15th Triennial World Congress, Barcelona, Spain

NEURO-FUZZY BASED FAULT DIAGNOSIS APPLIED TO AN ELECTRO-PNEUMATIC VALVE

Faisel J Uppal, Ron J Patton and Vasile Palade

Control & Intelligent Systems Engineering, University of Hull
Hull, HU6 7RX United Kingdom
E-mail: f.j.uppal@hull.ac.uk, r.j.patton@hull.ac.uk,
v.palade@eng.hull.ac.uk

Abstract: The early detection of faults (just beginning and still developing) can help avoid system shutdown, breakdown and even catastrophes involving human fatalities and material damage. Computational intelligence techniques are being investigated as extension of the traditional fault diagnosis methods. This paper discusses the properties of the TSK/Mamdani approaches and neuro-fuzzy (NF) fault diagnosis within an application study of an electro-pneumatic valve actuator in a sugar factory. The key issues of finding a suitable structure for detecting and isolating ten realistic actuator faults are described. *Copyright © 2002 IFAC*

Keywords: Neuro-fuzzy networks, Fault Diagnosis, Fault Isolation, Fuzzy Hybrid Systems, Actuators

1. INTRODUCTION

The mathematical model used in the traditional Fault Detection and Isolation (FDI) methods is sensitive to modelling errors, parameter variation, noise and disturbance. Process modelling has limitations, especially when the system is complex and uncertain and the data are ambiguous i.e. not information rich.

Computational Intelligence (CI) methods (Neural Networks (NN), Fuzzy Logic (FL), Evolutionary Algorithms (EA) are known to overcome some of the above mentioned problems (Patton *et al* 2000). Neural networks are known to approximate any non-linear function, given suitable weighting factors and architecture. Moreover, on-line training makes it possible to change the FDI system easily in cases where changes are made in the physical process or the control system. NN can generalise when presented with inputs not appearing in the training data and make intelligent decisions in cases of noisy or corrupted data. However, the NN operates as a "black box" with no qualitative information available of the model it represents (Patton, 1994).

Fuzzy logic systems on the other hand have the ability to model a non-linear system and to express it in the form of linguistic rules making it more transparent i.e easier to interpret. They also have inherent ability to deal with imprecise or noisy data therefore making them suitable for fault diagnosis (Dexter, 1995). Neuro-fuzzy (NF) model is a combination of neural network and fuzzy logic to exploit the learning ability of NN and the reasoning ability of FL.

This paper provides a tutorial study of the use of NF structure identification and clustering methods with application to a non-linear model of an electro-pneumatic valve system. It is well known that for the non-linear systems the problem of discriminating between uncertain model behaviour and faults present a significant challenge. This paper describes a multiple-model strategy, taking care of multiple operating points through the NF modelling framework. Section 2 outlines the value of the NF approach to modelling, whilst Section 3 enters more into details of structure identification prior to describing the NF model construction strategy in

Section 4. The remainder of the paper is concerned with the fault diagnosis application problem.

2. WHAT IS A NEURO-FUZZY MODEL?

The NF model combines, in a single framework, both numerical and symbolic knowledge about the process. Automatic linguistic rule extraction is a useful aspect of NF especially when little or no prior knowledge about the process is available (Brown and Harris, 1994; Jang, 1995). For example, a NF model of a non-linear dynamical system can be identified from the empirical data. This model can give us some insight about the non-linearity and dynamical properties of the system.

The most common NF systems are based on two types of fuzzy models TSK (Takagi and Sugeno, 1985; Sugeno and Kang, 1988) and Mamdani (1975, 1976) combined with NN learning algorithms. TSK models use local linear models in the consequents, which are easier to interpret and can be used for control and fault diagnosis (Füssel, *et al* 1997; Ballé *et al* 1997). Mamdani models use fuzzy sets as consequents and therefore give a more qualitative description. Many NF structures have been successfully applied to a wide range of applications from industrial processes to financial systems, because of the ease of rule base design, linguistic modelling, application to complex and uncertain systems, inherent non-linear nature, learning abilities, parallel processing and fault-tolerance abilities. However, successful implementation depends heavily on prior knowledge of the system and the empirical data (Ayoubi, 1995).

NF networks by their intrinsic nature can handle a limited number of inputs. Moreover, NF models usually identified from the empirical data are not very transparent. Transparency in this context means that a more meaningful description of the process i.e. less rules with appropriate membership functions to be obtained. In ANFIS (Jang, 1993, 1995) a fixed structure with grid partition is used. Antecedent and consequent parameters are identified by a combination of least-squares estimates and gradient based methods, known as *the hybrid learning rule*. This method is fast and easy to implement for low dimensional input spaces. It is more prone to losing the transparency and the local model accuracy because of the use of *error back-propagation* that is a global and not locally non-linear optimisation procedure. One possible method to overcome this problem can be to find the antecedents & rules separately e.g. using clustering and constrain the antecedents prior to applying optimisation methods.

Hierarchical NF networks can be used to overcome the dimensionality problem by decomposing the system into a series of MISO and/or SISO systems called *hierarchical systems* (Tachibana and Furuhashi, 1994). The local rules use subsets of input spaces and are activated by higher level rules.

The criteria on which to build a NF model are based on the requirements for fault diagnosis and the system characteristics. The function of the NF model in the FDI scheme is also important i.e. Pre-processing data, Identification (Residual generation) or classification (Decision Making/Fault Isolation). For example a NF model with high approximation capability and disturbance rejection is needed for identification so that the residuals are more accurate. Whereas in the classification stage, a NF network with more transparency is required.

3. STRUCTURE IDENTIFICATION OF NF MODELS

For complexity reduction and transparency, *Structure Identification* methods can be applied to find appropriate input partition, rules & membership functions (MFs). Methods like Evolutionary Algorithms (EA), Classification and Regression Trees CART (Jang, 1994), Clustering and unsupervised NN (e.g. like the Kohonen feature maps) can be used. Once the structure is determined i.e. the rules and input membership functions, the consequent parameters can be identified by optimisation techniques like *Least-Squares Estimation*. The Product Space Clustering approach can be used (Babuska, 1998) for structure identification of TSK & Mamdani fuzzy models. For a MISO non-linear dynamic system with p inputs, the Product space $(X \times Y) \subset \Re^{p+1}$ is divided into subspaces in which linear models can approximate the non-linear system. The locally linear model tree LOLOMOT algorithm developed by Nelles can be used to identify a TSK fuzzy model with dynamic linear models as consequents. When using such structure identification techniques, a major issue is the sensitivity to uneven distribution of data. For example in most clustering algorithms, more clusters are created in regions with more data. A possible solution to this is problem may be to initialise the algorithm with large number of clusters.

Transparency of the NF models can be enhanced by tuning rules & MFs (Babuska, 1998). This approach is referred to as *structure simplification/ optimisation techniques*. To find the optimal number of rules, different cluster validity measures and methods like Compatible Cluster Merging *CCM* (Krishnapuram and Freg, 1992) can be used. At NF model level the rules are further simplified by merging similar fuzzy sets & removing fuzzy sets similar to the universal set. Setnes *et al*., (Setnes and Kaymak 1998) used a supervised fuzzy clustering algorithm that uses input-output data, orthogonal techniques and tuning for complexity reduction.

4. CONSTRUCTION OF NEURO-FUZZY MODELS

Two major classes of knowledge representation in fuzzy modelling are proposed by Takagi and Sugeno (Takagi and Sugeno, 1985) and Mamdani (Mamdani *et al* 1976).

In the linguistic or Mamdani fuzzy model both antecedents and consequents are linguistic fuzzy sets.

This model is mainly used to give a more linguistic description of the process

This fuzzy model can be represented as a multi-layer NF network (Fig.1) in which the input/output membership functions, rules, *normalisation* and *de-fuzzification* stages are expressed as neuron layers.

The i_{th} input fuzzy set in Layer-1 can be described by a Gaussian membership function with centres m_i and spread σ_i.

$$\mu_{Ai}(x) = \exp\left[-\left\{\frac{(x-m_i)^2}{\sigma_i}\right\}\right] \qquad (1)$$

The second layer consists of rule neurons. The firing strength of each rule is given by:

$$\alpha = \min([\mu_{A1}(x), \mu_{An}(x), \ldots \mu_{An}(x)]) \qquad (2)$$

or

$$\alpha = \prod \mu_A(x) \qquad (3)$$

Applying COA de-fuzzification (Layers-3,4,5) the output y can be calculated as (W Hauptmann and K Heesche 1995):

$$y = \frac{\sum_k M_k \cdot \sum_j (w_{kj} \cdot \alpha_j)}{\sum_k A_k \cdot \sum_j (w_{kj} \cdot \alpha_j)} \qquad (4)$$

where M_k and A_k are the moments and areas of the k_{th} output membership function and

w_{kj} = 1 if partial connection exists

0 otherwise

The rules can be expressed as:

R_1: if x_1 is A_1 and x_2 is B_1... then y is C_1
R_2: if x_1 is A_2 and x_2 is B_2... then y is C_2
…
R_n: if x_1 is A_n and x_2 is B_n... then y is C_n

Where (A_1,A_2…A_n, B_1,B_2…B_n,…) are the input fuzzy sets and (C_1,C_2…$C_{n\ldots}$) are the output fuzzy sets. The number of trainable parameters is usually large and some structure identification is needed to find the optimal network i.e. the number of rules, shape and position of the membership functions is determined. Evolutionary algorithms gradient based algorithms can be used for training. This kind of network is more transparent and close to human reasoning but the complexity is high and the training is difficult.

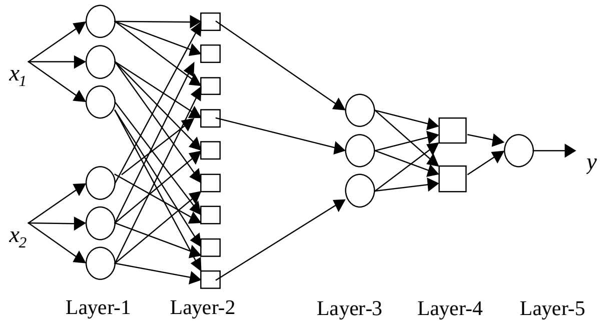

Fig. 1 Structure of Mamdani NF model.

TSK models with linear functions as consequents can be expressed as a non-linear function approximation by local linear models selected by fuzzy rules. In the TSK NF system (Fig. 2) the antecedents are similar to the Mamdani NF structure. The consequents can be any function describing the response of the model within the fuzzy region.

The same approach is applied as in the previous network, to express the TSK fuzzy system as a neural network. Fig.2 shows an example of such a network with two inputs, one output and two rules. Antecedent and consequent parameters can be optimised by Gradient-based optimisation algorithms. Hybrid learning rule by Jang (1993) can be applied for faster learning.

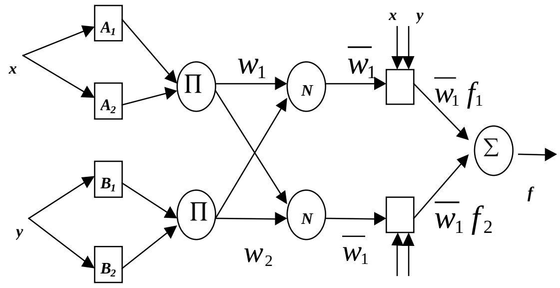

Fig. 2 Structure of TSK NF model.

The output of this model can be expressed as:

$$f = \frac{w_1 f_1 + w_2 f_2}{w_1 + w_2} = \overline{w}_1 f_1 + \overline{w}_2 f_2 \qquad (5)$$

where f_1 and f_2 are the outputs of the two sub-models ($f_I = a_i x + b_i y + c_i$; $i=1,2$) in which a_i,b_I *and* c_i are the linear parameters of i_{th} linear model. The rules are in the following form:

R_1: if x is A_1 and y is B_1 then f_1
R_2: if x is A_2 and y is B_2 then f_2

TSK fuzzy models are suitable for accurate modelling and identification but are less transparent than Mamdani models.

5. NF BASED FAULT DETECTION AND ISOLATION

Fig. 3 describes a FDI scheme in which several NF models are constructed to identify the faulty & the fault-free behaviour of the system.

$$r_i(k) = f\begin{pmatrix} u(k), u(k-1), \ldots u(k-n_u), \\ y(k), y(k-1), \ldots y(k-n_y) \end{pmatrix}; i = 1 \ldots n \qquad (6)$$

Each residual r_i in (6) is ideally sensitive to one particular fault in the system. In practice however, as a consequence of noise and disturbances, residuals are sensitive to more than one fault.

To take into account the sensitivity of residuals to various faults and noise we apply a NF classifier. A linguistic style (Mamdani) NF network is used which processes the residuals to indicate the fault.

This NF model is constructed with following set of rules:

If r_1 is small … r_r is large … r_n is small then $fault_r$ is large

Fuzzy threshold evaluation (7) is employed to take into account the imprecision of the residual generator at different regions in the input space.

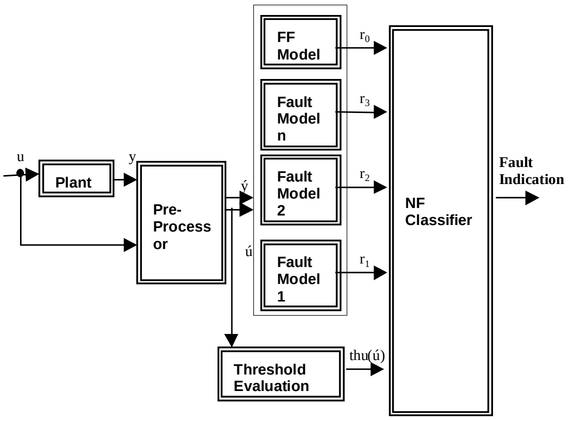

Fig. 3 Neuro-fuzzy based FDI scheme.

$$th_v(u) = \frac{\sum_{i=1}^{C} th_i \mu_i(u)}{\sum_{i=1}^{C} \mu_i(u)} \tag{7}$$

C is the total number of I/P regions with different sensitivity to faults and a multidimensional fuzzy set μ_i defines the fuzzy boundary of i_{th} such region. This approach depends heavily on the availability of the faulty and fault-free data and it is more difficult to isolate faults that appear in the dynamics.

Residuals can also be generated by a non-linear dynamic model of the plant that approximates a non-linear dynamic system by local linear models. Such a model can be obtained by *Product space clustering* (Babuska 1998), or tree-like algorithms (LOLIMOT algorithm by Nelles, 1996). Each local model is a linear approximation of the process in an I/P subspace and the selection of the local model is fuzzy. The output of such a model can be described by:

$$y = \frac{\sum_{i=1}^{C} \alpha_i(u_s) \cdot f_i}{\sum_{i=1}^{C} \alpha_i(u_s)} \tag{8}$$

where f_i is the i_{th} local linear model given by:

$$\begin{aligned} f_i &= b_{i,1}u(k) + b_{i,2}u(k-1) + \cdots + b_{i,nu}u(k-n_u) + \\ & a_{i,1}y(k) + a_{i,2}y(k-1) + \cdots + a_{i,ny}y(k-n_y) + c_i \end{aligned} \tag{9}$$

a_i , b_i and c_i are the parameters of the i_{th} model, u_s is the I/P subspace defining the operating point, α_i is the degree to which the i_{th} local model is valid at this operating point.

From a_i , b_i and c_i physical parameters like time constants, static gains, offsets etc (Füssel, 1997) can be extracted for each operating point and can be compared with the parameters estimated online. This approach heavily depends on the accuracy of the non-linear dynamic model described above. Also the output error should be minimum when operated in parallel to the system. Moreover, this method requires that there is sufficient excitation at each operating point for online estimation of parameters.

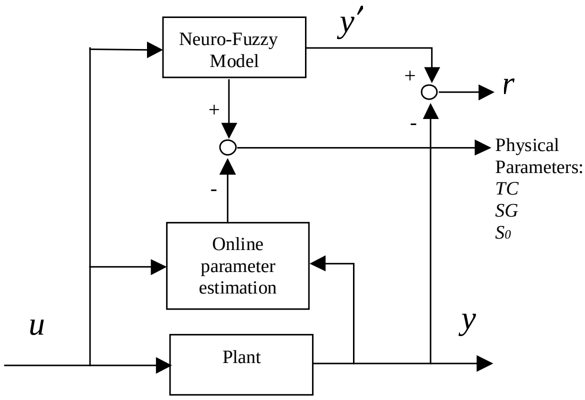

Fig. 4. TSK NF based FDI scheme.

6. CASE STUDY: FDI OF AN ELECTRO-PNEUMATIC VALVE ACTUATOR

The valve considered for FDI is an electro-pneumatic flow controller in the evaporisation stage of a sugar factory. A non-linear mathematical model of the valve has been constructed using SIMULINK and MATLAB. The model is then used to generate faulty/ fault-free data to evaluate the Neuro-fuzzy based fault isolation schemes presented in the previous sections. The whole valve assembly consists of 3 main parts (Fig. 5):

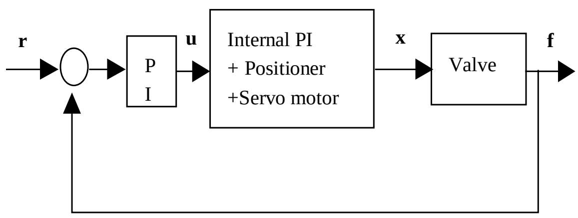

Fig. 5 Main parts of the valve assembly.

The PI controller controls the Positioner & Actuator output to regulate the flow through the valve.

Positioner and Actuator: Pneumatic pressure is applied to the servomotor diaphragm to control the stem position that changes the flow. The Positioner adjusts this pressure input to the servomotor to obtain the correct stem position of the actuator (See Fig. 6).

The Valve is the final element in the assembly that alters the flow according to the valve stem position (See Fig. 6).

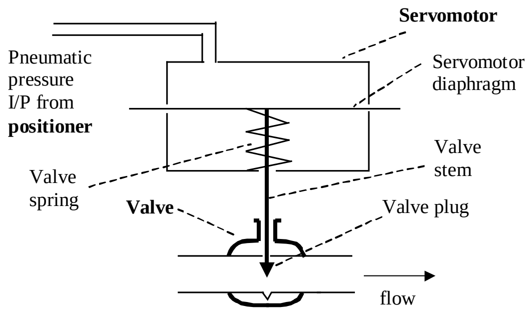

Fig. 6. The Servomotor and valve assembly.

The following list of faults is considered in the valve actuator assembly:

f_1 - External PI controller proportional gain fault
f_2 - External PI controller integral gain fault
f_3 - Increased friction of the servomotor
f_4 - Decreased elasticity of the servomotor
f_5 - Decrease of pneumatic pressure
f_6 - Internal PI controller fault
f_7 - Internal position sensor fault
f_8 - Valve clogging
f_9 - Valve leakage
f_{10} - Choked flow

Two neuro-fuzzy models are used here with transparent structure. A TSK structure with linear dynamic models as consequents is used to approximate the internal PI controller, the Positioner and servomotor. This system non-linearity is considered to be mainly in the dynamics i.e. a transparent TSK model is ideal for this case. The TSK model identified has three locally linear models as consequents. The time constants of these local models are 18sec, 12sec and 8sec, respectively which show that the system is faster at high values of flow and slower at the low values.

Figure 7 shows the performance of the TSK model in closed-loop, parallel to the system.

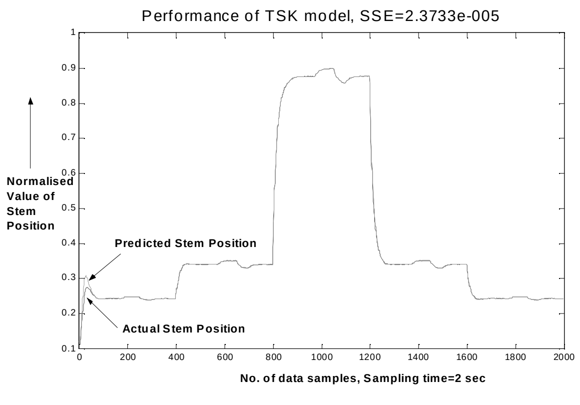

Fig. 7 Performance of the TSK model.

From the local models, and *RLSE* the changes in physical parameters e.g. time constant (r_{TC}), static gain (r_{SG}), static (r_{S0}) offset and settling time (r_{ST}) are computed. These changes are the residuals that can be used for fault isolation.

A Linguistic/Mamdani NF model is identified to approximate the valve. The model input is the valve stem position *x* and the output is the volumetric flow rate *f*. From the input set-point flow and measured flow, and by integrating and using *RLSE*, the control input u can be predicted. The GK-clustering algorithm (Gustafson & Kessel 1979) is used to partition the input space (Fig. 8), where clusters are projected onto the I/O space to find MF's. A gradient-based optimisation method is then used to fine-tune the MFs (Fig. 9).

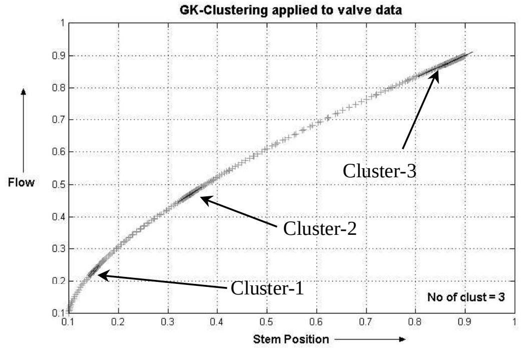

Fig. 8. Valve data clustered in three groups.

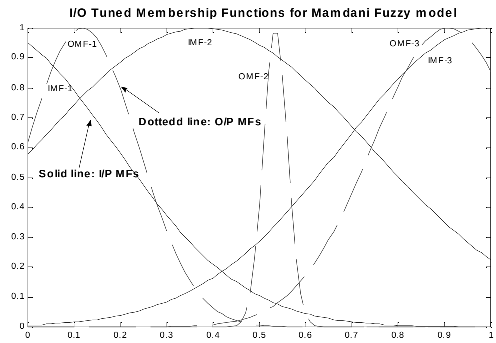

Fig. 9 Tuned I/O MFs for Mamdani model.

Table 1: Fault Isolation

	f_1	F_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	F_{10}
r_u	Op + Cl -	Op - Cl +	~	~	~	~	~	~	~	~
r_x	~	~	~	~	~	~	Ch	~	~	~
r_f	~	~	~	~	~	~	~	Op + Cl -	Op - Cl +	-
r_{st}	~	~	+	+	0	-	~	~	~	~
r_{tc}	~	~	+	-	+	-	~	~	~	~
r_{s0}	~	~	0	0	0	+	~	~	~	~

Op+ : Positive value when valve is being opened
Op- : Negative value when valve is being opened
Cl+ : Positive value when valve is being closed
Cl- : Negative value when valve is being closed
Ch : Changed

The predicted values u, x, f and the measured values are used to generate the residuals r_u , r_x , r_f. The fault isolation table given in Table-1 shows that some faults could only be detected during the time when the valve is being opened and closed. Moreover, choked flow could only be detected at high values of flow.

7. CONCLUSIONS

Neuro-fuzzy systems not only have powerful approximation abilities for modelling unknown dynamic non-linear systems, but a high level language description of the system can also be obtained. The transparent structure of NF is very useful to study the effect of faults on system characteristics. In this work different approaches for NF based fault diagnosis are studied. An approach is presented which uses TSK and Mamdani NF models to generate residuals. For structure identification GK-Clustering algorithm is used and ten realistic faults are diagnosed in the electro-pneumatic valve actuator model. The main challenges of NF based FDI methods are to minimise false alarms enhance detectability and isolability and minimise detection time by hardware implementation.

8. ACKNOWLEDGEMENTS

This work has been conducted through funding from the Framework 5 RTN *DAMADICS* in collaboration with the Lublin sugar factory and Warsaw University of Technology. Faisel Uppal acknowledges funding support through a CVCP ORS (Overseas Research Scholarship) award, together with the Hull University's Open Scholarship.

REFERENCES

Ayoubi M. (1995). Neuro-fuzzy Structure For Rule Generation And Application in The Fault Diagnosis of Technical Processes, *Proc. American Control Conference, Washington*, 2757-2761.

Babuska R. (1998). *Fuzzy modelling in control*, Kluwer Academic Publishers.

Ballé P., Nelles O. and Füssel D. (1997). Fault detection for nonlinear processes based on local linear fuzzy models in parallel and series-parallel model. *Proc. of IFAC symposium on fault detection, supervision and safety for technical process safeprocess'97*, The University of Hull, Kingston Upon Hull, UK.

Brown M. and Harris C. J. (1994). *Neuro-fuzzy adaptive Modelling and Control*, Prentice-Hall.

Füssel D., Ballé P., and Isermann R. (1997). Closed loop fault diagnosis based on a non-linear process model and automatic fuzzy rule generation, *Proc. of IFAC symposium on fault detection, supervision and safety for technical process safeprocess'97*, The University of Hull, UK.

Dexter A. L. (1995). Fuzzy model based fault diagnosis, *IEE Proc. Control Theory Appl.*, vol. 142 (6), 545-550.

Gustafson D. E. and Kessel W. C. (1979). Fuzzy clustering with a fuzzy covariance matrix, in *Proc. IEEE Conf. Decision Contr.*, San Diego, CA, pp761-766.

Hauptmann W. and Heesche K. (1995). A neural net topology for bidirectional fuzzy – neuro transformation, Proc. of *IEEE Int.Conf. on fuzzy systems*, vol. 3. March 20-24, 1995, Yokohama, Japan.

Jang J. S. R. (1993). ANFIS: Adaptive network based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665-684.

Jang J.S. and Sun R. (1995). Neuro-fuzzy modeling and control. *Proc. IEEE*. vol. 83(3), pp 378-405.

Jang J. S. R (1994). Structure determination in fuzzy modelling: a fuzzy CART approach. In *Proceedings of IEEE International Conf. On Fuzzy Systems*, Orlando, Florida, June 1994.

Krishnapuram R. and Freg C. P. (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging, *Pattern Recognition* , 25(4), 385-400.

Mamdani E. and Assilian S. (1975). An experiment in linguistic synthesis with fuzzy logic controller, *Int. J. Man-Machine Studies*, vol. 7, no. 1, pp. 1-13, 1975.

Mamdani E. (1976). Advances in the linguistic synthesis of fuzzy controllers, *Int. J. Man-Machine Studies*, vol. 8, pp. 669-678.

Nelles, O. and R. Isermann (1996). Basis function networks for interpolation of local linear models. IEEE Conference on Decision and Control (CDC), Kobe, Japan, pp. 470–475.

Patton R. J., Uppal F. J. and Lopez-Toribio C. J. (2000). Soft Computing Approaches To Fault Diagnosis For Dynamic Systems: A Survey, *Proc. of 4th IFAC Symposium on Fault Detection Supervision and Safety for Technical Processes, Budapest*, 14-16 June 2000, vol 1, pp. 298-311.

Patton R. J., Frank P. M., Clark R. N. (2000). *Issues of Fault Diagnosis for Dynamic Systems*, Springer-Verlag London, ISBN 3-540-19968-3.

Patton R. J., Chen J. and Siew T. M. (1994). Fault Diagnosis in Non-linear dynamic systems via neural-networks, *Proc. IEE Int. Conference Control '94*, Coventry, UK, 2, 1346-1351.

Setnes M. and Kaymak U. (1998). Extended fuzzy c-means with volume prototypes and cluster merging, In: *Proceedings EUFIT'98*, Aachen, Germany, Sept. 1998, pp. 1360-1364.

Sugeno M. and Kang G. (1988). Structure identification of fuzzy model, *Fuzzy Sets and Systems*, 26(1), 15-33.

Tachibana K. and Furuhashi T. (1994). A hierarchical fuzzy modelling method using genetic algorithm for identification of concise submodels, *Proc.* of *2nd Int. Conference on Knowledge-Based Intelligent Electronic Systems*, April 1994, Adelaide, Australia.

Takagi T. and Sugeno M. (1985). Fuzzy identification of systems and its applications to modelling and control, *IEEE Trans. Syst., Man, Cybern.*, vol. 15, 1985.