International Energy Agency

Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnostics

Technical Synthesis Report
IEA ECBCS Annex 25

Energy Conservation in Buildings and Community Systems

This report is part of the work of the IEA Energy Conservation in Buildings & Community Systems Programme (ECBCS).

Document TSR-05-1998
ISBN 1 902177 11 2
Participating countries of ECBCS:
Australia, Belgium, Canada, CEC, Denmark, Finland, France, Germany, Greece, Italy, Israel, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, Sweden, Switzerland, Turkey, United Kingdom, United States of America
Front cover photograph:
Plant Room, Eros School, Stuttgart, Germany.

Additional copies of this report may be obtained from:
ESSU
Unit 3A Sovereign Court
Sir William Lyons Road
Coventry CV4 7EZ
Great Britain
Tel: +44 (0)24 7669 2050
Fax: +44 (0) 24 7641 6306
e-mail: exco.support@ecbcs.org

Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnostics

Annex 25 Technical Synthesis Report based on Building and Optimisation Source Book. Contributing Authors: J. Hyvärinen, J.C. Visier, M. Stylianou, H. Yoshida, K. Sagara, N. Nakahara, K. Kamimura, T. Miyazaki, N. Nakamura, T.M. Rossi, J.E. Braun, W. Ray, T.I. Salsbury, P. Haves, P. Sprecher, L.G. Olsson, P. Isakson, M. Madjidi, Y. Jiang, H. Tsutsui, H. Vaezi-Nejad, K. Baer, M. Hinkelmann, M. Schöckle, A.S. Glass, A. L. Dexter, M. Benouarets, P. Gruber, K. Matsunawa, H. Iizuka, H. Itoh, F. Miyasaka, T.Takase, J. Tödtli, J .Seem, G. Kelly

Summary made by:
Martin W. Liddament
Head of Air Infiltration and Ventilation Centre/ExCo Support Services Unit
August 1999

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 the 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).

Energy Conservation in Buildings and Community Systems (ECBCS)

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.

The Executive Committee

Overall control of the programme is maintained by an Executive Committee, which not only monitors existing projects but also identifies new areas where collaborative effort may be beneficial. To date the following have been initiated by the Executive Committee (completed projects are identified by *):

1 Load Energy Determination of Buildings *
2 Ekistics and Advanced Community Energy Systems *
3 Energy Conservation in Residential Buildings *
4 Glasgow Commercial Building Monitoring *
5 Air Infiltration and Ventilation Centre
6 Energy Systems and Design of Communities *
7 Local Government Energy Planning *
8 Inhabitant Behaviour with Regard to Ventilation *
9 Minimum Ventilation Rates *
10 Building HVAC Systems Simulation *
11 Energy Auditing *
12 Windows and Fenestration *
13 Energy Management in Hospitals *
14 Condensation *
15 Energy Efficiency in Schools *
16 BEMS - 1: Energy Management Procedures *
17 BEMS - 2: Evaluation and Emulation Techniques *
18 Demand Controlled Ventilating Systems *
19 Low Slope Roof Systems *
20 Air Flow Patterns within Buildings *
21 Thermal Modelling *
22 Energy Efficient Communities *
23 Multi-zone Air Flow Modelling (COMIS) *
24 Heat Air and Moisture Transfer in Envelopes *

25 Real Time HEVAC Simulation *
26 Energy Efficient Ventilation of Large Enclosures *
27 Evaluation and Demonstration of Domestic Ventilation Systems
28 Low Energy Cooling Systems*
29 Daylight in Buildings
30 Bringing Simulation to Application
31 Energy Related Environmental Impact of Buildings
32 Integral Building Envelope Performance Assessment
33 Advanced Local Energy Planning
34 Computer-aided Evaluation of HVAC System Performance
35 Design of Energy Efficient Hybrid Ventilation (HYBVENT)
36 Retrofitting in Educational Buildings
37 Low Exergy Systems for the Heating and Cooling of Buildings

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

Summary

Operational faults in HVAC systems continue to impact on building energy performance and often result in adverse comfort conditions. The main objective of this study was to develop real time procedures to monitor and optimise HVAC performance. Activities covered:

- Evaluating modelling performance for the real time simulation of HVAC systems;
- The compilation of a database of important troubles and faults in existing systems;
- Demonstrating the implementation of Building Optimisation, Fault Detection and Diagnosis (BOFD).

Scope

This report contains a summary of ECBCS Annex 25 Real Time Simulation of HVAC Systems for Building Optimisation, Fault Detection and Diagnosis (BOFD). It is primarily aimed at building services practitioners, designers and policy makers who require background knowledge of BOFD approaches. It is designed to be accessible to the non-expert and to give an introduction to the benefits of BOFD. Other key Annex Reports are:

(i) Building Optimisation and Fault Diagnosis Source Book. Edited by Juhani Hyvärinen and Satu Kärki of the Technical Research Centre of Finland.
(ii) Technical Papers of Annex 25.

Annex 25 Participating Countries: Belgium, Canada, Finland, France, Germany, Japan, Netherlands, Sweden, Switzerland, United Kingdom, and United States of America.
Annex 25 Operating Agent: Professor Reijo Kohonen, ABB Installaatiot OY, Finland.

Contents

1.0 Introduction and Goals of Annex 25

Building optimisation and fault diagnosis (BOFD) is the process of minimising building operation costs (i.e. energy demand, maintenance costs etc.) while maintaining good indoor air quality and thermal comfort. It is aimed at achieving optimum control and early fault detection (and correction) of a building's services, heating, ventilation and air conditioning system. Such an approach can lead to a substantial reduction in unnecessary energy use. Already building services and HVAC systems are increasingly being automatically controlled through computers and software. Current systems, however, often do not adequately cope with the development of faults or service needs. The additional cost of incorporating BOFD need not be significant since many current building energy management systems already contain sufficient processing resources for fault detection as well as a large array of sensors.

Fortunately, automated fault detection techniques are well established in the areas of industrial automation and processes, as well as in other fields of application. Thus the intention has been to adapt existing methods to the problem of buildings and their systems rather than to develop new techniques. The purpose of this Annex has been to develop and evaluate BOFD concepts. The primary objectives have been to:

- Develop methodologies and procedures for optimising real-time performance, automating fault detection and fault diagnosis in HVAC processes and to develop BOFD prototypes that can be implemented in a building energy management (BEM) system. These BOFD methods and system prototypes were called applications;
- Evaluate the most suitable modelling approaches for real time simulation of HVAC systems;
- Determine the most basic approaches most suitable for fault analysis;
- Create a database of the most important problems and faults in HVAC systems;
- Demonstrate the implementation of the schemes in a real BEM system;
- Facilitate and promote the technology transfer to industry;
- Apply optimal control techniques to the problem of building optimisation.

Fault detection and diagnosis forms an essential part of building optimisation and it is on this aspect that Annex 25 has focused. Optimal control has previously been covered by other forums, for example Annexes 16 and 17 on Building Energy Management Systems (BEMS).

2.0 Fault Definition

Conventionally, a process fault is usually considered to cause a failure that results in the operation of the process coming to a halt. However, in the context of fault diagnosis, this definition is inadequate. Instead the concept of fault or defect must be applied to the deviation in operating performance of a process from its design or target performance. Initially, therefore, such a defect may not necessarily affect overall system performance because compensation in small deviations can be corrected automatically by controllers. By applying this concept, it then becomes possible to detect defects before they affect performance and thus enable fault prediction. In fault diagnosis, therefore, emphasis must be placed on monitoring the development of a defect and minimising the damage or loss caused by it.

A defect is hence the departure from the normal operating point of a process. Faulty operation may propagate to other parts of the process or sub-process. An example of such a defect is a slowly increasing blockage in a heat exchanger. The defect is small initially but becomes steadily more serious as the system becomes more blocked. Eventually the fault begins to propagate causing changes in inlet and outlet temperature and upsetting the performance of the rest of the system. If the development of a defect is rapid, then the potential for advance 'prediction' becomes difficult. With slowly developing defects, prediction is easier.

3.0 A Summary of the Components of BOFD

3.1 Introduction

The main purpose of a real time fault diagnosis is to monitor the operation of the (HVAC) system, its subsystems and components and to detect, locate and, if possible, even predict the presence of any defects causing faulty operation. Ideally the system should resolve the primary defect and give instructions for undertaking corrective action. In practice, this (latter) aspect is unlikely to happen at present, and much rests with the operator to use the information provided to locate the cause himself.

In essence fault detection and identification require the following:

- A 'real time' means of identifying that a fault has occurred. This is usually based on continuously monitoring key aspects of the system and comparing its performance with a 'reference' or 'expected' performance;
- Identify the reason and location of the fault;
- Assess the adverse nature of the fault (e.g. can it be tolerated to the next service interval? Alternatively, is immediate action required such as a change in control strategy or must the system be shut down?).

3.2 The Benefits of BOFD

Building Optimisation and Fault Detection methods offer the following advantages:

- Savings in energy and water consumption through identification of inefficient operation of processes;
- Improved indoor environment;
- Reduced maintenance costs, combined with an ability to schedule maintenance efficiently;
- Maintenance of health and safety.

3.3 Main Annex Steps

In performing the Annex, the following main steps were undertaken:

(i) BOFD System Concept

The first part of the project was to develop a general building optimisation and fault diagnosis system description. This gave a framework for BOFD methods developed by the Annex. Aspects covered:

- Definition of BOFD;
- Approaches to 'reasoning';
- Components of a BOFD system;
- System structure;
- BOFD system implementation.

Conceptually, a building must be considered in terms of:

- Systems;
- Sub-systems;
- Individual components.

Assuming that the building, HVAC system and associated controls are optimally balanced as part of the design and commissioning process, BOFD is concerned with detecting (and correcting) any degradation in building performance as a whole (i.e. in air quality, comfort, energy demand, etc.). For example, once it is detected that building energy consumption exceeds its anticipated level, then the cause must be established by identifying the sub-system or component that is at fault and correcting the problem.

(ii) HVAC System Components and Off-Line Database

To assist in the development of this framework, functional descriptions of typical HVAC systems were prepared and analysed. These covered:

- Heating systems;
- Air handling systems;
- Chillers;
- Heat pumps;
- Thermal storage systems.

(iii) Description and Selection of Methods appropriate to BOFD

Methods for fault detection and building optimisation in selected HVAC systems and processes were studied and the applicability of each assessed. Aspects covered:

- Requirements (e.g. measurement signals, accuracy, filtering etc.) of different BOFD methods;
- Applicability of method according to individual processes;
- The application of methods (e.g. fault diagnosis, optimisation, parameter estimation etc.);
- The measurement data needed by each method;
- Computational load.

(iv) Development of Methods

Development methods covered the development of the methods for BOFD and for the diagnosis of specific components and sub-processes of the HVAC process. It also covered the development procedures for the determination of the desired performance. Approaches included:

- Fault detection;
- Fault detection combined with identification of location and cause of fault.

(v) Evaluation of Methods to Detect and Locate the Fault in a Process and to Identify the Potential Fault Sources

The problem of diagnosing the location and cause of a fault was studied. Diagnosis is an essential part of the BOFD system because it binds together the building optimisation and fault detection approaches. It is this aspect of BOFD that formed the basis of this Annex.

4.0 Building Optimisation and Fault Diagnosis Methods

4.1 Introduction

Building optimisation and fault detection relies on comparing actual (monitored) performance against a 'reference' level. In the case of energy consumption, this reference point can be derived from design calculations, relevant standards and/or statistical data based on the building type and installed HVAC system. The tasks and objectives of building optimisation are summarised in Figure 1. In this figure the main steps are illustrated, namely:

- Detecting of the fault;
- Diagnosis based on the information provided by the detection mechanism combined with knowledge about the system;
- Determining what actions to take and applying them.

Building optimisation is performed by:

- Adjustment control parameters for optimum operation through a learning process;
- Applying optimal control and fault detection algorithms;

- Design of HVAC sub-systems through off-line simulations (such simulations are also useful during building operation since they have sufficient precision to predict the air conditioning load, energy consumption and the quality of the indoor environment over a range of external environmental parameters (air temperature, humidity, wind velocity etc.));
- Retrofitting, as necessary, some parts of the system.

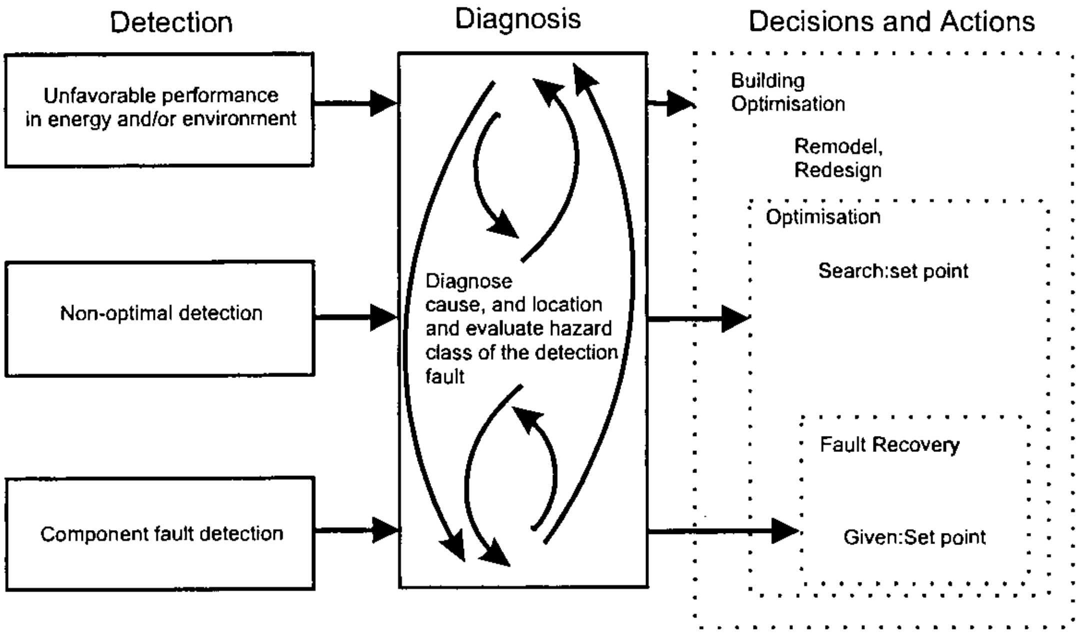

Figure 1 Tasks and Objectives of Building Optimisation and Fault Detection

4.2 Components of a BOFD System

Figure 2. illustrates a block diagram for the process of dealing with the development of a system fault.

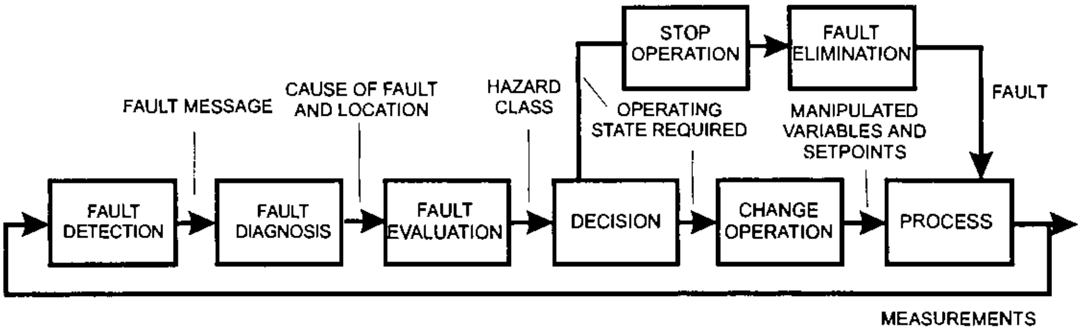

Figure 2 Stages in Automatic Fault Detection and Processing

If a process fault occurs, it needs to be detected as quickly as possible. This is achieved by monitoring the performance of key components and ensuring that they are operating within their acceptable band of tolerance. A deviation leads to a fault message. This monitoring is followed by fault diagnosis in which the fault is located and the cause established. The fault is then evaluated, i.e. its significance is determined.

The processing steps are illustrated in Figure 3.

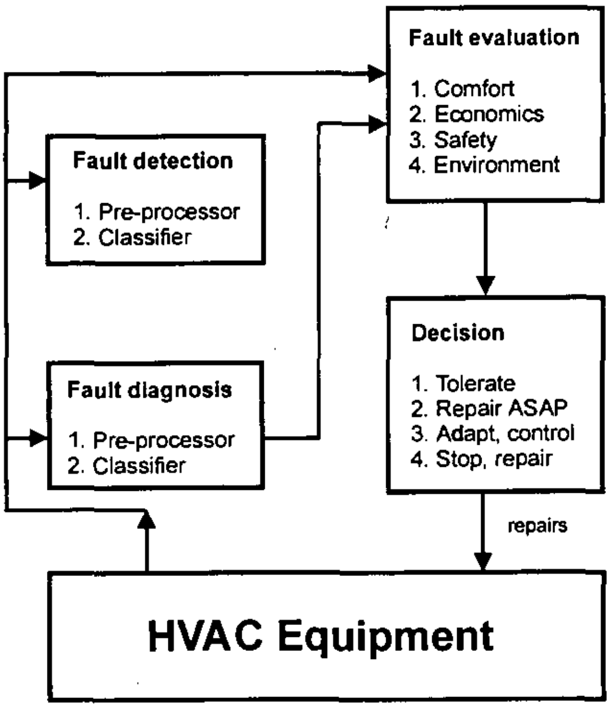

Figure 3 Supervision of HVAC and R Equipment

Pre-processing

Fault detection and diagnosis first undergoes some pre-processing that compares intended performance against what is happening. An indication of a fault is given by the deviation of the observed behaviour from the expected. In critical areas 'redundancy' may be used, i.e. multiple sensors or actuators may be used for the same purpose. A voting procedure is used to compare performance and faults are detected by majority rules. However, this technique is expensive, bulky and limited in ability, it is therefore unlikely to be used in HVAC applications. Instead, various modelling techniques (see Section 5) are applied for this purpose in which simulated results are compared with measurement output from sensors (e.g. temperature, flow rates etc.). As an example, a mathematical model can provide the expected value of measured thermodynamic states as a function of measured external driving conditions. The difference between expected and actual measurement values (residuals) will be zero when there is no fault. This model may be 'physical', i.e. actually representing the process, or 'empirical'.

Classifiers

Classifiers operate on the information provided by the pre-processor to decide if a fault actually exists in the system (detection) and, if so, then which component(s) are the cause (diagnostics). Much of the work of BOFD is to develop reliable classifiers that can accurately detect or validate a fault with the minimum of false alarms. Techniques are based on the use of expert systems and fuzzy logic.

In a broad sense, the classifier is an 'expert' system; i.e. it emulates the reasoning process of a human expert. It consists of a knowledge base and an 'inference' engine. The knowledge base contains expert knowledge about the system while the inference engine combines data about a particular problem with the expert knowledge to provide a solution. Complexity can vary from a few simple rules to a complex reasoning system (See Section 6.4).

4.3 Fault Evaluation

Fault evaluation assesses the impact of a fault on overall system performance. There are essentially two types of faults, namely hard failures and performance degradation. Hard failures are severe and abrupt faults for which 'evaluation' is not usually necessary. For example a broken fan belt would need to be replaced. On the other hand it is not always obvious when performance degradation (such as a fouling heat exchanger) has become severe enough to justify a service expense. This is when fault evaluation is necessary to make this decision, based on the impact of the fault on overall system performance.

Economic considerations alone can be used to determine if the cost of servicing is justified. Important costs include service, energy, downtime, safety hazards and environmental hazards (e.g. refrigerant leak). The latter three costs are difficult to quantify, although it is possible to eliminate the needs for these costs by assuming that these problems are small in comparison to the financial impact of the fault. Making this assumption leads to the following four fault evaluation criteria:

- Economic - service to minimise the costs of energy and service;
- Comfort - service the equipment when it is not capable of maintaining the control set-point;
- Safety - service equipment when its operating state is leading to premature component wear;
- Environmental Hazard - service the equipment when the environment is being harmed (e.g. refrigerant leak).

Separate modules can be designed to evaluate each of these criteria independently. Anyone of these can justify service alone. Another way to visualise these criteria is as a constrained (comfort, safety and environment) minimisation of energy and service costs.

4.4 Reaction Decision

Having detected, diagnosed and evaluated a fault, a decision must be made on how to react to it. Four possible alternatives are (in order of increasing severity):

- Tolerate (i.e. its impact on overall performance is not severe enough as determined by the evaluation criteria, e.g. mild heat exchanger fouling detected by a sensitive FDD);
- Tolerate but repair as soon as possible (hazard criteria reached [is this tolerable?]);
- Adapt the control (unsafe condition reached, adapt controls to prevent it continuing);
- Stop operation until repaired (unsafe condition reached, adapting controls not possible).

4.5 Performance Criteria for Evaluating FDD Systems

There are many different ways to design an FDD system. Furthermore, there is unlikely to be a 'best' solution to every problem because engineering designs always involve trade-offs between competing priorities. Classically, FDD systems have been judged by sensitivity, false alarm rate and detection speed. Decision thresholds are determined by balancing these three criteria. Many design improvements are intended to provide better sensitivity and/or detection speed without increasing the false alarm rate. In HVAC applications, detection speed is generally not an important criterion. However, false alarm rate is critical because users will quickly abandon the system that calls for service needlessly. Given an acceptable false alarm rate, the best possible sensitivity should be obtained for the given budget, available tools and development schedule for the project.

5.0 Approaches to Fault Diagnosis and Detection

5.1 Introduction

Fault detection relies on comparing the actual performance of a system with a 'reference' value and then assessing the cause of the fault by understanding the behavioural pattern or departure of actual operating condition from the reference condition. Alternatively, faults can be detected by using reference conditions that represent the output or influence of known faults. Much therefore rests with establishing the reference conditions against which performance can be compared. This is essentially accomplished by using a 'model' which represents ideal system performance (and/or faults conditions). Various reference models are available with each having their own range of applications, advantages and disadvantages. Unfortunately, no single model is universally applicable and thus it is important to understand the range of models that exist and their relevant applications and limitations. Common models are:

5.2 Physical 'White Box' Models

Physical models may also be referred to as 'white box', 'analytical', or 'first principles' models. These are quantitative mathematical models based on theoretical equations describing the physical processes occurring in the system. Physical models make as much use as possible of known performance information (a priori knowledge) about the system being modelled. A typical example would include a 'thermal' model which is able to predict the thermal environment in a space, for given envelope characteristics (insulation, thermal mass etc.), weather conditions (radiation, outdoor temperature, wind velocity) and HVAC operating conditions. The primary disadvantage of physical models is often their inherent complexity. This may result in slow operational speed and an inability to keep up with changing conditions in 'real' time.

5.3 Empirical 'Black Box' Models

Often, the physical model is replaced by an empirical 'black box' approach in which the physics of operation is not represented. Instead, non-physically based simplified equations are used which are demonstrated to represent the characteristics of the system. In this case the various parameters or coefficients have no physical meaning and, therefore, any changes in their value to adapt to fault conditions have no direct relationship to a particular quantity (i.e. the model parameters do not represent physical quantities such as air temperature, heat transfer coefficient etc.).

Black Box type models studied by the Annex included:

'ARX and 'ARMAX' models: These stand for Auto-Regressive eXogenous and Auto-Regressive Moving Average eXogenous models respectively. No knowledge of the inside of the structure is considered. Fault detection is achieved by carrying out statistical tests not only on the residuals between monitored and modelled behaviour (i.e. the level of deviation between observation and expectation) but also on the model parameters themselves (Figure 4).

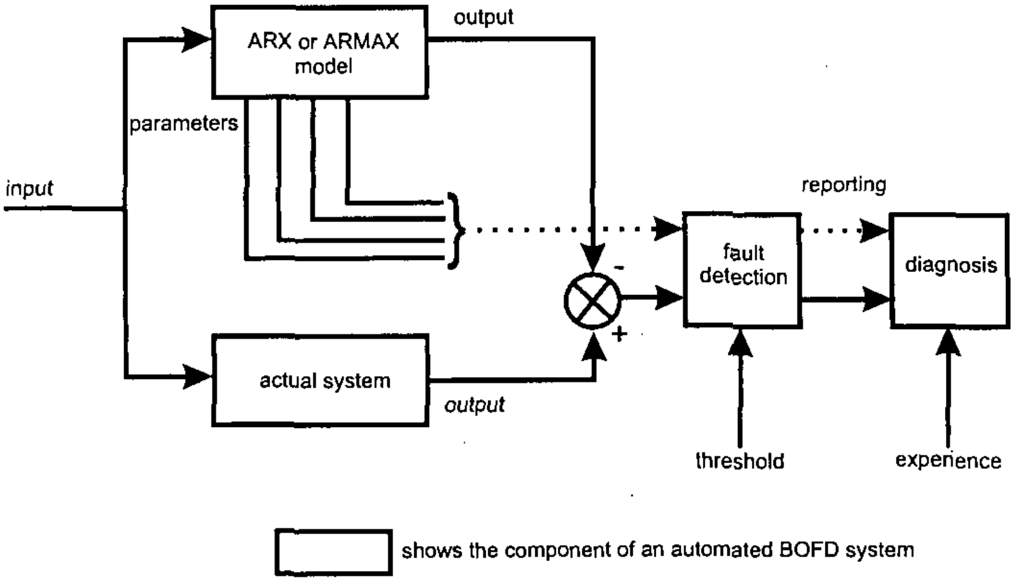

Figure 4 Fault Detection and Diagnosis with ARX and ARMAX Models

5.4 State Estimation (using Kalman Filters and Observers)

This is based on a plant model being driven by the same input as the plant (Figure 5). The model is used to estimate the plant states and, based upon these, the plant outputs. Any differences between the true outputs and the modelled outputs are fed back into the model to achieve convergence with respect to initial conditions. In this way the residuals serve as fault indicating signals and the changes to the parameter coefficients that are needed to restore modelled results to the observed results provides an evaluation of the fault.

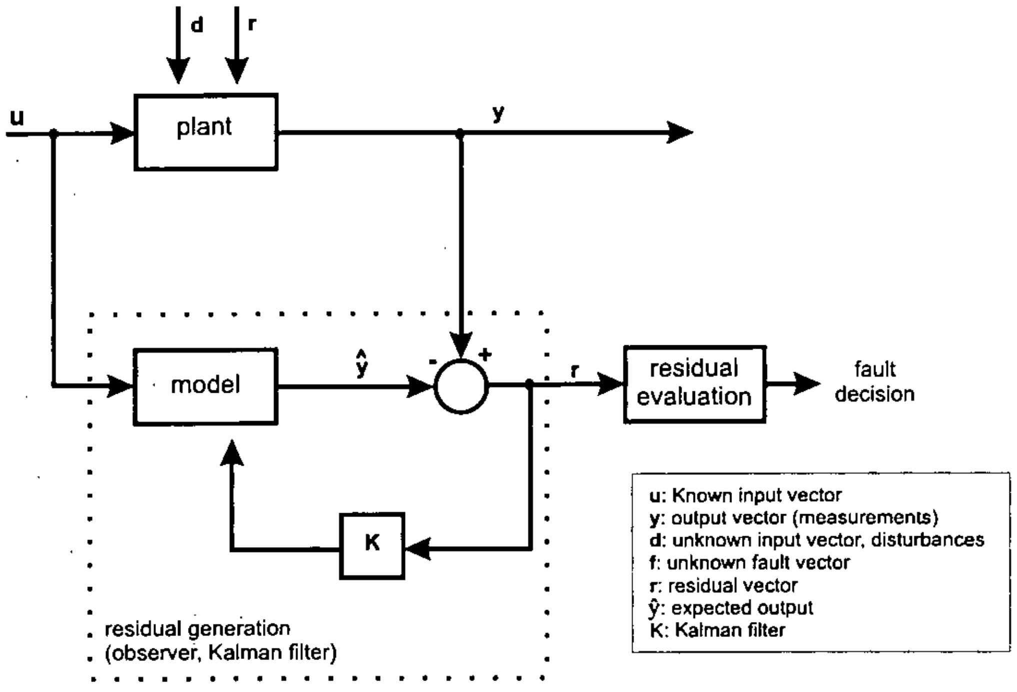

Figure 5 State estimation Technique Using Kalman Filters

5.5 'Training' a Model ('Learning' and 'Teaching')

In reality, many of the 'input' conditions to both white and black box models are themselves of an 'empirical' or uncertain nature (e.g. heat transfer coefficients, building air tightness, equipment characteristics, etc.). Such uncertainties could lead to substantial error in the 'reference' values, making FDD very unreliable. Before they can be used for fault detection and diagnosis, therefore, pre-processor models and classifiers must be 'taught' what to expect under normal and faulty conditions. This can be done in a controlled environment before commissioning the FDD system and/or on-line while the FDD system encounters natural faults. Controlled learning can be accomplished by experiment (e.g. by measuring the response of an actual piece of equipment or observing the response of a detailed model to imposed operating conditions and simulated faults). Such an approach can create a good reference for performance which does not necessarily assume that the unit being tested is operating properly at the beginning of the learning process. This technique can also be used to catch design errors. On-line learning occurs after the FDD system has been commissioned. When the system is expected to be performing well (e.g. directly after commissioning or servicing), the current measurements can be used to 'learn' normal performance. Also, immediately before service, when the system has known faults, fault models can be taught to replicate these faults.

Ideally the 'training' data should cover the whole operating range of the system. However, for applications such as HVAC systems, which are primarily driven by the effects of weather and occupancy, it may be difficult or expensive to cover the whole operational range within a realistic time frame.

More 'training' data is invariably necessary for 'black box' models than for 'white box' models because the parameters do not represent actual physical quantities. On the other hand, even if the training data are incomplete, the inherent resilience of physical models means that they are likely to be more accurate than other forms of models.

5.6 Applying FDD Models

FDD models can be used in the following two 'innovation' modes (where the term 'innovation' is applied because something 'new' is happening as a fault develops):

'Output Innovations'; actual system output is compared with that predicted by the reference model. (Figure 6).

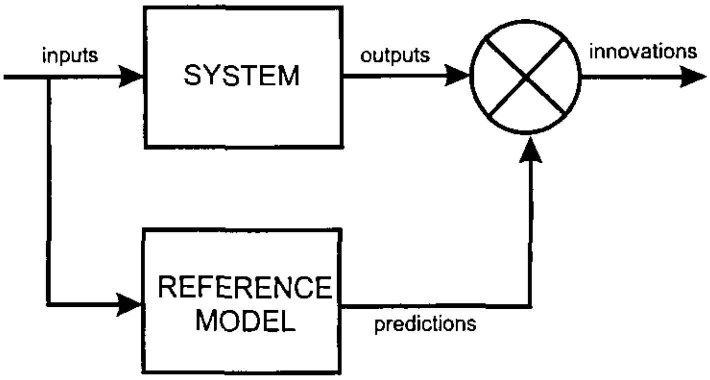

Figure 6 Output Innovations-Based Approach

This approach, therefore, looks for changes in measurable variables and, hence, detects that a fault condition has occurred (or is likely to do so). Analysing the way in which the departure from the reference value takes place can sometimes produce a diagnosis. In principle output innovations can be applied to all types of reference model. This approach can be a robust method for fault detection and it is capable of providing simple diagnostics. However, the potential to provide comprehensive diagnosis from these output deviations is limited.

'Parameter Innovations': The parameters used in the reference model are compared with those needed in an identical model to represent the currently observed behaviour of the real system (Figure 7). Hence, this approach looks for changes in properties that are not directly measurable, but, rather, is aimed at giving some indication of the actual nature of the fault by means of identifying the changes that need to be made to input parameters in order to reproduce the observed results. For example, the need to change the heat transfer resistance of a heat exchanger may be attributed to a build-up of fouling. Care is needed in selecting the number of parameters subjected to change. Too many will make the task impossible, while too few will result in an inadequate representation of possible fault areas. Hence it is important to identify the key potential fault parameters which, usually, will only represent a subset of the total number of parameters used in the model.

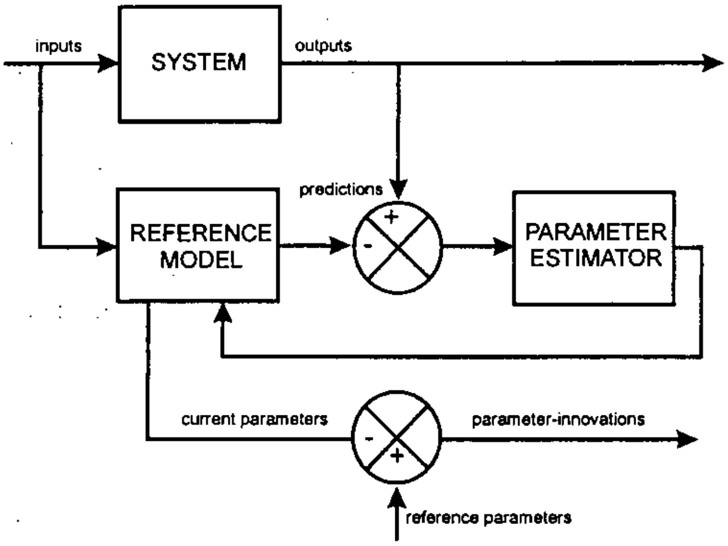

Figure 7 Parameter Innovations-based approach

6.0 Methods of Fault Diagnosis

6.1 Introduction

Once a fault has been detected, its cause must be diagnosed and located. Various techniques, which may be incorporated within the 'reference' model approach, are used to undertake this task. These include:

6.2 'Expert' Systems

An expert system is a means by which knowledge about a system and fault characteristics can be incorporated into the detection and diagnostic process. This is a 'rule' based approach and has proved effective for many applications.

If any fault occurs in a system it covers, the first stage is one of fault detection, either through the reference model approach or manually through a 'rule' based system. A fault warning is given and then, based on a knowledge base of faults and their causes, it automatically attempts to identify the cause of a fault.

The structuring of a fault detection and diagnosis expert system requires that the relationships between faults and the causes of faults and their combinations can be reduced to a 'knowledge' base in advance. This involves the input of experts (the acquisition of knowledge) combined with identifying how to 'structure' the knowledge (expression of knowledge) and how to implement the diagnosis (method of reasoning) (Figure 8).

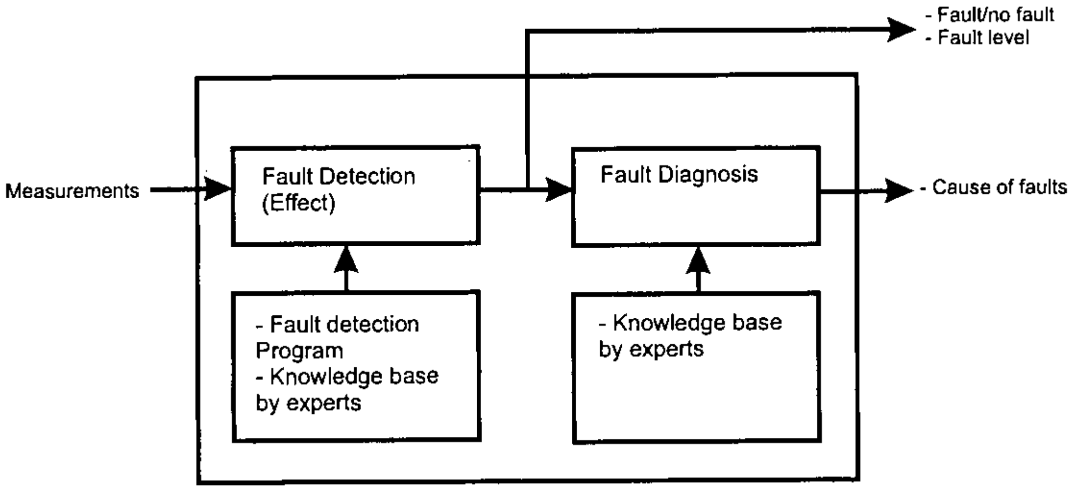

Figure 8 Fault Detection and Diagnosis System

6.3 Expression of Knowledge

Various methods are used to express the knowledge, these include:

IF - THEN Expression: This approach has long been proved. The description is made according to the rules in which 'attributes' (symptoms, test results, and measured values) make up the 'IF' part and conclusions (causes) make up the 'THEN' part (Figure 9). In the example, the first rule expresses the knowledge 'if the room is warmer' (the fault learned from the resident) and the supply air temperature is too high (the fault learned from a measured value), then either:

- The cooling water valve is malfunctioning;
- The size of the cooling water coil of the air-handling unit is insufficient.

There are hence two possible causes. Rules must be added to specify each of the causes. This approach has the advantage of being descriptive but is difficult to follow when too many rules have to be described.

If room is warmer AND supply temperature is too high THEN valve is malfunctioning

If room is warmer AND supply temperature is too high THEN size of coil is insufficient

If room is cooler AND VAV air volume is too much THEN controller is malfunctioning

Figure 9 Example of If-Then Expression

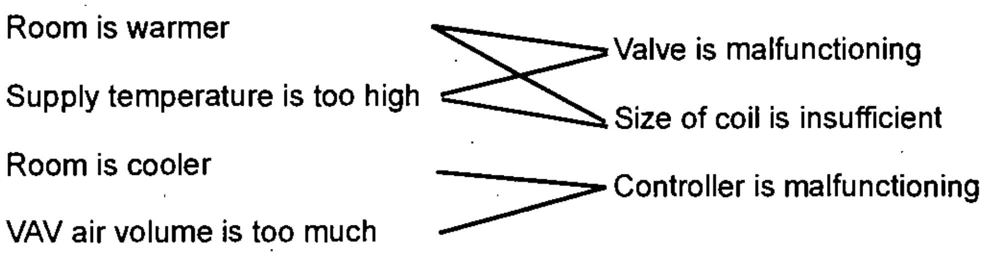

Figure 10 Example of Graphic Expression

Fault Tree: In this format, knowledge is expressed by a tree like structure (Figure 11) in which questions about attributes constitute nodes and choices about the attribute values constitute arcs generated from these nodes. Leaves of the tree represent conclusions. For example, if attributes are 'room' is 'warmer', the arc is traced to the question about the attribute 'air supply temperature'. If the answer to the question is 'too high' then the 'malfunction of the cooling water valve' or the 'insufficient coil size' is the conclusion. Rules must be added to specify each of the causes. This approach is analogous to a trouble shooting chart. Combinations of rules are difficult to represent.

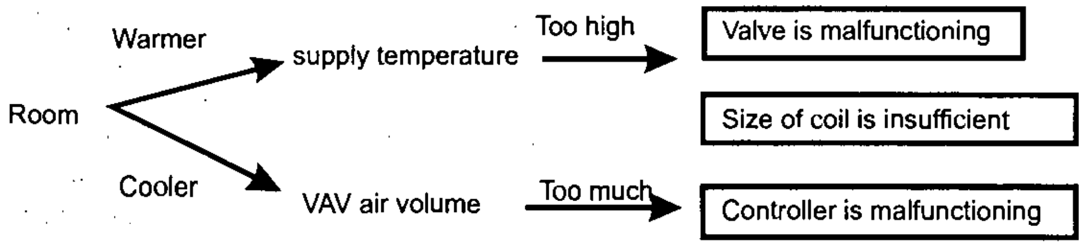

Figure 11 Example of Fault Tree Expression

Decision Table Expression: The relationships between attributes and conclusions are arranged in order in a two dimensional table and these relationships are described by using symbols (Figure 12). If, for example, attributes are the 'room' is 'warmer' and the 'air supply temperature' is 'too high', they correspond to the first and second lines. Thus the conclusion obtained is 'valve is malfunctioning' or 'the size of cooling coil is insufficient'. This format enables easy verification of the rules but it only gives a very poor approximation of the sequence of diagnosis.

	Room		Supply air temperature	VAV air volume
	Warmer	Cooler	Too high	Too much
Valve is malfunctioning	O	X	O	X
Size of coil is insufficient	O	X	O	X
Controller is malfunctioning	X	O	X	O

Figure 12 Example of Decision Table Expression

6.4 Reasoning or 'Inference' Methods

An integral part of optimisation and fault detection is the process of 'reasoning'. Expert systems for the reasoning of non-optimal states are a practical method of diagnosis. Such an approach relies on a good database of system, sub-system and component characteristics. Essentially two aspects must be considered; these are:

The 'forward inference' or 'top down' approach: this is concerned with identifying the cause of a problem which manifests itself at the 'building' level' (i.e. what has caused an unacceptable deviation in air quality, temperature or energy consumption?). In this approach, the observed 'deviation' becomes the 'objective function'. This means that the undesirable effect caused by the fault is known right from the start (e.g. it's too cold). The problem is one of identifying the specific subsystem and component that is causing the problem. The task becomes one of isolating the location of the fault by identifying the reason for its occurrence. This is achieved by proceeding according to the order of probabilities or some other corresponding order of importance. Localising the fault, therefore becomes one of proceeding through a hierarchical tree. As an example, a question is first made about the 'room' followed by a question about the air supply temperature'. The rules that match the result are applied and the conclusion part becomes 'true'. In the fault illustrated in Figure 13, if the room is 'warmer' and the 'air supply temperature is 'too high' then the malfunction is the 'cooling water valve' or the 'size of coil is insufficient' becomes true and is drawn as the conclusion. In this example, the conclusion is not drawn in a single stage. In the case of multi-stage rules, however, rules are applied one by one before a final conclusion is reached.

If room is warmer AND supply temperature is too high THEN valve is malfunctioning

If room is warmer AND supply temperature is too high THEN size of coil is insufficient

If room is cooler AND VAV air volume is too much THEN controller is malfunctioning

Figure 13 Forward Difference Approach to Reasoning

- ***The 'bottom up' approach:*** this is concerned with identifying how a particular component, sub-system or system fault will manifest itself at the 'building level'. Initially, a fault in a component or subsystem will not necessarily be apparent at the building level at the same time as at lower levels. Hence, the bottom up approach can be used to predict or give advance warning of a problem. This is accomplished by understanding the consequences of a fault in a specific process. If it is assumed, for example, that the conclusion is that the 'cooling valve is malfunctioning' the questions about the room and the air supply temperature are made as conditions for these assumptions to hold true. When the hypothesis for which all conditions is found to be true is satisfied, the result becomes a conclusion. Since it is not realistically possible to monitor every component, it is essential when using the bottom up approach to select such components and faults for investigation for which a failure is either probable or would have the greatest impact.

For the purpose of this Annex, the 'top down' approach was called 'building optimisation' since it starts with the need to specify 'optimal' conditions against which actual performance can be compared. Conversely, the 'bottom up' approach was called 'fault diagnosis' since it is concerned with identifying the impact of specific faults.

6.5 Advance 'A Priori' Knowledge

Advance knowledge about the operation of a system to be diagnosed is essential when designing in both the top down and bottom up approaches. In the top down approach, it is needed for designing the 'reasoning' process and, in the bottom up approach, it is needed for selecting the most important components.

6.6 Associative Networks (Semantic Networks)

Rule based approaches typically describe knowledge through conditional rules. Unfortunately, in the case of fault diagnosis, a fault may not immediately result in the occurrence of some of the symptoms. In an associative network, the symptom/fault relationship is stored in the form of direct links between symptom and fault. A small associative network describing symptoms and faults associated with a heat exchanger is illustrated in Figure 14.

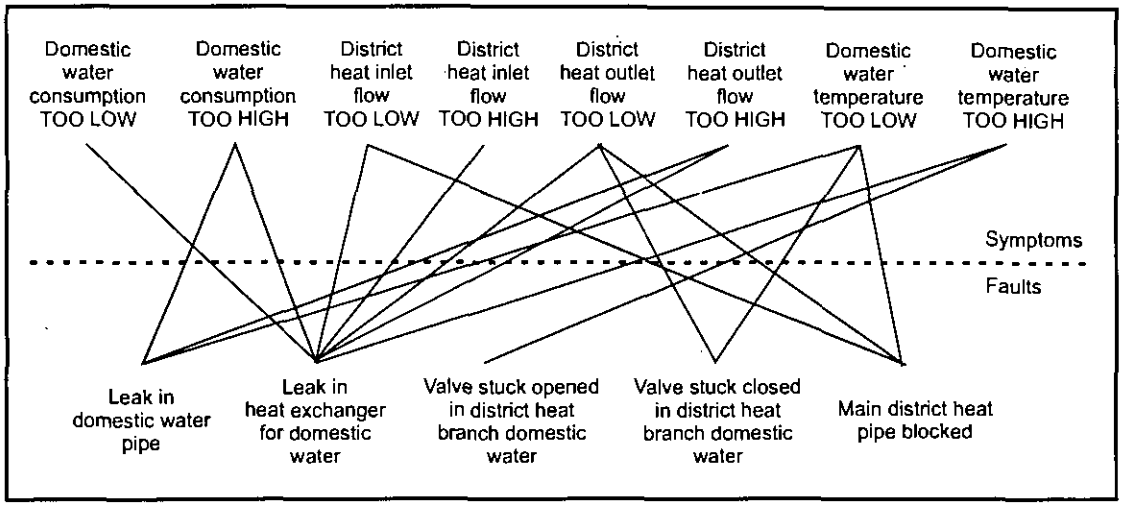

Figure 14 Associative network for the diagnosis of faults in a district heating subsystem

By means of an example, assume that the pre-processing programme has shown the following two symptoms:

- District heat outlet flow too low;
- Domestic water consumption too low.

There are three faults capable of causing the detected symptoms but only one fault that is capable of causing both symptoms (i.e. heat exchanger leak). If a third symptom is detected, e.g.:

- Domestic water temperature is too low;

then no single fault is capable of accounting for all of the symptoms present. Instead, there are three potential causes i.e:

- Blocked district heat pipe combined with heat exchanger leak;
- Leak in domestic hot water combined with heat exchanger leak;
- Valve stuck closed combined with heat exchanger leak.

Advantages: These include:

- A natural format for diagnostic knowledge in terms of fault/symptoms associations;
- Coding is more simplified than a rule based approach;
- Reasoning is speeded up because only links have to be processed;
- It can handle multiple faults and uncertainty.

Disadvantages: These include:

- Difficult to ensure completeness and correctness of knowledge base (although this is a problem of all network systems);

- The associative network has to be set up for each HVAC system separately and each system is unique;
- It can only list faults stored within the knowledge base and network.

7.0 Qualitative Approaches to FDD

7.1 Introduction

A difficulty with model based approaches is that that they pre-suppose that the normal, prescribed behaviour of the HVAC components and the associated control systems can be accurately modelled. However it is extremely difficult, if not impossible, to model the whole system accurately. Even if it can be assumed that the mathematical models themselves are correct, many of the relevant parameters will be unknown or might only be obtainable by means of difficult or time-consuming measurements.

It is known, however, that experienced service engineers can diagnose many faults without knowing all the system parameters. They often identify problems on the basis of a qualitative appraisal. The approach, therefore, is to adapt this concept to FDD. Broadly speaking, a qualitative description differs from a quantitative one in that the amount of information relevant to the description of a system is reduced. A typical quantitative model used in HVAC applications, for example, involves a description in terms of continuous state variables which satisfy systems of well determined differential equations (i.e. the physical state of the system is accurately defined by mathematical means). A qualitative description, on the other hand, either involves non quantitative attributes such as 'liquid', 'vapour', 'mechanism jammed' etc. or involves coarsely quantised quantities such as 'temperature above 100°C' or 'temperature below 0°C', etc. In essence this approach is based on the application of 'qualitative' physics, in other words, unlike black box approaches, the concepts and parameters applied have physical meaning. Qualitative Physics deals with states, transitions and behaviour. Generally this approach should take into account the following features:

- Physical or descriptive attributes assume a finite set of values;
- Qualitative values should encompass the full range of behaviour being modelled;
- It must be possible to interpret the results of qualitative analysis in a meaningful way;
- Qualitative values should be non over-lapping otherwise the results may be ambiguous;
- The coarseness of approach should be adaptable and appropriate for the parameter being represented. For example, intervals used in temperature ranges must be relevant to the phenomenon in hand;
- The various values assumed for qualitative values must be naturally ordered;
- A representation must include, as part of its overall formulation, both a description of the relevant phenomena and a set of operations to obtain solutions.

7.2 General Approaches

A number of general approaches have been investigated, the most significant to date are:

Device Structures: These are analogous to classical HVAC layout diagrams in which a number of components are interconnected in a system. Each component occupies a fixed position and responds to given inputs by yielding certain outputs.

Process Theory: This is closer to conventional dynamic systems in physics, e.g. the change in phase from liquid to gas.

Constraint Propagation: This is a mathematical modelling approach for tracing the propagation of algebraic equations from one part of a complex system to another.

Three possible approaches to qualitative methods for FDD are illustrated in Figure 15. The first is a pragmatic one in which general qualitative rules are derived from expert knowledge (including analytical methods) and incorporated as built-in rules in an FDD system. The second approach makes use of formal qualitative modelling methods to generate rules for incorporation into the previous approach. A long-term goal is to integrate qualitative methods into the FDD system itself. The aim is to provide a system capable of robust diagnosis, using qualitative methods in situations which do not lend themselves to the application of more precise methods.

A fault detection method using the second approach was successfully devised and tested under laboratory conditions as part of the Annex programme. The behaviour of a central air handling plant was studied under steady state conditions. It proved feasible to relate the controller behaviour to temperature conditions in such a way that component faults in the air-handling unit could be detected on the basis of qualitative information.

The behaviour of the system was modelled analytically. The device approach was used for qualitative modelling (i.e. the air handling unit was viewed as a system of components and connections). In addition, it proved particularly useful to identify suitable 'landmarks' although they were not supplied to model the dynamic evolution of the system. A sequential controller operates the heating coil, the dampers, and the cooling coil in sequence so as to attain the desired air temperature. The 'landmarks' used were transitions between heating and damper operation, in one instance, and damper operation and cooling in the other. These 'landmarks' were associated with particular combinations of outside air temperature together with the supply and return air temperatures. They allow discrepancies between the qualitative temperature states and the sequential controller states to be detected. Individual component faults can be analysed to obtain the rules leading to the detection of each particular fault. The results can be tabulated to aid fault diagnosis. In general, however, each 'signature' of qualitative fault symptoms can be associated with more than one type of fault. One would therefore need quantitative fault models to distinguish between them.

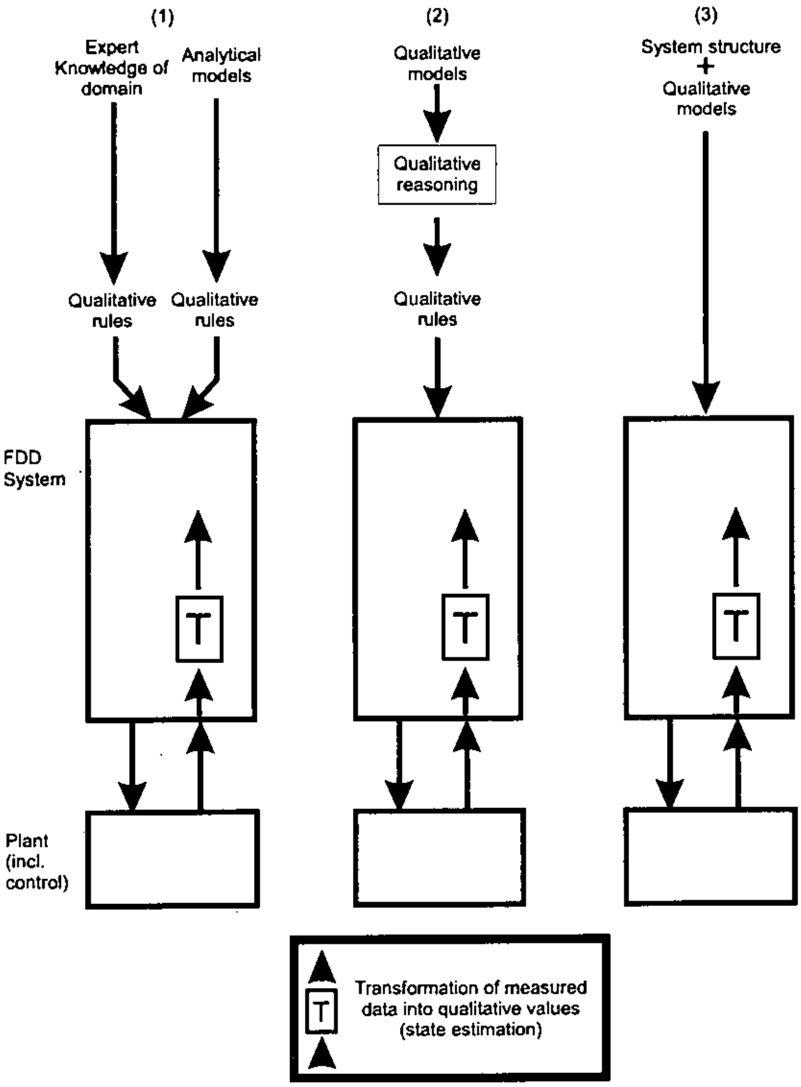

Figure 15 Qualitative Methods for Fault Detection

7.3 'Fuzzy' Logic Model Based Approaches

In practice, inferences about faults are often vague or incomplete. In other words, there is a level of uncertainty about the cause of a problem or the symptoms caused by a particular fault. To accommodate this, a level of uncertainty or degree of confidence is introduced, ranging from +1 to –1 (or +1 – 0). This is used to make as accurate inferences as possible from vague or incomplete attributes, knowledge and conclusions. Level +1 indicates that an event or conclusion is absolutely true whereas as -1 refers to absolutely false. Some methods include the use of 'probability' values or the use of 'fuzzy' inference. It is necessary to choose the optimal method for the problem in question.

The use of fuzzy qualitative models can take more realistic account of the uncertainties associated with describing the behaviour of a HVAC plant. It also, more easily, incorporates expert knowledge about symptoms of faults than other qualitative techniques. Methods can be divided into 'explicit' fuzzy logic which relate to descriptions of the plant (and contain detailed knowledge), and 'implicit' fuzzy logic in which symptoms are related to faults. Models are based on expert systems and/or training data.

An Explicit 'Innovations' Approach to Fuzzy Model Fault Detection: An innovations approach to fuzzy model based fault detection is illustrated in Figure 16.

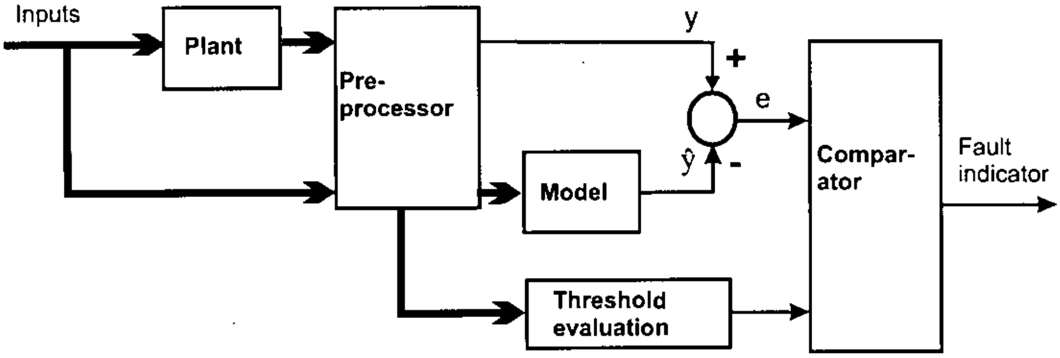

Figure 16 Innovations Approach to Fuzzy Fault Detection

An explicit fuzzy model is used to describe the behaviour of the correctly operating plant. Expert knowledge is used to choose the structure of the model that is best suited to describing the relationship between the measured variables. The main issues that need to be considered are whether a dynamic or steady state relationship is most appropriate and what level of precision is required. An important aspect is the level of confidence that can be placed in the model prediction.

To avoid problems that can arise when the model is not based on complete training data (for example when normal operating data must be used), the reference model is distributed in structure. This means that the model consists of a 'coarse-grained' set of fuzzy rules and a finer grained fuzzy model which is trained, on-line, using measured data from the plant. The fuzzy rules are generic, describing the behaviour of a class of plant of this type, and are based on expert knowledge or identified from data generated by simulating plants of similar design. The fuzzy model is specific, describing the observed behaviour of a particular plant. A measure of the local density of training data can be used to decide which of the two models offers the most credible prediction. In common with all innovations based schemes, faulty operation is detected by comparing the observed behaviour with the predictions of the fuzzy reference model using a threshold which takes into account the estimated modelling errors.

A Parameter Based Approach to Fuzzy Model-Based Fault Diagnosis: Explicit fuzzy models, based on pre-defined reference sets, are used as the reference models describing the symptoms of faulty and fault free behaviour at all possible operating points. Design information is used to normalise the measured data so that all values are in the range 0.0 to 1.0 or -1.0 to +1.0. A particular model is defined by specifying the values of the elements of an associated fuzzy relational array. Each element of the array is a measure of the credibility or confidence that the associated rule correctly describes the behaviour of the system around that operating point. Fuzzy identification is used to generate a partial fuzzy model (describing the behaviour around the current operating point) from operating data collected on-line from the actual plant under test. The degree to which a particular fault, or correct operation, is present is determined by comparing the rules of the partial fuzzy models with the rules of the reference models (see Figure 17). A steady state detector is used to detect and reject any transient data in cases where the reference models describe only the steady state behaviour of the plant.

The fuzzy reference models are based on expert knowledge, or, are generated off-line from training data produced by computer simulation of typical plant, with and without faults. The operation of several different plant designs must be simulated to generate representative training data so that the fuzzy reference models are sufficiently generic to capture the underlying characteristics of the behaviour of the actual plant. The advantages of this approach include:

- Avoidance of the problems associated with on-line training on the actual plant (for example, the assumption that the plant has been fully commissioned and is initially working correctly);
- Can be used when it is impossible to train on data collected from the actual plant (for example when the training data must be representative of faulty behaviour);
- Requires little modification to be used to identify faults in different designs of the same class of plant.

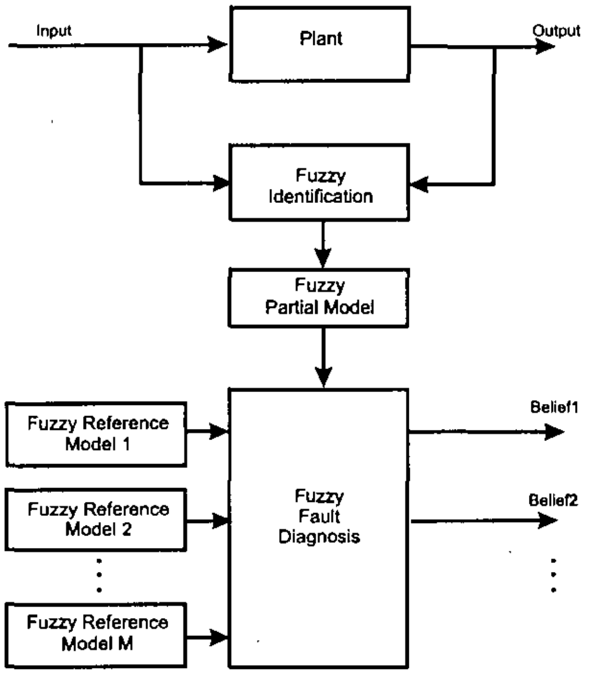

Figure 17 A Parameter Based Approach to Fuzzy Model Based Fault Detection

8.0 Knowledge Acquisition and Presentation

8.1 Introduction

BOFD techniques rely on a good database of knowledge. Various tools are available for such knowledge acquisition and presentation. Essential matters that must be considered are the need to:

- Identify those processes and components on which the development of fault detection should focus;
- Collect data on practical ways in which faults are detected;
- Collect data on how faults are remedied.

Methods of knowledge acquisition include:

Quantitative Methods: these require statistical data on the occurrence of defects and faults in the systems to be studied. These data can be collected or may already exist.

Qualitative Methods: This approach involves discussion in which a group of experts analyse the system in question according to an agreed model and the results of the discussion are recorded in tabular form. It is applied when no (or little) quantitative data are available and is appropriate to the collection of expert system data. When using qualitative data, the amount of data recorded tends to become very large. Experts must therefore indicate which items are most important for the operation of the system.

8.2 Availability Performance

Consideration of the 'availability performance' offers a starting point for the development of appropriate tools. This is a measure of the system's ability to perform a task without downtime. For example, 95% availability means that a system will be impaired by a fault for, say, 50 hours in a period of 1000 hours. The ability of a system to perform a task depends just as much on availability performance as it does technical performance. In general, availability performance can be subdivided into reliability performance, maintainability and supportability as illustrated in Figure 18. Key factors are:

Reliability Performance: describes the ability to operate without defects for a given time. A systematic reliability performance is necessary which determines which stages are given greater or lesser emphasis or are emitted entirely i.e. it must be decided which chains or events the analysis will be directed at. This requires a functional description including:

- A breakdown of the system into functional groups based on system descriptions, diagrams and lists of components;
- A description of modes and use functions;
- A description of the usage profile (e.g. performance curves, shutdowns);
- Possible capacity or output level requirements.

Maintainability: This describes how soon after a failure the faulty equipment can return to operating condition after maintenance has been carried out.

Supportability: Describes the maintenance organisation's ability to organise and complete the necessary maintenance.

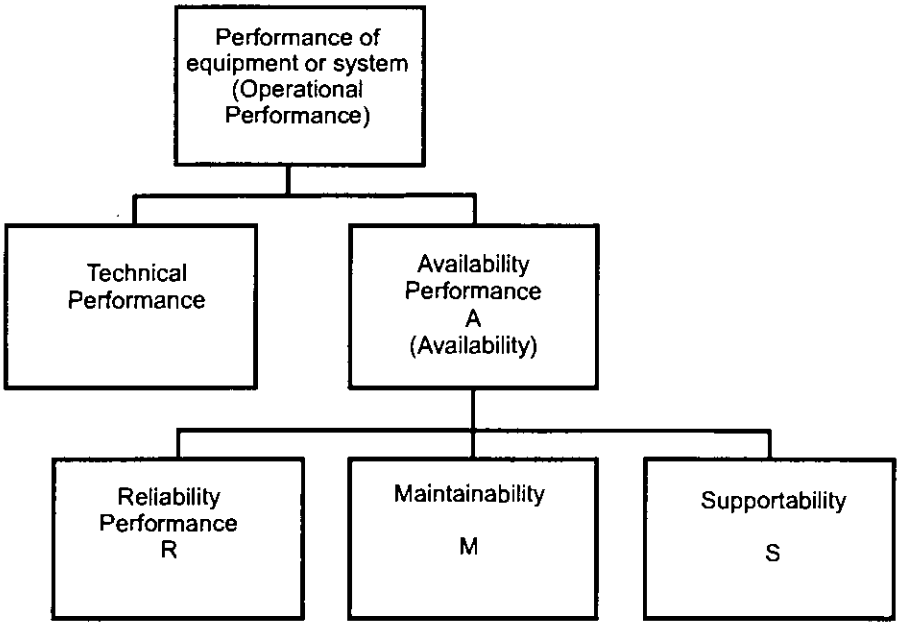

Figure 18 Operational Performance of a Technical System

8.3 Qualitative Availability Methods (Methods of 'Identification')

In carrying out design work, it is important to identify concealed or latent risk and problem points at the earliest stage and to find better solutions for them before the plant is installed or commissioned. Best adapted to this stage of analysis are so called methods of identification; these include:

Hazard and Operability Analysis (HAZOP): This approach was developed in the chemical industry to identify risks. Experts examine the process and develop a list of keywords (e.g. flow, temperature etc.) and a list of associated deviations (e.g. flow too large, too small, wrong flow direction etc. temperature too high, too low etc.). The experts reflect on the causes of deviations, their effects and possible actions that can be taken to reduce the risk. These results are recorded on a specific form.

Failure and Mode Effect Analysis (FMEA): This approach provides a systematic audit of the system's faults, the causes of them and their effects, by mode of failure, on the reliability performance and safety of the machine, system or plant. The purpose is to identify all the components and system parts that are important in causing a system shut down. FMEA does not call for preliminary information on reliability technology per se. The analysis is suited as a design tool because it is the design engineer who best knows the causes and effects related to equipment faults. The following criteria, among others may be relevant:

- Significance in terms of reliability, impact and safety;
- Deferability;
- Detectability;
- Possibility of continued system operation during repair.

Restrictions of FMEA include:

- Intrinsically this is a qualitative method. Quantitative data cannot be obtained on the system's overall reliability performance (although the analysis can be utilised in building a reliability model itself);
- In studying the safety of the system, data that are not essential from the standpoint of its safety are also examined;
- This approach does not encompass an examination of the effect of combinations of faults.

Maintenance and Effects Analysis (MEA): This approach concentrates on what happens after the fault has been diagnosed and the process should be repaired without significant downtime.

The Effects of Repairs and Maintenance on the System: This concentrates on how maintenance influences system operation.

The HAZOP and FMEA methods are well suited for pinpointing the important components of a HVAC system.

8.4 Classifying Qualitative Information

Since qualitative approaches can lead to large amounts of data, it is necessary that such data should be classified or prioritised. Evaluation and classifying involves the following stages:

Evaluation Criteria: The characteristics that are important for correct operation are listed on an 'evaluation grid' on which each of these characteristics is defined.

Weighting: The various characteristics are grouped according to an agreed hierarchy. Within each group, the significance of parameters are 'weighted' according to their importance by experts. The weights given to each parameter or characteristic are normalised such that the sum of the weights on each hierarchy level comes to 1 (or 100%).

Component Assessment: After having defined the criteria, their assessment grid and weights, the evaluation continues with an assessment of the value of a component's effect on the system characteristics in case of a fault. Each component or fault is assessed by asking, "What effect does a component or a specific fault have to the system?" This result is analysed and given a number from the assessment grid.

Priority Index: The last step is to calculate a weighting index that can be used for prioritisation.

8.5 Parameter Estimation

Parameters are specific quantities that are used in models which have been developed to simulate system (and fault) conditions. Accurate valuation of these parameters is necessary if a model is to perform correctly. In FDD there are basically two ways in which such models are used, i.e.:

Repeated Parameter Estimation: This is undertaken repeatedly during the monitoring phase. Parameter values are refined to enable accurate simulation of fault-free conditions. Subsequent changes in parameter values that may be necessary to simulate subsequent behaviour are then associated with specific faults.

Once Only Parameters Estimation: Parameters are estimated once and then kept constant. In general they are used to define some important system features, like fault-free states. The differences between measured and modelled conditions (i.e. the residuals) are generally the main focus of interest in fault detection.

It is useful to distinguish between these two approaches because the requirements on the parameters are quite different. In the latter case the choice of suitable model is the important task while, in the former case, other aspects must also be considered such as:

- Sensitivity of parameters to changes in system or component behaviour;
- Statistical features of parameters due to modelling and measurement errors. Statistical features of the parameter estimate because they have a direct impact on the possibility to detect and diagnose faulty performance.

8.6 Steady State (Fault) Detection

Many fault detection and diagnosis methods developed during the course of Annex 25 require the HVAC system to operate primarily in 'steady state'. Consequently, the ability to identify steady state conditions is essential to the operation of these FDD techniques. In practice, however, controlled HVAC systems responding to changes in weather cannot be expected to be found in steady states in the strictest sense of the definition. What is needed, therefore, is a practical criteria to identify when a HVAC system is quasi-stationary, i.e. operating in a sufficiently close approximation to steady state that the FDD method in question can be expected to yield reliable results. Filtering the BEMS data in order to identify such quasi-stationary states can be a key aspect of pre-processing. Conceptually, inputs and outputs of the operation or plant to be monitored are collected and pre-processed and then fed into a fault detector where decisions are made on correct and faulty behaviour of the plant (Figure 19).

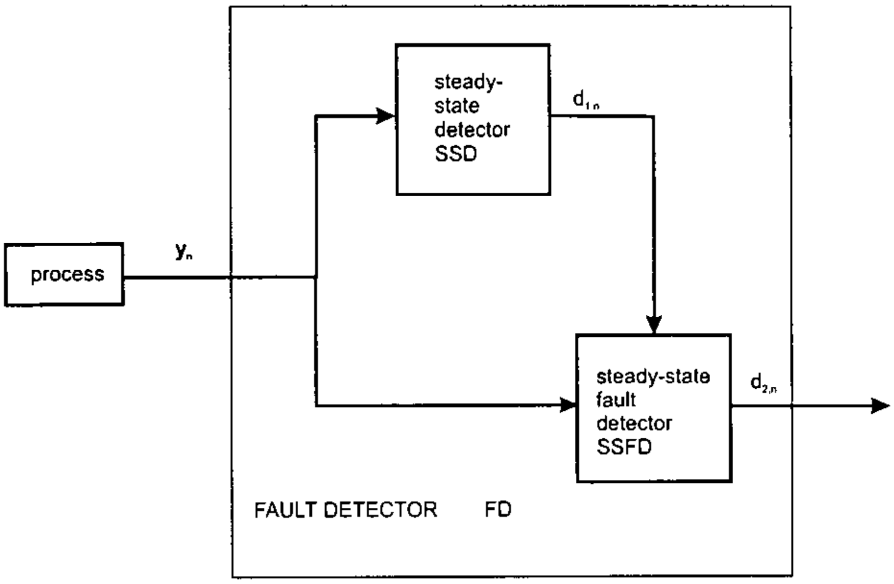

Figure 19 Block Diagram of the Fault Detection Process

If these decisions are based on the steady state behaviour of the plant, then the overall fault detector must incorporate a steady state detector (which outputs a steady state condition) and a steady state fault detector (which outputs a decision).

A classic problem in designing a successful fault detector is the necessity of striking a balance between two conflicting goals. i.e.:

- ***Maximum Sensitivity:*** This is aimed at ensuring that the probability that a genuine fault remains undetected is minimal.

- ***Maximum Reliability:*** The probability of false alarms should be minimal.

An analogous problem applies to the design of a steady state detector needed to identify those quasi-stationary states appropriate for a given fault detection procedure. i.e.:

- Those non-steady state conditions under which the FDD method might yield false alarms need to be excluded from the steady state detector;
- Steady state conditions should not be interpreted so strictly that they scarcely occur in practice.

Selecting a satisfactory compromise involves ‘tuning’ the relevant parameters in relation to the HVAC system and the FDD method being considered. The steady state processors described by Annex 25 comprise the following three processes and tuneable parameters:

1. The raw data from a particular sensor is subjected to low pass filtering to yield some sort of 'moving average'. The tuneable parameter is the effective 'time-window' used in the averaging procedure;
2. The available data are processed to obtain a measure of 'fluctuation'. The tuneable parameter is, again, the effective 'time-window', which may be chosen independently to that used in averaging;
3. The third parameter is a 'threshold'. Whenever the fluctuation measurement falls below the 'threshold', the sensor signal is deemed to be quasi-stationary.

In addition, the Steady State Fault Detector must be tuned in terms of 'thresholds'. Both types of thresholds may incorporate 'hysteresis' if it is desirable to avoid all too rapid changes in decision outputs. For example, the steady state detector can be programmed to switch on at a slightly lower threshold and switch off again at a somewhat higher value.

Moving Average: Two types of moving average have been the subject of Annex 25 investigation, these are:

- Moving averages using time windows of fixed length;
- Geometrically weighted moving averages.

Three principal measure fluctuations were used, i.e.:

- Moving average of the functional variation (i.e. the integral of the absolute value of the derivative);
- Moving variance using a time window of fixed length;
- Moving geometrically weighted variance.

9.0 Thresholds for Fault Detection, Diagnosis and Evaluation

9.1 Introduction

Fault detection indicates a deviation of performance from expectation, diagnosis determines the cause and evaluation assesses whether the impact is significant enough to justify service. In each of these steps it is necessary to define criteria or thresholds for establishing appropriate outputs. The output would be 'fault' or 'no fault' for fault detection, the type of fault for diagnosis and 'repair' or 'don't repair' for evaluation. Each of the thresholds associated with these outputs is unique and requires different estimation methods. These are described in further detail below.

9.2 Fault Detection Thresholds

Fault detection methods compare measurements with expectation to identify faults. The simplest methods involve range checking where the measurements are 'expected' to be within certain fixed bounds. These bounds must be set rather loosely to avoid false alarms. As a result, range checking can only be used to detect large changes in performance. Assuming that measurements are normally distributed about a mean, the bounds of range checking can be based on a specified confidence level (e.g. within 3 standard deviations).

The sensitivity for detecting faults is dramatically improved through the use of models of expected performance. The model predicts the outputs of a process for normal operation for given measured inputs. A fault is indicated when the difference between predicted and measured outputs (i.e. residuals) is greater than a threshold. Even under normal operation, the residuals are non-zero due to acceptable variations in model outputs, measurements etc. Figure 20 depicts the probability distribution for 'normal' and 'faulty' operation.

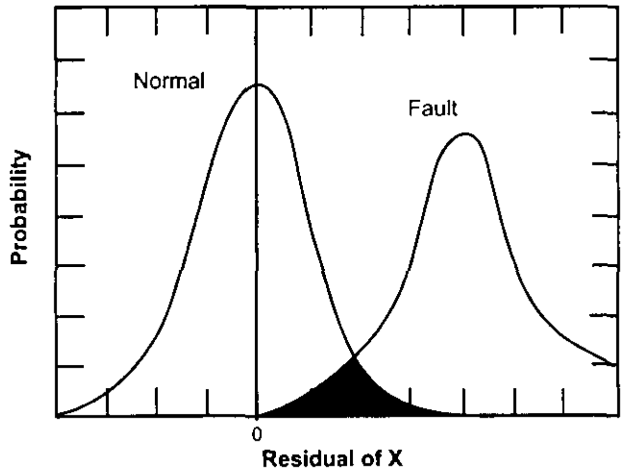

Figure 20 Fault Detection Example

The normal distribution has a zero mean, while the fault distribution moves away from zero as the fault severity increases. The shaded area, formed by the overlap of the two distributions, represents the probability that a fault diagnosis will be in error. The threshold should be set such that the shaded area is a small fraction of the area under the distributions. For example, a 99.9% confidence level for a correct fault diagnosis would be achieved when the shaded area is 10^{-3} of the total area.

9.3 Fault Diagnosis Methods

There are several different approaches to diagnosis once a fault has been detected; these include:

Fault Isolation: This is one of the simplest methods and is applied at the components level. In this case, the fault is diagnosed as soon as it is detected and no additional thresholds beyond fault detection thresholds are necessary. The main disadvantage is that many monitoring measurements are necessary.

Comparison of Measurements with Expectation: An alternative involves comparing physical parameters, as determined by measurement, with expected values for normal operation. For example, heat exchanger conductance could be estimated from the conditions at the inlet and outlet, to diagnose fouling.

Fault Models: A more common approach, requiring fewer measurements, involves the use of fault models. For each type of fault to be diagnosed, a fault model predicts the expected outputs associated with specific faults. A fault is diagnosed through the use of a classifier that attempts to find the fault model that gives the smallest error between predictions and measurements. With fault modelling, thresholds are generally necessary for whether the confidence in a particular diagnosis is strong enough to make a diagnostic decision. One approach would be to set a minimum acceptable limit on the probability that a correct diagnosis had been achieved. A second approach could be to set a minimum acceptable limit on the ratio of fault probabilities for the two most likely faults.

9.4 Service Thresholds

In general, equipment service should not be performed unless the benefit justifies the expense. For abrupt faults such as a motor failure, the benefit of service is often obvious and no additional thresholds are necessary for fault evaluation. However, for performance degradations such as heat exchanger fouling, a service decision should be based upon economic considerations.

In theory, economics alone could be used to determine if the cost of service is justified for all possible faults. The costs associated with faults that lead to down-time, safety hazards and environmental hazards are difficult to quantify but are generally much larger than the service costs required to repair them. With this assumption, the following four fault evaluation criteria may be used to identify the need for service:

1. **Comfort** - Service whenever the equipment cannot maintain comfort;
2. **Safety** - Service whenever equipment or personnel safety is compromised;
3. **Environment** - Service whenever the environment is adversely affected;
4. **Economic** - Service at intervals that minimises the combined costs of energy and service.

The thresholds associated with the first three criteria are relatively straightforward to specify. For comfort, the difference between the space temperature and the controller set point should be between bounds dictated by the 'expected' controller performance. For safety, thresholds should be based on experience of the particular equipment.

Thresholds associated with economic evaluation criteria are much more difficult to specify and additional work is necessary. Some simplified models have been developed which are aimed at optimising service schedules to minimise maintenance periods and limit excessive energy waste.

10.0 Typical Faults in HVAC Systems

10.1 Introduction

An important aspect of the Annex was to characterise typical faults in HVAC systems and to develop a 'fault' database. Of particular importance was the ranking of faults in order of significance. For each system, the structure took the form of:

- Defining a reference system;
- Identifying a fault list for the reference system;
- Ranking the significance of the faults;
- Producing a list of symptoms for the most important faults.

Systems included:

10.2 Heating Systems

Fault lists were produced for the following heating systems and associated components:

(i) Hydraulic Heating (and Domestic) Hot Water System

System:

- A hydraulic heating and hot water system for (1) a medium to large size residential building and (2) a non residential building;
- Multi-storey building;
- Heat generation using two gas boilers with a design temperature of 90°C;
- A distribution network split into two circuits (north facade and south facade). The temperature in each circuit is regulated using a three-way valve.

Associated Controls:

- Boiler control and sequencer;
- Scheduling of occupancy periods;
- Adaptive control of the flow temperature set-point of each secondary loop (weather compensation);
- Adaptation of start stop/time (optimum start/stop);
- Flow temperature control;
- Local temperature control (thermostatic valves).

(ii) District Heating Sub-distribution System

Heat energy is distributed from the district network to the radiator or domestic hot water network through a heat exchanger. Varying the water mass flow on the district heating side of the network controls the amount of heat transferred through the heat exchanger. The main processes of the sub-distribution system are;

Primary Circuit:

- Inlet;
- Domestic water branch;
- Heating branch;
- Return.

Secondary Circuit:

- Domestic Hot Water:
    - cold water;
    - circulated water;
    - hot water.

- Heating (Radiator Network).

Other Devices:

- Controls;
- Electrical devices;
- Supply energy equipment.

(iii) Oil Burner Systems

The burner burns oil to heat the boiler water. Oil is fed to the burner through oil feed lines where it is homogeneously mixed with air and ignited. The flue gas temperature in the chamber is 1400°C. The heat from the flue gas is transferred to the boiler water through chamber walls prior to the flue gas passing up the chimney.

Associated components include:

- Heating circuit;
- Domestic hot water circuit;
- Oil tank and oil circuitry;
- Air feeding circuit and the burner head;
- Furnace and flue gas channels;
- Burner control equipment (electrical circuits);
- Other control equipment (electrical circuits).

10.3 Chillers and Heat Pumps

Faults occurring in both vapour compression equipment and absorption refrigeration were analysed. Equipment and problems reviewed included:

Vapour Compression Refrigeration Equipment:

- Lack of refrigerant;
- Air in the refrigeration circuit;
- Liquid refrigerant in oil.

Absorption Refrigeration Equipment:

- Loss of vacuum;
- Clogging of the condenser tubes;
- Clogging of the evaporator tubes.

10.4 VAV Air Handling Units

A reference system was based on a single duct, pressure independent system. The air-handling unit has outdoor air and return air dampers, cooling and heating coils, an air filter section, a supply air fan, a return air fan and air plenum sections. The supply is ducted to three VAV boxes which supply conditioned air to three different zones. The return air fan takes air from the conditioned space and discharges it into the mixing section and/or the outside through a motorised exhaust air damper. Both the supply and return fans are fitted with variable frequency controllers for keeping static pressure constant in the air ducts. The pressure independent VAV system attempts to maintain a constant static pressure at the VAV box inlets by sensing and controlling the static pressure in the supply duct. A static pressure controller (with appropriate algorithm) sends a control signal to a variable frequency motor controller to vary the supply fan speed. The supply airflow rate is measured and the desired return airflow rate is calculated. The reference system also employs a dry-bulb type economiser cycle to save energy through the use of outdoor air free cooling.

Faults cover:

- Air mixing section;
- Filter coil section;
- Fan section;
- VAV boxes.

10.5 Thermal Storage Systems

Thermal storage systems are being widely used and contribute to the effective use of energy by reducing peak demand, heat recovery, solar energy utilisation and seasonal storage. Water ice, and other phase change material are used as the thermal storage media. The primary components and common faults are:

Thermal Storage Tank:

- Abnormal water level;
- Temperature problems.

Heat Source/Sink:

- Temperature problems;
- Flow rate of primary pump too low.

Variable Water Volume Piping and Control System:

- Temperature problems;
- Supply flow rate too high;
- Supply flow rate too low.

Constant Water Volume Piping and Control system:

- Temperature problems.

11.0 Conclusions

In developing fault diagnosis methods, a system concept was adhered to. This comprised two broad approaches. The first was based on monitoring an entire building or large system in which faults were detected and diagnosed. This was called optimising building use. The second was to deal with the detection of faults in individual components, devices or systems. These approaches, nevertheless, lead to the utilisation of the same methods of detection and diagnosis, and the difference between them is probably just philosophical.

The bulk of the work of this Annex was involved in the development of methods of detecting faults in components using process data obtained from system simulation. To some extent, development work was carried out with test rigs on a laboratory scale. Some fifty new methods for HVAC processes were developed or applied in this project.

The fault detection methods studied were based on the use of process models. There was much discussion on the classification of fault detection methods but a clear-cut recommendation was not reached although various ideas were presented. In addition to fault detection, a number of the methods that were developed also embodied diagnostic features. Both the fault detection and fault diagnosis methods that were developed can be used as part of a more extensive fault diagnosis system.

It pays to stick to simple process models. One way to achieve this is to model steady state conditions. However there must be pre-processing so that the tool is aware that the system is operating in a near steady state. One of the central problems is the setting of alarm limits. At this stage no general guidelines can be given but a practical solution is to set several limits that are based on different criteria and can be adjusted by the user.

The next step is the utilisation of results into practical applications. This is being undertaken by ECBCS Annex 34: Computer-Aided Fault Detection and Diagnosis.

The International Energy Agency (IEA) Energy Conservation in Buildings and Community Systems Programme (ECBCS)

The International Energy Agency (IEA) was established as an autonomous body within the Organisation for Economic Co-operation and Development (OECD) in1974, with the purpose of strengthening co-operation in the vital area of energy policy. As one element of this programme, member countries take part in various energy research, development and demonstration activities. The Energy Conservation in Buildings and Community Systems Programme has sponsored various research annexes associated with energy prediction, monitoring and energy efficiency measures in both new and existing buildings. The results have provided much valuable information about the state of the art of building analysis and have led to further IEA sponsored research