Research Article

Fault Identification in Industrial Processes Using an Integrated Approach of Neural Network and Analysis of Variance

Yuehjen E. Shao and Chia-Ding Hou

Department of Statistics and Information Science, Fu Jen Catholic University, 510 Chungcheng Road, Xinzhuang District, New Taipei City 24205, Taiwan

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

Received 21 November 2012; Revised 28 April 2013; Accepted 14 May 2013

Academic Editor: Jun Zhao

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 its importance in process improvement, the issue of determining exactly when faults occur has attracted considerable attention in recent years. Most related studies have focused on the use of the maximum likelihood estimator (MLE) method to determine the fault in univariate processes, in which the underlying process distribution should be known in advance. In addition, most studies have been devoted to identifying the faults of process mean shifts. Different from most of the current research, the present study proposes an effective approach to identify the faults of variance shifts in a multivariate process. The proposed mechanism comprises the analysis of variance (ANOVA) approach, a neural network (NN) classifier, and an identification strategy. To demonstrate the effectiveness of our proposed approach, a series of simulated experiments is conducted, and the best results from our proposed approach are addressed.

1. Introduction

Process personnel have always wanted to search for process faults in real time to significantly improve the underlying process. Statistical process control (SPC) charts have been successfully used to detect process faults for several decades. Because technological progress allows more and more advanced sensors to be used in a process, it has become popular to monitor multiple quality characteristics during a process. A multivariate process is simply defined as a process with two quality characteristics or more to be monitored. Due to having multiple quality characteristics, it is much more difficult to determine at what time a fault occurs in a multivariate process compared with that for a univariate process.

Multivariate statistical process control (MSPC) charts have been studied and developed [1–3]; however, their major function is basically to generate an out-of-control signal when process faults occur. It is extremely difficult to estimate or identify the beginning time of a fault using only MSPC charts.

In most situations, the beginning time of a fault contains most of the information behind the causes of the process fault. Rapidly and accurately estimating the beginning time of a fault would contribute much to the identification of the associated root causes of the fault and would significantly improve the process.

There have been many studies that investigated the beginning time of a process fault; however, most of the studies have focused on univariate processes [4–14]. In addition, most related studies have focused on the use of the maximum likelihood estimator (MLE) method [3–12]. However, the MLE method has a strict assumption: the underlying process distribution is known. Because the real-world process distribution is typically unknown, this strict assumption seriously restricts the range of the applicability of the MLE method. Besides, another problem that can be encountered is that there are a considerable number of explanatory variables when modeling a multivariate process with a considerable number of quality characteristics. To overcome the limitations of the MLE method and the difficulties when there

are too many explanatory variables in a multivariate process, this study focuses on a multivariate process with ten quality characteristics and considers process variance shifts as the underlying process faults. Additionally, this study assumes that the process covariance matrix has shifted from Σ_0 to Σ_1 when the process fault has occurred. There are 56 input variables considered in this study. It is not practical to use all 56 variables as inputs into the proposed neural network (NN) classifier. Consequently, this study uses a hybrid technique to select fewer but more significant explanatory variables. This is the first stage of building the proposed scheme. The chosen significant variables are then used as inputs into the proposed NN models. This modeling is the second stage of creating the scheme. After conducting the NN classification, an identification strategy is combined with the scheme to estimate the beginning time of a process fault.

The structure of this study is organized as follows. Section 2 addresses the problems with previous studies. The research gaps and the proposed methodologies used are discussed in Section 3. Section 4 discusses the experimental simulations, where the results and analysis for the typical and the proposed approaches are reported. The final section concludes this study.

2. Problems and Process Models

In this section, we discuss the difficulties that can be encountered in practice. Several research studies that have investigated determining the beginning times of faults will be addressed. In addition, this section presents the models of a generalized multivariate process and the process fault.

2.1. Problems Statement. In typical MSPC applications, an out-of-control signal would indicate that a process fault has occurred in the underlying process. At that moment, although we have evidence regarding the status of the underlying process, we would have difficulty determining the beginning time of the fault. In particular, if the effects of the underlying process faults are minor, the probability of triggering a signal at the beginning time of the fault would be extremely low. As a result, it is almost impossible to determine the beginning time of a process fault by only using the MSPC chart. For example, consider a multivariate process with ten quality characteristics monitored by $|\mathbf{S}|$, an MSPC chart. The process fault with a process variance shift has occurred at time 201. Due to the small magnitude of the fault, it is not detected until time 230. Observing Figure 1, it is apparent that the beginning time of the fault is not equal to the MSPC signal. Their difference gets larger when the magnitude of the fault gets smaller.

Several studies have been conducted to address the difficulty of determining the beginning time of a fault [13, 14]. Process faults have typically been divided into two types: process mean shifts and process variance shifts [15, 16]. These studies propose the MLE approach to estimate the beginning time of a fault when the process mean or variance has shifted to a univariate process. The MLE approach with the use of EWMA charts has been reported for a univariate process [17].

FIGURE 1: $|\mathbf{S}|$, an MSPC chart for the sample generalized variance.

Whereas most of the existing MLE approaches have focused on univariate processes, the study in [3] derived an MLE for a multivariate process. However, the performance of this MLE was not stable when the number of quality variables became large.

The MLE is criticized for its strict assumption that states that the underlying process distribution must be known. This assumption is not feasible for practical processes. As a result, machine learning (ML) methods have been used to determine the beginning time of a fault [13, 14]. However, the number of input variables from those studies is extremely small because of the simplicity of the process structure. There are few studies that have investigated how to identify the beginning time of a fault when considerable input variables are involved. Too many input variables generally result in a time-consuming training stage with the ML approach. However, the study in [18] considered a large number of inputs in their experiments and also assumed that all the quality variables were at faults in the process, which is a rare case in industrial applications. Accordingly, this study proposes an effective hybrid scheme that integrates ANOVA, NN techniques, and an identification strategy to overcome the aforementioned difficulties.

2.2. The Multivariate Process and the Fault Models. In contrast to the traditional multivariate normal distribution assumption for a process, this study considers a multivariate process that follows an unknown multivariate distribution. Assume that the multivariate process is initially in control and the sample observations are from an unknown distribution $F(\underline{\mu}, \Sigma)$ with a known mean vector $\underline{\mu}$ and covariance matrix Σ_0 . After an unknown time $\tau + 1$, we assume that the process covariance matrix changes from Σ_0 to Σ_1 . Let

$$\mathbf{X}_{\sim i,j} = [X_{i,j,1}, X_{i,j,2}, \dots, X_{i,j,p}]' \quad (1)$$

$p \times 1$ be a vector that represents the p characteristics on the j th observation in subgroup i with the unknown distribution function $F(\mu, \Sigma)$. Accordingly, we have

$$\begin{aligned} X_{\sim 1,1}, \dots, X_{\sim 1,n} &\stackrel{\text{iid}}{\sim} F(\mu, \Sigma_0), \\ X_{\sim 2,1}, \dots, X_{\sim 2,n} &\stackrel{\text{iid}}{\sim} F(\mu, \Sigma_0), \\ &\vdots \\ X_{\sim \tau,1}, \dots, X_{\sim \tau,n} &\stackrel{\text{iid}}{\sim} F(\mu, \Sigma_0), \\ X_{\sim \tau+1,1}, \dots, X_{\sim \tau+1,n} &\stackrel{\text{iid}}{\sim} F(\mu, \Sigma_1), \\ &\vdots \\ X_{\sim T,1}, \dots, X_{\sim T,n} &\stackrel{\text{iid}}{\sim} F(\mu, \Sigma_1), \end{aligned} \quad (2)$$

where n is the sample size, $\tau + 1$ is the change point, T is the signal time in which a subgroup covariance matrix exceeds the limits of the control chart $|\mathbf{S}|$, “ $\stackrel{\text{iid}}{\sim}$ ” means “independent and identically distributed,” and Σ_0 is the in-control covariance matrix, which is defined as follows:

$$\Sigma_0 = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,j} & \cdots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \vdots & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \sigma_{i,1} & \cdots & \vdots & \sigma_{i,j} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p,1} & \sigma_{p,2} & \cdots & \sigma_{p,j} & \cdots & \sigma_{p,p} \end{bmatrix}_{p \times p}. \quad (3)$$

Following the suggestion of [19], this study considers the following variance shift as the process fault:

$$\Sigma_1 = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \theta\sigma_{1,j} & \sigma_{1,j+1} & \cdots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \theta\sigma_{2,j} & \sigma_{2,j+1} & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ \theta\sigma_{i,1} & \theta\sigma_{i,2} & \cdots & \theta^2\sigma_{i,j} & \theta\sigma_{i,j+1} & \cdots & \theta\sigma_{i,p} \\ \sigma_{i+1,1} & \sigma_{i+1,2} & \cdots & \theta\sigma_{i+1,j} & \sigma_{i+1,j+1} & \cdots & \sigma_{i+1,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p,1} & \sigma_{p,2} & \cdots & \theta\sigma_{p,j} & \sigma_{p,j+1} & \cdots & \sigma_{p,p} \end{bmatrix}_{p \times p} \quad (4)$$

where θ is the inflated ratio. Let the sample variance-covariance matrix in subgroup i be defined as

$$\begin{aligned} \mathbf{S}_i &= \frac{1}{n-1} \sum_{j=1}^n \left(X_{\sim i,j} - \bar{X}_{\sim i} \right) \left(X_{\sim i,j} - \bar{X}_{\sim i} \right)' \\ &= \begin{bmatrix} S_{i,1,1} & S_{i,1,2} & \cdots & S_{i,1,p} \\ S_{i,2,1} & S_{i,2,2} & \cdots & S_{i,2,p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{i,p,1} & S_{i,p,2} & \cdots & S_{i,p,p} \end{bmatrix}_{p \times p}. \end{aligned} \quad (5)$$

To monitor a multivariate process variance shift, the sample generalized variances $|\mathbf{S}_i|$, $i = 1, 2, \dots$, and the following control limits are used [2]:

$$\begin{aligned} \text{UCL} &= |\Sigma_0| \left(b_1 + 3\sqrt{b_2} \right), \\ \text{LCL} &= \max \left(0, |\Sigma_0| \left(b_1 - 3\sqrt{b_2} \right) \right), \end{aligned} \quad (6)$$

where $|\Sigma_0|$ is the determinant of Σ_0 and

$$\begin{aligned} b_1 &= \frac{1}{(n-1)^p} \prod_{i=1}^p (n-i), \\ b_2 &= \frac{1}{(n-1)^{2p}} \prod_{i=1}^p (n-i) \left(\prod_{i=1}^p (n-i+2) - \prod_{i=1}^p (n-i) \right). \end{aligned} \quad (7)$$

3. The Proposed Scheme

In recent years, intelligent approaches, such as neural networks and support vector machines, have had an important role in the development of industrial technologies [20–22]. Although acceptable results may be obtained using traditional intelligent approaches, these approaches may not fulfill the particular needs of industrial applications. Recent studies have shown that hybrid intelligent approaches can help achieve a better performance for particular applications [18, 19, 23, 24]. In this study, we develop a hybrid scheme to effectively determine the change point of a multivariate process. The proposed scheme includes the ANOVA, an NN, and the identification strategy. The scheme can be used when the multivariate process distribution is unknown and when there are a large number of input variables. The following sections address these components.

3.1. ANOVA. The proposed hybrid two-stage method integrates the framework of ANOVA and an NN. In stage I, a one-way ANOVA test is applied to select important, influential variables. In stage II, the selected significant variables are taken as the input variables into the NN.

The purpose of performing a one-way ANOVA in stage I is to determine whether data from the “in-control” and “out-of-control” groups have a common mean, that is, to determine whether the measured characteristics from the “in-control” and “out-of-control” groups are actually different. Because matrix $\mathbf{S}_i$ is symmetric, only the elements

on and above the diagonal need to be examined by the one-way ANOVA. To simplify the notation, let $Y_{i,1} = S_{i,1,1}$, $Y_{i,2} = S_{i,1,2}, \dots$, $Y_{i,p} = S_{i,1,p}$, $Y_{i,p+1} = S_{i,2,2}$, $Y_{i,p+2} = S_{i,2,3}, \dots$, $Y_{i,N-1} = S_{i,p,p}$, and $Y_{i,N} = |S|$, where $N = 1 + p(p+1)/2$. Let $Y_{i,j,k,l}$ be the l th observation at the k th level of the factor (where level 1 represents an “in-control” group and level 2 represents an “out-of-control” group) for the variable $Y_{i,j}$ mentioned above, $i = 1, 2, \dots, T$; $j = 1, 2, \dots, N$; $k = 1, 2$; $l = 1, 2, \dots, n_{i,j,k}$. Let $\mu_{i,j}$ and $\tau_{i,j,k}$ be the corresponding overall mean and treatment effect, respectively. Accordingly, the linear equation for the one-way ANOVA model is

$$Y_{i,j,k,l} = \mu_{i,j} + \tau_{i,j,k} + \varepsilon_{i,j,k,l},$$

$$k = 1, 2, \quad l = 1, 2, \dots, n_{i,j,k}, \quad i = 1, 2, \dots, T, \quad j = 1, 2, \dots, N. \quad (8)$$

To identify significant variables, an F -test statistic is used to test the differences between the in-control and out-of-control groups. Those significant variables selected in this stage are then substituted into the NN to construct a two-stage model.

3.2. Neural Network. The purpose of using an NN is to classify the process output as either an in-control or out-of-control process. The identification strategy uses this information to activate its function. Accordingly, the beginning time of a process fault can be estimated in real time.

The structure of the NN can be briefly described as follows. The NN nodes are divided into three layers, which include the input, the output, and the hidden layers. The nodes in the input layer receive input signals from an external source, and the nodes in the output layer provide 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

$$\begin{aligned} \text{net}_j &= \sum_i w_{ji} \times o_i, \\ \text{net}_k &= \sum_j w_{kj} \times o_j, \end{aligned} \quad (9)$$

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 as

$$\begin{aligned} o_i &= \text{net}_i, \\ o_i &= \frac{1}{1 + \exp^{-(\text{net}_i + \theta_i)}} = f_i(\text{net}_i, \theta_i), \\ o_k &= \frac{1}{1 + \exp^{-(\text{net}_k + \theta_k)}} = f_k(\text{net}_k, \theta_k), \end{aligned} \quad (10)$$

where net_j (net_k) is the input signal from the external source to node $j(k)$ in the input layer and $\theta_j(\theta_k)$ is a bias. The transformation function shown in (10) is called the sigmoid

function and is the one most commonly used transformation function. Accordingly, this study uses the sigmoid function.

The generalized delta rule is the conventional technique used to derive the connection weights of the feedforward network. Firstly, a set of random numbers is assigned to the connection weights. Then, to obtain a pattern p with target output vector $t_p = [t_{p1}, t_{p2}, \dots, t_{pM}]^T$, the sum of the squared error to be minimized is given as

$$E_p = \frac{1}{2} \sum_{j=1}^M (t_{pj} - o_{pj})^2, \quad (11)$$

where M is the number of output nodes. By minimizing the error E_p using the gradient descent technique, the connection weights can be updated using the following equations:

$$\Delta w_{ji}(p) = \eta \delta_{pj} o_{pj} + \alpha \Delta w_{ji}(p-1), \quad (12)$$

where for the output nodes,

$$S_{pj} = (t_{pj} - o_{pj}) o_{pj} (1 - o_{pj}) \quad (13)$$

and for other nodes,

$$\delta_{pj} = \left(\sum_k (\delta_{pk} \times w_{kj}) o_{pj} (1 - o_{pj}) \right). \quad (14)$$

3.3. An Identification Strategy. This study uses an NN to classify the status of a process at a certain time C . When the output of the NN is classified as “0”, this indicates that the process fault has not occurred. When the output of the NN is classified as “1”, this indicates that a process fault has intruded into the underlying process. When an SPC chart is triggered at time T , we know a fault has intruded into the underlying process. The identification component is then activated, and the NN begins to classify the status of the process from time $T-1$ to 1 in a backward sequence.

If the NN output is “1” at time $T-1$, we may conclude that the beginning time of the fault has been confirmed at time $T-1$ instead of time T . Then, we can proceed to time $T-2$. If the NN output is “1” again at time $T-1$, we could conclude that the beginning time of the fault has been confirmed at time $T-2$ instead of time $T-1$. However, because all classifiers are not perfect, we could obtain misclassification results. That is, we may encounter a problem in which the NN output is 0 at time $T-1$ and the values of the outcome are all 1s from time $T-2$ to $T-C$ (where $1 \leq C \leq T-1$); one may ask what is the subsequent decision. The decision on the beginning time of a fault is not definitively made by observing only a single outcome.

In this study, because the NN outputs are either 1 or 0, we can consider them as the success or failure of a Binomial experiment, respectively. Accordingly, we can use the cumulative probability distribution of a Binomial experiment to determine the beginning time of a fault. If the NN has a good classification capability, we know that most of the output values from time T to $T-\tau$ should be classified as 1, which implies that the cumulative probability of the

Binomial distribution is near 1. Due to there being no perfect classifiers in reality, several misclassifications of NN outputs must be tolerated. Therefore, the cumulative probability of the Binomial distribution should be less than a certain threshold value. That is, if the value of the cumulative probability is greater than a threshold at a time $T - i$, we can conclude that the beginning time of a fault has occurred at time $T - i$. However, there is no theoretical threshold value. According to our experience and numerous simulations' results, we therefore estimate the threshold value as follows.

(1) *Determining the Threshold.* During the training and testing for the NN modeling phase, denoted previously as phase I, we can obtain an accurate identification rate (AIR) for the classification tasks. The AIR is equivalent to the probability of a successful rate (P_S) from the Binomial experiments. Because the number of successes must be an integer, the following relationship should be used:

$$[n_S] \geq n \times P_S, \quad (15)$$

where n_S is the number of successes in n Binomial experiments and $[n_S]$ is the smallest integer that is greater than or equal to the value of $n \times P_S$. The integer $[n_S]$ is used as a standard, and the corresponding cumulative probability is considered to be the threshold. As a result, the threshold is calculated as follows:

$$P(X_D \leq n_S) = \sum_{x_D=0}^{n_S} \binom{n}{x_D} p^{x_D} (1-p)^{n-x_D}, \quad (16)$$

where X_D is the accumulation of the Binomial trial outputs.

(2) *Performing the Confirmation Test.* To perform the confirmation data test, the new process data vectors were generated. For each confirmation data vector, the phase I NN model that classifies the confirmation data was used. This confirmation test is referred to as phase II. The accumulation of the NN outputs in phase II is denoted as X_{NN} . The number of successes of the NN outputs in phase II is denoted as n_A . At time t_i , the value of the cumulative probability can be calculated as the following:

$$P(X_{NN} \leq n_A) = \sum_{x_{NN}=0}^{n_A} \binom{n}{x_{NN}} p^{x_{NN}} (1-p)^{n-x_{NN}}. \quad (17)$$

(3) *Conducting the Decision Rule.* After performing steps (1) and (2), the decision rule can be set up as follows:

$$\text{If } P(X_{NN} \leq n_A) \geq P(X_D \leq n_S), \quad (18)$$

time t_i is the beginning time of a process fault.

4. Simulated Examples

This study performs a series of simulations to compare the existing single-stage NN method with the proposed hybrid scheme proposed in Section 3. The corresponding estimators of τ for these two methods are denoted as $\hat{\tau}_{ANN}$ and $\hat{\tau}_{AA}$, respectively.

4.1. *Assumptions.* Without loss of generality, we assume that each quality characteristic is sampled from a normal distribution with zero mean and one standard deviation. In addition, we assume that we monitor ten quality characteristics simultaneously (i.e., $p = 10$) and the in-control covariance matrix is as follows:

$$\Sigma_0 = \begin{bmatrix} 1.0 & 0.5 & \cdots & \cdots & 0.5 \\ 0.5 & 1.0 & 0.5 & \cdots & 0.5 \\ \vdots & 0.5 & 1.0 & 0.5 & 0.5 \\ \vdots & \vdots & 0.5 & \ddots & 0.5 \\ 0.5 & \cdots & \cdots & 0.5 & 1.0 \end{bmatrix}_{10 \times 10}. \quad (19)$$

For the out-of-control covariance structure, without loss of generality, we assume that a variance shift occurs at the first quality characteristic. Consequently, the following out-of-control covariance matrix is considered:

$$\Sigma_1 = \begin{bmatrix} \theta^2 \sigma_{1,1} & \theta \sigma_{1,2} & \cdots & \theta \sigma_{1,p} \\ \theta \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ \theta \sigma_{p,1} & \sigma_{p,2} & \cdots & \sigma_{p,p} \end{bmatrix}_{p \times p}. \quad (20)$$

In this study, the training data sets include 1000 data vectors for every possible parameter setting. Whereas the first 500 data vectors are all from an in-control state, the last 500 data vectors are from an out-of-control state. The structure of the testing data sets is the same as that of the training data sets; that is, the testing data sets involve 1000 data vectors. The first 500 data vectors are from an in-control state and the last 500 data vectors are from an out-of-control state.

This study considers four values of the inflated ratio θ : 1.1, 1.2, 1.3, and 1.4. In our proposed two-stage model, we have 7, 10, 10 and 10 input nodes for the ANOVA-NN models for $\theta = 1.1$, $\theta = 1.2$, $\theta = 1.3$, and $\theta = 1.4$, respectively. For all the models, there is only one output node. This output node indicates the classification results of the process status, where a value of 0 indicates that the process is in control and a value of 1 implies that the process is out of control. Furthermore, the change point of the process is assumed to be 201 ($\tau + 1 = 201$). For each data structure, we use a sample size (n) of 12 and repeat the simulation 5 times. The average of the estimates of each approach for the 5 simulation replicates was then recorded along with their standard errors.

4.2. *Modeling Results and Analysis.* In stage I, we use a significance level of 0.05 and apply a one-way ANOVA test to select the important, influential variables. The results are given in Table 1. The significant variables selected in this stage are then used as the input variables into the NN. In addition, from Table 2, it can be seen that between the two methods discussed above, the two-stage ANOVA-NN scheme tended to have a better performance than that of the existing single-stage NN method.

To evaluate the performance of the two estimators discussed above, the bias and the mean squared error (MSE) were used in this study. The bias of an estimator $\hat{\tau}$ is

TABLE 1: Significant variables selected using the one-way ANOVA.

θ	Significant explanatory variables
1.1	$Y_1, Y_2, Y_3, Y_6, Y_7, Y_9, Y_{10}$
1.2	$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}$
1.3	$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}$
1.4	$Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}$

TABLE 2: Average beginning time of a fault estimate and standard error for two estimators.

θ	Method	
	$\hat{\tau}_{ANN}$	$\hat{\tau}_{AA}$
1.1	242.50 (7.69)	238.17 (10.36)
1.2	248.17 (13.14)	223.50 (20.32)
1.3	242.00 (13.04)	223.67 (16.56)
1.4	230.50 (5.75)	207.50 (7.20)

the distance between the expected value of the estimator and the parameter being estimated. It is used to indicate the accuracy of the estimator and is defined as follows:

$$\text{bias}(\hat{\tau}) = E(\hat{\tau}) - \tau. \quad (21)$$

The MSE is the expected value of the squared errors and is defined as follows:

$$\text{MSE}(\hat{\tau}) = E(\hat{\tau} - \tau)^2. \quad (22)$$

It is used to indicate how far, on average, the collections of estimates are from the parameters being estimated. The effects of the inflated ratio θ on the biases and the MSE of the two estimators are shown in Figures 2 and 3, respectively. From Figure 2 it is found that the biases of the two estimators decrease as θ increases, and the bias of the two-stage scheme appears to be smaller than the one of the other method. On the other side, again, Figure 3 shows that the mean squared error of the two-stage scheme tends to be smaller than the one of single-stage NN method. Consequently, it seems that the proposed two-stage ANOVA-NN scheme is more efficient in detecting the actual change point than the existing single-stage NN method.

5. Conclusions

The objective of this work is to develop an effective scheme to identify the beginning time of a fault, specifically for a process variance shift in a multivariate process with a general distribution. On the basis of our numerical study, the two-stage procedure introduced here was generally more efficient in detecting the beginning time of a fault than that of the single-stage NN method. This work could be a useful guide to engineers attempting to search for the root cause of a process disturbance.

Based on our results, further studies can be expanded. For example, extensions of the proposed two-stage procedure to discrete multivariate processes or other statistical techniques are possible. Such work deserves further research and is our future concern.

FIGURE 2: Biases of the two estimators.

FIGURE 3: Mean squared errors of the two estimators.

Acknowledgment

This work is partially supported by the National Science Council of China, Grant no. NSC 99-2221-E-030-014-MY3 and Grant no. NSC 100-2118-M-030-001.

References

- [1] H. Hotelling, "Multivariate quality control," in *Techniques of Statistical Analysis*, C. Eisenhart, M. W. Hastay, and W. A. Wallis, Eds., McGraw Hill, New York, NY, USA, 1947.
- [2] F. B. Alt, "Multivariate quality control," in *Encyclopedia of Statistical Sciences*, N. L. Johnson and S. Kotz, Eds., vol. 6, John Wiley & Sons, New York, NY, USA, 1985.
- [3] C. D. Hou, Y. E. Shao, and S. Huang, "A combined MLE and generalized P chart approach to estimate the change point of a multinomial process," *Applied Mathematics and Information Sciences*, vol. 7, no. 4, pp. 1487–1493, 2013.
- [4] T. R. Samuel, J. J. Pignatiello, and J. A. Calvin, "Identifying the time of a step change with $\bar{X}$ control charts," *Quality Engineering*, vol. 10, no. 3, pp. 521–527, 1998.

- [5] T. R. Samuel, J. J. Pignatiello, and J. A. Calvin, "Identifying the time of a step change in a normal process variance," *Quality Engineering*, vol. 10, no. 3, pp. 529–538, 1998.
- [6] J. J. Pignatiello and T. R. Samuel, "Estimation of the change point of a normal process mean in SPC applications," *Journal of Quality Technology*, vol. 33, no. 1, pp. 82–95, 2001.
- [7] M. B. Perry, J. J. Pignatiello, and J. R. Simpson, "Estimation of the change point of a Poisson rate parameter with a linear trend disturbance," *Quality and Reliability Engineering International*, vol. 22, no. 4, pp. 371–384, 2006.
- [8] Y. E. Shao and C. D. Hou, "Estimation of the starting time of a step change disturbance in a γ process," *Journal of the Chinese Institute of Industrial Engineers*, vol. 23, no. 4, pp. 319–327, 2006.
- [9] Y. E. Shao and C. D. Hou, "Estimation of the change point of a uniform process using the EWMA chart and MLE," *ICIC Express Letters*, vol. 3, no. 3, pp. 451–456, 2009.
- [10] J. J. Pignatiello and T. R. Samuel, "Identifying the time of a step-change in the process fraction nonconforming," *Quality Engineering*, vol. 13, no. 3, pp. 357–365, 2001.
- [11] T. R. Samuel and J. J. Pignatiello, "Identifying the time of a change in a poisson rate parameter," *Quality Engineering*, vol. 10, no. 4, pp. 673–681, 1998.
- [12] R. Noorossana, A. Saghaei, K. Paynabar, and S. Abdi, "Identifying the period of a step change in high-yield processes," *Quality and Reliability Engineering International*, vol. 25, no. 7, pp. 875–883, 2009.
- [13] Y. E. Shao, H. Y. Huang, C. D. Hou, K. S. Lin, and J. E. Tsai, "Change point determination for an attribute process," *ICIC Express Letters*, vol. 5, pp. 3117–3122, 2011.
- [14] Y. E. Shao, "An integrated neural networks and SPC approach to identify the starting time of a process disturbance," *ICIC Express Letters*, vol. 3, pp. 319–324, 2009.
- [15] Y. E. Shao, C. J. Lu, and C. C. Chiu, "A fault detection system for an autocorrelated process using SPC/EPC/ANN and SPC/EPC/SVM schemes," *International Journal of Innovative Computing, Information and Control*, vol. 7, pp. 5417–5428, 2011.
- [16] 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.
- [17] Y. E. Shao and C. D. Hou, "A combined MLE and EWMA chart approach to estimate the change point of a gamma process with individual observations," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 5, pp. 2109–2122, 2011.
- [18] Y. E. Shao and C. D. Hou, "Change point determination for a multivariate process using a two-stage hybrid scheme," *Applied Soft Computing*, vol. 13, no. 3, pp. 1520–1527, 2013.
- [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] W. Bischoff and F. Miller, "A minimax two-stage procedure for comparing treatments: looking at a hybrid test and estimation problem as a whole," *Statistica Sinica*, vol. 12, no. 4, pp. 1133–1144, 2002.
- [21] 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.
- [22] W. Dai, Y. E. Shao, and C. J. Lu, "Incorporating feature selection method into support vector regression for stock index forecasting," *Neural Computing and Applications*, 2012.
- [23] Y. E. Shao, C. J. Lu, and Y. C. Wang, "A hybrid ICA-SVM approach for determining the fault quality variables in a multivariate process," *Mathematical Problems in Engineering*, vol. 2012, Article ID 284910, 12 pages, 2012.
- [24] 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.