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

Haris M. Khalid
Systems Engg. Department
K.F.U.P.M
Dhahran, Saudi Arabia
Email: mhariskh@kfupm.edu.sa

Amar Khoukhi
Systems Engg. Department
K.F.U.P.M
Dhahran, Saudi Arabia
Email: amar@kfupm.edu.sa

Fouad M. Al-Sunni
Systems Engg. Department
K.F.U.P.M
Dhahran, Saudi Arabia
Email: alsunni@kfupm.edu.sa

***Abstract* - In this paper, an efficient scheme to detect the unprecedented changes in system reliability and find the failed component state by classifying the faults is proposed using kalman filter and hybrid neuro-fuzzy computing techniques. 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 Adaptive 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.**

Keywords - Kalman filter, hybrid neuro-fuzzy, 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 become more and more popular in industrial applications of fault diagnosis. Process faults, if undetected, have a serious impact on process economy, product quality, safety, and productivity and pollution level. In order to detect, diagnose and correct these abnormal process behaviors, efficient and advanced automated diagnostic systems are 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 break down of the system after the occurrence of faults. This will improve the reliability and safety of the system, and avoid unnecessary and costly stoppages. Complete reliance on human operators to monitor the conditions of the systems is often difficult, especially as engineering systems are becoming more complex.

For example, in chemical processes several kinds of failures may compromise safety and productivity. In fact, the occurrence of faults may affect efficiency of the process (e.g., lower product quality) or, in the worst scenarios, could lead to fatal accidents (e.g., temperature run-away) with injuries to personnel, environmental pollution, equipments 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, FDI techniques are 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 from zero of the residuals indicates a fault has likely occurred [4]. Several model-based works on the detection and identification of faults and tuning parameters were considered in [7-8]. 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 in general are nonlinear and are becoming more complex [5][6].

To overcome these problems of fault diagnosis, soft computing is considered as an emerging approach, which parallels the remarkable ability of the human mind to reason and learn in circumstances characterized by uncertainty and imprecision. In contrast with hard computing methods that only deal with precision, certainty, and rigor, it is effective in acquiring imprecise or sub-optimal, but economical and competitive solutions to real-world problems. As we know, qualitative information from practicing operators may play an important role in accurate and robust diagnosis of motor faults at early stages.

The paper is organized as follows: Fault Diagnosis problem statement is considered in Section II. Section III discusses the implementation and simulation results. Section IV does the discussion for all the techniques implemented. Finally some conclusions are given Section V.

II. The Fault Diagnosis Problem Statement

Fault is an undesirable factor in any process control industry. It affects the efficiency of system operation and reduces economic benefit to the industry. The early detection and diagnosis of faults in mission critical systems becomes highly crucial for preventing failure of equipment, loss of productivity and profits, management of assets, reduction of shutdowns.

To have an effective fault diagnosis of highly non-linear systems, hybrid techniques have been introduced here by showing the genetic neuro-fuzzy Based- FDI as shown in figure 1.

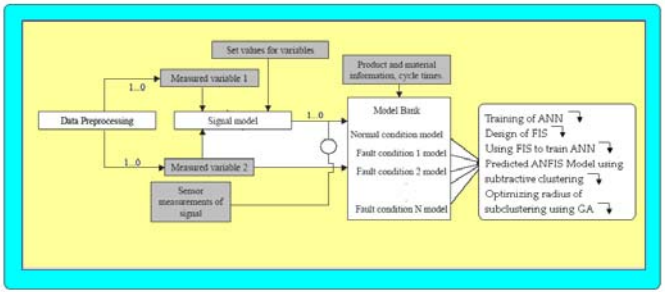


Figure 1. Implementation plan for the evaluation of the proposed scheme

A. System Description

A Benchmark laboratory-scale two-tank process control system has been used to collect data at a sampling rate of 50 milliseconds. The system is considered as a multi-input single-output (MISO) process with hydraulic height and liquid output flow-rate of the second tank being the two inputs while leakages fault level on a discrete scale of 1 to 4 being the output. The objective of the benchmark dual-tank system is to reach a reference height of 200ml in the second tank. To achieve this objective, a Proportional Integral (PI) controller works in a closed loop configuration. Data is collected by introducing leakage fault in the closed loop system. (See. Fig 2). This is done through the pipe clogs of the system using drainage valve between the two tanks. The PI controller tends to treat the introduced fault as a disturbance and acts to suppress it. The closed-loop nature of the experiment also tends to suppress the faults introduced in the system, thereby making it more difficult to detect these faults.

B. 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 3.

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.

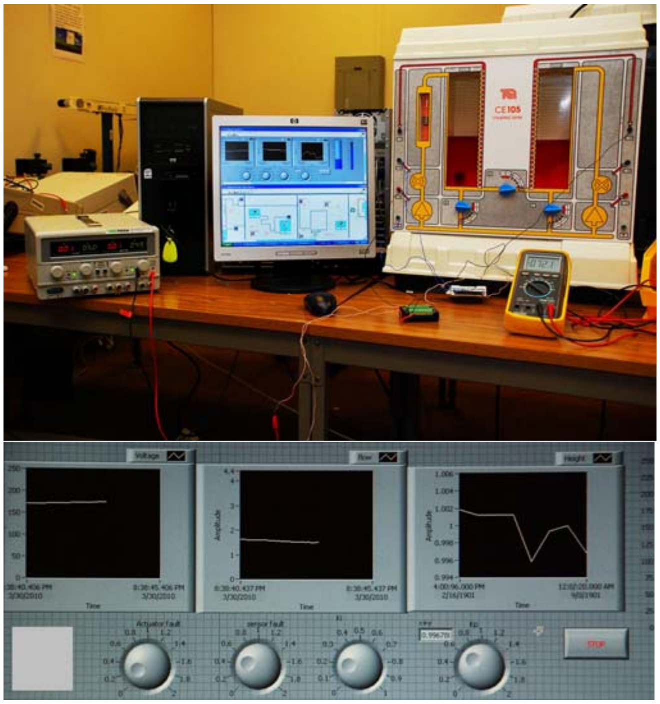


Figure 2 A – The two tank system interfaced with the Labview through a DAQ and the amplifier for the magnified voltage , B – The labview setup of the apparatus including the ciruit window and the block diagram of the experiment.

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:

$$\dot{Q}_i = -a_m Q_i + b_m \phi(u) \quad (1)$$

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:

$$Q_\ell = C_{d\ell}\sqrt{2gH_1} \quad (2)$$

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

$$A_1 \frac{dH_1}{dt} = Q_i - C_{12}\varphi\left(H_1 - H_2\right) - C_\ell \varphi\left(H_1\right) \qquad (3)$$

$$A_2 \frac{dH_2}{dt} = C_{12}\varphi\left(H_1 - H_2\right) - C_0\varphi\left(H_2\right) \quad (4)$$

where $\varphi(.) = sign(.)\sqrt{2g(.)}$, $Q_\ell = C_\ell \varphi\left(H_1\right)$ is the leakage flow rate, $Q_0 = C_0\varphi\left(H_2\right)$ 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. 3, 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 (5) and (6):

$$\frac{dh_1}{dt} = b_1 q_i - (a_1 + \alpha) h_1 + a_1 h_2 \quad (5)$$

$$\frac{dh_2}{dt} = a_2 h_1 - (a_2 - \beta) h_2 \quad (6)$$

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.

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 as:

$$\dot{x}_3 = e = r - h_2$$
$$u = k_p e + k_I x_3 \quad (7)$$

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

$$\dot{x} = Ax + Br \quad y = Cx \quad (8)$$

where

$$x = \begin{bmatrix} h_1 \\ h_2 \\ x_3 \\ q_i \end{bmatrix}, \quad A = \begin{bmatrix} -a_1 - \alpha & a_1 & 0 & b_1 \\ a_2 & -a_2 - \beta & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -b_m k_p & 0 & b_m k_I & -a_m \end{bmatrix}, \quad (9)$$

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

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. $q_\ell = \alpha h_1$, $q_o = \beta h_2$.

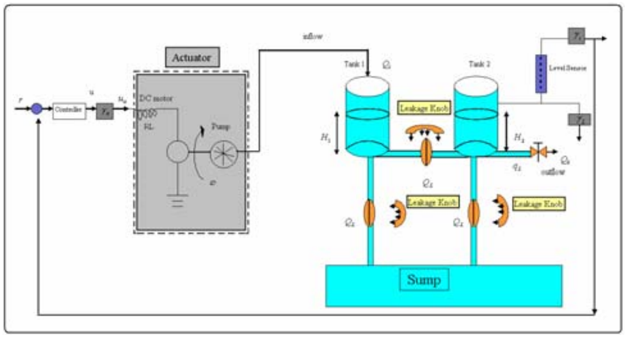


Fig. 3 Process control system: A Lab-scale two-tank system

III. Implementation and Simulation Results

A. Fault Detection Using Kalman Filter

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

$$x(k+1) = A_0 x(k) + B_0 u(k-d) + w(k)$$
$$y(k) = C_0 x(k) + \upsilon(k) \quad (7)$$

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:

$$Q = E\left[w(k)w^T(k)\right], \text{ and } R = E\left[v(k)v^T(k)\right] \quad (8)$$

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:

$$\hat{x}(k+1) = A_0\hat{x}(k) + B_0 u(k-d) + K_0\left(y(k) - C_0\hat{x}(k)\right)$$
$$e(k) = y(k) - C_0\hat{x}(k) \quad (9)$$

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

K_0 is the filter gain. The larger the 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.

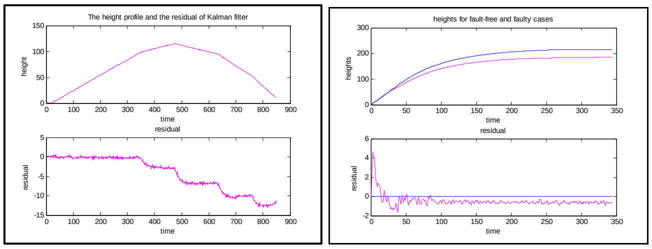
It is noteworthy that the fault-free model of the system is identified using a recursive least-squares identification scheme. The order of the estimated model was iterated to obtain an acceptable model structure using a combination of the AIC criterion and the identified pole locations.

The identified model is essentially a second-order system with a delay 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 .

$$\hat{x}(k+1) = A_0\hat{x}(k) + B_0 u(k-d) + K_0\left(y(k) - C_0\hat{x}(k)\right) \qquad (10)$$

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

The Kalman filter was evaluated under different fault scenarios for an on-off controller, a P controller, and a PI controller, as shown in Fig.4.



(a) (b)

Fig. 4. Kalman filter results for an (a) On-Off and (b) PI Controller: for Flow and Height under various leakage magnitudes

IV. ANFIS BASED FAULT DIAGNOSIS USING SUBTRACTIVE CLUSTERING

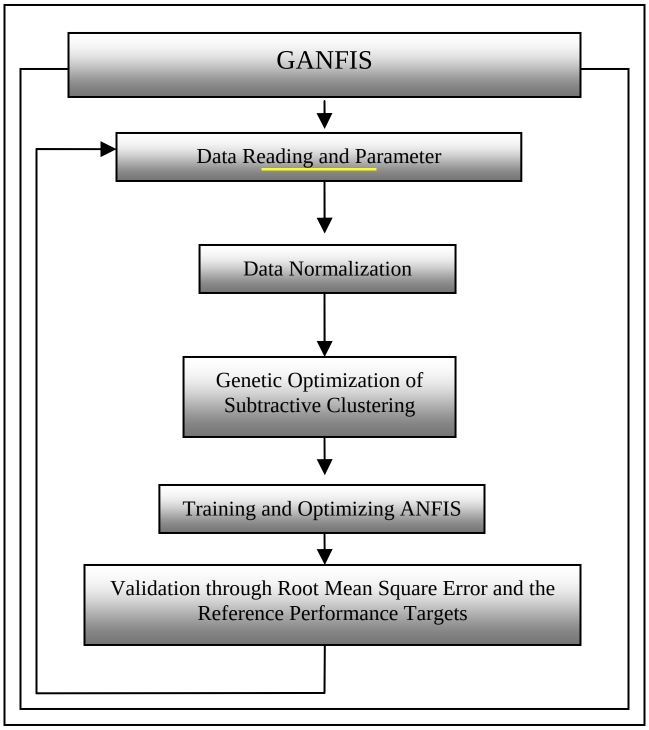


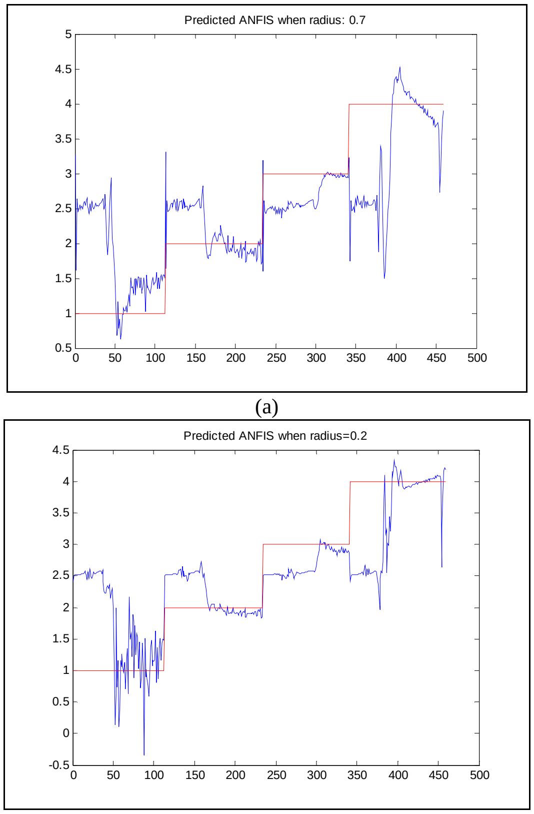
Figure 5. Implementation Scheme (GANFIS system)

The tasks of our fault diagnosis scheme, GANFIS system (Fig.5) are executed with an increasing precision accompanied with a more detailed fault picture. Firstly, the data collected from the plant has been normalized which comprises of the pre-processing of the data. Then, the optimal cluster has been tested through ANFIS using the subtractive clustering technique. Then, the genetic optimization of the subtractive clustering radius has been performed and the performance has been validated by checking the root mean square error and the performance targets of the performance targets.

A.

The Subtractive Clustering technique has been applied here in order to form hybrid versions of Neuro-Fuzzy. The procedure for the Subtractive Clustering proceeds by defining a cluster center based on the density of surrounding data points. All the data points within the RADII of this point form then a cluster. This process is repeated until all the data is clustered.

This scheme has been followed and employed to get a final trained ANFIS. It has been shown that the Predicted ANFIS is performing better in following the Original output when the radius is 0.7 rather than when the radius is 0.2 as shown in the figure 6.



(b)

Figure 6. (a)Predicted ANFIS using Subtractive Clustering when radius: 0.7, (b) Predicted ANFIS using Subtractive Clustering when radius: 0.2

As can be seen in the previous section that a slight change in the radius of subtractive clustering has showed in the change of results. So, the genetic optimization of the parameters is used in order to get the optimal shape of the fault prediction. Figure 7 shows the four various phases in which the performance of genetic optimization can been seen. The original graph is drawn in red which is showing step sizes indicating different levels of faults. The green graph is showing 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.

These functions when implemented in the genetic algorithm gives the best fitness function value as follows:

Fitness Function Value:0.187462

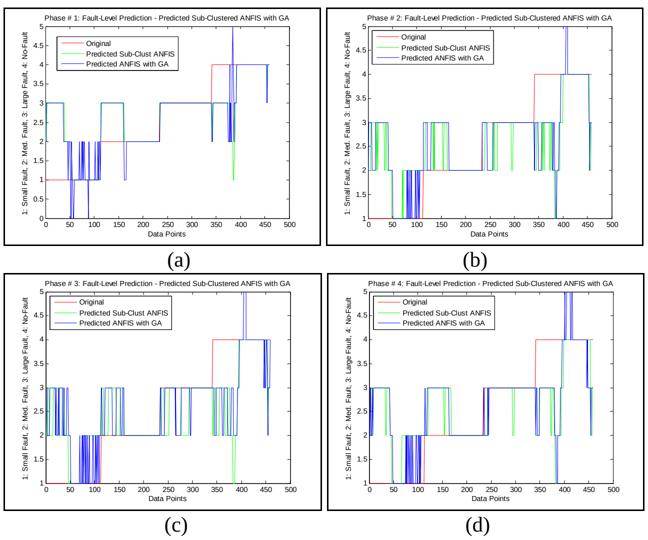


Figure 7. (a-d) Four random Phases for predicted subclustered ANFIS with GA

B. ANN-Based Fault Diagnosis

The analysis of the ANN is a difficult task and it requires an expert opinion with a hit and trial scenario by putting different types training sets and functions and finding the final outcome of the best possible hidden layers and activation layers as per according to the problem at hand. A generic model of the ANN in fault diagnosis is as follows in figure 8. and a generic activation evaluation on a sigmoid activation function is shown in figure 9.

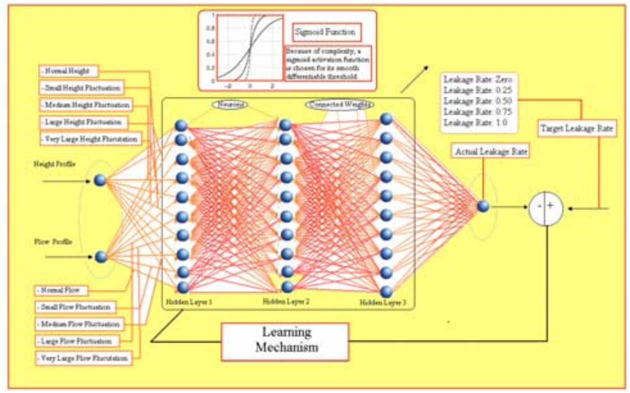


Figure 8. Evaluation of Neural Network-Based Fault Diagnosis

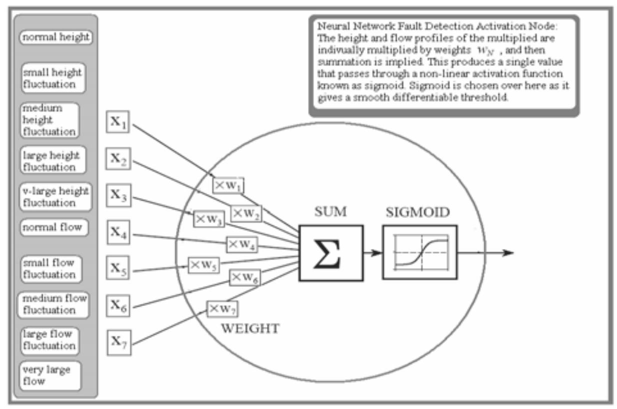


Figure 9. Neural Network Based Activation Function Evaluation

C. Discussion

In this paper, two modeling techniques have been used, a kalman filter-based approach for fault detection and the hybrid techniques have been implemented for the fault classification. A good comparison of the techniques can been in the histographs shown in figure 9 and figure 10. The chart in figure 10 shows the comparison with the error rates between fuzzy, ANN, ANFIS and the present method of genetic neuro-fuzzy. It is to no worthy that the error rate for the GANFIS is the least one because the genetic algorithm has performed well in optimization the subtractive clustering. In figure 11, it can be seen, that when the radius of the subtractive clustering chosen randomly, it is showing improvements in the results.

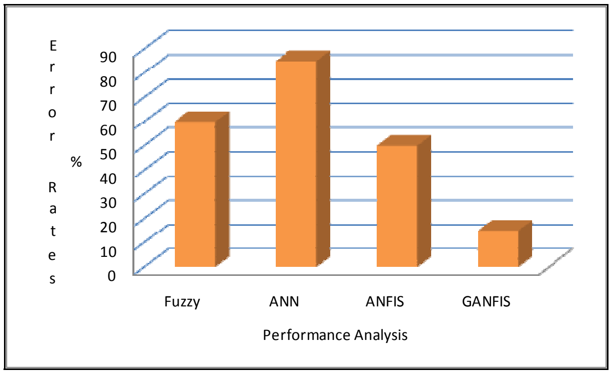


Figure 10. Comparison of error rates

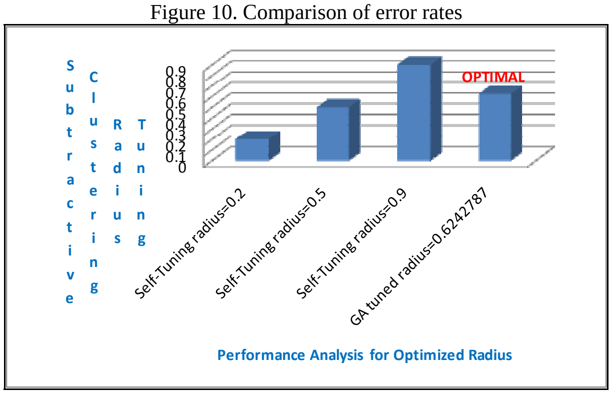


Figure 11. Comparison of subtractive clustering radius tuning

V. 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, steady-state value, and the coherence spectral changes 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 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.

REFERENCES

[1] J. C. Ag¨uero. System identification methodologies incorporating Constraints PhD thesis, School of Electrical Engineering and Computer Science, The University of Newcastle, Australia, 2005.

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

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

[4] 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.

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

[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.