International Journal of Innovative
Computing, Information and Control
Volume 7, Number 9, September 2011
 ISSN 1349-4198
pp. 5417–5428

A FAULT DETECTION SYSTEM FOR AN AUTOCORRELATED PROCESS USING SPC/EPC/ANN AND SPC/EPC/SVM SCHEMES

Yuehjen E. Shao[1], Chi-Jie Lu[2] and Chih-Chou Chiu[3]

[1]Department of Statistics and Information Science
Fu Jen Catholic University
No. 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 24205, Taiwan
stat1003@mail.fju.edu.tw

[2]Department of Industrial Engineering and Management
Ching Yun University
Jung-Li 320, Taoyuan, Taiwan
jerrylu@cyu.edu.tw

[3]Department of Business Management
National Taipei University of Technology
Taipei City 106, Taiwan
chih3c@ntut.edu.tw

Received May 2010; revised September 2010

Abstract. *The statistical process control (SPC) chart is effective in detecting process faults. One important assumption for using the traditional SPC charts requires that the plotted observations are independent to each other. However, the assumption of independent observations is not typically applicable in practice. When the observations are autocorrelated, the false alarms are increased, and these improper signals can result in a misinterpretation. Therefore, the use of engineering process control (EPC) has been proposed to overcome this difficulty. Although EPC is able to compensate for the effects of faults, it decreases the monitoring capability of SPC. This study proposes the combination of SPC, EPC and artificial neural network (SPC/EPC/ANN) and SPC, EPC and support vector machine (SPC/EPC/SVM) mechanisms to solve this problem. Using the proposed schemes, this study introduces a useful technique to detect the starting time of process faults based on the execution of the binomial random experiments. The effectiveness and the beneficial results of the proposed schemes are demonstrated through the use of series simulations.*
Keywords: Fault detection, SPC, EPC, Artificial neural networks, Support vector machine

1. Introduction.

Statistical process control (SPC) charts have been continuously developed and implemented in practice for more than 80 years. The primary function of SPC charts is to detect the presence of faults as soon as they intrude in the process. One important assumption for using the traditional SPC charts requires that the monitored observations are independent from each other. Otherwise, the false alarms are increased, and these improper signals result in misinterpretation and decrease capability of SPC charts.

However, autocorrelation commonly exists in real-life processes. The typical reasons include a shorter sampling interval, higher frequency of sampling time (i.e., they may be due to the use of automatic measuring systems or sensors), or the dynamic response time of the chemical materials. Consequently, autocorrelation occurs most frequently in the continuous and chemical processes [1-4]. The process personnel need to pay more attention while using traditional SPC charts on those autocorrelated processes.

Several studies have discussed this important issue. For example, [5] indicated that the false alarm signals would be increased even if a moderate autocorrelation exists among the observations. The algorithmic statistical process control (ASPC) which was proposed by [6] can be used to monitor an autocorrelated process. This is feasible since the controlled process would be independently and identically distributed (iid) as long as the proper control action is implemented and the original process model pertains. [7-9] all suggested to fit an appropriate ARIMA model to the autocorrelated observations and then the resulting independent residuals can be monitored by SPC charts. However, the ARIMA modeling could be a time-consuming task, [11,12]. In addition, [10] studied the performance of a Cusum control chart for monitoring the residuals. [6] discussed the problems and addressed the potential research direction about monitoring the autocorrelated observations. Some other research can also be referenced in [13-20].

The use of engineering process control (EPC) has been mainly suggested and implemented to deal with the difficulties of monitoring the autocorrelated process [21,22]. However, little work has been done on addressing the drawbacks of the integration of EPC and SPC. That is, although the good feature of fault compensation can be achieved by using EPC, it results in a decrease of insensitivity of the SPC monitoring capability. In this study, we propose the combination of SPC, EPC and artificial neural network (SPC/EPC/ANN), as well as SPC, EPC and support vector machine (SPC/EPC/SVM) mechanisms to solve the problem of insensitivity of SPC monitoring. Using the proposed mechanisms, we are able to effectively detect the starting time of a fault in an autocorrelated process. The superiority of the proposed approaches is also demonstrated with the use of simulated experiments.

The structure of this study is organized as follows. Section 2 discusses the difficulty of SPC/EPC system for monitoring an autocorrelated process. the basic concepts for the EPC are introduced in this section as well. Section 3 introduces the proposed approaches for detecting the fault of a process. Section 4 demonstrates a series of simulated experiments which are used to report the performance of the typical and the proposed approaches. Section 5 concludes this study.

2. **Problem Statement.**

The integration of SPC and EPC has become a promising area of research in recent years, and the benefits have been reported in several studies [18-20]. One of the important features of using proper EPC is to compensate for the effects of process faults, and consequently, the resulting process observations should be independent from each other. That is, the proper use of EPC results in a lack of autocorrelation among the observations. The traditional SPC chart is then able to monitor these independent observations. However, the method of using EPC to remove the autocorrelation has a drawback. When a proper EPC is used to fine-tune the process, the autocorrelation structure is removed; however, the effects of the fault are also obscured. This would cause the problem in which a fault can not be clearly detected by the SPC charts. Therefore, even when the autocorrelation is removed by a proper EPC, the SPC charts are still unable to effectively detect the process faults. As a consequence, the resulting average run length (ARL) would be larger than the exact ARL.

Consider a zero order process with an AR(1) noise [19]:

$$Y_{t+1} = qX_t + d_{t+1}, \quad d_{t+1} = \frac{a_{t+1}}{(1 - \phi B)} \tag{1}$$

where

Y_{t+1}: the output deviation at time $t + 1$,

X_t: deviation from the manipulate variable at time t,

q: the process parameter,
a_{t+1}: the white noise at time $t+1$, and they are iid with normal distribution.
d_{t+1}: the noise.

A minimum mean squared error (MMSE) control is commonly used to tune the process [14,15]. To compensate for the noise, d_t, it is can be shown that the following MMSE control action can be obtained:

$$\hat{X}_t = \phi X_{t-1} - \frac{\phi}{q} a_t = \phi X_{t-1} - \frac{\phi}{q} Y_t \tag{2}$$

In addition, substituting Equation (2) into Equation (1), the following equation holds:

$$Y_{t+1} = a_{t+1} \tag{3}$$

Equation (3) implies that the output deviations from the target would follow a sequence of white noise, and therefore, these output deviations from the target are independent to each other. Consequently, it is appropriate to use the SPC charts to monitor the output deviations when a suitable EPC control action is used to tune the process.

Now consider a step-change fault (or level shift) has been introduced in the process, and Equation (1) can be reformed as:

$$Y_{t+1} = qX_t + d_{t+1} + \delta \tag{4}$$

where δ stands for magnitude of the fault. When the MMSE is used to tune the process (i.e., Equation (4)), it can be shown that the following equation holds:

$$Y_{t+1} = \delta(1-\phi) + a_{t+1} \tag{5}$$

Equation (5) implies that the fault would be compensated by the MMSE control action. However, the fault becomes more difficult to detect since the magnitude of effects of the fault also becomes smaller.

Consider a process which is represented by Equation (1) with the following parameter settings: $q = 0.3$, $\phi = 0.8$ and the variance of the white noise is 1. Suppose that a fault, $\delta = 1$, is introduced at time 51. Figures 1 and 2 show the corresponding process outputs without and with use of MMSE control, respectively. Observing Figure 2, one can apparently notice that the fault is mainly compensated by the MMSE. However, from a

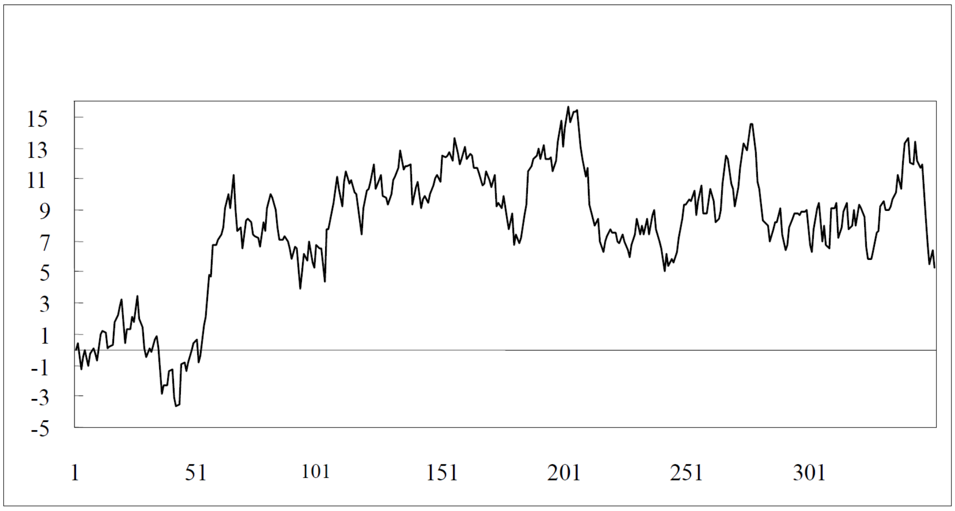

FIGURE 1. The process outputs (without use of MMSE) in the case of $\phi = 0.8$ and $\delta = 1$

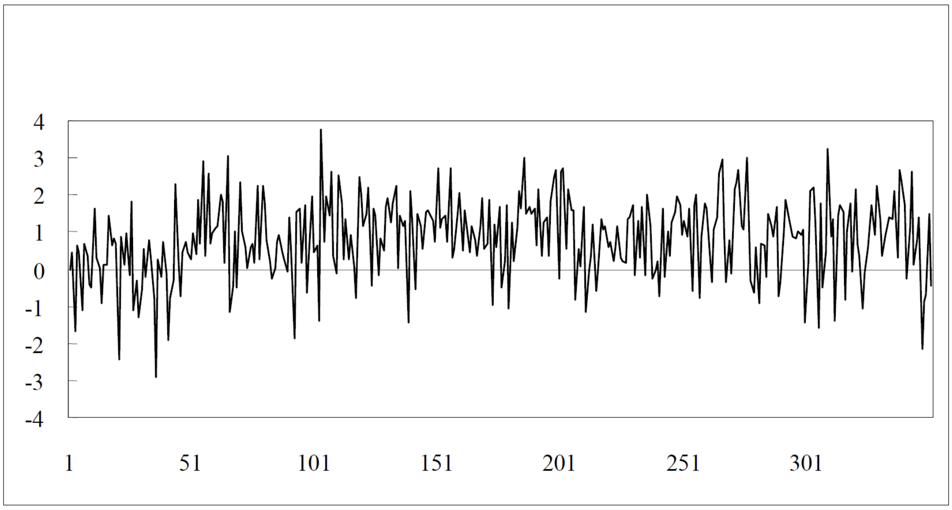

Figure 2. The process outputs (with use of MMSE) in the case of $\phi = 0.8$ and $\delta = 1$

monitoring point of view, the out-of-control signal needs more time to be triggered. That is, the out-of-control ARL becomes larger when EPC is used.

3. Proposed Methodologies.

In this section, we present the concept of ANN and SVM. In addition, the fault detection technique is addressed.

3.1. ANN.

An artificial neural network is a massively parallel system comprised of highly interconnected, interacting processing elements, or units that are based on neurobiological models. ANNs process information through the interactions of a large number of simple processing elements or units, also known as neurons. Knowledge is not stored within individual processing units, but is represented by the strength between units [23]. Each piece of knowledge is a pattern of activity spread among many processing elements, and each processing element can be involved in the partial representation of many pieces of information.

Owing to its associated memory characteristic and its generalization capability, ANN has increasingly found use in modeling nonstationary processes [24]. Recently, more and more computer scientists and statisticians have become interested in the computational potential of ANN algorithms. The capability of ANN and SPC charts in identifying a shift in the mean level were discussed by [25]. It concluded that an ANN can be designed which equal or exceed the performance of the standard $\bar{X}$ control chart. The ANN was trained to recognize shifts in process mean and variability values with a rational subgroup size of ten [26]. The SPC, EPC and ANN were combined to identify the types of underlying faults in a process [27]. The ANN was used to detect and classify three types of non-random faults [28]. The SPC/EPC was integrated with ANN to identify process faults [29]. The ANN and multivariate adaptive regression splines (MARS) were employed to classify the pattern of breast cancer [30]. For more related research issue, we refer readers to [31-36] and the references therein.

ANN can be classified into two different categories, feedforward networks and feedback networks [23]. The nodes in the ANN can be divided into three layers: the input layer, the output layer and one or more 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

$$net_j = \sum_i w_{ji} * o_i \text{ and } net_k = \sum_j w_{kj} * o_j,$$

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

$$\begin{aligned} o_i &= net_i \\ o_j &= \frac{1}{1+\exp-(net_j+\theta_j)} = f_j(net_j,\theta_j) \quad (6) \\ o_k &= \frac{1}{1+\exp-(net_k+\theta_k)} = f_k(net_k,\theta_k) \quad (7) \end{aligned}$$

where $net_j(net_k)$ 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 Equations (6) and (7) is called sigmoid function and is the one most commonly utilized to date. Consequently, sigmoid function is used in this study.

The generalized delta rule is the conventional technique used to derive the connection weights of the feedforward network [23]. Initially, a set of random numbers is assigned to the connection weights. Then for a presentation of a pattern p with target output vector $t_p = [t_{p1}, t_{p2}, ..., t_{pM}]^T$, the sum of squared error to be minimized is given by

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

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

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

where for output nodes

$$\delta_{pj} = (t_{pj} - o_{pj})o_{pj}(1 - o_{pj})$$

and for other nodes

$$\delta_{pj} = \left(\sum_k \delta_{pk} * w_{kj}\right) o_{pj}(1 - o_{pj})$$

Note that the learning rate affects the network's generalization and the learning speed to a great extent.

3.2. **SVM.** The concept about support vector machine was initially addressed by [37]. It became an attractive learning method due to the successful kernel-based framework [38]. Basically, the aim of SVM is to find a hyperplane in the middle of the most separated margins between two classes; so that, this hyperplane can be applied for classifying the new testing samples. In addition, the SVM has been successfully implemented in different areas, and they can be referenced by [39-44].

The use of SVM algorithm can be described as follows. Let $\{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, $\mathbf{x}_i \in R^d$, $y_i \in \{-1, 1\}$ be the training set with input vectors and labels. Here, N is the number of sample observations and d is the dimension of each observation, y_i is known target. The algorithm is to seek the hyperplane $\mathbf{w} \cdot \mathbf{x}_i + b = 0$, where $\mathbf{w}$ is the vector of hyperplane and b is a bias term, to separate the data from two classes with maximal margin width $2/\|\mathbf{w}\|^2$, and the all points under the boundary is named support vector. In order to obtain the optimal hyperplane, the SVM was used to solve the following optimization problem [45]:

$$\begin{aligned} &\text{Min } \Phi(\mathbf{x}) = \frac{1}{2}\|\mathbf{w}\|^2 \\ &\text{s.t. } \ y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1, \ i = 1, 2, ..., N \end{aligned} \tag{8}$$

It is difficult to solve Equation (8), and we need to transform the optimization problem to the dual problem by Lagrange method. The value of α in the Lagrange method must be non-negative real coefficients. The Equation (8) is transformed into the following constrained form [45]:

$$\begin{aligned} &\text{Max } \Phi(\mathbf{w}, b, \xi, \alpha, \beta) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^{N} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \\ &\text{s.t. } \ \sum_{j=1}^{N} \alpha_j y_j = 0, \ 0 \leq \alpha_i \leq C, \ i = 1, 2, ..., N \end{aligned} \tag{9}$$

In Equation (9), C is the penalty factor and determines the degree of penalty assigned to an error. It can be viewed as a tuning parameter which can be used to control the trade-off between maximizing the margin and the classification error.

In general, it could not find the linear separate hyperplane for all data. For problems that cannot be linearly separated in the input space, the SVM uses the kernel method to transform the original input space into a high dimensional feature space where an optimal linear separating hyperplane can be found. The common kernel function are linear, polynomial, radial basis function (RBF) and sigmoid.

3.3. **An illustrative example for the fault detection technique.** The concept of the fault detection technique is described as follows. When a set of input variables is applied to the proposed SPC/EPC/ANN or SPC/EPC/SVM mechanisms, we obtain a target output. In this study, the outcome of the output variable is classified into either 1 or 0. The values of 1 and 0 correspond to the success and failure of a binomial experiment. Typically, a binomial experiment possesses the following properties:

(1) There are two types of outcomes, success or failure, in each trail.
(2) The success rate of each trail is p and the failure rate of each trial is $1 - p$.
(3) Each experiment is mutually independent.

The decision about the starting time of a fault cannot be accurately made through observing the outcome of a signal trail. Instead, we should employ the cumulative probability distribution of a binomial experiment to determine the fault.

To clearly explain the concept of the proposed techniques, this study provides an illustrative example. Consider a process which can be represented by Equation (1), where $q = 0.3$, $\phi = 0.8$ and the variance of the white noise is 1. This autocorrelated process is tuned by Equation (2). After time 101, a fault has occurred in the process, and the process can now be represented by Equation (4). In addition, we initially applied SPC/EPC/ANN to obtain the outcomes of the process state. The detailed structures of the ANN will be described in the next section. An X control chart is used to monitor this autocorrelated process. In this example, an out-of-control signal is triggered at time 136.

Table 1 shows the outputs for performing the simulation. The first column of Table 1 is the sampling number which starts from time 1 to the signal time t (i.e., $t = 136$ in here). The second column of Table 1 is the backward number, recording list from the signal time t to 1 in a backward sequence. The value of this column can be deemed as the number of binomial experiments. The third column of Table 1 stands for the outcomes of the output variable in the SPC/EPC/ANN model. A value of 1 indicates that the process fault has occurred, while a value of 0 implies that the fault has not yet appeared. The fourth column of Table 1 is the total sum of cumulative occurrence of "1", starting from the time $t-1$. The last column of Table 1 is the binomial cumulative distribution probability. Now consider the value of the cumulative probability is 0.905 when the sampling number is 128. This means that the cumulative sum of "1's" (i.e., column 4) is 7 and the number of experiment (i.e., column 2) is 8, and thus, we can calculate the binomial probability by using Equation (10):

$$\begin{aligned} f(x) &= \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 1, 2, ...n \\ \binom{n}{x} &= \frac{n!}{x!(n-x)!} \end{aligned} \tag{10}$$

where
n: the number of binomial experiments (i.e., column 2 of Table 1),
x: the number of success in n experiments (i.e., column 4 of Table 1), and
p: the probability of obtaining the outcome of success in an experiment (i.e., the accurate identification rate of the ANN or SVM classifier).

By observing Table 1, we can see that the cumulative probability is 1 and 0.905 at sampling number 129 and 128, respectively. A higher value of cumulative probability indicates that a large proportion of "1's" has occurred in the experiments. This large proportion of "1's" implies that there is a very strong tendency for the presence of a process fault. If the value of cumulative probability is 1, we can be almost 100% sure that the fault has already occurred. In this example, since the value of the cumulative probability is 1 at time 129 (i.e., cumulative calculation from time 135 to time 129), we can be almost certain that the process fault has occurred at time 129, instead of the signal time 136. At time 128, the value of the cumulative probability is 0.905. Is the value of 0.905 large enough to draw the conclusion that a fault has occurred? There seems no theoretical answer for the question. According to our experience, the value of the cumulative probability would be set at 0.9 or higher. If the value of 0.9 is set to be our threshold, we may conclude that the process fault has occurred at time 128 (instead of time 136). One thing we have to understand is that the cumulative probability should be large in order to confidently determine the starting time of a fault.

4. **Simulated Experiments.** This study combines the methodologies of SPC/EPC/AN-N and SPC/EPC/SVM with the use of binomial experiments in order to establish fault detection mechanisms with the goal of overcoming deficiency of SPC/EPC systems. In order to show the efficiency of the proposed approaches, this study performs a series of simulations. The results are reported and discussed.

Again, suppose a process can be represented in Equation (1) in which $q = 0.3$, $\phi = 0.8$ and the variance of the white noise is 1. This autocorrelated process is fine-tuned with the use of MMSE control action, Equation (2). After time 101, a fault has intruded in the process, and the process has been represented by Equation (4).

To use the ANN and SVM, we need to design the structure of the ANN and SVM models. All training data sets include 20 simulation runs of data vectors. For a single simulation run, the first 100 observations are all from an in-control state (i.e., no fault involved), and after time 101, the remaining data are all from an out-of-control state (i.e., a fault intruded into the process after 101). The last collected observation in a simulation run is the time at which the SPC signal was triggered. Also, the faults consist of two different values of δ. In this study, two types of training data sets were developed for the autocorrelated process series. The first and the second training data sets were generated

Table 1. The use of fault detection technique

Sample number	Backward number	ANN output	Cumulative sum of "1"	Binomial cumulative distribution probability (P)
1	135	1	63	0.000
2	134	0	62	0.000
3	132	0	62	0.000
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
...	...	...	...	...
127	9	0	7	0.712
128	8	0	7	0.905
129	7	1	7	1
130	6	1	6	1
131	5	1	5	1
132	4	1	4	1
133	3	1	3	1
134	2	1	2	1
135	1	1	1	1
136				

in the cases of $\delta = 1$ and $\delta = 2$, respectively. These two different values of δ represent the small and moderate values of the faults. The large value of the process fault is ignored here since this large value could easily be detected.

The test data sets involve 10 simulation runs of data vectors. The inputs to the ANN and SVM were the values of Y and X. The ANN and SVM outputs consist of one node. This output node indicates the prediction of the process status. The value of 0 concludes that the process is no fault, and the value of 1 indicates that the process has faults. Table 2 shows the accurate identification rate (AIR) for both ANN and SVM models for the two different values of δ.

Table 2. Accurate identification rate (AIR) for ANN and SVM models

	$\delta = 1$	$\delta = 2$
ANN	0.801	0.745
SVM	0.816	0.753

In addition, this study performed another 100 simulation runs to confirm the usefulness of the proposed fault detection technique. Six threshold values were considered, i.e., $P = 0.7$, $P = 0.75$, $P = 0.8$, $P = 0.85$, $P = 0.9$ and $P = 0.95$. Although there is no theoretical aspect to choose the optimal value of P, we should set the P as a large number in order to have strong confidence for determining the starting time of a fault. Tables 3 and 4 demonstrate the simulation results in the case of $\delta = 1$ and $\delta = 2$, respectively. Observing Table 3, we can notice that the average out-of-control signal is given at time 386.64 when the typical SPC/EPC is used. The ideal or true out-of-control signal should be triggered at time period 101. Remember that the EPC action would decrease the SPC monitoring capability, and it causes a late signal time. In the case of $P = 0.7$, the proposed SPC/EPC/ANN detection technique provides average starting time of a fault at time 343.47. This is more close to the true value of signal time, 101. This could reach an 11.17% improvement. Additionally, the proposed SPC/EPC/SVM detection technique reports an average starting time of a fault at time 231.46. This improvement, 40.14%, is even better. The results apparently indicate that the proposed detection technique outperforms the typical SPC/EPC alone. The simulation results also seem to imply that the improvement is greater when P is smaller. However, it is not applicable in the case of $\delta = 2$.

Table 3. Performance comparison between SPC/EPC alone and the proposed approaches when $\delta = 1$

	$P = 0.7$	$P = 0.75$	$P = 0.8$	$P = 0.85$	$P = 0.9$	$P = 0.95$
SPC/EPC alone	386.64	386.64	386.64	386.64	386.64	386.64
Proposed SPC/EPC/ANN	343.47	348.94	353.29	356.16	359.18	363.94
Proposed SPC/EPC/SVM	231.46	241.87	249.51	259.33	265.17	277.70

Table 4. Performance comparison between SPC/EPC alone and the proposed approaches when $\delta = 2$

	$P = 0.7$	$P = 0.75$	$P = 0.8$	$P = 0.85$	$P = 0.9$	$P = 0.95$
SPC/EPC alone	293.05	293.05	293.05	293.05	293.05	293.05
Proposed SPC/EPC/ANN	84.05	89.95	92.24	96.25	102.12	119.39
Proposed SPC/EPC/SVM	74.36	79.42	84.18	88.13	90.85	101.12

Observing the case of $\delta = 2$ in Table 4, we can notice that the average out-of-control signal is triggered at time 293.05 when the typical SPC/EPC is used. However, in the case of $P = 0.7$, the proposed SPC/EPC/ANN and SPC/EPC/SVM detection techniques give the average starting time of a fault at time 84.05 and time 74.36, respectively. In our simulation, since the first 100 observations are all in a state of in-control, the ideal signal would occur at time 101. The setting of $P = 0.7$ seems to over-estimate the starting time of a process fault. When a larger value of P is chosen, the resulting average starting time of a process fault is much more reasonable. For example, in the case of $P = 0.95$, the proposed SPC/EPC/ANN and SPC/EPC/SVM techniques provide the average starting time of a fault at 119.39 and 101.12, respectively. In fact, in the case of $P = 0.95$, 59.26% and 65.49% of process improvements are achieved when the proposed two techniques are employed. It is a satisfactory achievement.

5. **Conclusions.**

The combination of SPC and EPC is an effective way to monitor and control a manufacturing process. Specifically, the EPC is typically used for controlling the autocorrelated process. However, the EPC compensation of the process fault also decreases the SPC monitoring capability. This study is motivated to overcome the deficiencies of an SPC/EPC system. This study combines the SPC/EPC/ANN, SPC/EPC/SVM with the use of a binomial distribution concept to effectively determine the starting time of a process fault. The effectiveness of the proposed approach is tested through a series of simulated experiments.

In this study, the new features of our proposed approaches and the main advantages of the results over other related research works are addressed. The apparent advantages include the less conservativeness of the process model used and the conditions imposed to derive the results. Also, the proposed approaches are simple to use, and the practical usefulness are clearly observed. From our research, we are able to draw the following conclusions. Firstly, when the magnitude of the fault is small, both proposed SPC/EPC/ANN and SPC/EPC/SVM detection techniques can achieve process improvements for all the values of P. Even when the threshold value of P is small, we can still obtain good results. Furthermore, it seems that the performance of SPC/EPC/SVM is better than that of SPC/EPC/ANN. Secondly, when the magnitude of the fault is moderate, we discovered that we have to pay more attention to the selection of P for both SPC/EPC/ANN and SPC/EPC/SVM detection techniques. In particular, $P = 0.95$ is strongly suggested. When $P = 0.95$ is chosen, 59.26% and 65.49% improvements are achieved. And finally, one important implication is that we should choose $P = 0.95$, for all possible magnitudes of a fault, in order to accurately obtain the starting time of a fault.

The proposed approach is simple to use and effective in categorizing the starting time of a fault. As a result, the proposed fault detection techniques would be of great help to process personnel in determining the root causes of a fault. Consequently, our proposed approaches result in a significant enhancement of overall process improvement. Nevertheless, this study discusses the case of mean shift faults, an attempt to categorize variance shift faults would also be a valuable contribution to this area of research.

Acknowledgment. This work is partially supported by the National Science Council of Taiwan, Grant No. NSC 99-2221-E-030-014-MY3. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] J. E. Jarrett and X. Pan, The quality control chart for monitoring multivariate autocorrelated processes, *Computational Statistics & Data Analysis*, vol.51, no.8, pp.3862-3870, 2007.
[2] B. K. Issam and L. Mohamed, Support vector regression based residual MCUSUM control chart for autocorrelated process, *Applied Mathematics and Computation*, vol.201, no.1-2, pp.565-574, 2008.
[3] C. M. Wright and M. Y. Hu, A note on detecting outliers in short autocorrelated data using joint estimation and exponentially weighted moving average, *Omega*, vol.31, no.4, pp.319-326, 2003.
[4] M. Pacella and Q. Semeraro, Using recurrent neural networks to detect changes in autocorrelated processes for quality monitoring, *Computers & Industrial Engineering*, vol.52, no.4, pp.502-520, 2007.
[5] L. C. Alwan and H. V. Roberts, Time series modeling for statistical process control, *Journal of Business and Economic Statistics*, vol.6, no.1, pp.87-95, 1988.
[6] F. W. Faltin, C. M. Mastrangelo, G. C. Runger and T. P. Ryan, Considerations in the monitoring of autocorrelated and independent data, *Journal of Quality Technology*, vol.29, no.2, pp.131-162, 1997.
[7] L. C. Alwan, Effects of autocorrelarion on control chart performance, *Communications in Statistics Theory and Methods*, vol.21, no.4, pp.1025-1049, 1992.
[8] S. A. Yourstone and D. C. Montgomery, Development of a real-time statistical process control algorithm, *Quality and Reliability Engineering International*, vol.5, no.4, pp.309-317, 1989.
[9] D. M. Wardrop and C. E. Garcia, Discussion of statistical process monitoring and feedback adjustment discussion, *Technometrics*, vol.34, no.3, pp.281-282, 1992.
[10] G. C. Runger, T. R. Willemain and S. Prabhu, Average run lengths for CUSUM control charts applied to residuals, *Communications in Statistics-Theory and Methods*, vol.24, no.1, pp.273-282, 1995.
[11] D. C. Montgomery and C. M. Mastrangelo, Some statistical process control for autocorrelation data (with discussion), *Journal of Quality Technology*, vol.23, no.3, pp.179-193, 1991.
[12] C. M. Mastrangelo and D. C. Montgomery, SPC with correlated observations for the chemical and process industries, *Quality and Reliability Engineering International*, vol.11, no.2, pp.79-89, 1995.

[13] J. F. MacGregor, On-line statistical process control, *Chemical Engineering Progress*, vol.84, pp.21-31, 1988.

[14] J. F. MacGtregor, T. J. Harris and J. D. Wright, Duality between the control of processes subject to randomly occurring deterministic disturbances and ARIMA stochastic disturbances, *Technomestrics*, vol.26, no.4, pp.389-397, 1984.

[15] C. W. Lu and M. R. Jr. Reynolds, CUSUM charts for monitoring an autocorrelated process, *Journal of Quality Technology*, vol.33, no.3, pp.224-316, 2001.

[16] J. Arkat, S. T. A. Niaki and B. Abbasi, Artificial neural networks in applying MCUSUM residuals charts for AR(1) processes, *Applied Mathematics and Computation*, vol.189, no.2, pp.1889-1901, 2007.

[17] T. Tsai, S. Wu, J. Lin and Y. Chen, Alternative estimation procedure in SPC when the process data are correlated, *Journal of Statistical Computation and Simulation*, vol.77, no.7, pp.575-583, 2007.

[18] W. T. Tucker, F. W. Faltin and W. A. Vander, Algorithmic statistical process control: An elaboration, *Technometrics*, vol.35, no.4, pp.363-375, 1993.

[19] Y. E. Shao, Integrated application of the cumulative score control chart and engineering process control, *Statistica Sinica*, vol.8, no.1, pp.239-252, 1998.

[20] Y. E. Shao, G. C. Runger, J. Haddock and W. A. Wallace, Adaptive controllers to integrate SPC and EPC, *Communications in Statistics-Simulation and Computation*, vol.28, no.1, pp.13-36, 1999.

[21] A. S. Matos, J. G. Requeijo and Z. L. Pereira, Integration of engineering process control and statistical control in pulp and paper industry, *Computer Aided Chemical Engineering*, vol.25, pp.399-404, 2008.

[22] C. Park and M. R. Jr. Reynolds, Economic design of an integrated process control procedure with repeated adjustments and EWMA monitoring, *Journal of the Korean Statistical Society*, vol.37, no.2, pp.55-174, 2008.

[23] B. Cheng and D. M. Titterington, Neural networks: A review from a statistical perspective, *Statistical Science*, vol.9, no.1, pp.2-30, 1994.

[24] D. F. Cook and C. C. Chiu, Using radial basis function neural networks to recognize shits in correlated manufacturing process parameters, *IIE Transactions*, vol.30, no.3, pp.227-234, 1998.

[25] G. A. Pugh, A comparison of neural networks to SPC charts, *Computers and Industrial Engineering*, vol.21, no.1-4, pp.253-255, 1991.

[26] A. E. Smith, X-bar and r control chart interpretation using neural computing, *International Journal of Production Research*, vol.32, no.2, pp.309-320, 1994.

[27] Y. E. Shao and C. C. Chih, Developing identification techniques with the integrated use of SPC/EPC and neural networks, *Quality and Reliability Engineering International*, vol.15, no.4, pp.287-294, 1999.

[28] R. Noorossana, M. Farrokhi and A. Saghaei, Using neural networks to detect and classify out-of-control signals in auto-correlated process, *Quality and Reliability Engineering International*, vol.19, no.6, pp.493-504, 2003.

[29] C. Chiu, Y. E. Shao, T. S. Lee and K. Lee, Identification of process disturbance using SPC/EPC and neural networks, *Journal of Intelligent Manufacturing*, vol.14, no.3-4, pp.379-388, 2003.

[30] 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.

[31] C. Lin, Y. Liu and C. Lee, An efficient neural fuzzy network based on immune particle swarm optimization for prediction and control applications, *International Journal of Innovative Computing, Information and Control*, vol.4, no.7, pp.1711-1722, 2008.

[32] Z. H. Mao and B. Jiang, Fault identification and fault-tolerant control for a class of networked control systems, *International Journal of Innovative Computing, Information and Control*, vol.3, no.5, pp.1121-1130, 2007.

[33] L. Mi and F. Takeda, Analysis on the robustness of the pressure-based individual identification system based on neural networks, *International Journal of Innovative Computing, Information and Control*, vol.3, no.1, pp.97-110, 2007.

[34] G. Vachkov, Growing neural models for process identification and data analysis, *International Journal of Innovative Computing, Information and Control*, vol.2, no.1, pp.101-123, 2006.

[35] 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(B), pp.4899-4906, 2009.

[36] Y. E. Shao, An integrated neural networks and SPC approach to identify the starting time of a process disturbance, *ICIC Express Letters*, vol.3, no.3(A), pp.319-324, 2009.
[37] V. N. Vapnik, *Estimation of Dependences Based on Empirical Data*, Nauka, Moscow, 1979.
[38] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
[39] N. Begum1, M. A. Fattah1 and F. Ren, Automatic text summarization using support vector machine, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.1987-1996, 2009.
[40] H. Tamura and K. Tanno, Midpoint validation method for support vector machines with margin adjustment technique, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(A), pp.4025-4032, 2009.
[41] G. Duan, Y.-W. Chen and T. Sukekawa, Automatic optical inspection of micro drill bit in printed circuit board manufacturing using support vector machines, *International Journal of Innovative Computing Information and Control*, vol.5, no.11(B), pp.4347-4355, 2009.
[42] G. Chen, T.-J. Wang, L.-Y. Gong and P. Herrera, Multi-class support vector machine active learning for music annotation, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.921-930, 2010.
[43] X. Song, S. K. Halgamuge, D. Chen, S. Hu and B. Jiang, The optimized support vector machine with correlative features for classification of natural spearmint essence, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.1089-1099, 2010.
[44] C. Wu, Y. Guo, X. Zhang and H. Xia, Study of personal credit risk assessment based on support vector machine ensemble, *International Journal of Innovative Computing, Information and Control*, vol.6, no.5, pp.2353-2360, 2010.
[45] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2nd Edition, Springer-Verlag, Berlin, 2000.