International Journal of Automation and Computing 1 (2006) 17-22

Product Quality Prediction by a Neural Soft-Sensor Based on MSA and PCA

Jian Shi, Xing-Gao Liu*

National Laboratory of Industrial Control Technology, Institute of Systems Engineering, Zhejiang University, Hangzhou 310027, PRC

Abstract: A novel soft-sensor model which incorporates PCA (principal component analysis), RBF (Radial Basis Function) networks, and MSA (Multi-scale analysis), is proposed to infer the properties of manufactured products from real process variables. PCA is carried out to select the most relevant process features and to eliminate the correlations of input variables; multi-scale analysis is introduced to acquire much more information and to reduce uncertainty in the system; and RBF networks are used to characterize the nonlinearity of the process. A prediction of the melt index (MI), or quality of polypropylene produced in an actual industrial process, is taken as a case study. Research results show that the proposed method provides promising prediction reliability and accuracy.

Keywords: PCA, RBF, MSA, Polypropylene, MI.

1 Introduction

Neural network systems have been widely used to model and control dynamic processes, because of their extremely powerful adaptive capabilities in response to nonlinear behavior. As a result, neural systems can mimic high-level cognitive tasks present in human behavior, and operate in ill-defined and time-varying environments with a minimum amount of human intervention. For example, they are capable of (i) learning from interaction with an environment, without restriction; to capture any kind of functional relationship between information patterns, if enough training information is provided, (ii) generalize learnt information to similar situations never seen before, and (iii) possess a good degree of fault tolerance, mainly due to their intrinsic massively parallel layout. These properties make neural computing appealing to many fields of engineering[1,2].

Complex systems are characterized by their hierarchical multi-scale nature, with respect not only to space, but also to time, showing dissipative structures induced by inherently nonlinear, and non-equilibrium interactions; and are stabilized by exchanging energy, matter, and information with their surroundings. A multi-scale approach is able to acquire much more information, and greatly reduce uncertainty in a system; which has led to an increase in the attention given to them from many branches of science and fields in engineering, including mathematics, chemistry, physics, astronomy, geology, biology, ecology and applied fields such as materials, mechanics, image analysis, computational methods, atmospheric science, and chemical engineering, etc[3,4].

The application of neural systems and MSA, is especially interesting for the control and optimization of chemical plants, since the kind of time-dependent problems dealt with in process engineering are highly nonlinear and uncertain, therefore, it is difficult to obtain detailed predictions from first principle models in real time[5~8]. A specific area of intrinsic interest in chemical manufacturing processes is the prediction of the quality of final products. This is even more vital in cases where it is difficult to implement reliable and fast on-line analyzers to measure relevant product properties, and to establish appropriate control strategies for production. Such situations can lead to the significant production of off-grades, especially during on-line operations involved to change product specifications. An alternative is to develop on-line estimators of product quality, based on available process information. One of the most powerful and increasingly used methodologies is inferential measurement. This method involves the forecasting of product quality, or difficult to measure process indicators, from other more reliable or easily performed plant measurements, such as pressure, flow rate, concentration or temperature[9].

The purpose of the current study is to develop a virtual soft-sensor to infer product quality from other more easily measured process variables, using neu-

Manuscript received February 28, 2005; revised September 6, 2005.

This work is supported by the National Natural Science Foundation of China (No.20106008), the National Development and Reform Commission of China (No.Fagai Gaoji-2004-2080), and the science fund for distinguished young scholars of Zhejiang University (No.111000-581645).

*Corresponding author. *E-mail address*: liuxg@iipc.zju.edu.cn

ral network architectures, as well as PCA and MSA. Until now, little literature concerning the integration of PCA, MSA and neural networks has been reported. The networks considered here are RBF networks, whose global convergence and short training time make them a primary choice for process product quality prediction[10]. The virtual soft-sensor developed in this work, has been applied to infer the Melt Index (MI) of polypropylene grades measured on-line in operating plants. Its performance is illustrated and evaluated with a case study of polypropylene quality inference. Obtained results are discussed, and concluding remarks about the design and implementation of the virtual sensor system are presented.

2 Data preprocessing

Inferential measurement systems based on neural networks are mostly developed using "data-based" methodologies, i.e. models used to infer the value of target variables are developed using 'live' plant data measured from a process plant. This implies that inferential systems are heavily influenced by the quality of the data used to develop their internal models. Consequently, the first step in building an inferential measurement system, is the preprocessing of data. During this stage, the selection of the most relevant information with which to develop a model, is performed. Moreover, as the values are heterogeneous, they must be scaled into a range suitable for data processing.

In complex industrial processes, such as chemical processing plants, the number of plant variables that can be measured is large, and the sampling rate used for these measurements may be high. This implies that variables may be highly co-related, and lead to the generation of large datasets containing a large amount of features. In such a situation it would be very useful to have an "intelligent system" capable of selecting the most relevant features, and examples, from all of the available information; with the purpose of optimizing resources needed to build an accurate and reliable model for the process under consideration. In the process of building an inferential measurement system, a reduction in the dimensions of the input space would simplify the input layer of the architecture, and reduce the time needed for training.

2.1 Selection of variables

A sensitivity analysis to reduce the dimensions of the input space could be performed by using projection techniques, which implies the definition of new combined variables. This approach typically uses statistical analysis (descriptive statistics, cross-correlations, factorial analysis, and PCA, etc.) to find relationships between variables, and to select new representative prototypes from subsets of related variable combinations. This process entails the projection of the input space into a lower dimension output space, without a loss of significant information. PCA is introduced in this work, for this purpose.

2.2 Training set

As some variables are more relevant than others, and also some input patterns of data may be more unique than others and should be considered for training, then the appropriate selection of patterns for training is one of the most important tasks in machine learning. Different strategies can be used for the selection of the most suitable training set of data, from all available process information. One method to construct a training set, consists of the selection of data from a time series of recorded plant data. While retaining the remaining data sequence, to test the performance of a sensor. This facilitates the representation of data, and evaluation, as if a sensor were operating under real plant conditions, predicting a target property sequentially over time.

3 Soft-sensor architecture

3.1 Multi-scale analysis

At first, two intimately related functions are defined, namely, the scaling function $\varphi(x)$ and its corresponding wavelet function $\psi(x)$. A multi-resolution analysis is defined using a nested sequence of closed subspace $\{V_j\}_{j\in Z}$. Each subspace V_j of scale j is spanned by a set of scaling functions $\{\varphi_{j,k}(x), \forall k \in Z\}$:

$$V_j = \{\varphi_{j,k}(x) | \varphi_{j,k}(x) = 2^{\frac{j}{2}}\varphi(2^j x - k)\}. \quad (1)$$

A mutually orthogonal complement of V_j in V_{j+1} is denoted by the subspace W_j, associated with multi-resolution analysis such that:

$$V_{j+1} = V_j + W_j \quad (2)$$

where + is a direct sum.

Similar to V_j, W_j is formed by another orthogonal basis:

$$W_j = \{\psi_{j,k}(x) | \psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)\}. \quad (3)$$

where $\psi(x)$ is the mother wavelet.

In 1989, Mallat proposed a fast decomposition algorithm to extract a wavelet coefficient based on MRA. $\forall f \in L^2(R)$, if $c_{j,k}$ is a projection coefficient from f to V_j, and $d_{j,k}$ is a projection coefficient from f to W_j, then:

$$f = \sum_{k\in Z} c_{j,k}\varphi_{j,k} = \sum_{k\in Z} c_{j-1,k}\varphi_{j-1,k} + \sum_{k\in Z} d_{j-1,k}\psi_{j-1,k} \quad (4)$$

$$c_{j-1,k} = \langle \sum_l c_{j,l}\varphi_{j,l}, \varphi_{j-1,k} \rangle = \sum_l \langle \varphi_{j,l}, \varphi_{j-1,k} \rangle c_{j,l}$$
$$= \sum_l h_{l-2k} c_{j,l} \quad (5)$$

$$d_{j-1,k} = \langle \sum_l c_{j,l}\varphi_{j,l}, \psi_{j-1,k} \rangle = \sum_l \langle \varphi_{j,l}, \psi_{j-1,k} \rangle c_{j,l}$$
$$= \sum_l g_{l-2k} c_{j,l} \quad (6)$$

and finally, the Mallat algorithm, or pyramid fast decomposition algorithm, is obtained[11]:

$$\begin{cases} C_{j-1} = HC_j \\ D_{j-1} = GC_j \end{cases}. \quad (7)$$

3.2 Architecture

Assume a data set $S = \{(X, Y)_k; k = 1, 2, \cdots, N\}$, where N is the number of sample data points, and $X = (x_1, x_2, \cdots, x_n)$, $Y = (y_1, y_2, \cdots, y_m)$, where n and m represent the numbers of input and output variables.

As shown in Fig.1, the first selection procedure is performed by PCA to obtain the most relevant features of input variables. All new variables are decomposed by wavelet theory to scale J, where A represents approximate information, while D represents detailed information. Each pair of scaled information is used to form a sub RBF network in the training procedure. Finally, multi-scale reconstruction is carried out to obtain a model output.

4 Case study: a soft-sensor for melt index in a propylene polymerization plant

4.1 Problem statement

The quality of polypropylene produced is determined essentially by the MI, which is measured by the flow rate of polymer through a die. The on-line measurement of this quantity is difficult, and requires close human intervention because the extrusion die often fools and blocks[12]. As a result, in most plants, the MI is evaluated off-line with an analytical procedure that takes almost 1 hour to complete in a laboratory, leaving the process without any real-time quality indicator during this period. Consequently, a model for estimating MI on-line would be very useful, both as an on-line sensor, and as a forecasting system. In addition, it would allow the supervision of the overall process, and avoid mismatches of product quality during product grade transitions. Fig.2 summarizes the main characteristics of polymer plants studied. Several sets of real data corresponding to several production cycles are analyzed. The MI values used are calibrated using off-line determinations. Changes in MI correspond to changes in the physical or chemical characteristics of the desired product. A pool of process information, formed by 9 process variables (pressure, flow rate, and temperature, etc.) was chosen to develop the virtual sensor.

4.2 Virtual sensor implementation

Fig.2 shows current sensor implementation within the process flow sheet. It can receive real-time readings of process variables, as well as the feedback signals of downstream on-line analyzers. Once trained, this virtual device uses only real time measurements of the selected process variables made by process sensors at any time; to infer the value of the product target property when leaving the reactor. Output can be redirected as information to a plant operator, or to a control system, to maintain optimal plant operation for a given

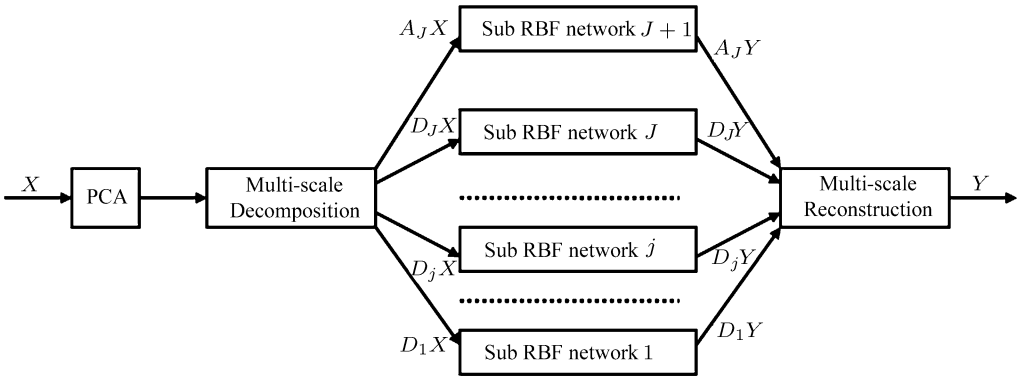

Fig.1 Soft-sensor architecture

product quality.

The aim of the current study is to implement a virtual sensor to predict the quality of polypropylene, i.e. the target variable MI, based on the state of the plant. The virtual sensor behaves as a black-box model, which relates the MI of the product to other process variables measured when the corresponding production cycle began.

4.3 Preprocessing of variables

Data used for training and testing the virtual sensor was acquired from historical logs, recorded in a real propylene polymerization plant for the 9 process variables and MI. Data was filtered to discard abnormal situations, and to improve the quality of the inference system. Input and output variables were normalized, with respect to their maximum operational values. Data from the time records of process variables, and MI, were separated into training, test, and generalization sets. The aim of the selection procedure was to preserve the time-series structure of recorded plant data. It should be noted that the test sets were obtained from the same batch as the training set, while the generalization set was derived from another batch. The PCA method mentioned earlier, for the selection of the most relevant features or input variables, was applied to reduce the size of the data set, and the time needed to adapt the virtual sensor.

5 Results and discussion

Models without PCA and MSA, were developed to compare the performance of the soft-sensor under consideration. The above mentioned data preprocessing techniques, we developed to reduce the number of variables, and were applied during the selection of the training set. Results obtained for all models considered, are summarized in the following section.

5.1 Models with reduced and complete sets of variables (PCA-MS-RBF and MS-RBF)

A reduction in the number of input variables has the advantage of (i) producing a lower dimension problem, (ii) cuts back any noise that could contaminate the measurement of discarded variables, and (iii) avoids variables that might provide conflicting information with respect to more relevant variables in relation to a target MI. The current PCA-MS-RBF model is built with 7 variables. Table 1 summarizes the performance of models developed with the reduced and complete sets of input process variables, after training with the sequential sets of data described previously. The absolute and relative mean errors listed in Table 1, indicate that a reduced set model functions better than a complete set model overall; with absolute mean error for the former being 0.0701, and for the latter being 0.1102. An increase of approximately 40% in prediction accuracy is obtained when the variable reduction technique is applied. Maximal errors, also listed in Table 1, are of the same order of magnitude as the mean errors for both models with sequential training. The

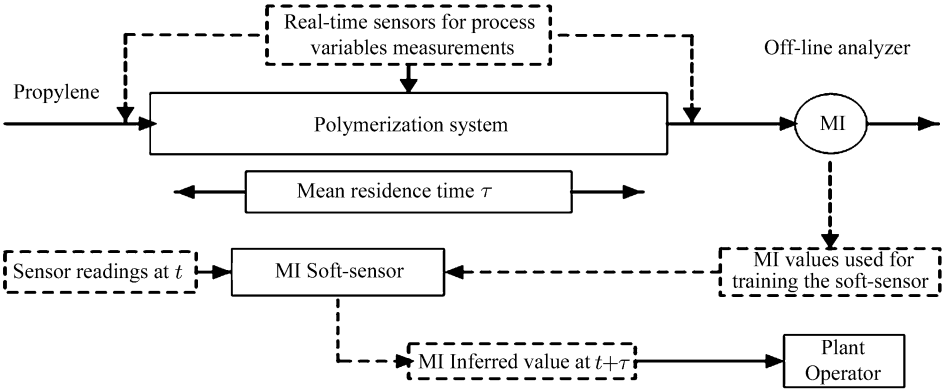

Fig.2 Propylene polymerization plant diagram

Table 1 Absolute and relative mean errors for the test set, predicted by models built with reduced and complete sets of process variables

	Mean error			Maximal error		
	Abs	Rel (%)	Comparison (%)	Abs	Rel (%)	Comparison (%)
PCA-MS-RBF	0.0701	2.87	62.66	0.1720	6.69	46.72
MS-RBF	0.1102	4.58	100.00	0.3065	14.32	100.00

better performance of PCA-MS-RBF is clearer in the time-sequence of the measured MI, as shown in Fig.3, where the model yields consistently good predictions.

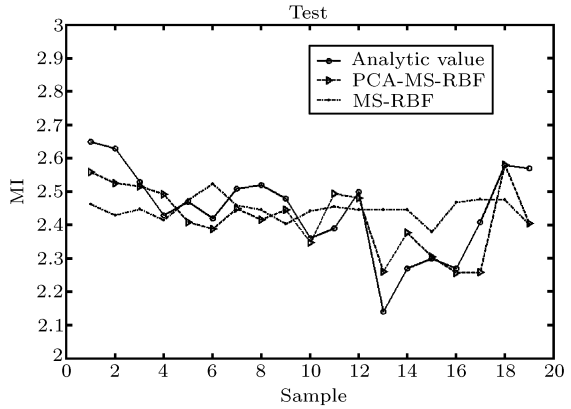

Fig.3 A comparison of measured MI time-records with predictions obtained using PCA-MS-RBF and MS-RBF

To illustrate the universality of the proposed model, a detailed comparison of the generalization set is presented in Table 2. It should be noted that its performance is consistent with the test result above, with a slight drop in predictive precision. The absolute mean error of PCA-MS-RBF is 0.1105, compared to 0.1641 for MS-RBF, which shows an increase of 30% in prediction accuracy. Maximal errors have a similar result of 0.1727 compared to 0.2417, in which accuracy increases by approximately 30%.

5.2 Models with and without MSA (PCA-MS-RBF and PCA-RBF)

Table 3 summarizes the results predicted by models developed with and without MSA. Multi-scale decomposition of information has the advantage of (i) obtaining much more information and (ii) reducing uncertainty in the system. Here, MSA is carried out using a Daubechies 3 wavelet, with a scale set to 5. The average performance indicates that PCA-MS-RBF yields better results than PCA-RBF, with the absolute mean error of the former being 0.0701, and the latter being 0.1016. An error decrease of approximately 30% is obtained by applying MSA. The maximal errors listed in Table 3 are of the same order of magnitude as the absolute mean errors. The effect of MSA is clearer in the tendency of the measured MI, as shown in Fig.4.

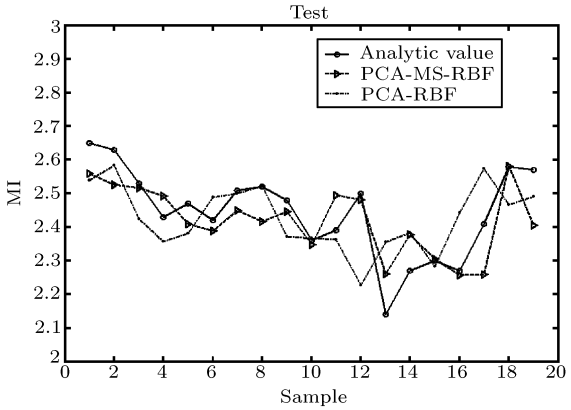

Fig.4 A comparison of measured MI time-records with predictions obtained using PCA-MS-RBF and PCA-RBF

Similarly, a detailed comparison of the generalization set is carried out. Table 4, accordingly shows results for the test set. PCA-MS-RBF yields predictions with an average absolute error of 0.1105, which is much better than the corresponding 0.1719 obtained

Table 2 Absolute and relative mean errors for a generalization set, predicted by models built with a reduced and complete set of process variables

	Mean error			Maximal error		
	Abs	Rel (%)	Comparison (%)	Abs	Rel (%)	Comparison (%)
PCA-MS-RBF	0.1105	4.33	67.55	0.1727	6.72	70.59
MS-RBF	0.1641	6.41	100.00	0.2417	9.52	100.00

Table 3 Absolute and relative mean errors for a test set predicted using models with and without MSA

	Mean error			Maximal error		
	Abs	Rel (%)	Comparison (%)	Abs	Rel (%)	Comparison (%)
PCA-MS-RBF	0.0701	2.87	68.82	0.1720	6.69	57.87
PCA-RBF	0.1016	4.17	100.00	0.2891	11.56	100.00

Table 4 Absolute and relative mean errors for a test set predicted using models with and without MSA

	Mean error			Maximal error		
	Abs	Rel (%)	Comparison (%)	Abs	Rel (%)	Comparison (%)
PCA-MS-RBF	0.1105	4.33	64.63	0.1727	6.72	60.87
PCA-RBF	0.1719	6.70	100.00	0.2859	11.04	100.00

from PCA-RBF. Similar behavior is observed in terms of the maximal error obtained.

6 Conclusions

A neural network multi-scale analysis based methodology, to design and build a virtual sensor to infer product quality from process variables, has been developed and tested. Since the out-performance of a soft-sensor with respect to the linear correlation of variables, is evident in this case, PCA has been proved to be effective in estimating the relevance of certain features; and to select from this quantitative information, the minimum number of variables needed as input to the sensor. The soft-sensor has been trained, using data selected in sequential order from the time records of variables within a pool of available plant information. As a proof-of-concept of the generic virtual sensor model, soft-sensors without the PCA and MSA procedure were also simultaneously tested. The PCA-MS-RBF model, with all process variables measured at the beginning of the production cycle considered as input, predicted MI with a relative mean error of approximately 3% when appropriately trained, compared to an average error of approximately 5 and 4% respectively, obtained from corresponding MS-RBF and PCA-RBF models. The obtained results, indicate that the proposed method provides prediction reliability and accuracy; and is capable of learning relationships between process variables measured at the beginning of a production cycle and the quality of a final product.

References

[1] G. Barto, R. S. Sutton, C. H. Anderson. Neuron-like adaptive elements that can solve difficult learning control problem. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, no. 5, 834–847, 1983.

[2] J. G. Taylor. Towards the networks of the brain: from imaging to consciousness. *Neural Networks*, vol. 12, no. 7-8, pp. 943–959, 1999.

[3] M. Basseville, A. Benveniste, K. Chou, S. Golden, R. Nikoukhah, A. Willsky. Modeling and estimation of multiresolution stochastic process. *IEEE Transactions on Information Theory*, vol. 38, no. 2, 766–784, 1992.

[4] J. H. Li, J. Y. Zhang, W. Ge, X. H. Liu. Multi-scale methodology for complex systems. *Chemical Engineering Science*, vol. 59, no. 8-9, 1687–1700, 2004.

[5] M. K. Gobbert, T. P. Merchant, L. J. Borucki, T. S. Cale. A multiscale simulator for low pressure chemical vapor deposition. *Journal of Electrochemical Society*, vol. 144, no. 11, 3945–3951, 1997.

[6] K. J. D. Sbarbaro.Hunt, R. Zbikowski, P. J. Gawthrop. Neural networks for control systems-a survey. *Automatica*, vol. 28, no. 6, 1083–1112, 1992.

[7] U. Levin, K. S. Narendra. Control of nonlinear dynamical systems using neural networks-part II: observability, identification and control. *IEEE Transactions on Neural Networks*, vol. 7, no. 1, 30–42, 1996.

[8] J. Li, M. Kwauk. Exploring complex systems in chemical engineering - the multi-scale methodology. *Chemical Engineering Science*, vol. 58, no. 3-6, 521–535, 2003.

[9] R. Rallo, J. Ferre-Gine, A. Arenas, F. Giralt. Neural virtual sensor for the inferential prediction of product quality from process variables. *Computers and Chemical Engineering*, vol. 26, no. 12, pp. 1735–1754, 2002.

[10] J. Park, I. W. Sandberg. Universal approximation using radial basis function networks. *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1990.

[11] S. Mallat. A wavelet tour of signal processing. *Academic Press*, San Diego, 1997.

[12] W. M. Chan, P. E. Gloor, A. E. Hamielec. Kinetic model for olefin polymerization in high-pressure autoclave reactors. *AIChE Journal*, vol. 39, no. 1, 111–126, 1993.

Jian Shi received his BSc degree in Mechanical Engineering from Zhejiang University, China, in 2003. He is currently pursuing his PhD degree in the Institute of Systems Engineering of Zhejiang University. His research interests include olefin polymerization modeling, multivariate statistical process control, and neural networks.

Xing-Gao Liu received his PhD degree in Control Science and Engineering, from Zhejiang University in 2000. He was a Post Doctoral Fellow in the Automation Department, at Tsinghua University from 2000 to 2002. Currently, he is an Associate Professor in Process Systems Engineering, at the Department of Control Science and Engineering, Zhejiang University. His research interests include process modeling, control and optimization, fault detection and diagnosis, neural networks in olefin polymerization, and air separation and distillation processes.