

Research Article

Hybrid Artificial Neural Networks Modeling for Faults Identification of a Stochastic Multivariate Process

Yuehjen E. Shao and Chia-Ding Hou

Department of Statistics and Information Science, Fu Jen Catholic University, No. 510, Zhongzheng Road, Xinzhuang, New Taipei City 24205, Taiwan

Correspondence should be addressed to Chia-Ding Hou; stat0002@mail.fju.edu.tw

Received 7 October 2013; Revised 5 November 2013; Accepted 5 November 2013

Academic Editor: Shen Yin

Copyright © 2013 Y. E. Shao and C.-D. Hou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to the recent rapid growth of advanced sensing and production technologies, the monitoring and diagnosis of multivariate process operating performance have drawn increasing interest in process industries. The multivariate statistical process control (MSPC) chart is one of the most commonly used tools for detecting process faults. However, an out-of-control MSPC signal only indicates that process faults have intruded the underlying process. Identifying which of the monitored quality variables is responsible for the MSPC signal is fairly difficult. Pinpointing the responsible variable is vital for process improvement because it effectively determines the root causes of the process faults. Accordingly, this identification has become an important research issue concerning recent multivariate process applications. In contrast with the traditional single classifier approach, the present study proposes hybrid modeling schemes to address problems that involve a large number of quality variables in a multivariate normal process. The proposed scheme includes multivariate adaptive regression splines (MARS), logistic regression (LR), and artificial neural network (ANN). By applying MARS and LR techniques, we may obtain fewer but more significant quality variables, which can serve as inputs to the ANN classifier. The performance of our proposed approaches was evaluated by conducting a series of experiments.

1. Introduction

A multivariate process monitors two or more quality variables. When a signal is triggered by the multivariate statistical process control (MSPC) chart, process personnel are typically only aware that the underlying process is in an unstable state. Identifying which of the monitored quality characteristics (or variables) is responsible for this MSPC signal is challenging. Accordingly, effective determination of the source of process faults becomes an important and challenging issue in MSPC applications, because these sources are associated with specific assignable causes that adversely affect the process.

Typically, a literature review has shown that there are different kinds of approaches to investigate on source identification of faults in a multivariate process. The first type of approach uses various graphical techniques, such as polygonal charts [1], line charts [2], multivariate profile charts [3], and boxplot charts [4] to assist in determining the quality variables at fault in a process. However, the operations of these graphical approaches are tedious and subjective.

The second type of approach uses the statistical decomposition techniques to interpret the contributors to an MSPC signal. Mason et al. [5] proposed the method to decompose the T^2 statistic into independent parts, each of which reflects the contribution of an individual quality variable. Since the decomposition of the T^2 statistic into p independent T^2 components is not unique, Mason et al. [6] provided a computing scheme that can reduce the computational effort. The same concept to decompose the T^2 statistics has been proposed by the studies [7, 8]. However, these approaches have not been analyzed in terms of the percentage of success in the classification of the variables that have actually shifted in the process [9, 10]. The study [11] investigated the method of principal components analysis (PCA) to determine the quality variables at fault in a multivariate process. The T^2 statistic is expressed in terms of normalized principal components scores of the multinormal variables. The normalized score with high values are detected when an MSPC signal is triggered. The contribution plots can then be used to

determine the variables which are responsible for the signal. In addition, the contribution plots were used by the studies [12, 13]. However, the PCA approach can be argued that the dimensionality of data may not be efficiently reduced by linear transformation. Also, the problem of the PCA consists in the fact that the directions maximizing variance do not always maximize information. More recently, the study [14] developed a statistical decomposition method to estimate the sources of process variance shifts in a multivariate normal process. Although the performance of the approach was acceptable, the decomposition method requires a large sample size, which may not be feasible for some practical applications.

The third type of approach employs the machine learning (ML) mechanisms, such as artificial neural networks (ANN) and support vector machine (SVM), to identify the quality variables which are responsible for the MSPC signal. A comparative study has been conducted by the studies [9, 10]. While the study [9] made a comparison between neural network approaches with the method of Mason et al. [5], the study [10] made a comparison between ANN and SVM with the method of Runger et al. [8]. Both studies [9, 10] concluded that ML methods are in general better than those obtained using the decomposition approach. The study [15] proposed a backpropagation-net based model which can identify the group of quality variables at faults and can classify the magnitude of the process shifts. The study [16] developed a two level-based model using T^2 control chart for detecting the signals and an ANN for identifying the sources of the signals. The study [17] proposed an ANN-based model to identify and quantify the mean shifts in bivariate processes. The authors [18] developed a neural-network-based identifier to detect the mean shifts and simultaneously to identify the sources of the shifts for a multivariate autocorrelated processes. They benchmarked the run-length performance of the proposed method against the Hotelling T^2 , the MEWMA, and the Z control charts. The authors [19] investigated the sources of process variance faults with the use of ANN and SVM; however, their considerations of process variance shifts were large. The authors [20] proposed a hybrid model for online analysis of MSPC signals in multivariate manufacturing processes. Their model consisted of two modules in which the first module used a SVM to recognize the unnatural pattern, and then, the magnitude of different shifts can be determined by using the second module, the NN models. The authors [21] also proposed a hybrid model for online analysis of MSPC signals in multivariate manufacturing processes. They also used the SVM to recognize the mean and variance shifts in the first module. In the second module, they employed two neural network models to recognize the magnitude of shifts for each variable simultaneously. The study [22] proposed a hybrid scheme which is composed of independent component analysis (ICA) and SVM to decide the fault quality variables when a step change disturbance existed in a multivariate process.

The literature review has shown that most of the existing studies are concerned with the determination of which variable or group of variables has caused the signal through

single step modeling. However, there is a difficulty that may not have been addressed yet. When the number of quality characteristics is large, the existing decomposition methods and/or machine learning methods may lack the capability to handle such a situation. In addition, because process faults are typically attributed to mean shifts and the multivariate normal process is one of the most widely used applications, the present study is motivated by addressing mean shift faults for a multivariate normal process with a large number of quality variables. A review of relevant literature also indicates that the application of ANN for process fault determination is promising; however, it suffers from the requirement of a large number of controlling parameters and the risk of model overfitting [23–25]. Consequently, contrary to the existing approaches, the present study proposes two-stage hybrid schemes to identify which quality variable or group of variables is responsible for process mean shift faults. The proposed schemes integrate multivariate adaptive regression splines (MARS), logistic regression (LR), and artificial neural networks, which are referred to as the MARS-ANN and LR-ANN schemes, respectively. The performance of the proposed approaches was examined by a series of computer simulations.

The rest of this paper is organized as follows. Section 2 provides brief overviews of process models and the proposed schemes. The various experimental conditions are addressed in Section 3. This study is concluded in Section 4.

2. Process Models and Methodologies

The structure of the process model is addressed. The proposed hybrid schemes are also described in this section.

2.1. Structure of the Process and the Mean Shift. This study considers the situation of process mean shifts and assumes that the multivariate process is initially in a normal state and the sample observations are derived from a k -dimensional multivariate normal distribution $N(\underline{\mu}_{\sim 0}, \Sigma_0)$, where

$$\underline{\mu}_{\sim 0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (1)$$

$$\Sigma_0 = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix}. \quad (2)$$

After a certain length of time, this study assumes that the mean vector changes from $\underline{\mu}_{\sim 0}$ to $\underline{\mu}_{\sim 1}$, where

$$\underline{\mu}_{\sim 1} = \underline{\mu}_{\sim 0} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_k \end{bmatrix}. \quad (3)$$

Let

$$\underset{\sim ij}{X} = [X_{ij1}, X_{ij2}, \dots, X_{ijk}]', \quad j = 1, 2, \dots, n \quad (4)$$

be a $k \times 1$ vector that represents k characteristics on the j th observation in subgroup i . The resulting sample mean vector is as follows:

$$\bar{\underset{\sim i}{X}} = \frac{1}{n} \sum_{j=1}^n \underset{\sim ij}{X}. \quad (5)$$

To detect a multivariate process mean shift, Hotelling [26] proposed the following chi-square statistic

$$\chi_i^2 = n \left(\bar{\underset{\sim i}{X}} - \underset{\sim 0}{\mu} \right)' \sum_0^{-1} \left(\bar{\underset{\sim i}{X}} - \underset{\sim 0}{\mu} \right). \quad (6)$$

This statistic is asymptotically distributed as a chi-square distribution with k degrees of freedom. The control chart that uses χ^2 as a monitoring statistic in (6) has the upper control limit

$$\text{UCL} = \chi_\alpha^2(k), \quad (7)$$

where $\chi_\alpha^2(k)$ is the upper α th percentile of the chi-square distribution with k degrees of freedom. If the plotted statistic χ_i^2 falls outside the UCL, the process is considered to be in an abnormal state, and our proposed method can be applied to identify the source of mean shifts. The proposed two-stage hybrid methods integrate the framework of MARS, LR, and ANN. In the initial stage, influencing variables are selected using multivariate adaptive regression splines or logistic regression. In the second stage, the significant influencing variables selected are taken as the input variables of the ANN. The following sections address these three components.

2.2. Logistic Regression. The purpose of performing logistic regression modeling in stage I was to identify important influencing variables and refine the entire set of input variables. The structure of the logistic regression model can be briefly described as follows. Let Y_1, Y_2, \dots, Y_n represent the dependent variables ($Y_i = 1$ denotes “the abnormal state” and $Y_i = 0$ denotes “the normal state”) and let

$$P_i = \Pr [Y_i = 1 | \bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{ik}] \quad (8)$$

be the conditional probability of event $\{Y_i = 1\}$ with a given series of independent variables $(\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{ik})$, where \bar{x}_{im} is the sample mean of the m th characteristic. The logistic regression model is then defined as follows:

$$\ln \left(\frac{P_i}{1 - P_i} \right) = \beta_0 + \sum_{j=1}^k \beta_j \bar{x}_{ij}. \quad (9)$$

Before screening significant independent variables, we performed the collinearity diagnosis procedure to exclude variables that exhibited high collinearity. After this diagnosis, the remaining variables served as independent variables for

logistic regression modeling and testing. The Wald forward method was applied to identify independent variables with significant influence on an abnormal state probability. These significant independent variables and the dependent variable were then substituted into the ANN to construct a two-stage model.

2.3. Multivariate Adaptive Regression Splines. The superior performance of the MARS has been reported in many applications [27–32]. MARS is typically capable of revealing important data patterns and relationships for the complex data structure that is often concealed in high-dimensional data [28, 31]. The MARS model can be represented as [33]

$$\hat{f}(x) = b_0 + \sum_{m=1}^M b_m \prod_{k=1}^{K_m} [S_{km}(x_{\nu(k,m)} - t_{km})], \quad (10)$$

where b_0 and b_m are the parameters, M is the number of basis functions (BF), K_m is the number of knots, S_{km} takes on values of either 1 or -1 and indicates the right or left sense of the associated step function, $\nu(k, m)$ is the label of the independent variable, and t_{km} is the knot location. The optimal MARS model is obtained in two steps. The purpose of the first step is to construct a large number of basis functions that initially fit the data. The purpose of the second step is to delete basis functions in order of least contribution using the generalized cross-validation (GCV) criterion. The variable importance measure was obtained by observing the decrease in the calculated GCV values when a variable was removed from the model. The GCV is described as

$$\text{GCV}(M) = \frac{1/N \sum_{i=1}^N [y_i - \hat{f}_M(x_i)]^2}{[1 - C(M)/N]^2}, \quad (11)$$

where N is the number of observations and $C(M)$ is the cost penalty measure of a model containing M basis functions.

2.4. The Artificial Neural Network. The ANN has been widely used in many SPC applications [34, 35]. The ANN is a parallel system comprised of highly interconnected processing elements that are based on neurobiological models. The ANN processes information through the interactions of a large number of simple processing elements called neurons.

Figure 1 illustrates that neurons in networks take inputs from the previous layer and send outputs to the next layer. Typically, ANN nodes consist of three layers: the input, output, and hidden layers. The nodes in the input layers receive input signals from an external source and the nodes in the output layers generate the target output signals. The output of each neuron in the input layer is the same as the input to that neuron. For each neuron j in the hidden layer and neuron k in the output layer, the net inputs are given by

$$\text{net}_j = \sum_i w_{ji} \times o_i, \quad \text{net}_k = \sum_j w_{kj} \times o_j, \quad (12)$$

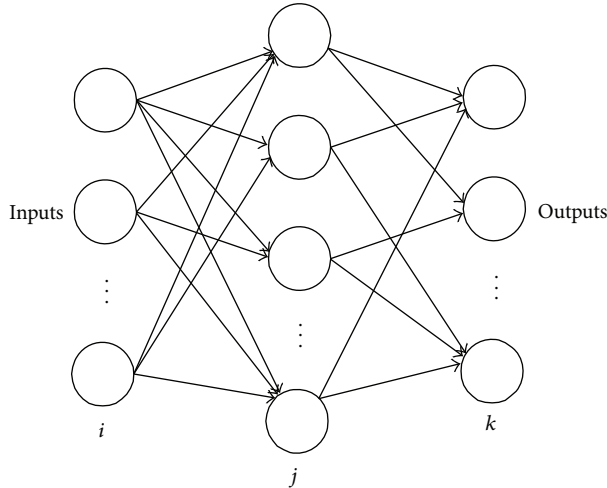


FIGURE 1: Structure of ANN model.

where $i(j)$ is a neuron in the previous layer, $o_i(o_j)$ is the output of node $i(j)$, and $w_{ji}(w_{kj})$ is the connection weight from neuron $i(j)$ to neuron $j(k)$. The neuron outputs are given by

$$o_i = \text{net}_i, \quad (13)$$

$$o_j = \frac{1}{1 + \exp^{-(\text{net}_j + \theta_j)}} = f_j(\text{net}_j, \theta_j), \quad (14)$$

$$o_k = \frac{1}{1 + \exp^{-(\text{net}_k + \theta_k)}} = f_k(\text{net}_k, \theta_k),$$

where $\text{net}_j(\text{net}_k)$ is the input signal from the external source to the node $j(k)$ in the input layer and $\theta_j(\theta_k)$ is a bias. The transformation function shown in (14) is called a sigmoid function and is the most commonly utilized function to date. As a result, this study used the sigmoid function.

3. Experiments and Analysis

3.1. The Parameter Settings. To evaluate the performance of the proposed approach, a series of simulations were conducted. Without loss of generality, this study assumed that each quality characteristic was initially sampled from a normal distribution with zero mean and one standard deviation. In addition, we assumed that twenty quality characteristics were monitored simultaneously (i.e., $k = 20$), and the covariance matrix was defined as in (2).

Because we considered 20 quality characteristics for the multivariate normal process, there are $2^{20} - 1$ possible types of mean shifts. They are represented by $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0), \dots$, and $(1, 1, \dots, 1)$, where 1 denotes a quality characteristic that is at fault and 0 denotes a quality characteristic that is not at fault. For an abnormal mean vector structure, we considered three types of mean shifts for demonstration: $(1, 0, \dots, 0)$, $(1, 0, 1, 0, \dots, 0)$ and $(1, 0, 1, 0, 1, 0, \dots, 0)$. This study also considered three different values of ρ : 0.1, 0.5, and 0.9. The sample size was assumed to be 10. Two values of δ_i were considered: 0.5 and 1.0. We repeated the simulation 500 times for each data structure. The structure of the ANN

is established as follows. When applying ANN in the single stage in this study, we had 20 input nodes and one output node in the ANN structure. The hidden nodes were set to the range $i - 2$ to $i + 2$, where i is the number of input variables. Thus, in the initial phase, the hidden nodes were 18, 19, 20, 21, and 22.

According to the suggestions of the study [36], the learning rates were set to 0.01, 0.005, and 0.001. After performing ANN modeling, we obtained the $\{20 - 20 - 1\}$ topology with a learning rate of 0.01, which provides the best result with the minimum test RMSE. Here, $\{n_i - n_h - n_o\}$ denotes the number of neurons in the input layer, number of neurons in the hidden layer, and number of neurons in the output layer, respectively.

3.2. The Results. For the hybrid LR-ANN model, this study calculated the variance inflation factor (VIF) to examine the presence of collinearity, used a 0.05 significance level, and employed logistic regression analysis to select important influencing variables in the initial stage. Values of VIFs greater than 10 were considered large enough to suspect serious multicollinearity [37–39]. As shown in Table 1, all of the VIFs are less than 10. Consequently, collinearity was not too high among the independent variables. The analysis results of LR modeling are summarized in Table 2. The significant variables selected in this stage served as the input variables of the ANN.

For the hybrid MARS-ANN model, we obtained the selection results of the variables after performing the MARS procedure. Tables 3, 4, 5, 6, 7, and 8 list the selection results for the MARS models for 6 different combinations of ρ and δ_i . In this selection procedure, the important explanatory variables were chosen; their relative importance indicators are listed in the last column of Tables 3 to 8.

When the first stage of hybrid modeling was completed, the ANN topology settings were established. Table 9 displays the corresponding ANN topologies for various types of hybrid models. The network topology with the minimum test RMSE was again considered as the optimal network topology. The learning rate of 0.01 was used for all of those models.

This study used the classical single stage of an ANN model and the proposed two-stage of MARS-ANN and LR-ANN models to determine the source of mean shift faults in a multivariate process. The experimental results are displayed in Table 10.

Table 10 reveals that the two-stage MARS-ANN and LR-ANN approaches exhibit better performance than the classical single-stage ANN method in many situations. Based on the results shown in Table 10, it is noted that when the type of mean shift is $(1, 0, 1, 0, 1, 0, \dots, 0)$, the LR-ANN approach exhibits the best performance in terms of accurate identification rates (AIR) for all (ρ, δ_i) combinations. The MARS-ANN approach was preferable to the single stage of the ANN in almost every case. The last two rows of Table 10 list the average and standard errors of the accurate identification rates. The proposed hybrid approaches, LR-ANN and MARS-ANN, outperformed the classical method, which is the single stage of the ANN. The proposed MARS-ANN approach had

TABLE 1: Collinearity diagnosis for LR models.

Variables	$\rho = 0.1$		$\rho = 0.5$		$\rho = 0.9$	
	$\delta_i = 0.5$	$\delta_i = 1.0$	$\delta_i = 0.5$	$\delta_i = 1.0$	$\delta_i = 0.5$	$\delta_i = 1.0$
\bar{x}_1	1.11	1.23	1.65	1.56	3.66	2.26
\bar{x}_2	1.09	1.09	1.95	1.95	9.85	9.85
\bar{x}_3	1.21	1.54	1.80	1.97	3.98	2.89
\bar{x}_4	1.07	1.07	1.91	1.91	9.54	9.54
\bar{x}_5	1.21	1.55	1.81	2.00	4.03	2.91
\bar{x}_6	1.07	1.07	1.88	1.88	9.31	9.31
\bar{x}_7	1.13	1.28	1.69	1.63	3.78	2.33
\bar{x}_8	1.11	1.11	1.98	1.98	9.82	9.82
\bar{x}_9	1.07	1.07	1.89	1.88	9.21	9.21
\bar{x}_{10}	1.05	1.05	1.81	1.80	9.21	9.19
\bar{x}_{11}	1.07	1.07	1.90	1.90	9.43	9.41
\bar{x}_{12}	1.07	1.07	1.89	1.89	9.60	9.59
\bar{x}_{13}	1.07	1.07	1.91	1.90	9.84	9.83
\bar{x}_{14}	1.06	1.06	1.83	1.82	9.13	9.11
\bar{x}_{15}	1.08	1.08	1.92	1.92	9.62	9.60
\bar{x}_{16}	1.07	1.06	1.93	1.92	9.95	9.93
\bar{x}_{17}	1.07	1.07	1.84	1.83	8.93	8.92
\bar{x}_{18}	1.09	1.09	1.96	1.96	9.78	9.78
\bar{x}_{19}	1.09	1.09	1.94	1.93	9.69	9.68
\bar{x}_{20}	1.09	1.09	1.97	1.96	9.64	9.63

TABLE 2: Significant variables selected by LR analysis.

Correlation ρ	Shift value δ_i	Significant explanatory variables
0.1	0.5	$\bar{x}_1, \bar{x}_3, \bar{x}_5, \bar{x}_9, \bar{x}_{12}, \bar{x}_{17}, \bar{x}_{18}, \bar{x}_{20}$
	1.0	$\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_5, \bar{x}_6, \bar{x}_{11}, \bar{x}_{12}, \bar{x}_{13}, \bar{x}_{16}, \bar{x}_{17}, \bar{x}_{20}$
0.5	0.5	$\bar{x}_1, \bar{x}_3, \bar{x}_4, \bar{x}_5, \bar{x}_8, \bar{x}_9, \bar{x}_{11}, \bar{x}_{12}, \bar{x}_{13}, \bar{x}_{16}$
	1.0	$\bar{x}_1, \bar{x}_3, \bar{x}_6, \bar{x}_8, \bar{x}_{12}, \bar{x}_{13}, \bar{x}_{15}, \bar{x}_{16}, \bar{x}_{20}$
0.9	0.5	$\bar{x}_1, \bar{x}_3, \bar{x}_8, \bar{x}_9, \bar{x}_{12}, \bar{x}_{13}, \bar{x}_{16}, \bar{x}_{20}$
	1.0	\bar{x}_1, \bar{x}_{13}

the smallest standard error, which implies the robustness of the mechanism. After comparing the performances of the LR-ANN and MARS-ANN approaches, we determined that the MARS-ANN approach is superior. The reason may be that population stratification in logistic regression analysis can lead to bias in estimates and test statistics. As a result, the results of the LR-ANN approach were somewhat unstable.

Table 11 summarizes the AIR with consideration of three different correlations, namely, the low, the moderate, and the strong correlations, respectively. The standard deviations for those AIR values are listed in parentheses. By observing Table 11, one is able to observe that the performance of the proposed hybrid models almost completely outperforms the classical single-stage ANN model. In particular, the proposed

MARS-ANN has the best and the most robust performance among those three modeling approaches.

Table 12 shows the overall improvement percentage of the proposed model in comparison with the classical single-stage model. The AIR improvements of the proposed LR-ANN model over the classical ANN model for three types of correlations are 18.73%, 10.67%, and -2.50%, respectively. Although there is a poor improvement for the case of $\rho = 0.9$, the average AIR improvement is 8.97%. In addition, the AIR improvements of the proposed MARS-ANN model over the classical ANN model for three types of correlations are 14.39%, 15.85%, and 6.96%, respectively. Accordingly, the average AIR improvement reaches 12.73%.

One important result is that our proposed approach is useful in dealing with difficulties of the smaller shifts for a multivariate process. The case of the smaller shift value (i.e., $\delta_i = 0.5$) drew particular attention from industries because it is very difficult to identify the sources of small mean shifts. Considering all the cases of $\delta_i = 0.5$, Table 10 illustrates that the 21.12% and 17.00% improvement in identification can be achieved when the proposed LR-ANN and MARS-ANN schemes are used. The improvements in identification are significant.

4. Conclusions

The ANN has been criticized for its long training process; however, the combination of LR/MARS and ANN is a good alternative for performing classification tasks. Accordingly,

TABLE 3: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.1, 0.5)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.501	1.070	1	\bar{x}_3	100.000
2	0.500	1.065	3	\bar{x}_5	98.993
3	0.339	0.963	1	\bar{x}_1	70.614
4	0.332	0.957	1	\bar{x}_7	68.366
5	0.131	0.873	1	\bar{x}_{13}	26.827
6	0.104	0.867	1	\bar{x}_{16}	20.879
7	0.099	0.866	1	\bar{x}_{14}	19.621
8	0.093	0.865	1	\bar{x}_{17}	18.071

TABLE 4: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.1, 1.0)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.517	0.372	3	\bar{x}_5	100.000
2	0.519	0.370	2	\bar{x}_3	99.356
3	0.395	0.328	2	\bar{x}_1	85.740
4	0.369	0.310	1	\bar{x}_7	79.157
5	0.066	0.211	1	\bar{x}_{20}	15.440

TABLE 5: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.5, 0.5)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.596	0.871	2	\bar{x}_3	100.000
2	0.582	0.864	2	\bar{x}_5	98.281
3	0.380	0.764	2	\bar{x}_7	68.774
4	0.369	0.763	1	\bar{x}_1	68.220
5	0.245	0.702	1	\bar{x}_{13}	40.739
6	0.234	0.700	1	\bar{x}_{16}	39.271
7	0.203	0.692	2	\bar{x}_{17}	33.836
8	0.195	0.691	1	\bar{x}_{14}	32.728
9	0.188	0.688	1	\bar{x}_{20}	30.632
10	0.186	0.688	1	\bar{x}_6	30.530

TABLE 6: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.5, 1.0)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.540	0.352	1	\bar{x}_3	100.000
2	0.535	0.344	2	\bar{x}_5	97.416
3	0.389	0.306	1	\bar{x}_1	83.637
4	0.380	0.294	1	\bar{x}_7	78.935
5	0.174	0.217	2	\bar{x}_{13}	35.183
6	0.167	0.216	2	\bar{x}_{16}	33.487
7	0.149	0.212	1	\bar{x}_{20}	29.230
8	0.143	0.211	1	\bar{x}_{14}	29.042

the proposed combination of the LR-ANN and MARS-ANN schemes was proven to be useful for determining the mean shift faults in a multivariate process.

The rationale behind the proposed schemes was initially to obtain fewer important explanatory variables by performing LR or MARS modeling. The resulting significant variables served as inputs to the designed ANN models. The proposed

LR-ANN and MARS-ANN models not only have fewer input variables but also possess better classification capabilities.

The proposed hybrid two-stage models in this study are not the only combination techniques; other artificial intelligence techniques, such as decision tree or genetic algorithms, can be integrated with neural networks or a support vector machine to further refine the structure of the classifiers and

TABLE 7: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.9, 0.5)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.789	0.363	2	\bar{x}_5	100.000
2	0.761	0.356	2	\bar{x}_3	98.013
3	0.570	0.304	2	\bar{x}_1	79.033
4	0.531	0.289	2	\bar{x}_7	72.994
5	0.412	0.225	1	\bar{x}_{13}	35.419
6	0.402	0.225	1	\bar{x}_{16}	34.692
7	0.349	0.221	1	\bar{x}_{14}	30.760
8	0.315	0.218	1	\bar{x}_6	27.641
9	0.317	0.218	1	\bar{x}_{10}	27.426
10	0.306	0.217	1	\bar{x}_{17}	26.965

TABLE 8: Basis functions and important explanatory variables for the MARS model with $(\rho, \delta_i) = (0.9, 1.0)$.

Function	Std. dev.	Cost of omission	Number of BF	Variable	Relative importance (%)
1	0.579	0.179	2	\bar{x}_5	100.000
2	0.565	0.176	2	\bar{x}_3	98.746
3	0.450	0.160	2	\bar{x}_1	91.803
4	0.434	0.151	2	\bar{x}_7	87.535
5	0.295	0.072	2	\bar{x}_{13}	31.978
6	0.293	0.072	2	\bar{x}_{16}	31.287
7	0.249	0.069	1	\bar{x}_{14}	26.538
8	0.222	0.067	1	\bar{x}_6	23.337

TABLE 9: ANN topology settings for different hybrid LR-ANN and MARS-ANN models.

Correlation ρ	Shift value δ_i	Type of mean shifts	ANN topology for LR-ANN	ANN topology for MARS-ANN
0.1	0.5	(1, 0, ..., 0)	{8-8-1}	{8-9-1}
		(1, 0, 1, 0, ..., 0)	{8-8-1}	{8-7-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{8-8-1}	{8-10-1}
	1.0	(1, 0, ..., 0)	{11-13-1}	{5-5-1}
		(1, 0, 1, 0, ..., 0)	{11-12-1}	{5-6-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{11-12-1}	{5-7-1}
0.5	0.5	(1, 0, ..., 0)	{10-10-1}	{10-10-1}
		(1, 0, 1, 0, ..., 0)	{10-10-1}	{10-10-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{10-10-1}	{10-10-1}
	1.0	(1, 0, ..., 0)	{9-9-1}	{8-9-1}
		(1, 0, 1, 0, ..., 0)	{9-7-1}	{8-8-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{9-11-1}	{8-9-1}
0.9	0.5	(1, 0, ..., 0)	{8-9-1}	{10-9-1}
		(1, 0, 1, 0, ..., 0)	{8-10-1}	{10-10-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{8-9-1}	{10-11-1}
	1.0	(1, 0, ..., 0)	{2-2-1}	{8-6-1}
		(1, 0, 1, 0, ..., 0)	{2-2-1}	{8-8-1}
		(1, 0, 1, 0, 1, 0, ..., 0)	{2-2-1}	{8-7-1}

improve classification accuracy. The applications of other process faults, such as variance shift faults, for a multivariate process should be further investigated.

The data-driven methods of multivariate statistical process control have been the subject of considerable interest from both the academic community and industry as an

important implement in the process monitoring area. Since the practical systems become more and more complicated and the physical models become extremely hard to obtain, considering the related topics within data-driven framework seems more meaningful in the current and future work to achieve more industrial oriented results [40–44]. In addition,

TABLE 10: Comparison of classification accuracy among the ANN, LR-ANN, and MARS-ANN models.

Correlation ρ	Shift value δ_i	Type of mean shifts	ANN	LR-ANN	MARS-ANN
0.1	0.5	(1, 0, ..., 0)	60.80%	61.20%	58.40%
		(1, 0, 1, 0, ..., 0)	60.80%	59.20%	62.40%
		(1, 0, 1, 0, 1, 0, ..., 0)	56.00%	70.40%	58.80%
	1.0	(1, 0, ..., 0)	58.00%	80.40%	85.60%
		(1, 0, 1, 0, ..., 0)	86.40%	84.80%	81.20%
		(1, 0, 1, 0, 1, 0, ..., 0)	56.00%	92.80%	86.00%
0.5	0.5	(1, 0, ..., 0)	38.80%	66.00%	59.60%
		(1, 0, 1, 0, ..., 0)	46.00%	55.60%	62.00%
		(1, 0, 1, 0, 1, 0, ..., 0)	46.40%	70.80%	64.80%
	1.0	(1, 0, ..., 0)	90.40%	91.60%	94.00%
		(1, 0, 1, 0, ..., 0)	94.40%	63.60%	94.40%
		(1, 0, 1, 0, 1, 0, ..., 0)	85.20%	96.40%	94.00%
0.9	0.5	(1, 0, ..., 0)	82.40%	93.20%	81.30%
		(1, 0, 1, 0, ..., 0)	94.40%	82.80%	94.40%
		(1, 0, 1, 0, 1, 0, ..., 0)	56.00%	96.80%	92.00%
	1.0	(1, 0, ..., 0)	98.80%	44.00%	100.00%
		(1, 0, 1, 0, ..., 0)	100.00%	100.00%	100.00%
		(1, 0, 1, 0, 1, 0, ..., 0)	98.40%	100.00%	99.20%
Average of accurate identification rates			72.73%	78.31%	81.57%
Standard error of accurate identification rates			0.05007	0.04092	0.03767

TABLE 11: AIR comparison of the classical-single stage and the proposed models.

	ANN	Proposed LR-ANN	Proposed MARS-ANN
$\rho = 0.1$	63.00% (0.1166)	74.80% (0.1344)	72.07% (0.1354)
$\rho = 0.5$	66.87% (0.2565)	74.00% (0.1632)	78.13% (0.1760)
$\rho = 0.9$	88.33% (0.1711)	86.13% (0.2161)	94.48% (0.0725)

TABLE 12: AIR improvement of the proposed models in comparison with classical ANN model.

	Proposed LR-ANN	Proposed MARS-ANN
$\rho = 0.1$	18.73%	14.39%
$\rho = 0.5$	10.67%	16.85%
$\rho = 0.9$	-2.50%	6.96%
Average improvement	8.97%	12.73%

real-time implementation of fault tolerant control system with performance optimization is an important issue in modern industries [45]. Extensions of the proposed procedures to data-driven design or real-time implementation of fault tolerant control system are possible. Such works deserve further research and are our future concern.

Acknowledgment

This work is partially supported by the National Science Council of the Republic of China, Grant no. NSC 102-2221-E-030-019 and Grant no. NSC 102-2118-M-030-001.

References

- [1] L. W. Blazek, B. Novic, and M. D. Scott, "Displaying multivariate data using polyplots," *Journal of Quality Technology*, vol. 19, no. 2, pp. 69–74, 1987.
- [2] N. Subramanyan and A. A. Houshmand, "Simultaneous representation of multivariate and corresponding univariate x-bar charts using a line graph," *Quality Engineering*, vol. 7, no. 4, pp. 681–682, 1995.
- [3] C. Fuchs and Y. Benjamini, "Multivariate profile charts for statistical process control," *Technometrics*, vol. 36, no. 2, pp. 182–195, 1994.
- [4] O. O. Atienza, L. T. Ching, and B. A. Wah, "Simultaneous monitoring of univariate and multivariate SPC information using boxplots," *International Journal of Quality Science*, vol. 3, no. 2, pp. 194–204, 1998.
- [5] R. L. Mason, N. D. Tracy, and J. C. Young, "Decomposition of T^2 for multivariate control chart interpretation," *Journal of Quality Technology*, vol. 27, no. 2, pp. 99–108, 1995.
- [6] R. L. Mason, N. D. Tracy, and J. C. Young, "A practical approach for interpreting multivariate T^2 control chart signals," *Journal of Quality Technology*, vol. 29, no. 4, pp. 396–406, 1997.
- [7] N. H. Timm, "Multivariate quality control using finite intersection tests," *Journal of Quality Technology*, vol. 28, no. 2, pp. 233–243, 1996.

- [8] G. C. Runger, F. B. Alt, and D. C. Montgomery, "Contributors to a multivariate statistical process control chart signal," *Communications in Statistics*, vol. 25, no. 10, pp. 2203–2213, 1996.
- [9] F. Aparisi, G. Avendaño, and J. Sanz, "Techniques to interpret T² control chart signals," *IIE Transactions*, vol. 38, no. 8, pp. 647–657, 2006.
- [10] Y. E. Shao and B.-S. Hsu, "Determining the contributors for a multivariate SPC chart signal using artificial neural networks and support vector machine," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 12, pp. 4899–4906, 2009.
- [11] T. Kourti and J. F. MacGregor, "Multivariate SPC methods for process and product monitoring," *Journal of Quality Technology*, vol. 28, no. 4, pp. 409–428, 1996.
- [12] J. A. Westerhuis, S. P. Gurden, and A. K. Smilde, "Generalized contribution plots in multivariate statistical process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 51, no. 1, pp. 95–114, 2000.
- [13] P. E. Maravelakis, S. Bersimis, J. Panaretos, and S. Psarakis, "Identifying the out of control variable in a multivariate control chart," *Communications in Statistics. Theory and Methods*, vol. 31, no. 12, pp. 2391–2408, 2002.
- [14] Y. E. Shao, C.-D. Hou, C.-H. Chao, and Y.-J. Chen, "A decomposition approach for identifying the sources of variance shifts in a multivariate process," *ICIC Express Letters*, vol. 5, no. 4A, pp. 971–975, 2011.
- [15] S. T. A. Niaki and B. Abbasi, "Fault diagnosis in multivariate control charts using artificial neural networks," *Quality and Reliability Engineering International*, vol. 21, no. 8, pp. 825–840, 2005.
- [16] R.-S. Guh, "On-line identification and quantification of mean shifts in bivariate processes using a neural network-based approach," *Quality and Reliability Engineering International*, vol. 23, no. 3, pp. 367–385, 2007.
- [17] L.-H. Chen and T.-Y. Wang, "Artificial neural networks to classify mean shifts from multivariate χ^2 chart signals," *Computers and Industrial Engineering*, vol. 47, no. 2-3, pp. 195–205, 2004.
- [18] H. B. Hwang and Y. Wang, "Shift detection and source identification in multivariate autocorrelated processes," *International Journal of Production Research*, vol. 48, no. 3, pp. 835–859, 2010.
- [19] C.-S. Cheng and H.-P. Cheng, "Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines," *Expert Systems with Applications*, vol. 35, no. 1-2, pp. 198–206, 2008.
- [20] M. Salehi, A. Bahreininejad, and I. Nakhai, "On-line analysis of out-of-control signals in multivariate manufacturing processes using a hybrid learning-based model," *Neurocomputing*, vol. 74, no. 12-13, pp. 2083–2095, 2011.
- [21] M. Salehi, R. B. Kazemzadeh, and A. Salmasnia, "On line detection of mean and variance shift using neural networks and support vector machine in multivariate processes," *Applied Soft Computing*, vol. 12, no. 9, pp. 2973–2984, 2012.
- [22] Y. E. Shao, C.-J. Lu, and Y.-C. Wang, "A hybrid ICA-SVM approach for determining the quality variables at fault in a multivariate process," *Mathematical Problems in Engineering*, vol. 2012, Article ID 284910, 12 pages, 2012.
- [23] L. Cao, "Support vector machines experts for time series forecasting," *Neurocomputing*, vol. 51, pp. 321–339, 2003.
- [24] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1506–1518, 2003.
- [25] C.-J. Lu, Y. E. Shao, and P.-H. Li, "Mixture control chart patterns recognition using independent component analysis and support vector machine," *Neurocomputing*, vol. 74, no. 11, pp. 1908–1914, 2011.
- [26] H. Hotelling, *Techniques of Statistical Analysis*, McGraw Hill, New York, NY, USA, 1947.
- [27] Q.-S. Xu, D. L. Massart, Y.-Z. Liang, and K.-T. Fang, "Two-step multivariate adaptive regression splines for modeling a quantitative relationship between gas chromatography retention indices and molecular descriptors," *Journal of Chromatography A*, vol. 998, no. 1-2, pp. 155–167, 2003.
- [28] S.-M. Chou, T.-S. Lee, Y. E. Shao, and I.-F. Chen, "Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines," *Expert Systems with Applications*, vol. 27, no. 1, pp. 133–142, 2004.
- [29] Y. Zhou and H. Leung, "Predicting object-oriented software maintainability using multivariate adaptive regression splines," *Journal of Systems and Software*, vol. 80, no. 8, pp. 1349–1361, 2007.
- [30] F. Vidoli, "Evaluating the water sector in Italy through a two stage method using the conditional robust nonparametric frontier and multivariate adaptive regression splines," *European Journal of Operational Research*, vol. 212, no. 3, pp. 583–595, 2011.
- [31] Y. E. Shao and C.-D. Hou, "Change point determination for a multivariate process using a two-stage hybrid scheme," *Applied Soft Computing Journal*, vol. 13, no. 3, pp. 1520–1527, 2013.
- [32] V. L. Pilla, J. M. Rosenberger, V. Chen, N. Engsuwan, and S. Siddappa, "A multivariate adaptive regression splines cutting plane approach for solving a two-stage stochastic programming fleet assignment model," *European Journal of Operational Research*, vol. 216, no. 1, pp. 162–171, 2012.
- [33] J. H. Friedman, "Multivariate adaptive regression splines," *The Annals of Statistics*, vol. 19, no. 1, pp. 1–141, 1991.
- [34] C.-C. Chiu, Y. E. Shao, T.-S. Lee, and K.-M. Lee, "Identification of process disturbance using SPC/EPC and neural networks," *Journal of Intelligent Manufacturing*, vol. 14, no. 3-4, pp. 379–388, 2003.
- [35] Y. E. Shao and C.-D. Hou, "Fault identification in industrial processes using an integrated approach of neural network and analysis of variance," *Mathematical Problems in Engineering*, vol. 2013, Article ID 516760, 7 pages, 2013.
- [36] B. Abbasi, "A neural network applied to estimate process capability of non-normal processes," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3093–3100, 2009.
- [37] F. A. Graybill and H. K. Iyer, *Regression Analysis*, International Thomson Publishing, 1994.
- [38] J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, *Multivariate Data Analysis*, Prentice-Hall, Upper Saddle River, NJ, USA, 5th edition, 1998.
- [39] Y. E. Shao, "Prediction of currency volume issued in Taiwan using a hybrid artificial neural network and multiple regression approach," *Mathematical Problems in Engineering*, vol. 2013, Article ID 676742, 9 pages, 2013.
- [40] H. Wang, T.-Y. Chai, J.-L. Ding, and M. Brown, "Data driven fault diagnosis and fault tolerant control: some advances and possible new directions," *Acta Automatica Sinica*, vol. 35, no. 6, pp. 739–747, 2009.
- [41] S. Yin and S. X. Ding, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012.

- [42] S. Yin, X. Yang, and H. R. Karimi, "Data-driven adaptive observer for fault diagnosis," *Mathematical Problems in Engineering*, vol. 2012, Article ID 832836, 21 pages, 2012.
- [43] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annual Reviews in Control*, vol. 36, no. 2, pp. 220–234, 2012.
- [44] S. Yin, S. X. Ding, A. H. A. Sari, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.
- [45] S. Yin, H. Luo, and S. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 2402–2411, 2013.