Development of a hybrid PCA-ANFIS measurement system for monitoring product quality in the coating industry

K. Warne*, G. Prasad, N.H. Siddique, L.P. Maguire
Intelligent Systems Engineering Laboratory, School of Computing and Intelligent Systems, University of Ulster
Northern Road, Derry, BT48 7JL, Northern Ireland, United Kingdom
*E-mail: { *k.warne, g.prasad, nh.siddique, and lp.maguire}@ulster.ac.uk*

Abstract – *In industry today many products are sold for their efficacy rather than their chemical composition. There are several key attributes within the coating industry such as, Anchorage, Seal strength etc., which characterize the quality of the final product and are features used by the company to promote the sale of the product. Such quality variables (dependent variables) however may involve measurement difficulties. The difficulties can be due to a variety of reasons, including: (1) Reliability of on-line sensors, (2) Lack of appropriate on-line instrumentation. In the coating process off-line laboratory tests determine product quality measurements. However, such laboratory analyses introduce delays in the measurement of key performance indicators. This can result in a significant economic loss* if the analysed product fails the quality control test. An improved monitoring system is required therefore to determine product quality online and minimise commercial wastage. To facilitate this, advanced monitoring and control or optimization techniques require inferred measurements, generated with correlations from readily available process variables (independent variables). Although inferential models are widely used in industry, only a few techniques for inferential model development are discussed in the open literature. This paper therefore will present an improved systematic approach for the development of inferential models using soft computing systems and demonstrate the methodology by inferring the 'Anchorage' of polymeric-coated substrates (i.e.Tyvek or paper) in the coating industry.*

Keywords: Inferential measurement, Principal Component Analysis (PCA), ANFIS.

1 Introduction

In striving towards computer automation, industries are constantly seeking effective tools to monitor and control increasingly complicated processes. In many chemical or manufacturing processes it is often difficult to measure or estimate some important process variables due to the limitations of process technology or measurement techniques. These hard-to-measure variables sometimes referred to as dependant variables frequently, represent product quality and are of direct economic interest. They are also the key indicators of process performance and are normally determined by off-line sample analyses in the laboratory. A significant delay (often several hours) however can be incurred by laboratory testing and although the product quality can be determined after this delay, it is often too late to make timely adjustments resulting in a significant economic penalty. Also automatic control and optimisation schemes cannot be implemented and as a result, process performance may be degraded. In such cases an estimate of the dependent variable may be obtained from an inferential model. Such models, also known as 'soft sensors' can be used to obtain a regression model between easily obtained measurements (independent variables) and quality variables (dependent variables).

Generally speaking, there are mainly three kinds of approaches to building inferential models: mechanistic modelling (first principles) [1], statistical regression methods [2] and artificial intelligence modelling [3]. More recently hybrid soft computing models, based on a combination of neural, fuzzy and evolutionary computation technologies, have been applied to a number of classification, prediction, and control problems [4, 5]. The successful development of inferential models, often referred to as 'soft-sensors', depends largely on the quality and nature of data used in model development. It has been shown that complex relationships in industrial processes can be identified through the use of artificial neural networks [6, 7]. However, in many processes there exist interdependencies between process variables and in such situations the multi-collinearity of the data needs to be examined. If the input data are highly collinear, which is typical in many manufacturing processes, the use of the original process data in model development will result in an ill-conditioned problem. Qin et al. [8] illustrates how collinearity results in a large prediction variance using neural networks. Another disadvantage of neural networks is the iterative procedure of network training and regression refinement. The accuracy of a neural network model may also be influenced by altering the topology (structure) and parameters of the network, therefore obtaining the best architecture may be time-consuming and dependant on the modeller's experience. A further problem is that of

* 0-7803-8566-7/04/$20.00 © 2004 IEEE.

overfitting which results in poor generalization. Evidently there are several pitfalls of the current inferential modelling techniques that need to be taken into careful consideration and addressed when developing an inferential measurement system. This paper aims to overcome the negative aspects highlighted above and provide a more systematic approach for the successful development of robust inferential models for industry.

The paper is organised as follows: Section 2 provides an overview of the methodology employed in this work. Section 3 discusses the development of a Principal Component Model. Section 4 discusses the Fuzzy Inference System developed and experimental results are presented. Finally, conclusions are drawn in Section 5.

2 Methodology

The proposed approach is addresses some of the problems that currently exist in the area of inferential modelling through the fusion of statistical and artificial intelligence models. A Principal Component Analysis (PCA) model is combined with a neuro-fuzzy model to generate a hybrid PCA-ANFIS approach for the development of inferential measurements for monitoring product quality in a manufacturing process. The neuro-fuzzy model employed incorporates the pattern recognition capabilities of neural networks with the transparent modelling advantages of fuzzy logic. This form of neuro-fuzzy approach provides a means of training a family of membership functions to emulate a complicated (e.g. nonlinear), multi-dimensional mapping function. The neuro-fuzzy application of choice for this paper is ANFIS [9]. To address the issue of multi-collinearity within the process data, a principal component analysis (PCA) algorithm is incorporated into the inferential model architecture.

The central idea of PCA is to reduce the dimensionality of a data set which consists of a large number of interrelated variables, while retaining as much variation as possible in the original data set [10]. This is achieved by transforming to a new set of variables, i.e. Principal Components (PCs), which are uncorrelated, and are ordered so that the first few retain most of the variation present in all of the original variables. In developing a PCA initialisation model, the process data can be adequately described using far fewer parameters than the original variables with no significant loss of information. The issue of collinearity in the process data is also eliminated. To date no formal procedure exists however for the fusion of these modelling techniques discussed. For example, the criterion for selecting the adequate PCs to retain from PCA are very much *ad hoc*, whose justification is still mainly that they are intuitively plausible. Therefore, in this paper a more concise statistical method for PC selection is adopted. Once the number of principal components is chosen that sufficiently represents the original data set, regression can be performed to develop a correlation model. As an alternative to using the original variables as inputs to ANFIS, the selected principal components are used from the PCA algorithm. The PCA-ANFIS architecture developed is illustrated in Figure 1.

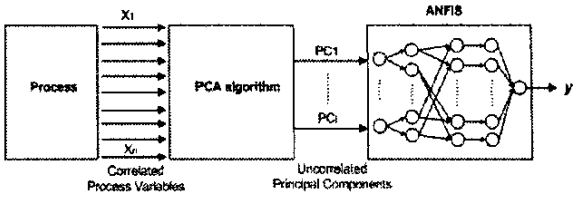

Figure 1. Architecture of PCA-ANFIS for inference of product quality.

In this case study the methodology is demonstrated by inferring the 'Anchorage' of polymeric-coated substrates (i.e.Tyvek or paper) in the coating industry. The source of training and testing samples is from the process data records, which are recorded from the supervisory control and data acquisition (SCADA) system, and the corresponding laboratory analysis. Several hours of process data was collected on all recorded variables. The dataset consisted of 11 variables and 600 observations representing the entire range of operating conditions for the product under consideration. Data on 'Anchorage' (i.e. quality variable) was based on measurements in the laboratory using a novel test method discussed in [11]. The data set was centered, scaled and checked for outliers. These unusual observations were detected using the statistics presented in [10]. Those observations whose test statistic deviated by more than three standard deviations from the mean were deleted. The total number of outliers deleted was 27. After deletion of outliers, the remaining data set had (573 obsersations x 11 variables) in the data matrix. The following sections discuss the development procedure.

3 Development of PCA Algorithm

The successful development of any inferential measurement system depends largely on the quality and nature of data used in the model development. In many industrial applications [12, 13] there exist interdependencies between process variables. In such situations the multi-collinearity of the data needs to be examined. If the input data are highly collinear, which is typical in many processes then using the original data for regression will result in an ill-conditioned problem. PCA is used to address multi-collinearity problems. It produces a compressed statistical model that gives linear combinations of the original variables that describe the major trends in a dataset. This is achieved by transforming to a new set of variables, i.e. Principal Components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables.

Mathematically, PCA relies upon an eigenvector decomposition of the covariance matrix or correlation matrix of the process variables. For a given dataset, **X**,

mean centered and scaled with m rows and n columns, with each variable being a column and each sample a row, the covariance matrix of $\mathbf{X}$ is estimated by $(n-1)^{-1}\mathbf{X}^T\mathbf{X}$. The NIPALS (noniterative partial least squares) algorithm was first introduced to compute the principal components in a sequential manner when the number of variables is large. The number of principal components that provide an adequate description of the data can be assessed using several methods [14]. PCA decomposes the data matrix $\mathbf{X}$ as the sum of the outer products of vectors $\mathbf{t}_i$ and $\mathbf{p}_i$ plus a residual matrix $\mathbf{E}$:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1^T + \mathbf{t}_2\mathbf{p}_2^T + \ldots + \mathbf{t}_k\mathbf{p}_k^T + \mathbf{E} \qquad (1)$$

Here k must be less than or equal to the smaller dimension of $\mathbf{X}$, i.e. $k \leq \min\{m, n\}$. The $\mathbf{t}_i$ vectors are known as principal component *scores* and contain information on how the samples relate to each other. The $\mathbf{p}_i$ vectors are known as *loadings* and contain information on how the variable relate to each other. In the PCA decomposition, the $\mathbf{p}_i$ vectors are eigenvectors of the covariance matrix, i.e. for each $\mathbf{p}_i$

$$\text{cov}(\mathbf{X})\,\mathbf{p}_i = \lambda_i\,\mathbf{p}_i \qquad (2)$$

where λ_i is the eigenvalue associated with the the eigenvector $\mathbf{p}_i$. The $\mathbf{t}_i$ form an orthogonal set ($\mathbf{t}_i^T\mathbf{t}_j = 0$ for $i \neq j$), while the $\mathbf{p}_i$ are orthonormal ($\mathbf{p}_i^T\mathbf{p}_j = 0$ for $i \neq j$, $\mathbf{p}_i^T\mathbf{p}_j = 1$ for $i = j$). Note that for $\mathbf{x}_i$ and any $\mathbf{t}_i$, $\mathbf{P}$ pair

$$\mathbf{x}_i\mathbf{P} = \mathbf{t}_i \qquad (3)$$

i.e., the score vector $\mathbf{t}_i$ defined by $\mathbf{P}$ is the linear combination of the original variables. Another way to look at this is that the $\mathbf{t}_i$ are the projections of $\mathbf{X}$ onto the $\mathbf{p}_i$. The $\mathbf{p}_i$ are arranged in descending order according to the associated λ_i. The λ_i are a measure of the amount of variance described by the $\mathbf{p}_i$. Because the $\mathbf{p}_i$ are in descending order of λ_i, the first pair capture the largest amount of information of any pair in the decomposition, and each subsequent pair capture the greatest possible amount of variance at that step. The variance of each pair can be accumulated and compared with a given constant (e.g. 85%, 95%) to choose the principal components of all the pairs. Using this PCA approach, the data can be adequately described using far fewer parameters than the original variables with no significant loss of information and the issue of collinearity in the data is eliminated. Once the number of principal components is chosen that sufficiently represents the original data set, regression can be performed to develop an inferential model. As an alternative to using the original variables as inputs to the inferential model, the selected principal components scores are thus used from the PCA algorithm.

During initial analysis of the process under study, it was identified from process knowledge that 11 candidate variables are related in some degree to the quality variable.

These 11 variables may contain an overlap of information on the final product quality and because they are possibly correlated to each other when these variables are used as inputs to build an inferential model, some information of no use is obtained from them. Thus, a great number of data sets may be compressed in a manner that retains the essential information and is more easily displayed than each of the process variables individually. Here, by referring to the 11 variables as the matrix $\mathbf{X}$ (size 573x11), then this procedure can be summarised as follows:

1) Perform PCA on matrix $\mathbf{X}$
2) Principal Component Selection from PCA

The chosen numbers of principal components together with the desired output y are employed in developing the inferential model. Table 1 lists the results of the PCA. The next section describes the rationale for choosing a subset of principal components for model development.

Table 1. The result of PCA

Percent variance captured by PCA model			
PC No.	Eigenvalue	Variance (%)	Total variannce (%)
1	8.297	75.431	75.431
2	1.935	17.591	93.022
3	0.349	3.181	96.203
4	0.216	1.963	98.167
5	0.093	0.843	99.010
6	0.057	0.523	99.533
7	0.027	0.253	99.786
8	0.011	0.096	99.883
9	0.007	0.063	99.946
10	0.005	0.049	99.996
11	0.0004	0.004	100

3.1 Principal Component Selection

Many of the rules used for choosing the number of principal components, m to retain, instead of p original variable are very much *ad hoc* rules-of-thumb, whose justification is still mainly that they are intuitively plausible, and that they work in practice [10]. The most obvious criterion is to select a (cumulative) percentage of total variation to which it is desired that the selected principal components (PCs) should contribute, for example 70% to 95%. The required number of PCs is then the smallest number of PCs for which this chosen percentage is exceeded. Principal components are successively chosen to have the largest possible variance, and the variance of the jth PC is l_j. Furthermore,

$$\sum_{j=1}^{p} l_j = \sum_{j=1}^{p} S_{jj} \qquad (4)$$

that is the sum of the variances of the PCs is equal to the sum of the variances of the elements of $\mathbf{X}$. The definition of 'percentage of variation accounted for by the first k PCs' is therefore

$$t_k = 100\sum_{j=1}^{k} l_j \Big/ \sum_{j=1}^{p} S_{jj} = 100\sum_{j=1}^{k} l_j \Big/ \sum_{j=1}^{p} l_j \qquad (5)$$

Choosing a cut-off, t^*, usually in the region between 70% and 95%, and retaining m PCs, where m is the smallest integer, k, for which $t_k > t^*$, provides a rule which, in

practice, preserves in the first m PCs most of the information in **X**. The best value for t^* will generally become smaller as p increases, or as n, the number of observations, increases [10].

The 'scree' graph discussed and named by Cattell [15] is another rule available for PC selection. The way in which Catell formulates the rule is to look for the point beyond which the scree graph defines a more-or-less straight line, not necessarily horizontal. The first point on the straight line is then taken to be the last component to be retained. Cattell's formulation, where we look for the point at which $l_{k-1} - l_k$ becomes fairly constant for several subsequent values is perhaps less subjective but still requires some degree of judgement.

Table 2. Eigenvalue analysis via Catell's formulation

Component Number, k	1	2	3	4	5	6
Eigenvalue, l_k	8.297	1.935	0.349	0.216	0.093	0.057
t_k	75.431	93.022	96.203	98.167	99.010	99.533
$l_{k-1} - l_k$	6.362	1.583	0.133	0.123	0.036	0.03

To give confidence that the correct number of principal components have been chosen, an approach similar to that of cross validation has been investigated in this study. The number of terms in the estimate for y, corresponding to the number of PCs, is successively taken as 1,2, ..., and so on, until overall prediction is no longer significantly improved by the addition of extra terms (PCs). The number of PCs to be retained, m, is then taken to be the minimum number necessary for optimal prediction. In this work a combination of all three techniques is adopted to decide the number of PCs to be retained. The first two rules are used to generate a subset of PCs. The subsets of PCs are then subjected to cross validation to select the number of PCs that give optimal prediction results when used in regression.

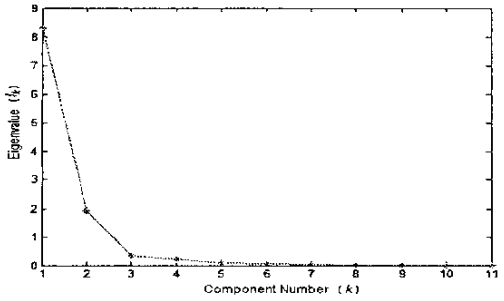

Figure 2. Scree graph from PCA

Using the criterion described above the adequate number of PCs to be used in model development can be determined. The first criteria, where a cut-off limit (t^*) corresponding to the cumulative percentage of total variation is decided, was adopted initially. The choice of t^* is often selected heuristically. One approach is to use Jolliffe's criteria [10] of $\lambda = 0.7$ (i.e. percentage of variation accounted for = 70%) to determine which PCs to retain. If this criterion is applied to the results of PCA in this work, it would suggest that one PC should be retained. It is unusual however that one PC contains enough information to adequately describe the original data set. Various cut-off limits are also suggested by others [16] ranging from 70%-95%, which when applied to this particular problem would affect the number of chosen PCs significantly. This highlights the subjectiveness of this approach. By adopting Cattell's formulation and plotting the scree graph, shown in Figure 2, the number of PCs to be retained increases.

From a general observation of the scree graph, it can be clearly seen that beyond three or four PCs lie on a more-or-less straight line. Table 2 shows the first six eigenvalue analysis using Catell's formulation and the results also indicate that after four PCs, $l_{k-1} - l_k$ becomes fairly constant for several subsequent values. Both three and four PCs therefore were retained and cross validation carried out to determine which set gives better prediction results. After cross validation it was concluded that three PCs gave optimal results, therefore three PCs were retained and used as inputs for the inferential model described in the next section. The new dataset used in the development of the fuzzy inference system consisted of (573 observations x 3 principal components).

4 Adaptive Network-Based Fuzzy Inference System (ANFIS)

The specific neuro-fuzzy system that is adopted in this work for prediction of product quality in the coating industry is ANFIS [9]. ANFIS has proven to be an excellent function approximation tool and uses a hybrid learning algorithm that combines the back propagation gradient descent and least squares methods to create a fuzzy inference system whose membership functions are iteratively adjusted according to a given set of input and output data.

The fuzzy model is based on a first order Takagi-Sugeno-Kang (TSK) architecture that is generally composed of k rules of the form

$$\text{Rule } i\text{: IF } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \ldots \text{ and } x_n \text{ is } A_{in}$$
$$\text{THEN } y = b_{i0}+b_{i1}x_1+b_{i2}x_2+ \ldots +b_{in}x_n$$

where $x_1, x_2, \ldots, x_n$ are antecedent variables, any y is the consequent variable. $A_{i1}, A_{i2}, \ldots,$ and A_{in} are fuzzy sets defined over the domains of the respective antecedents. $b_{io}, b_{i1}, \ldots, b_{in}$ are constant coefficients that characterize the linear relationship defined by the ith rule in the rule set, i=1 , 2 ,..., k. A TSK fuzzy model is a computationally efficient platform that is well suited for implementation of non-linear associations through the construction of many piecewise linear relationships [17]. It is beyond the scope of this paper to give a detailed discussion on the ANIFS architecture and its operation however in the literature many reviews and articles are available [18, 19].

As discussed earlier, an alternative to using the original variables as inputs to ANFIS, the selected principal components are used from the PCA algorithm. Figure 3

illustrates the proposed network structure. The remaining sub-sections discuss the model development phase and the validation results obtained.

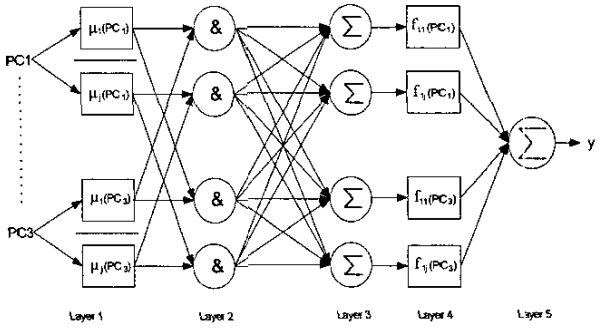

Figure 3. ANFIS Structure

4.1 Model development

To illustrate the performance advantage by the incorporation of a PCA model, an ANFIS model using the standard approach (i.e. original process variables as inputs) was developed to enable comparisons. First 60% of the dataset is used for training and the remaining used for validation. ANFIS attempts to minimise the mean squared error between the network outputs and the desired output as the data points in the training set are presented. The MSE is defined as:

$$\mathrm{MSE} = \frac{1}{\mathrm{N}} \sum_{n=1}^{N} (y_a - y_p)^2 \qquad (4)$$

where y_a and y_p are the actual and predicted responses, respectively, and N is the total number of predictions. Before training can occur, an initial FIS is created using the grid partition method [20].

As discussed earlier as an alternative to using the original variables as inputs to the inference system, the principal components from the PCA model are adopted. This PCA-initialised inference system is designed with three inputs: Principal component 1, Principal component 2, Principal component 3, and a single output-Anchorage. After several trials, use of three membership functions for all inputs is found to attain best results. Each input has generalized bell-shaped membership functions (MF). The results are presented in the following section.

4.2 Model Validation

To validate accuracy of the fuzzy models, validation data was applied to the models. The mean squared error (MSE) for training and validation data using the different approaches are illustrated in Table 1 and Figure 4. One of the first observations that can be made from adopting this PCA-ANFIS approach is the reduction in complexity. This type of PCA-initialised inferential measurement system results in a considerably less complex architecture as the number of inputs to the system will be significantly reduced. In this particular application the number of inputs to the system has been reduced by 75%. Due to the nature of ANIFS, the number of inputs to the network and number of membership functions chosen for each input has a significant effect on the number of modifiable parameters within the system.

Table 1. Performance measures for ANFIS and PCA-ANFIS models

Iterations	ANFIS		PCA-ANFIS	
	Training MSE	Validation MSE	Training MSE	Validation MSE
10	0.6416	3.2200	0.0790	5.1980
20	0.2905	1.5270	0.0676	2.0190
30	0.0756	0.6955	0.0590	0.8118
40	0.0449	0.3470	0.0529	0.4774
50	0.0361	0.3400	0.0475	0.3829
60	0.0302	0.3950	0.0424	0.2371
70	0.0296	0.4160	0.0384	0.1376
80	0.0292	0.5990	0.0361	0.0894
90	0.0292	0.6020	0.0345	0.0795
Training Time	1800 sec		12 sec	
Adjustable Parameters	2,126,863		135	

Using the grid partitioning method the system generates rules by enumerating all possible combinations of membership functions of all inputs; this leads to an exponential explosion even when the number of inputs is moderately large. This sometimes is referred to "curse of dimensionality" [21]. Clearly this issue has been eliminated with the incorporation of a PCA model as the number of adjustable parameters has been dramatically reduced. Evidently, training time will be significantly reduced also as the computational overhead associated with the PCA algorithm is negligible. It can be observed that the training time is 150 times quicker when the Principal Components are used as inputs to the network.

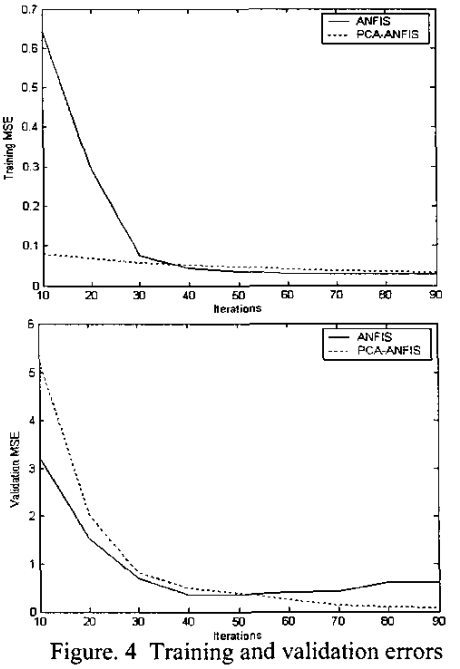

Figure. 4 Training and validation errors

Use of a smaller number of orthogonal latent variables (i.e. PCs), which are linear combinations of the original variables overcomes the problem of overfitting that is

common in neural network and ANFIS training. For the standard ANFIS approach the regression algorithm is over-fitting the training data after 50 iterations. For the PCA-ANFIS approach; no overfitting occurs. Another advantage of using this approach is the enhanced model accuracy. The optimal performance achieved resulted in a validation error of 0.0795 compared to that of 0.34 using the traditional ANFIS approach. This is a considerable improvement using the proposed approach. Therefore the advantages in both model performance and improved training ability by the incorporation of a PCA-initialisation model have been demonstrated successfully.

5 Conclusion

This paper presented a detailed systematic approach for the development of inferential measurements using a hybrid PCA-ANFIS approach and illustrated the improved effects on network performance using the approach. The advantages can be summarised as follows:

- Performance/accuracy of the model is significantly increased ;
- Model is more robust as collinearity affects are accounted for by incorporation of PCA algorithm ;
- Complexity of model significantly reduced; training time also significantly reduced.

A more concise method for fusion of statistical and computational intelligence models has been discussed. The application on which this methodology is demonstrated is unique. No such work in the literature to date has presented any inferential modelling strategies in the area of the coating industry. This strategy developed through the fusion of statistical and artificial intelligence modelling to generate a hybrid inferential measurement system has the potential to significantly improve the quality control monitoring system and reduce the economic loss encountered through the production of off-spec material. Future work will attempt to improve prediction accuracy further by investigating different types of computational intelligence models such as Self-Organizing Fuzzy Neural Networks (SOFNN). This work aims to automate much of the model development process while improving model accuracy concurrently.

References

[1] Ellis M., Taylor K., "On-line molecular weight distribution estimation and control in batch polymerisation", *AIChE Journal*, Vol. 40, pp. 445-462.

[2] Mosbah H., Zhand A., "Multivariable Inferential Control of Distillation Compositions Using Dynamic Principal Component Regression Models", Proceeding of American Control Conference, Colorado, June 2003, pp. 1974-1979.

[3] Yang X., Zhang J., Morris A., An Artificial Neural Network Approach for Inferential Measurement, International Conference on Neural Information, Vol. 1, pp. 485-488, 1995.

[4] Bonissone P, Chen Y-T, Goebel K and Khedkar P (1999b). Hybrid Soft Computing Systems: Industrial and Commercial Applications, *Proceedings of the IEEE*, pp 1641-1667, vol. 87, no. 9, September 1999.

[5] Bonissone P and Goebel K (2001a). Soft Computing for Diagnostics in Equipment Service, *Proceedings of IFSA/NAFIPS*, Vancouver, BC, July 2001.

[6] Bharitya S., Whiteley J., "Development of Inferential Measurements using Neural Networks", *ISA Transactions*, Vol. 40, pp. 307-323 (2001).

[7] Chen Y.T. and Bonissone P (1998). Industrial applications of neural networks at General Electric, Technical Information Series, 98CRD79, GE Corporate R&D, October, 1998.

[8] Qin et al., "Self-validating Inferential Sensors with application to Air Emission Monitoring", *Ind. Eng. Chem. Res*, Vol. 35, No. 5, pp. 1675-1685 (1997).

[9] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, No. 3, pp. 665-684, 1993.

[10] Jolliffe I.T., Principal component analysis, Springer-Verlag, NY, 1986.

[11] Warne et al. (2003) "A Novel Intelligent Approach to Anchorage Measurement using Electron Microscopy", *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1, pp. 3810-3816, Washington D.C..

[12] Zhong W, Yu J (2000) "MIMO soft sensors for estimating product quality with on-line correction", *Trans IchemE*, Vol 78, Part A.

[13] Fujii H, Shah S (1997) "Application of the PLS Technique to the Estimation of Distillation Tower Top Composition", *IFAC Symposium on Advanced Control of Chemical Processes (ADCHEM)*, pp. 295-300.

[14] J.E. Jackson, A Users Guide to Principal Components. New York: Wiley, 1991.

[15] Catell R.B (1966) The Scientific Use of Factor Analysis in Behavioural and Life Sciences, Plenum Press, New York.

[16] Dunteman G.H (1989) Principal Component Analysis, Quantatitive Applications in the Social Sciences, Sage Publications, London,

[17] J. Yen and R. Langari, ***Fuzzy*** *Logic: Intelligence,Control, and Information*. New York, NY: Prentice Hall, 1999.

[18] J.-S. R. Jang, Neuro-Fuzzy Modelling: Architectures, Analyses, and Applications,' Ph.D. Dissertation, EECS Department, Univ. of California at Berkeley, 1992.

[19] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence," Prentice Hall, 1996.

[20] Fuzzy Logic Toolbox User's Guide, The Mathworks Inc, 2001.

[21] Brown et al., High dimensional neurofuzzy systems: overcoming the curse of dimensionality, Proceedings of the International Conference on Fuzzy Systems, Japan, March 1995, pp. 2139-2146.