OPTIMIZATION OF PARAMETERS BY USING TAGUCHI'S DESIGN OF EXPERIMENTS APPROACH

B. Bratina[1], B. Tovornik[1]
[1]University of Maribor, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia

Corresponding Author: B. Bratina, University of Maribor, Faculty of Electrical Engineering and Computer Science, Smetanova 17, 2000 Maribor, Slovenia; bozidar.bratina@uni-mb.si

Abstract. The paper discusses applying Taguchi Design of Experiments approach for robust design i.e. selection of suitable parameters for various segments of fault detection and isolation (FDI) systems, to achieve better operation, e.g. less false alarms, quick response and ability to handle small faults (sensor bias, drift) under close loop conditions. Adequate system model should be reliable and exact to produce minimum number of false alarms and properly set up FDI scheme is important to operate under stable system control loop. In case of neural network models we try to find optimal structure and parameters of an auto-associate neural network dynamic model. In case of statistical multivariate methods such as principle components analysis, the number of principle components to be retained and not losing important properties can be a very subjective decision. In the field of observer based FDI, parameters can be set up to achieve better reconstruction of system outputs, thus less false alarms is produced, etc. Taguchi DoE of a neural network model will be discussed by simulations and real-time laboratory model results.

1 Introduction

Fault detection and isolation systems (FDI) are becoming widely used in many modern systems mainly due to evolved industrial IT equipment, which is raising the level of automation in various technical systems. First implementations of such systems were realized in economically justified systems e.g. nuclear power-plants, commercial airplanes, etc., however for smaller systems an economic perspective prevailed in order to prevent structural damages, plant down-times, predict maintenance intervals, etc. FDI systems offer operator support in decision making especially when quick reaction-time is expected to quickly deal with the issue causing unpredicted behaviour of the process. Poorly designed FDI system can cause more damage due to a bad model of the process or detection logic, which produces many false alarms and reduces the credibility of the FDI scheme [1].

FDI system is treated as a parallel system next to the monitoring (SCADA) platform therefore it should be implemented directly inside the system to be able to quickly detect process deviations. The procedure is detection, isolation and diagnosis, where a model of the process (or part of the system) can be used for reference/comparison purposes. Detection of faults can be difficult due to operating control loops, poor models, bad thresholds, therefore any badly designed part of the FDI scheme contributes to a higher false alarm rate, hence misleading the operator at his work. To achieve minimum number of false alarms, first a good model of the system is needed. Nonlinearities in the system are important to capture nonlinear properties of the system that is why a derivation of nonlinear model is preferred. After decades of model-based FDI approaches [2] including transparent first principles approach, statistical and data-driven approaches became possible for implementation [3]. Many of these uses parts of the data acquisition systems (SCADA, MES databases) for FDI and data processing e.g. principle components or neural networks, which can be put into practice, but need to be properly set up to achieve desired performance. The difficulty still resides in obtaining all the important data from the system to build the adequate model that can be used for FDI purposes. As nonlinear models are preferred, a Taguchi robust design of nonlinear principle components was used and realized by artificial auto-associative neural network (AANN) structure [4]. The structure is very useful and can be used in many modelling and FDI task [5]. In classic approach the structure and parameters of the network can be selected by trial and error approach where best results are considered for realization. On the other hand Taguchi's robust design technique uses Design of Experiments (DoE) approach upon which the best case parameters can be selected. The emphasis is on parameters that mostly affect the operation of the AANN especially (number of hidden neurons in encoding/decoding layer, number of neurons in bottle-neck layer, type of neuron's activation function, training method, etc). The Taguchi design [6] can be used to solve many optimization tasks, among others it has been shown it can improve industrial production procedures, be used for controller design [7], artificial neural networks (ANN) design, etc.

2 Auto-associative neural networks

In modern monitoring systems (SCADA, MES), databases can be used to obtain a proper model of the process behaviour, directly out of the recorded history. To be able to model the nonlinear properties hidden in the data, a nonlinear method should be used. Among many data-driven methods, statistical methods have many benefits nevertheless neural networks remain their main competition. In 1991, Kramer [2] presented nonlinear principle component approach where he proposed an auto-associative neural structure along with recommendations how to select optimal number of mapping layers, by using final prediction error and information theoretic criterion.

Such one-factor representation does not take all the key factors under consideration therefore also number of bottleneck nodes, transfer function, learning algorithm were included into design. To perform NLPCA as described by Kramer, the AANN in Fig. 1 shows the structure which contains 3 hidden layers of neurons between the input and output layers of variables.

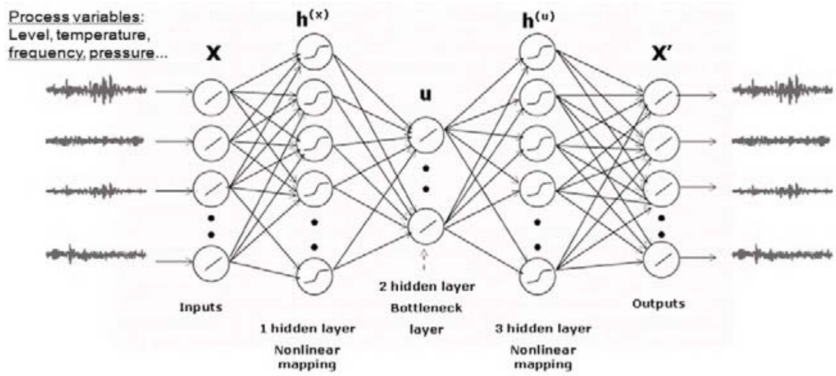

Figure 1. The structure of Autoassociative neural network

Next to the input layer there is a mapping (encoding) layer, followed by a bottleneck layer, which is then followed by a demapping (decoding) layer. A nonlinear function maps from the higher dimension input space to the lower dimension bottleneck space, followed by an inverse transform mapping from the bottleneck space back to the original space represented by the outputs. They are to be close to the inputs as possible by minimizing the cost function. Desired mean square error between the neural network output and the original data is thus minimized. The choice of the number of hidden layers in an encoding and decoding layer follows a general principle of parsimony, since more hidden layers increase the nonlinear modelling capability of the network, on the other hand that could also lead to over-fitted solutions. According to Kramer, a procedure and recommendations for designing the AANN structure upon the problem specifics, is given. He used a simple approach where he restricted the number of weight in the network to a fraction of the number of constraints imposed by the data set. The number of adjustable parameters in the network is defined by eq. 1. left:

$$n = (m + f + 1)(M_1 + M_2) + m + f \qquad M_1 + M_2 << m(n - f)/(m + f + 1) \tag{1}$$

where M_1 and M_2 are the number of nodes in the mapping and demapping layers, respectively, m is the number of nodes in the input and output layers, and f is the number of nodes in the bottleneck layer. The number of adjustable parameters is implying the right inequality. If the number of mapping and demapping nodes allowed by these inequalities is less than or equal to f, then there is not enough data to support extraction of f nonlinear factors, since the bottleneck by design occurs in the second hidden layer of the combined network. Instead, of Kramer's approach, a Taguchi DoE was tested where experimental results determine which combination of the structure and parameter settings are optimal for desired study case.

3 Taguchi Design of Experiments approach

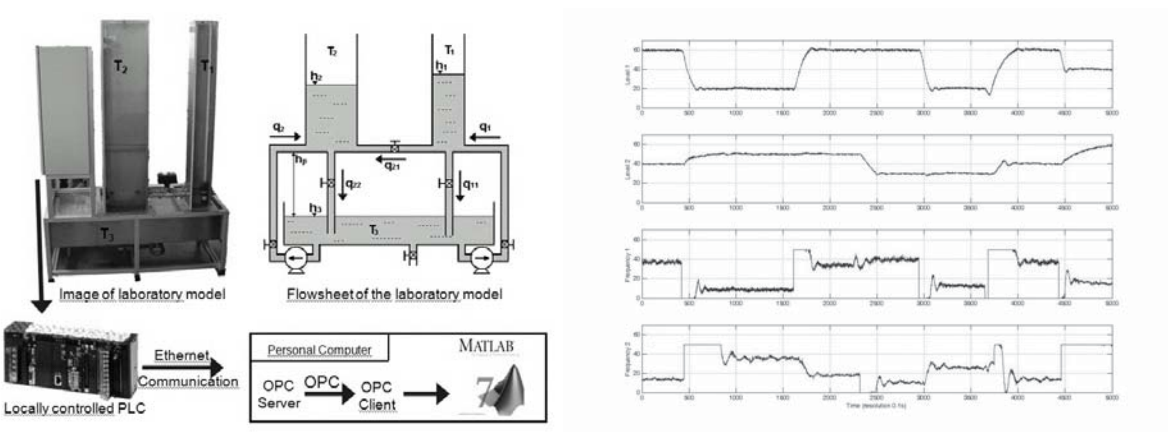

Figure 2. Laboratory hydraulic model (left) and fault-free operation (right)

The laboratory model (Fig. 2) has four system variables which are lead to the neural network input layer and the number of bottleneck layer neurons was selected as *m-1* (one less than the input layer). Nonlinear transfer function and Levenberg-Marquardt back-propagation algorithm was used for the training with 5000 training samples that were adequately pre-processed. According to Kramer's definition, the number of needed neurons in mapping layer should not be less than 7 (the error is approximately constant, Table 1). Well, to be able to reduce the number of mapping nodes if possible, a Taguchi design was proposed.

No of map. nodes	*adjustment parameters*	*Error*	*FPE*	*AIC*
1	23	16663	150,9336	2,4569
2	39	1290	6,8928	1,1165
3	55	410	1,5495	0,4683
4	71	655	1,9231	0,5621
5	87	42	0,1015	-0,8468
6	103	37	0,0750	-1,0743
7	119	25	0,0444	-1,0743
8	135	31	0,0486	-1,0348
9	151	24	0,0335	-1,1971
10	167	7	0,0097	-1,7350

Table 1. Results of one-factor representation

Factor	*Level 1*	*Level 2*	*Level 3*
A	5	10	20
B	1	2	3
C	Lin	Sig	Tanh
D	LM	GD	GDX
E	10%	30%	50%

Table 2. Planned variations of parameters and respective levels

Factor A: number of hidden neurons in encoding/decoding layer; Factor B: number of neurons in bottle-neck layer; Factor C: transfer function; Factor D: back-propagation learning method: LM (Levenberg-Marquardt), GD (gradient descent), GDX (momentum gradient descent with adaptation); Factor E: size of learning data against size of complete data

In the Table 2, variations of process variables were planned to conduct by by orthogonal array L_{27} (3^{13}). If a classic design of experiments (trial-and-error) would be used, 5^3 experiments would be necessary to conduct. To complete the task by Taguchi DoE, only 27 experiments were conducted a few times (3x27=81 experiments) in random order to achieve mean values of respective results. It is a subjective decision or case dependent how factors are entered into OA, as this defines the ability to detect correlation (interaction) between variables.

After conducting experiments, robustness measure in the form of Signal-to-Noise (S/N) ratio was calculated. Since a good process model is needed, neural network with smaller training error is preferred (S/N type "Smaller is better", [6]). Fig. 3 left shows training procedures (experiments) and calculation of S/N ratios, upon which best parameter selection can be achieved. Further analysis of cross-correlation and interaction between process variables is possible (ANOVA) and it depends how the data was inserted into OA (adequate columns).

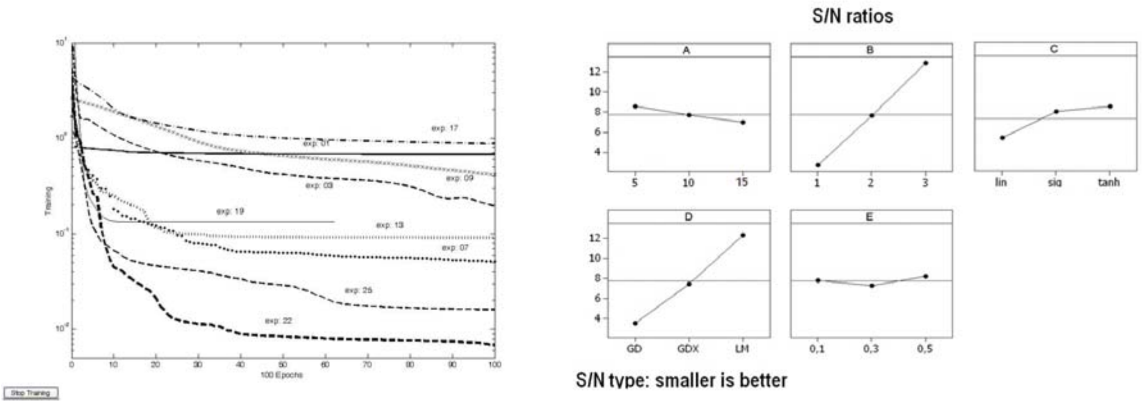

Figure 3. Results of AANN training according to planned experiments (left) and calculated S/N ratios (right)

Best result is thus achieved with combination of 5 mapping nodes, 3 bottleneck nodes, hyperbolic tangent sigmoid transfer function and Levenberg-Marquardt learning algorithm. AANN model outputs achieved by Kramer and Taguchi DoE are presented in Fig. 4., where a fault-free operation is simulated (level in the first tank).

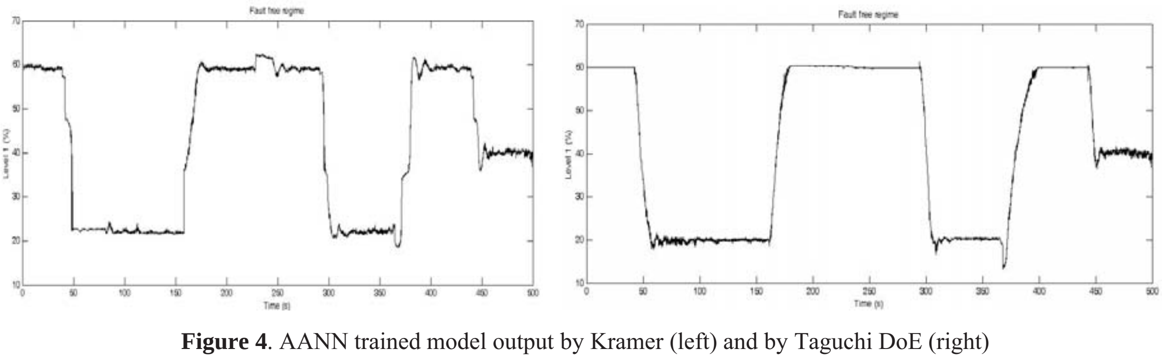

Figure 4. AANN trained model output by Kramer (left) and by Taguchi DoE (right)

4 Conclusion

The structure 4-5-3-5-4 for the AANN models was used to obtain faulty and fault-free regime models of the process. FDI scheme was realized in Matlab/ Simulink, where residual signals were generated according to deviation between measured and AANN outputs with a threshold function to detect faults. The performance was evaluated by several fault cases introduced to the three-tank laboratory model; displacement of the level sensors in the tanks and pipeline of the pumps can be partially clogged (closing the inlet valves). Additive faults were abruptly brought about.

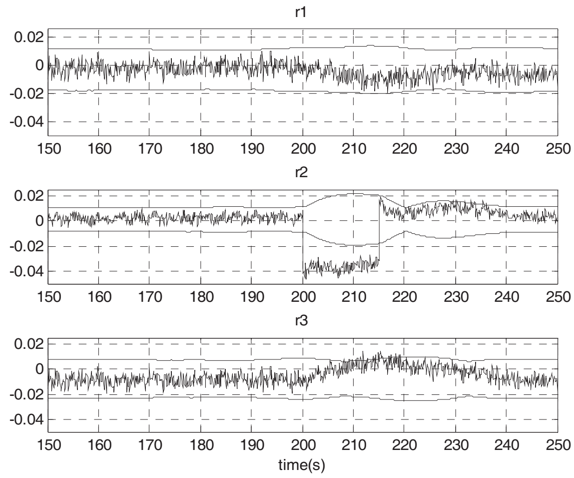

Figure 4. Detection of sensor bias fault in the second tank (level measurement)

By properly setting the isolation parameters shift detection on the level sensors was achieved capable of detecting small faults e.g. 4% variation of the measured signal (Fig. 5). Also a test for sensor drift was conducted, where it proved that drifts could be detected even under close loop conditions. The noise was approximately 2-3%, thus measurement signals were slightly filtered to improve detection results.

Many factors such as noise and unpredicted disturbances on measured signals can always produce unwanted responses in the system however sensitivity, reliability and accuracy of the FDI system can be greatly improved if suitable design is used. By applying Taguchi's robust design to various FDI system components (model, controller) we try to properly set up FDI scheme that all predicted faults can be detected. The same technique can be used in many different optimization or design tasks therefore an on-going research will discuss also the selection of parameters for optimal observer based FDI scheme, controller design, selection of the number of principle components in a classic principal component analysis optimization task, etc.

5 References

[1] Basseville, M., Nikiforov, I. V.: *Detection of abrupt changes: theory and applications*. Englewood Cliffs. Prentice Hall, New York, 1993.

[2] Garcia, E. A., Frank, P. M.: *Deterministic Nonlinear Observer-Based Approaches to Fault Diagnosis: A Survey.* Control Engineering Practice, 1997, Vol. 5, pp. 663-670.

[3] Jolliffe, I. T.: *Principal Component Analysis*, Second Edition. Springer, New York, 2004.

[4] Kramer, M. A. *Nonlinear principal component analysis using auto-associative neural networks*. AIChE Journal, 1991, Vol. 37, pp. 233-243.

[5] Bai, S., McLean, D. D., Thibault, J.: *Autoassociative Neural Networks for Robust Dynamic Data Reconciliation.* 2007, AIChE Journal Vol. 53, No. 2, pp. 438-448.

[6] Taguchi, G., Chowdhury, S., Wu, Y.: *Taguchi's Quality Engineering Handbook*. John Wiley & Sons, New Jersey, 2004.

[7] Vlachogiannis, J. G., Roy, R. K.: *Robust PID controllers by Taguchi's method.* The TQM Magazine, 2005, Vol. 17 No. 5, pp. 456-466.