Fault Diagnosis Using a Sequential Integration of Model-Free and Model-Based Approaches

BY

Muhammad Haris Khalid

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

Systems Engineering

April 2009

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA

This thesis, written by **MUHAMMAD HARIS KHALID** under the direction of his thesis advisor and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of MASTERS OF SCIENCE IN SYSTEMS ENGINEERING.

Thesis Committee

Dr. Rajamani Doraiswami
Thesis Advisor

Dr. Lahouari Cheded
Member

Dr. Naim Mohammad Faqir
Member

Dr. Fouad Al-Sunni
Department Chairman

Dr. Salam A. Zummo
Dean of Graduate Studies

28/4/09

Date

To my parents, brother, and sister.

Acknowledgements

With deep humility, I thank Almighty Allah, the merciful and beneficient, for his constantly reliable help and his everlasting inspiration.

I would like to express my deepest thanks and appreciation to my thesis advisor Dr. Rajamani Doraiswami, for his consistent help, guidance and attention that he devoted throughout the course of this work. His valuable suggestions and useful discussions made this work interesting to me. His support and words of encouragement gave a new life to my efforts in hard times. Without his help and valuable directions, this work would have never seen the light.

I am also grateful to my thesis committee member Dr. Lahouari Cheded who was just like a co-advisor to me. He deeply guided me on the critical aspects of the work. Thanks are also due to my other thesis committee member Dr. Naim Mohammad Faqir for his interest, cooperation and insightful feedback.

I am also thankful to the Chairman of my Department, Dr. Fouad Al-Sunni who has been always supportive and provided enormous help in this research. I wish to thank him from my heart.

The support and facilities provided by King Fahd University of Petroleum and Minerals for completing this work are highly appreciated.

Table of Contents

List of Tables

List of Figures

Thesis Abstract

Name: Muhammad Haris Khalid
Title: Fault Diagnosis Using a Sequential Integration of Model-free and Model-based Approaches
Major Field: SYSTEMS ENGINEERING
Date of Degree: April 2009

In this thesis, a diagnosis scheme is proposed for detecting and isolating incipient faults by using the sequential integration of two totally different approaches, namely model-free and model-based, ensuring thereby that crucial information about the presence or absence of a fault is monitored in the shortest possible time and the complete status regarding the fault is unfolded in time. Model-free approaches include limit and plausibility checks and knowledge-based analysis, while model-based approaches include the Kalman filter for fault detection and a parameter identification scheme for fault isolation. The model-free analysis indicates quickly a possible onset of fault, whereas the Kalman filter detects its presence/absence, and finally the identification scheme isolates any fault detected. To obtain an overall diagnostic picture, a reliable and accurate model is essential. Such a model is determined by the analysis of residual, comparison of frequency response, and the accurate determination of model structure based on location of poles of the identified model. A novel method to identify the structure of the model is based on the location of the poles. The proposed scheme is extensively evaluated on physical systems. The execution of the proposed combined scheme may also be used in designing an effective preventive maintenance strategy. The proposed scheme is evaluated on a physical fluid system exemplified by a benchmarked laboratory-scale two-tank system.

Master of Science Degree
King Fahd Petroleum and Minerals, Dhahran, Saudi Arabia
April 2009

ملخص الأطروحة

الاسم: مد حارث خالد

العنوان: تشخيص ا طأ باستخدام دمج متتابع للمناهج ا الية من وذج و ا ناهج ا عتمدة على وذج

التخصص الرئيس: هندسة النظم

تاريخ الدرجة العلمية: أبريل 2009

هذه الأطروحة، يقترح ططٌ للتشخيص للكشف عن الأخطاء الأولية وفرزها عن طريق الدمج ا تتابع لنوعي تلف كليا من ا ناهج: ا ناهج ا الية من وذج وا ناهج ا عتمدة على وذج، وبالتا ضمان مراقبة ا علومة ا ة عن غياب أو وجود ا طأ أقصر وقت كن، وأيضا كشف ا الة الكلية للخطأ مع الوقت. ا ناهج ا الية من وذج تشمل فحص ا د وفحص ا عقولية وأيضا التحليل القائم على ا عرفة. بينما ا ناهج ا عتمدة على وذج فتشتمل على "مرشح كا ان" (Kalman Filter) لكشف الأخطاء، وعلى التعرف على معاملات النموذج لفرز الأخطاء. التحليل ا ا من وذج يدل بسرعة على امكانية بداية خطإٍ ما، ح أن "مرشح كا ان" يكشف عن وجود أو غياب ا طأ، ويقوم التعرف على ا عاملات بفرز أي خطأ مكتشف. للحصول على صورة شاملة للتشخيص هناك حاجة أساسية لنموذج دقيق و موثوق. وذج كهذا يحدد عن طريق ليل الفرق ب قيم القياسات و قيم النموذج ، و عن طريق مقارنة الاستجابة الترددية ، و عن طريق التحديد الدقيق لبنية النموذج اعتماداً على موقع أقطاب النموذج ا ُعرف. طريقة جديدة للتعرف على بنية النموذج تعتمد على موقع الأقطاب. ا خطط ا قترح يقيم بشكل موسع على أنظمة فعلية. تنفيذ ا خطط ا قترح إ الا قد يستخدم تصميم استراتيجية فعالة للصيانة الوقائية. ا خطط ا قترح يتم على تقييم فعلي على نظام موائع متمثل بنظام مقياسي ذي خزان ا خت .

درجة ا اجست العلوم

جامعة ا لك فهد للبترول و ا عادن ، الظهران ، ا ملكة العربية السعودية

أبريل 2009

Chapter 1

Basic Definitions

Overview

This chapter discusses the definitions of the terminologies which are used in this thesis.

As a step towards a unified terminology, the IFAC Technical Committee SAFEPROCESS suggested preliminary definitions of some terms in the field of fault diagnosis. Some of these definitions are given here as a way to introduce the field. Most of these terms will be defined more formally later in this thesis.

The following list of definitions is a subset of the IFAC list:

- *Fault:* Unpermitted deviation of at least one characteristic property or variable of the system from acceptable/usual/standard behavior.
- *Failure:* Permanent interruption of a system's ability to perform a required function under specified operating conditions.
- *Fault Detection:* Determination of faults present in a system, and its time of detection.
- *Fault Isolation:* Determination of kind, location, and time of detection of a fault. Follows fault detection.
- *Fault Identification:* Determination of the size and time-variant behavior of a fault. Follows fault isolation.
- *Fault Diagnosis:* Determination of kind, size, location, and time of

detection of a fault. Follows fault detection. Includes fault isolation and identification. For the term *fault diagnosis*, one slightly different definition also exists in the literature, for example in Gertler (1991) and it says that *fault diagnosis* also includes *fault detection*.

- *Fault Evaluation:* to estimate the size and type or nature of the fault.

If fault detection is excluded from the term *diagnosis*, as in the SAFEPROCESS, one faces the problem of finding a word to describe the whole area. This is partly solved by introducing the abbreviation FDI (Fault Detection and Isolation), which is commonly used in many papers.

The relative importance of these tasks is obviously subjective. However, detection is an absolute must for any practical system, and isolation is almost equally important. Fault evaluation, on the other hand, may not be essential if no reconfiguration action is involved. Hence, in the literature, fault diagnosis is very often considered as fault detection and isolation, abbreviated as FDI.

1.1 Dictionary Definitions

In this context, it is also interesting to see how the word *diagnosis* is defined in a general dictionary such as Webster:

diagnosis

Etymology: New Latin, from Greek diagnōsis, from diagignōskein to distinguish, from dia- + gignōskein to know.
Date: circa 1681

1 a : the art or act of identifying a disease from its signs and symptoms *b : the decision reached by diagnosis*

2 a : investigation or analysis of the cause or nature of a condition, situation, or problem <diagnosis of engine trouble>

b : a statement or conclusion from such an analysis.

Chapter 2

Introduction

Overview

This chapter introduces the context of the thesis. It briefly describes related topics such as model-based and model-free approaches, system identification and fault diagnosis. The last sections present the summary of the proposed schemes, and the thesis organization will also be discussed in this chapter.

2.1 Literature Survey

Fault diagnosis is continuously gaining importance for process monitoring because of the increasing demand for higher performance as well as for increased safety and reliability of dynamic systems. Early diagnosis of process faults, while the system is still operating in a controllable region, can help avoid abnormal event progression. It can reduce (or possibly avoid) productivity loss, which in turn can help avoid major system breakdowns and catastrophes. Hence, fault diagnosis is a major research topic, attracting considerable interest from industrial practitioners as well as academic researchers. The survey [1] revealed that the US-based petrochemical industry could save up to $10 billion annually if abnormal process behavior could be detected, diagnosed, and appropriately dealt with. Studies suggest [2] that the petrochemical industry alone loses over $20 billion annually due to inappropriate reactions to abnormal behavior.

There is an abundance of literature on process fault diagnosis, ranging from analytical methods to artificial intelligence and statistical approaches. From a modeling perspective, there are methods that require accurate process models, semi-quantitative models, or qualitative models. At the other end of the spectrum, there are methods that do not assume any form of model information but rely only on historic process data. In addition, given the process knowledge, there are different search techniques that can be used for diagnosis. Fault diagnosis methods surveyed in [15], [16] and [17] can be classified into 2 general categories, model-based and model-free (also termed data-based or signal-based) depending upon the *a priori* knowledge of the process.. The hierarchy of fault diagnosis approaches is shown in Fig. 1.1.

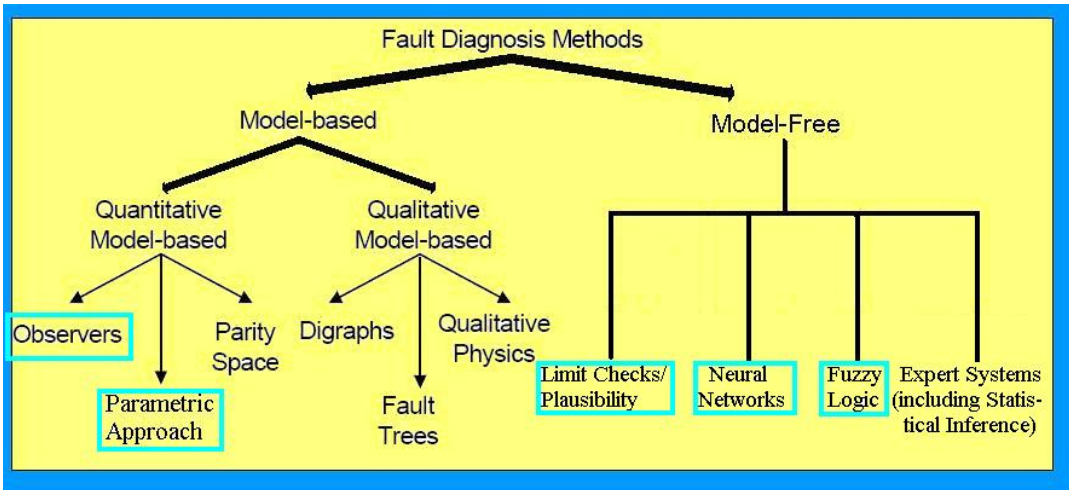


Figure 1 Taxonomy of Fault Diagnosis Methods

The type of *a-priori* process knowledge used is the most important distinguishing feature for classifying fault diagnosis systems. The basic *a-priori* knowledge that is needed for fault diagnosis is the set of faults and the relationship between the observations (symptoms) and the faults themselves. A diagnostic system may show these relationships explicitly (as in a table lookup), or it may infer them from some source of domain knowledge. The *a-priori* domain knowledge may be developed from a fundamental understanding of the process by using first-principles knowledge. Such knowledge is referred to in [18] as deep, causal or model-based knowledge. On the other hand, knowledge may be obtained from a past experience with the process. This knowledge is referred to as shallow, evidential or process history-based knowledge.

For model-based methods, the a-priori knowledge can be broadly classified as qualitative or quantitative. The model is usually based on fundamental understanding of the physics of the process. In quantitative models, this understanding is expressed in terms of mathematical functional relationships between the inputs and outputs of the system. In qualitative model equations, these relationships are expressed in terms of qualitative functions.

For model-free approaches, only the availability of a large amount of historical process data is assumed. There are different ways in which this data can be transformed and presented as a-priori knowledge to a diagnostic system. This is known as the feature extraction process from the process history data, and it is done to facilitate later diagnosis. This extraction process can proceed mainly as either a quantitative or a qualitative feature extraction process. Quantitative feature extraction can be either statistical or non-statistical.

There might be some overlap between the model-based and model-free approaches. This is just one classification, based on whether or not the knowledge about process characteristics is required.

2.1.1 Model-based Fault Diagnosis Approaches

2.1.1.1 Quantitative Model-based Methods

Most of the work on quantitative model-based approaches has been based on using general input-output and state space models to generate residuals. These approaches can be classified into observer, parity space and frequency domain methods. Good survey papers include [23], [24], and [25]. The mathematical model-based approach adopted in this thesis falls into the observer category.

- **Observer or filter-based approaches:** The basic idea behind the observer or filter-based approaches is to estimate the outputs of the system from the measurements (or a subset of measurements) by using either observers in a deterministic setting [26-31] or statistical filters (e.g. the Kalman filter) in a stochastic setting [32-35]. Then, the weighted output estimation errors (or innovations in the stochastic case) are used as the residuals. Depending on the circumstances, one may use linear [36] or nonlinear [37-39], full or reduced-order, fixed or adaptive observers (or Kalman filters) [40-41].
- **Parity space approaches:** Parity equations are rearranged and usually transformed variants of input-output or state space models of the plant [43]. The basic idea is to check the parity (consistency) of the plant models with sensor outputs (measurements) and known process inputs. The idea of this approach is to rearrange the model structure so as to get the best fault isolation. Dynamic parity relations were introduced by Willsky [44]. Redundancy provides freedom in the design of residual generating equations so that further fault isolation can be achieved.

Fault isolation requires the ability to generate residual vectors which are orthogonal to each other for different faults. Gertler et al. [45-46] suggested a so-called "orthogonal parity equation" approach in designing structured residual sets. The design of directional residual vectors using parity relations is not straightforward. The systematic approaches of designing parity equations with directional properties are presented in [47] and [48]. Chow and Willsky [49] proposed a procedure to generate parity equations from the state space representation of a dynamic system. Several researchers showed that some correspondence exists between observer-based and parity relation approaches. A full derivation of this structure equivalence can be found in [50].

- **Parametric Approach:** The parametric approach [8, 9, 13] is based on analyzing a feature vector which is computed by using on-line recursive identification. The feature vector usually represents the coefficients of a system transfer function. A failure is detected when the estimated value of the feature vector migrates from its nominal value. In practice the estimated value of feature vector may not converge properly as a result of model uncertainty and measurement noise. The choice of the order of the identified model and requirements for persistency of excitation are important considerations. Also, if the system represents an overall transfer function that consists of an interconnection of component transfer functions, then a change in one of the (diagnostic) parameters within the subsystem will, in general, affect all of the elements of the feature vector. Therefore the relationship between the diagnostic parameter vector and the feature vector [10] must be known or determined beforehand if such a fault is to be isolated.

2.1.1.2 Qualitative Model-based Methods

Based on various forms of qualitative knowledge used in fault diagnosis, qualitative model-based approaches can be classified into digraphs, fault trees and qualitative physics methods.

- **Causal model approaches using digraphs:** Cause-effect relations or models can be represented in the form of signed digraphs (SDG). A digraph is a graph with directed arcs between the nodes and SDG in which the directed arcs have a positive or negative sign attached to them. The directed arcs lead from the 'cause' nodes to the 'effect' nodes. SDGs provide a very efficient way of representing qualitative models graphically, and they are the most widely used form of causal knowledge for process fault diagnosis. Iri et al. [54] were the first to use SDG for fault diagnosis. From SDG, they derived what is called a cause-effect graph (CE graph). Umeda et al. [55] showed how SDG can be obtained from differential algebraic equations for the process. Shiozaki et al. [56] addressed the issue of conditional arcs in their SDG representation. Shiozaki et al. [57] also extended the idea of SDG to include five-range patterns instead of the usual three-range pattern used in the standard SDG. Kokawa et al. [58] used partial system dynamics, statistical information about equipment failure, and digraphs to represent the failure propagation network for identifying fault location.

  Rule-based methods using SDG have been used for fault diagnosis by

Kramer and Palowitch [59]. An important work in the field of steady-state qualitative simulation (QSIM) using SDG has been presented by Oyeleye and Kramer [60]. In recent years, Wilcox and Himmelblau [61] [62] have approached the problem of fault diagnosis using what is called possible cause and effect graph (PCEG) models. Vaidhyanathan and Venkatasubramanian [63] have used digraph-based models for automated HAZOP analysis. Use of SDGs for multiple fault detection is demonstrated by Vedam and Venkatasubramanian [64]. Improvement of fault resolution in SDG models through the use of fuzzy set theory is discussed by Han et al. [65]. Genovesi et al. [66] have presented a framework for process supervision using fuzzy logic-based fault diagnosis. Li and Wang [67] have shown how fuzzy digraphs can be used for qualitative and quantitative simulation of the temporal behavior of process systems.

- **Fault trees approaches:** Fault trees are used in analyzing system reliability and safety. Fault tree analysis was originally developed at Bell Telephone Laboratories in 1961. A fault tree is a logic tree that propagates primary events or faults to the top level event or a hazard. The tree usually has layers of nodes. At each node different logic operations such as AND OR are performed for propagation. Fault-trees have been used in a variety of risk assessment and reliability analysis studies by Kelly and Lees [68].
- **Qualitative physics approaches:** Qualitative physics knowledge in fault diagnosis has been represented in mainly two ways. The first approach is to derive qualitative equations from the differential equations termed as confluence equations. Considerable work has been done in this area of

qualitative modeling of systems and representation of causal knowledge, including [69], [70] and [71]. The other approach in qualitative physics was the derivation of qualitative behavior from the ordinary differential equations (ODEs). These qualitative behaviors for different failures can be used as a knowledge source. Sacks [72] examined piece-wise linear approximations of nonlinear differential equations through the use of a qualitative mathematical reasoner to deduce the qualitative properties of the system. Kuipers [73] predicted qualitative behavior by using qualitative differential equations (QDEs) that are an abstraction of the ODEs that represent the state of the system. In terms of applications of qualitative models in fault diagnosis, qualitative simulation (QSIM) and qualitative process theory (QPT) are the popular approaches. Examples of research work in QSIM include [74], [75], and [76]. Examples of using the QPT framework in process fault diagnosis include [77], [78] and [79].

2.1.2 Model-Free Fault Diagnosis Approaches

- **Expert system approaches:** Rule-based feature extraction has been widely used in expert systems for many applications. An expert system is generally a very specialized system that solves problems in a narrow domain of expertise. Initial attempts at the application of expert systems for fault diagnosis can be found in the work of Henley [80] and Niida [81]. Structuring the knowledge-base through hierarchical classification can be found in [82]. Ideas on knowledge-based diagnostic systems based on the task framework can be found in [83]. A rule-based expert system for fault diagnosis in a cracker unit is described in [84]. More work on expert systems in chemical process fault diagnosis can be found in [85] and [86].

A number of other researchers have also worked on the application of expert systems to diagnostic problems. Basila et al. [87] developed a supervisory expert system that uses object-based knowledge representation to represent heuristic and model-based knowledge. Zhang and Roberts [88] presented a methodology for formulating diagnostic rules from the knowledge of system structures and component functions. Becraft and Lee [89] proposed an integrated framework comprising of a neural network and an expert system. Tarifa and Scenna [90] proposed a hybrid system that uses signed directed graphs (SDG) and fuzzy logic. Zhao et al. [91] presented a wavelet sigmoid basis neural network and expert system based integrated framework for fault diagnosis of a hydrocracking process. Wo et al. [92] presented an expert fault diagnostic system that uses rules with certainty factors. Leung and Romagnoli [93] presented a probabilistic model-based expert system for fault diagnosis. An expert system approach for fault diagnosis in batch processes was discussed in Scenna [94].

- **Neural Networks approaches:** Considerable interest was shown in the literature regarding the application of neural networks for fault diagnosis. In general, neural networks that was used for fault diagnosis can be classified along two dimensions: (i) the architecture of the network such as sigmoidal, radial basis and so on; and (ii) the learning strategy such as supervised and unsupervised learning.

The most popular supervised learning strategy in neural networks has been the back-propagation algorithm. A number of papers address the problem of fault diagnosis using back-propagation neural networks. In chemical engineering, Watanabe et al. [95], Venkatasubramanian and Chan [96], Ungar et al. [97] and Hoskins et al. [98] were among the first researchers to demonstrate the usefulness of neural networks for fault diagnosis. A detailed and thorough analysis of neural networks for fault diagnosis in steady-state processes was presented by Venkatasubramanian et al. [99]. This work was later extended to utilize dynamic process data by Vaidyanathan and Venkatasubramanian [100]. A hierarchical neural network architecture for the detection of multiple faults was proposed by Watanabe et al. [101].

Most of the work on improvement of performance of standard back-propagation neural networks for fault diagnosis is based on the idea of explicit feature presentation to the neural networks by Fan et al. [102], Farell and Roat [103], Tsai and Chang [104], and Maki and Loparo [105]. Modifications to the selection of basis functions have also been suggested to the standard back-propagation network. For example, Leonard and Kramer [106] suggested the use of radial basis function networks for fault diagnosis applications. Kavuri and Venkatasubramanian ([107]; [108]; [109]) generalized radial units to Gaussian units and proposed methods to solve the hidden node problem.

[117] discuss the integration of wavelets with ART networks for the development of diagnostic systems.

- **Fuzzy Logic approaches:** In recent years the application of fuzzy logic to model-based fault diagnosis approaches has gained increasing attention in both fundamental research and application. *Symptoms* can be generated using observers based on the estimation of the output from the system. The first methods used fuzzy set theory to express cause-effect relations in expert systems. The key idea of model-based methods is the generation of signals, termed *residuals*. These are usually generated by using mathematical methods (based on state observers, parameter estimation or parity equations). The models correspond to the monitored system (Chen & Patton, 1999).

  Residuals are signals representing inconsistencies between the model and the actual system being monitored, but the deviation between the model and the plant is influenced not only by the presence of the fault but also modeling uncertainty. One solution is for the observer and controller parameters to be tuned via estimation from the real system for fault isolation and threshold adaptation (Schneider & Frank, 1994). The introduction of fuzzy logic can improve the decision-making, and in turn it will provide reliable and sufficient FDI, suitable for real industrial applications. However, difficulties arise in the training of the algorithm in the inference mechanism, where knowledge is hidden in large amounts of data and embedded in trained neural networks (Chen *et al*., 1997).

  A fuzzy feed-forward neural network (FNN) is applied to extract rules from an existing data base. Frank et al. (Frank & Kuipel, 1993; Frank 1993; Frank 1994a; Frank 1996; Schneider & Frank, 1996; Frank &

Köppen- Seliger 1997) use fuzzy logic for residual evaluation. This can be an important way of taking into account modeling uncertainty at decision-making rather than during residual generator design. By applying a fuzzy rule-based approach, the fault decision process can be made robust to the uncertainties so that false and missed alarm rates can be minimized. Considering *supervisory control* (Linkens et al., 1993, Frank & Kuipel, 1993) with tasks such as system management, process monitoring, identification, fault detection, diagnosis and adaptive capability reduces to lower level models for developing simpler structures for observers and controllers using TS fuzzy models.

- **Multivariate Statistical approaches:** The successful applications of multivariate statistical methods to fault diagnosis such as Principal Component Analysis (PCA) and Partial Least Squares (PLS) have been extensively reported in the literature. Overviews of using PCA and PLS in fault diagnosis and in process analysis and control were given by MacGregor et al. [118] [119], and Wise and Gallagher [120].

  In earlier work, Kresta et al. [121] presented the basic methodology of using the multivariate SPC procedure to handle large numbers of process and quality variables for continuous processes. Later, Nomikos and MacGregor [122] extended the use of multivariate projection methods to batch processes by using multiway PCA. To deal with nonlinearity, Qin and McAvoy [123] proposed a neural net PLS approach that incorporated feed forward networks into the PLS modeling. In order to handle nonlinearity in batch processes, Dong and McAvoy [124] utilized a nonlinear PCA method. Raich and Cinar [125] [126] proposed an integral statistical methodology combining PCA and discriminate analysis techniques, using distance and angle-based discriminants.

Chemometrics is defined as the science of relating measurements made on a chemical system to the state of the system via application of mathematical or statistical methods. Chemometric techniques have been applied in recent years to chemical engineering processes [127-128]. Dunia et al. [130], and Qin and Li [129] used PCA for sensor fault detection, identification and reconstruction. Dunia and Qin [131] looked at PCA from a geometric point of view and presented a methodology that analyzed the fault subspace for process and sensor fault detection.

A major limitation of conventional PCA monitoring is that the PCA model is time-invariant, while most real processes are time-varying. Hence the PCA model should also be recursively updated. An adaptive monitoring approach using recursive PLS has been presented by Qin [132], and a similar recursive PCA approach was proposed by Li et al. [133]. Another promising variant of the PCA approach is the multi-scale PCA (MSPCA) approach which integrates PCA and wavelet analysis [134-137].

- **Statistical Classifier approaches:** Fault diagnosis is essentially a classification problem and hence can be cast in a classical statistical pattern recognition framework [10]. Fault diagnosis can be considered as a problem of combining, over time, the instantaneous estimates of the classifier using knowledge about the statistical properties of the failure modes of the system ([138]; [139]).

2.2 Thesis Summary

In this thesis, a novel approach to sequential integration for fault diagnosis is developed on the basis of [160].

The faults considered here include sensor, actuator and leakage faults, and can be classified broadly as either parametric or additive faults. An additive fault manifests itself as an additive exogenous signal in the measured data, while a parametric fault induces a variation in the system parameters.

A fault signature manifests itself as an abrupt jump or a change either in the signal profile, or in the signal spectral characteristics (coherence spectrum), or in the signal statistics or in any other signal characteristics used. Knowledge-based methods employ knowledge gained from experts, data history records, extensive experimentation, and physical laws governing the physical system (acknowledged from [163]).

In general, there are two broad classes of fault diagnosis: model-free and model-based. The former class includes tools based on limits checking, plausibility analysis, neural networks, fuzzy logic, signal coherence spectra, statistical inference and artificial intelligence. A model-free approach is capable of detecting a possible fault quickly, unraveling its root cause(s) and isolating it. Being free from the use of a model, it has an equally attractive freedom from the usual model-related difficulties, such as identifying the required model, dealing with the presence of nonlinearities and structural complexities. However, these advantages are realized at a cost that may have various facets depending on the tool used. For neural networks, there is a lack of transparency, a need for a sufficient training data covering all or most operational scenarios, and a possibly lengthy training time.

Fuzzy logic techniques, though less opaque than neural networks, suffer from the difficulty of deriving precise rules which distil an expert's knowledge of the application domain and which are necessary to drive the fuzzy inference engine [147] (acknowledged from [163]).

On the other hand, given the availability of an appropriate model, the model-based method is transparent, and it provides a complete and accurate diagnostic picture by exploiting a wealth of readily available and powerful tools for analysis and design. Fortunately, the well-known difficulties in identifying a system model, due to its structural complexities and nonlinearities that may render its mathematical analysis intractable and its processing slow, can, for a vast number of practical systems, be mitigated by resorting to simple linearized models that are quite adequate in capturing most of the system dynamics of interest and whose predictive and inferential power can be enhanced by a rich repertoire of powerful linear analytical tools (acknowledged from [163]).

The model-based approach is based on the use of Kalman filtering [148-151], parity equation [152] system identification [153], and diagnostic model [153, 154]. In critical applications such as those involving hazardous leaks, it is important to ensure that a fault is detected quickly and reliably [155,156].

Therefore, the key advantage of our proposed sequential integration scheme is to harness the advantages of both approaches, which confer upon it an enhanced capability that neither of the two approaches enjoys. More specifically, the model-based approach provides not only a confirmation of the quick fault diagnosis made by the model-free approach but also an accurate unfolding-in-time of the finer details of the fault, thus completing the overall diagnostic picture of the system under test (acknowledged from [163]).

In this thesis, a novel scheme which combines the advantages of both model-free and model-based approaches to diagnose an incipient fault quickly and accurately is proposed using the framework of [154]. This scheme hinges on the sequential integration of model-free and model-based approaches.

When identifying a model of a physical system, the structure of the model may not be identical to that of the mathematical model derived from the physical laws due to various factors including the presence of noise and fast dynamics [154, 157]. A simple approach to estimating the mismatch between these two models is to choose as a measure of the goodness-of-fit between the outputs of the system and its mathematical model to be the sum of the squares of the residuals, defined as the difference between these two outputs. The selection of the model order based on this measure may be unsatisfactory. Increasing the order of the system and hence the number of estimated parameters may improve the goodness-of-fit. However, choosing a higher order model may result in overfitting the data, and consequently may include noise artifacts. To overcome the problem of overfitting, a number of model order selection criteria have been proposed. These criteria are based on penalizing not only the sum of the squares of the residuals but also the model order itself. Commonly-used criteria include Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Minimum Description Length (MDL) [149]. These are based on statistical decision theory requiring *a priori* knowledge of the probability distribution function (PDF) of the noise. In practical systems, it may not be possible to estimate the required PDF, and hence a Gaussian PDF is generally assumed [150-157]. In many cases, the application of these criteria may not always give the correct model order as the estimated model may still contain some artifacts due to noise and other causes.

This work adopts the novel approach to model order selection that was proposed in [161] and which directly verifies the presence of various artifacts in the estimated model. These artifacts manifest themselves by the presence of extraneous poles in the identified model. It is shown here that, for systems exhibiting a low-pass nature, which is the case for most practical systems, if the sampling frequency is chosen larger than four times the system's bandwidth (i.e. twice the Nyquist rate), the system poles will then be located in the right-half of the z-plane. This *a-priori* knowledge is exploited here to distinguish the system poles from the extraneous ones which fall in the left-half of the z-plane. The model order is selected so that all the poles of the identified model are in the right half of the z-plane, and the resulting identified model will correctly reflect the 'true model' (acknowledged from [163]).

A most popular approach to fault detection is based on the residuals generated by the Kalman filter [148-151]. A suboptimal steady-state Kalman filter, whose structure is similar to that of an observer, is used as it is computationally simple and has also been successfully used in a plethora of practical fault diagnosis applications [148-151]. A statistical decision-theoretic approach using the generalized likelihood ratio is employed to detect the presence or absence of a fault. The mean of the residual vector or the mean of its correlation is compared with a specified threshold value, which is determined from both the pre-specified value of the false alarm rate, and the variance of the measurement noise.

Next, the fault isolation problem is considered. The overall system modeled as an interconnection of subsystems, $G_i(z)$, i = 1, 2, … n_f, is shown in Fig. 1. Each subsystem, $G_i(z)$, is a transfer function that may represent a physical entity such as a sensor, actuator, controller or any other system component that is subject to a fault [157]. Each subsystem may be affected by some noise or disturbance inputs, w_i, as illustrated in Fig. 2. The feature vector of a particular $G_i(z)$ forms a $(q_i \times 1)$ vector, γ_i. The diagnostic parameter, γ, is a $(q \times 1)$ vector that augments the feature vectors of all subsystems, γ_i, i = 1, 2, …, n_f.

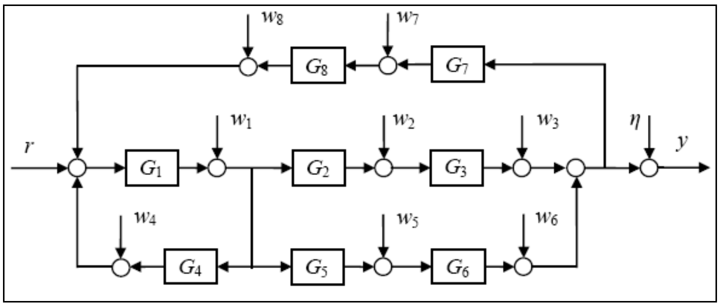


Figure 2 Interconnection of Subsystems

The Kalman filter is generally not suited for parametric fault isolation. There are two approaches to fault isolation. One is based on parameter identification where the feature vector, made of the coefficients of the system transfer function, is first estimated, and then the known relationship between this feature vector and the diagnostic parameter is used to derive the latter [155]. The other method is based on a diagnostic model which relates directly the input-output data to the diagnostic parameters [156,157]. In this thesis, we employ the diagnostic model approach. The proposed scheme is evaluated extensively on a simulated system as well as on a benchmarked laboratory-scale two-tank system [158].

The main contributions of the thesis are the sequential integration of model-free and model-based approaches, the new model order selection criterion and finally its impact on both the accuracy and reliability of fault detection based on a Kalman filter and on fault isolation using a diagnostic model approach.

2.3 Thesis Organization

The thesis is organized as follows:

A. **Chapter 1** gives the Basic Standard Definitions for the terminologies used.

B. **Chapter 2** is the Introduction. It gives a summary of the previous work reported in the literature and an overview of the topics discussed in the thesis.

C. **Chapter 3** describes the proposed sequential integration approach.

D. **Chapter 4** presents the proposed model-free approach for fault detection.

E. **Chapter 5** presents the proposed model-based approach for Fault Detection and Isolation (FDI) which includes a novel model order selection criterion.

F. **Chapter 6** addresses the evaluation of the proposed sequential integration of the fault diagnosis scheme on a benchmark two-tank system.

G. **Chapter 7** gives the performance analysis of the proposed scheme for fault diagnosis.

H. **Chapter 8** presents the Thesis contributions, conclusions and recommendations for future work.

Chapter 3

Proposed Sequential Integration Approach for Fault Diagnosis (FDI)

Overview

This chapter gives the main theme of the Proposed Sequential Integration Scheme for Fault Diagnosis and discusses its two building blocks, namely model-free and model-based approaches, and their various components.

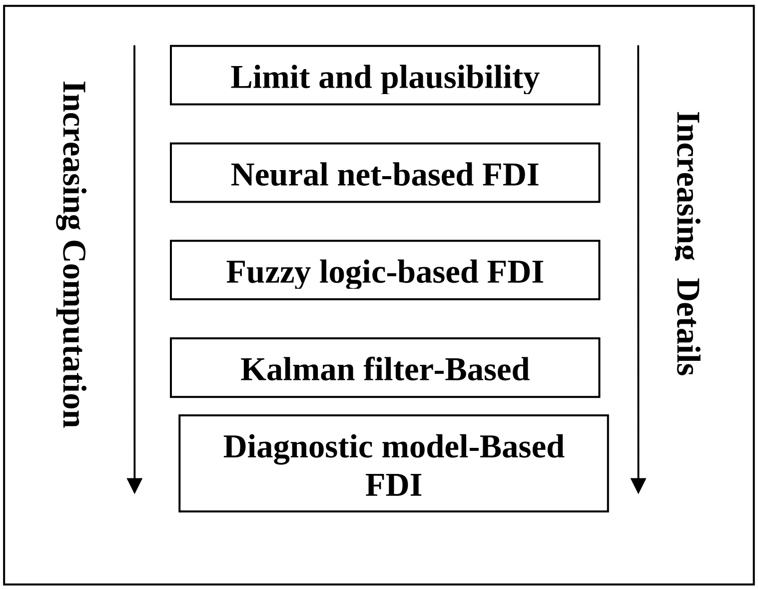


Figure 3 Sequential Execution of Tasks

In this thesis, a sequential integration of both model-free and model-based approaches is employed so as to meet the requirements for a quick and reliable fault detection and isolation scheme. The tasks of our fault diagnosis scheme (Fig.3) are executed with an increasing precision accompanied by a gradual unfolding-in-time diagnostic picture that reveals the presence of an incipient fault [159]. The multistage scheme of Fig. 1 involves three model-free stages followed by two model-based ones. This scheme starts with limit checks and a plausibility analysis, then a neural network stage for a quick fault classification, followed by a fuzzy logic block to unravel the real cause(s) of the fault. The last two model-based stages are used for the twofold purpose of capturing any incipient fault(s) that the first three model-free stages may have missed out as well as confirming their diagnosis. These two final stages involve a Kalman filter for fault detection and a diagnostic parameter identification scheme for fault isolation, thus completing the overall diagnostic picture.

Chapter 4

Proposed Fault Detection Scheme using Model-Free Analysis

Overview

This chapter introduces the proposed fault detection scheme using the model-free analysis. It gives a comprehensive view about the famous techniques of Fuzzy and Neural Networks and describes the Proposed Scheme of Model-Free Fault Detection Isolation comprising of Limits and Plausibility Checks, Fuzzy Logic Based FDI, and Neural Networks Based FDI.

4.1 Introductory Background

Model-based fault diagnosis uses mathematical models derived from physical principles and is based on parameter estimation or state estimation techniques. Unfortunately, comprehensive and robust models for complex processes are difficult or impossible to develop and validate in both normal and fault modes. The modeling and fault diagnosis development is usually extremely time-consuming, and often the models are limited and can not characterize the process with all possible faults. Therefore, the Model-Based approach is generally limited to those processes that are well understood or lend themselves to a combination of physical/empirical modeling approaches.

Recently, soft computing methods, integrating quantitative and qualitative modeling information, have been developed to improve FD reasoning capabilities.

In order to develop accurate and robust process control, model-based modern control methods and efficient adaptive and learning techniques are required. The adoption of effective fault diagnosis techniques is becoming crucial to ensure higher levels of safety and reliability in automated plants and autonomous systems. Process control is an efficient means of improving the operation of a process, the productivity of the plant, and the quality of the products. In process engineering, even a small improvement in the operation of the process can have great economic and environmental influences. Control problems in the industry are dominated by nonlinear and time-varying behaviour, many sensors that measure all kinds of variables and many loops and interaction among the control loops. The extraction of (fuzzy) information out of raw data is very important and it can potentially save time for industrial applications. Fuzzy control can be based on human experience and can mimic actions of human operators.

During recent years, the developments in these fields have introduced new tools for use in control engineering: neuro-fuzzy systems, guided random search techniques, predictive control, model reference control, etc. In process engineering, these new tools have found applications in non-linear process modeling and control, plant optimization, monitoring, scheduling, etc. The application area of control engineering methods can be extended also to systems beyond the realm of traditional process engineering. Modern techniques for control system design, including robust design for stochastic and nonlinear systems as well as intelligent control, are expected to lead to an increase in quality and productivity of manufacturing processes. The manufacturing and industrial sectors of economy are increasingly called to produce higher throughput and better quality while operating their processes at maximum yield. As manufacturing facilities become more complex and highly sophisticated, the quality of the production phase has become more crucial. The manufacture of such typical products as textiles and fibres, aircraft, automobiles, appliances, etc, involves a large number of complex processes, most of which are characterized by highly nonlinear dynamics comprising a variety of physical phenomena in the temporal and spatial domains. It is not surprising, therefore, that these processes are not well understood and their operation is "tuned" by experience rather than through the application of scientific principles. Machine breakdowns are common, thus limiting the uptime in critical situations. Failure conditions are difficult and, in certain cases, almost impossible to identify and localize in a timely manner. Scheduled maintenance practices tend to reduce machine lifetime and increase downtime, resulting in loss of productivity. Recent advances in instrumentation, telecommunications and computing are making available to manufacturing companies new sensors and sensing strategies, plant-wide networking and information technologies that are helping to improve substantially the production cycle.

In many practical situations, uncertainty in the process can affect the performance of the system significantly, no matter how the uncertainty is described (vagueness or ambiguity). This realization provides the motivation for a possible fuzzy logic approach to FDI. This has the ability to directly describe the potential failure modes in the parameters while handling a class of nonlinear systems.

Various approaches utilizing measurement data (model-free) have been proposed for fault diagnosis in process operations. Please refer to section 1.1.2 for a detailed literature survey about model-free fault diagnosis approaches.

4.2 Proposed Model-Free Scheme for Fault Detection and Isolation

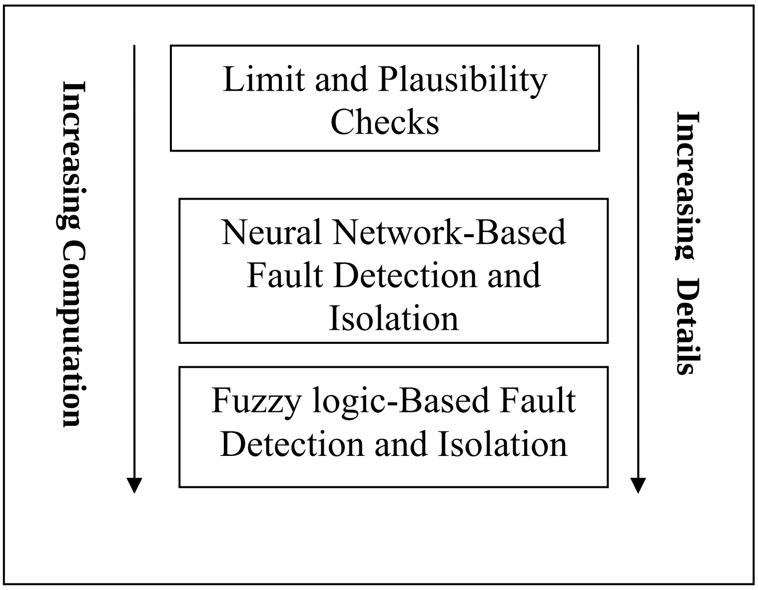


Figure 4 Proposed Model-Free Scheme for FDI

The model-free approach includes limit checks, plausibility analysis, steady-state values and settling time of the measurements, estimation of power spectral density and coherence function between data from the fault-free case and the measured data. A fault may be detected quickly by analyzing whether a measured value has violated its upper or lower limit (limit checks), and if so, whether it is meaningful when compared with the other measurements (plausibility), and whether its settling time or steady-state value or coherence spectral content is different from those arising from the fault-free case. An onset of a fault may also be detected by a change in the profile of the measured value of the sensor signal.

The fault diagnosis scheme can be carried out using neural network and fuzzy techniques or a combination of both. A pictorial view of the model-free scheme is shown in the Figure 4. Driven by the coherence spectral data, the neural network is used to capture the degree of the mismatch between the dynamics of the possibly fault-bearing system and the fault-free one. As such, the neural network provides a quick and accurate model-free fault classification scheme. However, it lacks transparency and cannot be used to unravel and point out to the root causes of the fault. Such a deficiency is then remedied by the use of a fuzzy logic scheme. Unlike the neural network, this scheme is driven by the steady-state value of the residual, defined by the difference between the measured and the fault-free sensor output, and captures the degree of the mismatch in the steady-state behavior of the possibly fault-bearing system and the fault-free one. By the very nature of its inner inferential workings necessary for the information extraction through the processing of intricate rules, the fuzzy logic scheme is slower but more transparent than the neural network, and provides not only a confirmation of any fault classification arrived at by the neural network but also a backtracking process aided by the fuzzy rules, leading to the root causes of the fault. The fuzzy if-and-then rules are derived and used by a process closely resembling the use of the residuals produced by a bank of static Kalman filters for fault classification. The synergistic value of this combination will no doubt provide a powerful fault diagnosis scheme.

The model-free approach is geared neither for the estimation of the fault magnitude nor for the diagnosis of incipient faults. The model-based approach is then used to provide these two important features that will complete the overall diagnostic picture of the fault.

4.2.1 Fuzzy Logic-Based FDI

Fuzzy logic is a form of multi-valued logic derived from fuzzy set theory to deal with reasoning that is approximate rather than precise. In binary sets with *binary logic*, in contrast to fuzzy logic named also *crisp logic*, the variables may have a membership value of only 0 or 1. In fuzzy set theory with fuzzy logic, the set membership values can range (inclusively) between 0 and 1. Similarly, in fuzzy logic the degree of truth of a statement can range between 0 and 1 and is not constrained to the two truth values {true (1), false (0)} as in classic predicate logic. And when *linguistic variables* are used, these degrees may be managed by specific functions.

The fuzzy diagnostic system takes features as inputs, and then it outputs any indications that a failure mode may have occurred in the plant. The fuzzy logic system structure is composed of four blocks: fuzzification, the fuzzy inference engine, the fuzzy rule base, and defuzzification, as shown in the Figure 5.

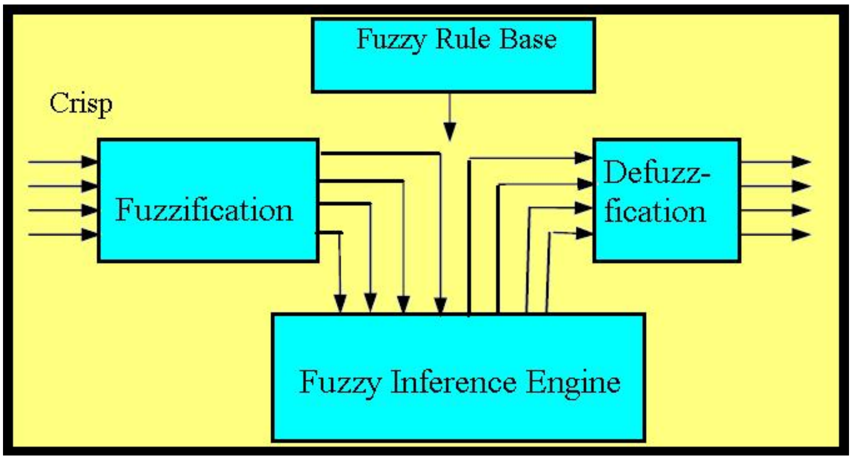


Figure 5 Fuzzy Logic Diagnostic System

The fuzzification block converts features to degrees of membership in a linguistic label set such as low, high, etc. The fuzzy rule base is constructed from symptoms that indicate a potential failure mode.

An example of Fuzzy Reasoning: Fuzzy Set Theory defines Fuzzy Operators on Fuzzy Sets. The problem in applying this is that the appropriate

Fuzzy Operator may not be known. For this reason, Fuzzy logic usually uses IF-THEN rules, or constructs that are equivalent, such as fuzzy associative matrices.

Rules are usually expressed in the form:

IF *variable* IS *property* THEN *action*

For example, an extremely simple temperature regulator that uses a fan might look like this:

IF temperature IS very cold THEN stop fan

IF temperature IS cold THEN turn down fan

IF temperature IS normal THEN maintain level

IF temperature IS hot THEN speed up fan.

Notice there is no "ELSE". All of the rules are evaluated, because the temperature might be "cold" and "normal" at the same time to different degrees [165].

4.2.2 Fuzzy Logic (FL) Fault Classifier

In this work, the system whose fault is to be diagnosed is expressed in the form of a sensor network shown in the Fig. 6. Consider a cascade-feedback combination of N systems, G_i with sensor outputs, y_i as shown in the figure. Let sensor gains and bias be k_{si} and v_i respectively.

The fuzzy fault diagnosis scheme uses the steady-state values of the sensor outputs y_i and the input u which are denoted by y_i^{ss} and u^{ss} respectively. A change in the gain k_{si} or in the steady-state gain of the transfer function G_i, denoted by g_i, is indicative of a fault in the i-th sensor and i-th subsystem respectively. In the case of sensors, the presence or absence of a fault is defined below (using a fuzzy-like definition of $k_{si} = 1$ for a no-fault case and $k_{si} \leq 1$ for a faulty case):

$$\begin{cases} |k_{si} - 1| \leq \sigma_{th} & no\ fault \\ |k_{si} - 1| > \sigma_{th} & a\ fault \end{cases} \tag{1}$$

where σ_{th} is some pre-specified threshold value. Similarly, in the case of subsystems, fault-free and faulty cases are defined as follows:

$$\begin{cases} |g_i - g_i^0| \leq \sigma_{th} & no\ fault \\ |g_i - g_i^0| > \sigma_{th} & a\ fault \end{cases} \tag{2}$$

where g_i^0 is the fault-free steady-state gain of the subsystem G_i .

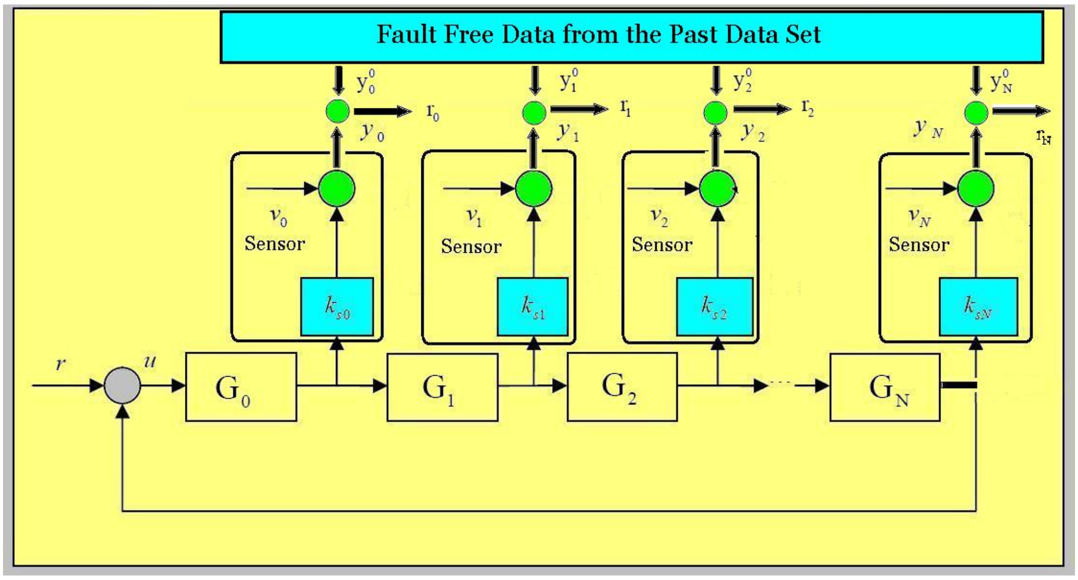


Figure 6 Sensor Network

An expression for the steady-state *p-th* sensor output, y_p^{ss} , in terms of the steady-state input u^{ss} is:

$$y_p^{ss} = \prod_{i=0}^{p} g_i k_{sp} u^{ss} \tag{3}$$

The expression of the fault-free sensor output is given by:

$$y_p^{ss0} = g_{0p}^0 u^{ss} \tag{4}$$

where $g_{0p}^0 = \prod_{i=0}^{p} g_i^0$ is the fault-free steady-state equivalent gain for the *p* cascaded subsystems from G_0 to G_p since the sensor gains are all unity for a fault-free case. Assuming that the noise term v_p is subsumed in the fuzzy membership function the deviation in the in steady-state output is:

$$\Delta y_p^{ss} = \left(\prod_{i=0}^{p} g_i k_{sp} - g_{0p}^0 \right) u^{ss} \tag{5}$$

Let us now define two fuzzy sets, namely *Z* (*zero)* and *NZ* (*non-zero)*. For simplicity, we will consider the case of a single fault, i.e. when only one device can be faulty at any given time. In this case, the fuzzy rules may take the following form:

Rule 1: If $\Delta y_i^{ss} \in NZ$, then there is a fault in the subsystem G_i or the sensor k_{si}.

Rule II: If $\Delta y_i^{ss} \in Z$, then there is no fault in the subsystem G_i or in the sensor k_{si}.

Rule III: If $\Delta y_i^{ss} \in Z$ and $\Delta y_{i+1}^{ss} \in NZ$ then there is a fault in the subsystem G_{i+1} or the sensor $k_{s(i+1)}$.

Rule IV: If $\Delta y_i^{ss} \in NZ$ and $\Delta y_{i+1}^{ss} \in Z$ then there is a fault in the sensor k_{si}.

Proposition 1:

If there are (N+1) subsystems, $G_i \ i = 0,1,2,...N$***, in a sensor network, then a single fault occurring in any one of the (N+1) subsystems at any one time can be detected and isolated.***

If multiple faults occurring simultaneously in the sensors, that is k_{si}, $i \in [i_1 \quad i_2]$ ***are faulty where*** i_1 ***and*** $i_2 \geq i_1$ ***are integers, then all the multiple sensor faults can be detected and isolated if the sensor following the last faulty sensor namely*** $k_{s(i_2+1)}$ ***is not faulty.***

None of the subsystems are faulty.

<u>Proof:</u> Define binary variables α_i and β_i to indicate a fault in a subsystem and a sensor respectively according to the following definitions:

$$\alpha_i = \begin{cases} 0 & if\ \Delta y_i^{ss} \in Z \\ 1 & if\ \Delta y_i^{ss} \in NZ \end{cases} \qquad (6)$$

$$\beta_i = \begin{cases} 0 & if\ \Delta y_i^{ss} \in Z \\ 1 & if\ \Delta y_i^{ss} \in NZ \end{cases} \qquad (7)$$

Consider a cascade connection of $(p+1)$ subsystems formed of $G_i, i = 0,1,2,..p$ with the corresponding gains g_i. Then the steady-state output y_p^{ss} of the last subsystem G_p is given by:

$$y_p^{ss} = \prod_{i=0}^{p} g_i k_{sp} u^{ss} \qquad (8)$$

In view of the assumptions that (a) only a single fault can occur at the sensor output, and (b) y_p^{ss} is a cascade connection of g_i *such that* $0 \leq i \leq p$, a fault in a subsystem G_p will affect the p-th sensor output y_p^{ss} but also all subsequent sensor outputs y_j^{ss} *such that* $j = p+1, p+2, ...N$. That is, if there is a fault in G_p then the change in the steady-state value of the sensor output Δy_i^{ss} will belong to either the fuzzy set Z or the fuzzy set NZ as given below:

$$\Delta y_i^{ss} \in \begin{cases} Z & i < p \\ NZ & p \le i \le N \end{cases} \tag{9}$$

In this case, the subsystem binary fault indicator α_i becomes:

$$\alpha_i = \begin{cases} 1 & i \ge p \\ 0 \text{ or } X & i < p \end{cases} \tag{10}$$

However, as the sensors, unlike the subsystems, are not connected in cascade, the sensor binary fault indicator β_i becomes:

$$\beta_i = \begin{cases} 1 & i = p \\ X & i \ne p \end{cases} \tag{11}$$

where X is a don't care value (0 or 1).

Let us express the results of our analysis of the sensor outputs in a tabular form as shown below in Table 1 . The columns of Table 1 give the binary values of α_i and β_i while the rows indicate the status of the fault. For example, when i-th sensor output y_i^{ss} is analyzed, and if $\Delta y_i^{ss} \in NZ$, then a 1 will be entered in the i-th row and in the two columns corresponding to α_i and β_i. The rest of the elements of the i-th row will take on the (don't care) binary value X, as we cannot decide on their fault status. For a clearer display of the structure of Table 1, we will restrict the number of subsystems to N=5 . We will consider first a subsystem fault and then a fault in a sensor. The table will have (N+1) =6 rows and 2(N+1)=12 columns.

Case 1: A single fault in a subsystem G_2. Table 1 below gives the status of both the subsystems' and sensors' faults:

The columnwise intersection of the (N+1) elements of the above binary matrix amounts to a columnwise logical ANDing of their binary values. This yields the following vector *d*:

d = [0 0 1 1 1 1 0 0 1 1 1 1]

From an analysis of the first half of the vector *d*, it can be deduced that there is a fault in G_2 or G_3 or G_4 or G_5 or k_{s2} or k_{p3} or k_{s4} or k_{s5}. Exploiting (1) the fact that, because of the cascade connection of the subsystems in the sensor network, a single fault occurring in any particular subsystem G_p will always propagate from this subsystem onwards and (2) the assumption that only one fault can occur at any given time instant, we can then reach the irrevocable conclusion that G_2 is the only faulty subsystem in the entire sensor network.

Table 1 Subsystem and sensor fault indicators when a subsystem G_2 is faulty

	status of a fault in the subsystem						status of a fault in the sensor					
	α_0	α_1	α_2	α_3	α_4	α_5	β_0	β_1	β_2	β_3	β_4	β_5
$\Delta y_0^{ss} \in Z$	0	X	X	X	X	X	0	X	X	X	X	X
$\Delta y_1^{ss} \in Z$	X	0	X	X	X	X	X	0	X	X	X	X
$\Delta y_2^{ss} \in NZ$	X	X	1	X	X	X	X	X	1	X	X	X
$\Delta y_3^{ss} \in NZ$	X	X	X	1	X	X	X	X	X	1	X	X

$\Delta y_4^{ss} \in NZ$	X	X	X	X	1	X	X	X	X	X	1	X
$\Delta y_5^{ss} \in N$	X	X	X	X	X	1	X	X	X	X	X	1

Table 2 Subsystem and sensor fault indicators when sensors k_{s2} , k_{s3} and k_{s4} are faulty.

	status of a fault in the subsystem						status of a fault in the sensor					
	α_0	α_1	α_2	α_3	α_4	α_5	β_0	β_1	β_2	β_3	β_4	β_5
$\Delta y_0^{ss} \in Z$	0	X	X	X	X	X	0	X	X	X	X	X
$\Delta y_1^{ss} \in Z$	X	0	X	X	X	X	X	0	X	X	X	X
$\Delta y_2^{ss} \in NZ$	X	X	1	X	X	X	X	X	1	X	X	A
$\Delta y_3^{ss} \in NZ$	X	X	X	1	X	X	X	X	X	1	X	X
$\Delta y_4^{ss} \in NZ$	X	X	X	X	1	X	X	X	X	X	1	X
$\Delta y_5^{ss} \in N$	X	X	X	X	X	1	X	X	X	X	X	0

***Case 2*:** Multiple faults in sensors: k_{s2} , k_{s3} and k_{s4} are faulty. Table 2 gives the status of faults:

The column-wise intersection of the (N+1) elements of the above binary matrix amounts to a columnwise logical ANDing of their binary values. This yields the following vector *d*:

d = [0 0 1 1 1 1 0 0 1 1 1 0]

From an analysis of the first half of the vector *d*, it can be deduced that there is a fault in G_2 or G_3 or G_4 or G_5 or k_{s2} or k_{p3} or k_{s4} or k_{s5}. Exploiting (1) the fact that, because of the cascade connection of the subsystems in the sensor network, a single fault occurring in any particular subsystem G_p will always propagate from this subsystem onwards and (2) the assumption that only one fault can occur at any given time instant, we can then reach the irrevocable conclusion that the sensors k_{s2}, k_{s3} and k_{s4} are faulty in the entire sensor network. For detection of multiple faults in sensors, the last sensor k_{s5} cannot be faulty.

4.2.3 Neural Network Based FDI

Artificial neural networks are made up of interconnecting artificial neurons (programming constructs that mimic the properties of biological neurons). Artificial neural networks may be used either to gain an understanding of biological neural networks, or for solving artificial intelligence problems without necessarily creating a model of a real biological system. The real biological nervous system is highly complex and includes some features that may seem superfluous based on an understanding of artificial networks. A simple neuron is shown in the figure (See Fig. 6) [165].

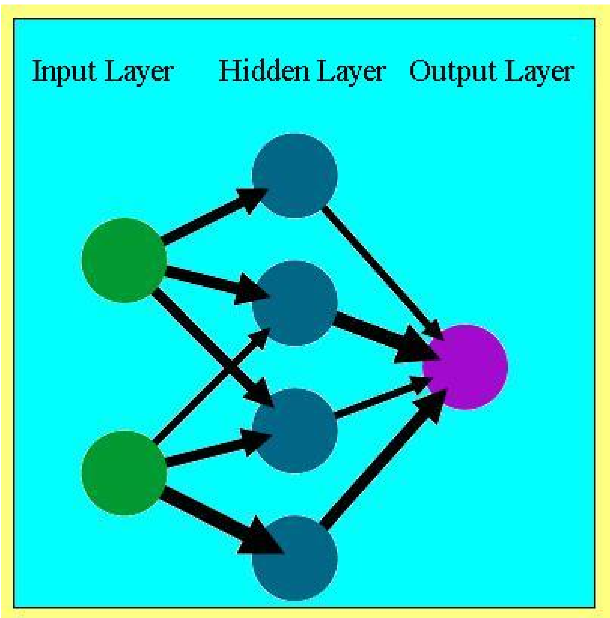


Figure 7 Simple Neural Network

An artificial neuron model is inspired from our understanding of biological nervous systems in its simplest form.

$$x = f\left(\sum_{i=1}^{n} w_i u_i + b_0\right) \qquad (12)$$

where $\{w_i\}$ are the synaptic weights, b_0 is the bias or a firing threshold, and x is the output of the neuron, and $f(.)$ is an activation function. The function $f(.)$ is some nonlinear such as threshold, Gaussian, sigmoid, or other. An artificial neuron model is completely described by the weights, the bias, and the nonlinear function [165].

An artificial neural network is an interconnection of a number of artificial neurons. A biological neuron is viewed as an elementary unit for information processing.

For convenience of representation, the bias term is included in as one of the weights with a unit input.

$$x = f\left(\sum_{i=0}^{n} w_i u_i\right) \quad where \quad b_0 = w_0 u_0, \quad u_0 = 1$$

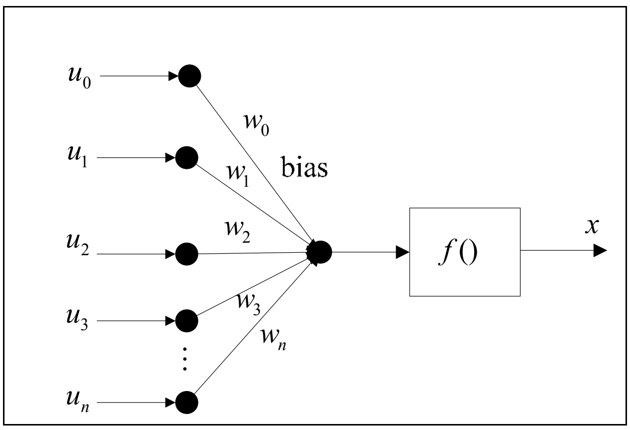


Figure 8 Artificial neuron model with the bias terms replaced by an Input and a Synaptic Weight

Generically, the multilayer feedforward ANN is restricted with the following structure:

- Input layer
- One hidden layer
- Output layer

The neurons forming the hidden layer have nonlinear activation functions, whereas the neurons forming the output layer has linear activation function, $f(.)=1$ [165].

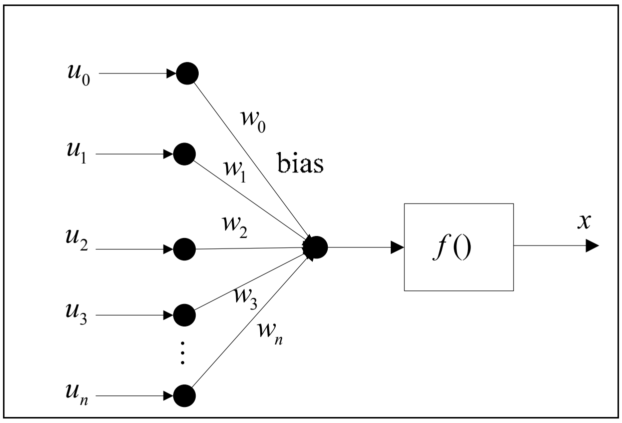


Figure 9 Neuron forming the hidden layer has a nonlinear activation function

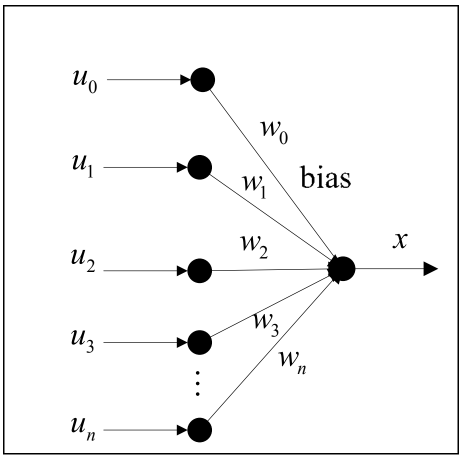


Figure 10 Neuron forming the output layer has a linear activation function $f(.)=1$

A typical multilayer feedforward ANN has one input, one hidden and one output layer. A detailed picture of the multilayer feedforward ANN is shown below:

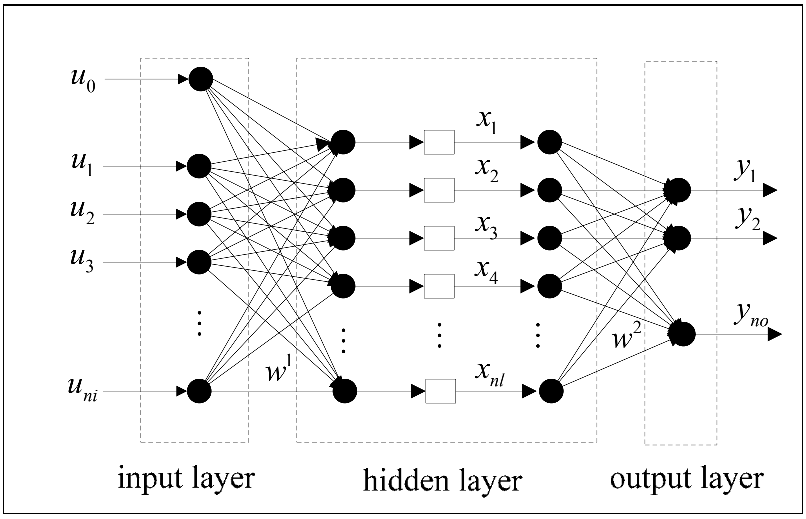


Figure 11 Neural Network with its input, hidden and output layers

The weights of the ANN are determined by presenting the ANN with a number of known input-output pairs as training set,

$$T = \{u^q, d^q, q = 1, 2, ..., N\}$$

$$u^q = \begin{bmatrix} u_1^q & u_2^q & u_3^q & \dots & u_{n_i}^q \end{bmatrix}^T$$

$$x^q = \begin{bmatrix} x_1^q & x_2^q & x_3^q & \dots & x_{n_l}^q \end{bmatrix}^T$$

$$y^q = \begin{bmatrix} y_1^q & y_2^q & y_3^q & \dots & y_{n_o}^q \end{bmatrix}^T$$

$$d^q = \begin{bmatrix} d_1^q & d_2^q & d_3^q & \dots & d_{n_o}^q \end{bmatrix}^T$$

where u^q, x^q, y^q and d^q are respectively the input, hidden layer output, the output of the ANN and the desired output of the ANN. The number of neurons in the input, hidden and the output layers are respectively n_i , n_l and n_o. The weights $w^1 = \{w_{ij}^1\}$ are associated with the n_i input nodes and nl hidden layer nodes and the weights $w^2 = \{w_{ij}^2\}$ are associated with nl hidden layer nodes, and the no output nodes [165].

Mathematical model

Let us relate the hidden layer states x and the input u in terms of the weights, $w^1 = \{w_{ij}^1\}$

$$x_i = f\left(\sum_{j=0}^{ni} w_{ij}^1 u_j\right), \quad i = 1, 2, .., nl$$

Expressing in matrix form we get

$$x = f\left(W^1 u\right)$$

where W^1 is a (*nl x ni*) matrix formed of the weights $w^1 = \{w_{ij}^1\}$ and f(.) is activation function.

Let us relate the output y to the hidden layer states x in terms of the weights, $w^2 = \{w_{ij}^2\}$

$$y_i = f_i\left(z_i\right), \quad z_i = \sum_{j=1}^{nl} w_{ij}^2 x_j \quad i = 1, 2, ...no$$

Expressing in matrix form we get

$$y = W^2 x$$

where W^2 is a (*nm x nl*) matrix formed of the weights $w^2 = \{w_{ij}^2\}$ and *f*(.) is activation function [165].

Artificial Neural Networks (ANNs) have been intensively studied during the last two decades and successfully applied to dynamic system modelling and fault diagnosis [3][4][5][6][7][8][9]. Neural networks stand for an interesting and valuable alternative to the classical methods, because they can deal with very complex situations which are not sufficiently defined for deterministic algorithms. They are especially useful when there is no mathematical model of a process being considered. In such situations, the classical approaches, such as observers or parameter estimation methods, cannot be applied. Neural networks provide excellent mathematical tools for dealing with nonlinear problems [10][11][12]. They have an important property whereby any nonlinear function can be approximated with an arbitrary accuracy using a neural network with a suitable architecture and weight parameters. For continuous mappings, one hidden layer based ANN is sufficient, but in other cases two hidden layers should be implemented. ANNs are parallel data-processing tools capable of learning functional dependencies of the data. This feature is extremely useful for solving various pattern recognition problems. Another attractive property is the self-learning ability. A neural network can extract the system features from historical training data by using a learning algorithm, requiring little or no *a priori* knowledge about the process. This makes ANNs nonlinear modelling tools of a great flexibility. Neural networks are also robust with respect to incorrect or missing data. Protective relaying based on ANNs is not affected by a change in the system operating conditions. Neural networks also have high computation rates, substantial input error tolerance, and adaptive capability. These features allow neural networks to be applied effectively to the modeling and identification of complex nonlinear dynamic processes and fault diagnosis [13] [14].

4.2.4 Artificial Neural Network (ANN) Fault Classifier

A fault in the sensor k_{si} and/or in a subsystem G_i can also be diagnosed by using a ANN. The inputs to the ANN are the spectrum of the coherence between the fault-free and measured sensor outputs.

$$c\left(y_i^0(j\omega), y_i(j\omega)\right) = \frac{\left|y_i^0(j\omega)\, y_i^*(j\omega)\right|^2}{\left|y_i^0(j\omega)\right|^2 \left|y_i(j\omega)\right|^2} \qquad (13)$$

where ω is the frequency in rad/sec, and $c\left(y_i^0(j\omega)y(j\omega)\right)$ is the coherence spectrum and the output of the ANN will the fault type, i.e. either a fault in a subsystem or in a sensor.

If there is no fault at all, then $c\left(y_i^0(j\omega)y(j\omega)\right)=1$ for all frequencies. If the measured and fault-free outputs are incoherent with each other at some frequencies, then the coherence spectrum will be less than 1 at those frequencies. Unlike the case of a FL classifier, the dynamic characteristics of the sensor outputs are employed in the fault diagnosis. Hence it can detect and classify a fault when the dynamic behavior deviates from that of a fault-free case.

Chapter 5

Proposed Fault Diagnosis (FDI) Scheme using Model-Based Analysis

Overview

This chapter introduces the proposed fault detection scheme using the model-Based analysis. This approach consists essentially of the following three components: model order selection to pick out the most appropriate model from a class of candidate models, Kalman filtering for fault detection, and fault isolation using diagnostic model.

Model-based fault diagnosis means to perform fault diagnosis by using models. An important question is how to use the models to construct a diagnosis system. To develop a theory for this is important as it is one of the main components in the Sequential Integration Approach.

In this chapter, the objective detection and diagnosis of certain faults is tackled by using the model-based approach. A fault can be defined as an unexpected deviation of at least one characteristic property or parameters of the system from the acceptable, usual or standard condition. Three types of faults can be encountered in a system given by the three parts in which a system can be split.

- *Actuator Faults*, which can be viewed as malfunction of the equipment that actuates the system, e.g. a malfunction in a solenoid valve.
- *System Dynamics Faults/Leakage Faults (or Component Faults),* which occur when some changes in the system make the dynamic relation invalid, e.g. leak in a tank in a two-tank system.
- *Sensor Faults,* which can be viewed as serious measurements variations.

This chapter starts, in Section 5.1, by giving an introductory background and a general motivation to the field of fault diagnosis. In Section 5.2, some fundamental definitions are reviewed. Then Section 5.3 contains an overview to some present approaches to fault diagnosis. Finally, Section 5.4 summarizes the thesis and explains the main contributions.

5.1 Introductory Background

From a general perspective, including for example process control, medical and technical applications, fault diagnosis can be explained as follows. For a process there are observed variables or behavior, for which there is knowledge of what is expected or normal. The task of fault diagnosis is to use the observations and the knowledge in order to generate a *diagnosis statement*, i.e. to decide whether there is a fault or not and also to identify the fault. Thus the basic problems in the area of fault diagnosis are: the procedure for generating the diagnosis statement, the parameters or behavior that are relevant to study, and the way to derive and represent the knowledge of what is expected or normal.

This thesis focuses on diagnosis of technical systems, and typical faults considered are for example leakage faults, sensor faults and actuator faults. The observations are mainly flow signal, height signal and output signal obtained from the sensors, but can also be observations made by a human, such as level of noise and vibrations. The knowledge of what is expected or normal is derived from commanded inputs together with models of the system. The term *model based* fault diagnosis refers to the fact that the knowledge of what is expected or normal is represented in an explicit model of the system.

Model based diagnosis of technical systems has gained much industrial interest lately. The reason is that it has possibilities to improve for example safety, environment protection, machine protection, availability, and repairability.

Some important applications that have been discussed in the literature are:

- Nearly all subsystems of aircrafts, e.g. aircraft control system, navigation system, and engines.
- Emission control systems in automotive vehicles.
- Nuclear power plants.
- Chemical plants
- Gas turbines
- Industrial robots
- Electrical motors

Manual diagnosis of technical systems has been performed as long as technical systems have existed, but automatic diagnosis started to appear first when computers became available. In the beginning of the 70's, the first research reports on model based diagnosis were published. Some of the earliest investigations were on chemical plants and aerospace applications. The research on model based diagnosis has since then been intensified during both the 80s and the 90s. Today, this is still an expansive research area with many unsolved questions. Some references to books in the area are Patton, Frank and Clark, 1989; Basseville and Nikiforov, 1993; Gertler, 1998; Chen and Patton, 1999.

Up to now, numerous methods for doing diagnosis have been published, but many approaches are more ad hoc than systematic. It is fair to say that few general theories exist, and a complete understanding of the relations between different methods has been missing. This is reflected in the fact that few books exist and that no general terminology has yet been widely accepted. However, the importance of diagnosis is unquestioned. This can be exemplified by the computerized management systems for automotive engines. For these systems, as much as 50% of the software is dedicated to diagnosis, and the other 50% is for other purposes such as control.

5.1.1 Traditional vs. Model Based Diagnosis

Traditionally diagnosis has been performed mainly by limit checking. For example, when a sensor signal level leaves its normal range, an alarm is generated. The normal range is predefined by using thresholds. This normal range can be dependent on the operating conditions. For example, in an aircraft, the thresholds for different operating points, defined by altitude and speed, can be stored in a table. This use of thresholds as functions of some other variables can actually be viewed as a kind of model based diagnosis.

Another traditional approach is duplication (or triplication or more) of hardware. This is usually called *hardware redundancy* and the typical example is to use redundant sensors. At least three problems are associated with the use of hardware redundancy: hardware is expensive, it requires space, and adds weight to the system. In addition, extra components increase the complexity of the system which in turn may introduce extra diagnostic requirements.

5.1.2 Model-Based Fault Diagnosis

Increased usage of explicit models in fault diagnosis has a large potential to have the following advantages:

- Higher diagnosis performance can be obtained, for example smaller and also more varied faults can be detected, and the detection time is shorter.

- Diagnosis can be performed over a larger operating range.
- Diagnosis can be performed passively without disturbing the operation of the process.
- Increased possibilities to perform isolation.
- Disturbances can be compensated for, which implies that high diagnosis performance can be obtained in spite of the presence of disturbances.
- Reliance on hardware redundancy can be reduced, which means that cost and weight can be reduced.

The model can be of any type, from logic based models to differential equations. Depending on the type of model, different approaches to model based diagnosis can be used, for example statistical approaches, AI-based approaches, or approaches within the framework of control theory. It is sometimes believed that model based diagnosis is very complex. This is not true since, for example, traditional limit checking is also a kind of model based diagnosis.

The disadvantage of model based diagnosis is quite naturally the need for a reliable model and possibly a more complex design procedure. In the actual design of a model based diagnosis system, it is likely that the major part of the work is spent on building the model. This model can however be reused, e.g. in control design. Someone may argue that a disadvantage of increasing the usage of models is that more computing power is needed to perform the diagnosis.

However, this conclusion is not fair. Actually, for the same level of performance it can be the case that an increasingly used model is *less* computationally intensive than the traditional approach.

The accuracy of the model is usually the major limiting factor of the performance of a model based diagnosis system. Compared to the area of model based control, the quality of the model is much more important in diagnosis.

The reason for is that the feedback, used in closed-loop control, tends to be forgiving against model errors. Diagnosis should be compared to open-loop control since no feedback is involved. All model errors propagate through the diagnosis system and degrade the diagnosis performance.

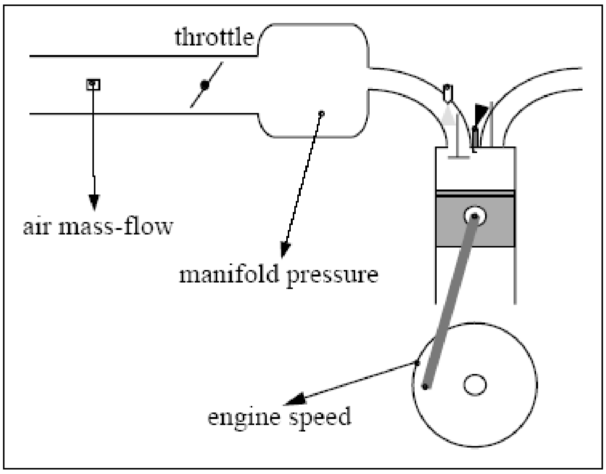


Figure 12 Spark Ignited Combustion Engine

Following is an example of a successful industrial application of model-based diagnosis.

Example:

Consider Figure 12, containing an illustration to the principles of a spark-ignited combustion engine. The air enters at the left side, passes the throttle and the manifold, and finally enters the cylinders. The engine in the figure has three sensors measuring the physical variables air mass-flow, manifold pressure, and engine speed.

The air flow $\dot{m}$ into the cylinders can be modeled as a function of manifold pressure p and engine speed n, i.e. $\dot{m} = g(p\ , n)$. The physics behind the function g is involved and it is therefore usually modeled by a black-box model. In engine management systems, one common solution is to represent the function g as a lookup-table. By using this lookup-table, an estimation of the air mass-flow can be obtained. When the measured air mass-flow significantly differs from the estimation, it can be concluded that a fault must be present somewhere in the engine. The fault can, for example, be that one of the three sensors is faulty or that a leakage has occurred somewhere between the air mass-flow sensor and the cylinder. This is an example of model based diagnosis that is commonly used in the production of cars today.

5.2 Main theme behind Model-Based Technique

In model-based diagnosis (de Kleer and Williams, 1987), a library of models can be used to perform the diagnosis of a system (Struss, 2007). The core objective of the model-based diagnosis is to find candidate diagnoses that explain observations (de Kleer, 2006). In system identification the aim is to find the state of the system (whether there is a fault or not), but the model-based diagnosis objective is to diagnose the system, i.e. find the problem (Balakrishnan and Honavar, 1998). However, diagnosis goes beyond the task of finding the problem. As written in Struss (2007), *"Diagnosis is only relevant if it supports a decision [...] "*. Thus, the final aim of diagnosis is not only to identify the problem, but also to find a possible remedy. Examples of remedy are replacement of components, reconfiguration, etc. Examples of applications of model-based diagnosis are automotive industry (Struss and Price, 2004), autonomous mobile robots (Steinbauer and Wotawa, 2005) and software debugging (Kob and Wotawa, 2004).

A possible way to perform fault diagnosis is through parameter estimation (Isermann, 1993). Using static and dynamic process models as well as measurements, relationships and redundancies are used to detect faults.

System identification is a complex process and can, for example, be supported by qualitative reasoning-based approaches (Trave-Massuyes et al., 2003). However, in the latter, a single model is iteratively updated. The paper by Addanki et al. (1991) introduces the concept of graphs of models. However, in their work, models are generated manually and they work with only one model at a time. Traditionally, system identification is treated as an optimization problem in which the difference between model predictions and measurements is minimized. Values of model parameters for which model responses best match measured data are determined by this approach[1]. This approach is not reliable because different types of modeling and measurement errors are present (Banan et al., 1994; Sanayei et al., 1997; Catbas et al., 2007). Moreover, they can compensate each other such that the global minimum indicates models that are far away from predictions of the model representing the correct state of the system (Robert-Nicoud et al., 2005c).

[1] Approach: In conventional system identification, a suitable model is identified by matching measurement data with model predictions. Model calibration involves minimization of the difference between predictions and measurement data through identification of good values of model parameters. This strategy is based on the assumption that the model that best fits the observations is the most reliable model. This assumption is flawed; there are several factors that could cause the best fit to be the wrong model. Errors influence the reliability of system identification. Various types of errors may compensate each other such that bad model predictions match measured values. The following definitions are used in this description: measurement error (e_{meas}) is the difference between real and measured quantities in a single measurement. Modeling error (e_{mod}) is the difference between the prediction of a given model and that of the model that accurately represents the real behavior. Modeling errors have three principal sources e_1, e_2 and e_3 (Raphael and Smith, 2003b). Source e_1 is the error due to the discrepancy between the behavior of the mathematical model and that of the real structure. Source e_2 is introduced during the numerical computation of the solution of the partial differential equations representing the mathematical

model. Source e_3 is the error due to the assumptions that are made during the simulation of the numerical model. Typical assumptions are related to the choice of boundary conditions and model parameters such as material properties, for example E and I. All these errors as well as the abductive aspect of the system identification task justify the use of a multiple model approach since many models may have equal validity under these conditions.

Therefore, instead of optimizing one model, a set of candidate models is identified, such that their prediction errors lie below a certain threshold value. A model is defined in Robert-Nicoud et al. (2005c) as a distinct set of values for a set of parameters. The threshold is computed using an estimate of the upper bound of errors due to modeling assumptions (e_{mod}) as well as measurements (e_{meas}). The set of candidate models is iteratively filtered using subsequent measurements for system identification. This approach could generate either a unique model for the structure or a set of models which are equally capable of representing the structure. This depends on parameters chosen for the identification problem and errors. Modeling assumptions define the parameters for the identification problem. The set of model parameters may consist of quantities such as elastic modulus, connection stiffness and moment of inertia. Each set of values for the model parameters corresponds to a model of the structure. An objective function is used to evaluate the quality of candidate models. An exemplary objective function E is defined as follows:

$$E = \left\{ \begin{matrix} \varepsilon \text{ if } \varepsilon > \tau \\ 0 \text{ if } \varepsilon \leq \tau \end{matrix} \right. \text{with } \varepsilon = \sqrt{\Sigma(\mathrm{m_i} - \gamma_\mathrm{i})^2} \Bigg\}$$

where ε is the error which is calculated as the difference between predictions γ_i and measurements m_i. The threshold value τ is evaluated from measurement and modeling errors in the identification process. The set of models that have E = 0 form the set of candidate models for the structure. An important aspect of the methodology is the use of a stochastic global search and optimization algorithm for the selection of a population of candidate models whose predictions are close to measurements (Robert-Nicoud et al., 2000). Mathematical optimization techniques that make use of derivatives and sensitivity equations are not used because search is performed among sets of model classes that contain varying numbers of parameters and multiple local minima have been observed in the search space.

5.3 Present Approaches to Model-Based Fault Diagnosis

This section is included for two reasons. The first is to point out some problems with present approaches to fault diagnosis. The second is to give newcomers to the field of fault diagnosis a short background to some of the approaches present in literature.

By reading recent books (Gertler, 1998; Chen and Patton, 1999) about fault diagnosis of technical processes, or survey papers (Patton, 1994; Gertler, 1991; Frank, 1993; Isermann, 1993), one can come to the conclusion that the two most common systematic approaches to fault diagnosis are either *residual view* or *parameter estimation*. Below these two approaches are presented shortly.

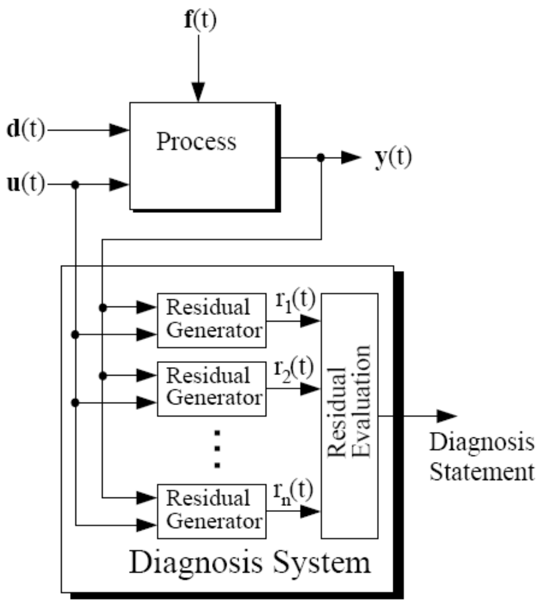


Figure 13 Diagnosis system based on the "residual view"

5.3.1 The "Residual View"

With this approach, faults are modeled by signals $f(t)$. Central is the residual $r(t)$ which is a scalar or vector signal that is 0 or small in the fault free case, i.e. $f(t)=0$, and is $\neq$ 0 when a fault occurs, i.e. $f(t) \neq 0$. The diagnosis system is then separated into two parts: *residual generation* and *residual evaluation*.

This view of how to design a diagnosis system is well established among fault diagnosis researchers. This is emphasized by the following quotation from the most recent book in the field (Chen and Patton, 1999):

"Chow and Willsky (1984) first defined the model-based FDI as a two-stage process: (1) residual generation, (2) decision making (including residual evaluation). This two-stage process is accepted as a standard procedure for model-based FDI nowadays."

A large part of all fault-diagnosis research has been to find methods to design residual generators. Of this large part, most results are concerned with linear systems.

A limitation with this approach to fault diagnosis is that faults are modeled as signals. This is very general and might therefore seem to be a good solution. However, the generality of this fault model is actually its drawback. Many faults can be modeled by less general models, and we will see in this thesis that to facilitate isolation this is necessary in many situations [164].

5.3.2 Parameter Estimation

The other main approach to model-based fault-diagnosis is to model faults as deviations in constant parameters. To illustrate the concept, consider a system with a model $\mathrm{M}(\theta)$, where θ is a parameter having the nominal (i.e. fault-free) value θ_0. By using general parameter estimation techniques, an estimate $\hat{\theta}$ can be formed and then compared to θ_0. If $\hat{\theta}$ deviates too much from θ_0, then the conclusion is that a fault has occurred [164].

The most severe limitation with this approach is its quite restricted way of modeling faults. To model many realistic faults, more general fault models must be used.

Another limitation is that when the number of diagnosed faults grows, the parameter vector θ grows in dimension. This is a serious problem because the computations needed to calculate $\hat{\theta}$ can become quite difficult [164].

5.4 Proposed Model-Based Approach Scheme

Most physical systems are structurally complex and nonlinear with their modeling, analysis and design requiring sophisticated tools which may be neither available nor possible because of the mathematical intractability involved in their study. This in effect is what gives the prime motivation for the use of the model-free approach based on neural networks or fuzzy logic. However, the fact that a large number of these complex and nonlinear systems can be linearized around some operating points gives us access to a vast and rich repository of linear analysis and design tools for these systems, while preserving most of their dynamical features of interest. This is the key reason why the study of the model-based approach to fault diagnosis and isolation is undertaken here. This approach consists essentially of the following three components: model order selection to pick out the most appropriate model from a class of candidate models, Kalman filtering for fault detection, and fault isolation using diagnostic model. These are described next.

5.4.1 Model-Order Selection Criteria

5.4.1.1 Introduction

Science is the systematic study of the universe—through observation and experiment—in the pursuit of knowledge that allows us to generalize. Although considered bad form in the current climate of political correctness (Lind 1998, 2004; Ellis 2004; Browne 2006; Sewell 2007b), the ability to generalize is a distilled version of what science is all about. Given some data, there will always be an infinite number of models or hypotheses that fit the data equally well and without making further assumptions there is no reason to prefer one model or hypothesis over another. Therefore, one is forced to make assumptions that provide an inductive bias.

Model selection is the task of choosing a model with the correct inductive bias, which in practice means selecting parameters in an attempt to create a model of optimal complexity for the given (finite) data. For a good book on model selection, see Burnham and Anderson (2002). Many methods of model selection employ some form of parsimony: that is, if they fit the data equally well, they prefer a simpler model (see Zellner, Keuzenkamp and McAleer (2001)). For example, Occam's razor advises us that, when competing theories have equal predictive power, one should choose the theory that introduces the fewest assumptions. For more details on Occam's razor, see Hoffmann, Minkin and Carpenter (1997) and the references therein. Bayesians use probability to choose among hypotheses, (hypothesis|data, background information) (Howson and Urbach 1989). Popperians choose among hypotheses that are equally consistent with the observations by preferring those which are more falsifiable (Popper 1934, 1959). Likelihoodists understand the plausibility of a hypothesis in terms of evidential support and they consider (data|hypothesis) (Edwards 1992). Minimum description length (MDL) (Rissanen 1978) is a technique from algorithmic information theory which dictates that the best hypothesis for a given set of data is the one that leads to the largest compression of the data. One seeks to minimize the sum of the length, in bits, of an effective description of the model and the length, in bits, of an effective description of the data when encoded with the help of the model. Classical Neyman–Pearson hypothesis testing considers (data|null hypothesis) but this method is flawed (see Atkins and Jarrett (1979); Minka (1998); Gabor (2004); Sewell (2008b)). The Akaike information criterion (AIC) (Akaike 1973) proposes that one should trade off the complexity of the model with its goodness of fit to the sample data. The model with the lowest AIC should be preferred. AIC = −2 log L + 2k, where log L is the maximum log-likelihood and *k* is the number of parameters.

5.4.1.2 New Model-Order Selection Criteria

In recent years there has been a strong emphasis on model evaluation criteria. This is recognized as one of the important area in model identification [13, 14]. It consists of choosing a criterion and using it to select the best approximating model among a class of competing models for a given data set.

Criteria based on statistical decision theory require an *a priori* knowledge of the probability distribution function (PDF) of the residuals. In practical systems, it may not be possible to estimate this PDF and hence a Gaussian PDF is generally assumed [13]. In many cases, the application of these criteria, assuming a Gaussian PDF, may not always give the correct model order as the estimated model may still contain some artifacts such as noise nonlinearities, sampling rate selection and pole-zero cancellation effects.

In [12], a two-stage identification scheme is proposed. First a high-order model is employed to capture both the system dynamics and any artifacts (from noise or other sources). Then in the second stage, these artifacts are removed by using a frequency-weighted estimation scheme. A different two-stage approach is proposed herein. It quickly verifies the presence of any artifacts directly from the estimated model. In the first stage, a conventional model structure selection criterion such as AIC, BIC or MDL is employed to select an initial model order. In the second stage, the presence of any artifacts in the selected model order is then quickly verified. If any artifacts are present in the estimated model order, the selected order is then discarded and a lower order chosen instead. The extraneous poles may arise due to a number of causes including an improper selection of the sampling rate, the presence of colored noise, pole-zero cancellation when the selected order is large, and oscillations due to nonlinearity in the physical system, e.g. hysteresis in the valve [15].

The models of most physical systems are continuous while the identified ones are discrete. We shall now derive a necessary and sufficient condition which guarantees that the discrete poles belong to the right half of the z-plane.

Proposition 2: *Given that the process to be identified is of a lowpass nature, then the poles of its discrete-time equivalent will lie in the right half of the z-plane* (Z^+) *if and only if the sampling frequency is more than twice the Nyquist rate.*

Proof: The poles of the discrete-time model (i.e. λ_d) are related to those of the continuous-time model (i.e. λ_c) by :

$$\lambda_d = e^{\lambda_c T_s} \qquad (14)$$

where $T_s = 1/f_s$ is the sampling period and f_s is the corresponding sampling frequency. Given that:

$$\lambda_c = \alpha_c + j\omega_c \quad where \quad \omega_c = 2\pi f_c \qquad (15)$$

where f_c is the frequency of oscillation, we now get:

$$\lambda_d = \cos(2\pi f_c / f_s) e^{\alpha_c T_s} + j\sin(2\pi f_c / f_s) e^{\alpha_c T_s} \quad (16)$$

From the above equation, we deduce the following:

$$\lambda_d \in Z^+ \Leftrightarrow \cos(2\pi f_c / f_s) e^{\alpha_c T_s} \geq 0 \Leftrightarrow -\frac{1}{4} \leq \frac{f_c}{f_s} \leq \frac{1}{4} \quad (17)$$

Or equivalently,

$$\lambda_d \in \begin{cases} Z^+ & -\frac{f_s}{4} \leq f_c \leq \frac{f_s}{4} \\ Z^- & else \end{cases} \quad (18)$$

This shows that the discrete-time poles lie in the right half plane if the sampling rate (f_s) is more than twice the Nyquist rate ($2f_c$). Otherwise, the poles lie in the left-half of the z-plane Z^- if the inequality $-f_s/4 \leq f_c \leq f_s/4$ is violated.

5.4.2 Fault Detection using Kalman Filter

The Kalman filter is designed for the normal fault-free operation. The fault-free model of the system, which is obtained from the system identification process described in the previous section, is given by:

$$x(k+1) = A_0 x(k) + B_0 u(k-d) + w(k)$$
$$y(k) = C_0 x(k) + u(k) \quad (19)$$

where $y(k)$ is the output, e.g., the height of the water in a tank, (A_0, B_0, C_0) are obtained from the discretized model of (A, B, C) for the ideal fault-free case $w(k)$

and $v(k)$ are zero-mean white plant and measurement noise signals respectively, with covariances:

$$Q = E\left[w(k)w^T(k)\right], \text{ and } R = E\left[v(k)v^T(k)\right] \qquad (20)$$

The plant noise $w(k)$ is a mathematical artifice introduced to account for the uncertainty in the *a-priori* knowledge of the plant model. The larger the covariance Q is, the less accurate the model (A_0, B_0, C_0) is and vice versa.

The Kalman filter is given by:

$$\begin{aligned} \hat{x}(k+1) &= A_0\hat{x}(k) + B_0u(k-d) + K_0\left(y(k) - C_0\hat{x}(k)\right) \\ e(k) &= y(k) - C_0\hat{x}(k) \end{aligned} \qquad (21)$$

where d is the delay and e (k) the residual.

The system model has a pure time delay which is incorporated in the Kalman filter formulation. The Kalman filter estimates the states by fusing the information provided by the measurement $y(k)$ and the *a priori* information contained in the model (A_0, B_0, C_0). This fusion is based on the *a priori* information of the plant and the measurement noise covariances *Q* and *R* respectively. When *Q* is small, implying that the model is accurate, the state estimate is obtained by weighting the plant model more than the measurement one. The Kalman gain K_0 will then be small. On the other hand, when *R* is small implying that the measurement model is accurate, the state estimate is then obtained by weighting the measurement model more than the plant one. The Kalman gain, K_0, will then be large in this case.

The larger K_0 is, the faster the response of the filter will be and the larger the variance of the estimation error becomes. Thus, there is a trade-off between a fast filter response and a small covariance of the residual. An adaptive on-line scheme is employed to tweak the *a- priori* choice of the covariance matrices so that an acceptable trade-off between the Kalman filter performance and the covariance of the residual is reached.

A statistical decision-theoretic approach was used to decide between two hypotheses [1-5]. If the absolute mean of the residual is less than a specified threshold value thr , then a fault is asserted. The threshold value is calculated from the pre-specified false alarm rate, and the variance of the residual.

5.4.3 Fault Isolation using Diagnostic Model

There are two approaches to the estimation of the diagnostic parameters, namely the system identification and the diagnostic model approach. In [6], a fault is isolated by using a two-stage approach. First, the feature vector θ is estimated. Then, the diagnostic parameter γ is estimated from the identified θ using the *a priori* known relation, $\theta = \varphi(\gamma)$ where φ is some nonlinear function [6].

The second approach [7], employed in this paper, is based on a diagnostic model, which directly relates the diagnostic parameters to the input and output. The diagnositic parameters are identified offline by performing a number of experiments. The diagnostic model relating the reference input r the diagnostic parameter γ and the residual $e(k)$, is given by:

$$e(k) = y(k) - y^0(k) = \sum_{i=1}^{q} \psi^T(k-1)\theta_i^{(1)}\Delta\gamma_i + v(k) \quad (22)$$

where, $\Delta\gamma_i = \gamma - \gamma_i^0$ is the perturbation in γ; $y^0(k)$ and γ_i^0 are the fault-free (nominal) output and parameter, respectively, $\theta_i^{(1)} = \dfrac{\delta\theta}{\delta\gamma_i}$, and ψ is the data vector formed of the past outputs and past reference inputs. The gradient $\theta_i^{(1)}$ is estimated by performing a number of offline experiments which consist of perturbing the diagnostic parameters, one at a time. The input-output data from all the perturbed parameter experiments is then used to identify the gradients $\theta_i^{(1)}$. The hypothesis H_i corresponding to the perturbation of the i^th^ diagnostic parameter is given by:

$$H_i: \quad e(k) = \psi^T(k-1)\theta_i^{(1)}\Delta\gamma_i + v(k) \qquad (23)$$

If *v* (*k*) is assumed to be a zero-mean Gaussian random variable, then the Bayes strategy suggests that the *most likely* hypothesis H_j is the one whose index satisfies

$$j = \arg\min_i \left\{ \left\| \ e(k) - \psi^T(k-1)\theta_i^{(1)}\Delta\gamma_i \ \right\|^2 \right\} \qquad (24)$$

Since the size of the fault, denoted by the perturbation $\Delta\gamma_j(k)$, is unknown, a composite hypothesis testing scheme is used in which we substitute the unknown $\Delta\gamma_j(k)$ by its least-squares estimate. Substituting the estimate of $\Delta\gamma_j(k)$ and simplifying the fault isolation strategy yields:

$$j = \arg\max_i \left\{ \cos^2\varphi_i \right\} \quad where \quad \cos\varphi_i = \frac{\langle e, \psi^T\theta_i^{(1)} \rangle}{\|e\| \left\| \psi^T\theta_i^{(1)} \right\|} \qquad (25)$$

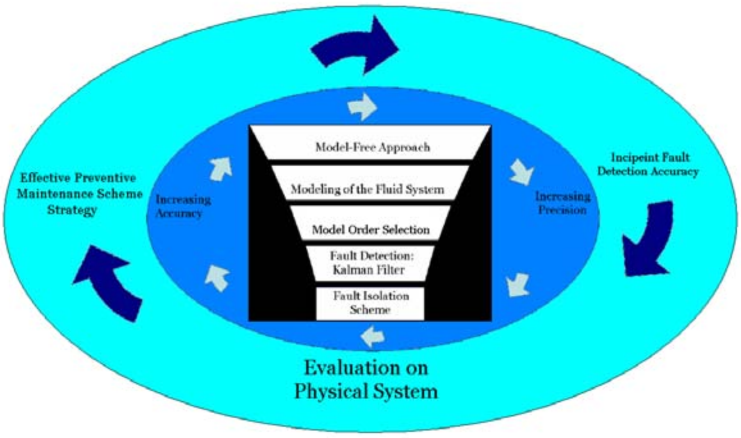
That is, γ_j is asserted to be faulty if the measured residual $e(k)$ and its hypothesized residual estimate $\psi^T(k-1)\theta_j^{(1)}$ are maximally aligned. A measure of isolability of faults in γ_i and γ_j is defined by the cosine of the angle between $\theta_i^{(1)}$ and $\theta_j^{(1)}$ denoted by $\cos\theta_{ij}^{(1)}$. The smaller $\cos\theta_{ij}^{(1)}$ is, the larger the isolability gets.

Chapter 6

Evaluation of the Proposed Sequential Integration Approach for Fault Diagnosis

Overview

This chapter demonstrates the evaluation of the proposed sequential integration approach for fault diagnosis. The evaluation of both the Proposed Model-Free and Model-Based Approaches has been done on a benchmark laboratory-scale process control system.



An evaluation of the proposed scheme for fault diagnosis was performed on a benchmark laboratory-scale process control system using National Instruments LABVIEW as shown below in Fig 4.

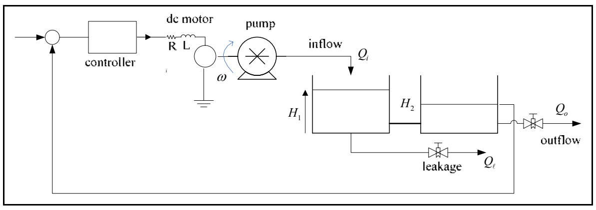


Figure 14 Two-tank fluid System

Fault diagnosis in fluid systems has become increasingly important in recent years from the points of view of economy, safety, pollution, and conservation of scarce resources [9, 16]. The proposed scheme is used to detect and isolate a fault by a sequential integration of model-free and model-based approaches. The sensor fault is simulated by including a gain term γ_s in the measured output $y_m = \gamma_s y$. The actuator fault is simulated by including a gain term γ_a in the control input to the actuator, namely the motor-pump sub-system, $u_a = \gamma_a u$. Finally the leakage fault is simulated by controlling the amount by which the drain valve in the tank is opened. This is equivalent to introducing a gain term γ_l to the height of the tank, $q_l = \gamma_l h$. A fault-free case corresponds to $\gamma_s = \gamma_a = 1$ and $\gamma_l = 0$ as is easily shown in Fig. 5.

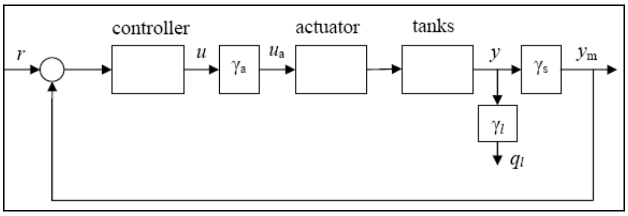


Figure 15 Fault Simulation in the Fluid System

6.1 Fault diagnosis using Model-Free Approach

A sequential integration of an artificial neural network (ANN) and a fuzzy logic (FL) scheme is employed here to isolate faults.

***Fuzzy-logic approach**:* The features were chosen to be the steady-state values of the control input, u^{ss}, measured flow flw^{ss} and height h^{ss} values and their corresponding ones in the fault-free case, namely u^0, flw^0 and h^0. Fuzzy logic rules given in Section 3.2 are used to detect and isolate faults. The Adaptive Network Fuzzy Inference System (ANFIS), based on Sugeno's method, is employed here to implement the fuzzy classifier [17].

The elements of the output vector $d = [d_1 \quad d_2 \quad d_3 \quad d_4]$ are the decisions defined by:

$$d_i = \begin{cases} 1 & i^{th} \ device \ is \ faulty \\ 0 & else \end{cases} \tag{26}$$

where the values of the index i =1, 2, 3, and 4 correspond to the actuator, flow sensor, level sensor and the leakage fault. The fuzzy rules using Sugeno's method [17] are:

If $\Delta flw^{ss} \in non-zero$ and $\Delta h^{ss} \in non-zero$ then $d_1 = 1$

If $\Delta flw^{ss} \in non-zero$ and $\Delta h^{ss} \in zero$ then $d_2 = 1$

If $\Delta flw^{ss} \in zero$ and $\Delta h^{ss} \in non-zero$ then $d_3 = 1$

If $\Delta u^{ss} \in small-positive$ then $d_4 = 1$

If $\Delta flw^{ss} \in zero$ and $\Delta h^{ss} \in zero$ then $d_i = 0 \ for \ all \ i$

where $zero$, $small-positive$ and $non-zero$ are fuzzy sets.

A neural network, driven by the coherence spectrum between the measured height and the fault-free height, produces classes of four possible faults, namely a fault in the actuator, the level sensor, the flow sensor, and a leakage fault.

The fuzzy approach is then integrated sequentially with the neural network-based fault classification to complete the required fault isolation scheme. The ANN-based classifier precedes the FL-based one, with the former providing a fast classification and the latter both confirming this classification and providing a way of pinpointing the real root cause(s) for the occurrence of this fault.

The fault magnitude is qualitatively estimated by the changes in the settling time $\Delta t_s = t_{ss}^0 - t_{ss}$ whereas its onset is indicated by the location of the change in the height profile.

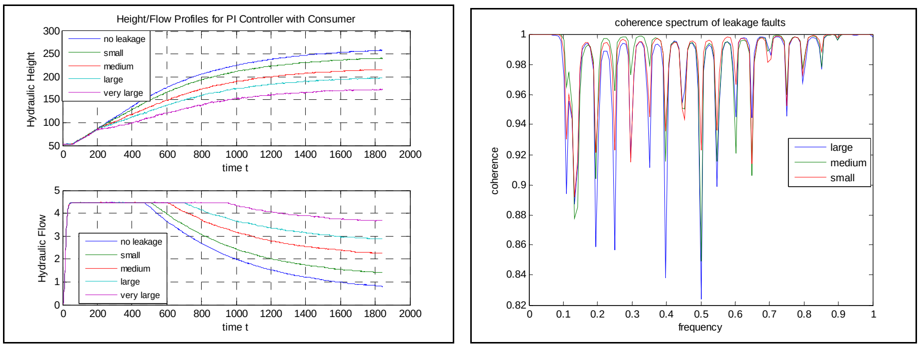


Figure 16 Height/Flow Profile and Coherence Spectrum under Leakage faults

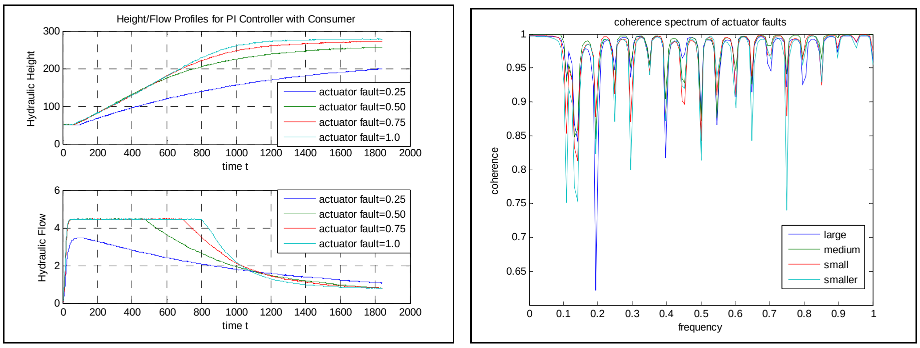


Figure 17 Height/Flow Profile and Coherence Spectrum under Actuator Faults

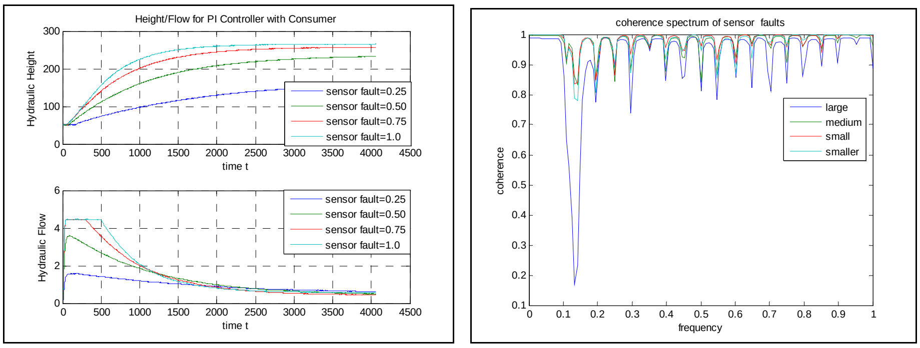


Figure 18 Height/Flow Profile and Coherence Spectrum under Sensor Faults

Figs 16-18 give the profiles of the flow and height and the coherence spectra, whereas Fig. 19 shows height profiles in the presence of leakages of different magnitudes occurring when the fluid level system is operated in a closed-loop or open-loop configuration. For the open-loop case, one can readily deduce both the onset and amount of the leakage from the height/flow profile. The leakage flow has five sections corresponding to the following five degrees of no-leakage, small, medium, large and very large leakage. However, by virtue of its control design objective, the closed-loop PI controller will hide any fault that may occur in the system and hence will make it difficult to detect it.

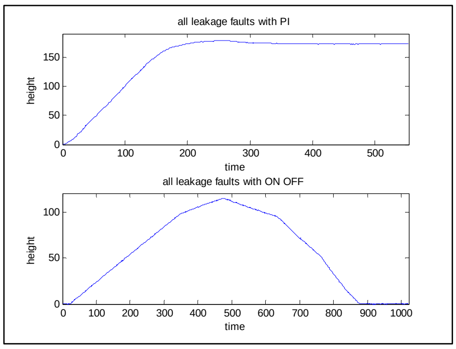


Figure 19 Plots of Various leakages with closed and open loop control

6.2 Model of the Fluid System

A benchmark model of a cascade connection of a dc motor and a pump relating the input to the motor u and the flow Q_i is a first-order time-delay system expressed by:

$$\dot{Q}_i = -a_m Q_i + b_m \phi(u) \qquad (27)$$

where a_m and b_m are the parameters of the motor-pump system and $\phi(u)$ is a dead-band and saturation type of nonlinearity. The Proportional and Integral (PI) controller is given by:

$$\begin{aligned} \dot{x}_3 &= e = r - h_2 \\ u &= k_p e + k_I x_3 \end{aligned} \qquad (28)$$

where k_p and k_I are gains and r is the reference input.

With the inclusion of the leakage, the liquid level system is modeled by [16]:

$$A_1 \frac{dH_1}{dt} = Q_i - C_{12}\varphi(H_1 - H_2) - C_\ell \varphi(H_1)$$

$$A_2 \frac{dH_2}{dt} = C_{12}\varphi(H_1 - H_2) - C_0 \varphi(H_2) \qquad (29)$$

where $\varphi(.) = sign(.)\sqrt{2g(.)}$, $Q_\ell = C_\ell \varphi(H_1)$ is the leakage flow rate,

$Q_0 = C_0 \varphi(H_2)$ is the output flow rate,

H_1 is the height of the liquid in tank 1,

H_2 is the height of the liquid in tank 2,

A_1 and A_2 are the cross-sectional areas of the 2 tanks,

g=980 $cm/\sec^2$ is the gravitational constant,

C_{12} and C_o are the discharge coefficient of the inter-tank and output valves, respectively.

The linearized model of the entire system formed by the motor, pump, and the tanks is given by:

$$\dot{x} = Ax + Br \quad y = Cx \qquad (30)$$

$$x=\begin{bmatrix}h_1\\h_2\\x_3\\q_i\end{bmatrix}, A=\begin{bmatrix}-a_1-\alpha & a_1 & 0 & b_1\\ a_2 & -a_2-\beta & 0 & 0\\ -1 & 0 & 0 & 0\\ -b_m k_p & 0 & b_m k_I & -a_m\end{bmatrix},$$

$$B=\begin{bmatrix}0 & 0 & 1 & b_m k_p\end{bmatrix}^T, C=[1 \quad 0 \quad 0 \quad 0] \qquad (31)$$

Where q_i ,q_ℓ , q_0 , h_1 a nd h_2 are respectively the increments in Q_i ,Q_ℓ ,Q_o , H_1^0 and H_2^0, whereas a_1, a_2, α and β are parameters associated with linearization, α is associated with leakage, $q_\ell = \alpha h_1$, and β is the output flow rate, $q_o = \beta h_2$.

6.3 Evaluation of the Proposed Model-Order Selection Criteria

The model of the physical system, based on the physical laws, is given in Section 6.2. It is of a fourth order for a PI controller, of a third order for a P controller, and of a second order for an on-off controller.

Various orders of the model of the fluid system ranging from 1 to 6 were initially selected, and for each order the corresponding model was identified using a least-squares method. The following quantities were computed:

- Poles of the identified model
- The loss function, $J=\frac{1}{N}\sum_{k=0}^{N-1}[y(k)-\hat{y}(k)]^2$ where $\hat{y}$ is the estimate of the system output y, and N is the number of data samples.
- AIC measure

6.3.1 Case 1: PI controller

Table 3 below shows that, for the selected model orders 1 to 3, all the poles are on the right half plane. The vital question arises now is how to select the appropriate model order out of these three order values. The loss function and the AIC measures are used only as initial guidelines. We selected a second order as it was found to be the smallest order which yielded an acceptable performance, as shown in Fig. 10.

Table 3 PI Controller: Poles of the identified model for different '1','2',..,'6' selected orders

order 1	order 2	order 3	order 4	order 5	order 6
0.9850	0.9847	0.9847	0.9845	0.9845	0.9845
	0.0712	0.0340 ± j0.5910	0.0915	0.3301 ± j0.5549	0.4726
			-0.0189 ± j0.6197	-0.3058 ± j0.5938	0.2429 ± j0.6847
					-0.4400 ± j0.5933

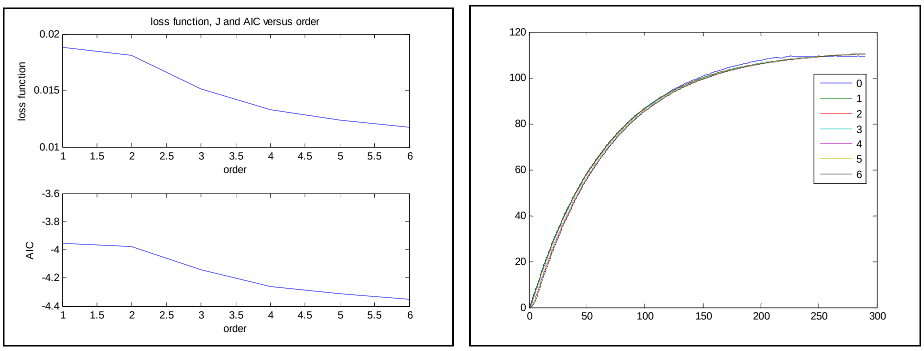


Figure 20 Loss Function and AIC versus Order and Step Response Versus Order. Order '0' = actual data and '1','2',..,'6'= selected orders.

6.3.2 Case 2: On-off controller

Table 4 below shows that, for the selected model order of 1, all the poles are on the right-half plane. Figure 11 shows that the performance related to model order 1 is acceptable.

Table 4 On-Off Controller: Poles of the identified model for different '1','2',..,'6' selected orders

order 1	order 2	order 3	order 4	order 5	order 6
0.9996	0.9996	0.9996	0.9996	0.9996	0.9996
	-0.1839	-0.1230 ± j0.5358	0.1147	0.4118 ± j0.6558	0.5400
			-0.1731 ± j0.5491	-0.5280 ± j0.6373	0.2898 ± j0.7560
					-0.6182 ± j0.6456

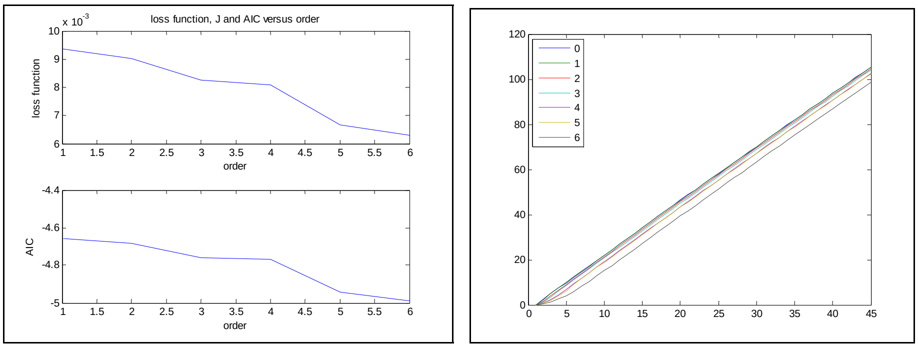


Figure 21 Loss Function and AIC versus Order and Step Response versus Order. Order '0' = actual data and '1','2',..,'6'= selected orders.

The proposed model order selection criterion (Section IV), which was thoroughly tested through extensive simulation runs and an evaluation on the physical system, was found to be very reliable. It has the ability to capture the input-output dynamic behavior, and not the dynamics resulting from the effect of noise and other artifacts.

6.4 Evaluation of Fault Detection Using Kalman Filter

First the fault-free model of the system is identified using a recursive least-squares identification scheme. An acceptable model order was then selected using the proposed model order selection criterion.

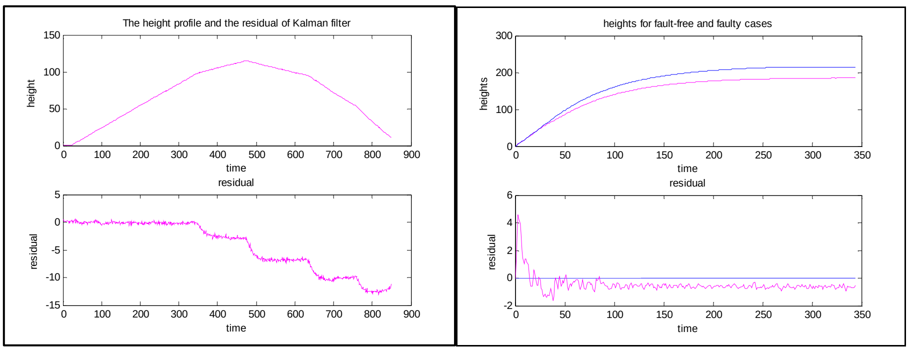


Figure 22 Kalman Filter results for an ON-OFF and PI Controller used for Flow and Height under various leakage magnitudes

The identified model is essentially a second-order system with a delay although the theoretical model is of a fourth order. Such a discrepancy is due to the inability of the identification scheme to capture the system's fast dynamics, especially in low-SNR scenarios. Using the fault-free model together with the covariance of the measurement noise R, and the plant noise covariance, Q, the Kalman filter model was finally derived. As it is difficult to obtain an estimate of the plant covariance Q a number of experiments were performed under different plant scenarios to tune the Kalman gain K_0.

$$\hat{x}(k+1) = A_0\hat{x}(k) + B_0u(k-d) + K_0\left(y(k) - C_0\hat{x}(k)\right)$$

$$e(k) = y(k) - C_0\hat{x}(k) \qquad (32)$$

The Kalman filter was evaluated under different fault scenarios for an on-off controller and a PI controller, as shown in Fig.12.

Comments: The model of the fluid system is nonlinear, complex and stochastic. A simplified linearized model which contains only the dominant poles (as it was difficult to identify the fast dynamics) was used in the design of the Kalman filter. Results from the evaluation on the physical system show that the Kalman filter is robust in modelling uncertainties including nonlinearities and neglected fast dynamics, while at the same time being sensitive to incipient faults.

6.5 Evaluation of Fault Isolation Scheme

The diagnostic model of the fluid system becomes:

$$e(k) = \sum_{i=1}^{3} \psi^T(k-1)\theta_i^{(1)}\Delta\gamma_i + v(k) \qquad (33)$$

where $\psi^T(k) = [-y(k-1) \quad -y(k-2) \quad u(k-1) \quad u(k-2)]$,

$\gamma_1 = \gamma_\ell$, $\gamma_2 = \gamma_a$ and $\gamma_3 = \gamma_s$

A number of experiments were performed offline by varying the diagnostic parameters, one at a time. Each of the γ parameters was varied one at a time, spanning three different values of 0.25, 0.5 and 0.75 of their maximum. From these experiments, the gradients $\theta_i^{(1)}$ were estimated:

The measure of isolability $\cos\theta_{ij}^{(1)}$ is given below:

$$\cos\theta_{12}^{(1)} = 0.8560, \cos\theta_{13}^{(1)} = 0.8379, \cos\theta_{23}^{(1)} = 0.7757 \quad (34)$$

Using the composite hypothesis testing scheme, a fault is isolated by determining which hypothesis gives the maximal alignment between the estimated and measured residuals. The results of the isolation scheme are encouraging.

Comments: The physical two-tank fluid system is nonlinear with a dead-band nonlinearity and fast dynamics. The identified model order is different from that of the model derived from the physical laws. The conventional two-stage identification scheme [6], based on first identifying θ and then deriving $\gamma = \varphi^{-1}(\theta)$, is not possible because of the irreversible collapse of the model structure from a fourth-order one to a second-order one. This difficulty is avoided by adopting the scheme proposed in [7] wherein a number of offline experiments on the physical system are performed by varying the diagnostic parameters so as to capture the influence of the diagnostic parameters on the input-output behaviour reliably. This in essence mirrors the use of a neural network in approximating a nonlinear map.

Chapter 7

Performance Analysis of the Proposed Scheme

Overview

This chapter gives the performance analysis of the proposed fault diagnosis scheme. The fault signatures are analyzed from the step response parameters, namely settling time and steady-state error, power spectral density profiles, model-order selection criteria, Kalman filter graphs for a complete fault diagnostic scheme.

Index Diagram:

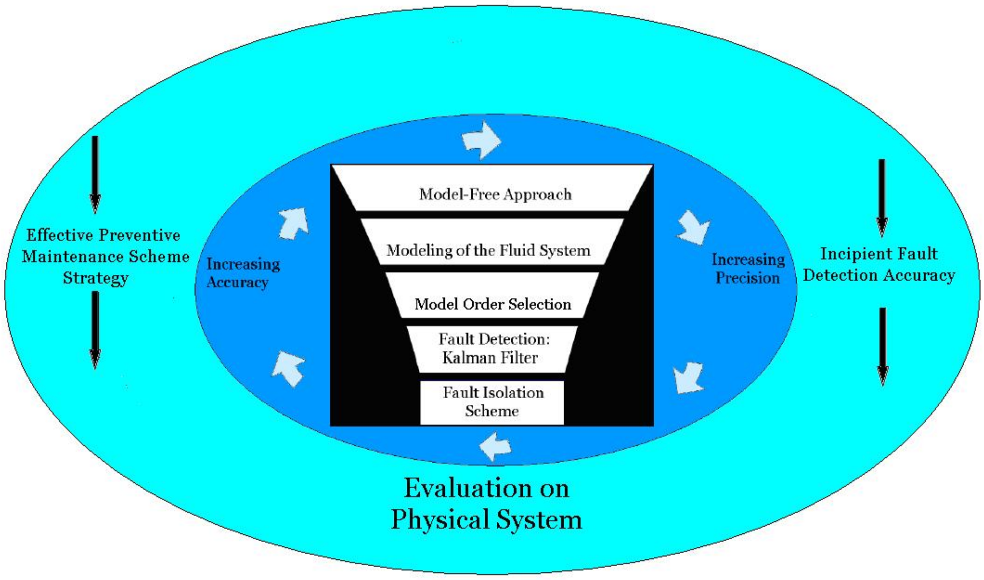


Figure 23 Evaluation on Physical System

7.1 Height Profiles of Various Faults

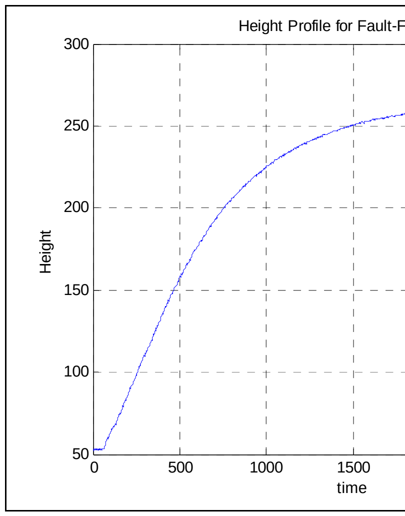
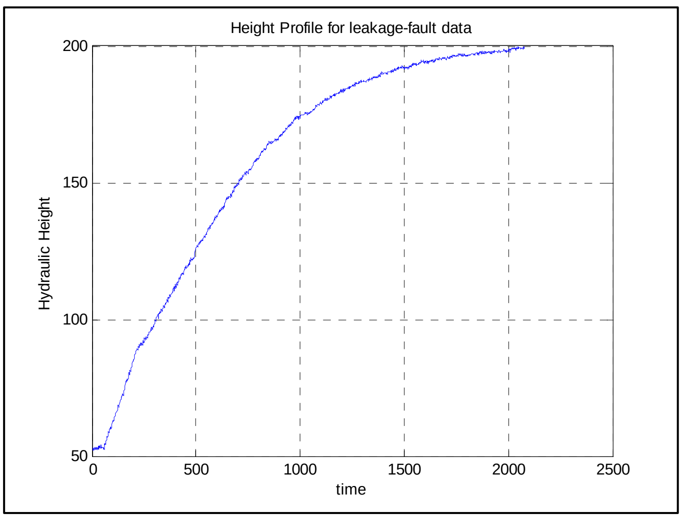


Figure 24 Height Profile of a Leakage-



Fault Data

Figure 25 Height Profile of a leakage-fault data

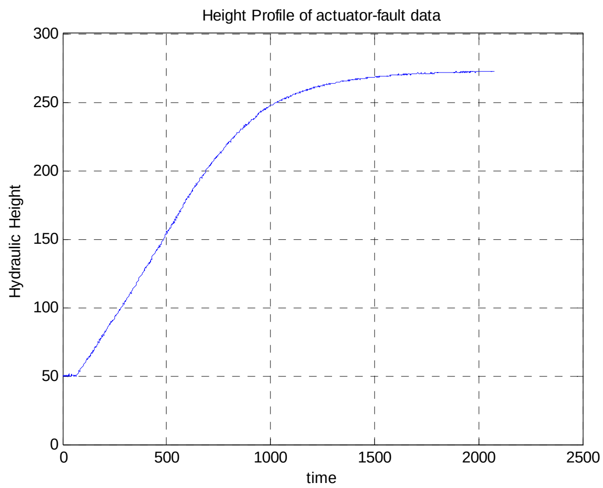


Figure 26 Height Profile of a Actuator-Fault Data

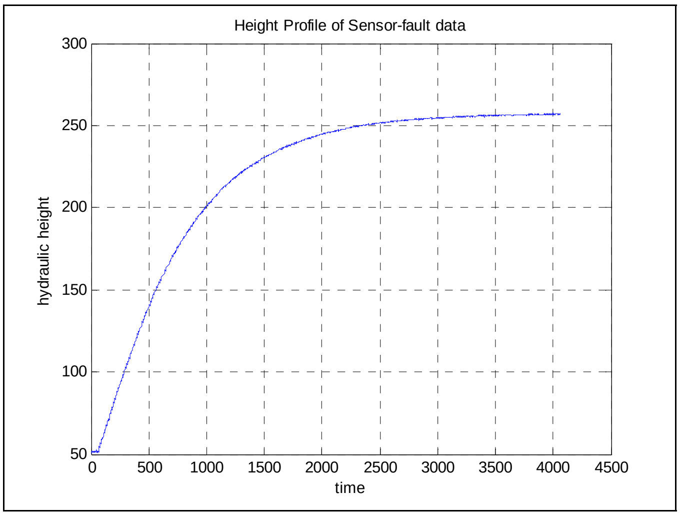


Figure 27 Height Profile of a Sensor-Fault Data

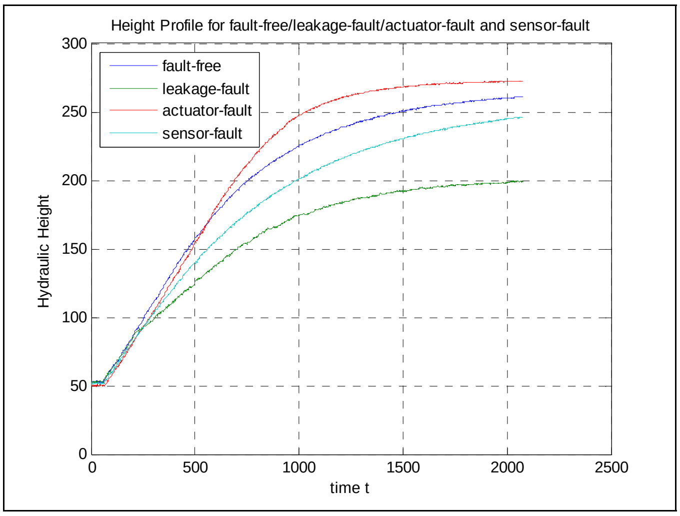


Figure 28 Height Profile of fault-free/actuator-fault and sensor-fault Data

7.2 Power Spectral Density Step Response

The fault signatures are analyzed from the step response parameters, namely settling time and steady-state error, and the power spectral density. Figures 19-22 show the step response power spectral density profiles under various types of faults when a proportional controller is employed. There is a change in the time constant, steady state value of the step responses and the power spectral density when the system is subjected to a fault.

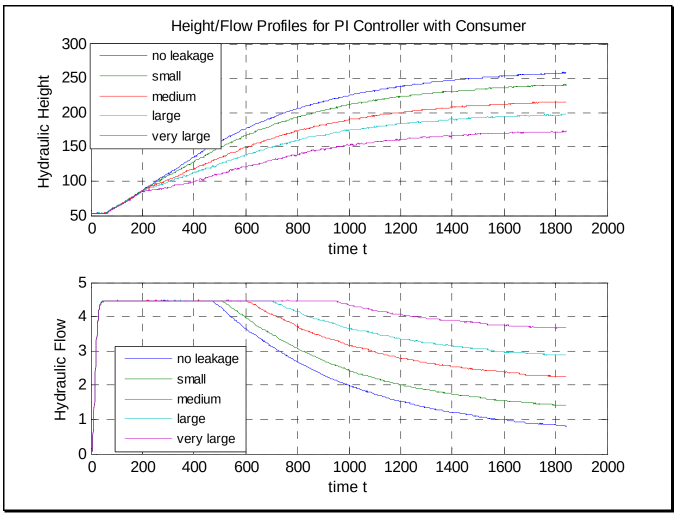


Figure 29 Height/Flow Profile under various Leakage magnitudes

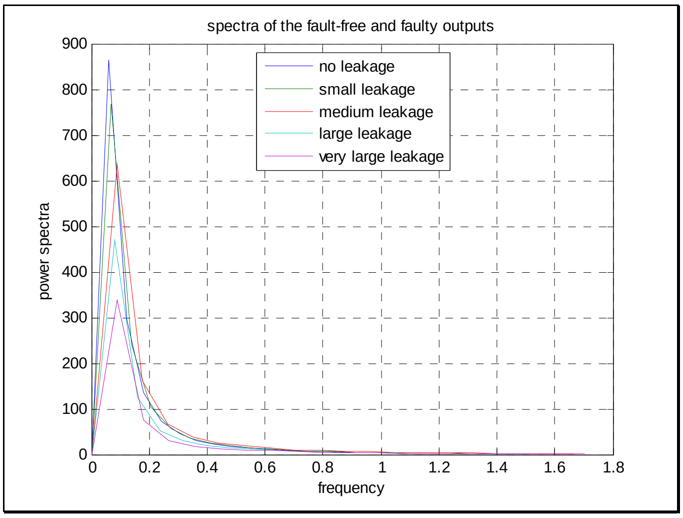


Figure 30 Spectra under various Leakage magnitudes

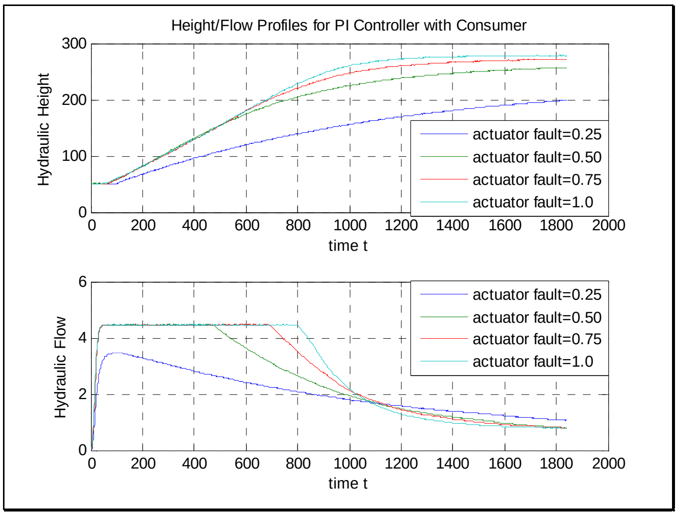


Figure 31Height/Flow Profile under various actuator faults

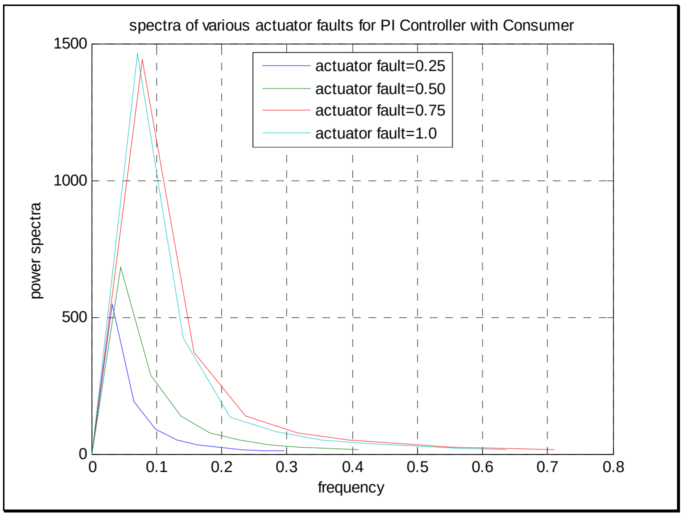


Figure 32 Spectra under various actuator faults

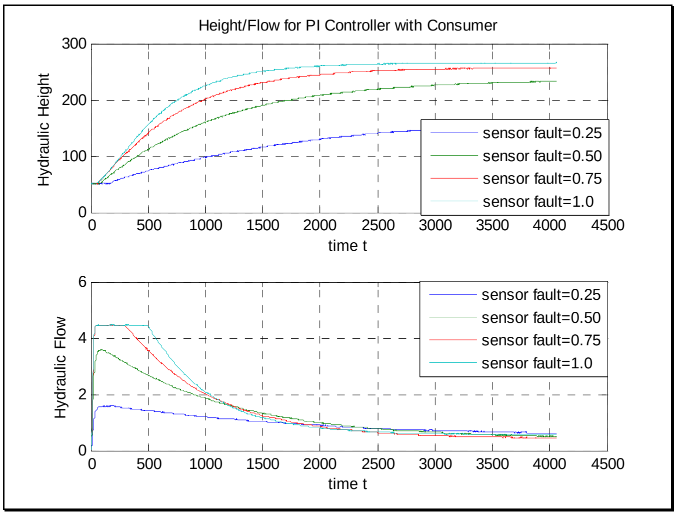


Figure 33 Height/Flow Profile under various sensor faults

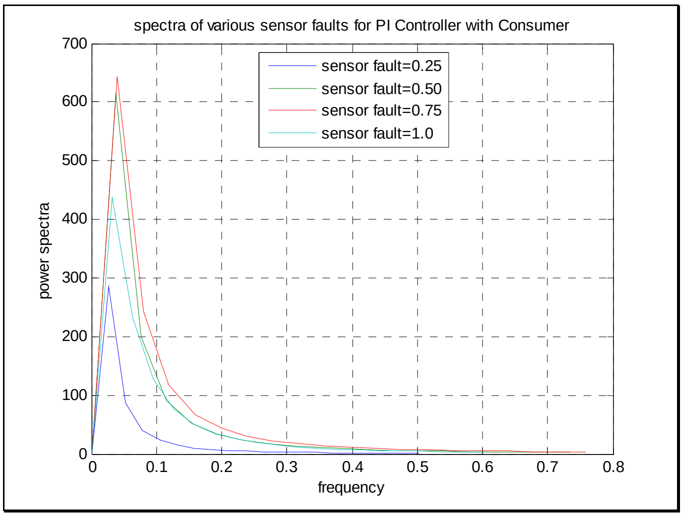


Figure 34 Spectra under various sensor faults

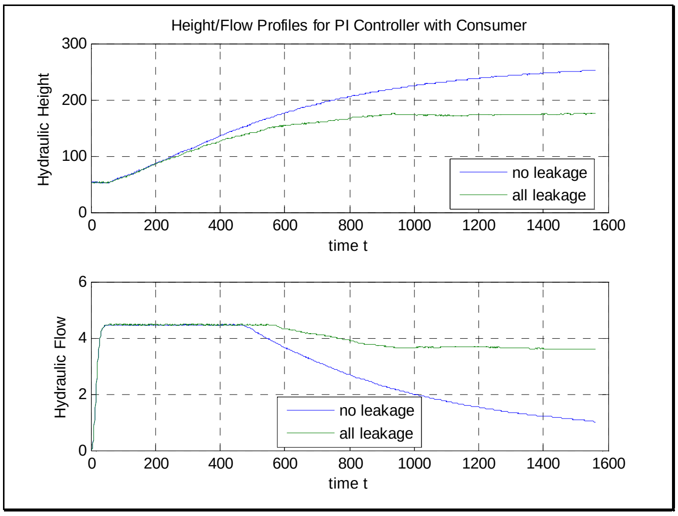


Figure 35 Input flow rate and tank height under various degrees of leakage

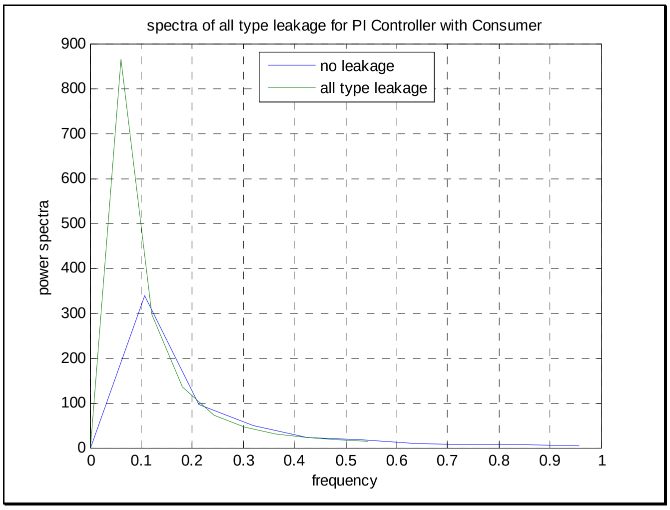


Figure 36 Input flow rate and tank height under various degrees of leakage

7.3 Coherence Spectral Density Step Response

Figs 33-35 give the profiles of the flow and height and the coherence spectra, whereas Fig. 23 shows height profiles in the presence of leakages of different magnitudes and when the fluid level system is operated in both open-loop and closed-loop configurations. For the open-loop case, one can readily deduce both the onset and amount of the leakage from the height/flow profile. The leakage flow has five sections corresponding to the following five degrees of no-leakage, small, medium, large and very large leakage. However, by its very nature, the closed-loop PI controller hides the fault and hence makes it difficult to visually detect the fault.

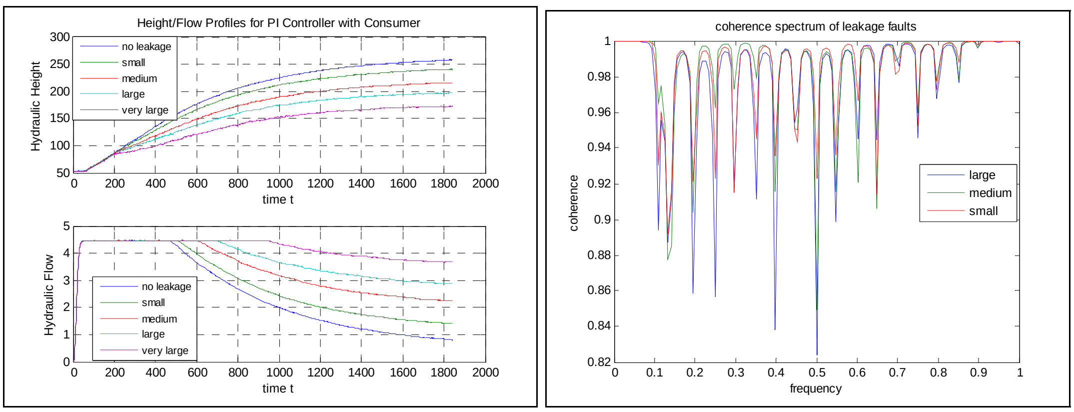


Figure 37 Height/Flow Profile and coherence under leakage faults

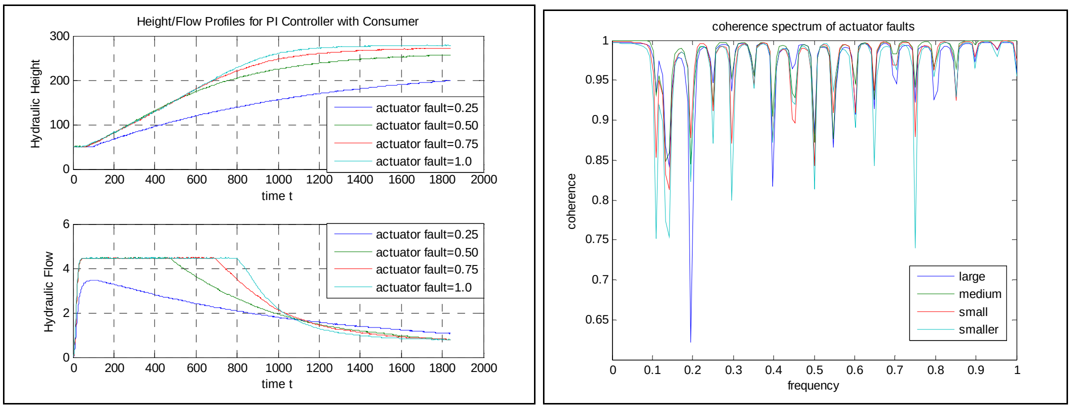


Figure 38 Height/Flow Profile and coherence under actuator faults

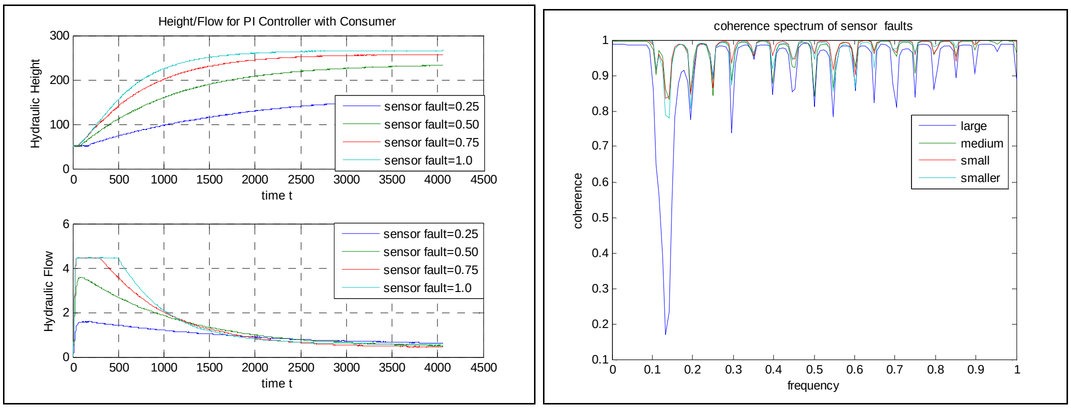


Figure 39 Height/Flow Profile and coherence under sensor faults

7.4 Model Order Selection Criteria: AIC Measure

Case 1: PI Controller

Table 5 PI Controller: Poles of the identified model for different selected orders

order 1	order 2	order 3	order 4	order 5	order 6
0.9850	0.9847	0.9847	0.9845	0.9845	0.9845
	0.0712	0.0340 ± j0.5910	0.0915	0.3301 ± j0.5549	0.4726
			-0.0189 ± j0.6197	-0.3058 ± j0.5938	0.2429 ± j0.6847
					-0.4400 ± j0.5933

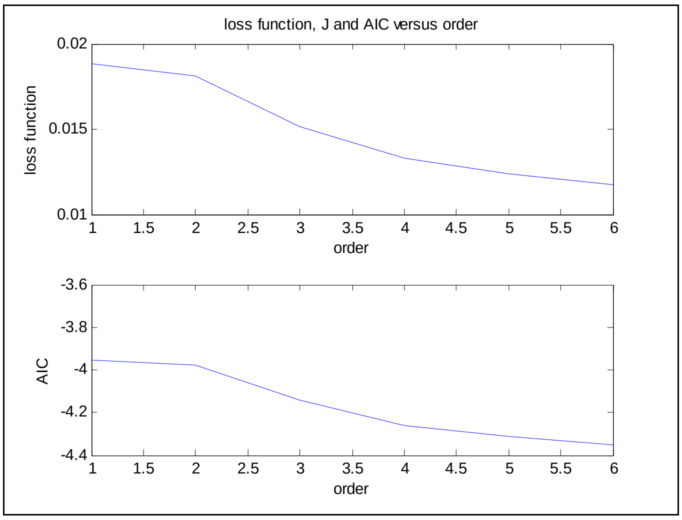


Figure 40 Loss function and AIC versus Order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

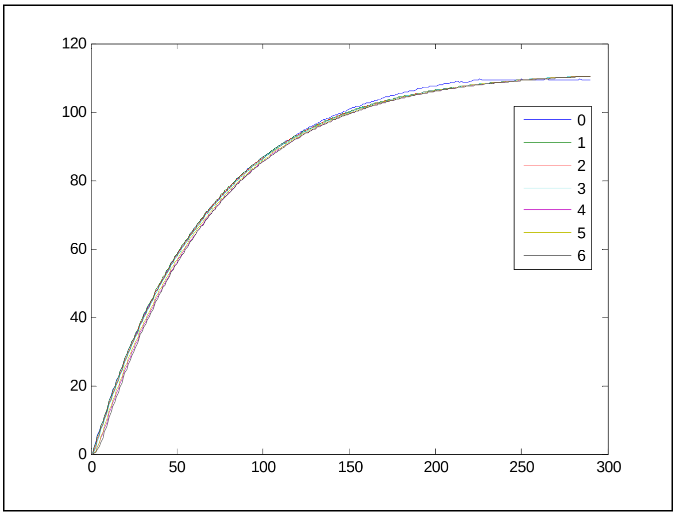


Figure 41 Step response vs. order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

For the selected model orders 1 to 3, all the poles are on the right half plane. The question arises as how to selected an appropriate order. We use the loss function and the AIC measures as guide lines. We selected a second order as it was found to be of minimal order which yielded acceptable performance.

Case 2: Proportional controller

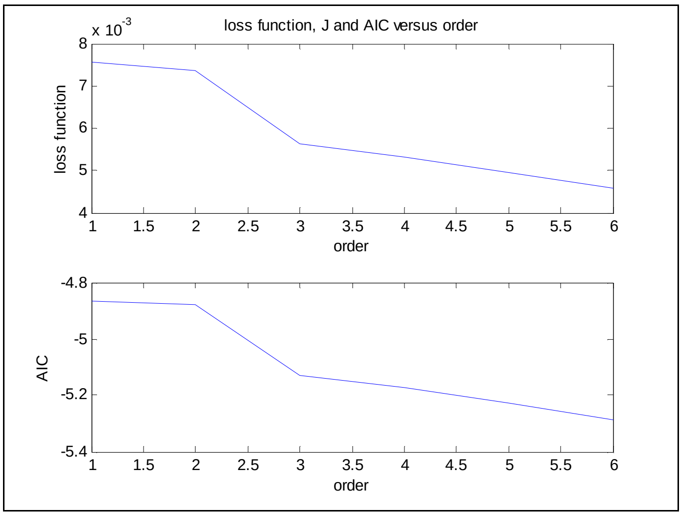


Figure 42 Loss function and AIC versus order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

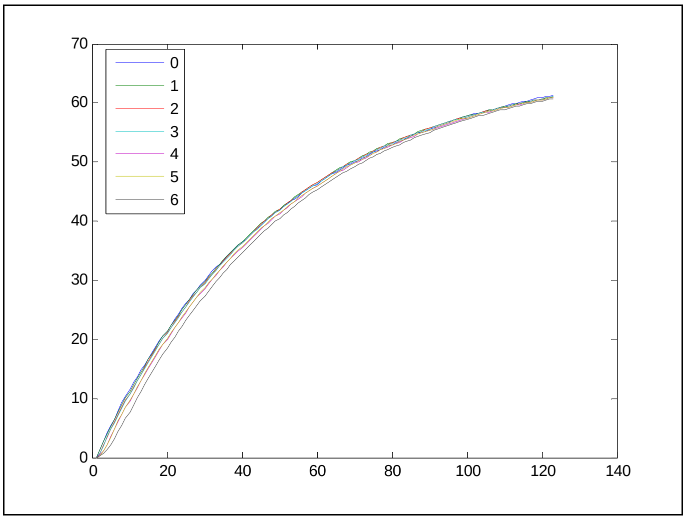


Figure 43.Step response vs. order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

Table 6 Proportional Controller: Poles of the identified model for different selected orders

order 1	order 2	order 3	order 4	order 5	order 6
0.9797	0.9796	0.9798	0.9800	0.9799	0.9799
	-0.1022	-0.0596 ± j0.6064	-0.2215 ± j0.6608	-0.3701 ± j0.6607	0.6519
					0.1858 ± 0.7008i
					-0.5087 ±j0.6557

Case 3: On-off controller

Table 7 On-Off Controller: Poles of the identified model for different selected orders

order 1	order 2	order 3	order 4	order 5	order 6
0.9996	0.9996	0.9996	0.9996	0.9996	0.9996
	-0.1839	-0.1230 ± j0.5358	0.1147	0.4118 ± j0.6558	0.5400
			-0.1731 ± j0.5491	-0.5280 ± j0.6373	0.2898 ± j0.7560
					-0.6182 ± j0.6456

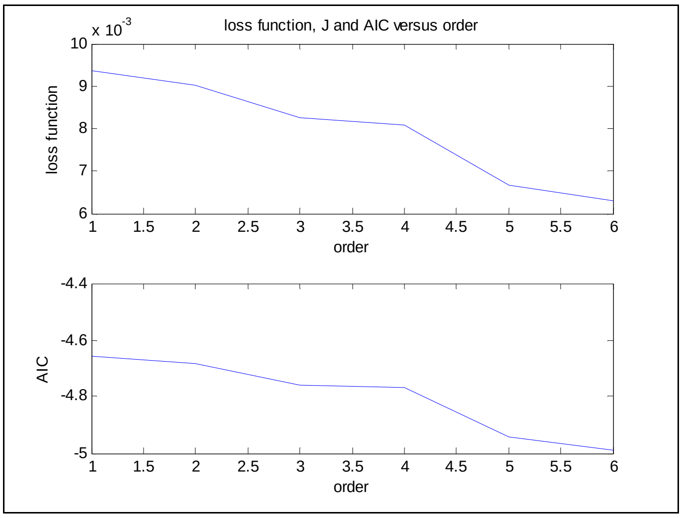


Figure 44 Loss function and AIC versus order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

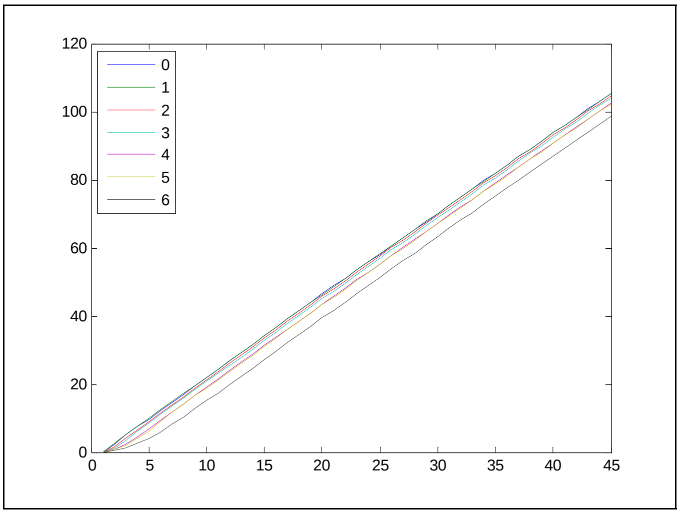


Figure 45 Step response vs. order. Order '0' indicates the actual data and '1','2',..,'6' indicates the selected order.

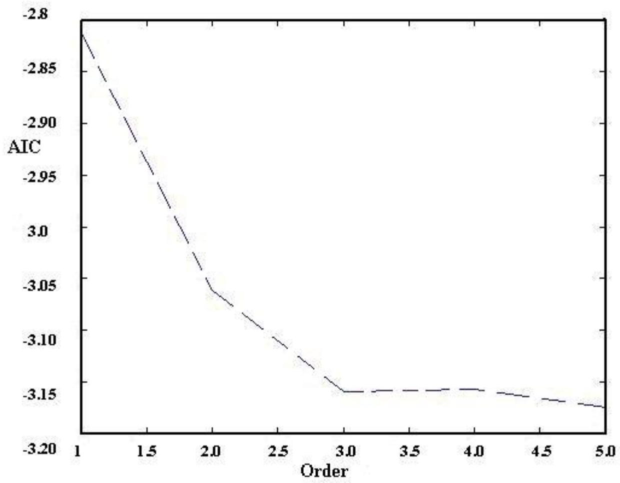


Figure 46 AIC: On/Off Controller

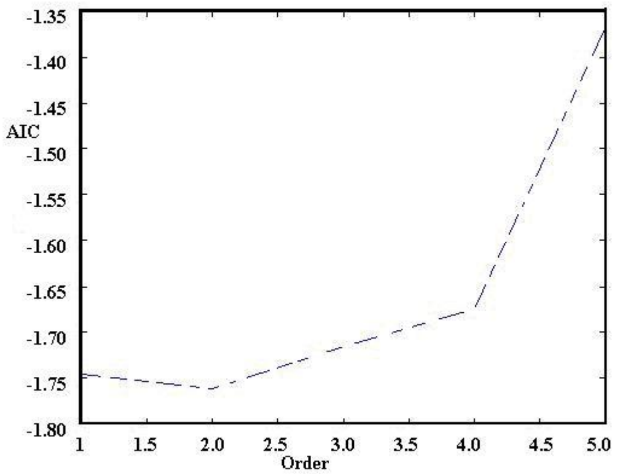


Figure 47 AIC: P Controller

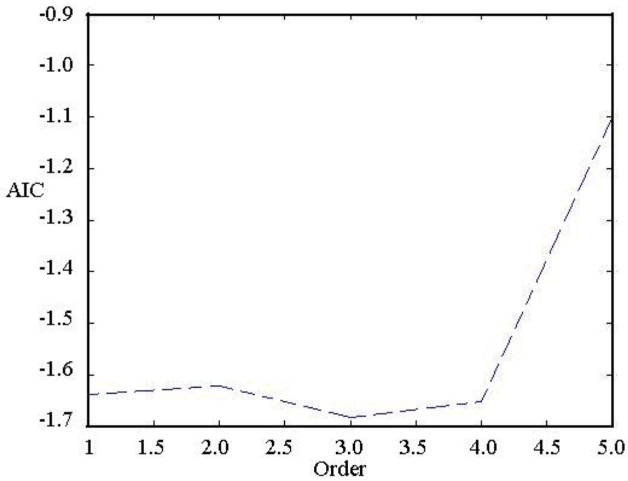


Figure 48 AIC: PI Controller

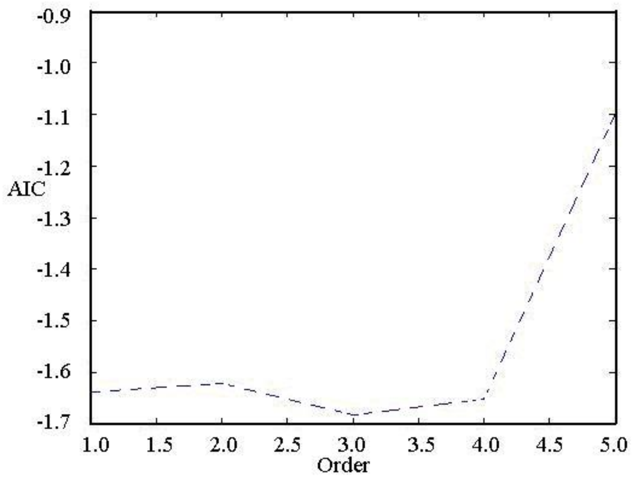


Figure 49 AIC: PID Controller

7.5 Kalman Filter Evaluation

The Kalman filter was evaluated under different fault scenarios for an on-off controller, a P controller, and a PI controller. (See Fig.13-14).

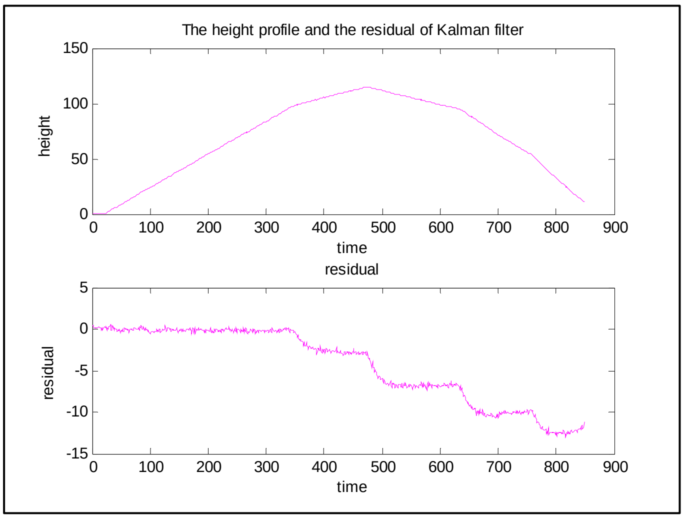


Figure 50 Kalman filter results for On-Off Controller: for Flow and Height under various leakage magnitudes

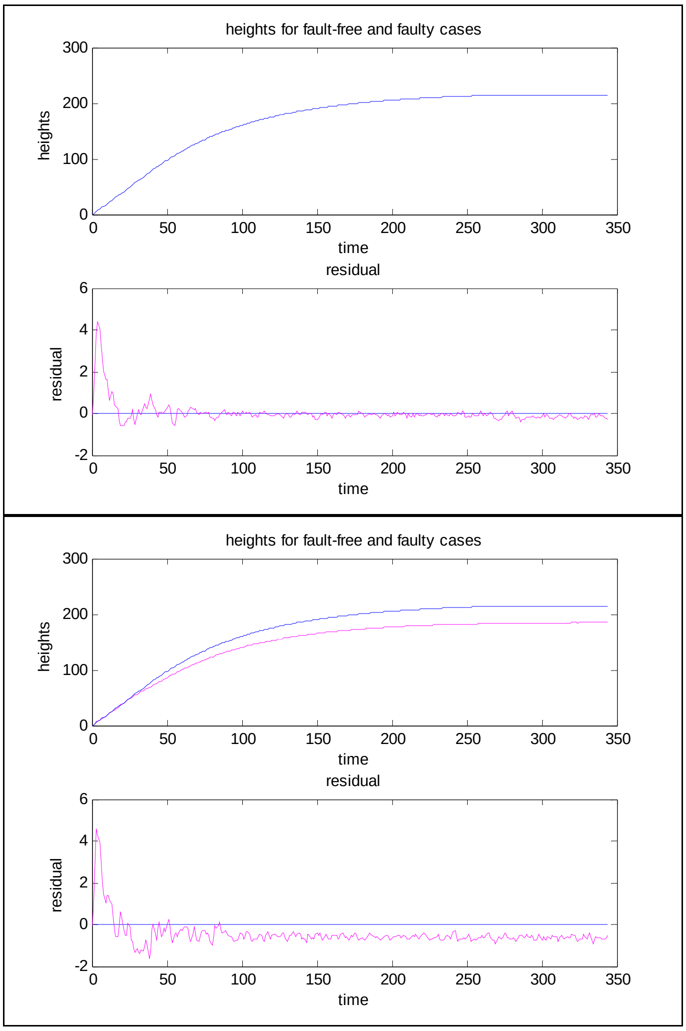


Figure 51 Height profiles and the residual when there is no fault and when there is a leakage fault.

Chapter 8

Conclusion

Overview

This chapter starts with the discussion of results from this thesis. It discusses the conclusion drawn from the research done.

The proposed fault diagnostic scheme based on (i) a sequential integration of model-free and model-based approach and (ii) the use of a new model selection criterion for system identification, was found promising when applied to a benchmarked laboratory-scale two-tank system. Through an integration of ANN and FL, the model-free approach quickly and reliably detects a presence of a possible fault from the profiles of the sensor outputs. The ANN is driven by the coherence spectrum of the residuals whereas the FL is fed with steady-state sensor output values. The model-free approach is also capable of providing a quick visual detection of the onset of any fault from the changes in the fault signatures such as settling time, steady-state sensor output values, and the coherence spectrum of the residuals. The fault indications obtained by the model-free approach are subsequently confirmed by the model-based approach which, through the use of a Kalman filter followed by a fault isolation scheme, provides a further necessary stage for capturing any faults, especially incipient ones, which may have escaped capture by the ANN-FL combination due to either insufficient training or incomplete fuzzy rules. Based on extensive simulations and an evaluation on a physical system, the proposed model order selection criterion was shown to be reliable and efficient. It has the ability to capture the input-output dynamic behavior accurately.

8.1 Recommendations for Future Work

- ✓ Results from the evaluation on the physical system shows that the Kalman filter is robust in modeling uncertainties including nonlinearities and neglected fast dynamics, while retaining its sensitivity to incipient faults. The fault isolation scheme, based on offline perturbed parameters experiments, was also found promising. Moreover this scheme bears a close resemblance to a neural network-based fault isolation scheme. However, this resemblance, though interesting, is currently undergoing further analysis.
- ✓ Hybrid model-based and soft computing techniques will be implemented to fault diagnosis problem. The model-based is built upon a pure but representative model of the plant. Based on the outcomes of this step, a neuro-fuzzy system is built. Once the neuro-fuzzy structure (rules number and premise and consequence membership function parameters) is identified and optimized, it is used in a generalization phase to achieve near-optimal online detection and identification with a reasonable computational complexity.

REFERENCES

[1] S. Nimmo, "Adequately address abnormal operations," Chemical Engineering Progress, vol. 91, pp. 36–45, 1995.

[2] H. Vedam and V. Venkatasubramanian, "PCA-SDG based process monitoring and fault detection," Control Engineering Practice, vol. 7, pp. 903–917, 1999.

[3] K.S Narendra and K. Parthasarathy (1990). Identification and control of dynamical systems using neural networks, *IEEE Transactions on Neural Networks* **1**(1): 12–18.

[4] P.M Frank and B. Köppen-Seliger (1997). New developments using AI in fault diagnosis, *Engineering Applications of Artificial Intelligence* 10(1): 3–14.

[5] B. Köppen-Seliger and P. M. Frank (1999). Fuzzy logic and neural networks in fault detection, *in* L. Jain and N. Martin (Eds.), *Fusion of Neural Networks, Fuzzy Sets, and Genetic Algorithms*, CRC Press, New York, NY, pp. 169–209.

[6] J. Korbicz, Ko´scielny, Z. Kowalczuk and W. Cholewa (2004). *Fault Diagnosis. Models, Artificial Intelligence, Applications*, Springer, Berlin.

[7] J. Korbicz (2006). Fault detection using analytical and soft computing methods, *Bulletin of the Polish Academy of Sciences: Technical Sciences* 54(1): 75–88.

[8] R. Isermann, " Fault diagnosis of Machines via parameter estimation and knowledge processing", Automatica, Vol. 29, No.4, pp. 825-825, 1993.

[9] R. Doraiswami, C.P.Diduch and Jiong Tang, "A Diagnostic Model For

Identifying Parametric Faults" to be published in IEEE Transactions on Control System Technology in 2009.

[10] R.Doraiswami, "Modelling and identification for fault diagnosis: a new paradigm" Proceedings of the 10th International Conference on Control Applications, September, 2001.

[11] S. Haykin (1999). *Neural Networks. A Comprehensive Foundation, 2nd Ed.*, Prentice-Hall, Englewood Cliffs, NJ.

[12] O. Nelles(2001). *Nonlinear System Identification. From Classical Approaches to Neural Networks and Fuzzy Models*, Springer-Verlag, Berlin.

[13] R. Doraiswami, Sevenson, M.; Diduch, C.P ,"Autonomous control systems: monitoring, diagnosis and tuning", IEEE Transactions on Systems, Man and Cybernetics, Part A, Volume 26, Issue 5, Sept.1996 pp. 646 – 655.

[14] Patan K. and Parisini T. (2005). Identification of neural dynamic models for fault detection and isolation: The case of a real sugar evaporation process, *Journal of Process Control* **15**(1): 67–79.

[15] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri. A review of process fault detection and diagnosis part i: Quantitative modelbased methods. Computers and Chem. Eng. 27, 293—311 (2003).

[16] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri. A review of process fault detection and diagnosis part II: Qualitative models and search strategies. Computers and Chem. Eng. 27, 313—326 (2003).

[17] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri. A review of process fault detection and diagnosis part III: Process history based methods. Computers and Chem. Eng. 27, 327—346 (2003).

[18] R. Milne. Strategies for diagnosis. IEEE Trans. on Syst., Man and Cyber. 17(3), 333—339 (1987).

[19] R. J. Patton. Robustness in model-based fault diagnosis: the 1995 situation. A. Rev. Contr. 21, 103—123 (1997).

[20] J. J. Gertler. Survey of model-based failure detection and isolation in complex plants. IEEE Contr. Syst. Mag. 8(6), 3—11 (1988).

[21] P. M. Frank and X. Ding. Survey of robust residual generation and evaluation methods in observer-based fault detection systems. J. of Process Contr. 7(6), 403—424 (1997).

[22] R. V. Beard. "Failure Accommodation in Linear System Through Self Reorganization". PhD thesis, Massachusetts Institute of Technology (1971).

[23] R. J. Patton. Robustness in model-based fault diagnosis: the 1995 situation. A. Rev. Contr. 21, 103—123 (1997).

[24] J. J. Gertler. Survey of model-based failure detection and isolation in complex plants. IEEE Contr. Syst. Mag. 8(6), 3—11 (1988).

[25] P. M. Frank and X. Ding. Survey of robust residual generation and evaluation methods in observer-based fault detection systems. J. of Process Contr. 7(6), 403—424 (1997).

[26] R. V. Beard. "Failure Accommodation in Linear System Through Self Reorganization". PhD thesis, Massachusetts Institute of Technology (1971).

[27] R. N. Clark. The dedicated observer approach to instrument failure detection. In "Proc. of The 18th IEEE Conf. on Decision & Contr.", pages 237—241, Fort Lauderdale, FL, USA (1979).

[28] P. M. Frank. Advanced fault detection and isolation schemes using nonlinear and robust observers. In “10th IFAC World Congress”, Munich, Germany (1987).

[29] P. M. Frank. Enhancement of robustness in observer-based fault detection. In “Proc. IFAC/IMACS Sympo. SAFEPROCESS’91”, volume 1, pages 275—287, Baden-Baden (1991).

[30] R. J. Patton and S. M. Kangethe. “Fault Diagnosis in Dynamic Systems, Theory and Application”, chapter 4, pages 99—154. Prentice Hall (1989).

[31] R. J. Patton and J. Chen. Observer-based fault detection and isolation: Robustness and applications. Contr. Eng. Practice 5(5), 671—682 (1997).

[32] R. K. Mehra and J. Peschan. An innovation approach to fault detection and diagnosis systems. Automatica 7, 637—640 (1971).

[33] A. S. Willsky. A survey of design methods for failure detection in dynamic systems. Automatica 12(6), 601—611 (1976).

[34] A. Zolghadri. An algorithm for real-time failure-detection in kalman filters. IEEE Trans. on Automat. Contr. 41(10), 1537—1539 (1996).

[35] M. S. N. Sa¨id, M. E. H. Benbouzid, and A. Benchaib. Detection of broken bars in induction motors using an extended Kalman filter for rotor resistance sensorless estimation. IEEE Trans. on Energy Conversion 15(1), 66—70 (2000).

[36] J. O’Reilly. “Observers for Linear Systems”. Academic Press (1983).

[37] P. M. Frank. Advanced fault detection and isolation schemes using nonlinear and robust observers. In “10th IFAC World Congress”, Munich, Germany (1987).

[38] R. Seliger and P. M. Frank. Fault diagnosis by disturbance decoupled nonlinear observers. In “Proc. 30th IEEE Conf. on Decision & Contr.”, volume 3, pages 2248—2253, Brighton, England (1991).

[39] R. Seliger and P. M. Frank. Robust residual evaluation by threshold selection and a performance index for nonlinear observer-based fault diagnosis. In “Int. Conf. on Fault Diagnosis (Tooldiag’93)”, Toolouse, France (1993).

[40] G. Bastin and M. R. Gevers. Stable adaptive observers for nonlinear timevarying systems. IEEE Trans. on Automat. Contr. pages 650—658 (1988).

[41] F. Hamelin and D. Sauter. Robust residual generation for FDI in uncertain dynamic systems. In “Proc. 34th IEEE Conf. on Decision & Contr.”, New Orleans, USA (1995).

[42] O. O. Oyeleye and M. A. Kramer. Qualitative simulation of chemical process systems: Steady state analysis. AICHE J. 34(9), 1441—1454 (1988).

[43] J. Gertler and D. Singer. A new structural framework for parity equation-based failure detection and isolation. Automatica 26(2), 381—388 (1990).

[44] A. S. Willsky. A survey of design methods for failure detection in dynamic systems. Automatica 12(6), 601—611 (1976).

[45] R. Dunia, S. J. Qin, T.F. Edgar, and T.J. McAvoy. Identification of faulty sensors using principal component analysis. AICHE J. 42(10), 2797—2812 (1996).

[46] P. Eide and P. S. Maybeck. A failure detection system for the f-16. IEEE Trans. on Aerospace and Electronic Systems 32(3), 1125—1136 (1996).

[47] J. Gertler and R. Monajemy. Generating directional residuals with dynamic

parity relations. Automatica 31(4), 627—635 (1995).

[48] J. Gertler. Fault detection and isolation using parity relations. Contr. Eng. Practice 5(5), 653—661 (1997).

[49] E. Y. Chow and A. S. Willsky. Analytical redundancy and the design of robust failure detection systems. IEEE Trans. on Automat. Contr. (7), 603—614 (1984).

[50] R. J. Patton and J. Chen. A review of parity space approaches to fault diagnosis. In "Proc. IFAC/IMACS Sympo. SAFEPROCESS'91", volume 1, pages 239—255, Baden-Baden (1991).

[51] X. Ding and P. M. Frank. Fault detection via factorization approach. Syst. Contr. Lett. 14(5), 431—436 (1990).

[52] P. M. Frank and X. Ding. Frequency domain approach to optimally robust residual generation anf evaluation for model-based fault diagnosis. Automatica 30(4), 789—804 (1994).

[53] J. Chen and R. J. Patton. "Robust Model-Based Fault Diagnosis for Dynamic Systems". Kluwer Academic Publishers, Boston/Dordrecht/London (1999).

[54] M. Iri, K. Aoki, E. OShima, and H. Matsuyama. An algorithm for diagnosis of system failures in the chemical process. Computers and Chem. Eng. 3(1-4), 489—493 (1979).

[55] T. Umeda, T. Kuriyama, E. Oshima, and H. Matsuyama. A graphical approach to cause and effect analysis of chemical processing systems. Chem. Eng. Science 35(12), 2379—2388 (1980).

[56] J. Shiozaki, H. Matsuyama, E. OShima, and M. Iri. An improved algorithm for diagnosis of system failures in the chemical process. Computers and

Chem. Eng. 9(3), 285—293 (1985).

[57] J. Shiozaki, H. Matsuyama, K. Tano, and E. OShima. Fault diagnosis of chemical processes by the use of signed, directed graphs. extension to five-range patterns of abnormality. Int. Chem. Eng. 25(4), 651—659 (1985).

[58] M. Kokawa, M. Satoshi, and S. Shigai. Fault location using digraph and inverse direction search with application. Automatica 19(6), 729—735 (1983).

[59] M. A. Kramer and B. L. Palowitch. A rule based approach to fault diagnosis using the signed directed graph. AIChE J. 33(7), 1067—1078 (1987).

[60] O. O. Oyeleye and M. A. Kramer. Qualitative simulation of chemical process systems: Steady state analysis. AICHE J. 34(9), 1441—1454 (1988).

[61] N. A. Wilcox and D. M. Himmelblau. Possible cause and effect graphs (PCEG) model for fault diagnosis i: Methodology. Computers and Chem. Eng. 18(2), 103—116 (1994).

[62] N. A. Wilcox and D. M. Himmelblau. Possible cause and effect graphs (PCEG) model for fault diagnosis II: Applications. Computers and Chem. Eng. 18(2), 117—127 (1994).

[63] R. Vaidhyanathan and V. Venkatasubramanian. Digraph-based models for automated HAZOP analysis. Reliability Eng. and Syst. Safety 50(1), 33—49 (1995).

[64] H. Vedam and V. Venkatasubramanian. Signed digraph based multiple fault diagnosis. Computers and Chem. Eng. 21, S655—S660 (1997).

[65] C. Han, R. Shih, and L. Lee. Quantifying signed directed graphs with the fuzzy set for fault diagnosis resolution improvement. Indust. and Eng.

Chemistry Research 33(8), 1943—1954 (1994).

[66] A. Genovesi, J. Harmand, and J. P. Steyer. A fuzzy logic based diagnosis system for the on-line supervision of an anaerobic digestor pilot-plant. Biochemical Eng. J. 3(3), 171—183 (1999).

[67] R. Li and X. Wang. Qualitative/Quantitative simulation of process temporal behavior using clustered fuzzy digraphs. AIChE J. 47(4), 906—919 (2001).

[68] B. E. Kelly and F. P. Lees. The propagation of faults in process plants: 2. fault tree synthesis. Reliability Eng. 16(1), 39—62 (1986).

[69] S. Soylemez and W. D. Seider. A new technique for precedence-ordering chemical process equation sets. AICHE J. 19(5), 934—942 (1973).

[70] J. de Kleer and S. Brown. A qualitative physics based on confluences. Artificial Intelligence 24(1-3), 7—83 (1984).

[71] Y. Iwasaki and H. A. Simon. Causality in device behavior. Artificial Intelligence 29(1), 3—32 (1986).

[72] E. Sacks. Qualitative analysis of piecewise linear approximation. J. of Artificial Intelligence in Eng. 3(3), 151—155 (1988).

[73] B. Kuipers. Qualitative simulation. Artificial Intelligence 29(3), 289—338 (1986).

[74] O. Raiman. Order of magnitude reasoning. In “Proceedings of AAAI-86”, pages 100—104 (August 1986).

[75] M. Mavrovouniotis and G. Stephanopoulos. Reasoning with order of magnitudes and approximate relations. In “Proceedings of AAAI-87” (July 1987).

[76] H. Kay and B. Kuipers. Numerical behavior envelopes for qualitative models. In “Proceedings of AAAI-93”, pages 606—613, Menlo Park, CA (1993).

[77] S. D. Grantham and L. H. Ungar. A first principles approach to automated troubleshooting of chemical plants. Computers and Chem. Eng. 14(7), 783—798 (1990).

[78] S. D. Grantham and L. H. Ungar. Comparative analysis of qualitative models when the model changes. AICHE J. 37(6), 931—943 (1991).

[79] B. Falkenhainer and K. Forbus. Compositional modeling: Finding the right model for the job. Artificial Intelligence 51, 95—143 (1991).

[80] E.J. Henley. Application of expert systems to fault diagnosis. In “AIChE Annual Meeting”, San Francisco, CA (1984).

[81] K. Niida. Expert system experiments in processing engineering. In “Inst. Of Chem. Eng. Symposium Series”, pages 529—583 (1985).

[82] T. S. Ramesh, S. K. Shum, and J. F. Davis. A structured framework for efficient problem-solving in diagnostic expert systems. Computers and Chem. Eng. 12(9-10), 891—902 (1988).

[83] T. S. Ramesh, J. F. Davis, and G. M. Schwenzer. Knowledge-based diagnostic systems for continuous process operations based upon the task framework. Computers and Chem. Eng. 16(2), 109—127 (1992).

[84] V. Venkatasubramanian. CATDEX, knowledge-based systems in process engineering: Case studies in heuristic classification. Technical Report, The CACHE Corporation, Austin, TX (1989).

[85] T. E. Quantrille and Y. A. Liu. “Artificial Intelligence in Chemical Engineering”. Academic Press, San Diego, LA (1991).

[86] C. RojasGuzman and M. A. Kramer. Comparison of belief networks and rule-based systems for fault diagnosis of chemical processes. Eng. Applications of Artificial Intelligence 3(6), 191—202 (1993).

[87] M. Basila, G. Jr., Stefanek, and A. Cinar. A model-object based supervisory expert system for fault tolerant chemical reactor control. Computers and Chem. Eng. 14(4-5), 551—560 (1990).

[88] J. Zhang and P. Roberts. Process fault diagnosis with diagnostic rules based on structural decomposition. J. of Process Contr. 1(5), 259—269 (1991).

[89] W. Becraft and P. Lee. An integrated neural Network/Expert system approach for fault diagnosis. Computers and Chem. Eng. 17(10), 1001—1014 (1993).

[90] E. Tarifa and N. Scenna. Fault diagnosis, directed graphs, and fuzzy logic. Computers and Chem. Eng. 21, S649—654 (1997).

[91] J. Zhao, B. Chen, and J. Shen. A hybrid ANN-ES system for dynamic fault diagnosis of hydrocracking process. Computers and Chem. Eng. 21, S929—S933 (1997).

[92] M. Wo, W. Gui, D. Shen, and Y. Wang. Export fault diagnosis using role models with certainty factors for the leaching process. In "Proc. Of the Third World Congress on Intelligent Contr. and Automation", volume 1, pages 238—241, Hefei, China (28 June-2 July 2000).

[93] D. Leung and J. Romagnoli. Dynamic probabilistic model-based expert system for fault diagnosis. Computers and Chem. Eng. 24(11), 2473—2492 (2000).

[94] N. J. Scenna. Some aspects of fault diagnosis in batch processes. Reliability Eng. and Syst. Safety 70(1), 95—110 (2000).

[95] K. Watanabe, I. Matsura, M. Abe, M. Kubota, and D. M. Himmelblau. Incipient fault diagnosis of chemical processes via artificial neural networks. AICHE J. 35(11), 1803—1812 (1989).

[96] V. Venkatasubramanian and K. Chan. A neural network methodology for process fault diagnosis. AICHE J. 35(12), 1993—2002 (1989).

[97] L. H. Ungar, B. A. Powell, and S. N. Kamens. Adaptive networks for fault diagnosis and process control. Computers and Chem. Eng. 14(4-5), 561—572 (1990).

[98] J. C. Hoskins, K. M. Kaliyur, and D. M. Himmelblau. Fault diagnosis in complex chemical plants using artificial neural networks. AICHE J. 37(1), 137—141 (1991).

[99] V. Venkatasubramanian, R. Vaidyanathan, and Y. Yamamoto. Process fault detection and diagnosis using neural networks i: Steady state processes. Computers and Chem. Eng. 14(7), 699—712 (1990).

[100] R. Vaidyanathan and V. Venkatasubramanian. Representing and diagnosing dynamic process data using neural networks. Eng. Applications of Artificial Intelligence 5(1), 11—21 (1992).

[101] K. Watanabe, S. Hirota, L. Iloa, and D. M. Himmelblau. Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks. AICHE J. 40(5), 839—848 (1994).

[102] J. Y. Fan, M. Nikolaou, and R. E. White. An approach to fault diagnosis of chemical processes via neural networks. AICHE J. 39(1), 82—88 (1993). 1

[103] A. E. Farell and S. D. Roat. Framework for enhancing fault diagnosis capabilities of artificial neural networks. Computers and Chem. Eng. 18(7), 613—635 (1994).

[104] C. S. Tsai and C. T. Chang. Dynamic process diagnosis via integrated neural networks. Computers and Chem. Eng. 19, s747—s752 (1995).

[105] Y. Maki and K. A. Loparo. A neural-network approach to fault detection and diagnosis in industrial processes. IEEE Trans. on Contr. Syst. Technology 5(6), 529—541 (Nov 1997).

[106] J. A. Leonard and M. A. Kramer. Diagnosing dynamic faults using modular neural nets. IEEE Expert 8(2), 44—53 (1993).

[107] S. N. Kavuri and V. Venkatasubramanian. Representing bounded fault classes using neural networks with ellipsoidal functions. Computers and Chem. Eng. 17(2), 139—163 (1993).

[108] S. N. Kavuri and V. Venkatasubramanian. Using fuzzy clustering with ellipsoidal units in neural networks for robust fault classification. Computers and Chem. Eng. 17(8), 765—784 (1993).

[109] S. N. Kavuri and V. Venkatasubramanian. Neural network decomposition strategies for large scale fault diagnosis. Int. J. of Contr. 59(3), 767—792 (1994).

[110] J. Y. Fan, M. Nikolaou, and R. E. White. An approach to fault diagnosis of chemical processes via neural networks. AICHE J. 39(1), 82—88 (1993).

[111] A. E. Farell and S. D. Roat. Framework for enhancing fault diagnosis capabilities of artificial neural networks. Computers and Chem. Eng. 18(7), 613—635 (1994).

[112] C. S. Tsai and C. T. Chang. Dynamic process diagnosis via integrated neural networks. Computers and Chem. Eng. 19, s747—s752 (1995).

[113] Y. Maki and K. A. Loparo. A neural-network approach to fault detection

and diagnosis in industrial processes. IEEE Trans. on Contr. Syst. Technology 5(6), 529—541 (Nov 1997).

[114] S. N. Kavuri and V. Venkatasubramanian. Representing bounded fault classes using neural networks with ellipsoidal functions. Computers and Chem. Eng. 17(2), 139—163 (1993).

[115] S. N. Kavuri and V. Venkatasubramanian. Using fuzzy clustering with ellipsoidal units in neural networks for robust fault classification. Computers and Chem. Eng. 17(8), 765—784 (1993).

[116] S. N. Kavuri and V. Venkatasubramanian. Neural network decomposition strategies for large scale fault diagnosis. Int. J. of Contr. 59(3), 767—792 (1994).

[117] X. Z. Wang, B. H. Chen, S. H. Yang, and C. Mcgreavy. Application of wavelets and neural networks to diagnostic system development. 2. an integrated framework and its application. Computers and Chem. Eng. 23(7), 945—954 (1999).

[118] J. V. Kresta, J. F. MacGregor, and T. E. Marlin. Multivariate statistical monitoring of process operating performance. Can. J. of Chem. Eng. 69(1), 35 (1991).

[119] J. F. MacGregor and T. Kourti. Statistical process control of multivariate processes. Contr. Eng. Practice 3(3), 403—414 (1995).

[120] B. M. Wise and N. B. Gallagher. The process chemometrics approach to process monitoring and fault detection. J. of Process Contr. 6(6), 329—348 (1996).

[121] J. V. Kresta, J. F. MacGregor, and T. E. Marlin. Multivariate statistical monitoring of process operating performance. Can. J. of Chem. Eng. 69(1),

35 (1991).

[122] P. Nomikos and J. MacGregor. Monitoring batch processes using multiway principal component analysis. AICHE J. 40(8), 1361—1375 (1994).

[123] S. J. Qin and T. J. McAvoy. Nonlinear PLS modeling using neural networks. Computers and Chem. Eng. 16(4), 379—391 (1992).

[124] D. Dong and T. J. McAvoy. Batch tracking via nonlinear principal component analysis. AICHE J. 42(8), 2199—2208 (1996).

[125] A. Raich and A. Cinar. Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. AICHE J. 42(4), 995—1009 (1996).

[126] A. Raich and A. Cinar. Diagnosis of process disturbances by statistical distance and angle measures. Computers and Chem. Eng. 21(6), 661—673 (1997).

[127] M. H. Kaspar and W. H. Ray. Chemometric methods for process monitoring and high performance controller design. AICHE J. 38(10), 1593—1608 (1992).

[128] M.J. Piovoso, K.A. Kosanovich, and J. P. Yuk. Process data chemometrica. IEEE Trans. on Instrumentation and Measurement 41(2), 262—268 (1992).

[129] S. Qin and W. Li. Detection, identification, and reconstruction of faulty sensors with maximized sensitivity. AICHE J. 45(9), 1963—1976 (1999).

[130] R. Dunia, S. J. Qin, T.F. Edgar, and T.J. McAvoy. Identification of faulty sensors using principal component analysis. AICHE J. 42(10), 2797—2812 (1996).

[131] R. Dunia and S. J. Qin. A subspace approach to multidimensional fault identification and reconstruction. AIChE J. 44, 1813—1831 (1998).

[132] S. J. Qin. Recursive PLS algorithms for adaptive data modeling. Computers and Chem. Eng. 22, 503—514 (1998).

[133] W. Li, H. H. Yue, S. Valle-Cervantes, and S. J. Qin. Recursive PCA for adaptive process monitoring. J. of Process Contr. 10, 471—486 (2000).

[134] B. R Bakshi. Multiscale PCA with application to multivariate statistical process monitoring. AIChE J. 44, 1596—1610 (1998).

[135] R. Luo, M. Misra, S. J. Qin, R. Barton, and D. M. Himmelblau. Sensor fault detection via multiscale analysis and dynamic PCA. Ind. Eng. Chem. Res. 37, 1489—1495 (1999).

[136] H. Zhang, A. K. Tangirala, and S. L. Shah. Dynamic process monitoring using multiscale PCA. In "Proc. Of 1999 Canadian Conf. On Electrical and Comput. Eng.", pages 1579—1584 (1999).

[137] H. B. Aradhye, B. R. Bakshi, R. A. Strauss, and J. F. Davis. Multiscale SPC using wavelets - theoretical analysis and properties. AIChE J. 49, 939—958 (2003).

[138] R. Rengaswamy and V. Venkatasubramanian. A fast training neural network and its updation for incipient fault detection and diagnosis. Computers and Chem. Eng. 24(2-7), 431—437 (2000).

[139] H. Ocak. "Fault Detection, Diagnosis and Prognosis of Rolling Element Bearings: Frequency Domain Methods and Hidden Markov Modeling". PhD thesis, Case Western Reserve University (2003).

[140] N. J. Scenna. Some aspects of fault diagnosis in batch processes. Reliability

Eng. and Syst. Safety 70(1), 95—110 (2000).

[141] J. T. Cheung and G. Stephanopoulos. Representation of process trends part i. a formal representation framework. Computers and Chem. Eng. 14(4-5), 495—510 (1990).

[142] M. Janusz and V. Venkatasubramanian. Automatic generation of qualitative description of ProcessTrends for fault detection and diagnosis. Eng. Applications of Artificial Intelligence 4(5), 329—339 (1991).

[143] R. Rengaswamy and V. Venkatasubramanian. A syntactic pattern recognition approach for process monitoring and fault diagnosis. Eng. Applications of Artificial Intelligence 8(1), 35—51 (1995).

[144] H. Vedam and V. Venkatasubramanian. A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis. In “American Contr. Conf.”, pages 309—313 (1997).

[145] H. Vedam, V. Venkatasubramanian, and M. Bhalodia. A b-spline based method for data compression, process monitoring and diagnosis. Computers and Chem. Eng. 22, S827—S830 (1998).

[146] R. Rengaswamy, T. Hagglund, and V. Venkatasubramanian. A qualitative shape analysis formalism for monitoring control loop performance. Eng. Applications of Artificial Intelligence 14(1), 23—33 (2001).

[147] F. O. Karry and C. W. De Silva, “Soft Computing and Intelligent Systems Design: Theory, Tools and Applications”, Addison Wesley, 2004.

[148] S.X. Ding, “ Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools” Springer-Verlag 2008.

[149] S. Simani, C. Fantuzzi and R J Patton, Model-based Fault Diagnosis using

Identification Techniques, Advances in Industrial Control, Springer Verlag, 2003.

[150] R.J. Patton, P. M. Frank, and R. N. Clark, Issues in Fault Diagnosis for Dynamic Systems, Springer-Varlag, 2000.

[151] J. Chen, and R.J. Patton, Robust Model-based Fault Diagnosis for Dynamic Systems, Kluwer Academic Publishers, 1999.

[152] J. J. Gertler Fault Detection and Diagnosis in Engineering Systems, Marcel Dekker, Inc, 1998.

[153] R. Isermann, " Fault diagnosis of Machines via parameter estimation and knowledge processing", Automatica, Vol. 29, No.4, pp. 825-825, 1993.

[154] R. Doraiswami, C.P.Diduch and Jiong Tang, "A Diagnostic Model For Identifying Parametric Faults", IFAC World Congress, July 2008. Also accepted in the IEEE transaction on Control Systems Technology, 2009.

[155] M. Ferrente and B. Brunone, "Pipe system diagnosis and leak detection by unsteady-state tests", Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongking, China, June 25-27, 2008.

[156] R. Doraiswami, Stevenson, M. Diduch, C.P, "Autonomous Control Systems: Monitoring, diagnosis and tuning", IEEE Transactions on Systems, Man. And Cybernetics, Part A, Vol. 26, Issue 5, September 1996, pp 645-655.

[157] R. Doraiswami, "Modelling and identification for fault diagnosis: a new paradigm" Proceedings of the 10th International Conference on Control Applications, September, 2001.

[158] E. A. Astrom," Control of Complex Systems", Springer-Verlag, 2001.

[159] M.A.A. Shoukat Chaudhury, Mrudil Jain and Sirish L. Shah, 'Detection and Quantification of Valve Stiction" American Control Conference, Minneapolis, Minnesota, USA June 14-16, 2006.

[160] R. Doraiswami, M. Sevenson, C.P. Diduch, "Autonomous control systems: monitoring, diagnosis and tuning", IEEE Transactions on Systems, Man and Cybernetics, Part A, Volume 26, Issue 5, Sept.1996 pp. 646 – 655.

[161] R. Doraiswami ; L. Cheded, "A new model order selection criterion with applications to physical systems", Internal report, 2009 (also submitted for publication to IEEE Journal on selected topics in signal processing).

[162] Q. Yang, "Model-Based and Data Driven Fault Diagnosis Mehotds with Applications to Process Monitoring", PhD Thesis, Case Western Reserve University, May 2004.

[163] R. Doraiswami, L. Cheded and Haris M. Khalid, "Fault Diagnosis using Sequential Integration of Model-free and Model-Based Approaches", Internal report, 2009 (also submitted for publication to IEEE Conference on Decision and Control (CDC), 2009).

[164] R. Doraiswami, "Modeling and System Identification", Class notes, 2009.

[165] R. Doraiswami, "Intelligent Control Systems", Class notes, 2009.