Fault Detection and Classification Using Kalman Filter and Hybrid Neuro-Fuzzy Systems

Amar Khoukhi, Haris M. Khalid, Lahouari Cheded and Rajamani Doraiswami

* Correspondence should be addressed to Haris M. Khalid, mhariskh@kfupm.edu.sa

Abstract— In this paper, an efficient scheme to detect and classify faults in a system using Kalman filtering and hybrid neuro-fuzzy computing techniques, respectively, is proposed. A fault is detected whenever the moving average of the Kalman filter residual exceeds a threshold value. The fault classification has been made effective by implementing a hybrid genetic neuro-fuzzy Inference system (GANFIS). By doing so, the critical information about the presence or absence of a fault is gained in the shortest possible time, with not only confirmation of the findings but also an accurate unfolding-in-time of the finer details of the fault, thus completing the overall fault diagnosis picture of the system under test. The proposed scheme is evaluated extensively on a two-tank process used in industry exemplified by a benchmarked laboratory scale coupled-tank system.

Index Terms — ***Kalman filter, soft computing, ANN, genetic algorithm, ANFIS, GANFIS, fault detection, fault isolation, benchmarked laboratory scale two-tank system.***

I. Introduction

Challenging design problems arise in modern fault diagnosis systems. Unfortunately, the classical analytical techniques often cannot provide acceptable solutions to such difficult tasks. This explains why soft computing techniques such as fuzzy logic, neural networks and evolutionary algorithm have become more and more popular in industrial applications of fault diagnosis. Process faults, if undetected, have a serious impact on process operation and efficiency~~,~~ product quality, safety, productivity and pollution level. In order to detect, diagnose and correct these abnormal process behaviors, the use of efficient and advanced automated diagnostic systems is of great importance to modern industries. The main objective of fault detection and isolation (FDI) is to provide early warnings to operators, such that appropriate actions can be taken to prevent the breakdown of the system caused by the occurrence of faults. This will improve the reliability and safety of the system, and avoid unnecessary and costly downtimes. Complete reliance on human operators to monitor the conditions of the systems is often difficult, especially as engineering systems are becoming ever more complex.

For example, in chemical processes, several kinds of failures may compromise safety and productivity. In fact, the occurrence of faults may reduce the efficiency of the process (e.g., lower product quality) or, in the worst scenarios, could lead to fatal accidents (e.g., temperature run-away) leading to injuries to personnel, environmental pollution and equipment damage. Major failures to be considered in chemical processes are: actuator failures (e.g., electric-power failures, pump failures, valves failures), process failures (e.g., abrupt variations of some process parameters, side reactions due to impurities in the raw materials) and sensor failures. To tackle these difficulties, Fault Diagnosis and Isolation (FDI) techniques need to be developed.

The model-based approach is popular for developing FDI techniques [1][2]. It mainly consists of two stages [3]. The first one is to generate residuals by computing the difference between the measured output and the estimated output obtained from the model of the system. Any departure of the residuals from zero indicates that a fault has likely occurred [4].

However, these methods are developed mainly for linear systems assuming that a precise mathematical model of the system is available. This assumption, however, may be difficult to satisfy in practice, especially as engineering systems are becoming more complex and are in general nonlinear [5][6]. Several model-based studies on the detection and identification of faults and tuning parameters were considered in [7-8].

Soft computing techniques offer an emerging alternative approach to the classical FDI based on analytical techniques and can deliver economical and competitive solutions to real-world problems. Their operation parallels the remarkable ability of the human mind to reason and learn in circumstances characterized by uncertainty and imprecision. As such and with their ability to acquire and use imprecise (incomplete) knowledge of the process under study, they complement the ability of classical FDI techniques which require precision, certainty and rigor. For example they allow the use of qualitative information from practicing operators which may play a vital role in achieving an accurate and robust diagnosis of faults in system components (e.g. motors) at early stages.

The paper is organized as follows: Section II reviews some related studies in this area, whereas Section III states the necessary theory behind the model-based approach to the fault diagnosis to the problem at hand. Section IV discusses the genetic neuro-fuzzy framework which provides a model-free-based approach to the fault diagnosis problem at hand Section V discusses the integration of these two approaches. Section VI discusses the implementation of this integrated approach and the simulation results obtained thereof. Finally some conclusions are given Section VII.

II. Overview of FDI Techniques

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/filter-based, parity space and frequency domain methods. Good survey papers include [9-13]. The mathematical model-based approach adopted in this paper falls into the observer category. 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 or statistical filters (e.g. the Kalman filter) in a stochastic setting. 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 or nonlinear, full or reduced-order, fixed or adaptive observers (or Kalman filters) in the generation of residuals.

Soft computing techniques are used to develop models required in FDI. Through the use of expert knowledge, rules and training, these techniques can provide models for a wide class of nonlinear systems with arbitrary accuracy. Among these techniques, neural networks are well recognized for their learning ability which allows them to approximate a wide class of nonlinear functions with an arbitrary accuracy [14]. For these reasons, they have been applied to many engineering problems [15-18], and used as models to generate residuals for fault detection [19]. However, these networks are inadequate in isolating faults, as they are black boxes in nature. Further, it is also desirable that a fault diagnostic system should be able to incorporate the experience of the operators [20] which cannot be completely represented by a dynamical model. Fuzzy reasoning allows symbolic generalization of numerical data by fuzzy rules and supports the direct integration of the experience of the operators in the decision-making process of FDI in order to achieve a more reliable fault diagnosis [21]. An up-to-date application of FDI techniques to motor fault detection and isolation was recently published in a special section in [22]. In what follows, a brief literature review is presented under three sections comprising of fault diagnosis using expert systems, fuzzy logic, neural network and Genetic Algorithm.

In recent years, the application of fuzzy logic to model-based fault diagnosis has gained increasing attention in both fundamental research and application. Rule-based feature extraction has been widely used in expert systems for many applications. Initial attempts at the application of expert systems for fault diagnosis can be found in the work of Henley [23] and Niida [24]. Structuring the knowledge-base through hierarchical classification can be found in [25]. Ideas on knowledge-based diagnostic systems based on the task framework can be found in [26]. A rule-based expert system for fault diagnosis in a cracker unit is described in [27]. More work on expert systems in chemical process fault diagnosis can be found in [28] and [29]. Wo et al. [30] presented an expert fault diagnostic system that uses rules with certainty factors. Leung and Romagnoli [31] presented a probabilistic model-based expert system for fault diagnosis. An expert system approach for fault diagnosis in batch processes was also discussed in Scenna [32]. Expert systems possess attractive features as they are knowledge based systems generated using input/output data and an inference engine built upon a set of rules defined by experts in the problem field. Nevertheless, major drawbacks of these systems include adaptivity to new and changing environment, the increasing of the number of rules.

Soft computing have been seen by many researchers as an alternative to classical expert systems paradigm, and a considerable interest was shown in the literature regarding the application of soft computing to fault diagnosis. A number of papers address the problem of fault diagnosis using back-propagation neural networks.

In chemical engineering, Watanabe et al. [33], Venkatasubramanian and Chan [34], Ungar et al. [35] and Hoskins et al. [36] 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. [37]. This work was later extended to utilize dynamic process data by Vaidyanathan and Venkatasubramanian [38]. Hierarchical neural network architecture for the detection of multiple faults was proposed by Watanabe et al. [39]. Most of the work on the improvement of the 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. [40], Farell and Roat [41], Tsai and Chang [42], and Maki and Loparo [43]. Modifications to the selection of basis functions have also been suggested to the standard back-propagation network with a view to improving both the accuracy and training time. For example, Leonard and Kramer [44] suggested the use of radial basis function networks for fault diagnosis applications.

Genetic Algorithms (GAs) have been implemented for a wide variety of problems, both real-world (e.g. fault diagnosis and fault tolerant systems) and abstract (e.g. solving NP-complete problems [45]). The bulk of the GA literature is concerned with practical applications. For a very complete bibliography, see [45], which contains a comprehensive survey.

However, soft computing techniques remain inadequate for fault isolation when compared to their model-based counterparts. This prompted us, in this paper, to propose a hybrid genetic neuro-fuzzy approach to the FDI problem, which meets the requirements for a quick and reliable fault detection and isolation scheme. The proposed scheme has been evaluated on a laboratory-scale two-tank system. It is a widely used prototype used to illustrate typical processes used in the wastewater treatment plants, petro-chemical plants, and oil/gas systems. The main contribution of this paper is therefore to illustrate how the fusion of different soft computing techniques into a single hybrid FDI system, namely a genetic neuro-fuzzy system, can provide accurate and reliable practical FDI systems.

III. MODEL-BASED APPROACH

System model: The state-space model of the system (A, B, C) is given by (See equation (1-2)):

$$x(k+1) = Ax(k) + Bu(k) + E_w w(k) \quad (1)$$

$$y(k) = Cx(k) + \upsilon(k) \quad \textbf{(2)}$$

where $x(k)$ is $nx1$ state vector, $y(k)$ is the scalar output; A is nxn, B is $n \times 1$, E_w is $n \times 1$ and C is $1 \times n$ matrices; $w(k)$ and $v(k)$ are zero-mean white noise processes representing the disturbance signal and measurement noise respectively, with covariances given by

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

The disturbance $w(k)$ may represent any input that affects the system but cannot be directly manipulated (unlike the input u including the gravity load, electrical power demand, fluctuations in the flow in a fluid system, wind gusts, bias, power frequency signal, dc offset, crew or passenger load in vehicles such as spacecrafts, ships, and helicopter and planes, and faults. The disturbance $w(k)$ is a mathematical entity that accounts for the uncertainty in the *a-priori* knowledge of the plant model. The larger the covariance Q is, the less accurate the model (A,B,C) is and vice versa. The matrix E_w is termed as disturbance distribution matrix.

Assumptions: It is assumed that (a) (A,E_w) is controllable so that the disturbance affects all the states, and (b) (A,C) is observable so that knowing the input and the output will allow the states may be estimated.

Kalman filter model**:** The Kalman filter is a copy of the fault-free (or nominal) state-space model of the system denoted (A_0,B_0,C_0) driven by the residual $e(k)$, which is the error between the system output $y(k)$ and its estimate $\hat{y}(k)$, i.e. the residual is $e(k) = y(k) - \hat{y}(k)$. The Kalman filter becomes (See equation (3-5)):

$$\hat{x}(k+1) = A_0\hat{x}(k) + B_0u(k) + K_0\left(y(k) - C_0\hat{x}(k)\right) \quad (3)$$

$$\hat{y}(k) = C_0\hat{x}(k) \quad (4)$$

$$e(k) = y(k) - C_0\hat{x}(k) \quad (5)$$

where K_0 is $nx1$ vector termed Kalman gain; $\hat{x}(k)$ is the estimates of the state $x(k)$. 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 but 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.

Model-based fault diagnosis: A fault in system on a system may manifest as (a) a change in the parameters of subsystems and/or (b) as an additive exogenous input to the system resulting in a mismatch between the actual system model (A,B,C) and Kalman filter model (A_0,B_0,C_0). The residual $e(k)$ will be a zero-mean white noise process if and only if there is no fault, (there is no model mismatch). If there is a fault, the residual will not be a zero-mean white noise process. In this case the residual is modeled by including a bias term. Let H_0 and H_1 be the two hypotheses indicating respectively the absence and the presence of a fault. The corresponding residual models under these hypotheses are given below (See equation (6)):

$$e(k) = \begin{cases} e_0(k) & : H_0 \\ e_f(k) + e_0(k) & : H_1 \end{cases} \quad (6)$$

where $e_0(k)$ is a zero-mean white noise process and $e_f(k)$ is a non-zero bias term. A Bayes statistical decision-theoretic approach is used to decide between two hypotheses. The Bayes decision strategy takes the general form (See equation (7)):

$$t_s(\boldsymbol{e}) \begin{cases} \le \eta & \textit{no fault} \\ > \eta & \textit{fault} \end{cases} \quad (7)$$

where $\boldsymbol{e}$ is a $N \times 1$ residual vector $\boldsymbol{e} = \left[e(k) \quad e(k-1) \quad e(k-2) \quad . \quad e(k-N+1)\right]^T$ and $t_s(\boldsymbol{e})$ is the test statistics of the residual $\boldsymbol{e}$ and η is the threshold value computed taking into account the variance of the noise term $e_0(k)$, prior probabilities of the two hypotheses, the cost associated with a correct or wrong decision, and the probability of a false alarm. Test statistics if the reference input $r(k)$ is a constant, a sinusoid of frequency f_0 and an arbitrary signal are listed below (See equation (8)):

$$t_s(\boldsymbol{e}) = \begin{cases} \left|\dfrac{1}{N}\displaystyle\sum_{i=k-N+1}^{k} e(i)\right| & r(k) = \text{constant} \\ P_{ee}(f_0) & r(k) \text{ is a sinusoid} \\ \dfrac{1}{N}\displaystyle\sum_{i=k-N+1}^{k} e^2(i) & r(k) \text{ is an arbitrary signal} \end{cases} \quad (8)$$

IV. Genetic Neuro-Fuzzy Inference Systems

Clustering is dividing a given data set into several groups (called clusters) so that all the data elements in a group are (a) similar or (b) dissimilar to the other elements in the group or in the rest of the other groups. The similarity and dissimilarity measures are appropriately defined such that elements in a group characterize a particular fault scenario: e.g. Euclidian distance between 2 vectors. Commonly used clustering techniques include K-means clustering, Fuzzy C-means clustering, Mountain clustering, and Subtractive clustering. From a machine learning viewpoint, clustering can be viewed as unsupervised learning of concepts.

For the GANFIS system proposed herein, subtractive clustering technique is used to formulate an ANFIS as shown in Fig. 4 . The subtractive clustering (SC) approach divides the data into clusters by defining a cluster center based on the density of surrounding data points. A radius for each cluster is chosen. All the data points within the radial distance of this point are then removed in order to determine the next data cluster and its center. This process is repeated until all the data is within the radial distance of a cluster center. SC algorithm finds optimal data point to define a cluster center based on the density of surrounding data points.

The subtractive clustering technique has been applied here in order to form hybrid versions of Neuro-Fuzzy. The GANFIS (GA+ANFIS) method is made up of six layers, as shown in Fig. 4. The first layer is an input layer, characterizing the crisp inputs. The second layer performs the fuzzification of the crisp inputs into linguistic variables, through Gaussian transfer functions. The third is the rule layer, which applies the product t-norm to produce the firing strengths of each rule. This is followed by a normalization layer, at which each node calculates the ratio of a rule's firing strength to the sum of all rules firing strengths. The fifth layer performs the inversion. The last layer does the aggregation and defuzzification, where an output is obtained as the summation of all incoming signals. The procedure for the Subtractive Clustering proceeds by defining a cluster center based on the density of the surrounding data points. All the data points within the Radii of this point form then a cluster. This scheme was repetitively implemented so as to get a final trained ANFIS. The selection of the radius is crucial: a slight change in the radius of subtractive clustering may result in a noticeable change in the results. So, the genetic optimization of the parameters is used in order to get a close-to-optimal value of the clustering radius value resulting in an optimal data clustering leading to a better fault isolation.

Comments: Both approaches have advantages and limitations. The GANFIS is a fast online method but is computationally burdensome to determine the clusters off line. The performance is critically dependent upon the set of fault data employed in creating the clusters. For better detection accuracy, it should include all fault scenarios that may occur.

The Kalman filter does not require any past fault data. It can estimate the fault status for any types of faults including unforeseen fault types if the Kalman filter model is accurate. The Kalman filter may not be able to capture accurately the system model when the operating regime is in a highly nonlinear region. Hence, to obtain a fast and reliable fault diagnosis scheme, both approaches are combined to get the best of both approaches.

V. Sequential Integration of Two Approaches

In this paper, a sequential integration of both model-free (GANFIS) and model-based (Kaman filter) 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 are executed with an increasing precision accompanied by a gradual unfolding-in-time diagnostic picture that reveals the presence of an incipient fault. The two-stage scheme of Fig. 1 involves a model-free stage based on GANFIS followed by a model-based one based on Kalman filter. The GANFIS scheme is fast and provides quickly the fault status: a fault is present or absent so that an appropriate action that is short of shutting down the system being monitored, may be taken. After some delay due to the computational burden involved in the model-free approach, the Kalman filter then asserts its fault status. If both approaches assert that there is a fault, a drastic action is taken to prevent the fault from adversely affecting the system's performance. The GANFIS scheme merely states qualitatively the presence or absence of fault. The Kalman filter scheme gives the quantitative picture of the fault status.

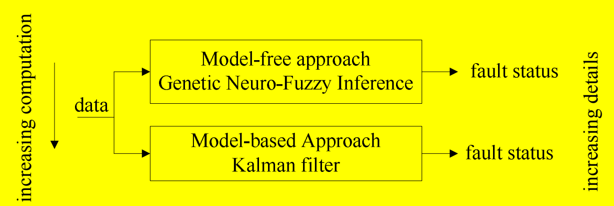


Figure 1: Sequential integration of GANFIS and Kalman filter

VI. Evaluation on a Physical Process Control System

A. System Description

The Benchmarked laboratory-scale process control system was used to collect data. The data was collected at a sampling time of 50 milliseconds. Different data sets were generated for the PI-based water level control system. Different fault scenarios were also been considered for the generation of the data sets.

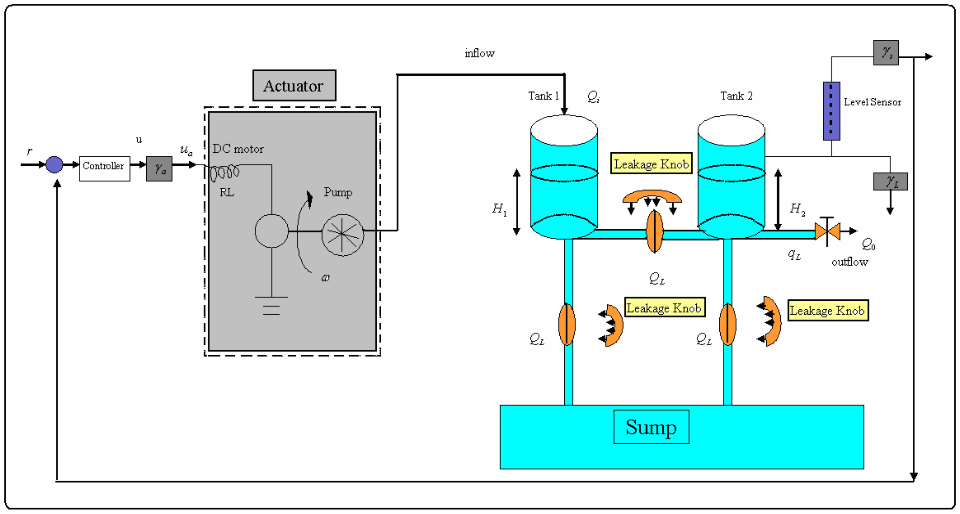


Figure 2: Two-tank diagram

The proposed scheme was evaluated on the above- cited process control system. The scheme is carried out by jointly interpreting model outputs. The implementation plan for the proposed scheme is shown in the Fig. 1.

B. Experimental Setup

The process Data was generated through an experimental setup as shown in Fig. 3. A two-tank system was used in order to collect the data with the introduction of actuator, and sensor faults through the system as can be seen in the labview circuit window. An amplified voltage of 18 volts was used to handle the controller effectively for the changes/fluctuations produced in the system. So, the fault diagnosis was done here in a closed-loop identification setup where the controller tends to suppress the faults while it is performing its feedback control task.

C. Process Data Collection and Description

The process data was collected at 50-millisecond sampling time. The main objective of the benchmarked dual-tank system is to reach a reference height of 200 ml in the second tank. During this process, several faults were introduced such as leakage faults, sensor faults and actuator faults. Leakage faults were introduced through the pipe clogs of the system, knobs between the first and the second tank, etc. Sensor faults were simulated by introducing a gain in the circuit as if there was a fault in the level sensor of the tank. Actuator faults were simulated by introducing a gain in the setup for the actuator that comprises of the motor and pump. A PI controller was employed in order to reach the desired reference height. Due to the inclusion of faults, the controller was finding it difficult to reach the desired level. For this reason, the power of the motor was increased from 5 volts to 18 volts in order to provide it with the maximum throttle to reach the desired level. In doing so, the actuator performed well in achieving its desired level but it also suppressed the faults of the system. So, it made the task of detecting the faults. After the collection of data, techniques of estimating key parameters such as settling time, steady- state value, and coherence spectra can be used to help us get a useful insight into the fault.

D. Model of the Coupled Tank System

The physical system under evaluation is formed of two tanks connected by a pipe. The leakage is simulated in the tank by opening the drain valve. A DC motor-driven pump supplies the fluid to the first tank and a PI controller is used to control the fluid level in the second tank by maintaining the level at a specified level, as shown in Fig. 2.

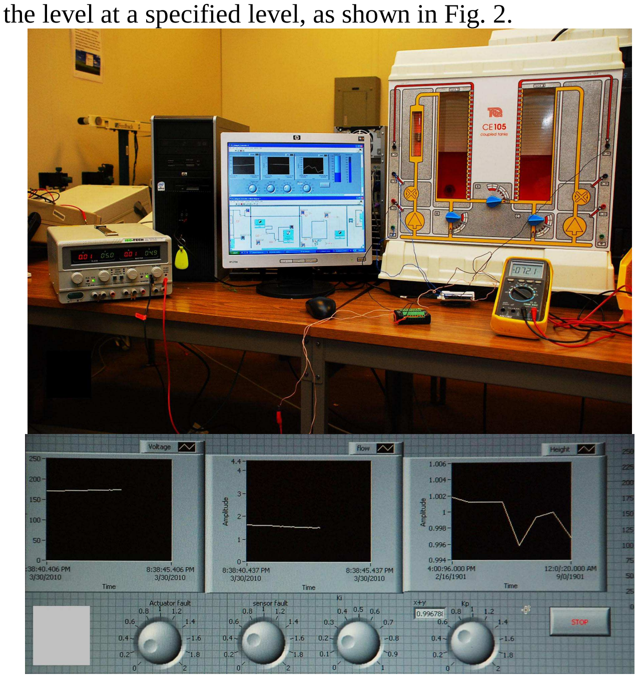


Figure 3: (Top) – The two tank system interfaced with the Labview through a DAQ and the amplifier for the magnified voltage, (Bottom) – The labview setup of the apparatus including the ciruit window and the block diagram of the experiment.

A step input is applied to the dc motor- pump system to fill the first tank. The opening of the drainage valve introduces a leakage in the tank. Various types of leakage faults are introduced and the liquid height in the second tank, H_2 and the inflow rate, Q_i, are both measured. The National Instruments LABVIEW package is employed to collect these data.

Appendix 1 gives the full details of the mathematical model, including the linearized version, of dual-tank process that is diagrammatically depicted in Fig. 2.

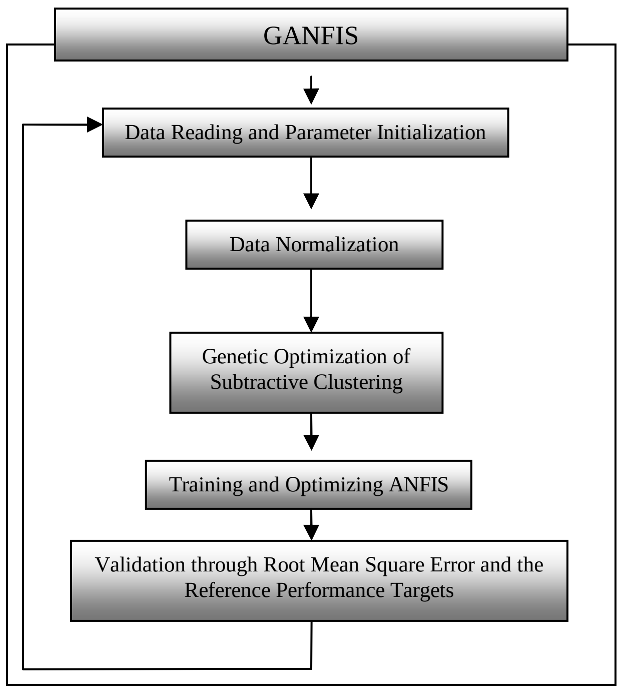


Figure 4. Implementation Scheme (GANFIS system)

Subtractive Clustering ANFIS-Based Fault Diagnosis

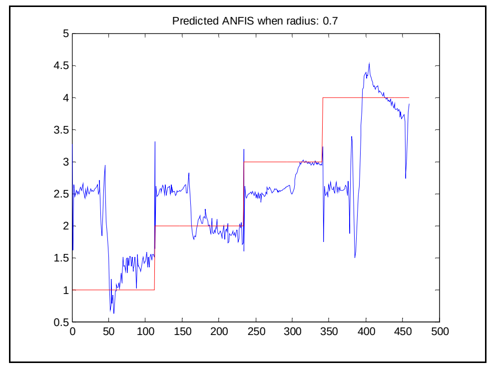
The first step is to compute offline clusters using subtractive clustering approach. A set of data comprising height profiles of the two-tank system under the normal or fault-free and various leakage fault scenarios were collected. The data set include all possible leakage fault scenarios. Each data element represents *N x 1* vector of height of the tank. Entire data is an *N* x *M* matrix where *M* is the number of experiments performed to obtain the height profile for different operating scenarios. This data matrix was divided into three groups such that vector elements in each group have similar properties and dissimilar to the elements in other groups. For example data elements associated with large, medium and small leakages were separated in to different groups. To achieve this, a subtractive clustering based on Neuro-Fuzzy technique is employed.

ANFIS performs better in following the Original output when the radius is 0.7 rather than when the radius is 0.2 as shown in Fig. 5 where the red line is showing the fault classification rung for faults and the predicted ANFIS fault classification is shown by blue line. As can be seen from the previous section that a slight change in the radius of subtractive clustering yields noticeable a change in the results. So, the genetic optimization of the parameters is used in order to get a close-to-optimal value of the clustering radius value that will lead to a better tracking of

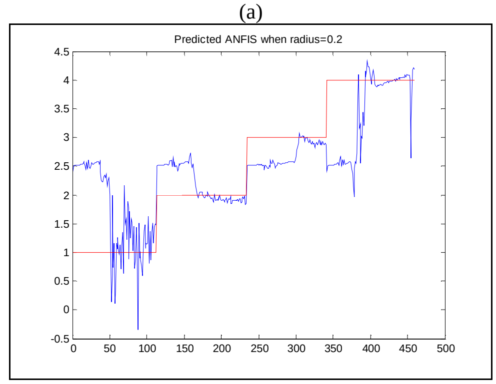
A PI controller, with gains k_p and k_I, is used to maintain the level of the Tank 2 at the desired reference input r .

faults.

Fig. 6 shows the four various process phases in which the performance of the genetic optimization scheme was assessed. The original graph is drawn in red which is showing step sizes that indicate different levels of faults. The green graph shows the prediction with the subtractive clustering technique followed the blue graph in each of all the four phases which is showing the genetic optimization of the radius parameter of the sub-clustering technique. When implemented in the genetic algorithm, these functions give the best fitness function value as follows: *Fitness Function Value: 0.187462*

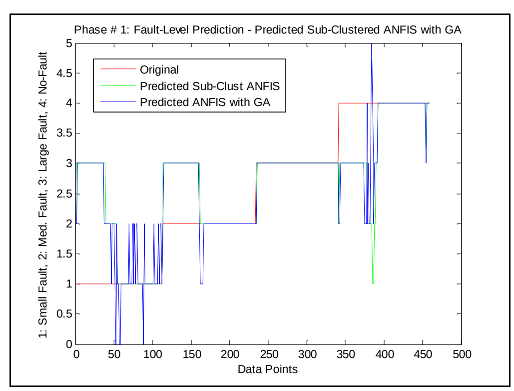


(a)

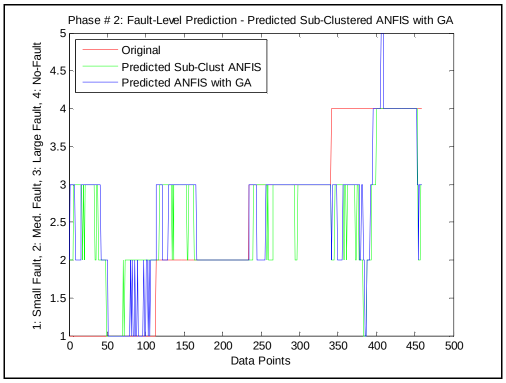


(b)

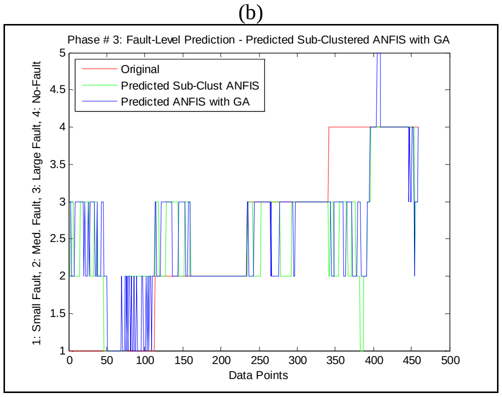
Figure 5: (a) Predicted ANFIS using Subtractive Clustering when radius: 0.7, (b) Predicted ANFIS using Subtractive Clustering when radius: 0.2



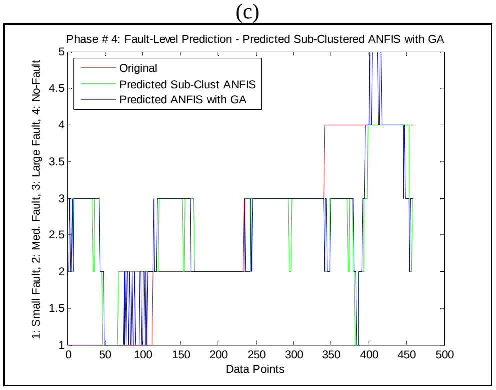
(a)



(b)



(c)



(d)

Figure 6: (a)(b)(c)(d) Four random Phases for trained sub-clustered ANFIS with GA

Kalman Filter-based Fault diagnosis

A fault-free model of the system is identified using a closed loop identification scheme [45]. The order of the estimated model was iterated to obtain an acceptable model structure using a combination of the AIC criterion and the identified pole locations. The identified model is essentially a second-order system even though the theoretical model is of a fourth order. 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. The reference input was a step input. The Kalman filter was evaluated under different fault scenarios for a PI controller. The Bayes strategy based on the absolute value of the mean of the residual is employed as the reference input was constant. The residual and the absolute mean of the residual are shown in sub-figures (A-D) of Fig. 8. A better visual picture of the fault status is shown in Fig where auto-correlations of the residuals clearly indicate presence or absence and the size of the fault. A fault is present if the auto-correlation is not delta function (indicating that the residual is not a zero mean white noise process).

A number of faults leakage faults were introduced. Fig. 7 shoes the normal and leakage faults of different degree severity. The Fig. 8 shows the residuals and their auto-correlations

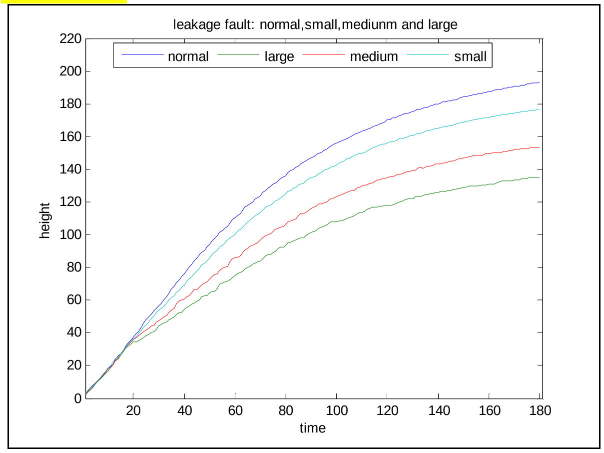


Figure 7: Large, medium and small leakage faults and fault-free heights

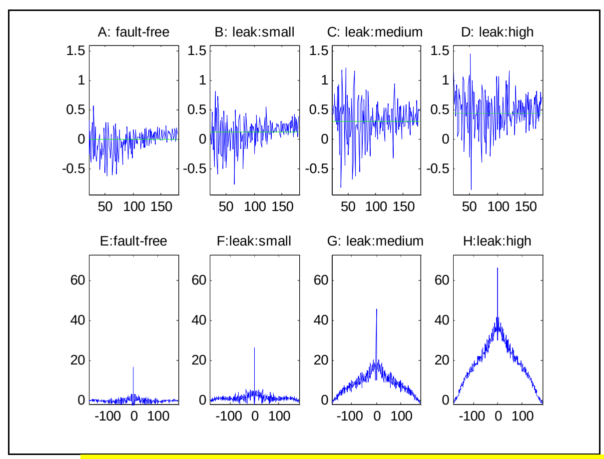


Figure 8: Sub-figures A, B, C and D show the residuals and their absolute means for no-leak, small, medium, and high leakage faults while sub-figures E, F, G and H show the corresponding autocorrelations

Fig. 8 clearly shows the presence of fault and the sizes of the fault via (a) the absolute mean value of the residual and (b) the auto-correlation of the residual.

In all of the cases, the fault status asserted by GANFIS and the Kaman filter were identical. The GANFIS is fast but provides only qualitative picture,

E. ANN-Based Fault Diagnosis

The analysis of the ANN is a difficult task and it requires wide experience in selecting the training sets, activation functions and the number and size of the hidden layers for the application at hand. A generic model of the ANN in fault diagnosis is as shown in Fig. 8, and a generic activation evaluation on a sigmoid activation function is shown in Fig. 9.

F. Discussion

In this paper, a hybrid FDI scheme, involving a model-based Kalman filter approach and a model-free genetic neuro-fuzzy approach, was proposed and evaluated on a lab-scale benchmarked process control system exemplified by a two-tank system. A good comparison of the techniques can be seen in the histograms shown in Figures 9 and 10. The chart in Fig. 9 shows the comparison with the error rates between fuzzy, ANN, ANFIS and the proposed genetic neuro-fuzzy method. It is important to note here that the error rate for the GANFIS is the least one because the genetic algorithm has performed well in the optimization of the subtractive clustering. In Fig. 10, it can be seen, that when the radius of the subtractive clustering is chosen randomly, it leads to improvements in the results.

I. CONCLUSION

In this paper, we presented a model-free approach to the fault diagnosis problem, based on a combination of different learning strategies like ANN, adaptive neuro-fuzzy and ANFIS. This model-free approach detects a presence of a possible fault from the profiles of the sensor outputs. Changes in the fault signatures such as settling time, and the steady-state value, give a quick indication that a fault may be in the making. An abrupt change in the sensor output profile indicates a possible onset of a fault. As such, this model-free approach can be made an effective part of an overall integrated approach that tackles both fault detection and isolation where the isolation part would be handled by an additional section using a model-based approach.

ACKNOWLEDGMENT

The authors wish to acknowledge the support of KFUPM in carrying out this work under the grant of SB100017.

APPENDIX 1

The mathematical model of 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 system (See equation A1):

$$\dot{Q}_i = -a_m Q_i + b_m \phi(u) \quad \text{(A1)}$$

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. It is assumed that the leakage Q_ℓ occurs in tank 1 and is given by (See equation A2):

$$Q_\ell = C_{d\ell}\sqrt{2gH_1} \qquad \text{(A2)}$$

With the inclusion of the leakage, the liquid level system is modeled by (See equation A3 and A4):

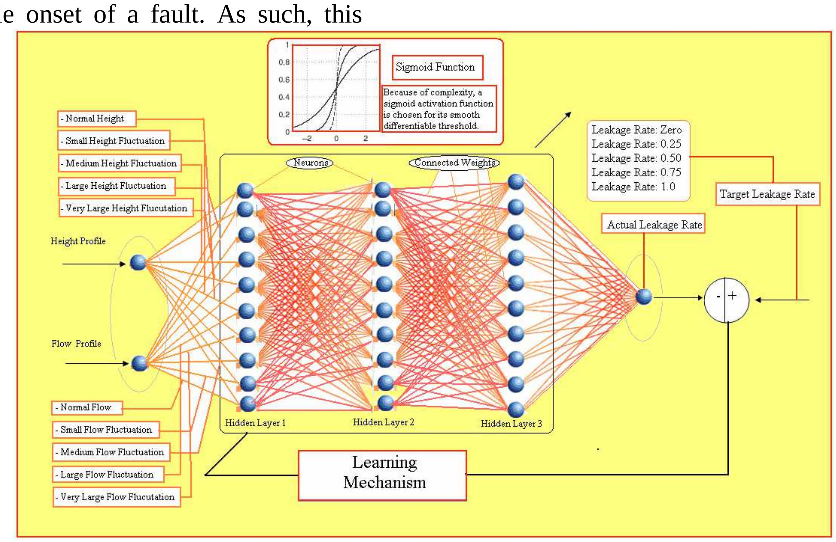


Figure 8: Evaluation of Neural Network-Based Fault Diagnosis

$$A_1\frac{dH_1}{dt} = Q_i - C_{12}\varphi(H_1 - H_2) - C_\ell\varphi(H_1) \quad \text{(A3)}$$

$$A_2\frac{dH_2}{dt} = C_{12}\varphi(H_1 - H_2) - C_0\varphi(H_2) \qquad \text{(A4)}$$

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 model of the two-tank fluid control system, shown above in Fig. 2, is of a second order and is nonlinear with a smooth square-root type of nonlinearity. For design purposes, a linearized model of the fluid system is required and is given below in (A5) and (A6):

$$\frac{dh_1}{dt} = b_1 q_i - (a_1 + \alpha)h_1 + a_1 h_2 \qquad \text{(A5)}$$

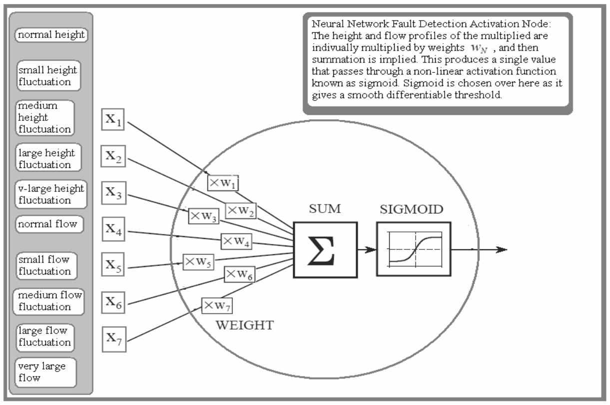


Figure 9: Neural Network Based Activation Function Evaluation

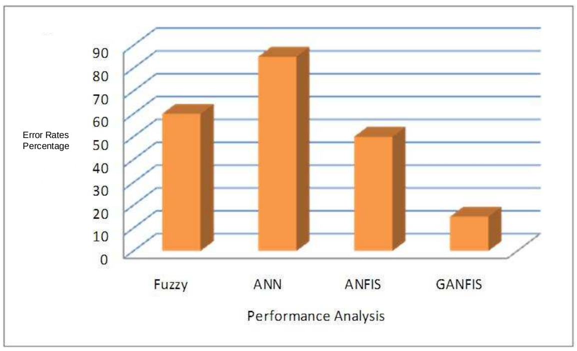


Figure 10: Comparison of error rates

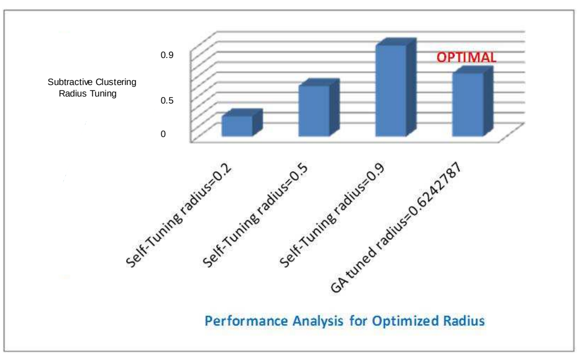


Figure 11: Comparison of subtractive clustering radius tuning

$$\frac{dh_2}{dt} = a_2 h_1 - \left(a_2 - \beta\right) h_2 \quad \text{(A6)}$$

Where h_1 and h_2 are the increments in the nominal (leakage-free) heights H_1^0 and H_2^0:

$$b_1 = \frac{1}{A_1}, \quad a_1 = \frac{C_{db}}{2\sqrt{2g(H_1^0 - H_2^0)}}, \quad \beta = \frac{C_0}{2\sqrt{2gH_2^0}},$$

$$a_2 = a_1 + \frac{C_{do}}{2\sqrt{2gH_2^0}} \quad \alpha = \frac{C_{d\ell}}{2\sqrt{2gH_1^0}}$$

and the parameter α indicates the amount of leakage rate, where q_i, q_ℓ, q_0, h_1 and h_2 are the increments in Q_i, Q_ℓ, Q_o, H_1^0 and H_2^0, respectively, the parameters a_1 and a_2 are associated with linearization whereas the parameters α and β are respectively associated with the leakage and the output flow rate, i.e. $q_\ell = \alpha h_1$, $q_o = \beta h_2$.

REFERENCES

[1] Simani, S., Fantuzzi, C., Patton, R.J., 2003. Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques. Springer, London.

[2] Isermann, R., 2005. Model-based fault-detection and diagnosis—status and applications. Annual Reviews in Control 29, 71–85.

[3] Chow, E.Y., Willsky, A.S., 1984. Analytical redundancy and the design of robust failure detection systems. IEEE Transactions on Automatic Control 29 (7), 603–614.

[4] Gertler, J.J., 1998. Fault Detection and Diagnosis in Engineering Systems. Marcel Dekker, New York.

[5] Frank, P.M., Ko¨ ppen-Seliger, B., 1997. New developments using AI in fault diagnosis. Engineering Applications of Artificial Intelligence 10 (1), 3–14.

[6] Frank, P.M., Ding, S.X., Ko¨ ppen-Seliger, B., 2000. Current developments in the theory of FDI. Proceedings of IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, vol. 1, Budapest, Hungary, pp. 16–27.

[7] Haris M. Khalid, R. Doraiswami and L. Cheded, "Intelligent Fault Diagnosis using a Sensor Network", ICINCO, Milan, Italy, July 2-5, 2009.

[8] Rajamani Doraiswami, Lahouari Cheded, Haris M. Khalid, Qadeer Ahmed and Amar Khoukhi, "Robust Control of a Closed loop identified system with parametric/model uncertainties and external disturbances ", International Conference on Systems, Modeling and Simulations, January 27-29, 2010, Liverpool, UK.

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

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

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

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

[13] 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)

[14] Zhang, G.Q.P., 2000. Neural networks for classification: a survey. IEEE Transactions on Systems, Man and Cybernetics: Part C—Applications and Reviews 30 (4), 451–462.

[15] A. Khoukhi, S. Albukhitan, “*PVT Properties Prediction Using Hybrid Genetic Neuro*□*Fuzzy Systems*”, ”, International Journal of Oil Gas and Coal Technology, vol. 4, n0 1, 2011 pp: 47-63

[16] A. Khoukhi: “*Data-Driven Multi-Objective Motion Planning of Parallel Kinematic Machines*”, IEEE Trans. On Control System Technology, Vol. 18, No. 6, Nov. 2010 pp. 1381-1389,

[17] A. Khoukhi, K. Benfreha: “*Application of multi-layered neuro-fuzzy networks to collision detection in CAD-CAM processes*”, Trans. of the Canadian Society of Mech. Eng. Vol. 28, (3-4), (2004) pp. 431-443.

[18] Wang, Y., Chan, C.W., Cheung, K.C., 2001b. Intelligent fault diagnosis based on neuro-fuzzy networks for nonlinear dynamic systems. Proceedings of IFAC Conference on New Technologies for Computer Control 2001, Hong Kong, China, pp. 101–104.

[19] de Miguel, L.J., Bla´ zquez, L.F., 2005. Fuzzy logic-based decision-making for fault diagnosis in a DC motor. Engineering Applications of Artificial Intelligence 18 (4), 423–450.

[20] Mo-Yuen Chow, “Special Section on Motor Fault Detection and Diagnosis”, IEEE Transaction on Industrial Electronics. 47 (5) (2000) 982–1107.

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

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

[23] 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).

[24] 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).

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

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

[27] 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).

[28] 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).

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

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

[31] 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).

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

[33] 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).

[34] 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).

[35] 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).

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

[37] 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).

[38] 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

[39] 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).

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

[41] 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).

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

[43] K.A. De Jong and W.M. Spears, Using genetic algorithms to solve NO-complete problems, in Proceedings of the Third International Conference on Genetic Algorithms, ed. By J. D. Schaffer. Morgan Kaufmann: 1989.

[44] D.E. Goldberg, K. Zakrzewski, B. Sutton, R. Gadient, C. Chang, P. Gallego, B. Miller and E. Cantu-Paz, Genetic Algorithms: A Bibliography, illegal Report no. 97011, Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign

[45] Forssell, U., & Ljung, L. (1999). Closed Loop Identification Revisited. *Automatica*, 1215-1241.