess is controlled separately. A preferred time for an ODT is a non-occupation period of the building, because changes in internal heat sources are minimal. Disturbances caused by outdoor conditions are made negligible for most subprocesses by setting the dampers in recirculating position.

4 FAULT DETECTION

4.1 CHOSEN FAULT DETECTION AND ISOLATION METHOD

On-line diagnostic tests share model-based diagnostic methods but they may also utilize other characteristic information of the process. While modelling the process behaviour, one could resort to traditional quantitative methods like parameter estimation, state estimation or parity equations of the process [5], [6], [7], [8], [9].
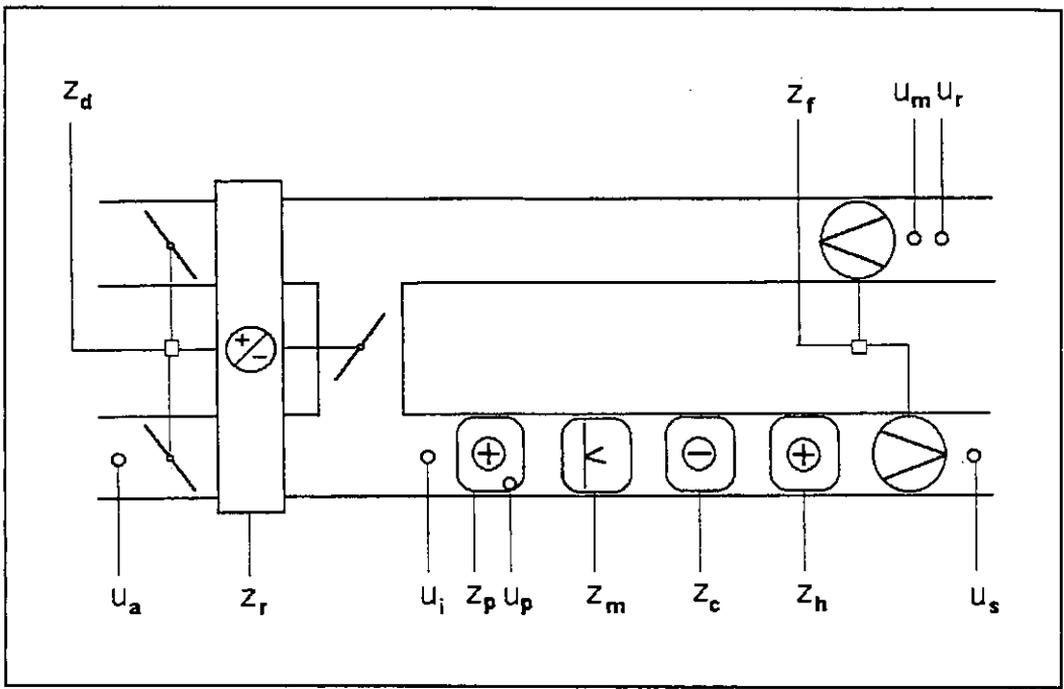
Figure 3. Simplified schematic of the air handling unit, control signals and measurements.

Methods must be selected carefully, because some lead to complicated procedures and therefore to robustness problems. Signal analysis is one choice [10], and features of this will be used. Many successful diagnostic tools are expert systems, which apply several kinds of information extracted from the process or given by the user. Although a knowledge based approach is partly included, the following pages do not concentrate on knowledge-level problem solving. In addition to these, during the recent years, qualitative modelling has become a real alternative for describing system behaviour [11], [12], [13], [14]. Many of these principles are applied in this paper.

4.2 QUALITATIVE APPROACH

Qualitative modelling and simulation is not always a straightforward technique. There are a number of individualized concepts and representations, without any common language [14], which makes a comparison of methods and results complicated. Another difficulty is that qualitative differential equations with boundary conditions may give extraneous solutions [15], which cause spurious behaviour [16].

The FDI method, presented in this paper enables integration of qualitative and quantitative knowledge in the same process model. The approach also eliminates the possibility of extraneous solutions. The basic difference between the original qualitative approach and the method of this paper is that the operating procedure of the diagnostic test is always the same and the tests are always excited by the same, known signals, producing the same envisionment every time, without any abnormal transitions. Neither is there a need to solve qualitative differential equations, but only to learn the quantity spaces, directions and landmarks of the qualitative variables.

A qualitative model of an HVAC process is created by examining the behaviour of exciting and response signals of one ODT (Figure 4). Each signal j represents a function f_j, which has a finite set of distinguished time-points:

$$A_j = \{t_0, t_1, t_2, \cdots, t_m \mid t_0 = a, t_m = b, t_i < t_{i+1}\} \qquad (1)$$

and a corresponding, finite set of landmark values:

$$B_j = \{l_1, l_2, \cdots, l_k \mid l_i < l_{i+1}\} . \qquad (2)$$

which represent some characteristic points of the signals. The landmarks are points, where

$$f_j(a) = l_1, \quad f_j(b) = l_k \qquad (3)$$

or

$$f_j(t_n) = l_i, \quad \forall f_j{}'(t_n) = 0, \ t_n \in A_j; \ l_i \in B_j . \qquad (4)$$

Landmarks defined by equation 4 are maximum and minimum points of the signal. When a signal gets its landmark point, the corresponding points of all other signals reach certain values specified as intervals between two adjacent landmarks. Either a landmark or an interval is described as the qualitative magnitude of the variable. In addition to the interval, the direction of change of the signal can also be identified. These definitions allow the setting of logical constraints, which describe signal and process dynamics. Table 1 presents some examples of such constraints. Comparison of the model and the real process behaviour, gives symptoms for fault detection.

4.3 QUANTITATIVE APPROACH

The quantitative part of the model benefits all available measured signals of process input and output. For one signal, j, this means the sampled values:

$$G_j = \{u_i = u(t_i) \mid t_i = a + ih, i = 0, \ldots, n; \ a + nh = b\}, \qquad (5)$$

where h denotes sampling time, and a and b are starting and ending time points. Repeating an ODT a number of times reveals that measured points u_i can be described by a stochastic variable $\underline{u}_i$. All the variables $\underline{u}_i$, $i \in (a,b)$ are not utilized in the model. Most important are the points, that correspond to the landmarks of the qualitative model. Due to the stochastic character of the process, the actual landmark points of repeated ODTs do not occur at same time instants. Thus, the time points of these landmarks $\underline{t}_i$ also are stochastic variables. For one signal, j, all the previous points are defined as
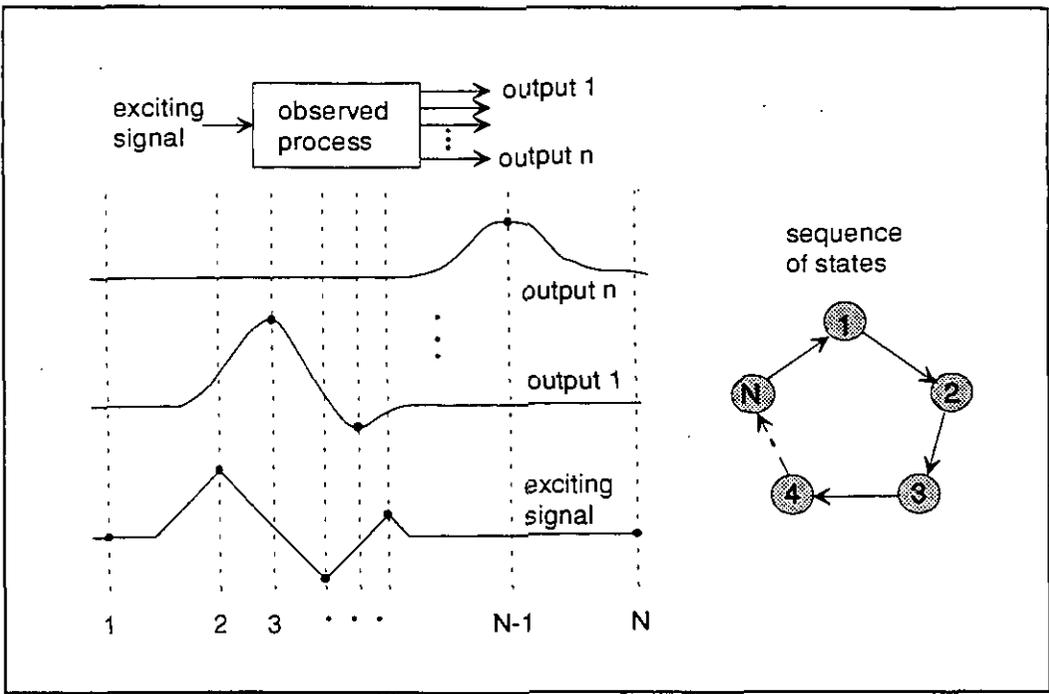
Figure 4. Qualitative states of exciting and response signals.

$$H_j = \{\underline{u}_i, \underline{t}_i \mid \underline{u}_i \sim \phi_{ui}(u_i), \underline{t}_i \sim \phi_{ti}(t_i), u_i \in \overline{B}_j, t_i \in \overline{A}_j, t_i \neq a, b\}, \qquad (6)$$

where Φ_{ui} and Φ_{ti} refer to a probability density functions. Notation $\overline{B}_j$ refers to the landmarks corresponding to the set B_j, but its elements are chosen from the set G_j. A similar definition concerns the set $\overline{A}_j$. It is assumed that arbitrary sample points $\underline{u}_i, \underline{t}_i$ i=0,1,2,.... are normally distributed stochastic variables, $\underline{u}_i \sim N(\mu_u, \sigma_u^2)$ and $\underline{t}_i \sim N(\mu_t, \sigma_t^2)$, where unknown parameters μ and σ are approximated from the data.

The measured signals, their discrete sampling points and the defined stochastic variables enable the setting of many kinds of logical constraints concerning the behaviour of the process. The constraints need not depend only on the points of one signal but they can connect several input and output signals in one constraint. By combining the qualitative and quantitative constraints, S_i, an integrated model M is produced:

$$M = \{S_i, i = 1, ..., n, S_i = \{x \mid x \in A_j, B_j, G_j, H_j, j = 1, ..., m\}\}, \qquad (7)$$

where the letter m equals the number of signals. So, the model is a collection of constrains which tie up qualitative time-points A_j, the corresponding landmarks B_j, quantitative sampled values of the signals G_j and the points defined by the probability density functions in H_j. The model M enables the prediction of the process dynamics during an ODT. A fault is detected by comparing predicted and real behaviour of the process.

The above method provides some benefits compared to the purely qualitative approach. The integrated model allows checking gains, delays and skewnesses of the responses, which may help to identify specific faults. All

the essential information of the model can be presented in compact, numeric and symbolic form. This is a benefit if the method is applied to larger entities.

Table 1. Examples of quantitative and qualitative constraints.

Constraint	Description
$\|\bar{u}(t_i) - u(t_i)\| < 1.96\,\sigma_{ui}, \; \forall u(t_i) \in \overline{B}_j, t_i \neq a,b$	Signal values at peak landmark points must be within specified limits
$\|\bar{t}_i - t_i\| < 1.96\,\sigma_{ti}, \; \forall t_i \in \overline{A}_j, t_i \neq a,b$	Time-points of peak landmark points must be within timing window
$\sum_i \|\bar{u}_i - u_i\| < 1.96\,\sigma_s, \; \forall u_i \in G_j$	Sum of signal values during control phase must be within limits
$\|[\bar{t}_i - \bar{t}_k] - [t_i - t_k]\| < 1.96\,\sigma_{di}, i \neq k; \; i,k \in \overline{A}_j;$ $t_i, t_k \neq a,b$	Time delays between peak landmark values must be within limits
$\{l_1, l_2, \cdots, l_k \mid k = Constant\}$	Number of landmarks for one signal is constant
$l_i < l_j, \forall i,j = (\alpha, \beta, \cdots, \omega), \; i,j \in [1,k]$	Mutual magnitude of landmarks for one signal must not change
$u(t_i) > u(t_{i-1}), \; \forall t_i \in C, \; C \text{ is given}$	Signal must increase monotonically
$u(t_i) < u(t_{i-1}), \; \forall t_i \in D, \; D \text{ is given}$	Signal must decrease monotonically

4.4 ADDITIONAL FEATURES

Any additional qualitative or quantitative feature suitable for the test procedure and giving new symptoms of faults can be included in the model above. The new details depend largely on the applied test procedure, monitored and exciting signals. For instance, if the quality of temperature control is examined by a step change of the set value, the proper characteristic quantities might be settling time, percent overshoots, peak response, peak time and/or some performance index, like the integral of the square of the error (ISE). Some other test procedures could check the operating time or full opening time of the equipment, the difference from a mean

value of a physical quantity or difference from a summed mean value of a signal, delays, dynamic or static limits of monitored variables, signal models etc.

Before diagnosis, the original set of constraints is decomposed into *elementary constraints*, which allow creation of characteristic fault patterns, i.e., structured residuals [8]. This is a basis for fault isolation. An elementary constraint may tie up one, several or all points of each signal, combine points from several signals, utilize timing and functional relationships and quantitative and qualitative information.

5. FAULT ISOLATION

5.1 HEURISTIC CLASSIFICATION

In detecting a fault, a diagnostic test generates a selection of symptoms characteristic of the fault. One choice for diagnosis is to match the symptoms and their potential faults with each other. Such a classification is characterized as heuristic, because relations are uncertain, based on empirical, a-priori knowledge of the system behaviour. It is typical of heuristic classification to skip intermediate relations between symptoms and faults. These relations may be unobservable or poorly understood [17], [18]. A general inference structure of heuristic classification is illustrated in figure 5. For many problems, solutions are not inferred directly from data but through data abstractions. The general structure also includes refinement of the solutions after the heuristic match.
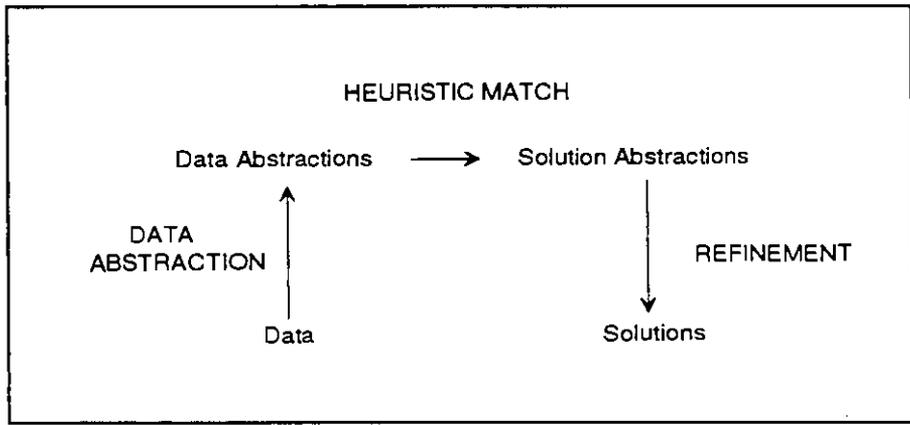
Figure 5. Inference structure of heuristic classification [17].

When a diagnostic test is performed, symptoms can be directly related to inferred faults without data abstractions. This is a domain-specific, non-hierarchical association between symptoms and faults. Results of the heuristic match are actually solution abstractions, and refinement is performed otherwise. This is due to incomplete knowledge, allowing only a rough interpretation of the fault to be made. But the straightforward

relationship between the faults and their symptoms also means an uncomplicated inference structure of the classification.

Symptoms resulting from the chosen diagnostic tests and referring to a fault should be easily extracted by a-priori knowledge of the system behaviour. This is successful if one concentrates only on abrupt and significant changes in process behaviour. Thus, an experienced engineer can create a classification relation between test results and faults while designing the system. Heuristic classification serves as a first aid reasoning method, giving a usable description of the fault before the system has learned enough to make more accurate decisions.

Feature vectors, extracted by diagnostic tests, definitely contain fault specific information, allowing more precise determination of actual faults than the heuristic classification. However, such knowledge is hardly available a-priori. Precise categorizing of the faults would require empirical knowledge of the observed process. This is not yet available at the time of BEM installation. So, heuristic classification makes the fault isolation, not perfect, but applicable, right after BEM installation and commissioning of the HVAC system.

5.2 ADAPTING CLASSIFICATION

If diagnostic tests generate relevant feature vectors, fault selectivity can be improved. The idea is based on accumulating knowledge, which consists of real, verified faults and their feature vectors. Using this knowledge, data from a new diagnostic test is classified and the most potential fault is chosen. If the name of a fault suggested by the system turns out to be erroneous, the user can change it and verify the results as a new solution. Collecting data of old solutions and using them as examples is a common learning paradigm. The learning procedure accumulates information in a case library, making decisions of the classifier gradually more precise through acquired experience. Figure 6 presents the principle of the whole isolation method.
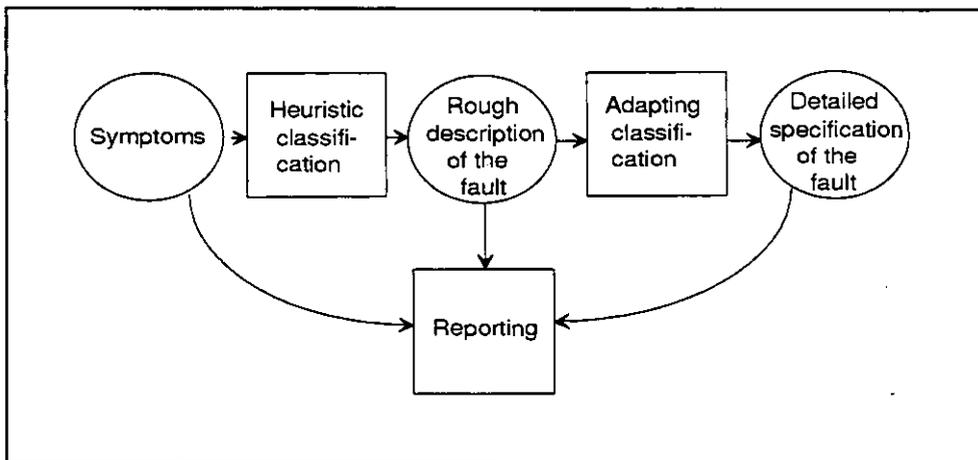

Figure 6. Principle of the applied reasoning method.

The fault diagnosis of the ODT resembles a pattern recognition system. Sensors transduce measurements of the process and logical constraints resulting from the process model, extract specific features from the data. Finally, a classifier evaluates the features and makes a decision. However, the learning principle is not similar to supervised learning, typical of pattern recognition, where the learning data is previously assigned to classes. In this case, a new class can always be created if the verified fault does not exist in the case library. Thus, the approach described above has elements familiar to case-based reasoning, which adapts old solutions to new problems or applies old cases to evaluate or justify new situations [19].

The learning ability is one way to increase the robustness of the fault isolation. Although an erroneous suggestion of a fault is made once, the probability of making a correct decision regarding the fault increases if a similar fault is encountered again. It is incorrect to say that all problems of fault diagnosis are solved by classification. But the learning ability makes the method flexible and enlarges its conventional areas of application.

5.3 PARALLEL OPERATION OF CLASSIFIERS

In a real BEM, heuristic and adapting classification can operate in parallel using the same knowledge base and retrieval technique. The results of both methods are presented to the user (Figure 6). The heuristic classification suggests a few possible faults or presents a rough explanation of the fault. Similarly, the adapting classification may display several faults, highlight the most probable one or place faults in a ranking list. Classification and retrieval obey the rules of discrimination or indexing. When the knowledge base is increasing, faults brought up by the classifier become more accurate. Parallel operation of both methods makes the system structure uncomplicated and clear. Both methods are necessary and support the total reasoning concept.

The reasoning concept is justified by the diagnostic test procedure, which concentrates on one subprocess at a time. Because the fault is already isolated in the subprocess, a few fault alternatives given by the heuristic classifier are enough to discriminate some typical fault cases. The resulting fault descriptions, although rough or unspecified, conduct the user or serviceman sufficiently close to the real fault. The learning principle makes the system more accurate. A similar concept applied simultaneously to the whole AHU would be more problematic.

The knowledge of the classifiers can be transferred to another BEM-system. Transferring means that only faults and their logical constraints are relocated. Numerical parameters, such as probability limits and other characteristic data are defined separately in the new system. One choice of transferring is to copy the rules of the heuristic classifier and to change their logical constraints according to the experience gained from the earlier application. Another choice is to relocate the whole case library to the new system and utilize the old system directly as a reasoner. However, both approaches can still be considered as heuristic, because the old system always differs from the new one and represents a-priori knowledge. After each fault case, the information received from the old database is checked and corrected in a new database. Gradually, the new data base replaces the old one. The transfer becomes more successful, if the structure, operation and instrumentation of

both the old and new AHU are sufficiently close to each other and the situation is the best if the AHUs are of same type, made by one manufacturer.

5.4 DISCRIMINATION PRINCIPLE

Discrimination refers to the process of deriving classification rules using samples of known fault cases. Usually, this approach leads to multivariate statistical representations and requires numerical fault data. If features are described both numerically and symbolically, and diagnosis problems are not fuzzy, non-statistical techniques may be appropriate for organizing fault cases. These methods are related to case-based reasoning and referred to indexing methods [20]. However, the indexing methods can be regarded as simplified versions of discrimination based on multivariate statistical representation. The following equations illustrate one approach of discrimination techniques and give an insight to the problem. Actual implementation will be a reduced version of the presentation.

Verified faults and their data represent known classes. Every new test data is compared with them using the following techniques. Let $\Omega = \{\omega_1, \omega_2, ..., \omega_c\}$ be a finite set of classes. A classifier finds the best category ω_i, $i = 1,2,3,...,c$ for the data, i.e., the most probable fault, by computing a discriminant function $g_i(x)$, where x is a vector-valued random variable. The classifier assigns the feature vector x to class ω_i if $g_i(x) > g_j(x)$, with all $j \neq i$. If the decision of the fault turns out later to be inaccurate, and no such fault exists, a new class ω_{c+1} is created and added to the knowledge base. On the other hand, if discrimination and decision is successful, but the feature vectors do not match, both of them are assigned to the same fault name.

It is a known fact that the maximum value of the discrimination function $g_i(x)$ corresponds the minimum conditional risk [21], which is achieved by replacing $g_i(x)$ with $P(\omega_i|x)$. Here $P(\omega_i|x)$ represents a posteriori probability, describing the probability of ω_i, in case x is already occurred. $P(\omega_i|x)$ can be computed by Bayes Rule from the a-priori probability $P(\omega_i)$ and from the probability $P(x|\omega_i)$ for x, conditioned on ω_i being the class of fault.

5.5 DISCRIMINATION FUNCTION

Because components of each feature vector consist of binary and discrete values, some non-parametric discrimination must be applied. Due to the various sources of data, the components of the feature vectors are also unnormalized, complicating their comparison. That is why the following discrimination techniques is more closely discussed.

All the components of the feature vectors are handled as binary-valued. Each component is assigned either a true (T) or a not true (N) value, later also referred to values 1 and 0. This is accomplished by presenting the behaviour of the integrated model as elementary, logical constraints, reducible to binary form. This concerns both quantitative and qualitative knowledge. While transforming to binary format, some information is lost. This loss

of information can be partly replaced by increasing the number of logical constraints. On the other hand, binary data is easy to handle and this format makes classification a compact procedure.

To be more specific, the probability p_{ij} of each feature component $x_i \ni x$, conditioned on class $\omega_j \ni \Omega$, is defined as

$$p_{ij} = Prob(x_i = 1 \,|\, \omega_j), \quad i = 1,2,\cdots,d \quad j = 1,2,\cdots,c. \tag{8}$$

In order to illustrate the relationships of the $g_j(x)$, it is assumed that the components of the feature vector are conditionally independent. Then $P(x|\omega_j)$ can be written as

$$P(x\,|\,\omega_j) = \prod_{i=1}^{d} p_{ij}{}^{x_i}(1 - p_{ij})^{1-x_i}, \tag{9}$$

which results directly from rules of probability [21]. Using Bayes Rule the discriminant function can be written as

$$g_j(x) = P(x\,|\,\omega_j)P(\omega_j), \tag{10}$$

which gets its final form by inserting $P(x|\omega_j)$ from (9)

$$g_j(x) = \sum_{i=1}^{d} x_i \log \frac{p_{ij}}{1 - p_{ij}} + \sum_{i=1}^{d} \log(1 - p_{ij}) + \log P(\omega_j). \tag{11}$$

Although it is an approximation of the real situation, the equation 11 describes how the discrimination function consists of a linear function of the feature components x_i and a constant. The latter one consist of the probabilities p_{ij} and the a-priori probability $P(\omega_i)$. The criteria $g_i(x) > g_j(x)$, with all $j \neq i$ determine the best choice for the fault. In practice decision of the fault class ω_j must be made using inaccurate values for the probabilities p_{ij} and without knowing any values of the a-priori probability $P(\omega_i)$. Thus, the role of the weighted combination of the feature components x_i becomes dominant while choosing the right class.

5.6 BINARY FORMAT FAULT PATTERNS

Table 2 presents a case, where the test results of a preheating process are converted into binary format. The column on the left describes the elementary constraints typically met when a fault is detected. The equations are decomposed from the integrated model. The other columns indicate, in binary format, if this requirement is satisfied by the specified fault. The collection of all logical states forms a pattern, that is typical for the fault. The number of the logical constraints is somewhat arbitrary, but correlates the number of faults. Increasing the number of fault cases would allow addition of more constraints, but the robustness of each new constraint must also be considered.

Table 2. Fault patterns of the process in binary format. The columns in the left presents elementary constraints decomposed from the process model. Other columns illustrate how these requirements are satisfied in each fault case.

Elementary constraints	Fault N:o 1	Fault N:o 2	Fault N:o 3	...	Fault N:o n
a	T	T	N	...	N
b	N	N	T	...	T/N
c	N	T	T	...	N
d	N	T/N	N	...	T
⋮	⋮	⋮	⋮		⋮
z	T/N	N	N	...	N

Most of the columns on the right are denoted as T or N and correspond to high $(p_{ij} \cong 1)$ or low $(p_{ij} \cong 0)$ probabilities of p_{ij}. Some elements, denoted as T/N, indicate that both choices may be possible and thus their p_{ij} may be closer to 0.5. This means that the corresponding feature component is more or less indifferent. The table presents a case, where each fault is verified and classified only once. If they were detected many times, there would be a distribution of probabilities p_{ij}. It is obvious that part of the detected faults, primarily abrupt and distinct changes in the process, satisfy some of the logical constraints in a similar way almost every time, leaving their probabilities p_{ij} close to one or close to zero.

5.7 CLASSIFICATION OF FAULTS

Heuristic classification means that part of the information of table 2 is introduced as a-priori data. The designer of the system can set some constraints typical of one fault or a group of faults. They are particular features, clearly seen in response to the test and repeated similarly in every occurrence of the fault. For a great part of the features, such a conclusion is not possible and they are regarded as indifferent states, denoted as T/N. Thus, the diagnostic test procedure, concentrating on one subprocess at a time and on abrupt and distinct faults, enables partial submission of the information needed in the probability $P(x|\omega_i)$ of equation 10, which is essential for fault isolation. The a-priori information, given by designer, does not allow the observed fault to be named exactly, but does make it possible to categorize a group of possible faults or suggest, which part of the process is faulty.
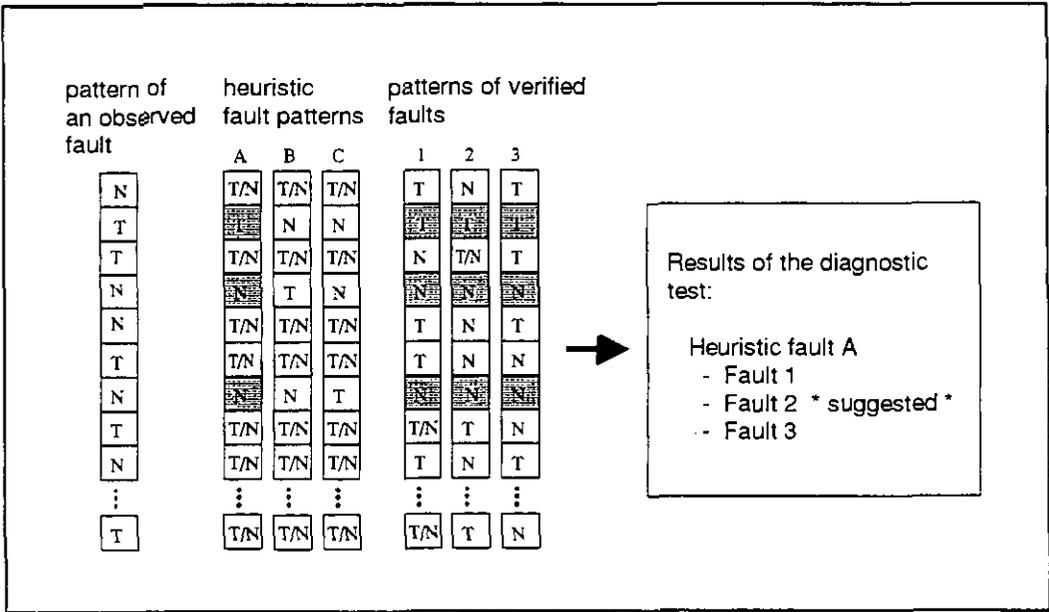
Figure 7. An example of classification. An observed fault is classified using heuristic and verified fault patterns.

Figure 7 presents a classification example, in which the heuristic approach is combined in a pattern recognition procedure. The pattern of the observed fault is compared with heuristic fault patterns and previously verified fault patterns, if such exist. Because only a few constraints of the heuristic patterns are fixed, a heuristic choice, matching the observed fault apparently will be found. When compared to fault patterns of the case library, there may be a fault matching exactly with the observed fault pattern, or which is closely related to the observed fault. While reporting the test results, the system may output the heuristic choice and possible other choices, highlighting the most probable fault, as shown in the figure.

If the user verifies that the observed fault is the same as the suggested one and the patterns of the both faults still differ, the system can replace or add the new features to the knowledge base. Combining old and new fault patterns will change the probabilities p_{ij} or turn some logical state indifferent, depending on the approach. However, pattern recognition and classification of the fault is more straightforward the next time.

When the most probable fault is selected, an exact match is not always found. Thus, one must choose the closest pattern, where the difference between the patterns is minimum. Equation 11 represents one alternative for calculating an approximate solution for the problem. Another alternative is to apply the notation of the table 2 and figure 7 and neglect the probabilities p_{ij}. This easily leads to nearest neighbor indexing and associative retrieval, which allows specification of match and mismatch weights to the features. Further simplification is also possible. A basic approach is template retrieval, which is only a variation of conventional database retrieval.

6 SUMMARY

One of the original ideas was to create a practical FDI method. This means that the method should be robust and consistent with the pragmatic aspects set by the process and its environment. Conclusions after the development and field tests are confirming that the ODT has features, conceivable for a robust and realistic diagnostic method. The applied process model is uncomplicated but at prescribed points it still represents a moderate replica of the process dynamics. It is evident that robust diagnosis cannot be resulted from a poor model. Distinct and abrupt changes in process behaviour usually causes exceeding of several thresholds during one ODT. Thus, fault detection is not based on crossing one limit. The ODT method allows user intervention, which provides the possibility of naming new faults and correct misclassifications. Reasoning utilizes heuristic and adapting classification of the same compact knowledge base, which is easy to implement as an embedded knowledge base system. The learning ability enlarges the areas of application and permits inheritance of system knowledge to similar HVAC systems.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Kurki, M. 1995. Model-based fault diagnosis for mechatronic systems. Espoo 1995, Technical Research Centre of Finland, VTT Publications 223. 116 p.

[2]   Steels, L. 1990. Components of Expertise. AI Magazine, Summer 1990, pp. 28 - 49.

[3]   Leitch, R. & Gallanti, M. 1992. Task classification for knowledge-based systems in industrial automation. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 22, No. 1, pp. 142 - 152.

[4]   Koch, G. G. 1992. Knowledge-based coordination of qualitative on-line diagnostic tests. In: IFAC Symposium on Intelligent Components and Instruments for Control Applications, Malaga Spain 20 - 22 May 1992. 6 p.

[5]   Willsky, A. S. 1976. A survey of design methods for failure detection in dynamic systems. Automatica, Vol. 12, No. 6, pp. 601 -611.

[6]   Isermann, R. 1984. Process fault detection based on modeling and estimation methods -a survey, Automatica, vol. 20, No. 4, pp. 387 - 404.

[7]   Isermann, R. 1994. Integration of fault detection and diagnosis methods. In: IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes -SAFEPROCESS, Espoo, Finland 13 - 16 June 1994. Helsinki University of Technology. pp. 597 - 612.

[8]   Patton, R., J. 1994. Robust model-based fault diagnosis: The state of the art. In: IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes - SAFEPROCESS, Espoo, Finland 13 - 16 June 1994. Helsinki University of Technology. pp. I - 24.

[9]   Pau, L. F. 1981. Failure diagnosis and performance monitoring. New York: Marcel Dekker. 427 p.

[10] Basseville, M. 1988. Detecting changes in signals and systems - A survey. Automatica, Vol. 24, No. 3, pp. 309 - 326.

[11] Kuipers, B. 1986. Qualitative simulation. Artificial intelligence 29, pp. 289 - 338.

[12] Kuipers, B. 1993. Reasoning with qualitative models. Artificial Intelligence 59, pp. 125 - 132.

[13] De Kleer, J. Brown, J. S. 1984. A qualitative physics based on confluences. Artificial intelligence 24, 1 - 3, pp. 7 - 83.

[14] De Kleer, J. 1993. A view on qualitative physics. Artificial intelligence 59, pp. 105 - 114.

[15] Dexter, A. L., Glass, A.S. 1993. The use of qualitative models in fault detection and diagnosis. In: Building Optimisation and Fault Diagnosis System Concept. J. Hyvärinen and R. Kohonen (eds.), Technical Research Centre of Finland, Laboratory of Heating and Ventilation, Espoo, Finland, pp. 79 - 95.

[16] Fouche, P., Kuipers, B. J. 1992. Reasoning about energy in qualitative simulation. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 22, No. 1, January/February 1992, pp. 47 - 63.

[17] Clansey, W., J. 1985. Heuristic classification. Artificial Intelligence, Vol. 27, No. 3, pp. 289 - 350.

[18] Isermann, R., Freyermuth, B. 1991. Process fault diagnosis based on process model knowledge. Part 1: Principles for fault diagnosis with parameter estimation. Journal of Dynamic Systems, Measurement, and Control, Vol. 113, pp. 620 - 626.

[19] Kolodner, J., Mark, W. 1992. Case-base reasoning, IEEE Expert, October 1992, pp. 5 - 6.

[20] Harmon, P. (ed.) 1991. Case-Based Reasoning II. Intelligent Software Strategies, vol. VII, No. 12, pp. I -9.

[21] Duda, R. O., Hart, P. E. 1973. Pattern classification and scene analysis. New York: John Wiley & Sons. 482 p.

FAULT SIMULATION OF A HVAC SYSTEM WITH THERMAL STORAGE TANK

Hideharu Niwa, Dr. Engineer
Nikken Sekkei Ltd., Japan
Nobuo Nakahara, Professor
Nagoya University, Japan

ABSTRACT

This report introduces some cases of fault detection and diagnosis of a HVAC system with a thermal storage tank using the simulation program HVACSIM[+]. From simulation results, and the influence of different kinds of fault at each sub-system level on the temperature profiles of the thermal storage tank are shown.

1 INTRODUCTION

HVACSIM[+] is a simulation tool that can be used to generate normal and fault data for HVAC systems [1][2]. In order to generate the required and reriable fault data for a BOFD study, the authors have added several new TYPEs of subroutine and tried to simulate dynamic behavior for various system alternatives. So far it has been possible to simulate the HVAC system with a thermal storage tank. The HVAC system with a thermal storage tank is one of the most popular and useful system in Japan, for the purpose of reducing erectlic demand and taking advantage of off-peak energy charge.

As shown in Figure 1, the storage system can be regarded as an interface which combines the HVAC sub-system with the heating or cooling PLANT sub-system. Therefore, it will be affected by the condition of both of them.
Figure 2 is a diagram of the influence of sub-systems which contain some kind of faults on the storage system. Since any fault of a sub-system is likely to affect the storage system, most of the faults may be detected by monitoring the temperature profiles of the storage tank.

The aim of this study is to show some cases of fault simulation of the HVAC total system with a storage tank, and to confirm the influence of faults which take place at each sub-system level on the storage system.
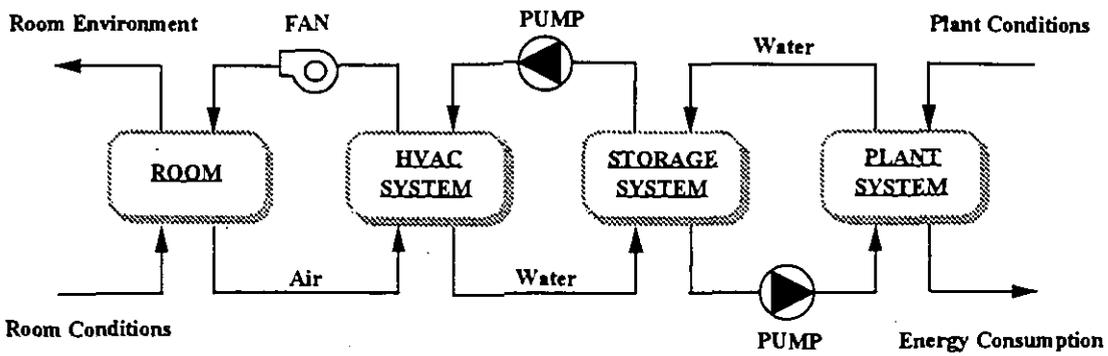
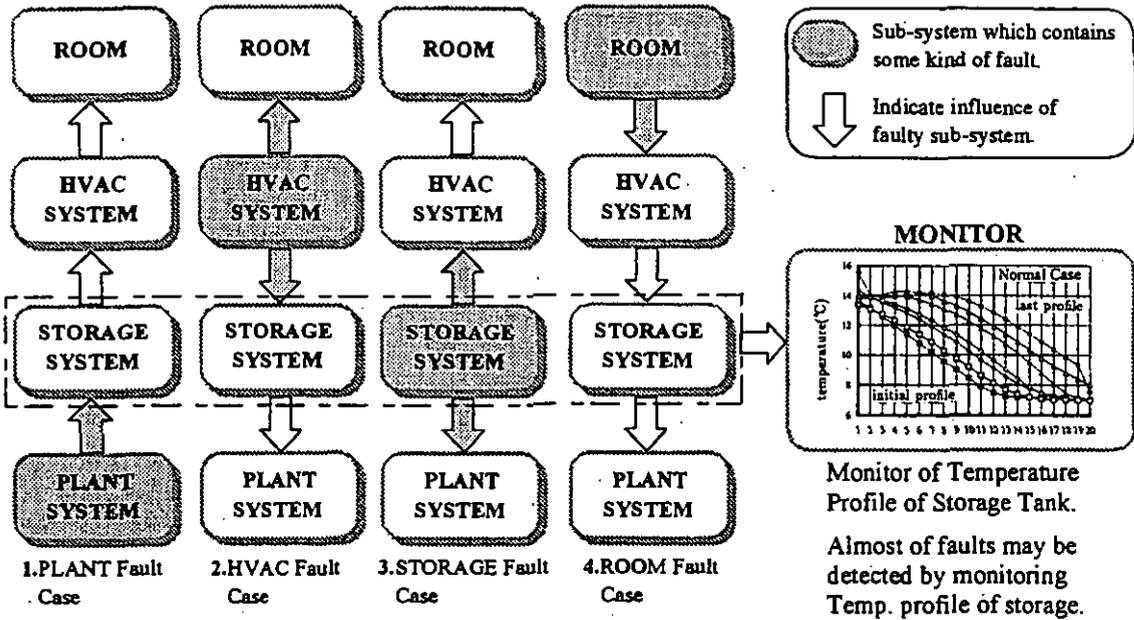Figure 1   General Diagram of HVAC Total System with thermal storage

Figure 2   Influence of Faulty Sub-System for the Storage System

2   DESCRIPTION OF HVAC SYSTEM WITH THERMAL STORAGE

2.1   SYSTEM AND BUILDING DESCRIPTION

Figure 3 shows the diagram of the HVAC system with thermal storage. The 3-room VAV HVAC system is designed for simulation. A set of multi-connected complete mixing storage tanks and a chiller are combined with the HVAC system. In this study, the outlet water temperature of the cooling tower is fixed at a constant value. Table 1 shows the design parameters for each piece of equipment in this system. The hot water boiler is installed to supply hot water to the reheater and heating coil. But the heating coil is not used in this study, because the simulation is only for a peak day in the summer.

Figure 4 shows the plan of the rooms and the main structure of the building simulated in this study.

The internal heat gain in these rooms are assumed as shown in Table 2 in accordance with the schedule as shown in Figure 5. Figure 6 shows the outdoor air temperature and horizontal solar radiation for the simulated day in the summer.

Table 1 Design Parameters of Equipments for HVAC Total System with thermal Storage Tank

| CHILLER | Type of compressor | : Piston |
	Cooling capacity	: 13 KW
	Compressor power	: 3.7KW
	Design evaporator outlet water temperature	: 7 ℃
	Design condenser outlet water temperature	: 24 ℃
	Design evaporator water mass flow rate	: 0.6Kg/s
	Design condenser water mass flow rate	: 0.9Kg/s
STORAGE TANK	Type : Multi-connected complete mixing type	
	Water volume : 2.5㎥×20 tanks	
AIR HANDLING UNIT	Fan : Design air flow rate	: 2.7Kg/s
	Cooling coil : Coolong capacity	: 42 KW
	Design water flow rate	: 2.0Kg/s
VAV UNIT	Design air flow rate	: 0.9Kg/s
	Minimun damper opening	: 0.4
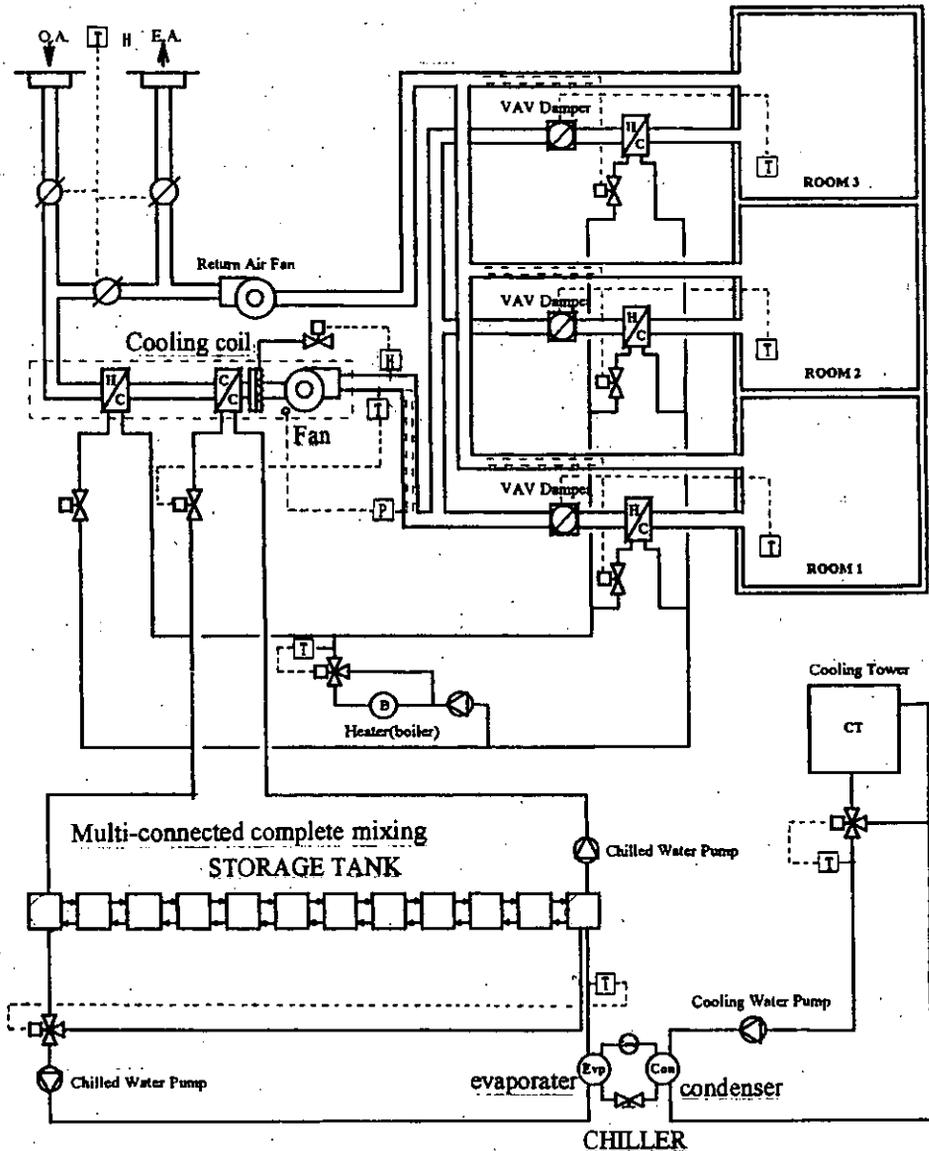
Figure 3   Diagram of a HVAC Total System with thermal Storage Tank

Table 2 Building Heat Gain and Thermal Performance

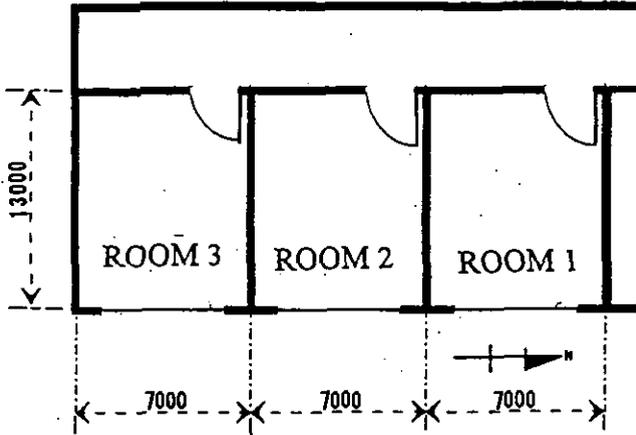| ROOM	Q_{equip} (KW)	Q_{light} (KW)	A_{floor} (m²)	A_{window} (m²)
1	1.0	1.0	98.0	7.2
2	2.25	1.2	98.0	7.2
3	3.5	1.47	98.0	7.2

Figure 5 Indoor Heat load profile

Figure 4 Building Model for Simulation
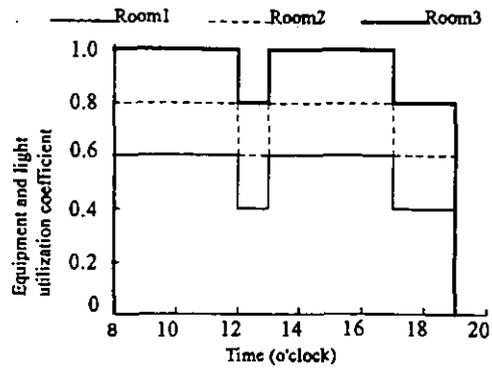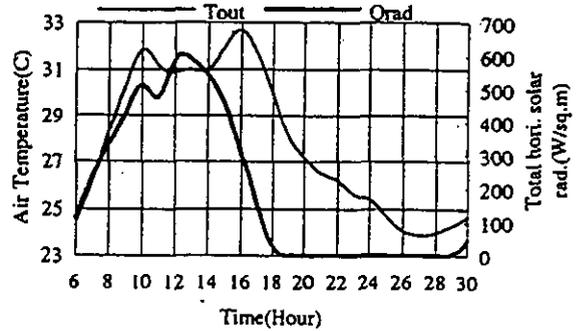

Figure 6 Outside air temperature
and Total Horizontal Solar Radiation

2.2 CONTROL STRATEGIES

The purpose of the control strategies in the summer mode for this system is to control the air temperatures in the three zones within the range 22 ° C- 25.5 ° C. The control strategies for control of the HVAC system during the summer are briefly described as follows,

1) The supply fan is controlled to keep the supply air gauge pressure at 0.249kPa(after main supply duct), and the return fan is modulated to provide air flowrate 0.44kg/s less than that of supply fan.

2) The cooling coil is controlled to keep supply air temperature at the setpoint. The setpoint is reset higher with in the range 16- 20 ° C when the maximum opening of VAV dampers is lower than 80% and lower when the maximum VAV damper opening is higher than 90%. The reset speed of the reseter is 0.0005 ° C/s. This value was selected in order to prevent hunting of the room air temperature.

3) The VAV damper opening changes within the range of 40-100% linearly when the indoor temperature is in the range of 21.9-25.2 ° C as shown in Figure 7.

4) The reheater is modulated by PID controller to prevent the zone air temperature from falling below 22 ° C. The supply hot water temperature *Thw* is reset linearly when outside air temperature *Text* is between -1.1 and 12.8 ° C. When *Text* <-1.1, Thw is set at 45 ° C, and when *Text*>12.8 ° C, *Thw* is set 35 ° C.

5) The minimum outside air(OA) flowrate is controlled by the OA damper at 40% of the total air flow rate when OA temperature is higher than 22.2 ° C.

6)As shown in Figure 8 two kind of operation schedule for chiller is assumed for the ON-PEAK day. Schedule A which takes into account the peak demand window(13:00 ∼ 16:00). Schedule B requires the chiller to operate continuously for 24 hours.

7)The inlet 3-way valve is controlled to keep Chiller outlet water temperature at set point 7 ° C. The chiller input is kept at a constant value.(13KW for schedule A,11KW for schedule B)

8)The outlet water temperature of cooling tower is assumed fixed at 24 ° C.
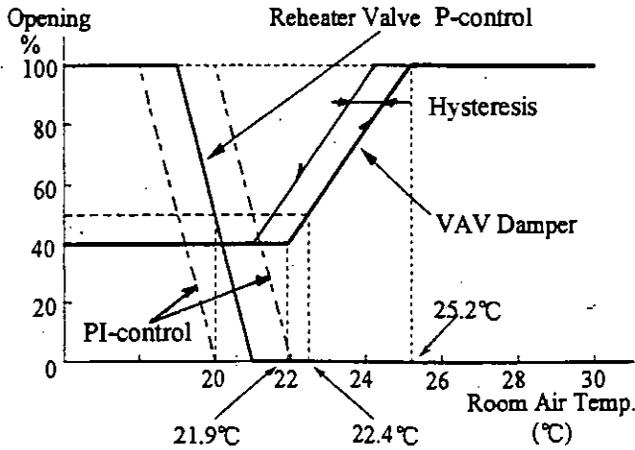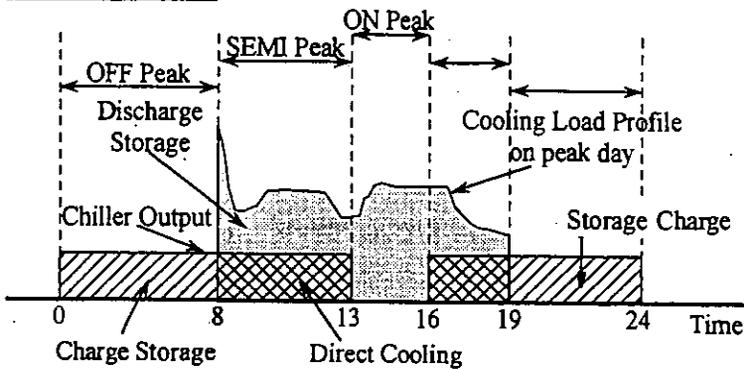
Figure 7   Control of VAV Damper and Reheater Valve
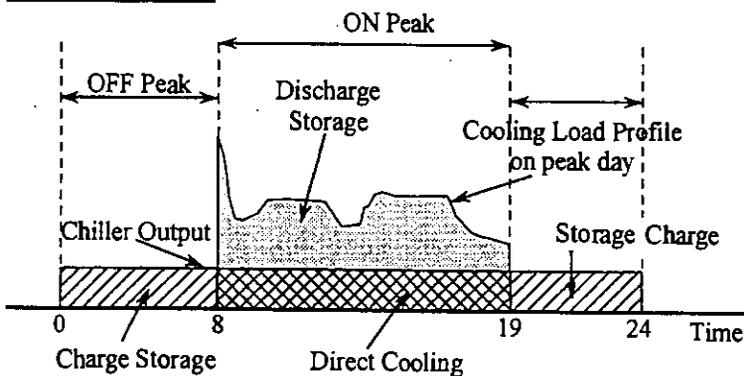
Figure 8   Schedule of Chiller Operation

2.3 NEWLY ADDED MODELS FOR TOTAL SYSTEM SIMULATION

Some new component types were added by the authors to simulate a HVAC system with thermal storge. Table 3 is the list of newly developed TYPEs. The multi-connected complete mixing storage tank is same as the model wich was intoroduced in a separate paper[3].

Table 3   Newly Developed TYPEs List

TYPE 43	Flowrate and pressure balance calculation
TYPE 31	PID controller(set point kick preventing)
TYPE 37	Pressure difference sensor
TYPE 38	Flowrate difference sensor
TYPE 42	Damper or Valve with motorized actuator
TYPE 44~46	Signal resetter

3   PROPOSED FAULTS FOR HVAC TOTAL SYSTEM

The proposed faults considered in this paper are of four kind of typical faults in the sub-systems. They are described below.

1) PLANT Fault   (a fault which takes place at plant system level)

Set point error of the outlet water temperature of the chiller.

The set point is kept higher at 12 ° C.

2) HVAC Fault   (a fault which takes place at HVAC system level)

Set point error of supply air temperature

Set point is kept at a constant 14° C. Therefore the coil outlet water temperature will be lower and water mass flowrate will be higher than in normal case.

3) STORAGE Fault   (a fault which takes place at storage system level)

Lack for water volume in the storage tanks.

Because of leakage of water in the storage tanks, the effective water volume of each tank decreases to 1.5 m^3

4) ROOM Fault   (a fault which takes place at room level)

Excessive internal heat generation

The internal heat gain in the Room3 is excessively high(5.5KW)

4 SIMULATION RESULTS

4.1 BEHAVIOR OF HVAC SYSTEM

Figure 9 and Figure 10 show the simulation results of five cases for schedule A and B, respectively. The difference between shecdule A and B is not so significant, that the description below is appliable to both of them. The fine cases are discussed below.

4.1.1 NORMAL Case

Each of design parameters and the control parameters are properly adjusted for the normal case. The room air temperature is almost maintained inside the control band between 22 and 25.5 ° C. The inlet water temperature to the cooling coil, that is to say outlet temperature of the storage system, increses slightly to about 8 ° C from 7 ℃, while the outlet water temperature changes from about 12 to 16 ° C according to changes in the cooling load . The outlet tempertaure of the chiller is kept at steadily at about 7 ° C by the inlet three way valve of the chiller. The supply air temperature changes from 16 ° C to 20 ° C by the reset schedule as described in 2.2.

4.1.2 Case of PLANT fault

Because the set point of outlet water temperature of the chiler is kept higher at 12 ° C, the outlet water temperature of the chiller goes out of control and increses to about 10 ° C. As the result, the inlet water temperature of the cooling coil is higher than in the normal case, but its effecte on the room air temperature is not serious for this day. However, as shown in Figure 11, the water temperature of the lower parts of tanks are so high that the effect will be carreied over the room air temperature on the next day.

4.1.3 Case of HVAC fault

In this case the supply air temperature set point is kept at 14 ° C. Therefore, during occupied period the VAV damper opening for room3 is kept its minimum and the air temperature of room3 becomes lower than 22 ° C between 17 and 19 o'clock. The outlet water tempetature of the cooling coil is lower than in the normal case and the water flow rate through the cooling coil increases almost to its maximum. As a result of the increase in water flow rate, the difference between inlet and outlet water temperature of cooling coil decreases to about 3 ° C in the afternoon.

4.1.4 Case of STORAGE fault and ROOM fault

These case show similar behavior. In the case of the storage fault, because of a lack for storage capacity, the water temperature of lower tempetature sections of tanks and the inlet water temperature of cooling coil increase in the afternoon. In the case of the ROOM fault, the room air temperature of room3 exceeds the upper limit of the control band. In both of these cases, the efect on the room air temperature is not so serious on this day. However in case of the ROOM fault shown in Figure 11, the temperature of the storage tank at 8 o'clock is so high that it has a significant infuluence on the system behavior the next day.

From these results it is confirmed that the simulation program HVACSIM [+] is capable of simulating faulty condition of a HVAC system with thermal storage, and to give reasonable results.
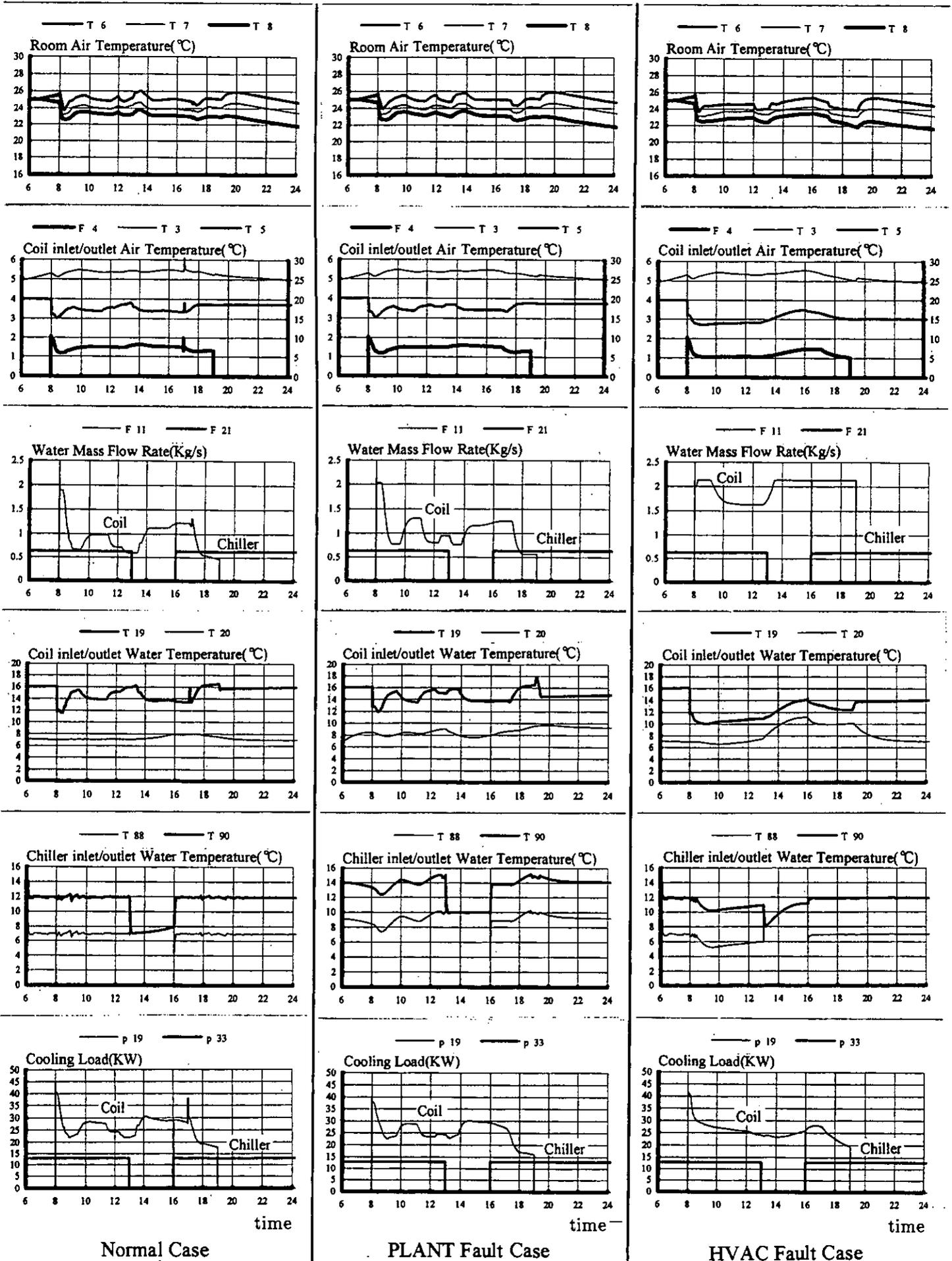
Figure 9 Simulation Results for Schedule A

282

STORAGE Fault Case (left column)

Room Air Temperature(°C) — T 6, T 7, T 8

Coil inlet/outlet Air Temperature(°C) — F 4, T 3, T 5

Water Mass Flow Rate(Kg/s) — F 11, F 21
Coil
Chiller

Coil inlet/outlet Water Temperature(°C) — T 19, T 20

Chiller inlet/outlet Water Temperature(°C) — T 88, T 90

Cooling Load(KW) — p 19, p 33
Coil
Chiller

ROOM Fault Case (right column)

Room Air Temperature(°C) — T 6, T 7, T 8

Coil inlet/outlet Air Temperature(°C) — F 4, T 3, T 5

Water Mass Flow Rate(Kg/s) — F 11, F 21
Coil
Chiller

Coil inlet/outlet Water Temperature(°C) — T 19, T 20

Chiller inlet/outlet Water Temperature(°C) — T 88, T 90

Cooling Load(KW) — p 19, p 33
Coil
Chiller

STORAGE Fault Case   ROOM Fault Case

Figure 10　Simulation Results for Schedule B
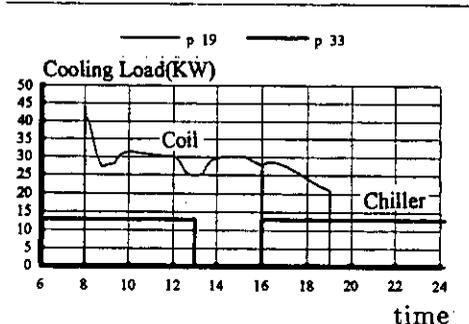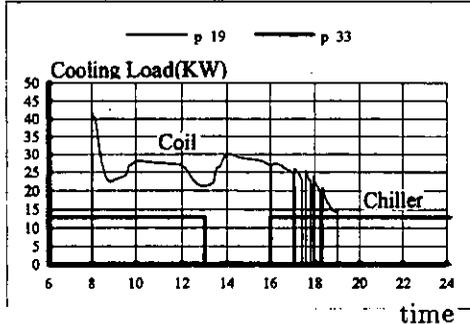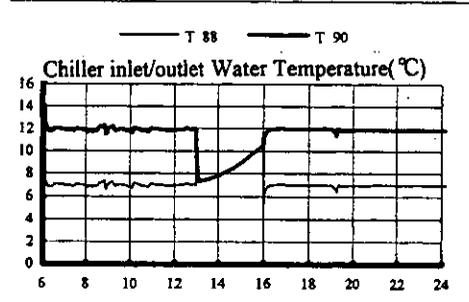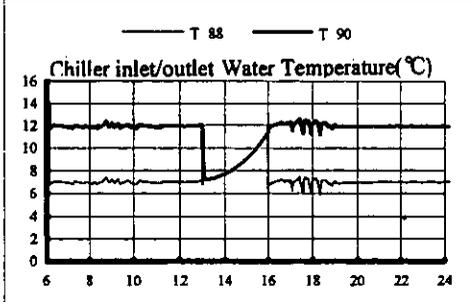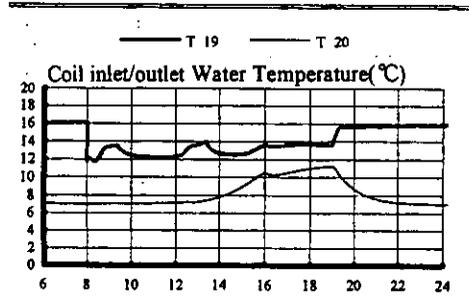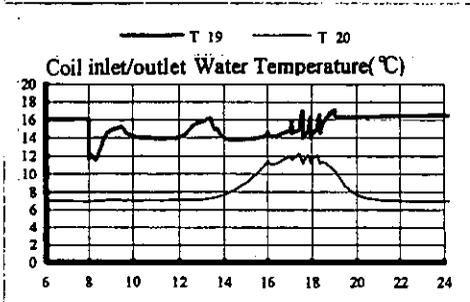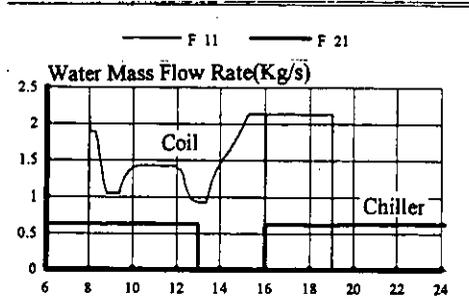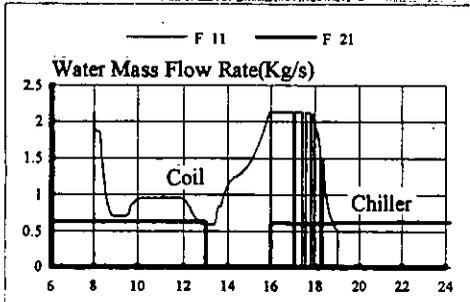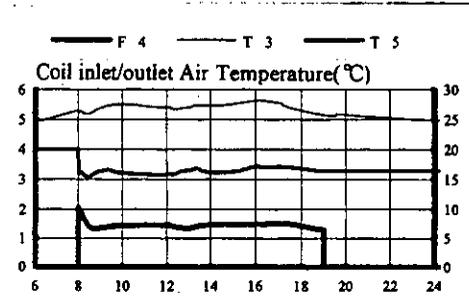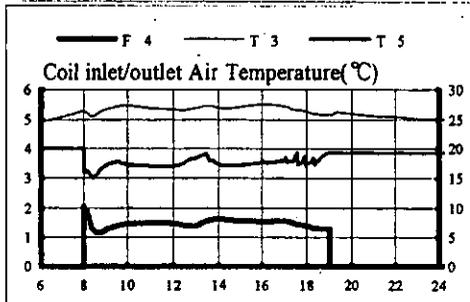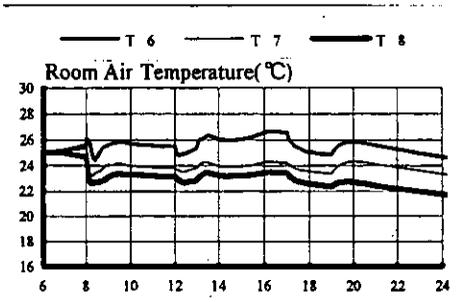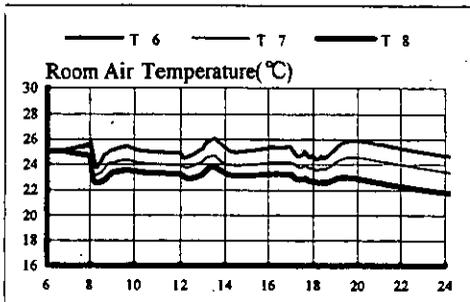
STORAGE Fault Case

ROOM Fault Case

4.2 TEMPERATURE PROFILES IN THE STORAGE TANK

Figure 11 shows the water temperature profiles in the thermal storage tank at each three hours. The horizontal axis is the number of tanks. As shown in this figure, it is confirmed that a different characteristic of the temperature profiles exist for each fault case. The principal difference between these profiles is as follows,

1)In NORMAL Cases the profiles is smoothly sloped and are nearly parallel at any time increment. Water temperature of both ends of tanks is maintained at about 7 ° C and 14 ° C all daylong, since the storage tank is designed optimally at the condition of the normal peak cooling load.

2)In the case of the PLANT fault the profiles wave and cross each other at the lower temperature sections. At 8 o'clock the temperature difference between both ends of tanks is smaller than in the normal cases.

3)In the case of the HVAC fault the profiles wave and cross each other at the higher temperature sections. However, because of the normal charging operation the profile has recovered by 8 o'clok on the next day.

4)In the case of the STORAGE fault, because of a lack for storage capacity the water temperature of the storage tanks rapidly increases. As the result, the temperature of storage tanks are uniformly high at about 14 ° C at the end of occupied period.

5)In the cases of the ROOM fault, the pattern of the temperature profiles is similar to cases of the storage fault. However the maximum temperature of the storage tanks is a little lower than in the case of the storage fault. The most significant difference is that the profile at 8 o'clock on the next day is much higher than profile at 8 o'clock on this day.

From these results it is clear that for the kind of faults studied in this paper all have some effect on the temperature profiles of the storage tanks. Thease kinds of faults were further studied using pattern recognition methods and reportedin a separate paper[4].

5 CONCLUSION

This report only shows several cases of fault simulation of a HVAC system with thermal storage tank. The results demonstrate the influence that faults at the sub-system level have on the thermal storage system. The method of fault detection by monitoring temperature profiles of the storage tank shows considerable promise and should be studied further in real systems
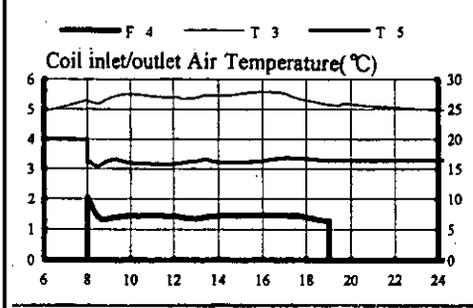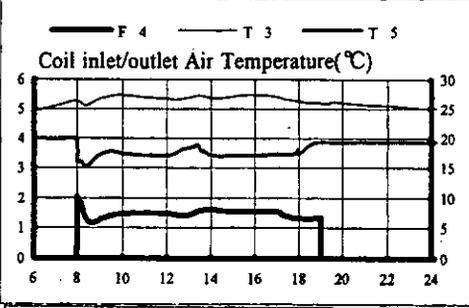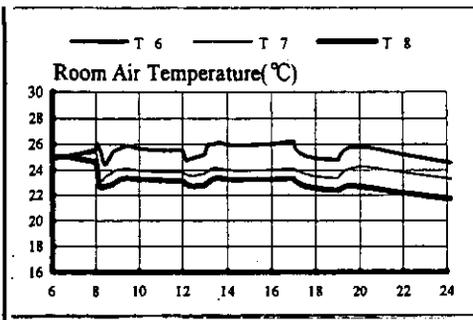
Reference

[1]Park C.,Clark D.R & Kelly G.E., An overview of HVACSIM [+] , a Dynamic Buil/HVAC/Control System Simulation Program,proc.1st Annual Building Energy Simulation Conference, Seattle,WA,USA.1985
[2]Clark D.R., HVACSIM [*] Building System and Equipment Simulation Program: Reference Manual, NBSIR 84-2996
[3]Kazunobu Sagare & Nobuo Nakahara, Fault Detection in Thermal Storage Tank Using Physical Model
[4]Mingjic Zheng & Nobuo Nakahara, Study on Fault Detection and Diagnosis of Thermal Storage Control System with Pattern Recognition
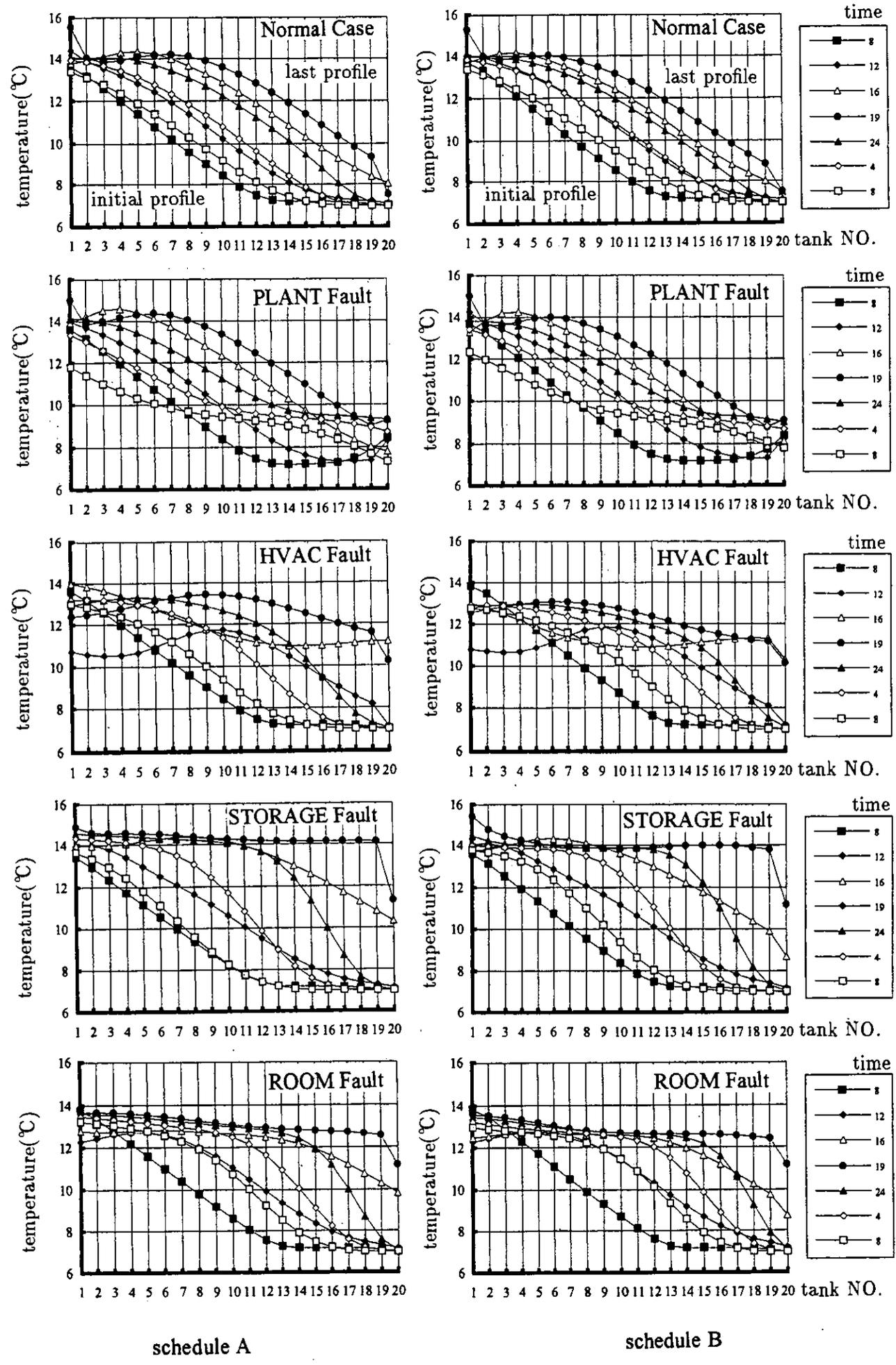
Figure 11 Temperature Profiles of Storage Tank

Fault Detection of Thermal Storage System by Expert System using Fuzzy Abduction

Koichi YAMADA

Advanced Technology Center,
Yamatake-Honeywell Co., Ltd.
YBP West Tower, 134 Gohdo-cho,
Hodogaya-ku, Yokohama 240,
JAPAN
E-mail: yamada@atc.yamatake.co.jp

Kazuyuki KAMIMURA

Building Systems Division,
Yamatake-Honeywell Co., Ltd.
Totate-International Bldg. 2-12-19,
Shibuya, Shibuya-ku, Tokyo 150,
JAPAN

ABSTRACT

Fuzzy abduction is a procedure for deriving fuzzy sets of hypotheses which explain a given fuzzy set of events using a set of rules with a truth value. The derived fuzzy sets of hypotheses are called fuzzy explanations. This paper starts with discussion about diagnosis using conventional expert systems and that using fuzzy relational equations. Then, it proposes a new approach using a fuzzy abduction, and applies the technique to fault detection of a thermal storage system.

1 INTRODUCTION

Diagnosis using expert systems has long been a major topic of research. However, the difficulty of knowledge acquisition has not yet been overcome, though various approaches have been intensively studied in the area of artificial intelligence.

Limiting the problem to diagnosis, one reason for the difficulty of knowledge acquisition is the "directionality" of production rules /1/. Most expert knowledge takes the form: "if some cause happens then certain symptoms will arise," while conventional expert systems use the opposite directionality: "if <symptoms> then <cause>." This means that constructing a knowledge base is almost equivalent to creating knowledge to solve inverse problems in a general manner.

To avoid this difficulty, another approach, called abduction /1-5/, has been proposed. It is a type of reasoning that derives a set of hypotheses (causes) which explain a given set of events (symptoms) using some knowledge (causal relations). Adoption of this approach prevents us from having to solve the inverse problems, and enables us to create a knowledge base more easily.

Another problem in conventional expert systems is the treatment of "intensities", or degrees of symptoms. One way to deal with the problem is to assign different symbols to different degrees of a symptom, as if they were different symptoms. For example, you may assign three different symbols, "extremely-too-high", "very-too-high", and "too-high" to a symptom for water temperature depending on the degree, when it is too high. This approach, however, produces a large number of combinations of "pseudo symptoms" for a cause, and eventually forces us to construct a vast number of rules. To cope with the problem, fuzzy logic, which introduces a value between 0 and 1 to treat a degree of truth or membership, could be used to express the intensities of symptoms.

In this paper, we try to apply fuzzy abduction /4,5/ to fault detection of a thermal storage system. Fuzzy abduction, as you imagine from the discussion above, can avoid the above two problems, directionality and intensity. It employs rules as knowlege to express causal relations between causes and symptoms. Those rules have a truth value

between 0 and 1 that gives a degree of causal intensity. Then, fuzzy abduction is defined as the procedure for deriving fuzzy sets of hypotheses (causes) which explain a given fuzzy set of events (symptoms) under the causal relations. The derived fuzzy sets of hypotheses are called fuzzy explanations.

2 FUZZY RELATIONAL EQUATIONS

Before we go into the details of fuzzy abduction, we should mention a conventional approach using fuzzy relational equations /6-10/, because it apparently solves the above two problems. This approach expresses causes, symptoms and their causalities as a fuzzy set $\underline{P}$ on possible causes $P = \{P_i\}$ (i=1,...,n), a fuzzy set $\underline{Q}$ on possible symptoms $Q = \{Q_j\}$ (j=1,...,m), and a fuzzy relation R on $P \times Q$, respectively. Then, the fuzzy relational equation is given in the following Max-Min composition:

$$q_j = \bigvee_i (p_i \wedge r_{ij}), \tag{2.1}$$

where p_i, q_j and r_{ij} are memberships of P_i in $\underline{P}$, Q_j in $\underline{Q}$ and $<P_i, Q_j>$ in R, respectively.

Then, the inverse problem of fuzzy relational equations–which is a problem to derive p_i given q_j and r_{ij}–could be considered as a kind of abductive reasoning. Furthermore, the memberships p_i, q_j and r_{ij} seem to be interpreted as intensities of causes, symptoms and causal relations, respectively.

However, this interpretation raises a question about the relation between p_i and q_j. In diagnosis, the more (less) the intensity of a cause is, the more (less) those of its symptoms tend to be. Considering an "i" in eq. (2.1), however, the value of q_j can not be greater than that of r_{ij}, even if p_i is large. Such characteristics of the equation might not be adequate for expressing intensities of causes and symptoms.

Based on possibility theory /11,12/, the fuzzy relational equations should be considered as dealing with possibilities rather than intensities of occurrences. If r_{ij} is regarded as a conditional possibility $\pi(Q_j | P_i)$, $p_i \wedge r_{ij}$ is a combinational possibility of "P_i" and "Q_j when P_i is present" /12/. Then, the obtained q_j should be considered as the possibility that Q_j arises, not as the intensity of occurrence of Q_j.

However, again, the interpretation of membership is also strange, when they are directly used in diagnosis. Because, in fuzzy set $\underline{Q}$, memberships should clearly be interpreted as intensities of symptoms, not possibilities, because those symptoms are actually observed, and the possibilities must be 1.0. In studies such as /6,7/, memberships of $\underline{Q}$ are understood as intensities, while those in the obtained fuzzy set $\underline{P}$ as possibilities or certainties. In this interpretation, as you easily notice, possibilities and intensities are confusingly mixed.

In addition, the inverse problem of fuzzy relational equations has cases for which there are no solutions. To cope with the cases, there have been proposed several methods to get approximate solutions /8-10/. However, those methods share a common problem from the practical point of view; they include iterative calculation where the number of iterations is unknown before the trial.

Fuzzy abduction has also cases where there are no solutions (fuzzy explanations). In this approach, however, we can easily obtain approximate fuzzy explanations, as shown later.

3 FUZZY ABDUCTION

3.1 Strong Fuzzy Abduction /4/

In this section, we briefly describe strong fuzzy abduction. We suppose that there are three sets, $P = \{P_i\}$ (i=1,...,n), $Q = \{Q_j\}$ (j=1,...,m) and $R = \{R_{ij}\}$, where $P \cap Q = \emptyset$ and R_{ij} is a rule given as:

$$R_{ij} : P_i \to Q_j. \tag{3.1}$$

P_i, Q_j and R_{ij} have truth values p_i, q_j and r_{ij} between 0 and 1, respectively. Then, the strong fuzzy abduction is defined as follows:

Definition 1
Strong fuzzy abduction is the procedure for deriving fuzzy sets $\underline{P}$ on P that explain a given fuzzy set $\underline{Q}$ on Q under a given set of rules R.

In the above, memberships of P_i in $\underline{P}$ and Q_j in $\underline{Q}$ are equivalent to truth values p_i and q_j, respectively. The derived fuzzy sets $\underline{P}$ are called strong fuzzy explanations. Furthermore, "$\underline{P}$ explains $\underline{Q}$ under R" is true, if the next equation holds:

$$\begin{aligned} q_j &= \max_{i,\, q_{ij}^* \neq \emptyset} (q_{ij}^*) \\ &= \max_{i,\, p_i + r_{ij} \geq 1} (p_i + r_{ij} - 1), \end{aligned} \tag{3.2}$$

where q_{ij}^* is called the CTV (Causal Truth Value) and is given as follows;

$$q_{ij}^* = \begin{cases} p_i + r_{ij} - 1, & \text{if } p_i + r_{ij} \geq 1, \\ \emptyset, & \text{otherwise} \end{cases} \tag{3.3}$$

The CTV given in eq. (3.3) is the smallest truth value that is consistent with modus ponens in Lukasiewicz Infinite-valued logic ($L_{ALEPH-1}$). This means that CTV is the smallest truth value supported by p_i and r_{ij}. $\emptyset$ means that there is no value that is consistent with p_i and r_{ij} according to the definitions of implication "$P_i \to Q_j$" in $L_{ALEPH-1}$.

Then, the necessary and sufficient conditions of existence of strong fuzzy explanations are given below.

$$\forall j \in [j \mid q_j > 0], \exists i; \forall j' \in [j' \mid r_{ij'} - q_{j'} \geq 0], r_{ij} - q_j \geq r_{ij'} - q_{j'} \geq 0. \tag{3.4}$$

When the above conditions are satisfied, it is guaranteed that in general there is only one largest strong fuzzy explanation (SLFE) and multiple minimal strong fuzzy explanations (SMFEs). Fuzzy sets between the SLFE and a SMFE are also strong fuzzy explanations. The SLFE is given as follows:

$$\underline{P}^{max} = \lozenge_{ij} (p_{ij}^{\wedge} / P_i), \tag{3.5}$$

$$p_{ij}^{\wedge} = \begin{cases} p_{ij}^*, & \text{if } p_{ij}^* \neq \emptyset, \\ 1.0, & \text{otherwise.} \end{cases} \tag{3.6}$$

where $p_{ij}{}^*$ is the ITV (Inverted Causal Truth Value) calculated from r_{ij} and q_j by the next equation:

$$p_{ij}{}^* = \begin{cases} q_j - r_{ij} + 1, & \text{if } r_{ij} \geq q_j, \\ \varnothing, & \text{otherwise.} \end{cases} \tag{3.7}$$

ITV is the largest truth value that is consistent with modus tolens in $L_{ALEPH-1}$. This means that ITV is the largest truth value supported by q_j and r_{ij}. Then, $\lozenge$ is an operator defined as;

$$a/A \lozenge b/B = \begin{cases} a/A + b/B, & \text{if } A \neq B, \\ \min(a,b)/A, & \text{if } A = B. \end{cases} \tag{3.8}$$

Furthermore, fuzzy sets given in eq. (3.9) are SMFEs, if and only if they are included in the SLFE and do not include other $\underline{P_k}{}^{min}$ ($k' \neq k$).

$$\underline{P_k}{}^{min} = \sum_{j,\, q_j \neq 0} \left(\Delta_{i \in \phi(j)} (p_{ij}{}^* / P_i) \right), \tag{3.9}$$

where $\phi(j)$ means a set of "i" such that $p_{ij}{}^* \neq \varnothing$ ($r_{ij} - q_j \geq 0$) for the given j, and Δ_i is an operator to pick a term among those with different "i".

For the case where eq. (3.4) is not satisfied, a method to obtain approximate strong fuzzy explanations without iterative calculations has been proposed /4/.

3.2 Fuzzy Abduction for Diagnosis /5/

The strong fuzzy abduction shown above can be used to reason causes $\underline{P}$ of observed symptoms $\underline{Q}$ using knowledge of causal relations R. However, this fuzzy abduction is too strict to use in real-world diagnoses, because they require that the observed symptoms $\underline{Q}$ must be equal to what causes $\underline{P}$ derives under the causal relations R. In the real world, it is risky to assume that operators can find all of the symptoms caused by $\underline{P}$. Therefore, another fuzzy abduction customized for diagnostic problems was proposed /5/.

Now, let us express the fuzzy set on Q given by eq. (3.2) as $E(\underline{P}, R)$. The strong fuzzy abduction defined in **Definition 1** is the procedure to derive causes $\underline{P}$ that satisfy $E(\underline{P}, R) = \underline{Q}$ for given symptoms $\underline{Q}$. The new definition of fuzzy abduction is given below using the concept of "covering" /1/.

Definition 2
Fuzzy abduction is the procedure for deriving fuzzy sets $\underline{P}$ on P that "cover" a given fuzzy set $\underline{Q}$ on Q under a given set of rules R.

The only difference between **Definition 1** and **2** is that between the words, "explain" and "cover". "$\underline{P}$ covers $\underline{Q}$ under R" is defined in the next equation:

$$E(\underline{P}, R) \supseteq \underline{Q} \tag{3.10}$$

In this paper, the term *fuzzy explanation* is taken to mean fuzzy sets $\underline{P}$ satisfying eq. (3.10). Fuzzy explanations that are not strong are termed "weak fuzzy explanations".

The necessary and sufficient conditions of existence of weak fuzzy explanations are given below:

$$\forall \ j, \ \exists i; \ r_{ij} \geq q_j. \qquad\qquad\qquad (3.11)$$

The conditions in eq. (3.11) are always satisfied, if those in eq. (3.4) are. If these conditions are satisfied, there is also only one largest fuzzy explanation (LFE) and multiple minimal fuzzy explanations (MFEs). In this case, the LFE is equivalent to the universe of discourse P. MFEs are all $\underline{P}_k{}^{min}$ (given by eq. (3.9)) which do not include other $\underline{P}_{k'}{}^{min}$ ($k' \neq k$). That is, only $\underline{P}_k{}^{min}$ included in $\underline{P}^{max}$ are SMFEs; other $\underline{P}_k{}^{min}$ are weak MFEs (WMFEs).

4 DIAGNOSIS OF THERMAL STORAGE SYSTEM

We employed the fuzzy abduction for a prototype expert system that diagnoses an imaginary thermal storage system. The knowledge of causal relations was created referring to a collaboration by the Thermal Storage System W.G. of Japan BEMS/BOFD Committee /13/. A part of the knowledge is shown in Table 4.1. The causal intensities were determined intuitively based on our experience.
The algorithm used in the system is as follows:

Step 1: First, the system checks the conditions of eq. (3.11). If they are not satisfied for one or more "j", the system looks for the "i" where q_j - r_{ij} is the smallest, and modifies r_{ij} to be $r_{ij} = q_j$ for all "i" at current "j".

Step 2: The system checks the conditions of eq. (3.4). If they are satisfied, SMFEs are derived from eq. (3.9), and the system shows them as the candidates of causes.

Step 3: If the conditions of eq. (3.4) are not satisfied, WMFEs are derived and the system shows them as the candidates of causes.

The above algorithm adopts only minimal fuzzy explanations (SMFEs and WMFEs) as the candidates of causes. This is because many causes rarely happen at the same time. From the point of diagnosis, the largest fuzzy explanation is understood as the worst case that happens only in an imaginary world.
Then, some results of the diagnosis are shown below. The cases are all imaginary ones devised carefully to demonstrate the potential capability of the algorithm. The intensities of observed symptoms are assumed to be given directly by the user, or through a word expressing a degree of intensity. It could be also possible to calculate intensities of some symptoms using membership functions, if they are measurable symptoms.

(1) Case 1

This is a very simple case. Suppose that the operator found the following three symptoms:

Q_1 : Abnormal change of temp. in tank during off operation hours
at intensity level $q_1 = 0.4$
Q_6 : Heat loss   at intensity level $q_6 = 0.8$
Q_7 : Condensation on slab   at intensity level $q_7 = 1.0$

First, we check whether eq. (3.11) is satisfied or not. In this case, it is not satisfied at j=6 and there are no fuzzy explanations. So, we modify $r_{2,6}$ from 0.7 to 0.8 according to the algorithm shown above. As the result, the modified problem satisfies not only eq. (3.11) but also eq. (3.4). Therefore, there are strong fuzzy explanations, and in this case, only a SMFE is obtained as follows:

1) SMFE : $p_2 = 1.0$, other p_i are 0.0.

In words, the cause is "Damage of insulation". Degree of 1.0 is the intensity of the cause.

Table 4.1 Example of causes of faults - faults knowledge base for thermal storage tank in cooling process

P_i \ Q_j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.0	0.7	0.4	0.4	0.4													
2	0.4					0.7	1.0											
3					0.4			1.0	0.7	0.4						1.0		
4					0.4			1.0	1.0	0.4						1.0		
5					0.4			0.4	0.4	0.4	1.0					0.4		
6											0.7	1.0						
7											0.7		1.0					
8											0.7			1.0				
9											0.7				1.0		1.0	1.0
10		0.4	0.4	0.7	1.0													
11		0.4	0.4	0.7	1.0													

Causes

P1 : Temp. sensor in tank malfunction
P2 : Damage of insulation
P3 : Water level sensor malfunction
P4 : Water supply unit malfunction
P5 : Damage in water proof
P6 : Scale on heat pump evaporator
P7 : Scale in heat pump inlet 3-way valve
P8 : Scale in primary and secondary water heat exchanger
P9 : Scale in piping and foot valve
P10 : Tank volume is too small
p11 : Connection pipe size between tanks is too big

Symptoms

Q1 : Abnormal change of temp. in tank during off operation hours
Q2 : Temp. in coolest side of tank is too high
Q3 : Room temp. and humidity are too high
Q4 : Inlet water temp. of heat pump is too high
Q5 : Heat storage efficiency is insufficient
Q6 : Heat loss
Q7 : Condensation on slab
Q8 : Increase of overflow
Q9 : Increase of water supply
Q10 : Temp. in tank is too high
Q11 : Deterioration in water quality
Q12 : Heat pump COP falls
Q13 : Heat pump inlet 3-way valve malfunction
Q14 : Heat exchanger efficiency is insufficient
Q15 : Water leakage in piping
Q16 : Abnormal water level
Q17 : Flow rate of primary pump is too small
Q18 : Water evacuation

(2) Case 2

Suppose that the operator found the following symptoms:

Q_8 : Increase of overflow   at intensity level $q_8 = 0.8$
Q_9 : Increase of water supply   at intensity level $q_9 = 0.5$
Q_{10} : Temp. in tank is too high   at intensity level $q_{10} = 0.3$

This case satisfies the conditions of eq. (3.11), but not of eq. (3.4). We can derive approximate strong fuzzy explanations using the approach proposed in /4/. However, in this paper, we obtain weak fuzzy explanations assuming that the operator has failed to find all of symptoms.

In this case, there are three causes, P_3, P_4, and P_5 which might cause the observed symptoms. However, P_5 is excluded, because it cannot cause Q_8 at intensity of 0.8 and Q_9 at 0.5. Actually, eq. (3.9) gives us the following two weak MFEs:

1) WMFE 1: $p_3 = 0.9$, other p_i are 0.0.
2) WMFE 2: $p_4 = 0.9$, other p_i are 0.0.

Therefore, the diagnosis could be as follows: "Water level sensor malfunction" (MFE1) or "Water supply unit malfunction" (MFE2). From Table 4.1, we can easily see that the symptoms the operator found are just part of the totality of symptoms that must exist, whichever the true cause is. This proves that the proposed approach can cope with the operator's failures to find some symptoms.

(3) Case 3

Suppose that the operator found the following two symptoms:

Q_{10} : Temp. in tank is too high   at intensity level $q_{10} = 0.6$
Q_{11} : Deterioration in water quality   at intensity level $q_{11} = 0.4$

In this case, conditions of eq. (3.11) is not satisfied at j=10. So, in order to obtain fuzzy explanations, we modify $r_{3,10}$, $r_{3,11}$, and $r_{3,12}$ from 0.4 to 0.6. However, the modified problem still does not satisfy the condition of eq. (3.4), because there is neither P_i nor combination of P_i that causes only Q_{10} and Q_{11}. So, using eq. (3.9), we obtain WMFEs that causes other symptoms as well as Q_{10} and Q_{11}.

1) WMFE 1: $p_5 = 1.0$, other p_i are 0.0.
2) WMFE 2: $p_3 = 1.0$, $p_6 = 0.7$, other p_i are 0.0.
3) WMFE 3: $p_3 = 1.0$, $p_7 = 0.7$, other p_i are 0.0.
4) WMFE 4: $p_3 = 1.0$, $p_8 = 0.7$, other p_i are 0.0.
5) WMFE 5: $p_3 = 1.0$, $p_9 = 0.7$, other p_i are 0.0.
6) WMFE 6: $p_4 = 1.0$, $p_6 = 0.7$, other p_i are 0.0.
7) WMFE 7: $p_4 = 1.0$, $p_7 = 0.7$, other p_i are 0.0.
8) WMFE 8: $p_4 = 1.0$, $p_8 = 0.7$, other p_i are 0.0.
9) WMFE 9: $p_4 = 1.0$, $p_9 = 0.7$, other p_i are 0.0.

In this example, if the cause is just one, it must be "Damage in water proof". Otherwise, two causes are arising at least. One is "Water level sensor malfunction" or "Water supply unit malfunction". The other is one of "Scale on heat pump evaporator", "Scale in heat pump inlet 3-way valve", "Scale in primary and secondary water heat exchanger", and "Scale in piping and foot valve".

5 CONCLUSIONS

A new approach of diagnosis using fuzzy abduction was proposed and applied to diagnosis of a thermal storage system. Diagnosis is a field that attracts much interest among researchers in many areas, and has been studied intensively for long time. However, conventional approaches both from knowledge engineering and fuzzy theory have problems dealing with causal relations with intensities of causes and symptoms, while the treatment of intensities is essential in some diagnoses dealing with numerical and/or sensual values such as temperature, level, quality, etc. The diagnosis of a thermal storage system is one of such examples.

In the paper, we discussed the problem in detail, and showed that fuzzy abduction is effective to deal with it. Then, our prototype system demonstrated that the approach could apply to the diagnosis of a thermal storage system successfully.

6 REFERENCES

1. Peng, Y., Reggia, J. A. Abductive Inference Models for Diagnostic Problem-Solving. Springer-Verlag. 1990.
2. Pople Jr., H.E. On the mechanization of abductive logic. IJCAI-73. 1973. pp.147 - 152
3. Poole, D. A Logical Framework for Default Reasoning. Artificial Intelligence. 1988. Vol. 36. pp. 27 - 47.
4. Yamada, K., Mukaidono, M. Fuzzy Abduction Based on Lukasiewicz Infinite-valued Logic and Its Approximate Solutions. FUZZ-IEEE / IFES '95. 1995. Vol. 1. pp. 343 - 350.
5. Yamada, K. Fuzzy Abductive Reasoning for Diagnostic Problems. IFSA '95. 1995. Vol. 1. pp. 649 - 652.
6. Terano, T., et al. Diagnosis of Engine Trouble by Fuzzy Logic. IFAC 7th World Congress. 1978. pp. 1621 - 1628
7. Mo, H., et al. Improvement of Reasoning Model for Fuzzy Diagnosis System. Journal of Japan Society for Fuzzy Theory and Systems. 1993. Vol. 5. No. 3. pp. 565 - 576 (in Japanese)
8. Pedrycz, W. Numerical and Applicational Aspects of Fuzzy Relational Equations. Fuzzy Sets and Systems. 1983. Vol. 11. pp. 1 - 18.
9. Pedrycz, W. Inverse Problem in Fuzzy Relational Equations. Fuzzy Sets and Systems. 1990. Vol. 36. pp. 277 - 291.
10. Arnould, T., et al. Backward-chaining with fuzzy "if... then..." rules. 2nd IEEE Inter. Conf. Fuzzy Systems. 1993. pp. 548 - 553.
11. Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems. 1978. Vol. 1. pp. 3 - 28.
12. Nguyen, H. T. On conditional possibility distributions. Fuzzy Sets and Systems. 1978. Vol. 1. pp. 299 - 309.
13. Sagara, K., Nakahara N., Kamimura, K., Miyazaki, T., Nakamura, M. Second draft of the Thermal Storage System W.G. of Japan BEMS/BOFD Committe. Nov. 15. 1993.

APPLICATION OF BLACK-BOX MODELS TO HVAC SYSTEMS FOR FAULT DETECTION

H.C. Peitsman, V.E. Bakker
TNO Building and Construction Research
Department of Indoor Environment, Building Physics and Systems
Delft, The Netherlands

Abstract

The paper describes the application of black-box models for fault detection and diagnosis (FDD) in HVAC systems. In this study, Multiple Input- Single Output (MISO) ARX models and Artificial Neural Network models are used.
The ARX models are examined for different processes and subprocesses and compared with each other. Two types of models are established, called *system models* and *component models*. In case of system models, the HVAC system as a whole is regarded a black box, instead of a collection of component models. With the component model type, the components of the HVAC system are regarded as separate black boxes.

Keywords: Fault detection, Fault diagnosis, HVAC systems, System Identification Methods

1 INTRODUCTION

The purpose of system identification techniques is to describe the behaviour between the inputs and outputs of a process mathematically. There are numerous techniques available to describe a dynamical process with a mathematical model. A special category of models is the black box models such as *ARX models* and artificial *Neural Networks*.

This paper presents the results of the work done in the context of the IEA-ANNEX 25 Workgroup (Kohonen and Hyvärinen 1991). The considered processes are (1) a laboratory chiller and (2) a simulated VAV system.

The theory of black-box modelling is briefly explained. A learning data set is analyzed data, collected from a laboratory chiller. Deviations of faulty data can be detected by the models because the models are trained under healthy conditions. This is illustrated by an example of a Chiller system and a Variable Air Volume (VAV) system. Detection of faults can be achieved by comparison of the process (measured) value with the estimated value.

ARX models and Artificial Neural Networks (ANN) give the user a good opportunity to detect faults and deviations of several processes.

The use of Artificial Neural Networks (ANNs) for fault detection is still in an experimental phase. The major results come from Dexter and Hepworth (1993), who compared fuzzy and neural methods. Kreider et al. (1992), Curtiss et al. (1993) Seem and Braun (1991) demonstrated how ANN controllers can be used for the same

purpose as proportional, integral, and derivative (PID) algorithms. Previous results (Duyvenvoorde 1993) showed that expert systems enhanced with neural networks appear to be particularly promising for efficient and semi-automatic supervision of HVAC systems for commercial buildings. ANNs models represent a new methodology in HVAC system prediction. There are no strict rules regarding the amount of hidden layers and neurons. There are some guidelines (Fogel 1991) available.

2 SYSTEM IDENTIFICATION

The purpose of system identification (Ljung 1987) is to describe the behaviour between the inputs and outputs of a process mathematically. ARX models and ANN models can be categorized as black box models because they require less physical knowledge of the process. The process is considered as a black box with the relationship between the inputs and the outputs described by a mathematical formula. The mathematical model is derived as accurately as possible with the available input and output data of the considered process.

The procedure to determine a proper black box model from observed input-output data involves three basic elements:

1. Measurement of input-output data
2. Selection of model structure and estimation of parameters
3. Validation of selected model

Input and output values of the process are collected in a dataset. This dataset can consist of design data or measured data obtained under "healthy" conditions. Such measured datasets require sufficient dynamical conditions of the examined process.

Final selection of the model is the most significant as well as the most difficult step to make. The selection is influenced by proper insight and a prior knowledge of a process.

The accuracy of a model is tested by comparison of the forecasted value with the measured values. This means that the measured input variables are presented to the derived model whereafter the computed (forecasted) values are compared with the measured values. Also some statistical properties are examined to validate the obtained models such as reliability, average error, standard deviation of residuals, auto-correlation and cross validation.

2.1 PARAMETER IDENTIFICATION FOR ARX MODELS

The parameters of a model can be estimated by employing a system identification method. In this study, one of the two types of model used was the Multiple Input- Single Output (MISO) Auto Regressive with eXogenous inputs (ARX) model.

Consider a process with an input signal $u(t)$ and output signal $y(t)$. Probably the most simple input-output relationship is obtained by describing it by a linear equation:

$$y(t)+a_1y(t-1)+...+a_{na}y(t-na) = b_1u(t-1)+...+b_{nb}u(t-nb)+e(t) \qquad (1)$$

Since the white-noise term e(t) here enters as a direct error in the difference equation, the model is often called an *equation error model* (structure). The white-noise is assumed to be constant and is captured in the signal y(t). ARX models deal well with white noise and are therefore used for construction of the models. Other models gave no improved results.

2.2 ADDITION OF EXTRA INPUTS FOR THE MODEL

The equation (1) for the ARX model indicates that ARX models describe a process by a linear equation. Real processes seldom have a direct linear relation between in- and output of the process. Certain capacities and delays also give the process time dependencies.

The structure of the ARX model takes account of the probable existence of time dependencies by offering the possibility of using the previous values of the inputs and output(s) in the calculation. However, the relation between these previous values and the output still is linear.

In reality the relationship between the inputs and the output of a process is much more complex. Due to the assumption of linearity, the model has become less complex and the initialisation (fitting) of the model has become easier, but the performance of the model is poorer because of this assumption. To achieve an improved ARX model, extra inputs are added to the model. An example with three physical inputs and six artificial inputs is presented in Figure 1. The extra inputs are squares and/or multiplications of the input values.
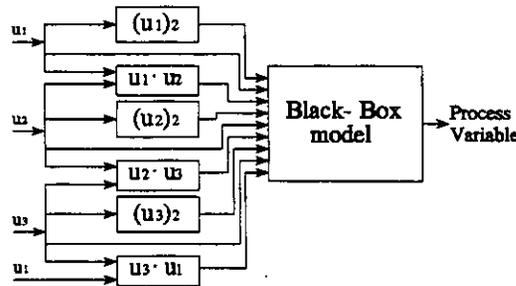
Figure 1 Artificially added inputs to the ARX model.

Another possibility is to use historical measured model inputs (t-1, t-2, ..) and predicted historical outputs as new model inputs. Using historical outputs as new model inputs will cause the model to be less sensitive to spikes in the input signals. The new predicted output is taken into account with the history of the process; it acts like a filter.

Also, the relations between in- and outputs are no longer linear but of a second degree.

2.3 REMOVAL OF THE AVERAGE VALUE

The performance of the model can also be improved by subtracting the mean values from the input and output signals. Note that the constants to be subtracted from the signals have only to be determined during the training (fitting) of the model. One of the aims of the Fault-Diagnosis-System is to detect trends in signals (slow shifts in the mean), so that faults, which develop gradually in the system, can be detected.

3 MODEL STRUCTURES

3.1 THE STRUCTURE OF THE ARX MODEL

For instance, consider an ARX model not fitted with one input signal but nine. An example (van Galen 1993, Peitsman 1994) of the structure of the input signal u(t) of ARX model could be:

$$u(t)=[u1 \ u2 \ u3 \ u1^2 \ u2^2 \ u3^2 \ u1*u2 \ u2*u3 \ u1*u3]$$

The example model has three physical inputs but the total number of inputs are artificially upgraded to nine. So, the process value y(t) is polynomial equation with 19 arguments and 19 polynomial values.

The equation for the forecasted process variable y(t) is:

$$y(t) = a_1 y(t-1) + b_1 u(t) + b_2 u(t-1) \tag{2}$$

Where y(t) is the predicted process value, u(t) is the vector of input signals and b_i the corresponding parameter vector. The structure of b_i is described as:

$$b_i=[bi_1, \ bi_2, \ bi_3, \ bi_4, \ bi_4, \ bi_5, \ bi_6, \ bi_7, \ bi_8, \ bi_9], \quad i= 1,2.$$

The number of orders of historical-terms of u(t) and y(t) can be varied. In this example, the number of historical-terms gives the best results for u(t) is 2 and y(t) is 1. This means a mathematical model structure according equation (2).

The complete structure of the example ARX model is given by:

$$
\begin{aligned}
A1(q) &= 1 + a_1 q^{-1} + a_2 q^{-2} + \ldots + a_n q^{-n} \\
B1(q) &= b_{11} q^{-1} + b_{12} q^{-2} + \ldots + b_{1n} q^{-n} \\
B2(q) &= b_{21} q^{-1} + b_{22} q^{-2} + \ldots + b_{2n} q^{-n} \\
B3(q) &= b_{31} q^{-1} + b_{32} q^{-2} + \ldots + b_{3n} q^{-n} \\
&\ldots\ldots \\
&\ldots\ldots \\
B9(q) &= b_{91} q^{-1} + b_{92} q^{-2} + \ldots + b_{9n} q^{-n}
\end{aligned}
\tag{3}
$$

3.2 THE STRUCTURE OF THE ARTIFICIAL NEURAL NETWORK (ANN)

A neural network model is usually built up out of three layers, the input layer, the hidden layer and the output layer (Nelson and Illingworth 1991). The three layers are connected in a non-linear way (see Figure 2).
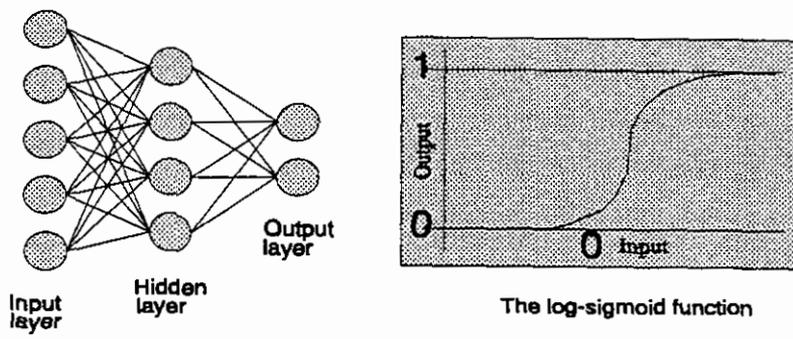
Figure 2 The neural network structure and the log-sigmoid function

In this study the back propagation algorithm is used to train multi-layered feed-forward networks with a Log-Sigmoid transfer function (equation 4).

$$y_i = \frac{1}{1 + e^{(\sum_j W_{ij} x_j - b_j)}}$$

(4)

3.3 THE EXAMINED CHILLER

The examined refrigerant system is a reciprocating chiller. It is a laboratory chiller plant (van Galen 1993, Peitsman 1994) equipped with the following measurement instruments:

1. Temperature sensors.

2. Pressure transducers.

3. Water flow metering devices for measurement of the incoming cooling water flow.

4. Heatmeters for measurement of heat flows and for interpreting the quantity of the water flow.

5. Power consumption metering device for measurement of the electrical input to the motor of the compressor.

6. Adjustment of signal from the extant mass flow transmitter which measures the refrigerant mass flow.

The Process Instrumentation Diagram of the chiller plant is presented in Figure 3.
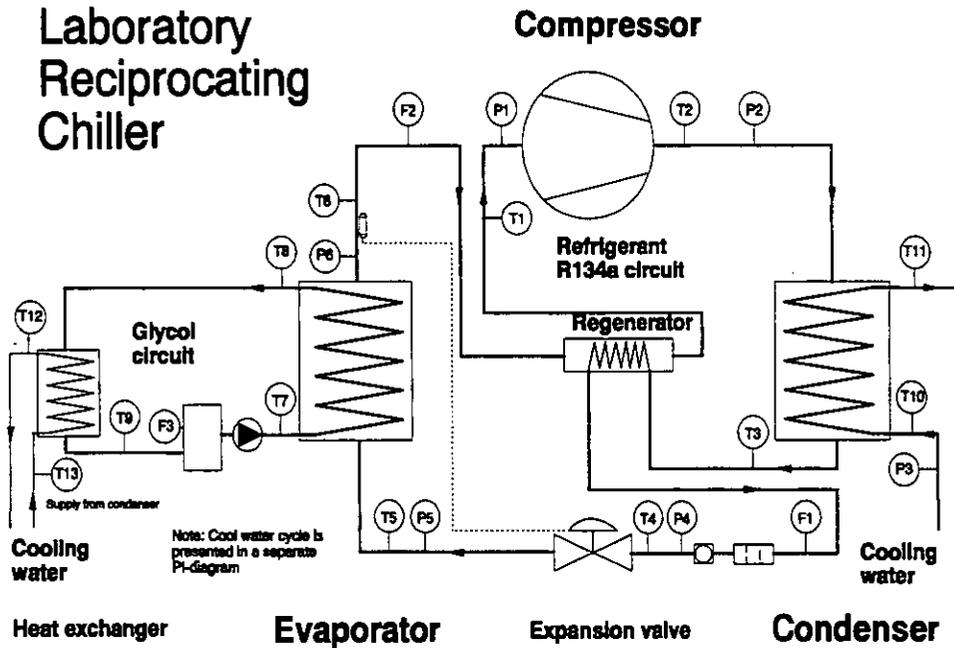
Figure 3 PID of chiller plant

3.4 BLACK BOX MODELS OF THE CHILLER

Two types of models are created with the measured learning data:

1. *System models*; the chiller is considered as a black box where three independent inputs estimate the outputs.

2. *Component models*; major components of the chiller are considered as a black box.

The datasets were collected with a data acquisition unit. After analysing and removing trend measurement faults from the data, a dataset is available to train the models. This means that the model is fitted using healthy data.

3.5 SYSTEM MODELS

Independent variables are necessary to generate a system model. The condenser and evaporator supply temperatures are limited by external conditions, respectively the outside air temperature and the internal load of a building. The instantaneous power of the compressor is a measure of the internal load of the chiller. The model parameters are determined for every process variable with mentioned inputs, as shown in Figure 4.
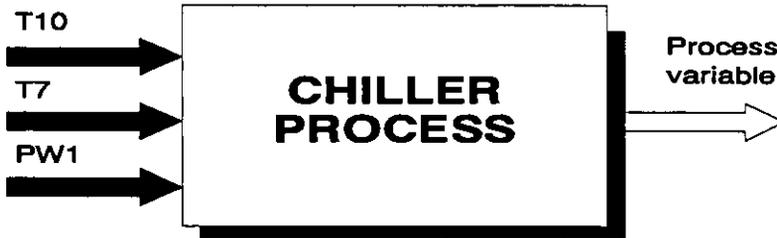
Figure 4 Model scheme

Observation data was collected from a data-logger every 10 seconds. This sample-time determines the time step of the models. The sensor readings that are used as **inputs** for the ARX- and ANN system models are:

- T10   : condenser supply water temperature
- T7   : evaporator supply glycol temperature
- PW1   : instantaneous power of compressor
- F4   : flow cooling water condenser (ANN only)

These magnitudes were selected because their universal availability in Building Automation Systems (BAS).

The 14 estimated process **output** varables of the system models are:

T1   : Suction refrigerant gas temperature

T2   : Discharge refrigerant gas temperature

T3   : Leaving refrigerant liquid temperature condenser

T4   : Entering refrigerant liquid temperature expansion valve

T5   : Entering refrigerant liquid/vapour temperature evaporator

T6   : Leaving refrigerant vapour temperature evaporator

T8   : Outlet chilled glycol temperature

T11   : Leaving cooling water temperature condenser (inlet heat exchanger)

F1   : Refrigerant mass flow

P1   : Suction refrigerant pressure

P2   : Discharge refrigerant pressure

P4   : Refrigerant pressure before expansion valve

P5   : Refrigerant pressure after expansion valve

P6   : Refrigerant pressure after evaporator

3.6 THE ARX SYSTEM NETWORK CONFIGURATION

The total amount of inputs used for deriving the **system** model of the chiller system is 9. The model has three physical inputs but, by combining the physical inputs, the total number of model inputs is artificially increased to nine. So, the process value y(t) is given by a polynomial equation with 19 arguments and 19 coefficients.

The **derived inputs** are:

$T10^2$ $T10*T7$ $T10*PW1$

$T7^2$ $T7*PW1$

$PW1^2$

The **historical inputs** are :

Inputs: the values from the previous time step (t-1) of all the 9 inputs are used.

Outputs: the predicted value from the previous time step (t-1) of the output is used as input.

Each process output variable has its own model. This means that fourteen output models are needed for the defined fourteen process variables. For FDD applications a selection of models can be used.

3.7 THE NEURAL NETWORK CONFIGURATION

The ANN model with 10 neurons gave very poor results, and so 15 neurons were used for the *component* models and 20 neurons for the *system* models.

The same observation data was used as for the ARX models (time step 10 seconds).

The total number of inputs used for deriving the **system** model of the chiller system is 8.

The **measured inputs** are:

T10 T7 F4 PW1

The **historical inputs** are :

Inputs: the values of all 4 inputs from the previous time step.

Outputs: the predicted values for the output from the two previous time steps.

The Network predicts all system model outputs in one run. This is possible because the same inputs are used for all outputs.

3.8 COMPONENT MODELS

The configuration of the component models is achieved after studying the structure of refrigerant systems. This makes the model a little bit "grey"(some physical knowledge) instead of black (with no physical knowledge). The laboratory chiller consists of a compressor, condenser, evaporator, expansion valve and a subcooler. The filter-drier is not considered in the model. Every component has its own inputs and outputs based on the physical inputs and outputs of the selected components in the chiller system. The configuration of the refrigerant components of the chiller is shown in Figure 5.
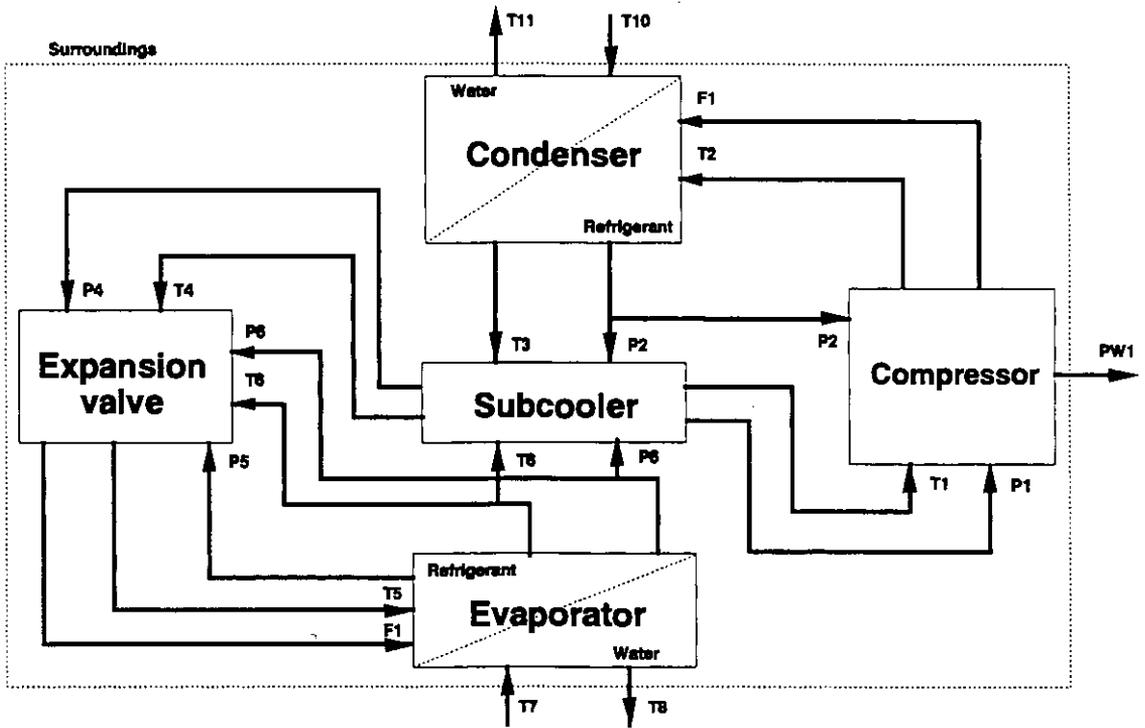
Figure 5 Configuration of the refrigerant components

Table 1.   The selected inputs and outputs of the component models

	Compressor	Condenser	Evaporator	Exp. valve	Subcooler
Inputs	T1,P1,P2	F1,T2,T10	T5,T7,F1	P4,P5,P6,T4,T6	T3,P2,P6,T6
Outputs	F1,T2,PW1	T11,T3,P2	P5,P6,T6,T8	T5,F1	P1,T1,P4,T4

Selected inputs and outputs of the component models are present in table 1.

The structures of the ARX- and ANN inputs of the component models are the same as described for the system models. For each component, different Network models were used because each component has different physical inputs.

4 MODEL VALIDATION

4.1 RELIABILITY COEFFICIENT

Determination coefficient (r^2) is a measure for reliability of the obtained resulting black box model. The coefficient has a value between 0 and 1. If r^2 is equal to 0 then the model is fitted inadequately.
The mathematical description is:

$$r^2 = 1 - \frac{SS_e}{SS_y} \tag{5}$$

where SS_e is the sum of the squares of the differences between the measured and the estimated output Y,

$$SS_e = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{6}$$

SS_y is the sum of the squares of the differences between the measured and average output value Y.

$$SS_y = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{7}$$

If r^2 is equal to 1 then the model fits the measured output perfectly.

4.2 STANDARD DEVIATION AND AVERAGE ERROR OF RESIDUALS

For σ the sample estimation σ is used, based on the deviations $x_i = X_i - X_{mean}$ of the sample observations from the sample mean.
The average error is defined as:

$$AE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{8}$$

The outcome of the average error gives an indication of the accuracy of the model by comparing the model output with the measured output.

4.3 AUTO-CORRELATION OF RESIDUALS

The auto correlation coefficient is a measure of the closeness of relationship between a time series and the same time series one timelag back. For large time series the denominator becomes $\sum x^2(t)$.

$$r = \frac{\sum(x(t)*x(t-1))}{\sqrt{(\sum x^2(t))(\sum x^2(t-1)}} \tag{9}$$

4.4 CROSS VALIDATION

To check wheather the selected model is correct, a separate dataset is set apart to test the model. This dataset is used to calculate the residual by simulation of the selected model. *The cross validation error can be defined as the sum squared error of the residual of the test data.* Generally if the sum squared error of a model is minimized during the learning procedure with the learning dataset, the cross validation error will be minimized as well. This holds until a certain point, where the sum squared error on the learning

Figure 6 Error/Time Relation

data set still can be further minimized, at the cost of an increasing cross validation error. The learning dataset will become over-fitted and the model will lose its generalizing ability.

Because general applicability of the model is desired, the best point to stop the training of a model is the point where the cross validation error has found its minimum.

4.5 VISUAL OBSERVATION OF THE RESIDUE

A good method to see if the model is proper is a visual observation of the residue. A proper model gives a residue which is uncorrelated with the used inputs. This means the model has used all the information that was present. An inproper residual implies that the model found a local minimum, or some important inputs lacked. Other starting values of the learning weights or even other inputs should result in a better model. With ANN's this still is a matter of trial and error.

The residual is defined by:

$$Y_{res} = y_{for} - y \tag{10}$$

with y_{for} as the predicted and y as the measured (real) value.
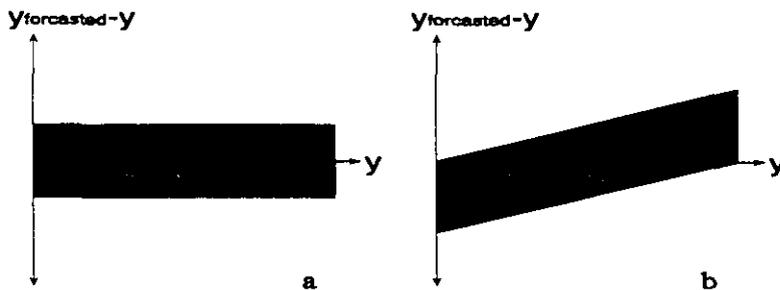
Figure 7 Model validation with residual test

A **proper** model is a model where residual values as a function of time are uncorrelated with the inputs, refer to figure 7a. An **improper** model has an oblique shape of residual plots, the deviation depends on the magnitude of the real value, refer to figure 7b. This implies that certain factors are not considered in the fitted model. Extra input parameters have to be added or the model structure has to be changed.

4.6 VALIDATION RESULTS OF THE CHILLER MODELS

A total of 14 system models and 16 component models were fitted. An example of the discharge pressure P2 black-box models (ARX+ANN) is presented in Figure 8a and 8b. The residual test for the discharge pressure P2 is presented in Figure 8c and 8d. An example of fault detection using discharge pressure P2 when there is air in the system is presented in Figure 8e and Figure 8f.

Table 2:   Model Validation of system- and component models

Outputs of Neural Networks and ARX System Models							
Property	Reliability		Std_residuals		Average error		Auto corr.	
	ARX (%)	ANN (%)	ARX (kPa)	ANN (kPa)	ARX (kPa)	ANN (kPa)	ARX (%)	ANN (%)
P1	88	98.5	8.0	3.00	6.0	2.20	99	80.2
P2	99	99.5	13.0	8.60	9.0	6.60	93	62.4
Property	ARX (%)	ANN (%)	ARX (K)	ANN (K)	ARX (K)	ANN (K)	ARX (%)	ANN (%)
T11	99	99.4	0.56	0.396	0.4	0.299	92	69.4

Outputs of Neural Networks and ARX System Models							
Property	Reliability		Std_residuals		Average error		Auto corr.	
	ARX (%)	ANN (%)	ARX (kPa)	ANN (kPa)	ARX (kPa)	ANN (kPa)	ARX (%)	ANN (%)
P1	100	99.8	0.6	1.00	0.20	0.70	26	11.5
P2	99	99.8	12.0	5.60	8.00	4.20	91	22.5
Property	ARX (%)	ANN (%)	ARX (K)	ANN (K)	ARX (K)	ANN (K)	ARX (%)	ANN (%)
T11	99	99.7	0.45	0.272	0.33	0.168	89	37.6

The results of the model validation of three models are presented in Table 2.
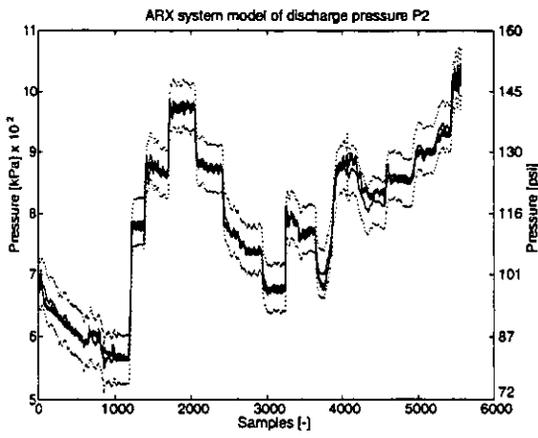
8a ARX model discharge pressure P2
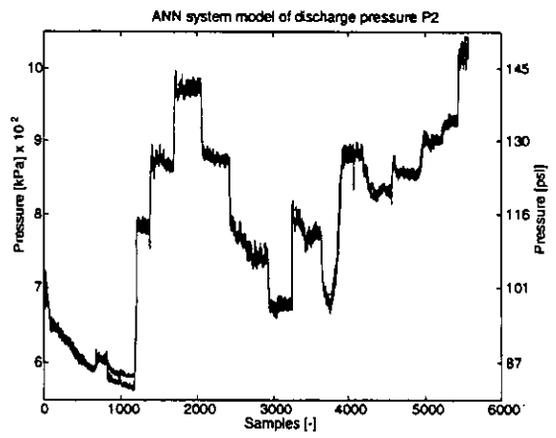
8b ANN model discharge pressure P2
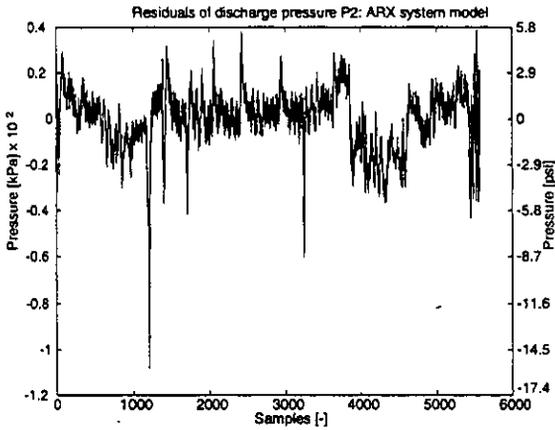
8c Residue of ARX model for P2
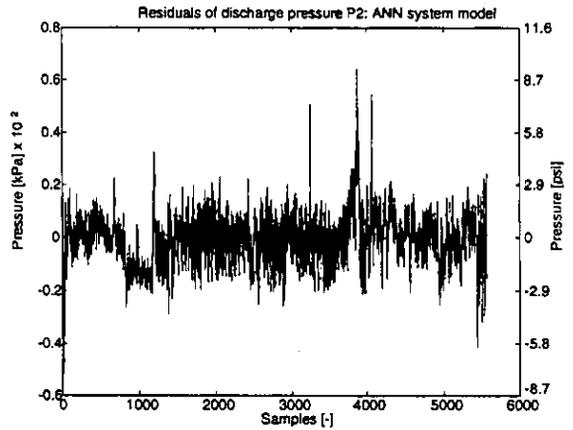
8d Residue test neural network for P2
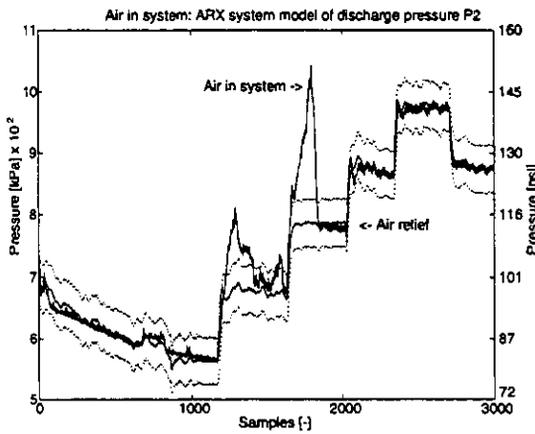
8e Fault detection with ARX model
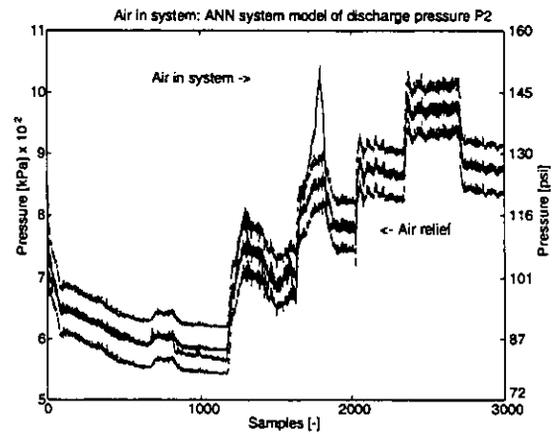
8f Fault detection with ANN

Figure 8 ARX and ANN learning model results

The threshold limits in Figure 8 are determined by using statistical methods. A 3-sigma limit (3 standard deviations) is used as a threshold value [Ljung, 1987].

The reliability of the ARX and ANN models (equation to forecast the outputs) is good, approximately 97 percent.

One condition for the generation of a reliable model is a healthy and dynamical dataset which is measured under optimal or good operation conditions of the chiller over a large working range. A second condition for system identification is the availability of a dataset with a fixed time step, because the regression models use values of previous time steps of the inputs and outputs to predict the actual value. Whenever a dataset meets these two conditions, black box modelling techniques can be applied.

The fault detection examples in figure 8 show how faults may be detected.

With black box models, it is possible to describe a process variable without a lot of physical knowledge about the considered process. However, fitted models are only valid for examined reference systems so each system must be identified anew. The results of the cross validations appear to be reasonable because the standard deviation and the average error of the residuals are comparable with the results of the models.

The quality of the available dataset is of vital importance for a reliable model. Black box modelling is a proper process method to predict the process variables of a chiller. Two types of black box modelling have been applied to the reciprocating chiller; ARX modelling and ANN modelling.

Black box models do not require a lot of physical knowledge to operate, although some physical knowledge of the process is necessary to obtain the best inputs and delays for the fitted model.

4.7 COMPARISON OF NEURAL NETWORKS WITH ARX MODELS

The use of ANNs in fault detection showed that ANNs can be very promising for modelling whole chillers. ANNs gives a slightly better result on system models and component models than the ARX models. The performance of the ANN components models would be poorer if the output at t-1 and t-2 were not be used as inputs for the model.

An advantage of ANNs is that, because of their nonlinear nature, they are more capable of predicting process variables in a nonlinear process.

Important for the training of ANNs is the choice of the initial state, the point where the ANNs start to learn. Wrong initial weight factors and offsets of a Neural Network often results in a impropriate learned network, though a proper set of input-output pairs is used. Crucial is to find the *global minimum* of the residue that is the smallest sum of squared errors. Often a wrong initial state causes the ANN to find a local minimum. In this case the ANN models required a longer time to learn than ARX models.

The auto-correlations of the residuals (table 2) of ANN models are better than the auto-correlations of the residuals of the ARX models. The higher auto-correlation of the ARX model indicates the presence of an independent process variable in the residue, whereas the best residuals consist of white noise, e.g. no dependence of the residue on process variables (figure 8). The non-linearity of the refrigerant system makes it difficult to fit an ARX model.

5 THE VARIABLE AIR VOLUME SYSTEM (VAV)

The VAV exercise is based on ARX models. A schematic description of the regarded VAV system is given in figure 9.
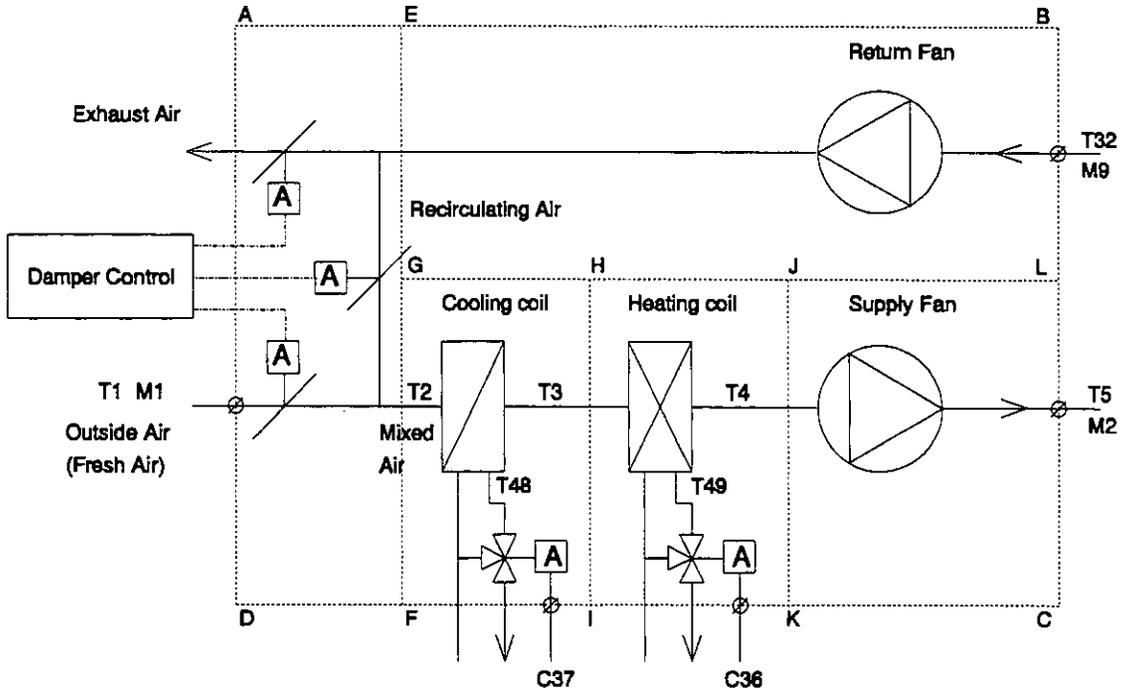
Figure 9 A simple PI-diagram of the AHU

The next models can be defined:

ABCD = System model AHU of the VAV system.

AEDF = Component model Mixing box.

GHFI = Component model Cooling coil.

HJIK = Component model Heating coil.

JLKC = Component model Supply fan.

EBGL = Component model Return fan.

The total number of inputs used for deriving the **system** model of the AHU is 24.

The **6 measured inputs** are:

- T1   Fresh air temperature

- T32   Air temperature before return fan

- M1   Fresh air mass flow rate

- M9   Return air mass flow rate

- C36   Heating coil demand (0-1)

- C37   Cooling coil demand (0-1)

The **18 derived inputs** are:

T1^2	T1*M1	T1*T32 T1*M9 T1*C36		T1*C37	
M1^2	M1*T3	2M1*M9	M1*C36	M1*C37	
T32^2	T32*M9 T32*C36	T32*C37			
M9^2	C36^2	C37^2			

The **historical inputs** are the values of the 24 inputs from the four previous time steps (t-1, t-2, t-3, t-4). To check the working condition of the AHU system the following **outputs** are predicted:

model 1: T5 Outlet air temperature after the supply fan.

model 2: M2 Supply air mass flow rate.
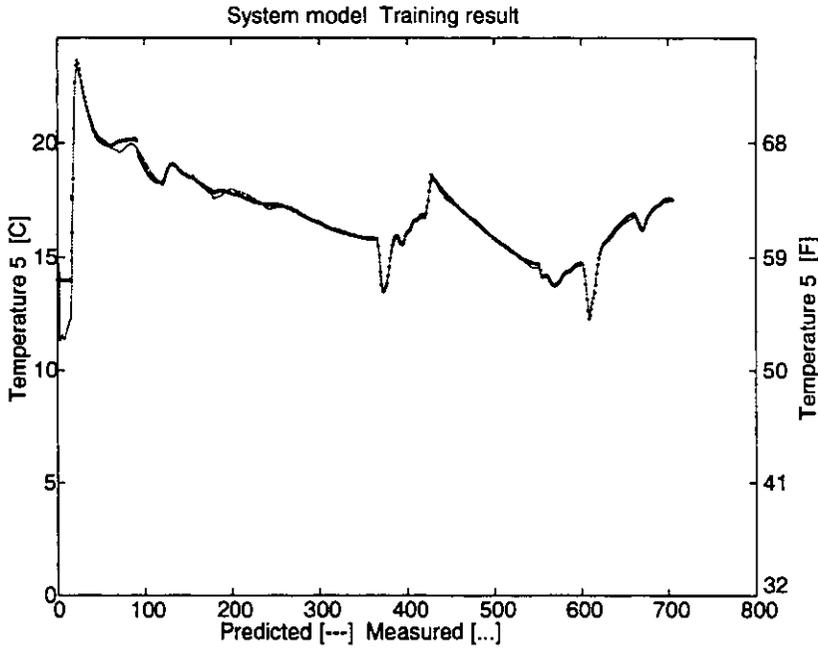
Figure 10a The results of training the ARX model of the supply air temperature T5

A healty data set is used to train the models. The training results of model 1 are presented in figure 10a. It shows that the model fits the system well. A faulty data set is used for fault detection. Figure 10b presents the results with the faulty data set. It shows that the forecasted temperature is much lower than the "measured" temperature. This indicates a malfuntion.

To discover the possible malfunctioning of a specific component, new models need to be made on the component level. The system model of the AHU can be divided into seperate component models. These component models are the Mixing box, Cooling coil, Heating coil, Supply fan and Return fan.

An example of a component model is presented for the cooling coil. The following procedure can be followed to make a good choice in selecting a system model or a component model:
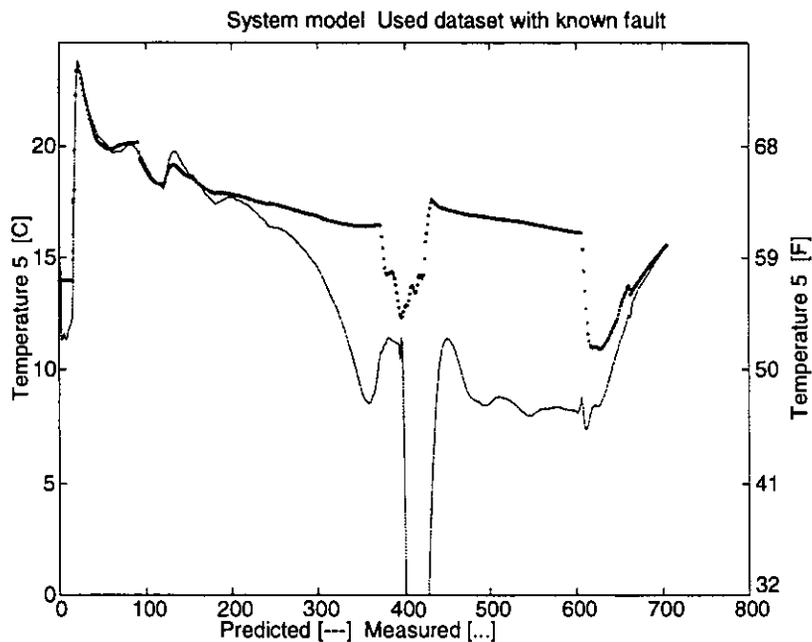
Figure 10b Comparison of the predicted supply air temperature with that measured in the presence of a fault

5.1 USE OF A SYSTEM MODEL

In normal circumstances this model will check the performance of the whole AHU system. If a deviation occurs between the forecasted and measured output as calculated by the system model, the system must be functioning incorrectly.

When the possibility of a malfuntion is raised, the next step is to use a component model to localize this malfuntion.

5.2 USE OF A COMPONENT MODEL

A component model will be used after the system model has detected a malfuntion. The reason for using component models is that a component model can pinpoint the cause of the malfuntion with greater accuracy than the system model. When the system model could not identify to a specific area of the malfuntion, then each component model must be used to find the cause of the malfunction.

5.3 COOLING COIL MODEL

The structure of the black-box model is dependent of which measured values are available. In the case of the cooling coil, two different types of models are possible to predict the air and water outlet temperatures T3 and T48. The Model of the cooling coil is presented in figure 11.

313
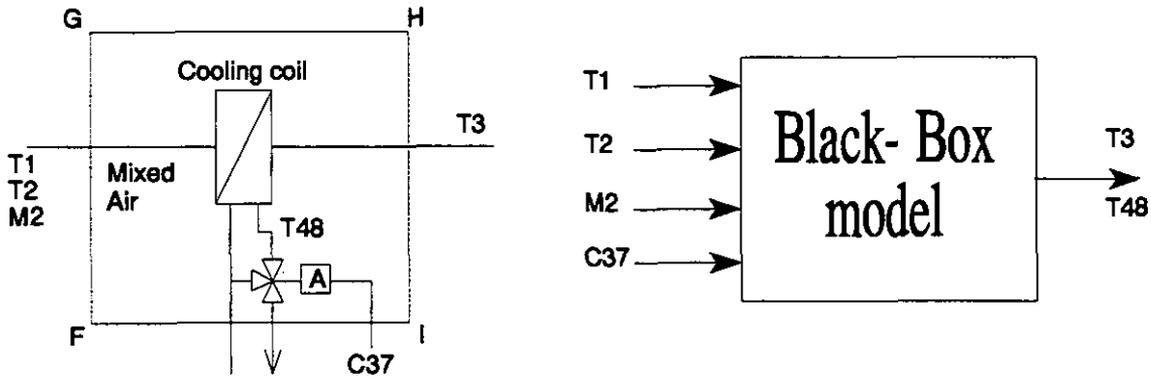
Figure 11 Cooling Coil model

The total number of inputs used for deriving the cooling coil model is 12.

The **measured values** used to derive the inputs for the second model of the cooling coil are:

- T1 Fresh air temperature

- M2 Air mass flow rate through the cooling coil

- T2 Air temperature before the cooling coil

- C37 Cooling coil demand (0-1)

The **8 derived inputs** are:

T1^2   M2^2   T2^2   C37^2

C37*T1   C37*T2   C37*M2   T2*M2

The **historical inputs** are :

The values of the 12 inputs from the five previous time steps (t-1, t-2, t-3, t-4, t-5).

The **outputs** predicted by the cooling coil are:

model 1: T3 Outlet air temperature of Cooling coil

model 2: T48 Outlet water temperature of cooling coil

The training results in figure 12 show that the model fits the cooling coil well.

314
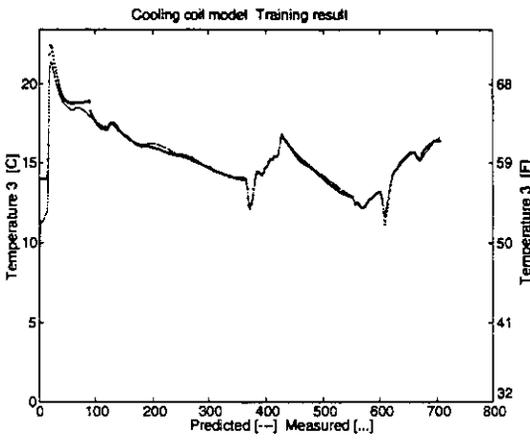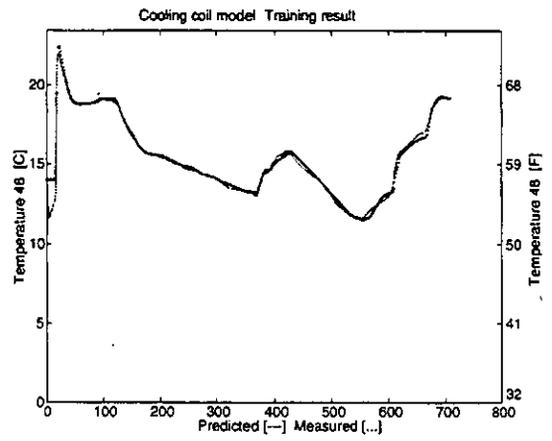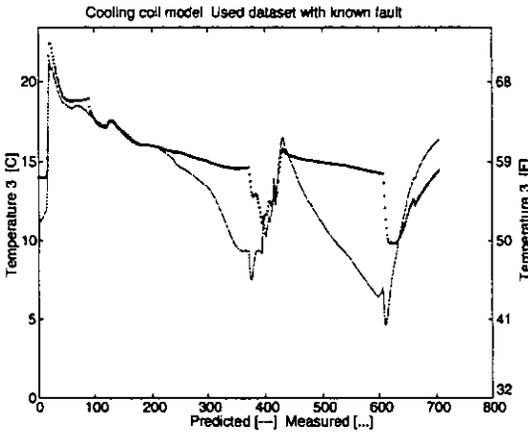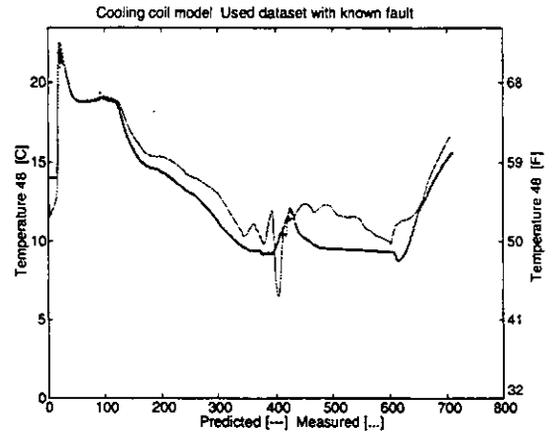
Figure 12 Training results of the component model

5.4 DATA ANALYSIS AND FAULT DETECTION

To detect a fault in the VAV system, the predicted output of the model must be compared to the measured output of the system. If the value of the measured output is not within the bounds of the predicted output, there can be an error in the system, or an incorrect model has been used. If the predicted output of the model in the training period is quite similar to the measured output of the VAV system it can then be concluded that the model fits the process well.

5.5 RESULTS OF THE VAV SYSTEM MODELS

Figure 10 presents the system model forecasting results and the predicted and "measured" air temperature after the supply fan T5. It shows that the predicted temperature is much lower than the measured value. There is though a connection between those two values. The connection is that the predicted temperature has the same trend as the measured temperature but the prediction reacts much more strongly, which causes a lower output

temperature. This difference in temperature indicates a fault in the cooling coil section. This symptom can be explained by three possibilities.

Situation 1: <u>No water mass flow through cooling coil.</u>

This situation can take place when the coil is blocked while the valve is opened to the maximum. With this given valve position the model will forecast a lower output temperature because of the expected larger water mass flow rate.

Situation 2: <u>High chilled water temperature.</u>

The temperature of the supply cooling water is already near the temperature of the outside air before it even enters the coil. This is due to a defect in the cooling water circuit.

The difference between cooling coil supply water temperature and inlet air temperature is not large enough to cool the air sufficiently.

Situation 3: <u>A fouled cooling coil.</u>

In this situation the cooling capacity of the coil will decrease because the heat exchange between the water and the air decreases. To compensate for this decreased cooling capacity, the actuator signal will be increased, dependent on the fouled level of the coil and the valve will be opened further, thus increasing flow of chilled water. The model was trained using the mass flow rate through the coil. Taking into account the new valve position, the model will forecast a larger temperature drop than really occurs.

5.6 RESULTS OF THE VAV COMPONENT MODELS

Figure 12 present the results of the cooling coil component model. Figure 12a and 12b presents the training results of T3 and T48. The models fit the cooling coil well. Figure 12c and 12d presents the forecasted results of T3 and T48. The predicted supply air temperature T3 is much lower than the measured value. The predicted output water temperature T48 is higher then the measured value. This shows a decrease of transported heat to the cooling coil. This proves that the malfunction takes place in the cooling coil section. The most relevant diagnosis is a fouled cooling coil, as described for situation three.

A black box model for the AHU of the VAV system was used. The *system* model was used to predict the supply air temperature (T5) and the supply air mass flow rate (M2).

The predicted output was compared with the "measured" output. The diagnoses of the predicted results with the faulty dataset pointed to a malfunction in the cooling coil section.

To pinpoint the causes of the fault, a component model of the cooling coil section was used. The *component* model of the cooling coil section pinpointed the fault.

This example is based on a simulation exercise of a VAV system. A lot of sensors are available. For practical applications sensors should be used that are standard available in a BAS. Depending on the faults to be detected, additional sensors could be required on component level.

The exercises are based on models of healthy operation. No models were used for faulty operation.

6 CONCLUSIONS

The presented examples illustrate the possibilities of black-box models for fault detection. The use of system models makes it possible to detect faulty behaviour of a system, whereafter component models can be used to locate the defective component.

For non-linear systems, ANN models fit better than ARX models. This is also illustrated by the auto-correlations of the residuals (table 2), which are better for ANN models than for ARX models.

ARX and ANN can be used as black-box identification techniques if a dataset is available that meets the two following conditions:
1) The dataset must be healthy and dynamical, as well as measured under optimal or good operation conditions of the chiller in a large working range.
2) The dataset has got to have a fixed time step

The goal of system identification by using black-box models is to enable fault detection in HVAC systems. Further development of the black-box modelling techniques is necessary to realize reliable models that can be used in practical applications. Since the design of models for fault detection in this stadium is a dynamic research proces, causing repeated adjustment of the models, it is difficult to present up to date results; therefore, the results of the models as presented in this paper might already be improved.

Apart from fault detection, fault diagnosis is also a point of research. Black-box models can be used in model-based reasoning systems (Davis and Hamscher 1987, van Soest 1993), which are capable of precise localization of defective components. Combination of black-box fault detection with reasoning systems will increase the reliability of FDD.

ACKNOWLEDGEMENTS

REFERENCES

Curtiss, P.S., J.F. Kreider, M.J. Brandemuehl. 1993. *Adaptive control of HVAC Processes using Predictive Neural Networks*. ASHRAE Transactions 99(1)

Davis, R. and W. Hamscher. 1987. *Model-based reasoning: Troubleshooting, in Exploring Artificial Intelligence; survey talks from the National conferences on Artificial Intelligence*, Morgan Kaufman Publishers, San Mateo, pp 297-346

Dexter, A.L. and S.J. Hepworth. 1993. *A Comparison of Fuzzy and Neural Methods of Detecting Faults in an Air-handling Unit*. IEA Annex25 report, Oxford.

Duyvenvoorde, A.M. van. 1993. *Experimental fault detection and diagnosis at a laboratory chiller*. IEA Annex 25 report, TNO Building and Construction Research, Delft, The Netherlands.

Fogel, B.F. 1991. *An Information Criterion for Optimal Neural Network Selection*. IEEE Transactions on Neural Networks, Vol. 2, No. 5.

Galen van P.J.M., H.C. Peitsman, A.M. van Duyvenvoorde. 1993. *Fault detection to a real chiller plant in an office building using black model identification*. IEA Annex 25 report, TNO Building and Construction Research, Department of Indoor Environment and Systems, The Netherlands.

Kohonen, R. and J. Hyvärinen. 1991. *Real time simulation of HVAC systems for building optimisation, fault detection and diagnosis*. IEA Annex 25 workplan, Technical research centre of Finland, Laboratory of heating and ventilation, Espoo, Finland.

Kreider, J.F. et al. 1992. *Expert systems, neural networks and artificial intelligence applications in commercial building HVAC operations*. Automation in Construction 225-238, Elsevier Science Publishers.

Ljung, L., 1987. *System Identification - Theory for the User*. Prentice Hall, Englewood Cliffs, New York, USA.

Nelson, Marilyn McCord and W.T. Illingworth. 1991. *A Practical Guide to Neural Nets*. Addison-Wesley Publishing Company, Inc.

Peitsman H.C., A.M. van Duyvenvoorde, W.O.J. Böttger. 1994. *Application of ARX-models and Neural networks to a reciprocating chiller for fault detection*. IEA Annex 25 report, TNO Building & Construction Research, Cambridge, MA, U.S.A.

Seem, J.E. and J.E. Braun. 1991. *Adaptive methods for real-time forecasting of building electrical demand.* ASHRAE Transactions 97(1).

Soest, D.C. van. 1993., *Modelling for model-based diagnosis*, Thesis University of Twente, Enschede,The Netherlands.

Typical Faults of Air-Conditioning Systems and Fault Detection by ARX Model and Extended Kalman Filter

Harunori Yoshida, Dr. Eng.
Member ASHRAE

Tatsuhiro Iwami

Hideki Yuzawa

Masami Suzuki

ABSTRACT

Since faulty operation of heating, ventilating, and air-conditioning (HVAC) systems is detrimental to energy conservation, and maintenance experts are no longer able to detect faults due to the sophistication of current air-handling units (AHUs), automated fault detection and diagnosis (FDD) is increasingly important. In the present study, the results of a survey about typical faults that are commonly encountered in air-handling systems are summarized, and two methods of finding abrupt faults are described.

To investigate the development of automated fault detection schemes, two methods to detect an abrupt fault are tested, and the effectiveness of the methods is analyzed. Both are based on a mathematical model of system dynamics. The first one is an autoregressive exogenous (ARX) model and the second is based on an extended Kalman filter. It is shown that faults that are difficult to detect by a simple limit checker method can be detected in both cases on the basis of computer simulation by HVACSIM+.

INTRODUCTION

In most HVAC systems, highly reliable operation or a low probability of malfunction are not important issues. Only specific types of facilities, such as building spaces housing telecommunication systems or mainframe computers for processing bank business, laboratories for radioactive material handling, etc., are exceptions that require highly reliable operation in order to avoid substantial economic loss and/or environmental damage. This is because the main purpose of most HVAC systems is to provide human comfort, and a breakdown of the system does not result in damage as serious as would result to those facilities previously mentioned. Although occupants will complain about the undesirable thermal environment due to the breakdown, they are generally reconciled to it and will reluctantly wait for the system to recover. According to the results of one survey, occupants can wait for 30 to 60 minutes without many complaints (OSTEC 1990).

Summarizing the results, it can be said that prompt detection of the cause of a breakdown and prompt repair are more important than high reliability. In other words, system availability, which is defined as the ratio of operable duration to the lifespan, is more important than reliability in most HVAC systems.

Under these circumstances, there is a strong need to develop an automated and real-time fault detection and diagnosis (FDD) system. In addition to mechanical malfunctions, the detection of undesirable or faulty operation of HVAC systems in terms of energy conservation is also an important issue. One of the main possibilities for implementing the system is in using the surplus power of existing computers in HVAC control systems or building energy management systems (BEMS) or at least connecting an FDD system to existing systems.

There is a mutual understanding about the requirements of FDD, and a great deal of research work has been done within Annex 25 of the International Energy Agency (IEA 1995). Since there are many variations in system configurations and components of HVAC systems, it seems unlikely that research to seek a general FDD method will be successful without focusing on system characteristics. For this purpose, several reference systems were defined in Annex 25—an air-handling unit (AHU), a heating system, chillers and heat pumps, and a thermal storage system. This paper is based on the reports written by the authors on activities related only to AHU reference systems. In the present study, results of a questionnaire survey about typical faults that are encountered in AHU systems are summarized, and two methods of finding faults that occur suddenly are proposed and their capabilities are investigated.

TYPICAL FAULTS OF REFERENCE SYSTEM

In Annex 25, an AHU reference system was proposed and refined by one group. The performance of all the basic components is described in detail, including control strategies. The configuration of the system is shown in Figure 1. The system has a standard AHU unit with a supply air volume controller, three automatically controlled air dampers for the economizer cycle, and, for each room, one terminal unit with a VAV box and a reheater coil (three units altogether). Although the reference system may vary in detail from country to country, it was accepted

Harunori Yoshida is an associate professor in the division of Global Environment Engineering of the Faculty of Engineering at Kyoto University, Kyoto, Japan. Tatsuhiro Iwami is a manager at Sinryou Corporation, Yokohama, Japan. Hideki Yuzawa is a mechanical engineer at Nikken Sekkei Ltd., Tokyo, Japan. Masami Suzuki is a researcher at Sinryo Corporation Research and Development Center, Tsukuba, Japan.
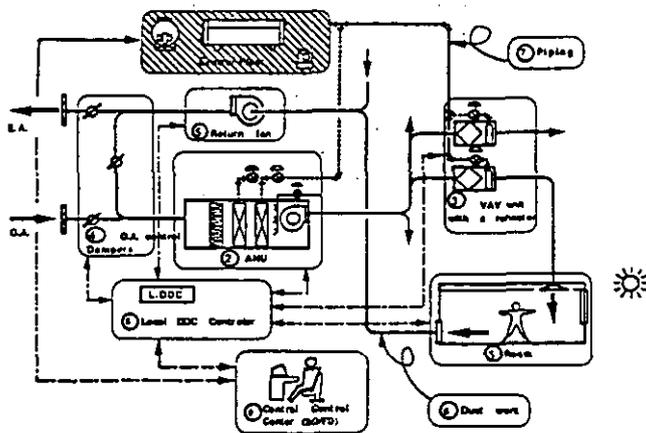
Figure 1 *Configuration of AHU reference system used for questionnaire.*

as the common reference system among participants from more than 12 countries.

Since there are many different faults related to AHU systems, it is not cost-effective to find all possible faults. Before a FDD system can be developed, knowledge about typical or important faults is necessary. Therefore, a survey was conducted by the authors. A questionnaire was prepared to collect information on a range of faults, from design faults to user-level faults (such as a thermostat blocked by furniture), using the reference system diagram. The consequences of a fault were classified into four cases: those relating to poor environment, energy consumption, reference or design value, and physical damage. Prominent experts from three professional fields—design engineers, fabricating engineers, and maintenance engineers—were asked to be the respondents.

Table 1 shows the data that were collected for 10 important faults ranked by all experts. Many of them are faults at the stage of design and fabrication. In other words, most common faults can be eliminated by careful design and commissioning. Therefore, using

Figure 2 *Reasons given for fault importance.*

an FDD system at the commissioning stage is preferable. Since the importance of a fault differs slightly among professions, it is important to establish for whom an FDD system is provided. Furthermore, it should be understood that pure mechanical faults or sophisticated operational faults can be hidden behind these trivial faults.

In the questionnaire, the grounds for judging a particular fault to be important were investigated by selecting one of several predefined reasons: (1) degradation of environment, (2) energy conservation, (3) relation to damage, (4) frequency of a fault, (5) difficulty of detection, (6) difficulty in recovering, and (7) expenditure required to fix. The results are shown as percentages in Figure 2. Environmental degradation and energy conservation were selected as the most important grounds. In developing an FDD system, one should refer to the results to eliminate less important faults from the list of those to be detected.

To investigate the nature of the pure hardware faults in more detail, 10 important faults for the each of the seven reasons mentioned were selected, and then only the hardware malfunctions were listed in Table 2 with rank. A simple maintenance

TABLE 1   Ten Important Faults Selected By All Professions

Rank	Point	Subsystem	Process Variable Deviation	Component	Component Defect	Stage
1	39	room	poor air quality	occupants	smoking	U
2	38	piping	water leakage	piping	condensation due to improper thermal insulation	F
3	37	room	room air temperature deviation	occupants	excessive internal heat generation	U
4	35	room	room air temperature deviation	air diffuser	inadequate air-flow rate	F,C
5	31	VAV unit	too much or less air volume	VAV unit damper	failure in adjustment	F,M,C
6	29	AHU	excessive pressure deference across an air filter	air filter	being stuffed	M
7	27	air duct	abnormal noise or vibration	duct-work	insufficient noise control	D,F
7	27	room	room air temperature deviation	air,diffuser	inadequate positions of diffusers	D,F
7	27	local DDC control	false opening signal to a VAV unit	room air thermostat	improper location	D,F
10	23	piping	room air temperature	piping	insufficient flow rate due to	F,C

The abbreviations in the column of labeled stage (A) a failure in the control algorithm, (C) a failure in commissioning, (D) a failure in the design stage, (F) a failure at the construction or fabrication stage, (M) a failure at the maintenance stage, (S) a failure in the control hardware and software, and (U) a failure caused by users.

TABLE 2   List of Hardware Faults
(Rank means the rank in 10 faults selected for each reason)

Reason Why a Fault is Important	Rank	Hardware Fault
1) Environmental degradation and occupant complaints	8	—air filter stuffing
	9	—humidifier malfunction
2) Increased energy consumption	1	—air filter stuffing
	2	—fouling on cooling coil fins
	4	—OA damper malfunction
	10	—scaling in coil tubes
	10	—air leakage through ductwork
3) Serious secondary damage	2	—no alarm due to computer malfunction
	3	—water leakage
	4	—false alarm
	7	—accumulation of dust inside ductwork
	7	—malfunction of fan motor
4) Frequent occurrence	1	—air filter stuffing
5) Difficult detection	2	—no alarm due to computer malfunction
	6	—accidental fire damper shutting off
	7	—OA damper malfunction
	7	—water strainer stuffing
	9	—accumulation of dust inside ductwork
	9	—scaling in coil tubes
6) Lengthy repair time	8	—malfunction of fan motor
	9	—accumulation of dust inside ductwork
7) Costly repair	8	—accumulation of dust inside ductwork

fault, the blocked air filter, is most common; however, mechanical faults (such as coil and damper malfunction) are seen widely. Judging by the number of faults, the most important are those associated with energy loss, secondary damage, and difficulty in detection. It is significant that the number of faults in category 5, which are difficult to detect, is larger than that of the others. This may suggest that an FDD system should tackle these types of faults.

According to the results of our survey, the authors conclude that the aims of an FDD system should be (a) to predict the possible deterioration of materials or components as quickly as possible, (b) to detect inefficient operation in terms of energy loss, and (c) to find the fault that cannot be detected even by an expert (for example, a fault in computer control software), faulty operation that results in energy loss, mechanical faults that have no immediate effect on room environmental condition, etc.

DETECTION OF SUDDEN FAULTS

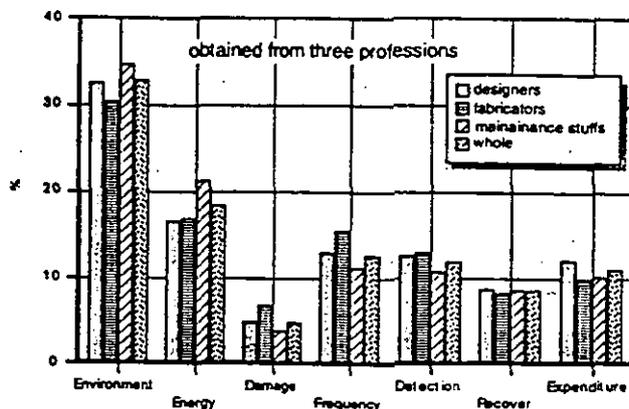Many kinds of generic methodologies for FDD are available, but there is a wide variety of approaches, aims, and performances: some are qualitative and some are quantitative, some are based on a physical model and some on a black-box model, etc. (IEA 1995). However, since faults in HVAC systems include all sorts of the kinds as described above, we cannot arbitrarily select one FDD method as the best.

A fault may develop slowly or suddenly: cooling coil clogging and control device malfunctioning are examples of each category. In the former, a fault can be detected by checking whether a change in component or system performance has or has not occurred. In the latter, a fault can be detected by checking at regular intervals whether a value exceeds a fixed limit or not or whether the dynamic characteristics of a component or a subsystem change or not. In some cases, a fault may already exist at the time of completion of building construction. Contradictory motion of air dampers used for an economizer cycle is an example. This results from faulty installation that was not detected during commissioning. For detecting this kind of fault, the reference performance must be known. This indicates that we must choose an appropriate FDD method by considering the nature of the faults, especially when they occur and how they develop.

In this paper, the authors concentrate on the detection of a sudden fault. The easiest detection method is to measure an appropriate state variable and check whether it exceeds a higher or lower limit. Room air temperatures and supply air temperatures are often checked this way. However, a fault cannot be detected by this simple method when the state variables stay within the limits even when the fault is present. Furthermore, there is no objective way of determining the values of the limits.

To detect a sudden fault, more sophisticated methods can be used. One of them is based on the dynamic performance, that is, a fault is detected by checking the difference between the measured state variables and those obtained from a dynamic model (Figure 3). This method is well developed and many research studies have been reported (Willsky 1976; Isermann 1984); however, relatively little work is reported in the field of air-conditioning engineering. Usolo et al. (1985) deal with the detection of an abrupt fault (a change in the bias on the output of a temperature sensor) using an extended Kalman filter methodology. The parameters of the model are assumed to be known even though in many components installed in real buildings they are not known. In this study, therefore, two models—an autoregressive and exogenous (ARX) model and an extended Kalman
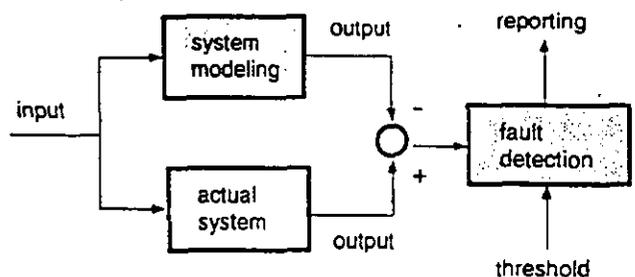
Figure 3   Fault detection using a model.

filter model in which parameters are estimated—are used to detect faults in a VAV unit and an AHU cooling coil subsystem.

The effectiveness of the methods is examined using simulation, where the state variables of the reference system are generated by the computer code HVACSIM+ (Clark and May 1985; Clark 1985). HVACSIM+ was developed by a U.S. organization and some modifications were made to allow faults to be simulated and to stabilize the numerical computation.

THE ARX MODEL APPROACH
TO FAULT DETECTION

The following fault was implemented: a VAV unit has a malfunction that causes the opening ratio to remain at the value just before the fault. The dynamic performance of a VAV unit including the controller is modeled by an ARX model as follows:

$$y_n = -\sum_{i=1}^{p} a_{n-i}\ y_{n-i} + \sum_{j=0}^{q} b_j z_{n-j} + v_n \qquad (1)$$

where

y = air mass flow rate through the VAV unit,

z = room air temperature signal from the temperature sensor,

v = random variables (normally distributed),

a = autoregressive parameters (order p), and

b = exogenous parameters (order q).

The variables y and z must both be measured in actual situations. A disadvantage of the present method is that it is necessary to measure y using an anemometer. If the air-conditioning load of a room can be measured, it can be used instead of y but this is more difficult to measure. However, since recent VAV units are often equipped with an anemometer, this assumption can be acceptable.

Only faults in the part of the plant between the input and output of the model can be detected (for example, if the control signal is used as the output instead of the mass airflow rate, faults in the damper actuator cannot be detected).

The model relating output and input or extra (exogenous) variables is linear. In a strict sense, the performance of a VAV unit is nonlinear; however, the approximation can be used for the present purpose, namely, fault detection. The structure of the model is defined, but the parameters (a and b) and the order of the model (p and q) are unknown. These have to be identified using simulated variables; y, z, and the recursive least-squares algorithm can be applied for the identification. The model order (p and q) can be determined by minimizing AIC (an information criterion), which is defined as follows (EITIJ 1993; Akaike 1972):

$$AIC = N\ln\sigma_y^2 + 2(p+q), \qquad (2)$$

$$\sigma_y^2 = \sum_{k=n}^{n-N+1} (\hat{y}_k - y_k)^2 / N \qquad (3)$$

where

$\hat{y}$ = estimated y based on the model and

N = data window length used for model identification.

Before performing the on-line identification, the value of N must be empirically determined and the model order (p and q) must be selected as the optimum values by trial and error using test data sets. During the operation of the HVAC system, the parameters are identified and updated at every sample time. If a fault occurs abruptly, the difference between the predicted output and the measured output becomes very large; then, by evaluating the value S (defined by Equation 4) using the latest data obtained, the occurrence of a fault can be statistically tested because the value follows a χ^2 distribution.

$$S = L\sigma_y^2 = \sum_{k=n}^{n-L+1} (\hat{y}_k - y_k)^2 \qquad (4)$$

where L is the data window length for fault detection. It should be understood that the estimated parameters will gradually adapt to the fault condition and detection will no longer be possible.

The simulation was performed as follows:

- The simulation is carried out under cooling operations.

- The internal cooling load decreases during lunch.

- The time at which the fault occurs is set at 10:00 and 14:00.

- Three typical days with different weather conditions are simulated: clear, half-cloudy, and cloudy days that represent high, medium, and low cooling loads are used for the simulation to check the performance of the fault detection system under different operating conditions.

In Figure 4 the simulation results (sensor output of room temperatures, air mass flow rate, VAV opening rate, zone temperature controller output) are shown for the case when the fault occurs at 10:00. It will be seen that the fault cannot be detected by a limit-checker method because the room air temperature variations are very little for a long time after the fault occurrence. This is also observed in the case of a fault at 14:00 (not shown).

The model order was ARX($p = 3$, $q = 7$). The selected sampling interval was 2.5 minutes, and the values of the window length were $N = 72$ and $L = 8$. The optimal values were found using test data sets. In Figure 5 the effectiveness of the fault detection scheme is shown for three days with different weather conditions: clear, half-cloudy, and cloudy, which represent high, medium, and low cooling loads. The dotted lines are the 99.5% and 95.0% confidence thresholds used in the statistical test. (Faulty behavior seems to occur before fault occurrences. The reason is that the time to evaluate S is defined by the time of the first datum, n, not by the last datum, $n - L + 1$, in Equation 4; therefore, faulty behavior occurs after the fault occurrences.) It is concluded that the method performs well if the cooling load is not too small. This is because the VAV unit is not active enough to detect the fault under the small load conditions.

4

THE EXTENDED KALMAN FILTER
APPROACH TO FAULT DETECTION

The air-handling unit of the reference system supplies air at a constant temperature under control of a PI controller. The main components of the subsystem that make up the control loop are a cooling coil, a water control valve, a PI controller, and a temperature sensor (Figure 6). In the present study the aim is to detect a malfunction in this control loop. The variables are the signals from the temperature sensor and the control signal to the valve (valve opening). Both are easy to monitor.

By regarding each component as a linear system, the entire control loop can be considered to be linear. In Figure 7, the characteristics of each component are shown as Laplace transform transfer functions. The governing equations are written as follows:

$$
\left.
\begin{aligned}
\dot{\phi} &= \frac{1}{T_I}\varepsilon_s \\[2mm]
\dot{\varepsilon}_s &= -\frac{1}{\tau}\varepsilon_s + \frac{1}{\tau}\varepsilon_o \\[2mm]
\dot{\varepsilon}_o &= \left(\frac{C_w K_V}{K_c}\cdot K_p\right)\Delta\theta_w\phi + \left(\frac{C_w K_V}{K_c}\cdot K_p\right)\Delta\theta_w\varepsilon_s \\[2mm]
&\quad -\left(\frac{K_a}{K_c}\right)m_a\varepsilon_o + \left(\frac{K_a}{K_c}\right)m_a\varepsilon_i + \left(\frac{C_w m_w}{K_c}\right)\Delta\theta_w
\end{aligned}
\right\}
\tag{5}
$$

where

ϕ	=	defined by the first equation of Equation 4,
ε_s	=	$\theta_s - \theta_{Set}$,
ε_o	=	$\theta_o - \theta_{Set}$,
ε_i	=	$\theta_i - \theta_{Set}$,
$\Delta\theta_w$	=	$\theta_{w,o} - \theta_{w,i}$,
θ_i	=	inlet air temperature,
θ_o	=	outlet air temperature,
θ_s	=	air temperature measure by the sensor,
θ_{Set}	=	set temperature of supply air,
$\theta_{w,i}, \theta_{w,o}$	=	inlet and outlet water temperature,
C_w	=	specific heat of water,
K_a	=	thermal capacity of coil,
K_c	=	integral coefficient of PI controller,
K_p	=	proportional coefficient of PI controller,
K_V	=	proportional coefficient relating water mass flow rate to the control signal,
m_a	=	air mass flow rate,
m_w	=	water mass flow rate at 50% valve opening,
T_I	=	integral time of PI controller,
τ	=	time constant of temperature sensor
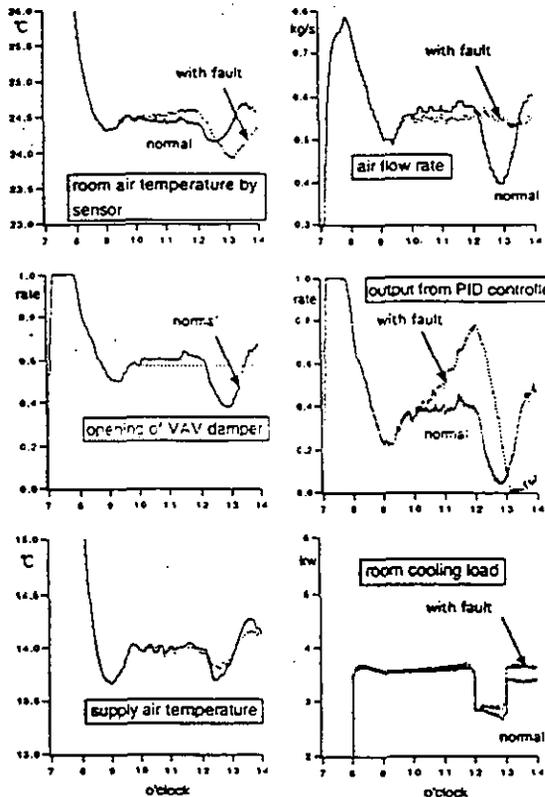
Figure 4 *Simulated state variables with fault occurrence at 10:00.*
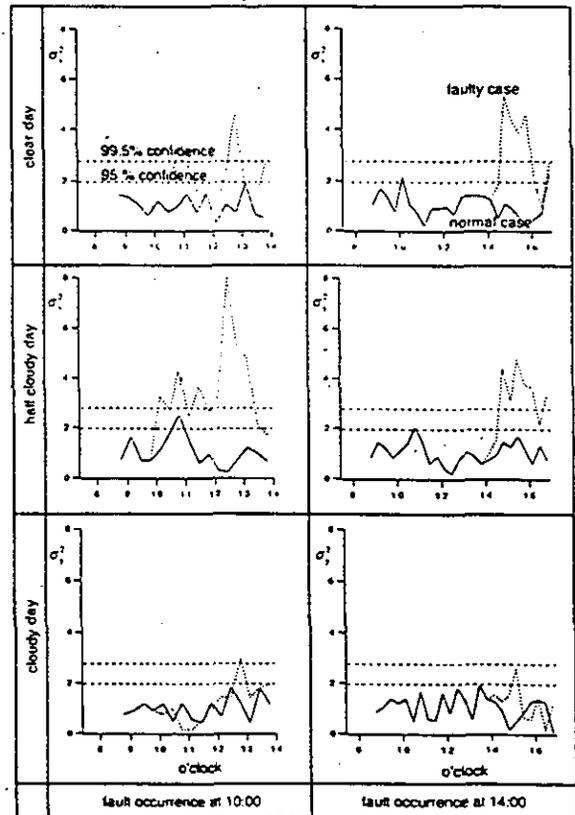
Figure 5 *Performance of the fault detector for the six cases.*

PI controller

Figure 6   System diagram.

Figure 7   Block diagram of the cooling coil control loop.

by defining state variables (x), input (u), observed variables (y), and parameter vector (a) as follows:

$$x = (x_1, x_2, x_3)^T = (\phi, \varepsilon_S, \varepsilon_o)^T.$$

$$u = (u_1, u_2, u_3)^T = (\Delta\theta_w, m_a, \varepsilon_i)^T$$

$$a = (a_1, a_2, a_3, a_4, a_5, a_6) = \left(\frac{1}{T_I}, \frac{1}{\tau}, K_p, \frac{C_w K_v}{K_c}, \frac{K_a}{K_c}, \frac{C_w m_w}{K_c}\right). \tag{6}$$

The parameters can be determined from the component performance; however, in actual situations it is difficult to obtain values for them. In addition, since they tend to change with time, they should be taken as unknown parameters. By defining a new vector z, the state space equation can be written as follows:

$$z = (x:a)$$

$$\left.\begin{array}{l} \dot{z} = f(z, u) + w \\ y = h(x) + v \end{array}\right\} \tag{7}$$

where

z = combined new state variable,

w, v = white noise vectors,

y = vector of the measured data, the valve control signal of opening c and ε_S, namely,

$$y = (y_1, y_2)^T = (c, \varepsilon_S)^T = (K_p(\varepsilon_S + \phi), \varepsilon_S)^T.$$

This equation must be further converted into a discrete-time form to utilize the simulation results. The final equations are nonlinear; however, they can be solved by applying the extended Kalman filter algorithm. Fault detection is performed using the same method as that used for the ARX model. Further details of Equation 7 and the process used to solve the equations are reported in Yoshida et al. (1995).

Two types of fault were implemented:

- The actuator of the control valve of the AHU becomes out of order at 11:45 and the current valve position is then held constant throughout the remainder of the test.

- The temperature sensor becomes out of order at 11:45 and the signal to the controller is held constant at the same value throughout the remainder of the test.

Two cases were tested. In the first case, the sampling interval is 2.5 minutes and the evaluation interval is 15 minutes. In the other they are 5 and 30 minutes, respectively. The results are shown in Figure 8. The shorter sampling time is better for fault detection. The fault can be detected in 30 minutes in the first case, but it cannot be detected in the second case. This is because no significant change in state variables takes place even if the temperature sensor is damaged. When the control signal changes to a considerable extent due to the integral action of the controller, the detection is possible, but this can happen just by chance. This is a disadvantage of the present method.

DISCUSSION

Even though the aim of AHU systems is to provide human comfort, some faults that result in discomfort are difficult to define as true faults. It is easy to eliminate them from the list of faults to be detected, but if this is done, some important aspects of the problem may disappear. As defining typical faults is the starting point for developing a fault detection system and we tend to arbitrarily choose a fault to test the performance of the FDD method, we must always consider what faults are most common and important.

Although it was mentioned that rapid fault detection is more important than constructing a highly reliable system in the case of AHU systems, prompt detection of a malfunction is not easy, even for experts in the maintenance of HVAC systems, for several reasons.

1. Modern systems are so sophisticated that a malfunction may not necessarily cause an easily detectable defect. For example, in an air-conditioning system with VAV units and a constant supply air temperature controller, the malfunction of the controller may not affect the room air temperature directly. The reason is that VAV units can compensate for the inadequate supply air temperature by controlling the air volume.

2. The installation of direct digital control (DDC) systems and computer control systems has become common, and most components are electronic products involving integrated circuit technologies. As a result, the system tends to become a black box. Under these situations it is difficult for normal maintenance engineers to detect and diagnose a malfunction.

3. Since the number of components of an HVAC system has increased along with the recent rapid expansion of the building size, engineers cannot devote attention to each component. In other words, careful maintenance cannot be achieved any more.

4. Sophisticated control strategies based on computer software are now commonly used as well as standard control algorithms such as PID. This can cause malfunctions to occur in the software itself that cannot be detected by engineers utilizing their experience or know-how.

Recognition of these issues emphasizes the necessity for developing automated and real-time FDD systems. Future work must assess the cost-effectiveness, and field tests are required.

In the present study, two methods of detecting abrupt faults have been investigated. The first method, which uses a black-box model, can be applied to any system if input and output data are available. However, since the parameters have no physical meaning, the method is restricted to fault detection only and it requires the optimization of the model order, which is a difficult
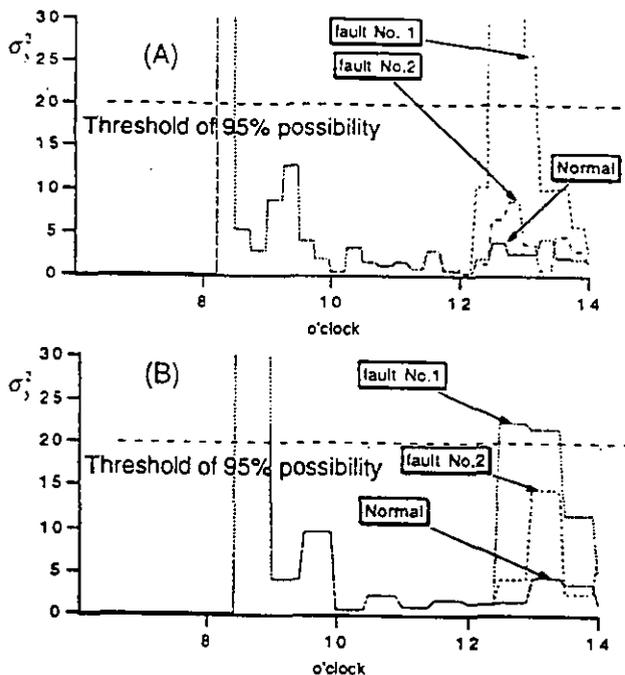
(A) Time step is 2.5 minutes and fault-checking interval is 15 minutes

(B) Time step is 5 minutes and fault-checking interval is 30 minutes.

Figure 8   Performance of the extended Kalman filter method.

process to automate. These are the disadvantages of the method. The model used in the second method is based on the physical structure of the plant and therefore fault diagnosis is possible.

There are many weak points associated with our methods. For example, an anemometer is required for the VAV fault, and the effectiveness of detecting a fault using the extended Kalman filter approach depends on the type of fault. These issues will be addressed in a future work by the authors.

CONCLUSIONS

In the present study, the results of a survey about typical faults that are commonly encountered in AHU systems were summarized and two methods of finding abrupt faults were tested.

In most HVAC systems, highly reliable operation or a low probability of malfunction are not important issues. However, since faulty operation of HVAC systems is detrimental to energy conservation and because maintenance experts are no longer able to detect faults (mainly because the AHU systems have become too sophisticated or too computerized), automated fault detection and diagnosis (FDD) is now important.

According to the results of the survey, many faults originate at the design and commissioning stage and most of them are very simple, such as the clogging of filters. Important faults can be obscured; therefore, classification of faults is necessary in order to find the important faults to develop an efficient FDD system. The survey conducted by the authors should provide a good reference.

Two methods to detect a sudden fault have been proposed and the effectiveness of the methods has been tested. Both are based on a mathematical model of system dynamics. The first one is an autoregressive exogenous (ARX) model, and the second is based on an extended Kalman filter. Using computer simulation by HVACSIM+, it has been shown that faults that are difficult to detect by a simple limit checker can be detected using both methods. Further study or field testing is required to ensure the methods are feasible in practice.

ACKNOWLEDGMENTS

REFERENCES

Akaike, H., et al. 1972. *Statistical analysis and control of dynamic systems* (in Japanese). New York: Science Publishing Co.

Clark, D.R. 1985. *HVACSIM+ building systems and equipment simulation program—Reference manual.* NBSIR 85-2996. Gaithersburg, Md.: National Institute of Standards and Technology.

Clark, D.R., and W.B. May. 1985. *HVACSIM+ building systems and equipment simulation program—User's guide.* NBSIR 85-3243. Gaithersburg, Md.: National Institute of Standards and Technology.

EITIJ. 1993. *Digital signal processing handbook* (in Japanese). Tokyo: OHM Publishing Co.

Hyvarinen, J., ed. 1995. *IEA ANNEX 25, Building optimization and fault diagnosis source book* (1st draft). Paris: International Energy Agency.

Isermann, R. 1984. Process fault detection based on modeling and estimation. *Automatica* 20(4): 387-404.

OSTEC. 1990. Report of maintenance system with artificial intelligence for intelligent buildings (in Japanese). Osaka: Osaka Science and Technology and Engineering Center.

Usolo, P.B., et al. 1985. An innovation-based methodology for HVAC system fault detection. *J. of Dynamic Systems, Measurement, and Control, Transactions of ASME* 107: 284-289.

Willsky, A.S. 1976. A survey of design methods for failure detection dynamic systems. *Automatica* 12: 601-611.

Yoshida, H., et al. 1995. Fault detection of HVAC systems by an extended Kalman filter. Tsinghua-HVAC-'95, 2nd International Symposium on Heating, Ventilation and Air Conditioning. Beijing: National Natural Science Foundation of China.

An application of fault detection observers

AN25/CH/080493/1

Peter Sprecher, Landis & Gyr Building Control, Switzerland

Abstract

A design method is considered that aims for observers being robust against disturbances but sensitive to faults. This method is applied to a simulated boiler. One step of the design procedure, the choice of two weighting matrices, is seen to be quite difficult. The resulting optimal observers perform well as long as the design assumptions are met. Significant performance deteriorations occur in case of certain deviations from design assumptions as e.g. parameter estimation errors neglected in the design.

1 Introduction

Fault detection observers (FDOs) are based on a mathematical plant model driven by the same inputs as the plant. The simplest form of a FDO uses directly the difference between the plant output and the model output as fault indicating signal, called residual. This residual will deviate from zero in case of faults and should be closed to zero otherwise. For a survey on observer based fault detection methods see e.g. Frank [1].

A serious problem of observer based fault detection is that the residual deviates from zero not only in case of faults but also in case of model inaccuracies and disturbances. A design goal is therefore to make the residual insensitive to model inaccuracies and disturbances but sensitive to faults. A design method with that goal was developed by Frank and Wünnenberg [3] and is briefly described in sect. 2. This design method is applied to the example introduced in sect. 3. Simulation results are reported and discussed in sect. 4.

2 Optimal fault detection observers

Consider the system structure in fig. 1.

u : known input vector
y : measurement vector
d : unknown input vector:
    external disturbances and modeling errors
f : unknown fault vector
r : residual vector
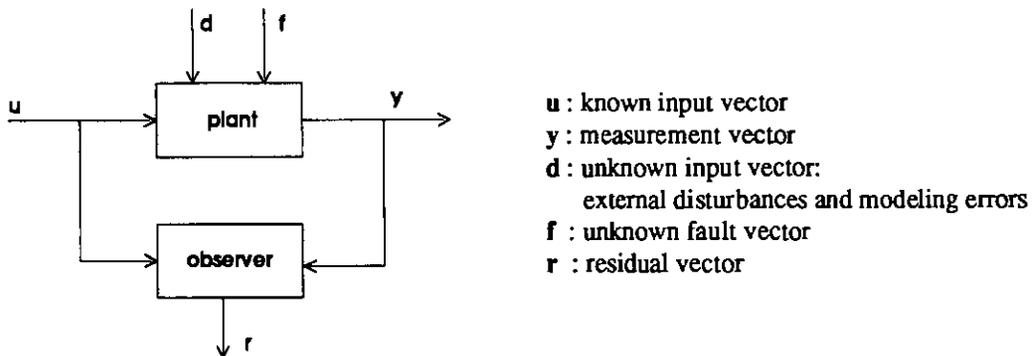
Fig. 1 Principle of observer based fault detection

Assume that the plant can be described by

$$\dot{x}(t) = Ax(t) + Bu(t) + E_d d(t) + K_d f(t) \tag{1a}$$
$$y(t) = Cx(t) \qquad + E_m d(t) + K_m f(t), \tag{1b}$$

where x is the state vector and all matrices are known and of appropriate dimensions. To generate a residual that is insensitive to the disturbances d but sensitive to the faults f Frank and Wünnenberg [3] have developed so called optimal fault detection observers (OFDOs). Their approach is based on the generalized observer

$$\dot{z}(t) = Fz(t) + Ju(t) + Gy(t) \tag{2a}$$
$$r(t) = L_1 z(t) + L_2 y(t). \tag{2b}$$

Observing (1) and leaving out effects of initial conditions the residual r can be decomposed as

$$r = r_d + r_f, \tag{3}$$

where r_d and r_f represent the residuals caused by the disturbances d and the faults f respectively. Assuming a scalar residual for simplicity and random variables d and f the design goal can be formulated as minimization of the performance index (PI)

$$P = \frac{E[r_d^2]}{E[r_f^2]} . \tag{4}$$

Instead of minimizing P its inverse, $P_i = 1/P$, will be *maximized* and used as PI here.
P_i is found to be

$$P_i = \frac{w^T V H_3 C_f H_3^T V^T w}{w^T V H_2 C_d H_2^T V^T w}, \tag{5}$$

where:

V is a basis for the so-called parity space, the left nullspace of $Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^P \end{bmatrix}$,

p is the observer order,

$$H_2 = \begin{bmatrix} E_m & 0 & .. & & 0 \\ CE_d & E_m & & & \vdots \\ CAE_d & CE_d & E_m & & \\ \vdots & \vdots & & & \\ CA^{P-1}E_d & CA^{P-2}E_d & .. & CE_d & E_m \end{bmatrix}, \tag{6}$$

$$H_3 = \begin{bmatrix} K_m & 0 & .. & & 0 \\ CK_d & K_m & & & \vdots \\ CAK_d & CK_d & K_m & & \\ \vdots & \vdots & & & \\ CA^{P-1}K_d & CA^{P-2}K_d & .. & CK_d & K_m \end{bmatrix}, \tag{7}$$

$$C_d = E\left[\left[\tilde{d}^T, .. D^P \tilde{d}^T \right]^T \left[\tilde{d}^T, .. D^P \tilde{d}^T \right] \right], \tag{8}$$

$$C_f = E\left[\left[\tilde{f}^T, .. D^P \tilde{f}^T \right]^T \left[\tilde{f}^T, .. D^P \tilde{f}^T \right] \right], \tag{9}$$

D is the differentiation operator,

$\tilde{d}$ and $\tilde{f}$ are the disturbances and faults low-pass filtered by $1/N(s)$,

$N(s) = a_0 + a_1 s + .. + a_{p-1} s^{P-1} + s^P$ is the characteristic polynomial of the observer,

w has to be found such that P_i is maximal.

P_i is maximized by solving the generalized eigenvalue-eigenvector problem

$$VH_3C_fH_3^TV^Tw = \lambda VH_2C_dH_2^TV^Tw \qquad (10)$$

and taking as w_{opt} the eigenvector corresponding to the largest eigenvalue λ_{max}. The observer matrices are then designed as:

$$F = \begin{bmatrix} 0 & .. & .. & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & . & & : \\ : & & & & : \\ 0 & .. & .. & 1 & -a_{p-1} \end{bmatrix} \qquad (11)$$

$$J = T\,B \qquad (12)$$

$$G = -\begin{bmatrix} v_0^T \\ : \\ v_{p-1}^T \end{bmatrix} + \begin{bmatrix} a_0 \\ : \\ a_{p-1} \end{bmatrix} v_p^T \qquad (13)$$

$$L_1 = -[0, .. 0,1], \text{ p-dimensional} \qquad (14)$$

$$L_2 = v_p^T, \qquad (15)$$

where

$$\begin{bmatrix} v_0^T, & .. & v_p^T \end{bmatrix} = w_{opt}^T V, \qquad (16)$$

$$T = \begin{bmatrix} v_p^T CA^{p-1} + v_{p-1}^T CA^{p-2} + .. + v_1^T C \\ : \\ v_p^T CA + v_{p-1}^T C \\ v_p^T C \end{bmatrix}. \qquad (17)$$

The design procedure is summarized to:
1) Choose the observer order p.
   Suggestion based on Frank and Wünnenberg [3]: p=n-m+1 (n=dim(x), m=dim(y))
2) Choose the observer polynomial N(s).
3) Choose V such that VQ=0.
4) Estimate the weighting matrices C_d and C_f.
5) Maximize P_i by solving (10).
6) Setup the observer according to (11)-(17).

Wünnenberg [4] extends his approach to the class of nonlinear systems described by

$$\dot{x}(t) = Ax(t) + B(u(t), y^*(t)) + E_d d(t) + K_d f(t) \qquad (18a)$$
$$y(t) = Cx(t) \qquad\qquad\quad + E_m d(t) + K_m f(t), \qquad (18b)$$

where

$$y^* = N_m y = N_m Cx \qquad (19)$$
$$N_m E_m = 0 \qquad (20)$$
$$N_m K_m = 0. \qquad (21)$$

3 Application example: boiler model

The control system in fig. 2 is considered.
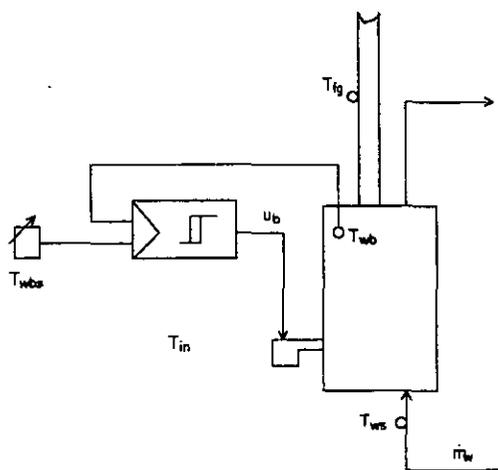
known or measured:
- u_b: burner control signal (0/1)
- T_{ws}: supply water temperature
- T_{wb}: boiler water temperature
- T_{wbs}: boiler water temperature setpoint
- T_{fg}: flue gas temperature

unknown:
- $\dot{m}_w$: water mass flow
- T_{in}: heating room temperature

Fig. 2 Boiler control system

The boiler model is taken from Gass and Hopkirk [5] and represented in fig. 3. It was developed at the university of Liege, Belgium.
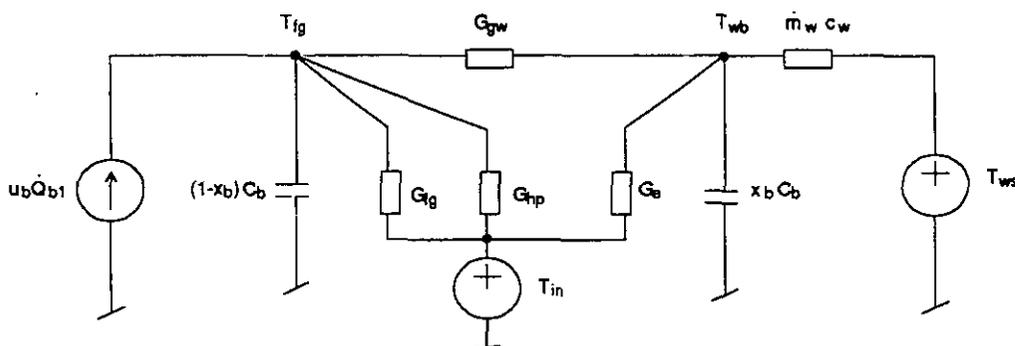
Fig. 3 Boiler model

- $\dot{Q}_{b1}$: burner power (burner on)
- G_{gw}: conductance for heat transfer 'gas - water'
- G_{hp}: conductance for heat transfer 'chimney inlet - plant room'
- G_e: conductance for heat transfer 'boiler water - plant room'
- G_{fg}: conductance for flue gas and ventilation losses
- $(1-x_b)C_b$: thermal capacity of the refractory wall
- $x_b C_b$: thermal capacity of the boiler (water and metal, without refractory wall)
- c_w: specific thermal capacity of water

The conductances G_{gw} and G_{fg} switch their values with u_b. All other parameters and $\dot{m}_w$ are assumed constant. In the original model all resistances depend on u_b. The numerical values for the parameters are chosen as follows (most of them taken from Gass and Hopkirk [5]):

G_e	= 47.6 W/K	G_{hp}	= 12.35 W/K
$G_{gw}(u_b = 0) = G_{gw0}$	= 136.1 W/K	$G_{gw}(u_b = 1) = G_{gw1}$	= 1108.6 W/K
$G_{fg}(u_b = 0) = G_{fg0}$	= 40.3 W/K	$G_{fg}(u_b = 1) = G_{fg1}$	= 110.7 W/K
C_b	= 580 Wh/K	x_b	= 0.98
c_w	= 1.16 Wh/kgK	$\dot{Q}_{b1}$	= 158.36 kW

The fault free and undisturbed system can be brought into the following form:

$$\dot{x} = Ax + B(u,y) \tag{22a}$$
$$y = x, \tag{22b}$$

where:

$$x = \begin{bmatrix} T_{wb}, & T_{fg} \end{bmatrix}^T \tag{23}$$

$$u = \begin{bmatrix} u_b, & T_{ws} \end{bmatrix}^T \tag{24}$$

$$A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}, \quad a_{11} = -\frac{G_e + \dot{m}_w c_w}{x_b C_b}, \quad a_{22} = -\frac{G_{hp}}{(1-x_b)C_b} \tag{25}$$

$$B(u,y) = \begin{bmatrix} b_1(u,y) \\ b_2(u,y) \end{bmatrix} = \begin{bmatrix} 0 & \dfrac{\dot{m}_w c_w}{x_b C_b} \\ \dfrac{\dot{Q}_{bl}}{(1-x_b)C_b} & 0 \end{bmatrix} u +$$

$$\begin{bmatrix} \dfrac{G_e}{x_b C_b} \\ \dfrac{G_{hp}+G_{fg}(u_b)}{(1-x_b)C_b} \end{bmatrix} T_{in} + \begin{bmatrix} \dfrac{-G_{gw}(u_b)}{x_b C_b} & \dfrac{G_{gw}(u_b)}{x_b C_b} \\ \dfrac{G_{gw}(u_b)}{(1-x_b)C_b} & \dfrac{-G_{gw}(u_b)+G_{fg}(u_b)}{(1-x_b)C_b} \end{bmatrix} y \tag{26}$$

$\dot{m}_w$ and T_{in} are assumed constant at $\dot{m}_w$=4000 kg/h and T_{in}=25 °C and treated as parameters. To enable the implementation of an observer estimates for all parameters are assumed available. Some of these estimates may be erroneous. The influence of such errors on the residual will be modeled by disturbance signals d.

The implementation of an observer would be eased by a change to a discrete time system description. However, to design the observer robust with respect to some selected parameter estimation errors, the continuous time description is advantageous (see [6]).

The following modeling errors, external disturbances, and faults will be considered:
- modeling errors:
  parameter estimation errors in
  • water flow $\dot{m}_w$
  • conductances for flue gas and ventilation losses G_{fg0} and G_{fg1}
  • capacity ratio x_b
- external disturbances:
  measured errors in
  • boiler temperature T_{wb}
  • flue gas temperature T_{fg}
- faults:
  growing deviation of the conductances G_{gw0} and G_{gw1} from their initial values due to soot

The residuals of the OFDOs will be compared with the ones of two Luenberger-like observers (LLOs) of the form

$$\dot{z}_i = a_{ii0}z_i + b_{i0}(u,y) + h_i(y_i - z_i), \quad i \in \{1,2\}, \tag{27}$$

where the index '0' denotes nominal parameter values, and the feedback gains $h_{1..2}$ are chosen to have the poles at p_{o1}=-1/120sec and p_{o2}=-1/30sec. This choice gives acceptable settling times for the residuals.

4 Simulation experiments

The simulation experiments reported here are executed using the parameter values from sect. 3 apart from specific parameter changes specified for individual experiments. Additional settings are:
- burner controller switching difference: 10K
- supply water temperature: $T_{ws}=50\,^{\circ}C$
- boiler water temperature setpoint: $T_{wbs}=65\,^{\circ}C$

The experiments are summarized in tables 1 and 2, sect. 4.5.

4.1 Experiment group 1: Correlated parameter estimation errors

Let

$$\dot{m}_w = \dot{m}_{w0} + \Delta\dot{m}_w \tag{28}$$

$$G_{fgi} = G_{fgi0} + \Delta G_{fgi}, \quad i=1,2 \tag{29}$$

$$G_{gwi} = G_{gwi0} + \Delta G_{gwi}, \quad i=1,2. \tag{30}$$

The last index '0' always denotes nominal parameter values assumed for observer design and 'Δ' a deviation of the true parameter value. Nominal values are the ones defined in sect. 3. Bringing the system description into form (18) leads to

$$\dot{x}(t) = A_0 x(t) + B_0(u(t), y(t)) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_1(t) \\ d_2(t) \end{bmatrix} + \begin{bmatrix} -\gamma \\ 1 \end{bmatrix} f(t) \tag{31a}$$

$$y(t) = x(t), \tag{31b}$$

where:

$$d_1(t) = \frac{\Delta\dot{m}_w c_w}{x_b C_b}(u_2(t) - x_1(t)) \tag{32}$$

$$d_2(t) = \frac{\Delta G_{fg}(u_b)}{(1-x_b)C_b}(T_{in} - y_2(t)), \quad \Delta G_{fg}(u_b = i) = \Delta G_{fgi} \tag{33}$$

$$f(t) = \frac{\Delta G_{gw}(u_b)}{(1-x_b)C_b}(y_1(t) - y_2(t)), \quad \Delta G_{gw}(u_b = i) = \Delta G_{gwi} \tag{34}$$

$$\gamma(t) = \frac{1-x_b}{x_b}. \tag{35}$$

Follow now the design procedure in sect. 2:

1) Observer order: p=1.

2) Observer polynomial: $N(s) = a_0 + s$, $a_0 = 1/36sec$.

3) Basis for parity space:

$$V = \begin{bmatrix} -a_{11} & 0 & 1 & 0 \\ 0 & -a_{22} & 0 & 1 \end{bmatrix}$$

4) Weighting matrices:
   Simplifying assumptions:
   • regular boiler operation with medium load

$$\bullet \ d_1(t) = -\frac{\Delta\dot{m}_w c_w}{x_b C_b}(15K + 5K\sin(\omega_b t + \varphi)) = \Delta\dot{m}_w(a_1 + b_1\sin(\omega_b t + \varphi)) \tag{36}$$

$$\bullet \ d_2(t) = -\frac{\Delta\overline{G}_{fg}}{(1-x_b)C_b}(115K + 55K\sin(\omega_b t + \varphi)) = \Delta\overline{G}_{fg}(a_2 + b_2\sin(\omega_b t + \varphi)) \tag{37}$$

$$\bullet \ f(t) = -\frac{\Delta \overline{G}_{gw}}{(1-x_b)C_b}(75K + 50K\sin(\omega_b t + \varphi)) = \Delta \overline{G}_{gw}(a_3 + b_3\sin(\omega_b t + \varphi)) \tag{38}$$

$$\bullet \ \omega_b = 2\pi / 12 \min \tag{39}$$

$\bullet \ \Delta \dot{m}_w, \ \Delta \overline{G}_{fg}, \ \Delta \overline{G}_{gw},$ and φ are r.v. with:

$$E[\Delta \dot{m}_w] = 0, \quad \sqrt{\text{Var}[\Delta \dot{m}_w]} = 0.1 \ \dot{m}_{w0} \tag{40}$$

$$E[\Delta \overline{G}_{fg}] = 0, \quad \sqrt{\text{Var}[\Delta \overline{G}_{fg}]} = 0.1 \ \overline{G}_{fg 0} = 0.1 \ (G_{fg 00} + G_{fg 10})/2 \tag{41}$$

$$E[\Delta \overline{G}_{gw}] = 0, \quad \sqrt{\text{Var}[\Delta \overline{G}_{gw}]} = 0.05 \ \overline{G}_{gw0} = 0.05 \ (G_{gw 00} + G_{gw 10})/2 \tag{42}$$

$$\varphi \sim U(0, 2\pi), \quad \text{independent of } \Delta \dot{m}_w, \ \Delta \overline{G}_{fg}, \text{ and } \Delta \overline{G}_{gw} \tag{43}$$

Filtering $d_1(t)$, $d_2(t)$, and $f(t)$ by $1/N(s)$ leads to

$$\tilde{d}_1(t) = \Delta \dot{m}_w(\tilde{a}_1 + \tilde{b}_1\sin(\omega_b t + \tilde{\varphi})) \tag{44}$$

and corresponding expressions for $\tilde{d}_2$ and $\tilde{f}$, where

$$\tilde{a}_i = \frac{a_i}{a_0}, \qquad \tilde{b}_i = \frac{b_i}{\sqrt{a_0^2 + \omega_b^2}}, \qquad \tilde{\varphi} = \varphi - \text{arctg}\left(\frac{\omega_b}{a_0}\right), \qquad i = 1..3. \tag{45}$$

C_d and C_f now follow as

$$C_d = \begin{bmatrix} E[\Delta\dot{m}_w^2](\tilde{a}_1^2 + \frac{\tilde{b}_1^2}{2}) & E[\Delta\dot{m}_w\Delta\overline{G}_{fg}](\tilde{a}_1\tilde{a}_2 + \frac{\tilde{b}_1\tilde{b}_2}{2}) & 0 & 0 \\ E[\Delta\dot{m}_w\Delta\overline{G}_{fg}](\tilde{a}_1\tilde{a}_2 + \frac{\tilde{b}_1\tilde{b}_2}{2}) & E[\Delta\overline{G}_{fg}^2](\tilde{a}_2^2 + \frac{\tilde{b}_2^2}{2}) & 0 & 0 \\ 0 & 0 & E[\Delta\dot{m}_w^2]\frac{\tilde{b}_1^2}{2}\omega_b^2 & E[\Delta\dot{m}_w\Delta\overline{G}_{fg}]\frac{\tilde{b}_1\tilde{b}_2}{2}\omega_b^2 \\ 0 & 0 & E[\Delta\dot{m}_w\Delta\overline{G}_{fg}]\frac{\tilde{b}_1\tilde{b}_2}{2}\omega_b^2 & E[\Delta\overline{G}_{fg}^2]\frac{\tilde{b}_2^2}{2}\omega_b^2 \end{bmatrix} \tag{46}$$

$$C_f = \begin{bmatrix} E[\Delta\overline{G}_{gw}^2](\tilde{a}_2^2 + \frac{\tilde{b}_2^2}{2}) & 0 \\ 0 & E[\Delta\overline{G}_{gw}^2]\frac{\tilde{b}_2^2}{2}\omega_b^2 \end{bmatrix} \tag{47}$$

At this point *by mistake*

$$E[\Delta\dot{m}_w\Delta\overline{G}_{fg}] = \sqrt{E[\Delta\dot{m}_w^2]}\sqrt{E[\Delta\overline{G}_{fg}^2]} \tag{48}$$

was assumed, i.e. observing (40) and (41) full positive correlation of $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$. Although this assumption is generally not met, this experiment group is reported here, because the resulting OFDO shows an interesting sensitivity property.

5) Maximize the performance index P_i (see (10)):
P_{imax} is found to be $P_{imax}=1150$, what means very good amplification of the faults and suppression of the disturbances.

6) Observer matrices according to (11)-(15):
(shown in rounded form, implemented in double precision)

$$
\begin{aligned}
F &= -0.0278 \ \text{sec}^{-1} \\
J(u,y) &= TB(u,y) = [-0.988 \quad 0.157]B(u,y) \\
G &= [-90.6 \quad 15.6] \ \text{sec}^{-1} \\
L_1 &= -1 \\
L_2 &= [-0.988 \quad 0.157]
\end{aligned}
\qquad (49)
$$

This completes the design of the OFDO.

Exp. 1.1: Disturbances and faults exactly as assumed for design

In (36)-(38) set

$$\Delta \dot{m}_w = -0.1 \ \dot{m}_{w0} \qquad (50)$$

$$\Delta \overline{G}_{fg} = -0.1 \ \overline{G}_{fg0} \qquad (51)$$

$$\Delta \overline{G}_{gw} = -0.05 \ \overline{G}_{gw0} \qquad (52)$$

$$\varphi = 0, \qquad (53)$$

and apply the resulting disturbance and fault signals to the nominal plant.

The following figures with simulation results always show the 3 cases

(1) no disturbances, no faults (full line),

(2) disturbances, no faults (dashed line), and

(3) no disturbances, faults (dotted line).

Fig. 4 and 5 show the boiler water temperature and the flue gas temperature. Note that the course of these temperatures satisfies well the assumptions (36) - (39) concerning mean, amplitude, and frequency of the relevant temperature differences. As expected the disturbance signal d_1 corresponding to a decrease of $\dot{m}_w$ causes a faster rise and a slower drop of the boiler water temperature. The rise and the steady state of the flue gas temperature are little sensitive to the disturbance and fault signals. However, its drop is slower with the fault f signal corresponding to a decrease of $\Delta \overline{G}_{gw}$ due to less cooling by the boiler water.
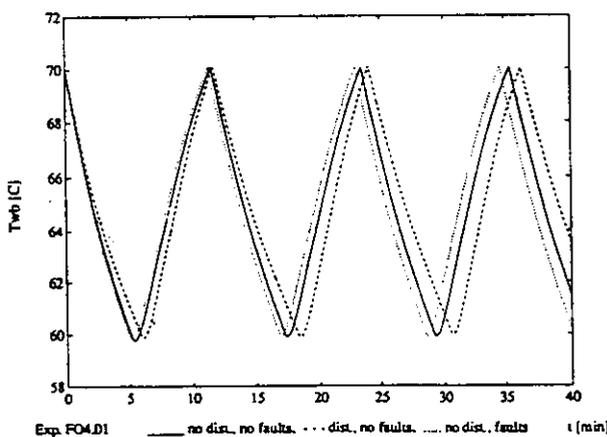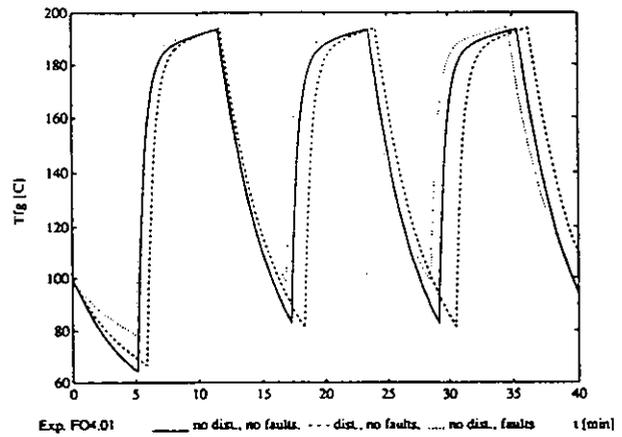
Fig. 4   Boiler water temperature

Fig. 5   Flue gas temperature

Fig. 6 and 7 show the residuals generated by the OFDO and the two LLOs (27). Clearly the OFDO is the best one. It fulfils what the design promised in that the measured performance index P_{im} deviates by less than 0.01% from the expected value. The LLO for the boiler water temperature is obviously not suitable in the current situation and will not be considered further.
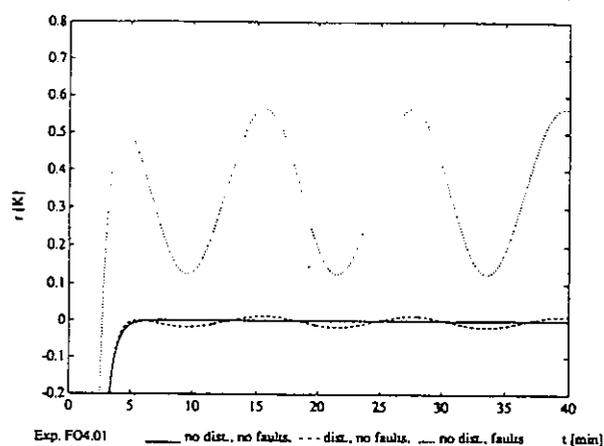
Fig. 6 Residuals of optimal fault detection observer (OFDO) $(P_{im}=1150)$
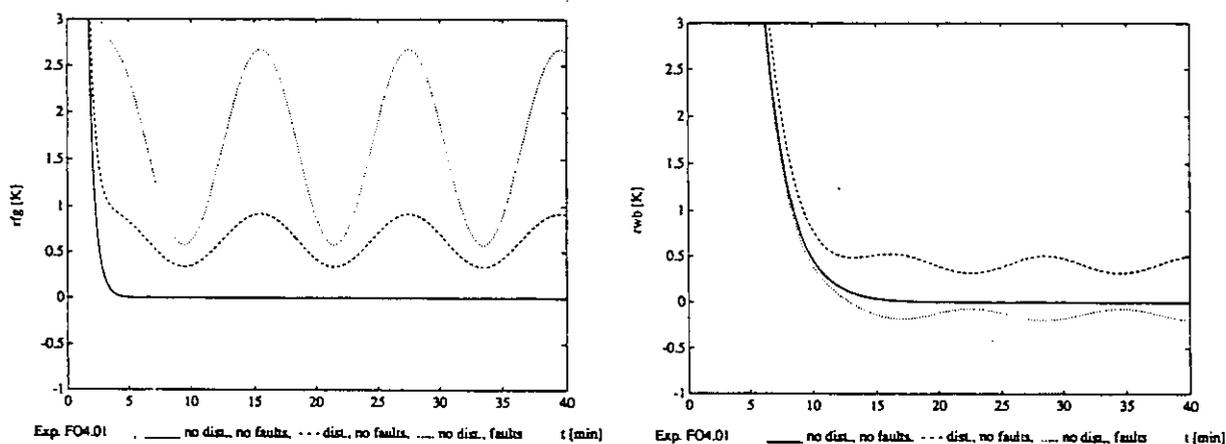
Fig. 7 Residuals of Luenberger-like observers (LLOs)
   a) for the flue gas temperature $(P_{im}=7.3)$ b) for the boiler water temperature $(P_{im}<1)$

Exp. 1.2: Disturbances and faults realized by constant parameter errors

Instead of the external signals $d_1(t)$, $d_2(t)$, and $f(t)$ implement in the plant the parameter errors

$$\Delta \dot{m}_w = -0.1\, \dot{m}_{w0} \tag{54}$$

$$\Delta G_{fg}(u_b=0) = \Delta G_{fg}(u_b=1) = \Delta \overline{G}_{fg} = -0.1\, \overline{G}_{fg0} \tag{55}$$

$$\Delta G_{gw}(u_b=0) = \Delta G_{gw}(u_b=1) = \Delta \overline{G}_{gw} = -0.05\, \overline{G}_{gw0}. \tag{56}$$

Fig. 8 and 9 show the residuals of the OFDO and the LLO for T_{fg}. The OFDO is still better, but its PI has become about 4 times lower than in exp. 1.1, while the LLO could slightly improve its PI.

337
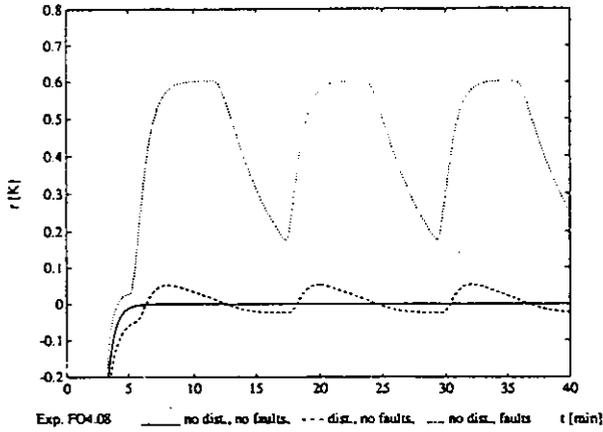
Fig. 8 Residuals of the OFDO (P_{im}=255)
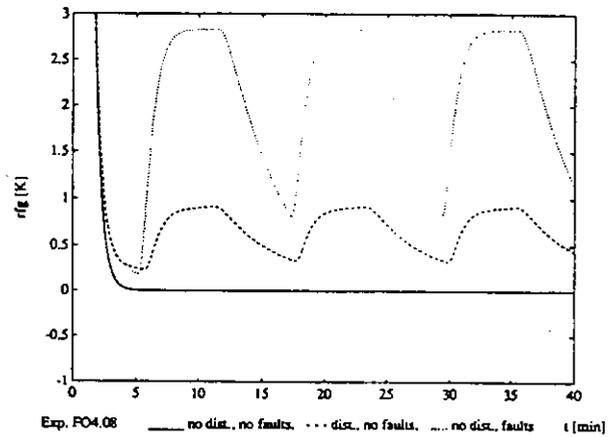
Fig. 9 Residuals of the LLO for T_{fg} (P_{im}=10)

Exp. 1.3: Parameter estimation error $\Delta\dot{m}_w$ with opposite sign

Parameter errors as in exp. 1.2, but $\Delta\dot{m}_w = +0.1\ \dot{m}_{w0}$.

Fig. 10 and 11 show a significant deterioration of the OFDO performance w.r.t. exp. 1.2, while the LLO for T_{fg} is insensitive to the sign switch in $\Delta\dot{m}_w$. The sign of $\Delta\dot{m}_w$ enters the OFDO design only in the correlation with $\Delta\overline{G}_{fg}$. In the design $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$ were assumed to have zero means and full positive correlation. Then the event, that $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$ have different signs, is highly unlikely. It would be even impossible, if the full correlation would be caused by $\Delta\dot{m}_w$ being a linear function of $\Delta\overline{G}_{fg}$ satisfying (40) and (41). But the design assumptions are not necessarily met, and thus the selected event may occur vey well. Therefore it is concluded, that in the current situation the PI is highly sensitive to deviations from the assumed correlation. This correlation will be dropped in the next experiment group.
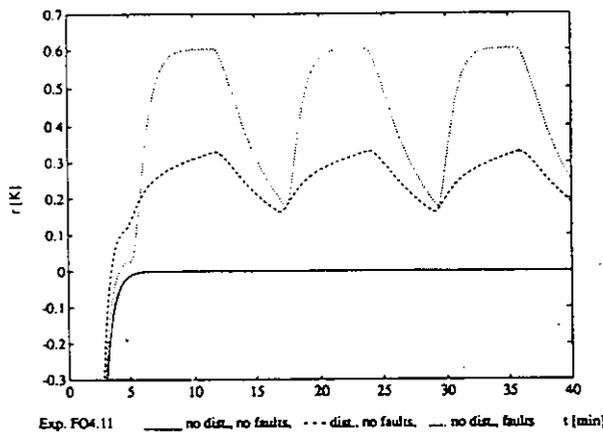
Fig. 10 Residuals of the OFDO (P_{im}=3.2)
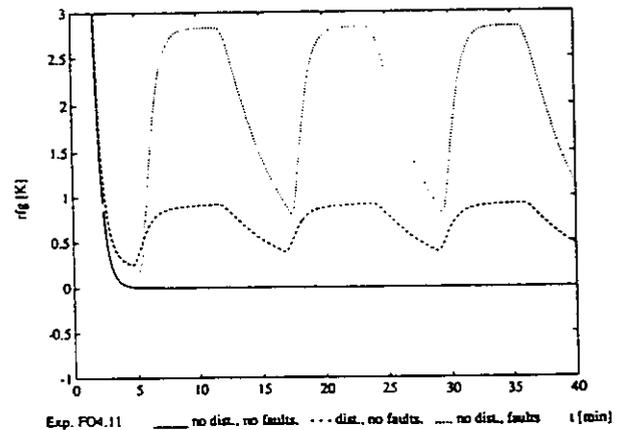
Fig. 11 Residuals of the LLO for T_{fg} (P_{im}=8.7)

4.2 Experiment group 2: Uncorrelated parameter estimation errors

Consider the same parameter estimation errors $\Delta\dot{m}_w$, $\Delta G_{fg0.1}$, and $\Delta G_{gw0.1}$ as in exp. group 1, but assume $\Delta\dot{m}_w$ uncorrelated with ΔG_{fg0}, and ΔG_{fg1}.

OFDO design:

1) - 3) Exactly as in exp. group 1.

4) Weighting matrices:
As in exp. group 1, but let now $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$ be uncorrelated. Then all nondiagonal elements of C_d and C_f vanish.

5) Maximize P_i (see (10)):
P_{imax} is found to be $P_{imax}=7.4$ and is about 160 times lower than in exp. group 1, what means a significant deterioration.

6) Observer matrices according to (11)-(15):
(shown in rounded form, implemented in double precision)

$$
\left.
\begin{aligned}
F &= -0.0278 \text{ sec}^{-1} \\
J(u,y) &= TB(u,y)=[-0.629 \quad 0.777]B(u,y) \\
G &= [-57.7 \quad 76.9]\text{ sec}^{-1} \\
L_1 &= -1 \\
L_2 &= [-0.629 \quad 0.777]
\end{aligned}
\right\}
\tag{57}
$$

Exp. 2.1: Constant parameter errors as assumed for design

Parameter errors as in exp. 1.2.

The resulting OFDO residual in fig. 12 shows the expected modest performance and is similar to the LLO residual in fig. 9. At least the deviations from design conditions due to true parameter errors didn't cause a lower PI than promised by the design.

Exp. 2.2: Parameter estimation error $\Delta\dot{m}_w$ with opposite sign

Parameter errors as in exp. 1.2, but $\Delta\dot{m}_w = +0.1\ \dot{m}_{w0}$.

Comparing fig.12 and fig.13 shows that the OFDO residual is now little sensitive to the sign of $\Delta\dot{m}_w$. Further experiments show that it is also little sensitive to the size of $\Delta\dot{m}_w$.

Exp. P06.07 _____ no dist., no faults, --- dist., no faults, ... no dist., faults   t [min]

Exp. P06.08 _____ no dist., no faults, --- dist., no faults, ... no dist., faults   t [min]

Fig. 12 Residuals of the OFDO (P_{im}=13)
        parameter errors as assumed for design
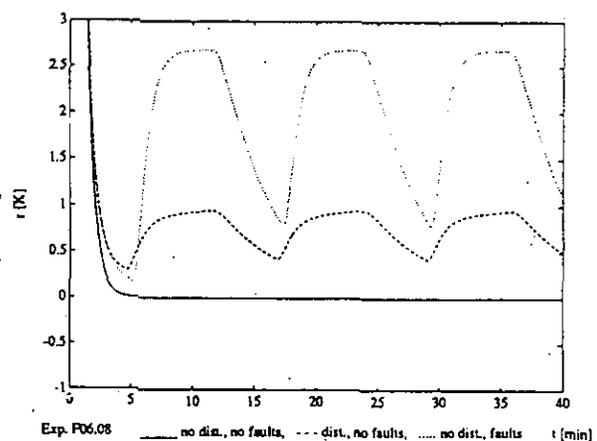
Fig. 13 Residuals of the OFDO (P_{im}=7.3)
        $\Delta\dot{m}_w$ with opposite sign

Exp. 2.3: Input dependent parameter estimation error ΔG_{fg}

Parameter errors as in exp. 1.2, but ΔG_{fg} refined to input u_b dependent values:

$$\Delta G_{fg} = \Delta G_{fg}(u_b) = \begin{cases} \Delta G_{fg0} = -0.1 \, G_{fg0} & \text{if } u_b = 0 \\ \Delta G_{fg1} = -0.1 \, G_{fg1} & \text{if } u_b = 1 \end{cases} \tag{58}$$

The OFDO shows similar performance as in exp. 2.2 (P_{im}=8.6).

Exp. 2.4: Parameter estimation errors $\Delta G_{fg0/1}$ with doubled magnitude

Conditions as in exp. 2.3, but $\Delta G_{fg0/1}$ with doubled magnitude.

The residuals in fig.14 and 15 show bad performance for both the OFDO and the LLO. It is no longer possible to distinguish the assumed disturbances and faults. In the selected event $\overline{\Delta G}_{fg}{}^2$ is 4 times higher than expected in the design. Correcting the expectation and redesigning the OFDO for 4-fold $\text{Var}\left[\Delta \overline{G}_{fg}{}^2\right]$ gives P_{imax}=1.9, i.e. the same value that the original OFDO has reached already. Therefore it is concluded that, given observer order p=1, one can't do much better. Observing (31a) and the small value of γ (0.02), the bad performance is not surprising. Obviously it must be difficult to separate the effects of $d_2(t)$ and $f(t)$, because mainly their sum acts on the system.
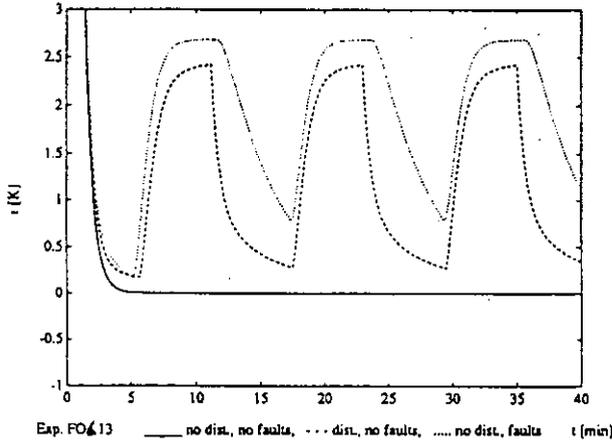
Fig. 14 Residuals of the OFDO (P_{im}=1.9)
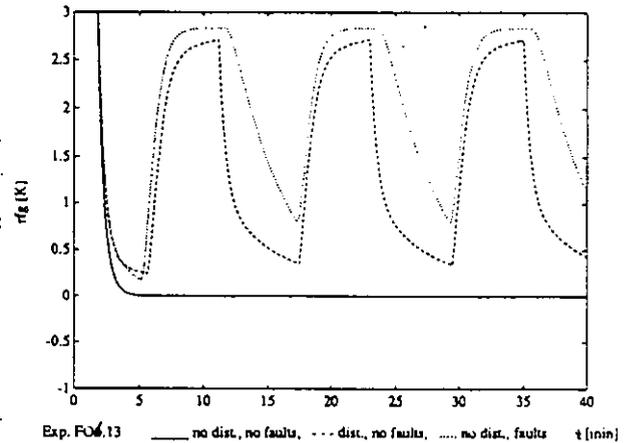
Fig. 15 Residuals of the LLO (P_{im}=1.7)

Exp. 2.5: Input dependent parameter deviation ΔG_{gw}

Parameter errors as in exp. 1.2, but ΔG_{gw} refined to input u_b dependent values:

$$\Delta G_{gw} = \Delta G_{gw}(u_b) = \begin{cases} \Delta G_{gw0} = -0.05 \, G_{gw00}{}^2 / G_{gw10} & \text{if } u_b = 0 \\ \Delta G_{gw1} = -0.05 \, G_{gw10}{}^2 / G_{gw10} & \text{if } u_b = 1 \end{cases} \tag{59}$$

The sizes of ΔG_{gw0} and ΔG_{gw1} are based on the assumption that the resistance of the soot film on the combustion chamber wall has increased by 5%.

Both, the OFDO and the LLO perform well in this case. The PIs are 30 for the OFDO and 23 for the LLO. The improvement with respect to exp. 2.1 is due to the increase of ΔG_{gw} in case u_b=1.

340

Exp. 2.6: Unconsidered parameter estimation errors

Parameter errors as in exp. 1.2, and in addition an estimation error in the capacity ratio x_b, i.e.

$$x_b = x_{b0} + \Delta x_b = x_{b0} - 0.05(1 - x_{b0}).$$
(60)

Note that this error was not considered in the OFDO design.

Fig. 16 and 17 show that both, the OFDO and the LLO residuals are rather sensitive to the error Δx_b. However, the PIs are not that bad because in case of disturbances large residuals appear only for short periods.
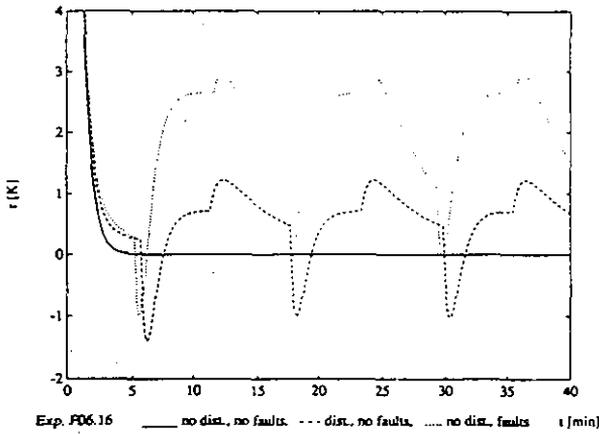
Exp. F06.16 _____ no dist., no faults. - - - dist., no faults. ..... no dist., faults   t [min]
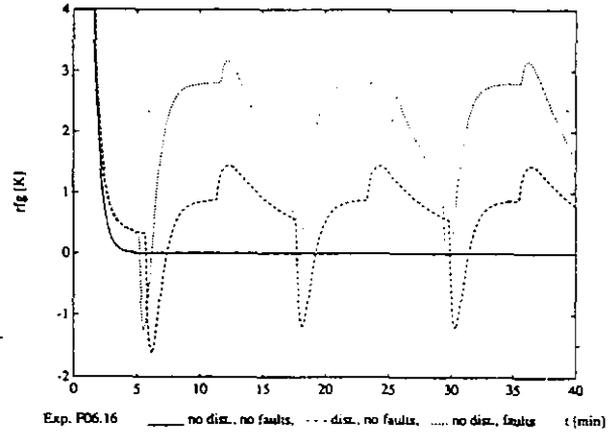
Exp. F06.16 _____ no dist., no faults. - - - dist., no faults. ..... no dist., faults   t [min]

Fig. 16 Residuals of the OFDO (P_{im}=7.7)   Fig. 17 Residuals of the LLO (P_{im}=6.2)

It is likely that all estimated parameters deviate more or less from the true values. However, to consider all these possible errors in the OFDO design would complicate the design too much and probably lead to an OFDO that is insensitive also to faults. A difficulty is therefore to find those parameters that have significant influence on the residual and the estimates of which are most uncertain.

4.3 Experiment group 3: 2nd order observer

As the performance of the 1st order observer in exp. 4.2 is rather modest, an improvement is now attempted by increasing the observer order. The same parameter estimation errors are assumed as in exp. group 1.

OFDO design:

1) Observer order: p=2.
2) Observer polynomial: $N(s) = (s-p_0)^2 = a_0 + a_1 s + s^2$, p_0=-1/36sec.
3) Basis for parity space:

$$V = \begin{bmatrix} -a_{11} & 0 & 1 & 0 & 0 & 0 \\ 0 & -a_{22} & 0 & 1 & 0 & 0 \\ -a_{11}^2 & 0 & 0 & 0 & 1 & 0 \\ 0 & -a_{22}^2 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4) Weighting matrices:

$$C_d = E\left[\left[\breve{d}^T, D\breve{d}^T, D^2\breve{d}^T\right]^T \left[\breve{d}^T, D\breve{d}^T, D^2\breve{d}^T\right]\right], \quad C_f = E\left[\left[\breve{f}^T, D\breve{f}^T, D^2\breve{f}^T\right]^T \left[\breve{f}^T, D\breve{f}^T, D^2\breve{f}^T\right]\right]$$

Proceeding with the same simplifying assumptions as in exp. group 1 leads to the filtered signals

$$\tilde{d}_1(t) = \Delta \dot{m}_w (\tilde{a}_1 + \tilde{b}_1 \sin(\omega_b t + \tilde{\phi}))$$

and corresponding expressions for $\tilde{d}_2$ and $\tilde{f}$ (see eqs. (37) and (38)), where

$$\tilde{a}_i = \frac{a_i}{a_0}, \qquad \tilde{b}_i = \frac{b_i}{\sqrt{\left(a_0 - \omega_b^2\right)^2 + a_1^2 \omega_b^2}}, \qquad \tilde{\phi} = \phi - \arctan\left(\frac{a_1 \omega_b}{a_0 - \omega_b^2}\right), \qquad i = 1..3. \tag{61}$$

The submatrices $C_d(1:4,1:4)$ and $C_f(1:2,1:2)$ can be taken from exp. group 1 (see eqs. (46) and (47)) with $\tilde{a}_{1..3}$ and $\tilde{b}_{1..3}$ as defined in (61) and nondiagonal elements cleared because of assuming uncorrelated parameter estimation errors. It turns out here that all other elements of C_d and C_f do not enter the performance index and therefore may be set to 0.

5) Maximizing P_i (see (10)):
P_{imax} is found to be $P_{imax}=13$, what means a modest improvement by a factor 1.8 compared to the 1st order observer under the same premises (see design in exp. group 2).

6) Observer matrices according to (11)-(15):
(shown in rounded form, implemented in double precision)

$$\left.\begin{aligned}
F \quad &= \begin{bmatrix} 0 & -0.000772\sec^{-2} \\ 1 & -0.0556\sec^{-1} \end{bmatrix} \\
J(u,y) \quad &= TB(u,y) = \begin{bmatrix} 0 & 0 \\ -0.844 & 0.536 \end{bmatrix} B(u,y) \\
G \quad &= \begin{bmatrix} -0.000651\sec^{-2} & 0.000414\sec^{-2} \\ -0.0450\sec^{-1} & 0.0296\sec^{-1} \end{bmatrix} \\
L_1 \quad &= \begin{bmatrix} 0 & -1 \end{bmatrix} \\
L_2 \quad &= \begin{bmatrix} -0.844 & 0.536 \end{bmatrix}
\end{aligned}\right\} \tag{62}$$

Exp. 3.1: Constant parameter estimation errors as assumed for design

Parameter errors as in exp. 1.2.

The resulting OFDO residual in fig. 18 shows only a slightly better performance than the one of the 1st order observer in exp. 2.1, fig. 12. Its magnitude in case of faults is 0.35 K only and is about 8 times smaller than in exp. 2.1. Therefore, observing (2b) and the values of L_1 and L_2 in (62), it is obvious that measurement errors would significantly corrupt this small residual.

Exp. 3.2: Input dependent parameter estimation error ΔG_{fg}

Parameter errors as in exp. 1.2, but ΔG_{fg} refined according to (58).

Fig. 19 shows a significant performance deterioration w.r.t. exp. 3.1. The absolute and relative decrease of the PI due to refinement of ΔG_{fg} is larger than in case of the first order observer (see exp. 2.1 and 2.3).
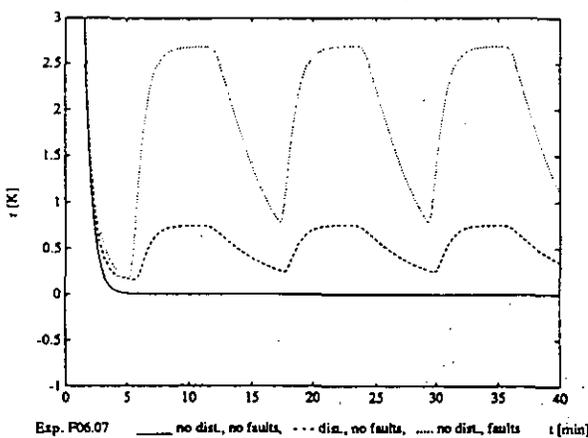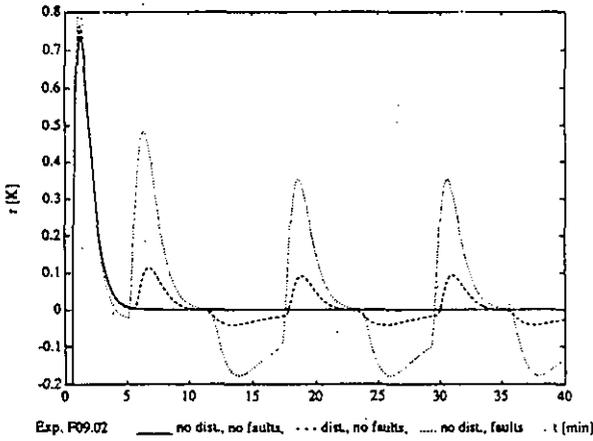
Fig. 18 Residuals of the OFDO (P_{im}=16)
parameter errors as assumed for design
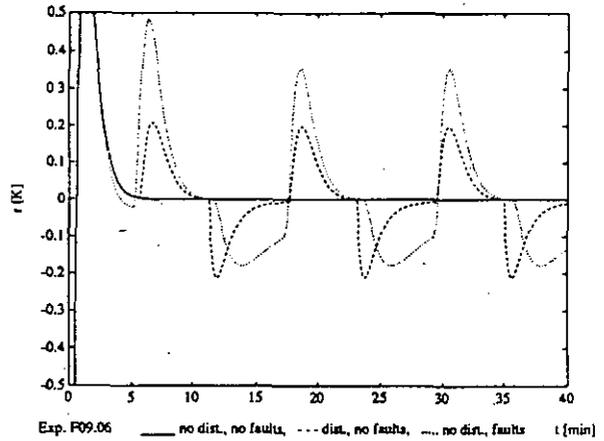
Fig. 19 Residuals of the OFDO (P_{im}=2.6)
Input dependent estimation error ΔG_{fg}

Exp. 3.3: Unconsidered parameter estimation errors

Parameter errors as in exp. 1.2, and in addition an estimation error in the capacity ratio x_b according to (60).

The PI drops from 16 in exp. 3.1 to 1.5. So its decrease due to the error Δx_b is again larger than in case of the 1st order observer (see exp. 2.1 and 2.6).

4.4 Experiment group 4: Stochastic measurement errors

The small residuals in exp. 3.1-3 would get corrupted significantly in case of measurement errors. To avoid this risk measurement errors are now considered in the observer design. As the OFDO residuals in exp. group 2 and 3 are little sensitive to $\Delta \dot{m}_w$, this estimation error is now dropped to simplify the OFDO design.

As in exp. group 1 let

$$G_{fgi} = G_{fgi0} + \Delta G_{fgi}, \quad i = 1,2$$
$$G_{gwi} = G_{gwi0} + \Delta G_{gwi}, \quad i = 1,2.$$

In addition introduce the stochastic measurement errors e_{wb} and e_{fg} for the boiler water temperature and the flue gas temperature. Bringing the system description into a form close to (18) leads to

$$\dot{x}(t) = A_0 x(t) + B_0(u(t),x(t)) + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1(t) \\ d_2(t) \\ d_3(t) \end{bmatrix} + \begin{bmatrix} -\gamma \\ 1 \end{bmatrix} f(t) \qquad (63a)$$

$$y(t) = x(t) + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1(t) \\ d_2(t) \\ d_3(t) \end{bmatrix}, \qquad (63b)$$

where:

$$d_1(t) = \frac{\Delta G_{fg}(u_b)}{(1-x_b)C_b}(T_{in} - y_2(t)), \quad \Delta G_{fg}(u_b = i) = \Delta G_{fgi} \qquad (64)$$

$$d_2(t) = e_{wb}(t) \qquad (65)$$

$$d_3(t) = e_{fg}(t) \qquad (66)$$

$$f(t) = \frac{\Delta G_{gw}(u_b)}{(1-x_b)C_b}(y_1(t) - y_2(t)), \quad \Delta G_{gw}(u_b = i) = \Delta G_{gwi}$$

343

$$\gamma(t) = \frac{1 - x_b}{x_b}.$$

Note that form (18) cannot be reached exactly, because no matrix N_m exists such that

$$x = N_m y \quad \text{and} \quad N_m \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0. \tag{67}$$

To still use the OFDO design of Wünnenberg [4] assume the term $B_0(u,x)$ in (63a) replaced by $B_0(u,y)$. This will lead to a suboptimal fault detection observer (SOFDO).

To generate the measurement errors $e_{wb}(t)$ and $e_{fg}(t)$ use in both cases a *discrete* 2nd order butterworth filter, driven by a white gaussian noise sequence and followed by a ZOH. The following parameters are used (nominal values):

- sampling time: $T_s = 1\text{sec}$

- filter cutoff frequency: $\omega_c = 1/30\text{sec}$

- means and standard deviations:

$$\mu_{d_1} = \mu_{e_{wb}} = E[e_{wb}] = 0.2\,K \tag{68}$$

$$\sigma_{d_1} = \sigma_{e_{wb}} = \sqrt{\text{Var}[e_{wb}]} = 0.1\,K \tag{69}$$

$$\mu_{d_2} = \mu_{e_{fg}} = E[e_{fg}] = 0.5\,K \tag{70}$$

$$\sigma_{d_2} = \sigma_{e_{fg}} = \sqrt{\text{Var}[e_{fg}]} = 0.25\,K \tag{71}$$

OFDO design:

1) - 3) See exp. group 3.

4) Weighting matrices:

$$C_d = E\left[\left[\tilde{d}^T, D\tilde{d}^T, D^2\tilde{d}^T\right]^T \left[\tilde{d}^T, D\tilde{d}^T, D^2\tilde{d}^T\right]\right], \quad C_f = E\left[\left[\tilde{f}, D\tilde{f}, D^2\tilde{f}\right]^T \left[\tilde{f}, D\tilde{f}, D^2\tilde{f}\right]\right]$$

For $d_1(t)$ and $f(t)$ the same simplifying assumptions as in exp. group 1 are used. Keeping the index numbers from exp. group 1 this leads to

$$\tilde{d}_1(t) = \Delta\overline{G}_{fg}(\tilde{a}_2 + \tilde{b}_2 \sin(\omega_b t + \tilde{\varphi})) \quad \text{and} \tag{72}$$

$$\tilde{f}(t) = \Delta\overline{G}_{gw}(\tilde{a}_3 + \tilde{b}_3 \sin(\omega_b t + \tilde{\varphi})), \tag{73}$$

where $\tilde{a}_{2/3}$, $\tilde{b}_{2/3}$, and $\tilde{\varphi}$ are defined in (61).

$d_1(t)$ .. $d_3(t)$ are assumed to be uncorrelated in pairs. Then C_d can be constructed from the elements of the 3 matrices

$$C_{d_i} = E\left[\left[\tilde{d}_i, D\tilde{d}_i, D^2\tilde{d}_i\right]^T \left[\tilde{d}_i, D\tilde{d}_i, D^2\tilde{d}_i\right]\right], \quad i = 1..3, \tag{74}$$

by simply inserting these elements in the correct places in C_d:

$$C_d((j-1)3+i, \ (k-1)3+i) = C_{d_i}(j,k), \quad i,j,k = 1..3. \tag{75}$$

Elements of C_d not set in (75) vanish. C_{d_i} follows as

$$C_{d_1} = E[\Delta \overline{G}_{fg}^{\;2}] \begin{bmatrix} (\tilde{a}_2^{\;2} + \dfrac{\tilde{b}_2^{\;2}}{2}) & 0 & -\dfrac{\tilde{b}_2^{\;2}}{2}\omega_b^{\;2} \\[2mm] 0 & \dfrac{\tilde{b}_2^{\;2}}{2}\omega_b^{\;2} & 0 \\[2mm] -\dfrac{\tilde{b}_2^{\;2}}{2}\omega_b^{\;2} & 0 & \dfrac{\tilde{b}_2^{\;2}}{2}\omega_b^{\;4} \end{bmatrix} \qquad (76)$$

The same with $\Delta \overline{G}_{fg}$ replaced by $\Delta \overline{G}_{gw}$ holds for C_f.

C_{d_2} and C_{d_3} are estimated experimentally using the system shown in fig. 20:

- generate series $v(k)$ and $w(k)$ of length 10'000

- determine $s_v^{\;2} = \hat{V}ar[v(k)]$ and $S_w = \hat{V}ar\left[[w(k) \quad \Delta w(k)/T_s \quad \Delta^2 w(k)/T_s^{\;2}]^T\right]$,

  where '^' denotes empirical estimates

$$- \; C_{d_i} = \begin{bmatrix} \mu_{d_i}^{\;2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \dfrac{\sigma_{d_i}^{\;2}}{s_v^{\;2}} S_w, \; i=2,3 \qquad (77)$$
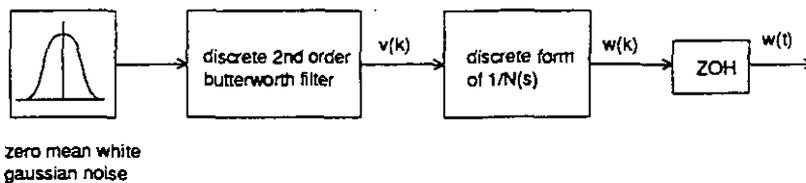
Fig. 20 System to simulate measurement errors filtered by 1/N(s)

5) Maximizing P_i (see (10)):
   P_{imax} is found to be P_{imax}=8.6, what means again a rather modest performance.

6) Observer matrices according to (11)-(15):
   (shown in rounded form, implemented in double precision)

$$\begin{aligned}
\mathbf{F} \;\; &= \begin{bmatrix} 0 & -0.000772\text{sec}^{-2} \\ 1 & -0.0556\text{sec}^{-1} \end{bmatrix} \\[2mm]
\mathbf{J(u,y)} \;\; &= \; \mathbf{TB(u,y)} = \begin{bmatrix} 0.872\text{sec}^{-1} & -0.0860\text{sec}^{-1} \\ -0.469 & 0.120 \end{bmatrix} \mathbf{B(u,y)} \\[2mm]
\mathbf{G} \;\; &= \begin{bmatrix} -0.00236\text{sec}^{-2} & 0.000118\text{sec}^{-2} \\ -0.897\text{sec}^{-1} & 0.0926\text{sec}^{-1} \end{bmatrix} \\[2mm]
\mathbf{L_1} \;\; &= [0 \;\; -1] \\[1mm]
\mathbf{L_2} \;\; &= [-0.469 \;\; 0.120]
\end{aligned} \qquad (78)$$

Exp. 4.1: Constant parameter estimation errors as assumed for design

As in exp. 1.2 implement the parameter errors

$$\Delta G_{fg}(u_b = 0) \; = \; \Delta G_{fg}(u_b = 1) \; = \; \Delta \overline{G}_{fg} = -0.1 \, \overline{G}_{fg\,0} \qquad \text{and}$$

$$\Delta G_{gw}(u_b = 0) \; = \; \Delta G_{gw}(u_b = 1) \; = \; \Delta \overline{G}_{gw} = -0.05 \, \overline{G}_{gw\,0}$$

Exp. 4.1: Constant parameter estimation errors as assumed for design

As in exp. 1.2 implement the parameter errors

$$\Delta G_{fg}(u_b = 0) = \Delta G_{fg}(u_b = 1) = \Delta \overline{G}_{fg} = -0.1\,\overline{G}_{fg\,0} \qquad \text{and}$$

$$\Delta G_{gw}(u_b = 0) = \Delta G_{gw}(u_b = 1) = \Delta \overline{G}_{gw} = -0.05\,\overline{G}_{gw\,0}$$

To generate the measurement errors use the nominal parameters $(T_s, \omega_c, \mu_{d1..2}, \sigma_{d1..2})$ as assumed for design.

As measurement errors are small compared to the ranges of the temperature measurements, they show almost no visible effects in plots as shown in fig. 4 and 5. However, the residuals in fig. 21 and 22 are rather sensitive to these errors. Note in fig. 21 that by considering measurement errors in the design, the residual has a far larger magnitude than in exp. 3.1 (compare to fig. 18). So the design procedure has reacted to possible measurement errors in the desired way. The PI is as expected rather modest and similar to the one of the LLO for T_{fg} in fig. 22.
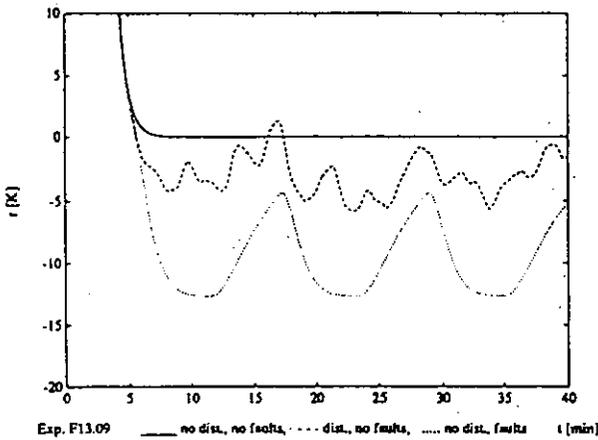
Fig. 21 Residuals of the SOFDO $(P_{im}=7.6)$ Fig. 22 Residuals of the LLO $(P_{im}=5.4)$
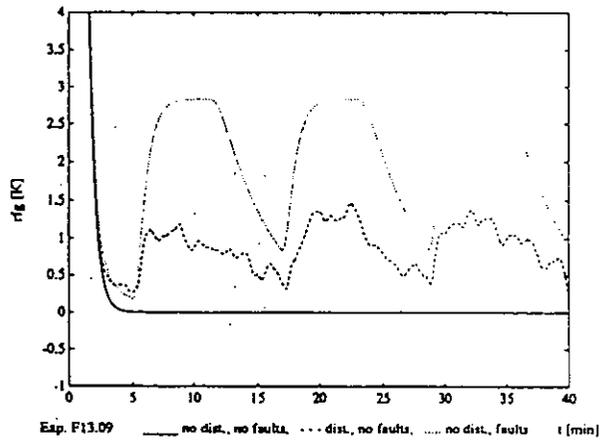
Exp. 4.2: Input dependent parameter estimation error ΔG_{fg}

Parameter errors as in exp. 4.1, but ΔG_{fg} refined according to (58).

As shown in fig. 23 the residual shows little change compared to exp. 4.1.

Doubling the errors $\Delta G_{fg\,0}$ and $\Delta G_{fg\,1}$ decreases the PI to 1.8, i.e. a rather bad performance.However, redesigning the observer assuming doubled $\Delta \overline{G}_{fg}$ gives a PI of 3.3 only. Therefore with doubled $\Delta G_{fg\,0/1}$ one can't do much better than what the original observer has reached.

Exp. 4.3: Changing parameters of the measurement errors

Parameter errors as in exp. 4.2, and in addition different deviations of the true parameters ω_c, μ_{fg} and σ_{fg} from their nominal values.

There is little change in the performance of the SOFDO, if ω_c is halfed ($\to P_{im}=6.2$), if ω_c is doubled ($\to P_{im}=4.8$), or if σ_{fg} is doubled ($\to P_{im}=5.2$). A considerable deterioration occurs if μ_{fg} is doubled, as shown in fig. 24. This deterioration is probably due mainly to the error introduced by replacing in the observer the term $B(u,x)$ by $B(u,y)$.
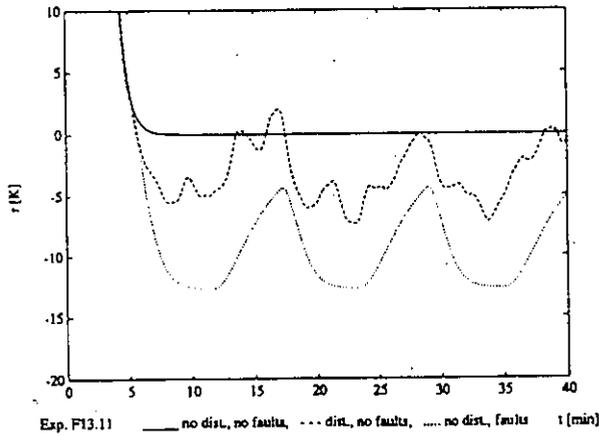
Fig. 23 Residuals of the SOFDO (P_{im}=5.5)
input dependent parameter error ΔG_{fg}
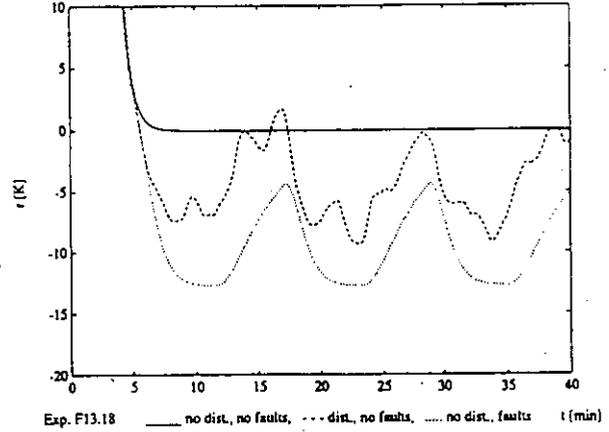
Fig. 24 Residuals of the SOFDO (P_{im}=3.3)
doubled bias in T_{fg}

4.5 Summary of the experiments

Tables 1 and 2 summarize the experiments. The first row of each experiment group summarizes the design assumptions.

exp. no.	disturbances taken into account		neglected	fault	performance		remarks
	$\Delta \dot{m}_w$	$\Delta \overline{G}_{fg}$	Δx_b	$\Delta \overline{G}_{gw}$	PI	subjective valuation	
1	external sine wave	external sine wave		external sine wave	1150		fully correlated disturbances
1.1	external sine wave	external sine wave		external sine wave	1150	very good	
1.2	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$		$-.05\,\overline{G}_{gw0}$	255	good	
1.3	$+.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$		$-.05\,\overline{G}_{gw0}$	3.2	bad	highly sensitive to violation of the assumed correlation of $\Delta \dot{m}_w$ and $\Delta \overline{G}_{fg}$
2	external sine wave	external sine wave		external sine wave	7.4		uncorrelated disturbances
2.1	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$		$-.05\,\overline{G}_{gw0}$	13	acceptable	
2.2	$+.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$		$-.05\,\overline{G}_{gw0}$	7.3	acceptable	little sensitive to sign of $\Delta \dot{m}_w$
2.3	$-.1\,\dot{m}_{w0}$	$-.1\,G_{fg00}$ $-.1\,G_{fg10}$		$-.05\,\overline{G}_{gw0}$	8.6	acceptable	
2.4	$-.1\,\dot{m}_{w0}$	$-.2\,G_{fg00}$ $-.2\,G_{fg10}$		$-.05\,\overline{G}_{gw0}$	1.9	very bad	according to redesign for doubled $\sigma\!\left(\Delta \overline{G}_{fg}\right)$ no improvement possible
2.5	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$		$-.05\,\dfrac{G_{gw00}^2}{G_{gw10}}$ $-.05\,G_{gw10}$	30	good	
2.6	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg0}$	$-.05(1-x_{b0})$	$-.05\,\overline{G}_{gw0}$	7.7	modest	

Table 1 Experiments with 1st order observers

347

exp. no.	disturbances taken into account		neglected	fault	performance		remarks
	$\Delta\dot{m}_w$	$\Delta\overline{G}_{fg}$	Δx_b	$\Delta\overline{G}_{gw}$	PI	subjective valuation	
3	external sine wave	external sine wave		external sine wave	13		
3.1	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg\,0}$		$-.05\,\overline{G}_{gw0}$	16	good	$\overline{|r|}$ too small
3.2	$-.1\,\dot{m}_{w0}$	$-.1\,G_{fg\,00}$ $-.1\,G_{fg\,10}$		$-.05\,\overline{G}_{gw0}$	2.6	bad	$\overline{|r|}$ too small
3.3	$-.1\,\dot{m}_{w0}$	$-.1\,\overline{G}_{fg\,0}$	$-.05(1-x_{b0})$	$-.05\,\overline{G}_{gw0}$	1.5	very bad	$\overline{|r|}$ too small

Table 2a Experiments with 2nd order observers, without measurement errors

exp. no.	disturbances taken into account					fault	performance		remarks
	$\Delta\overline{G}_{fg}$	e_{wb}		e_{fg}		ω_c	$\Delta\overline{G}_{gw}$	PI	subjective valuation	
		μ [K]	σ [K]	μ [K]	σ [K]	[rad/s]				
4	external sine wave	.2	.1	.5	.25	1/30	external sine wave	8.6		
4.1	$-.1\,\overline{G}_{fg\,0}$	.2	.1	.5	.25	1/30	$-.05\,\overline{G}_{gw0}$	7.6	modest	
4.2a	$-.1\,G_{fg\,00}$ $-.1\,G_{fg\,10}$	.2	.1	.5	.25	1/30	$-.05\,\overline{G}_{gw0}$	5.5	modest	
4.2b	$-.2\,G_{fg\,00}$ $-.2\,G_{fg\,10}$	.2	.1	.5	.25	1/30	$-.05\,\overline{G}_{gw0}$	1.8	very bad	redesign for doubled $\sigma(\Delta\overline{G}_{fg})$ predicts little improvement
4.3a	$-.1\,\overline{G}_{fg\,0}$	.2	.1	.5	.25	1/60	$-.05\,\overline{G}_{gw0}$	6.2	modest	
4.3b	$-.1\,\overline{G}_{fg\,0}$	.2	.1	.5	.25	1/15	$-.05\,\overline{G}_{gw0}$	4.8	rather bad	
4.3c	$-.1\,\overline{G}_{fg\,0}$	.2	.1	.5	.50	1/30	$-.05\,\overline{G}_{gw0}$	5.2	modest	
4.3d	$-.1\,\overline{G}_{fg\,0}$	.2	.1	1.0	.25	1/30	$-.05\,\overline{G}_{gw0}$	3.3	rather bad	

Table 2b Experiments with 2nd order observers, with measurement errors

In exp. group 1 good performance is achieved. However, it turns out that this is possible only due to the full correlation between $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$, which was assumed *by mistake*. The residual is highly sensitive to violations of this design assumption.

Assuming $\Delta\dot{m}_w$ and $\Delta\overline{G}_{fg}$ to be uncorrelated in exp. group 2, the performance decreases significantly, but is now more robust to deviations from the design assumptions. The bad performance in exp. 2.4 is due to the fact that the effects of $\Delta\overline{G}_{fg}$ and $\Delta\overline{G}_{gw}$ are hard to separate.

Stepping to a 2nd order observer in exp. group 3 gives modest improvement as long as the design assumptions are met. In case of faults the residual is too small and obviously would get corrupted by measurement errors. The sensitivity to deviations from design assumptions is higher than for the 1st order observers in exp. group 2.

The size of the residual in case of faults is significantly increased in exp. group 4 by taking measurement errors into account. The sensitivity to deviations from the design assumptions concerning the measurement errors is rather low. The performance stays modest because the effects of $\Delta\overline{G}_{fg}$ and $\Delta\overline{G}_{gw}$ are hard to separate.

5 Discussion of the design method

The fault detection method introduced by Frank and Wünnenberg [3] and Wünnenberg [4] is a sound approach to the robustness problem in observer based fault detection. The method leads to a fault detection observer that is optimal with respect to the defined performance index. The following points may impede its application in practice:

- A parametric process model must be set up and its parameters must be identified.

- Among the nonlinear systems only a restricted class can be handled.

- The computations involved in the design are heavy and may not suitable for on-line execution. Therefore adaptation of the observer to a changing operating point is often not possible.

- An analytical determination of the weighting matrices is difficult because it requests signal models for the disturbance and fault signals. To get such models many assumptions and gross simplifications are often necessary. Frank [2] suggested determining the weighting matrices experimentally by simulation. This suggestion is followed in exp. group 4 when determining the weigthing matrix entries stemming from the measurement errors. To experimentally determine here also the entries stemming form a parameter estimation error, say $\Delta\dot{m}_w$, one could proceed as follows:

  • Assume a distribution $f(\Delta\dot{m}_w)$.

  • Assume that $\Delta\dot{m}_w$ changes slowly with time. (Note that no on-line parameter estimation is assumed.)

  • Choose a large representative sample of $f(\Delta\dot{m}_w)$. Let N be ist size.

  • For each observation $\Delta\dot{m}_w(i)$, i-1..N, do a simulation to get a sample function

$$d_{1i}(t) = \frac{\Delta\dot{m}_w(i)\, c_w}{x_b C_b}(u_2(t) - x_1(t)), \quad t \in [0,T], \tag{79}$$

(see eq. (32)). Note that $d_{1i}(t)$ is a sample function of a *non-ergodic* stochastic process $\{d_1(t,\omega)\}$, where ω is an elementary event, to which a random variable $\Delta\dot{m}_w(\omega)$ is assigned: $\omega \to \Delta\dot{m}_w(\omega)$. Therefore the desired correlations cannot simply be determined by time averaging.

  • Filter each $d_{1i}(t)$ by $1/N(s)$, to get $\tilde{d}_{1i}(t)$ and its derivatives up to order p.

  • Determine the empirical correlations of $\tilde{d}_1 ... D^p\tilde{d}_1$ by averaging over time and over all N sample functions.

In case of correlated parameter estimation errors a multidimensional distribution would have to be assumed. Strictly speaking, the signals $d_1(t)$ and $d_2(t)$ as defined in (32) and (33) are correlated through the state vector x, even if the corresponding parameter estimation errors $\Delta\dot{m}_w$ and ΔG_{f_g} are uncorrelated, because the states depend on the parameter errors.

- The performance of the resulting observer may be sensitive to deviations from design assumptions, as e.g. deviations in mean and variance of a parameter estimation error, an unexpected parameter estimation error not considered in the design, or unmodeled process features. The robustness to such deviations should therefore be checked by simulations. This may lead to a large testing work.

6 Conclusions

The design method of Frank and Wünnenberg leads to *optimal* fault detection observers, whose residuals are maximal robustness against disturbances and maximal sensitive to faults. As the method requires a heavy design effort, one may often prefer simpler fault detection methods. In such cases the optimal fault detection observers can provide useful benchmarks.

References

[1] Frank P.M.: Enhancement of robustness in observer based fault detection, in Isermann R. (ed.): IFAC /IMACS-Symposium on fault detection, supervision, and safety for technical processes, Safeprocess 91, Baden-Baden (FRG), 1991, preprints, Vol.1, p.275-87.

[2] Frank P.M.: Personal communication, Sept. 1992.

[3] Frank P.M., Wünnenberg J.: Robust fault diagnosis using unknown input observer schemes, in Patton R.J., Frank P.M., Clark R.N.: Fault diagnosis in dynamic systems, theory and applications, Prentice Hall, 1989.

[4] Wünnenberg J.: Observer based fault detection in dynamic systems, VDI-Fortschrittberichte, Reihe 8, Nr.222, Düsseldorf, 1990.

[5] Gass J., Hopkirk N.: Analysis of the boiler, IEA-Annex 10, system simulation, Exercises on the building 'La chaumiere', 1985.

[6] Sprecher P.: Unknown input fault detection observers, report AN25/CH/920327/1, IEA-Annex25, 1992.

FAULT DETECTION IN THERMAL STORAGE TANK USING PHYSICAL MODEL

Kazunobu Sagara

Department of Architecture, Faculty of Engineering,

Mie University, Kamihama, Tsu 514, Japan

Nobuo Nakahara

Department of Architecture, Faculty of Engineering,

Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464, Japan

Abstract

A method of fault detection in thermal storage tank and results of applying to real thermal storage systems are presented. In the fault detection, a physical model is used for a water thermal storage tank of multi-connected complete mixing type, and unknown parameters of the physical model are identified from measured data. In this report, simulation study was conducted in which several faults in storage tank were tried to detect. Furthermore, some primitive operation faults were discussed, and analysis results of identified parameters of the model are presented for measured data in real buildings. The identified model parameters varied widely because of modeling accuracy and measuring error, so statistical methods were applied in order to detect fault conditions.

1 INTRODUCTION

Thermal storage systems are being widely used in Japan, the U. S. and other countries. They contribute to the effective use of energy; peak shift of electrical demand, heat recovery, solar energy utilization and seasonal storage. Water, ice and other phase change materials are used as the thermal storage media. Water thermal storage has a long history in Japan and most of the large scale applications have used water[1][2][3].

A method of fault detection in thermal storage tank and results of applying to real thermal storage systems are presented in this report. In the fault detection, a physical model is used for a water thermal storage tank of multi-connected complete mixing type, and unknown parameters of the physical model are identified from measured data. Faults can be detected by comparing value of identified parameters with value of the identified parameters under normal condition. In this report, simulation study was conducted in which some faults in storage tank such as insulation damage of tank wall or abnormal water level in tank were tried to detect. Furthermore, some primitive operation faults appearing just after completion of building were discussed, and analysis results of identified parameters of the model for measured data in real buildings are presented.

351
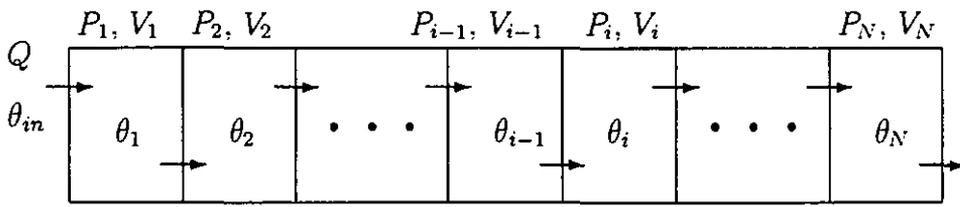
Fig. 1 Thermal storage tank consisting of divided tanks connected in series

Heat gain from tank wall in the physical model for parameter identification is divided into two parts as cyclic and constant heat gain for shorter simulation time of parameter identification, while unsteady one-dimensional heat conduction in concrete tank wall and ambient soil was applied in a physical model generating fault data for simulation study.

2 FAULT DETECTION METHOD

2.1 PHYSICAL MODEL OF THERMAL STORAGE TANK

Type of water storage tank is roughly classified into two type; the multi-connected complete mixing tank and the temperature-stratified tank. The former type is more popular in Japan because it was the first that was developed in Japan, though storage performance is below the latter type. Many buildings in Japan have space under the basement floor which is divided by high tie beams for protection against earthquakes. The divided space has been used for water thermal storage of the multi-connected complete mixing type. So, the physical model for the multi-connected complete mixing type is used in this study. Assumption in modeling is as follows.

- In each divided tank, one part is useful to thermal storage (effective tank volume) and the other part is useless (dead zone), and the effective volume ratio, P, is defined as the ratio of effective tank volume to tank volume in each divided tank.
- Water mixed completely in the effective tank volume enters into the next divided tank.
- Water density and specific heat are constant regardless of water temperature.
- Heat gain from tank wall is divided into two parts as cyclic and constant heat gain.

Under this assumption, water temperature in the ith divided tank of thermal storage tank is calculated with the following equation (Fig. 1).

$$\rho c_p P_i V_i \frac{d\theta_i}{dt} = \rho c_p Q(\theta_{i-1} - \theta_i) + U_i A_i (\theta_{amb,i} - \theta_i) + q_i A_i \qquad (1)$$

where

$$\theta_0 = \theta_{in}, \qquad V = \sum_{i=1}^{N} V_i, \qquad i = 1, \cdots, N \qquad (2)$$

and i is tank number, ρ water density, c_p water specific heat, P_i effective volume ratio of ith divided tank, V_i volume of ith divided tank, θ_i water temperature in ith divided tank, t time, Q flow rate, $\theta_{amb,i}$ ambient temperature of ith divided tank, U_i over-all coefficient of heat transfer of ith divided tank wall, A_i surface area of ith divided tank wall, q_i constant heat gain from wall in ith divided tank, N number of divided tanks, V volume of the whole tank system, θ_{in} water temperature entering the end divided tank. This model equation is non-linear due to variable flow rate (Q).

In this mixing model, cyclic heat gain is assumed to be proportional to difference between ambient temperature and water temperature in each divided tank. The cyclic heat gain is related to charging and discharging of heat into concrete tank wall due to cyclic change of water temperature. Ambient temperature, $\theta_{amb,i}$, is assumed to be equal to daily average of measured water temperature in each divided tank. Constant heat gain, i.e. the last term of Equation (1), is related to heat gain from ceiling of tank and heat gain difference between daily average water temperature and actual ambient temperature. Three parameters, P_i, $U_i A_i$ and $q_i A_i$, of the model have to be identified for fault detection.

2.2 PARAMETER IDENTIFICATION OF PHYSICAL MODEL

In this study, P_i, $U_i A_i$ and $q_i A_i$ are assumed to have the same value in all divided tanks to realize identification only with measured water temperature in several divided tanks as follows.

$$
\begin{array}{ccccccc}
P_1 & = & P_2 & = & \cdots & = & P_N & = & P \\
U_1 A_1 & = & U_2 A_2 & = & \cdots & = & U_N A_N & = & UA \\
q_1 A_1 & = & q_2 A_2 & = & \cdots & = & q_N A_N & = & qA
\end{array}
$$

Then, P, UA and qA which realize the closest temperature profile to the measured profile are obtained by using the modified Powell method[4][5].

Identification procedure is as follows.

1. Initial water temperature in each divided tank is interpolated from water temperature of several measured divided tanks.

2. Initial value of model parameters is set. In this study, $P = 1.0$ and $UA = qA = 0.0$.

3. Water temperature in each divided tank at each measuring time scan is calculated using the physical model of storage tank under measured input temperature and flow rate data.

4. Temperature errors in measured divided tanks are calculated and summed up as:

$$
\text{Error} = \sum_{\substack{\text{Measured} \\ \text{tanks}}} \sum_{\substack{\text{Scan} \\ \text{times}}} (\text{Measured temperature} - \text{Calculated temperature})^2
$$

5. New value of P, UA and qA is set with the modified Powell method.

6. Above procedure is repeated until minimum error is achieved.

3 SIMULATION STUDY

3.1 SYSTEM AND OPERATION CONDITION

The thermal storage system for simulation study is assumed to be composed of a water thermal storage tank of the multi-connected complete mixing type, a chiller and an air handling unit (AHU). The storage tank is divided into 26 sub-tanks connected in series and its total volume is 400 m³. The thermal storage system is shown in Fig. 2 schematically.

In the storage tank, the ceiling, bottom and outer wall are insulated, but partition walls between divided tanks have no insulation. Water to the chiller is pumped up from the lower and higher temperature end tanks, and chilled water is put into the lower temperature end tank. Water to the AHU is pumped up from the lower temperature end tank and returns to the higher temperature end tank.

Assumption in system simulation is as follows.
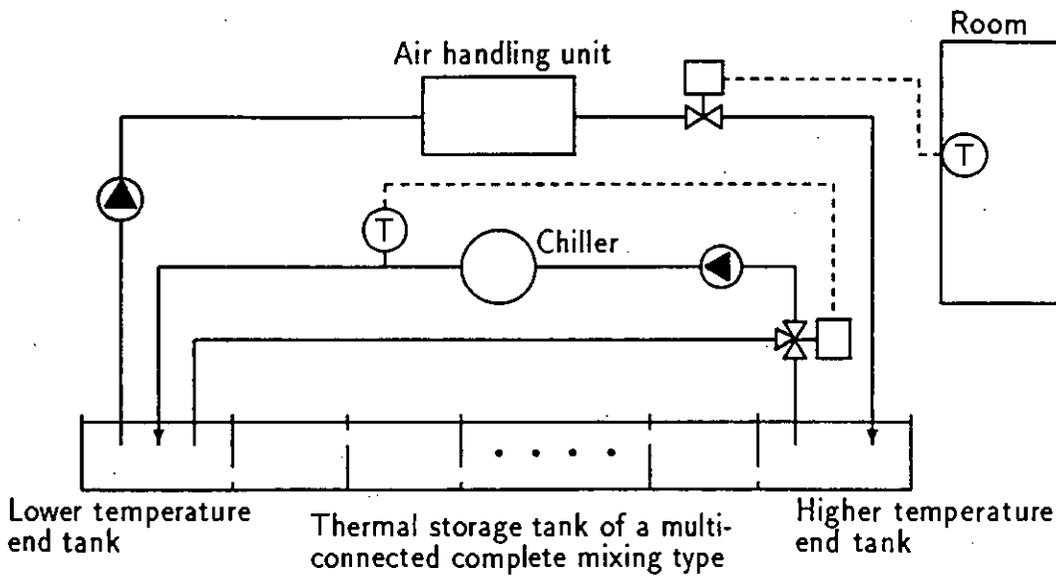
Fig. 2 Thermal storage system for simulation study

- Input/output temperature difference of chiller is always 5 °C regardless of the chiller COP.
- Input temperature of chiller is normally controlled to 11 °C, and output temperature of chiller is 6 °C normally.
- Input/output temperature difference of AHU is always 7 °C regardless of the cooling coil characteristics, and output temperature of AHU is calculated from the input/output temperature difference.
- Water temperature in each divided tank for no operation time also changes due to heat gain from tank wall.

3.2 DATA GENERATION FOR SIMULATION STUDY

Assumption in modeling of storage tank for data generation is the same as modeling for fault detection except of heat gain from tank wall following as:

- Heat gain from tank wall is unsteady heat transfer, and one-dimensional heat conduction in concrete tank wall and ambient soil is applied in each divided tank.
- Heat capacity of insulation material and air is neglected, and thermal resistance between water and surface of tank wall is neglected.

Under this assumption, water temperature in the ith divided tank of thermal storage tank is calculated with the following equation.

$$\rho c_p P_i V_i \frac{d\theta_i}{dt} = \rho c_p Q(\theta_{i-1} - \theta_i) + \sum_{j=1}^{6} q_{ij} A_{ij} \tag{3}$$

where j is wall number in each divided tank, q_{ij} heat gain from jth wall of ith divided tank, A_{ij} surface area of jth wall of ith divided tank.

Temperature in concrete tank wall and ambient soil is assumed to be governed by the unsteady one-dimensional heat conduction equation. Temperature in concrete and soil was calculated with the finite difference method for each wall of each divided tank under constant soil temperature at 1 m distance from concrete wall.
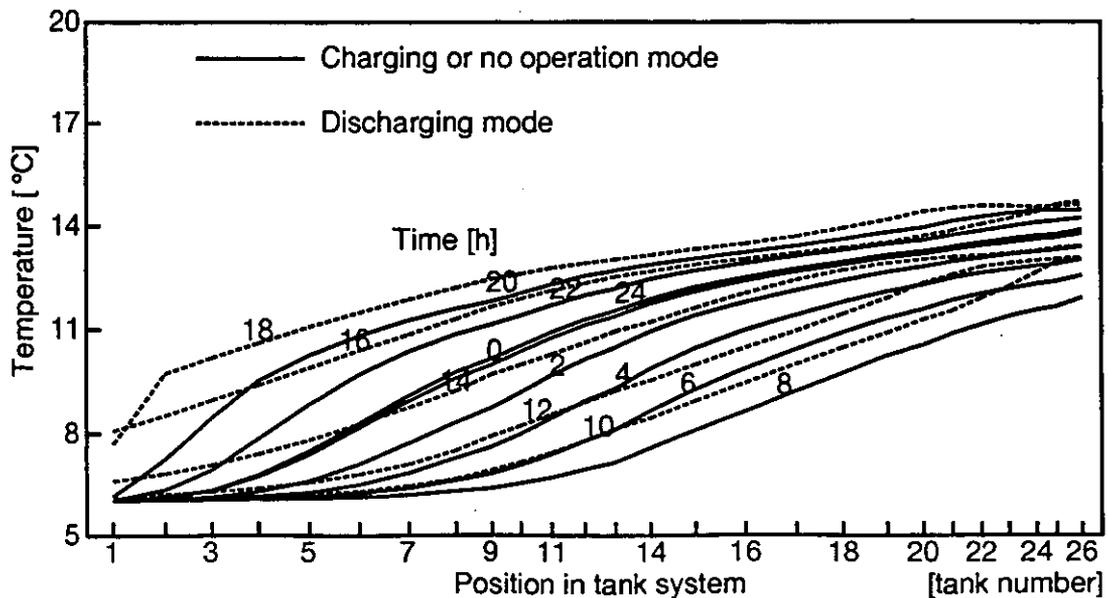
Fig. 3 Temperature profile in storage tank under normal condition

Generated data to apply fault detection is the 8th day data of the simulation started with 11 °C uniform temperature in the whole storage tank. In this thermal storage system, the chiller is operated in 21 hours from 0:00 to 13:00 and from 16:00 to 24:00, and the AHU is operated in daytime from 8:00 to 18:00. Fig. 3 shows the change of temperature profile in the storage tank under normal condition.

Generated data are assumed to be available only at measuring scan time of every one hour. So, input conditions between the measuring scans are interpolated linearly in the simulation for parameter identification. The divided tanks measuring water temperature are assumed to be the 1st, 4th, 7th, 11th, 15th, 17th, 21st and 26th divided tanks in this simulation study.

3.3 PARAMETER IDENTIFICATION UNDER FAULT CONDITION

The result of parameter identification under normal condition is the reference for fault detection, which is compared with identification results under fault conditions. If the system has any faults, the identified value of model parameters will change.

Variation of fault conditions in this simulation study is as follows.

- Damage of all insulation in the 10th divided tank
- Ceiling insulation damage in all divided tanks
- Damage of all insulation in all divided tanks
- Abnormally low water level ($P = 0.80$)
- Abnormal measured temperature in the 11th divided tank due to sensor damage (the measured temperature is correct temperature + 3°C)

Results of parameter identification are listed in Table 1. Daily average water temperature in the whole tank system, θ_{ave}, is also shown in this table. The result of insulation damage only in the 10th tank is not so different from that of normal condition, and it is difficult to detect insulation damage in one divided tank. In case of ceiling insulation damage in all divided tanks, the constant heat gain, qA, increases remarkably. The cyclic heat gain, UA, in damage of all insulation in all tanks is larger than that in only ceiling insulation damage, but the damage effect of side wall and bottom insulation is not so much

Table 1 Parameter identification results

Condition	P	UA [*1]	qA [*2]	θ_{ave} [*3]
Normal condition	0.953	505	66	10.5
Damage of all insulation in 10th tank	0.953	512	75	10.5
Ceiling insulation damage in all tanks	0.952	530	282	10.7
Damage of all insulation in all tanks	0.965	683	266	10.8
Abnormal water level($P = 0.80$)	0.802	437	62	10.5
Abnormal measured temperature in 11th tank[*4]	0.912	818	171	10.6
No heat gain	0.933	317	33	10.4

*1 : [kcal/h°C]   *2 : [kcal/h]   *3 : [°C]   *4 : correct temperature + 3°C

because only outer side wall has insulation and partition wall between divided tanks has no insulation. Abnormal water level is detected easily because water level is related to the effective volume ratio, P, only. Abnormal measured temperature in one divided tank due to sensor damage has similar effect to increase of heat gain from tank wall due to insulation damage, but sensor damage can be distinguished by checking water temperature in each divided tank. In case of no heat gain, it is expected that $P = 0.95$, $UA = 0$ and $qA = 0$, and their differences are the result of identification error. Identification error may be attributed to modeling of heat gain, convergence error, interpolated initial temperature profile, and non-linear characteristics of the model equation.

When this FDD method is applied to real thermal storage systems, the value of identified parameters varies widely in general because of measuring error, so statistical judgment will be effective to detect fault conditions.

4 APPLICATION TO REAL THERMAL STORAGE SYSTEMS

Measured data in real thermal storage systems were used to test the performance of FDD method. Analysis results of identified parameters of the physical model are presented for an office building and a hospital building in this study.

4.1 OFFICE BUILDING

4.1.1 HVAC system

The data used in this study are a part of data recorded in an existing office building constructed in 1988. Outline of the building and HVAC system are listed in Table 2. The thermal storage system is composed of a water thermal storage tank of the multi-connected complete mixing type, three heat pumps and three air-conditioning systems.

Output temperature of two heat pumps is controlled to be constant with two two-way valves, and output temperature of chiller is controlled with a two-way valve (Variable water volume control, VWV). Control of AHU circuit is also VWV control, and the flow rate is controlled with a two-way valve according to supply air temperature or room air temperature. Additional heat pump operation is determined according to temperature in the lower temperature end tank. The tank is divided into 26 sub-tanks connected in series and its total volume is 400 m³.

Measured data for every fifteen minutes are input/output temperature and flow rate for each heat pump and each air handling unit, and water temperature in 8 divided tanks among 26 tanks. Data calculated

Table 2 Outline of the office building and HVAC system

Site	Tokyo, Japan
Structure	Steel and reinforced concrete
Building area	2404 m^2
Total floor area	9784 m^2
Number of stories	4 floors and 1 basement floor
Main building use	Office
Heat pump	Air source heat pump (heat recovery type), 80 h.p.
	Air source heat pump, 80 h.p.
	Air source chiller for cooling, 80 h.p.
Air conditioner	Single duct VAV systems for ordinary rooms
	Single duct + fancoil unit system for computer room
Storage tank for chilled water	Multi-connected complete mixing type
	26 divided tanks
	400 m^3

from the measured data are input temperature of the end tank, and total flow rate to heat pumps and air handling units.

Since flow rate and input temperature are available only for every fifteen minutes, these input conditions between the measuring scans are interpolated linearly.

Operation condition in 1 August 1989 is shown in Fig. 4. The cooling load consists of daytime load for ordinary rooms and 24 hours load for computer room. Heat pump operation in nighttime has priority over daytime operation because of cheaper night electric cost. Additional operation of heat pump started at 15:00 in this case because of temperature rise in the lower temperature end tank. This is the standard operation condition under full cooling load for the thermal storage system of this office building.

4.1.2 Some primitive operation faults

Operation conditions in 21 June 1989 are shown in Fig. 5 as an example of some primitive operation faults. These faulty conditions appeared because the thermal storage system was insufficiently controlled just after completion of the building, and the system was controlled successfully from July 1989. Primitive operation faults which can be found in this operating condition are as follows.

- High input temperature from heat pumps at the start of heat pumps
- Excessive stored chilled water under low cooling load condition
- Frequent daytime operation of heat pumps and insufficient nighttime operation

Input temperature from heat pump should be adequately low during charging mode. In this system, the thermal storage tank was settled under the basement floor, and the heat pumps were installed on the roof. So, input temperature from the heat pumps is kept high at the start of heat pumps because of heat capacity of long pipes. It is difficult to improve this fault, but, if heat pumps don't repeat a start and stop action so often, the effect of this fault on temperature of stored chilled water can be neglected.

Chilled water was stored too much in the tank, and only a little stored water was used for cooling. Excessive stored chilled water causes unnecessary energy loss through tank wall.

The heat pumps were operated frequently in daytime in spite of full charge, and nighttime operation is insufficient in spite of cheaper electric cost. Set point of temperature for additional operation of heat pumps may be too low.
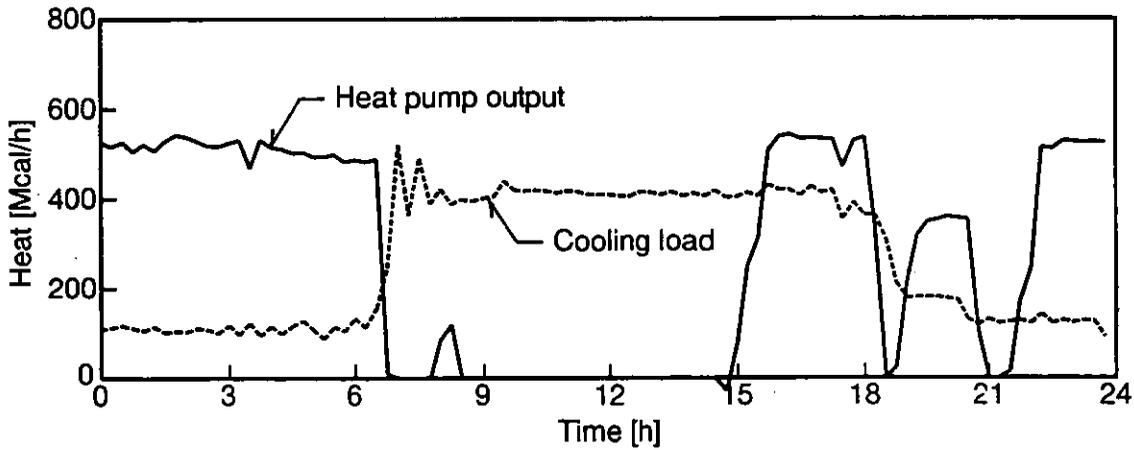
a. Heat pump output and cooling load
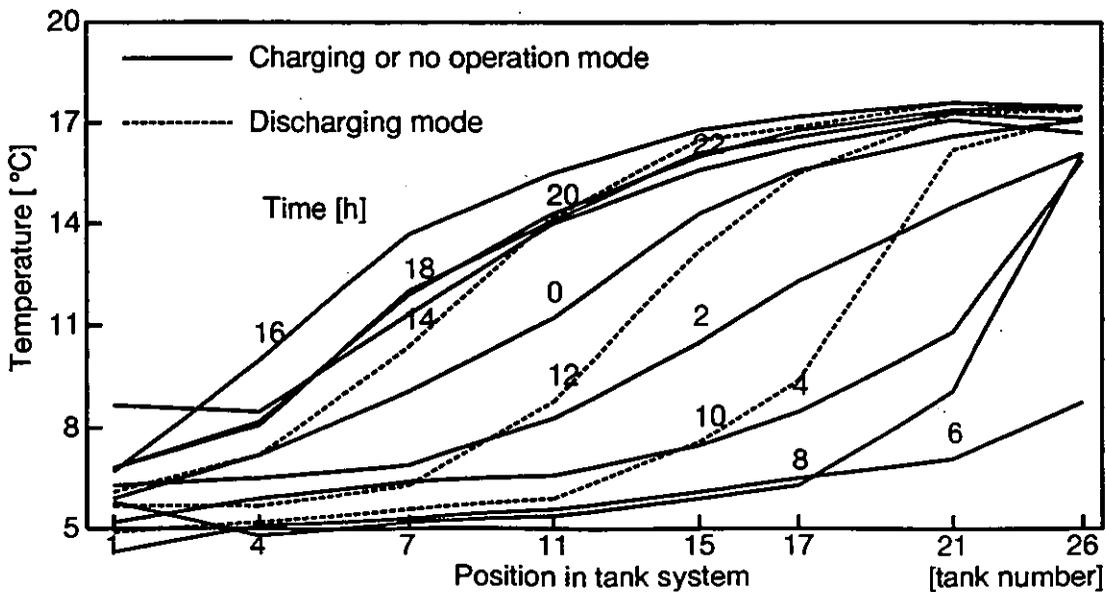
b. Temperature profile in storage tank

Fig. 4 Operation condition under full cooling load (1 Aug., 1989)

These primitive operation faults may be detected easily by checking input temperature from heat pump, stored heat in the tank and ratio of nighttime stored heat to total stored heat.

4.1.3 Analysis of identified parameters

Model parameters were identified using the measured data of 92 days from 1 July to 30 September in 1989. Fig. 6 shows temperature change in each measured divided tank with identified parameters in 1 August 1989. This figure shows comparison of measured temperature change with calculated temperature change under identified parameters.

The distribution curve of the identified parameter was assumed to be a normal distribution with the mean and variance obtained from identification results. Histograms of each identified parameter are shown in Fig. 7, in which the value of parameters on the outside of both 1 % are excluded. The excluded parameter
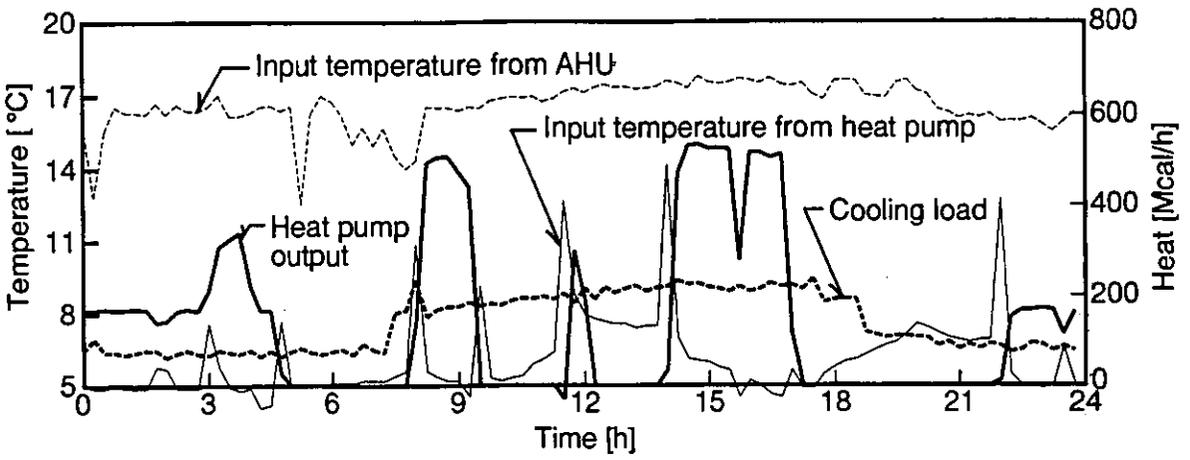
a. Input temperature, heat pump output and cooling load

b. Temperature profile in storage tank

Fig. 5 Operation condition before adjustment (21 Jun., 1989)

values are listed in Table 3. Every identified parameters of these abnormal days have extreme value, and especially, qA has extreme value. These extreme values were found to be attributed to abnormal measured data of temperature in tank and inlet except an unknown day.

Histograms of each identified parameter separated in Sunday/holiday and weekday are shown in Fig. 8. It is found that the identified parameters in Sunday/holiday have different value from that in weekday. So, t test was tried to judge whether this is true or not. The t test is a statistical tool for judging whether the two normal distributions have the same mean value or not in this case. Result of the t test is shown in Table 4, and it is found that there is a significant difference between the two mean values, that is, P in Sunday and holiday is smaller than that in weekday and qA in Sunday and holiday is larger than that in weekday.

Cooling load is generally low in Sunday and holiday, and correlation analysis was tried to examine the effect of some operation conditions on the value of P and qA. These results are shown in Fig. 9. These

Fig. 6 Temperature change in each measured divided tank with identified parameters (1 Aug., 1989)
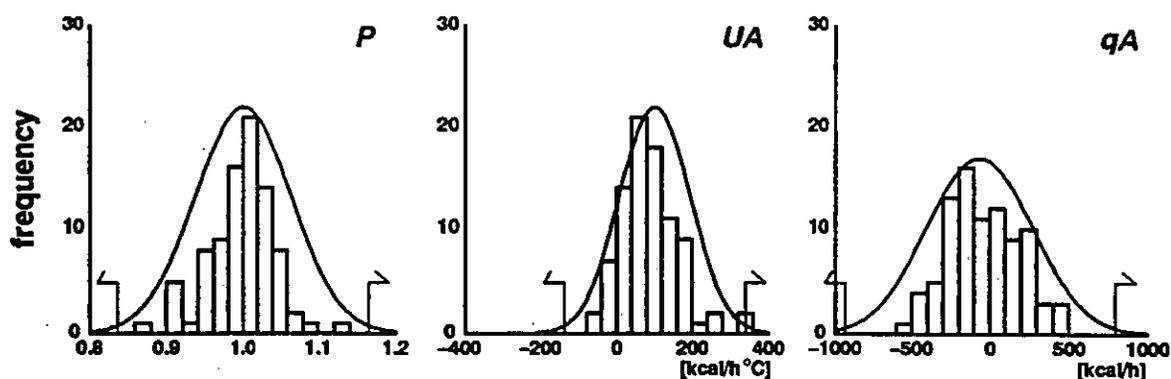
Fig. 7 Histograms of identified P, UA and qA in office building

Table 3 List of extreme parameter values

Date	P	UA	qA	Cause
14 Jul.	1.205	305	-1082	Abnormal measured data
31 Jul.	0.800	498	1273	Unknown (false alarm)
14 Sep.	1.185	191	-1352	Abnormal measured data
15 Sep.	1.233	397	-1137	Abnormal measured data

correlation coefficients have comparatively high value, and it is found that P is smaller in case of low cooling load and large difference of storage temperature, and qA is larger in case of low storage temperature. The small P can be attributed to the lower mixing level in each divided tank because of strong buoyancy caused by low flow rate and large temperature difference. The large qA can be attributed to larger heat gain from tank wall because of large temperature difference.
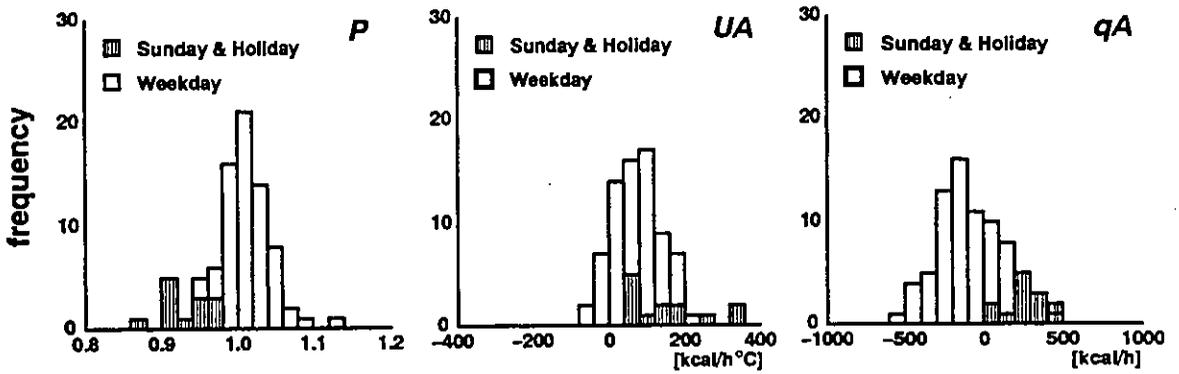
Fig. 8 Histograms of identified P, UA and qA in Sunday/holiday and weekday

Table 4 Result of t test for Sunday/holiday and weekday

| Parameter | Mean | | t value | Significance |
	Sunday/holiday	Weekday		level
P	0.932	1.010	8.09	1 %
qA	269	-96	6.34	1 %
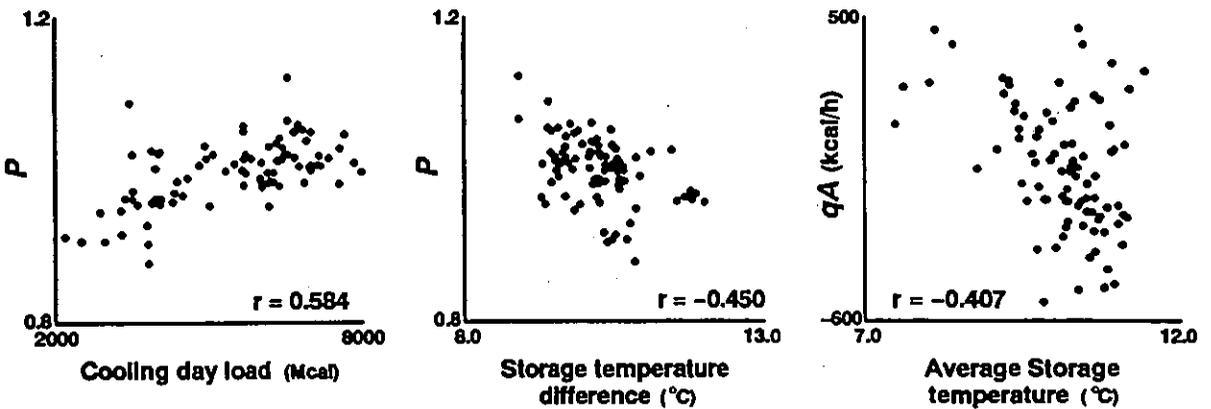
Fig. 9 Relationship between some operation conditions and identified parameters

4.1.4 Fault detection of insulation damage

Fault data of insulation damage or abnormal water level were unavailable from the recorded data in this office building. In this study, fault data of insulation damage were generated from the measured data. Generally, insulation damage causes temperature rise in each divided tank, so temperature in each divided tank is assumed as follows.

$$\theta_i^f = \theta_i^m + 0.1 \qquad [^\circ C] \qquad (4)$$

where θ_i^f is temperature in ith divided tank under insulation damage condition, and θ_i^m is measured temperature. This temperature rise corresponds to the ceiling insulation damage in all divided tanks as shown in Table 1.
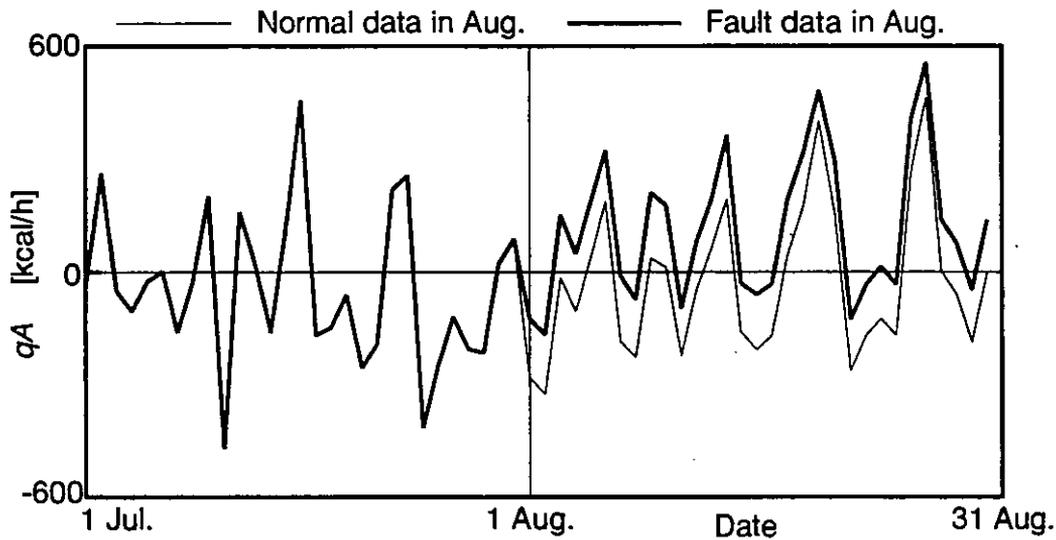
Fig. 10 Change of qA identified under normal condition and insulation damage condition
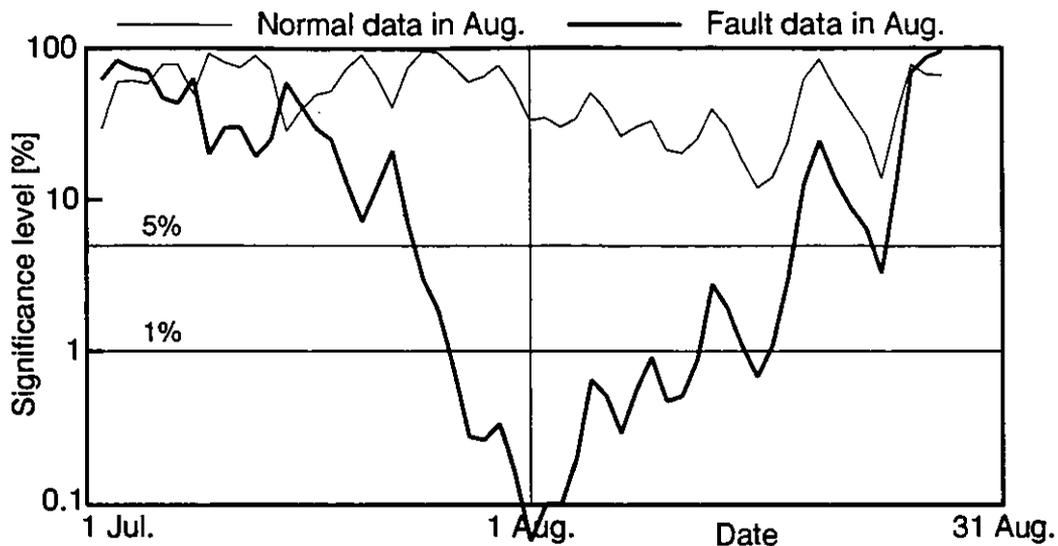
Fig. 11 Change of the significance level of t test for identified qA

Fig. 10 shows the change of identified qA in which the data in July are kept normal and the data in August are assumed to be under insulation damage condition. Identified qA in August under normal condition is also shown in this figure. The difference between qA under normal condition and that under insulation damage condition is smaller than the day change of qA. So, it is difficult to detect the insulation damage by checking the value of qA itself.

Statistical approach is useful in this case. The data of identified qA were divided into two groups before and after a specific boundary day, and whether the two groups have the same mean value or not was judged using t test. The significance level of t test was calculated from the identified qA data divided at the boundary day of two groups. The change of the significance level is shown in Fig. 11 according to the boundary day. The significance level has the peak value in 1 August when insulation damage happened.

362

Table 5 Outline of the hospital building and HVAC system

Site	Tokyo, Japan
Structure	Steel and reinforced concrete
Total floor area	17,000 m^2
Number of stories	6 floors and 2 basement floor
Main building use	Hospital (200 beds)
Chiller	Turbo chiller (300 USRT) $\times$ 2
Air conditioner	AHU + FCU for sickrooms
	AHU of intake outside air + terminal AHU for consulting rooms
	AHU + reheater for operating rooms
Storage tank for chilled water	Multi-connected complete mixing type 20 divided tanks 780 m^3

4.2 HOSPITAL BUILDING

4.2.1 HVAC system

The data used in this study are recorded in an existing hospital building constructed in 1979. Outline of the building and HVAC system of the hospital are listed in Table 5. The thermal storage system is composed of a water thermal storage tank of the multi-connected complete mixing type, two chillers and various air-handling systems and fan-coil units (FCU).

Output temperature of two chillers is controlled to be constant with a three-way valve. Control of AHU and FCU circuit is VWV control. The flow rate is controlled with a two-way valve according to supply air temperature or room air temperature, and supply water temperature to FCU is controlled to be constant with a three-way valve. The thermal storage tank is divided into 20 sub-tanks connected in series and its total volume is 780 m^3.

The data used in this study were measured from 1 May to 20 August 1995, though the data from 19 June to 5 July were excluded because the data collection system was broken down.

Measured data for every ten minutes are input/output temperature and flow rate for each chiller and the secondary circuit, and water temperature in 14 divided tanks among 20 tanks. Data calculated from the measured data are input temperature of the end tank and total flow rate to chillers, AHU and FCU.

The sensor mounting fault for input temperature of the end tank from two chillers and secondary circuits was found through inspection of the measured temperature. The sensor mounting fault was recovered at 22 July except for one chiller. The sensor of input temperature from this chiller is kept under fault condition. So, the input temperature for the simulation of parameter identification was assumed to be equal to the input temperature from the another chiller.

Operation condition in 31 July 1995 is shown in Fig. 12. The cooling load consists of daytime load and 24 hours load, and chiller operation in nighttime has priority over daytime operation. Additional operation of chiller started at 9:00 and 15:00 in this case because of stored heat shortage in tank. This is the standard operation condition under full cooling load for the thermal storage system of this hospital building.

4.2.2 Analysis of identified parameters

Model parameters were identified using the measured data of 95 days from 1 May to 20 August in 1995. Fig. 13 shows temperature change in several measured divided tanks with identified parameters in 31 July
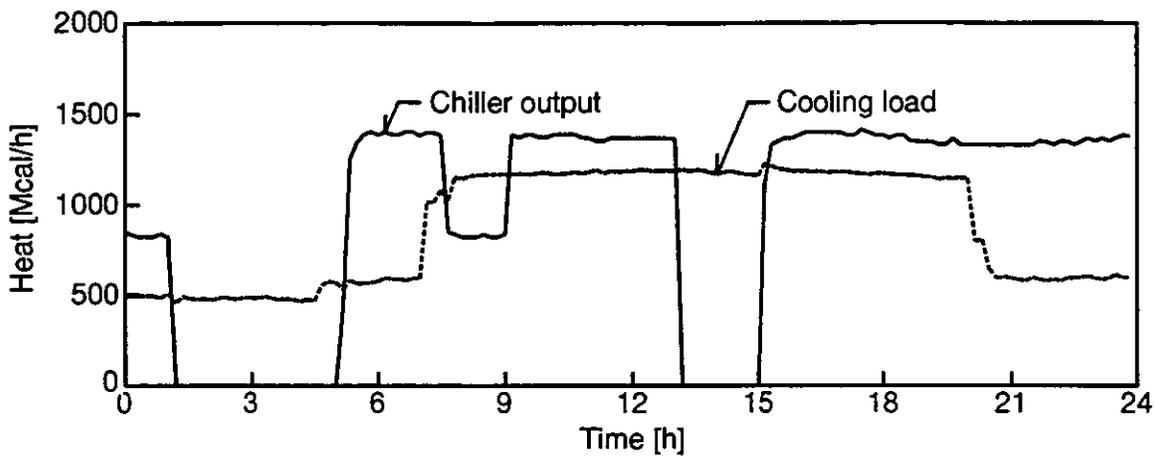
a. Heat pump output and cooling load
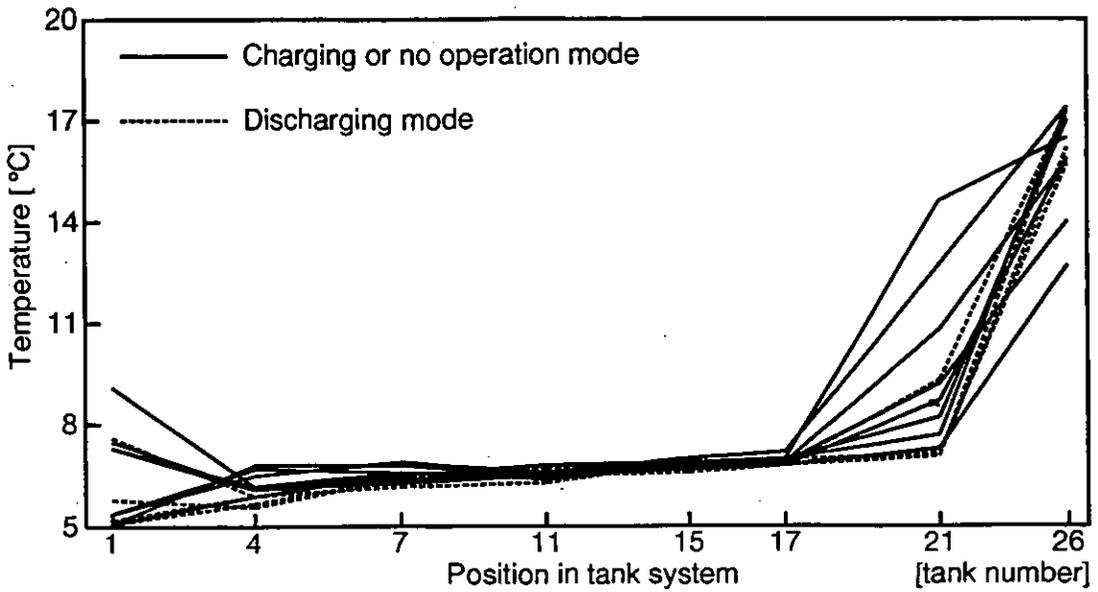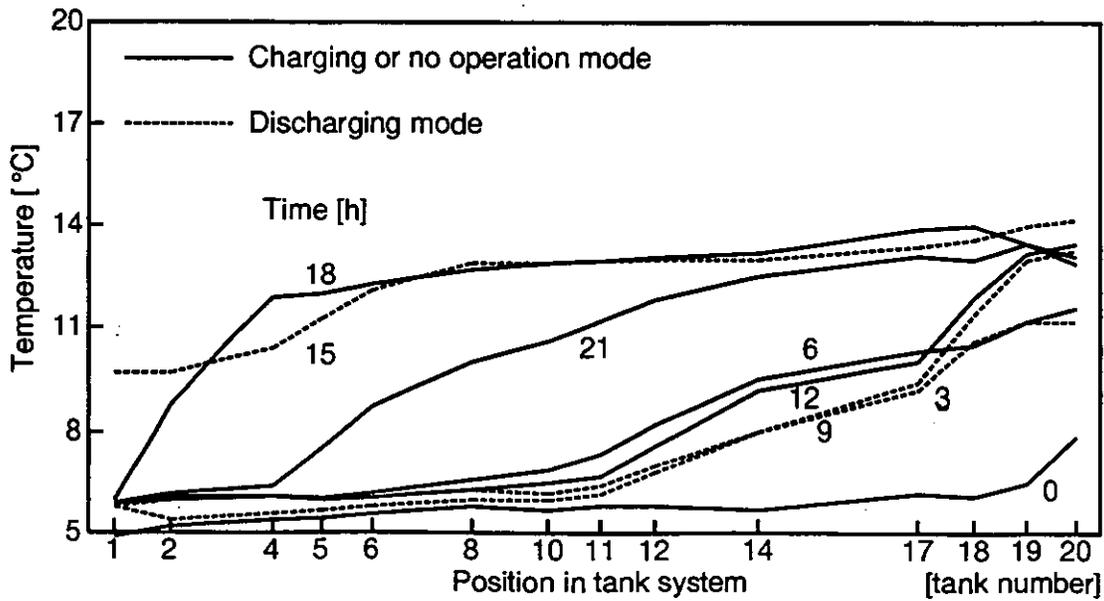

b. Temperature profile in storage tank

Fig. 12 Operation condition under full cooling load (31 July, 1995)

1995.

Histograms of each identified parameter are shown in Fig. 14. The value of identified parameters for some days were unavailable because the data transmission fault was happened and the simulation could not be conducted due to abnormal value of temperature and flow rate.

The value of identified model parameters varies widely because of modeling accuracy and measuring error. So, statistical tests were applied in order to detect fault conditions in this case. Thompson test is a statistical test in which a particularly far value can be detected in the identified value of each parameter. Histograms of significance level of Thompson test for each identified parameter are shown in Fig. 15. The significance level higher than 1% or more strict than 1% were obtained in 14 June in case of P, in 19 and 20 July in case of UA and no extreme value in case of qA.

Table 6 shows a list of extreme parameter values. The cause of the extreme value of 3 days in the list is unknown because neither abnormal water level nor insulation damage of tank wall happened. It may
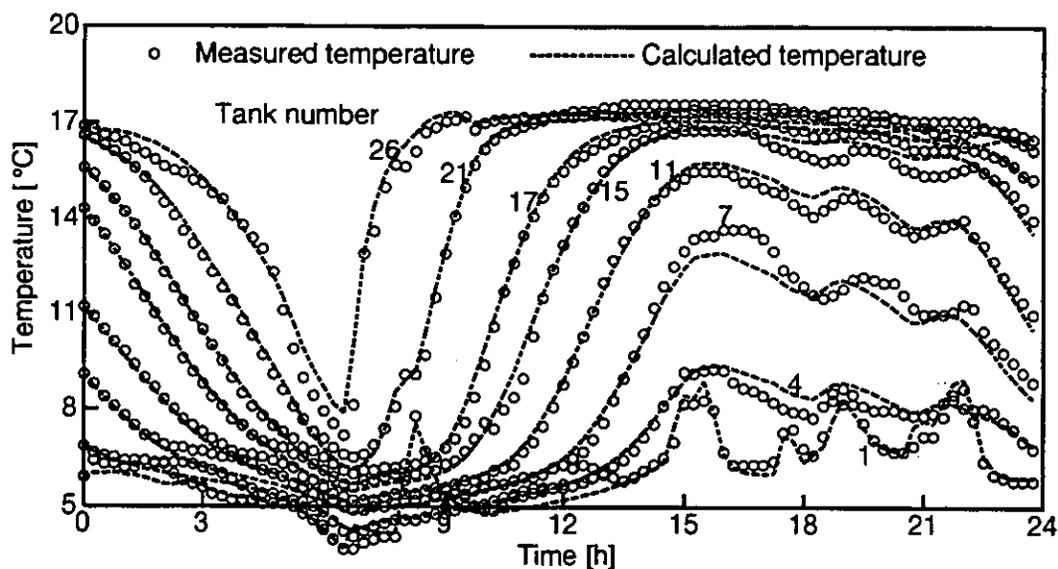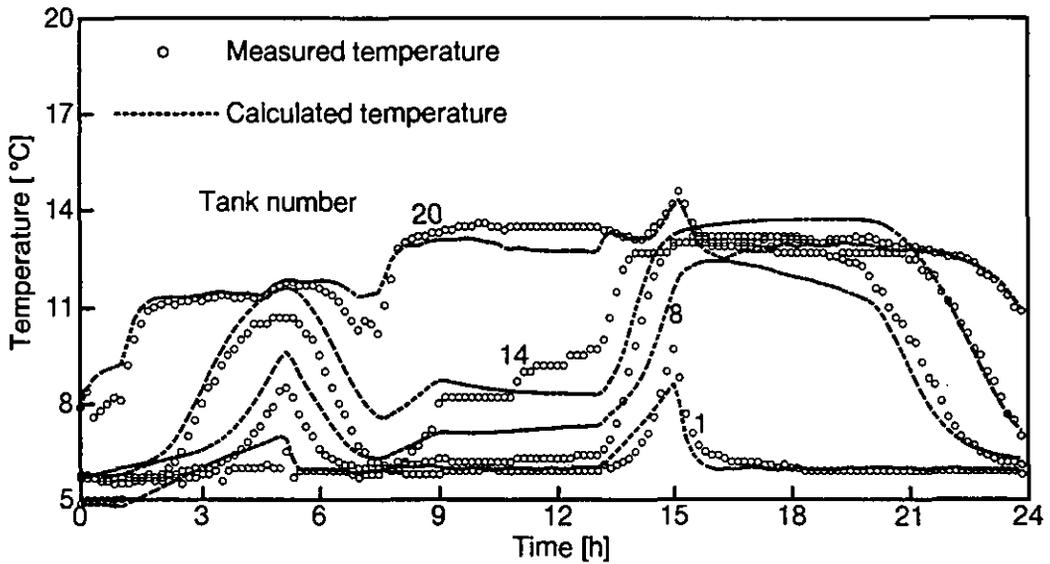
Fig. 13 Temperature change in each measured divided tank with identified parameters (31 July, 1995)
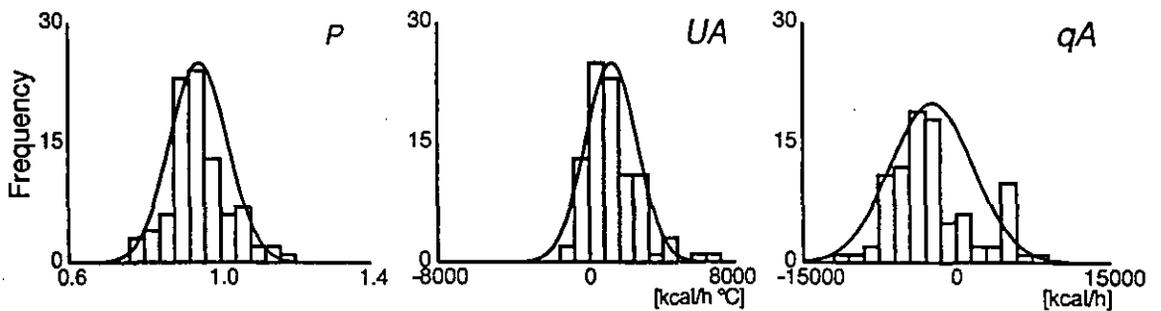
Fig. 14 Histograms of identified P, UA and qA in hospital building

attribute to the sensor mounting fault before 22 July when the fault was partially recovered. The value of identified parameter is unavailable in 4 days because of the data transmission fault. The simulation for these 4 days cannot be conducted due to abnormal flow rate and temperature, and so this fault could be detected.

4.2.3 Fault detection of insulation damage

Fig. 16 shows the change of identified qA related to constant heat gain from tank wall. And, the shift of identified qA suggests the insulation damage of tank wall. The identified qA has negative value before 22 July and positive value after 22 July. But, any insulation damage of tank wall did not happen in this day. The 22 July is the day when the sensor mounting fault for input temperature into the storage tank was recovered. The shift of the input temperature apparently resulted in the shift of heat flow through tank wall. P and UA did not shift in 22 July. So, the shift of input temperature into storage tank due to sensor fault affects the value of identified qA only.
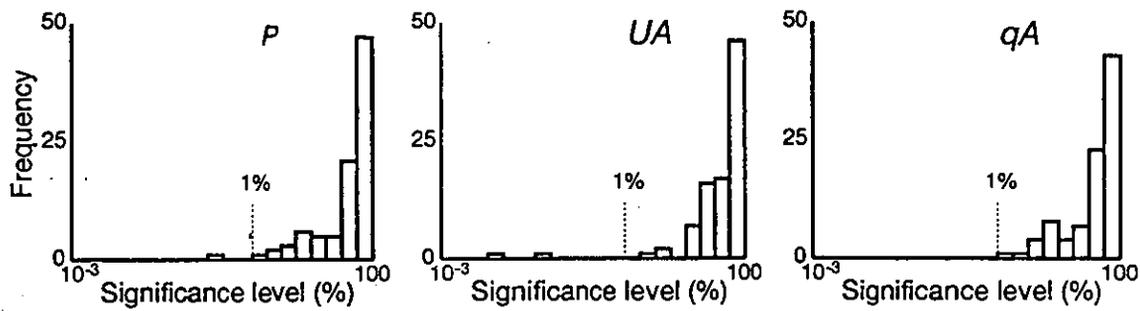
Fig. 15 Histograms of significance level of Thompson test for each identified parameter

Table 6 List of extreme parameter values

Date	P	UA	qA	Cause
6 Jun.	—	—	—	Abnormal measured data
14 Jun.	1.172	4077	-5634	Unknown (false alarm)
18 Jul.	—	—	—	Abnormal measured data
19 Jul.	0.913	6555	-4769	Unknown (false alarm)
20 Jul.	0.989	5929	-8416	Unknown (false alarm)
29 Jul.	—	—	—	Abnormal measured data
3 Aug.	—	—	—	Abnormal measured data

5 CONCLUSION

1. A method of fault detection in thermal storage tank and results of applying to a real thermal storage system are presented.

2. In the fault detection, a physical model is used for a water thermal storage tank of multi-connected complete mixing type, and unknown parameters of the physical model are identified from measured data. Faults can be detected by comparing value of identified parameters with value of the identified parameters under normal condition. In the physical model for fault detection, heat gain is divided into two parts as cyclic and constant heat gain.

3. Simulation study was conducted in which some faults in storage tank were tried to be detected, and it was confirmed that various insulation damages of tank wall and abnormal water level can be detected successfully from parameter identification.

4. Some primitive operation faults were discussed and analysis results of identified parameters are presented for measured data in a real system of office building. The distribution curve of the identified parameter was assumed to be a normal distribution with the mean and variance obtained from identification results. The value of parameters on the outside of both 1 % are judged to be under fault condition, and it is found that these extreme values are attributed to abnormal measured data of temperature in tank and inlet except an unknown day. Fault data of insulation damage was unavailable from measured data, so the fault data of insulation damage were generated as 0.1 °C higher than the measured data. The data of identified parameter were divided into two groups before and after a specific boundary day, and whether the two groups have the same mean value or not was judged using t test. The significance level of t test has the peak value in the day when
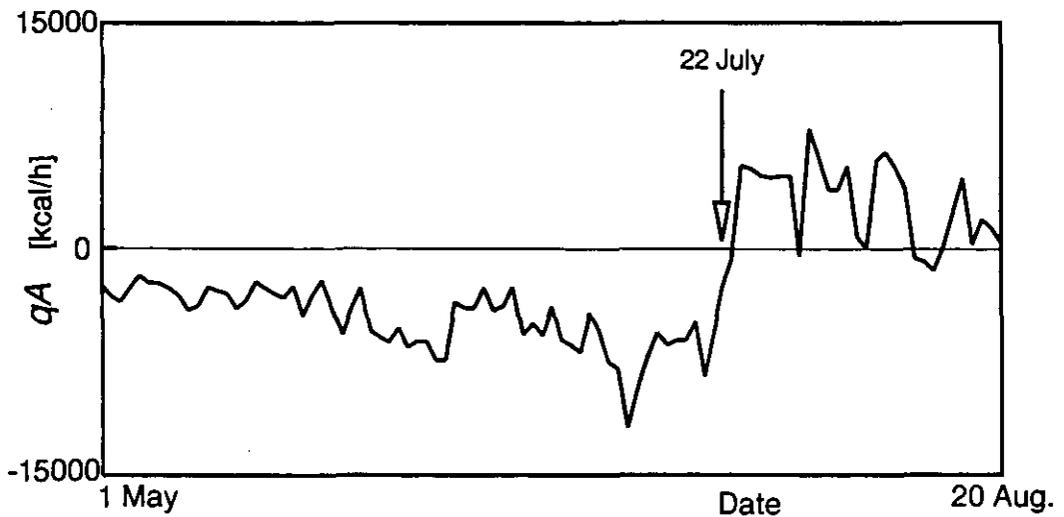
Fig. 16 Change of identified qA of the hospital building

insulation damage happened.

5. Analysis results of identified parameters are presented for measured data in a real system of hospital building. The fault condition such as abnormal water level or insulation damage of tank wall did not happen from May to August 1995. But, the data transmission fault was detected. This is the result that the simulation for parameter identification could not be conducted due to abnormal data of temperature and flow rate. Some false alarms were observed, but the causes are unknown. The shift of identified qA related to constant heat gain from tank wall was observed, but any insulation damage of tank wall did not happen. This is because the identified value of qA was affected by the shift of measured input temperature into storage tank due to sensor fault.

REFERENCES

1. N. Nakahara and K. Sagara:Water Thermal Storage Tank Part 1 Basic Design Concept and Storage Estimation for Multi-connected Complete Mixing Tanks, *ASHRAE Transactions*, Vol.94, Part 2, 1988

2. N. Nakahara, K. Sagara and M. Tsujimoto:Water Thermal Storage Tank Part 2 Mixing Model and Storage Estimation for Temperature-stratified Tanks, *ASHRAE Transactions*, Vol.94, Part 2, 1988

3. N. Nakahara and M. Okumiya:Water Thermal Storage Tank Part 3 Storage Estimation, Applications of Self-balanced Stratified Tanks, *ASHRAE Transactions*, Vol.94, Part 2, 1988

4. M. J. D. Powell : An Effective Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives, *Computer J.*, Vol. 7, 1964

5. W. I. Zangwill : Minimizing a Function without Calculating Derivatives, *Computer J.*, Vol. 10, 1967

A Fault Detection and Diagnosis Method Based on First Principles Models and Expert Rules

T.I. Salsbury, P. Haves and J.A. Wright

Department of Civil & Building Engineering,
Loughborough University of Technology
Loughborough, Leics. LE11 3TU, UK.

Abstract

A fault detection and diagnosis method based on the use of first principles models is described. A single reference model of correct operation is used to generate innovations and thereby detect faults. Diagnosis is performed by analysing the variation in the innovations over the operating range using a rule-based classifier. The method has been tested on the cooling coil subsystem in a simulated air-handling unit. Results of these tests are presented to illustrate the operation of the method.

1: Introduction

Fault *detection* involves the determination that the observed behaviour of the target system is unacceptably different from the expected behaviour. The unacceptable behaviour may occur over the whole operating range or be confined to a limited region. Fault *diagnosis* involves determining which of the possible causes are consistent with the observed behaviour. It may be possible to identify the nature of the fault unambiguously, but often it is only possible to eliminate some of the possible causes. The process of diagnosis requires that the most important possible causes of faulty operation have been identified in advance and that these different causes give rise to behaviours that can be distinguished with the available instrumentation. Unambiguous diagnosis imposes progressively greater demands on both the instrumentation and the sophistication of the diagnosis method as the number and similarity of the faults to be considered increases.

Faults in HVAC systems can be divided into two classes: *binary* or *abrupt* faults, e.g. broken fan belt, and *degradation* faults, e.g. coil fouling. Abrupt faults are easier to detect, since they generally result in a sudden failure of some part of the plant, although they are not necessarily easier to diagnose. In the case of degradation faults, it is necessary to define a threshold, below which the fault is considered insignificant and above which it is considered desirable to detect the fault. The difficulty of detecting and diagnosing degradation faults depends on the threshold adopted. In principle, this threshold must be determined by some kind of cost-benefit analysis. The possible benefits of detecting a particular fault include energy savings, avoidance of occupant discomfort or illness, and avoidance of damage to the building fabric and contents or other components of the HVAC system. The costs of detecting a particular fault include the cost of any additional instrumentation, computer hardware and software, and any human intervention. Both the costs and the benefits will depend on the particular building and application and must be determined on a case-by-case basis.

2: Description of Method

The fault detection and diagnosis (FDD) method presented here is illustrated in Figure 1. It consists of two basic components: a *preprocessor* (for fault detection) and a classifier (for fault diagnosis). The preprocessor consists of a two static models: a first principles model and a radial basis function (RBF) model. The parameters of the first principles model are estimated from test data obtained from the system being monitored. The radial basis function network is used to compensate for any structural inadequacies in the first principles model. Together, the two models represent the correctly operating system and they are used to predict the steady-state output of the system from the measured inputs. The prediction is passed through a first order

low-pass filter, which is used to model the dynamics of the system, and the predicted dynamic output is compared with the measured output. Any differences between the predicted and measured outputs (*innovations*) are indicative of changes in the system. However, innovations may also be due to modelling errors and noise in the measurements. To allow for these errors, the RBF model is also used to predict a confidence interval for the output predictions. The confidence limits act as a detection threshold. If the innovations exceed the threshold for a sustained period then a fault is deemed to have occurred.

The *classifier* consists of three 'bins' which are used to store a moving average of the innovations generated by the preprocessor in different regions of the operating space. In the example presented here, there are three different operating regions, corresponding to low, medium and high coil duty. A rule-based classifier is used to analyse the contents of these bins and produce a diagnosis. The number of bins, the operating regions that they describe and the rules have been specifically designed to distinguish between three common degradation faults in cooling coil subsystems: valve leakage, coil fouling and supply air temperature sensor drift.
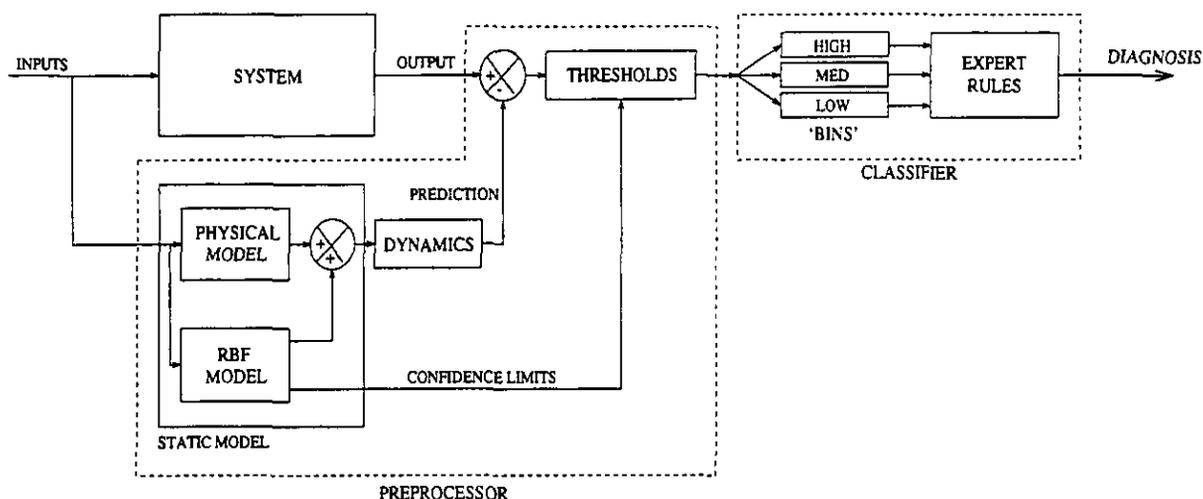
Figure 1: The fault detection and diagnosis scheme

2.1: First Principles Model

The first principles model of the cooling coil subsystem has two parts: a valve model and a coil model. The valve model is necessary in order to predict the flow rate of the water entering the coil from the control signal to the valve actuator since this is not usually measured. The model represents the installed characteristic of an equal-percentage three-port valve by means of a modified exponential function whose curvature is determined by a single parameter. The model of the coil is based on the effectiveness-NTU relationship for a counter-flow heat exchanger. A single parameter, UA, represents the overall conductance of the coil. The dynamic response of the system is approximated by a first-order low-pass filter. The model does not treat the non-linear dynamics associated with rate limited actuators, the variation in coil time constant with operating point or the dynamic response of the supply air temperature sensor. Details of the valve and coil models can be found in [1].

The model of the valve and coil is used to predict the off-coil air temperature from measurements of the on-coil air temperature, the supply air mass flow rate (calculated from the velocity measurements) and the control signal to the valve. Other quantities such as the chilled water supply temperature and mass flow rate of water in the primary circuit are not generally mea-

sured by the BEMS, and are assumed to remain constant at their design values. Any changes to the values of these quantities therefore constitute unmeasured disturbances, and significant enough changes will generate false alarms. These inputs to the model that are assumed to remain constant can be treated as parameters. Table 1 shows all the parameters required by the valve and coil model, together with their method of identification.

Table 1: The parameters of the first principles model

parameter	identification method
valve curvature coefficient	training data
overall conductance of coil	training data
water mass flow rate into valve	design data
chilled water supply temperature	design data
maximum supply air flow rate	design data
maximum temperature rise across fan	commissioning data
time constant	expert knowledge

The last two parameters in Table 1 are used to enable the temperature rise across the supply fan to be estimated in the case where the fan is located between the cooling coil and the supply air temperature sensor.

Design and manufacturers' data are used to provide initial estimates of the valve curvature coefficient and the overall conductance of the coil. Training data obtained from tests on the system are then used to refine these parameter estimates by minimising the mean square difference of the model predictions and the measured system outputs. Since the model is not linear in these two parameters, their optimum values cannot evaluated analytically for the training data. A non-linear direct search method is therefore used [2].

2.2: Radial Basis Function Model

The first principles model necessarily includes a number of simplifications and approximations that limit the accuracy of the model in some regions of the operating space. In regions where there is training data, the errors in the predictions of the first principles model can be determined from a direct comparison of the predictions and the training data. These errors are then compensated for using an RBF network [3], which uses a set of weighted Gaussian distributions to model the errors. The centres of the basis functions are positioned in the input space of the models using the 'k-means clustering' algorithm [4]. Further weights are calculated to allow the RBF to model the density of training data in the input space and the confidence limits for the combined model.

An important property of RBFs is that the resulting model is *local*; its contribution to the combined model tends to zero outside the region covered by the training data. Another important property is that the model is linear-in-the-parameters, allowing a confidence interval for the predictions of the model to be estimated as part of the least squares parameter estimation procedure. In the regions where there is a high density of training data, this confidence interval is used as the confidence interval for the whole model. In regions where the density of training data is low, the confidence interval for the combined model is estimated from the differences between the predictions of the first principles model and the training data. This assumes that the distribution of the training data within the operating space is sufficient to ensure that the errors in the predictions of the first principles model are not significantly greater outside of the training data regions.

2.3: Classifier

Three 'bins' are employed to store a moving average of the amount by which the innovations exceed the 99% confidence interval. The bins correspond to low, medium and high coil duty,

as determined by the control signal to the coil valve actuator, u_c. The low bin is updated if $0 \leq u_c < 0.2$, the medium bin if $0.2 \leq u_c < 0.6$ and the high bin if $0.6 \leq u_c \leq 1.0$. A measure of the age of the data in each bin is also maintained. If this value exceeds a threshold, the innovation data in the corresponding bin is considered too old and the bin is emptied and its value reset.

Boolean rules derived from expert knowledge are used to analyse the content of the bins. A bin can have one of four states:

State	Condition
positive	$b > +\Delta$ AND $a < T$
negative	$b < -\Delta$ AND $a < T$
small	$|b| < |\Delta|$ AND $a < T$
invalid	$a \geq T$

where b is the moving average value of the innovation, Δ is the threshold for b, a is the effective age of the bin value and T is the threshold for a. An example of a rule for the valve leakage fault is:

IF *low bin* IS *positive*
AND *mid bin* IS *small*
AND *high bin* IS *small*
THEN *fault* IS *leakage*

An evidence value is generated with each diagnosis that is proportional to the number of valid bins. If the states of the bins match the rules for more than one fault type then the evidence is shared equally between them.

3: Results

The FDD scheme was tested on a simulated air-conditioning system to assess its ability to detect fouling and leakage faults in the cooling coil of an air-handling unit. The test data are generated from an HVACSIM+ simulation of a three zone VAV air-conditioning system. The main component models used in the simulation are described in [5]. The structure of the plant and the control scheme is similar to that described in [6] but the sizing of the equipment is taken from the detailed design of a recently completed office building in London. The simulated building has three zones, each having time-varying occupancy, equipment and lighting loads. In addition, each zone is subject to significant, highly variable, solar gains. Test data were generated by the simulation using weather data measured on two days: a March day, during which the cooling demand remains relatively low for most of the occupancy period, and a June day in which the system is subjected to a wide range of operating conditions. Figure 2 shows the behaviour of the correctly operating cooling coil on the first and second days.

Training data were generated by performing open and closed loop tests on the simulated plant, as shown in Figure 3. Tests of this sort could be carried out as part of the commissioning process in a real building [7]. Constant loads and inlet conditions were maintained during the tests. In the first phase, all closed loop control was disabled and the control signal to the cooling coil valve actuator was varied in a series of steps at four different air flow rates. In the second phase, local loop control was restored, and the set-point for the supply air temperature varied in a series of steps. Three series of steps were performed, the first with the VAV boxes under closed loop control, the second with maximum air flow rate through the VAV boxes and the third with minimum air flow rate through the VAV boxes. Due to the inaccuracies associated with the measurement of mixed air temperature, the training data are restricted to the part of the operating range of the plant in which the mixing box dampers are set to provide full outside air. The supply temperature exceeds the outside air temperature when the valve is closed due to the heat produced by the fan.
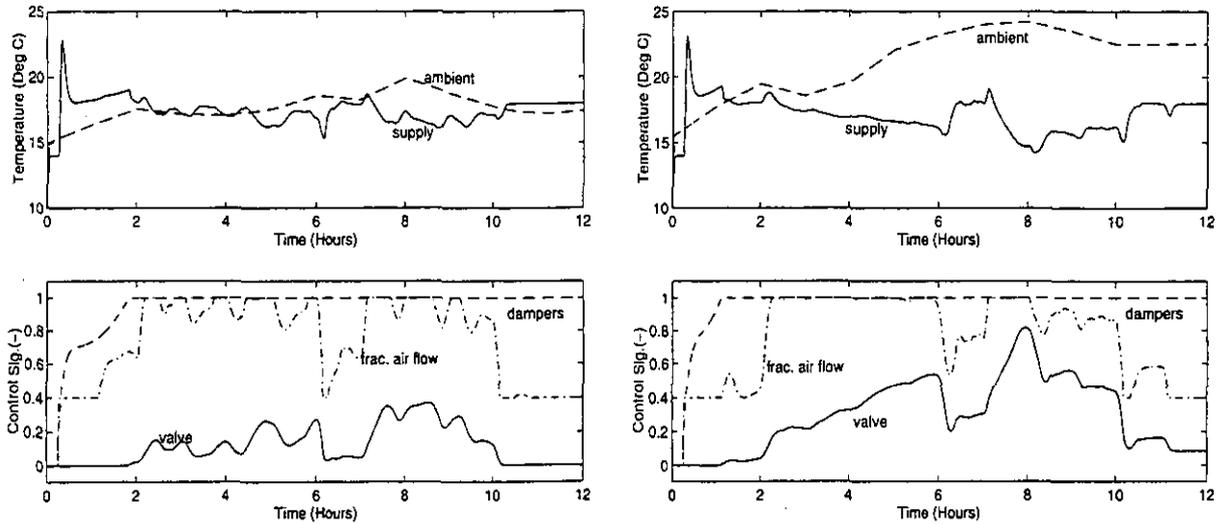
Figure 2: Test data on the March day (left) and June day (right) with no faults. The upper graph shows the supply air temperature and the outside air temperature and the lower graph shows the fractional air flow rate and the mixing box and cooling coil control signals.

The FDD scheme was tested using the correct operation case and two different fault cases: 3 % flow rate leakage through the control valve and 1 mm of calcium carbonate fouling on the tubes. The results of applying the scheme to these three cases on each of the test days are presented in Figures 4 and 5. Figure 4 shows the variables associated with the preprocessor for each of the test days. Two graphs are presented for each day: the upper graph shows the innovation as a solid line and the confidence interval as a symmetrical band centred about zero on the innovation axis and the lower graph shows the control signal to the valve, which is used to determine the appropriate bin for any innovations. Figure 5 shows the variables associated with the classifier. Four graphs are presented for each day: the three upper graphs show the bin values for the different regions of the operating range and the lower graph shows the final output which is the evidence for a particular fault. The thresholds that are shown as a band centred about zero were determined from the training data.

The scheme did not generate any false alarms on either of the test days despite the existence of modelling errors. The modelling errors can be observed as non-zero innovations for the fault free test cases. The confidence limits prove to be a reliable measure of model accuracy and the majority of innovations are contained within the confidence interval. However, the confidence limits were exceeded on the March day when the coil was operating at very low duty but the low bin value did not exceed the maximum value experienced for the training data and a false alarm was not generated. The modelling errors that are apparent may arise in either the static or the dynamic submodels. The confidence limits that are estimated are a measure of the inaccuracy of the static submodel. Since the dynamic submodel is rather crude, sustained transient activity can be expected to lead to false alarms. However, it is assumed that, in practice, most systems do not undergo sustained periods of high frequency transient activity and that they spend most of their time at, or near, steady-state.

The scheme was unsuccessful at detecting fouling on the March day, due to the plant not entering the high duty regions of the operating range where the fault is apparent. Also, the leakage fault was not diagnosed on this day despite the coil spending most of its time at low duty. Reference to Figure 4 shows that the control signal was rarely very close to zero. For most of the time, the control signal fluctuated around a value of 0.2, which is too high for the leakage fault to have any effect on performance. There are two periods during this test where the valve was nearly closed (at 1 hour and 4 hours). The innovations do exceed the confidence limits during these periods, but this is not sustained for long enough to cause the bin values to exceed the

373
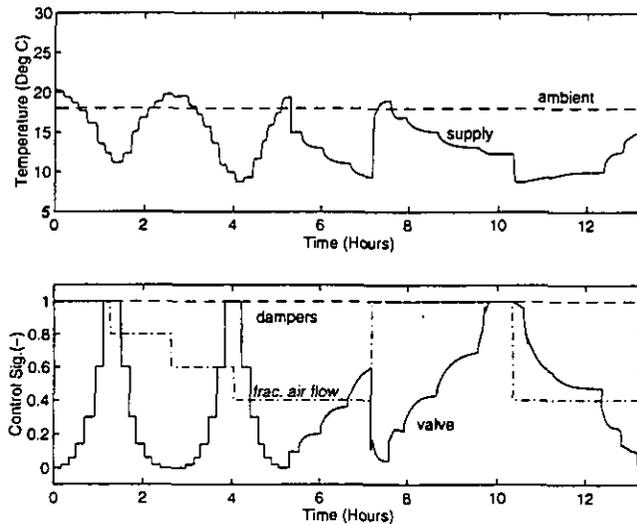
Figure 3: The training data. The upper graph shows the supply air temperature and the outside air temperature and the lower graph shows the fractional air flow rate and the mixing box and cooling coil control signals.

thresholds. The main reason for the lack of data is that the model is not used when the dampers are modulating due to the difficulty associated with obtaining a reliable measurement of on-coil air temperature in real systems. Since the coil usually has a high gain where the valve is very nearly closed, the closed loop control will result in rapid movement through the region where leakage is apparent, resulting in only small amounts of useful data.

On the June Day the scheme managed to detect both of the faults. The fouling fault was detected due to the coil being exercised at high duty where the effects are most apparent. The leakage fault was detected at the end of the test when the valve was almost fully closed.

4: Conclusions

The work has demonstrated that an FDD scheme based on the use of a reference model of correct operation coupled with expert rules is capable of identifying faults in HVAC plants. The combination of a first principles model and a local black box model has been proposed as a means of maximising the use of both a priori information and the information contained within training data. The two-tier modelling approach that has been adopted also allows statistical techniques for linear systems to be employed to calculate confidence limits.

Simplifications in both the static first principles submodel and the dynamic submodel lead to wider estimates of the confidence intervals and hence a less sensitive scheme. Although the RBF model can compensate for the structural inadequacies of the first principles model, this will only occur within the regions where training data were obtained. As the model is used outside these regions, the confidence limits will widen to reflect the underlying inaccuracy of the first principles model. Dynamic inaccuracies are not catered for by the estimated confidence interval and the mismatch between the approximated and true dynamic behaviour will lead to false alarms if there are sustained periods of transient activity. Future work will involve assessing whether enhancements to the first principles model can improve the effectiveness of the scheme (e.g. by including the effects of the fluid capacity rates on the overall conductance of the coil). The problem of dynamic inaccuracies will be addressed by analysing whether first principles relationships can be used to extend the dynamic model.

Acknowledgements

References

[1] M. Benouarets, A. L. Dexter, R. S. Fargus, P. Haves, T. I. Salsbury, and J. A. Wright. Model-based approaches to fault detection and diagnosis in air conditioning systems. In *Proceedings of system simulation in building 94*, 1994.

[2] M. J. Box. A new method of constrained optimisation and a comparison with other methods. *The Computer Journal*, 1965.

[3] S. A. Billings S. Chen. Neural networks for non-linear dynamic system modelling and identification. *International Journal of Control*, 56:319–346, 1992.

[4] C. J. Darken J. Moody. Fast learning in networks of locally tuned processing units. *Neural Computing*, 1:281–294, 1989.

[5] P. Haves. Component-based modelling of vav systems. In *Proceedings of system simulation in building 94*, 1994.

[6] A. L. Dexter and P. Haves. The influence of tuning on the performance of building control systems. In *Proceedings of system simulation in building 90*, 1994.

[7] P Haves, D R Jørgensen, T I Salsbury, and A L Dexter. Development and testing of a prototype tool for HVAC control system commissioning. *submitted for publication in ASHRAE Transactions*.
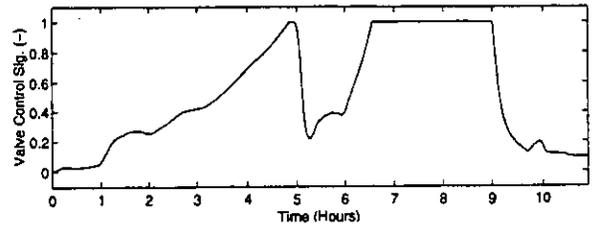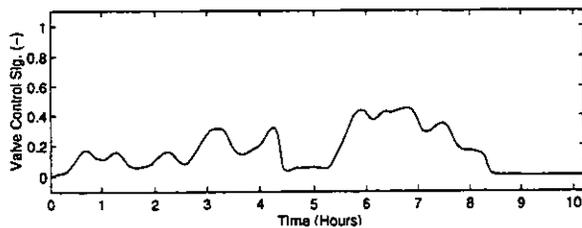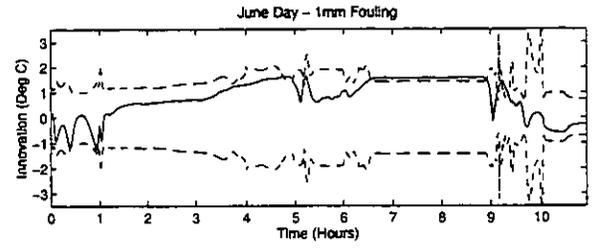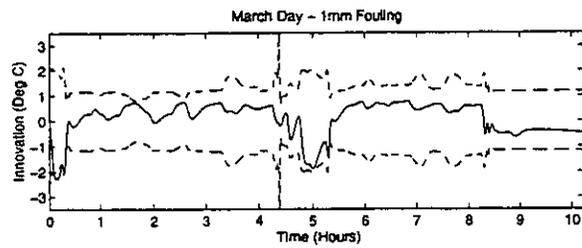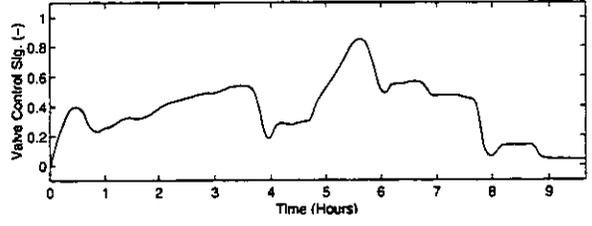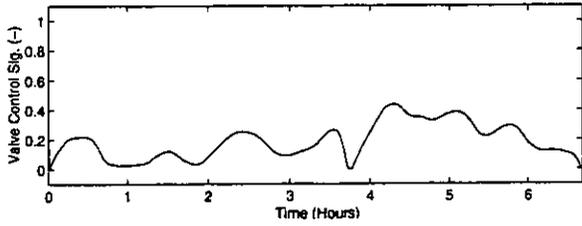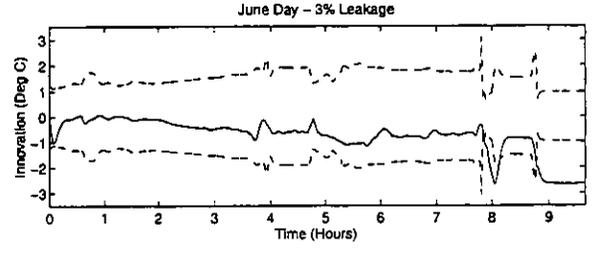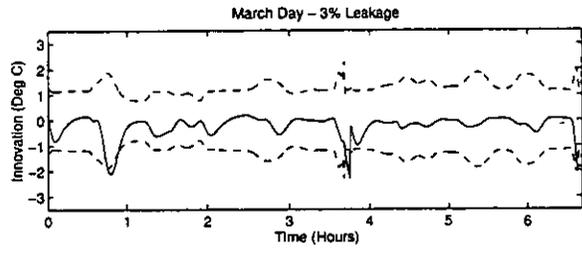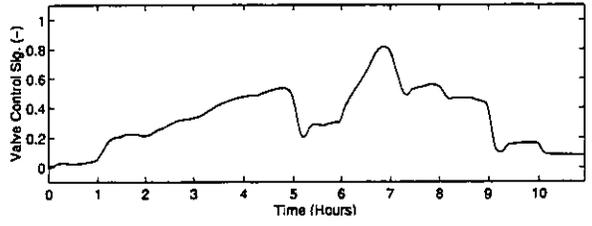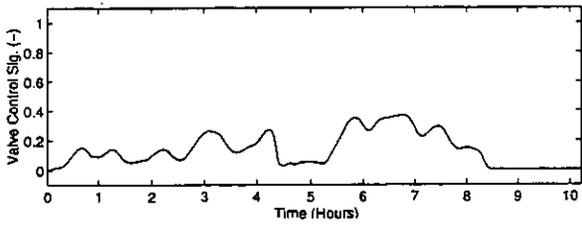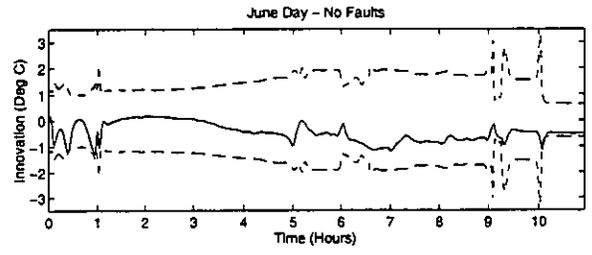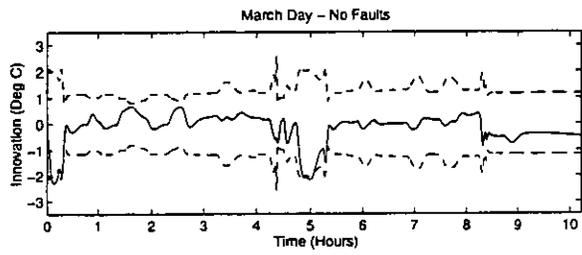
Figure 4: Test Results : Preprocessor Data

376
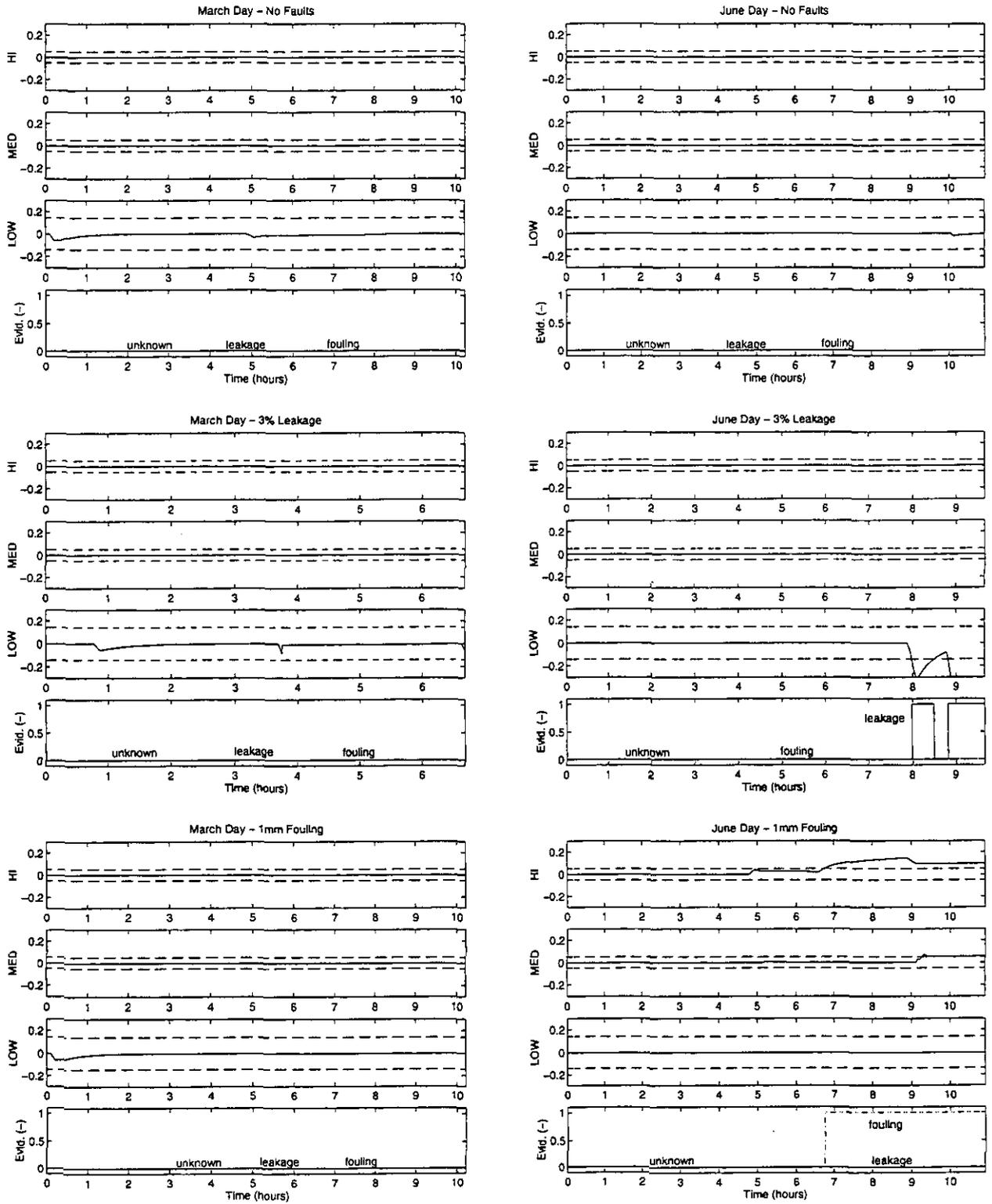
Figure 5: Test Results : Classifier Data

CONDITION MONITORING IN HVAC SUBSYSTEMS USING FIRST PRINCIPLES MODELS

P. Haves   T.I. Salsbury   J.A. Wright

Member *ASHRAE*

ABSTRACT

The paper describes a condition monitoring scheme based on first principles models. The scheme involves estimating the values of model parameters that are expected to change in the event of a fault. The first principles models are, in general, not linear in the parameters, and recursive estimation of the parameters of these models is avoided by estimating the parameters of an intermediate model that is, linear in the parameters. This intermediate model, which takes the form of a radial basis function network, is used periodically to generate data covering the complete operating range of the system. These data are then used in the estimation of the parameters of the first principles model. The paper describes the techniques used and presents results from applying the method to the task of detecting two faults in the cooling coil subsystem of an air-handling unit.

1 Introduction

The sensor and control signals in an HVAC system contain potentially valuable information about the state of the system, and energy management and control systems (EMCS) have the ability to monitor and store these signals. In practice, the only checks that are carried out are to verify that set points are being maintained and that certain critical variables remain with predetermined limits. This may allow the detection of certain abrupt or catastrophic faults, but the task of identifying the underlying cause of the failure often requires a detailed manual

---

[0]P. Haves is a Senior Lecturer, T.I. Salsbury is a Research Assistant and J.A. Wright is a Lecturer in the Department of Civil and Building Engineering, Loughborough University of Technology, UK.

analysis of trend data up to the point of failure or a test of the system so that a diagnosis may be made. The slow changes in performance caused by degradation faults may be even more difficult to observe manually, and these faults may remain undetected until their effects are quite severe. One reason for this is that feedback control tends to reduce the effect of such faults on the ability of the system to maintain comfort, even though other aspects of the performance of the system, such as energy consumption, may be significantly impaired.

An alternative approach is to use the sensor and control signal data collected by the EMCS to monitor the state of the system and infer the nature and extent of any faults using an automatic fault detection and diagnosis (FDD) system. Fault *detection* involves the determination that the observed behavior of the target system is unacceptably different from the expected behavior. The unacceptable behavior may occur over the whole operating range or be confined to a limited region. Fault *diagnosis* involves determining which of the possible causes are consistent with the observed behavior. In some cases, it may be possible to identify the nature of the fault unambiguously, but often it is only possible to eliminate some of the possible causes. The process of diagnosis requires that the most important possible causes of faulty operation have been identified in advance and that these different causes give rise to behaviors that can be distinguished with the available instrumentation.

Fault detection and diagnosis in HVAC systems has been explored as part of an International Energy Agency collaborative research project (Annex 25) and by a number of other researchers (e.g. Usoro et al. 1985, Pape et al. 1991). Significant research has been carried out by other industries, such as chemical processing. All of the approaches require the use of models of some sort. These models may be qualitative (e.g. Kaler 1990, Dexter and Benouarets 1995a) or quantitative (e.g. Iserman 1984) or a combination of the two (e.g. Yu and Lee 1991). For fault detection only a single model of correct operation is required; the system is deemed to be faulty if its behavior does not match that of the model of correct operation. Diagnosis of different faults requires either an analysis of how the difference between the observed and predicted behaviors varies with operating point or the use of models of the different types of faulty operation (Benouarets et al. 1994).

The approach presented here involves identifying a model of the system on-line and analyzing its parameter values to ascertain whether a fault has occurred. The paper describes an FDD method designed to diagnose multiple faults and presents results of tests performed on a simulated cooling coil subsystem with fouled coil and a leaking valve.

2 Overview of the FDD Method

Figure 1 is a schematic of the condition monitoring scheme which comprises a radial basis function (RBF) model and a first principles ('physical') model. The RBF models the local behavior of the HVAC system and is updated using a recursive gradient-based estimator when the system is (approximately) in steady state. In order to avoid estimator wind-up (or *over-training*), the RBF is only updated when the difference between the predicted and measured output exceeds a certain threshold, indicating that the performance of the system has changed. The RBF is then exercised over the operating range of the system and the data generated used in the estimation of the parameters of the physical model using a direct search technique (Box 1965). The parameters that are estimated for the physical model are physically meaningful and represent a tangible measure of the system performance. Determination of detection thresholds is therefore greatly simplified and they can be set to suit each particular system and its performance criteria.

This indirect method of estimating the parameters of the physical model has been adopted because physical models are not, in general, linear in the parameters and are, therefore, unsuited to recursive parameter estimation. Because the RBF is a local model, it provides an estimate of the most recently observed behavior of the system in different parts of the operating space, responding relatively quickly to changes in the behavior of the system.

The rest of the paper describes the application of the method to the detection and diagnosis of coil fouling and valve leakage in a cooling coil.

3 First Principles Model

The model represents the principle static characteristics of cooling coil subsystems of the type found in air-handling units. The air-side approach of a coil is defined by:

$$\alpha = \frac{T_{ai} - T_{ao}}{T_{ai} - T_{wi}} \tag{1}$$

The NTU-effectiveness relationship for counter-flow operation of a dry coil is used to estimate the approach from the overall conductance, UA, and the air- and water-side capacity rates C_a and C_w:

$$\alpha = \varepsilon \frac{min(C_a, C_w)}{C_a} \tag{2}$$

where:

$$\varepsilon = \frac{1 - exp(-NTU(1 - \omega))}{1 - \omega exp(-NTU(1 - \omega))} \tag{3}$$

and $NTU = \frac{UA}{min(C_a, C_w)}$ and $\omega = \frac{min(C_a, C_w)}{max(C_a, C_w)}$.

The capacity rate of the air can be calculated directly since the air mass flow rate is usually available (or can be calculated from the velocity), but the water mass flow rate through the coil is not usually measured and needs to be inferred from the control signal to the actuator of the control valve. To achieve this, a model has been developed to approximate the behavior of a typical equal percentage three port valve. The model consists of a modified exponential function, characterized by the parameter c. The relationship between the fractional water mass flow rate into the coil, $\frac{\dot{m}_w}{\dot{m}_{w,max}}$, and the valve actuator control signal, u_c is given by:

$$\frac{\dot{m}_w}{\dot{m}_{w,max}} = \frac{exp(cu_c) - 1}{exp(c) - 1} \tag{4}$$

The modification avoids the need to switch to a different function in the lower part of the range in order to avoid the significant opening at zero stem position that would occur with a pure exponential. The model also treats leakage resulting from restriction of the travel of the stem due to foreign matter on or near the valve seat. The leakage parameter, l, specifies the fractional flow when the valve is nominally closed. The fractional water flow into the coil is then:

$$\frac{\dot{m}_w}{\dot{m}_{w,max}} = max(\frac{exp(cu_c) - 1}{exp(c) - 1}, l) \tag{5}$$

Four parameters ($\dot{m}_{w,max}$, c, l, UA) are required for the combined valve and coil model. For simplicity, the coil model uses a constant value of UA, and therefore does not treat the effect of varying fluid flow rate on the corresponding heat transfer coefficient. However, it is possible for the variation in the water-side heat transfer coefficient to be partly compensated for by a change in the value of the valve characteristic, c. An extension to the model would be needed to treat the corresponding effect of variations in the air flow rate, e.g. in VAV systems.

The approach that is estimated by the model is compared with the approach calculated from the measured temperatures using Equation 1. There are different associated with the measurement of each of these temperatures. The temperature of the air entering the coil, T_{ai}, is equal to the temperature of the air leaving the mixing box. However, this temperature is difficult to measure accurately due to imperfect mixing of the outside and recirculation air streams in many systems. The approach adopted here is to treat as invalid any measurements made when

the mixing box dampers are not in either of their extreme positions. The temperature of the air leaving the coil, T_{ao}, is generally not measured in a draw-through system, so the supply (discharge) air temperature is used, after correcting for the temperature rise across the supply fan, ΔT. This temperature rise may vary with the flow rate through the fan, e.g. in a VAV system. The relationship between the temperature rise and the flow rate depends both on the characteristics of the fan and the way it is controlled. Since the rise is relatively small ($\sim 1\,K$, $2\,°F$), an approximate correction has been assumed:

$$\Delta T = \frac{\dot{m}_a}{\dot{m}_{a,max}} \Delta T_{max} \quad . \tag{6}$$

The maximum temperature rise across the fan, ΔT_{max} can be measured as part of commissioning and the maximum air mass flow rate, $\dot{m}_{a,max}$ can be estimated from the system specifications.

3.1   Initialization of the First Principles Model

The parameters of the physical model are initialized in two stages. Design information and manufacturers' performance data are used to produce initial estimates of the parameter values and to define the regions of feasibility for each parameter. Training data collected when the plant is deemed to be operating correctly (e.g. at commissioning time) are then used to estimate the parameter values corresponding to the fault-free system. Significant differences between the initial values and the values identified from the training data indicate differences in performance between the system as-designed and the system as-built.

Since the physical model is not linear in the parameters, analytical techniques cannot be used to locate the optimum parameters values for a given data set. Problems were experienced with gradient-based methods as a result of discontinuous derivatives in the valve model. Box's complex direct search method has therefore been adopted since it does not require derivative information. The search is also effective in handling parameter bounds. Each of the parameters of the physical model is normalized so that its feasible range maps onto a range of zero to one, in an effort to produce circular objective function surface contours which can help to increase robustness. The objective function used is the mean of the squares of the errors (MSE) of the physical model predictions at each of the training data samples. The search process is terminated when the objective function value fails to improve for a number of iterations.

3.2 Updating the Parameters of the First Principles Model

During operation of the condition monitoring scheme, a subset of the parameters of the physical model (fault parameters) is updated so as to provide the best fit to data generated by the RBF model. The fault parameters that are included in the search are the UA of the coil and the fractional flow leakage l; all other parameters remain fixed at their initial values. The RBF model is used to generate outputs at points distributed uniformly within the input space and these are compared with the outputs of the physical model in order to calculate the MSE. The complex method is then employed to minimize the MSE by searching for the optimum fault parameters.

4 Radial Basis Function Model

Radial basis function models are mathematical tools for approximating multi-dimensional surfaces using local non-linear functions. The most common function used is the Gaussian function. A model is constructed by centering a number of Gaussian functions at selected positions within the input space and selecting widths so that the tails of neighboring functions overlap, see Figure 2 which shows two Gaussian functions in a single input dimension. The output of a particular function (its *activation*) decreases with the Euclidean distance between the input to the model and the center of the function, giving the model its local properties. The output of the network, y_p, at a particular point, x, is the sum of the activations, $\phi_i(x)$, multiplied by the corresponding weights, w_i:

$$y_p = \left[\begin{array}{cccc} \phi_1(x) & \phi_2(x) & \cdots & \phi_c(x) \end{array} \right] \left[\begin{array}{c} w_1 \\ w_2 \\ \vdots \\ w_c \end{array} \right] \tag{7}$$

The weights are estimated initially from the training data and are then updated using approximately steady-state operating data from the system. An important property of RBF networks is that, if the number and shape of the basis functions are fixed prior to training, estimation of the weight vector is a linear optimization task which can be solved analytically using conventional least squares methods.

For this application, the RBF network is used to model the static relationship between the system inputs (the fractional air mass flow rate through the coil, $\frac{\dot{m}_a}{\dot{m}_{a,max}}$ and the control signal to the valve u_c) and the system output (the air-side approach calculated from the temperature

measurements, α). The weight vector is updated using a simple linear gradient descent method. Usually, if an RBF model is trained to approximate a system off-line from a finite data set, then the positioning of the centers and the degree of overlap between the Gaussian functions can be optimized to suit the system characteristic by employing non-linear optimization methods. However, since the system characteristic will change, these attributes cannot be fixed *a priori* and on-line refinement would make the model non-linear-in-the-parameters. A general, fixed, configuration of basis functions is therefore required that will allow the model to approximate arbitrary non-linearities by optimizing the weights alone. To achieve this, it is ensured that the range of each input is normalized to be between 0 and 1, the centers are then positioned on a uniform grid in this normalized input space. The widths are selected to be equal to the distance between adjoining centers thereby providing a parsimonious activation surface. It was determined that the non-linearity of the cooling coil subsystem could be accurately modeled by having eight centers in each dimension (giving a total of sixty-four centers for the two dimensional input space).

4.1 Initialization of the RBF

The RBF weights are initialized using a conventional least-squares estimation procedure. This procedure is sensitive to outlier points and any 'holes' in the training data may lead to an ill-conditioned solution. To avoid these problems and to produce consistency between the two models at the outset, the RBF model is initialized using training data generated from the calibrated physical model. The data are generated on a fine, regular grid of points that span the operating space.

If a series of n data points is generated by the physical model, the multiple input, single output, RBF model can be written as:

$$y_p = \Phi \mathbf{w} + \varepsilon \tag{8}$$

where

$$y_p = \begin{bmatrix} y_p(1) \\ y_p(2) \\ \vdots \\ y_p(n) \end{bmatrix}, \Phi = \begin{bmatrix} \phi_1(x(1)) & \phi_2(x(1)) & \cdots & \phi_c(x(1)) \\ \phi_1(x(2)) & \phi_2(x(2)) & \cdots & \phi_c(x(2)) \\ \vdots & \vdots & & \vdots \\ \phi_1(x(n)) & \phi_2(x(n)) & \cdots & \phi_c(x(n)) \end{bmatrix}$$

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_c \end{bmatrix}, \text{ and } \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$y_p(i)$ is the ith prediction of the physical model, $\mathbf{w}$ is the vector of RBF weights, Φ is the matrix of basis functions at each input point, x, and ε is the vector of prediction errors between the RBF and the physical model. The least squares criterion can be applied to calculate the unbiased estimate of $\mathbf{w}$ that has minimum variance, which is:

$$\hat{\mathbf{w}} = (\Phi^T \Phi)^{-1} \Phi^T y_p \tag{9}$$

4.2 Updating the RBF

Once the RBF has been initialized by training it to represent the correct operation state of the system it is updated using normal operating data from the system. A gradient-based recursive parameter estimation method known as normalized least mean squares (Åström and Wittenmark 1989) was chosen for its modest data storage and processing requirements. For the scalar output RBF described in the previous section, the prediction at sample i, $\hat{y}(i)$ in response to an input vector, $x(i)$, is given by:

$$\hat{y}(i) = \phi^T(i)\hat{\mathbf{w}} \tag{10}$$

where $\phi^T(i) = [\phi_1(x(i)) \ldots \phi_c(x(i))]$ is the vector of basis function activations and $\hat{\mathbf{w}}$ is the vector of weights. If the weight vector at sample i-1 is used to generate the model prediction then a reasonable criterion of how well the model performs is:

$$V_i(\mathbf{w}) = \frac{1}{2}\left[y(i) - \phi^T(i)\hat{\mathbf{w}}(i-1)\right]^2 \tag{11}$$

where $y(i)$ is the measured output of the system. Differentiation with respect to $\hat{\mathbf{w}}(i-1)$ shows that the gradient of this criterion is:

$$\begin{aligned} V_i'(\mathbf{w}) &= -\phi\left[y(i) - \phi^T(i)\hat{\mathbf{w}}(i-1)\right] \tag{12} \\ &= -\phi e(i) \tag{13} \end{aligned}$$

where $e(i)$ denotes the prediction error. The projection algorithm involves moving the parameter estimate in the direction of the negative gradient by an amount κ such that:

$$\hat{\mathbf{w}}(i) = \hat{\mathbf{w}}(i-1) + \kappa\phi(i)e(i) \tag{14}$$

If $\hat{\mathbf{w}}(i)$ is assumed to be the correct weight vector at sample i then:

$$y(i) = \phi^T(i)\hat{\mathbf{w}}(i-1) + \kappa\phi^T(i)\phi(i)e(i) \tag{15}$$

Re-arranging for κ:

$$\kappa = \frac{1}{\phi^T(i)\phi(i)} \tag{16}$$

Hence, the updating formula is:

$$\hat{\mathbf{w}}(i) = \hat{\mathbf{w}}(i-1) + \frac{\phi(i)}{\phi^T(i)\phi(i)}e(i) \quad . \tag{17}$$

In this form, the updating formula would converge on the optimum parameter values for a finite set of training data samples if the number of samples were equal to the number of parameters. In practice, where the number of presented samples is greater than the number of parameters a learning rate, λ, is introduced so that:

$$\hat{\mathbf{w}}(i) = \hat{\mathbf{w}}(i-1) + \frac{\lambda\phi(i)}{\gamma + \phi^T(i)\phi(i)}e(i) \tag{18}$$

where γ is a small number that is introduced to protect against division by zero for the case when $\phi(i) = 0$. In situations where the system parameters are time-invariant then the learning rate affects the speed of convergence and the accuracy of the solution (S. Abu el Ata-Doss et al. 1985). For this application, in which a system has time-varying parameters, the learning rate determines the tracking speed of the estimator. .

4.3 Data Preprocessing

Both the RBF and the physical model are static models. The dynamics of the system are ignored due to the difficulty of modeling them accurately using physical equations. A transient detector is used to prevent updating of the RBF when the measured system variables are varying significantly. Firstly, a discrete-time, low-pass, filter is used to reduce the effect of noise. The

activity of each variable, defined as the absolute change from one time step to the next, is then averaged using another discrete-time, low-pass, filter. Finally, the averaged *activity* is normalized and compared with a threshold value. When the activity is below the threshold, the system is deemed to be sufficiently close to steady state to update the RBF. A description of a similar detector is given by Dexter and Benouarets (1995b).

5   Results

The results of tests on a simulated air-conditioning system are presented to demonstrate the ability of the scheme to detect both fouling and leakage faults in the cooling coil of an air-handling unit. The test data are generated from an HVACSIM+ simulation of a three zone VAV air-conditioning system. The main component models used in the simulation are described by Haves (1994). The structure of the plant and the control scheme is similar to that described by Dexter and Haves (1990) but the sizing of the equipment is taken from the detailed design of a recently completed office building in London. The simulated building has three zones, each having time-varying occupancy, equipment and lighting loads. In addition, each zone is subject to significant, highly variable, solar gains. Figure 3 and Figure 4 show the behavior of the correctly operating cooling coil on two different days.

Training data were generated by performing closed loop tests on the simulated plant. Tests of this sort could be carried out as part of the commissioning process in a real building. Constant loads and inlet conditions were maintained during the tests. The supervisory control scheme was disabled and the set-point for the supply air temperature varied in a series of steps. Three tests were performed, the first with the VAV boxes under closed loop control, the second with the VAV boxes demanding maximum air flow and the third with the VAV boxes demanding minimum air flow. The training data are shown in Figure 5. Due to the problems associated with the measurement of mixed air temperature discussed in Section 3, the training data are restricted to the part of the operating range of the plant in which the mixing box dampers are set to provide full outside air.

Table 1 shows the values of the parameters estimated for the physical model by using the complex method to optimize the fit of the model to the training data.

The condition monitoring scheme was configured to estimate the UA of the coil and the leakage through the coil valve. The scheme was then tested using three different fault cases: 3 % flow rate leakage through the control valve, 1 mm of calcium carbonate fouling on the tubes, and both faults present at the same time. The results of applying the scheme to these three cases on each of the test days are presented in Figure 6. There are three graphs for each of the

fault cases. The upper graph shows the difference between the measured approach and the approach predicted by the physical model with the correct operation parameters, which gives an indication of system performance at the current operating point. The middle graph shows the value of the UA parameter estimated from the data generated by the RBF and the lower graph shows the estimated leakage. The breaks in the lines indicate times when data were rejected by the transient detector. The parameter estimates at the end of each run for each of the two days are given in Table 2. The value of the learning rate, λ, was 0.1. The UA value for the fault-free case is slightly different to that given in Table 1 because it was estimated from data generated by the RBF instead of directly from the measured training data.

A comparison of the UA values observed for the different system states on the two days indicates a modest but significant reduction when the coil is fouled. Two factors give rise to a smaller reduction than would be expected. Firstly, the coil operates at only moderate duty, particularly on the first day, and hence its performance is less sensitive to fouling. Secondly, if combinations of valve position and air flow rate (the variables that span the operating space) that occurred in the training data do not occur in the test data, the corresponding basis functions will not be updated. Since all parts of the RBF contribute with equal weight to the re-estimation of the parameters of the physical model, the change in the estimated value of the UA will only be a fraction of the actual change. This problem is being addressed in further work.

A similar, but more subtle, effect is observed for the leakage parameter. The leakage faults are not detected on the first day, in spite of the coil operating at lower duty than on the second day. This is due to very little steady-state data being collected at the operating point where the fault is apparent (i.e. when the valve is fully or nearly closed). There are two reasons for this. Firstly, the model is not used when the dampers are modulating due to the uncertainty involved in estimating the temperature of the air entering the coil from the measurements of the fresh air and return air temperatures. Secondly, the region where the valve is nearly closed has a high gain and the data in this region will usually be changing relatively quickly, and hence will be rejected by the transient detector. On the second day, there happened to be a number of steady-state samples collected at the point where the valve was nearly closed, and the scheme managed to estimate over 1% leakage on both of the days with leakage.

For both of the test days, it can be observed that the graph showing the difference between the prediction of the correct operation physical model and the measured output of the system gives a good indication of the presence of a fault. For the fouling faults, the residual is positive at high coil duty, whereas for the leakage faults, it is negative at low duty. The correlation between the nature of the residual and the type of fault can be exploited for fault diagnosis as described by Benouarets et al. (1994). However, this approach requires the model to be accurate for the complete range of operation, as modeling errors will lead to false alarms. The effect of

modeling error can be observed in the graph showing the fouling fault on the second day. On this occasion, the residual has a negative value when the coil is at low duty. Reference to the raw data revealed that the control signal to the valve at this point is $\sim$10 %; at other times during the test when the control signal had a similar value, but the air mass flow rate was different, there is no corroborating evidence of leakage. This suggests a modeling error within this operating region. Because of the lack of corroborative leakage evidence, the scheme is best able to fit the physical model to the evidence by decreasing the estimated value of UA.

6 Conclusions

A condition monitoring scheme based on physical models has been described and its ability to detect the presence of valve leakage and water-side coil fouling within the cooling coil subsystem of an air-handling unit demonstrated. The UA of the coil and the fractional flow rate of water leaking through the coil were the quantities reported by the scheme. These quantities could be transformed into more tangible measures of performance, such as the reduction in coil capacity and the increase in energy usage, thus simplifying the task of setting detection thresholds. No detection threshold values have been proposed in this paper; ideally their magnitude should be tailored to each particular system based on the level of deterioration in performance that could be tolerated by the building owner and/or occupier.

Estimation of the parameter values of the physical model is a non-linear optimization task and there will always be a danger of converging to a local minimum, regardless of the technique used. If the data used for the optimization covers only a small region of the operating space then the chance of the method failing to converge to the true (global) optimum is increased. Use of the RBF model enables data to be generated across the range of operation and this helps to make the non-linear optimization more robust. However, the method requires the use of two parameter estimation procedures, each having potential inaccuracies. It is difficult to quantify the estimation inaccuracies that are incurred with the present method. The algorithm that is used to estimate the parameters of the RBF does not allow confidence intervals to be calculated directly, but this could be achieved by using a more computationally intensive estimator, such as recursive least squares, that would make the covariance matrix directly available at each time step. However, since the system is assumed to have time-varying parameters then forgetting would need to be employed which would serve to keep the covariance matrix high, leading to over-estimated levels of parameter uncertainty. Inaccuracies also occur when estimating the parameters of the physical model from the RBF. The non-linear form of the physical model prevents analytical evaluation of confidence intervals for each of the parameters but an assessment of the overall fit of the physical model to the RBF can be made from the MSE.

The accuracy of the estimated parameters will also depend on the structural adequacy of the physical model. The model that has been used in the paper is a simplified representation of a cooling coil subsystem that has been developed with the objective of capturing the principle static characteristics of the system while keeping the number of parameters to a minimum. One of the approximations made by the model is that of a fixed UA; in reality the UA will vary as a function of the air and water flow rates. The overall effect of the change in the UA due to variations in the water flow rate can be compensated for to a certain extent by the valve model. The effective value of the valve characteristic parameter, c, is determined by the combined effects of the inherent characteristic of the valve, the valve authority and the variation in UA with water flow rate. The variation in UA with air flow rate is not accounted for within the structure of the physical model. If the UA value were estimated using data from different regions of the operating range at different times, the estimated value would vary due to the variations in the flow rates. However, because the RBF is used to generate data throughout the whole range of operation, the physical model optimization process yields an effective UA which is a good indicator of coil performance.

The method that has been described is best suited to the tracking of degradation faults, where the system continues to be operated throughout its range of operation but with changed characteristics. Failure faults typically result in the system saturating at one operating point, which would make it impossible to build a global model of the new, faulty, characteristic. Failure faults could be detected by observing changes in the weights of the RBF, but diagnosis of failure faults may require the use of test signals to acquire more information about the system.

Acknowledgements

Nomenclature

T_{ai}	Air temperature entering the coil
T_{wi}	Chilled water supply temperature
T_{ao}	Air temperature leaving the coil
$\dot{m}_w$	Mass flow rate of water into coil
$\dot{m}_a$	Mass flow rate of air onto Coil
c_{pw}	Specific heat of water
c_{pa}	Specific heat of air
C_a	Capacity rate of air
C_w	Capacity rate of water
UA	Overall heat transfer conductance
ε	Effectiveness of the coil
NTU	Number of transfer units for the coil
ω	Ratio of fluid capacities
c	Valve model curvature parameter
l	Valve model leakage parameter
ΔT	Temperature rise across fan
u_c	Control signal to cooling coil valve actuator
h_w	Water-side convective heat transfer coefficient
h_a	Air-side convective heat transfer coefficient
$\dot{Q}$	Heat transfer rate of coil
$\dot{Q}_{max}$	Maximum possible heat transfer rate of coil
α	Actual air-side temperature approach
$\hat{\alpha}$	Predicted air-side temperature approach
ΔT_{max}	Maximum temperature rise across fan
y_p	Physical model prediction of approach
w	RBF weight
$\phi(.)$	RBF activation
x	RBF input vector
λ	RBF learning rate
γ	A small number

REFERENCES

Abu el Ata-Doss, S., J.L. Estival, J. Richalet. 1985. "Dynamics of Identification for the Normalized Least Mean Square Algorithm." Proceedings of the 7th IFAC/IFORS Symposium on

Identification and System Parameter Estimation. Swansea, Wales, UK. p. 1237.

Åström, K.J., B. Wittenmark. 1989. "Adaptive Control." Addison-Wesley.

Benouarets, M., A. L. Dexter, R. S. Fargus, P. Haves, T. I. Salsbury and J. A. Wright. 1994. "Model-based Approaches to Fault Detection and Diagnosis in Air-conditioning Systems." Proceedings of the 4th International Conference on System Simulation in Buildings. Liege, Belgium. p. 529.

Box, M.J. 1965. "A New Method of Constrained Optimization and a Comparison with Other Methods." The Computer Journal. Vol 8, p. 42.

Dexter, A.L., M. Benouarets. 1995. "Generic Modelling of HVAC plant for Fault Diagnosis." Proceedings of the IBPSA conference, Madison, Wisconsin, USA.

Dexter, A.L. and M. Benouarets. 1995. "Transient Detection." IEA Annex 25, Building Optimization and Fault Diagnosis Source Book. VTT, Espoo, Finland. p. 259.

Dexter, A.L. and P. Haves. 1990. "The Influence of Tuning on the Performance of Building Control Systems." Proceedings of the 3rd International Conference on System Simulation in Buildings. Liege, Belgium.

Haves, P. 1994. "Component-Based Modelling of VAV Systems." .Proceedings of the 4th International Conference on System Simulation in Buildings. Liege, Belgium. p. 31.

Iserman, R. 1984. "Process Fault Detection based on Modeling and Estimation Method - A Survey." Automatica, Vol. 20, p. 387.

Kaler, G.M. 1990. "Embedded Expert System Development for Monitoring Packaged HVAC Equipment." ASHRAE Transactions, Vol. 96, Pt. 2, p. 733.

Pape, F.L.F., J.W. Mitchell and W.A. Beckman. 1991. "Optimal Control and Fault Detection in Heating, Ventilating, and Air-conditioning Systems." ASHRAE Transactions, Vol. 97, Pt. 1, p. 729.

Usoro, P.B., I.C. Schick, and S. Negahdaripour. 1985. "An Innovation-based Methodology for HVAC System Fault Detection." ASME Transactions, Vol. 107, p. 284.

Yu, C.C., C. Lee. 1991. "Fault Diagnosis Based on Qualitative/Quantitative Process Knowledge." AIChE Journal, Vol. 4, p. 617.

Table 1: The values of the parameters estimated from the training data

Parameter	Value
UA (kW.K^{-1}, kBTU.hr^{-1}.°F^{-1})	5.19 (9.84)
$\dot{m}_{w,max}$ (kg.s^{-1}, lb.s^{-1})	6.20 (13.67)
c	2.61
l (%)	0.03

Table 2: The values of the parameters estimated from the up-dated RBF

	Day 1			Day 2		
System state	UA		l	UA		l
	kW.K^{-1}	(kBTU.hr^{-1}.°F^{-1})	%	kW.K^{-1}	(kBTU.hr^{-1}.°F^{-1})	%
fault-free	5.19	(9.84)	0.03	5.19	(9.84)	0.03
3% leakage	5.13	(9.73)	0.11	5.19	(9.84)	1.09
1 mm fouling	5.04	(9.56)	0.03	4.90	(9.31)	0.27
3% leakage + 1 mm fouling	4.93	(9.35)	0.18	4.76	(9.03)	1.37
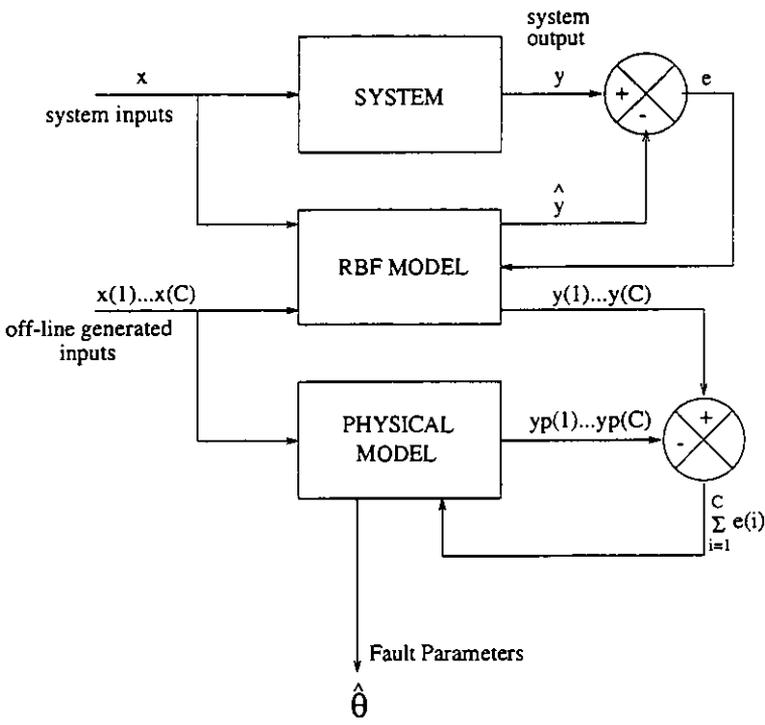
Figure 1: The condition monitoring scheme
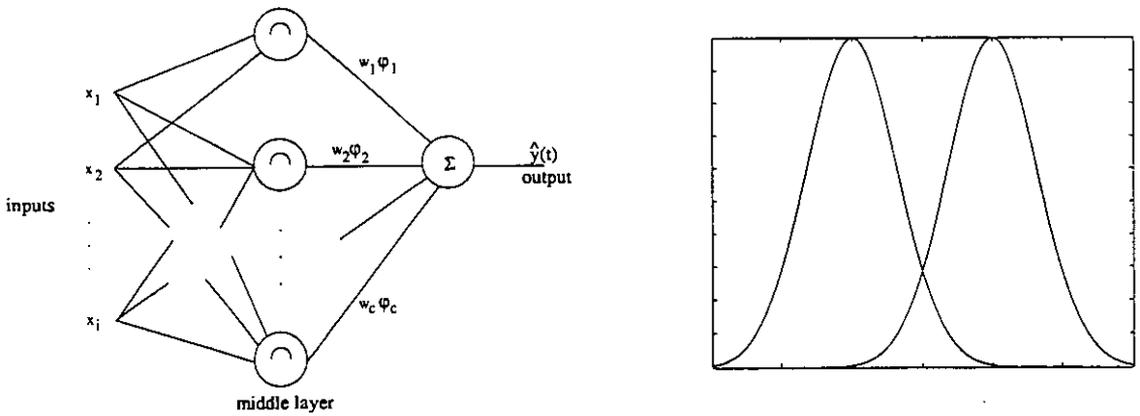
Figure 2: An RBF network and Gausssian functions
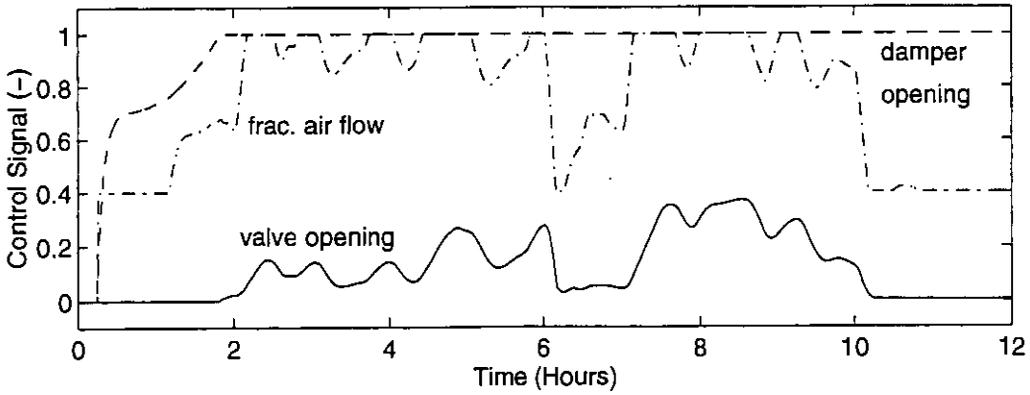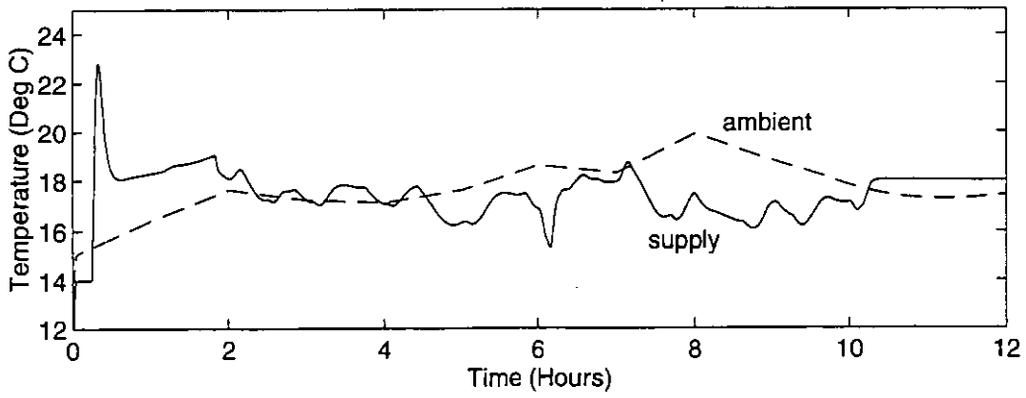
Figure 3: Test data on first day with no faults
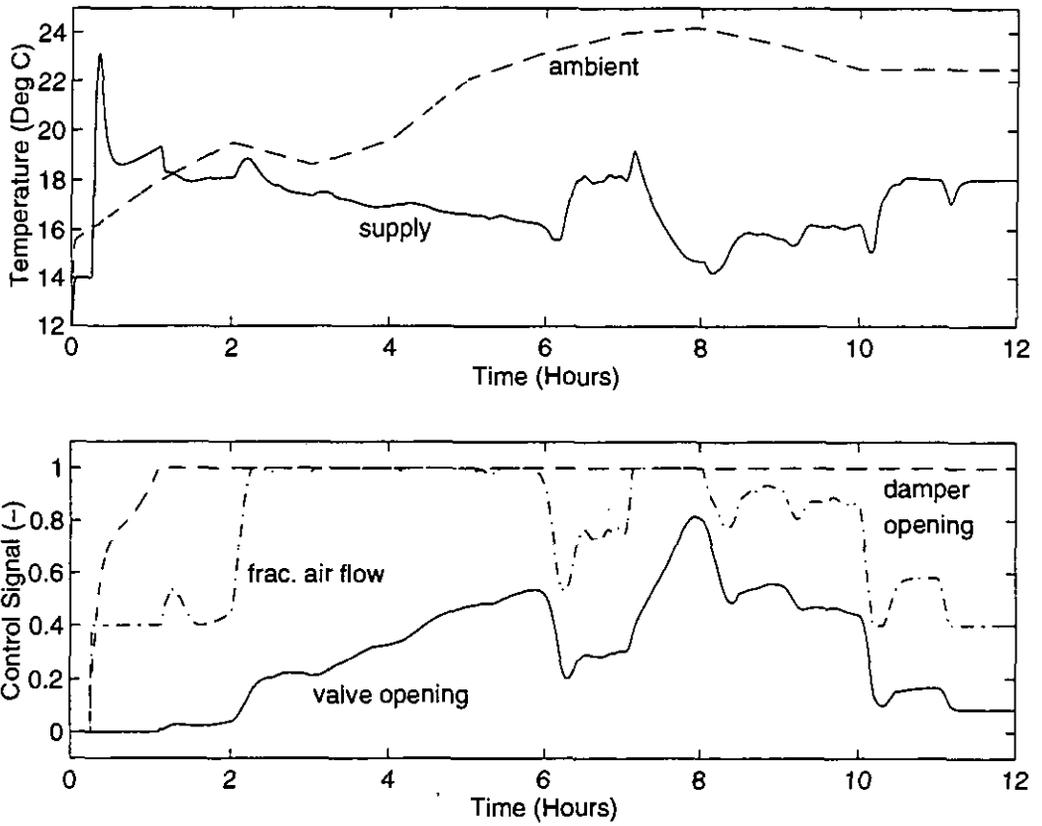
Figure 4: Test data on second day with no faults
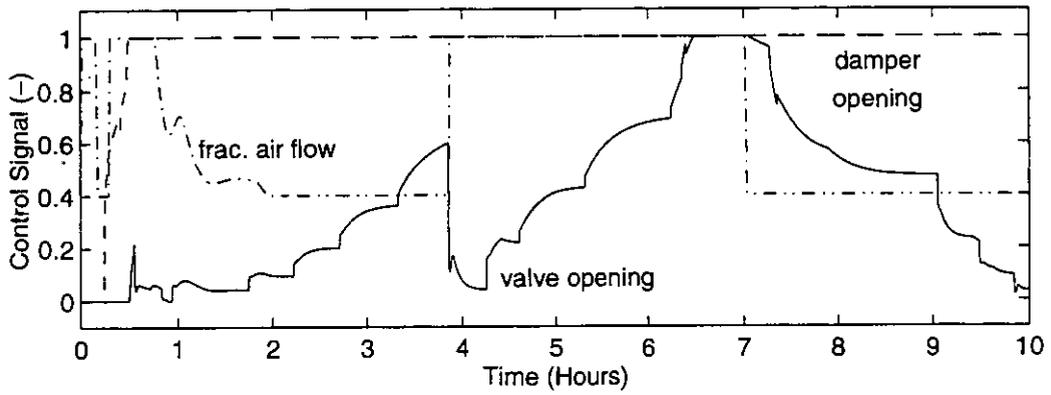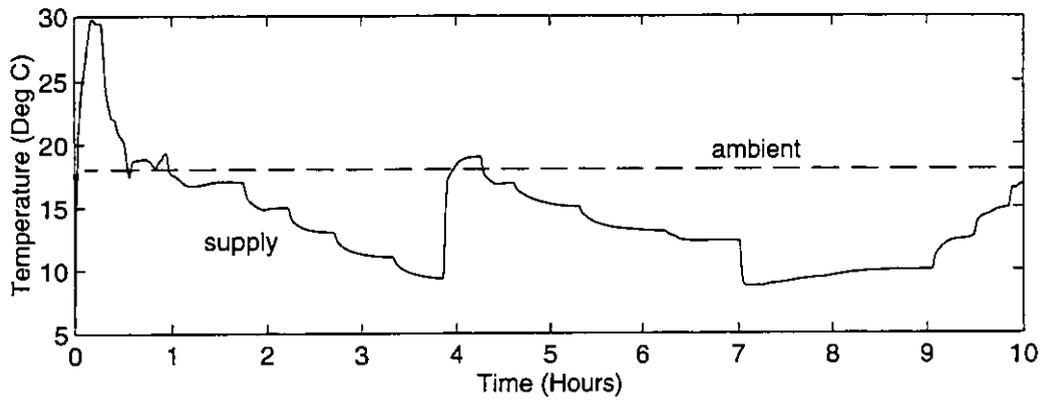
Figure 5: The training data
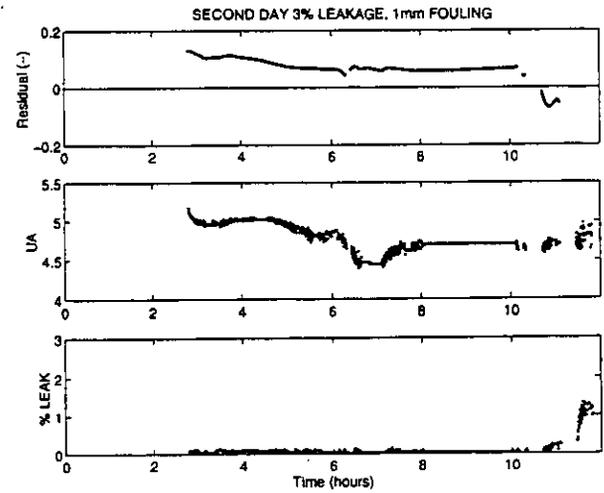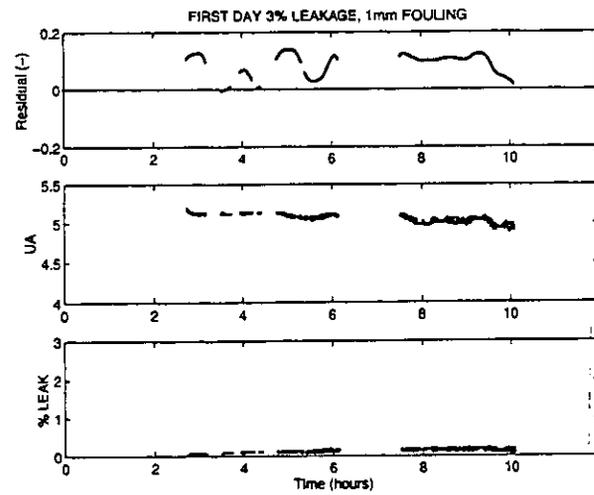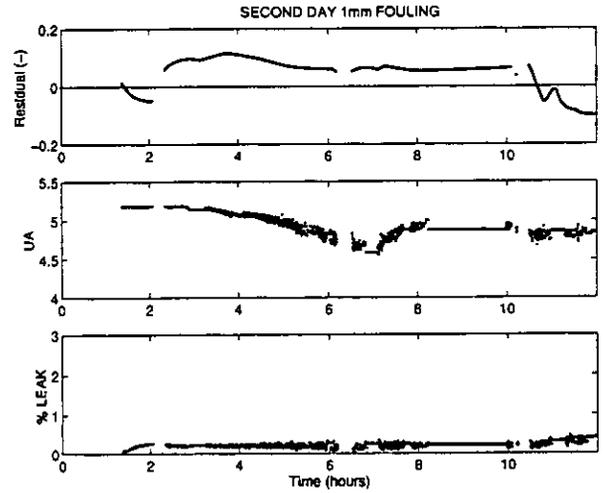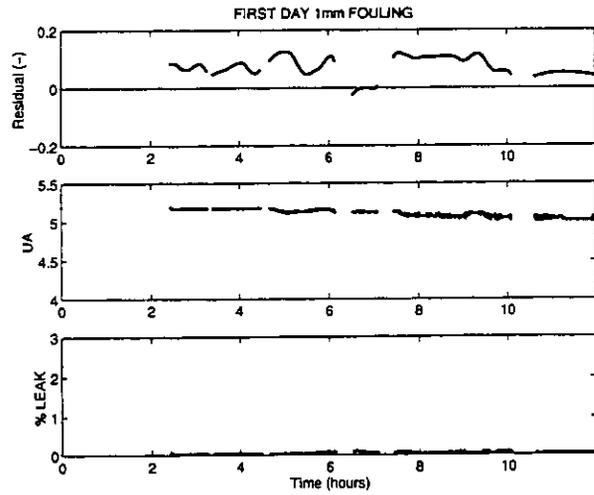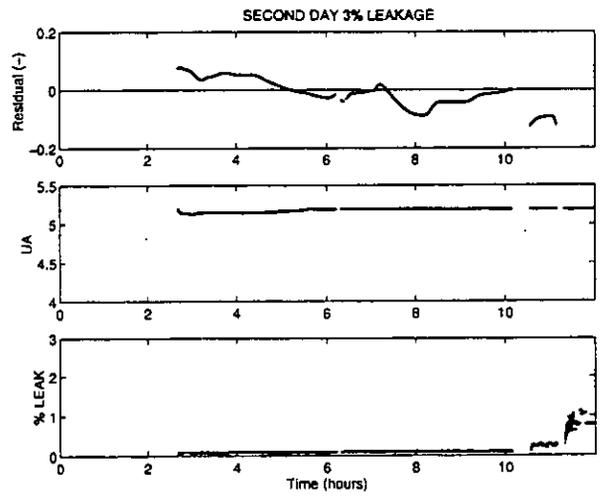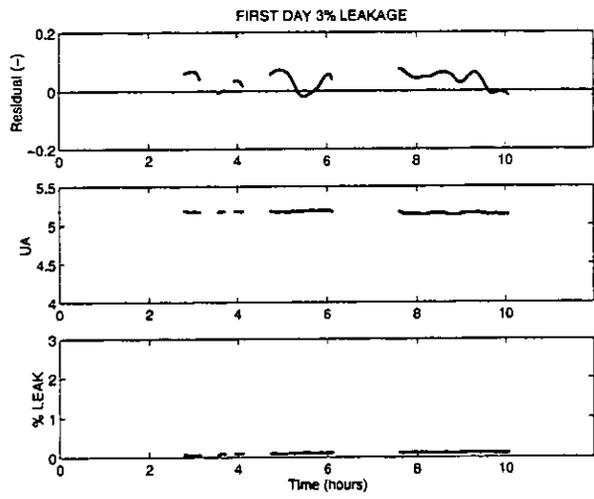
Figure 6: Test results

Fault Detection in an Air-Handling Unit
Using Residual and Recursive Parameter Identification Methods

Won Yong Lee
Korea Institute of Energy Research, Taejon, Korea

. **Cheol Park**
George E. Kelly
National Institute of Standards and Technology, Gaithersburg, Md. USA

ABSTRACT

A scheme for detecting faults in an air-handling unit using residual and parameter identification methods is presented. Faults can be detected by comparing the normal or expected operating condition data with the abnormal, measured data using residuals. Faults can also be detected by examining unmeasurable parameter changes in a model of controlled system using a system parameter identification technique. In this study, Auto Regressive Moving Average with eXogenous input (ARMAX) and Auto Regressive with eXogenous input (ARX) models with both Single-Input/ Single-Output (SISO) and Multi-Input/Single-Output (MISO) structures are examined. Model parameters are determined using the Kalman filter recursive identification method. This approach is tested using experimental data from a laboratory Variable Air Volume (VAV) air handling unit operated with and without faults.

INTRODUCTION

Fault detection and diagnosis of Heating Ventilation Air-Conditioning (HVAC) systems

is an important part of maintaining proper performance, reducing energy consumption, increasing reliability, and availability of the system. One of main purposes of on-line monitoring and diagnosis is the early detection of failures of equipment and sensors used in the control of HVAC system.

Studies on fault detection are extensive and various approach have been proposed. Willsky (1976) examined statistical techniques for the detection of failures in stochastic dynamic systems. Isermann (1984) surveyed existing fault detecting and diagnosing methods based on the estimation of unmeasurable process parameters and state parameters. Patton et al. (1989) also provided an overview of various fault detection and diagnosis methods by presenting research which included many references to application case studies. Frank (1990) reviewed the state-of-the-art fault detection and isolation in automatic process using analytical redundancy.

In recent years, several schemes for fault detection in HVAC systems have been investigated. Liu and Kelly (1989) proposed a rule-based diagnostic method for fault detection. Anderson et al. (1989) studied statistical analysis preprocessors and rule-based expert systems to monitor and diagnose HVAC system faults. Pape et al. (1991) developed a methodology for fault detection in HVAC systems based on optimal control. In order to detect faults in system operation, deviation from optimal performance were sensed by comparing the measured system power with the power predicted by using the optimal control strategy. Norford et al. (1993) presented a method for diagnosing fault in HVAC systems using the parametric models of consumed electric power.

In this paper, faults and symptoms were studied using changes in physical quantities, such as the deviation of temperature, pressure, or flow rate from their normal operating points. When a process operates under normal conditions, the process parameters should be at their normal values. A fault in the system can be detected by observing the residual value, which is the difference between the normal (or expected) data and abnormal operating data. If some physical change in the equipment causes a deviation from the normal state, the model parameters of the process will also deviate from their normal values. These parameters can be estimated for fault-free and fault-containing systems using parameter identification methods.

Faults are detected when a specified threshold was exceeded. The threshold can be determined by using statistical methods. A three-sigma limit (three standard deviations) is often used as a threshold value (Montgomery et al. 1994, Rose et al. 1993, Farnum 1992, Fasolo et al. 1992).

There are two types of faults: complete (or abrupt failures) and performance degradations. Complete failures are severe and abrupt faults. Performance degradations are gradually evolving faults. Although there are many kinds of potential faults in an air-handling unit, the eight different equipment and instrumentation faults shown in Figure 1 were studied in this study, based on experimental testings.

SYSTEM UNDER TEST

Air-handling Unit

The variable air volume (VAV) used in this study was based on a reference system (Kelly

1992) developed by the International Energy Agency (IEA) Annex 25. A simplified system layout diagram of the air-handling unit (AHU) is shown in Figure 2. The unit consists of fans, dampers, a cooling coil, sensors and controllers. The static pressure in the main supply duct is controlled to maintain a constant static pressure at each VAV box inlet by sensing the static pressure and controlling the speed of the supply fan. The flow difference between the return fan and the supply fan is controlled by a return fan with variable speed. The supply air temperature is controlled by the chilled water control valve to maintain a constant reference temperature. Heating and preheating of the outdoor air are not considered in this study.

Controllers

A proportional-integral-derivative (PID) controller using the velocity algorithm was designed to control the supply air temperature. Two other controllers were designed to control the static pressure in the supply duct and the difference between the supply fan and return fan flow rates. The PID velocity control algorithm is expressed as:

$$U(i) = U(i\text{-}1) + K_P[E(i) - E(i\text{-}1)] + K_I T_S E(i) + \frac{K_D}{T_S}[E(i) - 2E(i\text{-}1) + E(i\text{-}2)] \tag{1}$$

where $U(i)$ denotes the control signal at the ith sampling instant, $E(i)$ is the error at the ith sampling time defined by the difference between the set point value and the measured value, and T_S is the sampling period. The parameters K_P, K_I, and K_D are the proportional, integral, and derivative gains of the PID velocity control algorithm, respectively. The sampling period was 10

404

seconds.

The supply air temperature was controlled at 14.5°C (58°F) using a three-way control valve. The supply duct static pressure controller maintained the static pressure at 249 Pa (1.0 in. H$_2$O) in the supply duct by modulating the supply fan speed. The return fan speed was controlled to maintain the return air flow rate at 0.472 m^3/s (1000 cfm) below the supply air flow rate.

In the present study, the controller gain was approximately determined using a simple first-order transfer function for the system with the delay term obtained from a step change in the set point. The transfer function is given by:

$$T(S) = \frac{K_S}{1 + T_C S} e^{-T_D S} \tag{2}$$

where K_S is the system gain, T_C is the time constant, T_D denotes the dead time, and S is the Laplace variable. From the experimental data, the K_S, T_C, and T_D values for the supply air temperature controller were determined to be 1.02 K/V (1.836°F/V), 80 seconds, and 20 seconds, respectively. Using this approximate transfer function, the PID controller gains were first adjusted to minimize the integral absolute error over time (Dorf 1980) and then modified by experiments on the actual system. The final PID gains used for this three-way valve controller were K_P = 1.5 V/K (0.8333 V/°F), K_I = 0.0157V/sK (0.0087 V/s°F), and K_D = 10.6 Vs/K (5.8889 Vs/°F), respectively. The controllers for the static pressure and the air flow rate

difference did not used derivative terms due to the fast response of the controlled variables. The proportional and integral (PI) controller gains were also determined by computer simulations and then modified by experiments (Lee et al. 1994). Normal operating conditions for the controller tuning and the fault detection tests are given in Table 1.

To smooth the measured data and reduce the effect of random noise, smoothing filters were applied to the measured supply duct pressure, the measured flow rates, and the supply air temperature. The smoothed values were then used by the controllers. The following equation was employed:

$$M_{ss}(i) = \alpha M_s(i) + (1-\alpha)M_{ss}(i-1)$$

(3)

where M_{ss} is the smoothed measurement, α is the smoothing weight factor, M_s is the actual measurement, and i is the current sampling instant. A value of α equal to 0.7 was employed.

TECHNIQUE OF FAULT DETECTION

Residual Method

Faults in a broad sense result in symptoms that involve the deviations of measured values from their normal operating points. A fault can be detected by observing residual values which are defined as the differences between actual measured values under a fault condition and the expected values under normal operation.

The residual of the supply air temperature, R_T, was defined as:

$$R_T = T_S - T_{S,SP} \qquad (4)$$

where T_S is the supply air temperature and $T_{S,SP}$ is the supply temperature setpoint.

The residual of the supply duct static pressure, R_P, was defined as

$$R_P = P_S - P_{S,SP} \qquad (5)$$

where P_S is the measured static pressure value and $P_{S,SP}$ is the static pressure setpoint value.

The residual of the flow difference between the supply and return fans, R_Q, was defined as

$$R_Q = Q_D - Q_{D,SP} \qquad (6)$$

where Q_D is the difference between the measured supply and return air flow rates and $Q_{D,SP}$ is the setpoint value.

The residual of the cooling coil control signal, R_U, was defined as

$$R_U = U_{CC} - U_{CC,SP} \qquad (7)$$

where U_{CC} is the control signal for the cooling coil valve as determined by Equation 1, and $U_{CC,SP}$ is the setpoint value or reference value at a normal condition.

The value of U_{CC} controls the supply water temperature for a three-way valve or the supply water flow rate for a two-way valve. At a normal operation, U_{CC} is the same as $U_{CC,SP}$. However, when a fault occurs, U_{CC} will deviate from $U_{CC,SP}$. A problem, however, arises because $U_{CC,SP}$ is not a fixed value, but varies with the load on the AHU. One possible way of handling this difficulty is to calculate the mean value and standard deviation of $U_{CC,SP}$ every sampling time using a number of data points (e.g. 20 data points) from previous time steps. This works well for systems subject to slowly varying loads and for quickly developing (complete) faults.

Another approach for determining $U_{CC,SP}$ is to use a reference model that is developed under normal conditions. The reference model is a function of load change and environmental conditions, such as outdoor air temperature and humidity. The residual is the calculated difference between the measured value and the estimated value from the reference model. If there is no fault, the measured and the estimated values should be the same. Deviations between the measured and the estimated values indicate the presence of faults. This reference model approach is essential for detecting long-term performance degradation, such as the fouling of heating and cooling coils.

The residuals for the actuators are defined as the difference between the input control signal and the measured positions of the actuators or speed signals of the fans. The residuals of the supply fan speed, R_{NS}, and the return fan speed, R_{NR}, are given by

$$R_{NS} = N_s - U_s \tag{8}$$

$$R_{NR} - N_R - U_R \qquad (9)$$

where N_S and N_R are the measured values of the supply and the return fan speeds, and U_S and U_R are the control signals for the supply and return fans, respectively.

The residual of the cooling coil valve position was defined as

$$R_V - V_P - U_{CC} \qquad (10)$$

where Vp is the three-way cooling coil valve position determined by monitoring a variable resistor on the valve stem. At normal operation, R_{NS}, R_{NR}, and R_V are approximately zero. However, these values deviate from zero, when an actuator fault occurs.

Parameter Identification Methods

When a process operates under normal conditions, the parameters in a continuously updated model of the process will be at their normal values. If some physical changes in the system causes deviation from the normal state, some or all of the model parameters will deviate from these normal values. The fault condition can then be detected as shown in Figure 3.

The parameters of a model can be estimated by employing a system identification method. In

this study, Multi-Input/Single-Output (MISO) and Single-Input/ Single-Output (SISO) Auto Regressive Moving Average with eXogenous input (ARMAX) models and Auto Regressive with eXogenous input (ARX) models were used , and model parameters were recursively identified using Kalman filters.

The general structure of the SISO or MISO ARMAX models (Ljung 1987) is given by

$$A(q)y(t) - B(q)u(t-nk) \cdot C(q)e(t) \tag{11}$$

where n is the number of time step delays from input to output. $A(q)$ and $C(q)$ are polynomials in terms of the time delay operator q^{-1}

$$
\begin{aligned}
A(q) &- 1 + a_1 q^{-1} + ... + a_{na} q^{-na} \\
C(q) &- 1 + c_1 q^{-1} + ... + c_{nc} q^{-nc}
\end{aligned}
\tag{12}
$$

$$
B(q) - \begin{vmatrix}
b_{11} & b_{12} & .... & . \\
b_{21}q^{-1} & b_{22}q^{-1} & .... & . \\
. & . & .... & . \\
b_{nb1}q^{-nb+1} & b_{nb2}q^{-nb+1} & .... & b_{nbnu}q^{-nb+1}
\end{vmatrix}
\tag{13}
$$

where $B(q)$ is an nb x nu matrix. The quantities na, nb, and nc are the orders of the polynomials, and nu is the number of input variables. For the SISO model, $nu = 1$.

For the first order MISO model with a delay of two sampling times, equation 11 becomes

$$y(t) = -a_1 y(t-1) + b_{11} u_1(t-2) + b_{12} u_2(t-2) + \ldots + b_{1n} u_n(t-2) + e(t) + c_1 e(t-1) \tag{14}$$

where n is the number of input variables.

As a special case of the ARMAX model, the ARX model structure is given by

$$A(q)y(t) = B(q)u(t-nk) + e(t) \tag{15}$$

This equation can also be written explicitly for a first order model with a delay of two sampling times as

$$y(t) = -a_1 y(t-1) + b_{11} u_1(t-2) + b_{12} u_2(t-2) + \ldots + b_{1n} u_n(t-2) + e(t) \tag{16}$$

Recursive Parameter Estimation Using Kalman Filter

The typical recursive parameter identification algorithm (Ljung 1987, 1991; Johansson 1993) is given by

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(y(t) - \hat{y}(t)) \tag{17}$$

$$y(t) = \psi(t)^T \theta_0 + e(t) \tag{18}$$

$$\theta_0(t) = \theta_0(t-1) + w(t) \tag{19}$$

411

where $\hat{\theta}(t)$ is the parameter estimate at time t. $\psi(t)$ is the regression vector which contains old values of observed inputs and outputs, $y(t)$ is the observed output at time t, and $\hat{y}(t)$ is the prediction of the value $y(t)$ based on observations up to time $t-1$ and the current model at time $t-1$. θ_0 represents the true description of the system, $e(t)$ is the noise source with the variance, $R_2 - E[e^2(t)]$, and $w(t)$ is assumed to be white Gaussian noise with covariance, $R_1 - E[w(t)w^T(t)]$.

The gain $K(t)$ determines how the current prediction error $[y(t) - \hat{y}(t)]$ updates the parameter estimate. It is typically chosen as

$$K(t) - Q(t)\psi(t)$$

(20)

The Kalman filter algorithm is given by

$$\hat{y}(t) - \psi^T(t)\hat{\theta}(t-1)$$

(21)

$$Q(t) - \frac{P(t-1)}{R_2 + \psi^T(t)P(t-1)\psi(t)}$$

(22)

$$P(t) - P(t-1) + R_1 - \frac{P(t-1)\psi(t)\psi^T(t)P(t-1)}{R_2 + \psi^T(t)P(t-1)\psi(t)}$$

(23)

An optimal choice of $Q(t)$ is computed from Equation 17 through 23.

Threshold Checking

The proper choice of the threshold values are important for detecting faults. The thresholds can usually be determined from statistical properties of the process. The concept of statistical method is very straightforward. If a measurement is greater than an upper limit threshold limit or is lower than a lower threshold limit, the process is said to be out of the normal state and a fault is presumed to have occurred.

In this study, a three-sigma limit was used as the threshold value. If the measurable characteristic x of an item is normally distributed with the mean $\bar{X}$ and the standard deviation σ, it is possible to find the probability that x will lie within a fixed interval. The probability that x will fall within the interval $[\ \bar{x} - 3\sigma,\ \bar{x} + 3\sigma\ ]$ is 0.9973. The threshold for a measured variable x was specified as $|x - \bar{x}| - 3\bar{\sigma}$, where $\bar{X}$ denotes the assumed mean and $\bar{\sigma}$ denotes the assumed standard deviation. Typically, $\bar{X}$ and $\bar{\sigma}$ are calculated from a set of test data(Fasolo and Seborg 1992, Farnum 1992). When residual method is used, $|X - \bar{X}|$ is the value of the residual, and when the parameter identification method is used, $|X - \bar{X}|$ is the difference between the estimated value and the mean value at normal condition.

TEST RESULTS AND DISCUSSION

As previously mentioned, a fault in the system can be detected by observing the residual values. When an input-output model is used for the system description, a fault can be detected by looking for changes in the model parameters, which are estimated by using model identification methods.

The first order system models were used in this study to estimate the model parameters before and after a fault occurs. Pseudo-linear ARMAX equations and linear ARX equations were employed. The structure of the multi-input ARMAX system model is given by

$$P_S(t) - - a_{P1}P_S(t\text{-}1) + b_{P1}U_P(t) + b_{P2}U_Q(t) + b_{P3}\theta(t) + e(t) + c_{P1}e(t\text{-}1) \tag{24}$$

$$\begin{aligned}T_S(t) - - a_{T1}T_S(t\text{-}1) + b_{T1}U_{CC}(t\text{-}2) + b_{T2}Q_S(t\text{-}2) + b_{T3}T_M(t\text{-}2)\\ + b_{T4}H_M(t\text{-}2) + e(t) + c_{T1}e(t\text{-}1)\end{aligned} \tag{25}$$

$$Q_D(t) - - a_{Q1}Q_D(t\text{-}1) + b_{Q1}U_P(t) + b_{Q2}U_Q(t) + b_{Q3}\theta(t) + e(t) + c_{Q1}e(t\text{-}1) \tag{26}$$

where P_S is the static pressure at the supply duct, Q_D is the flow difference between the supply and return fans, and T_S the supply air temperature. The variable θ is the angle that the recirculating air damper makes with a plane perpendicular to the direction of flow, Q_S is the supply air flow rate, T_M is the mixed air temperature, and H_M is the mixed air humidity ratio. The subscripts P, Q, and T denote the supply air static pressure, the flow difference between the supply and return fans, and the supply air temperature, respectively.

If noise is not explicitly taken into account, Equations 24 through 26 become an ARX model. A disturbance influences the output and this output changes the feedback signal to the controller, which in turn changes the controller output. Since the control signal includes information on disturbances in the SISO model, only the control signal and the output need to be considered. The structure of the simplest SISO ARX model becomes

$$P_S(t) = -a_{P1}P_S(t\text{-}1) + b_{P1}U_P(t) + e(t)$$

(27)

$$Q_D(t) = -a_{Q1}Q_D(t\text{-}1) + b_{Q1}U_Q(t) + e(t)$$

(28)

$$T_S(t) = -a_{T1}T_S(t\text{-}1) + b_{T1}U_T(t\text{-}2) + e(t)$$

(29)

where $e(t)$ is the equation error.

Four different identification methods were compared by using average absolute errors (AAE) defined by

$$AAE(y) = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

(30)

where y and $\hat{y}$ are the observed and the predicted values.

The HVAC system was tested for the parameter identification method under the load condition shown in Table 1. It is important to note, however, that system identification parameters may change with load changes. Since load conditions often vary slowly in actual building systems,

one might expect that dramatic changes in the identified parameters would indicate quickly developing (complete) faults.

Table 2 shows the AAEs of T_S, P_S, and Q_D calculated for four ARMAX and ARX models. It can be seen that all the estimated values are close to each other for the case of a constant load on the AHU and no external disturbances. Since the result of the SISO ARX model is almost the same as the other results, only the results from the model corresponding to Equations 27 through 29 will be discussed below.

For system parameter identification, normalized input values are used. The supply air temperature is divided by room air temperature and the control signals and actuator signals are normalized to make their maximum value unity. All the faults were introduced after 1,500 second in operation. If complete faults occur, the control and the measured signals change significantly. To detect these kind of faults, it is necessary to use feed-back signals from the system and the controllers. It should be noted that those signals which were *momentarily* out of bound of the given thresholds during the observation periods were ignored in this study.

Fault 1 is a complete failure of the return fan. The return fan was changed from normal operation to an abruptly shut off condition. Since the return fan was controlled to maintain the return fan air flow rate below the supply air flow rate by a fixed amount, the return fan fault caused the return fan flow to change dramatically. The best variables for detecting this fault are the return fan rotational speed and the air flow rate difference between the supply and return flow rates.

Figures 4a shows system variables such as supply air temperature, air flow rate difference, and pressure at the supply air duct. Figures 4b through 4h show the residuals of the supply air pressure, the flow rate difference between supply and return air fans, the supply air temperature, the three-way valve control signal, the supply fan rpm, the return fan rpm, and the three-way valve position, respectively. Residual values in Figure 4c shows that the return fan failure causes the flow rate difference to jump suddenly, while the supply air pressure and temperature are maintained constant. The significant fault signature can be seen in the residual values of the return fan speed (Figure 4g). If the return fan is stopped, the controller attempts to compensate by increasing the control signal. However, the fan is not controlled and the fan failure generates a big change in the residual value of the fan rotational speed.

As shown in Figure 5, this fault can also be detected by the parameter identification method. The identification parameters of the flow difference are greatly changed and deviated significantly from the threshold due to the return fan fault (Figure 5d), while the parameters of supply air pressure stays within the threshold range (Figure 5b) and parameters of the supply air temperature deviates only slightly from the threshold (Figure 5f).

The residual values and the changes in the model parameters for this fault and the other seven faults given in Figure 1 are summarized in Table 3 and Table 4, respectively.

Fault 2 is a complete failure of the supply fan. The supply fan was changed from normal operation to an abruptly shut off condition. Since the static pressure at the supply duct is

controlled at a certain value by modulating this supply fan speed, the failure in the supply fan significantly influences the static pressure at the supply duct. From Table 3, it is seen that the supply pressure abruptly decreased to a zero value (residual value -0.249 kPa [-1.0 in. H_2O]), and also the flow rate difference is decreased to zero value. This failure causes the supply fan controller to increase the output control signal to its maximum value in an attempt to increase the static pressure in the main supply duct.

To keep the flow difference between the supply and return fan flows positive, the return fan rotational speed is also decreased to zero. Since there is no air flow through the cooling coil, the air temperature in the supply duct slowly increases and thus the cooling coil control signal also increases. The best variables chosen for fault detection of this fault are the supply fan rotational speed residual and the static air pressure residual. The fault can also be detected through the parameter identification scheme, as shown in Table 4. The identification parameters of pressure and flow rate difference are greatly changed but those for supply air temperature changed little.

Fault 3 is a complete failure of the chilled water-circulation pump. The pump was changed from normal operation to an abruptly shut-off condition, as might result from mechanical or electrical problems. From Figure 6, it is seen that the supply air temperature is changed temporarily and then returns to normal. If the pump fails, the mixed water flow rate through the three-way valve is immediately decreased and thus the supply air temperature is increased and the error signal to the controller is increased. This error signal is reduced by increasing the three-way valve opening position. From Figure 6e and Table 3, it is seen that the cooling coil valve control signal is

increased above the threshold value to compensate for the pump failure. The identification parameters of the supply air temperature are changed due to the pump fault, while the parameters of supply pressure and flow difference are within the thresholds as seen in Figure 7.

Fault 4 is the fault condition where the cooling coil control valve sticks in a certain position. In this case, the residual values after the fault do not change significantly in spite of fault occurrence. If there is no external disturbance, the output condition should be unchanged. But, there is a small difference between the normal or expected signal and the measured value. In the case of noise and external disturbances, such as the load change and fresh air temperature change, the supply air temperature may be slightly changed. As time goes by, this small change of temperature and the small difference between the setpoint value in the normal case and the measured value cause the control signal to change continuously due to the integral term of the controller. It can be said that it is difficult to detect this fault from the supply air temperature residual, but this fault can be detected over time from the change of input control signal . From Table 3, the profile of the control signal residual of the cooling coil valve can be seen to be slightly different from the one without the fault. However the residuals related to the supply air pressure and the air flow difference are not changed. From Table 4, it can be seen that the parameters for the supply air temperature change slightly, while the parameters of the flow difference do not change. The parameters of the supply air pressure are within the threshold values.

Fault 5 is the case when a temperature sensor undergoes a complete failure. From a given set of

sensor readings, a normal operating range for each temperature sensor can be established based upon expert knowledge about the process, sensor characteristics, and historical data bases. Once the range of each measurement is selected, it can be determined whether the measurement is within a normal range or not.

If the temperature sensor is disconnected, sometimes the measured temperature oscillates randomly. If the supply air temperature range is out of the normal operating range, typically between 0°C (32°F) and 40°C (104°F), when the system operates in cooling mode, the temperature sensor is known to be at fault. The temperature signal can be set to zero to make the output signal constant at the fault condition. The input signal is abruptly changed and the controller attempts to compensate by increasing the control signal. However, the temperature signal is not changed and the temperature sensor fault generates a big change in the cooling coil valve control residual, while the supply air pressure and the flow difference residuals are not changed as seen in Table 3. It is seen from Table 4 that this results in the parameters for the supply temperature changing greatly, while the parameters of the flow difference are unchanged. The parameters of the supply air pressure are within the threshold values.

Fault 6 is a complete failure of the static pressure transducer in the air supply duct. For this fault, the output of the pressure transducer is abruptly changed to zero due to electrical or mechanical problems. The pressure transducer generates the feed back signal to the supply fan controller. The failure in the pressure transducer significantly influences the static pressure in the supply duct. From Table 3, it is seen that the pressure residual is greatly changed, but the

--

supply temperature and the flow difference due to step change are not significantly changed. The impulse response values in Table 3 can be ignored, since for fault detection, it is best to consider only step and ramped values. Because the feed back pressure signal is zero (actual value is not zero), this controller makes the supply fan control signal maximum to try to maintain the feed back pressure signal at the reference value. Unlike the supply fan failure, the supply fan operates at its maximum rotational speed and the flow difference and the supply temperature are controlled normally after some transient changes.

This fault can be detected through the parameter identification scheme as shown in Table 4. The identification parameters of the pressure are significantly changed, while those of the flow rate difference are not changed. The parameter changes of the supply air temperature are insignificant.

Fault 7 is a failure of the supply fan flow station. The output signal is abruptly changed from its normal value to zero due to a differential pressure transducer failure or a mechanical fitting problems. Since the return fan was controlled to maintain a constant flow difference using the flow station signal, the flow station failure causes the return fan control signal to change, which is proportional to rotational speed. Residual values in Table 3 show that the supply fan flow station failure causes the flow rate difference residuals to jump suddenly, while the supply air pressure and temperature residuals are constant. The significant fault signature appears in the residual values of the flow rate difference. As the supply flow station output signal is zero, the return fan controller attempts to compensate by sending a lower control signal to decrease the return fan flow rate, which decrease the fan speed. The actuator residual values are not changed.

As shown in Table 4, the fault can be also be detected through the parameter identification method. The identification parameters of the flow difference are greatly changed due to the supply flow station fault, while the parameters of the supply air temperature are slightly changed. The parameters of the supply air pressure are within the threshold values.

Fault 8 is a failure of the return fan flow station. The output signal is changed abruptly from its normal value to zero due to the same problems as in fault 7. If the return flow station output signal is reduced to zero, the flow difference signal is increased and the return fan controller attempts to compensate by sending a higher control signal to increase the return flow and reduce the flow difference. Compared with fault 7, this fault has the opposite effect on the residual values. As shown in Table 4, the fault can be detected through the parameter identification method. The identification parameters of flow difference are greatly changed due to the return fan flow station fault, while the parameters of the supply air pressure and temperature are slightly changed. The parameters of the supply air pressure are changed but not nearly as much as those of the flow difference.

For the eight complete faults discussed above, the two fault detection methods can be used to detect the faults of a VAV air-handling unit. The residual method requires less computing time to calculate the residuals, but requires more sensors than parameter identification methods. The residuals after the faults display the unique fault signatures seen in Table 3. Thus, not only fault detection but also fault diagnosis is possible. The later is the subject of a companion paper (Lee et al. 1996).

SUMMARY

Residual and parameter identification methods were employed for fault detection in an air-handling unit of a building HVAC system. For parameter identifications, ARMAX and ARX models were employed with MISO and SISO structures to estimate model parameters recursively using the Kalman filter. Eight complete faults of equipment and sensors were tested under constant load conditions and for short time periods. These faults were examined using both residual and parameter identification methods using the laboratory measured data. The test results show that both methods can be used for detecting the presence of faults in the air-handling unit.

Faults were detected when residuals and identification parameters change significantly and thresholds were exceeded. Momentary indication of a fault was not accounted for, but continuous presence of fault signature for a reasonable time period was considered. The work was done for one load on the AHU. If building loads change rapidly, these methods may not detect the faults.

The proposed approach can be applied to practical problems when observation is made in a short time period under the assumption that the load remains constant. However, further investigation is needed for the load change cases. Fault signatures were developed for eight complete faults. The use of these signatures to diagnose a particular fault is the subject of a second paper.

REFERENCES

Anderson, D., L. Graves, W. Reinert, J.F. Kreider, J. Dow, and H. Wubbena. 1989. A quasi-real time expert system for commercial building HVAC diagnostics. *ASHRAE Trans.* 95(2).

Dorf, R. C. 1980. *Modern Control Systems, 3rd Ed.*

Farnum, N.R. 1992. Control charts for short runs: non-constant process and measurement error. *Journal of Quality Technology* 24(3): 138-144.

Fasolo, P.S. and D.E. Seborg. 1992. Fault detection in HVAC control system using statistical quality control charts. Dept. of Chemical and Nuclear Eng. University of California Santa Barbara, CA.

Frank, P.M. 1990. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - A survey and some new results. *Automatica* 26(3): 459-474.

Isermann, R. 1984. Process fault detection based on modeling and estimation method - A survey. *Automatica* 20: 387-404.

Johansson, R. 1993. *System Modeling and Identification*, Prentice-Hall.

Kelly, G.E. 1992. Description of a reference air-handling system. IEA Annex 25 working paper, Liege meeting.

Lee, W.Y., J.M. House, C. Park, and G.E. Kelly. 1996. Fault diagnosis of an air-handling unit using artificial neural networks. ASHRAE *Transactions* 102.

Lee, W.Y., Park, C., and Kelly, G.E. 1994. Fault detection of dampers and fans of an air-handling system. IEA Annex 25 working paper, MIT meeting.

Liu, S.T. and G.E. Kelly. 1989. Rule-based diagnostic method for HVAC fault detection. *Proceedings of Building simulation '89*, Vancouver.

Ljung, L. 1987. *System Identification, Theory for the User*, Prentice-Hall.

Ljung, L. 1991. *System Identification Tool Box-for use with MATLAB-User's Guide*. The Math Work, Inc.

Montgomery, D.C., J. B. Keats, G.C. Runger, and W.S. Messina. 1994. Integrating statistical process control and Engineering process control. *Journal of Quality Technology* 26(2): 79-87.

Norford, L.K. and R.D. Little. 1993. Fault detection and load monitoring in ventilation system. *ASHRAE Trans.* 99(1).

Pape, F.L.F., J.W. Mitchel, and W.A. Beckman. 1991. Optimal control and fault detection in heating, ventilating, and air-conditioning systems. *ASHRAE Trans.* 97(1): 729-745.

Patton, R., P. Frank, and R. Clark. 1989. *Fault Diagnosis in Dynamic Systems - Theory and Application*. Prentice Hall.

Rose, K.C.B. , R.J.M.M. Does, and Y. Schurink. 1993. Shewhart-type control charts for individual observation. *Journal of Quality Technology* 25(3): 188-198.

Willsky, A.S. 1976. A Survey of design method for failure detection in dynamic systems. *Automatica* 12: 601-611.

TABLE 1 Nominal Operating Conditions

Variables	Description	Nominal values
Q_S	Supply air flow rate (m^3/s)	1.5
T_{WI}	Inlet water temperature of the cooling coil (°C)	10.2
T_M	Mixed air temperature (°C)	22.0
H_M	Mixed air humidity (dew point temperature, °C)	12.5

TABLE 2 AAE Comparison of ARMAX and ARX Models (Pump Fault Condition)

Model structure		AAE(T$_S$) (°C)	AAE(P$_S$) (kPa)	AAE(Q$_D$) (m^3/s)
ARMAX	MISO	0.0243	0.0028	0.0988
	SISO	0.0282	0.0030	0.0997
ARX	MISO	0.0256	0.0028	0.0976
	SISO	0.0256	0.0030	0.1010

TABLE 3   Residual Values After Faults

	R_P (kPa)	R_Q (m³/s)	R_T (°C)	R_U (V)	R_{NS} (V)	R_{NR} (V)	R_V (V)
Fault 1	0	0.7256	0	0	0	-10.0	0
Fault 2	-0.249	-0.47	1.766(r)	4.9(r)	-10	0	0
Fault 3	0	0	1.5(i)	1.66	0	0	0
Fault 4	0	0	0	-0.06	0	0	0.11
Fault 5	0	0	-14.5	5.0(i)	0	0	- 6.0(i)
Fault 6	-0.249	0.55(i)	1.2(i)	1.33	-1.0(i)	0	0
Fault 7	0	-0.8	0	0.25	0	0	0
Fault 8	-0.65(i)	1	0	0	0	0	0

Note that (r) and (i) mean ramp and impulse responses after faults and other residual values are step changes.

TABLE 4   Changes in Identification Parameters After Faults (SISO ARX)

	a_{T1}	b_{T1}	a_{P1}	b_{P1}	a_{Q1}	b_{Q1}
Fault 1	0	-0.03	0	0	-0.4	0.42
Fault 2	-0.1797	-0.0427	-0.8432	0.0305	-0.954	0
Fault 3	-0.1778(i)	-0.05	0.05	0	0	0
Fault 4	-0.0551	-0.044	0.0713	0	0	0
Fault 5	0.7(i)	-3.586	0.05	0	0	0
Fault 6	0.25(i)	-0.05	-0.77	0.01	0	0
Fault 7	-0.05	0.05	0.05	0	-1.0	-0.15(i)
Fault 8	-0.04	0	0.2	-0.06(i)	-0.6	0.6
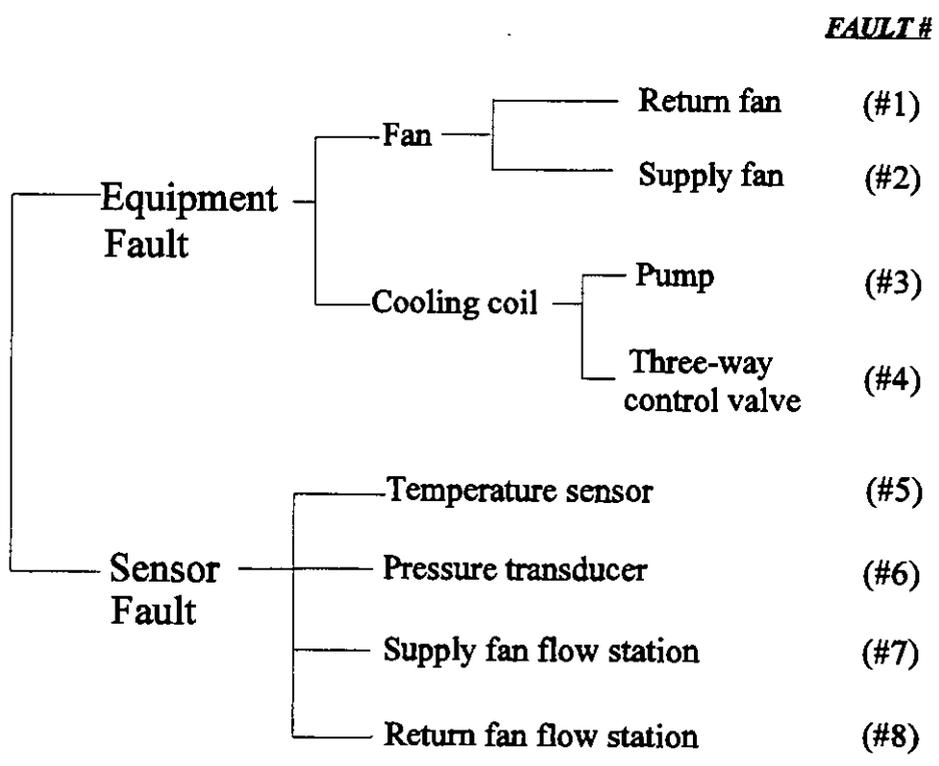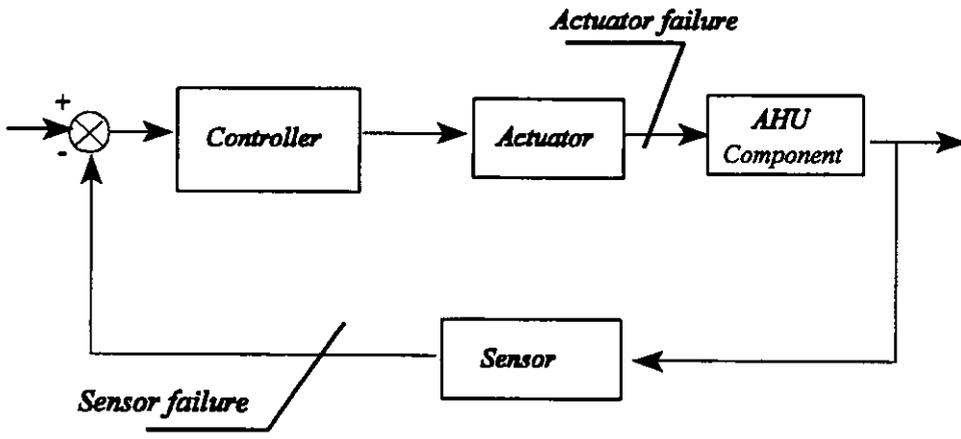
Figure 1 Fault situations.
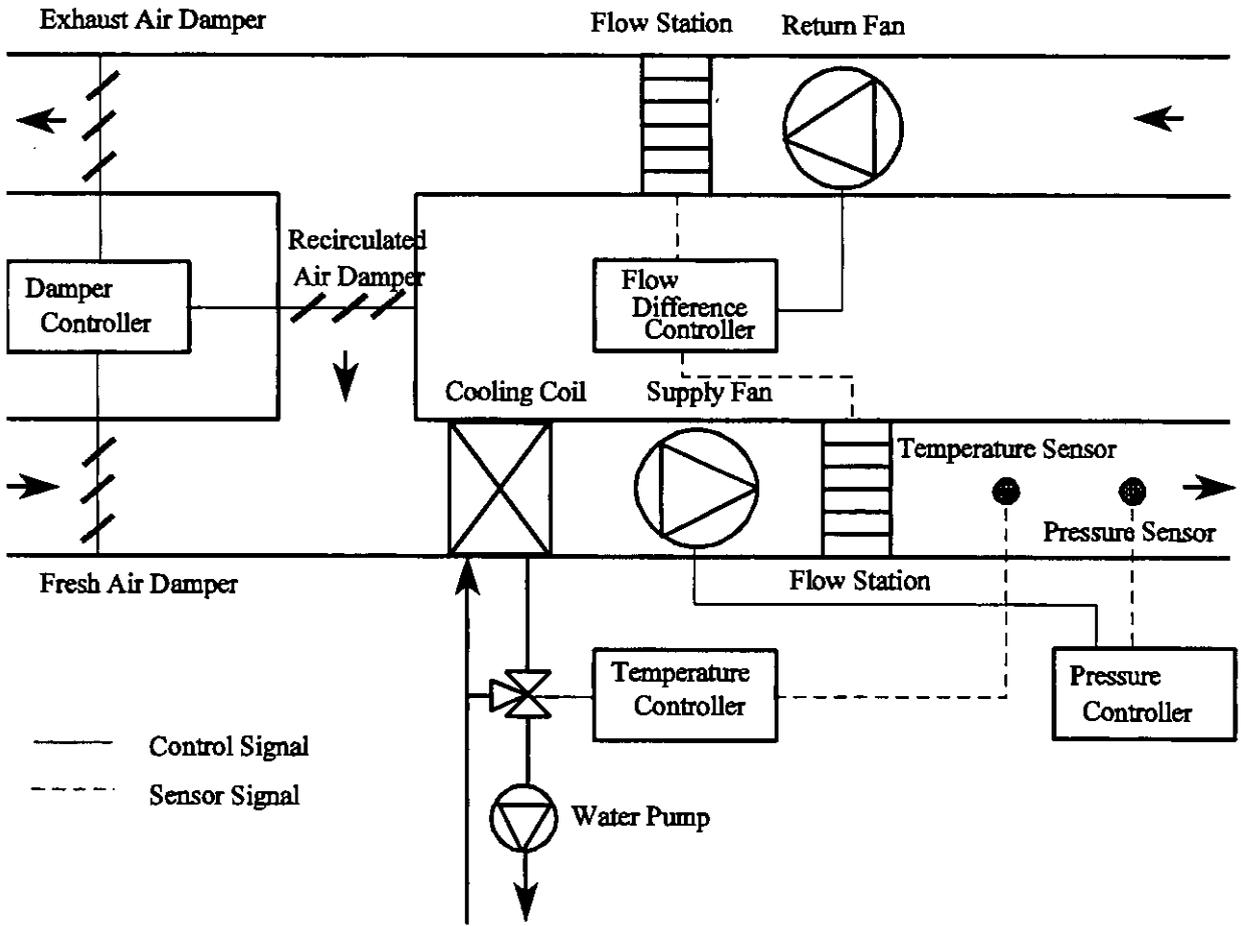
Figure 2 System layout diagram.
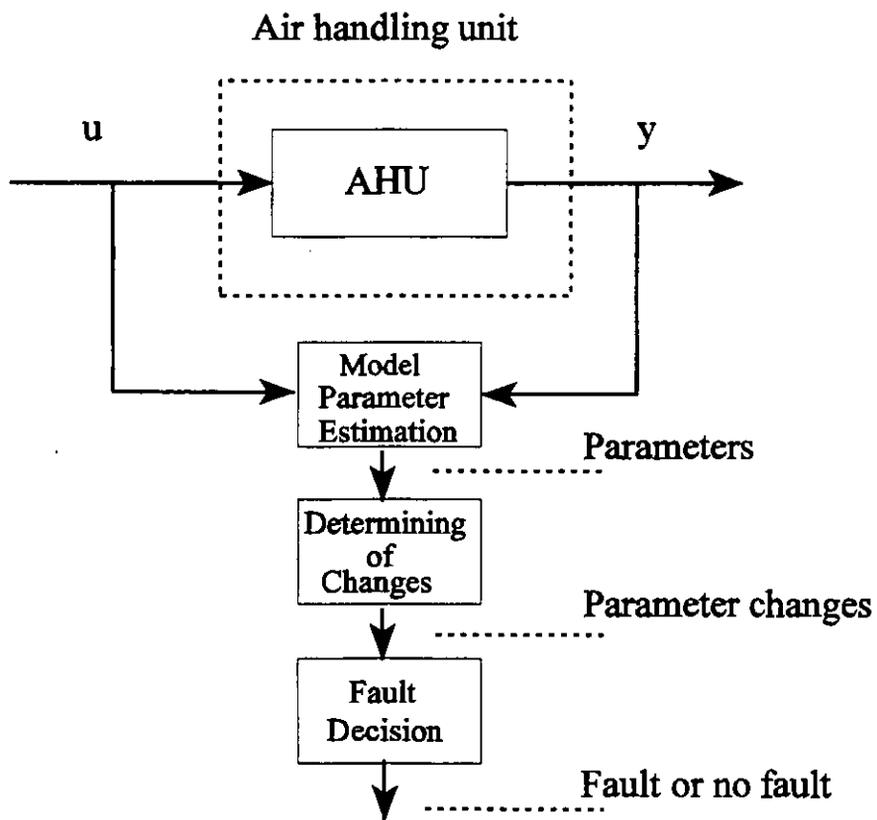
Figure 3   Fault detection using parameter estimation.
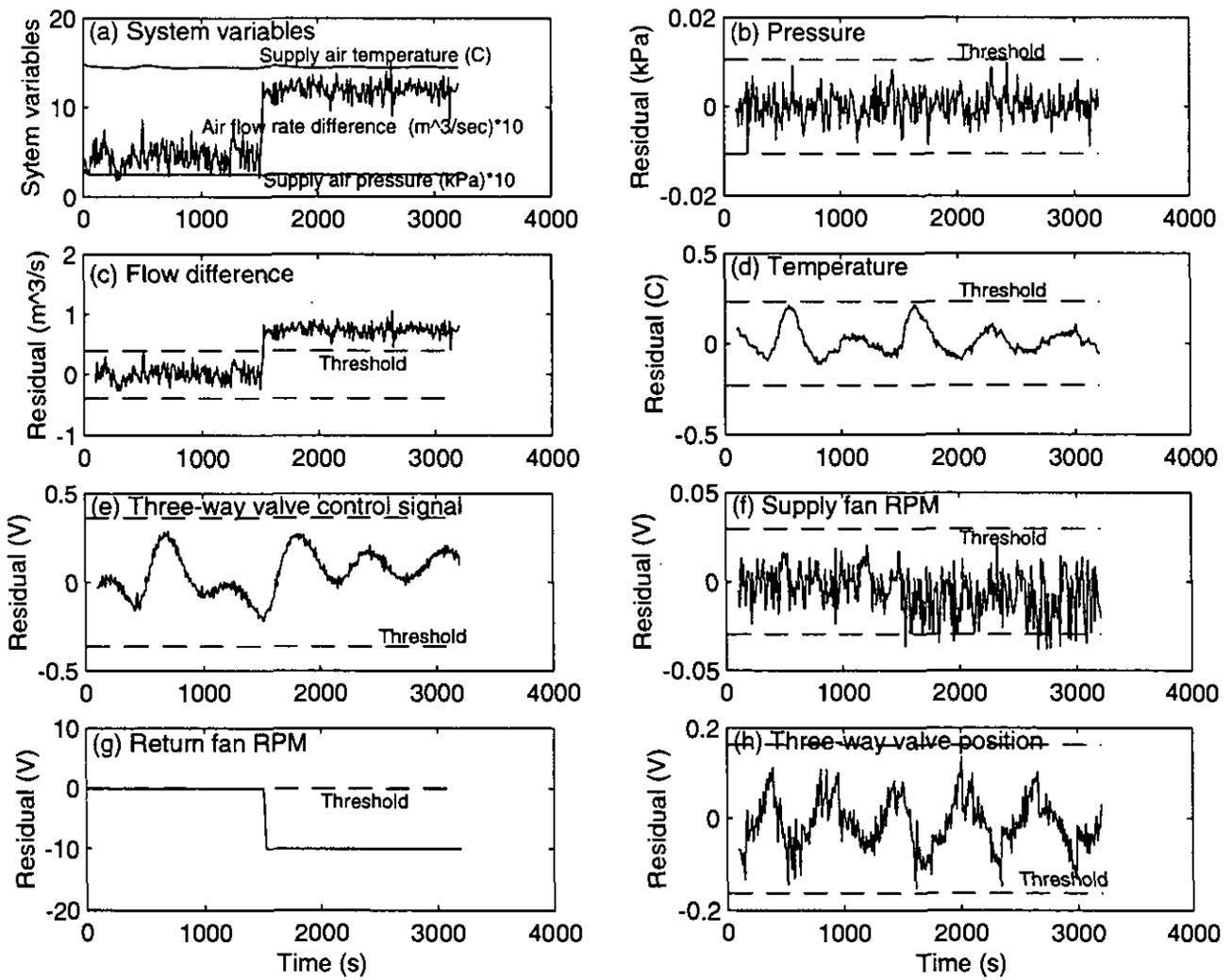
Figure 4 Residuals for fault 1.
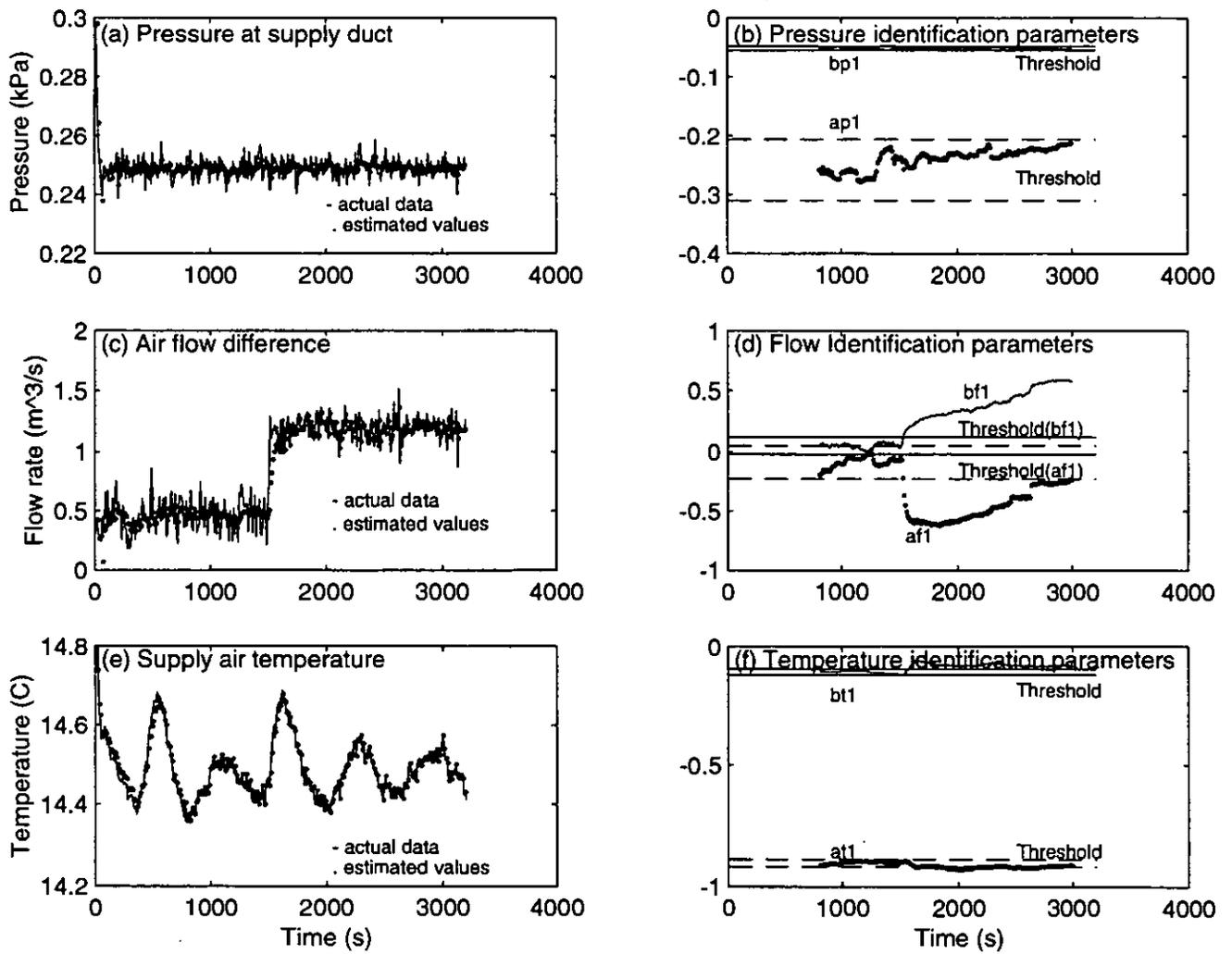
Figure 5   System identification result and identification parameters for fault 1.
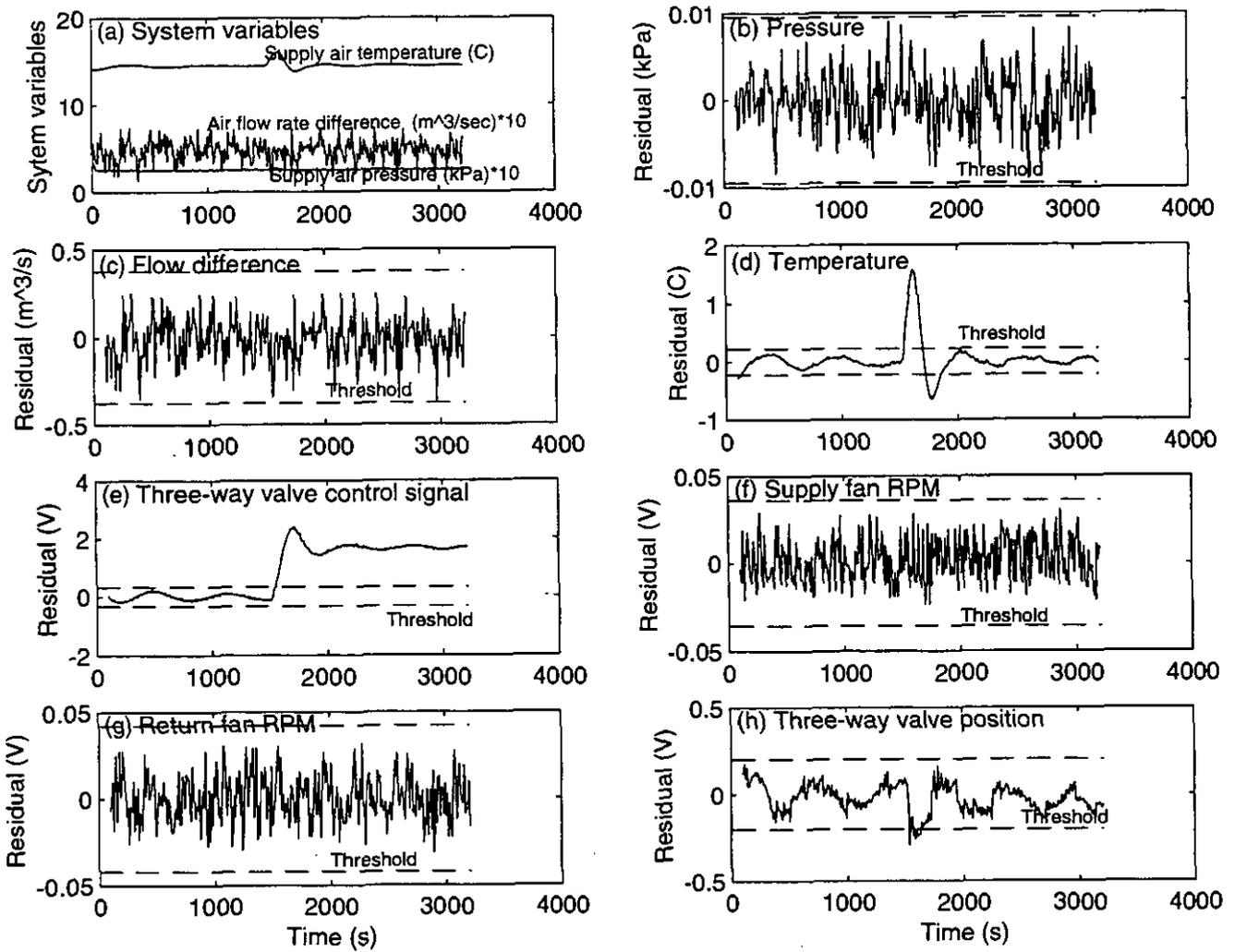
Figure 6 Residuals for fault 3.
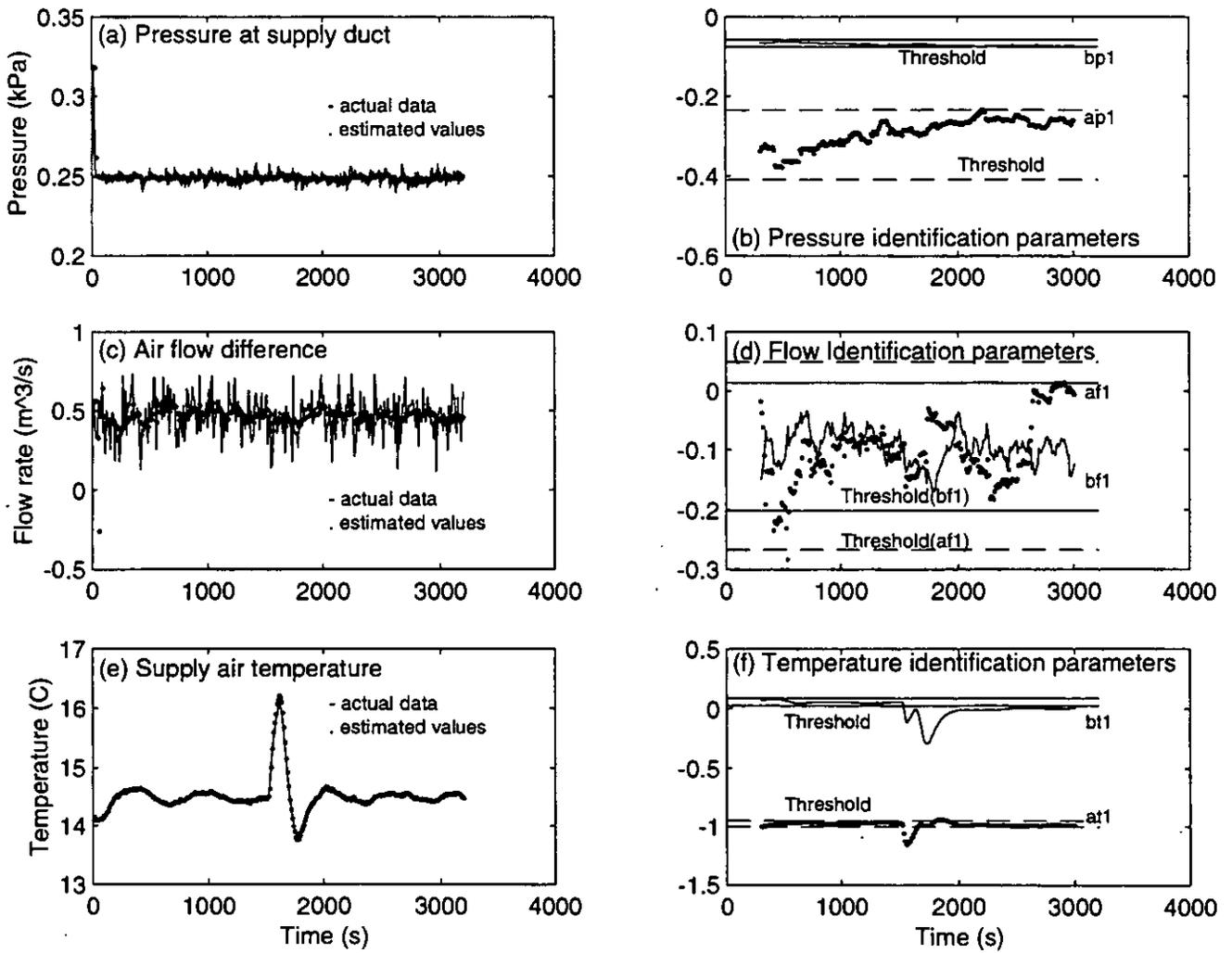
Figure 7 System identification result and identification parameters for fault 3.

FDD tests on a Coil subsystem using

Parameter Estimation in the Physical Model Approach

L. G. Olsson, P. Isakson and P. Blomberg

Div. of Build. Services Engineering, KTH

S-10044 Stockholm

Abstract

The potentials of a FDD method based on parameter estimation and physical models is tested on a coil subsystem in laboratory environment as well as in a simulated coil version. Basically the faultfree case is analysed because rather fundamental features in the parameter search procedure are of interest and these should not change if faults occur or not. However, some fault situations are also tested to study the sensitivity of the parameters in detecting and diagnosing applications. The statistical features of parameters are of particular interest considering
- variability of the signals
- impact of filter
- sampling rate and period

The coil system can be run under different options from externally operated fast moving valve to a rather slow valve motion with feedback loop for air temperature control (simulations).

The parameters in the model act as pre-defined fault indicators and features studied are, for instance,
- coil temperature effectiveness, flow rates and valve leakage
- effectiveness of air control loop (deviations from set-point and time delay)

The following general conclusions are drawn from these tests:
- large variability of the signals is necessary to keep statistical errors small and thereby also intercorrelation effects
- filter with time constant about equal to the thermal one for the coil should be used to reduce transient modelling and measurement errors
- to resolve dynamic features the sampling interval must be maximum 1/4 of the shortest relevant time constant of the system. The sampling period must be long enough to ensure sufficient signal variability

With these conditions fulfilled the present method may serve as a FDD tool, particularly suited for analysing degrading component functions, as indicated by the tests performed. The lower limit of faults to be resolved seems to be about 10 %, compared to correct behaviour. If faults occur that are not targeted by the parameters special care must be taken in the analysis by use of a priori learnt identification of parameter changes and such faults.

1. Introduction

In the present report a fault detection and diagnosis (FDD) approach based on parameter estimation and physical models for the component functions is analysed in more detail regarding fundamental features like

- accuracy of parameters
- data acquisition and use of filter
- sensitivity to changes in system
- applicability (On-line FDD to function control).

The capability to detect and diagnose faults rests critically on such features, most of them should be rather independent whether faults occur or not.

For this purpose a subsystem consisting of a coil, operated in the dry regime, and various types of control valve functions is analysed when it is running in its faultfree state as well as in some fault situations. This system has been implemented in two ways to generate test data:

- in laboratory with externally operated valves to establish stepwise changes of supply water temperature and flow rate
- by simulating the same system with a control valve and a feedback loop for the air side. Slightly different models compared to those used in the FDD routine are used.

In both cases rather short time sequences of signals but high sampling rates were exploited because the coil is quite small. A central point in this study is to check the overall uncertainties of various parameters and means to reduce these, particularly important for simultaneously occurring faults.

In section 2 the FDD procedure and application examples are briefly described. The test cases run are accounted for in sections 3, using measured data and 4, exploiting simulations. Both correct and faulty behaviour of the systems are tested. Aspects on data acquisition is discussed in section 5 and the potentials of the present procedure are commented upon in section 6.

2. Method and Application cases

The parameter estimation procedure makes use of measured data sets for various signals, relevant for the component or subsystem to be analysed, a system description based on physical models and targeted features that are essential for the function. In this section a brief account of the applied method and the data generation regarding a coil subsystem as well as its modelling for FDD and simulation is given.

2.1 Comments on FDD approach

An overview of the procedure is presented in the Source Book [1] section 4.3.1. The method is thus a computer program with administration and calculation routines. Their level of development and how they are applied in the present studies are further described in Appendix A of this report. Here a few points are further emphasised:

- the parameters may be of detecting or diagnosing types, to what extent is given by the models. However, the parameters must be defined so that they indicate a malfunction when changing from their reference values, the correct behaviour,

- the parameters may be seen as mean values with standard deviations (STD:s) and intercorrelations between pairs of parameters. In determining the accuracy all these effects have to be considered (simplified to quadratic addition, cf. section 5.2, [1]),

- a fault may also be identified by combining parameters within the same set of parameters or using data from different subsystems or components. Such procedures may be quite modelling dependent, needing further knowledge for interpretation.

2.2 Subsystems tested

Schematic pictures of the tested subsystems are shown in figures 2.1 and 2.2, indicating the components included and the signals used. Both the laboratory and simulated systems use similar coils but the latter one is controlled by a mixing valve with feedback loop and the former one with flipping valves that switch between two different supply water temperatures. The coil is quite small and consists of 24 tubes arranged in 3 parallel circuits and 3 rows. The air flow cross section is 0.09 m^2 with wavy fins. The nominal power is 10 kW.

Laboratory system in open loop arrangement:
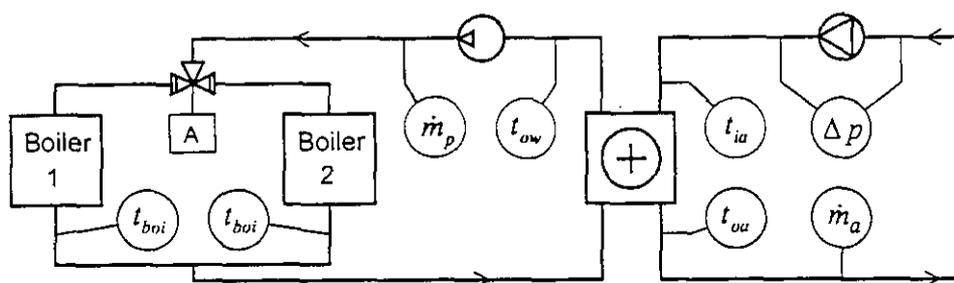
Fig. 2.1 Schematic layout of the experimental set-up with signals recorded. Denotations as ordered: t and $\dot{m}$ = temperature and flow rate & w and a = water and air & i and o = inlet and outlet & A, boi and set = valve stem position, boiler and set-point.

Input and output temperatures are measured as well as the flow rates. The data are stored in a PC which also controls valve switching. The measurements may be carried out over a rather short period (150 - 550 s) and sampling interval of 1 s, due to the short thermal time constant of the coil, about 5 s according to measurements. The types of signals that can be applied are:

- step up and step down of water inlet temperature or sequences of such steps while air inlet temperature is constant,

- step up and step down of the water flow rate while the air flow rate is kept constant during each measurement session.

Simulated system in closed loop operation:
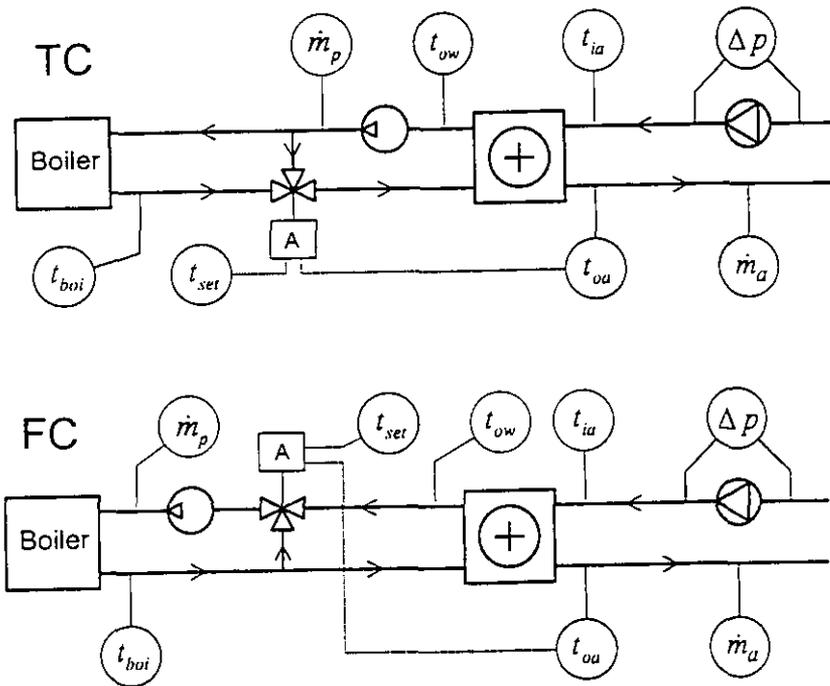
Fig. 2.2 Schematic layout of the subsystem with signals recorded. Denotations as in Fig. 2.1 but with TC and FC=temperature and flow control, respectively.

The recording of data was made on the same time scales as the experiments but the valve motion was set rather slow, 15 s from closed to open or vice versa, and a PI controller with the following rather arbitrarily chosen features:

- band width   1 C
- integral time constant   25 s
- valve characteristics   linear

The models used are almost similar to those applied for the FDD routines and will briefly be discussed in the following section. The simulations include both thermal and hydraulic modelling for the valve, connecting pipes and the coil, while the boiler or chiller is assumed to provide a pre-determined level of temperature. The pressure drops in the circuits are estimated by using valve authorities combined with a given pump curve. Air flow rate and supply temperature are ascribed values according to given time dependent functions.

2.3 Models and Parameterisation

2.3.1 FDD model

The subsystem model consists of

- the coil (at present running in the dry regime)

440

- connecting pipes between positions of temperature sensors and coil
- the control valve, given by K_v-numbers and flow characteristics

Apart from the by-pass circuit the pipes are hydraulically included in the coil piping but a minor temperature loss to the surrounding may be accounted for.

The measured and calculated output temperatures for air and water are used in the minimisation procedure. The model data are based on the steady state approximation, modified to include thermal capacitance and transport delays in the coil and pipes. The parameters indicating various faults or degrading function are defined as follows:

- temperature effectiveness of the coil (p1)

$$E_{act} = p1 \cdot E_m(x, DD) \tag{2.1}$$

where, x denotes input data and DD design data for the coil. Subscript act and m refer to actual and modelled values, respectively. IEA An 17 heating coil model has been utilised [2], however, with dynamics modified and an option to use polynomial fits for E.

- pump flow rate (p2)

$$\dot{m}_{act} = p2 \cdot \dot{m}_{w,x} \tag{2.2}$$

where, x refers to experimental data, directly measured or calculated from information about the pump (speed, energy consumption, pressure drop and/or pump characteristics). Often this flow rate is set to constant. Sometimes it is practical to omit p2 (fixed to 1).

- function of three port valve (p3)

$$\dot{m}_{by} = \dot{m}_{act} \cdot F(A \cdot p3, type) \tag{2.3}$$

where, F is a function that can be calculated, given the valve stem position A, valve characters type (lin, log or a combination thereof) and the size (K_v-numbers). Parameter p3 is a multiplier sensing the valve behaviour, i.e. mismatch between flow rate (FC) or supply water temperature (TC) and A. If p1 and p3 are free valve leakage can only be observed if A is small (nearly closed states). To suppress the effect of correlations when p2 is also free it might be necessary to ascribe *p3* a narrow validity range (max. 0.9-1.1). The STD might then give information about leakage, at least if it is large.

- controller function for air outlet temperature (p4 and p5)

$$t_{oa,act} = t_{set} \cdot p4 + G(\Delta t_{set}, \tau(x), p5, bd) \tag{2.4}$$

where, G is a function (exponential continued with a cosine) depending on steps in the set-point, Δt_{set} (if none put equal to bd/2 with bd denoting control band width) and the equivalent time constant, $\tau(x)$, for the coil.

In the present FDD approach the previous expressions are thus used to model the output temperatures for air and water in temperature or flow control mode operation of the coil. In the latter case the mixing temperature from the coil and by-pass circuits may also be used as option. The signals are converted to dynamic ones by following corrections:

- the transport delays on the water side are reproduced by means of flow dependent dead times and smoothed transferred signals. Accumulating water tanks are placed at the inlet and outlet of the coil and the volumes correspond to the pipe distance from the temperature sensors to coil midpoint. When these volumes are filled, after a number of time steps (> 1 but below a pre-set limit), the mean values of the temperatures are shifted in to and out from the coil.

- the heat balance is then evaluated in the steady state approximation and including the dynamics due to thermal capacitance of the coil (using one node). The outlet temperature for the air is then obtained while the one for water is stored as described above.

This information is then passed on to the parameter estimation program which performs the main calculations [3] . Briefly, the residuals between measured and calculated responses are evaluated, transformed to frequency domain and damped with a Gaussian filter function before being minimised utilising the Levenberg-Marquardt procedure (least square fitting), see [4].

Finally, if a reduced number of parameters are used it might be necessary to redefine some parameters in order to associate them with specific faults. As an example, carry out the fittings with p1 and p3 free while p2 is kept constant. Then, for

- FC using the mixed water temperature as response simple manipulations with coil and valve equations give (denotations according to list of symbols):

$$t_{mix} = t_{boi} - g(p3) \cdot crr \cdot p1 \cdot p3 \qquad (2.5)$$

$$t_{oa} = t_{ia} + g(p3) \cdot p1$$

To compare the indicators for temperature effectiveness obtained in the two fittings an equivalent parameter for water side results may be defined:: 'p1'= p1*p3.

- TC using the coil water outlet temperature as response gives similarly (due to the water feedback):

$$t_{ow} = t_{boi} - h(p3) \cdot crr \cdot p1 / p3 \qquad (2.6)$$

and the equivalent parameter 'p1'=p1/p3.

This option in using reduced number of parameters will also be tested below in order to simplify the interpretation.

2.3.2 Comments on the simulation model

The system in figure 2.2 is simulated in order to generate data for FDD analyses. To make the tests more easy to interpret the modelling is rather similar to the FDD model, but

- the pipes are described with specific models considering heat losses, thermal capacitance and transport delays for water flow

- the control valve with regulator and pump are also simulated by simple models, e.g., the pressure drop over the valve is considered by assuming a constant pump pressure but modified with the valve authorities.

Furthermore,

442

- all parameters are set to 1.0 when indicating the correct behaviour, while faults may be introduced by ascribing them other values
- coil model is based on measured values of the temperature effectiveness $E_m(..)$ in all cases.

3. Tests on Experimental Data with Large Variability

The laboratory set-up described in section 2.2 was used and these experiments were made with various step changes of the input signals, which should represent cases with maximal variability. It must be emphasised that such operation conditions can not be obtained in real plants but the analyses are nevertheless of interest from principle point of view, because the results serve as an upper limit of the capability of the method. The studies are restricted to the coil unit in heating mode.

3.1 Measurements and FDD model

Three sets of step input signals (1 in temperature and a sequence for the water flow rate plus a rapid change in air temperature) were combined to data sets of 512 sampling events (cf. Fig. B.1 in Appendix B). These data sets were used in all test cases and must be considered as a very favourable arrangement to obtain good variability of the signals. The duration of each step was made long enough to establish thermal equilibrium. Data were recorded each second for the inlet and outlet temperatures as well as the flow rates.

As no control devices are involved the FDD model was restricted to parameters checking the coil function, p1 and p2, as defined in section 2.3. There are three objectives with the present studies:
- impact of filter
- parameter correlation
- test impact of some faults

In all case the analyses are carried through using a starting period of 50 s, the skip time, where data are omitted in the fittings to ensure that differences that may occur in the initialisation conditions are damped out. On a 90 MHz PC each parameter estimation took about 30 s.

3.2 Results

Correct behaviour:

The resulting parameter values are given in Table 3.1 for different filter time constants, including correlations and the objective function. Denotations used:

px   = estimated value for parameter with number x
sx   = standard deviation (STD) for parameter with number x
a   = air side response
w   = water side response
R   = type of response (a or w)
CC/./= correlation coefficients for a parameter pair
/123 = STD in decimal digits
ObjF = objective function with residuals based on fluid outlet temperatures
τ_f = filter time constant

443

τ_f [s]	R [-]	p1/s1 [-]	p2/s2 [-]	CC/12/ [-]	ObjF [C]
1	a	0.934/002	1.128/001	-0.94	0.18
	w	1.078/007	1.188/011	0.86	0.71
5	a	0.934/003	1.129/011	-0.94	0.14
	w	1.020/018	1.080/033	0.95	0.53
10	a	0.934/003	1.130/013	-0.94	0.11
	w	1.019/024	1.078/043	0.96	0.43
25	a	0.934/003	1.130/012	-0.95	0.06
	w	1.030/038	1.097/051	0.87	0.25
50	a	0.934/003	1.128/012	-0.96	0.03
	w	1.051/032	1.138/059	0.98	0.10

As can be seen:

- the parameters are very constant for different filter time constants but differ for air and water side fittings as well as from 1. This is probably due to modelling defects.
- the insensitivity of the applied filter effect, apart from the STD:s for water side that increase slowly, may be explained by the fact that time periods after a step change are long compared to the thermal time constant. However, filter time constants about equal to the thermal one seem reasonable, i.e. here about 5 s.
- the STD is small in all cases as might be expected because of the variability of the data
- the correlation between p1 and p2 is strong but as long as STD:s are small this will not prohibit fault indication.

Faulty behaviour:

It might also be interesting to introduce some faults, which can easily be done for the flow rates by modifying the signals to the FDD model. Thus the following faults are introduced:

F2w = water flow rate is 15 % lower than the measured one (*1.15 on the signal to the model)
F2a = air flow rate is 10 % higher than the measured one (*0.90 on the signal to the model)
and combination of these two ones. Additional denotations are
F0 = faultfree case
sc = correlation error in parameters
N = parameters normalised to their reference values

Moreover, total errors in treated parameters are approximated by adding all contributions quadratically. The results are collected in the following table.

Table 3.2 Two free parameters and different flow rate faults

case	R	p1/s1	p2/s2	CC/12/	ObjF	sc
F0	a	0.934/003	1.129/011	-0.94	0.14	
	w	1.020/018	1.080/033	0.95	0.53	
F2a	a	0.908/003	1.049/011	-0.93	0.15	
	w	1.022/017	0.994/028	0.96	0.45	
N	a	0.972/004	0.929/015	-0.93	-	0.008
	w	1.002/025	0.920/043	0.96	-	0.033
F2w	a	0.935/002	0.979/008	-0.92	0.14	
	w	1.018/028	0.934/031	0.95	0.56	
N	a	1.001/004	0.867/014	-0.93	-	0.007
	w	0.998/033	0.865/046	-0.96	-	0.057
F2a+F2w	a	0.907/003	0.912/009	-0.93	0.15	
	w	1.020/015	0.870/022	0.95	0.45	
N	a	0.971/004	0.808/014	-0.96	-	0.007
	w	1.000/023	0.806/040	0.96	-	0.030

By normalising parameters to their reference values it is possible to identify faults. However, in case of air flow it is affecting both parameters and may add to errors in water flow. This puts requirements on statistical quality and fault magnitudes in order to resolve such effects. According to the definitions in eqs. (2..1-2) air flows, p0, may be tested by the relations

$$p0 = p1w / p1a$$
$$e = (ea^2 + ew^2)^{1/2}$$
$$ex = (sx^2 + scx^2)^{1/2}$$

(3.1)

where,

- x = a for air and w for water side data
- e = overall parameter uncertainty

In the present case p0=1.031/039, i.e. not statistically safe. Increasing the fault to 20 % gives the values 1.137/068, a more reliable indication (2 times the error *e*). The change in p2 for the water side ensures that the heat balance is fulfilled while p2 changing for the air side is not clear. Modelling errors may be the reason why air flow faults affect both p1 and p2 in this case.

3.3 Summing up

The present analysis has indicated that

- although the fittings are made on well controlled laboratory measurements and validated models, reference parameter values differ from 1 with 5 - 10 %. This is probably a realistic figure that will increase for in situ measurements
- the statistical quality of all data is good but correlations are large implying that these are embedded in the model structures
- the filter function has a small impact on the parameter values, probably due to the spectral shape of the signals. However, a slight filtering effect seems reasonable to reduce measurement noise

- faults can be observed by changes in parameters but for faults that affect more than one parameter there is a lower limit in magnitude for these to be resolved, depending on the quality of the signals.

4. Tests on Simulated data with Moderate Variability

In this case signals that are more realistic for normal operation of real plants are utilised in identifying the parameters. Both correct and faulty operations were simulated and tested with the same FDD routine as used in the previous example.

4.1 Generation of test data

The simulations were carried through for a short time period of 300-600 s and a time step of 0.2 s. Data were recorded each second, corresponding to the shortest coil time constant being resolved with 4 to 5 steps. The tests were performed with both temperature and flow control of the coil (TC and FC) and data were generated with a coil design similar to the laboratory set-up but the air supply temperature and flow rate were changed harmonically to mimic moving dampers in mixing and VAV boxes. Thus,

driving signals:

air inlet temperature	t_{ia}	(20-25 C)
air flow rate through fan	m_a	(0.012-0.027 kg/s)
water flow rate through pump	m_w	(0.010 kg/s)
boiler temperature	t_{boi}	(72 C for TC)
		(62 C for FC)
set-point for air temperature	t_{set}	(40 C)

recorded signals (to FDD):

outlet air and water temperatures:	t_{oa} and t_{ow}
valve stem position (relative):	A

To make the tests more realistic a random noise was applied on some signals, i.e.

temperatures	1 %
flow rates	5 %

of their amplitudes.

An example of simulated signals is shown in Appendix B. Faulty operation is obtained by giving the parameters in the simulation model values different from 1 and/or modifying the signals to the FDD routine.

4.2 FDD implementation

The fittings were performed in three steps with filter time constant of 5 s:

- analysis of water side responses giving p1, p2 and p3 or p1 and p3
- analysis of air side responses giving p1, p2 and p3 or p1 and p3
- analysis of controller function giving p4 and p5

The skip time was set to 30 s in all cases and the calculations were carried out on a PC, as earlier. Each fitting was performed within a minute. The signals accounted for in section 4.1 are all needed in the fitting procedures as well as design data for the system. The FDD and simulation models differ in treatment of the dynamic behaviour, as indicated in section 2.3.2.

4.3 Results and discussion

4.3.1 Correct behaviour

An example of the fittings is shown in Appendix B and the resulting parameters are collected in Tables 4.1 and 4.2, together with STD:s, cross correlations and objective function. Denotations used are the same as defined above. However, when parameters have been processed using eqs. (2.5) or (2.6) these data are given in single quotes. It should also be emphasised that in no case it was possible to let all three parameters be free simultaneously. One must be restricted or fixed to make the minimisation procedure to run properly.

Table 4.1a Results for temperature control mode based on 600 sampling intervals/data set. p3 restricted to 0.95-1.05. Fitting to coil outlet temperatures

R	p1/s1	p2/s2	p3/s3	CC/12/13/23/	ObjF
a	1.023/010	0.969/012	1.050/001	/-.72/.45/-.15/	0.53
w	0.974/016	0.969/015	1.040/002	/0.73/.71/.61/	0.88

Table 4.1b Results for temperature control mode based on 600 sampling intervals/data set. p3 restricted to 0.95-1.05. Fitting to set-point for air temperature

R	p4/s4	p5/s5	CC/45/	ObjF
a	0.989/010	0.992/055	-0.01	2.20

Tests with two free parameter, p1 and p2, gave a very irregular result as will be further discussed in 4.3.2.

In Tables 4.2 the corresponding data for the FC case are displayed for the option of mixed outlet water temperature and using two free parameters while p2 was fixed. The result is acceptable but the correlation between p1 and p3 is high, particularly for the air side. This is not quite surprising and not dangerous as long as the corresponding STD:s are small.

Table 4.2a Results for flow control mode based on 300 sampling intervals/data set. p2 is fixed to 1. Fitting to coil outlet air and mixed water temperatures

R	p1/s1	p3/s3	CC/13/	ObjF
a	1.029/015	0.930/036	-0.95	0.14
w	'0.987/010'	0.991/015	-0.59	0.25

Table 4.2b Results for flow control mode based on 300 sampling intervals/data set. Fitting to set-point for air temperature

R	p4/s4	p5/s5	CC/45/	ObjF
a	1.010/023	1.244/177	0.16	3.49

Fittings with parameters p1 and p2 free but p3 restricted were also tested with rather poor results. The air side fittings gave parameters in good agreement with the previous ones but for the water side no minimum could be indicated.

4.3.2 Fault tests

Finally, a few tests with minor system faults were carried out in order to analyse the sensitivity of the parameters. The TC and FC cases were tested by applying faults like

F1 = coil fouling corresponding to about 15 % increase in the overall air-to-water thermal resistance
F2w = water flow rate is 15 % lower than the measured one (*1.15 on the signal to the model)
F2a = air flow rate is 20 % higher than the measured one (*0.80 on the signal to the model)
F3 = valve leakage changed from 1 to 5 %,

and combination of those. Furthermore, F0 denotes faultfree case and the results are given in Tables 4.3 to 6, where the same denotations as defined above are used.

The TC case was analysed with two sets of parameters, viz.

- p1 and p2 free but p3 restricted and
- p1 and p3 free but p2 fixed.

Data sets of 600 samplings are used in all tests.

Table 4.3 Test of fault cases with TC. p3 restricted to 0.95-1.05.

Case	R	p1/s1	p2/s2	p3/s3	CC/12/	CC/13/	CC/23/	ObjF
F0	a	1.052/015	0.936/008	1.049/014	-.14	.78	.09	0.57
	w	1.014/077	1.023/104	1.044/049	-.88	.33	.-13	0.95
F1	a	0.954/024	0.920/021	1.048/012	-.81	-.54	.22	0.81
	w	0.864/032	0.948/018	1.050/040	.59	.94	.29	0.95
N	a	0.907/029	0.983/022					
	w	0.852/083	1.017/106					
F2w	a	0.984/035	0.893/041	1.050/051	-.48	-.51	-.47	0.57
	w	1.047/140	0.926/150	1.050/024	.99	.81	.73	0.95
N	a	0.935/038	0.954/042					
	w	1.032/156	0.905/183					
F1+F2w	a	0.877/068	0.928/117	1.050/026	-.98	-.59	-.47	0.82
	w	0.926/089	0.914/101	1.049/032	.95	.61	.35	0.95
N	a	0.834/092	0.991/117					
	w	0.913/115	0.893/143					
F3	a	0.955/031	1.103/053	1.045/065	-.96	-.64	.66	0.62
	w	0.945/080	0.963/072	1.045/058	.81	.70	.18	0.57

F1+F3 gave no significant difference from F1. As can be seen from the table

- faults are possible to indicate but with various degree of accuracy. It is particularly important to normalise with reference data
- the STD:s differ much from case to case and are sometimes quite large. The reason for this may be the TC operation, further discussed below
- the correlations change sign when faults occur which may indicate weaknesses in the model structure
- leaking valve is difficult to detect by increase of s3 unless the leakage is considerable, although a minor effect can be seen in s3.

It might also be of interest to study a 2 parameter fitting with same test data. It should be noted that in these cases the parameter p1 for water side fittings has been processed by means of eqs. (2.5) an (2.6), which has been indicated by giving the tabulated values in single quotes.

Table 4.4 Test of fault cases with TC. p2 is fixed to 1.

Case	R	p1/s1	p3/s3	CC/13/	ObjF	sc
F0	a	0.847/009	1.462/058	-0.87	0.39	
	w	'0.943/037'	1.082/048	0.96	0.94	
F1	a	0.809/015	1.377/101	-0.87	0.74	
	w	'0.842/035'	1.103/041	0.94	0.93	
N	a	0.955/017	0.942/117	-0.87	-	0.037
	w	0.893/051	1.019/063	0.95	-	0.053
F2w	a	0.779/008	1.537/059	-0.78	0.49	
	w	'1.079/018'	1.059/017	0.74	1.05	
N	a	0.920/012	1.051/082	-0.84	-	0.028
	w	1.144/042	0.979/051	0.85	-	0.035
F1+F2w	a	0.741/012	1.600/153	-0.79	0.83	
	w	'0.936/035'	1.061/043	0.89	1.07	
N	a	0.916/015	1.096/164	-0.83	-	0.041
	w	0.993/051	0.980/064	0.93	-	0.055

The results are very contradictory for this case, e.g.

- the reference values differ clearly from 1
- some faults may be indicated in the normalised parameters, others not
- statistical errors and correlations are large

The reason for this behaviour is probably that the feedback of coil outlet water makes the operation less distinct, implying that the parameters for correct as well as faulty behaviour will be difficult to identify, as already indicated. A solution to this problem would be to also measure the coil water inlet temperature and use it in the fittings as described in section 3. The valve function is then tested separately. In addition to that the measuring period could be increased.

The FC case was analysed with two free parameters as in 4.3.1 and tests with fault cases F2w, F3 and F2w+F3, using data sets of 300 samplings.

Table 4.5 Test of fault cases with FC. p2 is fixed to 1.

Case	R	p1/s1	p3/s3	CC/13/	ObjF	sc
F0	a	1.029/015	0.930/036	-0.99	0.14	
	w	'0.972/022'	0.991/015	-0.59	0.25	
F2w	a	1.036/015	0.812/031	-0.99	0.14	
	w	'0.967/020'	0.855/014	-0.64	0.25	
N	a	1.007/021	0.873/048	-0.99	-	0.031
	w	0.995/030	0.863/021	-0.62	-	0.026
F2a+F2w	a	1.027/048	0.606/070	-0.99	0.38	
	w	'0.983/023'	0.735/016	-0.85	0.20	
N	a	1.001/051	0.692/079	-0.99	-	0.063
	w	1.011/032	0.742/022	-0.72	-	0.023
F3	a	1.024/015	0.934/035	-0.99	0.15	
	w	'0.972/022'	0.994/025	-0.76	0.27	

F2w+F3 gave no significant difference from F2w. With this parameterisation both statistical errors and intercorrelation effects are reasonably small and the flow rate fault well indicated. It is clear that small valve leakage can not be observed unless the operation is such that the valve is in a nearly closed position. The same yields for TC cases. It is also worth noting that flow rate errors of about equal size but opposite direction have the same impact on p3 and need further information to be resolved.

4.4 Summing up

The present analyses were carried out on good models and data, apart from low variability of the signals, a situation rather opposite to the tests described in section 3. There are both similarities and differences in the results:

- correlations are still large supporting the conclusion that these features are embedded in the model structure
- reference data are close to 1, due to the modelling
- statistical uncertainties are in general larger, probably caused by limited richness of the signals, an effect that can be reduced by increasing variability and/or sampling number
- the procedure is capable of indicating pre-defined faults by change of parameter values even in more complex systems. However, parameterisation is dependent of detailed system layout and should be tested from case to case (by simulation)
- based on simulations faults smaller than 10 % (in terms of parameter values) can not be resolved, particularly if more than one fault occur simultaneously
- parameter changes for various types of faults should be checked with simulations in spite of their pre-definitions if the component function is complex, as is a coil with control valve, for instance.

5. Data acquisition

In the previous tests, sampling rate and filter time constants were chosen empirically, i.e. transients should be resolved with at least 4 steps for dynamic models. Likewise, the filter time constant should not exceed the thermal one but suppress outliers in the measured signals or modelling defects. Such conditions imply, however, rather large sampling rates and to ensure high variability also long sampling periods, requiring large data storage capacity. In the present section the consequences of reduced resolution of the measurements is studied using simulated data (cf. 4.2). The tests are restricted to FC case and mixed water side responses with p1 and p3 free. As before, parameter p1 is processed and given in single quotes in the tables.

5.1 Increased sampling interval

The time step used in the previous studies was set to 1 s, considerably shorter than the thermal time constant. Time steps of 5 and 10 s for the sampling routine were also tried. The filter time constant was put equal to the time step length or, tentatively, to 50 s. Moreover, the number of sampling intervals was kept constant, 300, implying that the measuring time was increased proportionally to the length of the time step. The results are given in Table 5.1.

Table 5.1 Results for flow control mode using different length of sampling interval but the same number of samplings (300).

step [s]	p1/s1 [-]	p3/s3 [-]	CC/13/ [-]	Objf [-]	τ_f [s]
1	'0.977/009'	0.992/007	-0.44	0.37	1
	'0.881/098'	0.817/082	-0.97	0.05	50
5	'1.088/079'	1.283/079	-0.76	0.48	5
	'0.760/206'	0.664/143	-0.99	0.09	50
10	'1.160/017'	1.336/015	-0.25	0.57	10
	'1.030/021'	1.099/019	-0.85	0.13	50

As can be seen the parameter values behave rather irregular, an effect increasing with length of sampling interval and filter time constant. For dynamic analyses these are very restricted.

5.2 Increased filter time constant

The same system as in the previous case was used but with sampling interval restored to 1 s and filter time constant considerably increased, up 10 times the thermal one. Overall sampling time was set to 300 s in the tests.

Table 5.2 Results for flow control mode using different filter time constants (τ_f). Thermal coil time constant is about 5 s.

τ_f [s]	p1/s1 [-]	p3/s3 [-]	CC/13/ [-]	ObjF [C]
1	'0.977/009'	0.992/007	-0.44	0.37
5	'0.978/018'	0.991/015	-0.59	0.25
10	'0.984/028'	0.992/024	-0.78	0.20
25	'0.971/035'	0.972/032	-0.70	0.11
50	'0.881/098'	0.817/082	-0.97	0.05

As can be seen the results are slowly deteriorating when the filter time constant is increased. τ_f should not be set larger than two times the thermal coil time constant.

5.3 Increased number of sampling intervals

The final quantity in the data acquisition procedure to study is the number of sampling intervals. The variability of the signals affects the possibility to resolve the parameters (reduced uncertainty) and increased length of the measuring period gives in general more information. τ_f is set to 5 s.

Table 5.3 Results for flow control mode using different number of sampling intervals (N).

N [-]	p1/s1 [-]	p3/s3 [-]	CC/13/ [-]	ObjF [C]
300	'0.978/018'	0.991/015	-0.59	0.25
600	'0.978/013'	0.992/011	-0.68	0.23
1200	'0.985/008'	1.006/007	-0.65	0.21

The parameter values are changing very little but in addition to the improved statistics there is a tendency for increased correlations when the samplings are increased. However, the statistical features of the parameters (STD + intercorrelations) are improved. This result is equivalent to, repeating a measurement series M times, the STD will reduce with $\sqrt{M}$, assuming the same boundary conditions.

5.4 Summing up

The empirical rules to determine length of sampling interval for dynamic system identification apply, i.e. transients should be resolved with at least 4 steps. Also the filter time constant should be restricted, suppress measurement noise but keep the essential features of the models. The number of samplings has a statistical impact on the results in that it reduces the STD:s of the parameters. The obvious drawback with long sampling periods is increasing measurement and computation times. In addition to that the possibility to resolve fluctuating faults is reduced.

6. Conclusions

Tests with the present method have been carried out on a coil in various applications, open or closed loop operation and in the latter case in both temperature and flow control mode. Both correct and faulty cases were studied by use of data from a laboratory set-up and a simulated system.

As regard to the fitting procedure it can be concluded that
- the estimation is quite sensitive implying that the modelling must be very realistic and the signal information rich, otherwise the number of free parameters must be reduced
- the parameter accuracy is set by the STD:s and intercorrelation effects. Correlations between parameters occur generally in the present test cases, indicating that they are embedded in the model structure by definition. The impact of the correlations can be kept small by high statistical quality of the measured signals
- the data acquisition should be carried out with large sampling rate and a slight filter effect when dynamic models are used, i.e. the standard (literature) values used in the present tests seem adequate.

The present studies, however, indicate that the adopted parameter estimation method has potential in FDD applications:
- minor faults were introduced in the measured signals (faulty sensors) and in simulations, which were mostly possible to identify by changes in parameters. However, sometimes the faults affect more than one parameter, particularly when these faults are not specifically targeted, making it urgent to carry out simulations with many different types of faults to map possible parameter combinations
- certain parameter sets did not produce useful results, probably due the model structure not giving a unique result within the uncertainty of the measured data (FC and TC coil). To cope with such cases the number of free parameters could be reduced, the variability of the signals increased by disturbing the system, introducing additional sensors to make the fittings more robust etc.

452

It seems as the present method is best suited for analyses of degrading component or system functions but other possibilities of exploiting it exist. At On-line application the parameter estimation procedure is repeatedly applied and a two-step procedure is proposed:

- detection analyses: few global parameters when low variability of signals occur, to check overall system function
- diagnostic analyses: when a fault has been observed special measurements are conducted with large variability of the signal to allow more free parameters

By such an arrangement the drawback with the delay in responses (computing time) can be reduced.

Finally, to make a more complete diagnosis of a HVAC system more components must be included in the analysis, like fan, pump, mixing and VAV boxes, and combined with the information from the coil subsystem. In this way the interpretation will also be more substantial.

List of Symbols

1. Denotations (SI units)

A = valve stem position
a = air side
bd = band width
CC = correlation coefficient
crr = thermal capacity ratio (air/water)
DD = design data (or characteristic curves)
Δ t = temperature difference
ex = error of type x
E(.) = temperature effectiveness
F(.) = function or fault
G(.) = function
g(.) = function
h(.) = function
i,j = indices

$\dot{m}$ = flow rate
N = number of samplings, normalised data
ObjF = objective function
pj = parameter j
R = type of response (air, water, controller)
sc = intercorrelation fault in p
sj = standard deviation (STD) for parameter pj
t = temperature
τ = time constant
type = valve characteristics
x = input signal

2. Subscripts

act = actual (observed)-
a = air side
boi = boiler
by = by-pass
c = control
f = filter
i = inlet
m = model based
mix = mixed coil and by-pass flows
o = outlet
p = pump
set = set-point value
w = water side
x = experimental

References:

1. J. Hyvärinen, editor: IEA An 25 BOFD Source book, 1996

2. M. J. Holmes, The Simulation of Heating and Cooling Coils for Performance analysis, Int. Conf. on System Simulation in Buildings, Liege 1982.

3. W. Spirkl, Dynamic System Testing Program, InSitu Scientific Software, Klein & Partners, Baaderstr. 80, D 80469 München, 1994 and Dynamic SDHW Testing. J of Solar Energy Eng., Transactions of the ASME, 112:98-101, 1990

4) W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Numerical Recipes, Cambridge University Press, 1989.

Appendix

A   Implementation of the method

The aim is to provide a software that automatically performs an analysis of a subsystem. However, in the present tests the software is in a rather preliminary stage of development but still arranged to mimic a real application. Thus the procedure is divided in two parts:

- data acquisition which is done in separate measurements or simulations providing necessary data sets

- the analysis, carried through in a sequence of batch runs for each defined component or subsystem

Below this second part is further described.

A.1 Program structure

Prerequisites for the test case as well as in a more final version:

- the user must provide compiled modules for the actual components

- information about design features, parameters and how to address in- and output spectra (text files of specified formats at present)

- long time storing of the results for each fitting procedure, at present performed manually with text editors,

arranged in an order as given by the following block diagrams (some routines not fully implemented).

```
1       menu
                -pre-processor
                -classifier/evaluation
_____

2       pre-processor:
                provide files with models,
                design data, free parameters,
                sets of measured signals and
                storage for results
_____

3       start parameter estimation:
                -call fitting program for the specified component
                -store results (parameters and residuals)
_____

4       more components to analyse ?    yes --> 2
_____

5       menu:
_____
```
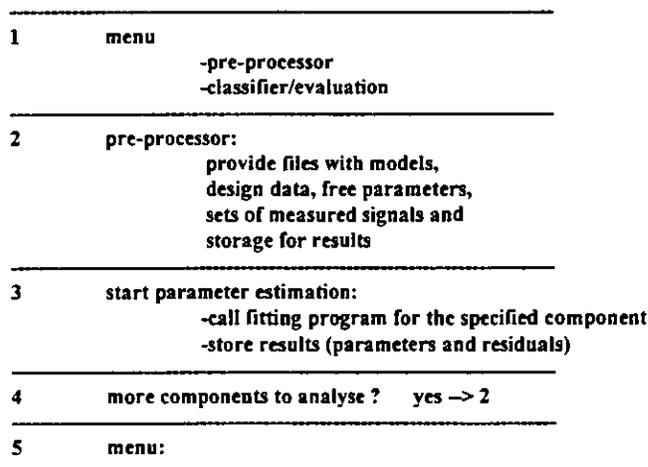
Fig. A/1. Tabular flow chart for pre-processing part of the FDD routine.

1	menu:
	....
	classifier/evaluation

2	classifier:
	read results from pre-processor files
	treat (refine)/normalise parameters
	store new results

3	more component to analyse? yes ——> 2

4	display results in tables or plots (for time evolution)
	with decision making aspects:
	-parameter values with statistics
	-diagnosis and proposed actions
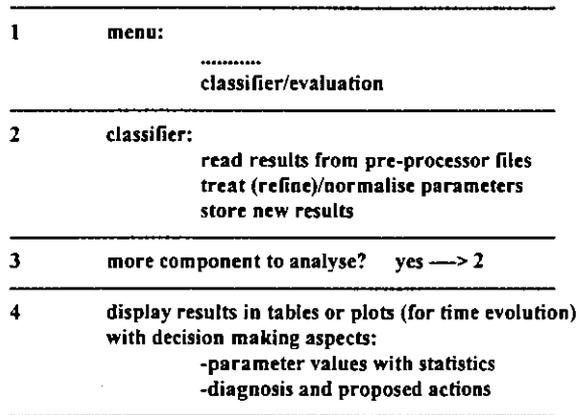
Fig. A/2. Tabular flow chart of the parameter analysis. All these routines are performed manually at present.

A.2 Software available.

The implementation of the programs according to the flow charts is thus partly completed:

- the supervisory program handler (menu) remains to be developed
- the main administration for supplying information (all types of data files and subroutines for models) to the parameter estimation program exists but minor modifications may be needed. This program is written in FORTRAN 77 and DOS command files
- the program package for parameter estimation [3] is commercially available (compiled), written in PASCAL and automatically included in the calculation routine
- modules for various models are to be included in the course of development of the FDD procedure
- the parameter analysis routine remains to be developed (cf. Fig A/2) and should thus include parameter refinement routines, error calculations, plotting facilities, comments on the results etc.

B Examples of measured signals and residuals

Example of signals used in the parameter estimation procedures based on experiments and simulations. Also the corresponded residuals for the air side are given which are rather representative for all tests. Time scale in seconds.
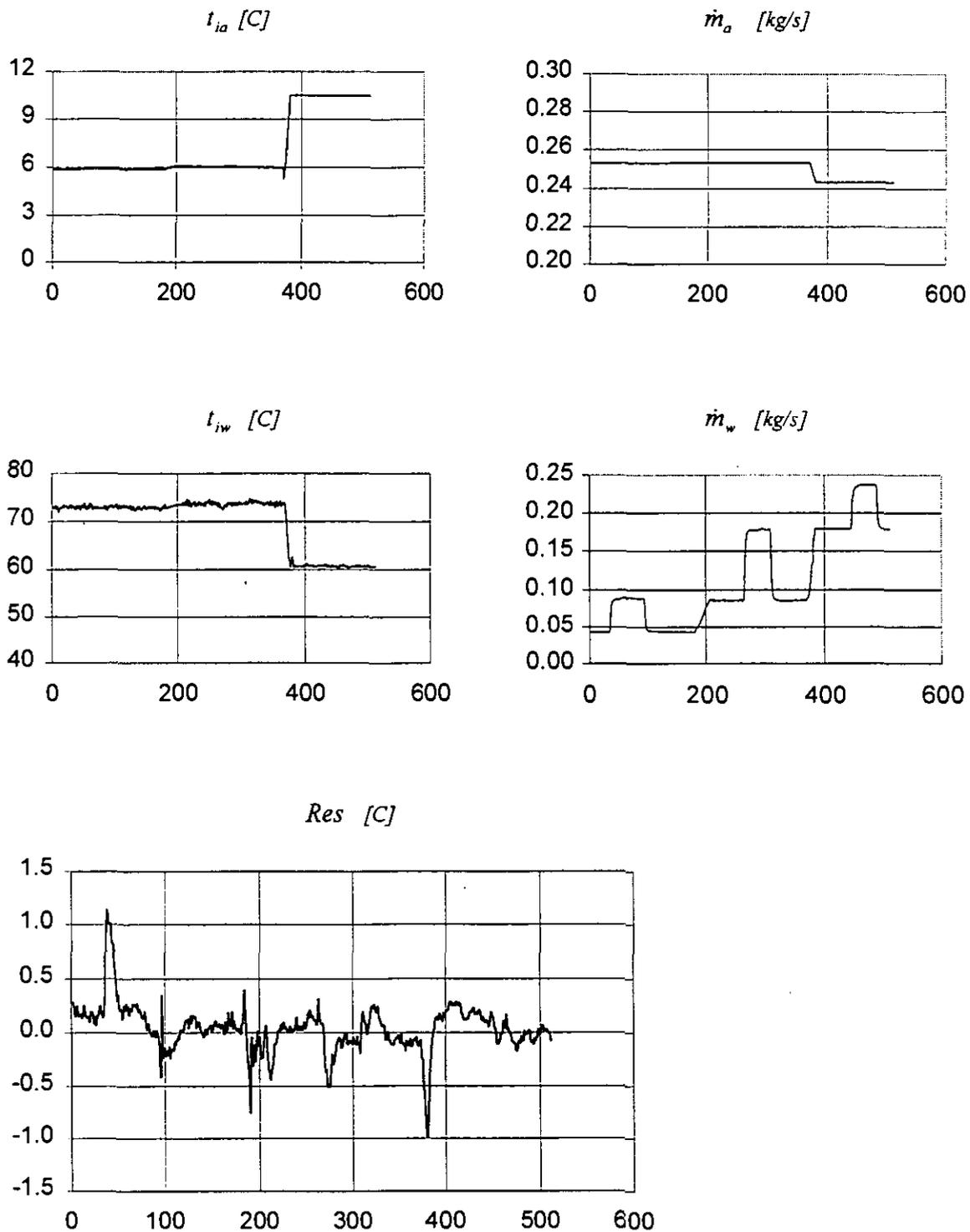
Fig. B.1 Example of results from tests with experimental data. From top: air and water inlet temperatures, air and water flow rates; at bottom the residuals for air side responses.
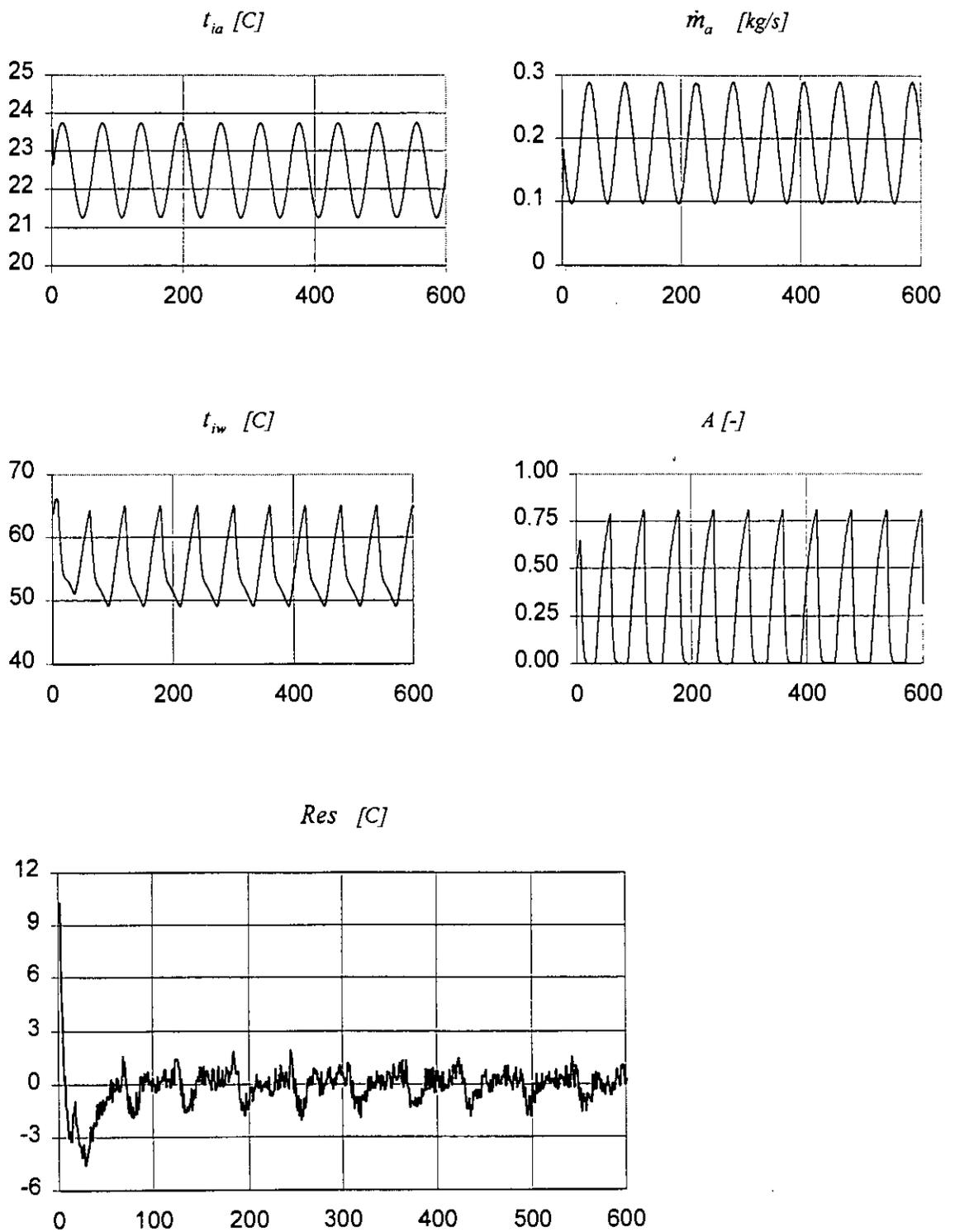
Fig. B.2 Example of input signals used in all tests with temperature control. From top to bottom: air and water inlet temperatures, air flow rate and valve stem position. Bottom figure gives the residuals for air side responses.

Monitoring energy consumption in heating systems

Peter Sprecher, Landis & Gyr Building Control, Switzerland

Summary

A 1st order model of a heating system is setup, including the most important heat gains and losses. Determination of its parameters is discussed. In case that no measurements of fault free operation are available the parameters of a simplified model can be estimated based on prior knowledge. Otherwise a more sophisticated model can be identified using the measurements available. Some possibilities of automatic model validation are compiled. Model based testing of new energy consumption measurements is then discussed. Emphasis is put on easy threshold determination. An economic threshold and a statistical one are distinguished. Finally a suggestion to reduce false alarms is added.

1 Introduction

The most common method for monitoring the energy consumption in heating systems is the energy signature method. The energy signature shows the energy consumption within a certain time period in dependence of the outdoor temperature. More precisely the energy consumption depends also on the room temperature setpoint schedule, the solar radiation, the wind speed, and internal heat gains. As a building has a thermal inertia the energy consumption depends also on the history of the mentioned variables. Therefore the energy signature method needs several extensions and guidelines for selecting appropriate measurement periods. This paper attempts to compile such extensions and guidelines. It further addresses the problem of *automtic* testing of identified models and of new measurements.

2 Basic approach

Fig. 1 illustrates the basic approach for testing the energy consumption.
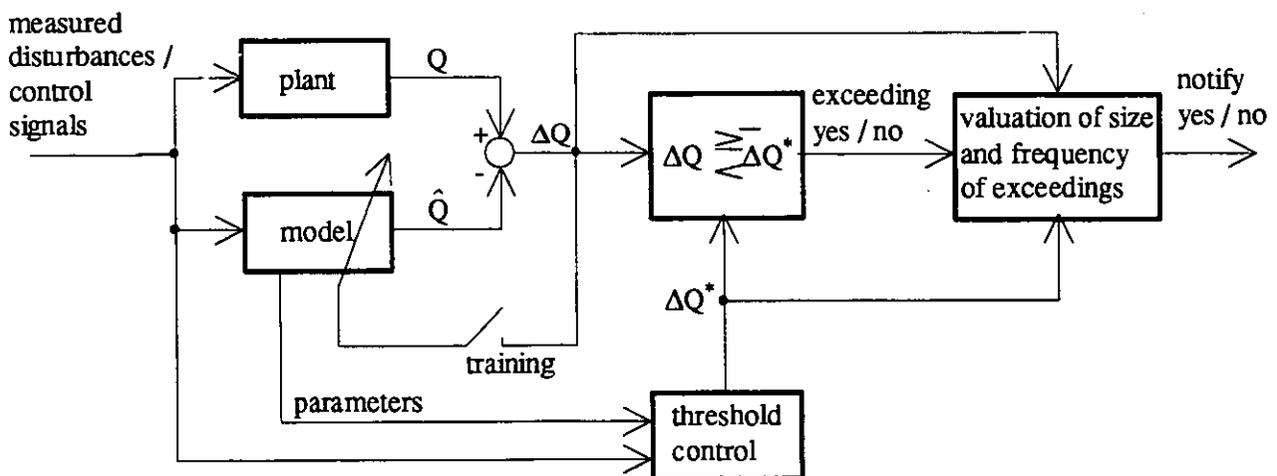
Fig. 1 System structure for testing the energy consumption Q

A model is used to determine the expected energy consumption $\hat{Q}$. The difference to the actual consumption is compared to a threshold which depends on the current inputs and the knowledge in the model. To decide on alarming the size and the frequency of the threshold exceedings are considered.

For a given installation the following 2 cases can be distinguished:

1. No energy consumption measurements have been collected and evaluated so far. New measurements have to be checked based on *prior knowledge*.

2. A representative set of energy consumption measurements has been collected and verified to represent fault free operation. New measurements can be checked based on these *reference measurements*.

In the first case the expected energy consumption for a given outdoor temperature ist determined using a *very simple model* based on prior knowledge. In the second case the expected consumption is determined using a *more sophisticated model* built from the reference measurements.

3 Models

Heating systems without mechanical ventilation are considered only.

3.1 Basic model with main heat contributions

Heat balance for the building:

The following first order model is considered to be sufficiently accurate in most cases:

$$C_B \dot{T}_R = \dot{Q}_H - G_B(T_R - T_O) + \dot{Q}_{GU} + \dot{Q}_{GoccN}\frac{\ell_{occ}}{\ell_{occN}} + \dot{Q}_{GM} + \dot{Q}_{Smax}\frac{I_S}{I_{Smax}} - \dot{Q}_{WN}\frac{T_R - T_O}{T_{RN} - T_{OD}}\frac{max(v_W - v_{W0},0)}{v_{WN} - v_{W0}}, \qquad (1)$$

where:

T_R, T_{RN} = room temperature and its nominal value (20°C)

T_O, T_{OD} = outdoor temperature and its design value

I_S, I_{Smax} = solar radiation and its maximum

v_W, v_{WN} = wind speed and its nominal value

v_{W0} = min. wind speed for significant contribution of wind (contributions of smaller wind speeds taken into account in G_B)

$\dot{Q}_H$ = heating power

$\dot{Q}_{GU}$ = unknown free heat gain (e.g. people in case of no automatic admission checking)

ℓ_{occ} = occupancy level [], expressing the degree to which the building is occupied (e.g. number of occupied rooms w.r.t. total number of rooms)

ℓ_{occN} = nominal value of ℓ_{occ}, to be chosen freely (typically $\ell_{occN} = 1$)

$\dot{Q}_{GoccN}$ = nominal value of occupancy dependent internal heat gain, appearing if $\ell_{occ} = \ell_{occN}$ (e.g. people, lighting)

$\dot{Q}_{GM}$ = measured internal heat gain (e.g. electricity, people in case of automatic admission checking)

$\dot{Q}_{Smax}$ = max. solar heat gain

$\dot{Q}_{WN}$ = nominal heat loss due to wind (= heat loss for $v_W = v_{WN}$, $T_R = T_{RN}$ and $T_O = T_{OD}$)

G_B = heat conductance of building enveloppe (for heat losses by transmission *and* ventilation)

C_B = heat capacity of building

If a large building with different room temperatures is considered, the average of these temperatures is taken as T_R.

The energy consumption in a time period d (e.g. 24h) is $\overline{Q}_H d$, where '⁻' denotes the average taken over d. This energy consumption can be computed from (1) as

$$\overline{Q}_H d = G_B(\overline{T}_R - \overline{T}_O)d - \overline{Q}_{GU}d - \dot{Q}_{GoccN}\frac{\overline{\ell}_{occ}}{\ell_{occN}}d - \overline{Q}_{GM}d - \dot{Q}_{Smax}\frac{\overline{I}_S}{I_{Smax}}d + \dot{Q}_{WN}\,\overline{f}(v_w,T_R,T_O)d + C_B\Delta T_R \,, \qquad (2)$$

where:

$$- f(v_w,T_R,T_O) = \frac{T_R - T_O}{T_{RN} - T_{OD}}\frac{\max(v_W - v_{W0},0)}{v_{WN} - v_{W0}}$$

- ΔT_R = room temperature change over d, i.e. $T_R(t) - T_R(t-d)$

Heat balance for the heat generator:

For simplicity an oil boiler is considered. Based on [1], sect. 3.1.3, the energy consumption in a time period d becomes approximately

$$Q_d = \frac{\overline{Q}_W}{\eta_{BIN}}d(1-r_0) + r_0\dot{Q}_{Bu}d \,, \qquad (3)$$

where:

$\overline{Q}_W$ = average power supplied to water
$\dot{Q}_{Bu}$ = burner power
r_0 = stand-by losses [%]
η_{BIN} = nominal boiler efficiency

Total heat balance:

Combining (2) and (3), observing an eventual domestic hot water energy demand $\overline{Q}_{DW}$, and neglecting distribution losses the total heat balance becomes

$$Q_d = \frac{\overline{Q}_H + \overline{Q}_{DW}}{\eta_{BIN}}d(1-r_0) + r_0\dot{Q}_{Bu}d \,. \qquad (4)$$

$\overline{Q}_{DW}$ is assumed to be independent of the outdoor temperature and to depend on the occupancy level $\overline{\ell}_{occ}$ only.

To have a model form more suitable for parameter estimation (4) is rewritten as

$$\overline{r} = \frac{\Delta t_{on}}{d} = \frac{Q_d}{Q_{dmax}} = p_1\frac{\overline{T}_R - \overline{T}_O}{T_{RN} - T_{OD}} - p_2 - p_3\frac{\overline{\ell}_{occ}}{\ell_{occN}} - p_4\frac{\overline{Q}_{GM}}{\overline{Q}_{GMN}} - p_5\frac{\overline{I}_S}{I_{Smax}} + p_6\,\overline{f}(v_w,T_R,T_O) + p_7\frac{d_N}{d}\frac{\Delta T_R}{T_{RN} - T_{OD}} \,, \qquad (5a)$$

461

where:

$\bar{r}$ = average burner running ratio over d

Δt_{on} = burner running time within d

$\dot{Q}_{GMN}$ = nominal measurable internal gain (may be chosen freely)

$Q_{d\,max}$ = $\dot{Q}_{Bu}d$, max. energy consumption within d

d_N = nominal value of time period d (may be chosen freely, d may vary)

$$p_1 = \frac{G_B(T_{RN} - T_{OD})}{\dot{Q}_{BIN}}(1 - r_0) \tag{5b}$$

$$p_2 = \frac{\bar{\dot{Q}}_{GU}}{\dot{Q}_{BIN}}(1 - r_0) - r_0 \tag{5c}$$

$$p_3 = \frac{\dot{Q}_{GoccN} - \dot{Q}_{DWN}}{\dot{Q}_{BIN}}(1 - r_0) \tag{5d}$$

$$p_4 = \frac{\dot{Q}_{GMN}}{\dot{Q}_{BIN}}(1 - r_0) \tag{5e}$$

$$p_5 = \frac{\dot{Q}_{Smax}}{\dot{Q}_{BIN}}(1 - r_0) \tag{5f}$$

$$p_6 = \frac{\dot{Q}_{WN}}{\dot{Q}_{BIN}}(1 - r_0) \tag{5g}$$

$$p_7 = \frac{C_B(T_{RN} - T_{OD})}{\dot{Q}_{BIN}d_N}(1 - r_0) \tag{5h}$$

$\dot{Q}_{BIN}$ = $\eta_{BIN}\dot{Q}_{Bu}$, boiler nominal power

$\dot{Q}_{DWN}$ = nominal value of the domestic hot water energy demand $\dot{Q}_{DW}$

$p_{1..7}$ are the parameters to be estimated. Basically each of them measures the ratio of a certain power and the boiler nominal power.

3.2 Simplified models and appropriate measurement periods for specific cases

Depending on the available measurements and the selected measurement periods $[t_i, t_i + d_i]$, i=1,2,.., the full model (5a..h) can be simplified. However simplifications due to missing measurements lead to less accurate models. Sometimes the measurement periods can be selected in such a way that the model inaccuracy manifests ifself little only. The following cases are considered:

1. **No measured internal heat gain $\dot{Q}_{GM}$:**

   - Drop term with $\bar{\dot{Q}}_{GM}$ in (5a).

   - Internal heat gains taken into account in $\bar{\dot{Q}}_{GU}$ or $\dot{Q}_{GoccN}$ only.

2. Occupancy level $\bar{\ell}_{occ}$ unknown or constant:

- Drop term with $\bar{\ell}_{occ}$ in (5a).

- Unmeasured internal heat gains taken into account in $\bar{\dot{Q}}_{GU}$ only.

- Redefine p_2 as $p_2 = \dfrac{\bar{\dot{Q}}_{GU} - \dot{Q}_{DWN}}{\dot{Q}_{BIN}}(1 - r_0) - r_0$.

3. No measurement of solar radiation I_S:

- Drop term with $\bar{I}_S$ in (5a).

- Solar heat gain taken into account in $\bar{\dot{Q}}_{GU}$.

- Try to select measurement periods where the average solar heat gain is small w.r.t. the heating power $\bar{\dot{Q}}_H$ (e.g. exclude periods where the outdoor temperature is close to or above the heating limit). A statistical method to exclude periods with high solar gain is reported in sect. 6.3.

4. No measurement of wind speed v_W:

- Drop term with v_W in (5a).

- Signifcant losses due to wind speed taken into account in term $G_B(\bar{T}_R - \bar{T}_O)$.

- A statistical method to exclude periods with high losses due to wind is reported in sect. 6.3.

5. No measurement of room temperature T_R:
Replace in (5a) $\bar{T}_R$ by the average setpoint $\bar{T}_{RS}$.

6. Large time periods d, or periods with small room temperature change ΔT_R:
If d is selected that large, that the energy required for the room temperature change ΔT_R is small w.r.t. the heating energy $\bar{\dot{Q}}_H d$, the term with ΔT_R in (5a) may be dropped (e.g. d = 1 week). The same simplification is possible, if begin and end of the period are chosen such that the change ΔT_R is small (e.g.; $\Delta T_R \leq 1K$). This simplification leads to a static model.

7. Electric heating system:
In (5a..h) set $r_0 = 0$ and replace $\dot{Q}_{Bu}$ and $\dot{Q}_{BIN}$ by the maximum electric heating power $\dot{Q}_{el,max}$. Q_d is directly measured.

8. Domestic hot water supplied by separate boiler:
Drop $\dot{Q}_{DWN}$ in (5d).

In [11], sect. 4.6, an approach called h-m-method is introduced for installations with *large solar heat gain*. The model used takes the terms with p_1, p_2, and p_5 of the full model (5a). Dividing by $(\bar{T}_R - \bar{T}_O)$ leads to a model of the form

$$h = H_0 - A_e m,\tag{6a}$$

where

$$h = \frac{Q_d/d + p_2\dot{Q}_{Bu}}{\bar{T}_R - \bar{T}_O} \quad \text{('heat demand')}, \qquad m = \frac{\bar{I}_S}{\bar{T}_R - \bar{T}_O} \quad \text{('meteo')}\tag{6b}$$

$$H_0 = G_B \frac{1 - r_0}{\eta_{BIN}}, \qquad\qquad A_e = \frac{\dot{Q}_{Smax}}{l_{Smax}} \frac{1 - r_0}{\eta_{BIN}} \qquad\qquad (6c)$$

Taking 'measurements' h and m the parameters H_0 and A_e are identified. A problem with this approach is that h is usually not measurable because it includes the parameter p_2. Therefore this h-m-method is not further considered here.

3.3 Data directed model search

Model (5a) is based on physical considerations. However this doesn't guarantee an appropiate model. An alternative is to search the model structure based on a representative set of measurements by so-called stepwise-type regression procedures (see Montgomery and Peck [6], chap. 7). This approach is followed e.g. in Abdel-Nabi et al [7]. A further alternative is given by MacKay [8], who uses the available data to train an ANN. These data based methods are not discussed in this paper because the author has no experience with them.

4 Expected energy consumption based on prior knowledge

Basically estimates of the parameters in model (5a) are determined from prior knowledge and then used to determine the expected energy consumption or a related figure. Distinguish whether a heating demand calculation is available or not:

Case 1: Heating demand calculation available:

From this calculation the following data are available or derivable:
- T_{OD} : design outdoor temperature
- $\hat{G}_B$: estimate of building envelope heat conductance

Eventually also:
- $\hat{\overline{Q}}_{GU}$: estimate of average free heat gain
- $\hat{\dot{Q}}_{DWN}$: estimate of nominal domestic hot water energy demand
- $\hat{\dot{Q}}_{GoccN}$: estimate of nominal occupancy dependent internal heat gain

Useful data for estimating $\overline{\dot{Q}}_{GU}$, $\dot{Q}_{DWN}$, and $\dot{Q}_{GoccN}$ are given in [2], sect. 5.

The boiler manufacturer is expected to supply the following data:

- $\hat{\dot{Q}}_{BIN}$: estimate of boiler nominal power
- $\hat{r}_0$: estimate of stand-by losses

Estimates for further parameters appearing in (5a..h) are unlikely to be available. Using this prior knowledge the following model can be setup:

$$\hat{r} = \hat{p}_1 \frac{\overline{T}_R - \overline{T}_O}{T_{RN} - T_{OD}} - \hat{p}_2 - \hat{p}_3 \frac{\overline{\ell}_{occ}}{\ell_{occN}} - \hat{p}_4 \frac{\overline{\dot{Q}}_{GM}}{\dot{Q}_{GMN}}, \qquad\qquad (7a)$$

where:

$\hat{r}$ = expected average burner running ratio

$$\hat{p}_1 = \frac{\hat{G}_B(T_{RN} - T_{OD})}{\hat{Q}_{BIN}}(1 - \hat{r}_0) \tag{7b}$$

$$\hat{p}_2 = \frac{\hat{\bar{Q}}_{GU}}{\hat{Q}_{BIN}}(1 - \hat{r}_0) - \hat{r}_0 \tag{7c}$$

$$\hat{p}_3 = \frac{\hat{Q}_{GoccN} - \hat{Q}_{DWN}}{\hat{Q}_{BIN}}(1 - \hat{r}_0) \tag{7d}$$

$$\hat{p}_4 = \frac{\hat{Q}_{GMN}}{\hat{Q}_{BIN}}(1 - \hat{r}_0) \tag{7e}$$

If no internal gains are measured drop term with $\overline{\hat{Q}}_{GM}$ in (7a). If the occupancy level $\bar{\ell}_{occ}$ is constant or unknown proceed as in sect. 3.2, case 2, i.e. redefine $\hat{p}_2$ as $\hat{p}_2 = (\hat{\bar{Q}}_{GU} - \hat{Q}_{DWN})/\hat{Q}_{BIN}(1 - \hat{r}_0) - \hat{r}_0$. If some of the estimates $\hat{\bar{Q}}_{GU}$, $\hat{Q}_{DWN}$, and $\hat{Q}_{GoccN}$ cannot be supplied drop them in (7c) and (7d). In the worst case the terms with $\hat{p}_3$ and $\hat{p}_4$ must be deleted and $\hat{p}_2 = -\hat{r}_0$.

Case 2: No heating demand calculation available

Assume this case only for *old* installations (<1980). From the building location and climate tables the design outdoor temperature T_{OD} can be determined. From the boiler type sign the expected boiler nominal power $\hat{Q}_{BIN}$ can be read. Based on [1], sects. 3.1.3 and 5.6.2, estimate the boiler stand-by losses as $\hat{r}_0 = 2\%$ and its oversize factor as $\hat{f}_{os} = 3$. With these minumum data setup the model

$$\hat{r} = \hat{p}_1 \frac{\overline{T}_R - \overline{T}_0}{T_{RN} - T_{OD}} - \hat{p}_2, \tag{8a}$$

where:

$$\hat{p}_1 = \frac{\hat{Q}_{BIN}/\hat{f}_{os}}{\hat{Q}_{BIN}}(1 - \hat{r}_0) = \frac{1 - \hat{r}_0}{\hat{f}_{os}}, \qquad\qquad \hat{p}_2 = -\hat{r}_0 \tag{8b}$$

It is also possible to have combinations of case 1 and 2. Example: Estimates $\hat{\bar{Q}}_{GU}$, $\hat{Q}_{DWN}$, and $\hat{Q}_{GoccN}$ can be supplied, but $\hat{G}_B$ not. To estimate p_1 use (8b) instead of (7b).

Note that the models (7a..e) and (8a..b) are static. Therefore they can be applied only to large time periods d or to periods with small room temperature change ΔT_R.

5 Expected energy consumption based on previous measurements

Basically estimates of the parameters in model (5a) are determined based on previous measurements and then used to determine the expected energy consumption or a related figure.

5.1 Parameter estimation

If measurements in (5a) are missing the corresponding parameters cannot be identified and the model must be simplified according to sect. 3.2. The remaining parameters can then be identified by a least squares procedure, either direct (DLS) or recursive (RLS), (see Tödtli [4]).

The recursive version has the following advantages:
- low storage and computing demand
- estimates early available
- can handle measurement series of unlimited length
and the following disadvantages:
- several design parameters required
- at an early stage difficult to distinguish between good and bad measurements

The direct version has opposite features. A good compromise would be to run first the direct version and to derive from its results design parameters for the recursive version to follow.

Both versions can be supplemented by an estimator of the equation error variance λ^2 (see later in sect. 6.3). Further the recursive version given in [3], sect. 2, is modified by substituting the matrices P_k and Q_k according to

$$\begin{aligned} P_k &= \lambda^2 \tilde{P}_k \\ Q_k &= \lambda^2 \tilde{Q}_k \end{aligned} \quad , k=0,1,.. \tag{9}$$

where $\tilde{P}_k$ and $\tilde{Q}_k$ are scaled matrices. The complete estimator then reads:

$$\hat{p}_0 = p_{00} \tag{10a}$$

$$\tilde{P}_0 = \tilde{P}_{00} \tag{10b}$$

For k = 1,2,..

$$\hat{p}_k = \hat{p}_{k-1} + \frac{\tilde{P}_{k-1} a_k}{1 + a_k{}^T \tilde{P}_{k-1} a_k} \left(b_k - a_k{}^T \hat{p}_{k-1} \right) \tag{10c}$$

$$\tilde{P}_k = \tilde{P}_{k-1} - \frac{\tilde{P}_{k-1} a_k a_k{}^T \tilde{P}_{k-1}}{1 + a_k{}^T \tilde{P}_{k-1} a_k} + \tilde{Q}_k . \tag{10d}$$

$$\hat{\lambda}_k^2 = \frac{1}{k} \left[(k-1)\hat{\lambda}_{k-1}^2 + \left(b_k - a_k{}^T \hat{p}_{k-1} \right)^2 \right] \tag{10e}$$

where:

k   = number of estimation step (here: number of measurement periods observed so far)
b_k = left side of measurement equation (here: $b_k = \bar{r}_k$)

466

a_k = measurement vector (here: weights of parameters $p_{1..7}$ according to (5a))

$\hat{p}_k$ = parameter estimates (here: $\hat{p}_k = [\hat{p}_1, .. \hat{p}_7]^T$ or subset thereof)

$\tilde{P}_k$ = scaled estimation matrix, $\tilde{P}_k = Var[p_k | a_1, .. a_k, b_1, .. b_k] / \lambda^2$

$\hat{\lambda}_k^2$ = estimate of equation error variance λ^2

Design parameters:

p_{00} = expected value of initial parameter vector p_0

$\tilde{P}_{00}$ = $Var[p_0] / \lambda^2$

$\tilde{Q}_k, k = 0,1, ..$ = variances of parameter changes at steps k=0,1,.., scaled by λ^2

For the assumptions underlying the complete estimator (10a..e) see [3], sect. 2. However, in contrast to [3] λ^2 is now assumed constant to enable its canceling in updating $\tilde{P}_k$. Note that basically no initial estimate for λ^2 is required, because $\hat{\lambda}_0^2$ is canceled in (10e). However $\hat{\lambda}_0^2$ might still be useful to estimate $\tilde{P}_{00}$ as $\hat{Var}[p_0] / \hat{\lambda}_0^2$.

Consider now the choice of the design parameters p_{00}, $\tilde{P}_{00}$, $\tilde{Q}_k, k = 0,1, ..$, and $\hat{\lambda}_0^2$:

It is reasonable to set $\tilde{Q}_k = 0$ for all k. If $\tilde{Q}_k \neq 0$ the estimator would follow parameter changes due to gradually appearing faults. Such faults couldn't then be detected by comparing the expected and the effective energy consumption.

With $\tilde{Q}_k = 0 \; \forall k$ the estimator assumes the parameter vector p to be constant, and p_{00} is its expected value based on prior knowledge. For $p_{1..4}$ prior knowledge based estimates are already given in sect. 4. For p_5 and p_6 such estimates are very difficult to supply. Therefore simply set $\hat{p}_5 = \hat{p}_6 = 0$. For p_7 assume that the building can be heated up from T_{OD} to T_{RN} in about 2 days. Then

$$\hat{p}_7 = \frac{\hat{C}_B(T_{RN} - T_{OD})}{\hat{Q}_{BIN}d_N}(1 - \hat{r}_0) = \frac{48h}{d_N}.$$

Use these prior knowledge based estimates as components of p_{00}. Note that they only serve as start values for the parameter estimator. The estimator also works with bad start values, e.g. all $\hat{p}_i = 0$. It then simply takes longer to get plausible values.

A good choice for $\tilde{P}_{00}$ is $\tilde{P}_{00} = \hat{P}_{00} / \hat{\lambda}_0^2 = diag(\hat{Var}[p_1], .. \hat{Var}[p_m]) / \hat{\lambda}_0^2$, m = number of parameters to estimate. However, the variances of the individual parameters are hard to estimate based on prior knowledge. A rough estimate can be given by observing that usually

- $p_i \in (0,1)$ for $i = 1,2,4,5,6$

- $p_3 \in (-1,1)$

- $p_7 \in (0,1)$ if $d_N = 48h$

Following Tödtli's suggestion [4], sect, 3.1, estimate $Var[p_i]$ equal to 1 for i=1..7 (square of range, disregarding larger range of p_3).

To estimate the equation error variance λ^2 one may proceed as follows:

1) Subjectively estimate the maximum error $\Delta \bar{r}_{max} = |\hat{\bar{r}} - \bar{r}|_{max}$ (e.g. 20 %).

2) Assume $\Delta \bar{r}_{max}$ to be 2σ of the normal distributed equation error. Then $\hat{\lambda}_0^2 = (\Delta \bar{r}_{max})^2 / 4$.

5.2 Selection of appropriate measurement periods

The full model (5a) and its simplifications are usually rough approximations only. It is therefore advisable to select the measurement periods in such a way that the effects of the model deficiencies are as little as possible. This holds for the identification of the model parameters as well as for the later use of the model for fault detection. The following selection rules are suggested:

1) Period d not significantly smaller than 24h. (Reason: Keep variation of $\overline{Q}_{GU}$ low.)

2) Drop sections with small energy consumption. (Reason: Expect dominant unknown free heat gain.)

In case of static models:

- If measurement series covers at least 2 heating periods:
  3.1) Select d = 1 week.

- Otherwise select period such that:
  3.2) $d \approx 24h$.
  3.3) Room temperature has equal values at period begin and end.
  3.4) Room temperature changes at period begin and end in same direction.
  3.5) Change of average daily room temperature w.r.t. previous day is small.
  3.6) Deviation of average daily room temperature w.r.t. its long term filtered value is small.

The rules 3.1 or 3.2-6 resp. aim to avoid errors due to the thermal inertia of the building. An alternative to the rather complicate rules 3.2-6 is suggested by Visier and Paillassa [5], where a separate model is introduced for each weekday. As long as the occupancy schedule is fix each of these 7 models can adapt to the specific conditions of the weekday concerned (occupancy, free heat gains, short term history). A drawback is that large measurement series are required as in case d = 1 week.

An application may have available measurements of the solar intensity I_S. But the modelling of its influence according to (5a) may be insufficient because the solar heat gain may vary strongly due to varying occupants behaviour. In this case it is appropriate to exclude measurement periods with significant solar intensity.

A stastical method to detect and discard outliers is reported in sect. 6.3.

5.3 Testing parameter estimates and resulting model

Parameter estimates may be unreliable, especially during the first few steps of RLS estimation. Further the model structure may be inappropriate (e.g. under- or over-defined). It is therefore necessary to perform some tests before using the model for fault detection.

Parameter *estimates* may be *unreliable* because of so called multicollinearity, i.e. near linear dependence among the regressors. Mulicollinearity can be detected by checking the variance inflation factors

$$VIF_i = \begin{cases} [(A^T A)^{-1}]_{ii} & \text{for DLS case} \\ \tilde{P}_{ii} & \text{for RLS case} \end{cases}, \ i=1..m, \tag{11}$$

where m = dim(p), and A is the nxm measurement matrix composed of measurement vectors a_j, j=1..n. According to Montgomery and Peck [6], chap. 8, multicollinearity is severe it VIF > 5. In this case one must try to collect further measurements in *new regressor regions*. If this is not possible (here: measurements cover

already a complete heating period) one has to simplify the model by dropping one of the nearly linear dependent regressors. If the model has only 1 independent variable (here the outdoor temperature T_O) a simple alternative to testing VIF's is to directly check the observed range of this variable for being sufficiently large (see Visier and Paillassa [5], sect. 4.1).

If the VIF's are acceptable, the model may be checked for having *insignificant parameters*. In case of DLS a simple test is the following t-test for each parameter p_i:

$$\left| \frac{\hat{p}_i}{se(\hat{p}_i)} \right| \begin{matrix} H_1 \\ \geq \\ < \\ H_0 \end{matrix} \; t_{\alpha/2,n-m} \; , \; i=1..m, \tag{12}$$

where:

$$se(\hat{p}_i) \;\; = \;\; \sqrt{\frac{1}{n-m} \sum_{j=1}^{n} (b_j - \hat{b}_j)^2 \; [(A^T A)^{-1}]_{ii}} \; , \; \text{standard error of } \hat{p}_i \; (\text{here: } b_j = \bar{r}_j)$$

H_1 / H_0 = hypothesis that p_i is / is not significant

$t_{\alpha/2,n-m}$ = threshold of t-statistic for significance level $\alpha/2$ and n-m degrees of freedom

For the RLS case a test statistic corresponding to the one in (12) has not yet been studied here. A more reliable test is offered by the so-called backward eliminination procedure (see Montgomery and Peck [6], chap. 7, suitable for DLS only). Insignificant parameters and corresponding regressors can be deleted.

A measure for the quality of the whole model is the coefficient of mulitiple determination

$$R^2 = 1 - \frac{SS_E}{S_{yy}} = 1 - \frac{\sum_{i=1}^{n}(b_i - \hat{b}_i)^2}{\sum_{i=1}^{n}(b_i - \bar{b})^2} \; , \; (\text{here: } b_i = \bar{r}_i). \tag{13}$$

In case of RLS R^2 can be determined approximately only because $\bar{b}$ has to be computed recursively and is not known exactly until all n measurements are processed. Based on some first experiments $R^2 \geq 0.8$ is suggested for the model to be acceptable.

Finally the parameter estimates should pass the simple plausibility test of being within their expected ranges specified in sect. 5.1.

6 Testing energy consumption

Basically the energy consumption is checked for exceeding its expected value by more than a tolerated amount. A key problem is the choice of the corresponding threshold.

6.1 Criteria for threshold selection

On the one hand the threshold has to observe the uncertainty the expected energy consumption is subject to. On the other hand it has to observe when the excess energy consumption becomes severe concerning economics and environment. Rossi and Brown [9] distinguish between a statistical threshold and a rule threshold. As rule threshold they suggest an economic threshold resulting from the minimization of a cost function containing energy and service costs. Minimization is subject to comfort and safety constraints.

6.2 Simple economic threshold

A simple economic threshold Δr_e^* can be derived by requiring that the costs for a single service c_s are amortized by the energy cost savings within a given time d_a, say 1 heating period:

$$\Delta r_e^* Q_{dmax} c_e \frac{d_a}{d} = c_s \tag{14}$$

c_e are the energy costs per kWh. (14) is a gross simplification in that the excess energy consumption $\Delta Q_d = \Delta r Q_{dmax}$ is assumed to be the same for each time period. More precisely one may assume that Δr depends on the load and consider a typical total load, specified e.g. in heating degree days per year. This would lead to a load dependent threshold.

6.3 Statistical threshold based on previous measurements

Let: A = nxm measurement matrix composed of n<k previous measurement vectors $a_1..a_n$
  $b_i = \bar{r}_i$
  $b = [b_1,..b_n]^T$, measurement vector corresponding to A

DLS case:

A threshold can be determined by using the following results of Montgomery and Peck [6], chap. 4:

- Let: A and a_k given, b and b_k *unknown* r.v.

- $$\Delta b_k = b_k - a_k^T \hat{p} \sim N\left(0, \lambda^2(1 + a_k^T(A^TA)^{-1}a_k)\right), \tag{15}$$

  where

  $$\hat{p} = (A^TA)^{-1}A^Tb, \tag{16}$$

- $$(n-m)\frac{MS_E}{\lambda^2} \sim \chi^2_{n-m}, \tag{17}$$

  where

  $$MS_E = \frac{1}{n-m}\sum_{i=1}^{n}\left(b_i - a_i^T\hat{p}\right)^2 \tag{18}$$

  is a bias-free estimator for λ^2.

- Δb_k and MS_E independent $\tag{19}$

- $$\Rightarrow \frac{\Delta b_k}{\sqrt{MS_E(1 + a_k^T(A^TA)^{-1}a_k)}} \sim t_{n-m} \tag{20}$$

Fixing a type 1 error probability α the threshold for $\Delta \bar{r}_k = \Delta b_k$ follows as

$$\Delta r_{s,k}^* = t_{\alpha/2,n-m}\sqrt{MS_E(1 + a_k^T(A^TA)^{-1}a_k)}. \tag{21}$$

RLS case:

In [3], sect. 3.2, it is shown that for given $a_1..a_k$ *and* $b_1,..b_{k-1}$ similar to (15)

$$\Delta b_k = b_k - a_k{}^T \hat{p}_{k-1} \sim N\left(0, \lambda^2(1 + a_k{}^T \tilde{P}_{k-1} a_k)\right). \tag{22}$$

Substituting in the direct λ^2-estimator (18) $\hat{p}$ by $\hat{p}_{k-1}$ and n by k-1 the resulting new estimator

$$\hat{\lambda}_{k-1}^2 = \frac{1}{k-1-m} \sum{}_{i=1}^{k-1} \left(b_i - a_i{}^T \hat{p}_{i-1}\right)^2 \tag{23}$$

enables recursive computation. To simplify writing consider now $\hat{\lambda}_k^2$ instead of $\hat{\lambda}_{k-1}^2$. Here only little is known about the statistical properties of this estimator. It is not bias-free because

$$E\left[\hat{\lambda}_k^2 | a_1..a_k\right] = \frac{\lambda^2}{k-m} \left(\sum{}_{i=1}^{k} a_i{}^T \tilde{P}_{i-1} a_i + k\right). \tag{24}$$

For large k the bias decreases because $\tilde{P}_{i-1}$ dies out. Dividing by k-m instead of k makes the bias even larger. Replacing therefore in (23) k-m by k the resulting recursive estimator reads

$$\hat{\lambda}_k^2 = \frac{1}{k}\left[(k-1)\hat{\lambda}_{k-1}^2 + \left(b_k - a_k{}^T \hat{p}_{k-1}\right)^2\right]. \tag{25}$$

From this biased estimates unbiased ones, $\hat{\lambda}_{ub,k}^2$, could be derived by

$$\hat{\lambda}^2{}_{ub,k} = \hat{\lambda}_k^2 \frac{k}{s_k + k}, \tag{26}$$

where

$$s_k = s_{k-1} + a_k{}^T \tilde{P}_{k-1} a_k, \quad s_0 = 0. \tag{27}$$

Here the distribution of Δb_k given in (22) is only known to hold, if in addition to $a_1..a_k$ the measurements $b_1,..b_{k-1}$ are also given. But in this case $\hat{\lambda}_{k-1}^2$ is not a r.v. at all. A reasonable way out is to replace in (22) λ^2 by its estimate $\hat{\lambda}_{k-1}^2$, and assume that approximately

$$\Delta b_k \sim N\left(0, \hat{\lambda}_{k-1}^2(1 + a_k{}^T \tilde{P}_{k-1} a_k)\right). \tag{28}$$

The statistical threshold for $\Delta \bar{r}_k = \Delta b_k$ follows as

$$\Delta r_{s,k}^* = h_n(\alpha)\sqrt{\hat{\lambda}_k^2(1 + a_k{}^T \tilde{P}_{k-1} a_k)}, \tag{29}$$

where $h_n(\alpha) = \Phi^{-1}(1 - \alpha/2)$. α is a type 1 error probability.

$\Delta r_{s,k}^*$ may serve two purposes:

1) Before entering a new measurement pair $(\bar{r}_k, a_k)$ into the RLS estimator it may be checked for being abnormal by the *two-sided* test

$$\left| \bar{r}_k - a_k^T \hat{p}_{k-1} \right| \underset{H_0}{\overset{H_1}{\underset{<}{\geq}}} \Delta r_{s,k}^* \quad , \tag{30}$$

where H_1 / H_0 are the hypotheses that the measurement pair is / is not abnormal. This test serves to detect outliers above or below the expected value $a_k^T \hat{p}_{k-1}$ (e.g. abnormally high energy consumption due to wind or fault, or abnormally low consumption due to sun). Discarding outliers, α becomes the probability that a measurement pair is not used in the estimator although it is a normal one.

2) A measurement pair $(\bar{r}_k, a_k)$ may be checked for showing expectionally high energy consumption by the *one-sided* test

$$\bar{r}_k - a_k^T \hat{p}_{k-1} \underset{H_0}{\overset{H_1}{\underset{<}{\geq}}} \Delta r_{s,k}^* \quad , \tag{31}$$

where H_1 / H_0 are the hypotheses that there is / is not a fault causing high energy consumption. $\hat{p}_{k-1}$ is to be replaced by $\hat{p}$ in case of DLS. Redefining $h_n(\alpha)$ in (29) as $h_n(\alpha) = \Phi^{-1}(1-\alpha)$, α becomes a false alarm probability.

The two cases show that it is not always possible to distinguish between faults and disturbances. For parameter estimation this doesn't matter but for fault detection it is a severe problem.

6.4 Statistical threshold based on prior knowledge

Starting from (29) the prior knowledge based statistical threshold reads

$$\Delta r_{s,k}^* = h_n(\alpha)\sqrt{\hat{\lambda}_0^2(1 + a_k^T \tilde{P}_{00} a_k)} \approx h_n(\alpha)\sqrt{\hat{\lambda}_0^2 + a_k^T P_{00} a_k} \quad . \tag{32}$$

The values supplied for P_{00} in sect. 5.1 may do for RLS estimation by they are not sufficiently accurate for fault detection. A good choice for P_{00} is $P_{00} = \mathrm{diag}\left(\hat{V}ar[p_1], ... \hat{V}ar[p_m]\right)$. A simple possibility to estimate $Var[p_i]$, $i=1..m$, is:

1) Subjectively estimate the maximum errors in the prior knowledge based component estimates of p_i.

2) Build thereof the maximum error $\Delta p_{i,max} = \left|\hat{p}_i - p_i\right|_{max}$.

3) Assume $\Delta p_{i,max}$ to be 2σ of the normal distributed p_i. Then $\hat{V}ar[p_i] \approx (\Delta p_{i,max})^2 / 4$.

472

6.5 Threshold selection

Given an economic threshold and a statistical threshold follow Rossi and Braun [9] and select the larger of the two as relevant threshold Δr_k^*:

$$\Delta r_k^* = \max(\Delta r_e^*, \Delta r_{s,k}^*),$$ (33)

where:

Δr_e^* : determined from (14)

$\Delta r_{s,k}^*$: $\begin{cases} \text{as given in (21) or (29)} & \text{if model based on previous measurements} \\ \text{as given in (32)} & \text{if model based on prior knowledge} \end{cases}$

If the statistical threshold is larger than the economic one it is not possible to detect moderate faults. In case of a large difference the model should be improved.

In case of RLS estimation the statistical threshold based on previous measurements will quickly fall below the one based on prior knowledge. This may lead to the assumption, that the measurement based model is now more accurate than the prior knowledge based one and might be used already. However, the measurement based model may still be unreliable and lead to wrong conclusions as illustrated by Visier and Paillassa [5], fig. 12, Chamberte school. Therefore this model must not be used for fault detection until it has passed the tests in sect. 5.3.

6.6 Detecting slowly developping deviations

Although the estimator has no forgetting property ($\tilde{Q}_k \equiv 0$), in the long term it will still adapt to slowly developping deviations. To detect such deviations one can take a copy of the identified model e.g. every 2 heating periods and analyze those copies manually. An automatic detector using the statistical threshold (29) can work with the *first* copy.

6.7 Notification in case of threshold exceeding

Comparing the energy consumption with predictions based on inaccurate models can lead to many false decisions. To reduce the number of false alarms one can base on the ideas of Pakanen [9] and issue a fault message only in case of *several threshold exceedings within short time*. Further issue the message the sooner the larger the exceedings are. Formally e.g.:

$$\sum_{i \in I_k} \Delta \bar{r}_i \underset{H_0}{\overset{H_1}{\underset{<}{\geq}}} \gamma L \overline{\Delta r_{I_k}^*} \ ,$$ (34a)

where

$I_k = \{k - L + 1 \leq i \leq k \,|\, \Delta \bar{r}_i \geq \Delta r_i^*\}$, index set of exceedings (34b)

$\overline{\Delta r_{I_k}^*} = \dfrac{1}{|I_k|} \sum_{i \in I_k} \Delta \bar{r}_i^*$, average of exceeded thresholds (34c)

$|I_k|$ = number of elements in I_k

L　　　= time window to remember exceedings (e.g. $L = 10\ d_N$)

H_1 / H_0 = hypotheses that there is / is not a fault

γ　　　= minimum percentage of exceedings of size $\overline{\Delta r_{I_k}^*}$ required for notification

To distinguish between cases where the presence of a fault is rather uncertain and those where it is rather certain, a small γ and a large one can be introduced. In the first case a warning message is appropriate whereas in the second case a fault message is due.

7　Outlook

The methods suggested in this work are only partly tested. Therefore the next step is to apply them to different installations and data sets.

References

[1] Bundesamt für Konjukturfragen: Handbuch Planung und Projektierung wärmetechnischer Gebäude-sanierungen, EDMZ, Best.-Nr. 724.500d, Bern, 1983.

[2] Erni H.P.: Energiemanagement, ein interdisziplinäres Thema, INFRASTRUCTA, Basel, 1994.

[3] Sprecher P.: Adaptive tolerance band for an energy signature, AN25/CH/240993/1.

[4] Tödtli J.: Recursive parameter estimation according to the method of least squares, associate optimization problems and direct solutions, in German, Scientra Electrica, Vol. 29, Nr. 4, Basel, 1983.

[5] Visier J.-C., Paillassa V.: Application of an energy signature method to 6 different schools, AN25/F/130494/1.0.

[6] Montgomery D.C, Peck E.A.: Introduction to linear regression analysis, Wiley, 1982.

[7] Abdel-Nabi D.Y. et al.: Regression analysis of residential air conditioning energy consumption at Dhahran, Saudi Arabia, ASHRAE Trans. 96, part 2, 1990, P. 223-32.

[8] MacKay D.J.C.: Bayesian non-linear modelling for the energy prediction competition, ASHRAE Transactions Symposium paper, 1994.

[9] Rossi T., Braun J.: Threshold selection and optimal maintenance, Annex 25 meeting, Boston, 1994.

[10] Pakanen J.: Classification of faults as a method to eliminate problems of threshold adjustment, Annex 25 meeting, Zürich, 1993.

[11] Schweizerischer Ingenieur- und Architekten-Verein, Bundesamt für Konjukturfragen: Leitfaden für Messungen an Gebäuden und Haustechnikanlagen, D027, EDMZ, Best.-Nr. 724.619d, Bern, 1989.

A COMBINED APPROACH FOR FDD IN HVAC SYSTEMS USING CHARACTERISTIC CURVES MODELLING AND KNOWLEDGE BASED DIAGNOSIS

Klaus BAER

Universität Stuttgart
Institut für
Kernenergetik und Energiesysteme
Pfaffenwaldring 31
70550 Stuttgart
GERMANY
E-mail: baer@ike.uni-stuttgart.de

Madjid MADJIDI

Universität Stuttgart
Institut für
Kernenergetik und Energiesysteme
Pfaffenwaldring 35
70550 Stuttgart
GERMANY
E-mail: madjidi@ike.uni-stuttgart.de

Abstract

In the presented approach, faults are detected through comparing reference against measured data upon pre-set thresholds. The reference data is generated through simulation based on characteristic curves. Threshold violation initiates knowledge based diagnosis. The diagnosis uses knowledge models of building, system, performance indices, symptoms and faults. In case of a threshold violation, the diagnosis tries to collect sensors of interest by applying rules to the knowledge models, to give a prognosis for those sensors for the next time step and to check them for threshold violation. Through that, a pattern of real and prognosticated symptoms is generated. It is used for generating fault hypotheses and a preliminary diagnosis, which has to be confirmed in the next time step.

1 INTRODUCTION

Optimizing buildings in terms of low energy consumption and low cost in spite of high comfort requires the use of new methods. They are usually supported by various software tools for the different phases in the building life-cycle. Big efforts have been made during the last years to represent on the computer all phases in the building life-cycle, from design over construction to operation phase and to integrate the corresponding software tools to take advantage of a continuous description of the building together with its technical equipment [4].

In this context, fault detection and diagnosis (FDD) in HVAC systems can be considered as one (sub-) task, which starts with the operational phase. If FDD is integrated in such a life cycle environment, a huge amount of knowledge about the building, its usage and environment, the system and the processes it runs, its components and their topology

has been collected (and validated!) until the operational phase. This enables FDD methods to operate on a wide knowledge base. In addition, the simulation models for building and system simulation are available already from previous design phases. These simulation models can be used to compute the system performance as a reference for the systems behaviour. Thus, the fault detection is reduced on comparing simulated (= reference data) and measured (possibly faulty) values. This simplifies fault detection decisively. Through the fact, that most of the time there will be a difference between reference and measured values, it is necessary, to define thresholds. The violation of such a threshold indicates a faulty process. The thresholds can be either determined through systematic, precursory off-line example simulations, experience, constraints or in relation to the given set points.

The available knowledge about the building and the system has to be extended through knowledge about faults, symptoms causing those faults and fault/symptom relations. Major technical considerations in describing this knowledge are how to formulate the relations (rules) on such an abstraction level, that they become as independent as possible from a specific system so that they can be used in different projects without variation. The presented approach tries to achieve this by describing symptoms and faults in a component independent way and by formulating the rules upon either basic physical equations or symptom patterns.

2 CONCEPT

The fundamental structure of the discussed approach for a FDD-system is shown in fig. 1. In analogy to the two steps fault detection and fault diagnosis, the system is divided into two parts (see [6]): the preprocessor supervises the system with regard to threshold violations and starts diagnosis if necessary. The diagnosis is managed by the classifier. The core of the classifier is the knowledge base describing following knowledge domains:

- knowledge about the performance indices,

- knowledge about possible symptoms,

- knowledge about the controlled system including system topology ,

- knowledge about the building, environmental conditions and buildings usage,

- knowledge about possible faults.

The knowledge is described through hierarchically ordered objects collecting data in object attributes and processing information in methods attached to the objects. The hierarchies give the possibility to reason on different levels of abstraction and to help limiting searchspace for faults. Thus, all advantages of object-oriented programming can be used (classification, hierarchies, (multiple) inheritance, message passing, polymorphism, ...) and in addition, declarative programming (rules) can be closely intermixed with
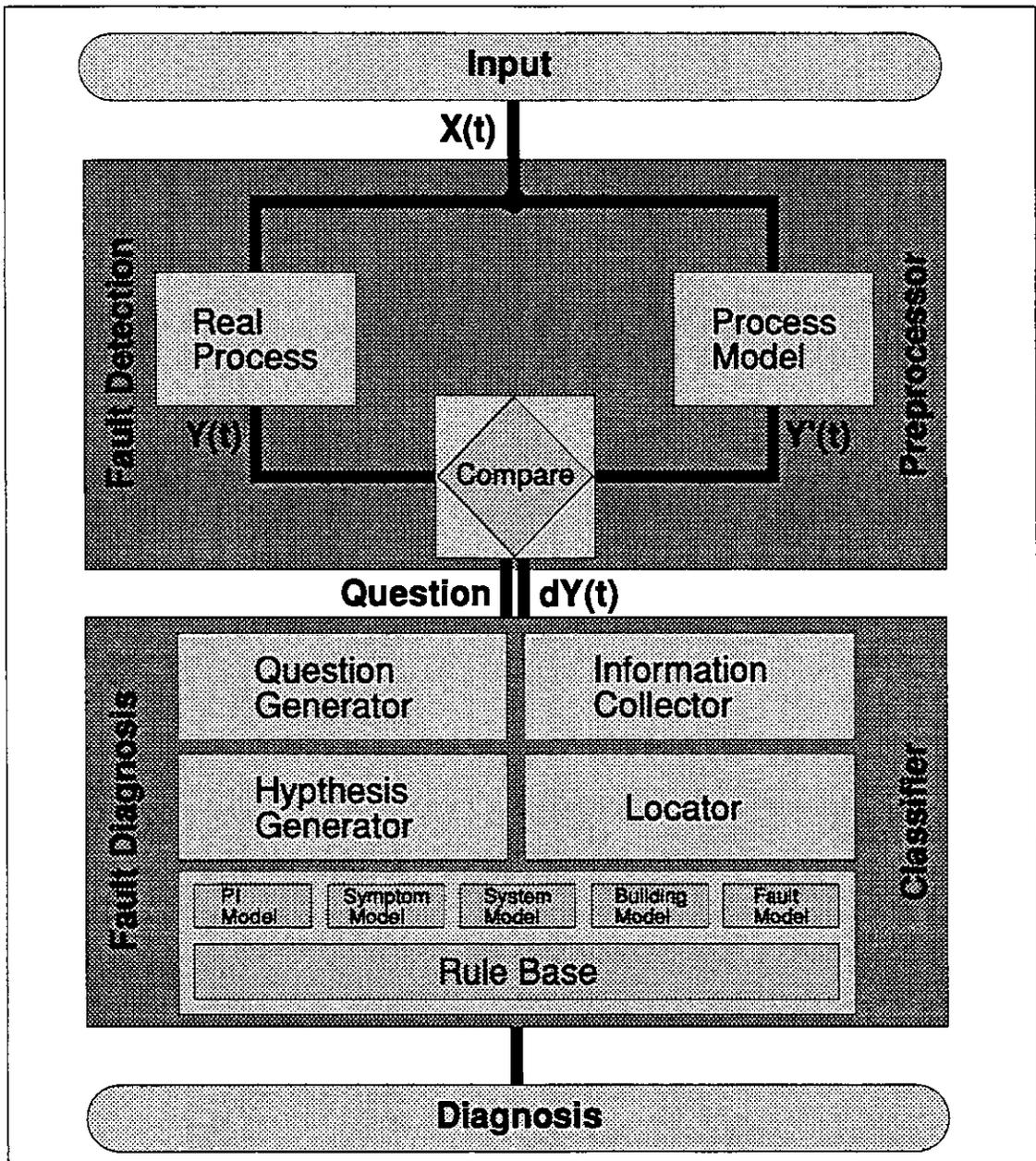
Figure 1: Structure of a knowledge based FDD-System

procedural programming (methods). Therefore the best fitting paradigm can be chosen for a given subtask (e.g. calculations by methods, diagnosis by rules).

The rulebase mainly represents the relations between the models. If the preprocessor initiates a request for diagnosis, the inference process collects the information, which is directly available in this moment and tries to isolate the location of the fault (e.g. on branch level). Upon this basis the inference process builds hypotheses, which have to be verified through the rule base. If this verification requires additional information to

the measured or simulated data, the system generates corresponding questions to the preprocessor. The information returned by the preprocessor will be used additionally by the diagnosis.

3 FAULT DIAGNOSIS KNOWLEDGE MODELS

3.1 PERFORMANCE INDEX MODEL

A performance index (PI) is determined according to a specific calculation method [6]. It is required for the calculation method, that an intervall can be specified, so that for values of the PI within the intervall it is allowed to assume normal operation. The simplest form of a PI is the difference between desired and measured value. Other examples are integrated values like efficiency or heat demand, statistcal values like mean or maximum values of time series or mass flow or energy balances of the considered processes. The PI's characterize the system.
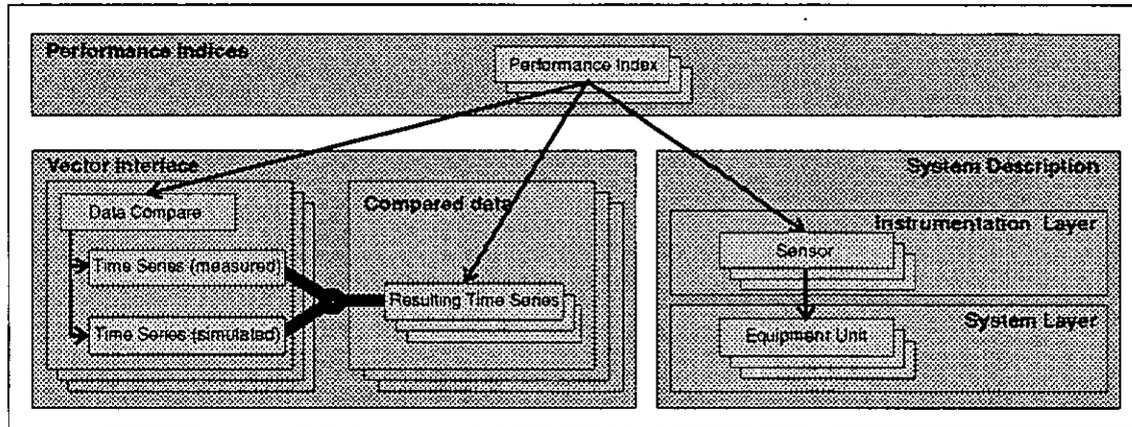
Figure 2: Performance Indices as interface between real system and its model in the diagnosis system

If any PI exceeds one of its predefined thresholds, a fault has probably occured. The performance indices are the central element for fault detection. As interface between the incoming data and the system model (see fig. 2), the PI's are responsible for initiating diagnosis. Each PI surveys a single variable at a specific location, where the variable may be directly a process variable (pressure, flow, temperature, or valve position) or inferred variables like trends and trend rates etc. The attributes of the PI are shown in table 1. At least some of the system characteristic PI have to be checked periodically to ensure a precocious fault detection.

Table 1: Attributes of the performance indices

Attribute	Meaning
Value	Current numerical value of the PI
Sensor	Reference to the sensor the PI is attached to
Input data	Reference to the simulated and measured data
Resulting data	Reference to the compared data
Upper threshold (absolute)	Maximum value (regarding absolute data)
Upper threshold (difference)	Maximum value (regarding difference between desired value and measured value)
Lower threshold (absolute)	Minimum value (regarding absolute data)
Lower threshold (difference)	Minimum value (regarding difference between desired value and measured value)
Methods	
Calculate PI	Method for calculation of the value of the PI
Check PI	Verification of threshold violations

3.2 SYSTEM MODEL

The system description represents the basic knowledge for fault detection and diagnosis. Besides the layout data of the components, the system description provides the information about the topological relations between the components (see table 2). The predeccessor/successor attributes represent the sequence of components in a branch of a process. The direction is given through the flow direction of the medium (water, air) in that branch.

In addition, the description of the system instrumentation is necessary. Measuring instruments are special system components. They do not influence the system process directly. Instead, they are measuring a process variable in a specific location of the process under control. Malfunction of an instrument does not affect any process variable except through other system components controlled by the instrument (like e.g. a valve). For fault diagnosis, the location of the measuring signal is important. Therefore, the same topolgy describing attributes shown in table 2 are relevant also for measuring instruments. As the sequence of instruments is not important, the predeccessor and successor attributes are used to determine the location of an instrument between two 'real' system components.

Attribute	Meaning
Branch	Branch(es) of which the current component is a component of
Components	Subcomponent(s) of current component
Predeccessor	Preceeding component(s)
Successor	Succeeding component(s)

As an example, in fig. 3 the structure of the system description for a fully instrumented district heater substation as described in [1] is shown including attached PI's. The part-of relation between the system components is shown in fig. 4. Each system component is instanciated of a corresponding kind of generic system components wherefrom it inherits its layout attributes and behaviour (methods). At the stage of FDD, all those attributes
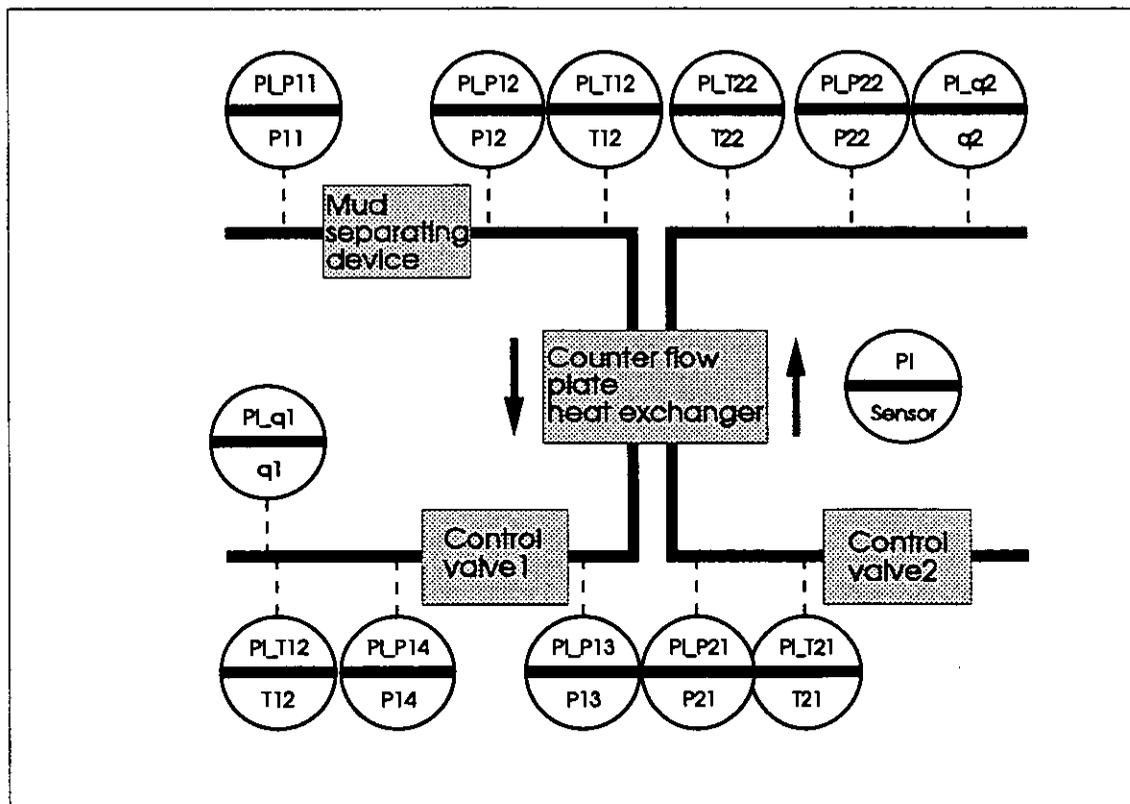
Figure 3: System structure of a fully instrumented district heating substation including PI instances (joint evaluation exercise of the heating group)
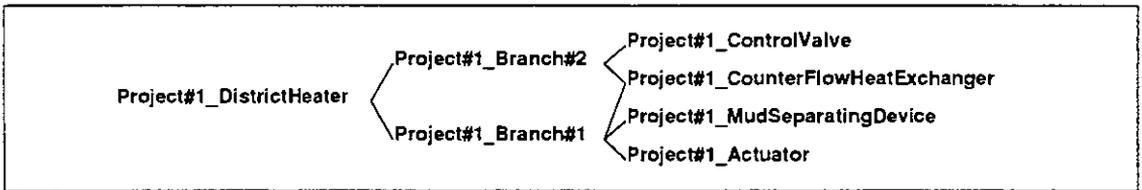
Figure 4: Example of part-of-relation of system components

should have been calculated or requested during preceeding design phases.

3.3 BUILDING MODEL

Important informations of the building description for FDD are the required indoor temperatures together with the allowed ranges for all rooms, the knowledge about usage schedule, the inner heat sources as well as the size and the direction of windows. Those variables determine the variation in load demand. For fault diagnosis it is important to have access to these data, because the simulated reference behaviour of the system probably will not be able to represent every possibly abrupt variation in these variables (e.g. variations in solar gain on a partly cloudy day). Hence it will result in false alarms and the fault diagnosis has to be put in the position to identify them as false alarm. Therefore the above mentioned data is used. In addition, the building model is required to have access to the room-equipment relation. The used building model is described in more detail in [2]. For the discussed DHS example, no knowledge about the building description is required except the load demand.

3.4 SYMPTOM MODEL

Symptoms are the effects of faults. Several kinds of symptoms may be distinguished:

- symptoms, whose deviation can be measured

- symptoms, whose deviation can be calculated

- symptoms, which can be observed

Most symptoms can occur in same or different expression in any of the system processes (heat supply, energy supply, domestic hot water supply). The measurable symptoms will be recorded either on the inlet or the outlet side of a device (single component, whole branch) (see fig. 5), because the measuring instrument normally are positioned there. 'Device' in this context may be any aggregation of components, that is either the whole district heater substation, the primary or secondary loop as a whole or a single component in one of the loops.
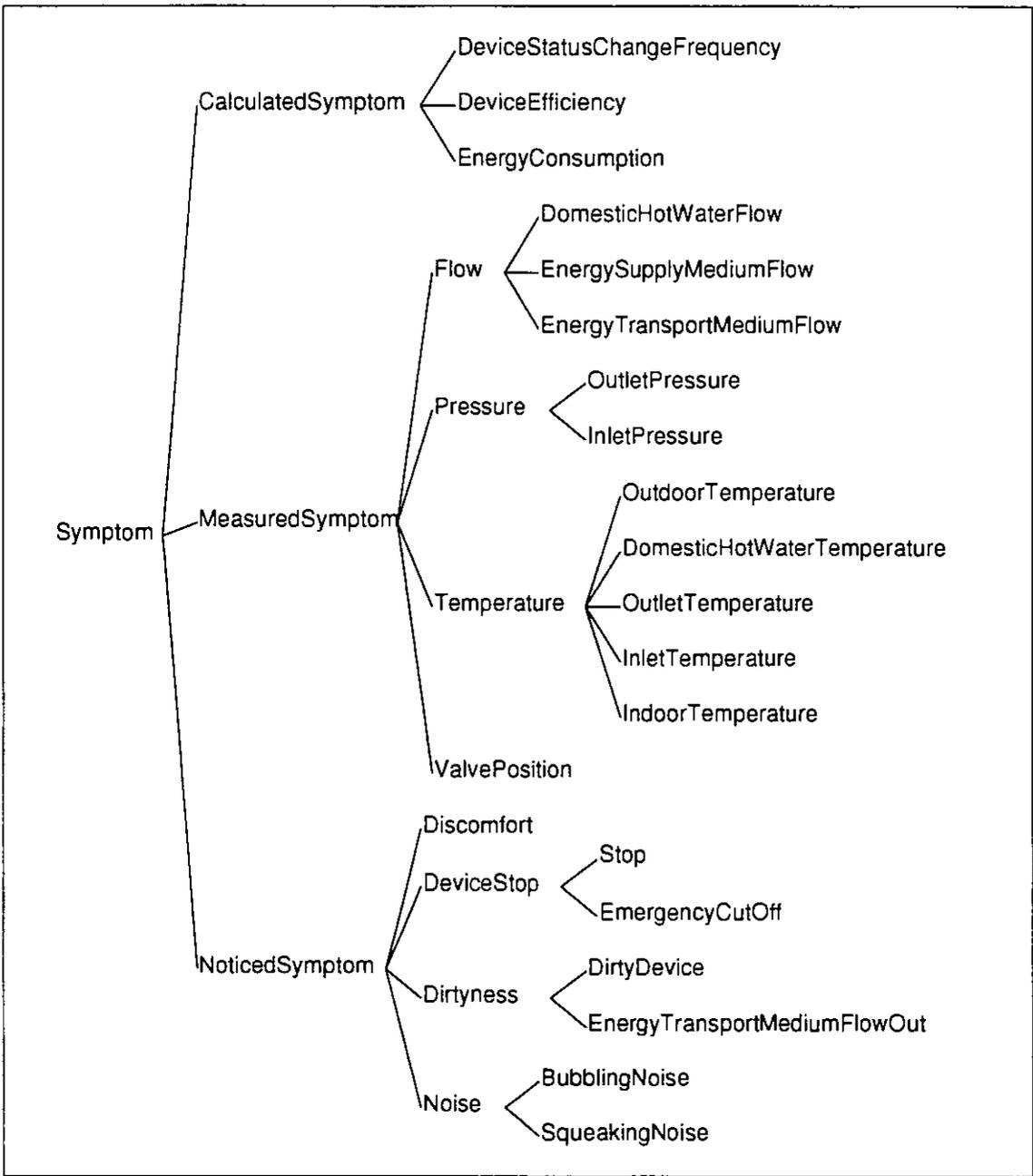
Figure 5: Structure of symptoms
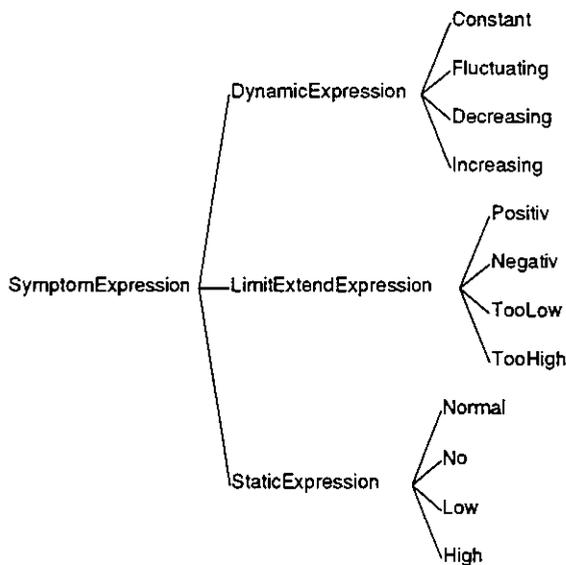
Figure 6: Possible symptom expressions

Typically, symptoms are described in qualitative form through expressions as shown in fig. 6. The expressions may be structured into static, dynamic and comparative expressions. The distinction between symptoms and their expressions makes the handling of symptoms and creation of hypotheses more easy. Symptoms can change their expression during fault diagnosis without the need to instantiate a new symptom, e.g. from fluctuating to decreasing. In addition, they can be assumed to be present even with the expression 'Normal' what means, that the corresponding process variable is in normal operation conditions.

As shown in fig. 2, PI's are the interface between the time series (measured, simulated and compared) and the computer model of the system. Each PI monitors a distinct process variable over time. If a monitored variable exceeds the pre-set thresholds, the PI signals this as a symptom. The symptom expression is determined upon the kind of threshold violation (e.g. exceeding upper or lower threshold).

3.5 FAULT MODEL

Similar to the symptom model, the faults can also be classified independendly from the system components (see fig. 7). For each 'fault family', it is possible to formulate specific rule-sets and methods, especially to investigate faults of the actual fault family. So, search-space will be reduced: after hypothesising a specific fault only the corresponding rulesets have to be examined. In addition, such a structure allows to infer on different levels of abstraction. First, the type of fault is isolated (e.g. operation fault, wear fault

GenericFault

OperationFault
- CoordinationFault
  - InsufficientSurroundingConditionsConsideration
  - InsufficientPowerDemandCoordination
- MeasuringFault
  - DefectSensor
  - DeviceInfluencedSensor
  - SensorBiasOrDrift
- ControlFault
  - WrongSetPoint
  - UnfavorableScheduling
  - UnfavorableSettings

Failure
- DefectDevice — JammedDevice
- BreakDown
  - ElectricalBreakDownFault
  - MechanicalBreakDownFault
  - ThermalBreakDownFault
    - OverheatingFault
    - IcingFault
- CutOff
- Blockage
  - PartlyBlockageFault
  - TotalBlockageFault
- Leakage
  - BlockageFault
  - LeakageToOtherBranchFault
  - LeakageFault

WearFault
- EnergyTransportMediumFault
  - DirtyEnergyTransportMediumFault
  - AirInWaterCircuitFault
- SoilingFault
- DepositionFault

Figure 7: Structure of faults

or breakdown) and then the kind of fault is determined (e.g. Blockage, Leakage or Cutoff).

4   REASONING MODEL

The problem of finding generally applicable rules can be reduced by assuming fully instrumented systems. In this case, each process variable at each device inlet and outlet (temperature, pressure, flow, valve position) will be measured and be available for diagnostic purposes. This simplifies the rules considerably. Of course, additional rules are required, which handle the normal case of not fully instrumented systems. Strategies in such cases are to extend the focus from component level to branch level in the hope to

Table 3: Fault attributes

Attribute	Meaning
Location	Location where fault occured
Equipment type	Type of equipment components where actual fault can occur in
Correlated symptoms	List of syptoms indicating this fault

have the missing variable on that level or to make asumptions on missing variables.

4.1   STRUCTURING OF RULES

The rule base is structured in rules for controlling the inference process, component related rules, symptom related rules and fault related rules. For symptom and fault related rules it is advisable to structure the rules according to the symptom and fault knowledge bases (see fig. 8).

While component related rules express component specific relationships like dependencies between the primary and secondary loop in the heat exchanger, symptom related rules may be understood as some kind of bottom-up rules, checking basic relations between process variables. In contrast, the fault related rules describe more complex relations and symptom patterns, which have to occur simultaneously to indicate a fault.

4.2   RULES

The rules describe the mutual relations between the system and symptom model as well as the symptom/fault relation. In general, rules can be implemented as forward or backward chaining rules. Backward chaining is more efficient, if the possible results are known and the total number of possible results is limited [3]. This is usually true for diagnostic problems. In our case, the possible results are the faults shown in 7. The number of possible faults is limited, because they are abstract, i.e. device independend, faults. Of course the total number of all faults in a whole plant may be very high. This is one reason to divide the inference process into a part for collecting sensors and correlated components of interest and into a part to build and check fault hypotheses for each of the components of interest separately. In that case, the whole diagnostic problem is splitted up into smaller, component related diagnostic problems with limited search space. The task of collecting sensors and components of interest is solved preferably through forward chaining, because the result is topology dependend and can not be described in advance.
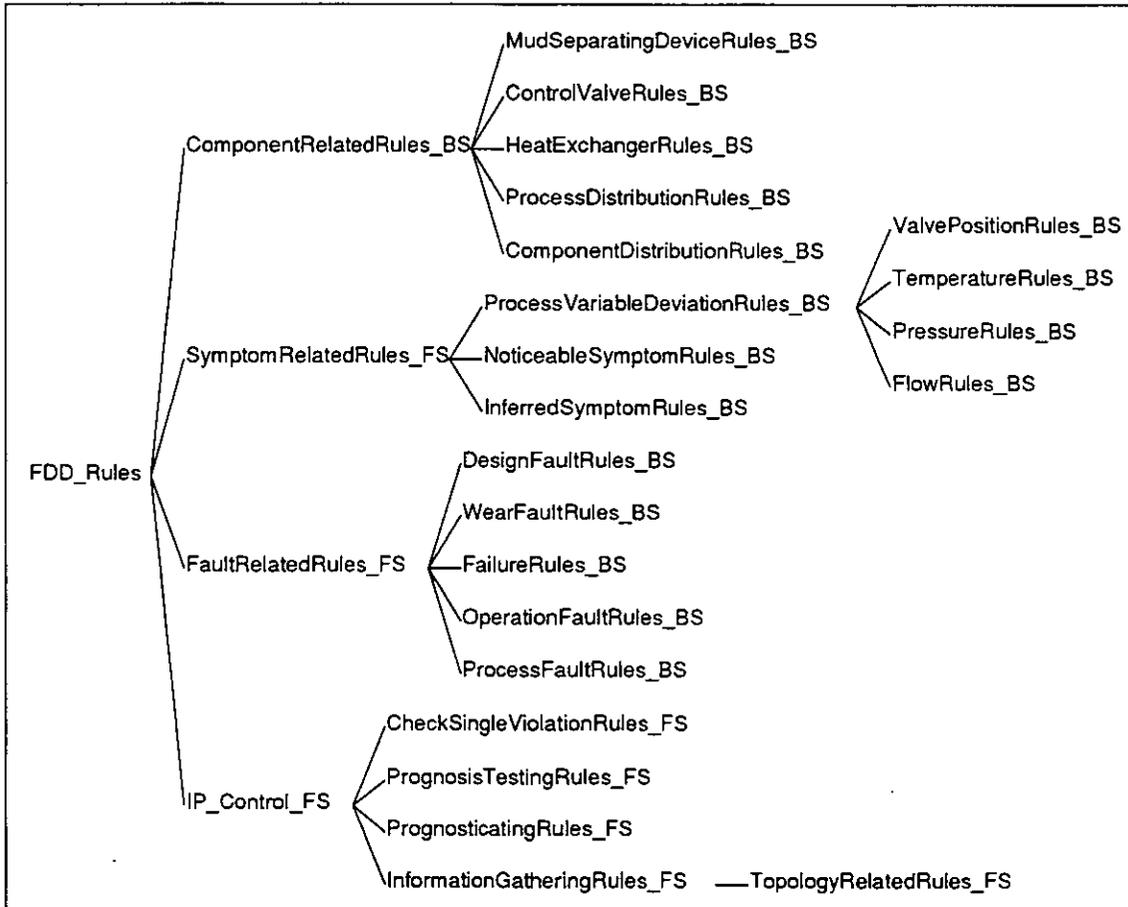
Figure 8: Structure of rule-sets

The rules shown in table 4 are developed from fundamental physic equations from the point of view of symptoms related to process variable deviation. They are written in the form

IF $< process\ variable\ deviation/fault >$ THEN $< symptom >$

as the causal dependency between fault and symptom is. In general, those rules lead to the assumption, that a specific fault could cause the symptom. Table 5 exemplarily shows symptom patterns which have to occur simultaneously, if the assumed fault really causes the initiating symptom.

4.3 DIAGNOSIS SCHEME

The diagnosis process will be started through one of the instanciated PI which has detected a threshold violation. First, the system collects the available information which includes:

Table 4: Fundamental rules inferred from basic physic equations

IF	THEN
Leakage	Flow(Device, In) $\neq$ Flow(Device, Out) $(\Longrightarrow \dot{m}_1 = \dot{m}_2)$
PressureDrop(Device) too low	Flow(Device, Out) too low $\vee$ Resistance too low $(\Longrightarrow \Delta p = \dot{m}^n \cdot R)$
PressureDrop(Device) too high	Flow(Device, Out) too high $\vee$ Resistance too high
Leak(Device) $\vee$ Blockage(Device)	Flow(Device, Out) too low
Total Blockage (Device)	Flow(Device, In) == 0 Flow(Device, Out) == 0
Cross section (Device) too big	Resistance too low
Cross section (Device) too small	Resistance too high
Leak(Device) $\vee$ Breakage(Device)	Cross section (Device) too big
Dirtyness(Device) $\vee$ Blockage(Device)	Cross section (Device) too small
Pump problems	Flow(Device, In) too low $\wedge$ Flow(Device, Out) too low $\wedge$ Flow(Device, In) == Flow(Device,Out)

- the sensor who triggered the PI,
- the branch of which the sensor is a component of,
- the type of the PI (value, trend, trend rate, etc.),
- the type of the variable under control (temperature, flow, etc),
- the expression of the variable deviation (too high, too low, etc),
- the component predeccessor and successor of firing sensor.

Next, the diagnosis process tries to limit the search-space by checking only sensors of interest. Sensors of interest are the sensors next to the device the first deviation was detected at or sensors using the firing variable for controlling another device. This reflects the simple assumption, that the fault will be found with good probability in the same device, in which the fault expressed itself in one or several symptoms. If the measured data of the initiating sensor is used for controlling another device (e.g. in fig.3, T_{22}) controls the valve position of control valve 1), this establishes a connection between the

Table 5: Fault/symptom patterns of leakage faults

Fault	$T_{D,In}$	$T_{D,Out}$	ΔT_D	$P_{D,In}$	$P_{D,Out}$	ΔP_D	$\dot{m}_{D,In}$	$\dot{m}_{D,Out}$
Leakage								
Leakage					too low	too high	constant (decreasing)	$\leq \dot{m}_{D,In}$
LeakageToOtherBranch								
Different medium					too low	too high	normal	$\leq \dot{m}_{D,In}$
Same medium								
1.) $P_{PS,In} = P_{SS,In}$								
$\vee T_{PS,In} > T_{SS,In}$								
PS	normal	too low	too high				normal	
SS	normal	too high	too low				normal	
$\vee T_{PS,In} < T_{SS,In}$								
PS	normal	too high	too low				normal	
SS	normal	too low	too high				normal	
2.) $P_{PS,In} > P_{SS,In}$								
$\vee T_{PS,In} > T_{SS,In}$								
PS	normal	too low	too high	decreasing	decreasing	too high	normal	decreasing
SS	normal	too high	too low	normal	increasing	too low	normal	increasing
Blockage								
Total							0	0
Partly				too high	too low	increasing	decreasing	decreasing

PS = Primary Side
SS = Secondary Side

two devices. Therefore, the sensors attached to both devices have to be regarded as sensors of interest. For those sensors of interest prognoses for the next time step will be estimated on the basis of extrapolation of the simulated and measured data. This will result in a pattern of prognosed threshold violations.

In the most unfavorable case, this pattern will be exactly the initiating symptom. In this case, the available information is to small to make a diagnosis. The system gives a message upon the detected threshold violation and will check after the next time step, if it is possible to prognose another symptom pattern with the new data.

If a symptom pattern can be prognosed, basic rules as shown in table 4 will be applied and corresponding fault hypotheses will be put up. Those hypotheses have to be checked with rules as shown in table 5. If this does not lead to a diagnosis result, another list of sensors of interest has to be collected, e.g. all sensors in the actual branch. Again, they have to be checked in the same way as described. If it is still not possible to give a diagnosis, it will be considered as a false alarm and in the next time step another try of a diagnosis will take place.

An overview over the presented models and their correlations as well as the diagnosis scheme is given in fig. 9.

5   SYSTEM APPLICATION

For the joint evaluation exercise a simulation model was built with the characteristic curves method [5]. The model input and output parameters are shown in fig. 10. The used equations are shown in fig. 11.

In the following, the test case 9 'jammed control valve 1' is discussed. Fig.12 shows the temporal curves of the model input parameters: inlet temperatures of primary and secondary side, flow of secondary side, valve throw of control valve 1 and the inlet pressure in front of the mud separating device. Fig.13 shows the temporal curves of the model output parameters: outlet temperatures of primary and secondary side, flow in primary side and inferred parameters like temperature differences at the heat exchanger in both branches and the pressure drop at the control valve. In each diagram, the measured and simulated curves as well as the differences are shown. All curves show filtered values (filter length 20 time steps).

The adjustment of the thresholds of the monitored PIs is shown in fig. 14. Fig. 15
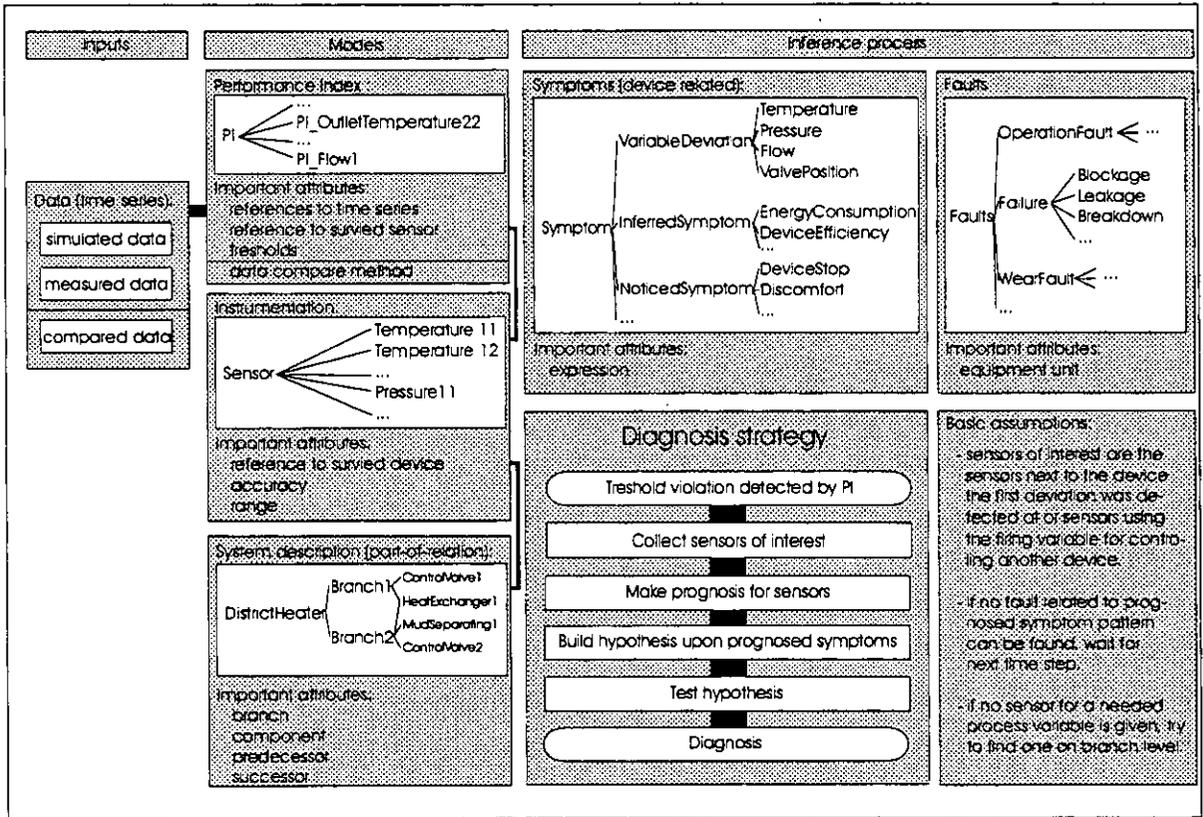
Figure 9: Overview on model dependencies and the diagnosis scheme of the presented approach
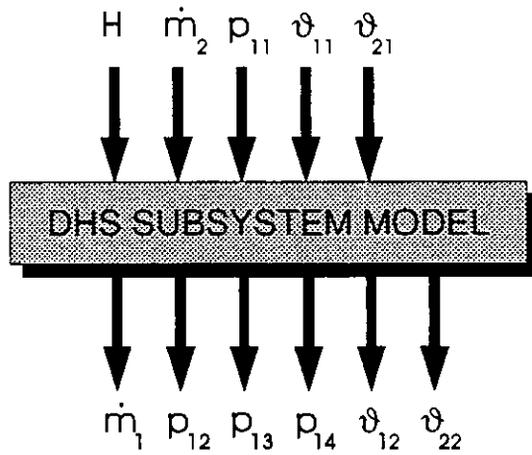
Figure 10: Input and output parameters of the DHS simulation model

shows the attributes of the monitored PIs at the time step, in which a threshold violation is detected the first time ($\rightarrow$ 271, see attribute *TresholdViolationTimestep*). The PI monitoring the pressure drop at the valve detects this threshold violation (P_{13} - P_{14}). The detection is marked through the value *TooLow* in the attribute *Treshold-*

Valve and Motor:	$\dot{m}_1 = \dfrac{a}{H^b} + c$	(1)
Hydronic:	$p_{12} = p_{11} - \dfrac{R_{12}}{\dot{m}_1^{n_{12}}}$	(2)
	$p_{13} = p_{12} - \dfrac{R_{23}}{\dot{m}_1^{n_{23}}}$	(3)
	$p_{14} = p_{13} - \dfrac{R_{34}}{\dot{m}_1^{n_{34}}}$	(4)
Heat Exchanger:	$\epsilon = c_1 \cdot \dot{m}_1^2 + c_2 \cdot \dot{m}_1 +$ $c_3 \cdot \dot{m}_1 \cdot \dot{m}_2 + c_4 \cdot \dot{m}_2 +$ $c_5 \cdot \dot{m}_2^2 + c_6$	(5)
	$\vartheta_{12} = \vartheta_{11} - \epsilon \cdot (\vartheta_{11} - \vartheta_{21})$	(6)
	$\vartheta_{22} = \vartheta_{21} + \dfrac{\dot{m}_1}{\dot{m}_2} \cdot (\vartheta_{11} - \vartheta_{12})$	(7)

Figure 11: District heating substation model based on characteristic curves

ViolationExpression. The initiating PI is marked through the value *False* in attribute *IsPrognosedViolation*. As described above, a list of sensors of interest will be collected. On the basis of the values of the last 20 time steps a prognosis for the value in the next time step for each sensor of interest is estimated. These estimated values will be checked for Threshold violation. As can be seen in fig. 15, a threshold violation is prognosed for the PIs T_{12}, P_{14}, and T_{11}-T_{12}. The values and expressions can be found in the attributes *ThresholdViolationValue* and *ThresholdViolationExpression*.

The pattern of found symptoms and their expressions together with the system topology will be investigated further with rules as discussed above. The symptoms *temperature difference at heat exchanger too low, pressure drop at valve too low* and *pressure after valve too high* all indicate, that the flow in primary side is lower, than the flow should be according to the reference. Too low flow can be caused through partial blockage, leakage or wrong measuring values in the actual branch. Leakage can be eliminated, because in that case, the pressure would decrease. Too low pressure drop over the valve may be caused through a partial blockage after the valve or through discrepancy between the measured and the real valve position. A preliminary diagnosis on basis of the prognosticated values in time step 271 is *partial blockage after the valve or jammed control valve*. This diagnosis causes an alarm and will checked in future time steps. Depending on the really measured values, the diagnosis will be confirmed or corrected. As can be seen in fig. 13 (upper right diagram), depending on the thresholds adjustment, a difference between measured and simulated primary side flow is detected in the next time steps. Considering the input valve position (see fig. 12, lower right), the simulated flow increases according to the valve opening, where the measured flow is nearly constant. This is produced through either wrong measuring or jammed valve which is the final diagnosis.

6 CONCLUSIONS

The major advantage of the presented approach of a knowledge based diagnosis system is its general applicability. In principal, the presented system can be applied to a wide variety of HVAC systems without changes (asuming the rule base is complete). Only the system description has to be adopted to the system under control. But the diagnosis uses a lot of knowledge not only about the system but also about the building, the usage, the environmental conditions, etc. Because of the huge amount of required knowledge, such an approach is preferably recommendable in the context of a life cycle environment as mentioned in the introduction.

A general drawback of such knowledge based approaches is that it is difficult to ensure completeness and correctness as well as to check consistency in the underlying knowledgebases. Other advantages and disadvantages are summarized in table 6.

Table 6: Summary of advantages and disadvantages of the presented knowledge based approach

Advantages	Disadvantages
• explanation facility • attempt to model the reasoning of a human diagnostician (duplication of inference process possible) • general approach (only system description changes) • can be easily combined with other approaches (e.g. associative networks, neural networks)	• big effort required to build the system • weak theoretical foundation (as far as abduction is concerned) • difficult to maintain the rulebase (knowledge engineer and domain expert required)

Bibliography

[1] Ahonen, M.: *Test Data for a Fault Detection Method of a Water to Water Heat Exchanger.* Working Report, Technical Research Center, Espoo, 1994.

[2] Baer, K.; Schmidt, F.: *Knowledge Based Pre-Design of HVAC-Systems.* In: Kohonen, R.; Laitila, P. (Ed.), *INTEGRATION OF ADVANCED TECHNICAL SYSTEMS, Future Buildings Forum, Workshop*, Espoo, Finland: Technical Research Centre of Finland, June 15-17, 1994.

[3] Harmon, P.; King, D.: *Expert Systems - Artificial Intelligence in Business.* : John Wiley & Sons, 1985.

[4] Hinkelmann, M.; Kopf, M.: *Proceedings of ECPPM '94 - Building management data aquisition using a product model based design process.* In: Scherer, R. (Ed.), *Product and Process Modelling in the Building Industry*, Dresden: Balkema, Rotterdam, 1995. ISBN 90 5410 584 8.

[5] Madjidi, M.: *Characteristic Curves.* In: Hyvärinen, J. (Ed.), *Building Optimization and Fault Diagnosis Source Book, IEA Annex 25 Report*, Espoo: Technical Research Centre of Finland, 1995. 4.3.2.

[6] Rossi, T.; Braun, J.: *Classification of fault detection and diagnostic methods.* In: Hyvärinen, J.; Kohonen, R. (Ed.), *Building Optimization and Fault Diagnosis System Concept, IEA Annex 25*, Espoo: VTT Laboratory of heating and ventilation, 1993. 42–46, ISBN 952-9601-16-6.
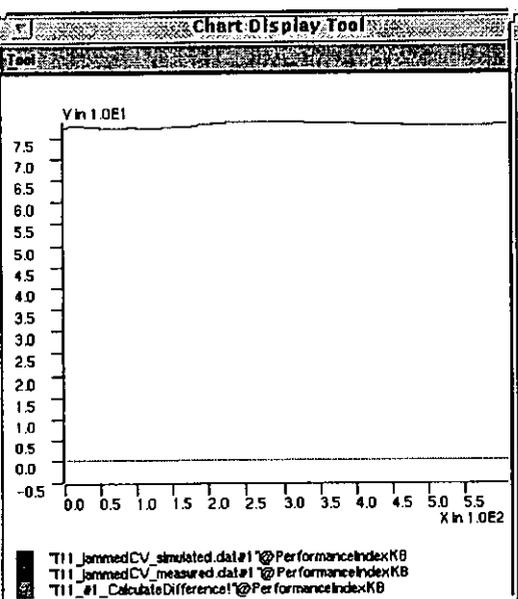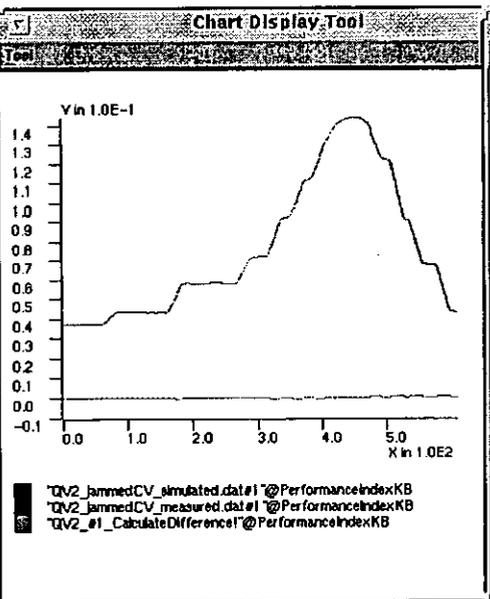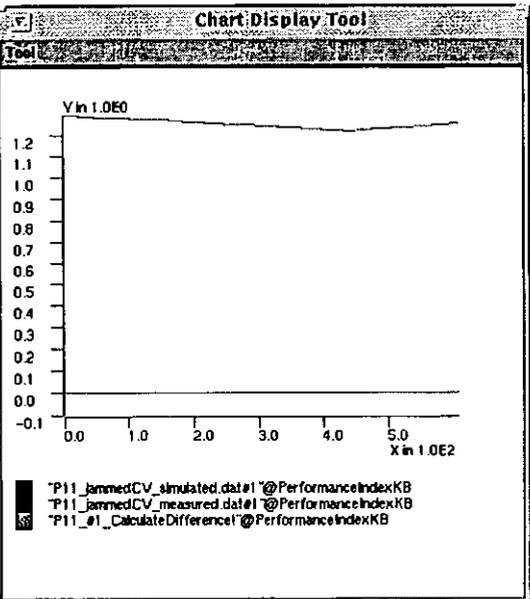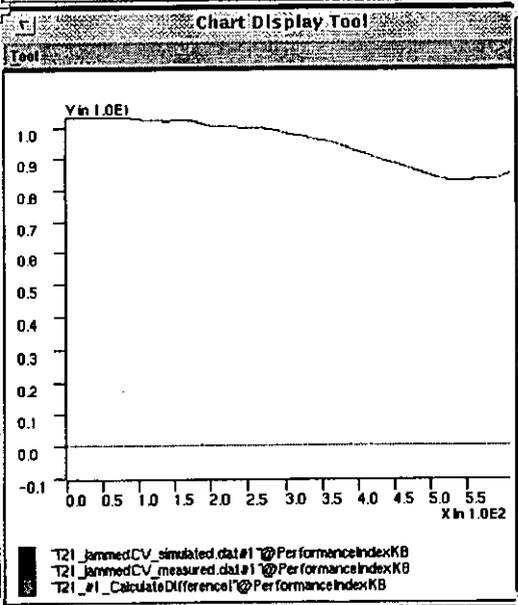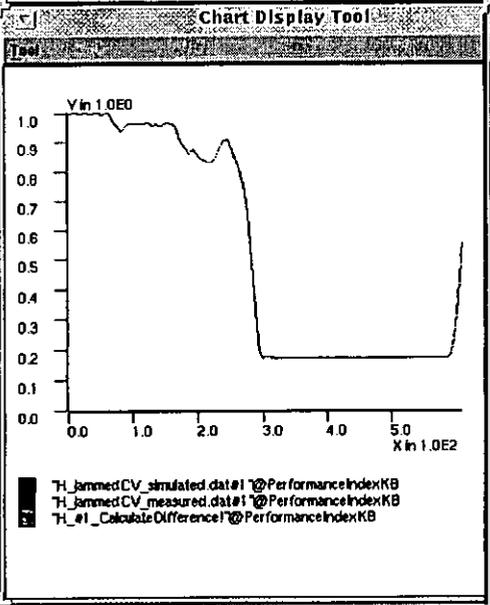
Figure 12: Temporally development of the model input parameters for the test case 'jammed control valve'

493
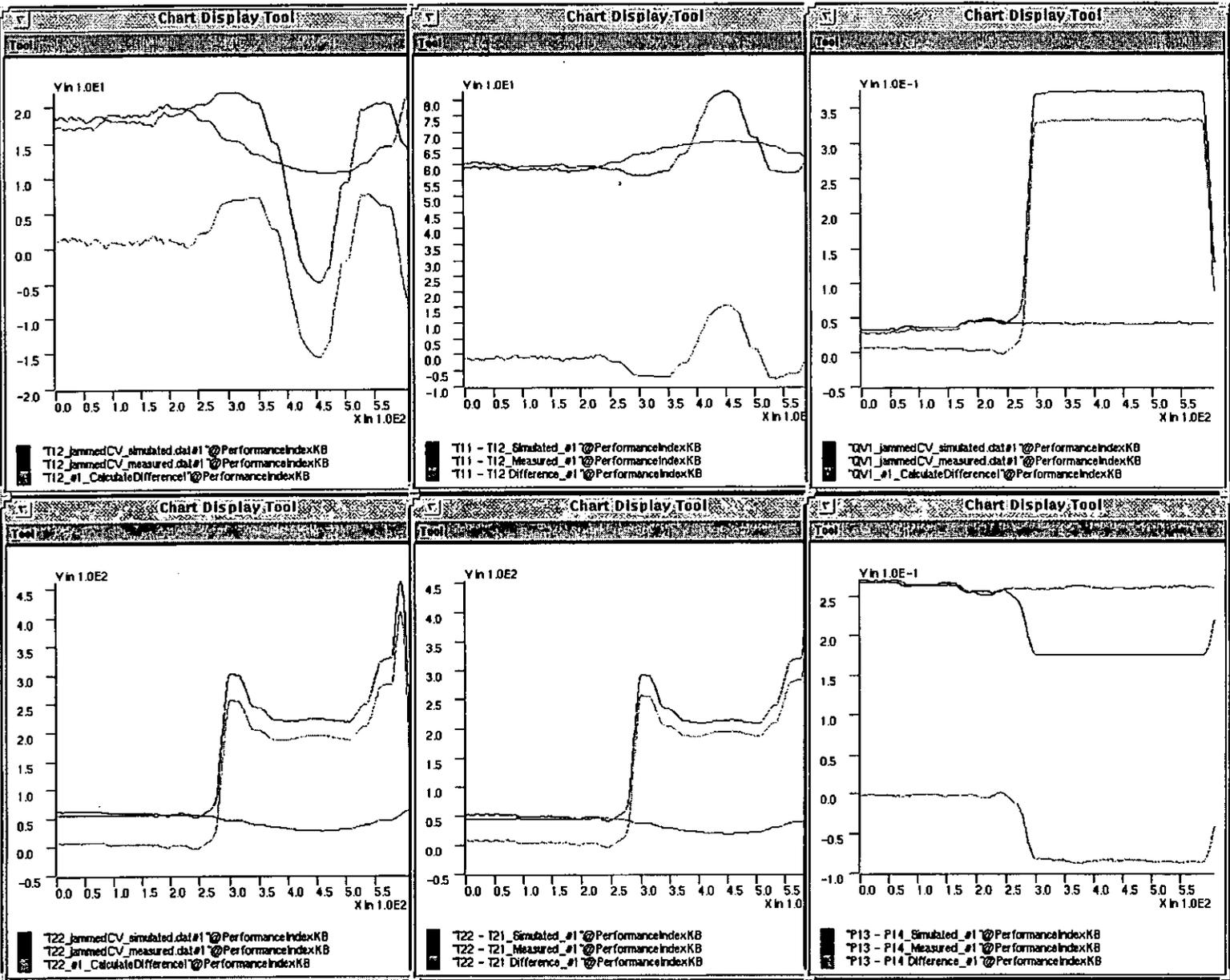
Figure 13: Temporally development of the model output parameters for the test case 'jammed control valve'

	TresholdAbsolutMax	TresholdAbsolutMin	TresholdDifferenceMax	TresholdDifferenceMin
P14 #1	1.5	0.8	0.01	-0.01
P13 #1	1.5	0.8	0.005	-0.005
P12 #1	1.5	0.8	0.005	-0.005
P11 #1	1.5	0.8	0.005	-0.005
QV2 #1	1.0	0.0	0.01	-0.01
QV1 #1	1.0	0.0	0.01	-0.01
T22 #1	90	10	10	-10
T21 #1	15	5	10	-10
T12 #1	90	10	2.5	-2.5
T11 #1	90	60	2.5	-2.5

Figure 14: Adjustment of the thresholds of the instantiated PIs

	T12 #1	T22 #1	QV1 #1	P13 #1	P14 #1	P13 - P14 #1	T22 - T21 #1	T11 - T12 #1
Branch	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
Components	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
Correlated Symptom	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
Data Compare Instance	#[Instance:	#[Instance:	Dr[Instance:	#[Instance:	#[Instance:	P13 - P1#[Instance:	T22 - T2#[Instance:	T11 - T12 DC #1]
Data Compare Method	Dir."Calculate	Dir."Calculate	Dir."Calculate	Dir."Calculate	Difference "Calculate	Difference "Calculate	Difference "Calculate	"Calculate Difference!
Data Source Column Number	2	4	5	9	10	PrkNull	PrkNull	PrkNull
Data Source Measured	uttgart.de/user/fabi1257/.dev/inhead/emmed CV: measured.dat					PrkNull	PrkNull	PrkNull
Data Source Simulated	uttgart.de/user/fabi1257/.dev/inhead/emmed CV_simulated.dat					PrkNull	PrkNull	PrkNull
IsPrognosedViolation	True	PrkNull	PrkNull	PrkNull	True	False	PrkNull	True
Predecessor	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
Process Variable	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
ResultVector	Difference[]	Difference[]	Difference[]	P1_Cal[Instance: P13	P1#[Instance: T22	T2#[Instance: T11	T12 Difference _#1]	
Sensor	tance#[Instance:	Dr[Instance: CFPHE	#[Instance: CFPHE_Pr	#[Instance: CFPHE_Tem	pance: CFPHE_TemperatureSensor#2]			
Successor	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
Timestep Position	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull	PrkNull
TresholdAbsolutMax	90	90	1.0	1.5	1.5	0.05	90	90
TresholdAbsolutMin	10	10	0.0	0.8	0.8	0.1	0	50
TresholdDifferenceMax	2.5	10	0.01	0.005	0.01	0.01	10	2.5
TresholdDifferenceMin	-2.5	-10	-0.01	-0.005	-0.01	-0.01	-10	-2.5
TresholdViolationExpression	TooHigh	PrkNull	PrkNull	PrkNull	TooHigh	TooLow	PrkNull	TooLow
TresholdViolationTimeStep	271	PrkNull	PrkNull	PrkNull	271	271	PrkNull	271
TresholdViolationValue	4.00831	PrkNull	PrkNull	PrkNull	0.012604	-0.010755	PrkNull	-4.01086

Figure 15: State of the monitored PIs in the moment of a threshold violation

Fault detection of a subprocess consisting of flow route and control valve

Juhani Hyvärinen
VTT Building Technology
P.O. Box 1804 (Lämpömiehenkuja 3)
FIN-02044 VTT (Espoo, Finland)

1. ABSTRACT

A fault detection method is described and applied to two test cases. The one case is typical for district heating subdistribution systems where mass flow of heating water is controlled with an exponential control valve. The other case is the oil feeding subprocess of an oil burner. There, the oil mass flow is controlled with a linear control valve. The fault detection method is based on simple first principles physical models and the processes are considered to be in steady state. The detection occurs when the model output deviates from the measured behaviour of the process. The diagnosis can be based on the model parameter values and their deviations from the initial values. The method does not suggest any specific diagnosis strategy but provides quantities that can be used in diagnosis.

The model parameters are estimated using least squares estimation algorithm. The threshold values for the fault detection are calculated statistically. In the case of the exponential control valve the unlinearity is solved by using the characteristic curve of the process. The characteristic curve is determined from the tuning phase data. The structure of the unlinearity is known and the parameters are estimated with least squares algorithm.

In the paper, the model parameter estimation is carried out with off-line data but especially in the case of linear control valve recursive estimation is possible. The process models and the method are simple and computationally effective so that they could be utilised in the near future in field level application specific controllers in real time applications.

2. INTRODUCTION

In this paper, a method is described that was developed as a part of technical work carried out for developing fault detection methods /1, 2/ for two typical heat production units: district heating subdistribution system /3/ and oil burner /4/. The method is based on a simple physical first principles static process model and on the use of analytical redundancy /5/ of the system under consideration. In other words, the redundancy contained in the static relationships among the system inputs and measured outputs is exploited for fault detection. The procedure of evaluation of the redundancy given by a mathematical model of some system equations can be roughly divided into two steps: generation of residuals, and decision and isolation of the faults. Here the residuals are generated based on static parallel model approach, and on parameter estimation approach in which the consistency of the mathematical equations of the system is checked by using actual measurements. Because of the static process model, direct redundancy relations can be used. The decision is based on a simple threshold. The cause of the fault can be diagnosed based on parameter estimates of the process.

The main emphasis in the development work has been on simplicity and on practical rather than on theoretical aspects of the method. The reason for this is that it is assumed that in the near future field level application specific controllers will be capable of simple diagnostic tasks but with limited computational recources only. For example, the computational speed, numerical and measurement accuracy, and amount of memory are limited for practical and, in the end of the day, for economical reasons.

Practical application of a fault detection method requires that the model structure is known and its parameters can be estimated reliably. Also it is required that the threshold values and parameter deviations from their nominal values for fault detection can be determined before the method is taken into use. For these purposes and especially for processes with linear control valve least squares estimation algorithms provide a good methodology and applying the method is quite straightforward. For processes with exponential control valve also the unlinearity must be modelled before the required model parameters and threshold values can be solved.

3. FAULT DETECTION METHOD

3.1 OUTLINE OF THE PROCESS AND THE METHOD
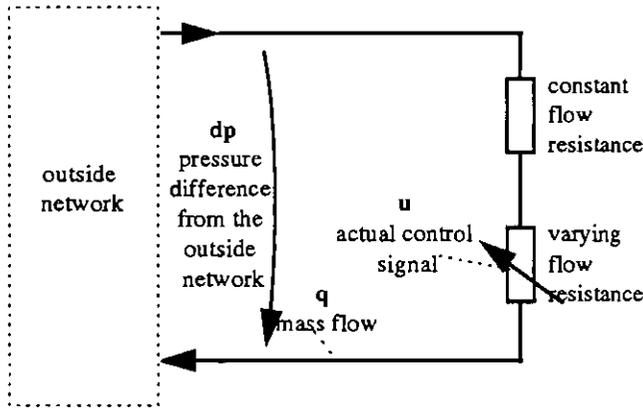
Figure 1. Electrical circuit analogy of the flow route and control valve.

The simplified process can be approximated to consist of control valve and a flow route the flow of which the control valve affects (Figure 1). The control valve represents a varying flow resistance. The flow route consists of a pipeline and other components, and it represents a constant flow resistance. The process is connected to an outside network which generates the pressure difference over the process. In an oil burner the pressure is generated with an oil pump in the burner, and in a district heating subdistribution system the distribution network can be seen as a constant pressure source.

The control valve can be described as a flow resistance the value of which changes as a function of actual control signal. The structure of the function of the varying flow resistance is usually known for each specific valve and it is typically characterised either by exponential or linear valve equation. In case of water or oil as media the process is fast and it can be considered to be in a steady state condition all the time. The process dynamics need not to be considered.

The basic idea of the method is to use a process model that relates the three variables: pressure difference dp, mass flow q, and actual control signal u. Process faults can then be detected either by calculating one of the process variables from the others, and comparing the calculated value to actual measured value, or by estimating the process model parameters and comparing them to nominal values. The threshold values and nominal parameter values are estimated from a non faulty process condition.

In the following, the two applications are described. In case of an oil burner the control valve is linear and in case of a district heating subdistribution system the control valve is exponential. For both applications, the process is described first briefly, then the process model is derived and last the applied method is described.

3.2 OIL FEEDING SUBPROCESS OF AN OIL BURNER

The subprocess under consideration is a part of the oil feeding line of an oil burner /4/. An oil pump (314) sucks oil from the oil tank and maintains constant pressure in the pipe (315) leading to the nozzle (317, 318). The pressure at the nozzle is controlled with a control valve in the return pipe line. Low pressure at the nozzle causes low mass flow to the furnace and high pressure causes high mass flow. Low and high pressure is gained with control valve position open and closed respectively.

The subprocess can be approximated to consist of a control valve, pipeline including constant flow resistances, and nozzle (figure 2). The pipeline consists of those pipes, curves, and components that are in the oil feeding line between the two pressure sensors PI1 and PI2 in figure 2.

The oil pump represents a pressure generator which generates a constant pressure over its ports. The analogy to electrical circuits is presented in figure 3.

498
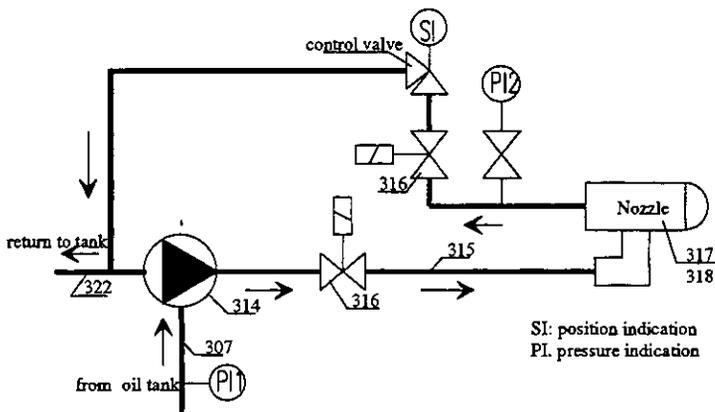
Figure 2. Oil burner oil feeding subprocess.

Figure 3.   Electrical circuit analogy of an oil pump - control valve -subprocess.

Because the oil pump generates a constant pressure there is no need to measure it directly. The pressure drop over the constant resistances of the feeding and return lines can be used to represent the oil mass flow. Thus the pressure difference (PI2-PI1) over the pump and the pipelines can be used as the output of the process. The control variable, i.e. the position of the control valve (SI), is used as the input of the process. If the oil mass flow was measured it could be used as an additional redundant information.

During normal operation the temperature of the oil can be considered as a disturbance. The effect of the oil temperature is assumed to be small. In a fault situation faults can be considered as disturbances. These kind of faults are malfunctioning of the valve, and the change of constant flow resistances due to fowling or blockage of pump, filters or pipeline components. Also big changes in oil viscosity can be considered as faults. The viscosity change may be an effect of too low oil temperature.

3.2.1 Process model

In figure 4 the principal of composing the process characteristic curve is presented. The characteristic curve between input u and output dp is composed from the characteristic curves of the pump and the control valve. The figure shows also a reason why the oil flow measurement is not necessarily needed.
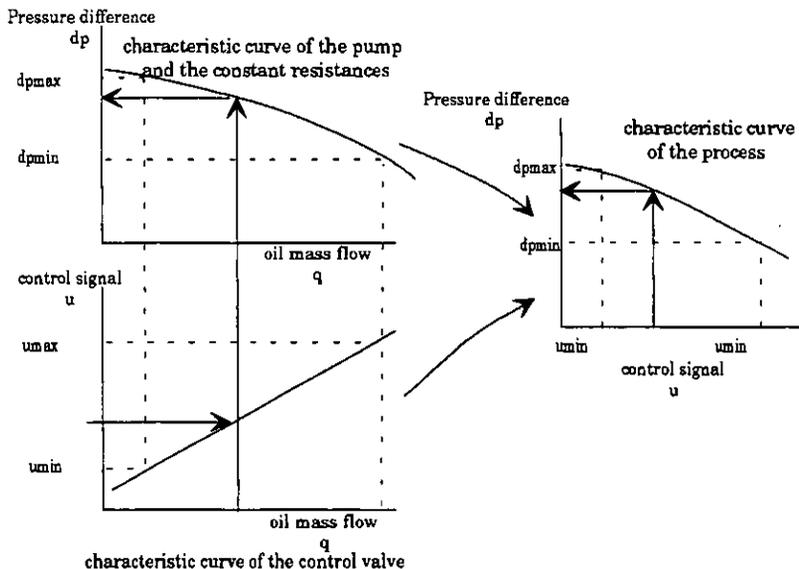
Figure 4.   Composing the characteristic curve of the subprocess.

The characteristic curve of figure 5 is composed from the process measurements and it describes the process operation.
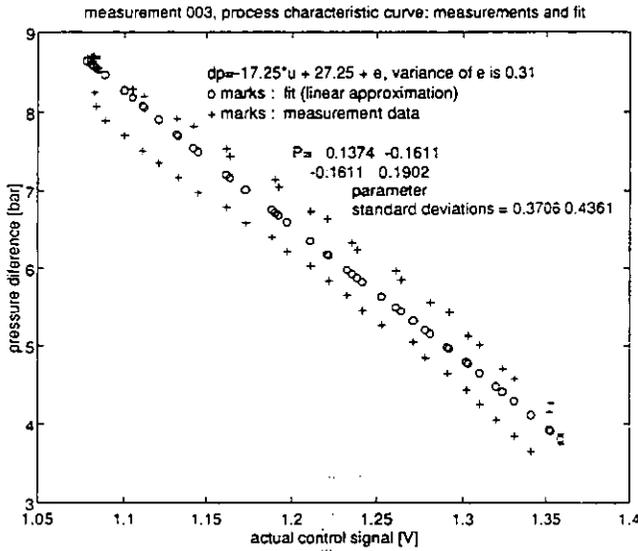
Figure 5.   The characteristic curve of the process. The measurements (+ marks) and the linear approximation (solid line).

The oil pressure difference over the pump is drawn as a function of the control signal u. The curve is almost linear and can be approximated with a polynomial form of

$$dp = P2 - P1 = c_1 * u + c_2 \qquad (1)$$

and is used to model the process.

The process quantities that are needed for modelling the oil feeding subprocess described above are:

quantity	computation/ measurement	unit	meaning
dp	PI1-PI2	kPa	pressure difference
u	TV	V	valve shaft position

The measured values must be instantaneous values. The sampling time does not play any role. It can be of varying length and chosen arbitrarily. The parameters c_1 and c_2 are estimated from the measurements.

3.2.2 Method description

Process faults can be detected either by using the model output error (figure 6) or the parameter deviations from their nominal values (figure 7). The model output error is the residual between the pressure difference given by the model (equation 1), and the actual measured value. If the parameter deviations are used for fault detection the parameters of equation 1 are estimated on line and compared to those estimated during tuning phase.
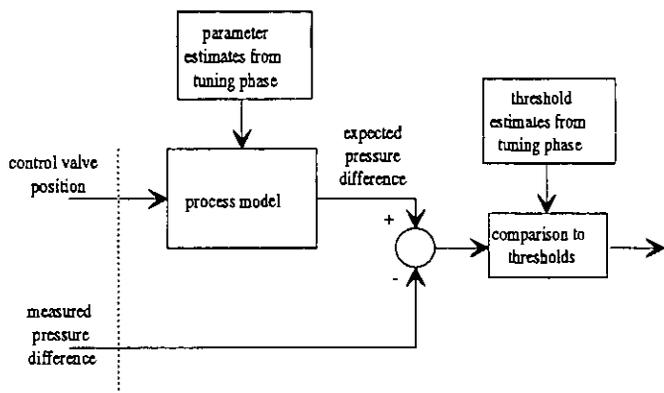
Figure 6.   Model output error in fault detection.
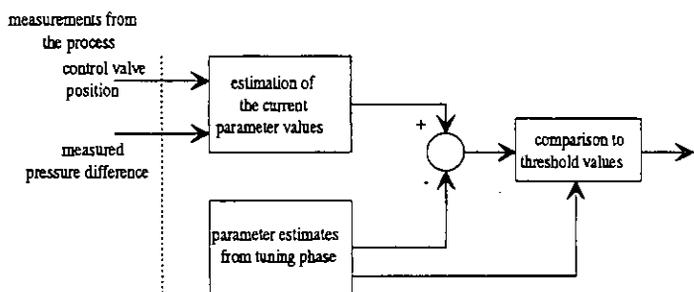
Figure 7.   Parameter deviations in fault detection.

When model output error is used the pressure difference over the oil pump given by equation 1 is compared to the corresponding pressure difference measured from the process (figure 6). The parameter values of equation 1 and the threshold values for fault detection are estimated during the tuning phase and kept fixed during the operation phase.

When parameter deviations are used for fault detection the parameters of the equation 1 are estimated periodically or recursively in real-time and the estimates are compared to values estimated during the tuning phase (figure 7).

The method is tuned following the steps below:

1. Measuring the tuning data

2. Prehandling of the measurement data

3. Parameter and threshold value estimation

It is important that the input signals to the process vary in a large enough area i.e. the input signals must be exciting enough. It would be best if all the normal operation points could be measured during the tuning phase for a long enough time. The considered process is nearly linear and for that reason already a few operating points should give a good result. The process is assumed to be in non-faulty condition during the measurement of the tuning data.

The control signal of the oil feeding valve can be varied quite freely. One has to take care that no harm is caused to the other parts of the process or user. For example, the burner heat effect may not be increased too high in order to avoid damages to the boiler. The control signal should be varied between 0 % and 100 % if possible. The changes should be slow enough so that the valve actuator and real position of the valve have enough time to reach each operation point.

During the tuning and operation phases unfiltered measurement data is utilised. No filtering or such is needed for prehandling of data but if possible the so called outliers should be removed and also those operating points where the process has clearly operated erroneously should be removed. For example, near open and close situations the process may saturate due to erroneous tuning of the actuator causing erroneous information of the process operation.

Equation 1 can be presented in the following form

$$ y = c_1 * x_1 + c_2 * x_2 + e \hspace{4cm} (2) $$

where
$y=dp$, $x_1 =u$, and $x_2 =1$ are signals
c_1 and c_2 are parameters the covariance of whose is P
and e is assumed to be a zero-mean white Gaussian noise the variance of which is λ^2

The parameter values c_1 and c_2,, their covariance matrix P, and the variance λ^2 of the white noise e are estimated using least squares method /6/. The estimation result is better if the averages are removed before the estimation. This, however, requires off-line estimation method to be used. The parameters can be solved by using some recursive estimation method in which case the estimation can be done in real-time. The initial values needed for recursive methods are gained for example through off-line methods.

Analytically, it is not possible to choose a threshold value for fault detection that gives 0 % false alarm rate. The thresholds are chosen in such a way that an alarm does not necessarily mean a fault but may be caused by a normal measurement noise, too. After the alarm event, it must be decided if the alarm is caused by a fault.

501

The model output error is the residual between process model and corresponding measurements. When used in fault detection the threshold value calculated from the variance λ^2 of the noise e is used to test this residual. Threshold value is chosen as in the following equation

$$\bar{e} = k * \lambda \qquad (3)$$

in which the coefficient k defines the confidence limits for the model output according to table 1.

Table 1. Confidence limits.

coefficient k	confidence limit [%]
1,65	90
1,96	95
2,58	99

In the same way, the threshold values for the deviations of the parameter c estimates are calculated from the diagonal elements of the covariance matrix P. Assuming these deviations to be normally distributed their thresholds are

$$\bar{c_i} = k * \sqrt{P_{ii}} \qquad (4)$$

where

$\bar{c}_i$ = threshold for $\hat{c}_i - c_i^0$

$\hat{c}_i$ = parameter c_i estimate

$\hat{c}_i^0$ = estimate from tuning phase, assumed to be the true value

P_{ii} = i^{th} diagonal element of matrix P

P = $\text{cov}\left(\hat{c} - c^0\right)$

and coefficient k (table 1) defines the confidence limits for the parameter estimates.

The output error can be used to detect the following faults:

1. block in the oil flow route

2. pump faults

3. control valve faults

Applying the method requires reasoning in two ways. Firstly, the way the threshold values are selected does not guarantee 0% false alarm rate and thus the decision of the fault situation can be made better if the decision function is based also on other properties, e.g., such as presented in /7/.

Secondly, the estimation of the model parameters produces a set of test quantities the changes of which indicate some faults in the process. Parameter estimates give some additional information of the process. The changes in the two parameters of equation 2 estimated from the measurements represent 9 different fault modes: either one of the parameters can increase, decrease or remain constant independently of the other parameter. If both parameters remain constant, the process is in non-faulty situation. Any other state out of the 9 possible states may mean some fault.

The nine different states are shown in table 2. The faults leading to different states must be located by simulation, experiments, or some other methods. By studying the effect of different faults to the characteristic curves it can be derived what is the effect of each fault to the parameter estimates of c_1 and c_2. The examples in table 2 are not verified by experiments and are the result of qualitative reasoning only.

Table 2.   The effect of the faults to the values of the parameters.

		parameter c_1		
		-	0	+
para-meter c_2	+		atomiser worn	
	0	oil filter is missing, nozzle ballast worn, valve fault	no faults	filter blocked, nozzle ballast blocked, valve fault
	-		atomiser blocked	pump worn

3.3 DISTRICT HEATING SUBDISTRIBUTION SYSTEM
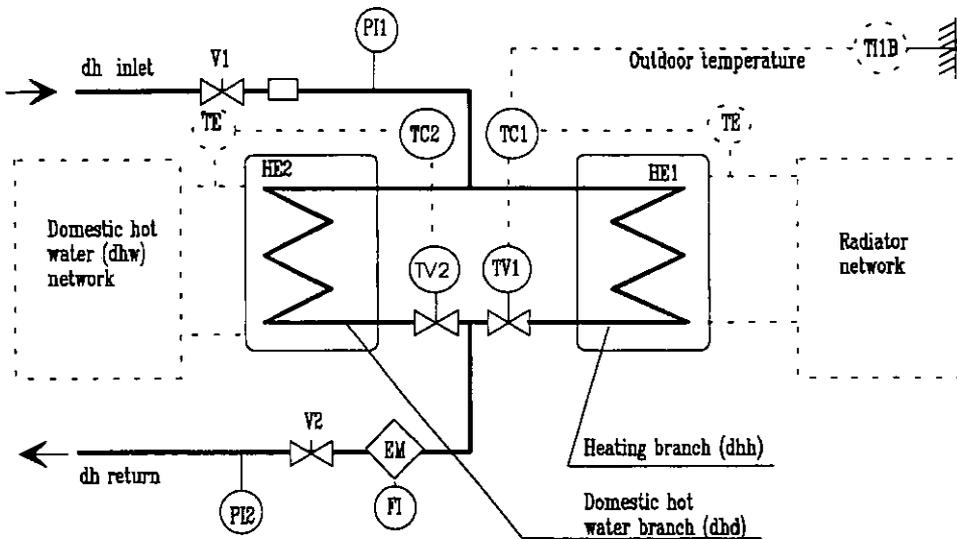
Figure 8.   District heating subdistribution system.

The subprocess studied is a part of the district heating subsystem described in /3/ and is located in the primary side of the district heating subdistribution system (DHS) both in the heating branch and domestic hot water branch. In figure 8 a rough diagram of the subdistribution system is shown. The flow route in consideration is drawn with solid thick line. It is assumed that only one of the two branches is in use at a time. For the domestic hot water branch this is the usual case during a summer period when no heating is needed and the heating branch is shut down with manual shut-off valves. The domestic hot water branch is used always but there are some moments when the domestic hot water demand is very low. For example during night there are moments and longer periods of time when it can be said that only heating branch is operational.

The simplified subprocess can be approximated to consist of a control valve, pipeline including constant flow resistances, and heat exchanger. The analogy to electrical circuits is shown in figure 9. The pipeline consists of those pipes, curves, and components that are in the line between the two pressure measurement sensors.

The district heating network provides the pressure over the flow route. The pressure difference between inlet pipe and return pipe remains steady for short time periods but varies because of seasonal conditions. During winter when energy need is high the inlet pressure is typically higher than in summer. The massflow is measured for energy billing purposes and can be used for other purposes in modern automation systems. Also the actual control valve position is usually measured or known. During normal operation the temperature of the water can be considered as a disturbance. The effect of the water temperature is assumed to be very small. The faults to be detected are malfunctioning of the valve, and the change of constant flow resistances due to fowling of heat exchanger or pipeline.
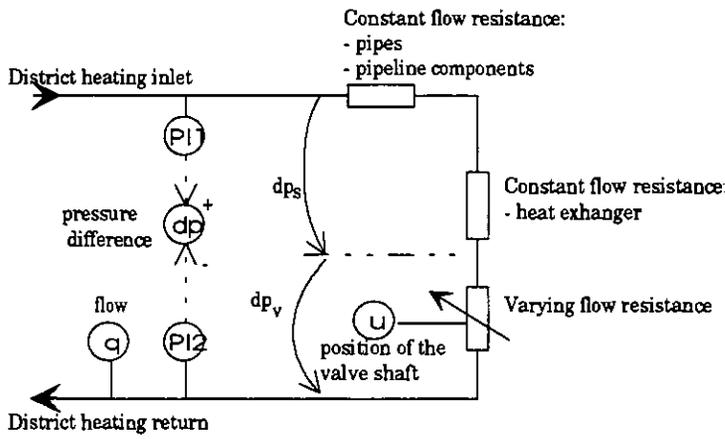
Figure 9.   Simplified diagram of the subprocess.

3.3.1 Process model

The flow and pressure equations of the simplified subprocess are

$$q^2 = kv^2 * dp_v \qquad (5)$$

$$dp = dp_s + dp_v \qquad (6)$$

$$dp_s = Z_s * q^2 \qquad (7)$$

where Z_s is constant flow resistance of the heat exchanger and pipeline

The dependency of the valve kv-value from the shaft position in general can be modelled with the following equation

$$kv = kv_c * g(u) \qquad (8)$$

where the function g(u) describes the valve unlinearity.

Applying equations 6, 7 and 8 to 5 the following equation is obtained

$$\underline{q^2} = kv_c^2 * \underline{g(u)^2 * dp} - kv_c^2 * Z_s * \underline{g(u)^2 * q^2} \qquad (9)$$

and further when arranging for g(u)²

$$g(u)^2 = \frac{q^2}{kv_c^2 * dp - kv_c^2 * Z_s * q^2} \qquad (10)$$

Taking a derivative of 9 with respect to u and supposing that dp does not change as a function of u, the following equation is obtained.

$$\frac{\partial q^2}{\partial u} = kv_c^2 * dp * 2 * g(u) * g'(u) - kv_c^2 * 2 * g(u) * g'(u) * Z_s * q^2 - kv_c^2 * g(u)^2 * Z_s * \frac{\partial q^2}{\partial u} \qquad (11)$$

504

If the valve unlinearity is characterised with the following equation

$$g'(u) = C * g(u) \tag{12}$$

the equation 11 can the be presented as below

$$\frac{\partial q^2}{\partial u} = kv_c^2 * dp * 2 * C * g(u)^2 - kv_c^2 * 2 * C * g(u)^2 * Z_s * q^2 - kv_c^2 * g(u)^2 * Z_s * \frac{\partial q^2}{\partial u} \tag{13}$$

By applying equation 10 to 13 and solving it with respect to dp*d(q²)/du the following is obtained

$$dp * \frac{\partial q^2}{\partial u} = 2 * C * \underline{q^2 * dp} - 2 * C * Z_s^2 * \underline{q^4} \tag{14}$$

Now, equations 9 and 14 form the equation set that models the subprocess. In case of a linear valve is utilised instead of an unlinear one, equation 9 alone forms the model. The underlined parts represent signals that are measured and calculated from process signals and the other parts represent model parameters that are estimated during tuning and operation phase.

Unlinear valves that fulfils the requirement of equation 12 are, for example exponential valves characterised with following equations.

$$g(u) = e^{n_e * u} \qquad\qquad C = n_e \tag{15}$$

$$g(u) = \left(1 - \frac{n_p}{100}\right)^{(100-u)} \qquad\qquad C = \ln\left(1 - \frac{n_p}{100}\right) \tag{16}$$

where n_p and n_e are manufacturer specific parameters used to describe the unlinearity.

Finally the subprocess model is

$$q^2 = \frac{kv_c^2 * g(u)^2 * dp}{1 + kv_c^2 * Z_s g(u)^2} \tag{17}$$

The process quantities that have to be measured are:

quantity	computation/ measurement	unit	meaning
dp	PI1-PI2	kPa	pressure difference
q	FI	m³/h	water volumetric flow
u	TV	%	valve shaft position

Often, the valve position is not measured and only the control signal value is known. This causes extra work during the tuning of the method because the valve shaft position as a function of control signal must be modelled. The control signal value from the controller can not be used in place of shaft position because the control signal changes stepwise and position rampwise and there is a difference between these two signals. The valve position must be used instead.

The measured values must be instantaneous values and they may not be filtered or averaged in any way. The sampling time does not play any role. It can be of varying length and chosen arbitrarily. The signals (underlined terms) in equation 9, however, can be filtered with any linear filter. They can, for example, be cumulated over some time interval.

3.3.2 Method description

Process faults can be detected either by using a model output error (figure 10) or parameter deviations from their nominal values (figure 11). The model output error is the residual between the model output, q^2, given by equation 17, and the actual measured value. If the parameter deviations are used for fault detection the parameters of equation 9 and 14 are estimated periodically and compared to those estimated during tuning phase.

The square of flow given by equation 17 is compared to the square of the flow measured and calculated from the process. The value of the function $g(u)$ is calculated using equation 15 or 16. The parameter values of equations 9 and 14 and the threshold values for fault detection are estimated during the tuning phase and fixed during the operation phase.

Figure 10.   Block diagram of the output error method.

When parameter deviations are used for fault detection, the parameters of the process model (equation 9) are estimated periodically and the results are compared to values estimated during the tuning phase. The parameters of the model describing valve unlinearity (equation 14) are estimated only during tuning phase and are fixed during normal operation period.

If the subprocess is linear i.e. there is no such a component that causes some unlinear feature like the unlinear valve does, then the parameters can be estimated on-line instead of periodical estimation.

Figure 11.   Block diagram of the parameter error method.

The method is tuned following the steps below
1. Measuring the tuning data
2. Prehandling of the measurement data
3. Parameter and threshold value estimation

It is important that the input signals to the process vary in a big enough area i.e. the input signals must be exciting enough. It would be best if each of the normal operation point could be measured during the tuning phase for a long enough time. For practical reasons this is usually not possible and the tuning must be done with less information. The process is assumed to be in non-faulty condition during the measurement of tuning data. The method is "taught" the non-faulty operation of the process during tuning phase. When, at the operation phase, the measured process operation deviates from the taught (modelled) operation, it can be assumed that there is a fault in the process.

In the DHS, the valve control signal can be varied freely. One has to take care that no harm is caused to the other parts of the process or user. For example, the domestic hot water temperature may not rise too high if the water is used during the tuning period. The control signal should be varied between 0 % and 100 % if possible. The changes should be slow enough so that the valve actuator and real position of the valve have enough time to travel to each operation point.

The pressure difference over the subprocess can not be affected by the operator. The pressure difference varies according to the heating load of the external district heating network. At a minimum, one should get measurement data for one pressure difference value over all the valve position values.

During the tuning and operation phases only unfiltered raw data are utilised. Outliers should be removed if possible. Also those operation points where the process has clearly operated erroneously should be removed. For example, in a near open and close situation the process may saturate due to erroneous tuning of the actuator causing erroneous information of the process operation.
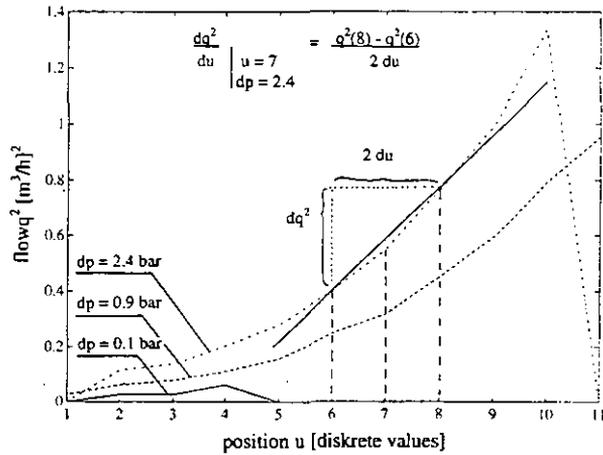
Figure 12. Process characteristic curves and calculation of partial derivative of q^2 with respect to valve shaft position.

In equation 14 the output signal (regressed variable) is the partial derivative of the square of flow with respect to the valve position. Solving this on-line from measurements is sensitive to measurement noise which causes the result to be unreliable. For this reason the partial derivative is calculated off-line from a characteristic curve of the process.

A characteristic curve is formed from the process (figures 12 and 13), in which the valve position u is on the x-axis and square of the flow on the y-axis. The different curves in figure 12 are isoclines having different values of pressure difference. The partial derivative is calculated for each discretized position of u on each isocline curve using the following equation.

$$\frac{\partial q^2}{\partial u} = \frac{q^2 (u + du) - q^2 (u - du)}{2 * du}$$
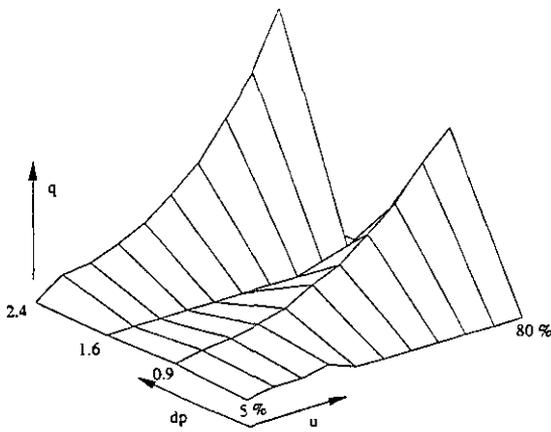
Figure 13. Process characteristic curves of figure 12 plotted as a surface (mesh-plot). The flow value remains in zero at dp value 1.6 because there has not been any measurements at that value. In reality the value of q increases continuosly with value of dp.

3.3.3 Parameter and threshold value estimation

Equation 14 can be represented in the following form

$$y = x_1 * a_1 + x_2 * a_2 + e \tag{19}$$

where

$$y = dp \frac{\partial q^2}{\partial u}, \quad x_1 = 2*q^2*dp, \text{ and } x_2 = 2*q^4 \text{ are signals}$$

a_1 = C and a_2 = -C*Z

and e is assumed to be zero-mean white Gaussian noise the variance of which is λ^2

The parameter values a_1 ja a_2 are estimated using least squares method /6/. Using parameter a_1 the parameter describing the valve unlinearity n_c or n_p can be solved.

Because the partial derivative of square of the flow is solved from a characteristic curve the parameters must be estimated off-line. Furthermore, the estimation result is better if the averages are removed before the estimation. This, too, requires an off-line estimation method to be used.

Equation 9 can be represented in the following form

$$y = x_1 * b_1 + x_2 * b_2 + e \tag{20}$$

where

$y = q^2$, $x_1 = g(u)^2*dp$, and $x_2 = g(u)^2 * q^2$ are signals

b_1 and b_2 are parameters, with covariance of P

and e is assumed to be a zero-mean white Gaussian noise the variance of which is λ^2

The parameter values b_1, b_2, their covariance matrix P, and the estimate of the variance λ^2 of noise e, are estimated using the least squares method. In this case, too, the result is better if the averages are removed from the signals before estimation.

It is not possible to choose a threshold value for fault detection that gives 0 % false alarm rate. The thresholds are chosen in such a way that an alarm does not necessarily mean a fault but may be caused by normal measurement noise, too. After an alarm occurs, it must be decided if the alarm is caused by a fault.

Fault detection method output is the residual between subprocess model and corresponding measurements. As a threshold for this residual a value calculated from the variance λ^2 of the noise e is used. The threshold values are chosen in the same way as in case of oil burner.

Applying the method requires reasoning in two ways. Firstly, the way the threshold values are selected does not guarantee a 0% false alarm rate and thus the decision of the fault situation must be done for example using a classifier presented in /7/. Secondly, the method produces a set of test quantities the changes of which indicate some faults in the process. The test quantities that can be used in addition to the residual, are the parameters a_1, a_2, b_1 and b_1.

4. RESULTS

4.1 OIL FEEDING SUBPROCESS OF AN OIL BURNER

The oil feeding subprocess was tested in a laboratory test rig. The burner was of a standard modulating type. The tests shown in figure 14 were done with the subprocess for tuning the model parameters. The measurements were used to build the characteristic curve of the process and for parameter estimation. The pressure difference is presented in the upper half of the figure. The valve shaft position is presented in the lower half of the figure.

The test takes place between the time of 45 s. and 135 s. During the test the burner effect is ramped up to the maximum value and then down to the minimum value. For tuning some data was removed as being invalid. The criteria for valid measurements was that the pressure difference is in the interval of [4 .. 9] bar and the control signal in the interval of [1 .. 1.4] V.
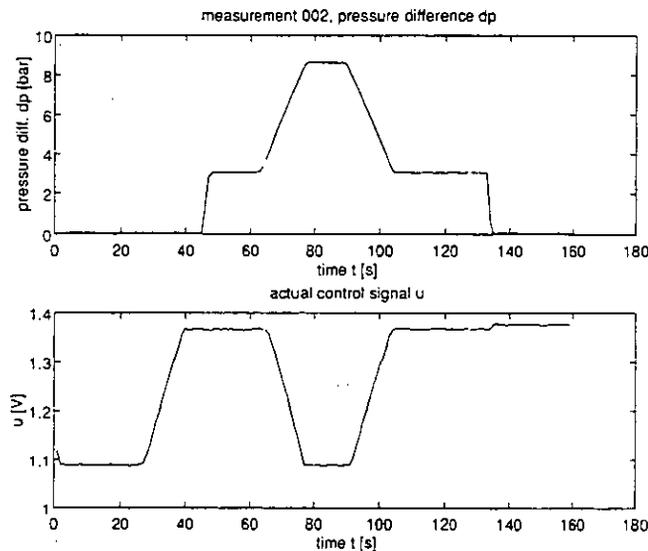
Figure 14. The tuning data for the oil feeding process. Pressure difference dp and actual control signal u.
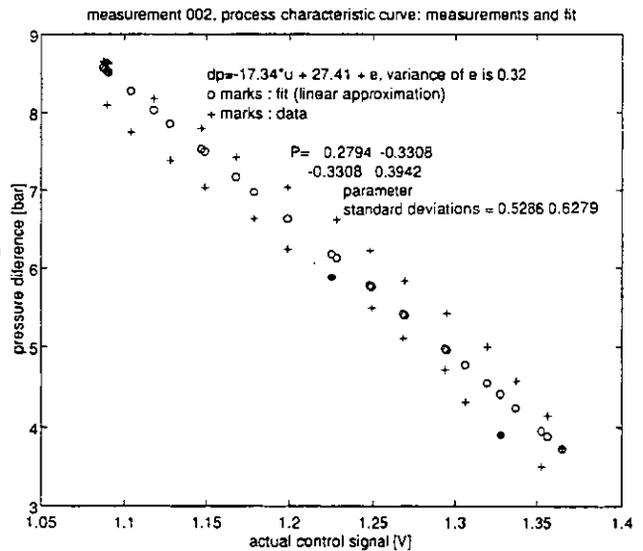
Figure 15. Linear approximation of the model constructed from the tuning phase data (o-marks). The measurement points marked with + mark.

Figure 15 shows the valid process values and the linear approximation of the process according to equation 2. The parameter values, the covariance matrix P and the standard deviations of the parameter estimates are shown as well.

The 95 % confidence limits of parameters c_1 and c_2 are respectively

$$c_1 \pm \Delta c_1 = -17.34 \pm 1.96 * \sqrt{(P_{11})} = -17.34 \pm 1.96 * \sqrt{0.28} = -17.34 \pm 1.96 * 0.53 = -17.34 \pm 1.59 \qquad (21)$$

$$c_2 \pm \Delta c_2 = 27.41 \pm 1.96 * \sqrt{(P_{22})} = -27.41 \pm 1.96\sqrt{0.39} = -27.41 \pm 1.96 * 0.63 = 27.41 \pm 1.89 \qquad (22)$$
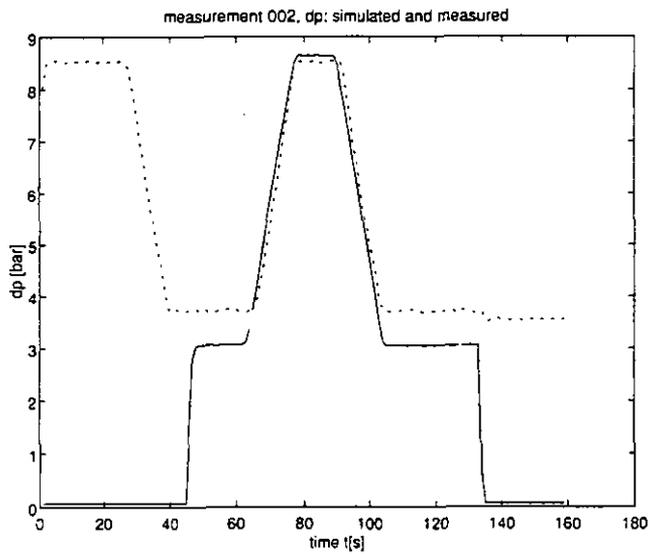
Figure 16. Simulated and measured pressure difference.

Figure 16 shows the simulation of pressure difference (dotted line) and the corresponding process measurement (solid curve). The process operating point has remained outside of the validity area most of the time. The simulated pressure has been high in the beginning of the test because the control signal has been ramped down manually before the test started.

The tests shown in figure 17 were done with the subprocess to test the fault detection method. The measurements were used to build the characteristic curve and for parameter estimation. The pressure difference is presented in the upper half of the figure. In the lower half of the figure the valve shaft position is presented
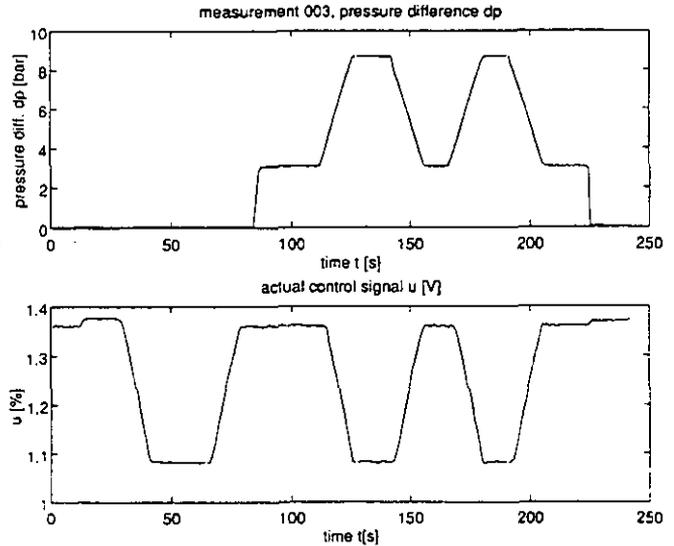
Figure 17. The data used to test the method.

In figure 19 the output of the process model constructed during the tuning phase (solid line) is compared to the measurements (plus marks) and to the output a new model (o-marks). The 95 % confidence limits of the model constructed during the tuning phase are shown with dotted lines. It can be seen that the process is operating all the time between the confidence limits, and also that the parameter values are within the confidence limits of the model parameters estimated during the tuning phase.

In figure 18 the measurements of a faulty process are compared to the output of a model of a non faulty process (figure 15). The fault has been that the oil pipe has been choked partially. It can be seen that the faulty process operates all the time outside of the confidence limits of the non-faulty process model. Also the parameter values are outside the confidence limits of the parameters of the non-faulty process.
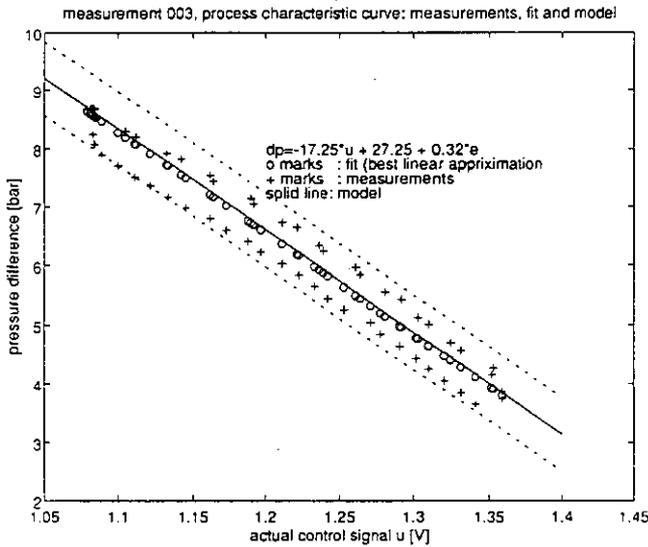
Figure 19. A linear approximation of the model from tuning phase (o-marks), model from test period (solid line and dotted lines for the confidence limit), and the measurements (+ marks).
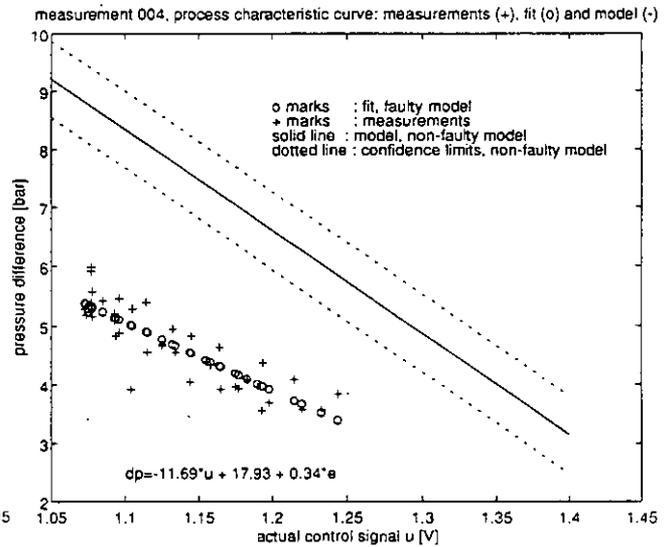
Figure 18. Process measurements(+), and the linear approximation (o-marks) of faulty process.

4.2 DISTRICT HEATING SUBDISTRIBUTION SYSTEM
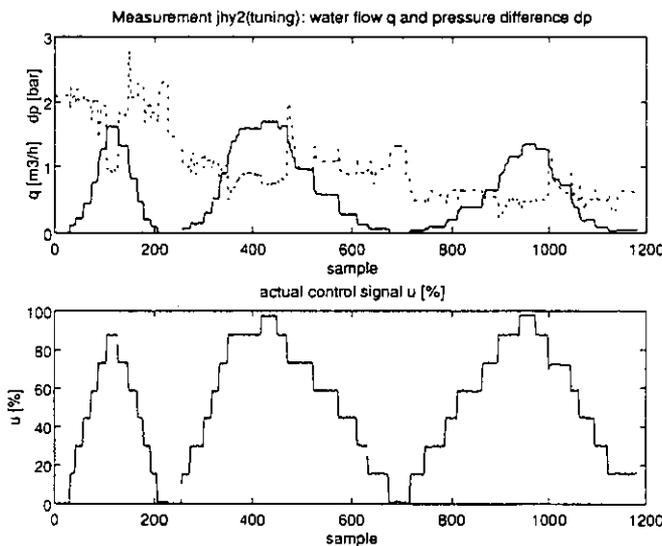
Figure 20.   DHS Tuning data.

The tests shown in figure 20 were done with the subprocess to tune the model parameters. The measurements were used to build the characteristic curve, and for parameter estimation. In the upper half of the figure both the flow (solid line) and the pressure difference (dashed line) are presented. In the lower half of the figure the valve shaft position is presented.

511

Figure 21 shows the simulated square of the flow (dashed line), the corresponding value calculated from the measurements (solid line), and the confidence limits. Only those operation points are included where the valve position has been in the interval of [5% .. 80%]. The simulation model of equations 15 and 17 has been used and parameters shown in table 3 were obtained. The error of the simulation is shown in figure 22.
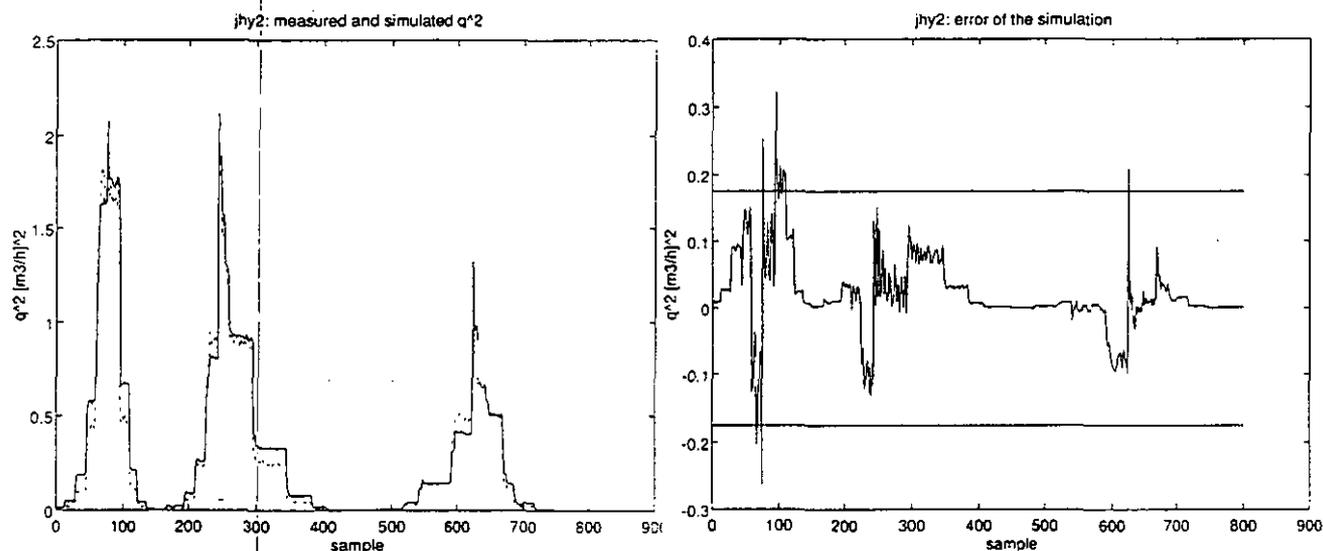
Figure 21. Simulation of q^2. Simulation done with tuning data.

Figure 22. Error of the simulation using tuning phase data.

Table 3. Parameter estimates obtained from tuning phase.

Equation 20	
b_1 b_2	1.0e-3* 0.3528 0.1054
λ^2	0.0046
P	1.0e-9* 0.65 0.67 0.67 0.71
n_e	0.0558

The results of applying the method to a non-faulty process are shown in figures 23 and 24. In 23 simulated (solid line) and measured (dashed line) q^2 are presented. In figure 24, the model residual (i.e. the error between measured and simulated process output) and its thresholds are presented. In the testing phase, only those measurements where the actual control signal remained in an interval of [5% .. 70 %] were taken into account. The other measurement values were considered to be out of the valid area of the method.

512
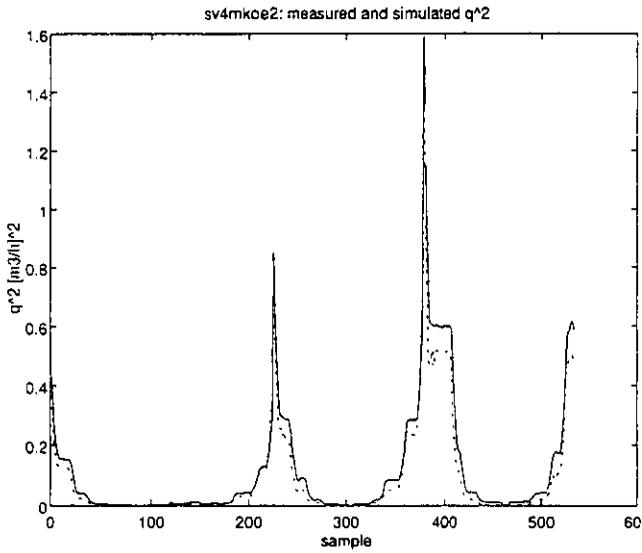
Figure 23. Simulation of q² using testing phase data.
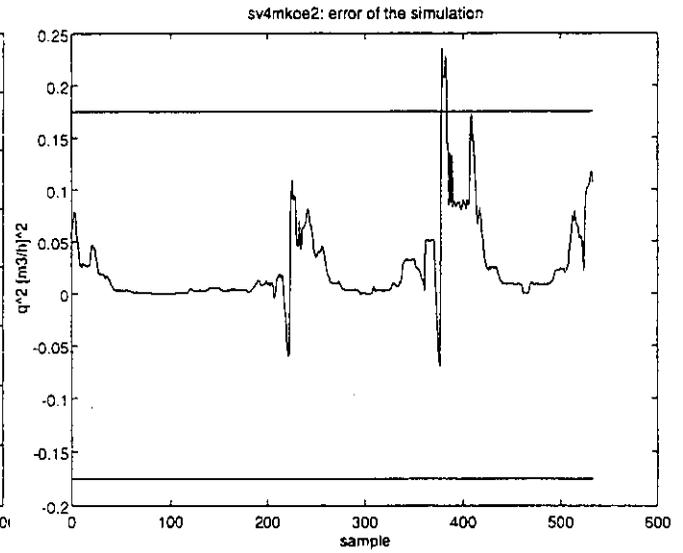
Figure 24. Residual of simulation of q². The residual is the difference between measured process value and respective simulated one.

5. CONCLUSIONS

The method operates well and can be utilised in practical applications according to test results for a non-faulty process. Fault simulations have and evaluation of the sensitivity to different faults must be carried out before practical application. The method utilises well known techniques and for that reason is easy to implement. Calculationally the method is light. The heaviest tasks in calculation are the parameter estimation and, in case of an unlinear valve, the calculation of an exponential term.

The method copes with a specific unlinear processes. The unlinearity assumed is typical in district heating systems and in HVAC applications. The unlinearity causes the performance of the method, if measured in terms of output error whiteness, to decrease, and the complexity of the method to increase. With a linear valve the method gives better results is less complex, and needs less calculations and memory.

The method does not give 0% false alarm rate but requires an additional classifier to make the decision whether the alarm is caused by a fault or not. The Method produces several test quantities, which all can be used in diagnosing the fault or in improving the fault detection. By combining the information from different test quantities the method gives relatively fault selective information of the process operation.

References

1   Hyvärinen, J. Fault detection method for district heating substation control valve based on static and physical model. IEA Annex 25. Working Paper AN25/FI/110494/1.0. Boston, 11.-13.4.1994.

2   Hyvärinen, J. Fault detection method for burner oil feeding line control valve based on characteristic curve. IEA Annex 25. Working Paper AN25/FI/260994/1.0. Stuttgart, 26.-28.9.1994.

3.   Hyvärinen, J., Karjalainen, S., Functional description of a district heating substation. In: Hyvärinen, J., Kohonen, R. (editors). Building optimisation and fault diagnosis system concept. Finland, Espoo: Technical research Centre of Finland, Laboratory of Heating and Ventilation, October 1993. pp. 142 - 148. (IEA Annex 25). ISBN 952-9601-16-6

4.   Hyvärinen, J., Tormonen, K., Siltanen, T. Functional description of an oil burner. In: Hyvärinen, J., Kohonen, R. (editors). Building optimisation and fault diagnosis system concept. Finland, Espoo: Technical research Centre of Finland, Laboratory of Heating and Ventilation, October 1993. pp. 142 - 148. (IEA Annex 25). ISBN 952-9601-16-6

5   Frank, P., M. Fault diagnosis in dynamic systems using analytical and knowledge based redundancy - A survey and some new results. Pergamon press, Great Britain. Automatica, Vol. 26, No. 3, pp 459-474, 1990.

6.   Söderström, T., Stoica, P. System Identification. University Press, Cambridge, UK: Prentice Hall International, 1989. p.60-95. ISBN 0-13-881236-5

7.   Pakanen, J. Classification of faults as a method to eliminate problems of threshold adjustment. IEA Annex 25, fourth meeting, Zürich, Switzerland, 27. - 29.9.1993. (working paper)

THE APPLICATION OF FIELD CHARACTERISTICS TO MONITOR HEAT PUMPS FOR FAULT DETECTION

Henri van Kuijk
Sulzer Energieconsulting SA
Rue de Langallerie 6
1000 Lausanne 4
Switzerland

Abstract

The present contribution shows the results of an investigation in the area of fault detection at heat pumps. It first demonstrates that the exergetic efficiency of a heat pump forms a field-characteristic capable of monitoring the performance of a heat pump plant over its full working range. This has been accomplished by analysing thoroughly one year of acquired data on a real air-to-water heat pump plant, on the one hand, and of laboratory tests performed on a water-to-water heat pump and on a chiller, on the other hand. The exergetic efficiency of a heat pump, shows a field-characteristic having two major, easy to measure, independent variables (temperature at evaporator, and temperature at condensor). It is furthermore demonstrated how this field-characteristic detects lack of performance of the monitored plants.

1. Introduction

The exergetic efficiency ε of a heat pump (HP) is expressed by the following equation:

$$\varepsilon = \frac{COP_{HP}}{COP'_{Carnot}} \qquad (1)$$

where COP_{HP} is the coefficient of performance of the HP expressed as:

$$COP_{HP} = \frac{Heat - rate\ at\ condensor}{Electrical\ power\ at\ compressor} \qquad (2)$$

and where COP'_{Carnot} is expressed as:

$$COP'_{Carnot} = \frac{T_{condensor}}{T_{condensor} - T_{evaporator}} \qquad (3)$$

In this relation (3), $T_{condensor}$ (Fig. 1, Pt 2w in heat sink medium) and $T_{evaporator}$ (Fig. 1, Pt 4w in heat source medium) are both measured at water side, which was easier to implement without losing any information. The real condensor and evaporator temperatures are more costly to measure and would not provide more pertinent values.

The first part of this investigation, was to study the behaviour of ε over the whole working range of the HP. The objective was to qualify the property of ε of being a stable function of COP'_{Carnot}.
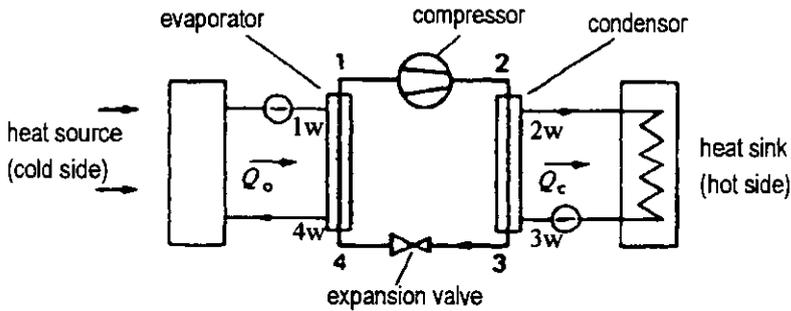
Fig. 1. Schematic representation of a heat pump

2. Acquisition of data

To analyse the behaviour of ε over the whole working range of heat pumps, measurements were performed on laboratory heat pump and chiller rigs as well as on real a HP plant. The laboratory equipment provided accurate performance data over the whole working range of heat pumps, whereas the real HP plant provided data in real conditions used to verify the 'theory'.

2.1 Data acquired on the ESH plant

In Winterthur, one of the Sulzer buildings, namely the Energiesparhaus (ESH), is equipped with a fully instrumented air-to-water heat pump. This real HP plant was used to acquire the needed operating data for this study. Fig. 2 displays a schematic representation of this plant.
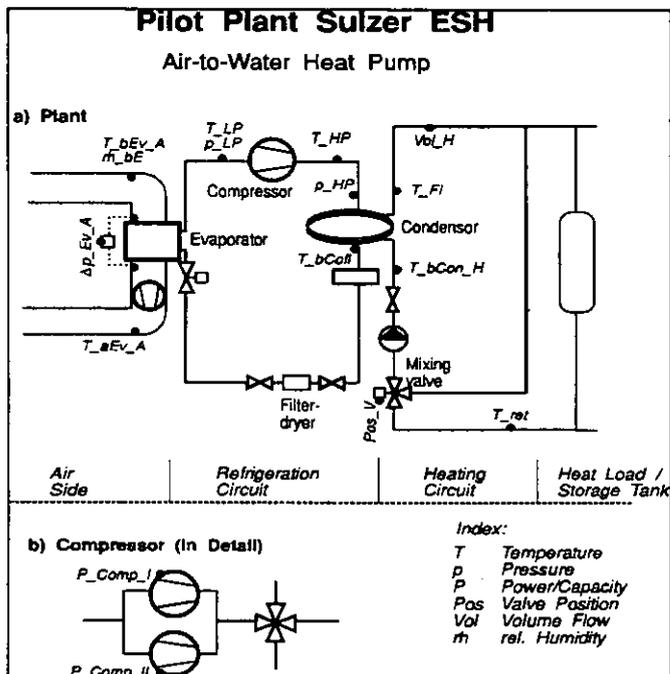
Fig. 2. Schematic representation of the air-to-water heat pump at ESH

Since this HP-plant is in operation, it was not possible to act upon the control settings. The condensor exhaust water temperature (Fig. 2, T_FI) was fixed (at 50°C). In consequence, only a part of the whole field-characteristic could be verified on this HP, but was still informative.

2.2 Data acquired on the EINEV plant

At the Ecole d'Ingénieurs de l'Etat de Vaud (EINEV) in Yverdon-les-Bains, a fully instrumented water-to-water laboratory heat pump rig could be used for this study. This test rig had the advantage of being run at different temperature levels (at both condensor and evaporator side) and allowed to investigate the whole operating range of the heat pumps.
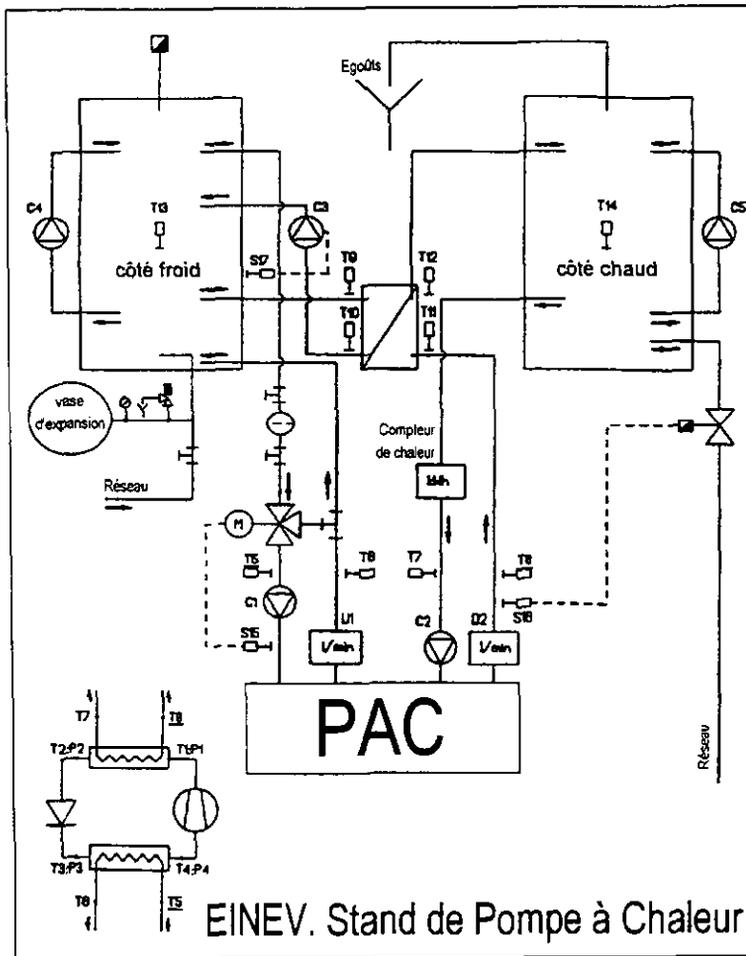
Fig. 3. Schematic representation of the EINEV HP laboratory plant

2.3 Data acquired on the CANMET plant (Canada)

Measurements performed on the CANMET test rig (Chiller) were made accessible to other IEA Annex 25 participants. The same scheme to find a field characteristic based on the exergetic efficiency as defined in (1) was applied. (Results see 3.3)

2.4 Steady-state conditions

A field-characteristic represents the performance of an equipment over a range of stable, or steady-state, conditions. To quantify the stability of a given magnitude we used a fixed-time window technique and calculated the standard deviation within this window (rolling values). When this value was less than a given threshold (e.g. 0.1K), the magnitude was considered of being stable. This technique was applied simultaneously to the different magnitudes involved in the computation of the exergetic efficiency as stated above. The corresponding selected values were then stored into a separate file containing only steady-state condition values. These 'steady-state points' were then used to determine the field-characteristic.

2.5 The field-characteristic of a heat pump or a chiller

To show the existence of a field-characteristic, all 3 data sets (ESH, EINEV and CANMET) were fitted to the following first order algebraic equation:

$$\varepsilon = a + b \cdot COP'_{Carnot} \tag{4}$$

with a and b as resulting coefficients of the curve-fit. These values a and b are characteristic to a given HP or chiller, but do not have any physical meaning. Important, of course, is the quality of the fit, that means the 'small' differences between measured and computed efficiencies. Repeated computations on these HP and chiller plant data, on the one hand, and discussions of the results of the curve-fits with HP specialists at Sulzer, ETHZ and EPFL, on the other hand, allow to state with confidence that the exergetic efficiency of heat pumps as defined in (1) does represent a reliable field-characteristic (Fig. 4).

In Figure 4 are represented both the measured ε (dots) and the fitted ε (web) performed on a water-water HP. Its shows a generally satisfying correlation between both magnitudes over the whole working range. Similar field-characteristics were obtained for the other plants.

The exergetic efficiency as a HP field-characteristic
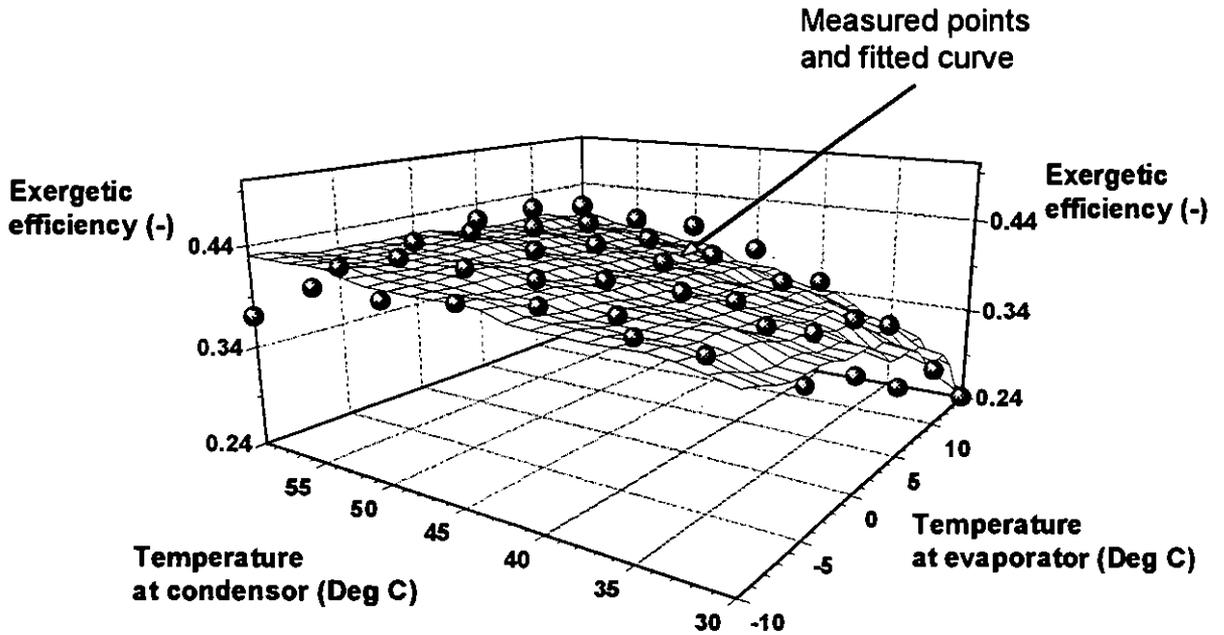
Fig. 4. The exergetic efficiency of a HP forms a field-characteristic

3. Deviations of acquired data from the characteristic

This paragraph presents graphically the deviations between the measurements (steady-state values over a given period) and the field-characteristic computed by the method described previously.

3.1   ESH Plant (air-to-water HP)

Figure 5 shows the absolute relative difference $(\varepsilon_1 - \varepsilon_2)/\varepsilon_2$ between computed and measured efficiency. The computed values represent values of the field-characteristic based upon the steady-state measured temperatures.
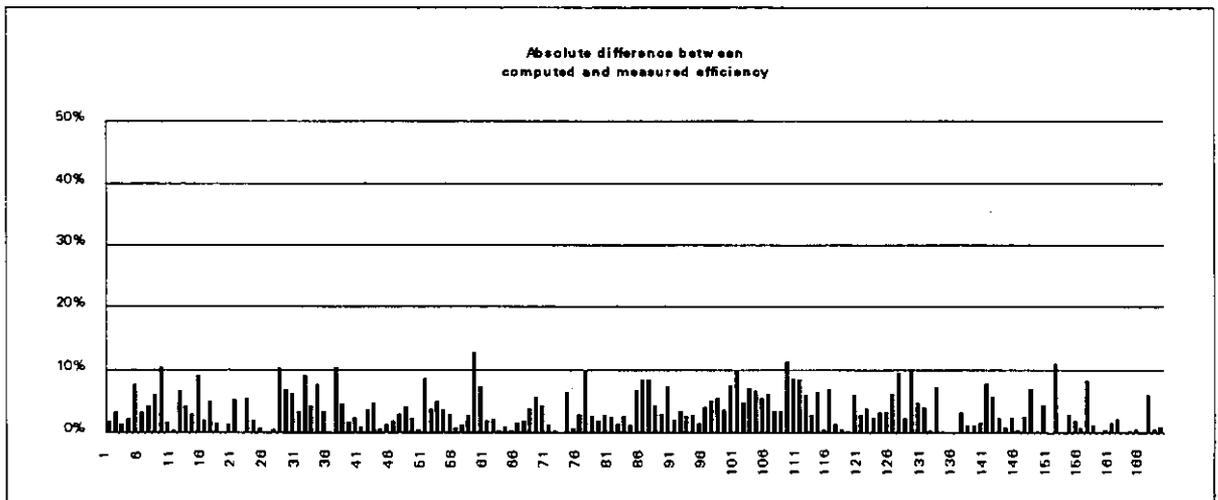
Fig. 5. Absolute relative difference $(\varepsilon_1 - \varepsilon_2)/\varepsilon_2$ between computed and measured efficiency

Deviations of no more than 13% were noticed. These relative small differences show that the field-characteristic represents correctly the performance of a HP plant working under 'normal' conditions.

3.2 EINEV Laboratory plant (water-to-water HP)

This laboratory water-to-water heat pump was tested over its full working range ($T_{evaporator}$ from -10°C to +15°C and $T_{condensor}$ from 30°C to 60°C). Each measured point in the graphs represents a stable (steady-state) condition.

Figure 6 shows the comparison between measurements and the fitted field-characteristic
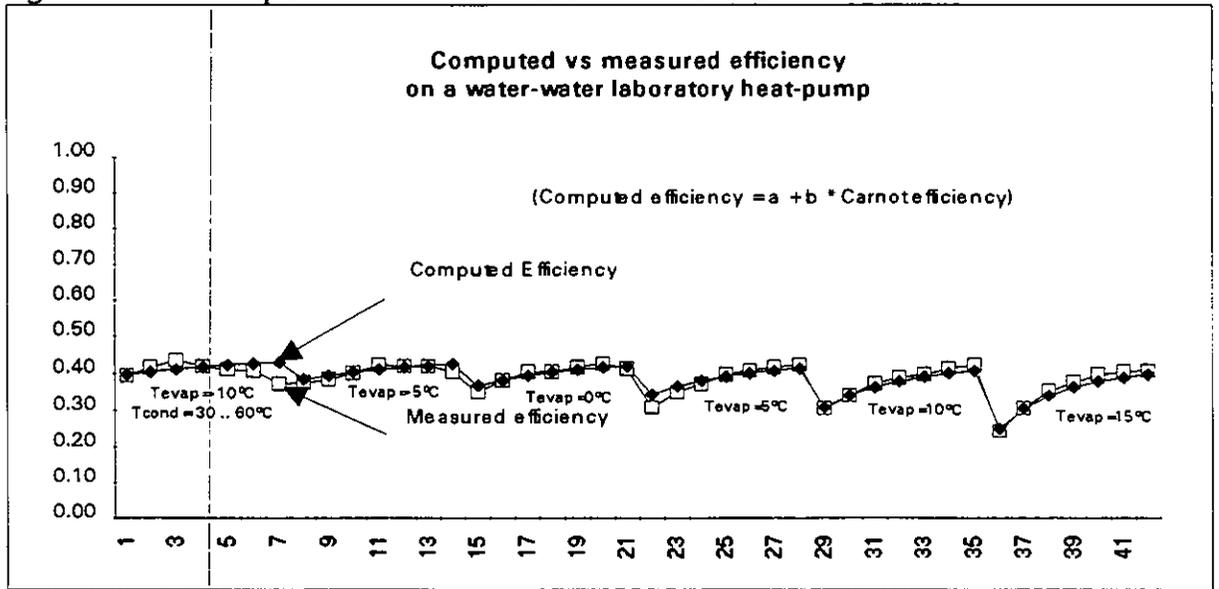
Fig. 6. Visualisation of the comparison between measurements and characteristic

(same values as used in fig. 4)

Figure 7 shows the relative deviations between the measured values and the EINEV-HP field-characteristic:
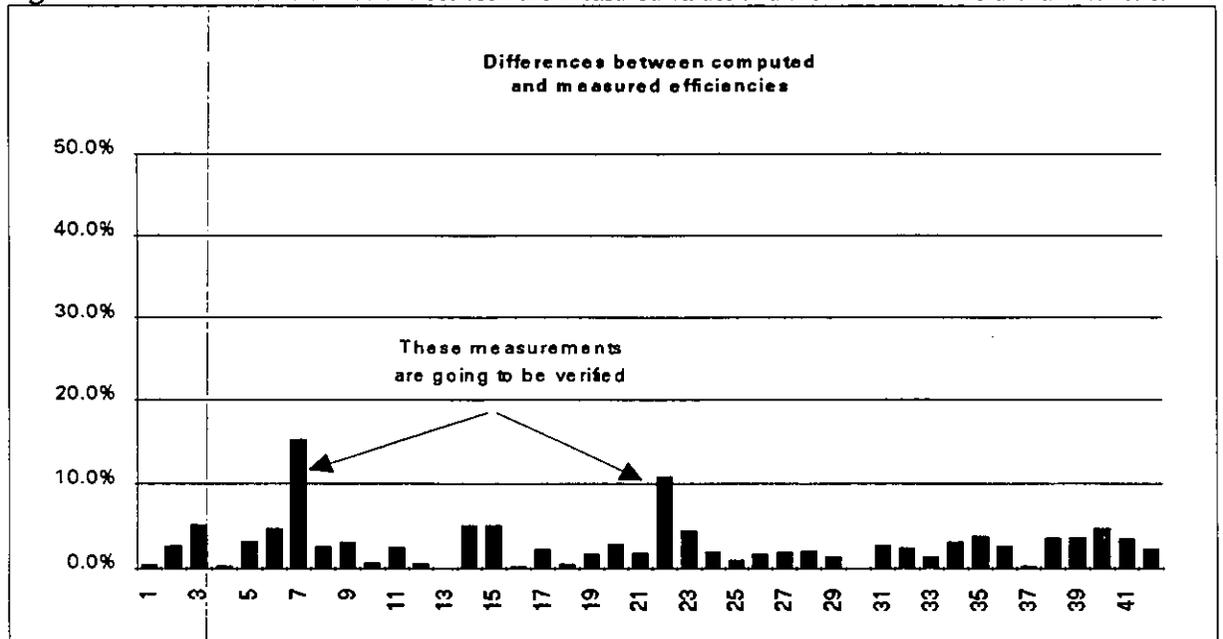
Fig. 7. Computed vs. measured efficiency (ε_1-ε_2)/ε_2 on the EINEV heat pump

Besides 2 exceptions, deviations of less than 5% were observed. This good agreement shows that the field-characteristic also represents correctly the performance of this water-to-water HP plant.

3.3 CANMET Laboratory plant (chiller)

Figure 8 shows a comparison between computed and measured efficiency values:
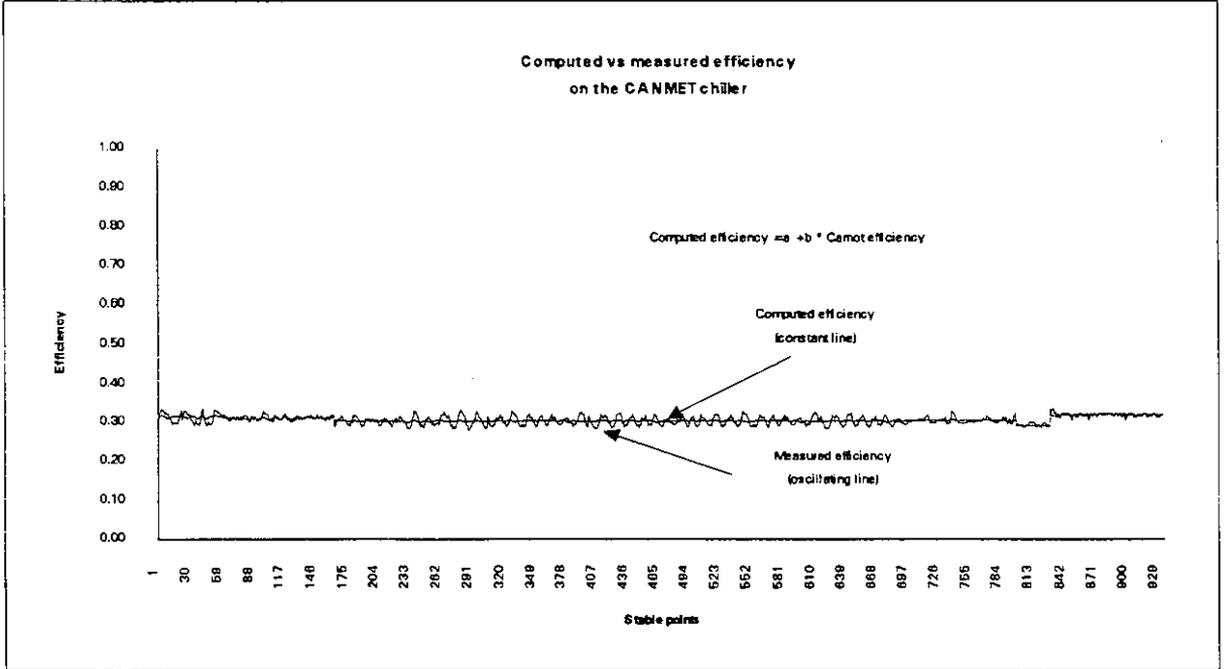
Fig. 8. Computed vs. measured efficiency on the CANMET chiller

A very good correlation could be observed. This good agreement shows that the field-characteristic also represents correctly the performance of a chiller plant. This result was expected, since a HP and a chiller are built on a similar thermodynamic basis.

Figure 9 shows the deviations between the measured values (steady-state values over a given period) and the exergetic field-characteristic at the CANMET plant:
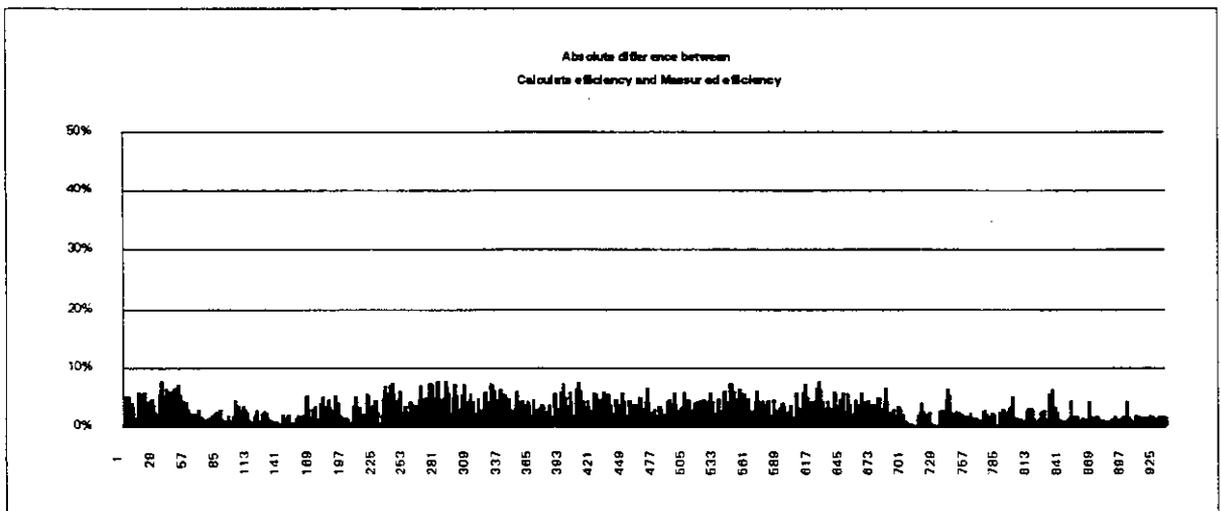
Fig. 9. Absolute relative difference $(\varepsilon_1 - \varepsilon_2)/\varepsilon_2$ between computed and measured efficiency

Deviations of less than 8% were observed, which show that the field-characteristic represents correctly the performance of a chiller plant working under 'normal' conditions.

4. The exergetic field-characteristic as a fault detector

CANMET made available to the IEA Annex 25 participants 2 sets of data: one set of 'normal data' containing operating magnitudes of a chiller working under normal, or non-faulty, conditions, and one set of 'faulty data', containing operating magnitudes of the same plant under 'abnormal' working conditions.

The first set of data was used, in this study, to determine the field-characteristic of the CANMET plant (see 3.3). This characteristic was then applied to the 'faulty data' to evaluate the pertinence of this field-characteristic to detect the specific faults.

Figures 10 to 13 present graphically the results of the comparison. The ordinate in the graphs represents the absolute ratio between the 'faulty' exergy and the 'normal' exergy from the field-characteristic as computed earlier (3.3). A short discussion of the results is presented at the end of this paragraph.

4.1 March 24th data

Figure 10 shows the detection of an abnormal performance of the chiller in the first part of the data set.
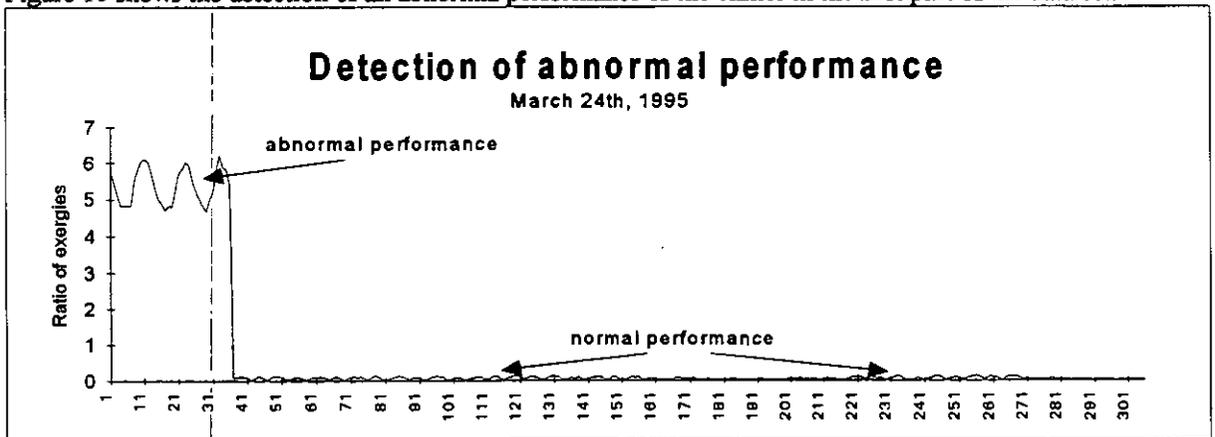
Fig. 10. An abnormal performance was detected in the first part of the data set.

4.2 March 26th data

Figure 11 shows the detection of a minor abnormal performance of the chiller (maybe due to start-up).
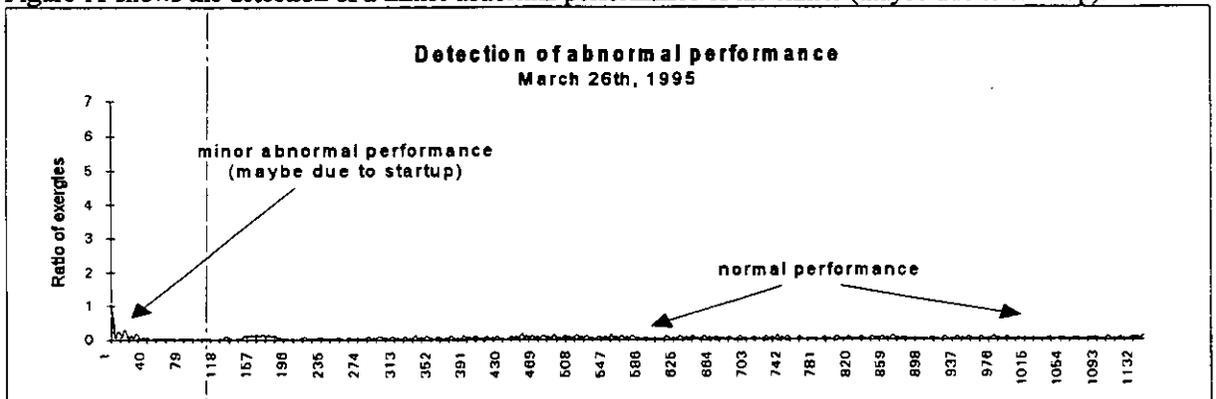
Fig. 11. A minor abnormal performance was detected in the first part of the data set.

4.3 March 28th data

Figure 12 shows the detection of an abnormal performance in the first part of the data set.
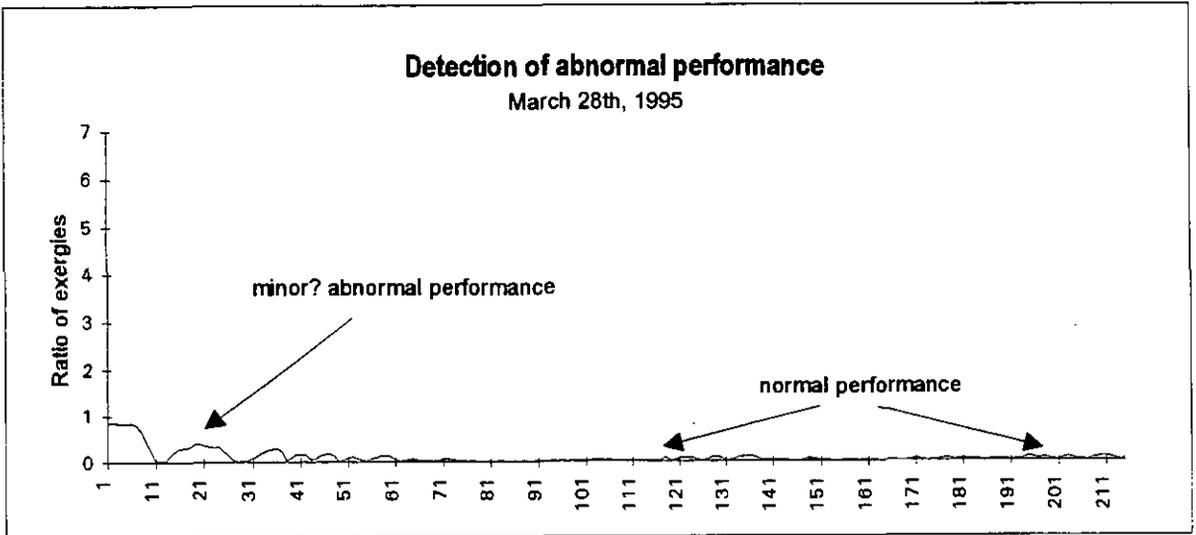
Fig. 12. A minor? abnormal performance was detected in the first part of the data set.

4.4 March 29th data

Figure 13 shows the detection of an abnormal performance in the middle of the data set.
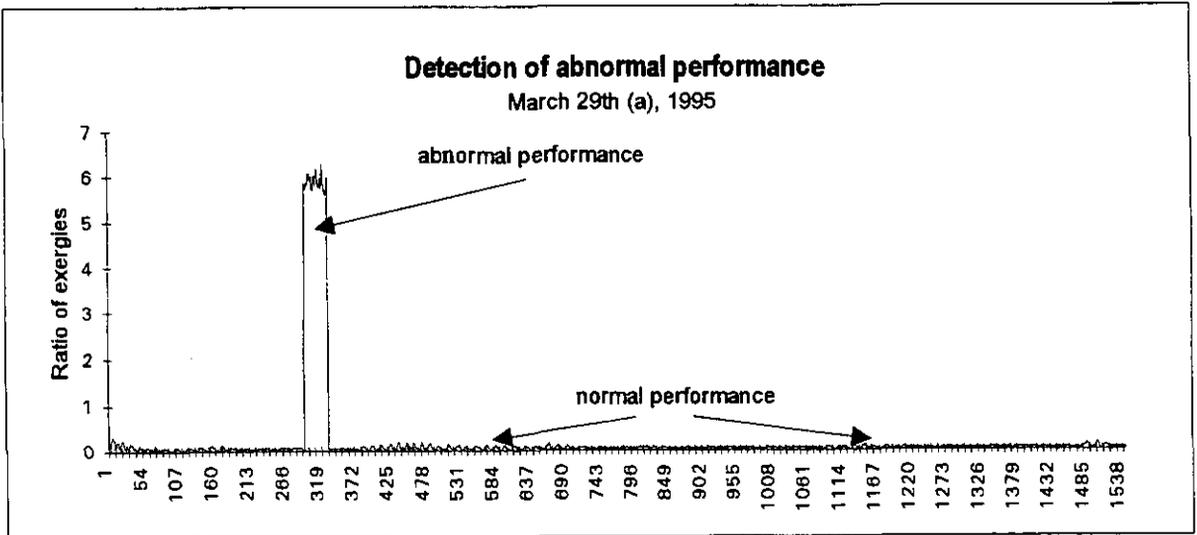
Fig. 13. An abnormal performance was detected in the middle of the data set.

4.5 Discussion of the results (CANMET 'faulty data')

Figures 10 to 13 show that the field-characteristic based upon the exergetic efficiency of a heat pump or a chiller (computed with normal operation data, see 3.3) detected several abnormal operating conditions in the 'faulty' data files. Giving an explanation of the cause of these 'faulty' conditions was not part of this detector,

but is rather part of the functionality of a separate diagnosis module, developed and presented earlier [1], which was not linked to the detection module for this study.

5. Conclusion

This paper shows the results of an investigation in the area of fault detection at heat pumps and at chillers. It could be demonstrated that the exergetic efficiency of a HP or of a chiller does form a field-characteristic capable of monitoring the performance of a that particular equipment over its full working range. Using this field-characteristic allows to detect quite easily any lack performance. Coupled to an appropriate diagnosis module, one obtains an efficient energy-management and preventive maintenance tool.

6. References

[1]   Daniel E. Maurer, Expertensystem für die Störungsdiagnose an wärmetechnischen Anlagen von Gebäuden, Forschungsbericht Nr 14, Laboratorium für Energiesysteme, ETH Zürich

FAULT DIRECTION SPACE PLUS IF-THEN METHOD FOR ON-LINE FDD IN AIR-CONDITIONING SYSTEM

Yi Jiang & Qianghua Zhou
Tsinghua University, P. R. China

ABSTRACT

This paper is basic research for on-line Fault Detection and Diagnosis (FDD) in air-conditioning system. Fault Direction Space plus if-then(FDS+) method is presented. The idea of FDS+ method is to avoid the system performance identification and prediction process as in normal fault detection procedure, and try to make the diagnosis procedure in more standard way. The Fault Direction Space is constructed by some Character Parameters (CPs) which are structured from measured data basing on the physical model of the components to be studied. The CPs are such selected that it can be expressed by the structure of the component so that the value of CPs should be constant during operation within a normal range for component fault free state. When a big change in CPs is observed, a fault is happened in components. The change of CPs becomes an indicator of fault. There can be a number of CPs in a system. Different type of faults may cause each CP change into different direction. All above phenomena can be expressed in Fault Direction Space: a fault free state becomes a point or a small region in the FDS which is called normal region; when there is a component fault happen, CPs is out of the normal region into abnormal region; different kinds of fault will be in different directions in the FDS, therefore the type of fault can be distinguished by comparing the measured direction of CPs with the standard fault directions. Since FDS method is based on component model and some faults for the air-conditioning system is difficult to be detected by this method, if-then method is presented to detection and diagnosis system faults.

INTRODUCTION

DDC control system for air conditioning is getting popular. It becomes very important to make full use of the computer of the DDC system. One of potential applications is on-line Fault Detection and Diagnosis(FDD). Discovering the malfunction of air-conditioning system can be helpful for maintaining the system at good performance and avoiding extra energy consumption. General procedure of on-lined fault detection can be divided into two steps: first, predicting the system performance for fault less state and comparing the forecast output with measured data. Symptoms can then be obtained from the differences. second, determining possible faults from the symptoms by a reasoning process. Physical models, ARMAX models and neural network models can be used for on-line prediction in first step. while a logical tree or Fuzzy match may be adopted for the reasoning procedure.

There are still some problems in practices. First of all, model on-line identification is necessary to predict the system performance. However the identification does need measured data from the object over different working states as the standard data, while determining what data is standard is the work of FDD. There exist some conflicts between the task of FDD and the process of FDD, thus the FDS+ method is presented to solve this problem.

Without complete information of the performance of each component, it is impossible to determine which one is in normal condition and which one is in fault. However, what we really want to known is if there is a great change on the performance of a component. If the system was accepted before and there is no great change in the performance, it should be acceptable at least , although there may be a fault at a small scale for a long time. From this point, the FDS+ procedure is put most attention on the detection of the change of performance of each component. A set of CPs(Character Parameter) is selected to express the component or system structure which the performance is dependent on. The FDD procedure can then base on the observing the variation of CPs. When a big change in CPs is observed, a fault happen in components. The change of CPs becomes an indicator of fault. There can be a number of CPs in a system. Different type of faults may cause each CP change into different direction. Take a water-air heat exchanger for example(figure 1):

heat balance equation:

$$Q = GC_a(t_{ao} - t_{ai}) = KF \frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}$$

$$Q = GC_a(t_{ao} - t_{ai}) = W'C_w(t_{wi} - t_{wo})$$

$$\frac{KF}{GC_a} = \frac{t_{ao} - t_{ai}}{\frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}} = CP_1$$

$$\frac{W'C_w}{GC_a} = \frac{t_{ao} - t_{ai}}{t_{wi} - t_{wo}} = CP_2$$

When there exist faults in water-air heat coil, different CPs will change in different direction as figure 9. At the first part of the priod, there is no fault, it can been seen that CPs changes within a small range; when faults

Figure 1 the heat transfer process(heat exchanger)

happen at 7500 sec., CPs change a lot and have different direction in different faults. CP_1 becomes larger and CP_2 becomes larger when fan speed decrease; CP_1 becomes smaller and CP_2 keeps constant when coil scale; CP_1 keeps constant and CP_2 becomes smaller when pressure of water system decrease.

If taking the change of CPs as axes to construct a fault direction space(FDS), the above fault symptom can be described conveniently in FDS. A fault free state becomes a point or a small region in the FDS, which is called normal region; when there is a component fault happened, whether fan speed decreasing, coil scale increasing or hot water system pressure decreasing, CPs is out of the normal region and into the abnormal region; different kinds of fault will change in different directions in the FDS: when fan speed decrease, the direction of CP is in the first quarter; when coil scale, the direction of CP is in the negative side of vertical axis; when hot water system pressure decrease, the direction of CP is in the negative side of horizontal axis(see figure 2). Therefore the type of fault can be distinguished by comparing the measured direction of the ΔCP with the standard fault directions.

Figure 2 Fault Direction in FDS

Since FDS method is based on component models while some system faults such as chilled water temperature too high or window opening can not be detected by this method, if-then method is presented to detect and diagnosis system faults. Therefore FDS+ is presented by combining FDS method with if-then method: detecting component faults by FDS method; detecting system fault by if-then method.

However, to put this method into real engineering application, there are still a few problems which need to be solved.

1. How to define a set of CPs to construct the FDS. The success of this procedure is highly dependent on the definition of the CPs.
2. How to set the size of normal region in the FDS. Due to the dynamic influence, non-linear effect and sensor errors in the real system, CPs vary within a small region rather than stay as constants even in a fault free state. To distinguish fault and normal state, it is necessary to define the normal state region.
3. How to determine the direction of different faults in the FDS. After the FDS for a typical component has been determined, does it suitable for all this kind of component? What is the universality of the FDS.

$$Q = GC_a(t_{ao} - t_{ai}) = KF \frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}$$

$$Q = GC_a(t_{ao} - t_{ai}) = W'C_w(t_{wi} - t_{wo})$$

$$\frac{KF}{GC_a} = \frac{t_{ao} - t_{ai}}{\frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}} = CP_1$$

$$\frac{W'C_w}{GC_a} = \frac{t_{ao} - t_{ai}}{t_{wi} - t_{wo}} = CP_2$$

Figure 1 the heat transfer process(heat exchanger)

When there exist faults in water-air heat coil, different CPs will change in different direction as figure 9. At the first part of the priod, there is no fault, it can been seen that CPs changes within a small range; when faults happen at 7500 sec., CPs change a lot and have different direction in different faults. CP_1 becomes larger and CP_2 becomes larger when fan speed decrease; CP_1 becomes smaller and CP_2 keeps constant when coil scale; CP_1 keeps constant and CP_2 becomes smaller when pressure of water system decrease.

If taking the change of CPs as axes to construct a fault direction space(FDS), the above fault symptom can be described conveniently in FDS. A fault free state becomes a point or a small region in the FDS, which is called normal region; when there is a component fault happened, whether fan speed decreasing, coil scale increasing or hot water system pressure decreasing, CPs is out of the normal region and into the abnormal region; different kinds of fault will change in different directions in the FDS: when fan speed decrease, the direction of CP is in the first quarter; when coil scale, the direction of CP is in the negative side of vertical axis; when hot water system pressure decrease, the direction of CP is in the negative side of horizontal axis(see figure 2). Therefore the type of fault can be distinguished by comparing the measured direction of the ΔCP with the standard fault directions.

Figure 2 Fault Direction in FDS

Since FDS method is based on component models while some system faults such as chilled water temperature too high or window opening can not be detected by this method, if-then method is presented to detect and diagnosis system faults. Therefore FDS+ is presented by combining FDS method with if-then method: detecting component faults by FDS method; detecting system fault by if-then method.

However, to put this method into real engineering application, there are still a few problems which need to be solved.

1. How to define a set of CPs to construct the FDS. The success of this procedure is highly dependent on the definition of the CPs.
2. How to set the size of normal region in the FDS. Due to the dynamic influence, non-linear effect and sensor errors in the real system, CPs vary within a small region rather than stay as constants even in a fault free state. To distinguish fault and normal state, it is necessary to define the normal state region.
3. How to determine the direction of different faults in the FDS. After the FDS for a typical component has been determined, does it suitable for all this kind of component? What is the universality of the FDS.

4. . The judgeability of the procedure. Because of dynamic influence, non-linear effect and sensor error, the fault vector do not remain in the same direction even in the same fault, but in a range of direction. Furthermore, direction range of different fault vectors may overlap,.thus it is important to define the classification ability of FDS.

STRUCTURE CPs

Finding a suitable set of CP is the key to make the FDS successful. The structure o f CP should be:

- it can be deduced from the physic model of component in air-conditioning system;
- it is expressed by the structure of components, thus it is in constant over the whole working range;
- it can be calculated from the measured data;
- it should include complete fault information so that each type of fault will appear in different direction in the FDS.

According to these principles, Cps for some components in air-conditioning system have been determined as shown in table 1.

<div align="center">

Table 1 the selection of CPs

</div>

No.	Components	CPs
1	water-air heat exchanger	$$cp_1 = \frac{KF}{GC_a} = \frac{t_{ao}-t_{ai}}{\dfrac{(t_{ai}-t_{wo})-(t_{ao}-t_{wi})}{\ln\dfrac{t_{ai}-t_{wo}}{t_{ao}-t_{wi}}}}$$ $$cp_2 = \frac{WC_w}{GC_a f_1(c)} = \frac{t_{ao}-t_{ai}}{(t_{wi}-t_{wo})f_1(c)}$$
2	dehumidify and cooling coil	$$cp_1 = \frac{WC_w}{G f_2(c)} = \frac{h_{ai}-h_{ao}}{(t_{wo}-t_{wi})f_2(c)}$$ $$cp_2 = \frac{KF}{G f_3(c)} = \frac{h_{ai}-h_{ao}}{\dfrac{(h_{ai}-h_{wo})-(h_{ao}-h_{wi})}{\ln\dfrac{h_{ai}-h_{wo}}{h_{ao}-h_{wi}}}f_3(c)}$$ $$cp_3 = e^{-\frac{KF-f_3(c)}{GC_a}} = \frac{1-\varphi_i}{1-\varphi_o}e^{\frac{f_3(c)}{GC_a}}$$
3	steam sprayer	$$cp_1 = \frac{W}{G f_4(c)}$$
4	steam -air heat exchanger	$$cp_1 = \frac{KF}{GC_a} = \frac{t_{ao}-t_{ai}}{t_{wi}^* - \dfrac{t_{ao}+t_{ai}}{2}}$$ $$cp_2 = \frac{t_{wi}-y}{f_6(c)}$$
5	co-moving dampers	$$cp_1 = \sqrt{\frac{S_{return}}{(S_{fresh}+S_{exhaust})f_9(c_2)}} = \frac{t_{mix}-t_{return}}{(t_{fresh}-t_{mix})\sqrt{f_9(c_2)}}$$
6	washing room	$$cp_1 = \frac{(1-\eta_2)}{f_7(c_1)} = \frac{t_{ao}-ts_o}{f_7(c_1)(t_{ai}-ts_i)}$$ $$cp_2 = \frac{WC_w}{G f_8(c_1)} = \frac{h_i-h_o}{(t_{wo}-t_{wi})f_8(c_1)}$$

note: $f_n(c)$ is a correction term of CPs for the open ratio of valves and it can be learned during operation.

A FDS can be easily constructed for any system by adopting above CPs, but there are some procedure to process when CPs are calculated on line. First, The CPs created is based on the static physic model of the system or components. However, the data to structure the CP are measured from real system in a dynamic process and include some variation due to sensor error. Moving average method is applied to decrease the effect. Second, what does the change in CPs mean or how to calculate the ΔCPs? It can be considered as the average change between the present period and the last, so that it can be defined as:

$$\Delta CP = \frac{\int_{t-T}^{t} \frac{\partial CP}{\partial \tau} d\tau}{\frac{1}{T} \int_{t-2T}^{t-T} CP d\tau}$$ (1)

T—Integral time

Figure 3 show the change of CP when a fault happen in t1; figure 4 show the change of ΔCP.
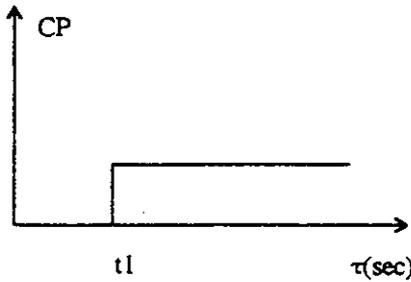
Figure 3 the change of CP   Figure 4 the change of ΔCP
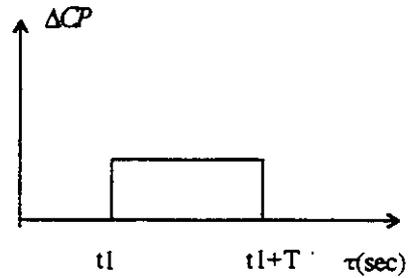
NORMAL REGION IN FDS

Although moving average method improves a lot for obtaining a stable CP, CP still changes within a region rather than stay at one value during a fault less state. This is due to dynamic influence, non-linear effect and sensor error in real system. Therefore it is necessary to define the normal state region(NSR). If the CP changes within this region, it can be considered as a fault less state.

Take a FDS of a water-air heat exchanger for example:

$$cp_1 = \frac{KF}{GC_a} = \frac{t_{ao} - t_{ai}}{\frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}}$$

$$cp_2 = \frac{WC_w}{GC_a f_1(c)} = \frac{t_{ao} - t_{ai}}{(t_{wi} - t_{wo}) f_1(c)}$$

Simulation is carried out by HVACSIM+ to discover the range of NSR, Figure 10 show the range of NSR for the FDS of heat exchanger at different faultless working state. It is clear that the NSR is operation state dependent. This makes the FDS method difficult to be used in practice. To avoid this an uniformed NSR is needed to be developed.

let $Dt_a = t_{ao} - t_{ai}$, $Dt_w = t_{wi} - t_{wo}$, $Dt_{ln} = \frac{(t_{ai} - t_{wo}) - (t_{ao} - t_{wi})}{\ln \frac{t_{ai} - t_{wo}}{t_{ao} - t_{wi}}}$ then

$$\frac{\Delta CP_1}{CP_1} = \sqrt{\left(\frac{\Delta Dt_a}{Dt_a}\right)^2 + \left(\frac{\Delta Dt_{ln}}{Dt_{ln}}\right)^2}$$

$$\frac{\Delta CP_2}{CP_2} = \sqrt{\left(\frac{\Delta Dt_a}{Dt_a}\right)^2 + \left(\frac{\Delta Dt_w}{Dt_w}\right)^2}$$

ΔDt_a, ΔDt_w, ΔDt_{ln} is the variation caused by above influence, approximately $\approx$ single $\Delta Dt_{ln} = \Delta Dt_a$, get the variation of CPs in different working state as following:

t_{ai}	t_{ao}	t_{wi}	t_{wo}	ΔCP_1	ΔCP_2
20°C	24°C	80°C	79°C	$0.08\Delta Dt_a$	$0.13\Delta Dt_a$
22°C	24°C	80°C	79.5°C	$0.16\Delta Dt_a$	$0.2\Delta Dt_a$

Thus CPs should be uniformed as:

$$\Delta CP_1' = \frac{\Delta CP_1}{\sqrt{\left(\frac{1}{Dt_a}\right)^2 + \left(\frac{1}{Dt_{1n}}\right)^2}}$$

$$\Delta CP_2' = \frac{\Delta CP_2}{\sqrt{\left(\frac{1}{Dt_a}\right)^2 + \left(\frac{1}{Dt_w}\right)^2}}$$

The uniformed NSR is shown in figure 11. It is less operation state dependent.

There is no clear boundary between fault and faultless state in most of the cases. The serious level of fault can be from 0, that is fault less state, to very heavy, what we more interest about the sensitivity is to known what level of fault can be discovered by the FDS+ procedure. Since there are many uncertainty influence such as error of sensors, dynamic influence as well as non-linear effect, the sensibility of the FDS+ procedure is the probability o being discovered for a certain level of fault . Figure 5 is an example of the probability of effect against the level of fault for an air heater when the fault is air flow rate decreasing. It can been seen that when the air flow rate decreased 10%, the probability of being discovered is only 31%. However when it decreased to 15%, in 90% cases it can be discovered.
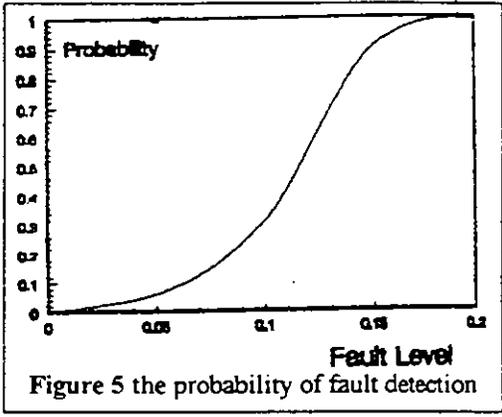
Figure 5 the probability of fault detection

FAULT DIRECTION IN FDS

Since CPs are deduced from physic model of components, the direction of fault can be get by analysing the physic process of fault. Also take heat exchanger for example, the direction of faults show on figure 2. The value of β can be deduced easily by analysing the physical process of fan speed decrease and β is equal to 45°. The information of fault direction can be store in the way of a vector with the same dimensions as the number of CP:

	$\Delta CP_1'$	$\Delta CP_2'$
fan speed decrease	0.71	0.71
heat coil scale	-1	0
water system pressure decrease	0	-1

Since what we pay great attention is the direction of the vector, CPs can be normalised to construct fault knowledge database.

$$\Delta CP_1' = \frac{\Delta CP_1}{\sqrt{\left(\Delta CP_1\right)^2 + \left(\Delta CP_2\right)^2}}$$

$$\Delta CP_2' = \frac{\Delta CP_2}{\sqrt{\left(\Delta CP_1\right)^2 + \left(\Delta CP_2\right)^2}}$$

The angle between the measured $\Delta CP'$ and the standard fault $\Delta CP'$ can be calculated as following during on-line process:

$$\psi = D\Delta$$

Where D is the matrix of standard fault direction, Δ is the vector constructed by the measured $\Delta CP'$ as

$$\Delta = \begin{vmatrix} \Delta CP_1^{'} \\ \Delta CP_2^{'} \\ \bullet \\ \Delta CP_n^{'} \end{vmatrix}$$

Each element in ψ is the $\cos\varphi$ of the angle between the measured Δ and the relative standard fault direction. When $\cos\varphi = 1$, the direction of the measured $\Delta CP'$ is exactly the same as the relative standard fault direction.

Due to the following reasons, the same type of fault in different operation states or at different levels can not be always at one direction in the FDS, but in a range and range of different faults may overlap, that is to say there exist a fuzzy zone and a clear zone in FDS for classifying different faults(see figure 6).

●the error caused by sensor can make the measured data deviate from the real state;

●the non-linear behaviour of the fault may cause the direction change when different level of fault occur;

●the non-linear behaviour of the component may also cause the direction change when it is in different working state;

●the dynamic influence is another contribution to the deviation of the fault direction.
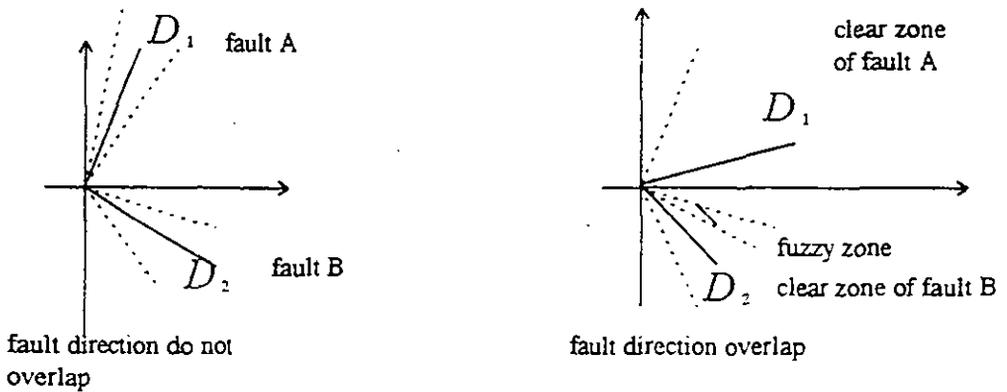
Figure 6 range of fault direction

$\cos\varphi_{min}$ is produced in fault state simulation to represent the range of fault direction. If $\cos\varphi < \cos\varphi_{min}$, this type of fault do not happen, however since fuzzy zone exist, if $\cos\varphi > \cos\varphi_{min}$, it might not mean this fault happens, it may be another fault. It is necessary to define the fuzzy region and clear region of different faults to finish the classification. The boundary angle of fuzzy zone and clear zone is get by different fault simulation. The membership in fuzzy zone is describe as following.
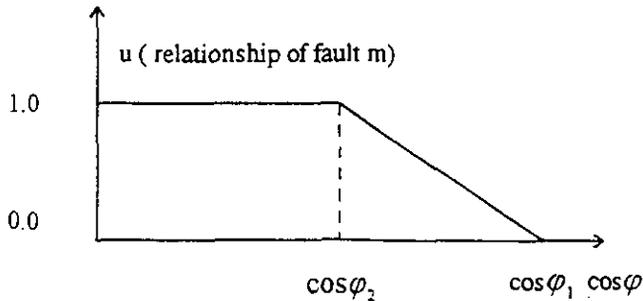
Figure 7 membership of fuzzy zone

$$u(\cos\varphi) = \begin{cases} 0 & \cos\varphi < \cos\varphi_1 \\ \dfrac{\cos\varphi - \cos\varphi_1}{\cos\varphi_2 - \cos\varphi_1} & \cos\varphi_1 \leq \cos\varphi < \cos\varphi_2 \\ 1 & \cos\varphi_2 \leq \cos\varphi \end{cases} \quad (2)$$

u —relationship of fault m
φ_1—boundary angle of fault m's fuzzy zone
φ_2—boundary angle of fault m's clear zone

GENERAL PROCEDURE OF FDS+ METHOD

The general process of FDD is divided into off-line study and on-line procedure. The whole procedure of FDS method is introduced through an example(see figure 8). Different FDD system is set up in different working mode to detect and diagnosis fault respectively. Take the one working mode for example: dehumidify and cooling coil on, heat exchanger off, steam sprayer off, minimum fresh air.

1. off-line study:
1)selection of CPs

$$cp_1 = \frac{W_3 C_w}{G f_4(c_2)} = \frac{h_{ai} - h_{ao}}{(t_{wo} - t_{wi}) f_4(c_2)}$$

$$cp_2 = \frac{KF}{G f_5(c_2)} = \frac{h_{ai} - h_{ao}}{\dfrac{(h_{ai} - h_{wo}) - (h_{ao} - h_{wi})}{\ln\dfrac{h_{ai} - h_{wo}}{h_{ao} - h_{wi}}} f_5(c_2)}$$

$$cp_3 = e^{-\frac{KF - f_5(c_2)}{GC_a}} = \frac{1 - \varphi_i}{1 - \varphi_o} e^{\frac{f_5(c_2)}{GC_a}}$$
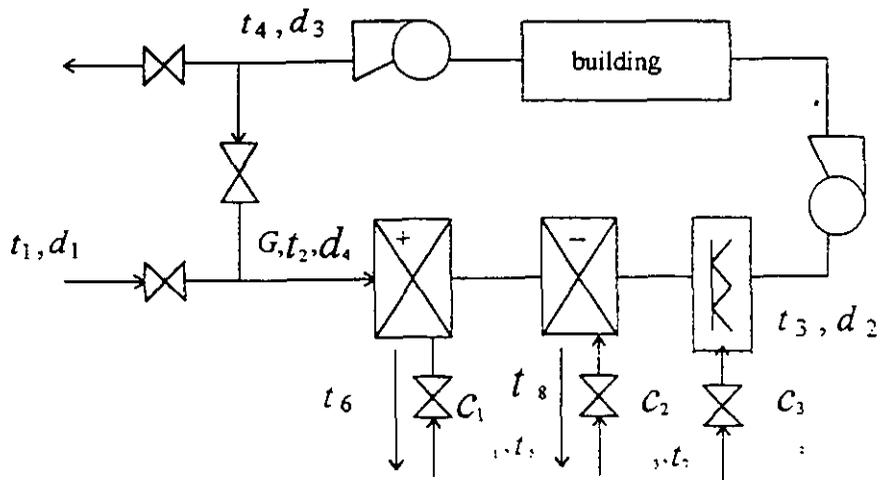
Figure 8 reference system

$$h_{ai} = h(t_2, d_4), h_{ao} = h(t_3, d_2), h_{wi} = h(t_7, d_{sai}), h_{wo} = h(t_8, d_{sai}).$$
$$\varphi_i = rh(t_2, d_4), \varphi_o = rh(t_3, d_2), t_{wi} = t_7, t_{wo} = t_8$$

where h_{wi} saturated air's enthalpy under input water temperature and h_{wo} saturated air's enthalpy under output water temperature

2) fault free state simulation. Uniformed NSR 0.9 is obtained through simulation.
3) fault matrix is get by off-line study (off-line analysis plus simulation validation):

	$\Delta CP_1'$	$\Delta CP_2'$	$\Delta CP_3'$	
fan speed decrease	0.73	0.46	-0.5	
cooling coil scale	0	-0.71	0.71	=D
pressure in cool water system fall	-0.41	-0.23	0.88	

4)fault state simulation. Classification information is get in the simulation.

Table 2 classification information

fault type	fault level	$\bar{D}_m$	$\cos\varphi_{min}$	integrate $\bar{D}_m$	integrate $\cos\varphi_{min}$	integrate $\cos\varphi_{max}$	clear zone	fuzzy zone
fan speed decrease	20%	(.70,.48,-.52)	.98	(.73,.46,-.50)	0.72	0.04	(0.72,1)	not exist
	25%	(.71,.47,-.52)	.99					
	30%	(.76,.43,-.48)	.75					
cooling coil scale	20%	(0,-.71,.71)	.16	(0,-.71,.71)	.13	.80	(.80,1)	(.13,.80)
	25%	(0,-.71,.71)	.13					
	30%	(0,-.71,.71)	.15					
water's pressure fall	20%	(-.46,-.26,.85)	.92	(-.41,-.23,.88)	.90	.90	(.90,1)	not exist
	25%	(-.40,-.21,.89)	.94					
	30%	(-.36,-.21,.91)	.97					

a. $\bar{D}_m$ —uniformed standard fault vector.

b. $\cos\varphi_{min}$ —minimum angle between the direction of state vector and standard fault vector, representing the region of this fault in FDS.

c. integrate $\bar{D}_m$ — uniformed standard fault vector in different fault level.

d. integrate $\cos\varphi_{min}$ — minimum angle between the direction of state vector and standard fault vector(include different fault level), representing the region of this fault in FDS.

e. integrate $\cos\varphi_{max}$ — maximum angle between the direction of state vector in other fault state and standard fault vector, representing the boundary of clear zone for this fault and showing classification ability of this fault in FDS.

5) rule set is set up through system structure analysis. There exist two system faults: building cool load high and temperature of cool water high. Under these fault state, state vector remain in NSR, that is to say FDS method can not detect these fault and a set of rules is applied to deal with these faults:

Table 3 rule set for system faults

fault type	load of cool coil	room temp.
building cool load high	max.	t>t0
cool water temp. high	max.	t>t0

a. t0 ----- set room temperature

The rules for building cool load high and cool water temperature high are the same, thus cool water temperature predictor is adopt to classify these two faults. Until now, FDD system has been set up off line and on-line FDD procedure will be simulated to verify the FDD system.

2. on-line detection and diagnosis.

Also in this mode, cool coil scale at 7500 sec. In on-line FDD, the variation of CPs are calculated firstly(see figure 12), then this variation is picked out by function 1(see figure 13), mod of uniformed state vector is found out of NSR in 8400 sec.(see figure 14). A fault is detected and diagnosis procedure processes successively. Fault matrix multiply operation is doing.

fan speed decrease	0.73	0.46	-0.5		$\Delta CP_1''$			x1
cool coil scale	0	-0.71	0.71	*	$\Delta CP_2''$	=		x2
water pressure fall	-0.41	-0.23	0.88		$\Delta CP_3''$			x3

$(\Delta CP_1'', \Delta CP_2'', \Delta CP_3'')$ is the normalised $(\Delta CP_1, \Delta CP_2, \Delta CP_3)$ and x1, x2, x3, x4 represent the difference between the direction of state vector and standard fault direction.

In fault state there exist normalised state vector (0.0144, -0.18638, -0.00367), after above matrix multiply get x2=0.69, fault is in fuzzy zone according table 3, relationship is 0.84(function 2). Of course in this example the other state vector show x2>0.8, cool water scale fault is clearly diagnosed.

CONCLUSION

The main character of the FDS method are:

1. The on-line identification of component or system model as required in most of the fault detection procedure is not needed. CP is used to play the rule of models instead. The work of model identification is replaced by the off-line CP determination.

2. The Fault Direction Space is used to replace the if-then logic tree used in most of fault diagnosis procedure. Although it may also be consider as a kind of logic tree, the representation of the FDS in a computer is in a way of a matrix rather than a group of rules and match can be done by a simple multiply operation rather than searching along the fault tree as it does in a if-than logic tree procedure. Furthermore, FDS method not only give the fault type but also give the relationship of this type.

NOMENCLATURE

c	openth of valve, damper	(--)	C_a	specific heat of air	(kJ/kg.°C)
C_w	specific heat of water	(kJ/kg.°C)	d_i	input air humility	(kg/kg.dry air)
d_o	output air humility	(kg/kg.dry air)	G	air flow rate	(kg/s)
h_{ai}	input air enthalpy	(kw/kg)	h_{ao}	output air enthalpy	(kW/kg)
K	heat transfer coefficient	(kW/m2.°C)	F	heat transfer area	,m2
Q	heat transfer power	(kW)	S	flow resistance coefficient	(1000/kg.m)
t_{ai}	input air temp.	(°C)	t_{ao}	output air temp.	(°C)
t_{wi}	input water temp.	(°C)	t_{wo}	output water temp.	(°C)
ts_i	input air wet bulb temp.	(°C)	ts_o	output air wet bulb temp.	(°C)
W	water flow rate	(kg/s)	φ_i	input air relative humility	(--)
φ_o	input air relative humility	(--)	τ	time	(sec.)
η_2	heat transfer coefficient for washing room	(--)	$f_n(c)$	correction term for the openth of valve or damper	(--)

REFERENCES

Hyvarinen J &Kohonen R, Building Optimisation and Fault Diagnosis system concept, IEA ANNEX 25 report, Painatuskeskus OY, Helsinki 1993
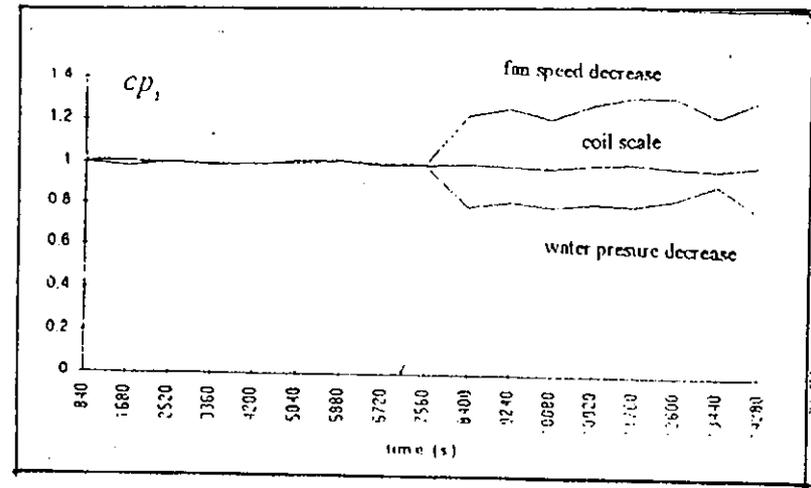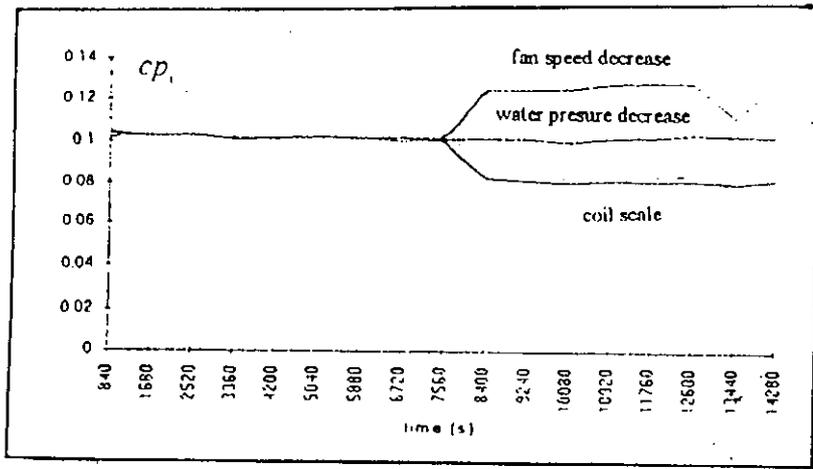
Figure 9 the change of CPs in different faults
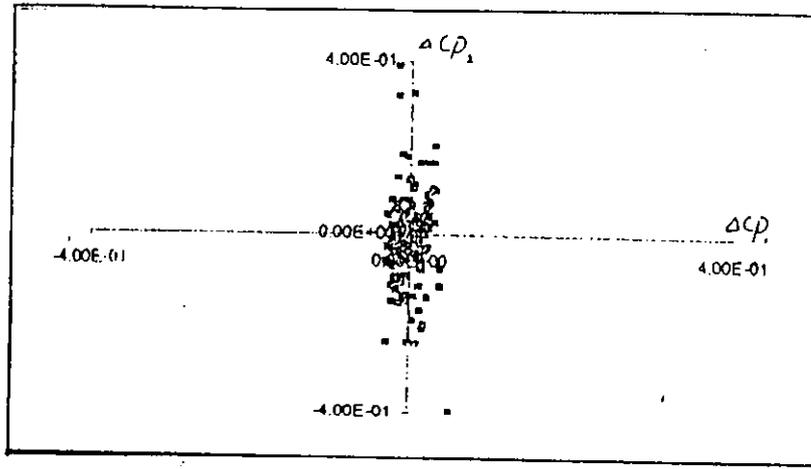
Figure 10 NSR without uniformation
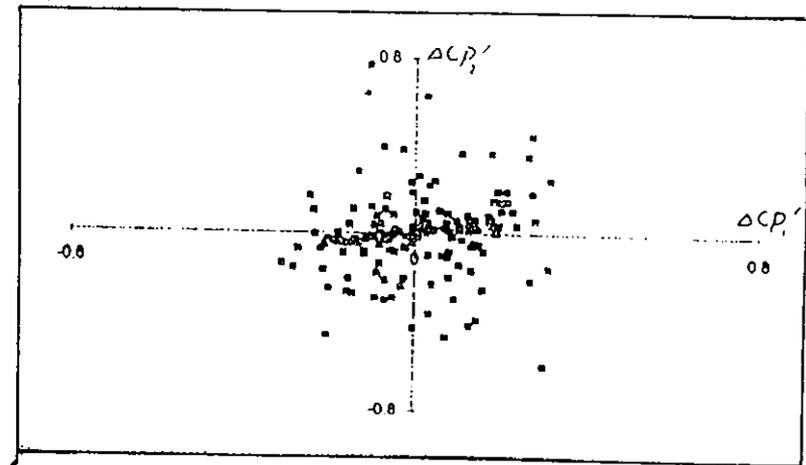
Figure 11 uniformed NSR
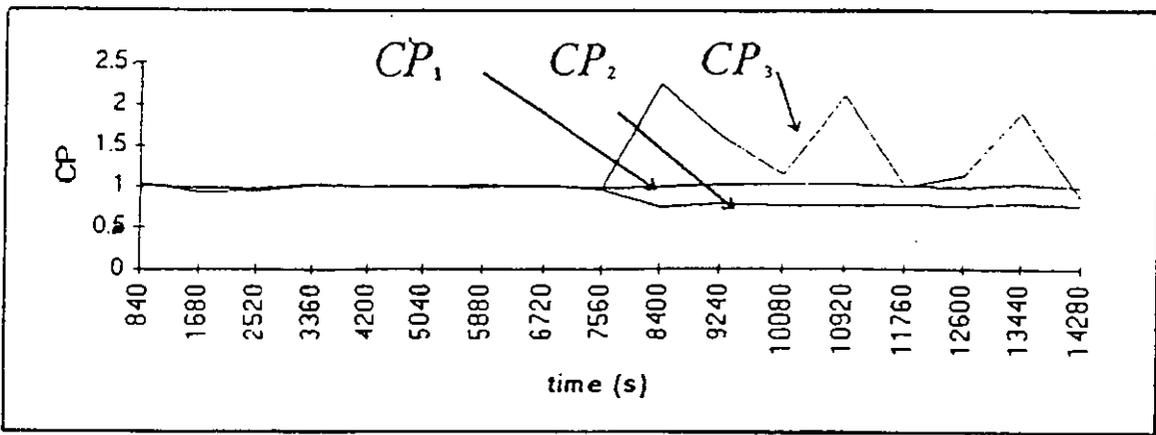
Figure 12 the change of CPs in operation
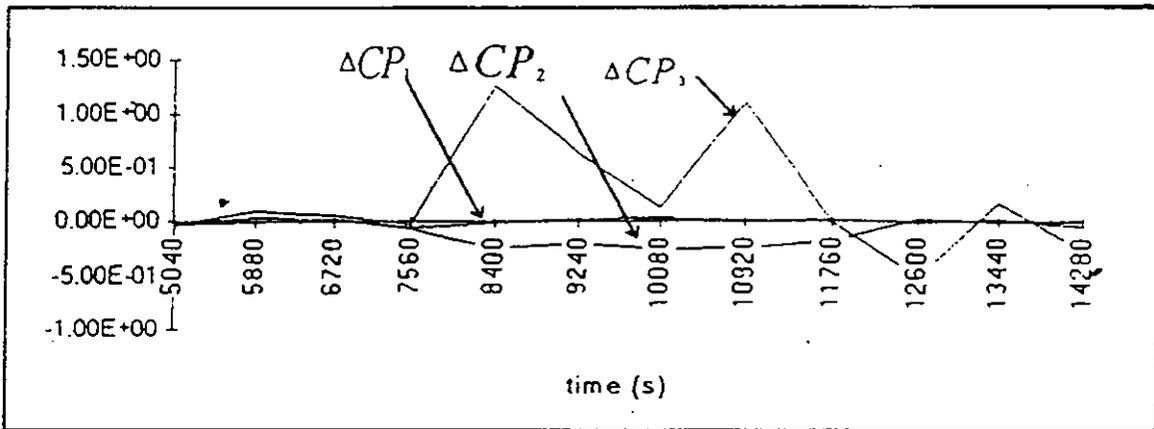
Figure 13 symptom of the change of CPs
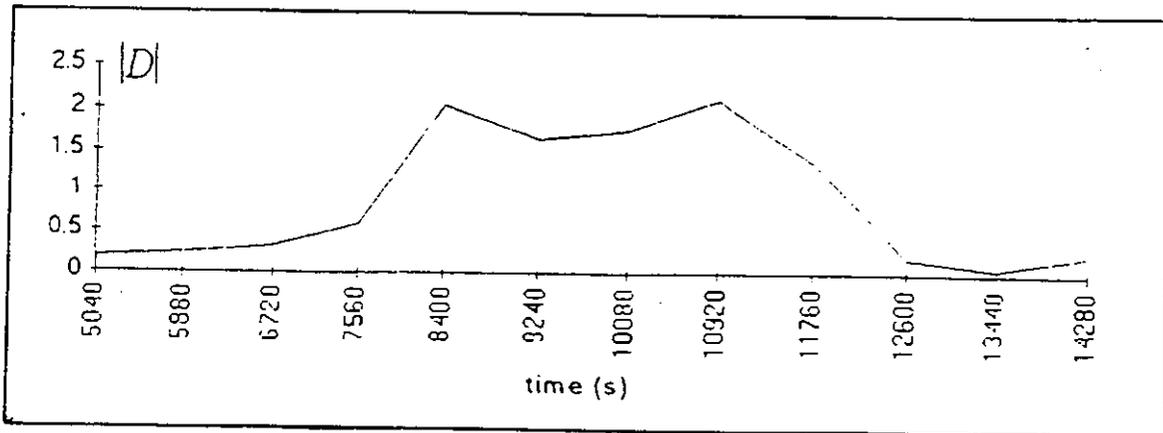
Figure 14 mod of uniformed Cps

A Methodology for Defect Detection in Heat Pumps

Roger Raeber

Energy Systems Laboratory, ETH Zentrum
CH-8092 Zurich, Switzerland

1   INTRODUCTION

Some faults in HVAC equipment don't occur instantaneously, they grow until they reach a level where the system performance becomes unacceptable or where the system breaks down. The period between the first indication of the system's performance deterioration to the system failure, i.e. the fault's *time constant*, can be minutes to months. Defect detection aims at detecting faults in an early phase so that preventive maintenance can be carried out before the system's breakdown.

The terms *defect*, *fault* and *failure* have been defined within the preparation phase of IEA Annex 25 [1]. They represent different phases of what is generally called a *fault*. The *defect* is the earliest phase of a fault.

Different approaches to detecting faults in vapor compression equipment have been investigated within the IEA Annex 25 [e.g. 2, 3]. Most of them use steady state models and so only certain faults can be detected and diagnosed. The method presented in this paper tries to take the system's dynamics into account so that faults leading to transient states can be detected as early as possible.

2   FAULTS

When a system tends to faulty operation, the system's behavior is changing. The transients caused by the faults have characteristic time constants. Faults with long time constants must be regarded as (quasi) steady-state conditions, whereas faults with short time constants are causing transients. *Short* time constants in the context of HVAC equipment stands for the duration of minutes up to hours. Faults causing significant transients in HPs are [3]:

- Refrigerant floodback
- Expansion valve faults
- Refrigerant leak
- Refrigerant line flow restrictions
- Compressor faults

Performance indices (PI) are used to detect the changes in the system's behavior as mentioned above. PIs are discussed later in section 3.2.1.

The method for defect detection as presented in this paper is not able to detect every possible defect; because it can only detect transients and their related defects. The work carried out until now was concentrated on the faults listed above. If a fault occurs, however, it must be possible to classify it, so that normal operation can definitely be distinguished from faulty operation. For this reason the models used must not only represent faulty operation, but also normal operation. This way, if a situation occurs, for which no model is implemented, then normal operation can be excluded and the defect detection system can inform the user about an unknown fault, that has been detected.

3 METHOD

3.1 CONCEPT

A performance index (PI) is a steady-state model of the corresponding system or component. It's a suitable indicator for the system's or component's current operation. The expectable range of such a PI is usually known from the system's/component's manufacturer and/or from laboratory experiments. Thus an expert is able to judge the system's condition by a suitable PI. A certain PI will not have the same value over the whole operation period. It will change due to changes of the operation set point of the system or due to a fault. Defect detection as presented here makes use of this fact. The system's dynamics make PIs change and these changes are used as indicators for defects.

The basic idea is, to choose suitable PIs for the system and/or its components, to determine characteristic patterns of normal and faulty operation based upon the PIs, to monitor the system's current PI patterns and to classify them by comparing them with the predetermined patterns. Defect detection thus makes use of the system's dynamics without actually simulating it dynamically, which is known to be very complex for systems like HVAC equipment.

Figure 1 shows a block diagram of the defect detection process. It is mainly divided into the preprocessing and the classification phase according to [2]. On the real system the necessary data points are measured in order to compute the operation patterns (OP) for each PI. The computation of the OPs will be descibed in section 3.2.2. The system's current OP is then compared with the predetermined ones (Pattern 1 to 3 in Figure 1) and the corresponding residuals are transmitted to the classification block. The predetermined OP most similar to the system's current OP, i.e. the OP with the smallest residual, will indicate the operation state (normal or faulty) with the highest probability.
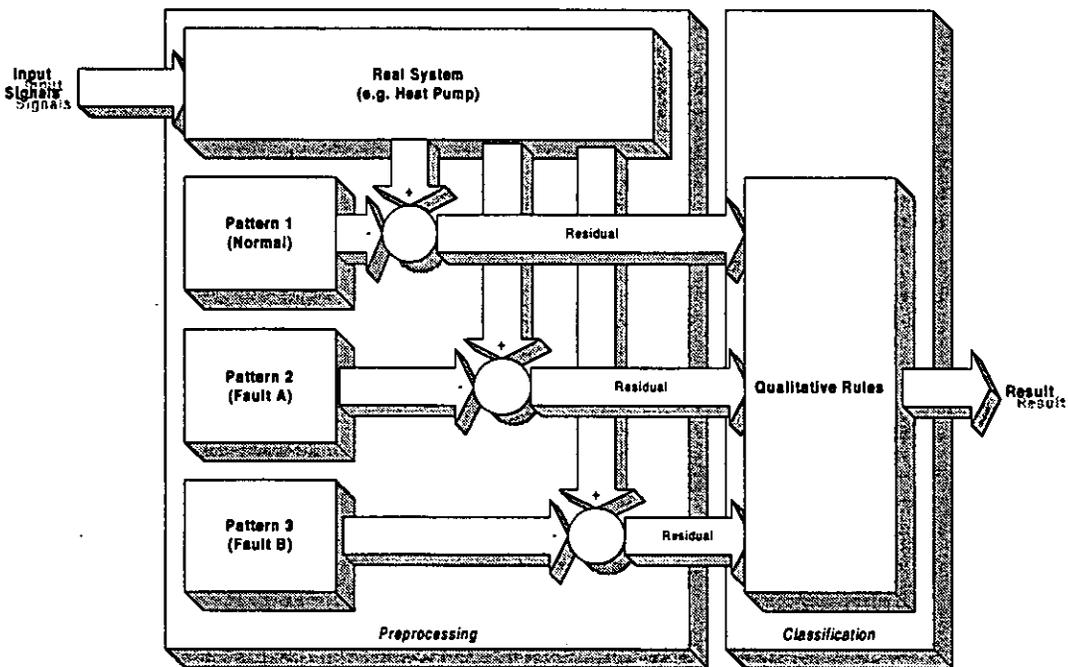
Figure 1: Defect detection preprocessing and classification

3.2 PREPROCESSING

The preprocessing part of the defect detection process consists of two main phases, the training phase and the operation phase. During the training phase the operation pattern database is set up. Data series of normal and faulty operation must be acquired from the HP in question, so that the OPs can be computed. This set of OPs then forms the knowledge database. The training phase must be run through only once. During the operation

phase the necessary data points are measured and the system's current OP is computed and classified. However, the OPs are always computed the same way.

3.2.1 Performance Indices

Before the preprocessing training phase can be started the PIs must be chosen. In general the PIs must be chosen according to the faults to be detected. So for each component affected by at least one fault one PI is chosen and one PI is added for the overall system. In the case of vapor compression equipment and considering the faults listed in section 2, the following PIs are suitable for defect detection purposes:

- Heat exchanger efficiency [6]:

$$\Phi' = \frac{t'_H - t''_H}{t'_H - t'_C}$$

- Compressor volumetric efficiency [6]:

$$\lambda = \frac{\dot{V}_{eff}}{\dot{V}_h}$$

- Chiller/HP exergetic efficiency [6]:

Chiller: *Heat Pump:*

$$\eta_{CC} = \frac{\varepsilon_C}{\varepsilon_{CC}}, \text{where} \qquad \eta_{CH} = \frac{\varepsilon_H}{\varepsilon_{HC}}, \text{where}$$

$$\varepsilon_C = \frac{\dot{Q}_0}{P_{el}} \text{ and } \varepsilon_{CC} = \frac{T_0}{T_c - T_0} \qquad \varepsilon_H = \frac{\dot{Q}_c}{P_{el}} \text{ and } \varepsilon_{HC} = \frac{T_c}{T_c - T_0}$$

The more PIs are taken into account, the more accurate the defect detection can be. Of course, every additional PI slows down the defect detection process. In addition every PI needs a certain number of sensors, in order to compute it. Sensors are expensive so the less sensors are needed the better. The number of PIs should be as small as possible to minimize sensor and control system cost, and great enough to meet a satisfactory accuracy and detection speed.

3.2.2 Operation Patterns and Signal Processing

The computation of the OPs includes the following steps:

1. Measurement of the required data points
2. Computation of the PIs based upon the measurements
3. Computation of the power spectral densities (PSDs) via FFT

The PIs that have been chosen determine the required data points. For the above listed PIs in a HP the following data points are needed:

Evaporator/Condenser:
Primary/secondary circuit inlet/outlet temperatures

Compressor:
Refrigerant circuit suction/liquid line pressures

System:
Compressor electric power consumption
Condenser secondary circuit inlet/outlet temperature
Refrigerant circuit evaporator/condenser temperature

All the other values required to calculate the PIs are constants and can be obtained from manufacturer information or from experimental knowledge.
The measurements are then carried out at a specific sampling rate. The choice of this rate is important, because the cut-off frequency f_c and thus the frequency range and the frequency resolution Δf of the PSDs are depending on it according to [4]

$$f_c = \frac{1}{2T} \text{ and } \Delta f = \frac{2f_c}{N}$$

where T is the sampling interval and N is the number of samples within the time window. The relations between frequency resolution, sampling interval, cut-off frequency and time window width are summarized in Table 1 for N = 1024 samples.

Time constant of defect	Time window width W [s]	Sampling interval T [s]	Cut-off frequency f_c [Hz]	Frequency resolution Δf [Hz]
1 s	0.5	0.0005	1024.0000	1.000000000
1 min	30.0	0.0293	17.0667	0.016666667
1 h	1800.0	1.7578	0.2844	0.000277778
1 d	43200.0	42.1875	0.0119	0.000011570
1 w	302400.0	295.3125	0.0017	0.000001653

Table 1: Data acquisition and signal processing parameters

To be able to detect defects that have time constants of about one day a sampling interval of about 40 seconds would be needed; with N = 1024 samples the time window would be 43200 seconds (12 hours) wide. The highest frequency shown in the corresponding PSD is 0.0119 Hz corresponding to a wavelength of 84 seconds. The time window width is the second important parameter for the later frequency analysis, because the wider the window is at a constant sampling rate, the better the frequency resolution will be.

Based on the measured data the PIs and the PSDs can be computed. The PSD S_{yy} is given as the square of the Fourier transformed PI signal Y according to [4] by

$$Y(j\omega) = \int_{-\infty}^{\infty} y(t)e^{j\omega t}dt \text{ and } S_{yy}(j\omega) = \frac{1}{NU}|Y(j\omega)|^2,$$

where N is the number of samples and U is a normalization factor. Y is computed using FFT [5]. The preprocessing phase results in diagrams like Figure 2.
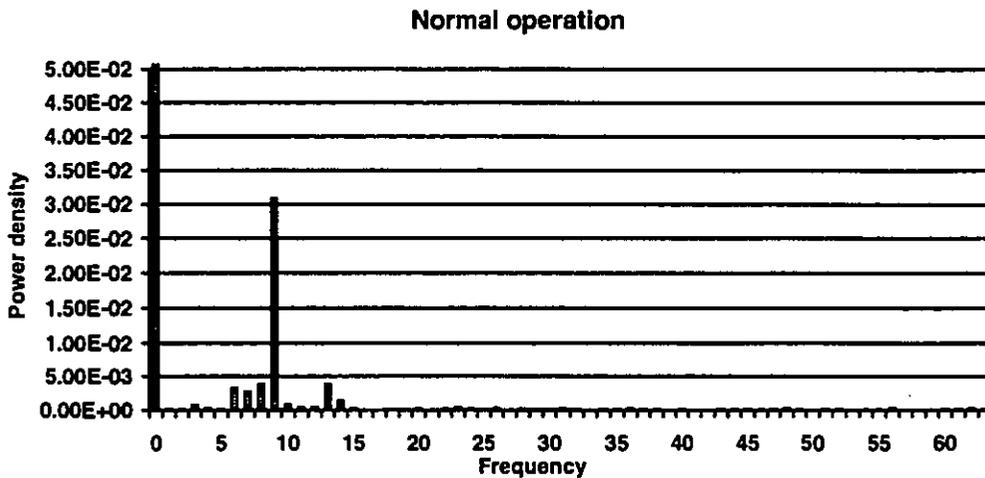
Normal operation

Figure 2: PSD example of normal operation based on the exergetic efficiency of a HP

To simplify matters the PSDs or periodograms will be called operation patterns (OP). OP examples of a HP's normal and faulty operation are displayed in Figure 2 and Figure 3 respectively. An OP can be represented as a n-element column vector or as a point in a n-dimensional feature space. This view of the OP will be used to illustrate the classification process (cf. section 3.3).

Faulty operation A
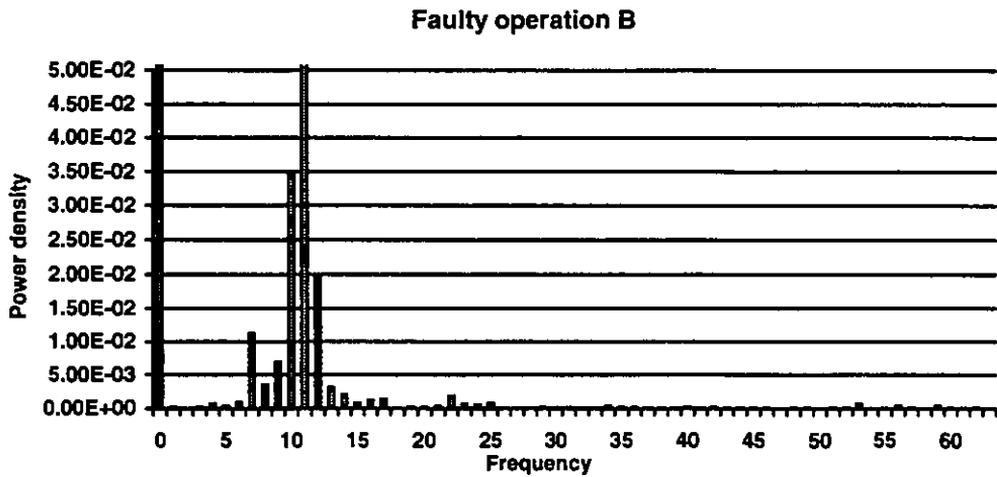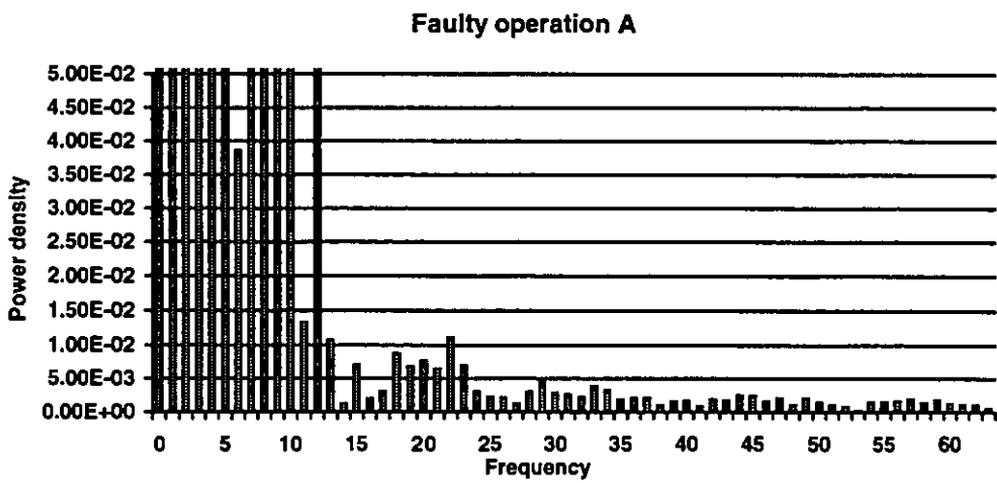
Faulty operation B

Figure 3: PSD examples of faulty operation based on the exergetic efficiency of a HP

3.2.3 Training Phase

During the training phase the defect detection system has to learn about the supervised process, in this case about a HP or chiller. Until now the training phase must be carried out once for each system, because it is not known yet, if and how far this method can be generalized. Training includes data acquisition of the system's normal and faulty operation, so that the required OPs can be computed based upon the PIs that have been chosen before. One OP per fault to be detected and at least one OP for normal operation is required to set up a useful knowledge database. After the training phase the defect detection system is ready for operation.

3.2.4 Operation Phase

In the same manner as during the training phase the supervised system's OPs are measured and computed. It is important, that the same measurement parameters (sampling interval, time window width, etc.) are used as during the training phase, so that the resulting OPs are comparable. Again the result of the preprocessing phase are the PSDs of various PIs, and again every OP is a n-dimensional column vector and can be represented as one point in a n-dimensional feature space.

3.3 CLASSIFICATION

The result of the classification is the assignment of an OP to its corresponding fault. Classification is here carried out in two steps:

1. *Classification of the system's current behavior*
   The defect detection system classifies the OP with respect to its time frame and makes a statement concerning the conditions the HP is currently in. It computes a ranking list of the faults with decreasing probability.

2. *Prediction of the system's future behavior*
   The defect detection system tries to predict the system's future behavior. After each defect detection run changes in the system's current behavior and in the fault probabilities are judged. The fault with the greatest probability gain becomes the most probable future fault.

The classification is based on the concept of the OPs representing points in a n-dimensional feature space. For the further discussion the following assumptions are made:

1. The feature space is 2-dimensional, the corresponding features are F_1 and F_2
2. The OP database consists of $N = 3$ OPs representing normal operation (b_1) and the faults A (b_2) and B (b_3)
3. The OPs in the knowledge database and the current OP are represented as points $(B_1, B_2, B_3,$ and $\hat{B}$) in the F_1-F_2 plane
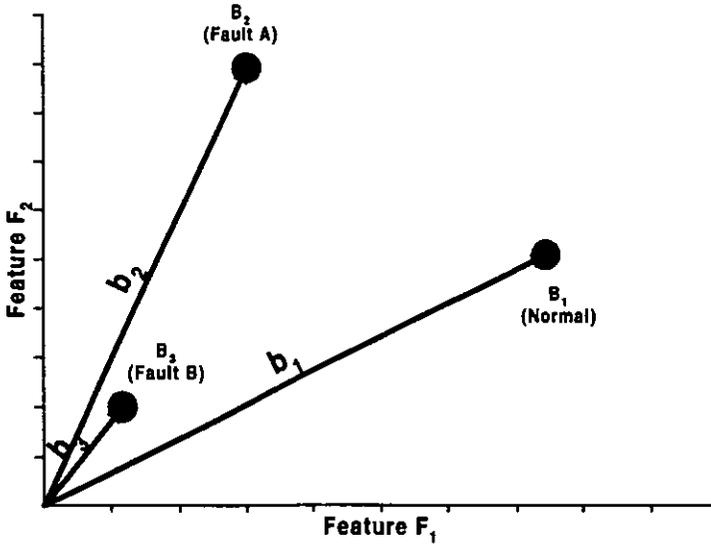
Figure 4: Classification feature space

3.3.1 Current Behavior

Classification of the current OP $\hat{B}$ is mainly a pattern recognition problem. In this case the OPs or feature vectors are classified by means of the distances between their ending points. If b_k is one known feature vector (OP) with ending point B_k and $\hat{b}$ is the current feature vector with ending point $\hat{B}$, then b_{ki} and $\hat{b}_i$ are the i-th of n elements of these vectors and the distance d_k between B_k and $\hat{B}$ is given by

$$d_k = \left|\overline{B_k \hat{B}}\right| = \left|b_k - \hat{b}\right| = \sqrt{\sum_{i=1}^{n}(b_{ki} - \hat{b}_i)^2} \ ,$$

A 2-dimensional example is given in Figure 5. The known OP with the smallest distance to the current OP indicates the system behavior (normal or faulty) with the greatest probability, and the smaller the distance d_k is, the higher the probability for the system behavior corresponding to OP B_k.
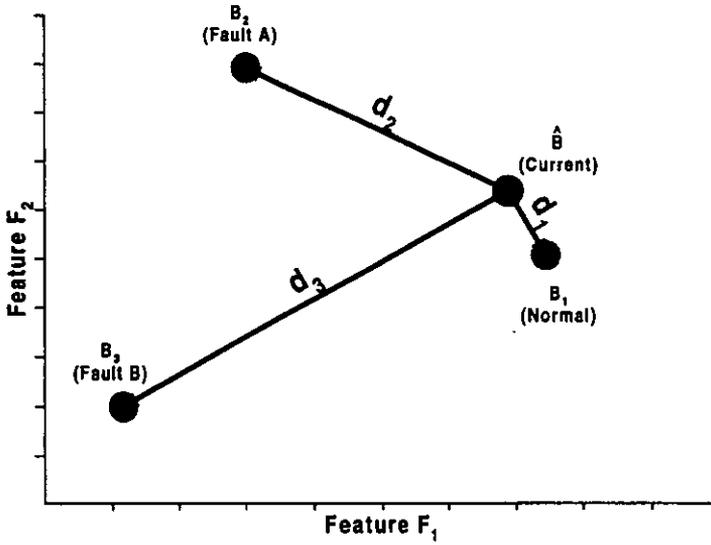
Figure 5: Classification of the system's current behavior

Although it would be possible to do so (cf. [7]), the statistical probabilities of the different OPs themselves are not quantified. A different probability quantity called *relative probability* is given instead. The relative probability $p_r(B_k)$, is a measure for the similarity of the current OP $\hat{B}$ to one particular known OP B_k. $p_r(B_k)$ is given by

$$p_r(B_k) = \frac{1}{d_k \sum_{j=1}^{N} \frac{1}{d_j}} \qquad \text{if } d_j, d_k \neq 0$$

if $\exists\, d_k = 0$, then $p_r(B_k) = 1$ and $p_r(B_l) = 0 \;\forall\; d_l$ with $l \neq k$

This similarity measure is called *relative probability*, because the sum of all p_r is always 1 (100%) and the relative probability of OP B_k is 1, when it is identical to the current OP $\hat{B}$. The distances between the current and the known OPs in Figure 5 and their corresponding relative probabilities are displayed in Table 2.

	Distance d_k to current OP $\hat{B}$	Relative Probability p_r
B_1	1	0.70588 (70.6%)
B_2	4	0.17647 (17.6%)
B_3	6	0.11765 (11.8%)

Table 2: Distances and relative probabilities (cf. Figure 5)

If there is a total of N OPs in a defect detection knowledge database and ν OPs represent normal operation and φ OPs represent faulty operation then the sum of the ν normal or φ faulty operation relative probabilities are indicating the overall relative probability for normal or faulty operation respectively.

543

3.3.2 Behavior Prediction

Under ideal normal conditions the point representing the system's current OP $\hat{B}$ will keep its location. In reality it will not exactly do so; the point will move due to noise in the measured data or due to transients in the system's behavior. Transients can be caused by normal changes of the operation set-point or by faults. However, the current OP's point $\hat{B}$ will move away from one normal OP's point towards an other normal or towards one of the faulty OP's points. The movements of the current OP's point can be used to predict the systems future behavior.

To avoid false alarms, movements due to noise must be distinguished from movements due to transients in the system. As can be seen in Table 1 the frequency range of a PSD for defects with a 1 day time constant is 0 to 0.0119 Hz, which is a relatively narrow band of very low frequencies. The relevant frequency components for defect detection are thereby in the lower half band of this range. Noise on the other hand lies in the higher half band and above and the corresponding frequency components are small. Tests with experimental data have shown, that normal OPs measured at different times have distances in the range of 1 to 1.5 from each other. Distance changes in this range are obviously caused by noise. Faulty OPs have distances of 40 to 60 from normal OPs. So noise can be neglected, because it is low compared with the distances between normal and faulty OPs. If noise is too high, then adequate filtering methods should be applied.

Figure 6 shows the original situation of the example from Figure 5 at time t_1 and a new situation at time $t_2 = t_1 + \Delta t$.
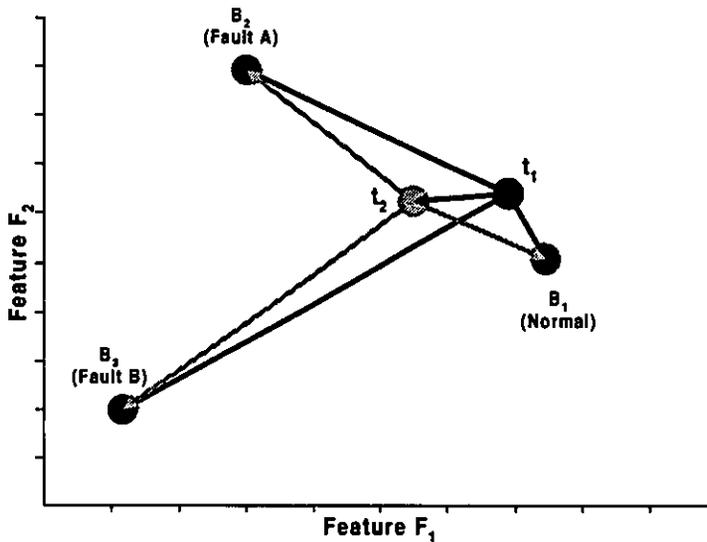
Figure 6: Prediction of the system's future behavior

The relative probabilities at t_2 have changed but still normal operation is most probable (B_1). However, normal operation is less probable at t_2 and the faults A (B_2) and B (B_3) have become more probable that before. Behavior prediction will now compute a ranking list of OPs which have become more probable, i.e. with relative probability gains. The OP with the highest gain denotes the most probable future behavior of the system.

	t_1		t_2		$t_1 \rightarrow t_2$
	d_r	$p_r(B_k)$	d_r	$p_r(B_k)$	$\Delta p_r(B_k)$
B_1	1	0.7059	2	0.4839	-0.2220
B_2	4	0.1765	3	0.3226	+0.1461
B_3	6	0.1176	5	0.1935	+0.0759

Table 3: Changes in relative probabilities used for behavior prediction

The distances and relative probabilities at t_1 and t_2 are summarized in Table 3. Normal operation has a relative probability loss of 22.2%. Among the OPs with relative probability gains fault A (B_2) has the highest gain (14.6%). So at the moment it is most probable that the system is tending to fault A.

The analysis of different possible system behaviors has shown, that for the user a relative probability threshold is required so that the importance of the relative probabilities and their gains can be interpreted more easily. For three known OPs sensible results could be obtained with a relative probability threshold of 0.5 and a relative probability gain threshold of +0.1; these thresholds have been found heuristically. However, the thresholds must be determined according to the number of known OPs. The more known OPs there are, the lower these thresholds will be.

4   VALIDATION

So far the defect detection method has been developed theoretically and the main statements could be verified with experimental data obtained from a laboratory chiller system. A more detailed data acquisition campaign is currently being done on a laboratory heat pump system. The main issues expected from this campaign are:

- Determination of signal processing parameters
- List of possible faults to be detected by the defect detection method
- Limits of the method (accuracy, sensitivity, robustness, etc.)

5   CONCLUSIONS

The analysis of experimental chiller data has shown, that the three basic conditions for defect detection on behalf of OPs are satisfied, i.e. that

1. OPs representing normal operation are similar to each other
2. OPs representing normal and faulty operation are significantly different from each other
3. OPs representing different faulty operations are significantly different from each other

This means, that the behavior of vapor compression equipment can be modeled with OPs as presented in this paper and that different OPs can be distinguished from each other. In addition to the preprocessing method implied by the OPs a method for classification and behavior prediction has been presented and verified with experimental chiller data. Further verification is currently being done to point out the limits of defect detection. No statements about the applicability of the OPs obtained from one system to another system of the same type or even of a different type can be made so far.

REFERENCES

[1]   J. Hyvärinen, R. Kohonen (Ed.):IEA ECB Annex 25: Building Optimisation and Fault Diagnosis System Concept, VTT Espoo (FIN), 1993

[2]   T. Rossi, J.E. Braun: Using Thermodynamic Impact for Detecting Refrigerant Leaks in Vapor Compression Equipment, Paper, Purdue University, West Lafayette, IN (USA), 1995

[3]   M. Stylianou, D. Nikanpour: Performance Monitoring, Fault Detection and Diagnosis of Reciprocating Chillers, EDRL-CANMET, Varennes (CAN), 1996

[4]   A.V. Oppenheim, R.W. Schafer: Zeitdiskrete Signalverarbeitung, Oldenbourg, München, 1992

[5]   E.O. Brigham: FFT, Schnelle Fourier-Transformation, 6. Auflage, Oldenbourg, München, 1995

[6]   H. Recknagel, E. Sprenger, E.R. Schramek (Ed.): Taschenbuch für Heizung und Klimatechnik, 67. Auflage, Oldenbourg, München, 1995

[7]   R.J. Schalkoff: Pattern Recognition, Statistical, Structural and Neural Approaches, Wiley, New York, 1992

STUDY ON FAULT DETECTION AND DIAGNOSIS OF THERMAL STORAGE CONTROL SYSTEMS WITH PATTERN RECOGNITION

Mingjie Zheng, Assistant Professor
Nobuo Nakahara, Professor
Nagoya University
Nagoya, Japan

Abstract

Automatic fault detection and diagnosis (FDD) is important process which has been ignored but has great effect to realize energy conservation in Heating, Ventilating and Air-conditioning (HVAC) systems. Present paper introduces a trial of FDD in the thermal storage control system by the pattern recognition method. It has been recognized that the temperature profiles is a proper indices of the thermal storage system performance. Mapping process based on the statistical theory with significant parameters derived from the Fourier Analysis of the temperature profiles showed a high performance of FDD for the two kinds of control faults using fault simulation data.

1. INTRODUCTION

The faulty phenomena observed in the thermal storage tank are affected not only by its proper faults but also by faults at any part of HVAC system. At the same time the thermal storage faults affects operational performance of HVAC system and the building environment. Experiences show that a peculiar control fault around the thermal storage tank gives a characteristic transition of the temperature profiles of the tank, which somehow provides an information on the existence of some kind of faults.

Nakahara, one of the authors, reported in [1] [2] the simulation results of two kinds of control faults for eighty-one cases using the dynamic storage simulation program. The combination of significant factors was determined by the design of experiments. The fault simulation results suggested that the cases which showed the typical temperature profile variations were only a part, and that more than a half of them showed uncertain. The result obtained from one kind of fault was sometimes similar to the result obtained from another kind fault.

Then, it becomes necessary that certain numeric indices to identify the faults should be developed. Present paper reports a successful result for identification of simulated control faults using Fourier Transform combined with a fault mapping.

Table 1 Estimation table of thermal storage tank efficiency of the multi-connected complete mixing tanks type

FACTOR (I) (significant only)		LEVEL (j)			FACTOR. EFFECT(FF) levels			INTERACTION (IR)				
I	factor name	1	2	3	1	2	3	comb. X×Y	level of X	level of Y 1	2	3
B	Min/Max Load ratio	0.8	0.5	0.2	2.85	0.94	-3.79	B×D	1	-5.92	0.18	5.73
									2	-0.30	-0.49	0.79
									3	6.22	0.31	-6.51
C	Yes/No of CDT for CWV	yes (2)	yes (1)	No	9.47	1.76	-11.23	C×D	1	-5.03	-0.32	5.35
									2	-1.06	-0.23	1.29
									3	6.09	0.54	-6.64
D	CWV Load Ratio to Total Load	0.2	0.5	0.8	3.63	0.20	-3.83	C×F	1	1.93	-1.29	-0.64
									2	2.49	0.35	-2.84
									3	-4.42	0.94	3.48
F	Limit Temp. Differential Ratio	0.4	0.3	0.2	7.19	0.76	-7.95	F×G	1	-1.01	-3.05	4.06
									2	0.96	1.03	-1.99
									3	0.05	2.01	-2.07
G	No. of Tanks	40	20	10	6.20	-0.21	-5.98	Storage	96.41	+ ΣFF		
H	Operate hours	0-24	18-12	22-8	5.50	-7.29	1.79	Efficiency		+ ΣIR	(%)	

Note: The interactionsat the right side is the correction part of the proincipal effects due to the non-linear relations among factors. For example , if the first level of the B factor and the third level of the D factor are chosen, the integrated effect is calculated as +2.85-3.83+5.73.
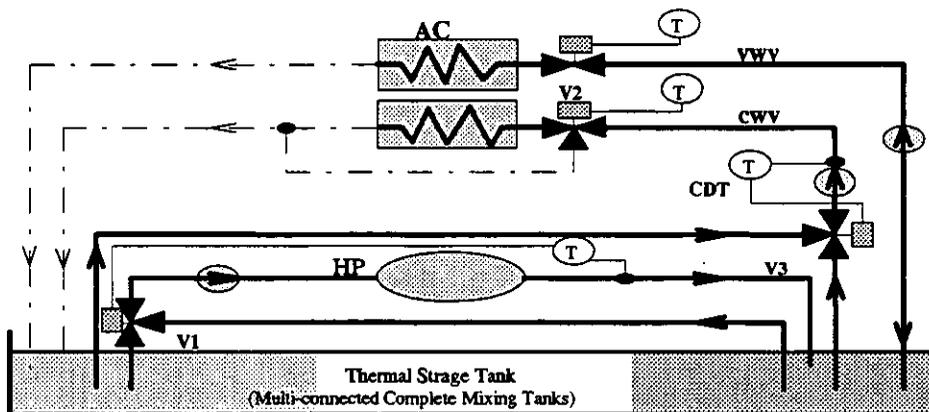
Figure 1 Simplified Thermal Storage Reference System Diagram

2. THERMAL STORAGE SYSTEM AND STORAGE EFFICIENCY

The reference thermal storage system was introduced at the meeting for Annex 25, ECBCS, IEA [3]. of which simplified diagram is shown in **Figure 1**. The estimation of the storage efficiency, composed of the significant factors only, is shown in **Table 1**. The thermal storage tank efficiency is defined as the actually available heat of storage with the limitation of maximum temperature rise at the delivery water tank over the nominal chiller outlet temperature to the nominal heat of storage in the water volume with constant temperature difference determined by the weighted coil design temperature difference at the peak load. The significant factors were analyzed and reduced into estimation tables [1][2] as shown in the table, for example, for the multi-connected complete mixing tanks type.

3 FAULT SIMULATION

3.1 SIMULATION ALGORITHM

The simulation algorithm of this type of tank was introduced elsewhere [1] by the author. Each tank is thought as having complete mixing characteristics and connected in a series.

3.1.1 Control algorithms

The principal control strategy is as follows. All the control loops were thought to have the PID action controller and no dynamic property was taken into account.

1) The V1 is manipulated to maintain constant chiller output temperature.

2) The capacity control of the chiller such as by the inlet vane is due only to the input power of the driver, not to the outlet temperature.

3) The maximum fall of the chiller inlet temperature below the designed temperature is determined as one of the parameters, called as the limited temperature difference ratio at the primary circuit, factor E, which is not significant.

4) The maximum rise of the chiller outlet temperature over the designed temperature should also be determined by the parameter called as the limited temperature difference ratio at the secondary circuit, the factor F.

5) The V2 is a temperature control valve due to the room air temperature, the outlet air temperature of the air handling unit (AHU), etc. Either two way or three way valves are used according to whether the variable water volume (VWV) system or the constant water volume (CWV) system are used. The factor B,C and D are concerned.

6) The valve V3, to which the factor C is concerned, is manipulated to maintain the delivery temperature to AHU in CWV system at a predetermined highest temperature, which are for keeping better storage performance. This control should not be used to VWV systems according to an expert knowledge.

3.1.2 Room temperature variations(RTV) and the heat exclusion rate(HER)

Simplified dynamic algorithm was introduced. The original program was developed to design an optimal storage

capacity at the given design conditions on the cooling load and the HVAC system and control design, so that it did not include the room model, and the room temperature was supposed to kept constant. It is correct as long as the water temperature for cooling is kept proper due to the premise of optimal design with no-fault operation. In order to make it possible to simulate the faulty condition, where the water temperature is expected to reach abnormal condition in which the heating and/or cooling

Figure 2 Simplified calculation model of RTV and EXR

load to keep the room temperature nearly constant cannot be excluded sufficiently, an simplified calculation model of the dynamic behavior was introduced as shown in **Figure 2**.

At the peak cooling condition the natural room temperature without any heat exclusion was supposed as $37°$ C in contrast to the normal cooling condition, that is $27°$ C. The RTV is then calculated in proportion to the ratio of the HER, which is calculated by the static coil model, to that of the peak one. When the load is not at the peak condition, the RTV is supposed as proportional to the HER ratio to the peak.

This model is not correct enough to simulate the RTV and HER themselves but precise enough to acquire the knowledge data for fault simulation based on the temperature profile transition.

3.1.3 Determination on optimal capacity and initial state for the fault simulation

Two days of simulation was supposed enough to determine the steady state and the limit temperature rise of the lowest temperature tank is examined if it reached the value just as predetermined[1]. The calculation is iterated using a kind of Newton Raphson Method until the error becomes within an allowable value.

After an optimal capacity of the tank is decided, two kinds of designated faulty conditions are given at the initial stage of simulation, that is at 0:00 in the morning, and the fault simulation begins only for one day for both conditions. The calculation results for each case are listed as follows.

    a) optimal capacity

    b) storage efficiency

    c) temperature profiles for the normal condition

    d) temperature profiles for two kinds of faulty conditions

    e) RTV and HER for each faulty condition

    f) inlet and outlet water temperatures across the chiller and the cooling coil

The fault simulation was carried out for the peak load day and a half load day in order to obtain wide range of knowledge data for annual operation.

3.2 FAULT CONDITIONS

The more a factor has a significant factorial effect, the more significant effects result when the factor goes out of order. The present paper pays attention to the importance of control strategies, because they have strong relations with the effects of these significant factors. Referring Figure 1 and Table 1, two kinds of faulty conditions are selected for examination as follows.

1) The valve V1 malfunctions to fix its position as sucking all of water from the last tank to which chilled water returns from AHU coils and no water flow from the initial tank. This fault is the typical control fault in the primary circuit and called as Fault ① in the following chapters.

2) Both the valve V2 for VWV system and V3, when it is provided, for the CWV system went out of order or installed another way by mistake. It was supposed, as the result, that the two way valve worked as if it was a three way valve and that the V3 did not work to bypass flow from the last tank. This is the typical fault in the secondary circuit and called as Fault ② in the following chapters.

3.3 SIMULATION RESULTS

The temperature profiles obtained from fault simulation for three cases among eighty one cases are shown in **Figure3** for the normal and two faulty conditions which are illustrated in both two dimensional and three dimensional coordinates. The time increment is shown as the parameter in the two dimensional graph, but it is shown on the third axis in the three dimensional graph. The most symptomatic phenomena for each fault are observed as follows.

3.3.1 Normal operation

In normal operation very few temperature profiles intersect each other. The peak temperature point of each profile is fixed at the position of the return tank, or the last tank. The highest temperature of the initial tank during a day is limited at the point of limited temperature rise following the design conditions.

3.3.2 Faulty operation ①

In the fault ① condition the profiles waves and the peak temperature points gradually move, so that the profiles intersects each other. Thus, the initial tank temperature rises as the peak point moves over the design condition. These symptoms do not always appear very clearly . When the design temperature difference for the primary side, i.e. the chiller side, and the secondary side, i.e. the AHU side, are similar, the symptom becomes unclear. So is the results when the CWV system load overwhelms the VWV system load.

Moreover, if the temperature rise in the initial tank appears in the evening when the cooling load profile is much alike triangle, that means small cooling load in the early morning and late evening, the faulty phenomena does not affect the room temperature as long as it is the first day of the faulty operation. If the fault is not detected and recovered, the
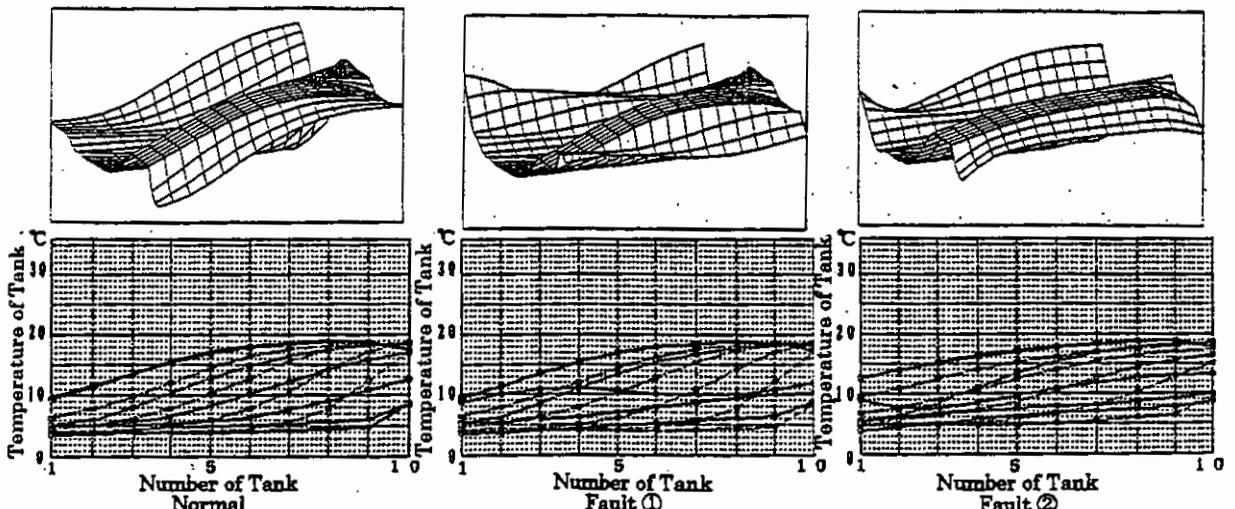
a)[Case18 (B1 C2 D2 F2 G3 H3)The effective volume of a tank:76.4m^3]
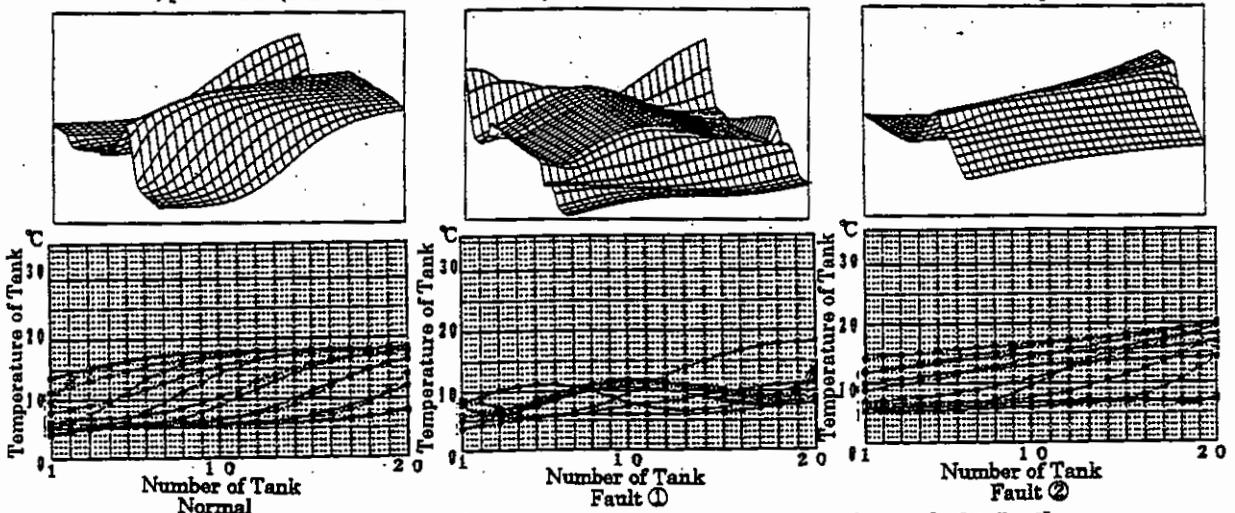
b)[Case49 (B2 C3 D2 F2 G2 H1)The effective volume of a tank:21.5m^3]
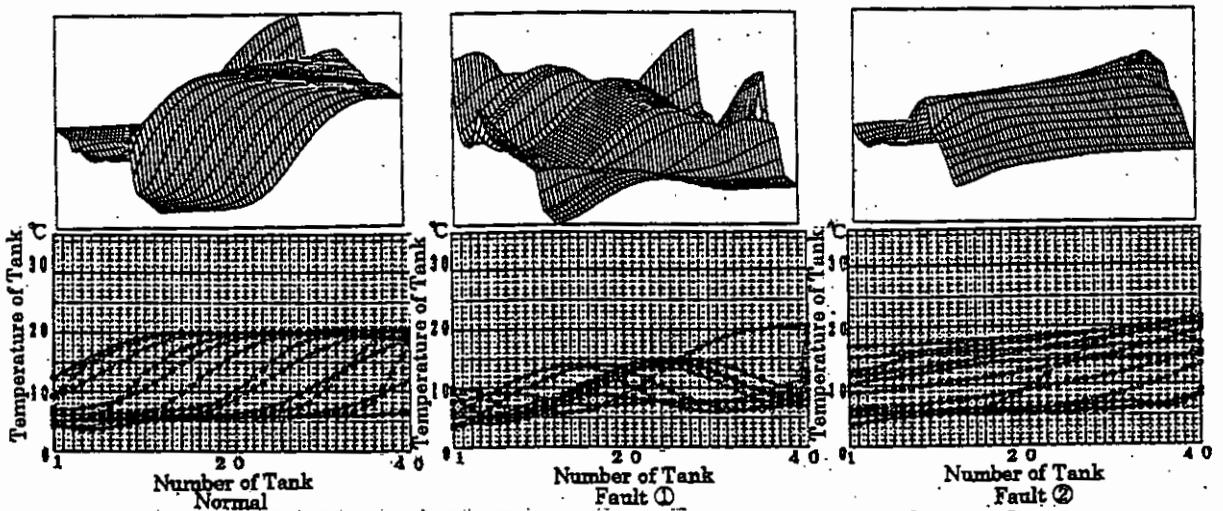
c)[Case35 (B2 C1 D2 F3 G1 H1)The effective volume of a tank:9.1m^3]

Figure 3 Temperature profiles of the tank in two and three dimensions

a)[Case18 (B1 C2 D2 F2 G3 H3)The effective volume of a tank:76.4m³]

b)[Case49 (B2 C3 D2 F2 G2 H1)The effective volume of a tank:21.5m³]

c)[Case35 (B2 C1 D2 F3 G1 H1)The effective volume of a tank:9.1m³]
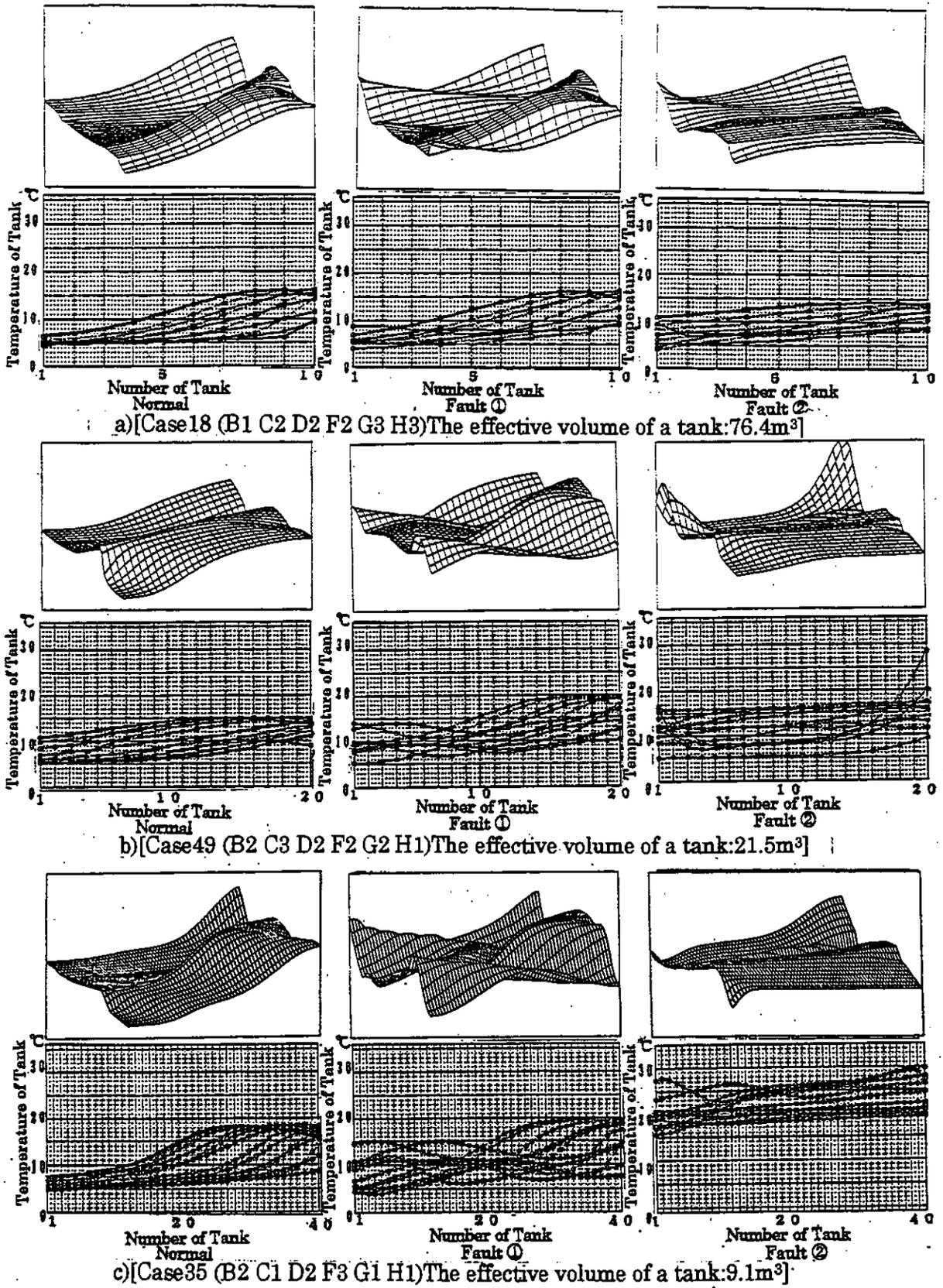
Figure 4 Temperature profiles of the tank in two and three dimensions for partial load

room condition will become vital in the following days.

3.3.3 Faulty operation②

In the fault ② condition the profiles become rather flat, resulting the temperature rise over the designated limited value in the initial tank. This is because the elevated temperature of return water at the partial load higher than the designed value, which is a typical phenomena of the VWV control using two way valve, does not appear. Again, it is obvious that the phenomena will not appear clearly if the cooling load has a flat pattern.

3.3.4 Partial load condition

In order to make sure the availability of the FDD measure during all seasons, a partial load (50%) condition was assumed and the same kind of fault simulations were executed. Some examples of the results are shown in Figure 4 for the same cases as shown in Figure 3. The storage capacity is large enough for partial load operation. The pattern of the temperature profiles are principally maintained for each kind of faulty case, though the temperature rise at the suction tank for the cooling coil is not high enough and the heat exclusion is almost fulfilled.

4. TWO DIMENSIONAL REAL FOURIER TRANSFORMATION

As seen in Figure 3 and Figure 4, the characteristics of the frequency and phase of the temperature profile pattern seem to have decisive factors to discriminate a fault from the other or from the normal. In order to quantify them the temperature profiles were analyzed into Fourier transformation. Two dimensional Fourier transform is expressed as Equation (1).

$$\begin{bmatrix} C \\ S \end{bmatrix}\begin{bmatrix} C \\ S \end{bmatrix}_{kl} = \frac{a_1 a_2}{NM} \sum_{r=0}^{N-1} \sum_{s=0}^{M-1} x_{rs} \begin{bmatrix} \cos\frac{2\pi k_r}{N} \\ \sin\frac{2\pi k_r}{N} \end{bmatrix}\begin{bmatrix} \cos\frac{2\pi l_s}{M} \\ \sin\frac{2\pi l_s}{M} \end{bmatrix} \qquad \begin{array}{l} k = 0, 1, 2\cdots N/2 \\ l = 0, 1, 2\cdots M/2 \end{array} \tag{1}$$

$$a_1 = \begin{cases} 2, & 0 < k < N/2 \\ 1, & k = 0, N/2 \end{cases} \qquad a_2 = \begin{cases} 2, & 0 < l < M/2 \\ 1, & l = 0, M/2 \end{cases}$$

The x is the temperature of the tank water [° C] and it was supposed to be two dimensional real periodic function with the time and number of tanks as variables. The C and S are sine and cosine components. The **figure 5** shows the three dimensional temperature profiles for three cases of transformed result for the same cases of simulated profiles in Figure3, the explanation how to read the figures is shown in **Figure 6**. The horizontal plane has degrees of sine and cosine components in X and Y axis, while the vertical axis is the Fourier value. The horizontal plane are divided into four zones each of which is dedicated to display the product of sine (S) and cosine (C) components shown as *CkSl*, for example, corresponding to Equation (1).

The results shows that composition of both the sine and cosine components are different between the normal the faulty for the identical cases but that the difference is not so large in Case 18. The difference mainly exists in the value, distribution and the degree of the sine and cosine components. This suggests a possibility of quantification using these factors as parameters for FDD.

5. PARAMETERS FOR QUANTIFICATION OF FAULTS

As shown above, factors describing the variation of the frequency and phase of the temperature profiles of the thermal storage tank are the value of sine and cosine components, their distribution and degree of the frequency of Fourier transform. Quantification trial using these factors were conducted as follows.

[case18] Normal	Fault ①	Fault ②
[case49] Normal	Fault ①	Fault ②
[case35] Normal	Fault ①	Fault ②
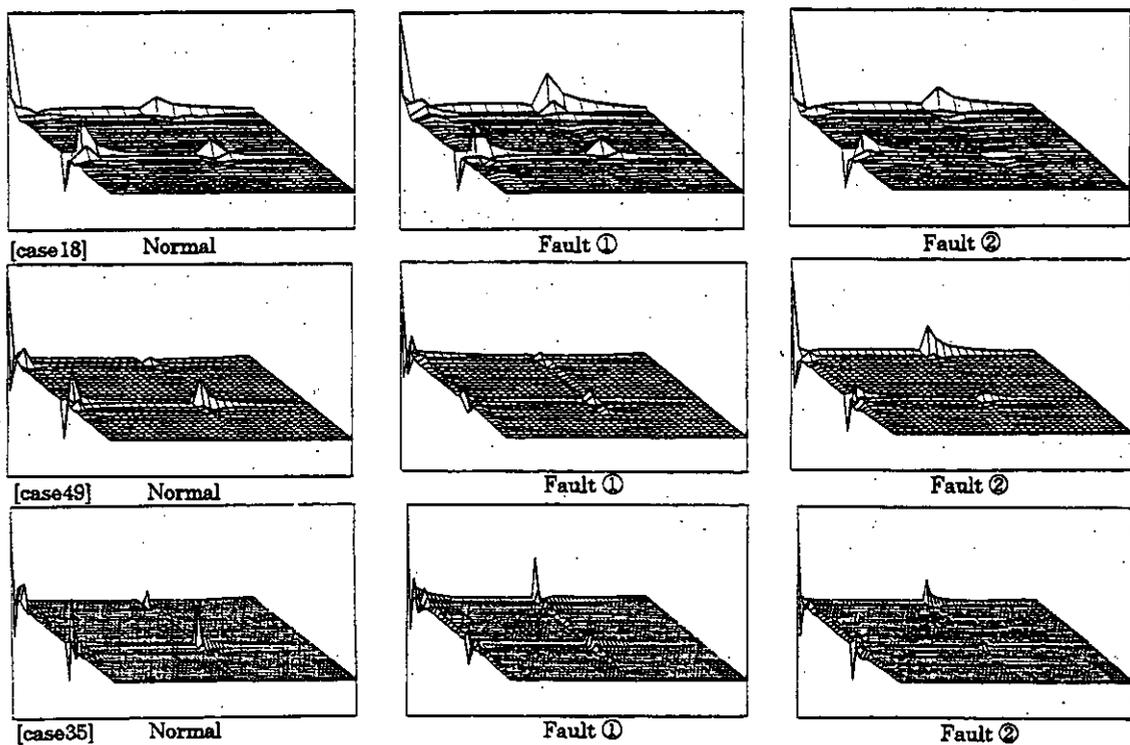
Figure 5 Graphical display of two dimensional Fourier Transformation

(a) 3-D Time domain

(b) 3-D Fourier domain
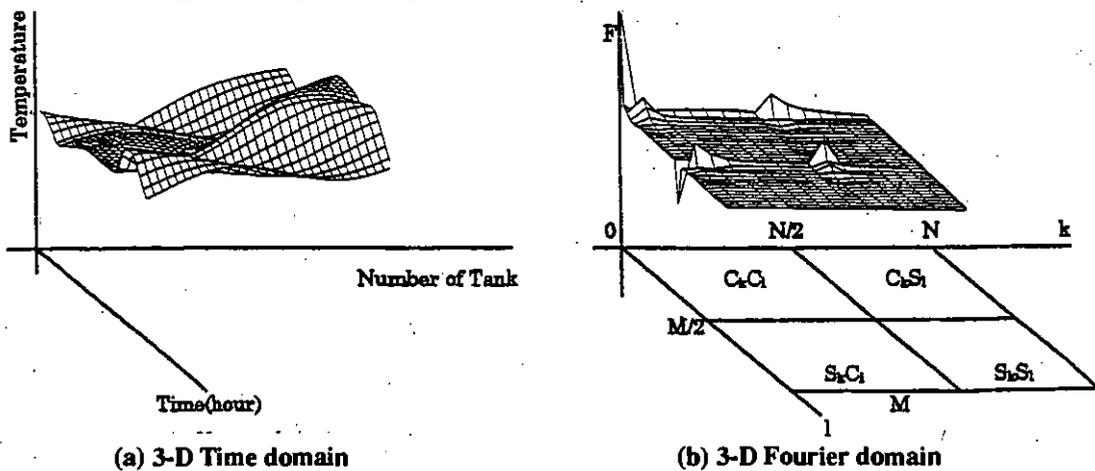
Figure 6 Understanding of the graph in Figure 3, Figure 4 and Figure 5

5.1 PARAMETERS

The twenty three parameters composed from real values as well as the Fourier values were supposed to check the significance for FDD. The quantity of the parameters was reduced by checking the internal correlations and statistical significance. Finally two significant parameters was derived as the most valuable for recognizing difference of patterns between the normal and the faulty ones. One of the significant characteristic parameters is a maximum component value of a part of the matrix F and another is the maximum of a converted matrix which is chosen as shown in 5.1.1.

The original two dimensional Fourier real transformation matrix F used here is composed of four zones, each of which corresponds to the graphical display shown in **Figure 5**, except for using N and M instead of $N/2$ and $M/2$.

5.1.1 Maximum value of Fourier transform

The difference among the normal and the faulty was observed especially at *the* C_kS_l zone in Figure 5. Then, the maximum value of the components of the zone defined as described in Equation (2) is selected as one of the parameters.

$$F_m = \text{Max} \ (F_{kl}) \qquad\qquad k = N/2, \cdots, N \qquad l = 1, \cdots, M/2 \tag{2}$$

5.1.2 Maximum value of frequencies over threshold along the X, the space, axis

It is observed that the profiles along the X axis, that is the space axis, varies considerably, if compared with the normal one, especially in the fault ② cases. Therefore, the maximum value F_l of the frequencies over a certain threshold value along the space axis was selected as a candidate of the significant parameter as shown in Equation (3) and (4).

The components F_{akl} of the converted matrix F_a are calculated by filtering with the threshold value b from the original Fourier matrix F and has the value of either k, $k-N/2$ or 0 depending on its magnitude as shown in Equation (4).

$$F_l = Max(F_{a_{kl}}) \qquad\qquad k=1,2,\cdots N, \qquad l=1,2,\cdots M \tag{3}$$

$$F_{a_{kl}} = \left(\begin{pmatrix} k & k < \dfrac{N}{2} \\ k-\dfrac{N}{2} & k \geq \dfrac{N}{2} \\ 0 & \end{pmatrix} \begin{array}{l} |F_{kl}| > b \\ \\ |F_{kl}| \leq b \end{array} \right) \qquad k=1,2,\cdots N \qquad l=1,2,\cdots M \tag{4}$$

5.2 MAPPING DATA FOR FDD

The **figure 7** and **Figure 8** show mapping results of all the data obtained from fault simulation on two dimensional planes, using the parameters described above, for 76 cases at peak cooling day and a half load day, respectively. The horizontal axis of it is F_m and the vertical axis is F_l. Both of them are normalized using the mean values of all the

cases in order to make applicable to any partial load conditions and to any design conditions.

The distributions of the coordinates among normal cases, fault ① cases and fault ② cases are expected to be different each other, which means successful FDD.

6. THE RESULTS AND PERFORMANCE OF FDD

How to group the data on the map for each normal and fault case is the next question. The performance of FDD depends upon the quality of grouping method. Two kinds of grouping was compared.
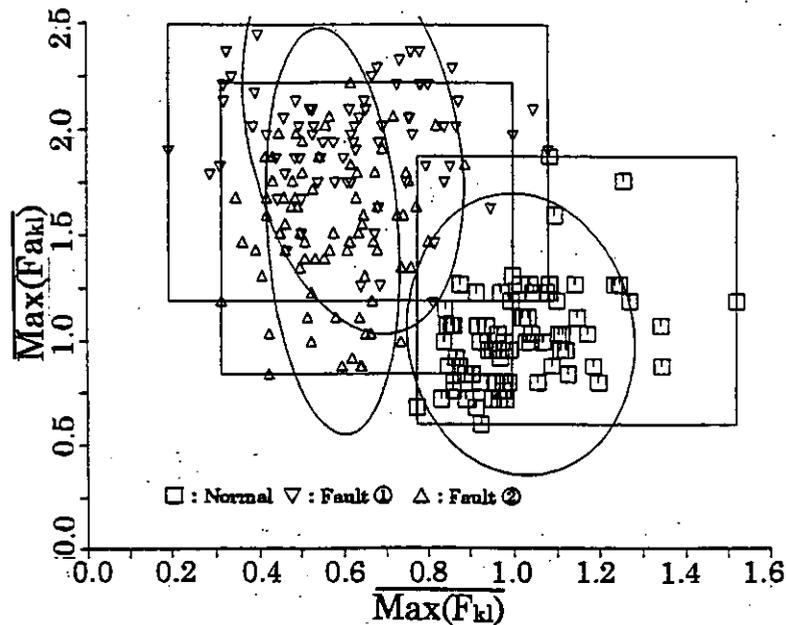
Figure 7 Mapping and Grouping results of Characteristic parameters for Peak Load
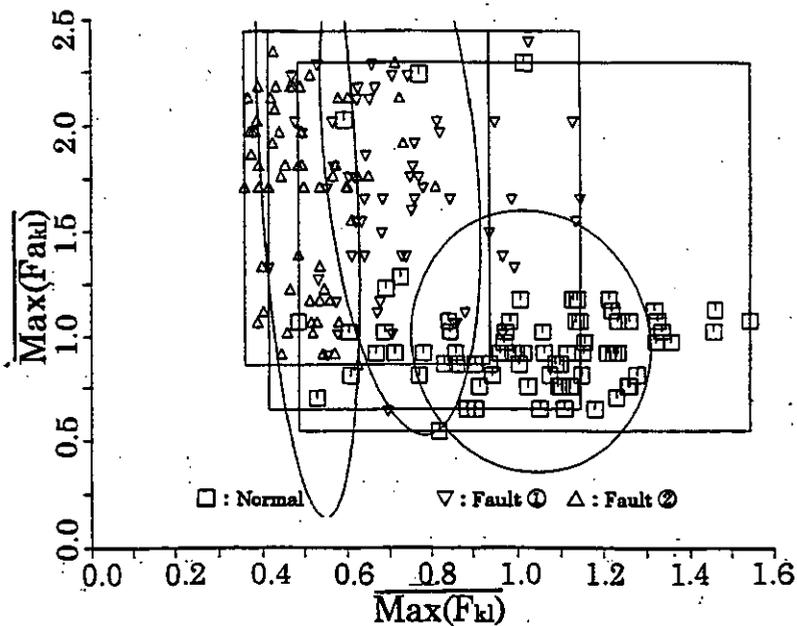
Figure 8 Mapping and Grouping results of Characteristic parameters for a Half Load

6.1 RECTANGULAR AREA GROUPING

The simplest way of grouping for such clear cases as seen in Figure 7 and Figure 8 for localizing faulty cases from the normal cases is to group with rectangular area on the plane by hand [4]. The area was decided to include all the data belonging to the same kind of case, that is, for the normal , for the fault ① and for the fault ②case. Overlapping among three areas is clearly seen and it seems difficult to classify each other among two kinds of faults in that overlapped zone. However, this method is very simple and has sufficient performance in the fault detection only as shown later.

6.2 MAHALANOBIS' GENERALIZED DISTANCE

The bivariate normal distribution $N(\mu, \sigma^2)$ was supposed for each three group data. The center of the distribution and the probability ellipse are calculated from the enough volume of the sampled data, every group of which has more than seventy data. The Mahalanobis's distance D is calculated by the following equation (5) and (6).

$$D^2 = (x-\mu)^T \Sigma^{-1}(x-\mu) \tag{5}$$

$$x = \begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}^T$$

$$\mu = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_p \end{bmatrix}^T$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & \sigma_1\sigma_p \\ \sigma_2\sigma_1 & \sigma_2^2 & \cdots & \sigma_2\sigma_p \\ \vdots & & & \vdots \\ \sigma_p\sigma_1 & \sigma_p\sigma_2 & \cdots & \sigma_p^2 \end{bmatrix} \tag{6}$$

The ellipses shown in these figures were drawn for a half of the standard deviation σ. The most probable answer of grouping for any data is the one that the Mahalanobis' distance is the shortest.

6.3 COMPARISON OF THE PERFORMANCE OF FDD

6.3.1 Effect of the threshold value b

The results of the performance obtained from the two grouping method for the peak load day data, shown in Figure 7, is compared in Table 2 for four cases of threshold value b in equation (4), that is, 0.1, 0.3, 0.5 and 1.0. The fault ② is discriminated from the normal on the mapped area with more than 99% probability of fault detection notwithstanding

Table 2 Performance of FDD for the Peak Load

Threshold	Fault ①		Fault ②	
b	Rectangular	Mahalanobis'	Rectangular	Mahalanobis'
0.1	0.92	0.99	0.97	1.0
0.3	0.91	0.93	0.96	1.0
0.5	0.87	0.89	0.96	1.0
1/0	0.89	0.99	0.96	0.99

the b value for the Mahalanobis' distance method and more than 96% for the rectangular area method.

In case of the fault ① the probability of detection reduces a little. Anyway it was disclosed that the FDD based on the Mahalanobis' generalized distance gives a better performance, because it has the statistical background.

6.3.2 Effect of the load ratio on the performance of detection

The difference of the performance detection due to the grouping method becomes clearer when the heat load is partial, as shown in Table 3. As it is clear from Figure 8, the reason is because the normal data in this case much more widely scatters on the map. However, the Mahalanobis' distance method shows almost complete detection in case of the fault ② and about 85 % successful even in the case of the fault ①, while the rectangular method has a very poor performance.

Table 3 Performance of FDD for a half load

Method	Rectangular Method				Mahalanobis' Distance Method			
Threshold d	Fault ①		Fault ②		Fault ①		Fault ②	
b	peak load	half load	peak load	half load	peak load	half load	peak load	half load
0.1	0.92	0.68	0.97	0.41	0.99	0.89	1.0	1.0
0.3	0.91	0.53	0.96	0.37	0.93	0.84	1.0	1.0
0.5	0.87	0.71	0.96	0.41	0.89	0.84	1.0	1.0
1.0	0.89	0.63	0.96	0.17	0.99	0.86	0.99	1.0

7. CONCLUSION

In order to identify selected thermal storage operation faults, mapping data using appropriate parameters based on Fourier transform on the two dimensional plane and grouping by the statistical principle have proved fairly effective in detecting and localizing faults as follows.

1)Two kinds of the control faults, one is malfunctioning of the three way valve at the primary circuit called as fault①

and the other is malfunctioning of the two way valve at AHU and the three way valve for the CDT control in CWV system called as **fault ②**, were supposed to take place at 0 o'clock after normal operation.

2) Dynamic fault simulation of a reference thermal storage system was carried out and data for FDD was acquired.

3) Variations of the phase, frequency and amplitude was analyzed by way of two dimensional real Fourier transform.

4) In order to quantify the temperature profiles for FDD, two significant parameters composed from Fourier value were reduced after checking twenty three parameters with the internal correlation and the statistical significance.

5) Two kind of grouping methods of data map were proposed and the performance of FDD was inspected. One is the simple rectangular area method and the other is the statistical method supposing the normal distribution.

6) The threshold value for the second parameter did not affect the performance for fault detection but it did affect and had a optimal value for localizing between the fault ① and fault ②.

7) Classifying by the Mahalanobis' generalized distance on the probability ellipse showed a satisfactory results of FDD compared with the simple rectangular grouping method.

8) The probability of FD from the normal was almost complete for both peak and a half load, if the Mahalanobis' distance method is applied, except for the fault ① at a half load.

9) The proposed method is only applicable to those faults which clearly affect the temperature profiles of the thermal storage tank. However, it is believed that most of thermal storage faults affect the temperature profiles, so that the pattern recognition method based on the temperature profiles is considered powerful.

Acknowledgement

References

[1]N. Nakahara, K. Sagara:WATER THERMAL STORAGE TANK:PART 1-BASIC DESIGN CONCEPT AND STORAGE ESTIMATION FOR MULTI-CONNECTED COMPLETE MIXING TANKS, ASHRAE TRANS-ACTIONS, Vol.94.2, 1988

[2]N. Nakahara, K Sagara, M. Tsujimoto:WATER THERMAL STORAGE TANK:PART 2-MIXING MODEL AND STORAGE ESTIMATION FOR TEMPERATURE-STRATIFIED TANKS, ASHRAE TRANSACTIONS, Vol.94.2, 1988

[3]IEA Annex 25: Building Optimisation and Fault Diagnosis system concept, IEA, 1993.10

[4]Proceedings of the Pan Pacific Symposium on Building and Urban Environmental Conditioning in Asia, Nagoya, 1995.3

APPLICATION OF CLASSIFICATION FUNCTIONS TO CHILLER FAULT DETECTION AND DIAGNOSIS

Meli Stylianou, Research Engineer, EDRL-CANMET

Abstract

This paper describes the application of a Statistical Pattern Recognition Algorithm (SPRA) to fault detection and diagnosis of commercial, reciprocating, chillers. The developed fault detection and diagnosis module has been trained to recognize five distinct conditions, namely, normal operation, refrigerant leak, restriction in the liquid refrigerant line and restrictions in the water circuits of the evaporator and condenser. The algorithm used in the development is described, and the results of its application to an experimental test bench are discussed.

Experimental results show that the SPRA provides an effective way of classifying patterns in multi variable, multi-class problems without having to explicitly use a rule-based system.

Introduction

Vapour compression refrigeration systems constitute the largest portion of commercial and industrial refrigeration capacity, accounting for an important portion of energy consumption in these sectors.
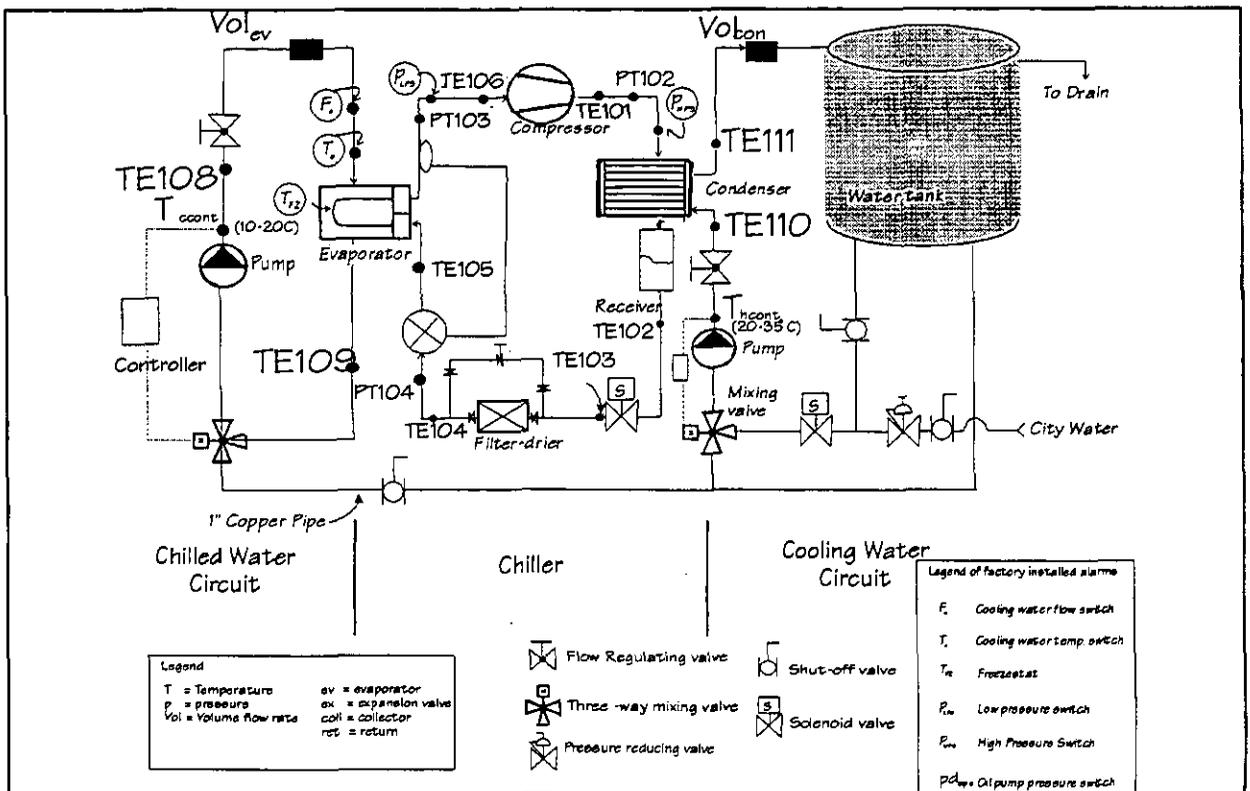
Figure 1 Testbench schematic

In large office buildings for example, it is estimated that 10% to 25% of the total electricity consumption can be attributed to cooling systems alone (Huang, Akbari, Reiner and Ritschard, 1991). Moreover, these percentages can be significantly higher if a cooling system is operating at low performance levels due to the presence of faults (Herzog and LaVine, 1992).

Fault detection systems for commercial chillers currently focus primarily on preventing mechanical failures, which is generally achieved through the use of switches that cut out the unit when temperatures and/or pressures exceed preset limits. Generally, these systems do not provide direct information as to the "health" of the chiller prior to its shutdown, resulting in unexpected periods of unavailability. In addition this luck of information leads to extended periods during which the unit operates under faulty conditions that lead to premature equipment failures and excessive energy consumption. However, as chillers become better instrumented, advanced lowcost fault detection and diagnosis modules become increasingly more attractive.

Fault detection and diagnosis in refrigeration equipment has been discussed previously in Grimmelius et al (1994), Rossi and Braun (1995) and Stylianou and Nikanpour (1996). The present paper describes a method that uses a statistical pattern recognition algorithm to identify a number of conditions commonly encountered in commercial chillers.

Test Unit
Experiments for the development of the SPRA were carried out on a refrigeration test unit, shown in Figure 1. It is based on a commercially available reciprocating chiller using refrigerant R22 with a cooling capacity of 17.6 kW (5 RT).

The chiller components include a two-cylinder semi-hermetic compressor, a cleanable shell and tube type condenser with water circulating through the tubes, and a direct expansion shell and tube evaporator. The chiller is equipped with a thermal expansion valve.

The test conditions for the chiller are achieved using a single storage tank and two three-way mixing valves serving two PID controllers. In order to have a controlled set point temperature at the inlet to the evaporator and condenser, the three-way valve supplying the evaporator mixes warm water from the storage tank with return water, while the valve serving the condenser mixes city water with storage tank water.

Instrumentation and Data Acquisition System
The instrumentation of the test bench is composed of eleven platinum Resistance Temperature Detectors (RTDs), four pressure transducers and two flow meters as given in Table 1.

With the exception of the RTD which measures the crankcase oil temperature (TE107), dry surface-mounted RTDs are used. This type of installation was chosen not only to avoid problems with refrigerant leaks, but also to duplicate the most likely way RTDs would be installed in the field. Pressures are measured using optical pressure transducers, mounted in the manner usually employed for pressure gauges, and as close to the desired point as conditions would allow. Flow rates are measured using positive displacement flow meters.

Data acquisition is carried out using a microcomputer-based system which enables the user to establish sampling frequencies of up to 1Hz. The output data files can be stored on the PC serving the test unit and can also be transferred to other platforms with specialized software applications for further analysis.

Table 1 Measured variables

Sensor	Measurement	Precision
TE101	Discharge temperature, °C	$\pm(0.25+0.0042T)$
TE102	High pressure liquid line temperature, °C	$\pm(0.25+0.0042T)$
TE103	High pressure liquid line temperature (before filter dryer), °C	$\pm(0.25+0.0042T)$
TE104	High pressure liquid line temperature (after filter dryer), °C	$\pm(0.25+0.0042T)$
TE105	Low pressure liquid line temperature, °C	$\pm(0.25+0.0042T)$
TE106	Suction line temperature, °C	$\pm(0.25+0.0042T)$
TE107	Crankcase oil temperature, °C	$\pm(0.25+0.0042T)$
TE108	Evaporator entering water temperature, °C	$\pm(0.25+0.0042T)$
TE109	Evaporator leaving water temperature, °C	$\pm(0.25+0.0042T)$
TE110	Condenser entering water temperature, °C	$\pm(0.25+0.0042T)$
TE111	Condenser leaving water temperature, °C	$\pm(0.25+0.0042T)$
PT101	Crankcase oil pressure, kPa	±6.5 kPa
PT102	Discharge pressure, kPa	±6.5 kPa
PT103	Suction pressure, kPa	±6.5 kPa
PT104	High pressure liquid line pressure, kPa	±6.5 kPa
FT101	Condenser water flow rate, l/s	±0.077 l/s
FT102	Evaporator water flow rate, l/s	±0.077 l/s

EXPERIMENTAL METHODOLOGY

The unit was used to map the normal range of operating conditions for commercial chillers and to simulate selected faults. The normal operating envelope of the unit was mapped by changing the setpoints for the entering water temperatures at the evaporator and the condenser, simulating different cooling and chilled water return conditions. These conditions included varying the cooling water temperature from 22°C to 34°C, and the chilled water entering temperature from 10°C to 15°C. The flow was set at 1.0 l/s for both the condenser and evaporator for all experiments.

A study (Stylianou and Scott, 1993) identified the most common faults occurring in commercial vapour compression units. Of these four were chosen to demonstrate the SPRA and are shown in Table 2.

Table 2: Fault Types

Fault Type	Fault Simulation
Refrigerant leak	Removal of refrigerant
Refrigerant line flow restriction (Plugged filter-drier, obstructions in the piping etc.)	Throttling of line after condenser
Condenser water side flow resistance (Pump fault, fouling, etc.)	Reduction of water flow
Evaporator water side flow resistance (Pump fault, fouling, etc.)	Reduction of water flow

Chiller Reference Model

The SPRA uses a statistical chiller model to predict the temperatures and pressures at different points in the refrigeration circuit (Table 3). This model is developed as a set of bilinear equations each representing a variable of interest. The set of equations is developed using operating data and fitting them using multiple linear regression.

This set of regression equations is of the form

$$y_i = \beta_0 + \beta_1 TE108 + \beta_2 TE110 \tag{1}$$

where y_i is the calculated reference value for the dependent variable, β_0, β_1 and β_2 are the regression coefficients. Estimates for β_0, β_1 and β_2 were calculated using a multivariate least-squares method (Jenrich, 1977a).

Table 3 Dependent Variables

1. TE101   Discharge temperature, °C (°F)
2. TE102   High pressure liquid line temperature, °C (°F)
3. TE105   Low pressure liquid line temperature, °C (°F)
4. TE106   Suction line temperature, °C (°F)
5. TE109   Evaporator leaving water temperature, °C (°F)
6. TE111   Condenser leaving water temperature, °C (°F)
7. PT102   Discharge pressure, kPa (psi)

The results of the regression analysis are shown in Table 4

Table 4 Results of Regression Analysis

Variable	Adjusted R^2	Std. Error of Estimate
TE101	0.99917	0.11280
TE102	0.99987	0.03416
TE105	0.99810	0.63781
TE106	0.99480	0.10531
TE109	0.99985	0.01872
TE111	0.99996	0.01969
PT102	0.99892	0.51470
PT103	0.99879	0.11641

Derivation of classification functions

In order to prepare the data for the derivation of the classification functions, the difference between the predicted and measured values for each of the variables in Table 3 is computed. These differences, also known as innovations, are the input to the SPRA and are used to generate the patterns which identify the appropriate faults.

Stylianou and Nikanpour (1996) have developed a pattern recognition module which used a rule-base as a means to recognize these patterns, but this approach was shown to have limited success in correctly classifying patterns resulting from small faults. This was primarily due to the "heuristic" nature of the discriminating functions.

In order to improve the performance of the module, the rule-based pattern recognition approach was replaced with a statistical one. This approach allows, under certain conditions, the optimal placement of the discriminant functions resulting in improved performance. The SPRA used is based on the Bayesian Decision Rule which states that for a set of classes whose *a priori* probability is q, a measurement vector x belongs to class (or fault) j if and only if:

$$q_j p_j(x) > q_k p_k(x) \tag{2}$$

where $p_j(x)$ is the conditional probability that x belongs to class j.

In order to determine the proper classification, the SPRA uses the statistical parameters of training data to derive a family of classification functions. The parameters extracted are

1. The mean vector for the s variables of each condition c

$$\overline{x}^c = (\overline{x}_{i=1}^c, \ldots, \overline{x}_{i=s}^c) \tag{3}$$

2. The covariance matrix

$$\hat{\Sigma} = \frac{\sum_{c=1}^{w} \sum_{t=1}^{n^c} (x_t^c - \overline{x}^c)^2}{(n - w)} \tag{4}$$

where w is the number of conditions
n is the total number of samples
n^c is the total number of samples for condition c

The statistical properties of the training data, defined by equations (3) and (4), are used to define a set of five classification functions, one for each condition of interest. The set of classification functions may be derived as follows:

Let d_{jk} be the hyperplane which separates class j from class k. For normally distributed data, this hyperplane is (Anderson, 1958):

$$d_{jk}(x) = \log \frac{p_j(x)}{p_k(x)} = [x - \tfrac{1}{2}(\overline{x}^j + \overline{x}^k)]' \hat{\Sigma}^{-1}(\overline{x}^j - \overline{x}^k) \tag{6}$$

If the *a priori* probabilities q are known, x is classified to the class j if

$$d_{jk}(x) = \log \frac{p_j(x)}{p_k(x)} > \log \frac{q_k}{q_j}, \qquad k = 1, \cdots, m, \quad k \neq j. \tag{7}$$

In the case where the *a priori* probabilities are equal, the right-hand side of the inequality (7) reduces to zero. x is then classified to the class with the highest *a posteriori* probability $p(x)$.

An alternative approach would be to let individual scoring for each class determine the appropriate classification. The classification functions that are used to estimate the scoring of each class are derived as follows. As a first step, d_{jk} is separated into two linear components:

$$d_{jk}(x) = d_j(x) - d_k(x) \tag{8}$$

From equation (7),

$$d_{jk}(x) = \left(d_j(x) - d_k(x)\right) > \log \frac{q_k}{q_j}, \qquad k = 1, \cdots, m; \quad k \neq j \tag{9}$$
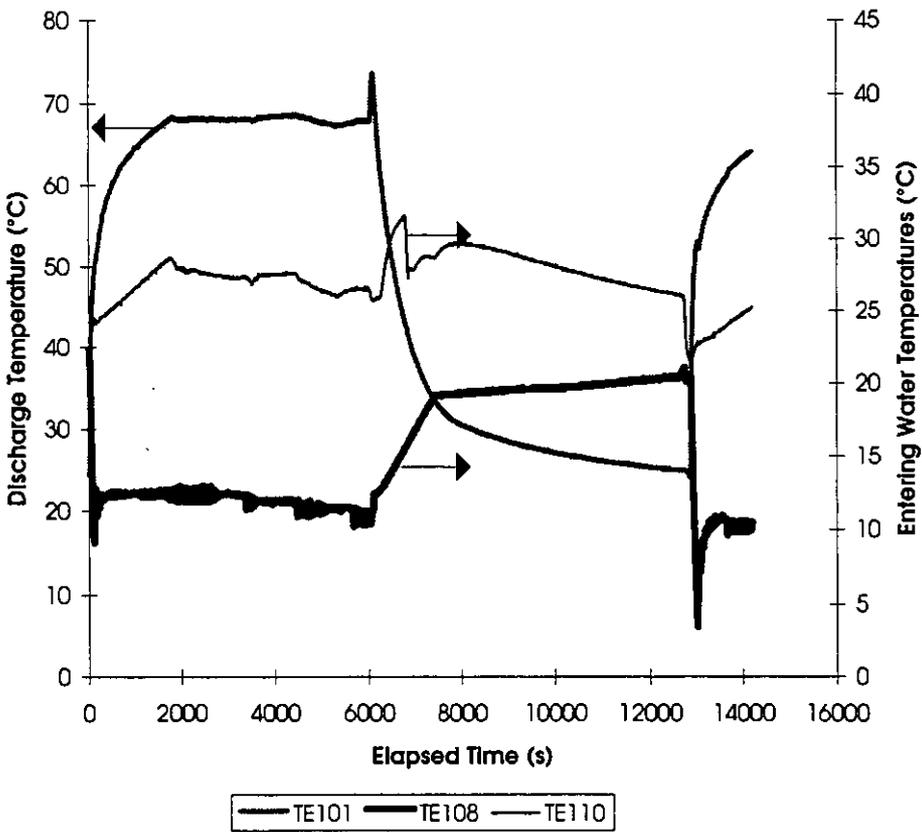
Figure 2 Selected chiller temperatures

Monitored Variables
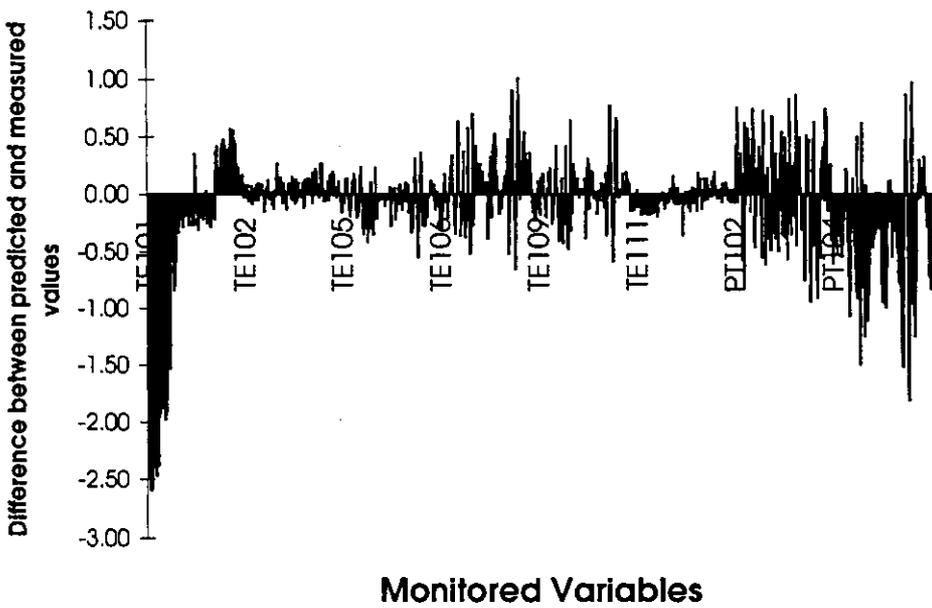
Figure 3 Input to the SPRA

$$\therefore d_j(x) > \log \frac{q_k}{q_j} + d_k(x), \qquad k = 1, \cdots, m; \quad k \neq j \tag{10}$$

Therefore, if the *a priori* probabilities are assumed equal, x is classified to the classification function with the highest score. From equations (6) and (8), the score for class j is calculated as

$$d_j(x) = (x - \tfrac{1}{2}\overline{x}^j)\hat{\Sigma}^{-1}\overline{x}^{j^T} \tag{11}$$

There is therefore a one-to-one correspondence between the classification functions and classes. These classification functions, shown to be a complete set (Jennrich, 1977), are therefore equivalent to the more conventional set of discriminant functions described by equation (6).

Experimental Results

Figure 2 displays the monitored values of the entering water temperatures at the condenser (TE108) and the evaporator (TE110), as well as the compressor discharge temperature (TE101) for a test during which the chiller was operating under normal conditions. The data from this test for each of the measurement points identified in Table 3 are the inputs to the Fault Detection and Diagnosis Module which, using a steady state detector (Stylianou and Nikanpour, 1996), extracts the samples identified as "steady state." These points are subsequently compared to the predictions of the statistical chiller model and the differences between the two are used as the input to the SPRA.

The input to the SPRA is shown in Figure 3, while the input's statistical properties are shown in

Table 5   Statistical Properties of Input to SPRA

	Mean	Min	Max.	Std.Dev.	Skewness	Kurtosis
ER101	-.046489	-2.61190	1.155498	.477772	-2.54940	11.26872
ER102	-.026796	-2.59047	.958786	.460570	-2.93833	13.31798
ER105	-.034628	-2.47138	.838622	.460088	-2.94957	13.25960
ER106	-.012320	-2.38847	.821476	.457748	-2.81272	12.57438
ER109	-.028462	-2.43343	1.146987	.469913	-2.46436	10.62305
ER111	-.047151	-2.48689	1.284156	.469586	-2.45818	10.76554
PT102	-.063986	-2.48689	.745764	.443109	-2.93600	13.16193
PT104	-.067563	-2.58546	1.132615	.425915	-2.71624	13.23883

Table 5. Since the chiller is operating in a normal mode, the input to the SPRA is expected to be white noise resulting from measurement and modelling errors. As indicated in Table 5, however, the distribution of the data deviates from this expectation as well as from the normality assumptions
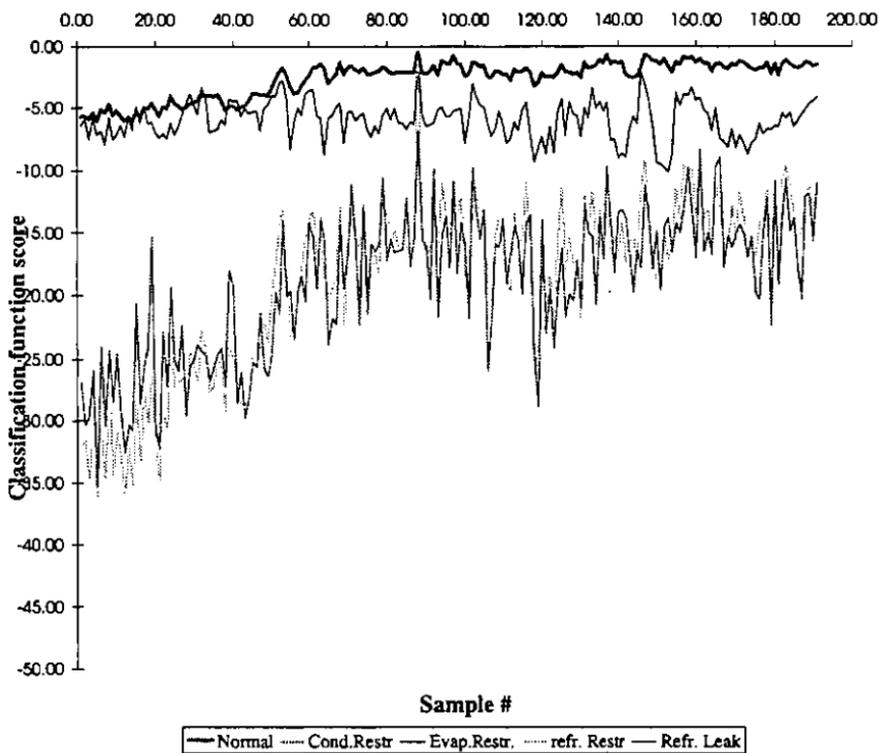
Figure 4a: Normal Operation: Classification Function Scores
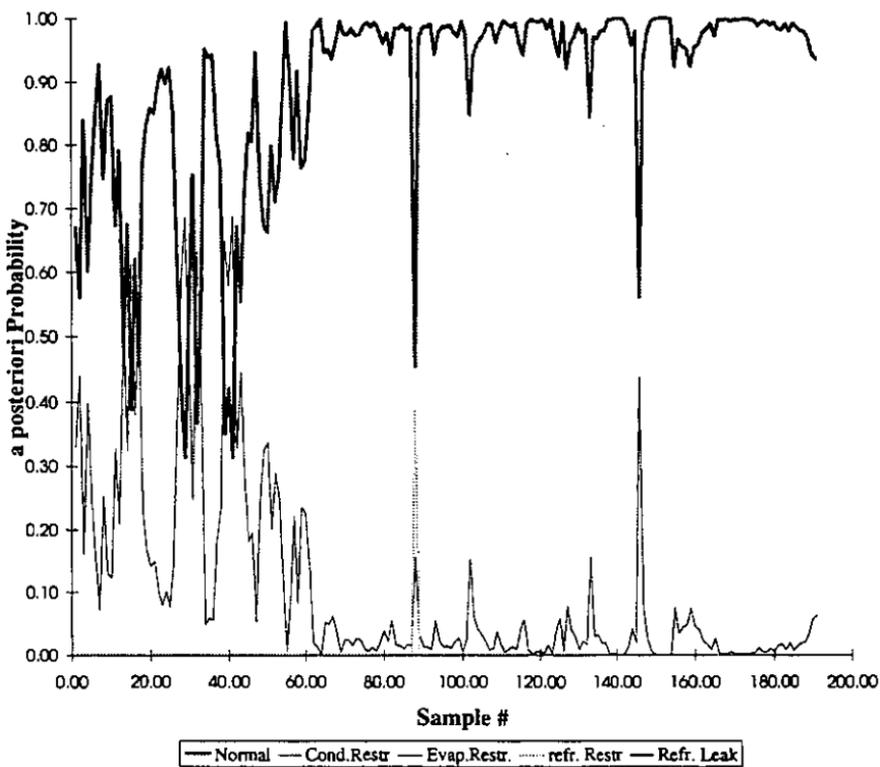
Figure 4b: Normal Operation: A posteriori Probabilities
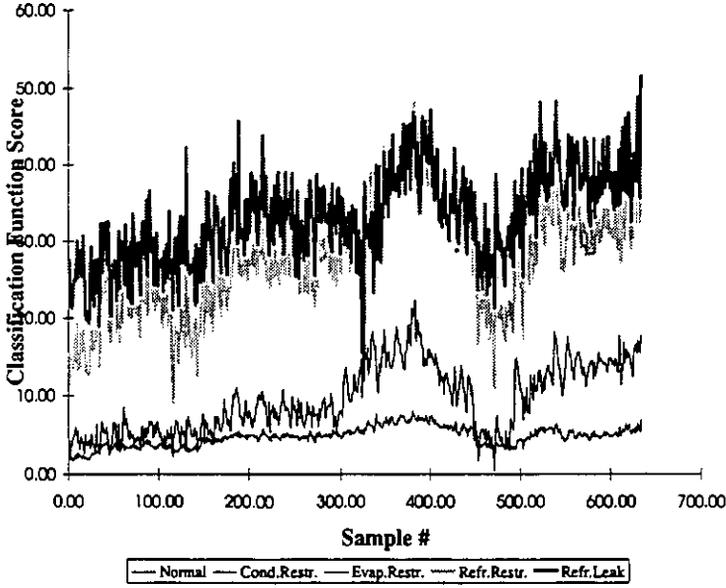
Figure 5a: Low Refrigerant Charge:Classification Function Score
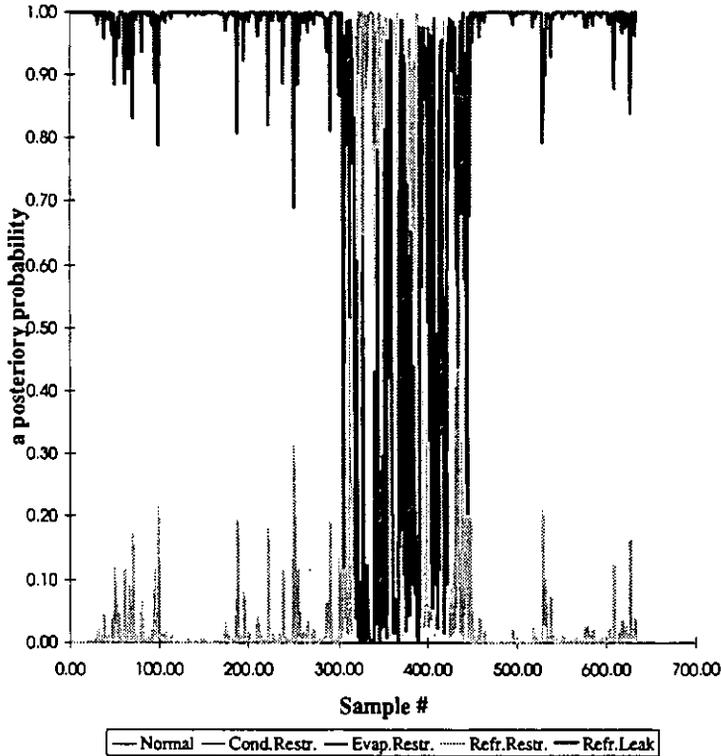
Figure 5b: Low Refrigerant Charge:A priori Probabilities

of the classification functions.

Where as typically normal distributions are symmetrical, the distribution of the input data appears to deviate from this symmetry as indicated by the skewness values. In addition, the values for the kurtosis identify the shape of the distribution as "sharper" than in the ideal curve. These deviations from the normality assumptions, however, do not appear to affect the performance of the classification functions as shown by the classification of this test in Figure 4.

Figure 4a displays the scores for each of the classification functions, while Figure 4b displays the equivalent probabilities, always under the assumption that the *a priori* probabilities for each of the conditions are equal. This assumption, however, biases the probabilities in favour of the faulty conditions resulting in some misclassifications (4.7%). This bias may be compensated for by including an appropriate threshold limit incorporating the *a priori* probabilities q of each of the conditions as shown in equations (7) and (8).

Figure 5a shows the result of applying the classification functions to test bench data generated when the chiller is operating under low refrigerant charge. As can be seen the SPRA correctly identifies the fault as refrigerant leak by assigning to it the largest classification function score, and the highest *a posteriori* probability (Fig. 5b). As expected the SPRA assigned the next highest probability to the fault class for restriction in the refrigerant line, and in some cases it misclassified the samples (11.3%).

The difference in magnitude of the probability between the two classes was significantly increased when a measure of the liquid line subcooling was included in the pattern. This measure, in the form of a third pressure measurement, placed after the condenser, increased the discriminating power of the SPRA and allowed the clear differentiation between faults caused by refrigerant leak and those caused, for example, by blocked filter/dryer. The impact of the additional pressure measurement is shown in Figure 6. As can be seen in this figure the probability (Fig. 6b) and the classification function scores (Fig. 6a) identify refrigerant leak as the faulty operation from where the samples were extracted.

The influence of the additional pressure measurement is also demonstrated by the performance of the SPRA when a gradual restriction in the refrigerant line is introduced. As seen in Figures 7 and 8, this gradual increase in restriction is reflected in both the case where the measure of subcooling is included and in the case where it is not. It is clear, however that in the case where the subcooling effect is taken into account, the probability assigned to the correct fault class is significantly higher than when it is not.

In addition to the above-mentioned faults, the SPRA is also capable of identifying water circuit problems in the cooling and chilled water circuits. Figure 9 demonstrates the gradual decrease in the flow rate across the evaporator. In this case the classification function score is steadily increased until it becomes the highest value. Similarly, Figure 10 displays the classification scoring for a decrease in the flow of water across the condenser. In this case it is the score for the fault associated with water flow problems at the condenser that increases until it surpasses all the other ones.
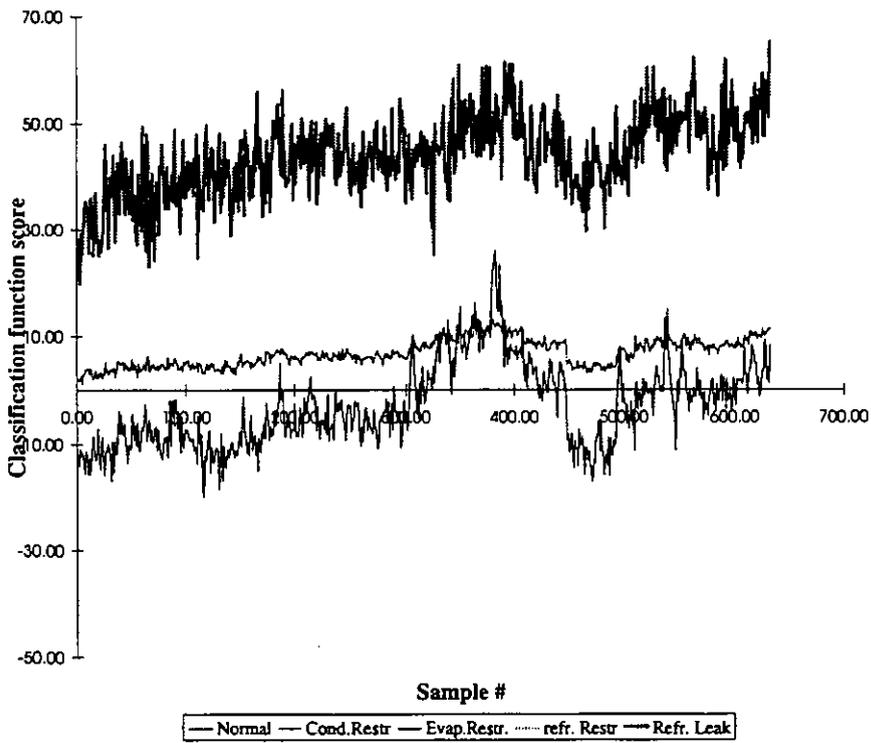
Figure 6a: Low Refrigerant Charge: Classification Scores with additional pressure measurement
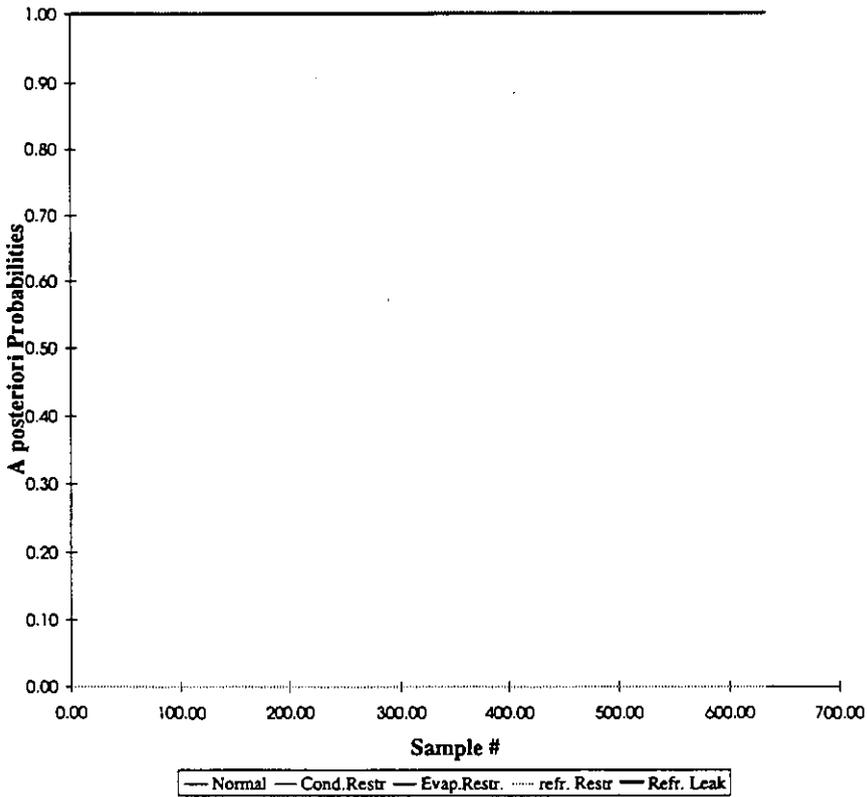
Figure 6a: Low Refrigerant Charge: A posteriori Probabilities with additional pressure measurement
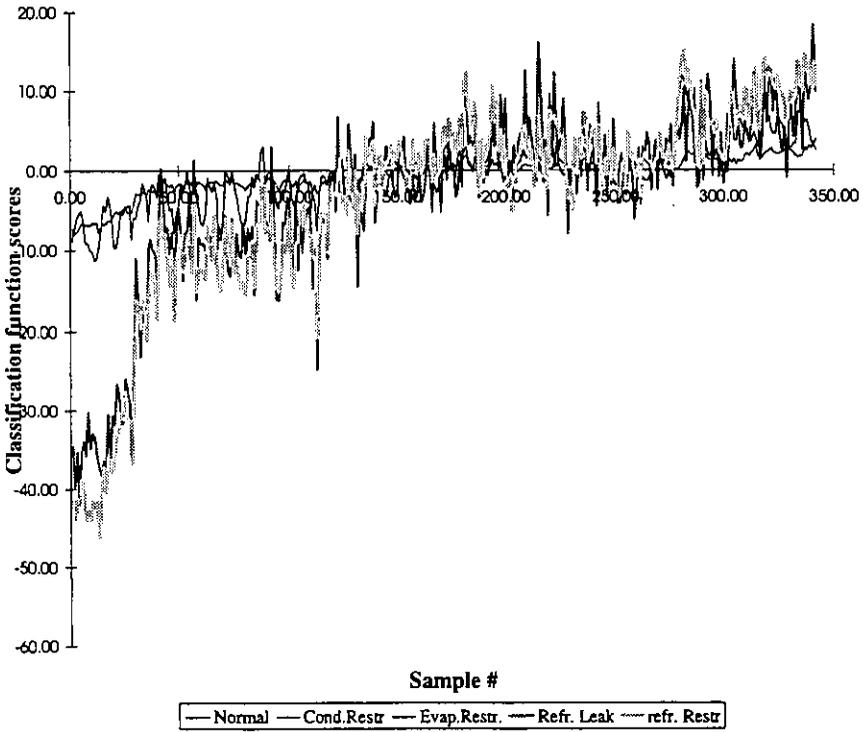
Figure 7a: Restriction in Refrigerant Line: Classification Function Scores
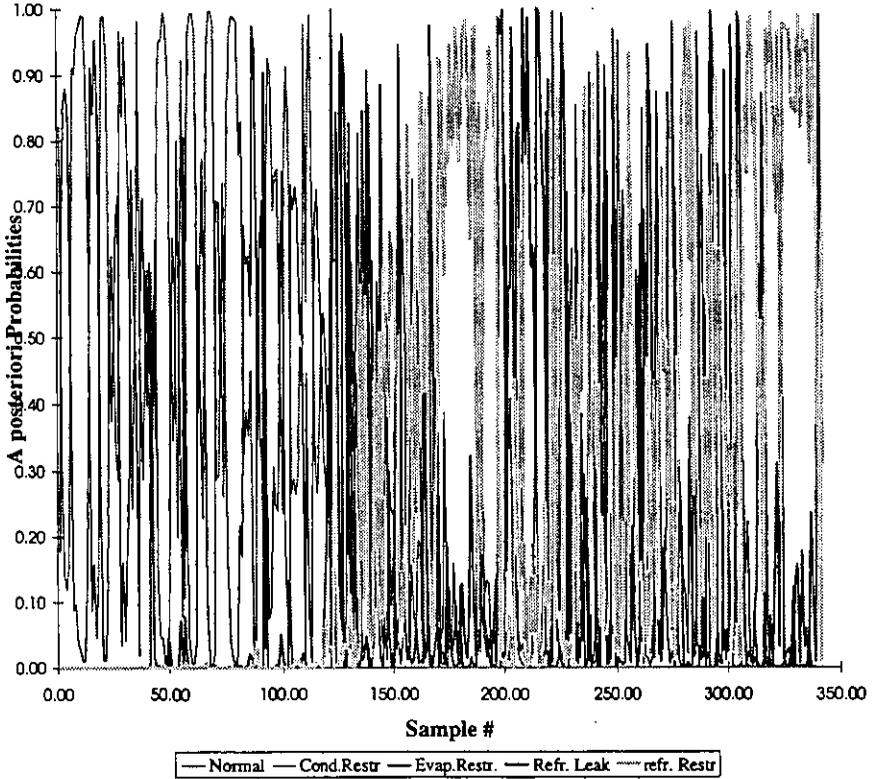
Figure 7b: Restriction in Refrigerant Line: A posteriori Probabilities

573
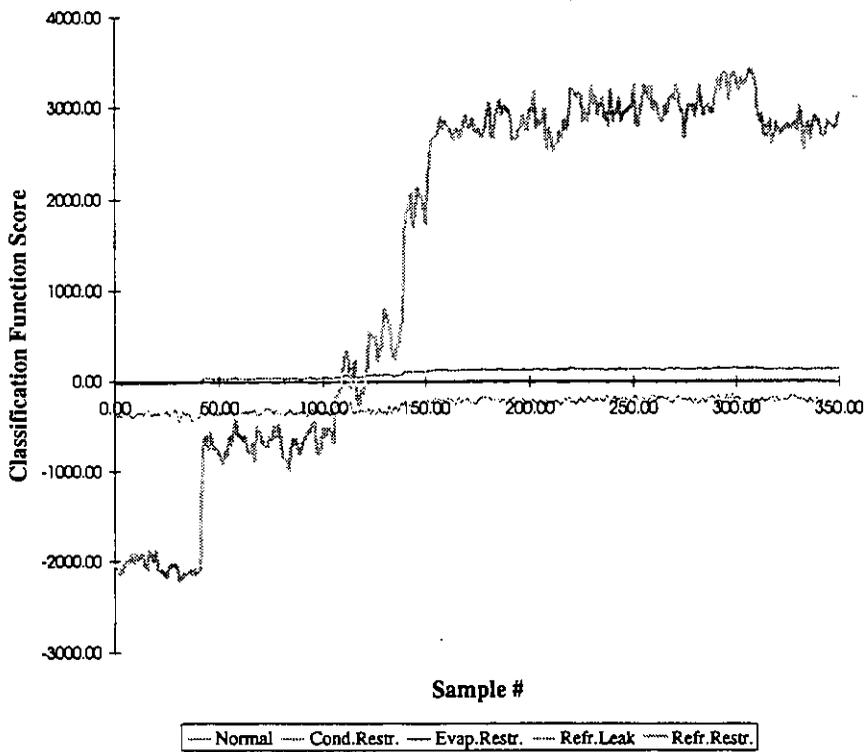
Figure 8a: Restriction in Refrigerant Line: Classification Function Scores with additional pressure measurement
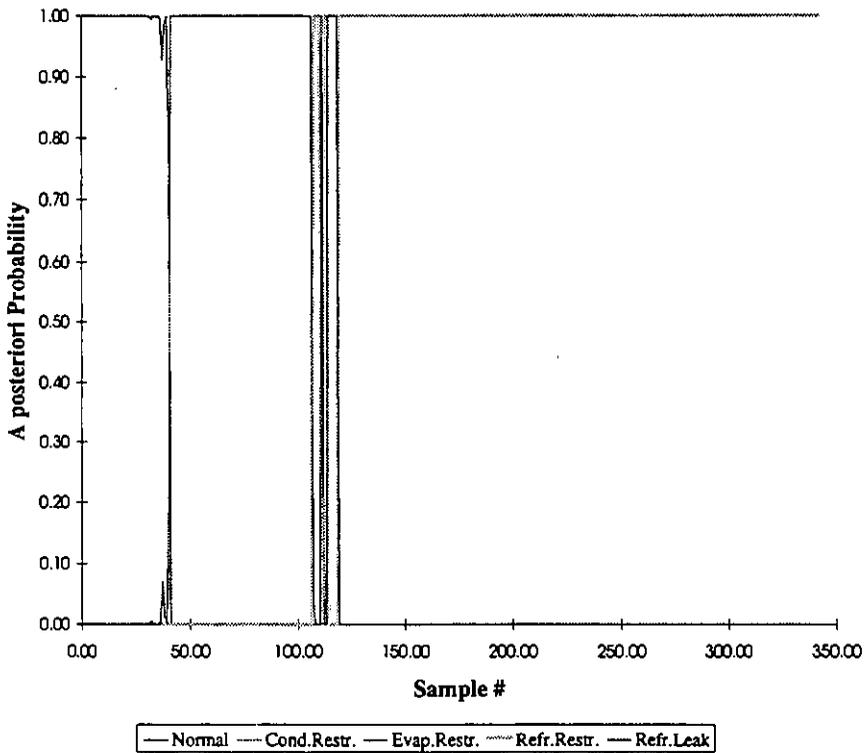
Figure 8b: Restriction in Refrigerant Line: A posteriori Probabilities with additional pressure measurement
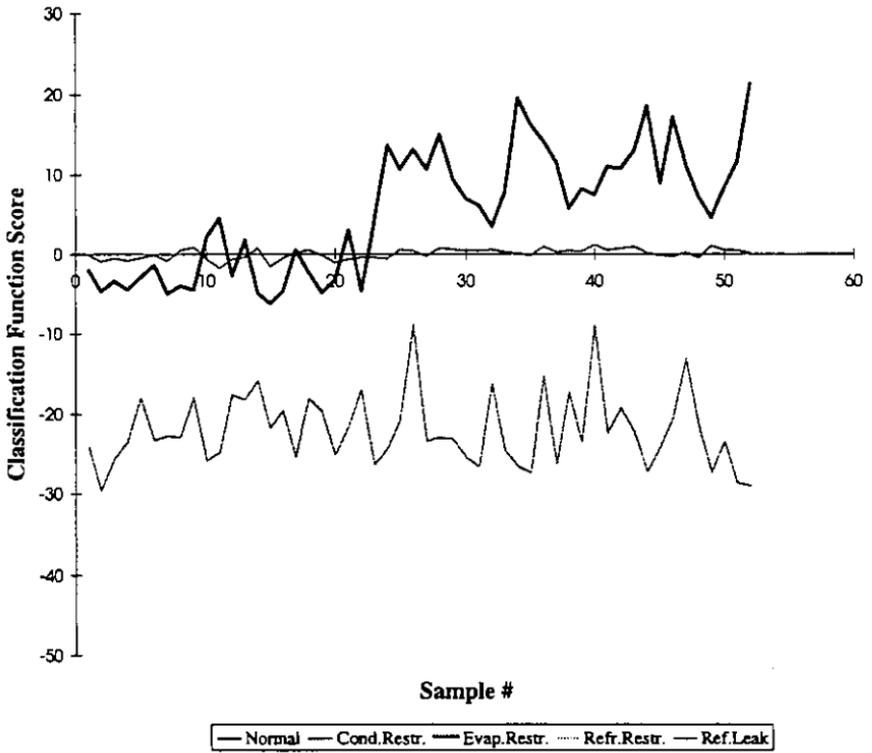
Figure 9: Restriction in the Chilled Water Circuit: Classification Scores
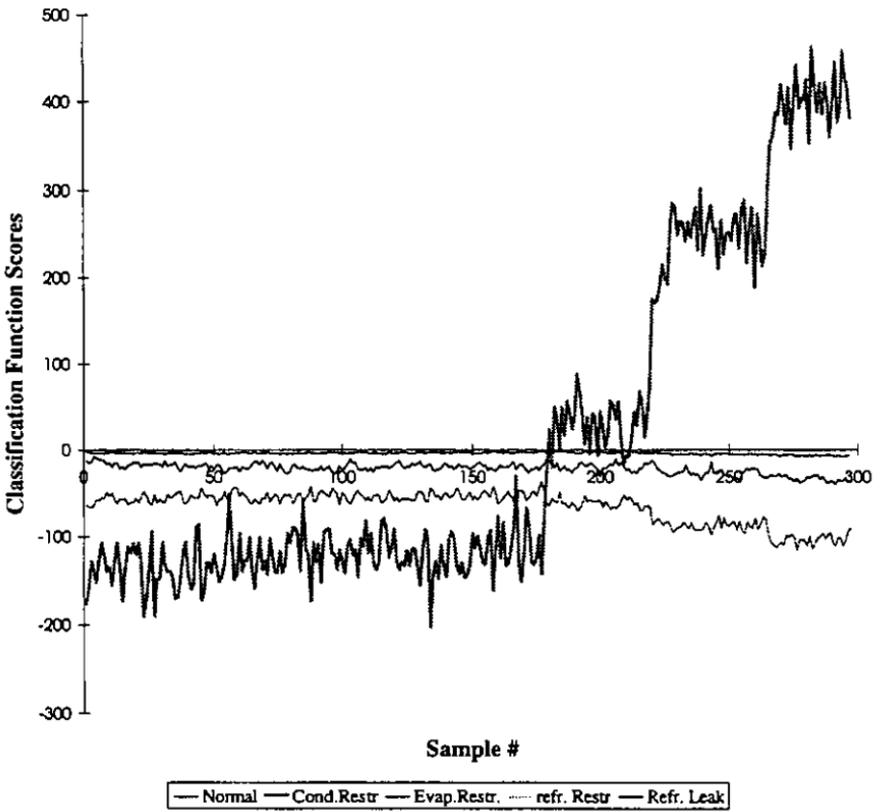
Figure 9: Restriction in the Chilled Water Circuit: Classification Scores

575

Conclusions

The SPRA described above provides an effective way of classifying patterns in multivariable, multi-class problems without having to explicitly use a rule-based system. This approach makes a number of assumptions including the normality of the distributions of the data. This assumption, as discussed above, is not strictly held. However, the experimental results show that its violation in this case does not affect the results significantly.

One assumption, however, that is made which cannot be verified and which could have a critical impact on the performance of the SPRA is the assumption regarding the *a priori* probabilities of each of the normal and four faulty operating conditions. These values are generally unknown and estimations are not based on the equipment in question, but rather on the experiences of the person or organization consulted.

For the experimental work discussed above, each of the five conditions was allocated equal *a priori* probability. This biases the results against the conditions that are known to generally occur more often, namely the normal operation and the refrigerant leak.

The impact of this bias is softened by the effect of the visualization of the results using the scores of the classification functions. These graphics offer an insight into the secondary patterns present in the samples which are not clearly evident in the case of the probability plots. Evidently the information content in the data for both the classification functions and the probability calculations is equivalent. However, in the case of large number of data points, it is helpful to have the advantage of effective visualization.

The above formulation of the chiller fault detection and diagnosis problem considers only five conditions: normal operation and four faults. The small number of faults detectable by the SPRA as described above, limits the applicability of the approach. However, it is possible to include a larger number of faults by identifying their specific patterns. In such a case if several faults exhibit similar patterns, additional measurements may be required to correctly diagnose them as was the case with the refrigerant leak and the restriction in the refrigerant line. Further work is required to justify the addition of more faults and to define their patterns.

The SPRA approach relies on the availability of training data for normal and faulty conditions. This requirement may pose difficulties if this is done while the unit is installed. However, it may be possible that patterns may be generated by the manufacturer and downloaded to the Building Energy Management System at the time of installation. The types of faults in such a case would be limited to faults that are intrinsic to the unit and which are not influenced by the surrounding plant. The four faults examined provide such a set, however, more testing is necessary to establish the applicability of the patterns to other, similar chillers.

References

Gimmelius, H.T., J. Klein Woud and G. Been 1995. On-line failure diagnosis for compression refrigeration plants. *Rev. Int. Froid* 18 (1): 31- 41.

Herzog, P., L. LaVine, 1992. Identification and quantification of the impact of improper operation of midsize Minnesota office buildings on energy use: A seven building case study. *Proc. ACEEE 1992 Summer Study on Energy Efficiency in Buildings*, pp.3.121-3.129

Huang,J., H. Akbari, L. Reiner and R. Ritschard 1991. *481 Prototypical commercial buildings for 20 urban market areas.* GRI Report GRI-90/0326.

Jenrich, R.I.(1977). "Stepwise Regression." In K. Enslein, A Ralston and H.S. Wilf (Eds.) *Statistical methods for digital computers.* New York: Wiley

Jenrich, R.I.(1977). "Stepwise Discriminant Analysis." In K. Enslein, A Ralston and H.S. Wilf (Eds.) *Statistical methods for digital computers.* New York: Wiley

Rossi, T., and J. Braun. 1995. "Thermodynamic impactfor detecting refrigerant leaks in vapour compression equipment" *1995 American Control Conference,* IEEE Control Systems Society, Seattle, Wash.

Stylianou M. and D. Nikanpour, 1996. "Performance Monitoring, Fault Detection and Diangosis of Reciprocating Chillers." *ASHRAE Transactions*, Vol. Part pp.

Stylianou M. and J. Scott, 1993. "Prioritization of Vapour Compression Equipment Faults" IEA-BCS Annex 25 Working Paper AN25/CA/140493.

Chiller Condition Monitoring Using Topological Case-Based Modeling

Hiroaki Tsutsui   **Kazuyuki Kamimura**

ABSTRACT

To increase energy efficiency and economy, commercial building projects now often utilize centralized, shared sources of heat such as district heating and cooling (DHC) systems. To maintain efficiency, precise monitoring and scheduling of maintenance for chillers and heat pumps is essential. Low-performance operation results in energy loss, while unnecessary maintenance is expensive and wasteful.

Plant supervisors are responsible for scheduling and supervising maintenance. Modeling systems that assist in analyzing system deterioration are of great benefit for these tasks.

Topological case-based modeling (TCBM) (Tsutsui et al. 1993; Tsutsui 1995) is an effective tool for chiller performance deterioration monitoring. This paper describes TCBM and its application to this task using recorded historical performance data.

INTRODUCTION

For large plants such as DHCs, which themselves consist of many subunits, it is difficult to precisely determine the system state under dynamically changing conditions. In isolated absorption chillers, for example, the relationship between steam output and cooling duty is linear, but when the chiller unit is embedded in a DHC system, its input/output relationships no longer follow the characteristics specified by the manufacturer, due to the dynamic nature of the system. Under these conditions, it is difficult to correctly determine the system state.

Usually, chiller performance is measured by a coefficient of performance (COP), which shows the ratio between input and output energy. However, as the COP is sensitive to changes in the external environment (such as the external temperature) and to various system states, the COP itself does not correctly reflect the performance of the chiller. Subsequent discussion will demonstrate that change of the COP ratio is, however, a useful metric for performance.

To overcome these problems, models are constructed that describe system changes through the relationships between a set of input variables and the output. For chiller performance monitoring, the output is the COP and the input variables are factors of the output.

When a model built for normal operation is used for performance monitoring, unit performance is characterized as the difference between the measured output value and the model output for the same inputs (Figure 1).

If the system includes third-party subunits or is already in use, it is difficult to generate ideal data for modeling, while it is often impossible to obtain complete and representative historical data from systems actually in use. When such incomplete historical data are converted to model parameters, it is difficult to determine whether observed differences between the measured and the model output for the same inputs are due to modeling errors or if they represent a real change in system behavior (Figure 2). As the measured output does not always reflect real system changes, it is difficult for plant supervisors to determine the nature of the real system changes and to choose appropriate maintenance periods (Tsutsui 1993).
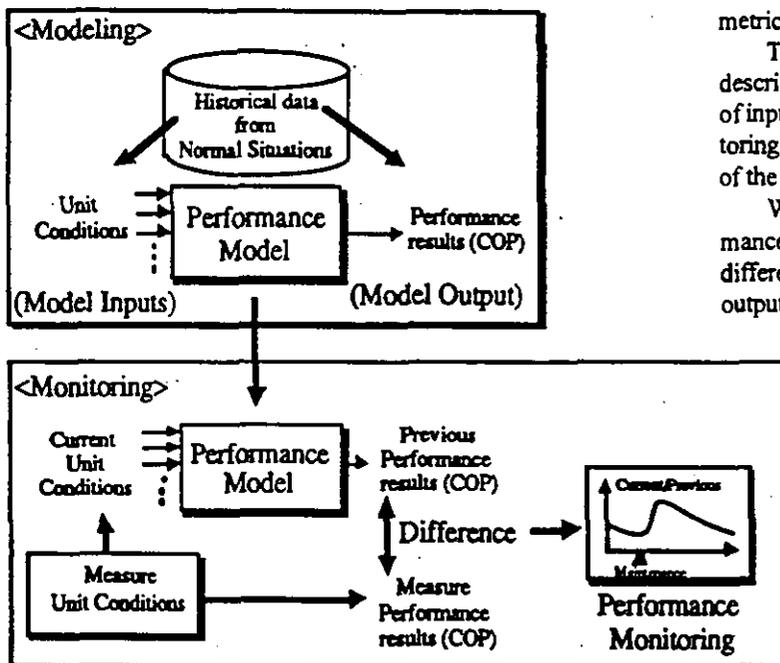
Figure 1   Model-based unit performance monitoring.

Hiroaki Tsutsui is with the Advanced Technology Center, Yamatake-Honeywell Co. Ltd., Yokohama, Japan. Kazuyuki Kamimura is with the Building Systems Division of Yamatake-Honeywell Co. Ltd., Tokyo, Japan.

- Modeling data (1 condition)
- — Actual Input/Output relationship
- •••• Input/Output Model
- ⊖ Reliable estimation by the model
- ⊘ Distrust estimation by the model

Input (factor1, factor2, ...)
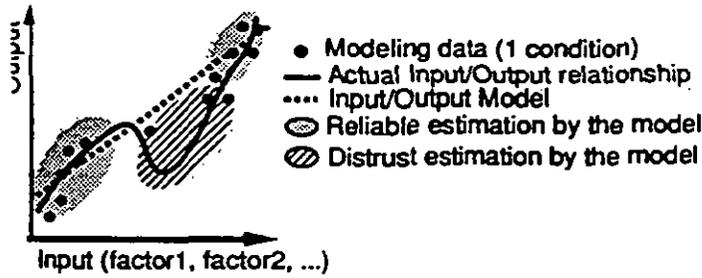
Figure 2 Input/Output relations modeled by historical data.

This paper compares the accuracy of TCBM with a conventional linear multiple regression model using the same set of input variables for an absorption chiller performance monitoring application, and demonstrates the advantages of TCBM over the standard model in an experimental situation.

TCBM PERFORMANCE MONITORING

TCBM Overview

TCBM Outline In TCBM, it is assumed that input/output data are extracted from a system involving a certain continuous functional relationship (of which TCBM does not require explicit description). The concept of a topological continuous mapping is used to define a system of neighborhoods in the input space that map neighborhoods in the output space with the desired precision.

For a given neighborhood in the output data, a locally continuous mapping is defined into a corresponding neighborhood in the input data. Using these neighborhoods, TCBM

defines an absolute size for the input space neighborhood that provides the required closeness of output. A measure of similarity between these input cases is defined, and a quantization-scaling method (the neighborhood system shown in Figure 3) is used to match new cases against historical data from the case base.

TCBM's localized formalism of cases makes the approach highly descriptive. Input variables are selected to guarantee local linearity of globally nonlinear behavior, and local formalism provides a resistance to local data fluctuations that is not present in modeling techniques based on global formalisms. Local calculations exploit a weighted integration using the outputs of similar input cases against new input cases (Tsutsui et al. 1993; Tsutsui 1995) (Figure 3).

TCBM treats historical data as a compressed case base according to the required closeness of the output, and estimates the output associated with new input cases using the following case-based reasoning (CBR) procedures:

- Record the past experiments (problems/answers) in the case base.

- Search in the case base for problems similar to the new problem.

- Modify the answer of similar problems to obtain an answer for the new problem.

- After returning the answer to the new problem, add the new case to the case base.

Conventional CBR (DARPA 1989) techniques define the neighborhood that provides the needed measure of similarity between cases using time-consuming ad hoc methods, with a difficulty of definition dependent on the application domain. In

Figure 3 TCMB overview.

AT-96-6-4

580
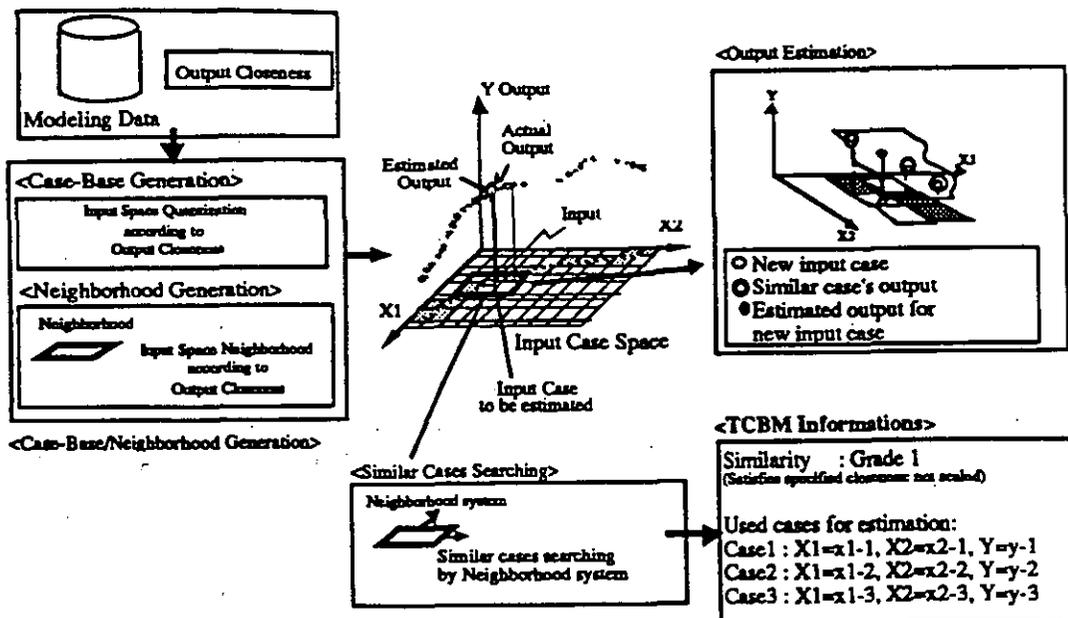
contrast, TCBM defines the neighborhood theoretically, based on the assumption that the input/output relationship is locally continuous.

Advantages of TCBM TCBM differs in several characteristics from general black-box modeling techniques. These differences are detailed here.

1. TCBM defines the input space neighborhood according to the specified output accuracy by topological continuous mapping. The neighborhood system, which is generated by rescaling ("zooming") the defined neighborhoods, can show the local relative relationship (similarity) between a new input case and recorded cases. The more the neighborhood has been scaled up, the more the inferred output becomes uncertain. Therefore, the zooming degree is a novel assurance index for inferred output according to the input situation.

2. TCBM regulates only input/output variables, without formalizing global input/output functional relationships.

3. TCBM stores data as cases—historical data are not converted to model parameters.

4. TCBM can identify the cases that are used for inferencing the new situation's output according to each new situation's input. Cases similar to the new input situation may be extracted from the case base using the neighborhood system.

5. Output estimation is performed locally by extracted cases using the neighborhood system.

These characteristics show advantages over general black-box modeling:

- General black-box models ensure output certainty by a single statistical value, such as mean square error, which does not consider the input situation. If the current input situation data are not used for modeling, the inferred output is uncertain. On the other hand, TCBM can evaluate output certainty according to the input situation (by no. 1 above) (Figure 2).

- TCBM can check the basis for inference, showing the cases that are used for inferencing and their closeness by zooming degree (by nos. 1,3, and 4).

- TCBM can construct complex nonlinear models without investigating input/output transfer function parameters or networks (by nos. 2 and 3).

- TCBM easily incorporates new cases or model improvements, facilitating simple online modeling (by no. 2).

- TCBM is suitable for modeling nonlinear systems (by no. 5).

Mathematical Basis of TCBM

To construct the case base, the input space is first partitioned into finite sets of input cases and quantized. The number of partitions is selected to give an output distribution for the same input case with the specified closeness. A single case is generated by converting the input historical data to a single integer input case. Here, for "n" inputs with partitioning number "m," a single input case is symbolized in an m-adic manner.

Case-Base Structure (Input variables are X1, . . ., Xn; the output variable is Y, the number of recorded data that belong to the same input case is k).

Input case $\{Xi\}(i = 1, . . ., n)$

Same input case number k

Average value of output for same input case: $Y = \sum_{j=1}^{k} Yj/k$

Average value of output differentiation for same input case:

$$\{\partial Y/\partial Xi = (\sum_{j=1}^{k} \partial Yj/\partial Xij)/k\} \ (i = 1, ..., n)$$

$$(\partial Yj/\partial Xij = \{1/ |xij(t+1)-xij(t-1)| /$$

$$\sum_{j=1}^{n} 1/ |xij(t+1)-xij(t-1)| \} \bullet$$

$$[yj(t+1) - yj(t-1)] / [xij(t+1)-xij(t-1)] \}$$

This equation is an approximation to simplify on-line calculation. The "t" in the equation is a time stamp.

Topology, Closeness and Neighborhoods

Topology In this paper topology is a concept that ensures a continuous relationship between input and output spaces and facilitates the definition of a concept of closeness for both input and output spaces. The set of values within a given closeness of a point constitutes the neighborhood of that point with the given closeness.

Continuous Mappings A mapping $f: X \geq Y$ is continuous if and only if $f^{-1}(z)$ is in a neighborhood of X for all z in the neighborhood of Y.

From this definition, given a neighborhood of a given closeness in Y and the existence of such a continuous mapping $f: X \geq Y$, it is possible to obtain the corresponding neighborhood in X.

In terms of TCBM, X corresponds to the set of input case data and Y to the set of output data.

The absolute size of a single neighborhood in X (the input space), which is a statistical value according to neighborhoods of the specified closeness in Y (interpreted in TCBM as granularity in Y), is a scalar quantity used as a measure of the local correlation between the particular variables in question. A smaller value of this measure corresponds to a higher correlation between X and Y.

The neighborhood system is generated by increasing or decreasing the quantization number of each variable of the neighborhood and is used to choose the cases that are used for inferencing, as described in the following section.

Inferencing Procedures TCBM's inferencing consists of three phases.

Phase 1: Searching—New input cases are compared with those stored in the case base to determine similarities, using the neighborhood system. If the system has n-inputs and one output, n+1 cases are selected. These n+1 cases can be represented by a local space of n+1 dimension.

Phase 2: Significant Weight Calculation—The weight of the selected cases for the new situation is calculated as follows. (There are two equations to calculate the weight; however, inferencing performance is almost same.)

1. Using Topological Distance:

Topological distance, $L = \sum_{i=1}^{n} \phi i \cdot |Xi^* - Xi|$, Xi^* is a new situation of Xi

($\phi i = Ri / \sum_{i=1}^{n} Ri$, where Ri is a correlation coefficient between Xi and Y).

Significant weight: $w = \exp(-L)$.

2. Using neighborhoods:

$$w = w(x1, ..., xn)$$

$$= 1/(\sqrt{2\pi})^n \cdot 1/(\sigma x1, ..., \sigma xn) \cdot e^{-1/2(x1^2/\sigma x1^2 + ... + xn^2/\sigma xn^2)}$$

σxi: the statistical distribution, xi, is $q - xi/2$ to the specified closeness in Y.

$q - xi$: the neighborhood of xi.

Phase 3: Inferencing—The selected $n + 1$ cases are used for inferencing the new situation Y^* from the new input set $X_1^*, ..., Xn^*$.

$$Y^* = \sum_{j=1}^{n+1} \left[wj \cdot \left\{ Yj + \sum_{i=1}^{n} (\partial Yj/\partial Xij) \cdot (Xi^* - Xij) \right\} \right] / \sum_{j=1}^{n+1} wj$$

where

Xij = value of ith input variable for jth selected case,

Yj = recorded value of output for jth selected case,

$\partial Yj/\partial Xij$= recorded value of output differentiation for jth selected case, and

wj = significant weight of jth selected case for new situation.

Updating the Case Base The update procedure is adaptive and utilizes an oblivion parameter as described below (updating procedure of $\partial Y/\partial Xi$ is the same as this procedure):

$$Y(new) = [kaY(old) + y]/[ka + 1]$$

where

$Y(new)$ = revised output,

$Y(old)$ = recorded value in the case base,

y = new output,

k = number of the same input case, and

a = oblivion parameter (determined from the system change rate; default value is 1).

Note: For the performance monitoring and fault detection problems, the updating function is not used.

Characteristics of TCBM for Performance Monitoring

Performance monitoring problems offer the following challenges: (1) Many variables must be considered to fully grasp the unit's characteristics. (2) Nonlinear modeling techniques are needed, with careful determination of the nonlinear aspects of the system. (3) It is necessary to determine whether differences between the measured output value and the model output for the

same inputs are due to modeling error or to real system changes. It is important to understand how much of the system's actual state-space has been recorded and can be used for modeling.

TCBM offers the following solutions to these problems.

- TCBM uses many variables simultaneously.

- TCBM is applicable for modeling nonlinear systems because output estimation is performed locally.

- A TCBM neighborhood can show the local relationship (similarity) between new and recorded cases. TCBM can eliminate estimation error by the use of a real input case that is identical in some sense (see "Estimations Only Using Quantized Cases" subsection) to the new input because in this situation stored historical data are used to generate the output, rather than the input from a global approximation such as would result from the use of, for example, a conventional linear regression model.

CHILLER PERFORMANCE MONITORING MODEL

Model Variables

Model variables were selected through statistical analysis and through knowledge of the mechanical principles involved in the system. Absorption chillers work as shown in Figure 4.

- The water coming from users flows inside a thermally conductive pipe. The cooling water removes heat from this pipe and returns chilled water to the users.

- The vaporized water is absorbed by the sorbent (LiBr).

- The heat of the vaporized water is discharged by cooling water, which is also affected by external temperatures.

- When the sorbent becomes saturated, it is heated to extract the water as steam.

Therefore, the higher the concentration of the sorbent (LiBr), the better the absorption performance. However, absorption chillers are affected by external temperatures because the cooling water is in contact with the surrounding environment. Considering these interactions, the model variables were selected as follows.

Model Output

The COP (see Introduction) was selected as the model output. The COP is calculated from historical data using the following equation:

$$COP = Q_e/Q_g$$
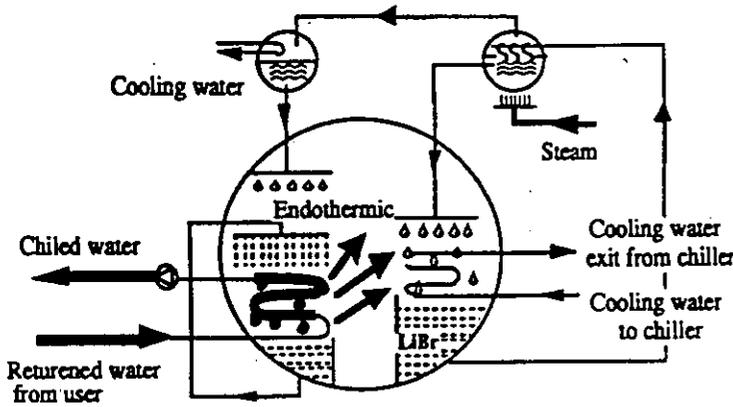
Figure 4 *Absorption Chiller.*

where

Q_e = (cooling quantity) = flow rate × (temperature of water returned from user - temperature of chilled water) / 1000 [Gcal/h {4.2 GJ/h}]

Q_g = (steam quantity) = steam flow rate × 489.46 /1000 [Gcal/h {4.2 GJ/h}].

Model Inputs

Five input variables were selected considering the chiller mechanism and the need to cover a wide range of chiller performance. The five selected input variables were temperature of cooling water input to chiller, temperature of cooling water output from chiller, temperature of water returned from user, temperature of chilled water, and steam flow rate.

Inside the chiller, the steam discharges its heat and becomes a mixture of liquid water and vapor. In this condition, the pressure should be considered as the pressure of saturation P. We assume the resistance between the chiller and the steam is R. The following equation is then satisfied:

$$P = P_0 + R \cdot Lg$$

(Lg is steam flow rate, P_0 is a constant).

The relationship between the temperature of the saturated steam, T, and its pressure is approximately linear under usual conditions. Thus,

$$T = T_0 + k^* \cdot P$$

(k is a coefficient, T_0 is a constant).

Combining the above two equations gives the following result:

$$T = a + b \cdot Lg$$

where a and b are constants. Therefore, Lg can be used in place of T.

External temperature was not selected as a model input because it is correlated with the five input variables. The coefficients of correlation (COR), using standard regression analysis,

between the external temperature and each of the five selected input variables were:

COR (external temperature, temperature of cooling water input to chiller) = 0.76

COR (external temperature, temperature of cooling water output from chiller) = 0.83

COR (external temperature, temperature of water returned from user) = 0.87

COR (external temperature, temperature of chilled water) = 0.27

COR (external temperature, steam flow rate) = 0.82.

Data Processing

Historical data recorded every 5 minutes were averaged every 30 minutes. As the historical data were recorded from a chiller actually in operation, they contained various kinds of noise, including local noise caused by the measurement equipment itself and global noise caused by the poor condition of the equipment being monitored.

For global noise elimination, all data outside of three standard deviations from the average were discarded on the assumption that the noise distribution followed a normal distribution. This eliminated transient erroneous data caused by temporary sensor failures or system failures.

For local noise elimination, all spike noise, regardless of noise frequency, was eliminated by fuzzy data smoothing techniques, which facilitate data smoothing similar to human intuitive judgment (Tanaka 1992). This preprocessing provided a more accurate data set for subsequent analysis.

Finally, as deviations due to real changes in system performance remain, data were averaged over 24-hour periods to provide a continuous set of data and reduce the effect of unknown (not measured) disturbances.

The fuzzy data smoothing and averaging processes serve to filter different noise frequency spectra, and together provide a data set representative of process performance to the analysis procedures that follow.

These procedures are necessary to make a precise model to detect the variances caused by subtle system changes such as deterioration. Note: If TCBM were to be applied to the problem of fault detection rather than performance monitoring, then this 24-hour averaging would not be applied. The unknown factor should be investigated and should be considered as the input of the model.

CHILLER PERFORMANCE MONITORING

Data for Evaluation

Historical data were recorded from July 1992 to May 1994. Data recorded from July 1992 to April 1993 were used for modeling, while data recorded from May 1993 to May 1994 were used for performance monitoring.

The TCBM system was trained once using the first data set. The modeling data contained a whole year's environmental changes and chiller operation patterns, and are thus taken to cover

the entire normal operating range of the chiller. The monitoring results show chiller performance changes from May 1993 to May 1994 based on the condition of the chiller from July 1992 to April 1993.

Model Precision

The data recorded from July 1992 to April 1993 were divided every two and one-half hours to generate two data sets, which were then used to confirm model precision and justify variable selection.

Figure 5 shows the results of linear modeling using a textbook standard linear regression model using the same set of input variables as the TCBM model. Figure 6 shows the results of TCBM. In this case, TCBM searches for only one similar case within the case base and indicates the COP of the similar input case. These results demonstrate that, in this case, TCBM was twice as precise as a standard linear multiple regression model on the basis of comparison of error averages over the entire data set.

Note: The sample numbers in Figures 5 through 8 represent the data sets after off-line and sensor failure intervals were removed. Thus there is not necessarily a linear relationship between sample number and calendar time.

Model Estimation and Monitoring

Using these models, the COPs from May 1993 to May 1994 were estimated. The estimations from each model are shown in Figures 7 and 8. For TCBM, the estimation process is the same as that described in the "Model Precision" subsection. All data from July 1992 to April 1993 were used for modeling.

Comparing these estimations with the model precisions in the "Model Precision" subsection, the difference between the estimated results and the model precision shows the actual changes in chiller condition. With the linear model, the mean difference between estimated results and the measured COP (~2.24%) is 2.5 times the precision of the model (~0.841%). In TCBM, the mean difference between the estimated results and the measured COP (~7.38%) is about 20 times the model precision (~0.352%).

There is a large difference between the results of the linear model (2.5 times) and TCBM (20 times). This difference demonstrates the differences between the modeling techniques—the linear model decides the model parameters of the hyperplane to fit the whole year's data, and the estimated value must lie on this hyperplane. On the other hand, the value estimated by TCBM is that of the most similar input case that shows the fact in the past.

Linear modeling is a statistical method and its hyperplane fits to the area of most frequent occurrence to minimize modeling errors. However, if the area of most frequent occurrence has changed due to system changes, the linear model may become very imprecise (Figure 9). When these cases are used for estimation, the reported difference
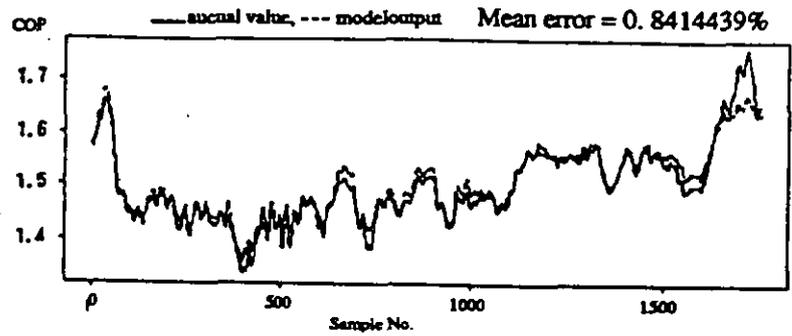
Figure 5 Linear model results.

between the linear model's output and the actual COP becomes smaller than the actual difference, although the chiller performance has changed.

The results given in this section do not always show the difference (for the same input) between the estimated and real values. Therefore, it is difficult to evaluate the output difference. TCBM can show the difference between the estimated and actual values only when the same input situation has occurred in the case base that shows the fact in the past, because TCBM can judge whether or not the same input situation that should be estimated is used for modeling. Figure 12 shows the result using this procedure. In this figure, the output differences have been treated as the ratio of the actual value to the value estimated by TCBM.

Explanation of Modeling and Monitoring

Modeling Figures 5 through 8 show that the range of COPs over the whole year was 1.3 to 1.8. An output accuracy of 1% was chosen to make a precise model. From this, the range of each variable in the input space was divided into 10 regions and quantized, as required by TCBM (see the "Mathematical Basis of TCBM" section). In this example, 10 regions are the minimum satisfying the required accuracy. These neighborhoods in Figure 10 were used for the testing represented in Figure 6.
Performance Monitoring

Estimation Using Neighborhoods The defined neighborhoods and quantization of input space in Figure 10 are used to estimate the current situation based on the previous year's conditions. These results are shown Figure 8.
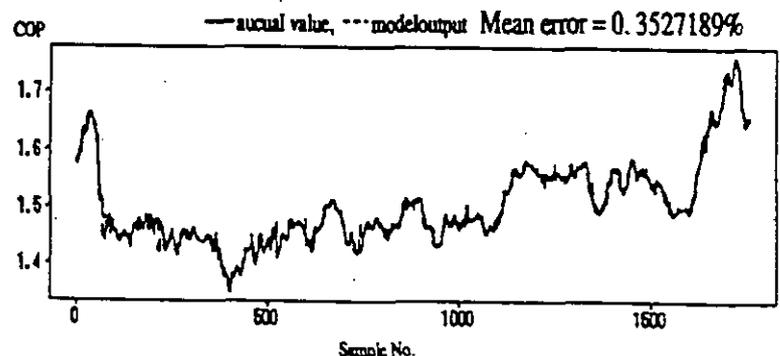
Figure 6 TCMB results.

Estimations Only Using Quantized Cases

The results given in the "Model Estimation and Monitoring" section do not always show the differences for the same input, so comparison with quantized cases is used, as shown in Figure 12. Input variable ranges are quantized in 10% intervals. Values within each of these regions are considered as the same case.

Changes in COP Ratio

In Figure 12, COP ratios of unity imply that the chiller acts at the previous year's performance level. Ratios greater than unity represent performance improved from the past year. Conversely, ratios of less than unity represent diminished performance. In the figure, the modeling data are derived from a prior period, and monitoring data show changes of COP for the same inputs (as described above).

Figure 12 shows that the COP ratio improves 10% after maintenance. The positive effects of maintenance are revealed when the chiller is used continually. After frequent use, chiller performance deteriorates. These results demonstrate the characteristic performance changes of this absorption chiller.

More precisely, the COP ratios fluctuate locally. These fluctuations seem to be due to the intermittent use of the chiller. In fact, performance also decreases after long periods of nonuse. The reasons for these changes are under investigation, but there seem to be two possibilities: that long-term nonuse actually does cause a performance drop, or the chiller is stopped for setpoint adjustment.

DISCUSSION

TCBM provides a practical method for monitoring absorption chillers, with an accuracy much higher than conventional modeling techniques. For practical application to performance monitoring, a threshold could be set, beyond which deterioration was considered to have occurred. Such a threshold should be set considering the cost benefits under each system's situation.

TCBM is a fast, efficient technique; its database of cases is compact enough for real-world applications (the application presented here required a case base of about 30 kbytes to store one year's data), and the methodology is robust enough not to require continual attention from operators.

CONCLUSIONS

Compared with a simple reference table of input and output values, TCBM provides the concepts of neighborhoods and closeness introduced in the "TCBM Performance Monitoring" section. Together, these concepts provide powerful determination of modeling accuracy and valuable information regarding the reasons for the model's behavior.

This paper has demonstrated that for experimental application to the problem of absorption chiller performance monitoring, TCBM clearly outperformed standard linear regression modeling techniques.
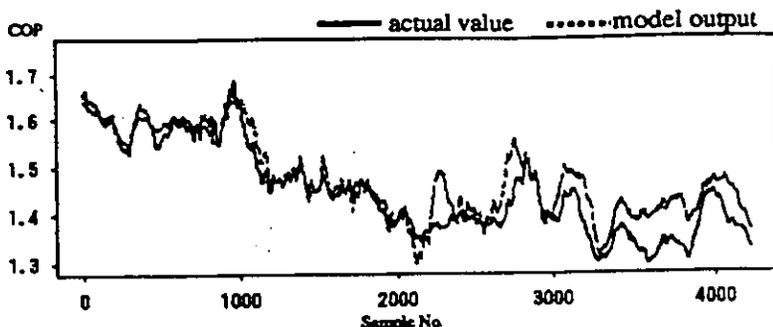
Figure 7 Linear model results [Difference between model and actual: 2.241458%].

The authors are now investigating the application of this technique to the analysis of other heat sources and air-conditioning units that are also difficult to model, with the goal of similarly observing the effects of control and operational changes on system performance.

REFERENCES

DARPA. 1989. Case-based reasoning. *Proceedings of a Workshop on Case-Based Reasoning*. Florida meeting. London: Defence Advanced Research Projects Agency Information Science and Technology Office.

Kyu, M. 1994. Research on the efficiency of DHC equipment: Part 1—Analysis of absorption chiller efficiency degeneration. Osaka: The Society of Heating, Air-Conditioning and Sanitary Engineers of Japan Conference (Japanese).

Tanaka, M. 1992. Fuzzy data smoothing. *Proceedings of ISA '92 Conference*, pp.1173-1184.

Tsutsui, H. 1993. Fault detection using modeling techniques by sensing data. *Proceedings (AN25/JP/270993/1.0) of IEA ANNEX25 Zurich Meeting*.

Tsutsui, H. 1995. Modeling technique using historical data "TCBM." *Instrumentation and Control Engineering* 38(5) (Japanese).

Tsutsui, H., K. Kamimura, and T. Matsuba. 1993. Chiller performance deterioration fault detection using topological case-based modeling. *Proceedings (AN25/JP/150493/2.0) of IEA ANNEX25 Tokyo Meeting*.
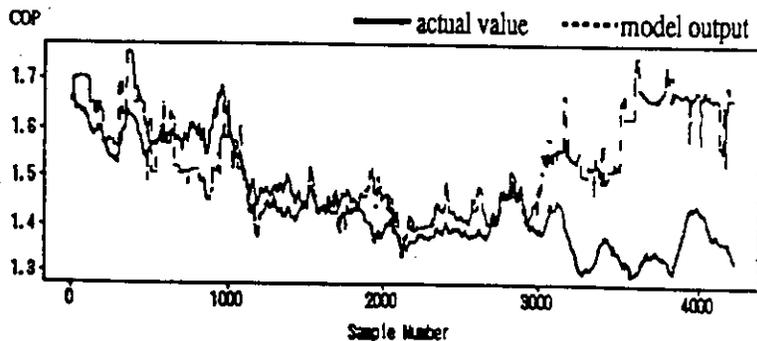
Figure 8 TCBM results [Difference between model and actual: 7.375945%].
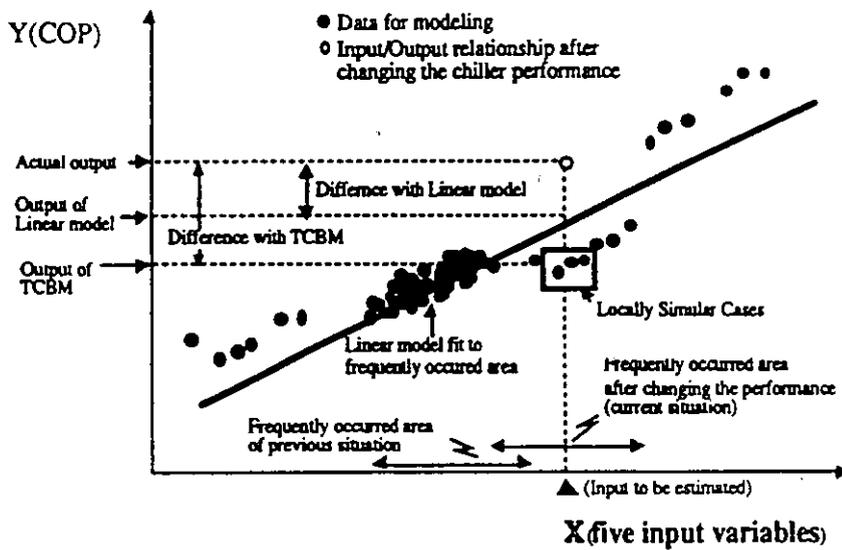
Figure 9   Model-based error estimate and system changes.
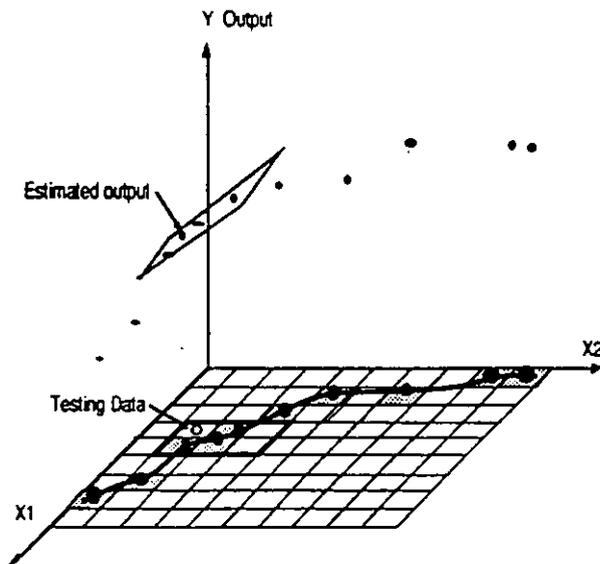
Figure 10 Previous situation with modeling.
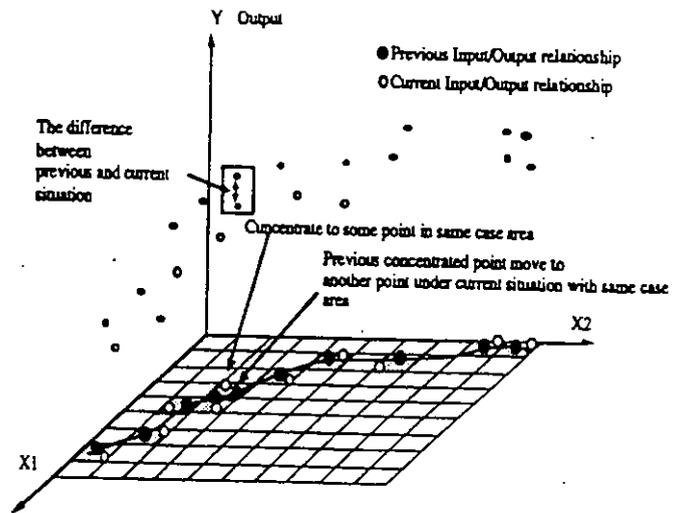
Figure 11 Current situation with estimation.
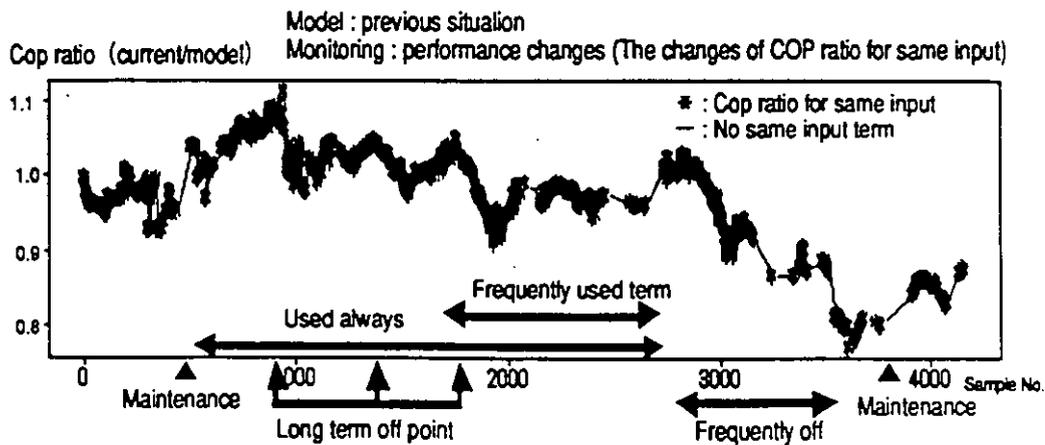
Figure 12 TCBM performance monitoring by the change iof COP ratio.

FAULT DIAGNOSIS OF AN AIR HANDLING UNIT

USING ARTIFICIAL NEURAL NETWORKS

Won Yong Lee is a Senior Researcher at the Korea Institute of Energy
Research in Taejon, Korea. John M. House and Cheol Park are
Mechanical Engineers and George E. Kelly is a Group Leader in the
Mechanical Systems and Controls Group, Building Environment Division,
Building and Fire Research Laboratory, National Institute of Standards and
Technology, Gaithersburg, Maryland.

Won Yong Lee (Ph.D.)

John M. House (Ph.D., Associate Member ASHRAE)

Cheol Park (Ph.D.)

George E. Kelly (Ph.D., Member ASHRAE)

ABSTRACT

The objective of this study is to describe the application of artificial neural networks to the

problem of fault diagnosis in an air handling unit. Initially, residuals of system variables that can

be used to quantify the dominant symptoms of fault modes of operation are selected. Idealized

steady-state patterns of the residuals are then defined for each fault mode of operation. The

steady-state relationship between the dominant symptoms and the faults is learned by an artificial

neural network using the backpropagation algorithm. The trained neural network is applied to

experimental data for various faults and successfully identifies each fault.

KEYWORDS

air handling, classification, maintenance, modeling, monitoring

FAULT DIAGNOSIS OF AN AIR HANDLING UNIT
USING ARTIFICIAL NEURAL NETWORKS

ABSTRACT

The objective of this study is to describe the application of artificial neural networks to the problem of fault diagnosis in an air handling unit. Initially, residuals of system variables that can be used to quantify the dominant symptoms of fault modes of operation are selected. Idealized steady-state patterns of the residuals are then defined for each fault mode of operation. The steady-state relationship between the dominant symptoms and the faults is learned by an artificial neural network using the backpropagation algorithm. The trained neural network is applied to experimental data for various faults and successfully identifies each fault.

INTRODUCTION

Modern buildings are being designed with increasingly complex operating systems that have seemingly limitless capabilities for monitoring and controlling the conditions in the building. Unfortunately, building operators are not always able to monitor and process the enormous amounts of data that are generated. Hence, there is a need for robust fault detection and diagnostic tools that can be used to assist the building operator and ensure that the system is operating in the manner in which it was designed. The benefits of a properly operating building system are numerous, including improved energy efficiency, improved occupant comfort and health, and longer equipment life.

In a companion paper Lee et al. (1995) describe methods for fault detection in an air handling unit (AHU). One approach used in that study is to define residuals which represent the difference between the existing state of the system and the normal state. Residuals that are significantly different from zero represent the occurrence of a fault. If the system that is being monitored is not too complex, the building operator should have little trouble isolating the source of the fault after the initial fault detection. However, for complex systems, isolating the fault can be very challenging and diagnostic tools are needed. This paper describes the use of artificial

neural networks (ANNs) for this purpose.

Several studies that examine the use of ANNs for fault diagnosis appear in the literature. Watanabe et al. (1994) and Fan et al. (1993) used ANNs for fault diagnosis of chemical processes. Watanabe et al. (1994) proposed a two-stage multilayer ANN. In the first stage the faults were diagnosed and in the second stage the degree of the fault was estimated. The study of Fan et al. (1993) was based on steady-state operating conditions. Koive (1994) reviewed studies that utilized ANNs for fault diagnosis and control and summarized the architectures most widely used in practice. The paper also summarized steady-state and dynamic fault diagnosis and control for a paper making machine.

Fault diagnosis can be thought of as pattern recognition and ANNs are well suited to this task. For example, ANNs using the backpropagation algorithm can be used for character recognition (Ananthraman 1995; Demuth and Beale 1992). The input patterns are matrix representations of the dark (1's) and light (0's) pixels of the 26 characters of the alphabet, and the output patterns are 26 bit strings of 1's and 0's that represent the various characters. A similar approach can be applied for fault diagnosis. Normal and fault modes of operation typically have operational signatures or distinguishing patterns for each mode of operation. ANNs can learn and exploit these patterns to diagnose the current operational mode of a system. The objective of this study is to describe the application of an ANN to the problem of fault diagnosis in an AHU. As an intermediate step, a second ANN is used as a process model for a cooling coil valve subsystem.

The first two sections of the paper provide a brief description of the AHU and the residuals used in the fault diagnosis. The eight faults and their corresponding symptoms and dominant residuals are then described. Next, a brief description of ANNs and the backpropagation algorithm is provided. Applications of ANNs to the development of a model of the cooling coil valve subsystem and fault diagnosis are then discussed. Finally, results of the fault diagnosis are presented and conclusions and recommendations for future work are discussed.

AIR HANDLING UNIT

A schematic diagram of the variable-air-volume (VAV) AHU utilized for this study is shown in Figure 1. The same system was used for a companion paper on fault detection (Lee et al. 1995). The AHU consists of fans, dampers, a cooling coil, sensors, and controllers. The static pressure in the main supply duct is maintained at a constant setpoint value of 249 Pa (1.0 in. of water) by sensing the static pressure and controlling the rotational speed of the supply fan. The supply air temperature is controlled by modulating the cooling water control valve to maintain a constant setpoint value of 14.5°C (58.0°F). The air flow rate difference between the supply and return air streams is controlled by the variable speed return fan to maintain a constant setpoint value of 0.472 m³/s (1000.0 ft³/min). A PID algorithm is used to control the cooling water valve and PI algorithms are used to control the supply and return fan speeds. Although not shown in Figure 1, a personal computer (PC) and data acquisition system (DAS) are used for purposes of computing control signals and logging data. The sampling period for control and data collection is ten seconds.

RESIDUAL DEFINITION

The approach used in this paper relies on the ability to identify patterns of residuals that can be used as signatures for various faults. Through laboratory testing it was determined that seven residuals are needed to identify the eight faults considered here (described in the next section). The first three residuals represent the difference between actual and setpoint values of the supply air temperature, supply air pressure, and the airflow rate difference between the supply and return ducts. The residuals are given by

$$R_T = T_S - T_{S,SP} \tag{1}$$
$$R_P = P_S - P_{S,SP} \tag{2}$$
$$R_Q = Q_D - Q_{D,SP} \tag{3}$$

where

R = residual value

T_S = supply air temperature

P_S = supply air pressure

Q_D = airflow rate difference in the supply and return ducts

$T_{S,SP}$ = setpoint value of T_S

$P_{S,SP}$ = setpoint value of P_S

$Q_{D,SP}$ = setpoint value of Q_D.

The cooling coil valve control signal can provide valuable insight into the operating status of the AHU. A residual is defined for the operation of the cooling coil valve and is given by

$$R_U = U_{CC} - U_{CC,EV} \tag{4}$$

where

U_{CC} = actual control signal to the cooling coil valve

$U_{CC,EV}$ = expected value of U_{CC}.

U_{CC} is determined by the PID controller for the supply air temperature, however, there is no obvious way to specify $U_{CC,EV}$. In this study $U_{CC,EV}$ is determined using an ANN model of the cooling coil valve subsystem. The model for $U_{CC,EV}$ is described in a later section.

Residuals for the operation of the supply and return fans are given by

$$R_{NS} = N_S - U_S \tag{5}$$
$$R_{NR} = N_R - U_R \tag{6}$$

where

N_S = measured value of the supply fan speed

N_R = measured value of the return fan speed

U_S = control signal for supply fan

U_R = control signal for return fan.

The final residual is based on a comparison of the actual cooling coil valve position and the expected value based on the actual cooling coil control valve signal. The residual is given by

$$R_V = V_p - U_{CC} \tag{7}$$

where

V_p = two-way cooling coil valve position.

Residuals such as R_U require the comparison of measured values to model values and can cause difficulties arising from the use of models. The most obvious problem is model error. Even if the model is accepted as being an accurate representation of the physical process, it may require the identification of parameter values specific to each physical system. In addition, the characteristics of the system can change over time and require that the parameters be identified periodically. However, these models and the residuals based upon them contain the underlying physics of the process(es) involved. They provide information on state variables, such as the rate of heat transfer from a coil, that can not be (easily) measured directly. Thus, it is not practical to think that such model based residuals can be eliminated. As stated previously, and as will become more apparent in the next section, R_U is a symptom of several of the faults studied here. Hence a reliable model of the operation of the cooling coil valve under normal operating conditions is needed and part of this study is devoted to a discussion of the use of an ANN for this purpose.

FAULT DESCRIPTION

Faults are typically classified as belonging to one of two categories, namely, faults due to a complete failure of a component or system and faults due to performance degradation. The main factor used in categorizing faults is the rate at which they occur. Complete failures typically occur abruptly, although they may be due to factors such as equipment wear that take place over years of use. Faults involving performance degradation evolve over periods of time that are typically measured in weeks, months, or years. Faults of this nature are difficult to detect in their early stages because only subtle changes occur in the component or system performance. The faults considered in this study represent complete failures of components of the system. Complete failures are considered because they can be easily introduced in a laboratory system and the fault symptoms can be observed almost immediately. Simulation studies are more appropriate for faults caused by performance degradations.

Eight faults representing complete failures of various components in the AHU are described

in the ensuing paragraphs. The dominant symptoms of each fault are also described. The faults are introduced when the system is operating at normal, steady-state conditions and the dominant symptoms correspond to the steady-state conditions after a fault has occurred. With the exception of the pump fault, all faults are simulated in the laboratory AHU by either sending faulty control signals from the PC to an actuator, or by overwriting sensor signals that are logged by the DAS with faulty values. In each case, the faulty control or sensor signal is equal to its minimum possible value (usually zero). The pump fault is introduced manually by reducing the pressure of the cooling water supplied to the cooling coil.

Fault #1 is a failure of the supply fan. During normal operation the supply fan is controlled to maintain a static pressure of 249 Pa (1.0 in. of water) in the supply air duct and the return fan is controlled to maintain a flow difference of 0.472 m^3/s (1000 ft^3/min) between the supply and return air ducts. The fault causes the supply fan rotational speed to decrease to zero, the supply air pressure to decrease to zero, and the control signal to the supply fan to increase to its maximum value in an attempt to offset the decreasing supply air pressure. The control signal for the return fan decreases to zero in an attempt to maintain the flow difference between the supply and return air ducts at the setpoint temperature, however, this condition can not be achieved due to the fault. Because there is no air flow, the supply air temperature gradually increases with the resultant effect of increasing the cooling coil valve control signal to its maximum value. Thus, the dominant residuals for fault #1 are R_P, R_Q, R_U, and R_{NS}.

Fault #2 is a failure of the return fan. The fault causes the return fan rotational speed to decrease to zero, the flow difference between the supply and return ducts to increase, and the control signal to the return fan to increase to its maximum value in an attempt to offset the increasing flow difference. Thus, the dominant residuals for fault #2 are R_Q and R_{NR}.

Fault #3 is a failure of a chilled water pump. It is assumed that more than one pump is used to deliver the chilled water to the AHU and, therefore, the fault causes the water flow rate to decrease, but not to zero. The decrease in the flow rate of cooling water causes the supply air temperature to increase initially. This causes the cooling coil valve signal to increase, thus

opening the valve. By opening the cooling coil valve it may be possible to bring the supply air temperature back to the setpoint value; however, the control signal to the cooling coil valve will be different from the normal condition. The dominant residual for fault #3 is R_U.

Fault #4 is a stuck cooling coil valve. After the fault is introduced, near normal operation continues and the residuals remain near zero until a disturbance occurs that calls for a significant change in the valve position. As an example, a load decrease in a zone causes the damper of the VAV box for that zone to close, thus increasing the static pressure in the supply duct. This causes the supply fan speed to decrease to bring the static pressure back down to the setpoint value and, consequently, the supply air temperature decreases. In an attempt to compensate for the decreasing supply air temperature, a signal is sent to the cooling coil valve to close further. However, because the valve is stuck, the position does not change. Over time, the integrator portion of the control algorithm causes the cooling coil control signal to decrease to its minimum value. The dominant residual for fault #4 is R_V.

Fault #5 is a failure of the supply air temperature thermocouple. A thermocouple failure typically results in a voltage signal that varies randomly between large positive and negative values. If the sensed value of the supply air temperature is outside of the range of normal operating conditions (0 to 40°C, for example), the temperature could be automatically set to zero so that the residual R_T would not fluctuate. This type of failure is simulated by overwriting the sensed supply air temperature with a value of 0°C. A zero supply air temperature signal causes the signal to the cooling coil control valve to decrease to its minimum value and thus close the valve in an attempt to raise the supply air temperature. The dominant residuals for fault #5 are R_T and R_U.

Fault #6 is a failure of the supply air pressure transducer. When this failure occurs, a zero reading is obtained for the supply air pressure (by overwriting the actual value). This causes the control signal to the supply fan to increase to its maximum value in an attempt to increase the supply pressure. The supply air flow rate increases for a short period of time (until the VAV boxes respond) and this makes it necessary for the cooling coil valve to open further to maintain

the supply air temperature at the setpoint value. The return fan control signal also increases for a short period of time in order to maintain the flow rate difference between the supply and return ducts at the setpoint value. The dominant residual for fault #6 is R_P.

Fault #7 is a failure of the supply fan flow station. When this fault occurs a zero reading is obtained for the supply flow station (by overwriting the actual value) and the return fan controller believes that there is no flow in the supply duct. Thus, the control signal to the return fan is decreased to its minimum value in an attempt to maintain the flow difference between the supply and return ducts at the setpoint value. However, the measured flow difference can only approach zero and does not reach the setpoint value because this would require a negative flow of 0.472 m^3/s (1000 ft^3/min) in the return duct. The dominant residual for fault #7 is R_Q.

Fault #8 is a failure of the return fan flow station. When this fault occurs a zero reading is obtained for the return flow station (by overwriting the actual value) and the return fan controller believes that there is no flow in the return duct. Thus, the control signal to the return fan is increased to its maximum value in an attempt to maintain the flow difference between the supply and return ducts at the setpoint value. However, the measured flow rate difference is unchanged by this compensation due to the presence of the fault. The dominant residual for fault #8 is R_Q.

ARTIFICIAL NEURAL NETWORKS

Introduction

The ANNs used in this study have a multilayer feedforward network structure and are trained using a backpropagation learning rule. Multilayer feedforward networks consist of an input layer, an output layer, and one or more hidden layers. A schematic diagram of a multilayer feedforward network with one hidden layer is shown in Figure 2. The inputs to the n_i input units are denoted x_1, x_2, ... x_{n_i}, the outputs of the n_o output units are denoted y_1, y_2, ... y_{n_o}, and outputs of the n_h hidden layer units are denoted h_1, h_2, ... h_{n_h}. The non-shaded units are bias units whose inputs are set equal to unity. The connections between the units of different layers of the network are weights and biases. The variable names assigned to particular weights and biases are given in Figure 2 and correspond to the dotted line connections in the figure. The ANN is trained to learn the functional

mapping of inputs to outputs using input/output training pairs. The output training data are referred to as the target output of the ANN. The goal is to train the network until the output of the ANN is suitably close to the target output (Hertz et al. 1991).

Consider initially the forward pass through the network. For a specific input pattern (set of input values), the output of the jth hidden layer unit is given by

$$h_j = f\left(\sum_{i=1}^{n_i} w'_{ji} x_i + b'_j\right) \tag{8}$$

where

f = activation function

w'_{ji} = strength of connection from ith input unit to jth hidden layer unit

b'_j = bias value for jth hidden layer unit.

The output of the kth output unit is given by

$$y_k = f\left(\sum_{j=1}^{n_h} w_{kj} h_j + b_k\right) \tag{9}$$

where

w_{kj} = strength of connection from jth hidden layer unit to kth output unit

b_k = bias value for kth output unit.

The backpropagation algorithm uses a gradient descent algorithm to update the weights and, therefore, the activation functions must be differentiable. The activation functions used for the ANNs in this study are

$$f(x) = x \tag{10a}$$

$$f(x) = \frac{1}{1 + e^{-x}} \tag{10b}$$

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{10c}$$

where the functions given by Equations (10a) to (10c) are referred to as the pure linear function,

the log-sigmoid function, and the tan-sigmoid function, respectfully. The result of the forward pass is the output pattern $y_1, y_2, ... y_{n_o}$.

As stated previously, training is continued until the output patterns are suitably close to the target patterns. Mathematically this is achieved by minimizing the sum of squares error (SSE) given by

$$SSE = \sum_{p=1}^{n_p} \sum_{k=1}^{n_o} (t_{k,p} - y_{k,p})^2 \tag{11}$$

where

$t_{k,p}$ = target value for the kth output unit of the pth pattern

$y_{k,p}$ = actual value for the kth output unit of the pth pattern

n_p = total number of training patterns.

From Equation (11), SSE is computed by summing over all n_o output values for all n_p training patterns.

The ANN is trained by updating the weights using a backpropagation learning rule. The change in weight w'_{ji} is based on the gradient descent rule and is given by

$$\Delta w'_{ji} = -\eta \frac{\partial(SSE)}{\partial w'_{ji}} \tag{12}$$

where

η = learning rate.

A more complete description of ANNs and the backpropagation algorithm is given by Hertz et al. (1991).

Application of an ANN to the Cooling Coil Valve Subsystem

To compute residual R_U, a model is needed to determine the expected value of the cooling coil valve control signal $U_{CC,EV}$. An ANN can also be utilized for this purpose. Curtiss et al. (1993) described the modeling of a heating coil using a neural network where the objective was to determine the load on the coil for the next time step. For this study the goal is to determine the

current value of U_{CC} for normal operating conditions. A schematic diagram of the cooling coil and the cooling coil valve subsystem is shown in Figure 3. T_M and ϕ_M are the mixed air temperature and relative humidity, respectively, Q_S is the supply air flow rate, and T_{WI} is the temperature of the cooling water at the inlet to the cooling coil. The other variables retain their previous definitions.

The ANN used to model the cooling coil valve subsystem has a single hidden layer with 10 units. Knowledge of the physical process and extensive training and testing of different combinations of input variables and network topologies were utilized to identify the inputs to the ANN. As discussed later in this section, the inputs represent a tradeoff between performance of the ANN under normal and faulty conditions. The input and output variables for training are:

Inputs $Q_S(i)$, $Q_S(i-1)$, $T_S(i)$, $T_S(i-1)$, $T_M(i)$, $T_M(i-1)$, $T_{WI}(i)$, $T_{WI}(i-1)$,

                $\phi_M(i)$, $\phi_M(i-1)$, $Q_S(i)$ $[T_M(i) - T_S(i)]$, $Q_S(i-1)$ $[T_M(i-1) - T_S(i-1)]$

Output $U_{CC}(i)$

where (i) refers to the current discrete time value and $(i-1)$ refers to the previous value. Inputs of the form $Q_S(i)$ $[T_M(i) - T_S(i)]$ are measures of the load on the coil at a particular time. The ANN is trained in a batch mode (off-line) using experimental data obtained as the system operates in a normal mode. The training data consists of 2271 input/output patterns and training proceeds until the average error for each training pattern is approximately 0.0015. A tan-sigmoid activation function is used for the hidden layer and a pure linear activation function is used for the output layer. A commercial ANN software package is used for the training (Demuth and Beale 1992).

The actual value of U_{CC} and the ANN model value of $U_{CC,EV}$ are plotted as a function of time (denoted t) in Figure 4 for a stuck valve fault (fault #4). The fault and a load decrease are at $t = 1800$ s. The load decrease at 1800 s causes U_{CC} to decrease to its minimum value (1 V in this case) in an attempt to bring the supply air temperature up to the setpoint temperature. $U_{CC,EV}$ also shows a decrease at $t = 1800$ s, however, this is due to the decrease in the supply air flow rate that occurs when the load decreases. A distinct difference in the two signals is observed and

this difference is used to compute R_U. This plot demonstrates that the ANN model responds to normal system changes in the appropriate manner; however, the model does not respond to the changes caused by the fault.

For all testing, the supply air temperature inputs to the ANN are replaced by the supply air temperature setpoint. This is done to try to avoid contaminating the ANN inputs with faulty data. In addition, $Q_S(i)$ is monitored so that if its value goes to zero, the input to the ANN model is modified so that $Q_S(i)$ is equal to its average value from the previous 20 time steps. Hence, faulty data associated with the supply fan fault (fault #1) and the supply fan flow station fault (fault #7) has only a minimal effect on the computation of the expected value of the cooling coil valve control signal, $U_{CC,EV}$.

In general the ANN model is susceptible to faulty input data, as would any model that uses real data as input. This is a key issue in the development of the ANN model because the goal is to predict the operation of the valve for normal conditions, not for fault conditions. For the latter case, a sufficiently well trained ANN would simply track the faulty control signal and the residual R_U would not indicate the presence of a fault. In this study this was avoided by not using past values of U_{CC} as inputs to the ANN model.

The input training data does not exhibit a great deal of variation for T_{WI}, ϕ_M, and T_M and, therefore, the ANN can only be used reliably for a relatively small range of these variables. Future effort will be devoted to collecting data over a wider range of conditions; however, the current set of training data is sufficient to demonstrate this application of ANNs. The need for a large training data set that covers the complete range of operating conditions for the process is a practical consideration that must be overcome to implement this model.

Application of an ANN to Fault Diagnosis

To utilize an ANN for fault diagnosis, the ANN must first be trained using data that is representative of the normal condition and of the various fault conditions. The inputs are seven normalized values of the residuals in Equations (1) to (7) and the outputs are nine values that constitute a pattern that represents the normal mode or one of the eight fault modes of operation.

Hence, nine input/output patterns are used to train the network. Actual measured data for normal operation may be available from historical databases or can be obtained as the system operates. However, this may not be the case for fault conditions. Introducing a fault to the system so that fault data can be collected may not be possible due to concerns for occupant comfort. Hence, an alternative method for obtaining patterns of residuals during fault modes of operation is needed.

In this study, idealized training patterns are specified by considering the dominant symptoms of each fault. Following the discussion in the section entitled FAULT DESCRIPTION, examples of dominant symptoms/residuals for several faults are:

IF Supply fan failure   **THEN** Supply fan RPM is zero.
Supply air pressure is zero.
Supply fan control signal is maximum.
Flow difference between supply and return ducts is zero.

IF Return fan failure   **THEN** Return fan RPM is zero.
Flow difference between supply and return ducts increases.
Return fan control signal is maximum.

IF Pump failure   **THEN** Cooling coil valve control signal changes.
Cooling coil valve position changes.

IF Cooling coil control valve failure   **THEN** Cooling coil valve control signal changes.
Cooling coil valve position does not change.

Using this type of reasoning it is possible to construct a pattern of dominant training residuals for each fault. The matching of dominant residuals to the various faults is depicted in Figure 5.

The residuals are normalized so that the dominant symptom residuals have the same magnitude for the different fault cases. A dominant symptom residual is assigned a value of ± 1 depending on the sign of the residual, and all other residuals are assigned a value of 0. The idealized input/output training patterns for the normal mode of operation and the eight faults are given in Table 1. The input patterns are based on conditions that are expected to exist after the system has reached steady state. Each output training pattern consists of eight values of 0 and one value of 1. The normal mode has a 1 as the output for the first unit, fault #1 has a 1 as the

output for the second unit, and so on.

For testing of actual data, a normalized residual is obtained by dividing a residual from Equations (1) to (7) by the absolute value of the maximum value obtained for this residual from measured fault data. Thus, the maximum value of the absolute value of R_T obtained for a particular fault is used to normalize the supply air temperature residual for all the considered faults. The normalized residual for the supply air temperature is:

$$R_T = \frac{T_S - T_{S,SP}}{\left|T_S - T_{S,SP}\right|_{max}} \tag{13}$$

where subscript "max" denotes maximum.

The ANN architecture is 7×5×9 where the first number is the number of inputs (residuals; n_i), the last number is the number of outputs (normal mode plus eight fault modes; n_o), and the middle number is the number of units in the hidden layer (n_h). A log-sigmoid activation function is used for both the hidden and output layers. The network is trained until the sum of squares error is less than 10^{-6} or until the number of training epochs exceeds 5000. A commercial ANN software package was used for the training (Demuth and Beale 1992).

The methodology described above is one example of how ANNs can be used for fault detection and diagnosis. Model-based approaches are a second example of the use of ANNs for this task. Model-based approaches compare ANN models of normal and faulty system or subsystem operation to the actual system operation. Diagnosis of the current state of the system is based on determining which model has the greatest degree of similarity to the actual system operation. The model-based approach is not used in this study.

Network Topology and Training

The selection of the appropriate number of hidden layers and the number of units in a layer is problem dependent and typically requires considerable engineering judgment (Schalkoff 1992). As is the case for most numerical algorithms, a tradeoff between accuracy and computational requirements may exist. For instance, by adding more hidden units and layers to a network, the

agreement between the actual and target outputs may be improved, but at the cost of increased training time and memory requirements. In addition, if too many hidden units are used, overfitting of the training data may occur and the generalization to new input patterns may be very poor. This is similar to the effect seen when curve fitting with too many free parameters (Hertz 1991).

Hecht-Nielson (1990) provides guidelines for training and testing ANNs. The basic approach is to divide the available data into a training and a testing set. Both sets should include data that covers the full range of operating conditions. The amount of training required to yield a sufficiently accurate ANN is also problem dependent. For most ANNs there exists an optimum number of training epochs which minimizes the error for the testing data. Additional training epochs will most likely yield lower training errors; however, the errors for the test data may increase. This phenomenon occurs commonly for ANNs trained with the backpropagation learning rule and is known as overtraining. Overtraining can result in an ANN that exhibits poor generalization because the ANN simply memorizes the input training patterns. In most cases, the appropriate network topology and number of training epochs can only be determined through an extensive trial and error process.

RESULTS AND DISCUSSION

Faults in the AHU are diagnosed by inputting residual vectors to the trained ANN. The residual vectors are obtained by introducing faults in the laboratory AHU and recording the subsequent response of the system. The system response is input in a batch mode to the trained ANN model of the cooling coil value subsystem to compute $U_{CC,EV}$. This computation could also be done on-line. Residuals are calculated using steady-state values of the system variables measured 900 s after a fault is introduced. Residuals for the normal and eight fault modes of operation are given in Table 2. Normalized residuals computed using expressions such as Equation (13) are given in Table 3 and are used as input to the ANN for fault diagnosis.

A more systematic approach to determining the steady-state residual values would be to develop and implement a steady-state detector. One possible method for detecting steady-state

conditions would be to use regression techniques to obtain linear equations that characterize the responses of variables such as T_S, P_S, and Q_D. Steady-state conditions would be indicated if the slopes of these lines were less than the threshold values defined for each signal. A steady-state detector such as this would be necessary for on-line implementation in a real system because the onset of a fault is not known a priori.

Results of the fault diagnosis are given in Table 4. From the training patterns in Table 1, a perfect diagnosis would yield values on the diagonal of unity, and all other values would be zero. The values on the diagonal in Table 4 are near unity (underlined), indicating that the ANN successfully diagnosed each condition. Thus, although the training was based on simple, idealized relationships between the symptoms and faults, the ANN accurately discriminates between the various faults and the normal condition when actual data is used. Because of their ability to generalize and to filter noise, ANNs appear to be useful tools for fault diagnosis. For the set of faults and the associated symptoms considered in this study, fault diagnosis methods based on IF-THEN rules, or pattern recognition techniques such as the nearest neighbor algorithm (Schalkoff 1992) are also expected to be effective. As the number of faults increases, however, implementation of the IF-THEN rules may become cumbersome. Extension of the ANN method for fault diagnosis to include other faults is expected to be straightforward. The major computational requirement for ANNs occurs during training (which can be performed off-line) and this is not expected to present a problem during on-line operation.

It should be noted that the success that was achieved for this set of data is not guaranteed in general. Because training is based on a small set of idealized data, generalization problems can occur when actual data is considered. This potential difficulty is easily envisioned for normalized residuals with values near ±0.5, as for fault #3 in Table 3. Although this fault is correctly diagnosed, this set of residuals could also have been identified as a normal operating condition. In this case the problem is linked to the severity of the fault. That is, a more severe pump fault would produce a larger value of U_{CC} and, therefore, a larger value of R_U. This would improve the likelihood of a correct diagnosis.

A second case where generalization may be imperfect relates to the state of the system when the fault occurs. As an example, consider the stuck valve fault. If the valve sticks in a position where it is roughly half open and the value used for normalization is based on the maximum possible difference between the actual and expected position, the corresponding normalized residual R_V will again be near ±0.5. Thus, the same kind of generalization problem as cited previously could be encountered.

The reliability of the ANN for diagnosing imperfect input data patterns can be improved in two ways. First, the input training data set can be extended to include patterns that account for less severe faults and faults related to the state of the system prior to the occurrence of the fault. For instance, the training patterns for the pump fault and valve fault could be extended as shown in Table 5. This kind of extension of the input training data is rather straightforward to envision once the patterns for the severe faults have been established. A second way in which the ANN can be taught to generalize more accurately is by training the network with noise added to the idealized input patterns. For instance, random noise that is normally distributed with a mean value of 0 and a variance of 0.1 can be added to the input patterns in Table 1. The training data then consists of the original idealized input patterns and the noisy patterns. Additional noisy input patterns with a different values of the variance can also be added to the training data set. Both approaches are currently being investigated to improve the robustness of the diagnosis.

It is also possible that the input to the ANN will represent a fault mode of operation for which the ANN was not trained. In fact, it seems very probable that this will occur occasionally and, therefore, must be accounted for in an actual implementation of the method. The most desirable output in this scenario would be a warning that the system is operating in some unknown fault mode. However, because the training data does not include this type of input, the output may be erroneous. Although this may appear to be a drawback of the ANN method, the reality is that this scenario is likely to cause problems regardless of the method used for fault diagnosis.

The diagnosis of experimental faults is based on steady-state or near steady-state conditions

and dynamic conditions are not considered. Further study is needed to determine how the method can be extended for use when dynamic conditions exist.

CONCLUSIONS AND RECOMMENDATIONS

The objective of this study was to describe the application of ANNs to the problem of fault diagnosis in an AHU. Initially, residuals of system variables were selected that could be used to quantify the dominant symptoms of fault modes of operation. Idealized steady-state patterns of these residuals were then defined for each mode of operation studied. The steady-state relationship between the dominant symptoms and the faults was learned by an ANN using the backpropagation algorithm. The trained ANN was applied to experimental data for various faults and successfully identified each fault.

An ANN was also used successfully as a model for a cooling coil valve subsystem. The output of the ANN was the expected value of the cooling coil valve control signal. Although the agreement between the actual and predicted control signal during normal operation was not perfect, the ANN model was adequate for identifying normal and fault modes of operation. The agreement for normal operating conditions could be improved by changing the inputs to the ANN, however, this may lead to a model that tracks the operation of the valve during fault conditions, rather than providing an estimate of the normal operation of the valve under normal operating conditions.

This study demonstrates the feasibility of using ANNs for diagnosis of faults in HVAC systems. Eight fault modes of operation were considered and all faults were of a severe nature. Hence, the symptoms of these faults are relatively easy to distinguish. Nonetheless, it is anticipated that, because of their abilities to learn complex, nonlinear relationships and to generalize, ANNs will also be effective for less severe faults.

The method can be extended in a straightforward manner to consider additional faults such as damper faults in the mixing box. It is likely that this will require the introduction of additional residuals to the analysis. As the complexity of the system and the number of faults considered grows, it may be desirable to develop separate ANNs for various subsystems and to use a

preprocessor to identify the appropriate subsystem to examine for the existence of a fault.

ACKNOWLEDGMENTS

REFERENCES

Ananthraman, S. 1995. Intelligent process optimization on the plant floor. *Industrial Computing*. Vol. 14 (4): 40-44.

Curtiss, P.S., J.F. Kreider, and M.J. Brandemuehl. 1993. Energy management in central HVAC plants using neural networks. *ASHRAE Transactions*. Vol. 99 (1): 496-504.

Demuth, H., and M. Beale. 1992. *Neural Network Toolbox: User's Guide*. Natick, Massachusetts: The MathWorks, Inc.

Fan, J.Y., M. Nikolaou, and R.E. White. 1993. An approach to fault diagnosis of chemical processes via neural networks. *AIChE Journal*. Vol. 39 (1): 82-88.

Hecht-Nielson, R. 1990. *Neurocomputing*. Reading, Massachusetts: Addison-Wesley.

Hertz, J., A. Krogh, and R.G. Palmer. 1991. *Introduction to the theory of neural computation*. Reading, Massachusetts: Addison-Wesley.

Koive, H.N. 1994. Artificial neural networks in fault diagnosis and control. *Control Eng. Practice* Vol. 2 (1): 89-101.

Lee, W.Y., C. Park, and G.E. Kelly. 1995. Fault detection of an air handling unit using residual and recursive parameter identification methods. Under review for *ASHRAE Transactions*.

Schalkoff, R.J. 1992. *Pattern recognition: statistical, structural, and neural approaches*. New York, New York: John Wiley & Sons.

Watanabe, K., S. Hirota, L. Hou, and D.M. Himmelblau. 1994. Diagnosis of multiple simultaneous faults via hierarchical artificial neural networks. *AIChE Journal*. Vol. 40 (5): 839-848.

Table 1 Normalized patterns for AHU fault diagnosis used in ANN training.

Net Inputs							Net Outputs	Fault Diagnosis
R_P	R_Q	R_T	R_U	R_{NS}	R_{NR}	R_V		
0	0	0	0	0	0	0	1 0 0 0 0 0 0 0 0	Normal
-1	-1	0	1	-1	0	0	0 1 0 0 0 0 0 0 0	#1 Supply fan
0	1	0	0	0	-1	0	0 0 1 0 0 0 0 0 0	#2 Return fan
0	0	0	1	0	0	0	0 0 0 1 0 0 0 0 0	#3 Pump
0	0	0	0	0	0	1	0 0 0 0 1 0 0 0 0	#4 Cooling coil valve
0	0	-1	-1	0	0	0	0 0 0 0 0 1 0 0 0	#5 Thermocouple
-1	0	0	0	0	0	0	0 0 0 0 0 0 1 0 0	#6 Pressure transducer
0	-1	0	0	0	0	0	0 0 0 0 0 0 0 1 0	#7 Supply flow station
0	1	0	0	0	0	0	0 0 0 0 0 0 0 0 1	#8 Return flow station

Table 2 Measured residuals for normal operation 900 s after the occurrence of a fault.

Fault	System Operation	R_P	R_Q	R_T	R_U	R_{NS}	R_{NR}	R_V
	Normal	-0.004	0.011	-0.080	0.130	0.037	0.065	0.011
#1	Supply fan fault	-0.249	-0.398	0.790	4.005	-9.983	0.017	-0.249
#2	Return fan fault	0.004	0.652	0.063	-0.097	0.071	-9.983	-0.042
#3	Pump fault	-0.004	0.079	0.013	2.145	0.043	0.072	-0.354
#4	Cooling coil valve fault	-0.003	-0.068	-1.095	-1.577	0.044	0.016	3.247
#5	Thermocouple fault	0.002	-0.027	-14.500	-3.697	0.063	0.056	-0.274
#6	Pressure transducer fault	-0.249	0.036	0.010	0.529	-0.003	0.079	-0.010
#7	Supply flow station fault	0.005	-0.868	-0.072	-0.086	0.055	0.017	0.014
#8	Return flow station fault	-0.001	1.072	0.000	-0.299	0.048	-0.011	0.099

Table 3 Normalized ANN input for fault diagnosis.

Fault	System Operation	R_P	R_Q	R_T	R_U	R_{NS}	R_{NR}	R_V
	Normal	-0.016	0.010	-0.006	0.032	0.004	0.007	0.003
#1	Supply fan fault	-1.000	-0.371	0.054	1.000	-1.000	0.002	-0.077
#2	Return fan fault	0.016	0.608	0.004	-0.024	0.007	-1.000	-0.013
#3	Pump fault	-0.016	0.074	0.001	0.536	0.004	0.007	-0.109
#4	Cooling coil valve fault	-0.012	-0.063	-0.076	-0.394	0.004	0.002	1.000
#5	Thermocouple fault	0.008	-0.025	-1.000	-0.923	0.006	0.006	-0.084
#6	Pressure transducer fault	-1.000	0.034	0.001	0.132	0.000	0.008	-0.003
#7	Supply flow station fault	0.020	-0.810	-0.005	-0.021	0.006	0.002	0.004
#8	Return flow station fault	-0.004	1.000	0.000	-0.075	0.005	-0.001	0.030

Table 4 Diagnosis results for the data in Table 3.

System Operation	Output Pattern								
Normal	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Supply fan fault	0.000	1.000	0.000	0.000	0.000	0.000	0.022	0.000	0.000
Return fan fault	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
Pump fault	0.017	0.001	0.000	0.927	0.000	0.000	0.000	0.000	0.000
Cooling coil valve fault	0.004	0.000	0.000	0.000	0.998	0.000	0.041	0.001	0.000
Thermocouple fault	0.000	0.000	0.102	0.000	0.000	1.000	0.000	0.000	0.000
Pressure transducer fault	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
Supply flow station fault	0.005	0.000	0.000	0.000	0.000	0.002	0.000	0.999	0.000
Return flow station fault	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Table 5 Additional training patterns for select faults.

Net Inputs							Net Outputs										Fault Diagnosis	
R_P	R_Q	R_T	R_U	R_{NS}	R_{NR}	R_V												
0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	#3	Pump
0	0	0	0.5	0	0	0	0	0	0	1	0	0	0	0	0	#3	Pump
0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	#4	Cooling coil valve
0	0	0	0	0	0	0.5	0	0	0	0	1	0	0	0	0	#4	Cooling coil valve
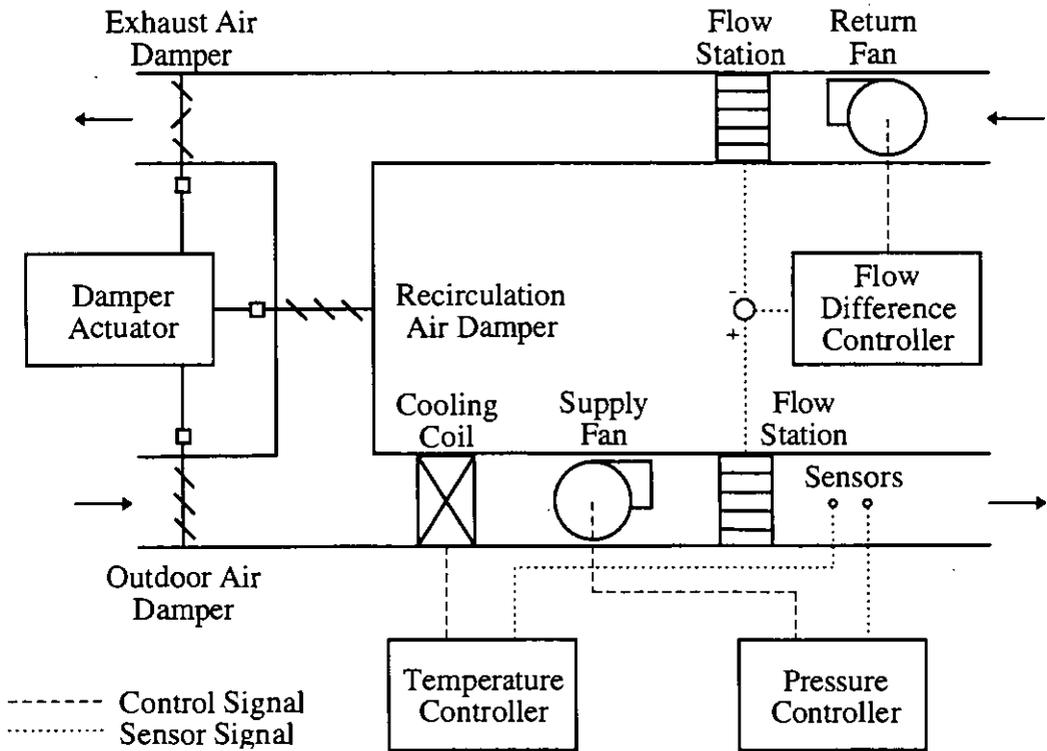
Figure 1 Schematic diagram of a variable-air-volume air handling unit.
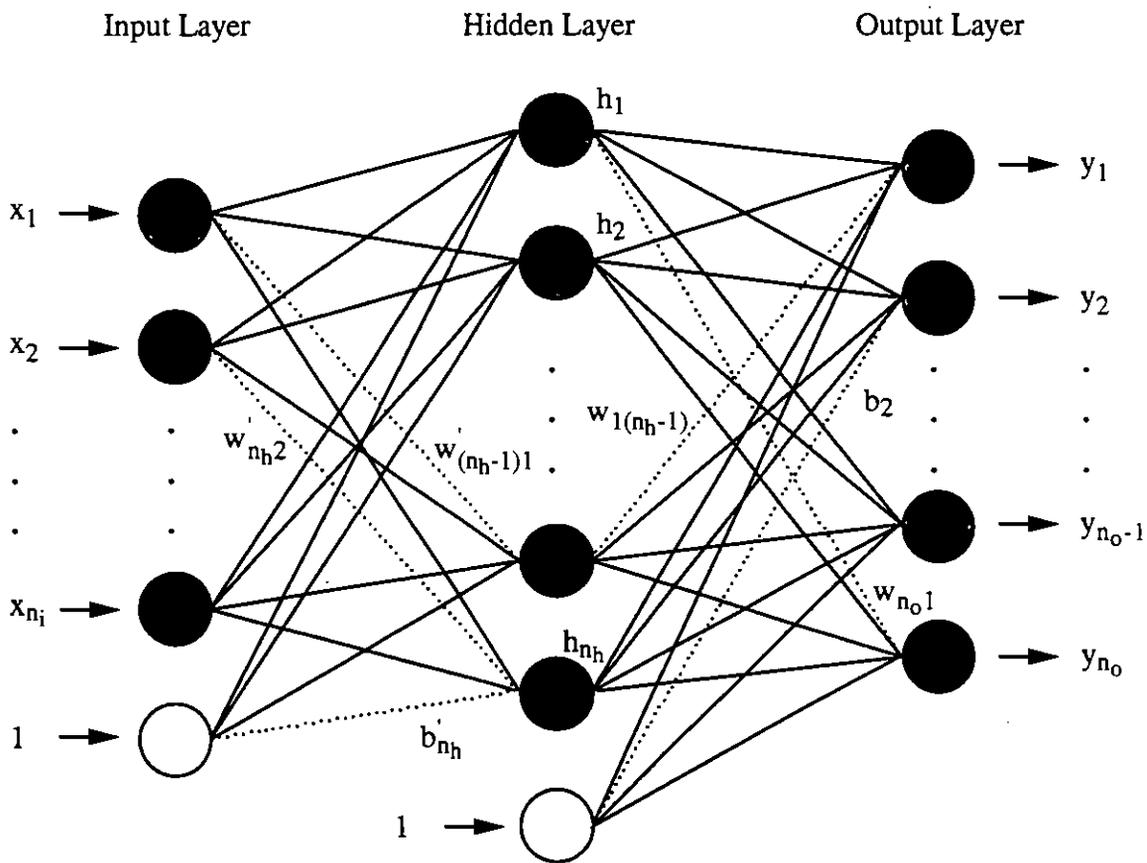
Figure 2 Two layer feedforward network.
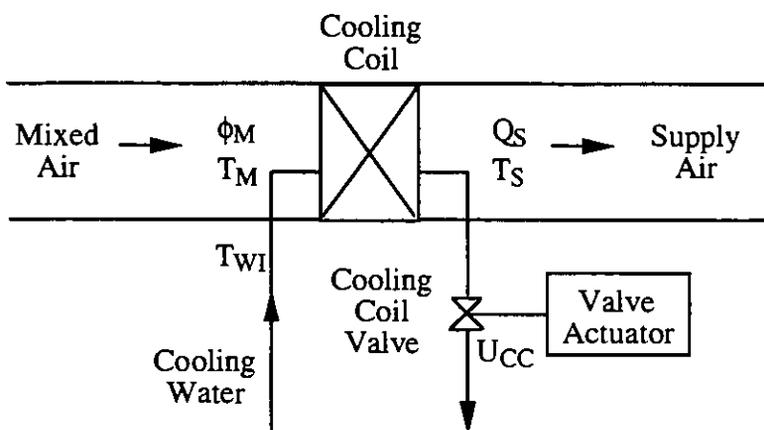
Figure 3 Schematic diagram of the cooling coil and cooling coil valve subsystem.
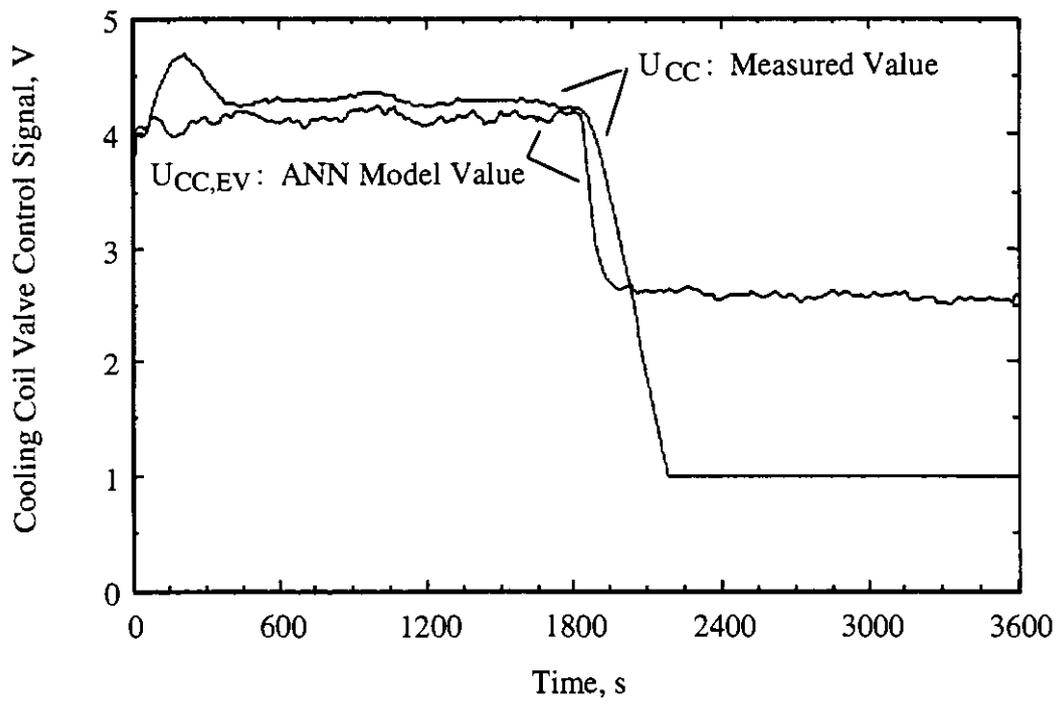
Figure 4 Actual and predicted cooling coil valve control signals.
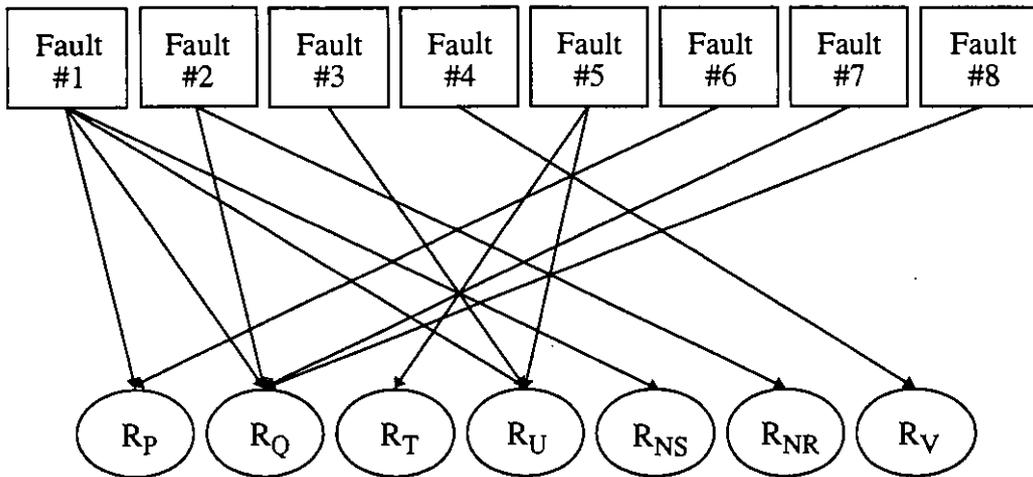
Figure 5 Matching of dominant residuals and faults.

FIGURE CAPTIONS

Figure 1 Schematic diagram of a variable-air-volume air handling unit.

Figure 2 Two layer feedforward network.

Figure 3 Schematic diagram of the cooling coil and cooling coil valve subsystem.

Figure 4 Actual and predicted cooling coil valve control signals.

Figure 5 Matching of dominant residuals and faults.

FAULT DIAGNOSIS AND TEMPERATURE SENSOR RECOVERY

FOR AN AIR-HANDLING UNIT

Won-Yong Lee, Dong-Ryul Shin, and John M. House *

Korean Institute of Energy Research
Taejeon Yuseonggu Jangdong 71-2, 305-343, South Korea

*National Institute of Standards and Technology
Gaithersburg, MD 20899-0001 USA

ABSTRACT

This paper describes the use of a two-stage artificial neural network for fault diagnosis in an air handling unit and discusses the use of an estimating equation to identify a failed temperature sensor. The stage one neural network is trained to identify the subsystem in which a fault occurs. The stage two neural network is trained to diagnose the specific cause of a fault at the subsystem level. A regression equation is derived to recover an estimate of the supply air temperature when the sensor measurement is determined to be erroneous. The estimated value is used for fault diagnosis and as a feedback value for control purposes. The regression equation and two-stage artificial neural network are tested successfully using simulation data.

Keywords: air handling, classification, maintenance, modeling, monitoring, sensor recovery

INTRODUCTION

The presence of faults and the influence that they have on system operation is a real concern in the heating, ventilating, and air-conditioning (HVAC) community. A fault can be defined as an inadmissible or unacceptable property of a system or a component. Unless corrected, faults can lead to increased energy use, shorter equipment life, and uncomfortable and/or unhealthy conditions for building occupants.

Faults are not a new problem in the HVAC industry; however, technological advances have helped create both a need and an avenue for the development of fault detection and diagnosis tools. The need has been created by the ever increasing complexity of the HVAC systems and control strategies that are being installed in present day buildings. In many cases the complexity of the systems exceeds the understanding of the building operator. When this occurs, faults may go undetected or, perhaps worse, may be "corrected" by introducing changes to the system that compensate for the fault rather than eliminating it. The latter scenario could lead to energy waste and possibly to subsequent faults that are related to the initial (and still existing) fault. Technological advances have also made it possible to monitor these complex systems, thus providing the information that is needed to characterize and understand the current operating status of the systems. Fault detection and diagnostic methods can provide a bridge between possessing information and understanding its meaning.

One of the main purposes of fault detection and diagnosis is to detect failures of actuators and sensors that are used in the control systems. To improve the operational reliability of systems in general, it is necessary to validate measured sensor data, isolate failed sensors, and recover the failed measurement. Hence, sensor recovery is an important aspect of comprehensive fault detection and diagnosis methods.

In previous papers Lee et al. (1996a and 1996b) describe methods for fault detection and diagnosis in a variable-air-volume (VAV) air-handling unit (AHU). One approach used in those studies was to define residuals that provide a measure of the difference between the existing state of the system and the normal state. Residuals that are significantly different from zero represent

the occurrence of a fault. Lee et al. (1996b) described the use of a single artificial neural network (ANN) trained on idealized residual patterns to diagnose faults in various subsystems of the AHU. Training a network such as this, which accounts for all considered faults, can require extensive computational resources due in part to the number of inputs, hidden neurons, and outputs necessary to discriminate each particular pattern. In addition, if a new fault is added to the existing set, it is necessary to retrain the ANN because the knowledge stored in the network (values of weights and biases) is probably not adequate to discriminate the new fault.

The objective of this paper is twofold. The first objective is to describe an architecture for a two-stage ANN for fault diagnosis that can alleviate, to a certain degree, the problems discussed in the previous paragraph. The second objective is to describe the use of regression equations for sensor recovery of failed temperature sensors.

The first sections of this paper provide a brief description of the AHU, simplified models of the components of the AHU, the residuals used in the fault diagnosis and the faults under consideration. The two-stage ANN and the regression analysis used to develop estimating equations for temperature sensors are then discussed. Results of the fault diagnosis are then presented, followed by simulation results for the proposed sensor recovery approach. Finally, conclusions and recommendations for future work are presented.

AIR HANDLING UNIT

A schematic diagram of the VAV AHU used for this study is shown in Figure 1. The same system was used for papers by Lee et al. (1996a and 1996b). The AHU consists of fans, dampers, a cooling coil, sensors, and controllers. The static pressure in the main supply duct is maintained at a constant set point value of 249 Pa (1.0 in. of water) by sensing the static pressure and controlling the rotational speed of the supply fan. The supply air temperature is controlled by modulating the cooling water control valve to maintain a constant set point value of 14.5°C (58.0°F). The air flow rate difference between the supply and return air streams is controlled by the variable speed return fan to maintain a constant setpoint value of 0.472 m³/s (1000.0 ft³/min). A PID algorithm is used to control the cooling water valve and PI algorithms are used to control

the supply and return fan speeds. The sampling period for control and data collection is ten seconds.

SIMPLIFIED AHU MODELS

The AHU model is a simplified dynamic model based on steady-state characteristic equations and approximate first order dynamics. Models developed in IEA ANNEX 17 (Wang, 1992) have been modified to fit the ANNEX 25 reference system using experimental data.

Cooling coil characteristics

The air temperature at the exit of the cooling coil T_{ao} is obtained by combining a steady-state model of the cooling coil with approximate first-order dynamics. The response of T_{ao} is approximated by

$$\frac{T_{ao}}{T_{aos}} = \frac{1}{1 + \tau s} \, exp(-\tau_d \, s) \tag{1}$$

where T_{aos} is the steady-state value of T_{ao}, τ is the process time constant, and τ_d is the process delay time. To determine T_{aos}, it is first necessary to compute the steady-state value of the moist air enthalpy at the exit of the cooling coil h_{aos}, which is given by

$$h_{aos} = h_{ai} + \varepsilon \, (h_{wi} - h_{ai}) \tag{2}$$

where h_{ai} is the moist air enthalpy at the inlet to the cooling coil, h_{wi} is the moist air enthalpy at the inlet chilled water temperature, and ε is the effectiveness of the cooling coil. T_{aos} is then calculated from the expression

$$h_{aos} = c_{pa} \, T_{aos} + w_{ai} \, (c_{pv} \, T_{aos} + h_{g,o}) \tag{3}$$

where c_{pa} is the specific heat of dry air, c_{pv} is the specific heat of water vapor, w_{ai} is the humidity ratio of the air at the inlet to the cooling coil, and $h_{g,o}$ is the enthalpy of saturated water vapor at a reference temperature T_o. For a single pass crossflow heat exchanger where the fluid having the minimum capacitance (air in this case) is mixed and the water is unmixed, the effectiveness is given by (Incropera and DeWitt, 1985)

616

$$\varepsilon = 1 - exp\left(-C_r^{-1}\left\{1 - exp\left[-C_r (NTU)\right]\right\}\right) \tag{4}$$

where C_r is the heat capacity rate given by

$$C_r = C_{min} / C_{max} \tag{5}$$

and NTU is the number of transfer units given by

$$NTU = UA / C_{min} \tag{6}$$

As stated previously, C_{min} and C_{max} are the capacitances for air and water, respectively. UA is the overall heat transfer coefficient of the coil.

Damper characteristics

The airflow rates through the exhaust, recirculation, outdoor and VAV box dampers are given by:

$$Q = K \sqrt{\Delta P} \tag{7}$$

where K is a resistance coefficient, Q is the volumetric airflow rate, and ΔP is the pressure drop across the dampers. The resistance coefficients are determined from least-squares regression of the experimental data for varying damper blade angles. The resulting third order polynomial equations are of the general form

$$K = C_0 + C_1\theta + C_2\theta^2 + C_3\theta^3 \tag{8}$$

where θ is the damper angle. Assuming the air density is constant throughout the system, conservation of mass yields

$$Q_{REC} = Q_{RET} - Q_{EXH} \tag{9}$$

$$Q_{REC} = Q_S - Q_{OUT} \tag{10}$$

$$Q_S = \sum_{i=1}^{n} Q_{VAV,i} \tag{11}$$

where subscript *REC* denotes recirculation air, *RET* denotes return air, *EXH* denotes exhaust air, *S* denotes supply air, *OUT* denotes outdoor air, and *VAV,i* denotes airflow through the ith VAV box, and where n is the total number of VAV boxes.

Sensor model characteristics

Sensors convert physical signals into electric signals. The response of the electric signal may exhibit a time delay from the change of the physical value. The sensor response is simulated using a first order model with a simple time constant, that is

$$\frac{T_o}{T_i} = \frac{1}{1 + \tau s} \tag{12}$$

Fan characteristics

Fan performance is characterized using the following dimensionless coefficients:

$$\Phi = \frac{Q}{D^3 N} \tag{13}$$

$$\Psi = \frac{\Delta P}{\rho (N D)^2} \tag{14}$$

where Q is the volumetric airflow rate of air through the fan, D is the fan impeller diameter, N is the rotational speed of the fan, ΔP is the pressure rise across the fan, and ρ is the density of air. The dimensionless parameters in Eqs. (13) and (14) are related by the polynomial expression

$$\Psi = \sum_{i=0}^{m} a_i \Phi^i \tag{15}$$

where the coefficients a_i ($i = 1$ to m) are determined from least-squares regression.

Air filter characteristics

The airflow rate and pressure drop across the air filter and the duct, including the coils, are related by

$$Q_f = K_f \sqrt{\Delta P_f} \tag{16}$$

where K_f is a friction coefficient that is treated as being constant. In real systems K_f will vary due

to fouling of the filter and the coils.

SOLUTION PROCEDURE

The simplified model equations are solved by considering the pressure and flow equations independently of the equations that characterize the cooling coil response. Initial values of the damper positions, valve positions and fan speeds are selected first. The pressure and flow equations are then solved simultaneously to determine the pressure and airflow rates throughout the system. Next, the equations governing the cooling coil are solved to determine the supply air temperature and the inlet and exit cooling water temperatures. The supply air temperature and pressure, and the volumetric airflow rate difference between the supply and return air streams are then used to compute the control signals to the fans and the cooling coil valve. Finally, the control signals are converted to new values of the supply and return fan speeds and the cooling coil valve position. The solution procedure is then repeated. The simulation requires solving 15 equations (12 for the pressure and flow characteristics, 3 for the cooling coil) for 15 unknowns.

RESIDUAL DEFINITION

The fault diagnosis method described by Lee et al. (1996b) identifies patterns of residuals that can be used as signatures for various faults. An ANN is trained with these patterns and then used to diagnose the status of the AHU for actual experimental data. The set of residuals used in this study is slightly different from Lee et al. (1996b). The residuals used here are:

$$R_{TS1} = T_S - T_{S,EV} \tag{17}$$

$$R_{TS2} = T_S - T_{S,SP} \tag{18}$$

$$R_{TM} = T_M - T_{M,EV} \tag{19}$$

$$R_{PS} = P_S - P_{S,SP} \tag{20}$$

$$R_{QD} = Q_D - Q_{D,SP} \tag{21}$$

$$R_{VCC} = V_{CC} - U_{CC} \tag{22}$$

where R denotes residual, T is temperature, P is pressure, Q is volumetric airflow rate, U is control signal, and V is valve position. Subscript M denotes mixed air parameters, D denotes volume matching parameters, CC denotes cooling coil parameters, SP denotes setpoint values,

and *EV* denotes estimated values calculated using regression equations.

FAULT DESCRIPTION

Eleven faults of various components in the AHU are described in the followed paragraphs. The dominant symptoms of each fault are also described. The faults are introduced when the system is operating at normal, steady-state conditions and the dominant symptoms correspond to the steady-state conditions after a fault has occurred.

Fault #1 is a failure of the supply fan. The fault causes the supply fan rotational speed to decrease to zero, the supply air pressure to decrease to zero, and the control signal to the supply fan to increase to its maximum value in an attempt to offset the decreasing supply air pressure. The control signal for the return fan decreases to zero in an attempt to maintain the flow difference between the supply and return air ducts at the setpoint temperature. The dominant residual for fault #1 is R_{PS}.

Fault #2 is a failure of the return fan. The fault causes the return fan rotational speed to decrease to zero, the flow difference between the supply and return ducts to increase, and the control signal to the return fan to increase to its maximum value in an attempt to offset the increasing flow difference. The dominant residual for fault #2 is R_{QD}.

Fault #3 is a failure of a local feed water pump. The fault causes the water flow rate to decrease, but not to zero, since the main supply pumps continue to operate normally. The decrease in the flow rate of cooling water causes the supply air temperature to increase, with the resultant effect being that the cooling coil valve opens further. By opening the cooling coil valve it may be possible to bring the supply air temperature back to the setpoint value; however, the control signal to the cooling coil valve will be different from the normal condition. The dominant residual for fault #3 is R_{TSI}.

Fault #4 is a stuck cooling coil valve. A load change occurring after the introduction of the fault will cause the control signal to the cooling coil valve to saturate at either the minimum or the maximum voltage because the valve is unable to respond to the control input. The dominant residual for fault #4 is R_{VCC}.

Fault #5 is a complete failure of the supply air temperature sensor. A temperature sensor failure typically results in a voltage signal that varies randomly between large positive and negative values. If this occurs, the temperature is automatically (and instantaneously) set to zero so that the temperature residuals given by Eqs. (17) and (18) do not fluctuate. Setting the supply air temperature signal to zero causes the cooling coil control valve to close in an attempt to raise the supply air temperature. The dominant residuals for fault #5 are R_{TS1} and R_{TS2}.

Fault #6 is a second type of failure of the supply air temperature sensor. In this case, the sensor drops from its supporting harness onto the floor of the duct, giving an incorrect temperature reading. Because the air duct surface is assumed to be at a temperature that is 5°C higher than the air flowing through the duct, the controller attempts to compensate by opening the cooling coil control valve. The supply air temperature at time i is given by

$$T_S(i) = T_S(j) + 5 [1 - exp(-[i - j] / 100)] \tag{23}$$

where $T_S(j)$ is the supply air temperature at time j when the fault occurs and the time constant for the response is assumed to be 100 seconds. The dominant residuals for fault #6 are R_{TS1} and R_{TS2}.

Fault #7 is a third type of supply air temperature sensor failure and is due to sensor drift. This type of failure is classified as a performance degradation rather than a complete fault and would be difficult to detect in its early stages. To simplify the discussion for this fault, the period of time over which the sensor degrades is not considered. The fault is simulated by introducing an instantaneous decrease in the supply air sensor reading of 1°C. The cooling coil valve controller attempts to compensate for this fault by closing the control valve. After some transient period, the sensed value of the supply air temperature reaches the setpoint value, although the actual temperature is approximately 1°C higher than the setpoint value. The dominant residuals for fault #7 are R_{TS1} and R_{TS2}.

Fault #8 is a failure of the supply air pressure transducer. When this failure occurs, a zero reading is obtained for the supply air pressure. This causes the control signal to the supply fan to

increase to its maximum value in an attempt to increase the supply pressure. The dominant residual for fault #8 is R_{PS}.

Fault #9 is a failure of the supply air flow station. When this fault occurs a zero reading is obtained for the supply flow station and the return fan controller decreases the return fan speed to its minimum value in an attempt to maintain the flow difference between the supply and return ducts at the setpoint value. The dominant residual for fault #9 is R_{QD}.

Fault #10 is a failure of the return fan flow station. When this fault occurs a zero reading is obtained for the return flow station and the return fan controller increases the return fan speed to its maximum value in an attempt to maintain the flow difference between the supply and return ducts at the setpoint value. The dominant residual for fault #10 is R_{QD}.

Fault #11 is a failure of the mixing box damper linkage. The fault causes a discrepancy between actual and expected values of the airflow rates for the air streams entering and exiting the mixing box. The discrepancy in the airflow rates leads to discrepancies in the actual and expected temperatures at the exit of the mixing box. It is assumed that the recirculation damper is closed when the fault occurs. The dominant residual for fault #11 is R_{TM}.

TWO-STAGE ARTIFICIAL NEURAL NETWORK

To use an ANN for fault diagnosis, the ANN must first be trained using data that represents the normal condition and the various fault conditions. Lee et al. (1996b) used a single ANN to classify the operating status of the AHU. Nine possible modes of operation were considered, namely, the normal mode and eight separate fault modes. Training a network such as that, which includes patterns for all considered faults, requires fairly extensive computational resources. In addition, if a new fault is added to the existing set, the ANN must be retrained because the knowledge stored in the network (values of weights and biases) is not adequate for discrimination of the new fault. To lessen or eliminate such problems, an architecture for a two-stage ANN is proposed here. Stage one is used to classify the subsystem where a fault is occurring. Stage two is used to diagnose the cause of a fault on the subsystem level. Using this architecture, less information is required for diagnosis at a given stage. In this study the subsystem classifications

are the cooling coil subsystem, the pressure control subsystem, the flow control subsystem, and the mixing box damper subsystem. Figure 2 shows the subsystem classification of the faults and Figure 3 shows a block diagram of the proposed two-stage network. Although four subsystems are considered in this study, stage two ANN results are presented only for the cooling coil subsystem. Hence, reference to stage two ANN parameters or results throughout the remainder of this paper will imply those of the cooling coil subsystem.

Idealized patterns of residuals are specified for training by considering the dominant symptoms of each fault (Lee et al., 1996b). Residuals are normalized so that the dominant symptom residuals have the same size for the different fault cases. A dominant symptom residual is assigned a value of 1 or -1 depending on the sign of the residual, and all other residuals are assigned a value of 0. The idealized input/output training patterns for the stage one ANN are given in Table 1, and training patterns for the stage two ANN to diagnose faults in the cooling coil subsystem are given in Table 2. The input patterns are based on conditions expected to exist after the system has reached steady state. Each output training pattern consists of four values of 0 and one value of 1 for the stage one ANN. For the stage two ANN, each output training pattern consists of six values of 0 and one value of 1.

To use the two-stage ANN for fault diagnosis, residuals obtained from test data must also be normalized. Normalization is performed by dividing each residual by the minimum non-zero value of this residual for all faults. One notable exception to this rule is that the residual value used for normalization must be significantly larger than the largest value obtained during normal operation of the system. This will reduce the possibility of a false alarm. Thus, with the cited exceptions, the minimum value of the absolute value of R_{TS2} obtained for all faults is used to normalize this residual for all faults. Stated mathematically, the normalized value of R_{TS2} is

$$\bar{R}_{TS2} = \frac{T_S - T_{S,SP}}{\left| T_S - T_{S,SP} \right|_{min}} \tag{24}$$

where subscript "min" denotes minimum. Normalized residual values greater than unity are set

equal to unity.

The stage one and stage two ANN architectures are 4×5×5 and 3×5×7, respectively, where the first number is the number of inputs, the last number is the number of outputs, and the middle number is the number of neurons in the hidden layer. The networks are trained until the sum of squares error is less than 10^{-6} or until the number of training epochs exceeds 5000. A commercial ANN software package was used for the training (Demuth and Beale, 1992).

SENSOR RECOVERY OF THE COOLING COIL VALVE SUBSYSTEM

When a critical sensor reading is found to be erroneous, it is necessary to estimate its true value using correlated measurements. A simple approach is to have one estimating relation for each sensor reading that needs to be recovered. Given the normal operation data set, the training is straightforward.

To identify an erroneous value of T_S and to compute residual R_{TSI}, a model is needed for the expected value of the supply air temperature. For this study, a regression equation is used to estimate the current value of T_S for normal operating conditions. A schematic diagram of the cooling coil and the cooling coil valve subsystem is shown in Figure 4. T_M and ϕ_M are the mixed air temperature and relative humidity, respectively, Q_S is the supply air flow rate, and T_{WI} is the temperature of the cooling water at the inlet to the cooling coil. The other variables retain their previous definitions.

The input and output variables for the regression equation are:

Inputs $Q_S(i), Q_S(i-1), Q_S(i)^2, Q_S(i-1)^2$

           $T_M(i), T_M(i-1)$

           $T_{WI}(i), T_{WI}(i-1)$

           $\phi_M(i), \phi_M(i-1)$

           $U_{CC}(i), U_{CC}(i-1), U_{CC}(i)^2, U_{CC}(i-1)^2$

           $Q_S(i) T_M(i), Q_S(i-1) T_M(i-1)$

Output $T_{S,EV}(i)$

where *(i)* refers to the current discrete time value and *(i-1)* refers to the previous value.

The coefficients of the regression equation are computed using simulation data obtained as the system operated in a normal mode. The training data consists of 1,000 points. A model for the expected value of the mixed air temperature $T_{M,EV}$ is obtained in a similar manner.

The actual value of T_S and the regression model value for $T_{S,EV}$ are plotted as a function of time in Figure 5 for a pump failure (fault #3). A load increase and the fault occur at t = 1000 s and 2,000 s, respectively. At normal conditions, it is seen that the regression equation estimates T_S sufficiently well. Since the fault is a failure of the local pump, the water flow rate decreases, but not to zero. The decrease in the cooling water flow rate causes the supply air temperature to increase initially. This causes the cooling coil valve signal to increase, thus opening the valve. By opening the cooling coil valve it is possible to bring the supply air temperature back to the setpoint value; however, the control signal to the cooling coil valve will be different from the normal condition. The change in U_{CC} causes a change in $T_{S,EV}$, while the real value of T_S returns to the setpoint value after a transient period of approximately 100 s. From Figure 5 a distinct difference in the two signals is observed and this difference is used to compute R_{TS1}.

Figure 6 shows the case in which the temperature sensor fails completely at t = 2,000 s. The value of the sensor measurement is set to zero. Since the control of the cooling coil valve is based on this sensor measurement, the control valve is closed to make the supply air temperature increase. This action causes the actual supply air temperature to increase. The estimated value follows the actual value closely at normal conditions and relatively closely after the sensor failure. Although the agreement between the actual and estimated value during faulty operation is not perfect, the estimated value is sufficiently close to recover the sensor output and can be used when the sensor fails. It is conceivable that the regression equation could provide more accurate estimates of the supply temperature for fault conditions if data for those conditions are collected and used to calculate the regression coefficients.

SIMULATION RESULTS AND DISCUSSION

Fault diagnosis

The training phase of the two-stage ANN used idealized patterns of normalized residuals as

inputs. In the testing phase, data obtained from a simulation program based on simplified AHU component models are used as inputs. Faults are introduced to the simulation program through modifications of the computer algorithm. Residuals (measured and normalized values) for the stage one and stage two ANNs are given in Tables 3 and 4, respectively. Values used for normalization are indicated with bold typeface.

Results of the fault diagnosis for the stage one and stage two ANNs are given in Tables 5 and 6, respectively. In the stage one ANN, each fault is classified as belonging to one of four subsystems shown in Figure 2. As an example, the pump, valve, and temperature sensor failures are classified as cooling coil subsystem faults. From Table 1 the idealized output pattern for a cooling coil subsystem fault is (0 0 0 1 0). From Table 5 it is seen that these failures are correctly classified by the stage one ANN, that is, the output pattern in Table 5 is also (0 0 0 1 0). The other faults are also properly classified according to subsystem by the stage one ANN.

In the stage two ANN, the cooling coil valve subsystem faults are further analyzed to determine the specific causes of the faults. The stage two ANN uses three residual values as inputs, namely, $\overline{R}_{TS1}$, $\overline{R}_{TS2}$, and $\overline{R}_{VCC}$. By comparing Tables 2 and 6, it is seen that the five faults associated with the cooling coil subsystem are correctly identified. Similar stage two ANNs could be developed for the other subsystems.

The results presented in this section demonstrate the capability of the two-stage ANN to correctly diagnose the faulty subsystem for eleven faults in an AHU, and to further diagnose the faulty component for the five faults that occurred in the cooling coil subsystem. The study by Lee et al. (1996b) demonstrated the capability of ANNs to generalize from idealized input data to noisy lab data. The two-stage approach simplifies the generalization by reducing a single ANN that encompasses all considered faults to a number of less complex ANNs, each one dealing with a subset of the residuals and symptoms associated with a complete diagnosis of all faults. As more faults are considered, more stages could be added. In addition, this kind of architecture would make it possible to limit retraining to only select ANNs. Retraining should also require fewer computational resources because the ANNs would not be as complex as in the case of a

single ANN used for all faults. Finally, this approach is beneficial from a standpoint of understanding the diagnosis and could be utilized by a building operator to step through the reasoning behind the fault diagnosis.

Temperature sensor recovery

For fault #5, the supply air temperature sensor experiences a complete failure. Real (simulated: T_S) and estimated values ($T_{S,EV}$) of the supply air temperature and the supply air temperature sensor signal are plotted as functions of time in Figure 6 for the case where the temperature sensor is not recovered. As described previously, when the sensor fails at t = 2000 s, the sensed value of T_S becomes 0°C and the cooling coil valve closes in an attempt to make the sensed supply air temperature increase to the setpoint value of 14.5°C. This action causes the actual supply air temperature to increase to approximately 18°C, while the estimated value increases to approximately 19°C. Figure 7 shows the stage one and stage two diagnoses for fault #5. During the initial 2000 s of the simulation, the stage one ANN output indicates that the system operation is normal (the first ANN output unit for the stage one ANN is nearly equal to unity and all other outputs are nearly equal to zero). After the fault, the value of the first ANN output unit for the stage one ANN quickly decreases to nearly zero, while the output for the fourth output unit of the stage one ANN and the second output for the stage two ANN increase to values near unity, indicating that fault #5 (temperature sensor failure) has occurred. Figure 8 shows that the failed sensor can be recovered within a very short time after the fault is detected. This is accomplished by switching the control of the cooling coil valve from the sensor output to the estimated value of the supply air temperature after the fault is detected.

Fault #6 is the case in which the supply air temperature sensor drops from its supporting harness to the floor of the air duct. Simulation results for this fault are shown in Figure 9. Because the air duct surface is assumed to be at a higher temperature than the air flowing through the duct, the cooling coil valve controller responds as if a large disturbance has upset the process. In particular, the controller attempts to compensate for the fault by opening the control valve. This causes the real and estimated values of the supply air temperature to decrease; however, the

wall temperature is assumed to be constant and the sensor temperature is equal to that of the wall. Figure 10 shows that the failed sensor can be recovered using the estimated value of the supply air temperature. That is, using $T_{S,EV}$ obtained from a regression equation as the feedback signal to the cooling coil valve controller, the real value of the supply air temperature (T_S) can be recovered to a value near the setpoint.

Fault #7 is a degradation type failure of the supply air temperature sensor. Although this type of failure occurs over a period of time, it is simulated as an instantaneous fault in this study. The response of the real, estimated, and sensor values of the supply air temperature are plotted in Figure 11 as functions of time. The fault occurs at t = 2,000 s and is simulated by decreasing the sensor measurement to a value 1°C less than the actual value. The controller then closes the cooling coil valve and the sensed temperature is raised to the setpoint value, while the actual temperature is raised to a value 1°C above the setpoint. Figure 12 shows that the failed sensor can be recovered after the fault is detected.

CONCLUSIONS

The objectives of this paper were to describe an architecture for a two-stage ANN for fault diagnosis and to describe the use of regression equations for sensor recovery of failed temperature sensors. The stage one ANN was trained to classify the subsystems where faults are occurring, and the stage two ANN was trained to diagnose the cause of faults at the subsystem level. The architecture can be extended in a straightforward manner to consider additional faults such as the faults in the VAV boxes, which can be accommodated with an additional stage two ANN. It is likely that this will require the introduction of additional residuals to the analysis.

To train the ANNs, residuals of system variables were selected that could be used to quantify the dominant symptoms of fault modes of operation. Idealized steady-state patterns of these residuals were then defined for each mode of operation studied and were subsequently used for training. The trained ANNs were applied to simulation data for various faults and successfully identified each fault.

A regression equation was used to recover an estimate for the supply air temperature when

the supply air temperature sensor yields erroneous measurements. Although the agreement between the actual and predicted temperature signal during faulty operation was not perfect, the regression model was adequate for identifying fault modes of operation. It was shown that the estimates of the sensor measurement can be used for control purposes.

Future work related to this study will include implementing a method for fault detection that can classify the system operation as either normal or faulty. In addition, the residual normalization procedure will be modified to accommodate the more realistic situation in which data for all types of faults may not be available. In this case, expert knowledge or statistical methods applied to normal operational data can be used to establish threshold values of the residuals that differentiate system operation as either normal or faulty. Another topic of future work is to modify ANN training patterns to accommodate diagnoses that represent partial belief in a given fault. The current method uses training patterns that represent complete belief that the operation is either normal or that a certain fault exists. A method that provides a measure of partial belief in a fault that increases as the fault worsens would be more intuitive to a building operator than the current approach. Finally, the two-stage ANN method must be implemented in real buildings to establish its capabilities, strengths, and weaknesses.

ACKNOWLEDGMENTS

REFERENCES

Demuth, H., and M. Beale. 1992. *Neural Network Toolbox: User's Guide.* Natick, Massachusetts: The MathWorks, Inc.

Incropera, F.P. and D.P. DeWitt. 1985. *Fundamentals of Heat and Mass Transfer.* Second Edition. New York, New York: Wiley and Sons.

Lee, W.Y., C. Park, and G.E. Kelly. 1996a. Fault detection of an air-handling unit using residual and recursive parameter identification methods. *ASHRAE Transactions* 102(1).

Lee, W.Y., J.M. House, C. Park, and G.E. Kelly. 1996b. Fault diagnosis of an air-handling unit

Table 1. Normalized stage one ANN training patterns for the AHU fault diagnosis.

$|\bar{R}_{PS}|$	$|\bar{R}_{QD}|$	$|\tilde{R}_{TS1}|$	$|\bar{R}_{TM}|$	Net Outputs					Fault Diagnosis
0	0	0	0	1	0	0	0	0	Normal
1	0	0	1	0	1	0	0	0	Pressure control subsystem
0	1	0	0	0	0	1	0	0	Flow control subsystem
0	0	1	0	0	0	0	1	0	Cooling coil subsystem
0	0	0	1	0	0	0	0	1	Mixing box damper subsystem

Table 2. Normalized stage two ANN training patterns for the cooling coil subsystem fault diagnosis.

$\bar{R}_{TS1}$	$\bar{R}_{TS2}$	$\bar{R}_{VCC}$	Net Outputs							Fault Diagnosis
0	0	0	1	0	0	0	0	0	0	Normal
-1	-1	0	0	1	0	0	0	0	0	Temperature sensor failure (fault #5)
1	1	0	0	0	1	0	0	0	0	Temperature sensor failure (fault #6)
-1	0	1	0	0	0	1	0	0	0	Temperature sensor failure (fault #7)
1	0	0	0	0	0	0	1	0	0	Pump failure (fault #3)
1	1	-1	0	0	0	0	0	1	0	Control valve failure (fault #4)
-1	-1	1	0	0	0	0	0	0	1	Control valve failure (fault #4)

Table 3. Measured and normalized[†] residuals for the stage one ANN.

Fault	System Operation	R_{PS}	$\lvert \overline{R}_{PS} \rvert$	R_{QD}	$\lvert \overline{R}_{QD} \rvert$	R_{TS1}	$\lvert \overline{R}_{TS1} \rvert$	R_{TM}	$\lvert \overline{R}_{TM} \rvert$
	Normal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
#1	Supply fan failure	-0.25	1.00	**-0.40**	1.00	1.32	1.00	**1.72**	1.00
#2	Return fan failure	0.00	0.00	0.40	1.00	0.00	0.00	0.25	0.25
#3	Pump failure	0.00	0.00	0.00	0.00	**0.50**	1.00	0.00	0.00
#4	Valve failure	0.00	0.00	0.00	0.00	0.50	1.00	0.00	0.00
#5	Temperature sensor failure	0.00	0.00	0.00	0.00	-18.00	1.00	0.00	0.00
#6	Temperature sensor failure	0.00	0.00	0.00	0.00	6.50	1.00	0.00	0.00
#7	Temperature sensor failure	0.00	0.00	0.00	0.00	-1.00	1.00	0.00	0.00
#8	Pressure transducer failure	**-0.25**	1.00	0.00	0.00	0.00	0.00	-2.00	1.00
#9	Supply flow station failure	0.00	0.00	-1.20	1.00	0.00	0.00	0.00	0.00
#10	Return flow station failure	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00
#11	Recirculation damper failure	0.00	0.00	0.00	0.00	0.00	0.00	2.00	1.00

[†] For the fault cases, measurements are made 900 seconds after the occurrence of a fault. Values in bold are used for normalization.

Table 4. Measured and normalized[†] residuals for the stage two ANN.

Fault	System Operation	R_{TS1}	$\overline{R}_{TS1}$	R_{TS2}	$\overline{R}_{TS2}$	R_{VCC}	$\overline{R}_{VCC}$
#3	Pump failure	**0.50**	1.00	0.00	0.00	0.00	0.00
#4	Valve failure	0.50	1.00	**0.50**	1.00	**-0.50**	-1.00
#6	Temperature sensor failure	-18.00	-1.00	-14.50	-1.00	0.00	0.00
#7	Temperature sensor failure	6.50	1.00	5.50	1.00	0.00	0.00
#8	Temperature sensor failure	-1.00	-1.00	0.00	0.00	0.00	0.00

[†] Measurements are made 900 seconds after the occurrence of a fault. Values in bold are used for normalization.

Table 5. Classification results by the stage one ANN for the data in Table 3.

Fault	System Operation	Output Pattern				
#1	Supply fan failure	0.00	**1.00**	0.04	0.00	0.00
#2	Return fan failure	0.01	0.00	**1.00**	0.00	0.00
#3	Pump failure	0.00	0.00	0.00	**1.00**	0.00
#4	Valve failure	0.00	0.00	0.00	**1.00**	0.00
#5	Temperature sensor failure	0.00	0.00	0.00	0.00	**1.00**
#6	Temperature sensor failure	0.00	0.00	0.00	**1.00**	0.00
#7	Temperature sensor failure	0.00	0.00	0.00	**1.00**	0.00
#8	Pressure transducer failure	0.00	0.00	0.00	**1.00**	0.00
#9	Supply flow station failure	0.00	**1.00**	0.00	0.00	0.00
#10	Return flow station failure	0.00	0.00	**1.00**	0.00	0.00
#11	Recirculation damper failure	0.00	0.00	**1.00**	0.00	0.00

Table 6. Classification results by the stage two ANN for the data in Table 4.

Fault	System Operation	Output Pattern						
#5	Temperature sensor failure	0.00	**1.00**	0.00	0.00	0.00	0.00	0.00
#6	Temperature sensor failure	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00
#7	Temperature sensor failure	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00
#3	Pump failure	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00
#4	Valve failure	0.00	0.00	0.00	0.00	0.00	**0.99**	0.00

Exhaust Air Damper

Flow Station

Return Fan

Damper Actuator

Recirculation Air Damper

Flow Difference Controller

Cooling Coil

Supply Fan

Flow Station

Outdoor Air Damper

------- Control Signal
············· Sensor Signal

Temperature Controller

Pressure Controller

Figure 1. System model for a VAV AHU.

Equipment and Sensors

Cooling
Coil
— Pump (Fault #3)

— Three-way Control Valve (Fault #4)

— Temperature Sensor (Fault #5, #6, and #7)

AHU

Pressure
Control
— Supply Fan (Fault #1)

— Pressure Transducer (Fault #8)

Flow
Control
— Return Fan (Fault #2)

— Supply Fan Flow Station (Fault #9)

— Return Fan Flow Station (Fault #10)

Mixing Box
Damper
— Dampers (Fault #11)

Figure 2. Subsystem fault classification.

Residual
Values

ANN
to Classify
Subsystem

First
Stage

Subsystem
Classification

Second
Stage

ANN
to Diagnose
Subsystem #1

ANN
to Diagnose
Subsystem #2

ANN
to Diagnose
Subsystem #3

. . . . .

ANN
to Diagnose
Subsystem #n

1 2 ... m_1 1 2 ... m_2 1 2 ... m_3 1 2 ... m_n

Figure 3. Two-stage ANN for fault diagnosis of an AHU.

Figure 4. Schematic diagram of the cooling coil and cooling coil valve subsystem.

Figure 5. Chilled water pump fault.

Figure 6.   Supply air temperature sensor fault: complete failure.

-x-x- Normal condition at the stage one ANN
-*-*- Cooling coil subsystem fault at the stage two ANN
-o-o- Temperature sensor complete fault at the stage two ANN

Figure 7.   ANN output for diagnosis of complete failure of the supply air temperature sensor.

Figure 8.   Sensor recovery for complete failure of the supply air temperature sensor.

Figure 9.   Supply air temperature sensor fault: 5°C offset due to contact with duct.

Figure 10. Sensor recovery for 5°C offset of the supply air temperature sensor.

Figure 11. Supply air temperature sensor fault: 1°C offset due to degradation.

Figure 12. Sensor recovery for 1°C offset of the supply air temperature sensor.

FIGURE CAPTIONS

Figure 1. System model for a VAV AHU.

Figure 2. Subsystem fault classification.

Figure 3. Two-stage ANN for fault diagnosis of an AHU.
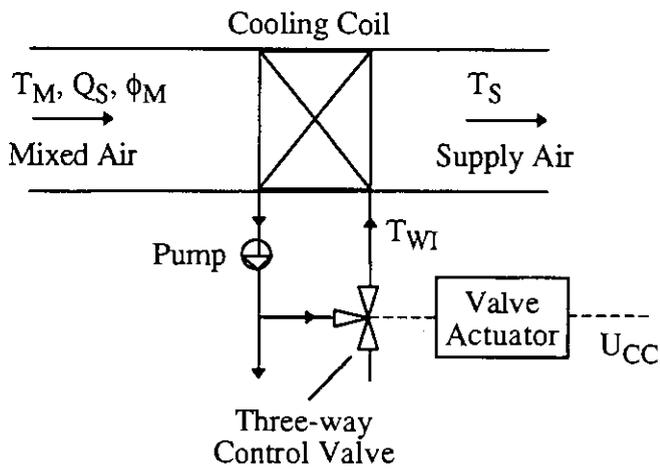
Figure 4. Schematic diagram of the cooling coil and cooling coil valve subsystem.
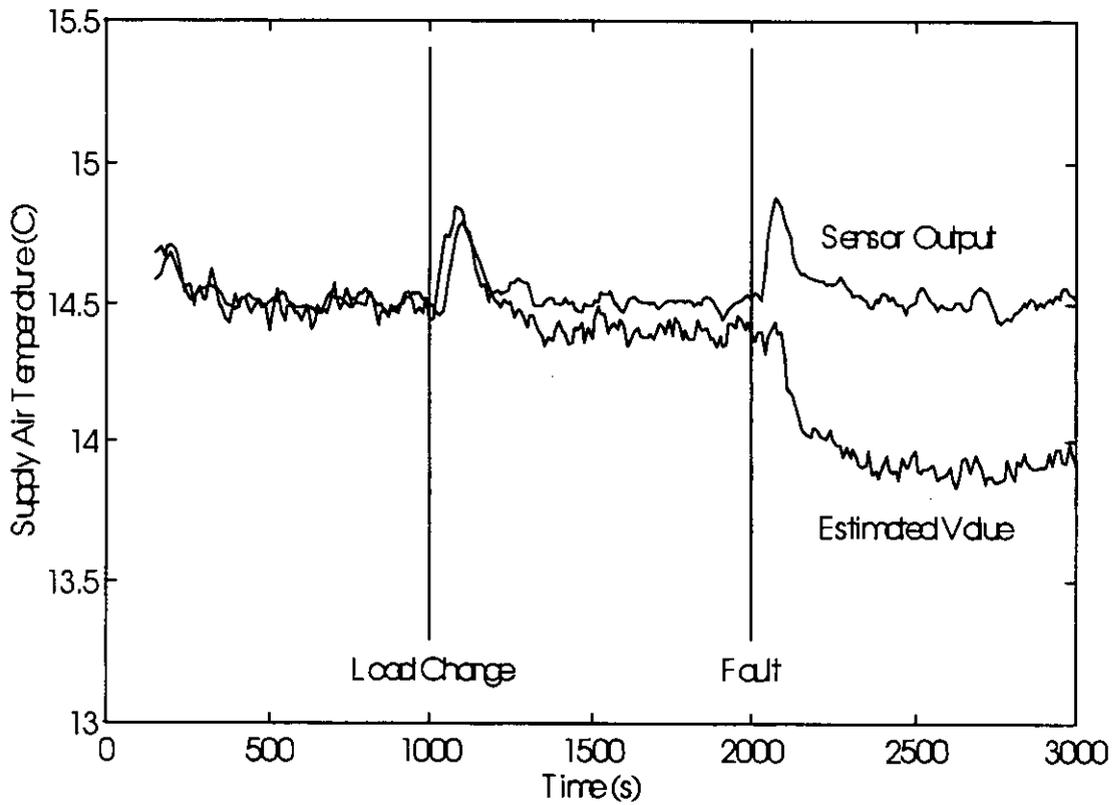
Figure 5. Chilled water pump fault.
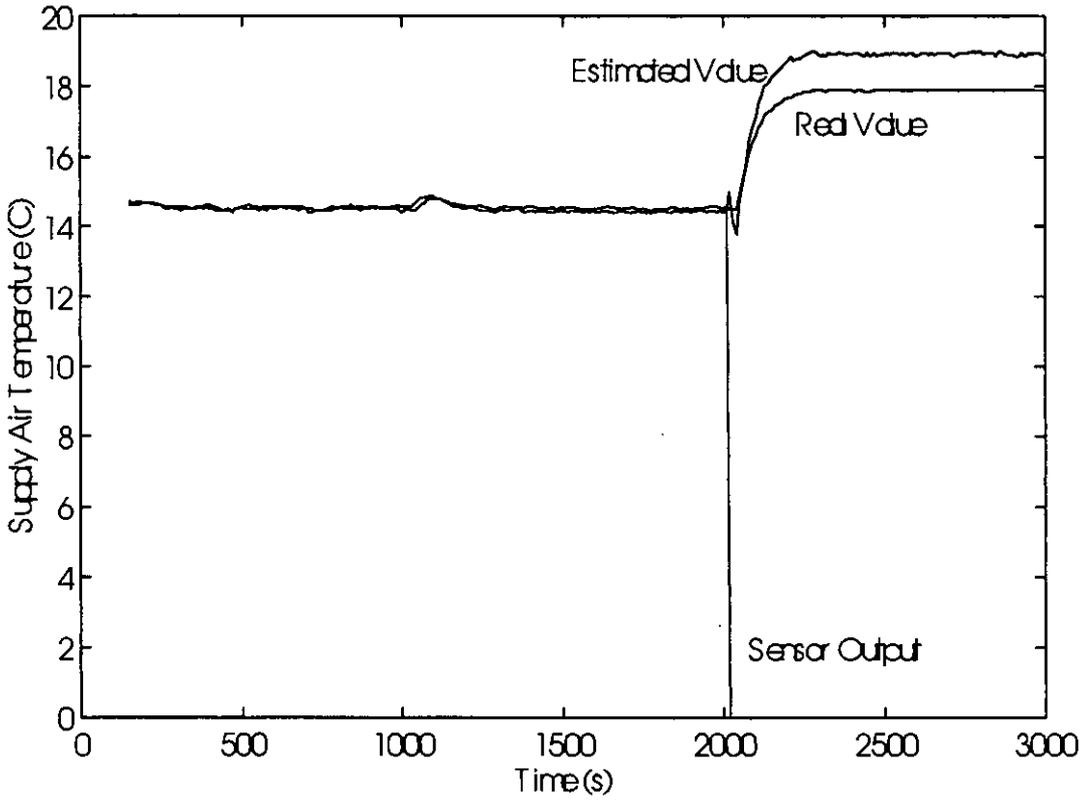
Figure 6. Supply air temperature sensor fault: complete failure.

Figure 7. ANN output for diagnosis of complete failure of the supply air temperature sensor.

Figure 8. Sensor recovery for complete failure of the supply air temperature sensor.

Figure 9. Supply air temperature sensor fault: 5°C offset due to contact with duct.

Figure 10. Sensor recovery for 5°C offset of the supply air temperature sensor.

Figure 11. Supply air temperature sensor fault: 1°C offset due to degradation.

Figure 12. Sensor recovery for 1°C offset of the supply air temperature sensor.

FAULT IDENTIFICATION IN AIR HANDLING UNITS USING PHYSICAL MODELS AND NEURAL NETWORKS

Ruxandra Dumitru and Dominique Marchio
Ecole des Mines de Paris
60, boulevard Saint-Michel
75272 Paris cedex 06

ABSTRACT

The main purpose of this paper is to *develop fault detection modules* for BEMS (Building Energy Management Systems), a software to aid building operators in detecting and diagnosing faults in HVAC systems. The fault detection modules proposed are based on two fundamentally different approaches based on component models: *physical models* and *neural networks*. These modules using these two approaches are illustrated for a cooling coil of an Air Handling Unit. The parameters to be observed, the threshold (the limit over which the fault is considered), the sampling interval and the values of energy overconsumptions are established. The modules are then tested using data bases (no fault, fault) obtained from simulations of an HVAC system.

1. INTRODUCTION & BACKGROUND

A "fault" is defined as being a failure or unacceptable change in a property of a system or one of its components. One can distinguish three major types of faults: complete failure, malfunction and degradation.

To address these three types of faults standard fault identification (FD) methodology is built on two independent foundations: fault detection and fault diagnosis. In the first stage, the presence of the fault is detected and in the second the fault is located.

More specifically in the first stage, all the components with largest percentage of faults are identified. In [1], the analysis of causes of failures in different stages of an HVAC installation (design, assembling, commissioning, plant, equipment, operation, maintenance and training of the operators) was presented. The conclusion of [1] was that 29% of the faults result during the design stage, 21% in the maintenance stage, 12% in the assembling stage; all other stages provide faults in a proportion inferior or equal to 10%.

The classification of the components according to the number of faults is made by splitting up the HVAC system into sub-systems: ventilation, mixing section, air conditioning plant, cooling and heating plant, recuperation, air distribution, control system, safety installations and the sub-systems in components. For example, the air conditioning is

further split up into: cooling coil, heating coil and humidifier. The conclusion was that 33% of failures were caused by the heating and cooling plant, 14% by safety installations, 11% by ventilation and 10% by mixing section, all other systems being inferior to 10%. Carrying on this analysis down to the component level, the results are as follows: 8% of faults are provided by the humidifier, followed by fans and chiller with 7% and finally the water system with 5%.

With the above classification the next problem is the choice of a fault identification scheme. The criterion for a particular FD scheme in BEMS should consider both robustness with respect to changes in the components and also generalize easily to other components and faults. In particular, the main issue is modeling the components. For example, in [12], fault identification using physical models supposes the description of the phenomena by heat and mass balance equations.

In an alternative approach, [3] *ANN (Artificial neural networks* are used for components for which detailed models are not available or are not easy to explain from a physical point of view.

1.1 PROBLEM STATEMENT

In this paper, two fault detection modules based on physical models and a ANN are developed and compared for an Air Handling System.

Comparison between the two methods of fault detection is made for two components of the Air Handling Unit: the mixing section and the cooling coil. In this paper, the methodology is illustrated only for the cooling coil.

2. SYSTEM DESCRIPTION

The Air Handling System (single zone fan system) is illustrated in *Fig. 1*. In its most basic configuration, the Air Handling System provides constant volume, forced air heating and cooling for a single zone. (containing a heating coil, a cooling coil, filters (not shown), and a supply fan.) Normal operation assumes a mixed flow during the occupation period, between 7 a.m. and 8 p.m.
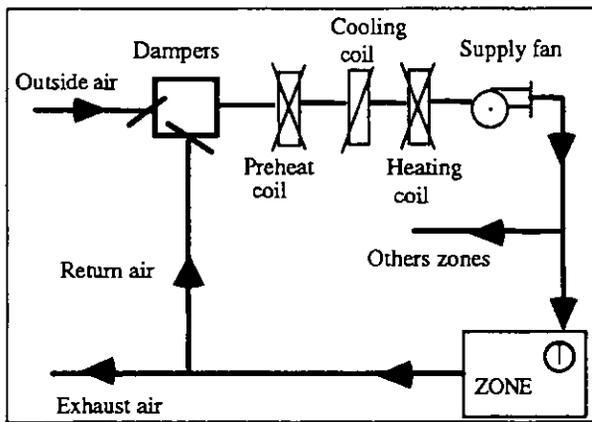
Figure 1. Schematic of Air Handling System.

It is further assumed that the quantity of outside air varies from "minimum required hygienic air", passing through "optimal mixing" to "all fresh air". The damper adjustment is driven by indoor temperature; its upper limit position is all-outside-air and its lower limit position is a mix of returned air with minimum required hygienic air.

The control of the cooling coil keeps the outlet temperature of the air constant for the same outdoor conditions.

3. FAULTS CONSIDERED

For the cooling coil, we considered two types of faults: **fouling on coil fins** and **scale in coil tubes.** The first fault leads to a reduction in overall heat transfer coefficient on the air side and the second fault to a reduction in overall heat transfer coefficient on the water side.

The cooling capacity decreases due to the fact that the heat transfer become gradually worse in both cases. As a result, the characteristic of the aeraulic system changes, hence changing the operation point. The principal change is in the fan, in the case of 'fouling on coil fins' or the pump, in the case of 'scale in coil tubes', will operate at part load ratio. Hence, on the air side the fan's electrical consumption increases or, on the water side, the pump's electrical consumption increases as does the total energy consumption in one case or in another.

4. FAULT IDENTIFICATION

Both fault identification modules use the same measured values: the inlet air conditions (mass flow rate, temperature, and humidity ratio), the inlet water conditions (mass flow rate, temperature) and the outlet conditions of air and water. However some differences do exists. For example, the fault identification method using neural networks uses the cooling load, calculated on water side, therefore a combination of water parameters; meanwhile the fault identification method using physical models uses the water parameters one by one

(independently). Another is the fault identification scheme using physical models needs information about fan and pump energy consumption.

In the case of ANN, it is necessary to condition the data first [11]. The results of usual sensor measurements in a AHU, which includes flow rates, temperatures, humidity ratios, pressures, can vary significantly in magnitude and therefore must be appropriately conditioned. Consequently, normalized deviation values were used for the parameters. Calculation of the normalized deviation of a variable consists of taking the difference of the actual value and the average value of the variable, then dividing it by the average value. This data conditioning method produces the desired effect of scaling all the process variables within a similar range, maintaining the qualitative relationship between them.

5. FAULT IDENTIFICATION USING PHYSICAL MODELS

5.1. Description of the method

The *fault identification* using *physical models* relies on the comparison between the predicted values of different variables and the measured values of the same variables. The predicted values are the outputs of a physical model using measured values. The fault identification uses "if-then rules" and information concerning the other components called "complementary information" (energy overconsumption for example: see Section 5.4). This kind of information is needed during the diagnosis stage. When the deviation between predicted and measured values (temperature, humidity ratio) is larger than a certain threshold, the presence of a fault is declared. For the diagnosis, this kind of information is too poor the complementary information is then used.

5.2. Assumptions

The basic assumption is that the coil is operating under steady state conditions i.e. the coil dynamics are much faster that the other system components. Statistical methods [6], [5], [13], [14] based on statistics tests of mean, variance and slope [4], [7] were used in order to establish the time intervals when the coil is operating under steady state conditions.

The algorithm calculates the outlet water temperature, air dry bulb temperature and humidity ratio for a coil with a completely wet fin surface. The coil is considered to be operating under "all wet" surface conditions if the surface temperature at the air inlet is lower than the inlet air dew-point temperature. Under wet conditions, the steady state air and water conditions can be determined using

standard heat exchanger effectiveness relationships based on enthalpies rather than temperatures.

The resistances to enthalpy transfer are computed in terms of conventional heat transfer coefficients, fin efficiencies, and "fictious enthalpy" parameters as described in Threlkeld [8].

5.3 Statistical method to determine steady - state

This statistical method implies three steps. First, each variable was considered separately with the null hypothesis that there was no difference in the variances from one period to the next. This was tested by the Fisher test on the ratio of the estimates of the variances from two successive periods. In cases where there was no significant difference between successive variances, the variances were pooled.

The Student test was then used to compare means of each measurement in successive periods, to find periods of apparent steady-state. Clearly it is sufficient in order to establish an unsteady-state to find a significant change in the mean of any measurement from one period to the next. To establish steady state, it is necessary but not sufficient to find a sequence of periods with no significant changes in the means of all measurements, since a trend could occur with small enough changes from one period to the next to escape statistical detection.

For each pair of time periods in succession, the means were tested against the null hypothesis that they are equal. If the null hypothesis could not be rejected, that pair was recorded as forming part of a sequence of periods in apparent steady state. As soon as a pair of means was found which had changed significantly-, if the earlier of the pair of periods already formed part of a sequence in apparent steady state, it marked the end of that sequence. Otherwise, the earlier period was flagged as unsteady sate and in either case, the next pair of periods was examined. Any sequence of two or more successive time periods which remained unflagged were then considered to be in apparent steady state.

For each sequence of three or more periods of apparent steady state, we found the slope of the means of each measurement versus time from the first period to each later period. We then tested the null hypothesis that each slope was zero by a Student test. If any test failed, the sequence was shortened by excluding the subset of periods with significant slope. The sequence was either reduced to zero duration, or there remained a sub-sequence which was deemed to be in steady state. The slope test was intended to prevent slow monotonic trends and cyclic patterns from being taken as steady states.

5.4. Algorithm

The inputs and outputs for the coil model are schematically represented in Fig. 2.
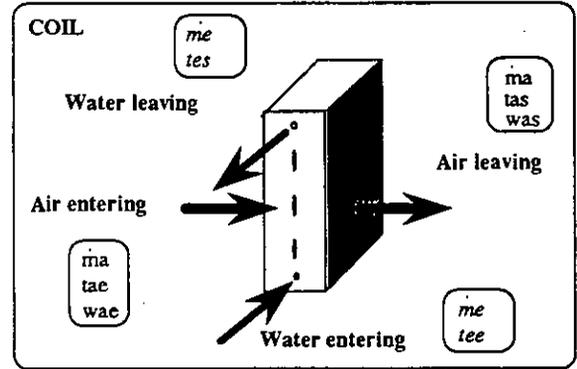
Figure 2. Coil model

Enthalpy based heat-transfer calculations for a wet surface use the fundamental relationship between heat transfer, enthalpy and capacity.

$$q_a = C_a \cdot (h_{ae} - h_{as})$$
$$q_e = C_e \cdot (h_{es,sat} - h_{ee,sat}) = \dot{m}_e c_e \cdot (t_{es} - t_{ee}) \quad (1)$$

where the capacity rate of the two fluid streams are:
$$C_a = \dot{m}_a$$

$$C_e = \dot{m}_e \cdot \frac{c_e}{c_{p,sat}} \quad (2)$$

$c_{p,sat}$ is calculate using the specific heat of dry air, the specific heat of water vapour and the humidity ratio of saturated air:

$$c_{p,sat} = c_a + w_{sat} \cdot c_v. \quad (3)$$

The outlet air enthalpy and the enthalpy of saturated air at the water temperature can be determined by modeling the coil as a counterflow heat exchanger [12]. However, since the heat transfer calculations are performed based on enthalpies, the overall heat transfer coefficient is based on enthalpy potential rather than temperature potential. Under this assumption, a wet coil , local heat transfer and the corresponding overall heat transfer coefficient between the air and water are calculated by the following:

$$q = KS_h \cdot (h_a - h_{e,sat}) \quad (5)$$

The fault identification model compares the predicted values of the outlet variables, calculated with the algorithm that had been presented, with the measured values of the same variables. The measured values are noted with an apostophy (') in the following identification scheme. Complementary information concerning the energy consumption (i.e. overconsumption of the circulation pump and overconsumption of the fan) is considered, as illustrated in Fig. 3.
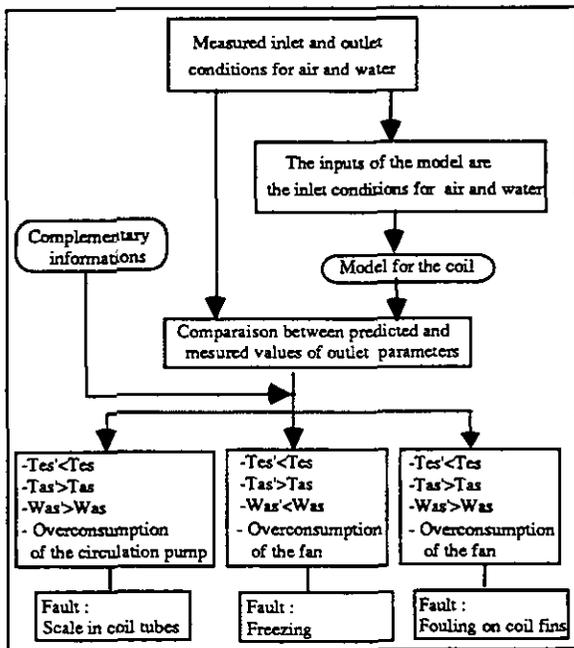
Figure 3. Fault identification model.

An example of one of the rules used in the above scheme is: the measured value of the air temperature exiting the coil, Tas is less than the expected value, Tas'(calculated with the algorithm). The explanation of this rule is that the considered faults (fouling, scaling or freezing) are sufficient to prevent the set-point being attained.

5.5. Application

This application of fault identification shows what kind of parameters are considered for fault detection. For example, outlet dry bulb temperature, the air's humidity ratio and cooling capacity were chosen to illustrate the consequences of each simulated fault. Hourly values of these parameters for a summer month (July), more precisely on 28th of July when the outside air temperature takes its maximal value 31°C at 4 p.m. are analyzed.

In the following examples, measurements for correct operation of the plant, called "reference values" (normal behavior), are compared to measurements for a non-optimal operation (failure). The last kind of operation is due to the appearance of the fault.

Both curves are obtained by simulation with DOE2. The DOE2 cooling coil model used is a bypass factor model.

The first simulated fault of the coil is the **scale in the tubes**. The effect of this fault is a decrease in the overall heat transfer coefficient between air and water, on the water side.

The reason is that scale in the tubes introduces a supplementary thermal resistance. The decrease in overall heat transfer coefficient leads to a decrease

in heat transfer. Therefore on the water side the mass flow rate decreases resulting in less heat absorbed by the water. On the air side the mass flow rate remains constant but the difference of enthalpy decreases increasing the outlet air temperature and humidity ratio (*Fig. 4*).
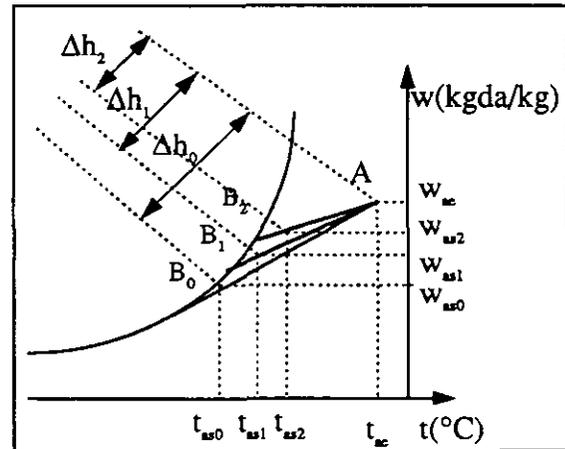
Figure 4: Influence of the overall heat transfer coefficient decrement.

In *Fig. 4*, the line AB_0 represents the cooling provided by the coil before scaling. In this case the difference of enthalpy is Δh_0. After the introduction of scaling, the quantity of cooling becomes AB_1 with a corresponding difference of enthalpy Δh_1. Consequently the outlet temperature and humidity ratio increase as does the surface temperature of the coil. Meanwhile as the scaling increases (line AB_2) the difference of enthalpy on air side, Δh_2 decreases which yields to a decreasing of the outlet air temperature and surface temperature of the coil.

The example in *Fig. 5* shows the variation of the outlet air humidity ratio in this case. Note that, generally, the measurements of humidity may be reliable within about ± 3% relative humidity for ordinary rooms temperatures under equilibrium conditions [8].
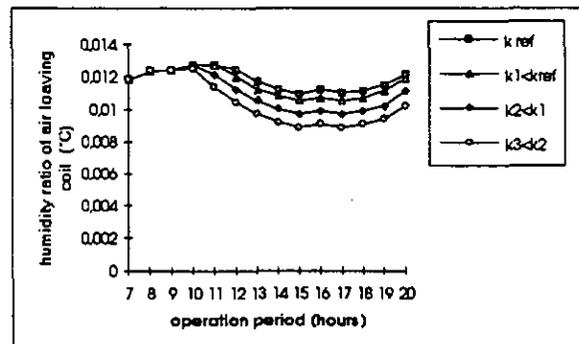
Figure 5. Scale in coil tubes -evolution of the humidity ratio of air leaving the coil.

The results concerning the observed parameters, thresholds, the sampling interval and values of energetic overconsumptions are given in *Table 1*.

644

The cooling capacity decreases and the energy consumption increases by 2% in the case of reduction in heat transfer coefficient of 6%, 5% for 16% reduction , 7% for 26%.

The second simulated fault is the **fouling on coil fins**. The heat transfer between air and water diminishes due to the decreasing of overall heat transfer coefficient on air side.

The parameters considered for illustrating the effect of the fault are: the outlet temperature and humidity ratio , the cooling coil capacity and the fan electrical consumption.

The hourly variations of the leaving air temperature is illustrated in the *Figure 6*.
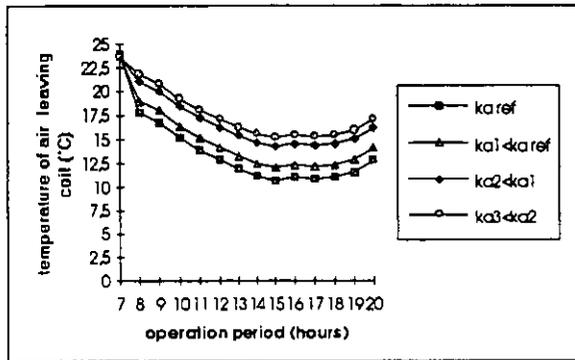
Figure 6. Fouling on coil fins - evolution of the air temperature leaving the coil.

A reduction in the overall heat transfer coefficient on air side of 20% (curve ka1) leads to a difference of temperature of 0.8°C between measured and reference values, which can not be detected. Meanwhile a reduction of 53% (curve ka2) of the overall heat transfer coefficient on air side determines a difference of temperature of 2.7°C, which is now measurable. A decrease of the overall heat transfer coefficient on air side of 66% induces a variation of temperature of 4°C (curve ka3). The temperature variation can be detected beginning by a reduction in the overall heat transfer coefficient on air side of 25%.

The fouling determines also a modification of aeraulic characteristic of the system, the operation point changes and the fan operates at part load ratio leading to a waste of energy at the fan level.

The results concerning the observed parameters, thresholds, the sampling interval and values of energy overconsumption are also given in *Table 1*.

FAULT	Effect of the fault		Observed parameters 3)	Threshold 5)	Sampling interval	Energy consumption 4) (in summer)
scale in coil tubes 1)	decrease of overall heat transfer co water side	6%	humidity ratio of air	no	1 hour	+2%
		16%	leaving cooling coil	yes	1 hour	+5%
		26%	(kg/kg dry air)	yes	1 hour	+7%
		6%	coil surface	no	1 hour	
		16%	temperature	yes C	1 hour	
		26%	(°C)	yes C	1 hour	
			cooling capacity (W)	yes	1 hour	
fouling on coil fins 2)	decrease of overall heat transfer on air side	20%	humidity ratio of air	no	1 hour	+2%
		53%	leaving cooling coil	yes	1 hour	+7%
		66%	(kg/kg dry air)	yes	1 hour	+10%
		20%	temperature of air	no	1 hour	
		53%	leaving cooling coil	yes	1 hour	
		66%	(°C)	yes	1 hour	
			cooling capacity (W)	yes	1 hour	

Table 1.

Obs: 1) The scale in coil tubes is a fault that can be detected starting by a diminution of 10% of the overall heat transfer coefficient on water side.

2) The fouling on coil fins is a fault that can be detected starting by a reduction of 25% of the overall heat transfer coefficient on air side.

3) Note that the number of observed parameters is larger than proposed in the identification scheme. Some of them can not be measured in practice. However, the fact that the data are provided by simulation allowed us to observe their evolution.

4) The values of the energy consumption are given only for the summer period. An augmentation of 10% in summer does not imply the same thing in winter.

5) It is considered that the temperature threshold is 1°C. Only the values superior to this threshold are taking into account for fault identification.

6. FAULT IDENTIFICATION USING ANN MODEL

6.1. Description of the method

The proposed fault identification methodology using ANN, supposes two stages [3], [9], [10].

First, the *ANN learns the normal operation of the equipment*. In this case, the ANN is trained and tested with data bases containing values of various parameters in the absence of fault. Note that this stage supposes two different operations: the learning and the testing of the network. Both are made with data bases representing the behavior without fault of the system, but not with the same. For example, in the coil case, the training is done with hourly values for the months of May and July and the testing with hourly values for the others months (August or September).

Second, the *ANN is tested with an actual measured data base*. The *fault detection* is based on the comparison between the outputs of the ANN for the normal operation with the outputs of the ANN for the actual measured data base. If the two sets of

outputs are different, we conclude that there is a fault.

After each test of the ANN, a representation of the predicted output versus real value is plotted. When the equipment is operating normally, the values from the testing data base are distributed in the proximity of the first bisectrix. When fault is present, the distribution of the testing data base values presents typical deviation from the first bisectrix (Fig. 7).
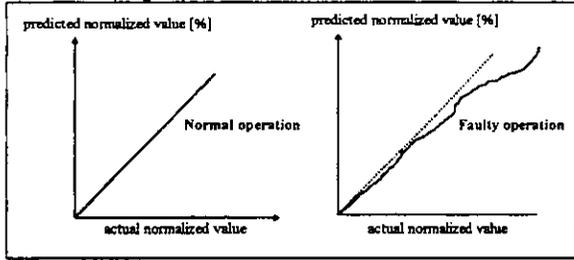
Figure 7. Representations of a predicted value versus a real value of an parameter.

6.2. Assumptions

The training of the ANN raises the problem of availability of data bases [9], [15]. In this approach, a minimum of 744 data pairs is needed for training. It is also important that these data corresponds to all possible situations of system operation. In the case of the cooling coil, the data used for training contain the hourly values of a month of mild weather. Suppose that only the month of August was chosen; in this case, the cooling coil operates in the most cases at maximum load. If the ANN is tested with the data bases for September, we note that the training was not efficient, because these data contains more operation situations than the training data.

6.3. Structure of the ANN

The ANN used for the fault detection in cooling coil has the following structure: 4 inputs, 3 outputs and 2 hidden layers containing 7, respectively 5 neurons (*Figure 8*). The inputs are the inlet air temperature, humidity ratio, mass flow rate and the cooling load. The outputs are the outlet air temperature and humidity ratio.
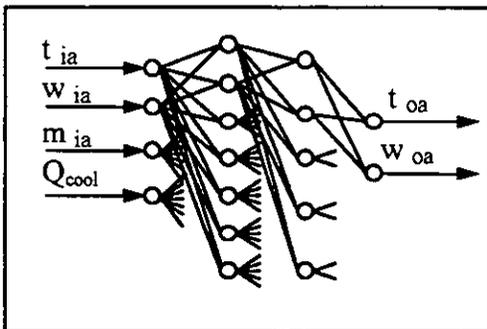
Figure 8. ANN for cooling coil.

Our generalized delta rule with backpropagation was used with a learning rate ε, of 0.9 and a momentum, μ of 0.1. A convergence criteria of 0.1, defined as the allowable difference between the actual output of a neuron in the output layer and the desired or target output, was used to determine the extent of further network training required.

6.4 Application

First, the ANN was trained under normal operation. For an efficient training, the data base must include the largest possible number of operating situations (the above mentioned months were chosen according to this purpose). The training data base contained hourly values for two months: May and July. The ANN was then tested using a data base for the month of August.

Next the ANN was tested using data bases representing faults: scale in coil tubes and fouling on coil fins, assuming various degrees of fouling and scaling. It may be seen that the deviation of the outputs with respect to the first bisectrix increases (*Fig. 9*).

646
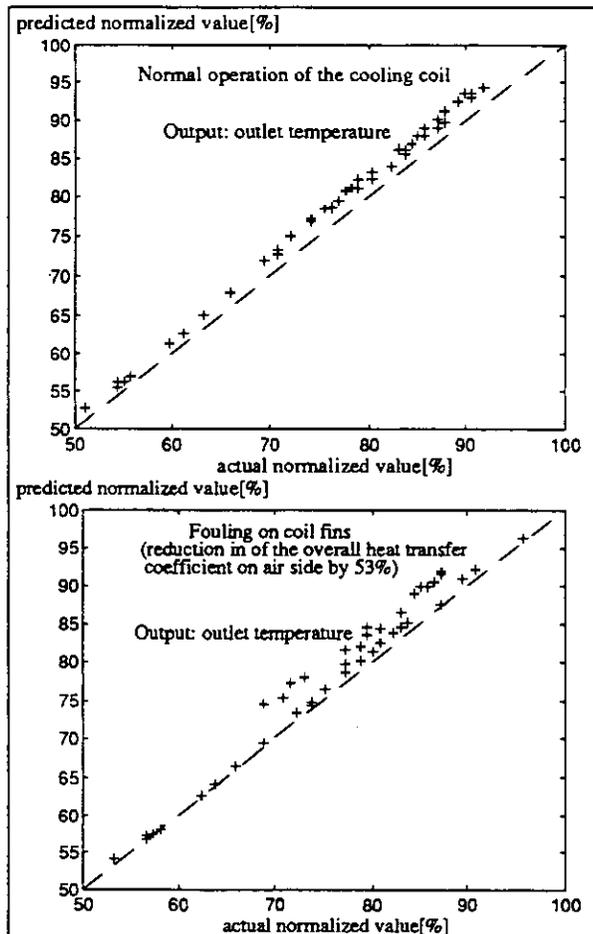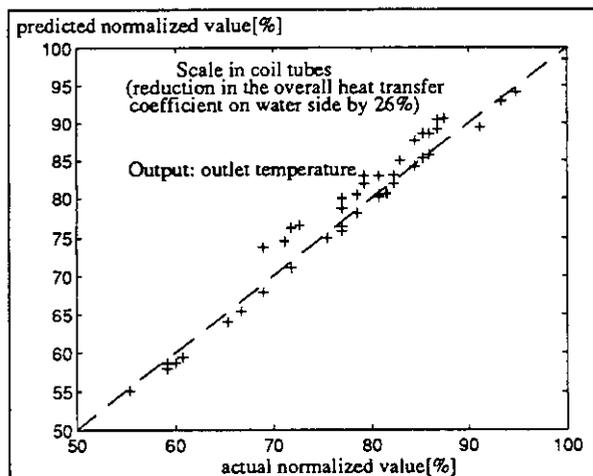
Figure 9. Increase of outputs deviation with respect to the first bisectrix due to the fault.

Fig. 9 is an example of the resulting graphical representations after the testing. The first graph represents distribution's of air temperature leaving the coil during the normal operation. In this graph the data base used for testing was the training data base. The distribution of the parameter is close to the first bisectrix (dashed line). The second graph represents the distribution of the same variable when fouling occurs determining a reduction in 53% of the overall heat transfer coefficient. Note that the distribution is no longer close to the bisectrix. The third graph represents the outlet air temperature distribution's when scaling occurs and a reduction in 26% of the overall heat transfer coefficient on water side is considered.

The relation between the evolution of the outlet parameters and the fault aggravation is given in the *Table 2*.

Faults	Outputs (outlet air conditions)	RMS error (%)	Average error (%)	Maximum error (%)
Fouling on coil fins k air decrease of 20%	temperature (°C)	1.8	-1.0	-6.0
	humidity ratio (g(dry air)/kg)	1.8	1.1	6.4
Fouling on coil fins k air decrease of 53%	temperature (°C)	2.1	-1.3	-6.5
	humidity ratio (g(dry air)/kg)	1.6	0.8	5.9
Fouling on coil fins k air decrease of 66%	temperature (°C)	2.1	-1.3	-6.5
	humidity ratio (g(dry air)/kg)	1.6	0.8	5.9
Scale in coil tubes k water decrease of 6%	temperature (°C)	1.8	-1.0	-6.0
	humidity ratio (g(dry air)/kg)	1.8	1.1	6.4
Scale in coil tubes k water decrease of 16%	temperature (°C)	1.3	-0.5	6.3
	humidity ratio (g(dry air)/kg)	1.5	0.9	5.4
Scale in coil tubes k water decrease of 26%	temperature (°C)	1.5	-0.6	6.3
	humidity ratio (g(dry air)/kg)	2.2	1.4	7.4

Table 2.

CONCLUSIONS

One of the limits of using a physical model for fault identification in the coil case is that detailed information about the coil geometry is needed (heat transfer coefficient). For the neural networks data bases are needed.

Our results for fault identification with physical models, show that 'scale in coil tubes' fault can be detected with a reduction in 10% of the overall heat transfer coefficient on the water side and the 'fouling on coil fins' for a reduction in 25% of the overall heat transfer coefficient on the air side.

Values of the energy consumption are given for the summer period for different degree of fault (*table 1*) with the precaution that an augmentation of 10% in summer does not imply the same thing in winter.

A fault detection method based on neural networks has been developed.

The training time consists of 400 000 iterations in the coil case. The time is given in terms of iteration rather than specific durations, as actual physical training times are dependent on the speed of the computer used for training.

These two kinds of approaches have been tested using simulated data. The simulations are provided with DOE2 [2]. There are two kinds of data: data representing the normal operation of the system and data representing faulty operation of the system. The simulation of faults is introduced in DOE2 input files modifying the suitable parameters. Long term simulation data (yearly hourly values) and short running time are some of the advantages of choosing DOE2.

FUTURE WORK

A method of diagnosis using neural networks is under study. The training in this case uses data bases representing faulty operation of the system. A fault type is assigned to a particular output neuron which value is use to determine fault magnitudes.

If data base provided by simulation tools are used, the simulation period for a certain value of overall heat transfer coefficient had to be increased, because in fact the fouling on the coil and the scale in tubes are long-lasting phenomena. Until now, the simulated data were "non-continuous" because the data bases were created for different degrees of the fault. A future issue will be to take into account data describing the time evolution of the fault.

ACKNOWLEDGMENTS

This research takes place also within a common international frame[1] of Annex 25 of IEA about fault identification in AHU systems.

This paper was first published in IBPSA proceedings of "Buiding Simulation '95".

REFERENCES

[1] SCHIEL H.F., "Questionnaire modules of ventilation & air conditioning plants", Annex 25, - *Building optimisation and fault detection*, mars 1992 - Helsinki

[2] DOE2, "Engineers Manual", Lawrence Berkely Laboratory, University of California, 1981

[3] CURTISS S.P. "Artificial Neural Networks for use in building systems Control and Energy Management", thesis University of Colorado, 1992

[4] CROWE C.M, HOLLY W., COOK R. (1989), "Reconciliation and rectification of process flow and inventory data", Canadian Journal of Chemical Engineering, vol.67, pp. 595-601, 1989

[5] NEUILLY M., CETAMA, "Modélisation et estimation des erreurs de mesure", Lavoisier, 1993

[6] RAGOT J., DAROUACH M., MAQUIN D., BLOCH G., "Validation de données et diagnostic", Editions Hermes, 1990

[7] MEYER M., "Validation de données sur des systèmes incompletement observés", thèse de doctorat de l'Institut National Polytechnique de Toulouse, juillet 1990

[8] THRELKELD J.L., "Thermal Environmental Engineering", 2nd Edition, Englewood Cliffs: Prentice Hall, 1970

[9] HOSKINS J.C., HIMMELBLAU D.M., "Artificial Neural Networks Models of knowledge representation in chemical engineering", Computers & Chemical Engineering, volume 12, no.10 ,1988 p.881-890

[10] KOSHIJIMA I., NIIDA K., "Neural Networks Approach to Fault Detection under steady state conditions", IFAC Symp. On - line Fault Detection and Supervision in the Chemical Process Iμndustries, 1992, Newark,DE

[11] BECRAFT W.R., LEE P.L., "An Integrated Neural Network/Expert system approach for fault diagnosis", Computers & Chemical Engineering, vol.17, no.10, pp. 1001-1014, 1993

[12] HVAC2 Toolkit, Algorithms and Subroutines for Secondary HVAC System Energy Calculations, ASHRAE, 1791 Tullie Circle N.E., Atlanta, 1993

[13] SAGE A. P., MELSA J.L., "Estimation Theory with Application to Communications and Control", New York:McGraw Hill, 1970

[14] VAN TRESS H.L., "Detection, Estimation and Modulation Theory", Part I and Part III, New York:John Wiley & Sons, 1971

[15] Yan H. H., Fischl R. and Chow J. C., "Indices for evaluation of neural network performance in power system security assessment," ISAP, Monttpellier, France, Sept. 5-9, 1994.

NOMECLATURE

- C_a = air-side capacity rate (kg/s)
- C_e = water-side capacity rate (kg/s)
- h_{ae} = entering air enthalpy (J/kg)
- h_{as} = leaving air enthalpy (J/kg)
- $\dot{m}_e$ = entering water mass flow rate (kg/s)
- $\dot{m}_a$ = dry air mass flow rate (kg(da)/s)
- t_{ae} = entering dry bulb temperature (°C)
- t_{as} = leaving air temperature (°C)
- W_{ae} = entering air humidity ratio (kg/kg da)
- W_{as} = leaving air humidity ratio (kg/kg as)
- t_{ee} = entering water temperature (°C)
- t_{es} = leaving water temperature (°C)
- c_a = specific heat of air = 1006 (J/kg°C)
- c_v = specific heat of water vapor = 1830 (J/kg°)C
- c_e = specific heat of water = 4186 (J/kg°C)
- KS_h = overall enthalpy heat transfer coefficient (kg/s)
- "e,sat"= "fictious enthalpy" (enthalpy of saturated air evaluated at the liquid chilled water temperature)

---

[1]Annex 25 "Real simulations of HVAC Systems for building optimisation, fault detection and diagnosis" is a part of the programme "Energy Conservation in Buildings and Community Systems" of the IEA, International Energy Agency.

ON-LINE FAULT DETECTION AND DIAGNOSIS USING FUZZY MODELS

M. Benouarets and A. L. Dexter

University of Oxford
Department of Engineering Science
Parks Road, Oxford OX1 3PJ, U.K.

Abstract: The paper describes a new fuzzy model-based method of fault detection and diagnosis. The method uses the principle of fuzzy matching to compare the rules of a fuzzy model identified on-line describing the actual operation of the system to each of the rules of a set of reference models which describe the fault-free and faulty operation of the system. The method accounts for any ambiguity, which may result from fault-free and faulty operation, or from different faults, having similar symptoms at some operating points. Results are presented which demonstrate the use of the scheme to detect and identify faults in a heating system in a large building.

Keywords: fault diagnosis, fuzzy logic, fuzzy models.

1. INTRODUCTION

Many of the existing fault detection and diagnosis (FDD) schemes use mathematical models to describe the behaviour of the plant. These models are used to predict the behaviour of the plant if it were in a number of possible operating states. The predictions are then compared with measurements of the plant's actual behaviour to produce residuals. However one problem associated with this approach is that, in practice, it is almost impossible to obtain a model that exactly matches the process behaviour. Mismatch between the model and the plant will lead to non-zero residuals which can cause false alarms unless appropriate thresholds are used. Also when the system under consideration is nonlinear, selection of appropriate thresholds may be quite difficult since the accuracy of the model will depend on the operating conditions and the value of the threshold must vary with operating point (Sauter, et al., 1994, Schneider and Frank, 1994).

Fuzzy systems or more specifically fuzzy models can be used to take account of the uncertainties

and imprecision associated with modelling the behaviour of a complex and ill-defined system (Sugeno and Yasukawa, 1993). A fuzzy model is a representation of the essential features of a system. It is based on the concept of fuzzy partitions of the information and operates with fuzzy sets instead of numbers. Fuzzy models can easily include whatever expert knowledge is available about the symptoms of faulty and fault-free operation of the system, and there has been increasing interest in their use in FDD schemes (Vachkov and Matsuyama, 1992; Terpstra, et al., 1992; Linden, et al., 1994; Maruyama et al., 1995).

In this paper, a fuzzy method of diagnosis which can be applied to ill-defined systems is presented. This method uses explicit fuzzy models, generated from training data, to describe the symptoms of both faulty and fault-free operation. The fuzzy models are produced by combining information extracted from measured data with whatever expert knowledge is available about the system. A fuzzy matching scheme is used to compare the parameters of an on-line fuzzy model generated from normal operating data with those of the

fuzzy reference models.

The organisation of the paper is as follows. Section 2 describes the fuzzy fault diagnosis scheme, introduces the concept of fuzzy matching and explains how it can be used to evaluate the degree of similarity of two fuzzy models. A technique for evaluating the ambiguity, which occurs when fault-free and faulty operation, or different faults, exhibit similar symptoms at some operating points, is given. A method of using the similarity and ambiguity measures to generate the strength of the evidence associated with each of the different operating states is also described. In section 3, results are presented which demonstrate the use of the fuzzy FDD scheme to detect and identify faults in the heating system of a large building. Some conclusions are given in the last section.

2. FUZZY FAULT DIAGNOSIS

The reference models used in the proposed FDD scheme (see Figure 1) are identified using expert knowledge and training data generated from a simulated plant. Expert knowledge is used to determine the structure of the fuzzy models, or, in the case where there is no expert knowledge and only input-output training data are available, the structure of the relationships between the variables, the shape of the membership functions and the number of the rules can be determined by clustering the data. Generic reference models which describe the underlying behaviour of a class of plants of similar design can also be used if it is impossible to obtain detailed information about the particular plant under test (Dexter and Benouarets, 1995a). One of the fuzzy reference models describes the fault-free operation of the system, the others describe the behaviour of the system in the presence of specific faults.

Each fuzzy reference model is a qualitative description of the relationship between input-output variables in the form of IF-THEN rules. To generate the fuzzy models, the elements of an associated relational matrix are estimated using a simple fuzzy identification scheme (Xu and Lu, 1987). Each entry in the relational matrix is a measure of the credibility (confidence) that the corresponding rule appropriately describes the behaviour of the system at a particular operating point.

The operating data are collected on-line from the actual plant, filtered using a moving-average filter and, if static fuzzy reference models are used, passed through an activity detector to locate portions of data where the system is in or near to steady-state (Dexter and Benouarets, 1995b). The data are then used to identify an on-line

fuzzy model which describes the current behaviour of the system. The rules of the identified on-line fuzzy model are compared to those of the fuzzy reference models using fuzzy matching. The reference models are also compared to each other to account for any ambiguity which may result from fault-free and faulty operation, or different faults, having similar symptoms at some operating points.
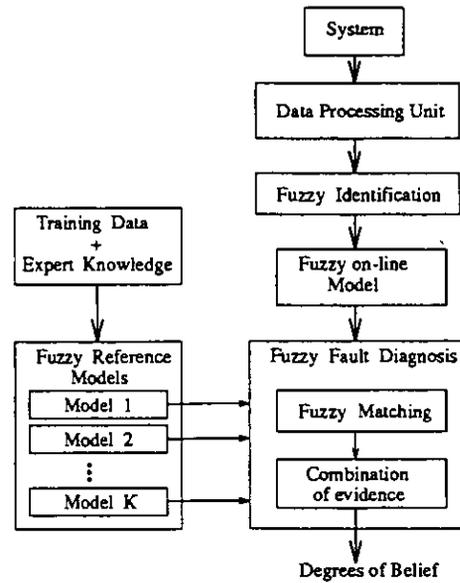
Figure 1: The fault detection and diagnosis scheme

2.1 Fuzzy matching and levels of ambiguity.

Lee et al., (1992) define the degree of matching $S(A, B)$ between two discrete membership functions A and B (see Figure 2) as the proportion of A that is contained in B where

$$S(A, B) = \frac{\sum_{-\infty}^{\infty} \mu_{A \cap B}(x)}{\sum_{-\infty}^{\infty} \mu_A(x)} \qquad (1)$$

and $\cap$ can be represented in fuzzy logic by the $'min'$ operator.
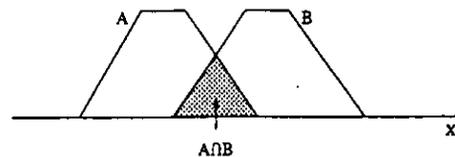
Figure 2: Fuzzy matching of two fuzzy sets

The degree of similarity of fuzzy models can also be evaluated using an expression similar to Equation 1 if the fuzzy models are regarded as level-2 fuzzy sets (Sosnowski and Pedrycz, 1992) with discrete membership functions given by the credibilities of their rules (see Figure 3). Thus, a measure of the degree of similarity between two fuzzy models M_i and M_j is expressed by the

following equation

$$S(M_i, M_j) = \frac{\sum_{n=1}^{N} min[c_{M_i}(n),\ c_{M_j}(n)]}{\sum_{n=1}^{N} c_{M_i}(n)} \quad (2)$$

where $c_{M_i}(n)$ and $c_{M_j}(n)$ are the credibilities of the nth rule in the fuzzy models M_i and M_j respectively, and N is the number of rules compared.
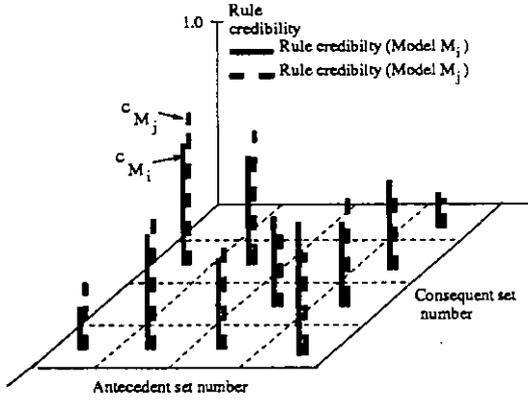
Figure 3: Fuzzy matching of two fuzzy models

The value $S(M_i, M_j)$ may be regarded as a measure of the extent to which the symptoms of the behaviour in the operating state represented by the fuzzy model M_i are similar to those of behaviour in the operating state represented by the fuzzy model M_j. Equation 2 can be used to evaluate the degrees of similarity between the on-line fuzzy model and each of the reference fuzzy models. Thus,

$$S(M_o, M_i) = \frac{\sum_{n=1}^{N} min[c_{M_o}(n),\ c_{M_i}(n)]}{\sum_{n=1}^{N} c_{M_o}(n)} \quad (3)$$

for $i = 1, 2, ..., K$ where M_o denotes the on-line fuzzy model and K is the total number of reference models (see Figure 1). The value $S(M_o, M_i)$ can be used as an indicator of the strength of the evidence that the system is in the same state as that described by the model M_i.

Since most physical systems are nonlinear to some extent, situations may occur where, for example, two or more reference models exhibit common symptoms at some operating points. In this case nearly equal degrees of similarity are associated with more that one state and it becomes practically impossible to distinguish between different faults or between correct and faulty operation at those operating points. To avoid this problem, additional sensors can be introduced to discriminate between different faults (Vachkov and Matsuyama, 1992). However, for economic and technical reasons, the installation of additional or different sensors is often not feasible in practice, and alternative solutions must be considered. The approach proposed here is to

evaluate the levels of ambiguity by calculating the maximum degree of similarity between the on-line model and a particular reference model, and each of the other reference models. This value can then be subtracted from the degree of similarity calculated in Equation 3 to eliminate the ambiguous component. Thus,

$$m(\{M_i\}) =$$
$$\frac{\sum_{n=1}^{N} [min\{c_{M_o}(n), c_{M_i}(n)\} - min\{c_{M_o}(n), c_{M_i}(n), c(n)\}]}{\sum_{n=1}^{N} c_{M_o}(n)}$$

where $c(n) = max_{j=1, j \neq i}^{K} c_{M_j}(n)$ and $m(\{M_i\})$ is the unambiguous strength of the evidence that the system has symptoms similar to those of model M_i. The above equation may be written in a more compact form as

$$m(\{M_i\}) = S(M_o, M_i) - Amb_{M_i} \quad (4)$$

where

$$Amb_{M_i} = \frac{\sum_{n=1}^{N} min[c_{M_o}(n), c_{M_i}(n), c(n)]}{\sum_{n=1}^{N} c_{M_o}(n)}$$

and is the total ambiguity associated with the state $\{M_i\}$.

The on-line model M_o will have similar symptoms to two reference models M_i and M_j if $S(M_o, M_i, M_j) \neq 0$ for any i or j such that $(i \neq j)$ where,

$$S(M_o, M_i, M_j) =$$
$$\frac{\sum_{n=1}^{N} min[c_{M_o}(n),\ c_{M_i}(n), c_{M_j}(n)]}{\sum_{n=1}^{N} c_{M_o}(n)}.$$

In this case, the unambiguous strength of the evidence $m(\{M_i, M_j\})$ that the system is in either the state described by the model M_i or by the model M_j is given by,

$$m(\{M_i, M_j\}) = S(M_o, M_i, M_j) - Amb_{\{M_i, M_j\}} \quad (5)$$

where the total ambiguity $Amb_{\{M_i, M_j\}}$ associated with the state $\{M_i, M_j\}$ is given by,

$$Amb_{\{M_i, M_j\}} =$$
$$\frac{\sum_{n=1}^{N} min[c_{M_o}(n), c_{M_i}(n), c_{M_j}(n), c(n)]}{\sum_{n=1}^{N} c_{M_o}(n)}$$

and $c(n) = max_{l=1, l \neq i, l \neq j}^{K} c_{M_l}(n)$. The degree of similarity between the on-line model and three or more reference models can be evaluated in the same way.

The values of the unambiguous strength of the evidence $m(.)$ can be used as normalised basic assignments (Klir and Folger, 1988) since the method of evaluation ensures that the following conditions are satisfied: $m(.) = 0$ indicates no evidence; $m(.) = 1$ indicates complete evidence;

$0 \le m(.) \le 1$ and $\sum_{\{M\} \in \Phi} m(\{M\}) \le 1$, where Φ is the frame of discernment.

2.2 Combination of evidence.

Consider the case where the fuzzy FDD scheme includes only three fuzzy reference models, M_1, M_2 and M_3. The similarity of the on-line fuzzy model with all possible *individual* and *composite* events must be evaluated, where *individual* events are associated with particular faults or correct operation and *composite* events are associated with the union of two or more *individual* events. In this example, the events are: $\{M_1\}$, $\{M_2\}$, $\{M_3\}$, $\{M_1, M_2\}$, $\{M_1, M_3\}$, $\{M_2, M_3\}$ and $\{M_1, M_2, M_3\}$.

The frame of discernment Φ when there are K reference models M_1, M_2, ..., M_K (where M_1 is the model of the correct operation and the rest $K - 1$ models each simulating the operation of the system in the presence of a particular fault) is given by,

$$\Phi = \{\{M_1\}, \{M_2\}, ..., \{M_K\}, \{M_1, M_2\}, \{M_1, M_3\},$$
$$..., \{M_{K-1}, M_K\}, ..., \{M_1, M_2, ..., M_K\}, \{M_U\}\}.$$

where the event $\{M_U\}$ is associated with observed symptoms that are dissimilar to those of any of the reference models (see Figure 4). The strength of evidence that the system is in an unrecognised state $\{M_U\}$ is given by

$$m(\{M_U\}) =$$
$$\frac{\sum_{n=1}^{N} [c_{M_o}(n) - min\{c_{M_o}(n), max_{j=1}^{K}\{c_{M_i}(n)\}\}]}{\sum_{n=1}^{N} c_{M_o}(n)}.$$

In those cases where all combinations of possible faults are not considered, $\sum_{\{M\} \in \Phi} m(\{M\})$ is strictly less that 1, and the remaining evidence is assigned to the event $\{M_1, M_2, ..., M_K\}$.

The total number of possible events, N_e, when K reference models are used is given by

$$N_e = \sum_{p=1}^{K} \frac{K!}{p!(K-p)!} \qquad (6)$$

where p is the number of reference models associated with each event. To avoid the computational complexity associated with combining evidence when a great number of fault models are considered, the diagnosis can be restricted to *individual* events or, if prior knowledge about the fault symptoms is available (Gertler and Anderson, 1992), restricted to include only *composite* events which are associated with faults that have similar symptoms.

Values of the normalised basic assignments are computed each time an on-line fuzzy model is identified. Since the model will be identified at different operating points, it may be assumed

that the evidence is obtained from different sources, and the Dempster rule of combination (Klir and Folger, 1988) can be used to combine new normalised basic assignments m_2 with values obtained previously m_1. Thus,

$$m_c(\{R\}) = \frac{\sum_{A \cap B = R} m_1(\{A\}) m_2(\{B\})}{1 - \sum_{A \cap B = \emptyset} m_1(\{A\}) m_2(\{B\})} \qquad (7)$$

where $m_c(\{R\})$ denotes the combined evidence associated with event $\{R\}$. The events $\{A\}$, $\{B\}$ and $\{R\}$ can be either *individual* or *composite*.

The diagnosis is based on the degree of belief, *Bel*, that the system is in a particular state or in one of a group of possible states, where

$$Bel(\{R\}) = \sum_{R_i \subseteq R} m_c(\{R_i\}). \qquad (8)$$
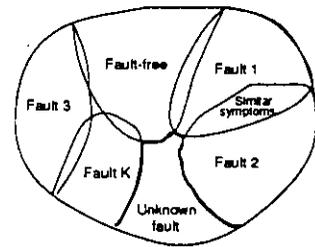
Figure 4: Representation of the fault space

3. A PRACTICAL APPLICATION

To demonstrate its applicability and effectiveness, the fuzzy fault diagnosis scheme is used to detect and identify seven different operating modes in a detailed simulation of the heating system in a large building (Li, et al., 1995). The following operating modes are considered:

- fault-free operation (F_1),
- bad combustion in the burner of the boiler (F_2),
- fouling in the heat exchanger of the boiler (F_3),
- bad tuning of the controller, heating curve too low (F_4),
- boost heating start time too early (F_5),
- boost heating start time too late (F_6),
- leakage in the three-way-valve (F_7).

Simulation data are generated, for fault-free operation and for operation in the presence of each fault, over a period of seventy days. The data are divided into two groups. The first group consists of data collected during the 35 odd days which are used to identify the fuzzy reference models. The second group of data collected during the even days are used for testing the FDD scheme. The raw data collected over each day are reduced to a

set of mean values which are representative of the operation of the heating system at different times of the day.

The inputs of the fuzzy models are: the mean temperature of the water supplied by boiler during the boost period (06:00-08:00), the mean temperature of the air extracted by the mechanical ventilation system at 06:00, the mean temperature of the air extracted by the mechanical ventilation system at 08:00, the mean temperature of the air in the building measured between 18:00-23:00 and the mean outside air temperature. The output of the fuzzy model is the mean temperature of the air extracted by the mechanical ventilation system during the occupancy period between 10:00-18:00.

Fuzzy C-Means clustering (Bezdek, 1991) is used to determine the fuzzy sets for the input and output spaces. The clustering is performed individually on each of the variables using the training data from all the operating modes. As can be seen in Figure 5, the number of clusters for each variable is limited to a maximum of four. Fuzzy reference sets are constructed by approximating the clusters by triangular membership functions.

The performance of the fuzzy FDD scheme is investigated using all seven test data sets. In this case, the diagnosis is restricted to *individual* events only. The tests are numbered T_1 to T_7 corresponding to the operating modes F_1 to F_7. Figures 6 and 7 show the variations of the normalised basic assignments and the degrees of belief for the test data corresponding to fault-free operation. The maximum values of the degrees of belief obtained during each of the tests are given in Table 1.

Table 1 Maximum values of the degrees of belief (%)

Test	Reference models						
	F_1	F_2	F_3	F_4	F_5	F_6	F_7
T_1	99.1	1.3	0.0	15.9	13.3	0.5	0.0
T_2	0.0	99.9	0.4	1.7	0.2	0.0	0.0
T_3	0.0	0.0	99.9	0.0	0.0	0.0	0.0
T_4	0.0	0.0	0.0	99.9	0.3	0.0	0.0
T_5	0.0	0.0	0.0	0.4	99.9	0.2	0.2
T_6	0.5	0.0	0.0	0.0	0.0	100	0.0
T_7	0.0	0.5	0.0	0.0	0.0	0.0	100

The fault condition is correctly diagnosed for each of the tests. However, the indoor temperature is unusually low on some of the days during test T_1, and the controller fails to predict the correct time to restart the boost heating. On these days, the FDD scheme generates some evidence that the behaviour of the plant is similar to operating modes F_4 and F_5, even though the operation of the plant is supposedly fault-free. As can be seen in Figure 7, the beliefs in operating modes F_4 and F_5 both decay to zero, as more evidence is collected

that the system is operating correctly (see Figure 6).
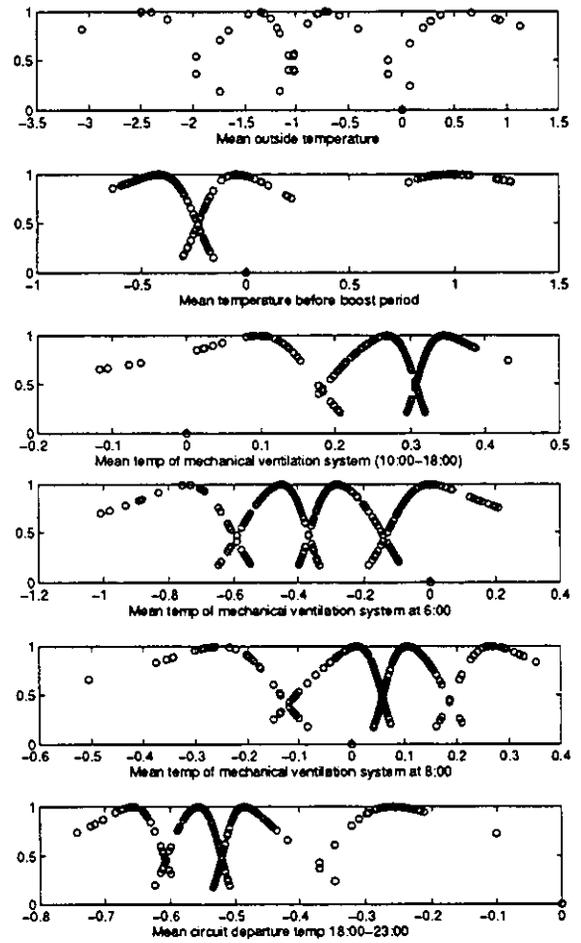
Figure 5: Clustering of the training data from all operating modes

4. CONCLUSIONS

A fuzzy diagnosis scheme which is suitable for on-line implementation in low-cost hardware is able to identify faults in ill-defined systems. Fuzzy matching can be used to generate normalised basic assignments if the ambiguity which arises from fault-free and faulty operation, or different faults, having similar symptoms is taken into account. Results have demonstrated that the scheme can successfully identify the different operating modes of a heating system in a large building.

ACKNOWLEDGEMENTS

REFERENCES

Bezdek, J.C. (1991). Self-Organization and Clustering Algorithms. *Proceedings of the Second Joint Technology Workshop on Neural Networks and Fuzzy Logic*, Vol. 1, 143-158.

Dexter, A.L. and M. Benouarets (1995a). Generic modelling of HVAC plant for fault diagnosis. *Paper to be presented at IBPSA conference, Madison, Wisconsin, August 14-16.*

Dexter, A.L. and M. Benouarets (1995b). Transient detection. In: *IEA ANNEX 25, Building Optimisation and Fault Diagnosis Source Book* (J. Hyvärinen, (Ed.)), pp. 259-263. Technical Research of Finland, P.O Box 206, FIN-02151 Espoo, Finland.

Gertler, J.J. and K.C Anderson (1992). An evidential reasoning extension to quantitative model-based failure diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, No. 2. 275-289.

Klir, G.J. and T.A. Folger (1988). *Fuzzy Sets, Uncertainty and Information.* Chap. 5. Prentice Hall Int, New Jersey, USA.

Lee, Keon-Myung, Kyoung-A. Seong and H. Lee-Kwang (1992). Fuzzy Matching and Fuzzy Comparison in Fuzzy Expert System. *Proceedings of the 2nd International Conference on Fuzzy Logic & Neural Networks*, Iizuka, Japan 17-22. 313-316.

Li, X., H. Vaezi-Nejad and JC. Visier (1995). Application of Neural Networks to the Fault Detection in HVAC Systems. *2nd International Conference on Industrial Automation*, Nancy, France.

Linden, S., H. Benkhedda, A.J. Wilkinson and N.A. Smith (1994). Fuzzy qualitative modelling of pressure die-casting for fault diagnosis. *Proceedings of the IEE Control*, University of Warwick. 1323-1328.

Maruyama, N., M. Benouarets and A.L. Dexter (1995). Detecting faults in nonlinear dynamic systems using static neuro-fuzzy models. *IEE Colloquium on Qualitative and Quantitative Modelling Methods for Fault Diagnosis*, York, Digest No: 1995/079, 8/1-8/10.

Sauter, D., N. Mary, F. Sirou and Thieltgen (1994). Fault Diagnosis in Systems Using Fuzzy Logic. *Proc. 3rd. IEEE Conference on Control Applications*, Glasgow, U.K., Vol. 2, 883-888.

Schneider, H. and P.M. Frank (1994). Fuzzy logic based threshold adaption for fault detection in robots. *Proc. 3rd. IEEE Conference on Control Applications*, Glasgow, U.K., Vol. 2. 1127-1132.

Sugeno, M. and T. Yasukawa (1993). A fuzzy-logic apprach to qualitative modelling. *IEEE Transactions of Fuzzy Systems*, Vol. 1, No. 1. 7-31.

Terpstra, V.J., H.B. Verbruggen, M.W. Hoogland and R.A.E Ficke (1992). A real-time, fuzzy, deep-knowledge based fault-diagnosis system for a CSTR. *Proc. IFAC Symp. on On-line Fault Detection and Supervision in the Chemical Process Industries*, Newark, Delaware, USA. 26-31.

Vachkov, G. and H. Matsuyama (1992). Fault diagnosis method by using fuzzy rule based models. *Proc. 2nd. Int Conf. on Fuzzy Logic and Neural Networks*, Iizuka, Japan. 385-388.

Xu, Chen-Wei and Yong-Zai. Lu (1987). Fuzzy model identification and self-learning for dynamic systems. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-17, No. 4. 683-689.

Sosnowski, Z.A, and W. Pedrycz (1992). *Fuzzy logic for the management of uncertainty.* (Zadeh L. and J. Kacprzyk, (Ed.)), pp. 501-505, John Wiley & Sons, Inc, New York, USA.
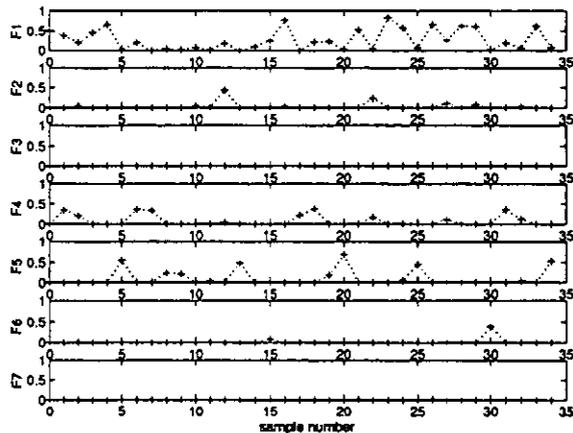
Figure 6: Variations of normalised basic assignments for test T_1
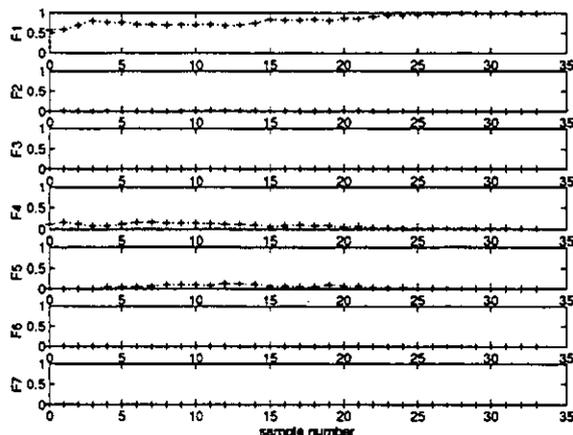
Figure 7: Variations of degrees of belief values for test T_1

ROUTINES FOR IMPLEMENTATION OF C EXPERT SYSTEMS

V.E. Bakker, H.C. Peitsman
TNO Building and Construction Research
Department of Indoor Environment and Systems
Delft, The Netherlands

Abstract

RICE stands for *Routines for Implemention of C Expert systems*. The general idea of the software is to provide a tool for easy implementing small, but powerfull, (fuzzy) expert systems within C or C++ programs. The software is supposed to take away all worries about the inference mechanism: the user provides the knowledge on a high level, but is still able to call plain C-code (or C++) from within the knowledge base. This means in fact that the user can build a program or a part of a program in which the execution order of plain C-code is completely determined by the knowledge base which is inferred.

In expert systems, the inference engine is not the critical part; designing a functional knowledge base is the hardest part, which is a task of the expert system user. This paper presents an example of an expert system based on RICE. The example is based on the results of a TRNSYS simulation of a VAV system plant.

1 INTRODUCTION

The introduction describes the general properties of expert systems; expert system logic and reasoning will be discussed. Several common reasoning methods will be explained. However, the RICE implementation of the methods falls out of the scope of this paper. The RICE program, including instruction manual, is available on World Wide Web.

Essential for the definition of experts systems is the seperation of knowledge and inference. Unlike other computer programs, an expert system has a predefined part, generally referenced to as the inference engine, and a variable part, referenced to as the knowledge base. An expert system can be described as

expert system = knowledge base + inference engine

In the following two sections, the knowledge base and inference engine will be discussed.

1.1 THE KNOWLEDGE BASE

The knowledge base of an expert system contains knowledge of a specific domain. For example, an expert system for medical diagnosis contains expert knowledge for a set of diseases. The knowledge is normally stored in the knowledge base using a symbolic representation, for example:

if there is smoke then there is fire

This representation is known as a production rule, or rule for short. Other representations are frames and objects. The main advantage of having a seperation between the knowledge base and the inference engine is the fact that an iterative development of the expert system is possible in a more natural way than in case of normal computer programs. The production rule is specificaly suitable for implementing heuristic knowledge.

It is possible that the knowledge base contains knowledge about knowledge within itself; this is called meta-knowledge. Examples of meta-knowledge are rules which, when fired, will focus on a certain part of the knowledge base (focus of attention). Also possible is knowledge which adapts other knowledge within the knowledge base, for example, a rule which changes the degree of certainty of another rule.

The knowledge base is filled with rules like the one in the example above and the inference engine will act on this knowledge base to obtain new information. The next section describes the inference engine in more detail.

1.2 THE INFERENCE ENGINE

The inference engine is the part of the expert system that is not changed by expert system developers. It is generally integrated with an expert system shell or toolbox. The main task of the inference engine is inferring new information using the knowledge base and (already) stored information. When we take the example from the previous section, the inference engine is able to infer the fact *there is fire* when the fact there is *smoke is true*, using the rule from the example.

To be able to obtain new information using existing information and knowledge from the knowledge base, the inference engine needs logic, to be able to perform the task in a logical sense, and reasoning methods to use the knowledge base in a consisting way. Both aspects will be discussed in the following sections.

1.3 REASONING AND SEARCH METHODS

In the design of expert systems several reasoning and search methods can be chosen. It depends on the application which methods and techniques are used. More detailed information about the design of expert systems can be found in [Luger and Stubblefield 1989].

In expert systems two main reasoning methods can be found: forward and backward reasoning. Forward reasoning, also known as data-driven or bottom-up reasoning or chaining, uses initial data and infers new data from the known data. The inference of new data is done by applying rules from the rule base. The inference of new knowledge stops in case some predefined goal is reached or no new knowledge can be infered. Backward reasoning, also called goal-driven or top-down reasoning or chaining, starts with certain goals and in order to solve (read: prove) these goals their subgoals are tried to be solved. This mechanism repeats until no subgoals are present or the initial goals are solved. The combination of forward and backward reasoning in one inference engine within an expert system is possible, but not often used in practice.

In addition to the above reasoning methods, three search methods can be distinguished: depth-first, breadth-first and heuristic search. The depth-first search method is normally used in combination with backward reasoning, the breadth-first method with forward reasoning. Whenever it is possible, depth-first search goes deeper into the search tree. The search stops in case no lower levels within the serach tree exist or a stop criterium is met. The breadth-first search method however, searches level by level within the search tree. The exploration of the next level of a search tree is started in case in the current level no more states can be solved or a certain stop criterium is met.
Heuristic search is applied to guide the search within the search tree and therefore can be seen as meta-knowledge: knowledge about knowledge. The best-first search method is an example of heuristic search. In this search method the 'most promising' branch of a search tree is chosen for further search. Confidence factors can be used to indicate which branch is more promising to lead to a satisfying result than others.

1.4 REAL-TIME ISSUES

Real-time behaviour is often easier to recognize than to define. As discussed in [O'Reilly 19??] many definitions of real time exist. Real-time is mostly related to 'fast': meaning that a system processes data quickly. A formal definition of real time is offered:

> *a hard real-time system is defined as a system in which correctness of the system not only depends on the logical results of a computation, but also on the time at which the results are produced.*

The most important item is the response time, if events are not handled timely, the process can get out of control. Thus, the feature that defines a real-time system is the system's ability to guarantee a response before a certain time has elapsed, where that time related to the dynamic behaviour of the system. If, given an arbitrary event or state of the system, the system always produces a response by the time it is needed, then the system is said to be real-time. Due to the real-time aspect the fuzzy expert system should be able to perform non-monotonic, temporal and progressive reasoning. Non-monotonic reasoning introduces the ability to change already proven data and all the infered data that depend on it. Assumptions which later on in the reasoning process appear to be erroneous can be changed and all the data which depend those assumptions must be reexamined and possibly withdrawn from the data base. The reexamination of the data base is done by a truth-maintenance system [Doyle 1979]. Non-monotonic reasoning is more like the reasoning of humans than monotonic reasoning is, in which previously added knowledge or data can not be reexamined.

In case the expert system deals with past, present and future, it has to be able to outdate previous and accept new data as the time window connected to the present time is moving. This principle is called temporal reasoning [Krijgsman et al. 1991] and an essential aspect of real-time expert systems.

Also an important aspect of the real-time expert systems is the progressive reasoning principle [Lattimer et al. 1986]. In case of progressive reasoning the rule base is divived in several layers, which are in fact also rule bases. The inference engine starts with the inference of the first layer and proceeds with the next (higher) layer in case the inference of the first layer is finished. Each layer is able to use the data inferred by the lower layers as well as the rules within that layer. In case the real-time environment aborts the reasoning by means of an interrupt, the data infered by the last layer which completed its inference can be used. In this way the more time available, the 'better' the conclusions get. This aspect has also close resemblance to human reasoning.

2 APPLICATION OF RICE TO A VAV SYSTEM PLANT

The purpose of this example is to establish fault detection using RICE 4.0. Data was generated from a TRNSYS simulation of a VAV system plant and controls that were similar to the IEA Annex 25 Reference Air Handling System (see Figure 1). The simulation was altered to generate faulty data. Ten hours of summer weather data was utilised, in which a TRNSYS error type was developed, causing the *Damper of the Mixing Box* to be fully opened after 15000 seconds from simulation start. Furthermore errors in the control strategy were introduced.
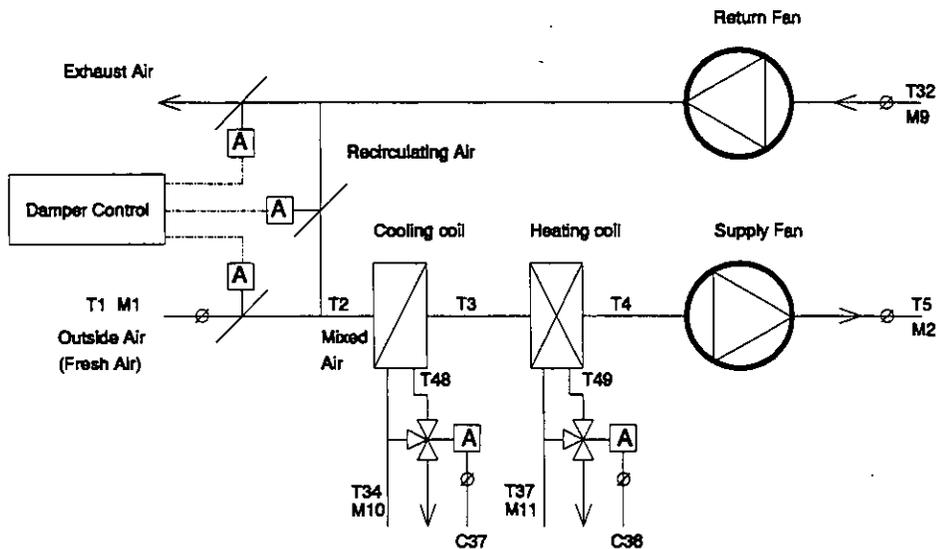
Figure 1 VAV Pl Diagram

2.1 DEFINITION OF THE TYPE OF FAULTS

The aim was to detect the following faults:

• Wastefull Heating (Heating not allowed)

- Wastefull Cooling (Cooling not allowed)
- Damper stuck in maximal open position
- Damper stuct in minimal open position
- Damper not responding to control signal (stuck in between)

2.2 DEFINITION OF THE RULES

The following rules were tried for fault detection:

- If the optimal damper position is not maximal open or closed, no heating or cooling is allowed.
- If the optimal damper position is maximal open and the estimated mixing temperature is higher than the setpoint of the supply air, cooling is allowed. If the estimated mixing temperature is lower, heating is allowed.
- If the mixing temperature is close to the estimated closed damper mixing temperature the damper is assumed to be closed.
- If the mixing temperature is close to the estimated open damper mixing temperature the damper is assumed to be opened.
- If the mixing temperature is between the estimated open damper and closed damper mixing temperatures the damper is assumed to be modulating.
- If the setpoint of the supply air is between the estimated open and closed mixing temperature the optimal damper position is modulating.
- If the optimal damper position is not modulating, the estimated open and closed mixing temperatures determine whether the optimal damper position is opened or closed.
- The control signals of the cooling and heating coil determined if cooling or heating was enabled.
- The control signal of the damper was compared to the estimated damper position
- The mixing temperatures were estimated using the simplyfied relation:

$$Tmix = (DampPos\text{-}Leakage)^*Tatm+(1\text{-}DampPos+Leakage)^*Tret$$

For details of the program example, see Appendix A & B.
The results are presented in Appendix C.

3 CONCLUSIONS

The purpose of the example was to establish fault detection using RICE 4.0. It is concluded to be very well possible to develop a simple expert system which is able to detect and diagnose faults of a particular application. RICE is very suitable for experimentation with expert systems, but the low-level straight-forward approach is not suited for a larger professional environment.

4 COPYRIGHT, WARRANTY AND UPDATES

4.1 INFORMATION

For information, please contact the RICE expert by:

René Jager
Wevershof 19
1483 XJ De Rijp
The Netherlands
tel   : + 31 2992 4297 (home)
E-mail   : rjager@simplex.nl

New version of RICE (4.1) is now available on World Wide Webb.

WWW http   : //simplex.nl/users/rjager/

Last version : 4.1
Output : source code + manual (45 pages)

4.2 LICENSE

The toolbox is free software; you can redistribute and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version. The toolbox is distributed in the hope that it will be useful but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this toolbox; if not, write to the Free Software Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA.

REFERENCES

[Doyle 1979]   J. Doyle. *Truth maintenance system.* Artificial Intelligence, vol. 12, 1979.

[Krijgsman et al. 1991]   A.J. Krijgsman, R. Jager, H.B. Verbruggen and P.M. Bruijn. *DICE: a framework for intelligent real-time control.* Proceedings 3th IFAC Workshop AIRTC '91, Napa (Ca), U.S.A., September 1991.

[Lattimer et al. 1986]   M. Lattimer Wright, M.W. Green, G. Fiegl and P.F. Cross. *An expert system for real-time control.* IEEE Software, pp. 16-24, March 1986.

[Luger and Stubblefield 1989]   G.F. Luger and W.A. Stubblefield. *Artificial intelligence and the design of expert systems.* The Benjamin/Cummings Publishing Company Inc., Redwood City (CA, USA), 1989.

[O'Reilly and Cromarty 19??]   C.A. O'Reilly and A.S. Cromarty. *Fast is not "Real Time" (Designing Effective Real-Time AI Systems).* Applications of Artificial Intelligence II 548, pp. 249-257, Bellingham, ???.

Appendix A : Source code of the test program

```c
/*
File          :     DR_AHU2.C
Date          :     17-02-1995
Author        :     V.E.Bakker
Company       :     TNO Building and Construction research
                    Department of Indoor Environment,
                    Building Physics and Systems
                    The Netherlands
E-mail        :     peitsman@bouw.tno.nl
Last modified :     24-09-1995

First trial for Air Handle Unit fault detection using the RICE 4.0 system.
*/


#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <conio.h>
#include "rice.h"
#include <dos.h>
#include "dr_ahu2.h"


int read_temps(FILE *, struct Temperatures *);
int read_stats(FILE *, struct Status *);

int theES;

long TimeStamp;
float Tatm, Tret, Tmix, Tsup, Tset;
float Vdamp, Verr, Dheat, Dcool, Chiller_On, Boiler_On;
float EstTmix, EstTmixMin, EstTmixMax, MaxMixTemp, MinMixTemp,
      MinDampNearest, OptDampPosMod,
      MaxDampPos = MAXDAMPPOS, MinDampPos = MINDAMPPOS, Leak = LEAKAGE;

FILE *OutputFile;

void printValues(void)
{
    gotoxy(1,1);
    printf("Simulation results:\n\n");
    printf("TimeStamp: %ld\n\n",TimeStamp);
    printf("Tatm .................... : %2.2f \n",Tatm);
    printf("Tretn ................... : %2.2f \n",Tret);
    printf("Tmix .................... : %2.2f \n",Tmix);
    printf("Tsup .................... : %2.2f \n",Tsup);
    printf("Tset .................... : %2.2f \n",Tset);
    printf("Vdamp ................... : %1.2f (input for error type)\n",Vdamp);
    printf("Verr .................... : %1.2f (output from error type)\n",Verr);
    printf("Dheat ................... : %1.2f \n",Dheat);
    printf("Dcool ................... : %1.2f \n",Dcool);
    printf("Chiller_on .............. : %1.1f \n",Chiller_On);
    printf("Boiler_On ............... : %1.1f \n\n",Boiler_On);
    printf("Estimated values:\n\n");
    printf("Estimated Tmix .......... : %1.2f \n",EstTmix);
    printf("Est.Tmix (Vdamp 0.19) .... : %2.2f \n",EstTmixMin);
    printf("Est.Tmix (Vdamp 1.00) .... : %2.2f \n",EstTmixMax); }

void fprintValues(void)
{
    fprintf(OutputFile,"\tSimulation results:\n\n");
    fprintf(OutputFile,"TimeStamp: %ld\n\n",TimeStamp);
    fprintf(OutputFile,"Tatm .................... : %2.2f \n",Tatm);
    fprintf(OutputFile,"Tretn ................... : %2.2f \n",Tret);
    fprintf(OutputFile,"Tmix .................... : %2.2f \n",Tmix);
    fprintf(OutputFile,"Tsup .................... : %2.2f \n",Tsup);
    fprintf(OutputFile,"Tset .................... : %2.2f \n",Tset);
    fprintf(OutputFile,"Vdamp ................... : %1.2f (Input of the Error
type)\n",Vdamp);
    fprintf(OutputFile,"Verr .................... : %1.2f (Output of the Error
type)\n",Verr);
    fprintf(OutputFile,"Dheat ................... : %1.2f \n",Dheat);
    fprintf(OutputFile,"Dcool ................... : %1.2f \n",Dcool);
    fprintf(OutputFile,"Chiller_on .............. : %1.1f \n",Chiller_On);
    fprintf(OutputFile,"Boiler_On ............... : %1.1f \n\n",Boiler_On);
    fprintf(OutputFile,"\tEstimated values:\n\n");
    fprintf(OutputFile,"Estimated Tmix .......... : %1.2f \n",EstTmix);
    fprintf(OutputFile,"Est.Tmix (Vdamp 0.19) .... : %2.2f \n",EstTmixMin);
    fprintf(OutputFile,"Est.Tmix (Vdamp 1.00) .... : %2.2f \n\n",EstTmixMax);
}

void showInfo(char *msg, float grade)
```

```
{
    fprintValues();
    fprintf(OutputFile,"\tConclusion of the Inference Mechanism:\n\n");
    fprintf(OutputFile,"%s : %1.1f\n\n",msg,grade);
    fprintf(OutputFile,"\tExplanation of the Inference Mechanism:\n\n");
    rice_ExplainHow(theES,msg,-1);
    fprintf(OutputFile,"\n\n");
}

void redirExplain(int kb, char* Explanation)
{
    fprintf(OutputFile,"%s",Explanation);
}

void redirInfo(int kb, int argc, char *argv[], float *grade)
{
    showInfo(argv[0],*grade);
}


RICE_Linker(myLinker)
{
    RICE_Link("Optimal Damper Position = Modulating",
    {
        *rice_grade = OptDampPosMod;
    });

    RICE_Link("Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint",
    {
        *rice_grade = MinDampNearest;
    });

    RICE_Link("Estimated Minimal Open Mixing Temperature < Tsetpoint",
    {
        if (EstTmixMin < Tset)
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Estimated Minimal Open Mixing Temperature > Tsetpoint",
    {
        if (EstTmixMin > Tset)
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Estimated Maximal Open Mixing Temperature < Tsetpoint",
    {
        if (EstTmixMax < Tset)
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Estimated Maximal Open Mixing Temperature > Tsetpoint",
    {
        if (EstTmixMax > Tset)
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Estimated Maximal Open Mixing Temperature = Tsetpoint",
    {
        if (EstTmixMax == Tset)
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Heating Enabled",
    {
        if ( Dheat > 0.02 )
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
    });

    RICE_Link("Cooling Enabled",
    {
        if ( Dcool > 0.02 )
            *rice_grade = 1.0;
        else
            *rice_grade = 0.0;
```

```
    });

    RICE_Link("Mixing Temperature = Maximal Open Mixing Temperature",
    {
        if (fabs(Tmix - EstTmixMax) <= 0.3)
                *rice_grade = 1.0;
        else
                *rice_grade = 0.0;
    });

    RICE_Link("Mixing Temperature = Minimal Open Mixing Temperature",
    {
        if (fabs(Tmix - EstTmixMin) <= 0.3)
                *rice_grade = 1.0;
        else
                *rice_grade = 0.0;
    });

    RICE_Link("Mixing Temperature = Between Maximal and Minimal Open Mixing Temperature",
    {
        if ((Tmix < MaxMixTemp) & (Tmix > MinMixTemp))
                *rice_grade = 1.0;
        else
                *rice_grade = 0.0;
    });

    RICE_Link("Control Signal of Damper = Maximal Open",
    {
        if (Vdamp > 0.98)
                *rice_grade = 1.0;
        else
                *rice_grade = 0.0;
    });

    RICE_Link("Control Signal of Damper = Minimal Open",
    {
        if (Vdamp < 0.22)
                *rice_grade = 1.0;
        else
                *rice_grade = 0.0;
    });
}


RICE_LINKER theLinker = myLinker;


void main(void)
{
    struct Temperatures Temps;
    struct Status Stats;
    FILE *tempsfile, *statsfile;

    if ((OutputFile = fopen(OUTPUTFILE,"w")) == NULL)
    {
        printf("Error could not open OUTPUT file! \n\n");
        exit(1);
    }

    if ((tempsfile = fopen(TEMPSFILE,"r")) == NULL)
    {
        printf("Error could not open TEMPERATURES file!\n\n");
        exit(1);
    }
    if ((statsfile = fopen(STATSFILE,"r")) == NULL)
    {
        printf("Error could not open STATUS file!\n\n");
        exit(1);
    }

    theES = rice_CreateES();
    rice_SetInteractive(theES,0);
    rice_RedirectInform(theES, redirInfo);
    rice_RedirectExplain(theES,redirExplain);
    rice_UseLinker(theES,1,&theLinker);
    rice_CompileKB(theES, THE_KB);


    clrscr();
    _setcursortype(_NOCURSOR);

    while ( read_temps(tempsfile, &Temps) & read_stats(statsfile, &Stats) )          {
        gotoxy(1,1);
        clreol();

        if (Temps.TimeStamp != Stats.TimeStamp)
```

662

```c
            {
                clrscr();
                printf("Status and Temperatures file not synchronized!\n\n");
                printf("Timestamp of Temperatures file: %ld\n",Temps.TimeStamp);
                printf("Timestamp of Status file       : %ld\n\n",Stats.TimeStamp);
                printf("Program stopped with exitcode: 1\n\n");
                exit(1);
            }

        TimeStamp   = Temps.TimeStamp;
        Tatm        = Temps.Tatm;
        Tret        = Temps.Tret;
        Tmix        = Temps.Tmix;
        Tsup        = Temps.Tsup;
        Tset        = Temps.Tset;
        Vdamp       = Stats.Vdamp;
        Verr        = Stats.Verr;
        Dheat       = Stats.Dheat;
        Dcool       = Stats.Dcool;
        Chiller_On  = Stats.Chiller_On;
        Boiler_On   = Stats.Boiler_On;

        EstTmix = (Vdamp-Leak)*Tatm+(1-Vdamp+Leak)*Tret;

        EstTmixMin = (MinDampPos-Leak)*Tatm+(1-MinDampPos+Leak)*Tret;

        EstTmixMax = (MaxDampPos-Leak)*Tatm+(1-MaxDampPos+Leak)*Tret;

        if ( fabs(Tset-EstTmixMin) < fabs(Tset-EstTmixMax) )
                MinDampNearest=1.0;
        else
                MinDampNearest=0.0;

        if ( EstTmixMin <= EstTmixMax )
        {
                MinMixTemp = EstTmixMin;
                MaxMixTemp = EstTmixMax;
        }
        else
        {
                MinMixTemp = EstTmixMax;
                MaxMixTemp = EstTmixMin;
        }

        if ( (Tset <= MaxMixTemp) & (Tset >= MinMixTemp) )
                OptDampPosMod = 1.0;
        else
                OptDampPosMod = 0.0;

        printValues();

        rice_ClearKB(theES);
        rice_InferKB(theES);

    }

    _setcursortype(_NORMALCURSOR);
    fclose(tempsfile);
    fclose(statsfile);
    fclose(OutputFile);

}

int read_temps(FILE *tempsfile, struct Temperatures *Temps)
{
    float timefloat;
    char  string[80];

    if ((fgets(string,80,tempsfile))==NULL)
        return(0);

    sscanf(string,"%f %f %f %f %f %f",&timefloat,&Temps->Tatm,&Temps->Tret,&Temps-
>Tmix,&Temps->Tsup,&Temps->Tset);
    Temps->TimeStamp=(long)(timefloat);

    return(1);
}

int read_stats(FILE *statsfile, struct Status *Stats)
{
    float timefloat;
    char  string[80];

    if ((fgets(string,80,statsfile))==NULL)
        return(0);
```

```
    sscanf(string,"%f %f %f %f %f %f %f",&timefloat,&Stats->Vdamp,&Stats->Verr,&Stats-
>Dheat,&Stats->Dcool,&Stats->Chiller_On,&Stats->Boiler_On);

    Stats->TimeStamp=(long)(timefloat);

    return(1);
}

/*
File    :        dr_ahu.h
*/

#define TEMPSFILE "err_temps.dat"
#define STATSFILE "err_stats.dat"
#define OUTPUTFILE "dr_ahu2.out"
#define THE_KB "dr_ahu2.kb"
#define MAXDAMPPOS 1.0
#define MINDAMPPOS 0.2
#define LEAKAGE 0.091


struct Temperatures
{
    long         TimeStamp;
    float        Tatm;
    float        Tret;
    float Tmix;
    float Tsup;
    float Tset;
};

struct Status
{
    long     TimeStamp;
    float    Vdamp;
    float    Verr;
    float    Dheat;
    float    Dcool;
    float    Chiller_On;
    float    Boiler_On;
};
```

Appendix B : The Knowledge Base

Layer Predictor

If Optimal Damper Position = Modulating
ThenNot Heating Allowed
ThenNot Cooling Allowed

IfNot Optimal Damper Position = Modulating
And Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint
Then Optimal Damper Position = Minimal Open

IfNot Optimal Damper Position = Modulating
AndNot Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint
Then Optimal Damper Position = Maximal Open

If Optimal Damper Position = Minimal Open
And Estimated Minimal Open Mixing Temperature < Tsetpoint
Then Heating Allowed
ThenNot Cooling Allowed

If Optimal Damper Position = Minimal Open
And Estimated Minimal Open Mixing Temperature > Tsetpoint
ThenNot Heating Allowed
Then Cooling Allowed

If Optimal Damper Position = Maximal Open
And Estimated Maximal Open Mixing Temperature < Tsetpoint
Then Heating Allowed
ThenNot Cooling Allowed

If Optimal Damper Position = Maximal Open
And Estimated Maximal Open Mixing Temperature > Tsetpoint
ThenNot Heating Allowed
Then Cooling Allowed

If Estimated Maximal Open Mixing Temperature = Tsetpoint
ThenNot Heating Allowed
ThenNot Cooling Allowed

IfNot Control Signal of Damper = Maximal Open
AndNot Control Signal of Damper = Minimal Open
then Control Signal of Damper = Between Maximal and Minimal Open

Layer Diagnose

Goal Diagnose Found

IfNot Heating Allowed
And Heating Enabled
ThenInf Warning Wastefull Heating
Then Diagnose Found

IfNot Cooling Allowed
And Cooling Enabled
ThenInf Warning Wastefull Cooling
Then Diagnose Found

If Mixing Temperature = Maximal Open Mixing Temperature
AndNot Control Signal of Damper = Maximal Open
ThenInf Error Damper stuck in Maximal Open position
Then Diagnose Found

If Mixing Temperature = Minimal Open Mixing Temperature
AndNot Control Signal of Damper = Minimal Open
ThenInf Error Damper stuck in Minimal Open position
Then Diagnose Found

If Mixing Temperature = Between Maximal and Minimal Open Mixing Temperature
AndNot Control Signal of Damper = Between Maximal and Minimal Open
ThenInf Error Damper not responding to control signal
Then Diagnose Found

Appendix C : Simulation Results

<u>Simulation results:</u>
TimeStamp: 16680

```
Tatm ..................... : 22.30
Tretn .................... : 19.37
Tmix ..................... : 22.18
Tsup ..................... : 24.53
Tset ..................... : 22.93
Vdamp .................... :  1.00  (Input of the Error type)
Verr ..................... :  1.00  (Output of the Error type)
Dheat .................... :  0.00
Dcool .................... :  0.03
Chiller_on ............... :  1.00
Boiler_On ................ :  0.00
```

Estimated values:

```
Estimated Tmix ........... : 22.03
Est.Tmix (Vdamp 0.19) .... : 19.69
Est.Tmix (Vdamp 1.00) .... : 22.03
```

Conclusion of the Inference Mechanism:

Warning Wastefull Cooling : 1.0

Explanation of the Inference Mechanism:

```
        [*] means unknown
        [0] means false
        [1] means true
```

IfNot Cooling Allowed [1]
And Cooling Enabled [1]
ThenInf Warning Wastefull Cooling [1]
Then Diagnose Found [*]

If Optimal Damper Position = Modulating [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [*]
ThenNot Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
And Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [0]
Then Optimal Damper Position = Minimal Open [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
AndNot Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [1]
Then Optimal Damper Position = Maximal Open [1]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [1]
ThenNot Cooling Allowed [0]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]

Then Cooling Allowed [*]

If Estimated Maximal Open Mixing Temperature = Tsetpoint [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]

Simulation Results:
TimeStamp: 20820

Tatm .... : 24.71
Tretn .... : 19.86
Tmix .... : 24.36
Tsup .... : 24.26
Tset .... : 24.29
Vdamp .... : 0.97 (Input of the Error type)
Verr .... : 1.00 (Output of the Error type)
Dheat .... : 0.00
Dcool .... : 0.33
Chiller_on .... : 1.00
Boiler_On .... : 0.00

Estimated values:

Estimated Tmix .... : 24.10
Est.Tmix (Vdamp 0.19) .... : 20.39
Est.Tmix (Vdamp 1.00) .... : 24.27

Conclusion of the Inference Mechanism:

Warning Wastefull Cooling : 1.0

Explanation of the Inference Mechanism:

IfNot Cooling Allowed [1]
And Cooling Enabled [1]
ThenInf Warning Wastefull Cooling [1]
Then Diagnose Found [*]

If Optimal Damper Position = Modulating [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [*]
ThenNot Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
And Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [0]
Then Optimal Damper Position = Minimal Open [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
AndNot Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [1]
Then Optimal Damper Position = Maximal Open [1]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [1]
ThenNot Cooling Allowed [0]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

667

```
If Estimated Maximal Open Mixing Temperature = Tsetpoint [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]
```

Simulation results:
TimeStamp: 20820

```
Tatm ..................... : 24.71
Tretn .................... : 19.86
Tmix ..................... : 24.36
Tsup ..................... : 24.26
Tset ..................... : 24.29
Vdamp .................... :  0.97 (Input of the Error type)
Verr ..................... :  1.00 (Output of the Error type)
Dheat .................... :  0.00
Dcool .................... :  0.33
Chiller_on ............... :  1.00
Boiler_On ................ :  0.00
```

Estimated values:

```
Estimated Tmix ........... : 24.10
Est.Tmix (Vdamp 0.19) .... : 20.39
Est.Tmix (Vdamp 1.00) .... : 24.27
```

Conclusion of the Inference Mechanism:

Error Damper stuck in Maximal Open position : 1.0

Explanation of the Inference Mechanism:

```
If Mixing Temperature = Maximal Open Mixing Temperature [1]
AndNot Control Signal of Damper = Maximal Open [1]
ThenInf Error Damper stuck in Maximal Open position [1]
Then Diagnose Found [*]
```

Simulation results:
TimeStamp: 16680

```
Tatm ..................... : 22.30
Tretn .................... : 19.37
Tmix ..................... : 22.18
Tsup ..................... : 24.53
Tset ..................... : 22.93
Vdamp .................... :  1.00 (Input of the Error type)
Verr ..................... :  1.00 (Output of the Error type)
Dheat .................... :  0.00
Dcool .................... :  0.03
Chiller_on ............... :  1.00
Boiler_On ................ :  0.00
```

Estimated values:

```
Estimated Tmix ........... : 22.03
Est.Tmix (Vdamp 0.19) .... : 19.69
Est.Tmix (Vdamp 1.00) .... : 22.03
```

Conclusion of the Inference Mechanism:

Warning Wastefull Cooling : 1.0

Explanation of the Inference Mechanism:

```
IfNot Cooling Allowed [1]
And Cooling Enabled [1]
ThenInf Warning Wastefull Cooling [1]
Then Diagnose Found [*]

If Optimal Damper Position = Modulating [0]
```

```
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [*]
ThenNot Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
And Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [0]
Then Optimal Damper Position = Minimal Open [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
AndNot Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [1]
Then Optimal Damper Position = Maximal Open [1]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [1]
ThenNot Cooling Allowed [0]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

If Estimated Maximal Open Mixing Temperature = Tsetpoint [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]
```

Simulation results:
TimeStamp: 20820

```
Tatm ....................... : 24.71
Tretn ...................... : 19.86
Tmix ....................... : 24.36
Tsup ....................... : 24.26
Tset ....................... : 24.29
Vdamp ...................... :  0.97  (Input of the Error type)
Verr ....................... :  1.00  (Output of the Error type)
Dheat ...................... :  0.00
Dcool ...................... :  0.33
Chiller_on ................. :  1.00
Boiler_On .................. :  0.00
```

Estimated values:

```
Estimated Tmix ............ : 24.10
Est.Tmix (Vdamp 0.19) .... : 20.39
Est.Tmix (Vdamp 1.00) .... : 24.27
```

Conclusion of the Inference Mechanism:

Warning Wastefull Cooling : 1.0

Explanation of the Inference Mechanism:

```
IfNot Cooling Allowed [1]
And Cooling Enabled [1]
ThenInf Warning Wastefull Cooling [1]
Then Diagnose Found [*]

If Optimal Damper Position = Modulating [0]
ThenNot Heating Allowed [*]
```

```
ThenNot Cooling Allowed [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [*]
ThenNot Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]
And Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [0]
Then Optimal Damper Position = Minimal Open [*]

If Optimal Damper Position = Minimal Open [*]
And Estimated Minimal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

IfNot Optimal Damper Position = Modulating [1]      .
AndNot Estimated Minimal Open Mixing Temperature = Nearest to Tsetpoint [1]
Then Optimal Damper Position = Maximal Open [1]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature < Tsetpoint [1]
Then Heating Allowed [1]
ThenNot Cooling Allowed [0]

If Optimal Damper Position = Maximal Open [1]
And Estimated Maximal Open Mixing Temperature > Tsetpoint [0]
ThenNot Heating Allowed [*]
Then Cooling Allowed [*]

If Estimated Maximal Open Mixing Temperature = Tsetpoint [0]
ThenNot Heating Allowed [*]
ThenNot Cooling Allowed [*]
```

   Simulation results:
   TimeStamp: 20820

```
Tatm ..................... :  24.71
Tretn .................... :  19.86
Tmix ..................... :  24.36
Tsup ..................... :  24.26
Tset ..................... :  24.29
Vdamp .................... :   0.97  (Input of the Error type)
Verr ..................... :   1.00  (Output of the Error type)
Dheat .................... :   0.00
Dcool .................... :   0.33
Chiller_on ............... :   1.00
Boiler_On ................ :   0.00
```

Estimated values:

```
Estimated Tmix ........... :  24.10
Est.Tmix (Vdamp 0.19) .... :  20.39
Est.Tmix (Vdamp 1.00) .... :  24.27
```

Conclusion of the Inference Mechanism:

Error Damper stuck in Maximal Open position : 1.0

Explanation of the Inference Mechanism:

```
If Mixing Temperature = Maximal Open Mixing Temperature [1]
AndNot Control Signal of Damper = Maximal Open [1]
ThenInf Error Damper stuck in Maximal Open position [1]
Then Diagnose Found [*]
```

Minimizing Operating Costs of Vapor Compression Equipment
with Optimal Service Scheduling

Todd M. Rossi and James E. Braun

Ray W. Herrick Laboratories, Purdue University

West Lafayette, IN 47907-1077, USA.

ABSTRACT

Dynamic programming was used to obtain optimal service schedules and costs for cleaning the condensers and evaporators of air conditioning equipment. Results were obtained for a range of service and energy costs, characteristic fouling times, and equipment sizes for a single building and location. Minimum operating costs were compared with regular service intervals (representative of current practice) and a strategy where service is only performed when a constraint is violated (e.g., a comfort violation). It was found that optimal service scheduling reduced lifetime operating costs by as much as a factor of two over regular service intervals and 50% when compared to constrained only service. For practical implementation, a simple near-optimal algorithm for estimating optimal service scheduling was developed that does not require on-line forecasting or numerical optimization and is easily implemented within a micro-controller. Over a wide range of cases tested, the near-optimal algorithm gave operating costs that were within 1% of optimal. This technique could also be applied to other systems where performance degradations are important such as large chillers and power plants.

NOMENCLATURE

Ce	=	*Energy cost ($/kWh)*
C_s	=	*Service cost ($)*
G	=	*Mass flux ($kg/m^2/s$)*
$H(t)$	=	$\int_0^t \bar{h}(t_r)dt_r$ *(kWh)*
J_0	=	*Cost function with fewest assumptions ($/year)*
$J_1 = J_0/C_e$	=	*Simplified cost function (kWh/year)*
$J_2 = J_1 * T_c$	=	*Dynamic programming cost function (kWh)*

J_3	=	Simplified cost function assuming the time between service tasks is fixed (kW)
J_4	=	Simplified cost function subtracting out the expected power consumption with no fouling (kW)
K_l^k	=	Energy consumed starting at time stage l after servicing and operating for k time stages (kWh)
N	=	Number of service tasks during equipment lifetime (dimensionless)
N_c	=	Number of service tasks during service cycle (dimensionless)
N_r	=	Hours of fan operation that would occur each year if there were no fouling (hours/year)
N_y	=	Number of hours per year (hours/year)
P	=	Equipment power consumption rate (kW)
PLF	=	Part-load fraction (dimensionless)
RC	=	Low pass filter time constant (h)
T	=	Threshold value that becomes negative to indicate when the simplified scheduler calls for service (kWh)
$T_c = n_c/n_s$	=	Duration of service cycle (years)
T_l	=	Equipment lifetime (years)
c_i	=	Dynamic programming cost for the ith decision stage (kWh)
f	=	Fouling state (dimensionless)
f^*	=	Fouling state immediately after service (dimensionless)
g_i	=	Dynamic programming cost-to-go at the ith decision stage (kWh)
$h(t)$	=	Extra power consumed because of fouling (kW)
$\bar{h}(t)$	=	$h(t)$ after low pass filter (kW)
n_c	=	Number of service opportunities per cycle (dimensionless)
$n_s = 1/\Delta t$	=	Number of service opportunities per year (1/year)
t	=	Time(h)
$\bar{t} = \{t_1, t_2, \ldots, t_N\}$	=	Set of N times service is done during equipment lifetime (h)
t_c	=	Calendar time (h)
t_f	=	Heat exchanger characteristic fouling time (h)
t_f^*	=	Normalized characteristic fouling time for heat exchangers (years)
t_i	=	Time of the ith service task (h)
t_r	=	Accumulated runtime since last service (h)
t_0	=	Runtime between service (h)
$\bar{x}(t) = \{x_0, x_1, x_2\}$	=	Equipment driving conditions
x_0	=	Ambient temperature (C)
x_1	=	Return air temperature (C)
x_2	=	Return air humidity ratio (kg/kg)

Greek Symbols

$$\Delta t = \frac{1}{n_s}$$ = *Minimum service interval (years)*

Δt_r = *Runtime during each time step (h)*

$\delta(\cdot)$ = *Dirac delta function (dimensionless)*

$\bar{\delta\tau}$ = *Number of time stages between service tasks (dimensionless)*

$\bar{\tau} = \{\tau_1, \tau_2, \ldots, \tau_N\}$ = *Dimensionless service schedule (dimensionless)*

$$\tau_i = \frac{t_i}{\Delta t \cdot N_y}$$ = *Time stage when the ith service is performed (dimensionless)*

Subscripts

ca   = *condenser air*

ea   = *evaporator air*

Superscript

max   = *maximum value*

min   = *minimum value*

INTRODUCTION

Automated fault detection and diagnostics (FDD) can improve reliability, reduce operating costs, and increase equipment safety in many applications. Most of the previous research and development of FDD techniques has focused on critical processes, such as aircraft engines [Onken & Stuckenberg 1979, Patton, Willcox & Winter 1987], nuclear power plants [Kitamura 1980], and the space shuttle main engine [Cikanek 1986, Duyar & Merill 1992]. However, the costs of applying FDD techniques are going down with reduced hardware and software costs and techniques are beginning to be applied to heating, ventilating, air conditioning, and refrigeration (HVAC&R) applications.

Many authors [Willsky 1976, Isermann 1984, Frank 1987, Basseville 1988, Gertler 1988, Frank 1990] have offered excellent review papers on fault detection and diagnostic techniques. Isermann [1984] presents the application of FDD techniques as a series of four steps termed "process supervision" that are illustrated in Figure 1. The first step is fault detection, in which a fault is indicated when the performance of a monitored system has deviated from expectation. The second step, diagnosis, determines which malfunctioning component is causing the fault. Following diagnosis, fault evaluation assesses the impact of

[McKellar 1987, Wagner & Shoureshi 1988, Wagner & Shoureshi 1992], and other vapor et al. 1985, Norford & Little 1993, Georgescu, Afshari & Bomard 1993b], refrigerators 1990, Pape, Mitchell & Beckman 1991, Georgescu et al. 1993a], air handling units [Usoro Graves, Reinert, Kreider, Dow & Wubbena 1989, Culp, Haberl, Norford, Brothers & Hall for building envelopes [Usoro & Schick 1985, Norford, Rabl & Spadaro 1987, Anderson, process supervision for HVAC&R systems. In particular, FDD techniques have been studied development and evaluation of techniques for the fault detection and diagnostic steps of There have been several papers published within the last ten years that describe the

Figure 1.   Supervision of HVAC&R equipment

the fault on system performance. Finally, a decision is made on how to react to the fault. This is usually a choice between tolerating the fault, repairing it as soon as possible, adapting the control, and stopping operation until repair is complete.

compression equipment [Kaler 1988, Culp 1989, Yoshimura & Noboru 1989, Kaler 1990, Kumamaru, Utsunomiya, Yamada, Iwasaki, Shoda & Obayashi 1991, Hiroshi, Matsuo, Fujiwara Yamada & Nishizawa 1992].

There appears to be no literature related to fault evaluation for HVAC&R systems. Fault evaluation is particularly important when the performance of a component is degrading slowly over time, such as occurs for heat exchanger fouling. In this case, it is possible to detect a fault well before it is severe enough to justify the service expense. In contrast, abrupt failures such as broken belts would not require the evaluation step when it is obvious that the fault should be repaired (e.g., the system no longer maintains comfort).

In principle, fault evaluation could be achieved by minimizing lifetime operating costs. An optimal service scheduler would purchase service when it contributes to reducing overall costs. Important costs to consider are: maintenance, energy, equipment down time (i.e. the cost of not maintaining comfort or refrigeration set point), premature component wear (avoidable service costs), and liability costs associated with injury to people, damaged facilities, or pollution of the environment. With the exception of maintenance and energy, the other costs are difficult to quantify. The optimization problem could be simplified by assuming that the comfort, premature wear, and liability costs are much larger than the service cost to repair them. This is equivalent to assigning an infinite economic penalty for these conditions (i.e. treating them as constraints) and minimizing the combined costs of energy and service. With these considerations in mind, the following four fault evaluation criteria are proposed for HVAC&R systems.

1. ECONOMIC CRITERIA - Service is required when it contributes to the reduction of the combined costs of energy and service over the lifetime of the unit.

2. COMFORT or REFRIGERATION SET POINT CRITERIA - Service is required when the equipment is not capable of maintaining building comfort (air conditioning applications) or a refrigeration set point (refrigeration applications).

3. SAFETY CRITERIA - Service is required when the operating state of the equipment could lead to damage or injury (e.g. liquid entering a compressor or head pressure above the tube burst pressure).

4. ENVIRONMENTAL CRITERIA - Service is required when the equipment is polluting the environment. This criteria is included for refrigerant leaks.

This paper presents results associated with maintenance scheduling that were determined using the fault evaluation criteria described above. The goals of the research were to evaluate the maximum cost savings associated with using optimal service scheduling for the cleaning of heat exchangers in air conditioning equipment and to develop a more practical, near-optimal scheduling technique. In order to evaluate the maximum savings potential, the simulated lifetime costs for service and energy associated with optimal maintenance scheduling were compared with results for regular service intervals and service dictated by the constraints only. Service at regular time intervals (e.g., 1 year) is characteristic of today's practice. Presumably, fixed service intervals are chosen based upon consideration of both energy and service costs. On the other hand, constrained only service represents a limiting case where the economic criteria for service, as described above, is ignored. The next three sections describe the development and application of the optimal, regular, constrained only, and near-optimal maintenance schedulers followed by a description of the system simulation and results of the comparisons.

OPTIMAL MAINTENANCE SCHEDULER

Neglecting inflation, the average annual cost of energy and service for vapor compression equipment can be expressed as

$$J_0 = \frac{1}{T_l} \int_{t_1}^{t_1 + T_l N_y} \left[C_e P(\bar{x}(t), f(t)) + C_s \sum_{i=1}^{N} \delta(t - t_i) \right] dt \qquad (1)$$

where T_l is the equipment lifetime, C_e is the constant cost of one kWh of energy, C_s is the constant cost of service to repair the fault, $\bar{t} = \{t_1, t_2, \ldots, t_N\}$ is the set of service times, N is the number of service tasks performed during the equipment lifetime, $\delta(\cdot)$ is the Dirac delta function, P is the instantaneous power consumption of the equipment, $\bar{x}(t)$ represents the natural driving functions controlling power consumption including ambient and load conditions, and $f(t)$ is the state of the degrading equipment components (e.g. fouling heat

exchangers and leaking compressor valves). The integral can be divided into the sum of integrals over the periods between service tasks to give

$$J_0 = \frac{1}{T_l}\left[C_e \sum_{i=1}^{N} \int_{t_i}^{t_{i+1}} P(\bar{x}(t), f(t))dt + NC_s \right] \qquad (2)$$

where $t_{N+1} = t_1 + T_l \cdot N_y$.

In the following development, the cost function of equation 2 is modified in order to simplify the numerical solution for the optimal maintenance schedule. First of all, the cost function is divided by C_e to get a cost function that gives the same optimal solution, but only depends upon the ratio of service to energy costs (C_s/C_e).

$$J_1 = \frac{J_0}{C_e} = \frac{1}{T_l}\left[\sum_{i=1}^{N} \int_{t_i}^{t_{i+1}} P(\bar{x}(t), f(t))dt + N\frac{C_s}{C_e} \right] \qquad (3)$$

The optimization problem is simplified by limiting the available opportunities to perform service to periodic instances of time (e.g. monthly), thereby restricting the size of the solution space. With this assumption, the integral portion of the cost function can only take a limited number of values that can be tabulated to eliminate repetitive calculations during the solution of the optimization problem. Each entry in the table is expressed as

$$K_{\tau_i}^{\tau_{i+1}-\tau_i} = \int_{t_i}^{t_{i+1}} P(\bar{x}(t), f(t))dt = \int_{\tau_i/n_s}^{\tau_{i+1}/n_s} P(\bar{x}(t), f(t))dt \qquad (4)$$

where n_s is the maximum number of potential service tasks that can be performed per year (e.g. $n_s = 12$ services/year for monthly service opportunities) and τ_i is the number of the time stage (each of duration $\Delta t = 1/n_s$ = minimum service interval) when the ith service task is performed. The subscript for K is the time stage number when equipment operation begins after a service task and the superscript is the number of time stages the unit runs before service is needed again. The simplified cost function is

$$J_1 = \frac{1}{T_l}\left[\sum_{i=1}^{N} K_{\tau_i}^{\tau_{i+1}-\tau_i} + N\frac{C_s}{C_e} \right] \qquad (5)$$

In order to eliminate the need to integrate out to potentially long and unknown equipment lifetimes, it is assumed that the service schedule eventually settles into a cycle with a period of n_c time stages (e.g. $n_c = 36$ for a 3.0 year cycle with $n_s = 12$ services/year). For a cycle to occur, the driving conditions for the cooling equipment (e.g. weather patterns, occupancy schedule, etc.) must be periodic. The length of the service cycle is $T_c = n_c/n_s$ and the total number of services in one service cycle is defined as N_c. In order to minimize the average operating cost, it is only necessary to integrate over one period of the cycle. This bounds the size of the table to n_c^2 elements $\left(K_l^k : 1 \leq k \leq n_c, \ 1 \leq l \leq n_c \right)$. The service schedule for one cycle is defined by $\bar{\tau}_c = \{\tau_1, \tau_2, \ldots, \tau_{N_c}\}$ and can also be represented by τ_l (time stage of the first service task in the schedule) and the set of the number of time stages between service tasks $\bar{\delta\tau} = \{\delta\tau_1, \delta\tau_2, \ldots, \delta\tau_{N_c-1}\}$, where $\delta\tau_i = \tau_{i+1} - \tau_i$ and $\delta\tau_{N_c} = n_c - \sum_{i=1}^{N_c-1} \delta\tau_i$. The average cost for the service cycle can then be expressed as

$$J_1 = \frac{1}{T_c}\left[\sum_{i=1}^{N_c} K_{\tau_i}^{\delta\tau_i} + N_c \frac{C_s}{C_e} \right] \tag{6}$$

The analysis can be further simplified by assuming that the longest period for the natural driving conditions $(\bar{x}(t))$ is one year. In this study, typical meteorological year (TMY) [Hall, Prairie, Anderson & Boes 1978] weather data were used in the simulations and the typical year's weather pattern reoccurred each year. With this simplification, the table only requires $n_s \cdot n_c$ elements $\left(K_l^k : 1 \leq l \leq n_s, \ 1 \leq k \leq n_c \right)$ and the duration of the service cycle (T_c) is an integer multiple of one year.

Given a model for determining $P(\bar{x}(t), f(t))$ (needed to calculate $K_l^k : 1 \leq l \leq n_s, \ 1 \leq k \leq n_c$), the cost ratio C_s/C_e, and the minimum service interval Δt, the optimization problem can be defined as minimizing J_1 with respect to $\{T_c, N_c, \tau_l\}$ and the number of time stages between service tasks $\left(\bar{\delta\tau}\right)$. This is an $N_c + 2$ dimensional optimization problem. The minimum service interval (Δt) can be the minimum reaction time

678

of the service organization or a sufficiently small quantity such that the minimum cost no longer depends on its value. In the later case, Δt would be decreased until the minimum cost became insensitive to Δt. For the cleaning of heat exchangers, a minimum service interval of 1 month is a realistic and sufficiently accurate interval for scheduling service that was utilized in this study.

The optimal maintenance scheduler minimizes the lifetime costs of energy and service while maintaining comfort, safety, and environmental protection as constraints. The comfort, safety, and environmental protection constraints are maintained by adding artificially high costs to the tabulated energy usage (K_l^k) for operating conditions that would result in constraint violations.

The numerical solution to the optimization problem is accomplished using a combination of two numerical techniques. Given values of $\{T_c,\ N_c,\ \tau_l\}$, the optimal set of time stages between service tasks $\bar{\delta\tau}^* = \{\delta\tau_1^*, \delta\tau_2^*, \ldots, \delta\tau_{N_c-1}^*\}$ is determined using dynamic programming [Bellman 1957]. An outer loop, containing the dynamic programming solution, is then used to find the optimal values of the three quantities: $\left\{T_c^*, N_c^*, \tau_1^*\right\}$.

Dynamic Programming

The use of dynamic programming for solving an optimal equipment replacement problem is discussed in Jardine [1973, section 4.4]. Dynamic programming determines a global minimum in an efficient manner for this type of problem by taking advantage of the fact that the number of possible solutions is restricted by the sequential nature of the decision process. A good description of dynamic programming can be found in Rao [1984].

Figure 2 is a black box diagram illustrating the sequence of decisions required to solve for the optimum placement of N_c service tasks among n_c time stages, starting at time stage τ_l. There are N_c-l decisions that have to be made and each one is referred to as a decision stage. The state information passed between decision stages is the time stage that service was last performed (τ_i, for the ith stage). The decision variable is the number of time stages to wait until the next service ($\partial\tau_i$, for the ith stage). The cost of each decision stage (c_i), except the

last, is the energy used while running the unit for $\partial \tau_i$ time stages starting at τ_i plus the service cost divided by the energy cost. The cost of the last decision stage $(c_{N_c\text{-}1})$ also includes the energy consumed while running the unit back to the first decision stage plus the cost of the first service task normalized by the energy cost. Finally, the output of each stage is the time of the last service plus the run time for that decision stage ($\tau_{i+1} = \tau_i + \partial \tau_i$). The sum of the costs of the decision stages is given by

$$J_2 = T_c J_1 = \left[\sum_{i=1}^{N_c} K_{\tau_i}^{\partial \tau_i} + N_c \frac{C_s}{C_e} \right] = \sum_{i=1}^{N_c - 1} c_i \tag{7}$$

$$c_1 = K_{\tau_1}^{\delta \tau_1} + \frac{C_s}{C_e} \qquad c_2 = K_{\tau_2}^{\delta \tau_2} + \frac{C_s}{C_e} \qquad c_{N_c - 1} = K_{\tau_{N_c - 1}}^{\delta \tau_{N_c - 1}} + K_{\tau_{N_c}}^{n_c - \sum_{i=1}^{N_c - 1} \delta \tau_i} + 2 * \frac{C_s}{C_e}$$
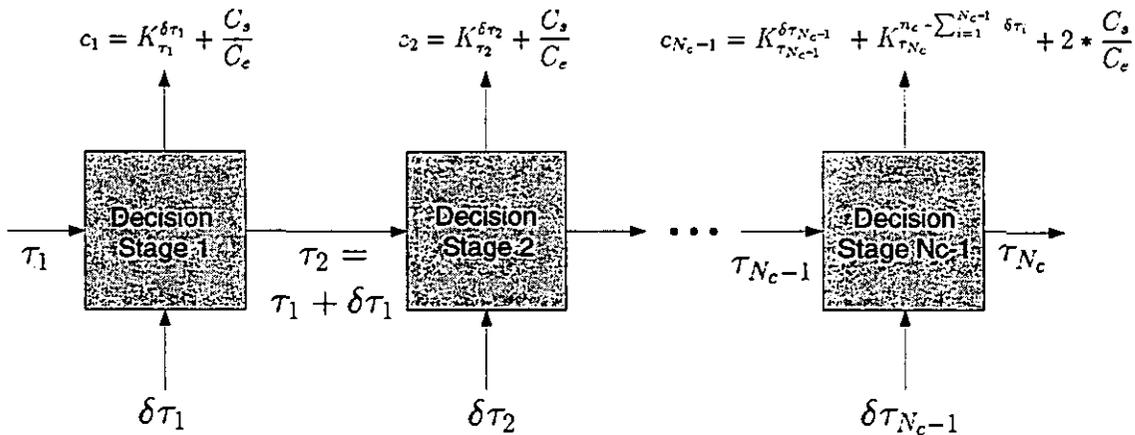
Figure 2.   Dynamic programming places N_c-1 service tasks to minimize J_1.

Dynamic programming is appropriate for this optimization problem because the performance of the vapor compression equipment does not depend on the chain of events that led to the most recent servicing. Figure 2 illustrates this point in that the decision to do service only requires knowledge of the time of the previous service task (represented by the state variable passed between decision stages). This property of the optimal servicing problem guarantees an optimal solution without the need for testing all possible service schedules.

Figure 3 illustrates an example dynamic programming problem for optimal maintenance scheduling. The abscissa contains the decisions stages and the ordinate contains the time stages when service can be done. In this example, there are $n_c = 12$ time stages in a $T_c = 1$ year service cycle (i.e. service opportunities are at monthly intervals and the cycle repeats

each year). There are $N_c = 4$ service tasks to place and the time stage of the first task is specified as $\tau_1 = 3$. The open circles indicate the available opportunities for doing service at each decision stage. Not all values are allowed for each stage because τ_{i+1} is restricted by definition to follow τ_i. This reduces the number of possible decisions to test reducing computation time. In general, the number of available time stages at each decision stage is n_c-N_c+1.
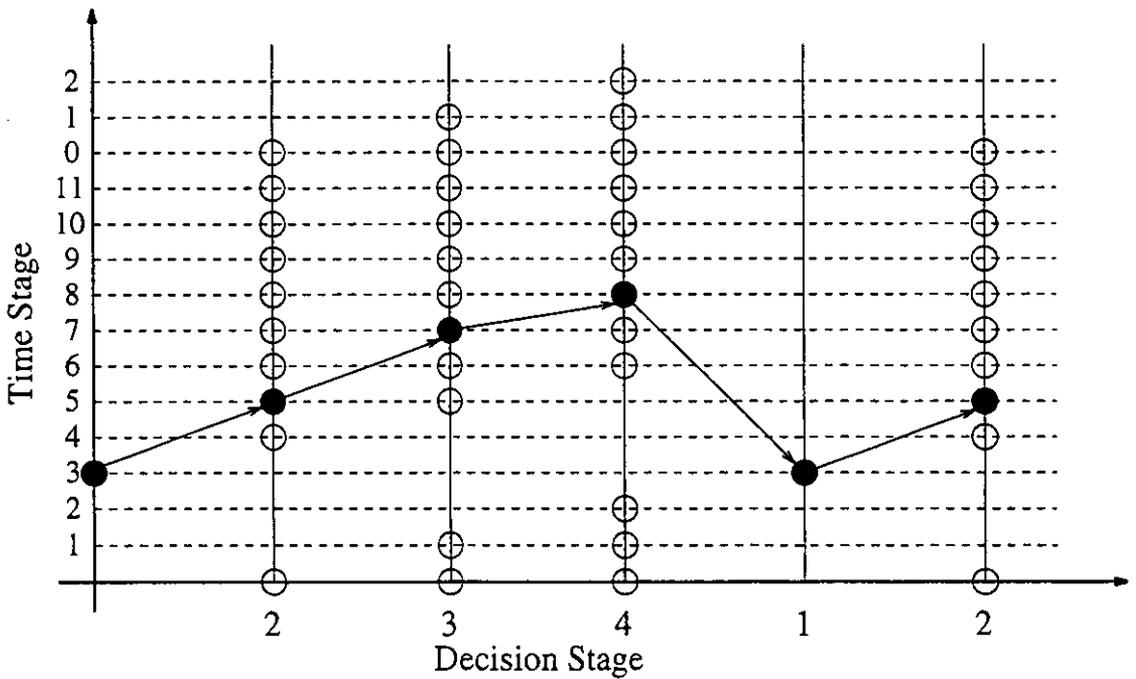
Figure 3.   Optimal trajectory through domain for an example problem

The example problem of Figure 3 is solved by starting at decision stage 1 and progressing backwards one decision at a time until decision stage 1 is reached again and the cycle repeats itself. Beginning at decision stage 1 (where service is performed at time stage 3), the costs of getting there from all 9 allowed time stages in decision stage 4 are computed and stored. Next, all possible time stages in decision stage 3 are considered. For each of these 9 time stages, the costs of moving to each possible time stage in decision stage 4 are computed. Then, the minimum costs-to-go from each of the 9 time stages at decision stage 3 to decision stage 1 are computed and stored with the associated path. This process is repeated until decision stage 1 is reached.

At each decision stage, a catalog of the minimum cost-to-go and the corresponding optimal trajectory to the end of the cycle is generated. The path associated with the filled circles connected by arrows in Figure 3 is the optimal trajectory for this example. In mathematical terms, the solution to the ith decision stage is

$$g_i^*(\tau_i) = \min_{\delta\tau_i} \left[c_i(\tau_i, \delta\tau_i) + g_{i+1}^*(\tau_i + \delta\tau_i) \right] \quad for \ 1 \le i \le N_c - 1 \tag{8}$$

where $g_{N_c}^* = 0$. The operator (min) refers to a one-dimensional optimization with respect to $\delta\tau_i$ (the number of time stages to wait before doing service again), while g_i^* is the minimum cost-to-go at the ith decision stage. The process begins at the last stage ($i = N_c\text{-}1$) and continues recursively until the first stage ($i = 1$).

Outer Loop Optimization

The minimization of J_1 with respect to its $N_c + 2$ independent variables is done using four nested optimization loops. The inner most loop solves the dynamic programming problem. The next loop solves for N_c^* using the golden section method [Rao 1984] and the outer most loops exhaustively search values of T_c and τ_1. An exhaustive search is used for determining T_c and τ_1 because J_1 is not guaranteed to be unimodal with respect to changes in these variables. All n_s possible values of τ_1 (e.g. 12 for a one month minimum service interval) and all values of T_c up to a practical limit (e.g. 12 years) that provide unique solutions are tested. It is not necessary to evaluate all values of T_c in order to test all service cycles. For example, a 12-year service cycle contains 1, 2, 3, 4, & 6 year cycles. Therefore, all cycles up to 12 years can be tested by considering T_c = 7, 8, 9, 10, 11, & 12 years. A non-exhaustive search can be used for N_c because the dependence of J_1 on N_c is unimodal for T_c and τ_1 fixed and all the service tasks optimally placed. Values of N_c greater than optimal will have increasingly higher cost due to excessive service and N_c less than optimal will have increasingly higher cost due to excessive energy consumption.

REGULAR AND CONSTRAINED SERVICE SCHEDULES

With regular maintenance scheduling, preventive maintenance (coil cleaning or filter changing) is performed at regular time intervals (e.g. annually, quarterly, or monthly). When the chosen interval is not sufficient to maintain the constraints on comfort and safety, then the costs associated with the schedule are set artificially high. The constrained service scheduler only performs service when a comfort or equipment safety constraint is violated. Constrained service is simpler to implement than optimal scheduling and could be considered as a viable alternative to regular service scheduling.

Costs are compared by contrasting total annual costs for the different maintenance schedulers. All schedulers use the tabulated values of K_i^k, but manipulate the values of the $N_c + 2$ service variables differently. In the case of regular service, the times for service are prescribed by the schedule. Service is performed just before the cooling season begins in the once per season schedule and an additional service is added in the middle of the cooling season in the twice per season service schedule. For constrained service, the schedule begins service on January 1 of the first year. Service is performed as needed for as many years as required to achieve a natural service cycle. A cycle occurs when service is performed at the same time segment in different years.

SIMPLIFIED NEAR-OPTIMAL SERVICE SCHEDULE

The simplified near-optimal scheduler captures the important features of the optimal scheduler with significantly reduced computational requirements and without the need for weather and load forecasters. The derivation of the simplified scheduler begins with the complete cost function given by equation 3.

Equation 3 can be approximated by fixing the time between service tasks ($t_0 \equiv t_{i+1} - t_i$, for all $1 \le i \le N$) and assuming that the cost is only a function of the runtime since the last service.

$$J_3 = \frac{N}{T_l}\left[\int_0^{t_0} P(\bar{x}(t), f(t))dt + \frac{C_s}{C_e}\right] = \frac{1}{t_0}\left[\int_0^{t_0} P(\bar{x}(t), f(t))dt + \frac{C_s}{C_e}\right] \tag{9}$$

where $T_i = N \cdot t_0$ and t_o is the equipment runtime between service tasks. The use of equipment runtime instead of clock time implies that the state of fouling is only a function of runtime after cleaning.

If J_l were independent of $\bar{x}(t)$, and f always had the same time dependence after service, then the assumption of a fixed service interval would be exact (i.e. $J_3 = J_l$). However, since J_l depends on seasonal changes in $\bar{x}(t)$, then the assumption of a fixed service interval improves when seasonal effects are reduced by subtracting the cost associated with no performance degradations. With this in mind, a new cost function is defined as

$$J_4 = \frac{1}{t_0}\left[\int_0^{t_0}\left[P(\bar{x}(t),f(t))-P(\bar{x}(t),f^*)\right]dt+\frac{C_s}{C_e}\right].\tag{10}$$

where f^* is defined as the value of f immediately after service. This cost function differs from J_3 by the baseline energy consumption of the unit during the time between service tasks with no fouling $\left(\frac{1}{t_0}\int_0^{t_0}P(\bar{x}(t),f^*)\right)$. The integral portion of equation (10) is the extra energy consumed in one service period due to degrading component performance and C_s/C_e is the amount of energy that could be purchased with the cost of one service task.

The service interval determined by minimizing J_4 is used as an approximation to the optimal solution that minimizes J_0. This approximation is less accurate when $\bar{x}(t)$ changes appreciably from one service cycle to the next. In general, performance is good when $t_0 \geq 1$ year, because each service cycle experiences similar weather conditions. It will be shown later that subtracting the power consumption with no fouling also provides a mechanism for considering multiple simultaneous faults.

The cost function J_4 can be minimized by determining where its derivative with respect to time is zero, providing the following solution:

$$\int_0^{t_0}h(t)dt+\frac{C_s}{C_e}=t_0\cdot h(t_0)\tag{11}$$

where

$$h(t) = P(\bar{x}(t), f) - P(\bar{x}(t), f^*) \tag{12}$$

is the extra power required to provide the necessary cooling due to the performance degradation. The solution of equation 11 for t_o is guaranteed to be a minimum when $h(t)$ is an increasing function. In order to filter diurnal fluctuations in $h(t)$, a low pass filter for $h(t)$ is used.

$$\bar{h}(t) = \int_0^t \frac{h(t') - \bar{h}(t')}{RC} dt' \tag{13}$$

The time constant (RC) should be much longer than one day and much shorter than the time between service tasks. For the simulation results in this paper, RC was set to 200 hours. Substituting $\bar{h}(t)$ into equation 11 results a classification rule used to evaluate the need for service.

$$\frac{C_s}{C_e} + \int_0^{t_0} \bar{h}(t)dt - t_0 \cdot \bar{h}(t_0) \underset{\omega_2}{\overset{\omega_1}{\underset{>}{\lessgtr}}} 0. \tag{14}$$

where ω_1 is the class "do service" and ω_2 is the class "no service". The runtime between service tasks (t_0) is determined with equation 14 as the time when the left hand side makes a transition from a positive to a negative quantity. The evaluation of this rule is based on information acquired since the last service only and requires no forecasting.

The following is a step-by-step procedure for using the simplified scheduler.

1. Identify the cost of energy and the cost of performing the service task to repair the performance degradation fault that is reducing efficiency.

2. Learn a model that tracks the expected power consumption with no performance degradation as a function of the measured driving conditions $\left(P(\bar{x}(t), f^*)\right)$.

3. Initialize three accumulator variables to zero. Integrated variables are $H(t) = \int_0^t h(t)dt$, $\bar{h}(t)$, and the accumulated runtime t_r.

4. Measure power consumption, driving conditions (e.g. ambient temperature), constraints (e.g. room temperature), and runtime for the current time step (Δt_r).

5. Update the accumulators using the following equations:

$$h(t) = P(\bar{x}(t), f) - P(\bar{x}(t), f^*)$$

$$\bar{h}(t) = \bar{h}(t) + \frac{h(t) - \bar{h}(t)}{RC} \Delta t_r$$

$$H(t) = H(t) + \bar{h}(t) \cdot \Delta t_r$$

$$t_r = t_r + \Delta t_r$$

$$T = \frac{C_s}{C_e} + H(t) - t_r \cdot \bar{h}(t)$$

where T is the threshold value used to decide if service is required or not and RC is the time constant for the low pass filter designed to average over diurnal variations in $h(t)$.

6. Compare T with 0 and evaluate the constraints. If $T<0$ or a constraint is violated, then perform the service task and reset the three accumulator variables.

7. Wait until the next time step and then return to step 4.

The procedure described in this section considers only one fault. However, a "single fault" could actually be a collection of faults that are always serviced together (e.g. condenser fouling and filter changing). With this assumption, only one decision is necessary at any time: whether to perform service or not. In order to consider multiple faults with different service times, it would be necessary to evaluate the impact of each combination of service tasks on the overall costs. It is possible to extend the simplified service scheduler for this case.

Consider the case where two performance degradations are developing at the same time. Define f_1 and f_2 as the fault states for the first and second degrading components, respectively (e.g. evaporator and condenser fouling). The degradation states are indicators of the severity

of each fault being considered. Each state tracks the evolution of only one fault to provide a means of determining the contribution of each fault to the overall degradation in performance. A degradation state is a general term that may be a single measurement or an estimate from many measurements. For example, the air pressure drop across a heat exchanger of filter is commonly used in buildings as an indicator of fouling. Another possibility is an estimate of the overall thermal conductance of a heat exchanger from many temperature and other measurements.

There are three rules that must be evaluated, along with the comfort constraint, to determine when to do service. They are given by

$$\frac{C_{si}}{C_e} + \int_0^{t_0} h_i(t)dt - t_0 \cdot h_i(t_0) \overset{\omega_1}{\underset{\omega_2}{\lessgtr}} 0 \tag{15}$$

for $i = 1, 2, \& 3$. C_{s1} is the cost of servicing component 1 only and

$$h_1(t) = P(\vec{x}(t), f_1, f_2) - P(\vec{x}(t), f_1^*, f_2). \tag{16}$$

C_{s2} is the cost of servicing component 2 only and

$$h_2(t) = P(\vec{x}(t), f_1, f_2) - P(\vec{x}(t), f_1, f_2^*). \tag{17}$$

Finally, C_{s3} is the cost of servicing components 1 & 2 together and

$$h_3(t) = P(\vec{x}(t), f_1, f_2) - P(\vec{x}(t), f_1^*, f_2^*). \tag{18}$$

In many cases, it may be possible to service both performance degradations at the same time for less money than both separately. Service is done when any of these rules fire. The results presented in this paper are only for a single fault developing at one time (evaporator or condenser fouling). More work would be necessary to determine approximate measures that

characterize degradations states in order to implement the procedure for multiple faults with different service times.

SYSTEM MODELING

Simulations were used to demonstrate the potential for optimal service scheduling and to evaluate the performance of the simplified near-optimal service scheduler. This section describes the models and the conditions considered.

Building Model

Building cooling requirements were determined using a TRNSYS simulation [Klein et al. 1990] of a three-zone office building in Nashville, TN, USA. The simulation model is a detailed representation of a building (TYPE 56 within TRNSYS) that considers coupling to the ambient (temperature and humidity), internal and solar gains, internal radiation and convective exchange, building mass, day and night occupancy schedules, and night setback control. The simulation model generated cooling requirements for every hour of a typical year using typical meteorological year (TMY) weather data [Hall et al. 1978], Figure 4 shows the average daily cooling load for the entire year and the hourly cooling load for a one-week period surrounding the peak cooling load day. The dashed lines enclose the cooling season during which the evaporator fans ran continuously during occupied periods for ventilation.

(a)   (b)
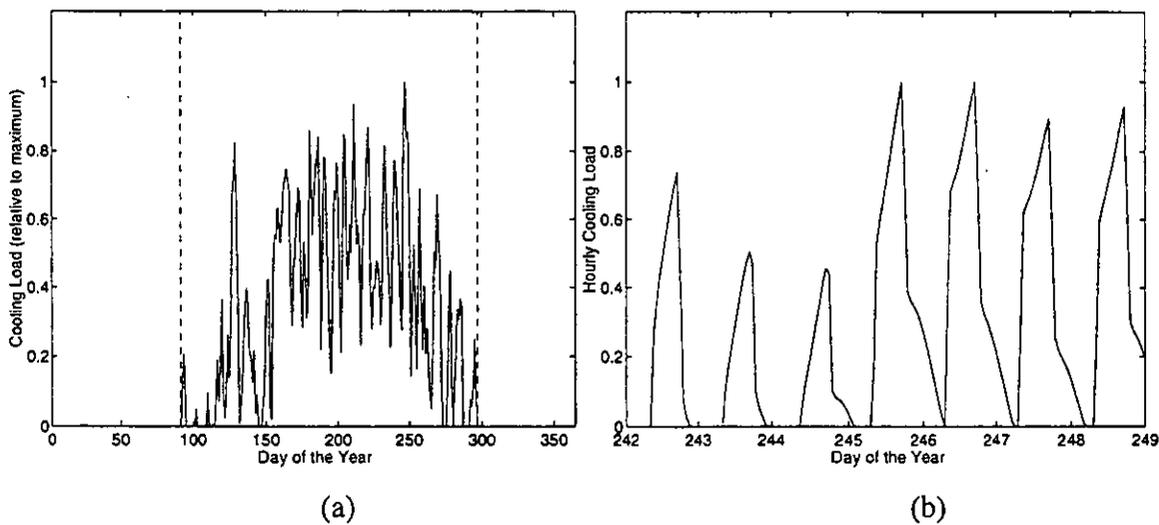
Figure 4.   Cooling load schedule for demonstration building.

The use of the building model was simplified by decoupling the building from the air conditioning unit. TRNSYS provided the sensible and latent cooling required to maintain fixed temperature and humidity set points for the building (24.4°C (76°F) and 60% RH). In the system simulations, the air conditioning unit was required to remove the total heat load whenever the sensible cooling requirement was greater than zero.

Air Conditioning Unit Model

The building was cooled by an air-to-air vapor compression air conditioning system using on/off control. A detailed physical model was used to predict the performance of a commercial rooftop air conditioning system [Rossi 1995]. The model provided the power consumption and total cooling capacity as a function of ambient temperature, evaporator filter fouling state, and condenser fouling state for return air conditions of 24.4°C (76°F) and 60% relative humidity. For simplicity, fouling was modeled only as an obstruction to air flow across the coils. The state of fouling is defined as the difference between the maximum and actual air mass flow rates divided by the maximum mass flow rates (0=no fouling and 1=completely obstructed flow):

$$f_{ea} = \frac{G_{ea}^{max} - G_{ea}}{G_{ea}^{max}}$$

$$f_{ca} = \frac{G_{ca}^{max} - G_{ca}}{G_{ca}^{max}}$$

The symbols f_{ea} and f_{ca} are the "fouling states" of the evaporator and condenser coils, respectively, G_{ea} and G_{ca} are the air mass fluxes through each coil, and G_{ea}^{max} and G_{ca}^{max} are the maximum values (unobstructed flow) of the respective air flow rates. For this demonstration system, $G_{ea}^{max} = 2.85$ kg/s/m^2 and $G_{ca}^{max} = 4.80$ kg/s/m^2, which corresponds to volume flow rates of 0.57 m^3/s (1200 cfm) and 1.65 m^3/s (3500 cfm).

A simulation was run over a uniformly spaced three-dimensional grid of ambient temperature ($10°C \leq T_{amb} \leq 35°C$), outside air mass flux ($0 \leq G_{ca} \leq 4.80$ kg/s/m^2), and inside air mass flux ($0 \leq G_{ea} \leq 2.85$ kg/s/m^2) and results for power consumption and total capacity

were tabulated. The lookup table was used by the service scheduling program to determine the amount of energy used to support the prescribed cooling load under the specified conditions and fouling states and to provide an upper limit on the cooling capacity in order to evaluate the comfort constraint.

Figure 5 shows example predictions of total capacity and power consumption predictions as a function evaporator and condenser air flow rates. When the evaporator air flow rate decreases, capacity begins to fall immediately and power consumption follows. When condenser air flow rate decreases, capacity remains relatively constant because increases in head pressure have a relatively small effect on evaporating conditions. However, the high head pressure does result in higher power consumption. Eventually, capacity suffers at extremely low condenser air flow rates.

(a)   (b)
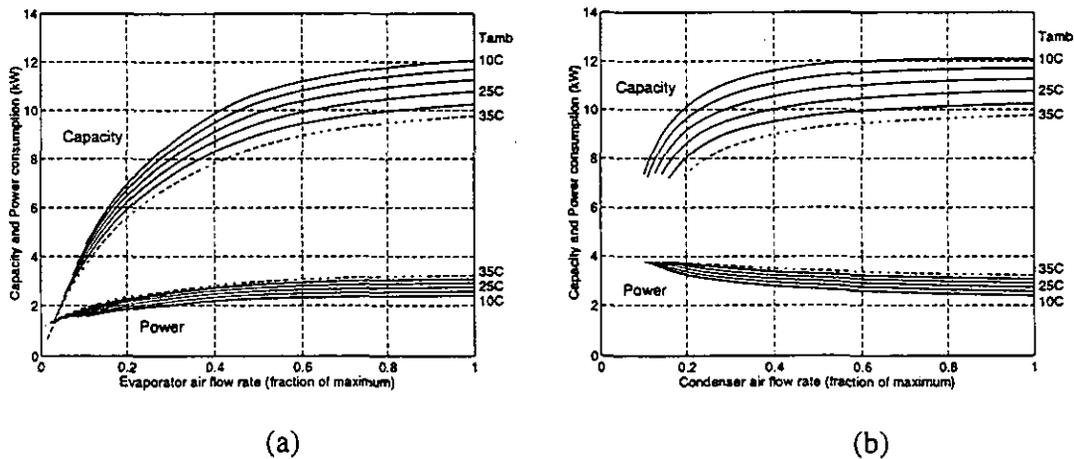
Figure 5.   Performance map for rooftop air conditioner

Fouling model

For simplicity, fouling was modeled as a linear function of fan runtime since the last servicing (t_r). It increased from 0, after cleaning, to 1 in a characteristic time t_f.

$$f = \begin{array}{ll} \dfrac{t_r}{t_f} & for \quad 0 < t_r \leq t_f \\[2ex] 1 & for \quad t_r > t_f \end{array} \tag{19}$$

The condenser fans ran only when the compressor was on and the evaporator fans ran continuously during occupied periods (12 hours/day) in the cooling season.

Constraints

Three of the four criteria that were defined for fault evaluation are constraints: comfort, safety, and environmental. Heat exchanger fouling can impact both comfort and equipment safety through its effect on the operating states of the air conditioner. The comfort constraint was handled by requiring the cooling capacity of the air conditioner to be greater than or equal to the cooling requirement at each hour of the simulation. The only safety constraints considered were high and low limits on the refrigerant operating pressures. The vapor compression cycle was required to conform to high head pressure (<3050 kPa (425 psig)) and low suction pressure (>275 kPa (25 psig)) safety constraints.

SIMULATION RESULTS

Given a building, an air conditioning unit, and a weather model, there are only two independent variables on which the optimal service schedule depends: the cost ratio C_s/C_e and the fouling time t_f. The first demonstration results in this section provide the costs of the optimal, regular, constrained, and simplified service schedules as a function of fouling time. Results are compared for evaporator and condenser fouling and for two different capacity-to-load ratios. The second set of results explores the cost ratio and capacity-to-load ratio dependence for fixed fouling time. The third set of results investigates the sensitivity of the simplified service schedule to the type of fouling model.

It is more meaningful to present results as a function of calendar time required for fouling rather than runtime. For this purpose, a normalized fouling time is defined as

$$t_f^* = \frac{t_f}{N_r} \qquad (20)$$

where N_r is the number of hours of fan operation that would occur each year if there were no fouling. For evaporator fouling, t_f^* is the actual number of years required for complete fouling because the evaporator fan runs continuously during occupancy, independent of the

fouling rate. For condenser fouling, t_f^* only approximates the actual number of years required for complete fouling because the compressor and condenser fan must operate longer as the condenser coil becomes fouled and the cooling capacity is reduced.

Cost vs. t_f^* for Evaporator Fouling

Figure 6 compares the combined costs of energy and service for different service schedules versus the fouling rate of the evaporator coil. The operating costs, plotted on the vertical axis, are normalized using the base operating cost with no fouling. The base cost is the total annual energy cost with no fouling for cooling the building and has a constant value of $857 for all results given in this paper. The normalized cost is the actual cost minus the base cost divided by the base cost and represents the fractional extra cost due to reduced efficiency and increased service associated with the fouled heat exchangers. The horizontal axis is the normalized characteristic fouling time of the heat exchanger (t_f^*).

All four plots of Figure 6 were generated for representative service and energy costs of C_s = $60 and C_e = 0.10 $/kWh. Figures 6(a) and 6(b) describe the main results, while 6(c) and (d) enlarge a portion of the constrained only solution with additional information for describing the unusual non-uniform nature of the curves. Figures 6(a) and 6(c) are for PLF_{max}=0.55 and 6(b) and 6(d) are for PLF_{max}=0.80. PLF_{max} (maximum part-load fraction) describes the size of the cooling equipment relative to the load and is the ratio of the design cooling capacity to the peak building load requirement (10 tons for this study).

As expected, Figures 6(a) & 6(b) demonstrate that the optimal solution always results in lower costs than the alternative service strategies. In both figures, the savings associated with optimal maintenance scheduling increase with fouling time. For optimal scheduling, the additional costs due to fouling asymptotically approach zero as the fouling time. In the limit, the fouling goes to zero and the optimal solution reduces the service frequency to zero. In contrast, the extra cost of the regular service schedules approaches the annual service cost as the fouling time increases. The constrained service schedule also approaches zero additional costs as the fouling time increases, but at a slower rate than the optimal scheduler. For a fouling time of 5 years in Figure 6(a), the minimum operating costs are about a third of those

for twice/season service and about two-thirds of those for once/season and constrained service.

(a)

(b)

(c)

(d)
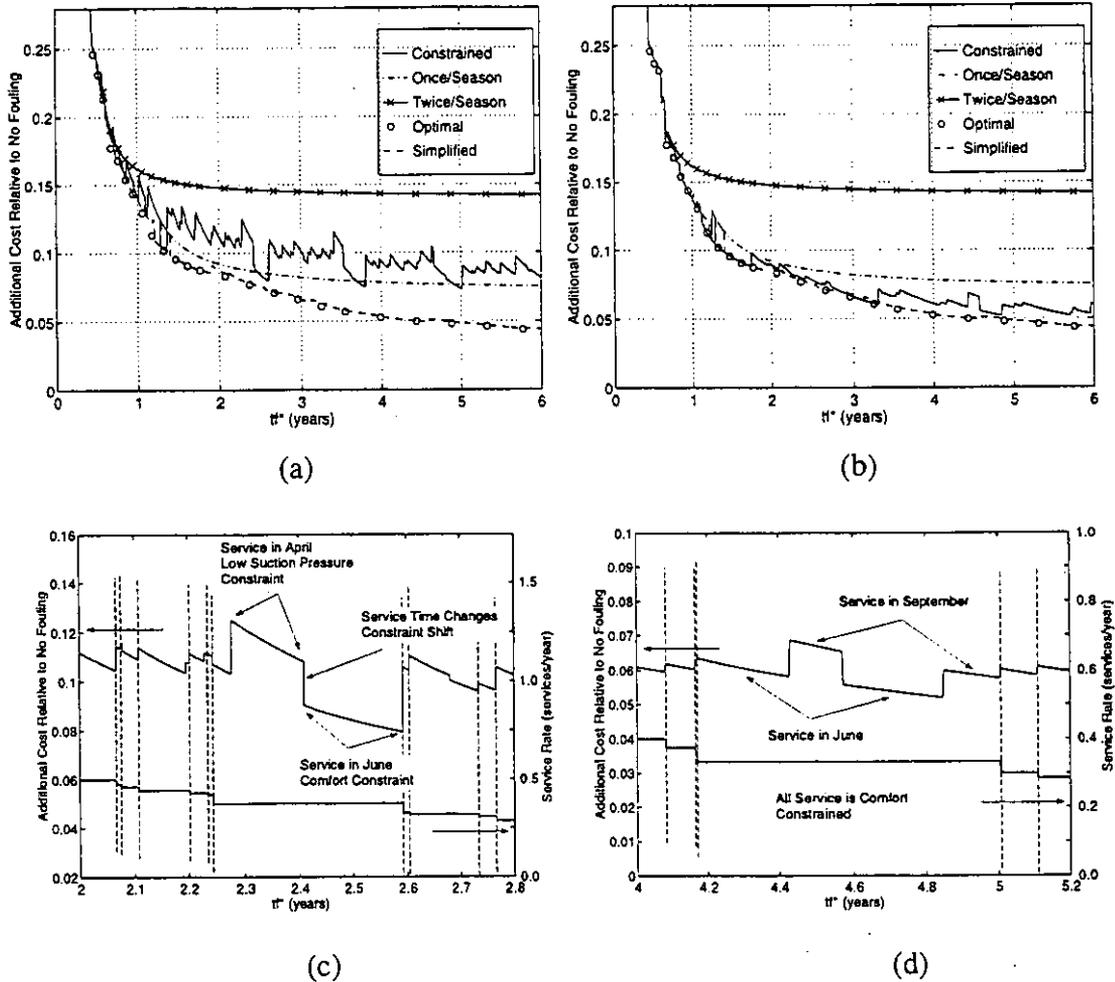
Figure 6.   Additional annual operating costs versus t_f^* for evaporator fouling.

The results of Figure 6(b) are similar to those for Figures 6(b) except that the minimum operating costs are not much lower than the costs for the constrained service scheduler. This behavior is not surprising since evaporator fouling has a direct impact on capacity and there is less extra capacity to meet the load when the heat exchanger fouls. In general, the constrained solution is nearly optimal when there is little extra capacity and is less efficient when there is extra capacity and longer fouling times.

The regular maintenance schedulers always provide excessive service whenever the service interval is short enough to ensure that comfort conditions are maintained. For the

results of Figures 6(a) and 6(b), the twice/season regular service scheduler maintains comfort for shorter fouling times than the once/season schedule at the expense of higher service costs for longer fouling times.

The costs associated with the simplified, near-optimal scheduler are close to those for the optimal solution under all conditions. The average difference between the optimal and near-optimal schedules is less than 0.2% for all results presented in Figures 6(a) and 6(b).

Figures 6(c) & 6(d) reproduce a portion of the costs of the constrained only schedules to help explain the irregular features. Curves providing the corresponding service rates are also included and are referenced to the right axes. The service rate is the average number of services per year computed over one service cycle. Dotted vertical lines indicate when the service rate changes. The primary causes of the irregular nature of the cost curves are discontinuities caused by sudden changes in service costs when the service rate changes followed by changing energy costs as the fouling time changes for a constant service rate. Secondary causes of irregular behavior in the cost curves are discrete changes in the month service is performed (with the same service rate) and changes in the constraint violation that triggers the service. Both of these secondary features are included in these figures. For example, when $PLF_{max} = 0.80$ and the service rate equals one service every three years, the total costs change when the service time changes from June (month 6) to September (month 9). For the case of $PLF_{max} = 0.55$ and the service rates equal to one service every 1, 2, & 3 years, the operating costs change rapidly when the service time changes from June to April because the constraint violation switches from comfort to low suction pressure.

Cost vs. t_f^* for Condenser Fouling

Figure 7 compares the combined costs of energy and service for different service schedules versus the fouling rate of the condenser. The trends and percent savings associated with optimal maintenance scheduling are similar to those reported for evaporator fouling. However, the scale of the vertical axis (extra cost) is larger due to the higher cost for condenser cleaning ($100) as compared to evaporator fouling ($60) resulting in larger absolute savings for optimal maintenance scheduling. In contrast to the results of Figure 6, the differences between the costs for the constrained only and optimal schedules do not depend strongly on PLF_{max}. Unlike evaporator fouling, condenser fouling does not have a

694

severe impact on capacity. The cooling capacity does affect the condenser fan runtime and therefore the rate of fouling. However, the results in Figures 7 are presented as a function of fouling relative the fan runtime.
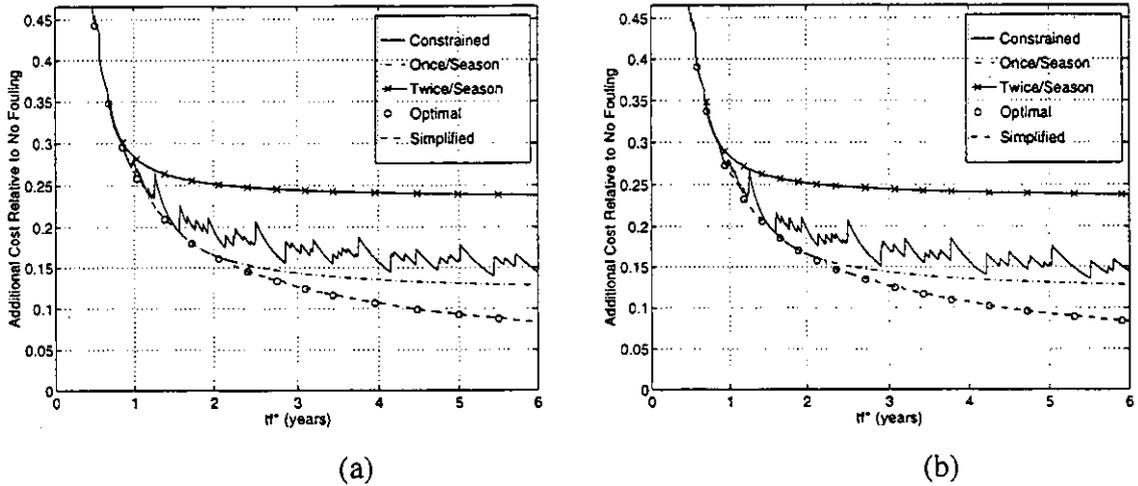
Figure 7.   Additional annual operating costs versus t_f^* for condenser fouling

Once again, the simplified, near-optimal scheduler gives operating costs that are within 0.2% of the optimal solution. The constrained only service schedules have the same irregular features observed in the evaporator fouling service schedule.

Cost vs. C_s and PLF_{max} for Evaporator Fouling

Figure 8 shows the costs of the four service schedules as a function of service cost and PLF_{max} for $t_f = 7830$ runtime hours ($t_f^* = 3.0$ years). When C_s was varied, PLF_{max} was fixed at 0.55 and when PLF_{max} was varied, C_s was fixed at $60. Figure 8(a) demonstrates that the costs associated with the regular and constrained schedules are linear functions of service cost, while the costs for the optimal schedule increase more slowly as service costs increase. When the service cost is small, the regular and optimal schedules converge because the cost of excessive servicing is small. As service costs increase, the cost penalties associated with the excessive service of the regular schedules increase. In contrast, the constrained only schedule is not providing enough service and its total cost is always greater than the optimal.

The differences become greater when the service costs are small and there is an incentive to service more often to reduce energy costs.
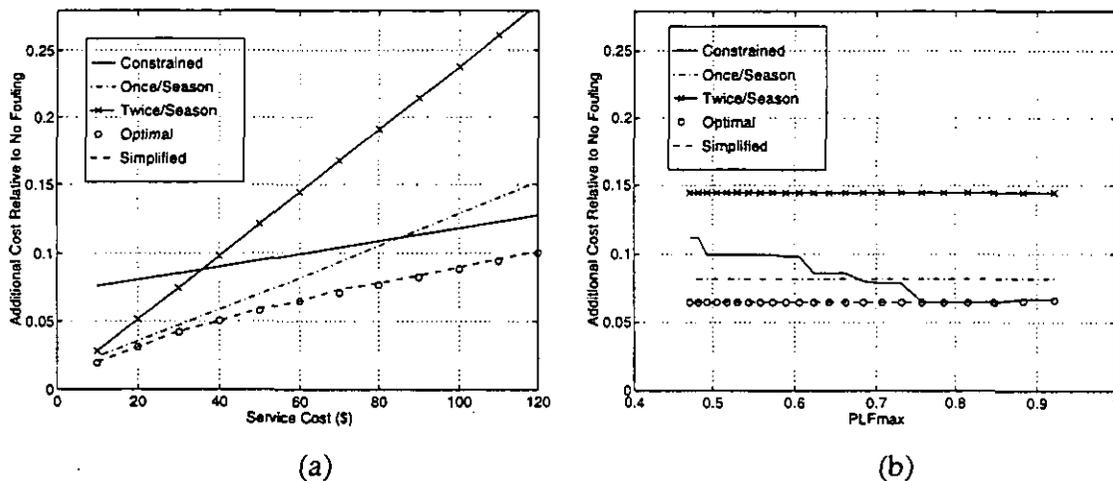
Figure 8.   Additional annual operating costs versus service cost (a) and PLF_{max} (b) for evaporator fouling.

Figure 8(b) shows that the costs of the optimal and regular service schedules are nearly independent of PLF_{max} for a fixed fouling rate. Since the fan runtime is fixed by the occupancy schedule, the coil fouls at the same rate, regardless of the cooling capacity. In addition, the efficiency of the on/off controlled compressor does not depend on its runtime. As a result, the same profile of cooling requirements will produce the same optimal schedule, regardless of cooling capacity if service decisions are not based on constraints. The constrained only schedule changes with PLF_{max} because as the capacity of the unit increases, it can operate with more fouling and still maintain the comfort constraint. In this case, the energy costs increase as the service rate decreases with additional cooling capacity.

The costs associated with the simplified, near-optimal scheduler are very close to the minimum costs under all conditions considered in Figure 8.

Cost vs. C_s and PLF_{max} for Condenser Fouling

Figure 9 shows the costs of the five service schedules as a function of service cost and PLF_{max} for condenser fouling with t_f = 1230 runtime hours. The fouling time of t_f = 1230 runtime hours corresponds to t_f^* = 2.75 years for PLF_{max} = 0.55 and t_f^* = 1.95 years for PLF_{max} = 0.80. When C_s was varied, PLF_{max} was fixed at 0.55 and when PLF_{max} was varied,

C_s was fixed at \$100. The results of Figure 9(a) are similar to those of Figure 8(a). For low service costs, the optimal and regular schedules give similar costs since they all provide enough service and there is a small penalty for over servicing. As service costs rise, the regular schedules cost more because they are providing excessive service. In contrast, the constrained schedule is under servicing resulting in larger cost penalties at low service costs.
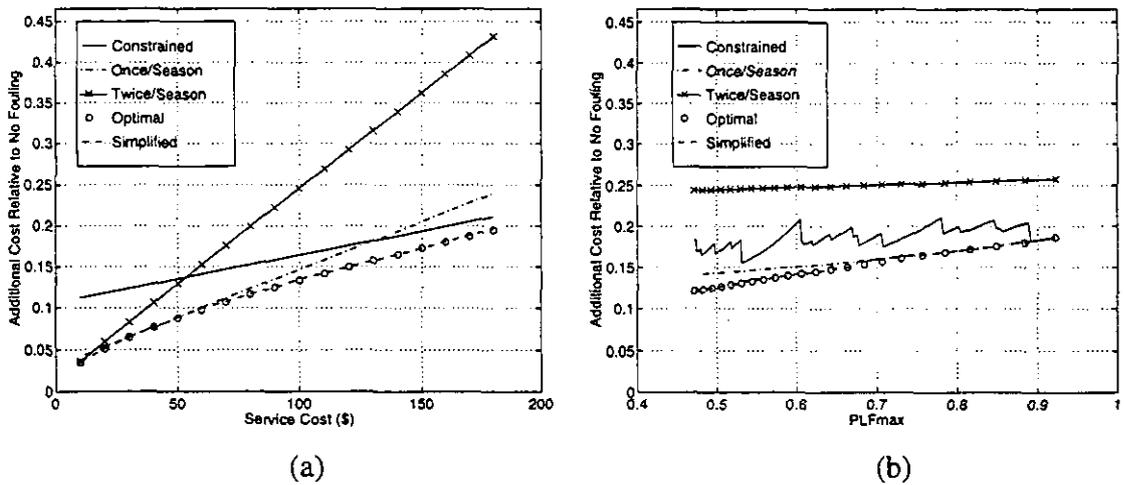
Figure 9.   Additional annual operating costs versus service cost (a) and PLF_{max} (b) for condenser fouling.

Unlike evaporator fouling, the costs of regular and optimal service schedules are not independent of PLF_{max} for a fixed fouling rate. When the equipment size becomes larger (smaller PLF_{max}), the condenser fans run less and fouling occurs more slowly (clock time), thereby decreasing the costs for all the schedules. The optimal schedule has the ability to reduce the service rate to less than once per year as the unit size increases, causing the costs of the once/year and optimal schedules to separate for low PLF_{max}.

Effect of the Fouling Model

For any point in time, an estimate of the optimal service time from the simplified service scheduler is based on the assumption that the current dependence of fouling on runtime will not change in the future. Under this assumption, the performance of this scheduler is extremely good as demonstrated in Figures 6 - 9. This section investigates the effect of using a fouling model that changes with time. The alternative fouling model consists of replacing portions of the constant fouling rate with randomly placed impulses while maintaining t_f as

the characteristic fouling time. The impulses model sudden unanticipated accumulations of dirt on the heat exchangers.

A mathematical description of the fouling model is:

$$f(t) = \int_0^t \left[\frac{1-\alpha}{t_f} + \frac{\alpha}{n \cdot t_c} w(\tau) \right] d\tau \tag{21}$$

where

$$w(\tau) = \delta(\tau - \tau_1) + \delta(\tau - \tau_2) + \ldots + \delta(\tau - \tau_n) \tag{22}$$

and $\bar{\tau} = \{\tau_1, \tau_2, \ldots, \tau_n\}$ is a set of uniformly distributed random runtimes in each cooling season when the impulse fouling occurs, and t_c is the calendar time required for the runtime to equal t_f in the absence of fouling. The two terms in the integrand are normalized such that complete fouling ($f = 1$) occurs approximately when the runtime accumulates to t_f, independent of α and n. The parameter α indicates the fraction of the fouling caused by impulses relative to a constant fouling rate. When $\alpha = 1$, the model is completely comprised of a series of impulses and when $\alpha = 0$, the model is completely linear. The parameter n is the number of impulses per calendar year. The impulses are randomly placed throughout the year, but reoccur at the same time each year.

Figure 10 shows the fouling state f as a function of runtime for a fouling time of $t_f = 7830$ hours (number of evaporator runtime hours in three cooling seasons). The vertical lines separate the cooling seasons. Figure 10(a) illustrates the dependence on n for $\alpha = 0.8$ and Figure 10(b) illustrates the dependence on α for $n = 3$. All fouling models attain complete fouling in the characteristic time t_f. The more nonlinear models are characterized by α near 1 and small n (e.g. $n=1$). In this case, fouling as a series of large step changes. As n grows larger, the evolution of fouling approaches a linear model, regardless of α.

Table 1 summarizes the costs for different fan control strategies and building capacity-to-load ratios for both evaporator and condenser fouling. The costs are the percent additional costs relative to optimal maintenance scheduling (actual - optimal cost over the optimal cost

times 100) averaged over 11 values of α between 0 and 1 and 5 values of n between 1 and 5. The fan control strategies are call for cooling (CFC), meaning that the fan operates only when the compressor is operating, and occupancy schedule (OCC), meaning that the fan runs during occupied periods for ventilation. For the cases considered, the simplified scheduler provided near-optimal operating costs. For evaporator fouling with $PLF_{max} = 0.80$ (little extra capacity), the constrained solution also provided performance comparable to the optimal schedule.

(a)   (b)
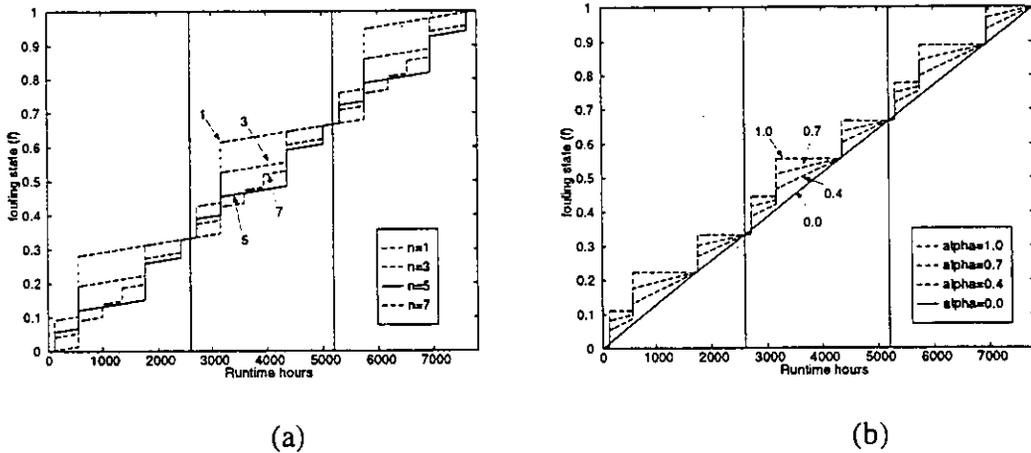
Figure 10.   Fouling state versus runtime hours.

Fouling Type	Control	PLF_{max}	% Additional Cost relative to Optimal			
			Simplified	Once/Season	Constrained	Twice/Season
Condenser	CFC	0.55	0.56	2.68	5.60	11.5
Condenser	CFC	0.80	0.53	1.85	4.42	9.93
Evaporator	OCC	0.55	0.46	2.11	3.74	7.88
Evaporator	OCC	0.80	0.46	2.11	0.65	7.88
Evaporator	CFC	0.55	0.36	1.81	2.61	7.58
Evaporator	CFC	0.80	0.31	1.21	0.21	6.63

Table 1.   Performance comparisons between different service schedules

CONCLUSION

This paper demonstrated that there is a significant opportunity for cost savings associated with optimal scheduling of condenser and evaporator maintenance. It was found that optimal service scheduling reduced lifetime operating costs for a rooftop air conditioner by as much as a factor of two over regular service intervals and 50% when compared to constrained only

service. For practical implementation, a simple near-optimal algorithm was developed for estimating optimal service times. In contrast to the optimal solution, this approach does not require on-line forecasting or numerical optimization and is easily implemented within a micro-controller. Over a wide range of cases tested, the near-optimal algorithm gave operating costs that were within 1% of the optimal results.

This study was performed for a small rooftop air conditioner. Greater opportunities for cost savings may be possible for larger equipment such as water chillers or electric power generating plants. These systems also experience performance degradations, but have significantly higher service and energy costs. Future work should extend the approaches described in this paper for determining optimal and near-optimal service times to these other systems. Implementation of the near-optimal maintenance scheduler should also be tested in the laboratory under controlled conditions.

REFERENCES

Anderson, D, L. Graves, W. Reinert, J.F. Kreider, J. Dow, and H. Wubbena. 1989. A quasi-real-time expert system for commercial building HVAC diagnosis. ASHRAE Transactions 95(2): 954-960.

Basseville, M. 1988. Detecting changes in signals and systems - a survey, Automatica 24(3): 309-326.

Bellman, R.E. 1957. Dynamic Programming. Princeton, New Jersey: Princeton University Press.

Cikanek, H.A. 1986. Space shuttle main engine failure detection. IEEE Control Systems Magazine 6(3): 13-18.

Culp, C.H. 1989. Expert systems in preventive maintenance and diagnosis. ASHRAE Journal 31(8): 24-27

Culp, C., J. Haberl, L. Norford, P. Brothers, and J.D. Hall. 1990. The impact of AI technology within the HVAC industry. ASHRAE Journal 32(12): 12-22.

Duyar, A., and W. Merill, 1992. Fault diagnosis for the space shuttle main engine. AIAA, Journal of Guidance, Control, and Dynamics 15(2): 384-389.

Frank, P.M. 1987. Fault diagnosis in dynamic systems via state estimation - a survey, Tzafestas, S., M. Singh, and G. Schmidt (eds), System Fault Diagnosis, Reliability, and Related Knowledge-Based Approaches 1: 35-98.

Frank, P.M. 1990. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results. Automatica 26(3): 459-474.

Georgescu, C., A. Afshari, and G. Bornard. 1993a. A model-based adaptive predictor fault detection method applied to building heating, ventilating and air-conditioning process. International Conference on Fault Diagnosis TOOLDIAG '93, Toulouse, France, April 5-7, 1993.

Georgescu, C., A. Afshari, and G. Bornard, 1993b. Qualitative fuzzy model-based fault detection and process diagnosis applied to building heating, ventilating, and air

conditioning. N. Piera Canete and M. G. Singh (Eds.), Qualitative Reasoning and Decision Technologies, Ocimne, Barcelona, 1993.

Gertler, J.J. 1988. Survey of model-based failure detection and isolation in complex plants. IEEE Control Systems Magazine December, 1988: 3-11.

Hall, I., R. Prairie, H. Anderson, and E. Boes, 1978. Generation of a typical meteorological year. Proc. of 1978 Annual Meeting, American Section of ISES, Denver 2(2): 669.

Hiroshi, I., H. Matsuo, K. Fujiwara, K. Yamada, and K. Nishizawa. 1992. Development of refrigerant monitoring systems for automotive air-conditioning systems. Society of Automotive Engineers: SAE Paper No. 920212: 29-39.

Isermann, R. 1984. Process fault detection based on modeling and estimation - a survey. Automatica 20(4): 387-404.

Jardine, A.K.S. 1973. Maintenance, Replacement, and Reliability. New York: Halsted Press, John Wiley and Sons Inc.

Kaler, G.M. 1988. Expert system predicts service. Heating. Piping, and Air Conditioning. November, 1988: 99-101.

Kaler, G.M. 1990. Embedded expert system development for monitoring packaged HVAC equipment. ASHRAE Transactions: 733-742.

Kitamura, M. 1980. Detection of sensor failures in nuclear plant using analytic redundancy. Transactions of the American Nuclear Society 34: 581-583.

Klein, S.A. et al. 1990. TRNSYS A transient system simulation program, 13.1 edn., Solar Energy Laboratory, University of Wisconsin - Madison, Madison, WI.

Kumamaru, T., T. Utsunomiya, Y. Yamada, Y. Iwasaki, I. Shoda, and M. Obayashi. 1991. A fault diagnosis system for district heating and cooling facilities. Proceedings of the International Conference on Industrial Electronics, Control, and Instrumentation, 17th, 1991, Kobe, Japan (IECON '91): 131-136.

McKellar, M.G. 1987. Failure diagnosis for a household refrigerator, Master's thesis, School of Mechanical Engineering, Purdue University.

Norford, L.K. and R.D. Little. 1993. Fault detection and load monitoring in ventilation systems. ASHRAE Transactions 99(1): 590-602.

Norford, L.K., A. Rabl, and G.V. Spadaro. 1987. Energy management systems as diagnostic tools for building managers and energy auditors. ASHRAE Transactions 93(2): 2360-2375.

Onken, R. and N. Stuckenberg, 1979. Failure detection in signal processing and sensing in flight control systems. Proceedings IEEE Conference on Decision and Control, San Diego: 449-454.

Pape, F.L.F., J.W. Mitchell, and W.A. Beckman, 1991. Optimal control and fault detection in heating, ventilating, and air conditioning systems. ASHRAE Transactions: Symposia, Part 1: 729-736..

Patton, R. J., S.W. Willcox, and J.S. Winter. 1987. A parameter insensitive technique for aircraft sensor fault analysis. AIAA Journal of Guidance, Control, and Dynamics: 359-367.

Rao, S. 1984. Optimization. 2nd ed., New Delhi, India: Wiley Eastern Limited.

Rossi, T. M. 1995. Detection, Diagnosis, and Evaluation of Faults in Vapor Compression Equipment, PhD thesis, Purdue University. Ray W. Herrick Laboratories, Purdue University, W.Lafayette, IN 47907 USA.

Usoro, P.B. and I.C. Schick, 1985. A hierarchical approach to HVAC system fault detection and identification. Dynamic Systems: Modeling and Control, Winter Annual Meeting of the ASME, Miami Beach, FL 1985, Nor 17-22 DSC v.1: 285-291.

Usoro, P. B., I.C. Schick, and S. Negahdaripour. 1985. An innovation-based methodology for HVAC system fault detection. Transactions of the ASME, Journal of Dynamic Systems, Measurement, and Control 107: 284-289.

Wagner, J. and R. Shoureshi. 1988. Observer designs for diagnostics of nonlinear processes and systems. ASME Journal, Winter Annual Meeting, Chicago, Illinois.

Wagner, J. and R. Shoureshi 1992. A robust failure diagnostic system for thermofluid processes. Automatica 28(2): 375-381.

Willsky, A.S. 1976. A survey of design methods for failure detection in dynamic systems. Automatica 12: 601-611.

Yoshimura, M. and I. Noboru 1989. Effective diagnosis methods for air-conditioning equipment in telecommunications buildings. INTELEC 89: The Eleventh International Telecommunications Energy Conference, October 15-18, 1989, Centro dei Congressi, Firenze, Conference Proceedings 21.1: 1-7.

DETERMINATION OF THE TUNING PARAMETERS OF A STEADY STATE DETECTOR FOR A CENTRAL AIR HANDLING UNIT

P.Gruber. Landis & Gyr, 6301 Zug, Switzerland

ABSTRACT

The dynamic behaviour (set point change, disturbances) of a central air handling unit (CAHU) is analyzed in detail, in order to determine the parameters of a steady state detector hooked to such a unit. It is shown that the dynamic response of the controlled unit to step set point changes is crucial for the determination of the time window needed for the steady state detection. The settling time of this response can be estimated by using only the steady state gains of heater and cooler and the control parameter of the PI controllers. The threshold are determined by evaluating the noise transfer functions between disturbance inputs and controlled variable T_s and by incorporating deterministic contributions to the deviation of T_s.

I. INTRODUCTION

The qualitative fault detection method which is used in the work of [1],[2] relies on the fact, that steady states or quasi-steady states of temperatures or other signals can be efficiently detected. One of the methods for the steady state detection is to compute recursively the running damped fixed window averages and corresponding variances for the relevant measured quantities [3]. Whenever the variances attain sufficiently low values, the system may be deemed to be in steady state. Tuning this type of steady state detector requires appropriate settings of an averaging window length or a damping parameter for the determination of the average and the variance of the signal and a threshold parameter for the variance.

In this work a method is derived which allows this window length to be estimated in terms of the settling times of the control loop. The settling times of the control loop are determined for step changes in set points, or in internal or external disturbances. These settling times are functions of the dynamics of the process and of the controller acting on the process. We assume that the parameters of the controllers are tuned such that a good transient behaviour is achieved. The tuning of the controller parameters is done by off- and on-line tuning methods. It will be shown that in the cases considered, the settling times can be estimated by the control parameters and the closed loop behaviour of the system without specific knowledge of the dynamics of the plant. Some ideas of how to choose the thresholds of the detector are also presented.

We concentrate on the central air handling unit (CAHU) only, as it was described in [4],[5]. First the control loop of the central air handling unit (CAHU) is described. Then the mathematical models of the CAHU and the control loop are derived in the third and fourth section. The fifth section presents the main results with the estimates of the settling times for the different modes of operation (pay-cooling, free-cooling, heating) and the different disturbances. In the sixth section the results are related to the parameters of the steady state detector.

2. THE CAHU CONTROL LOOP

In this work only the thermal behaviour of the system is considered.

The block diagram of the control loop of the CAHU basically looks as follows:
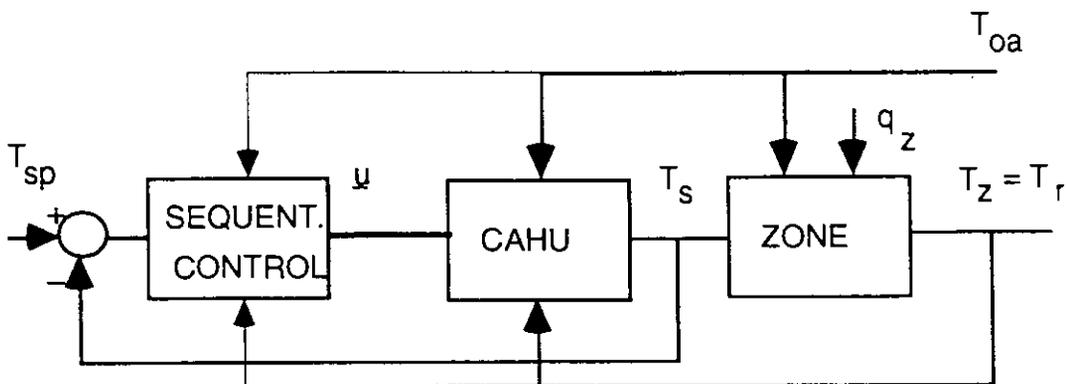
Fig.1: Temperature control of the supply air temperature of the CAHU

The **sequential controller** generates a control signal $\underline{u}$, which consists of three outputs: u_h, u_m, u_c. These outputs are controlled such that only one mode is active at a time. For each control variable u_h, u_m, u_c which is active in its corresponding mode (heating, damper, cooling), the controller has a different set of parameters. Each mode of the sequential controller consists of a PI-controller with some additional features. The two parameters of the PI-controller are first satisfactorily tuned for each mode. For the cooling and the heating mode the tuning is easy, because the process is more or less linear (except for the constraints imposed by the maximum heating and cooling capacity). For the damping mode however, the situation is different: the actuator gain is proportional to the difference of T_{oa} and T_r. Although the PI parameters for this case would actually be dependent on the operating conditions in the zone and the disturbance, they are set to constant values here.

The **CAHU** consists of the mixing box, cooling coil and heating coil including their dynamics and constraints. The output is the supply air temperature to the zones.

The **zone** includes the dynamics of the building and the sequential controllers acting via the VAV dampers and the reheating coils on the air flow rate and temperature in order to guarantee that the zone temperature T_z is equal to its setpoint T_{zsp}. q_z denotes the disturbances (heat gains and losses] of the zone.

The return air temperature T_r and the outdoor air temperature T_{oa}. usually act as slow disturbances on the control loop of the CAHU

3. THE CAHU CONTROL LOOP

We basically follow the system representation as it was used for the simulation studies in SIMULINK [FGGT93]. The CAHU block diagram of the system that is considered is shown in Fig.2
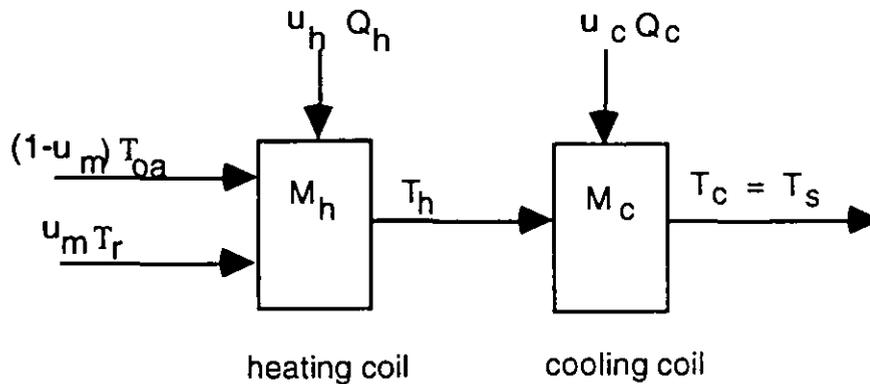
Fig.2: Block diagram of the CAHU

The mixing of the return and outside air temperature is assumed to be instantaneous

$$T_m = u_m T_r + (1 - u_m) T_{oa} \qquad (1)$$

$$u_m \in [0, u_{m\,max}], \qquad u_{m\,max} \leq 1$$

and the transport delays through the ducts are not taken into consideration too. Eliminating T_m leads to the model of the basic dynamics:

$$c M_h \dot{T}_b = c \dot{m}(u_m T_r + (1 - u_m) T_{oa} - T_b) + u_h \dot{Q}_b \qquad (2a)$$

$$c M_c \dot{T}_c = c \dot{m}(T_b - T_c) + u_c \dot{Q}_c \qquad (2b)$$

$$u_b \in [0,1], \qquad u_c \in [0,-1]$$

M_h=air mass in heater, M_c=air mass in cooler, c=specific heat capacity of air

That is a second order system with the states T_h and $T_c = T_s$, the disturbance inputs T_r and T_{oa}, and the control inputs u_h, u_d, and u_c. The equations are linear in the temperature and the heating and cooling control but nonlinear in the damper control. In this case the control input is a multiplication of the damper position and a temperature. In the s-domain one obtains the two equations become

$$(s\frac{M_h}{\dot{m}}+1)T_h(s) = \underbrace{(u_m T_r)(s)}_{u_r(s)} + \underbrace{((1-u_m)T_{oa})(s)}_{u_{oa}(s)} + \frac{\dot{Q}_h}{c\dot{m}}u_h(s) \qquad (3a)$$

$$(s\frac{M_c}{\dot{m}}+1)T_c(s) = T_h(s) + \frac{\dot{Q}_c}{c\dot{m}}u_c(s) \qquad (3b)$$

Solving the first equation for $T_h(s)$, substituting into the second equation and replacing T_c by T_s yields:

$$(s\frac{M_c}{\dot{m}}+1)T_s(s) = \frac{1}{(s\frac{M_h}{\dot{m}}+1)}(u_{oa}(s)+u_r(s)) + \frac{\frac{\dot{Q}_h}{c\dot{m}}}{(s\frac{M_h}{\dot{m}}+1)}u_h(s) + \frac{\dot{Q}_c}{c\dot{m}}u_c(s)$$

$$T_s(s) = \frac{1}{(s\frac{M_h}{\dot{m}}+1)(s\frac{M_c}{\dot{m}}+1)}(u_{oa}(s)+u_r(s)) + \frac{\frac{\dot{Q}_h}{c\dot{m}}}{(s\frac{M_h}{\dot{m}}+1)(s\frac{M_c}{\dot{m}}+1)}u_h(s)$$

$$+\frac{\dot{Q}_c}{c\dot{m}}\frac{1}{(s\frac{M_c}{\dot{m}}+1)}u_c(s) \qquad (4)$$

It is assumed that $T_{oa} < T_r$ holds, otherwise the damper reverses its direction.

4. THE CONTROL LOOP EQUATIONS

The PI-controller has the transfer function

$$G_r(s) = K_r\frac{(1+s\tau_n)}{s\tau_n}$$

with τ_n the integral time and K_r the proportional factor..
The parameters of the controllers for the three controller modes are denoted by:

$$K_r^c, K_r^d, K_r^h, \tau_n^c, \tau_n^d, \tau_n^h$$

The sequential controller determines which of the controller outputs are active, that means only one of the following mode is active.

pay - cooling mode: $u_h(s) = 0$, $u_m = 0$ $\Rightarrow$ $u_r(s) = 0$, $u_{oa}(s) = T_{oa}(s)$

free $-$ cooling mode: $u_h(s) = 0$, $u_c(s) = 0$

heating mode: $u_c(s) = 0$, $u_m = u_{mmax}$ $\Rightarrow$ $u_{oa}(s) = \underbrace{(1 - u_{mmax})}_{K_{oa}}T_{oa}(s)$,

$$u_r(s) = \underbrace{u_{mmax}}_{K_{re}}T_r(s) \qquad (5a,b,c)$$

$$0 \le K_{re} \le 1, \qquad 0 \le K_{oa} \le 1 \qquad (5d),(5e)$$

We can now consider the control loop for the supply air temperature in more detail. We make a distinction between the three controller modes

Pay-Cooling

$$K_c = \frac{\dot{Q}_c}{c\dot{m}}, \qquad \tau_c = \frac{M_c}{\dot{m}}$$

Heating

$$K_b = \frac{\dot{Q}_b}{c\dot{m}}, \qquad \tau_b = \frac{M_b}{\dot{m}}, \qquad K_{re} = u_{mmax}, \qquad K_{oa} = 1 - u_{mmax}$$
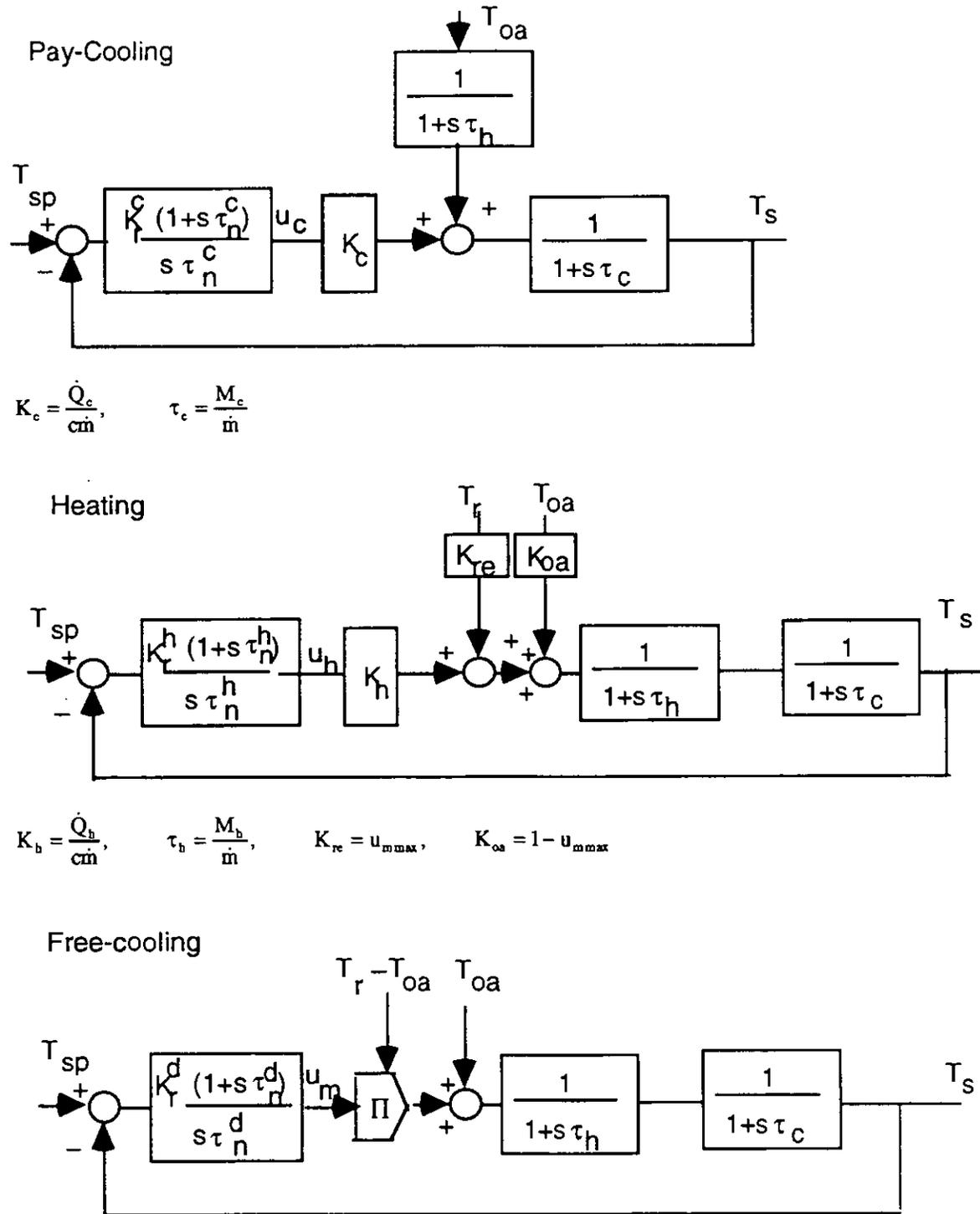
Free-cooling

Fig. 3a,b and c: The control loops for the three modes pay-cooling, free-cooling and heating for the case $T_{oa} < T_r$

706

5. THE CONTROL LOOP TRANSFER FUNCTION

A transfer function can be derived for each mode and each input if the system is linear. For the disturbance input in the free-cooling mode there is a nonlinear behaviour, such that some more assumptions have to be made. We start with the heating mode, continue with the pay-cooling mode and end with the free-cooling mode.

5.1. HEATING MODE

In the heating mode the dynamic of the system is of order three. In order to make things easy, we assume that the time constant τ_n of the PI-controller is tuned such it cancels the slowest mode of the system. In our case we set

$$\tau_n^h = \tau_c. \qquad (6)$$

With this assumptions we obtain for the set point and the two disturbance transfer functions the following expressions

Set point:
$$\frac{T_s(s)}{T_{sp}(s)} = G_{sp}(s) = \frac{\dfrac{K_r^h K_h}{\tau_n^h \tau_h}}{s^2 + \dfrac{s}{\tau_h} + \dfrac{K_r^h K_h}{\tau_n^h \tau_h}}$$

Disturbance T_{oa}:
$$\frac{T_s(s)}{T_{oa}(s)} = G_{oa}(s) = \frac{\dfrac{K_{oa}}{(1+s\tau_h)(1+s\tau_c)}}{1 + \dfrac{K_r^h(1+s\tau_c)}{s\tau_n^h}\dfrac{K_{oa}}{(1+s\tau_h)(1+s\tau_c)}}$$

$$= \frac{\dfrac{K_{oa}}{\tau_h}s}{\left(1+s\tau_c\right)\left(s^2 + \dfrac{s}{\tau_h} + \dfrac{K_r^h K_h}{\tau_n^h \tau_h}\right)}$$

Disturbance T_r:
$$\frac{T_s(s)}{T_r(s)} = G_{re}(s) = \frac{\dfrac{K_{re}}{\tau_h}s}{\left(1+s\tau_c\right)\left(s^2 + \dfrac{s}{\tau_h} + \dfrac{K_r^h K_h}{\tau_n^h \tau_h}\right)}$$

$$\rightarrow \quad \omega_0^2 = \frac{K_r^h K_h}{\tau_n^h \tau_h}, \qquad 2\xi\omega_0 = \frac{1}{\tau_h}, \qquad (7a,b,c,d,e)$$

We analyze the behaviour in response to both set point changes and disturbances.

5.1.1. Behaviour in response to a set-point change:

We assume that K_r^h and τ_n^h are chosen such that the dynamic behaviour for set point changes corresponds to well damped system, that means ξ lies in the range:

$$0.4 < \xi < 0.707 \qquad (8)$$

This corresponds to an overshoot of 25% to 5%.
With K_h known as well, one can eliminate τ_h from the equations (7d) and (7e) and gets for ω_0 and $\xi\omega_0$:

$$\omega_0 = \frac{2\xi K_h K_r^b}{\tau_n^b}, \qquad \xi\omega_0 = \frac{2\xi^2 K_h K_r^b}{\tau_n^b} \qquad (8a, b)$$

For an approximate five percent (5%) settling time τ_{settle},

one obtains with the above expressions for $\xi\omega_0$:

$$\tau_{settle,sp}^h \cong \frac{3}{\xi\omega_0} = \frac{3\tau_n^b}{2\xi^2 K_h K_r^b} \qquad (9)$$

To be on the safe side, the lower ξ must be used resulting in a longer settling time. It must be noted that for real poles instead of complex poles ($\xi>1$) the settling time increases with increasing ξ. The settling time obtained by the lower value of ξ of equation (8) holds therefore up to a ξ value of 1.5 which corresponds to a well damped system.

5.1.2 Behaviour in response to a disturbance:

For the assessment of the disturbance behaviour we start from the transfer function given in equation (7b) and (7c). We recognise that for both disturbances the influence on the output is the same except for the different gains. We look first at the outdoor influence. For a step disturbance of height ΔT_{oa}, the response looks like an impulse response due to the differential term. Transformed back into the time domain one obtains for the deviation from the operating point

$$\Delta T_s(t) = \frac{K_{oa}\Delta T_{oa}}{\tau_h \omega_0 \sqrt{1-\xi^2}} e^{-\xi\omega_0 t} \sin\left(\omega_0 \sqrt{1-\xi^2}\, t\right) \qquad (10)$$

With the known expression for τ_h, one can determine the time it takes until the deviation ΔT_s from the operating point falls under a chosen threshold, denoted by $\Delta T_{s,sp}$. This settling time is given by

$$\tau_{settle,dist}^b \cong \frac{1}{\xi\omega_0} \ln\left(\frac{\Delta T_{oa} K_{oa} 2\xi}{\Delta T_{s,sp}\sqrt{1-\xi^2}}\right) = \frac{\tau_{settle,sp}^b}{3} \ln\left(\frac{\Delta T_{oa} K_{oa} 2\xi}{\Delta T_{s,sp}\sqrt{1-\xi^2}}\right) \qquad (11)$$

For the disturbance $K_{oa}\Delta T_{oa}$ is replaced by $K_{rc}\Delta T_r$, T_r

The influence of the additional pole depends on a knowledge of the time constant τ_c. We can assume that this time constant is slower than τ_h, because the zero of the controller has cancelled it. Simulations show that the effect of a closed loop real axis pole at $-1/\tau_c$ on the settling time of the above system having complex pole with real parts at $-\xi\omega_0$ can be estimated by inspection of the ratio of the absolute value of the real part of the real pole to the absolute value of complex poles. If this ratio exceeds 1 and $\xi>0.4$ then the influence is negligible. The inequality for the estimation of the influence is thus given by:

$$\left|\frac{1}{\tau_c \xi\omega_0}\right| > 1 \quad \rightarrow \quad \left|\frac{\tau_n^b}{\tau_c 2\xi^2 K_h K_r^b}\right| > 1 \qquad (11a)$$

With the equality (6) we obtain for the above inequality:

$$\left|\frac{1}{2\xi^2 K_h K_r^b}\right| > 1 \qquad (11b)$$

Example: $\tau_n^h=2$, $K_r^b=0.025/0.5=0.05$,

                       $K_h=12900/(2.44*1005)=5.261$

                       $K_r K_h=0.263$

        $\xi=0.707$, $\tau_{settle,sp}^b=3*2/(2*0.5*0.263)=22.8\,sec$

                       $\omega_0=2*0.707*0.263/2=0.186$

$\xi=0.4$, $\qquad\qquad \tau^h_{settle,sp}=71.3sec$

Disturbance T_{oa}:

$\Delta T_{s,sp}=0.1$, $\Delta T_{oa}=1$ $u_{mmax}=0.8$, $K_{oa}=0.2$

$\xi=0.707$, $\qquad\qquad \tau^h_{settle,dist}=7.6\ln(1.414*0.2/(0.1*0.707))=10.5sec$

$\xi=0.4$, $\qquad\qquad \tau^h_{settle,dist}=23.76\ln(0.8*0.2/(0.1*0.92))=13.15sec$

Disturbance T_r:

$\Delta T_{s,sp}=0.1$, $\Delta T_r=1$ $u_{mmax}=0.8$, $K_{re}=0.8$

$\xi=0.707$, $\qquad\qquad \tau^h_{settle,dist}=7.6\ln(1.414*0.8/(0.1*0.707))=21.1sec$

$\xi=0.4$, $\qquad\qquad \tau^h_{settle,dist}=23.76\ln(0.8*0.8/(0.1*0.92))=46.1sec$

The settling times for a step disturbance of one degree and a decay to a tolerance band less than 0.1^o is always smaller than the settling time to a set point step and a 5% tolerance band.

Equations (9) and (11) give expressions for the settling times of the system, which are **independent on the knowledge of the dynamics of heating and cooling coils!**

5.2 PAY-COOLING MODE

The transfer function are given by the following equation:

$$\text{Set point:} \quad \frac{T_s(s)}{T_{sp}(s)} = G_{sp}(s) = \frac{K^c_r K_c(1+s\tau^c_n)}{s\tau^c_n(1+s\tau_c)+K^c_r K_c(1+s\tau^c_n)} = \frac{\dfrac{K^c_r K_c}{\tau^c_n\tau_c}(1+s\tau^c_n)}{s^2+(1+K^c_r K_c)\dfrac{s}{\tau_c}+\dfrac{K^c_r K_c}{\tau^c_n\tau_c}}$$

$$\text{Dist. } T_{oa}: \quad \frac{T_s(s)}{T_{oa}(s)} = G_{oa}(s) = \frac{\dfrac{1}{\tau_c}s}{(1+s\tau_b)\left(s^2+(1+K^c_r K_c)\dfrac{s}{\tau_c}+\dfrac{K^c_r K_c}{\tau^c_n\tau_c}\right)} \qquad (13a,b)$$

$$\rightarrow \quad \omega^2_0 = \frac{K^c_r K_c}{\tau^c_n\tau_c}, \qquad 2\xi\omega_0 = \frac{1+K^c_r K_c}{\tau_c} \qquad (13c,d)$$

5.2.1 Behaviour in response to a set point change:

We assume that $K_r{}^c$ and $\tau_n{}^c$ are again chosen such that the dynamic behaviour for set point changes corresponds to a well damped systems, e.g. ξ is in the range of 0.4 to 0.707. With K_C known as well, one can eliminate τ_c and gets for ξ_0 and $\xi\omega_0$:

$$\omega_0 = \frac{2\xi}{\tau^c_n}\frac{K_c K^c_r}{1+K_c K^c_r}, \qquad \xi\omega_0 = \frac{2\xi^2}{\tau^c_n}\frac{K_c K^c_r}{1+K_c K^c_r} \qquad (13e,f)$$

For an approximate five percent settling time $\tau_{settle,sp}$ one obtains with the above expression for $\xi\omega_0$:

$$\tau^c_{settle,sp} \cong \frac{3}{\xi\omega_0} = \frac{3\tau^c_n(1+K_c K^c_r)}{2\xi^2 K_c K^c_r} \qquad (14)$$

It remains to check the influence of the real zero at $-z_r$ of the transfer function. A similar consideration as for the additional real pole of the disturbance behaviour in the heating mode leads to the following inequality for the ratio of the absolute values of the real parts:

$$\left|\frac{z_r}{\xi\omega_0}\right| > 2 \quad \rightarrow \quad \left|\frac{1}{\tau_n^c\xi\omega_0}\right| > 2 \qquad (14a)$$

This conditions ensures that the dynamic behaviour,is not effected by the zero. It can be assumed, that τ_n^c is not much different from the slower time constant, in our case τ_c.

5.2.2 Behaviour in response to a disturbance

For the assessment of the disturbance behaviour we start from the transfer function of equation (13b). We recognise that for a stepwise disturbance of height ΔT_{oa},the response looks again like an impulse response due to the differential term. First we neglect the additional pole. Transforming the other part of equation (13b) back into the time domain one obtains

$$\Delta T_s(t) = \frac{\Delta T_{oa}}{\tau_c\omega_0\sqrt{1-\xi^2}}e^{-\xi\omega_0 t}\sin\left(\omega_0\sqrt{1-\xi^2}t\right) \qquad (15)$$

With the known expression for τ_c, one obtains for a settling time of ΔT_s under the threshold $\Delta T_{s,sp}$

$$\tau_{settle,dist}^c \cong \frac{1}{\xi\omega_0}\ln\left(\frac{\Delta T_{oa}2\xi}{\Delta T_{s,sp}\left(1+K_r^cK_c\right)\sqrt{1-\xi^2}}\right) = \frac{\tau_{settle,sp}^c}{3}\ln\left(\frac{\Delta T_{oa}2\xi}{\Delta T_{s,sp}\left(1+K_r^cK_c\right)\sqrt{1-\xi^2}}\right) \qquad (16)$$

It remains to check the influence of the pole at $-p_r$ of the transfer function. In order not to influence the dynamic behaviour the same inequality should hold as in the case of disturbance behaviour in the heating mode :

$$\left|\frac{p_r}{\xi\omega_0}\right| > 1 \quad \rightarrow \quad \left|\frac{1}{\tau_h\xi\omega_0}\right| > 1 \qquad (16a)$$

The time constant τ_h is not known, but in our example we assume that it is smaller than τ_c . In the heating mode the slower time constant was cancelled by the controller time constant τ_n^c. So we can use τ_n^c as upper limit for τ_h and plug it in the above equation.

Equations (14) and (16) give expressions for the settling time of the system, which are again independent on **the knowledge of the dynamics of the heating and cooling coils!**

Example: $\tau_n^c=2$, $\quad\quad\quad\quad$ $K_r^c=0.025/1.3=0.0192$,

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $K_c=40600/(2.44*1005)=16.56$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $K_r^cK_c=0.3184$

Setpoint

$\xi=0.707$, $\quad\quad\quad\quad\quad\quad$ $\tau_{settle,sp}^c=6*1.3184/(2*0.5*0.3184)=24.8sec$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $\omega_0=2*0.707*0.3184/(2*1.3184)=0.171$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ check for influence of zero:

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ $1/(2*0.171*0.707)=4.15>2$ --> negligible

$\xi=0.4$, $\quad\quad\quad\quad\quad\quad\quad$ $\tau_{settle,sp}^c=43.83sec$

Disturbance

$\Delta T_{s,sp}=0.1$, $\Delta T_{oa}=1$, $\tau_n^c=2$,

$\xi=0.707$, $\quad\quad\quad\quad\quad\quad$ $\tau_{settle,dist}^c=8.27\ln(1.414/(0.13184*0.707))=22.5s$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ check for pole influence: same result as for set

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ point

$\xi=0.4$, $\quad\quad\quad\quad\quad\quad\quad$ $\tau_{settle,dist}^c=14.6\ln(0.8/(0.13184*0.92))=27.5sec$

Again, the settling times for a step disturbance of one degree and a decay to a tolerance band of less than 10% is always smaller than the settling time to a set point step and a 5% tolerance band.

5.3 FREE-COOLING MODE

Looking at the control in the damper mode, one recognises immediately that this control loop is equivalent to the heating mode control loop with one important difference that there exist a time variable gain

$$K_d = (T_r - T_{oa}). \qquad (17)$$

For the moment we assume that K_d is constant and that

$$\tau_n^d = \tau_c \qquad (18)$$

Then the set point transfer function is given as

$$\text{Set point:} \quad \frac{T_s(s)}{T_{sp}(s)} = G_{sp}(s) = \frac{\dfrac{K_r K_d}{\tau_n^d \tau_b}}{s^2 + \dfrac{s}{\tau_b} + \dfrac{K_r K_d}{\tau_n^d \tau_b}}$$

$$\rightarrow \quad \omega_0^2 = \frac{K_r K_d}{\tau_n^d \tau_b}, \qquad 2\xi\omega_0 = \frac{1}{\tau_b}, \qquad (19)$$

The settling time is then according to equation (19)

$$\tau_{settle,sp}^d \cong \frac{3}{\xi\omega_0} = \frac{3\tau_n^d}{2\xi^2 K_d K_r} \qquad (20)$$

From equation (20) it is evident, that the settling time is inversely proportional to the varying gain K_d. One can expect a similar or smaller settling time in the damper mode for a set point change if

$$K_d \geq K_b \qquad (21)$$

In the above example this is the case as long as $K_d > 5^\circ C$.

For small differences between return air and outdoor air temperatures the gain becomes very small and the transient response very slow if the controller gain is not increased. If we assume that the controller has fixed coefficients, then one can conclude that **in the free cooling mode the steady state detection based on settling time is inappropriate if $T_r - T_{oa}$ gets small ($<2^\circ C$).**

If the influence of the disturbances is to be considered one has to make some further assumptions. We assume that the system is in steady state before a stepwise change in one of the disturbances occurs. In steady state ,the error is zero, which leads to the following steady state values:

$$u_m^0 (T_r^0 - T_{oa}^0) + T_{oa}^0 = T_{sp} \quad \rightarrow \quad u_m^0 = \frac{T_{sp} - T_{oa}^0}{T_r^0 - T_{oa}^0} < 1 \quad (22a)$$

$$\rightarrow \quad 1 - u_m^0 = \frac{T_r^0 - T_{sp}}{T_r^0 - T_{oa}^0} < 1 \quad (22b)$$

Next the effects of changes in the outdoor air temperature T_{oa} or the return air temperature T_r is investigated. For this purpose the deviations of the variables from the operating point are defined as:

$$\Delta T_s = T_s - T_{sp}, \quad \Delta T_{oa} = T_{oa} - T_{oa}^0, \quad \Delta T_r = T_r - T_r^0,$$

If the disturbance behaviour is linearized in this way, one obtains after some manipulations

$$\Delta T_s(s) = \frac{\dfrac{1}{\tau_b} s \left(\Delta T_{oa}(s)(1-u^0) + u^0 \Delta T_r(s) \right)}{(1 + s\tau_c)\left(s^2 + \dfrac{s}{\tau_b} + \dfrac{K_r K_d}{\tau_n^d \tau_b}\right)}$$

The disturbances ΔT_{oa} and ΔT_r are propagated basically in the same way as the disturbance in the heating mode. So the influence of a stepwise change is determined by equation (11) by substituting for K_{oa} new quantities :

$$K_{oa}^d = \Delta T_{oa} \frac{T_r^0 - T_{sp}}{T_r^0 - T_{oa}^0} \quad \text{or} \quad K_{re}^d = \Delta T_r \frac{T_{sp} - T_{oa}^0}{T_r^0 - T_{oa}^0} \qquad (22c,d)$$

Additionally K_h is replaced by K_d in equation (7b).

The factors K_{oa}^d and K_{re}^d with which the step heights ΔT_{oa} and ΔT_r are multiplied are due to the inequalities (22a, 22b) strictly less than 1. The influence is therefore comparable to the disturbance effects in the heating mode and the same estimates for the settling times are applicable.

The example follows very closely to the one of the heating mode and is left out.

5.4 MAXIMAL APPLICABLE DISTURBANCES

If we look at the result of the preceding sections one can conclude that for step wise disturbances of up to one degree Celsius and a settling time to 10% of the disturbance height, **the settling time of a set point change is good upper bound for the settling times which are to be expected for the disturbance inputs.** In the control loop configurations under investigation the disturbance transfer function always has a phase lead behaviour compared to the set point transfer function due to the PI controller.

$$G_{sp} = \frac{G_r G_s}{1 + G_r G_s} \qquad G_{dist} = \frac{G_s}{1 + G_r G_s} = \frac{1}{G_r} \frac{G_r G_s}{1 + G_r G_s} = \frac{1}{G_r} G_{sp}$$

The term $(G_r)^{-1}$ generates a phase lead effect.

We can compute now the maximum ratio of height of the step of the disturbance to the tolerance band width $\Delta T_{s,sp}$. This leads to the same settling time as the set point settling time with a 5% tolerance band:

$$\tau_{settle,dist} \cong \frac{\tau_{settle,sp}}{3} \ln\left(\frac{K\Delta T 2\xi}{\Delta T_{s,sp}\sqrt{1-\xi^2}}\right)$$

with $K = K_{oa}$ $\Delta T = \Delta T_{oa}$ heating mode

   $K = K_{oa}^d$ $\Delta T = \Delta T_{oa}$ free - cooling mode

   $K = K_{re}$ $\Delta T = \Delta T_r$ heating mode

   $K = K_{re}^0$ $\Delta T = \Delta T_r$ free - cooling mode

   $K = \dfrac{1}{1 + K_r K_c}$ $\Delta T = \Delta T_{oa}$ pay - cooling mode

$$\tau_{settle,dist} \le \tau_{settle,sp} \quad \rightarrow \quad \frac{\Delta T}{\Delta T_{s,sp}} \le \frac{e^3}{\dfrac{2\xi K}{\sqrt{1-\xi^2}}} \qquad (23)$$

Equality in equation (23) is obtained for:

$$\Delta T_{s,sp} = \frac{2\xi K}{\sqrt{1-\xi^2}} \frac{1}{e^3} \Delta T \qquad (23a)$$

numerical values: a) $\Delta T_{s,sp} = 0.1°C$ $K = 1$ b) $\Delta T_{oa} = 2°C$ $K = 1$

   $\xi = 0.707$ $\Delta T \le 1°C$ $\xi = 0.707$ $\Delta T_{s,sp} \cong 0.05°C$

   $\xi = 0.4$ $\Delta T \le 2.3°C$ $\xi = 0.4$ $\Delta T_{s,sp} \cong 0.1°C$

6. STEADY STATE DETECTOR PARAMETER DETERMINATION

The derivation of steady state detection schemes are described in [G&G96]. We consider here the detection using fixed window or geometrically weighted moving variances. The former is basically an averaging scheme which computes at each sampling time an estimate of the mean and the variance using a fixed number of data points. The latter uses a recursive low pass filter.

6.1. DETERMINATION OF THE TIME WINDOW

One signal that can be used to detect steady state is the error signal in the control loop. If this signal falls in absolute value to under 5% of the final value, we can assume quasi steady state conditions. The choice of the window length for the averager is determined as follows:

If the disturbance signal is effectively averaged over the length of the settling time and after this window length the variance is under the given threshold, then one can be reasonably certain that the influence on the error and so also on the steady state is less than the specified absolute error limits. Fig. 4 shows an illustrative example of this idea.
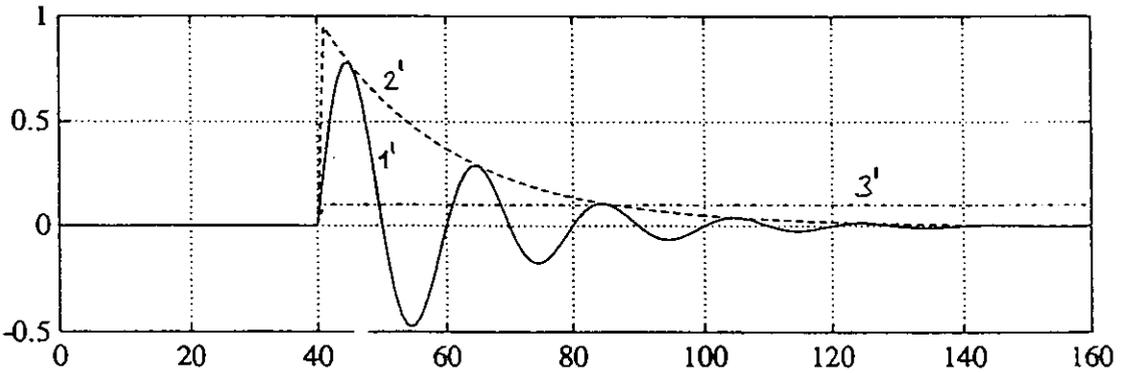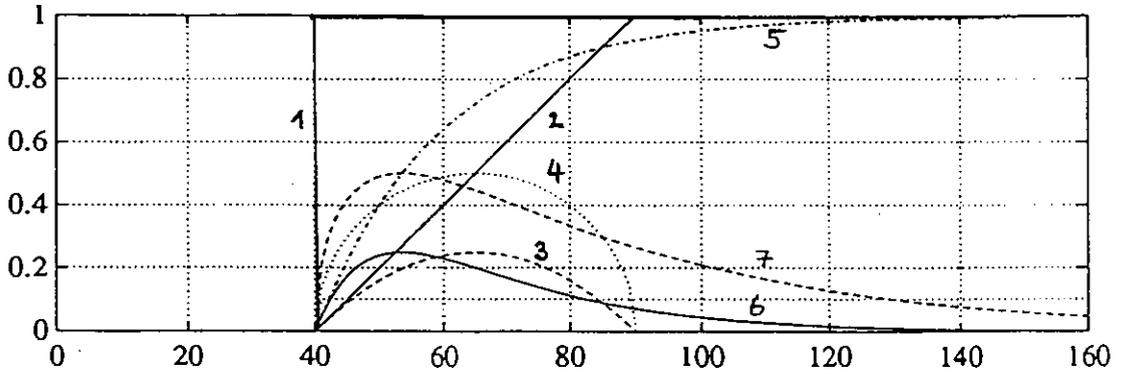
1: unit step disturbance at t=40
2: output of the fixed window mean estimator (averager of length 50)
3: output of the fixed window variance estimator (length 50)
4: standard deviation of the variance 3
5: output of the recursive mean estimator with damping factor 0.95
6: output of the recursive variance estimator with damping factor 0.95
7: standard deviation of variance 6

1': $\Delta T_s(t)$, transient response of the control loop due to step disturbance
2': exponential envelope
3': threshold $\Delta T_{s,sp}$

Fig.4a: Estimation of the mean and standard deviation for a stepwise disturbance input
Fig.4b: corresponding supply air temperature transients

713

It is assumed, that the step disturbance is a worst case signal with respect to the settling time.
A look at the Figures 4 shows the following: The settling time of the steady state detector should match the settling time of the control loop. Both variance estimators of Fig.4a reach approximately the same value for the variance estimate at the settling time of the control loop (t=90sec).
So the rule is as follows:

The averaging length T_w (or window length for a fixed time window) for the steady state detector should be chosen as

$$T_w \cong \tau_{settle,sp} \qquad (24)$$

$t_{settle,sp}$ is either the maximum of the settling times of the three controller modes or if the controller mode is known the corresponding settling time. For a geometrically weighted averaging procedure the correct exponential weighting factor is related to the time constant of the exponential transient behaviour of the low pass to a step input. This time constant is approximately $T_w/3$ [3].

6.2 THRESHOLD VALUES FOR THE STEADY STATE DETECTOR

It remains to derive a rule for the threshold and the hysteresis for the estimation of the variance or standard deviation of the disturbance (especially T_{oa})and of the control loop output, the supply air temperature T_s.

6.2.1 Supply air temperature:

The threshold values are set in response to two considerations. One is that the minimum threshold ε_{ss}^{Ts} pertaining to the supply air temperature T_s must take residual high frequency noise in the signal into account, whose standard deviation is denoted here by σ_{Ts}. A minimum standard deviation σ_{Ts} of T_s under stable operation can be found for instance by inspecting some recorded data. If deterministic changes in the outdoor air temperature, which act as disturbance on T_s, are taken into account as well, the standard deviation σ_{Ts}. has to be increased by a factor μ. μ can be related to the tolerance band $\Delta T_{s,sp}$, which was introduced in the deterministic disturbance behaviour (see influence of outdoor air temperature in next subsection..Secondly, in view of the fact that the response of the supply air temperature to step changes in its set point is approximately exponential, some further allowance has to be made for the residual effect of such a step change after the settling time, which is chosen as 5% of the original step. This residual amount is denoted by δ_{Ts}, and the resulting threshold ε_{ss}^{Ts} is given by:

$$\varepsilon_{ss}^{Ts} = \sqrt{\eta^2 \sigma_{Ts}^2 + \delta_{Ts}^2} \qquad (24a)$$

6.2.2 Outdoor air temperature

The same idea as before is used here. The noise contribution to the threshold is determined by using the standard deviation σ_{Ts} of the supply air temperature T_s.
If we assume a stationary random source with standard deviation σ_{Toa} as disturbance input we can calculate the standard deviation σ_{Ts} of the high frequency noise of the output signal T_s due to this input.
If we assume that we can neglect the additional zeros or poles, then all disturbance transfer functions are of the form

$$G_d(s) = \frac{2\xi\omega_0 K s}{s^2 + 2\xi\omega_0 s + \omega_0^2} = \frac{b_1 s}{s^2 + a_1 s + a_0} \qquad (25)$$

K is dependent on the controller mode and the disturbance input. However the following inequality holds (section V.d)):

$$K \leq 1 \qquad (26)$$

As an upper bound K can therefore be set to 1.
Now a stationary disturbance noise source with zero mean and variance σ_{Toa}^2 and a constant spectral density $S_{Toa}(\omega)=S_{Toa}$ is applied to the disturbance input:

$$\sigma_{Toa}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{Toa}(\omega)d\omega = \frac{1}{2\pi} \int_{-2\pi f_B}^{2\pi f_B} S_{Toa}(\omega)d\omega = S_{Toa} 2f_B$$

The bandwidth is normally half of the sampling frequency f_B at which the disturbances are measured: $f_B = f_s/2$. The sampling time is denoted by T_{sample}. Thus equation (25) gives the following σ_{Ts}^2 of the noise in steady state at the output of the control loop using Parseval's Theorem:

$$S_{Ts}(\omega) = |G_d(j\omega)|^2 S_{Ts}(\omega)$$

$$\sigma_{Ts}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_d(j\omega)|^2 S_{Ts}(\omega)\, d\omega = S_{Ts} \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_d(j\omega)|^2 d\omega$$

$$\sigma_{Ts}^2 = \frac{b_1^2 a_0}{2a_0 a_1} \frac{\sigma_d^2}{2f_B} = \frac{K^2 (2\xi\omega_0)^2 \omega_0^2}{2\omega_0^2 2\xi\omega_0} \sigma_{Toa}^2 = \frac{K^2 \xi\omega_0}{2} \frac{\sigma_{Toa}^2}{2f_B} \quad (27)$$

$$\sigma_{Ts} = \frac{K\sqrt{\xi\omega_0}}{\sqrt{2}} \frac{\sigma_{Toa}}{\sqrt{2f_B}} \quad (28)$$

If σ_{Ts}^2 is given instead of σ_{Toa}^2 then equation (27) has to be rewritten as:

$$\sigma_{Toa}^2 = \frac{2}{K^2 \xi\omega_0} 2f_B \sigma_{Ts}^2 = \frac{2}{K^2 \xi\omega_0} 2f_B \sigma_{Ts}^2 \quad (29)$$

$$\sigma_{Toa} = \frac{\sqrt{2}}{K\sqrt{\xi\omega_0}} \sqrt{2f_B} \sigma_{Ts} = \frac{\sqrt{2}}{K\sqrt{\xi\omega_0}} \sqrt{2f_B} \sigma_{Ts} \quad (30)$$

K can be set to 1 (equation 26)), $2f_B$ can be replaced by $1/T_{sample}$ and $\xi\omega_0$ by $3/\tau_{settle,sp}$ (equation (9)) resulting in:

$$\sigma_{Toa}^2 = \frac{2}{3} \frac{\tau_{settle,sp}}{T_{sample}} \sigma_{Ts}^2 \quad (31)$$

It remains to determine the effect of deterministic disturbance inputs. A step input is again regarded as the relevant signal. The residual effect of such a step change after the settling time for the variance estimator of T_{oa} is dneoted by δ_{Toa}, which is 5% of the disturbance step height. This leads to a resulting threshold ε_{ss}^{Toa} for the steady state estimator as:

$$\varepsilon_{ss}^{Toa} = \sqrt{\sigma_{Toa}^2 + \delta_{Toa}^2} \quad (31)$$

The disturbance step of height $20 * \delta_{Toa}$ has also effect on the deviation of the supply temperature. This influence after the settling time is assumed to be $\Delta T_{s,sp}$. The variance of T_s is thus increased by $\Delta T_{s,sp}^2$. Combined with σ_{Ts}^2 allows to determine μ:

$$\sigma_{Ts}^2 + \Delta T_{s,sp}^2 = \sigma_{Ts}^2 \left(1 + \frac{\Delta T_{s,sp}^2}{\sigma_{Ts}^2}\right) = \mu^2 \sigma_{Ts}^2$$

$$\mu^2 = 1 + \frac{\Delta T_{s,sp}^2}{\sigma_{Ts}^2}$$

6.2.3 Choice of Hysteresis

Because the variance estimation is evaluated at each sampling time instant a hysteresis type threshold has to be used to avoid rapid fluctuations of the steady state signal. This has to be applied to each variance estimation, that

means for each disturbance input and for the controlled signal. It is reasonable to choose the upper and lower values of the hystersis as a factor of the corresponding determined thresholds ε_{ss}^{Ts} and ε_{ss}^{Toa}. For a 10% factor one obtains:

$$\varepsilon_{ss,l}^{Ts} = 0.9\varepsilon_{ss}^{Ts} \qquad \varepsilon_{ss,u}^{Ts} = 1.1\varepsilon_{ss}^{Ts}$$
$$\varepsilon_{ss,l}^{Toa} = 0.9\varepsilon_{ss}^{Toa} \qquad \varepsilon_{ss,u}^{Toa} = 1.1\varepsilon_{ss}^{Toa}$$

As an example the hysteresis type threshold for the supply temperature is shown in Fig. 5
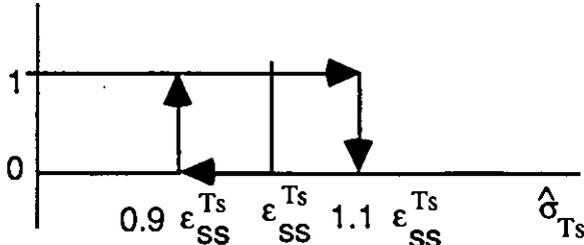
Fig.5: Example of thresholds for the supply temperature,

An example with test data is also treated in [6].

7 ACKNOWLEDGEMENTS

8. REFERENCES

[1] Fornera, L. Glass,A.S.,Gruber, P. and Tödtli, J..: "Fault Detection Based on Logical Programming Applied to the variable air volume air-handling unit, *Control Engineering Practice*, 4(1), 105-116, 1996

[2] Fornera,L. Glass,A.S.,Gruber, P. and Tödtli,J (1993).: "Qualitative Fault Detection Based on Logical Programming Applied to a Variable Air Volume Air Handling Unit", 2nd *IFAC Workshop on Computer Software Structures Integrating AI/KBS Systems in Process Control, Lund*, 1994

[3] A.S. Glass, P. Gruber : "Steady -state detection," *I.E.A Annex 25, Final report, Source book*, 1996

[4] Glass,A.S., Gruber, P., Roos,M. Tödtli,J.: Preliminary Evaluation of a Qualitative Model-Based Fault Detector for a Central Air Handling Unit", *3rd IEEE Conference on Control Applications*, Glasgow, 1994

[5] Glass, A.S., Gruber, P., Roos, M. & Toedtli, J.: "Qualitative Model-Based Fault Detection in Air-Handling Units", *IEEE Control Systems Magazine* 15(4), 11-22, August, 1995

[6] A.S.Glass, J. Toedtli: "Testing Qualitative Model-based Fault Detection for Air-Handling Units using Operational Building Data", *I.E.A Annex 25, Final report*, 1996

ANNEX

Summary of the equations needed:

Pay-Cooling

$\dot{m}$ = air mass flow

c = specific heat capacity of air

K_r^c, τ_n^c = PI control parameter

ξ = damping of controlled loop $(0.4 \leq \xi \leq 0.707)$

$$K_c = \frac{\dot{Q}_c}{c\dot{m}}$$

$$\tau_{settle,sp}^c = \frac{3\tau_n^c(1 + K_c K_r^c)}{2\xi^2 K_c K_r^c}$$

check: if $\dfrac{\tau_{settle,sp}^c}{3\tau_n^c} > 1$

Heating

$\dot{m}$ = air mass flow

c = specific heat capacity of air

K_r^h, τ_n^h = PI control parameter

ξ = damping of controlled loop $(0.4 \leq \xi \leq 0.707)$

$$K_h = \frac{\dot{Q}_h}{c\dot{m}}$$

$$\tau_{settle,sp}^h = \frac{3\tau_n^h}{2\xi^2 K_h K_r^h}$$

check: if $\dfrac{\tau_{settle,sp}^h}{3\tau_n^h} > 1$

Free-cooling

K_r^d, τ_n^d = PI control parameter

ξ = damping of controlled loop $(0.4 \leq \xi \leq 0.707)$

$$K_d = (T_r - T_{oa})$$

$$\tau_{settle,sp}^d = \frac{3\tau_n^d}{2\xi^2 K_d K_r^d}$$

check: if $\dfrac{\tau_{settle,sp}^d}{3\tau_n^d} > 1$

$\Rightarrow \quad \tau_{settle,sp}$

Threshold for controlled variable (supply air temperature)

$\sigma_{T_s}^2$ measured

$\delta_{T_s}^2$ and $\delta_{T_{oa}}^2$ estimated

K chosen

ξ chosen

$$\Delta T_{oa} = 20 * \delta_{T_{oa}}$$

$$\Delta T_{s,sp} = \frac{2\xi K}{\sqrt{1-\xi^2}} \frac{1}{e^3} \Delta T_{oa}$$

$$\eta = 1 + \frac{\Delta T_{s,sp}^2}{\sigma_{T_s}^2}$$

$$\varepsilon_{ss}^{T_s} = \sqrt{\eta^2 \sigma_{T_s}^2 + \delta_{det,sp}^2}$$

Threshold for disturbance variable (outdoor air temperature)

T_{sample} : sampling time chosen

$$f_B = \frac{f_s}{2} = \frac{1}{2T_{sample}}$$

$$\sigma_{T_{oa}}^2 = \frac{2}{3} \frac{\tau_{settle,sp}}{T_{sample}} \sigma_{T_s}^2$$

$$\sigma_{T_{oa}} \cong 0.82 \sqrt{\frac{\tau_{settle,sp}}{\Delta T}} \sigma_{T_s}$$

$$\varepsilon_{ss}^{T_{oa}} = \sqrt{\sigma_{T_{oa}}^2 + \delta_{T_{oa}}^2}$$

LIFE CYCLE EVALUATION OF PREVENTIVE MAINTENANCE USING FAULT PROBABILITY PROCESS

Hisahiro Ito, Associate Professor
Mie University
Nobuo Nakahara, Professor
Nagoya University

Abstract

Theoretical method for the preventive maintenance and the decision making of the investment for installation BEMS/BOFD system and the retrofit of components, subsystems and the total system has been developed introducing a transitional probability of incremental faulty stages. An example calculation shows the importance of the life cycle costing and the study to identify the transition probability.

1 INTRODUCTION

The present section reviews an innovative tool to identify the following issues:
  - reasonable investment to be spent for introducing BOFD system
  - *effective preventive maintenance strategy*
Present Value Based Life Cycle Cost Method is fundamentally applied as basic tool for their purposes. However, new ideas are developed to precisely evaluate the outbreak frequencies of fault and failure and the various kinds of damage by them. Thus, probability characteristics of fault's and failure's outbreaks are considered with "cost-converted" amount of damages in energy consumption and indoor environmental quality. The detail of this methodology has been described in references [1][2].

2 BASIC CONCEPT OF EVALUATION OF BOFD AND PREVENTIVE MAINTENANCE

BOFD system is quite alike fire-detection system. Both systems always supervise phenomena that could not happen in essential, so that these systems have ideally never done in the building's life. Thus, these systems involve contradictions that investments for them are thoroughly "worthless" when the ideals can be realized. However, the failure of fire-detection shall bring down enormous damages to human's life as well as building's life, so that the fire-detection system must be indispensable without any conditions. There is, therefore, little scope left to discuss whether it is needed or not from the viewpoint of cost/benefit.

How about the BOFD system? The failure of detecting a faulty operation in HVAC system will not inherently take human's life, or reduce building's life. The fault may decrease thermal comfort of occupant slightly. However, it may be trivial and unimportant problem if the global environment problems are taken into consideration as our commonsense. This is because we now live in a period when we must agree a slightly

uncomfortable indoor climate in our dairy life. Therefore, it is obvious that we endure slightly uncomfortable environment due to fault and/or failure in a HVAC system, since they happen entirely on rare occasions.

In fact, to install a BOFD system shall necessarily follow additional hardware, software, and sensors regardless of stand-alone type or BEMS-included type. The BOFD system must recover the investments for themselves completely in a building life. To improve the fault detecting accuracy could certainly increase investment to BOFD system itself. Thus, there shall be trade-off relationship between the investment for a BOFD system and the benefit due to it. On the other hand, building optimization based on the BOFD must be closely related to preventive maintenance strategy, i.e., whether a detected fault should be repaired or not. A minor fault may give no damage to thermal comfort and energy consumption, while the cost to repair the fault is inevitably needed. Thus, it is not always optimal from the viewpoint of cost/benefit to repair a fault whenever it is detected. It is very important to identify what level of fault should be repaired.

The following two standpoints must be discussed in all seriousness:
    1) How grade of BOFD system is reasonable to be applied in a building?
    2) How level of a fault should be repaired?
The latter level can be referred as "threshold" of repair/retrofit of a fault. Of course, both standpoints are closely related together. It is rational if they can be evaluated to minimize the total of the following costs.
    - investment for installing a BOFD system itself and adding/mending sensors
    - cost of preventive and/or after-failure maintenance of HVAC system
    - cost-converted damages due to fault and/or failure
The building's life should be taken as the period of evaluation. BOFD system is inherently to detect a fault in a HVAC system; a fault will follow damages to thermal comfort of occupants and energy consumption in HVAC system. Therefore, the benefit can be replaced as damages in the cost and benefit analysis. An investment for a BOFD system can be also interrupted as a kind of damage because the HVAC system is not complete. Thus, a BOFD system could be optimized to minimize the total life cycle damage in a building.

3 MODELING OF FAULT'S OUTBREAK, DETECTION AND MAINTENANCE PROCESS

3.1 BASIC CONCEPT

As have been discussed in the Annex 25, there are many kinds of target for BOFD; a unique machine such as pump, fan, and refrigerator, a subsystem such as thermal storage, and total HVAC system. On the other hand, fault in a machine or a system could shift in various states, from slight case to serious case. It will be rational that completely normal state and completely failure state can be also defined as extrapolation of "fault state". If these states can be shown as discrete value of fault level, change of fault state from a state, i, to a fault state, j, can be described by probability process, regardless of target of BOFD.

Fault detection and repair processes can be illustrated as **Figure 1**. The ordinate is fault level of a machine or a system and the abscissa is time. Fundamentally, a BOFD system always supervises fault state of a machine or a system every a time step, ΔT. A fault will surely either hold previous level or rise in level more than previous one. If the BOFD system can detect a fault at a level more than threshold, the system or machine will be repaired. The threshold means, of course, a fault level that a system or machine must be repaired. It should be determined to minimize the life cycle damage.

3.2 ASSUMPTIONS

The following rational hypotheses are proposed for fault detection and fault repair processes to model their process theoretically.

1) Fault state can be described discretely; L_0, L_1, $L2$, ..., L_i...., L_h, L_c. Here, L_0 is completely normal (no fault) state, L_h is a level previous to completely failure state, and L_c is completely failure state.

2) Fault will change from a level L_i to its arbitrary upper level L_j with a transmission probability, p_{ij}, every time step for detecting, ΔT. Here, the following constraint can be composed.

$$\sum_{j=i}^{c} p_{ij} = 1, \quad p_{ij} = 0 \ (j \prec i) \tag{1}$$

Thus, various kinds of fault progress such as drift and abrupt change to "failure" can be rationally reproduced. Transmission probability of the abrupt change, p_{ic}, will be very small but not zero.

3) If fault level is over L_e, a machine or system that a BOFD system is applied may waste energy comparing to no faulty operation. On the other hand, if fault level is less than L_e, no wastefulness may happen. The amount of wasted energy, which can be finally converted to money, will be assumed to be proportioned to its fault level. Thus, the energy damage when fault level is L_i ($L_i > L_e$) can be represented as D_{ei} ($/h).

4) If fault level is over L_a, a machine or system may not maintain indoor environmental quality within a comfort

L_o : completely nomal state, L_c : completely failure state, L_h : a state previous to L_c, L_e : lower limit of fault that give energy wastefulness, L_a: lower limit of fault that give indoor environmental aggravation, L_d : threshold where a fault should be repaired.
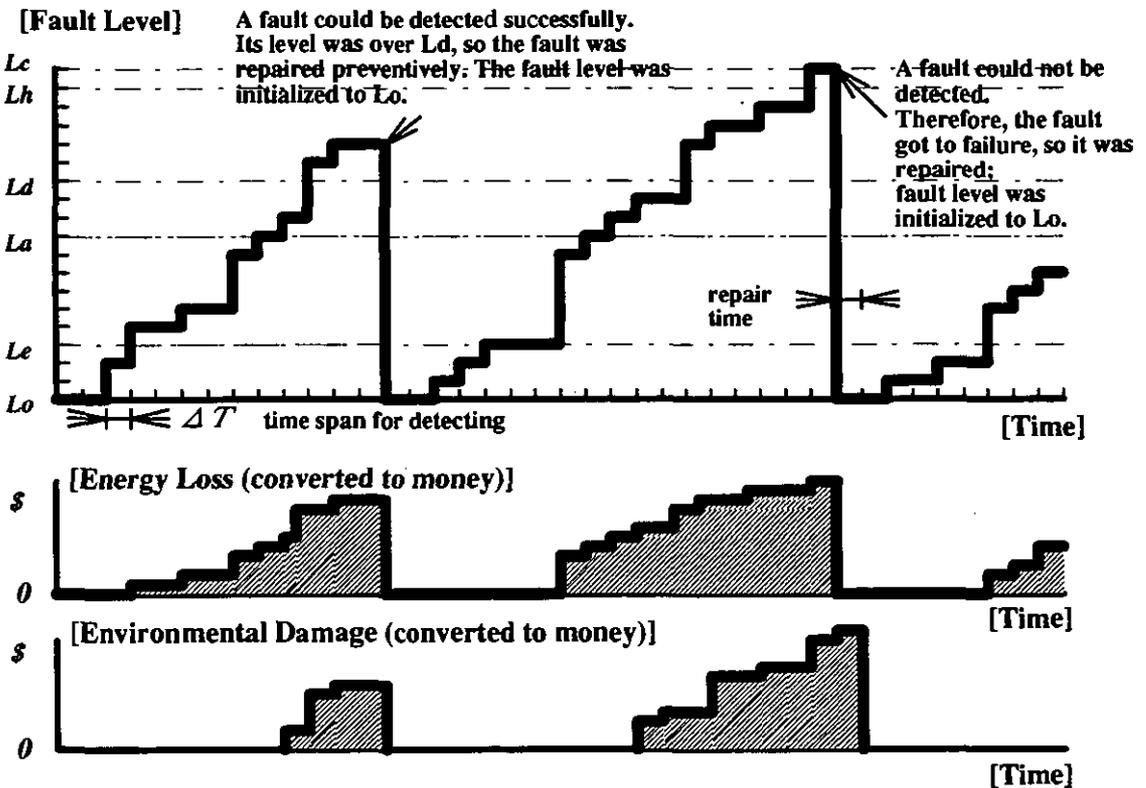

Figure 1 Sketch of process of fault detection and repair

criteria; concretely speaking, indoor air temperature and/or humidity may miss the comfort zone. If fault level is less than L_a, indoor air quality could be maintained within a comfort criteria because of automatic action of control devices. It is now assumed that the discomfort due to such a fault level can be converted to "money". Of course, if fault state gets "failure" with defeat of fault detection, it will bring huge damage to occupants' workability. The amount of discomfort will be dependent on the fault level. Thus, environmental damage when fault level is L_i ($L_i > L_a$) can be represented as D_{ai} (\$/h). In fact, it will be very difficult to identify D_{ai}. In the developing described later, it is assumed that the time consumed for recovering fault and failure can be ignored.

5) Accuracy for detecting a fault in a BOFD system will be not properly perfect. The probability to succeed detecting L_i level's fault is replaced as q_i. High-grade's BOFD system will have naturally high probability for q_i, while it will become expensive. Thus, the trade-off could be consisted between cost and benefit of a BOFD system. To identify the probability variable is important to evaluate a BOFD system precisely.

6) If the BOFD system can detect a fault more than level L_d, the system or machine is repaired. Thus, L_d is threshold, a system or machine must be repaired. To identify this value is very important, and it should be determined to maximize life cycle cost/benefit of a BOFD system. After the fault is repaired, the fault level could be restored to no fault level, L_0. Furthermore, the cost to repair a fault at level L_i is defined as D_{di}.

7) If the BOFD system can not detect a fault more than level L_d continuously, the system or machine will get to "failure", i.e., L_c. Consequently, the failure must be certainly repaired. After the failure is repaired, fault level can be restored to no fault level, L_0. Furthermore, the cost to repair a failure at level L_c is defined as D_c.

4 DEFINITION OF LIFE CYCLE DAMAGE

4.1 CONCEPT OF LIFE CYCLE DAMAGE

Energy wastefulness and environmental aggravation owing to a fault are definitely recognized as damages. Cost to repair a fault or failure can be also interpreted as a kind of damage. Investment for a BOFD system can be also interpreted so, since the investment is no more than loss of money for us. If high grade's BOFD system is applied, the first and second damages will decrease, while the third damage will increase. Therefore, we can perceive the trade-off relationship among those.

On the other hand, these damages should be evaluated in the life cycle of a building and/or HVAC system. Thus, a BOFD system can be evaluated by "Life Cycle Damage (LCD)", which is of course the concept extended from "Life Cycle Cost (LCC)". As stated earlier, it is very important to identify optimal grade of BOFD system and threshold to be repaired from the viewpoint of building optimization.

Such an identification can be attained to minimize the LCD. Basic concept to minimize LCD can be described as follows:

$minimize$: $Is + Ir + Ea + Ee + Ed + Er$ $(=LCD)$ (2)

    Is : Initial cost of a BOFD system (\$)

    Ir : Life cycle cost for maintenance of the BOFD system itself (\$)Ea : Life cycle damage of environmental aggravation (\$)Ee : Life cycle damage of energy wastefulness (\$)Ed : Life cycle cost for preventive maintenance with fault detection (\$) Er : Life cycle cost for post maintenance with failure (\$)

722

Is of a BOFD system generally consists of hardware, software, and sensors. When the BOFD is installed in BEMS, hardware cost may be neglected. *Ir* is mainly to calibrate and/or replace sensors. Durability of sensors is generally very low, so that they should be calibrated and/or replaced periodically many times for the building life. Thus, high reliable BOFD system must require higher *Is* and *Ir*.

4.2 INITIAL COST AND LIFE CYCLE MAINTENANCE COST OF BOFD SYSTEM

Is of a BOFD system will be closely related to the probability of succeeding fault detection, q_i. Higher reliable BOFD system will need greater cost for developing software and more sensors. Thus, *Is* can be represented as a function of q_i ($i=1,h$) as follows:

$$Is = Is\ (\ q_1,\ q_2,........,q_i,.....,q_h\ ) \tag{3}$$

$q_1, q_2,........,q_i,.....,q_h$ is synthetically described as "q_m" hereafter. It could be calculated by simple average, i.e., $q_m=(q_0+q_1+q_2+.......+q_i+.....+q_h)\ /\ h$. *Ir* is fundamentally for calibrating and/or renewing sensors, occasionally for improving hardware and/or software. Here, the following function could be composed.

$$Ir\ (q_m) = Dr\ (q_m)\ TL\ /\ Tr \tag{4}$$

    TL : life cycle of a building or HVAC system that a BOFD system is applied (h)

    Tr : mean time of proofing and/or replacing of sensors (h)

    Dr : costs for proofing and/or replacing of sensors ($/time)

Dr could be represented as a function of detecting reliability of BOFD system, q_m. *Tr* may be dependent on q_m since higher reliable system shall need shorter interval for replacing.

5 LIFE CYCLE DAMAGE BASED ON PROBABILITY PROCESS

5.1 PROBABILITY PROCESS

Time scale is discrete by the time step of fault detection, ΔT, and discrete time is represented as 1,2, ..., n, ... *TL*. Now, some probability variables are defined as follows in a discrete time, *n*:

    $p_i^{\ n}$: probability that a system or machine is at level, L_i, of fault state

    $q_d^{\ n}$: probability that fault is detected and repaired preventively

    $q_r^{\ n}$: probability that fault state get to completely failure and is repaired

    q_{ij} : probability that fault will change from level L_i to L_j after ΔT (independent on *n*)

    q_i : probability that BOFD system can detect fault of level L_i

$p_i^{\ n}, q_d^{\ n}$, and $q_r^{\ n}$ at *n* time are calculated using previous time's several probability variables as follows:

i) The probability that a system or machine is at level L_i of fault state at time $n, p_i^{\ n}$:

$$p_i^n = \left(q_d^{n-1} + q_r^{n-1}\right)p_{oi} + \left(1 - q_d^{n-1} - q_r^{n-1}\right)\sum_{j=0}^{i} p_j^{n-1} p_{ji} \qquad i=0,1,2,.....h \tag{5}$$

ii) The probability that fault is detected and repaired preventively at time $n, q_d^{\ n}$

$$q_d^n = \left(q_r^{n-1} + q_d^{n-1}\right)\sum_{i=d}^{h} p_{oi}q_i + \left(1 - q_r^{n-1} - q_d^{n-1}\right)\sum_{i=d}^{h}\sum_{j=0}^{i} p_j^{n-1}p_{ji}q_i \tag{6}$$

iii) The probability that fault state get to completely failure and is repaired at time n, q_r^n

$$q_r^n = \left(q_r^{n-1} + q_d^{n-1}\right)p_{oc} + \left(1 - q_r^{n-1} - q_d^{n-1}\right)\sum_{j=0}^{c} p_j^{n-1}p_{jc} \tag{7}$$

These probability process in relation to progress of a fault could be defined as a steady-state process. For example, the Markoff Chain theory can be applied to solve them.

5.2 DAMAGE DUE TO ENERGY WASTEFULNESS AND ENVIRONMENTAL AGGRAVATION

The damage as energy wastefulness at time n, E_e^n, can be calculated as the expected value using the probability variables as above defined and the damage rate per an hour, $D_{e\,i}$.

$$E_e^n = \left(q_d^{n-1} + q_r^{n-1}\right)\sum_{i=e}^{h} p_{oi}D_{ei}\Delta t + \left(1 - q_d^{n-1} - q_r^{n-1}\right)\sum_{i=e}^{h}\sum_{j=0}^{i} p_j^{n-1}p_{ji}D_{ei}\Delta t \tag{8}$$

The damage as indoor environmental aggravation at time n, E_a^n, can be also calculated as the expected value using the probability variables as above defined and the damage rate per an hour, $D_{a\,i}$.

$$E_a^n = \left(q_d^{n-1} + q_r^{n-1}\right)\sum_{i=a}^{c} p_{oi}D_{ai}\Delta t + \left(1 - q_d^{n-1} - q_r^{n-1}\right)\sum_{i=a}^{c}\sum_{j=0}^{i} p_j^{n-1}p_{ji}D_{ai}\Delta t \tag{9}$$

5.3 COSTS FOR REPAIRS OWING TO DETECTED FAULT AND UNDETECTED FAILURE

The cost of repairing a machine or system when a fault can be detected at time n, E_d^n can be calculated using the probability that fault is detected at time n and the cost rate per time, D_d.

$$E_d^n = q_d^n D_d \tag{10}$$

The cost of repairing a machine or system when a fault can not be detected at time n and it gets to the failure, E_r^n, can be calculated using the probability that fault is not detected and gets to the failure at n time and the cost rate per time, D_r.

$$E_r^n = q_r^n D_r \tag{11}$$

6 DECISION MAKING BASED ON LIFE CYCLE DAMAGE

6.1 LIFE CYCLE DAMAGE

Finally, the life cycle damage (LCD) can be basically represented, via total checking times throughout building's life by BOFD system, $L(=T\,L\,/\Delta T)$ as follows:

$$LCD = I_s + I_r + \sum_{n=0}^{L}\left(E_a^n + E_e^n + E_d^n + E_r^n\right) \tag{12}$$

Actually, interest rate and inflation rate about energy cost must be included in the equation (12). In the description, however, they are omitted because equation become to too complicate. Of course, they were involved in the case study described later.

6.2 DECISION MAKING PROCESS

To identify optimal grade of BOFD system to be installed and threshold to be repaired are important from the viewpoint of the building optimization. The former element can be represented by global probability to succeed detecting a fault, q_m. The latter can be also represented by a fault level that repair should be conducted, L_d. Therefore, we can obtain optimal values of q_m and L_d in order to minimize equation (12), i.e., LCD. If this thesis can be replaced by extremum problem, the optimum values may be derived conceptually from the following equation.

$$\frac{\partial LCD}{\partial q_m} = 0, \quad \frac{\partial LCD}{\partial L_d} = 0 \tag{13}$$

6.3 OTHER IMPORTANT SUBJECTS TO DEVELOP DECISION MAKING PROCESS

There are many subjects remained, which must be clarified to concretely develop the decision making process and realize it in actual application.

1) **Identifying transition probability, p_{ij}** : This subject is very important not only to reproduce fault advancement for calculating LCD but also to analyze basic properties of fault and/or failure in the Annex 25's work. For that purpose, field data will be able to be available.

2) **Identifying cost-converted damage due to environmental aggravation when level L_i of fault is generated, $D_{a\ i}$**. This factor may be most difficult to be identified, since it is closely related to human's thermal comfort. Thermal comfort is often hard to evaluate quantitatively and give strong incentive to evaluation of building/HVAC system. The trial of quantifying this was tried in the reference [1].

3) **Identifying cost-converted damage due to energy wastefulness when level L_i of fault is generated, $D_{e\ i}$** : This factor will be relatively easy to be quantified, since energy can be easily converted to money. Many "building and HVAC system simulation codes" such as HASP/ACSS/8501, DOE 2/3, and BLAST can be applied to estimate these values.

7 CASE STUDY OF OPTIMIZED PREVENTIVE MAINTENANCE

7.1 OUTLINE AND SCOPE OF CASE STUDY

Although there are many subjects to be clarified to concretely develop decision making process and realize it in actual application, case study was conducted to confirm mathematical propriety of the methodology to optimizing preventive maintenance. In addition, sensitivities of some important factors to the optimum grade (q_m) and the threshold (L_d) were analyzed. Standard conditions were given as shown in **Table 1**. Eight stages were assumed for discrete fault levels including completely normal and completely failure. The transition probabilities, money-converted damages, maintenance cost, and investments to BOFD system were assumed based on information of some skillful engineers. The life cycle damage was evaluated based on the Present Value Method, in which future's income and outcome are rebated by interest rate and inflation rate. The

equation (12) was also solved numerically under steady-state probability process.

7.2 AMOUNT OF LCD

First, variation of amount of LCD according to two parameters, q_m and L_d, was calculated under the conditions shown in **Table 1**. **Figure 2** shows its result. **Figure 3** also shows contents of LCD when q_m was 0.3. The differential between the highest LCD and the lowest LCD in this situation was approximately $180,000. Thus, combination of the grade of BOFD system (q_m) and the threshold for preventive maintenance (L_d) had large influence to amount of money in the building's life.

The optimum combination giving the lowest LCD was q_m=0.3 and L_d=2. This standard situation was when environmental damage and after-failure maintenance cost were estimated at relatively expensive values. In other words, to conduct preventive maintenance at relatively slight fault level was effective under such a heavy damage impacted case. So higher grade BOFD system was not needed under this situation.

8 CONCLUSIONS

The most important purpose of this paper was to appeal necessity of developing evaluation methods of BOFD system. Whatever great BOFD tool is quite unnecessary if building or HVAC system has high reliability and/or durability. Therefore, incentive of introducing BOFD system must be rationally evaluated based on frequencies of fault and/or fault and their damages.

Table 1 Calculation Conditions in Case Study

Outline of Building	Total Floor area : 10,000 m^2, Use for Office
Objective of BOFD	Motor-driven chiller (appropriate size and cost of chiller was assumed)
Discrete Fault Levels	Eight levels : $L_o \sim L_7$ (L_o : normal, L_7 : Failure)
Detective Interval	1 [hour]
Transition Probabilities of fault	$p_{i,i+1}$=3x10^{-8} , $p_{i,j}$=1x10^{-8} ($j \geq i+2$) , $p_{i,i}$=1-$p_{i,j}$(i>j),
Investment and Service Life of Fundamental Parts of BOFD system	Investment : 100,000 x q_m [$], Service Life : 50 years (Here, q_m is the mean probability of succeeding fault detection)
Investment and Service Life of Additional Parts of BOFD system	Investment : 15,000 x q_m [$], Service Life : 10 years (Here, q_m is the mean probability of succeeding fault detection)
Damages to Indoor Thermal Environment According to Fault Level (converted to money)	6,000 (L_7), 3,000 (L_6), 1,500 (L_5), 750 (L_4), 0(L_3 and less) [$/h]
Damages to Wasted Energy Consumption According to Fault Level (converted to money)	200 (L_6), 100 (L_5), 50 (L_4), 25 (L_3), 0 (L_2 and less) [$/h]
Cost needed for Preventive Maintenance	7,000 [$/time]
Cost needed for After-failure Maintenance	30,000 [$/time]
Building Life	100 [years]
Commercial Interest Rate	5 [%/years]
Inflation Rate	4 [%/years]
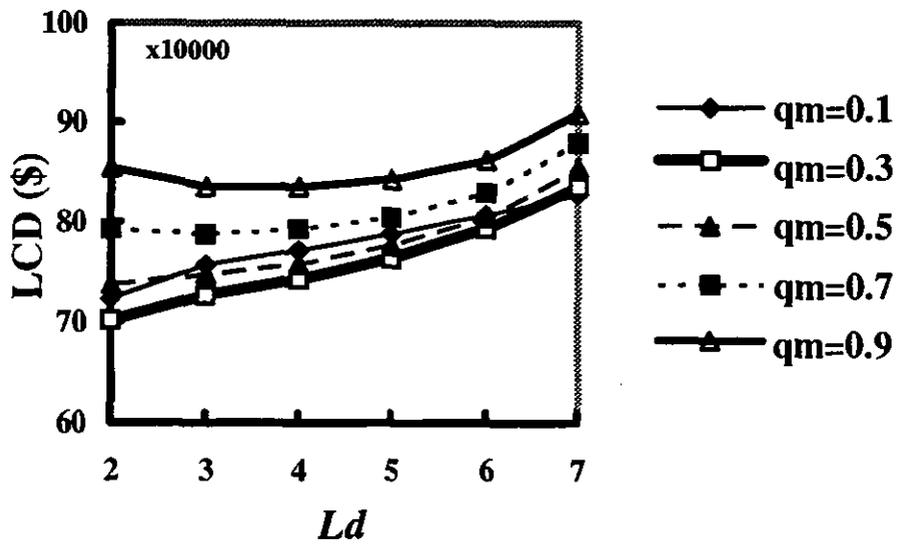
Figure 2 Calculation Result of LCD

Er : life cycle after-failure maintenance cost, Ed : life cycle preventive maintenance cost, Ea : life cycle environmental damage, Ee: life cycle energy damage, Ir: life cycle investment to additional parts in BOFD system, Is : life cycle investment to fundamental parts in BOFD system
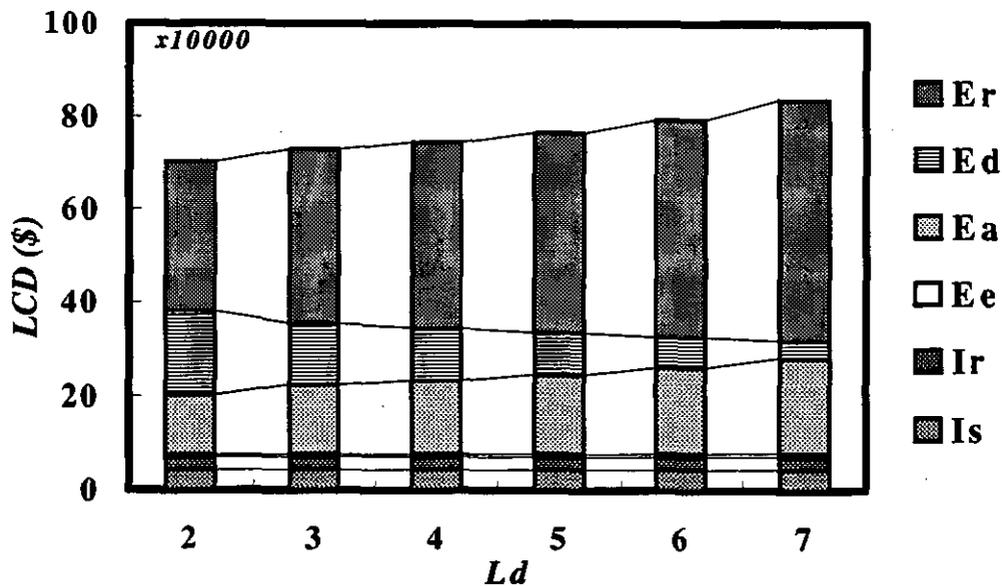
Figure 3 Contents of LCD (when q_m=0.3)

The present paper developed new philosophy and innovative and simple process for evaluating BOFD system and preventive maintenance strategy based on the life cycle damage, in which probability process of fault's outbreak was also considered. Furthermore, the present paper conducted a case study using the process. Although there are many subjects to be solved, this new methodology will become effective tool for decision making of introducing BOFD system and planning preventive maintenance strategy.

References

[1] H. Ito, "Identifying threshold of fault to be repaired based on life cycle building damage", IEA/Annex 25 6th meeting in Boston, 1993.

[2] H. lto, "Optimizing preventive maintenance based on fault detection/diagnosis in HVAC system", IEA/Annex 7th Meeting in Stuttgart, 1994

A Statistical, Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners

Todd M. Rossi and James E. Braun

Ray W. Herrick Laboratories
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907

This paper presents a method for automated detection and diagnosis of faults in vapor compression air conditioners that only requires temperature measurements, and one humidity measurement. The differences between measured thermodynamic states and predicted states obtained from models for normal performance (residuals) are used as performance indices for both fault detection and diagnosis. For fault detection, statistical properties of the residuals for current and normal operation are used to classify the current operation as faulty or normal. A diagnosis is performed by comparing the directional change of each residual with a generic set of rules unique to each fault. This diagnostic technique does not require equipment-specific learning, is capable of detecting about a 5 % loss of refrigerant, and can distinguish between refrigerant leaks, condenser fouling, evaporator fouling, liquid line restrictions, and compressor valve leakage.

INTRODUCTION

Automated fault detection and diagnostics (FDD) for HVAC systems has the potential to reduce energy and maintenance costs and improve comfort and reliability. Sensors are chosen and located strategically within the system so that their outputs are sensitive to the faults for which the FDD system is designed. The appropriate number and type of sensors results from a tradeoff between initial cost and performance of the FDD system. A micro-processor is used to process the measurements in order to provide the fault detection and diagnostic capability, and also may be used to provide communications with a central service facility.

FDD Background

As described by Isermann (1984) and depicted in Figure 1, a FDD system may include the following steps: fault detection, fault diagnosis, and fault evaluation. Fault detection indicates a deviation of performance from expectation, diagnosis determines the cause of the fault, and evaluation assesses whether the impact is severe enough to justify service. In each of these steps, it is necessary to define criteria or thresholds for establishing

appropriate outputs. The outputs would be fault or no fault for fault detection, the type of fault for diagnosis, and repair or don't repair for the fault evaluation step.

Fault detection is accomplished by comparing performance determined from measurements with some expectation of performance. If the deviation exceeds a threshold, then a fault would be indicated. As depicted in Figure 2, this process may be described in two steps: preprocessing and classification. The preprocessor takes measurements from sensors and manipulates them to generate features for classification. Classifiers then operate on the features to determine whether the system contains a fault.
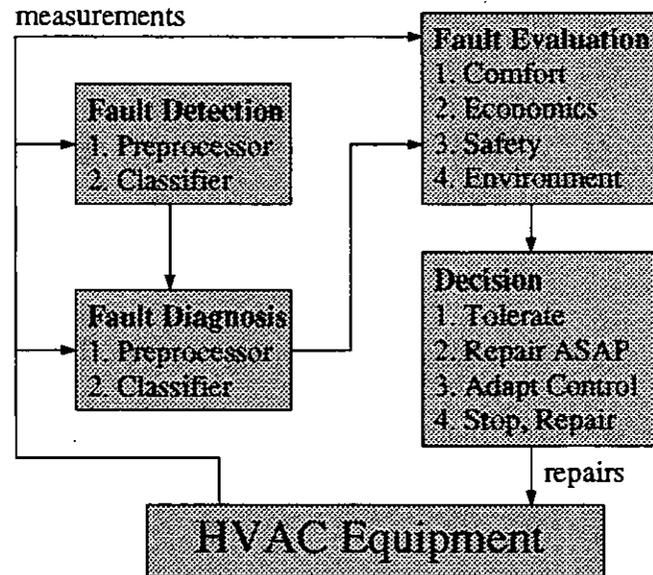
Figure 1. Supervision of HVAC&R Equipment

Figure 2. Sequential Steps in Fault Detection and Diagnosis

Simple transformations, characteristic quantities, and models are three types of processors. Simple transformations include the identity transformation (i.e., no preprocessing) and trend generation (i.e., time derivatives). Characteristic quantities are features that are computed directly from measurements and that are indicative of component performance. Examples include overall system efficiencies and heat exchanger

effectiveness. Model-based preprocessors utilize mathematical models of the monitored system to generate features. Model parameters could be learned from measurements when the system is operating normally, or determined using physical models. The features used by the classifier from model-based preprocessors could be the differences between measured and modeled performance (i.e., residuals) or physical parameters of the model.

In a broad sense, the classifier is an expert system. The knowledge necessary to make a fault decision can be stored in a number of forms, including: a set of production rules (i.e. IF, THEN, ELSE rules), a fault tree, and conditional probabilities for statistical pattern recognition classifiers. Typically, it is necessary to assign the thresholds for deviations between current and normal performance that constitute faults. In selecting thresholds, there is a tradeoff between detection sensitivities and false alarm rates. Tighter thresholds result in greater sensitivities (detection of smaller faults), but will lead to more false alarms (an indication of a fault that doesn't exist). Thresholds are often determined based upon heuristics, although better performance (lower ratio of false alarms to correct diagnoses) is achieved when statistical thresholds are employed.

In general, preprocessing simplifies the classification and improves overall performance of the FDD system. In the absence of any preprocessing (identity transformation), the FDD system is a classic expert system. All fault detection is then based upon rules that act directly on the measurements. Consider a vapor compression system that uses condenser head pressure to detect faults. Without preprocessing, a simple check for excessive head pressure could be: If the head pressure is greater than 425 psig (29.5 bar), then a fault exists. Since the head pressure varies under normal operation with the ambient temperature, the fault detection threshold must be greater than the highest head pressure associated with normal operation. A more complex expert system might contain a set of rules with different head pressure limits for different ambient temperatures.

Alternatively, a model-based preprocessor could model the relationship between head pressure and ambient temperature under normal operation. Then, a fault would be identified if the deviations between measured and modeled head pressures exceeded a specified threshold. The FDD system with the model-based preprocessor would be significantly more sensitive to abnormal behavior than the single rule system and easier to implement than the expert system with many rules. The thresholds for allowable deviations could be established by evaluating the statistical properties of the measurements, and how well the model for normal operation fits the measurements.

The structure of Figure 2 can also be used to describe fault diagnosis. Measurements are processed in order to simplify the classification required to identify the particular

component at fault. The overall classification problem is different for fault diagnosis than fault detection in that the decision is not binary (i.e., fault/ no fault): the classifier must choose the specific fault from a list of possibilities. However, the diagnostics problem can be reduced to a series of fault detection problems through fault isolation.

With fault isolation, fault detection methods are applied to individual components for which diagnoses are desired. For instance, condenser fouling in an air conditioner could be detected by estimating the heat exchanger effectiveness from measurements on the condenser. The fault is diagnosed as soon as it is detected and no additional classification is necessary. The disadvantage of fault isolation is the large number of measurements required. The diagnosis of heat exchanger fouling would require measurements of all states entering and leaving the heat exchanger.

Another diagnostic approach involves comparing physical parameters determined from measurements with values representative of normal operation. For instance, heat exchanger conductance could be estimated from entering and leaving conditions and used to diagnose fouling. Here again, fault detection and diagnosis are combined and no separate diagnostic classification is necessary.

A more common diagnostic approach that requires fewer measurements involves the use of fault models. For each type of fault to be diagnosed, a fault model predicts the outputs associated with the occurrence of that fault for a current set of inputs. The fault is diagnosed through the use of a classifier that attempts to find the fault model with the best representation for the current behavior. The advantage of fault modeling for diagnosis is that fewer measurements are required. However, it is necessary to have fault models for each fault and combinations of faults to be diagnosed.

Fault evaluation follows fault detection and diagnosis and requires an evaluation of the impact of a fault on system performance. Without this step, it must be obvious that the benefit of servicing the fault justifies its expense. This is the case for many "hard" failures, such as broken fan belts or seized compressors. However, fault evaluation is necessary for many performance degradations, such as heat exchanger fouling, where the fault could be detected and diagnosed well before the need for service. For HVAC applications, Rossi and Braun (1996) defined four criteria for evaluating the need for service: comfort, safety, environmental, and economic. In general, service should be performed whenever: 1) comfort cannot be maintained, 2) equipment or personal safety is compromised, 3) environmental damage is occurring (e.g., refrigerant leakage), or 4) reduced energy costs justify the service expense.

Applications of FDD to Vapor Compression Equipment

There is a large body of literature on fault detection and diagnostic techniques for applications in critical processes. As the cost of hardware (e.g., sensors, microprocessors) has gone down, interest in developing FDD systems for HVAC&R applications has increased. Most of the literature for HVAC&R applications has focused on "hard" failures for large central chilled water distribution systems and air handling systems. The literature for fault detection and diagnosis for vapor compression equipment is relatively sparse but includes contributions by McKellar (1987), Stallard (1989), Yoshimua and Noboru (1989), Kumamaru et al. (1991), Wagner & Shoureshi (1992), Hiroshi et al. (1992), Grimmelius et al. (1995), and Rossi & Braun (1996).

McKellar (1987) identified many of the common faults for home refrigerators and investigated the effects of several faults on thermodynamic measurements within the vapor compression cycle. The faults that he considered were compressor valve leakage, heat exchanger fan failures, frost on the evaporator, partially blocked capillary tube, and refrigerant charging failures. McKellar found that each of these faults had unique effects on three measures: suction pressure (or temperature), discharge pressure (or temperature), and discharge-to-suction pressure ratio and concluded that these measures were sufficient for developing a FDD system. He did not develop a general approach to characterizing expectations for these measurements (i.e., a model-based preprocessor) nor did he discuss thresholds for fault detection and diagnostic classifiers.

Based upon the work of McKellar, Stallard (1989) developed an expert system for automated FDD applied to refrigerators. Condensing temperature, evaporating temperature, condenser inlet temperature, and the ratio of discharge-to-suction pressure were used directly as classification features (i.e., no model-based preprocessor). Feature limit checking, the simplest of rule-based classifiers, was used for both detection and diagnostic classification. Fault diagnoses were performed by evaluating the direction in which classification features changed from expected values and matching these changes to expected directional changes associated with each fault (when they exceeded the fixed threshold). Different rules were used for each of three discrete ranges of ambient temperature.

Wagner and Shoureshi (1992) used a different approach to perform FDD for refrigerator faults. Dynamic, nonlinear state estimation techniques were used to generate residuals between current and expected states. Compressor shell temperature, condensing temperature, and compressor power were measured system responses, and ambient temperature was a measured model input. Experiments were used to develop dynamic

fault models for each of the faults considered. On-line measurements were statistically compared with normal and fault models in order to perform diagnostic classification.

Kumamaru et al. (1991) used characteristic curves to obtain quantitative expectations for heat pump performance as a function of cooling water temperature and loading. Diagnostics were performed using residuals as input features. The method did not utilize statistically based thresholds and did not detect performance degradations.

Yoshimua and Noboru (1989) used a combination of seven temperature and pressure measurements to perform FDD for packaged air conditioners. Their method used rules with fixed thresholds to perform detection and diagnosis. They did not utilize any preprocessing or statistical rule evaluation.

Hiroshi et al (1992) developed a refrigerant leak detection method for automotive air conditioners that used a measurement of the liquid to gas flow ratio in the liquid line. The method did not utilize any model-based preprocessing to account for the effects of ambient temperature and load conditions on expectations for this measurement and utilized fixed thresholds. As a result, the method could only detect refrigerant loss with a sensitivity of about 50% of full charge.

Grimmelius et al. (1995) used differences between measurements and outputs of steady-state models for expected behavior as input features for detection and diagnoses of chiller faults. The method used approximately 20 measurements, including temperatures, pressures, power consumption, and compressor oil level. Diagnoses were performed using a pattern recognition technique applied to the current residuals and a matrix of expected residual changes associated with each possible fault. The fault matrix was determined using experiments on a chilled water system to which faults had been introduced.

Rossi and Braun (1996) addressed the issue of fault evaluation for air conditioning equipment. They developed a simplified method for estimating the optimal service times that minimizes combined energy and service costs for cleaning condensers and evaporators in air conditioners. They also compared the costs associated with optimal maintenance scheduling with those associated with regular maintenance and with a procedure where service was performed based only on the comfort and safety criteria (constrained service). Savings of between 5 and 15% of operating costs were possible through optimal maintenance scheduling.

Although each of the previous studies provided contributions, all of the methods have limitations. In some cases, expensive measurements are required (e.g., mass flow rate, power consumption, pressure), while in others, fault detection and diagnostic sensitivity could be significantly improved with the use of model-based preprocessing and

statistically-based thresholds. The use of system specific fault models could require extensive experimentation for each possible fault, particularly when performance degradation faults are considered. None of the previous studies included results for sensitivities of the FDD methods in detecting and diagnosing faults.

Scope of this Study

This paper describes the development and evaluation of a new method for detecting and diagnosing faults in air conditioning equipment that only requires temperature and humidity measurements. The diagnostic approach is based on generic rules and does not require equipment specific experimentation. Thresholds for both fault detection and diagnosis are based upon statistical analysis of on-line measurements. Five distinct faults were considered: 1) refrigerant leakage; 2) liquid line restriction; 3) leaky compressor valves; 4) fouled condenser coil; and 5) dirty evaporator filter. Only the fault detection and diagnostic steps were considered. Further work would be necessary to establish methods for fault evaluation. One primary goal was to identify the sensitivity of the algorithm in detecting and diagnosing each of these faults.

Refrigerant loss detection could be an immediate application for the proposed FDD method, since fault evaluation is not necessary for this fault. A refrigerant leak should be repaired as soon as it is detected and diagnosed. As a step towards implementation, the impact of the number of sensors on the sensitivity for detecting and diagnosing refrigerant leakage was also studied.

The performance of FDD methods for this study was evaluated using a combination of simulations and laboratory experiments. Diagnostic rules were developed through simulation and checked within the laboratory. The sensitivities of the FDD method for detecting and diagnosing each of the five faults were estimated through simulation. Both simulations and experiments were used to investigate the impact of the number of sensors on FDD performance for detecting and diagnosing refrigerant leaks.

EVALUATION TOOLS

For the simulations, a vapor compression system model developed and validated by Rossi (1995) was utilized. It is a modular, steady-state model that solves mass, energy, and momentum balances for any set of entering air conditions. A steady-state model is appropriate because the fluid flow and heat transfer dynamics are generally much faster than the dynamics of the load and ambient conditions. The model allows the introduction of all of the faults considered in this study.

A fully-instrumented, three ton rooftop air conditioner was used for testing FDD algorithms. The system has fixed-speed condenser and evaporation fans, a fixed orifice expansion device, and a single-stage, on/off controlled reciprocating compressor. Figure 3 illustrates the manner with which faults were simulated in the test unit: 1) condenser fouling: paper was placed on the air-side of the coils, 2) evaporator filter fouling: paper was placed on the air-side filter, 3) leaky compressor valves: this effect was represented using a hot gas bypass line with a manual valve, 4) liquid line restriction: a manual valve was located in the liquid line before the expansion device, and 5) refrigerant leakage: refrigerant charge removal was controlled by a valve located on the high pressure side of the unit and monitored with a scale.

All refrigerant cycle temperature measurements were made using K-type thermocouples mounted to the exterior surfaces of refrigerant piping and insulated. On the air side, temperatures of entering and leaving air were measured using platinum RTDs for both the evaporator and condenser streams. The thermocouples and RTDs were calibrated by immersion in a mixed ice bath. The measured error for each sensor was used for an offset correction and to estimate the accuracy of the measurement. The error of all the thermocouple measurements was less than 0.4 C, while the RTDs were within 0.3 C. The relative humidity of air entering the evaporator was determined from measurements of dew point using a chilled mirror dewpoint hydrometer. This device is accurate to within 0.5 C and results in an error in relative humidity that is less than 0.05 for the range of conditions considered in this study.
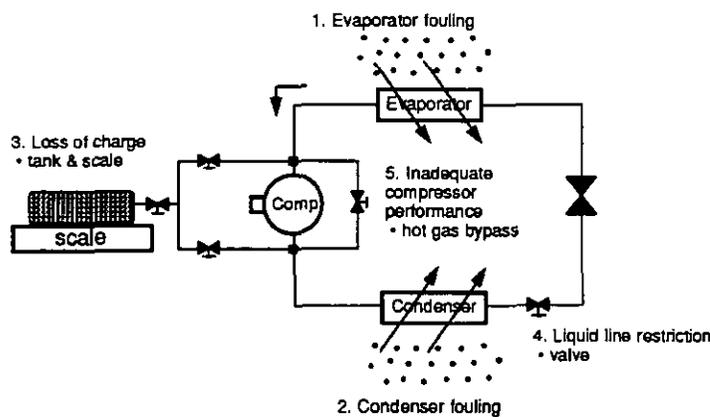

Figure 3. Rooftop Air Conditioner with Simulated Faults

Figure 4 shows a block diagram of the fault detection and diagnostic system in relation to input and output "measurements" from a vapor compression cycle. For fixed-speed fans, an air-to-air vapor compression cycle is driven only by the air inlet states into the evaporator and condenser (U). These states are characterized by the ambient temperature (T_{amb}) and the return air temperature (T_{ra}) and relative humidity (ϕ_{ra}) of air stream entering the evaporator. In steady-state, every internal state of a normally operating cycle depends solely on the driving conditions. In this investigation, the performance of the cycle was characterized using a vector of temperature measurements (Y) only. At most, seven temperatures were considered: 1) T_{evap}, evaporating temperature; 2) T_{sh}, suction line superheat; 3) T_{cond}, condensing temperature; 4) T_{sc}, liquid line subcooling; 5) T_{hg}, hot gas line or compressor outlet temperature; 6) ΔT_{ca}, air temperature rise across condenser; and 7) ΔT_{ea}, air temperature drop across evaporator.

The preprocessor portion of the FDD system contains two major components: a steady-state model and the preprocessor portion of a steady-state detector. The measured inputs are used by the model to predict each measured output for the vapor compression cycle under normal operation (i.e., no faults). The difference between measured and model predictions of the operating states (residuals) are used by the fault detection and diagnostic classifiers for decision making. A steady-state model was appropriate for the faults in this study because the fault development rate was generally slower than the equilibration time of the plant and it was easier to implement than dynamic modeling. Dynamic signatures would be useful for abrupt failures such as broken fan belts.

A steady-state detector is necessary to determine when the model's predictions are valid (e.g., predictions are not valid during startup and shutdown transients). The steady-state detector's preprocessor evaluates the variation in the output measurements for use by a classifier.

The classifier consists of fault detection, diagnostic, and steady-state classifiers. The fault detection classifier operates on the residuals and provides a binary output indicating when the current operating state deviates from expectation with greater than a prescribed statistical confidence. The diagnostic classifier provides the most likely cause of the fault after it is detected. The steady-state detection classifier provides a binary output that is an input to a switch (SW) that controls the output of the FDD system. The FDD system will only indicate a fault and provide a diagnosis when the system is in steady state.
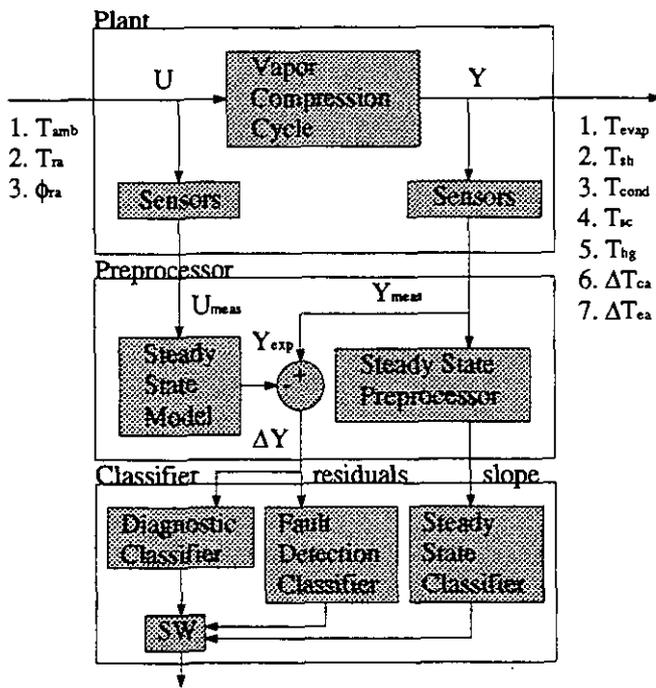
Figure 4. Fault detection and diagnostic system

This study has focused on the development and evaluation of the fault detection and diagnostic classifiers only. The steady-state model was a lookup table that produced perfect predictions of plant output measurements when given the correct measured inputs. Errors in measurements were propagated through the model and used to assess the sensitivity of the FDD system in detecting faults. In practice, imperfect modeling would decrease the sensitivity of the FDD system. Further work would be necessary to develop appropriate on-line models and assess their impact on overall FDD performance.

All simulations and experiments were performed at steady-state conditions in this study. Steady-state detection could be implemented using techniques developed by Glass et al. (1994). These methods are based on estimating the sample standard deviation about the mean of output measurements over a moving exponentially weighted window. Alternatively, the time rate of change in temperature measurements during a moving window could be used in steady-state detection.

Fault Detection Classifier

The features used for classification are the residuals between current state variable measurements and outputs from the models. The classifier identifies a fault when the current measurements are statistically different than the expected values. Figure 5 illustrates how this classifier works for a one-dimensional example. In this case, the

classifier feature is the residual of the suction line superheat. The probabilities of obtaining a specific residual are shown for both normal and faulty operation. Under normal operation, there is a distribution of residuals that results from measurement noise and modeling errors . In the absence of modeling errors and with random noise, the distribution for normal operation would have zero mean. The introduction of afault changes both the mean and/or standard deviation of the residuals.

If both normal and faulty distributions were known a priori, then this would be a classical classification problem. In this case, the current measurement residual would be compared to a threshold value that depended upon the normal and faulty distributions. The minimum error in classifying either normal behavior as faulty or fault behavior as normal would occur for a decision threshold located at the intersection of the two distributions. The probability of making an erroneous classification using this threshold, termed the classification error, would then be the overlap of the two probability distributions, shown as the shaded area in Figure 5. The classification error is the sum of the probability that a normal classification would be faulty and the probability that a faulty classification would be normal. It could be important to consider each type of classification error separately if the consequence of one is more serious than the other.
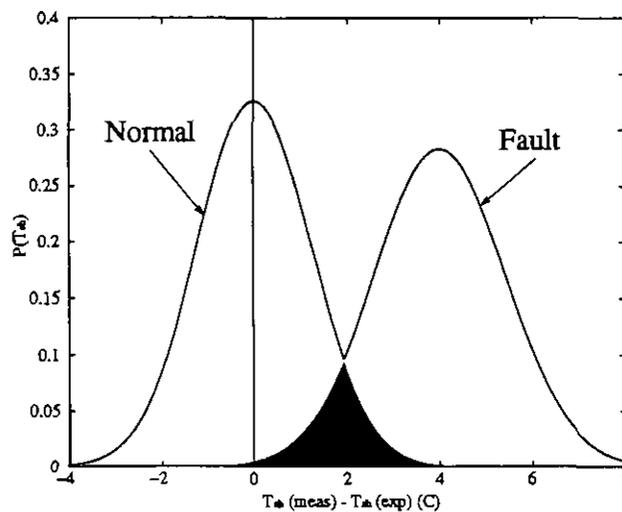

Figure 5. One dimensional example of the detection classifier

This problem differs from a classic classification problem in that the faulty distribution is not known a priori. As a fault progressively becomes worse, the current distribution of measurement residuals moves further away from the distribution for normal operation. A

fault should be indicated when the classification error (shaded area) is small enough for the false alarm rate to be acceptable.

An optimal linear classifier (Fukunaga, 1990) was used for estimating the classification error associated with current and normal distributions of residuals. This is one of several possible statistical classifier designs (Fukunaga, 1990) and assumes that measurement noise is caused by independent random processes, and has a Gaussian distribution. The residuals for both normal and current operation are fitted to a Gaussian model that is completely described in terms of a mean vector and covariance matrix. The assumption of a Gaussian distribution significantly reduces the memory and computational requirements for fault detection. A Bayes classifier for a Gaussian distribution is in general quadratic (Fukunaga, 1990, p.125) and results in the following inequality for evaluating whether a fault should be indicated.

$$\left(\mathbf{Y}-\mathbf{M}_N\right)^T \Sigma_N^{-1}\left(\mathbf{Y}-\mathbf{M}_N\right)-\left(\mathbf{Y}-\mathbf{M}_C\right)^T \Sigma_C^{-1}\left(\mathbf{Y}-\mathbf{M}_C\right)+\ln\frac{\left|\Sigma_N\right|}{\left|\Sigma_C\right|} \geq 0 \qquad (1)$$

where $\mathbf{Y}$ is a vector of current residuals, $\mathbf{M}_N$ and Σ_N are the mean vector and covariance matrix that describe the distribution of residuals in the absence of any faults (i.e., normal operation), and $\mathbf{M}_C$ and Σ_C are the mean vector and covariance matrix that describe the current distribution of residuals determined using recent measurements. For a perfect model, the mean of the residuals would be zero. However, an actual model would likely give biased estimates under certain operating conditions.

A fault would be indicated whenever the left-hand side of equation 1 was greater than zero. This formulation minimizes the probability of making a wrong decision, assuming that there is no built-in bias that favors a "faulty" or "normal" output. The quadratic classifier is reduced to a linear classifier by assuming that the separation between the distributions for current and normal operation is dominated by mean vector differences as opposed to covariance matrix differences. If it were assumed that $\Sigma = \Sigma_N = \Sigma_C$, then Equation 1 would reduce to

$$\left(\mathbf{M}_C-\mathbf{M}_N\right)^T \Sigma^{-1} \mathbf{Y}-\frac{1}{2}\left(\mathbf{M}_N^T \Sigma^{-1} \mathbf{M}_N -\mathbf{M}_C^T \Sigma^{-1} \mathbf{M}_C\right) \geq 0 \qquad (2)$$

In practice, an average covariance matrix, Σ, is determined as the weighted average of Σ_N and Σ_C according to

$$\Sigma = s\Sigma_N + (1-s)\Sigma_C \tag{3}$$

where s is determined by minimizing the classification error (i.e., probability of making an erroneous decision). The classification error is determined by integrating the overlapping areas associated with the multi-dimensional normal and fault distributions using (Fukunaga, 1990, p. 137)

$$\varepsilon = erfc\left(\frac{-\mathbf{V}^T\mathbf{M}_N - \mathbf{v}_o}{\sqrt{2\sigma_N^2}}\right) + erfc\left(\frac{-\mathbf{V}^T\mathbf{M}_C - \mathbf{v}_o}{\sqrt{2\sigma_C^2}}\right) \tag{4}$$

where,

$$\mathbf{V} = \left(s\Sigma_N + (1-s)\Sigma_C\right)^{-1}\left(\mathbf{M}_C - \mathbf{M}_N\right)$$

$$\mathbf{v}_o = -\frac{s\sigma_N^2\mathbf{V}^T\mathbf{M}_C + (1-s)\sigma_C^2\mathbf{V}^T\mathbf{M}_N}{s\sigma_N^2 + (1-s)\sigma_C^2}$$

$$\sigma_N^2 = \mathbf{V}^T\Sigma_N\mathbf{V}$$

$$\sigma_C^2 = \mathbf{V}^T\Sigma_C\mathbf{V}$$

In this study, a fault was indicated whenever the classification error (ε) was below a threshold equal to 0.001. This threshold for classification error gives a small false alarm rate and was found to provide acceptable fault detection sensitivities. The value of s that minimizes the classification error as determined with equation 4 was found using a golden-section search optimization method (Rao, 1984). The mean vector and covariance matrix that describe the distributions for normal and current operation were determined, as outlined in the section on testing and evaluation.

Diagnostic Classifier

Fault diagnosis is performed using the residual features as inputs to a rule-based classifier. The set of rules relates each fault to the direction that each measurement changes when the fault occurs. Table 1 gives the diagnostic rules for the five faults and seven measurements considered in this study. These rules were developed and tested through simulation over a range of operating conditions and tested using experiments at a single operating point. Each of the faults results in a different combination of increasing

or decreasing measurements. The rules are effectively fault models. However, they are generic for all similar types of air conditioners and do not require on-line learning.

Table 1. Rules for the diagnostic classifier

Fault	T_{evap}	T_{sh}	T_{cond}	T_{sc}	T_{hg}	ΔT_{ca}	ΔT_{ea}
Refrigerant Leak	↓	↑	↓	↓	↑	↓	↓
Compressor Valve Leakage	↑	↓	↓	↓	↓	↓	↓
Liquid-Line Restriction	↓	↑	↓	↑	↑	↓	↓
Condenser Fouling	↑	↓	↑	↓	↑	↑	↓
Evaporator Fouling	↓	↓	↓	↓	↓	↓	↑

Figure 6 is an illustration of the fault diagnostic classification method for two possible faults (refrigerant leakage and liquid line restriction) with two input features (superheat and subcooling residuals). The progression of changes in the contours of two-dimensional probability distributions are shown as the two different faults are slowly introduced. Normal operation is shown as the distribution centered at the zero point. As a fault develops, the contour moves along a curve. When the overlap between the normal performance distribution and the current distribution (as indicated by the classification error, ε) is small enough for the false alarm rate to be acceptable (e.g., $\varepsilon < 0.001$), then a fault is signaled by the fault detector. The different diagnostic classes are separated by the axis. The overlap of the current distribution with each of the modeled classes is calculated and represents the probability that the fault class is the correct diagnosis. A diagnosis is indicated when the probability (overlap) of the most likely class is larger than the second most likely class by a specified threshold. As the fault becomes more severe, confidence in the fault detection and diagnosis increases as the current distribution moves further from the normal distribution, and from the axis separating the classes. Again, the choice of a diagnositic threshold results from a tradeoff between diagnostic sensitivity and the rate of false diagnoses.
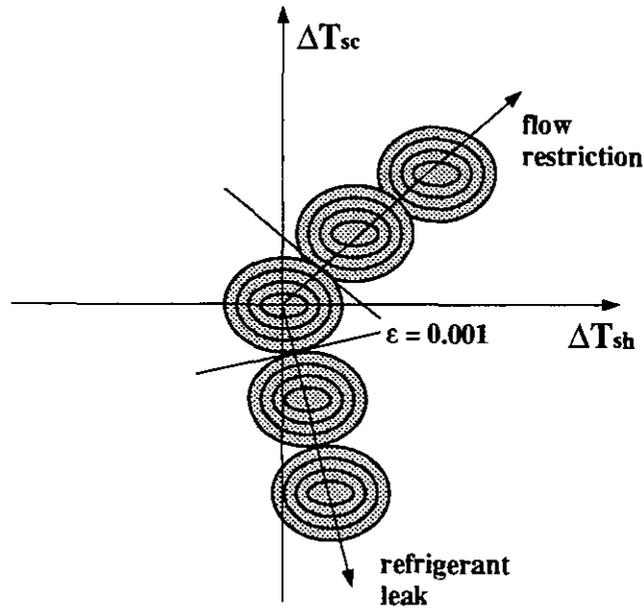
Figure 6. Fault Diagnostic Classifier (Two-Dimensional Example)

In order to perform the classification for diagnostics, the probability that each rule applies to the current operation is evaluated. The probability of each hypothesis is determined by the degree to which the distribution characterizing the current residuals overlaps each class. The overlap is evaluated by integrating the area under the m-dimensional Gaussian probability distribution that falls within each class's region of the domain as given by:

$$w_j = \int_0^{d_{jm}} \cdots \int_0^{d_{j2}} \int_0^{d_{j1}} p(\mathbf{M}_c - \mathbf{M}_N, \Sigma_c) dY_1 dY_2 \cdots dY_m \tag{5}$$

where d_{jk} is the integration limit associated with the domain of fault j in dimension k ($\pm\infty$), Y_k is the kth residual, and $p(\mathbf{M}_c\text{-}\mathbf{M}_N,\Sigma_c)$ is the m-dimensional probability distribution function for the current residuals. For diagnoses, the current distribution has been shifted in order to give zero mean for normal operation. A nonzero residual mean could occur for normal operation with an imperfect model.

The calculation of the overlap within each class is simplified by assuming that each dimension is independent. In this case, the probabilities in each dimension can be "ANDed" together such that:

$$w_j = \prod_{k=1}^{m} \int_0^{d_{jk}} p(\mathbf{M}_c(k) - \mathbf{M}_N(k), \Sigma_c(k,k)) dY_k \tag{6}$$

743

where $p(M_C(k)-M_N(k),\Sigma_C(k,k))$ is the probability distribution function for the current residuals in the kth dimension. For Gaussian distributions, the integration reduces to:

$$w_j = \prod_{k=1}^{m} \frac{1}{2}\left[1 + C_{jk}\, erf\left(\frac{M_C(k) - M_N(k)}{\sqrt{2}\,\Sigma_C(k,k)}\right)\right] \qquad (7)$$

where $C_{jk} = +1$ if $(M_C(k)-M_N(k))$ falls within the domain for the jth fault (i.e., $(M_C(k)-M_N(k))$ has the same sign as defined in Table 1 for the appropriate fault) and $C_{jk} = -1$, otherwise.

In this study, the probability associated with each diagnosis was estimated using Equation 7. A diagnosis was considered valid when the ratio of the probability of the most likely class to the second most likely class was greater than 2. This diagnostic threshold was found to provide good sensitivity and no misdiagnoses.

TESTING AND EVALUATION

In the simulation studies, outputs of the detailed physical model of an air conditioner were used to represent the plant. Sensors were modeled by adding zero mean, independent, and identically distributed Gaussian noise to known values of the plant inputs and modeled outputs. The standard deviations of the temperature and humidity measurements (σ_T and σ_ϕ) were inputs to the sensor model. For most of the results presented in this study, $\sigma_T = 0.5$ C and $\sigma_\phi = 0.05$.

In the experimental investigations, the driving conditions were fixed and the sensor values were recorded for no fault, and then for progressively larger degrees of fault introduction. The no fault measurement was used as the reference for generating the residuals for each fault level. Since the driving conditions were fixed and the system was allowed sufficient time to achieve steady state, there was no need for a steady-state detector.

The input and output measurements were also characterized using Gaussian distributions with both on-line measurements and manufacturers' specifications. The full extent of a sensor's variation is often not captured by sampling a sensor for several minutes to an hour. Examples of issues affecting sensor performance that may not be noticed within an hour include: ambient temperature variations, power cycling, long-term electrical noise and drift, and calibration errors. Therefore, measured variances are often much smaller than the specified accuracies of the sensors. To give more realistic measurement error estimates, the variances were modified by combining the estimated

variance and the specified accuracy as if they were independent, normally distributed noise sources, as follows:

$$\sigma^2 = \sigma^2_{meas} + \sigma^2_{spec} \tag{6}$$

where σ^2 is the overall measurement variance, σ_{meas} is the standard deviation of measurements estimated from the data, and σ_{spec} is the specified accuracy of each sensor. In most cases, σ_{spec} dominates σ_{meas}.

The preprocessor's steady-state model incorporated a lookup table that produced perfect plant predictions when given inputs with no measurement errors. With Gaussian distributed measurement noise, the distribution for the ith temperature residual (difference between measured and predicted outputs) is defined by

$$\Delta Y_i = Y_i(T_{amb}, T_{ra}, \phi_{ra}) - f_i(T_{amb} + w_T, T_{ra} + w_T, \phi_{ra} + w_\phi) + w_T \tag{7}$$

where Y_i is the ith output of the plant, f_i is the preprocessor model's prediction of the plant output for normal operation, w_T is zero mean noise added to the temperature measurements, and w_ϕ is zero mean noise added to the relative humidity measurement. Sensor noise was propagated through the steady-state preprocessor using a first-order Taylor series about the known operating point.

$$\Delta Y_i \approx Y_i(T_{amb}, T_{ra}, \phi_{ra}) - f_i(T_{amb}, T_{ra}, \phi_{ra}) - \frac{\partial f_i}{\partial T_{amb}} w_T - \frac{\partial f_i}{\partial T_{ra}} w_T - \frac{\partial f_i}{\partial \phi_{ra}} w_\phi + w_T \tag{8}$$

The mean vector and covariance matrix were estimated in order to characterize a normal distribution of the residuals. The ith entry in the mean vector is

$$E[\Delta Y_i] \approx Y_i(T_{amb}, T_{ra}, \phi_{ra}) - f_i(T_{amb}, T_{ra}, \phi_{ra}) \tag{9}$$

where E is the expected value operator. The covariance matrix can be determined by using a Taylor series approximation (Gelb, 1989). Diagonal entries in the ith row of the covariance matrix are

$$E\left[(\Delta Y_i - E(\Delta Y_i))^2\right] \approx \sigma_T^2 + \left(\frac{\partial f_i}{\partial T_{amb}}\right)^2 \sigma_T^2 + \left(\frac{\partial f_i}{\partial T_{ra}}\right)^2 \sigma_T^2 + \left(\frac{\partial f_i}{\partial \phi_{ra}}\right)^2 \sigma_\phi^2 \tag{10}$$

whereas off-diagonal entries are

$$E\left[\left(\Delta Y_i - E(\Delta Y_i)\right)\left(\Delta Y_j - E(\Delta Y_j)\right)\right] \approx \frac{\partial f_i}{\partial T_{amb}} \frac{\partial f_j}{\partial T_{amb}} \sigma_T^2 + \frac{\partial f_i}{\partial T_{ra}} \frac{\partial f_j}{\partial T_{ra}} \sigma_T^2 + \frac{\partial f_i}{\partial \phi_{ra}} \frac{\partial f_j}{\partial \phi_{ra}} \sigma_\phi^2 \quad (11)$$

and where

$$\sigma_T^2 = E\left[w_T^2\right]$$

$$\sigma_\phi^2 = E\left[w_\phi^2\right]$$

The Taylor series approximation was reasonable since only small measurement noise was propagated through the model. The partial derivatives in Equations 8, 10, and 11 were evaluated numerically using model predictions from the vapor compression system model described by Rossi (1995).

FDD SENSITIVITIES

In this section, simulated results are presented for FDD performance using all of the seven output temperature measurements depicted in Figure 4. Before a fault is detectable, the classification error between the current observation and the estimate provided by the model must be less than 0.001. Table 2 shows the effect of simulated charge leakage on the output of the fault detection and diagnostic system. When 2% of the charge is removed, the classification error is below 0.001 and a fault is indicated. The class probability indicating a refrigerant leak is already nearly a factor of 10 greater than the next most likely explanation. As the fault becomes worse, the certainty in the diagnosis improves dramatically.

Table 2. FDD performance for refrigerant leakage

Fault Size (%)	Classification Error (ε)	Diagnosis	Class Probabilities [leak comp rest cond evap]
0.1	3.71e-01	Normal	
0.5	1.21e-01	Normal	
1.0	1.72e-02	Normal	
2.0	1.59e-05	RefrigerantLeak	[0.26 0.00 0.03 0.00 0.00]
5.0	3.76e-15	Refrigerant Leak	[0.62 0.00 0.00 0.00 0.00]
10.0	0.00e+00	Refrigerant Leak	[0.91 0.00 0.00 0.00 0.00]

Table 3 shows the effect of simulated compressor suction valve leakage on the output of the fault detection and diagnostic system. Valve leakage was modeled as a decrease in the compressor's volumetric efficiency. A fault is indicated (classification error is below 0.001) when the efficiency is reduced by 5%. At this point, the class probability indicating valve leakage is already a factor of 7 greater than the next most likely explanation. Once again, the evidence becomes stronger as the fault level increases.

Table 3. FDD performance for compressor valve leakage

Fault Size (%)	Classification Error (ε)	Diagnosis	Class Probabilities [leak comp rest cond evap]
0.1	4.20e-01	Normal	
0.5	3.71e-01	Normal	
1.0	2.43e-01	Normal	
2.0	7.62e-02	Normal	
5.0	1.47e-04	Compressor Valve	[0.00 0.15 0.00 0.00 0.02]
10.0	8.13e-10	Compressor Valve	[0.00 0.48 0.00 0.00 0.01]
15.0	1.09e-15	Compressor Valve	[0.00 0.65 0.00 0.00 0.00]
20.0	3.19e-29	Compressor Valve	[0.00 0.77 0.00 0.00 0.00]
30.0	0.00e+00	Compressor Valve	[0.00 0.93 0.00 0.00 0.00]

The effect of a simulated liquid line restriction on the output of the fault detection and diagnostic system is shown in Table 4. The restriction was modeled by inserting a valve in the liquid line and decreasing its cross-sectional area. In this case, a fault is not detected until the diameter of the valve opening is reduced by about 80%. At this point, the class probability indicating liquid line restriction is overwhelming. The FDD system appears to be insensitive to liquid line restrictions because the restriction is modeled as a decrease in the diameter of the pipe just before the expansion device and the restriction is not noticeable when it is much larger than the expansion device opening. For the test unit, the diameter of the pipe is about 5 times larger than the expansion device diameter. In practice, a liquid line restriction would occur at the site of the expansion device or within the filter-dryer and would be detected with a much greater sensitivity than demonstrated in Table 4.

Table 4. FDD performance for liquid line restriction

Fault Size (%)	Classification Error (ε)	Diagnosis	Class Probabilities [leak comp rest cond evap]
5.0	4.44e-01	Normal	
10.0	4.47e-01	Normal	
15.0	4.42e-01	Normal	
20.0	4.40e-01	Normal	
30.0	3.85e-01	Normal	
40.0	2.87e-01	Normal	
60.0	1.96e-03	Normal	
80.0	0.00e+00	Restriction	[0.00 0.00 0.98 0.00 0.00]

Tables 5 and 6 show the effects of simulated condenser and evaporator fouling on the output of the fault detection and diagnostic system, respectively. Condenser and evaporator fouling were modeled as decreases in the air flow rate across the coils. For condenser fouling, a fault is indicated when the flow rate is reduced by 20%, whereas the evaporator flow rate is reduced by 40% before a fault is detected. However, the evidence for evaporator fouling is much stronger at the point when a fault is detected.

The sensitivities of the FDD system to the five faults would probably be adequate for this application. Of the five faults considered, only refrigerant leakage would require repair as soon as it was detected (environmental criteria). In this case, the FDD system detected changes of less than 2% in refrigerant charge. The other four faults would be serviced only if they affected comfort, economics, or safety. Since the cooling capacity and power consumption of an air conditioner are strongly coupled to the thermodynamic states used in the FDD method, the comfort and economic criteria are not likely to be violated if a fault is not detectable. Furthermore, unsafe operation would normally be associated with large deviations from normal operating states.

Table 5. FDD performance for condenser fouling

Fault Size (%)	Classification Error (ε)	Diagnosis	Class Probabilities [leak comp rest cond evap]
1.0	4.22e-01	Normal	
2.0	3.62e-01	Normal	
5.0	2.08e-01	Normal	
10.0	3.31e-02	Normal	
15.0	2.64e-03	Normal	
20.0	4.96e-05	Condenser Fouling	[0.00 0.00 0.00 0.14 0.00]
30.0	3.68e-10	Condenser Fouling	[0.00 0.00 0.00 0.26 0.00]
40.0	2.84e-19	Condenser Fouling	[0.00 0.00 0.00 0.32 0.00]
60.0	0.00e+00	Condenser Fouling	[0.00 0.00 0.00 0.35 0.00]

Table 6. FDD performance for evaporator fouling

Fault Size (%)	Classification Error (ε)	Diagnosis	Class Probabilities [leak comp rest cond evap]
5.0	3.50e-01	Normal	
10.0	1.73e-01	Normal	
15.0	9.53e-02	Normal	
20.0	3.80e-02	Normal	
30.0	2.25e-03	Normal	
40.0	4.20e-06	Evaporator Fouling	[0.00 0.00 0.00 0.00 0.74]
60.0	5.62e-19	Evaporator Fouling	[0.00 0.00 0.00 0.00 0.99]
80.0	0.00e+00	Evaporator Fouling	[0.00 0.00 0.00 0.00 0.98]
90.0	0.00e+00	Evaporator Fouling	[0.00 0.00 0.00 0.00 0.98]

SENSOR REQUIREMENTS FOR REFRIGERANT LEAKAGE

It is not necessary to use all of the seven output temperature measurements depicted in Figure 4 if the goal is to distinguish refrigerant leakage from the other possible faults. Figure 7 shows the simulated sensitivities for detecting refrigerant leaks as a function of the number of sensors for a measurement noise of $\sigma_T = 0.5$ C and $\sigma_\phi = 0.05$. For a given number of sensors, the combination providing the best sensitivity was selected. The numbers above each bar indicate which sensors were selected (see Figure 4 for a key). The results show that at least two measurements, T_{sh} and T_{sc}, are required to distinguish refrigerant leaks from the other four faults. However, adding a measurement of T_{hg} significantly improved sensitivity, while additional sensors did not provide much better performance. With three sensors (T_{sh}, T_{sc}, and T_{hg}), less than 2% reduction in charge was detected. When the measurement noise was reduced to $\sigma_T = 0.2$ C and $\sigma_\phi = 0.02$, then refrigerant loss of less than 0.7% of full charge was detectable.

Figure 8 shows the classification error and fault probability ratio versus percent charge reduction for three sensors. The fault probability ratio is the ratio of the most probable fault to the second most probable fault. The horizontal line in the center of the plot is the threshold for classification error (left axis) and fault probability ratio (right axis). A fault is indicated when the classification error is below the threshold and a valid diagnosis is indicated when the fault probability ratio is above the threshold. For this study, the minimum detectable fault was associated with both values exceeding their respective thresholds. This point occurs at the intersection of the vertical line separating the fault and normal regions and the abscissa in Figure 8, and is equal to the value of the three sensor bar in Figure 8.
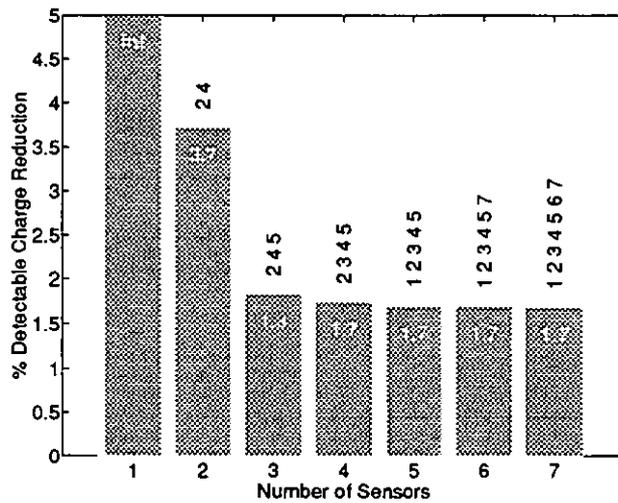
Figure 7. Refrigerant leak detection sensor sensitivity (simulation results; numbers indicate the sensors selected from the list in Figure 4)
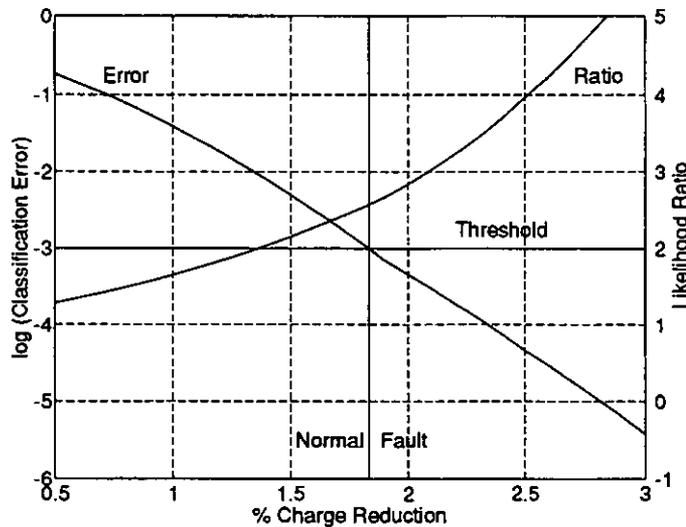

Figure 8. Refrigerant leak detection classification error and diagnostic fault probability ratio (3 sensors)

Figure 9 provides the minimum detectable charge reduction versus number of sensors for the experimental investigation. The results confirm that the simulation tool correctly predicted the relative importance of the output measurements. Three sensors were sufficient for detecting refrigerant loss in addition to the three input measurements required for the preprocessor model. However, the simulated performance was more

sensitive to refrigerant charge than actually occurs. The experiment indicates that a 5.0% reduction in charge can be detected.
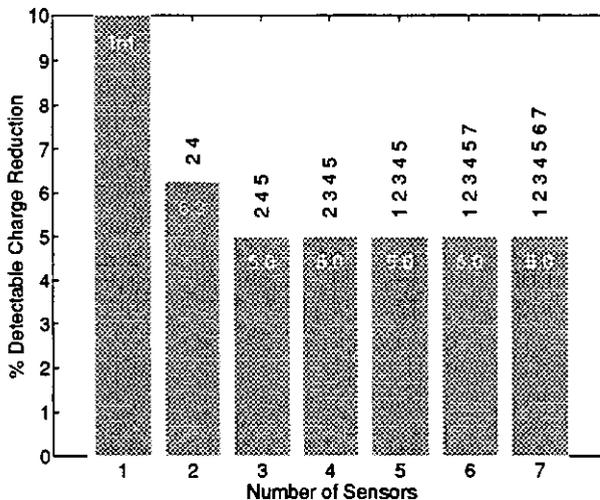

Figure 9. Refrigerant leak detection sensor sensitivity (experimental results; numbers indicate the sensors selected from the list in Figure 4)

CONCLUSIONS

A rule-based, statistical fault detection and diagnostic system was developed and evaluated for vapor compression equipment. The method only requires temperature measurements and one humidity measurement to distinguish between the following five faults: 1) refrigerant leakage, 2) liquid-line restriction, 3) compressor valve leakage, 4) condenser fouling, 5) evaporator fouling. The diagnostic approach is based on generic rules and does not require equipment specific experimentation. Thresholds for both fault detection and diagnosis are both based upon statistical analysis of on-line measurements.

The sensitivities of the FDD method for detecting and diagnosing each of the five faults were estimated through simulation in the presence of typical measurement errors. This method is most sensitive for detecting refrigerant loss. Simulated results showed that a 2% loss in refrigerant could be detected using five temperature measurements (superheat, subcooling, hot gas line, condenser air inlet, and evaporator air inlet) and one humidity measurement (evaporator air inlet) Experimental results confirmed the use of these measurements, but only a 5% reduction in refrigerant was detectable.

Refrigerant loss detection could be an immediate application for the proposed FDD method, since fault evaluation is not necessary for this fault. A refrigerant leak should be repaired as soon as it is detected and diagnosed. Since the other four faults should only be

serviced if they affect comfort, economics, or safety, less detection sensitivity is required for them.

In this study, the thresholds for fault detection (classification error) and diagnoses (fault probability ratio) were chosen heuristically to give a low false alarm rate. The sensitivity of the FDD method in detecting and diagnosing each fault was then evaluated for the fixed thresholds. Ideally, the selection of thresholds should consider the tradeoff between the sensitivity of the method and the false alarm rate. One would like to choose thresholds that allow the method to detect small faults (high sensitivity), but that rarely lead to an indication of a fault that doesn't exist (low false alarm rate). In theory, the principle of least risk could be used to help decide what are appropriate sensitivities and false alarm rates. It would be possible to evaluate the sensitivity and false alarm rate of the FDD method as a function of threshold values. This information could be used in conjunction with costs associated with misdiagnoses (choosing faulty behavior that's normal or normal behavior that's faulty) to determine appropriate thresholds. However, these costs would be very difficult to obtain for this particular application.

This study also did not consider the effect of modeling errors on the sensitivity of the FDD method. Modeling errors do reduce sensitivity and should be considered in future studies.

NOMENCLATURE

$E[\cdot]$	expected value operator
f_i	plant model for i^{th} output
M_C	mean vector that describes the current distribution of residuals determined using recent measurements
M_N	mean vector that describes the distribution of residuals determined using measurements for normal operation
$p(\cdot)$	probability density function
s	parameter that minimizes classification error for optimal linear classifier
T_{amb}	ambient temperature (inlet to condenser)
T_{cond}	condensing temperature
T_{evap}	evaporating temperature
T_{hg}	hot gas line temperature (compressor outlet)
T_{ra}	return air temperature (inlet to evaporator)
T_{sc}	liquid line subcooling
T_{sh}	suction line superheat
U	vector of inputs that affect plant performance

Y	vector of measured plant outputs
ΔT_{ca}	air temperature rise across condenser
ΔT_{ea}	air temperature drop across evaporator.
ΔY	vector of residuals between measured and modeled plant outputs
ε	classification error (probability of making an erroneous classification for normal or faulty behavior)
ϕ_{ra}	return air relative humidity (inlet to evaporator)
σ	standard deviation of measurements
Σ	weighted average of Σ_N and Σ_C: $\Sigma = s \Sigma_N + (1-s) \Sigma_C$
Σ_C	covariance matrix that describes the current distribution of residuals determined using recent measurements
Σ_N	covariance matrix that describes the distribution of residuals in the absence of any faults (i.e., normal operation)

Subscripts

exp	model prediction (expected performance predictions)
meas	measured
spec	manufacturers' specification
T	temperature
ϕ	relative humidity

REFERENCES

Fukunaga, K. 1990. *Introduction to Statistical Pattern Recognition*, Purdue University, W. Lafayette, IN: Academic Press.

Gelb, A. 1989. *Applied Optimal Estimation*, MIT, Cambridge, MA: M.I.T. Press.

Glass, A. S., P. Gruber, M. Roos, and J. Todtli. 1994. Preliminary Evaluation of a Qualitative Model-Based Fault Detector for a Central Air Handing Unit, *Proceedings of the Third IEEE Conference on Control Applications*, Glasgow, August 24-26, 1994, pp 1873-1882: IEEE Catalogue No. 94CH3420-7.

Grimmelius, H. T., J. Klein Woud, and G. Been. 1995. On-Line Failure Diagnosis for Compression Refrigeration Plants, *International Journal of Refrigeration*, 18(1): pp. 31 - 41.

Hiroshi, I. H., K. Matsuo, Fujiwara, K. Yamada and K. Nishizawa. 1992. Development of Refrigerant Monitoring Systems for Automotive Air-Conditioning Systems, *Society of Automotive Engineers*, SAE Paper No. 920212.

Isermann, R. 1984. Process Fault Detection Based on Modeling and Estimation - A Survey, *Automatica,* 20(4): 387-404.

Kumamaru, T., T. Utsunomiya, Y. Yamada, Y. Iwasaki, I. Shoda, and M. Obayashi. 1991. A Fault Diagnosis System for District Heating and Cooling Facilities, *Proceedings of the International Conference on Industrial Electronics, Control, and Instrumentation,* Kobe, Japan (IECON '91), pp. 131-136.

McKellar, M. G. 1987. Failure Diagnosis for a Household Refrigerator, M.S. Thesis, West Lafayette: Purdue University.

Rao, S. 1984. Optimization, Second Edition, New Delhi, India: Wiley Eastern Limited.

Rossi, T.M. and J.E. Braun. 1996. Minimizing Operating Costs of Vapor Compression Equipment with Optimal Service Scheduling, *International Journal of Heating, Ventilating, Air-Conditioning and Refrigerating Research,* 2(1): 23 - 47.

Stallard, L. A. 1989. Model Based Expert System for Failure Detection and Identification of Household Refrigerators, M.S. Thesis, West Lafayette: Purdue University.

Wagner, J. and R. Shoureshi. 1992. Failure Detection Diagnostics for Thermofluid Systems, *Journal of Dynamic Systems, Measurement, and Control,* 114(4): 699-706.

Yoshimura, M. and I. Noboru. 1989. Effective Diagnosis Methods for Air-Conditioning Equipment in Telecommunications Buildings, In *Proceedings of INTELEC 89: The Eleventh International Telecommunications Energy Conference,* October 15-18, 1989, Centro dei Congressi, Firenze, Vol.21.1: 1-7.