Research	*An Adaptive Dimension Reduction Scheme for Monitoring Feedback-controlled Processes*

Kaibo Wang[1] and Fugee Tsung[2, *, †]

[1]*Department of Industrial Engineering, Tsinghua University, Beijing 100084, People's Republic of China*
[2]*Department of Industrial Engineering and Logistics Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

Detecting dynamic mean shifts is particularly important in monitoring feedback-controlled processes in which time-varying shifts are usually observed. When multivariate control charts are being utilized, one way to improve performance is to reduce dimensions. However, it is difficult to identify and remove non-informative variables statically in a process with dynamic shifts, as the contribution of each variable changes continuously over time. In this paper, we propose an adaptive dimension reduction scheme that aims to reduce dimensions of multivariate control charts through online variable evaluation and selection. The resulting chart is expected to keep only informative variables and hence maximize the sensitivity of control charts. Specifically, two sets of projection matrices are presented and dimension reduction is achieved via projecting process vectors into a low-dimensional space. Although developed based on feedback-controlled processes, the proposed scheme can be easily extended to monitor general multivariate applications. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: adaptive chart; process dynamics; statistical process control; variable selection

1. INTRODUCTION

The importance of integrating statistical process control (SPC) with feedback-controlled processes to minimize process variance and maintain long-term quality has been widely addressed[1–5]. Under the integrated SPC/feedback-control framework, the primary role of a feedback controller is to compensate for process deviations by regulating controllable factors and to prevent process faults from affecting product quality. In contrast, the primary purpose of SPC is to discover underlying process shifts and to signal for operational intervention when necessary. The interaction between these two techniques has given rise to new challenges for SPC under various conditions.

*Correspondence to: Fugee Tsung, Department of Industrial Engineering and Logistics Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong.
†E-mail: season@ust.hk

One of the most significant changes that the feedback controllers generated is the presence of dynamic shift patterns. In the event of a sustained mean shift with a constant size, the observed shifts of process variables are time-varying. Intensive discussions have been seen in the literature concerning the monitoring of feedback-controlled processes[6−8]. Among them, the joint monitoring scheme proposed by Tsung et al.[9] has gained wide recognition and was further extended by Tsung and Apley[10] and Jiang[11]. This method suggests forming a multidimensional vector by stacking both input variables and output variables together; Hotelling's T^2 chart is then employed to monitor the process.

However, the application of a multivariate control chart runs the risk of including redundant variables, which is cost inefficient and is just as harmful as ignoring vital variables[12]. For example, the output of the proportional-integral (PI)-controlled process in Jiang[11] approaches zero in its steady state. Consequently, the joint monitoring of both the input and output streams is not advantageous to the monitoring of only the input stream. In general, adding interferential factors will reduce the sensitivity of a control chart. Therefore, it is crucial to conduct variable selection in a multivariate control chart. Identifying and monitoring variables that are truly contributing and informative are expected to help in the quick detection of process faults.

The use of variable selection techniques has been considered in different stages of multivariate control chart implementation. Early research emphasized the diagnosis of out-of-control alarms[12−14]. Variable selection techniques are utilized to search for variables that are responsible for an alarm. Once identified, such responsible variables can be traced to locate the source of assignable causes.

Variable selection also plays an important role in the design phase of a multivariate control chart. Encompassing all observable variables without evaluating their individual significance will lead to an oversized vector, which in turn slows down the detection of process faults. One of the solutions is to reduce dimensions by utilizing principle component analysis (PCA)[12,14,15]. The basic idea of PCA monitoring is to find efficient principle components (PCs), which are linear combinations of the original variables, and to monitor these PCs instead of the original variables. The number of PCs to be monitored is usually less than that of the process variables. However, the selection of essential PCs is a much debated subject. Some authors have suggested monitoring the first several PCs that correspond to large eigenvalues[12,14] and argued that if the assignable causes move the process mean vector outside the normal operating conditions, the last several PCs that correspond to small eigenvalues may be more sensitive to process failures. Moreover, the PCs are determined in the design phase and are not supposed to change in the running phase. This has obviously ignored dynamic features of feedback-controlled processes, as the importance of the original variables is not fixed. The PCs that are sensitive to small shifts are not necessarily instructive for large shift detection.

The purpose of this paper is to propose an adaptive scheme that adjusts the dimensions of a multivariate control chart on-line based on real-time information collected from a target process. Variables are adopted or dropped dynamically and only the most informative variables are monitored at each step. The dynamic optimization of the charting statistics is expected to improve the overall performance.

The remainder of this paper is organized as follows. The modeling and the conventional monitoring of feedback-controlled processes are reviewed in Section 2. In Section 3, an index that measures the efficiency of a multivariate control chart is introduced. An adaptive dimension reduction (ADR) procedure is proposed in Section 4 and related design issues are discussed. The performance of this procedure is studied and compared with existing methods in Section 5. Finally, Section 6 concludes this paper with a summary of major findings.

2. MODELING AND CONVENTIONAL MONITORING OF FEEDBACK-CONTROLLED PROCESSES

In order to gain quick detection of faults in a feedback-controlled process, it is fundamentally important to understand the way in which the process responds to common faults. In this section, a representative model for feedback-controlled processes is discussed and important features of responses to common process faults

are identified and analyzed. The review of existing methods is expected to discover opportunities for further improvement.

Consider a process with a single input and a single output. Without loss of generality, we assume that the target value of the process, T, equals zero. Let e_t be the measured deviation from the target, and let x_{t-1} be a controllable factor. The following equation models a collection of run-to-run (R2R) processes in the semiconductor manufacturing and chemical industries[5,9,10,16–19]:

$$e_t \equiv y_t - T + d_t = x_{t-1} + d_t \tag{1}$$

The process disturbance series d_t, reflects the uncertainty of the operational environment. In this paper, we only consider ARMA(1,1) models, as actual stationary industrial processes can often be presented by ARMA models with orders less than two[10]. Denote $d_t = \phi d_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$, where ε_t is a white noise series that follows a normal distribution with mean zero and variance σ_ε^2. The whole idea, however, can be extended to processes with other disturbance models without much modification. For example, if $\phi \to 0$, the disturbance becomes an MA(1) process; if $\theta \to 0$, it becomes an AR(1) model; if, in particular, $\phi = \theta$, the disturbance becomes a series of white noise.

Feedback controllers are usually set up to regulate the controllable factors continuously to maintain process output on target. In industrial practice, the PI controller has shown widespread popularity due to its simple form and efficient capability[19,20]. The general form of the PI controller is given as:

$$x_t = k_P e_t + k_I \sum_{k=0}^{t} e_k = k_P e_t + k_I \frac{1}{1-B} e_t \tag{2}$$

where B is a backshift operator such that $Be_t = e_{t-1}$.

The effectiveness of the PI controller is largely dictated by the tuning parameters k_P and k_I. Tsung et al.[19] presented an optimal design procedure of PI controllers and noted that the PI controller is more robust to model misspecification than the minimum mean square error controller. In this paper, only the PI controller is studied, and the optimal design procedure due to Tsung et al.[19] is followed. The main findings, however, can be extended to other control schemes without significant modification.

In industrial practice, a process failure usually leads to a shift in process means. Common failures include abrupt changes in raw materials[21] and malfunctions of sensors[22]. Denote this sudden-shift type of failure as μ_t. The measured deviation of the process is the additive effect of the regular output and the failure

$$e_t = x_{t-1} + d_t + \mu_t \tag{3}$$

where

$$\mu_t = \begin{cases} 0, & t \leq 0 \\ \delta, & t > 0 \end{cases} \tag{4}$$

is the sustained shift signal. It occurs starting from $t = 1$ and maintains a magnitude of δ thereafter.

For the monitoring of process (3), Tsung et al.[9] proposed a joint monitoring scheme that combines e_t and x_t to detect process shifts. Let

$$\mathbf{V}_t = [e_t, x_t]^T \tag{5}$$

be a vector of the latest observations. A bivariate Hotelling's T^2 chart is set up as follows:

$$T^2 = \mathbf{V}_t^T \mathbf{\Sigma}^{-1} \mathbf{V}_t > h_1 \tag{6}$$

where $\mathbf{\Sigma}$ is the variance–covariance matrix of the vector, $\mathbf{V}_t$, and it takes the form

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_e^2 & \sigma_{ex} \\ \sigma_{ex} & \sigma_x^2 \end{bmatrix} \tag{7}$$
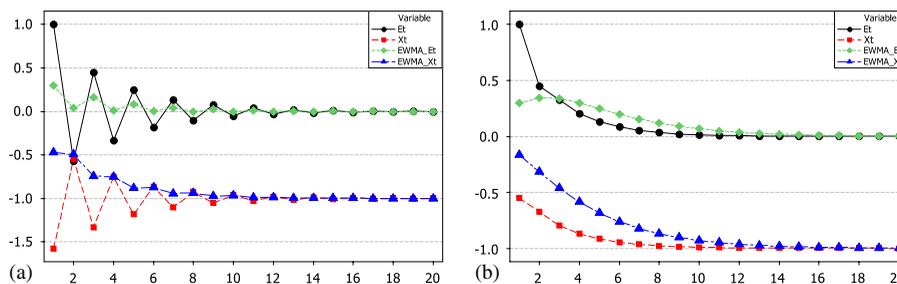
Figure 1. Responses of processes under PI controllers and their EWMA forecasts: (a) PI-controlled process 1: $\phi = 0.8$, $\theta = -0.7$, $k_P = -0.125$, $k_I = -1.448$ and (b) PI-controlled process 2: $\phi = 0.8$, $\theta = 0.3$, $k_P = -0.125$, $k_I = -0.427$

where σ_e and σ_x are the standard deviations of e_t and x_t, respectively. $\sigma_{ex} = \mathrm{cov}(e_t, x_t)$ is the covariance between e_t and x_t. The computation of $\mathbf{\Sigma}$ for PI-controlled processes is discussed by Tsung and Shi[23].

The parameter, h_1, is the control limit that achieves a desired in-control average run length (ARL). As both e_t and x_t follow normal distributions, vector $\mathbf{V}_t$ follows a bivariate normal distribution. Accordingly, T^2 follows a chi-square distribution with two degrees of freedom. The use of $h_1 = \chi_2^{-2}(1-\alpha)$ as the control limit guarantees a false alarm rate no larger than α[9], where $\chi_2^{-2}(\cdot)$ is the inverse cumulative density function (CDF) of the chi-square distribution with two degrees of freedom.

Although the above T^2 chart jointly monitors e_t and x_t, the dynamic responses caused by the feedback loop are not considered by this chart. Figure 1 shows the mean responses of two processes that are under PI controllers. As is seen, starting from the step at which a process shift occurs, the mean of the output stream, e_t, increases from zero to δ immediately. After going through several oscillations, it stabilizes at zero, $\lim_{t \to \infty} E[e_t] \to 0$. In contrast, the mean x_t decreases from zero first and then enters a negative steady level with $\lim_{t \to \infty} E[x_t] \to -\delta$. The transitional behavior of e_t and x_t are the direct result of the sudden shift and the feedback control loop. However, as the process enters its steady state, the information carried by e_t vanishes, leaving only x_t to show the abnormality of the process.

This clearly suggests the dynamic roles that e_t and x_t play. Even though the process shift has a constant size in (4), the measured shifts in both e_t and x_t are time-varying. In the transient stage, the shift signal is buried in both streams. In the steady-state stage, however, monitoring e_t is doubtful for the simple reason that its mean approaches zero. Furthermore, the pattern differs for different processes. In Figure 1(a), exhibits strong oscillations, whereas (b) shows a decaying trend.

In the following sections, we first introduce an index for measuring the efficiency of the multivariate charts. Based on this index, a dynamic dimension reduction scheme is proposed to make full use of the dynamic information for fault detection.

3. AN INDEX FOR MEASURING THE PERFORMANCE OF MULTIVARIATE CONTROL CHARTS

Let x be a variable that represents a normal, independent and identically distributed process, and denote its standard deviation as σ_x. When a two-sided Shewhart control chart on x is set up and a mean shift of magnitude μ_1 occurs, the probability that the shift is detected is

$$p_1 = \Phi(-h - \mu_1/\sigma_{\bar{x}}) + 1 - \Phi(h - \mu_1/\sigma_{\bar{x}}) \tag{8}$$

where h is the control limit and Φ is the CDF of the standard normal distribution. The corresponding out-of-control ARL of the chart is the inverse of p_1, p_1^{-1}.

Equation (8) suggests that, once the control limit h is given, the out-of-control ARL of the chart is solely determined by the standardized shift, μ/σ_x. Let

$$R = \mu/\sigma_x \qquad (9)$$

be the standardized shift; Jiang et al.[24] called this statistic the S/N ratio and noted its dominant role in the efficiency of control charts on x. The larger the S/N ratio, the faster the shift is detected. This index is also used by Jiang and Tsui[25] and Jiang[26] to measure the efficiency of a univariate control chart for monitoring autocorrelated processes.

As the S/N ratio provides a convenient index to measure the efficiency of control charts on x, it is worth trying to extend this index to a multivariate context. However, the performance of a multivariate chart is usually affected by more factors. At this instant, we will study the way in which the performance of a multivariate control chart is influenced by various factors and, furthermore, develop a feasible index to measure and evaluate it.

We first introduce the notation that is used in determining the performance of a multivariate chart. Let $\chi^2_{p,\delta}(x)$ be the CDF of a chi-square distribution with p degrees of freedom and a non-centrality parameter (NCP), δ. When $\delta = 0$, we abbreviate it to $\chi^2_p(x)$. Further let $\chi^{-2}_{p,\delta}(x)$ be the inverse CDF of the distribution and abbreviate it to $\chi^{-2}_p(x)$ if $\delta = 0$.

Without loss of generality, we assume that $\mathbf{M}_t$ is a p-dimensional vector that follows a multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}_M$. Hotelling's T^2 chart for detecting process mean shifts is defined as

$$T^2_M = \mathbf{M}^T_t \mathbf{\Sigma}^{-1}_M \mathbf{M}_t > h_2 \qquad (10)$$

It is known that the T^2_M statistic follows a chi-square distribution with p degrees of freedom. The control limit is $h_2 = \chi^{-2}_p(1-\alpha)$ given a false alarm rate of α.

When the mean of $\mathbf{M}_t$ shifts to $\mathbf{\mu}$, $\mathbf{M}_t \sim N(\mathbf{\mu}, \mathbf{\Sigma}_M)$, T^2_M in (10) follows a non-central chi-square distribution with NCP given as

$$\delta = \mathbf{\mu}^T \mathbf{\Sigma}^{-1} \mathbf{\mu} \qquad (11)$$

The out-of-control performance of (10) is solely determined by the magnitude of the NCP, δ[27]. The larger the δ, the faster the shift is likely to be detected. This also explains that an out-of-control ARL is always smaller than the in-control ARL, as the NCP equals zero when the process is in-control, while it becomes positive in the event of any shift. In addition, when $p = 1$, the matrix $\mathbf{\Sigma}_M$ reduces to a scalar, and Equation (11) is simplified to the square of the S/N ratio in Equation (9).

However, the NCP alone is not sufficient to determine the performance of a multivariate chart. For example, if two charts are used to monitor the same vector, the one with a higher false alarm rate is expected to detect a shift faster. Furthermore, if the NCPs of two charts are equal, the one with a lower dimension has a higher detection sensitivity. In turn, for two charts to achieve the same charting performance, the one with a higher dimension should have a larger NCP than the one with a lower dimension. These arguments are explained by the data in Table I.

In Table I, the dimensions of some T^2 statistics are listed in the first and the fifth columns. The other columns show the NCP of each chart. Each column achieves the same level of false alarm rate and detection power. Let the probability that a shift is detected be f. When the dimension increases, NCP increases accordingly. For example, for $\alpha = 0.10$, the probability that a shift with NCP equals one being detected by a one-dimensional T^2 chart is $f = 0.782$ (for simplicity, we treat the one-dimensional chart as a special T^2 chart). In order to achieve the same level of false alarm rate and detection power, the NCP increases to 1.389 for $p = 2$ and increases to 1.684 for $p = 3$.

In order to compare the performance of T^2 charts with different parameter settings, we propose a multivariate S/N ratio (MSN) function that comprises all influential factors and maps the sensitivity of the charts

Table I. Charts with equivalent sensitivity

p	$\alpha=0.100$ $f=0.782$ MSN$=0.2636$	$\alpha=0.050$ $f=0.867$ MSN$=0.1701$	$\alpha=0.002$ $f=0.988$ MSN$=0.0928$	p	$\alpha=0.100$ $f=0.782$ MSN$=0.2636$	$\alpha=0.050$ $f=0.867$ MSN$=0.1701$	$\alpha=0.002$ $f=0.988$ MSN$=0.0928$
1	1.000	1.000	1.000	11	3.132	3.245	3.393
2	1.389	1.411	1.427	12	3.266	3.387	3.548
3	1.684	1.721	1.752	13	3.395	3.523	3.697
4	1.931	1.980	2.026	14	3.519	3.654	3.841
5	2.147	2.208	2.268	15	3.638	3.780	3.981
6	2.342	2.413	2.488	16	3.754	3.902	4.115
7	2.521	2.601	2.692	17	3.866	4.021	4.246
8	2.688	2.777	2.882	18	3.975	4.136	4.374
9	2.844	2.941	3.061	19	4.080	4.248	4.498
10	2.992	3.097	3.231	20	4.183	4.357	4.619

to a common scale

$$\text{MSN} = 1 - \chi^2_{p,\delta}(\chi_p^{-2}(1-\alpha)) \tag{12}$$

where δ is the NCP defined in (11). The values of the MSN parameter are dominated by three parameters: the false alarm rate, α, the dimension of the charting statistic, p, as well as the NCP, δ. If either δ or α increases, MSN increases; whereas if p increases, MSN decreases. In fact, MSN is the probability that a particular shift with NCP δ will be detected by the T^2 chart (10), in the current step, $0 < \text{MSN} < 1$. Given all the information available, MSN is capable of evaluating the sensitivity of charts with different dimensions and false alarm rates. Therefore, it is reasonable to use MSN as an index to measure the performance of multivariate control charts. The larger the MSN, the more powerful the chart is in detecting a specific shift; while if two charts have equal MSN values, their sensitivities are also identical. For example, as shown in the second column of Table I, the probability that a shift with $\delta=1.684$ being detected by a three-dimensional chart equals the probability of a shift with $\delta=1.389$ being detected by a two-dimensional chart, as both charts have MSN$=0.264$. This confirms that in order to show equivalent detection power, a three-dimensional chart needs a higher NCP than a two-dimensional chart. Furthermore, if for the same shift, another three-dimensional chart with $\alpha=0.05$ is set up, it has MSN$=0.167$, which means that this chart has a relatively low false alarm rate, but consequently low detection power as well.

It is always desirable to maximize the MSN value by tuning design parameters. Among the influential factors, parameter α is habitually chosen based on the economic consideration of false alarms and the slow detection of real faults. It is therefore not practical and convincing to gain a higher MSN by sacrificing alarm accuracy. Another parameter is δ, which is determined by the real process status. It is believed that this parameter is the main channel that carries process shift information and is therefore critical to the successful detection of any process failures. If no process shifts occur, this parameter is expected to be zero. Finally, there is the dimension of a chart. For the same level of δ, a lower dimension gives a higher MSN. Although the number of process variables is fixed, employing dimension reduction techniques to trim down the dimension and, in the meantime, preserve useful information will improve the sensitivity of a control chart.

4. AN ADR MONITORING SCHEME

The MSN introduced in Section 3 has demonstrated the functional relationship between the dimension and the charting performance. To the extent that useful information is maintained, reducing the dimension is one way of improving sensitivity and increasing robustness. This is also the primary motivation of popular

dimension reduction techniques used in data analysis, regression analysis and pattern recognition[28]. In this section, we propose an ADR scheme to handle the dynamic shift patterns in a multivariate application. Related design issues are discussed as well.

4.1. The ADR framework

In multivariate control chart design, dimension reduction can be achieved via evaluating the importance of each variable and abandoning less important variables. However, this issue becomes more complicated in a dynamic context in which the contribution of a variable is difficult to quantify. Consider again the example in Figure 1. The output variable, e_t, shows obvious shifts in its mean during the transient stage, which suggests that it is important to fault detection. However, its mean approaches zero as time goes by, which turns it into a redundant position. Hence, with the presence of dynamic shift patterns, the contribution of any particular variable is not fixed. A dynamic decision procedure that can evaluate the contributions of variables in real-time is therefore needed for efficient control chart design.

Following the notation used in (6), we use $\mathbf{V}_t = [e_t, x_t]^T$ to denote the vector of the latest observations and define a new vector

$$\mathbf{W}_t = \mathbf{L}_t^T \mathbf{V}_t \tag{13}$$

where $\mathbf{L}_t^T$ is an $m \times 2$ projection matrix that transforms the original vector, $\mathbf{V}_t$, to $\mathbf{W}_t$. With $m \leq 2$, the transformed vector, $\mathbf{W}_t$, has a lower or equal dimension than that of the original vector, $\mathbf{V}_t$. The new vector, $\mathbf{W}_t$, can be monitored by the following T^2 chart:

$$T_W^2 = \mathbf{W}_t^T \mathbf{\Sigma}_w^{-1} \mathbf{W}_t > h_3 \tag{14}$$

where $\mathbf{\Sigma}_w$ is the covariance matrix of $\mathbf{W}_t$, $\mathbf{\Sigma}_w = \mathbf{L}_t^T \mathbf{\Sigma} \mathbf{L}_t$. In order to achieve a false alarm rate of α, the approximate control limit of this chart is given as $h_3 = \chi_m^{-2}(1 - \alpha)$.

As far as $m < 2$ holds, the T^2 chart (14) will reduce the dimensions of the chart in (6). The projection matrix, which is subscripted by a timestamp, t, plays an essential role in this scheme. If $\mathbf{L}_t$ performs in such a way that the contribution of each variable is evaluated at each step and redundant variables are abandoned dynamically, we call the chart in (14) an ADR chart.

Let $\mathbf{\mu}_t$ be the mean vector of $\mathbf{V}_t$ at step t. Equation $\mathbf{\mu}_t = 0$ holds if the process is in-control. Further denoting the mean of $\mathbf{W}_t$ as $\mathbf{\mu}_{wt}$. $\mathbf{\mu}_{wt}$ can be expressed as a function of $\mathbf{\mu}_t$:

$$\mathbf{\mu}_{wt} = E[\mathbf{W}_t] = E[\mathbf{L}_t^T \mathbf{V}_t] = \mathbf{L}_t^T \mathbf{\mu}_t \tag{15}$$

Consequently, the NCP of chart (14) is given as

$$\delta_{wt} = \mathbf{\mu}_{wt}^T \mathbf{\Sigma}_w^{-1} \mathbf{\mu}_{wt} \tag{16}$$

According to Equation (12), the MSN of the ADR chart becomes

$$\text{MSN}_w = 1 - \chi_{m,\delta_{wt}}^2(\chi_m^{-2}(1 - \alpha)) \tag{17}$$

The key part in implementing the ADR scheme is to find the optimal $\mathbf{L}_t$ to maximize MSN_w by projecting $\mathbf{V}_t$ to a lower dimension space. Suppose that there are K possible ways of reducing the dimension of $\mathbf{V}_t$. It is convenient to express all the possible dimension reduction methods by K projection matrices, $\mathbf{L}_t(k)$, $1 \leq k \leq K$. Selecting the best dimension reduction method is equivalent to searching through the K projection matrices for the optimal one that maximizes MSN_w in (17). As $m \leq 2$ always holds, the ADR chart will reduce, or in the worse scenario, maintain the dimension of the original vector, and maximize its sensitivity.

To evaluate the MSN of each projection method, the process mean vector, $\mathbf{\mu}_t$, needs to be known, which is rarely the case in practice. In this paper, the mean vector is obtained via an efficient forecasting algorithm. The details of forecasting $\mathbf{\mu}_t$ will be introduced in a later section.

The proposed ADR procedure provides a general framework for projection-based methods. The conventional T^2 chart can be treated as a special form of the ADR chart with $\mathbf{L}_t^T = \mathbf{I}$. We can also express the

covariance matrix, $\boldsymbol{\Sigma}$, by its eigen-decomposition, $\boldsymbol{\Sigma} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}}$, where $\mathbf{U}$ is a matrix of eigenvectors of $\boldsymbol{\Sigma}$, and $\boldsymbol{\Lambda}$ is a diagonal matrix with the corresponding eigenvalues on the diagonal. By removing insignificant eigenvalues and their corresponding eigenvectors, we obtain two reduced m by m matrices, $\boldsymbol{\Lambda}_m$ and $\mathbf{U}_m$, with $m < p$. If we further denote $\mathbf{L}_t^{\mathrm{T}} = \boldsymbol{\Lambda}_m^{-1/2}\mathbf{U}_m^{\mathrm{T}}$ and use it with (13). The ADR chart reduces to a PCA-based control chart.

4.2. Projection matrices for ADR

In the ADR procedure, the reduction of dimensions is achieved via the application of projection matrices. The performance of the resulting ADR chart depends heavily on the efficiency of the matrix being used. Therefore, the projection matrix, or the candidate matrices, should be chosen carefully. The fundamental criteria for choosing appropriate projection methods include: first, the projected statistic must have a lower dimension. In the ADR procedure, reduction in dimensions is one important way of improving sensitivity. Second, the projected variables must be instructive of fault detection. Reducing the dimensions should not be achieved at the cost of sacrificing significant variables. Third, the interpretability of out-of-control alarms must be meaningful. In the event of an alarm signal, it is desirable for the chart itself to suggest which one or group of underlying variables or components is responsible for this alarm.

In this part, two different dimension reduction strategies are studied. Correspondingly, two collections of projection matrices are proposed to implement the reduction. The first strategy is built on the original variables, whereas the second strategy is developed based on components obtained from an orthogonal decomposition algorithm.

In regression analysis, variable selection involves building a regression model between response variables and explanatory variables[28]. Significance testing of regression coefficients helps to remove insignificant explanatory variables. In the ADR procedure, although the objective of maximizing MSN is different from the purpose of conventional regression, the idea used in variable selection can still be borrowed. Each variable is either included in or expelled from the final model. In the first dimension reduction strategy, which is referred as ADR-1 in the remainder of this paper, we employ a naive enumerative algorithm. All possible combinations of the variables are tried, and the one with the highest MSN is chosen for control chart setup. Those variables that do not appear in the combination are naturally discarded. In the feedback-controlled process, two variables, e_t and x_t, are monitored; there are a total of three possible combinations, which are shown as follows:

$$\mathbf{L}(1) = [1, 0]^{\mathrm{T}}$$

$$\mathbf{L}(2) = [0, 1]^{\mathrm{T}}$$

$$\mathbf{L}(3) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{18}$$

The resulting statistics after applying the three matrices in (18) are e_t, x_t, and $\mathbf{V}_t$, respectively. As each variable has a clear physical meaning, the interpretation of an out-of-control signal becomes straightforward. The corresponding covariance matrix of each projection is

$$\Sigma(1) = \sigma_e^2$$

$$\Sigma(2) = \sigma_x^2 \tag{19}$$

$$\boldsymbol{\Sigma}(3) = \boldsymbol{\Sigma}$$

The ADR-1 scheme is analogous to the diagnostic method presented by Doganaksoy et al.[29], in which each variable of a vector is checked when an out-of-control signal is triggered. However, the purpose here is to reduce the dimension of the aggregate T^2 statistic. If a variable becomes insensitive to a process shift, such as the e_t in its steady state, monitoring the remaining variables is expected to be more efficient.

In the second strategy, which is referred to as ADR-2 in the remainder of this paper, independent components are firstly obtained via an orthogonal decomposition. The reduction of dimensions is then achieved by making a decision on which components are less informative and should be dropped.

Let $T_x = x_t/\sigma_x$, $T_{e \cdot x} = (e_t - x_t \sigma_{ex}/\sigma_x^2)/\sqrt{\sigma_e^2 - \sigma_{ex}/\sigma_x^2}$. Mason *et al.*[30] suggested that the T^2 statistic in (6) can be decomposed in two parts

$$T^2 = T_x^2 + T_{e \cdot x}^2 \tag{20}$$

which is called the MYT-decomposition of the T^2 (Reference[31]). The two components, T_x and $T_{e \cdot x}$, obtained from the decomposition are perpendicular. In order to fit the MYT-decomposed components into the projection framework, we define

$$\mathbf{L}(1) = [0, 1/\sigma_x]^{\mathrm{T}}$$
$$\mathbf{L}(2) = \mathbf{\Sigma}^{-1}[b_1, 0]^{\mathrm{T}}$$
$$\mathbf{L}(3) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{21}$$

where $b_1 = \sqrt{\sigma_e^2 - \sigma_{ex}^2/\sigma_x^2}$. The corresponding variance–covariance structure is

$$\mathbf{\Sigma}(1) = 1$$
$$\mathbf{\Sigma}(2) = 1 \tag{22}$$
$$\mathbf{\Sigma}(3) = \mathbf{\Sigma}$$

In the ADR-2 scheme, $\mathbf{L}(1)$ transforms the vector $\mathbf{V}_t = [e_t, x_t]^{\mathrm{T}}$ to the standardized input, x_t/σ_x, which is the square root of the first component in (20). Applying projection $\mathbf{L}(2)$ to $\mathbf{V}_t$ yields

$$\mathbf{W}_t = \mathbf{L}_t^{\mathrm{T}} \mathbf{V}_t = [b_1, 0] \mathbf{\Sigma}^{-1} \begin{bmatrix} e_t \\ x_t \end{bmatrix} = [b_1, 0] \begin{bmatrix} \dfrac{e_t - x_t \sigma_{ex}/\sigma_x^2}{\sigma_e^2 - \sigma_{ex}/\sigma_x^2} \\ \dfrac{x_t - e_t \sigma_{ex}/\sigma_e^2}{\sigma_x^2 - \sigma_{ex}/\sigma_e^2} \end{bmatrix} = T_{e \cdot x} \tag{23}$$

which is the square root of the second component in (20). Therefore, implementing $\mathbf{L}(1)$ and $\mathbf{L}(2)$ is equivalent to reducing the dimension via MYT-decomposition.

As noted by Mason *et al.*[30], there are other ways of decomposing the T^2 statistic, which implies that there are other options in choosing the projection matrices. In practice, good interpretability of decomposed components is one important criterion. In (21), the physical meaning of the projected components is straightforward. An alarm reported by $\mathbf{L}(1)$ means that abnormal signals are found in the input stream. An out-of-control signal triggered by $\mathbf{L}(2)$ suggests abnormalities in the conditional output when the status of the input is known to be in-control. Based on (20), the use of $\mathbf{L}(3)$ leads to the monitoring of the original vector, which means that neither x_t nor e_t has sufficient evidence to report an alarm, while their overall effect suggests the existence of a process failure.

It is interesting to note that the $\mathbf{L}(2)$ projection is equivalent to the U_0 statistic in Jiang[11]. The U_0 chart is designed for utilizing individually. The ADR-2 scheme, however, has aggregated it with other statistics to produce an adaptive framework. The automatic switches among different statistics are expected to take advantage of each one and to produce a much more powerful chart.

4.3. Extensions to general multivariate applications

Although the two projection schemes in (18) and (21) are given based on the feedback-controlled process in (1), it can be easily extended to a general environment with multiple variables.

In order to find the maximum of MSN, the projection scheme in (18) involves enumerating all possible combinations of variables. Suppose that there are p variables being measured and that the total number of feasible combinations is $2^p - 1$ (the combination with all variables excluded is ignored). The amount of computational power will increase exponentially with p. Therefore, the ADR-1 chart is only feasible for applications with small sets of variables.

In the ADR-2 procedure, the MYT-decomposition plays the central role in maximizing MSN. Let the p process variables be $z_1, z_2, \ldots, z_p$. Furthermore, we use $T_j.$ to denote the jth components of MYT-decomposition, which is obtained by adjusting the jth variable, z_j, with regard to all the $(j-1)$ variables, $z_1, z_2, \ldots, z_{j-1}$. The calculation of the MYT-decomposition is discussed by Mason *et al.*[30].

In a multivariate context, we still face the problem of choosing the best combination of the MYT-decomposed components. As the number of components equals the number of the original variables, the direct search needs to go through $2^p - 1$ possibilities again. However, the search procedure can be simplified by incorporating the orthogonal property of the components. The search procedure is listed as follows:

(1) For the p components, calculate their respective MSN values.
(2) Sort the components from high to low based on MSN values.
(3) Set $k = p$; calculate the MSN for the T^2 chart on the first k components, $(T_1., \ldots, T_k.)$.
(4) Calculate the MSN for the T^2 chart on the first $(k-1)$ components, $(T_1., \ldots, T_{k-1}.)$.
(5) If the MSN obtained in step 4 is larger than the MSN for the chart on the first k components, reduce k by 1 and go back to step 4; otherwise, set $k^* = k$, and choose to monitor the first k^* components.

The above procedure is analogous to the backward variable selection in conventional regression analysis[32], in which variables are sequentially dropped based on their significance to the response variable. In the ADR-2 procedure, components are added or dropped based on their contribution to the MSN. The components are first sorted based on their relative importance. Then, the smallest component is repeatedly dropped, until MSN stops increasing. This stopping rule implies that no more components can be dropped without reducing the MSN. Therefore, the maximum of the MSN will be achieved by monitoring the remaining components. As the complexity of this procedure increases linearly with p, it will significantly alleviate the computational demand of applications with large numbers of variables.

The sequence of the original variables, $z_1, z_2, \ldots, z_p$, influences the way that the components are generated. Therefore, a carefully selected sequence is desired for easy interpretation of future out-of-control signals. Analogous to the regression adjustment method in Hawkins[33], we recommend denoting upstream variables with lower subscripts and downstream variables with higher subscripts. Thus, the MYT-decomposed components will convey meaningful physical explanations.

4.4. Discussions on the mean vector estimation

In the implementation of the proposed ADR procedures, the real-time process mean vector is always needed when evaluating and comparing the MSN of different statistics. As the mean vector is rarely known in practice, we will employ forecasting algorithms to estimate the mean vector and set up the ADR charts for process monitoring.

In general, a forecasting method can be either model-based or model-free. A model-based method involves fitting a time-series model to a historical data set to gain a model for the process[34,35]. Then, a one-step-ahead estimation can be obtained from the model. The downside of this method is its heavy dependency on model accuracy. Estimating a time-series model for a process usually requires a large data set, especially when a high-order model is used (Box *et al.*[36]).

A model-free method, on the other hand, does not assume a particular model for the process. It uses weights or kernels to smooth out noise and to discover the running trend of a process. Among them, the exponentially weighted moving average (EWMA) procedure assigns exponentially falling off weights to historical data to obtain an estimation that balances both the latest and the historical data. Alwan and Roberts[37] pointed out that, for most cases, the EWMA procedure acts as a good approximation of a time-series process. Therefore,

the application of the EWMA procedure is recommended in this paper due to its simplicity and its reasonably good forecasting performance.

Let $\boldsymbol{\mu}_t = [\mu_e, \mu_x]^T$ be the forecasted mean vector of $\mathbf{V}_t$ at time step t. The mean vector at $t+1$ is given as

$$\boldsymbol{\mu}_{t+1} = \lambda\boldsymbol{\mu}_t + (1-\lambda)\mathbf{V}_t \tag{24}$$

where λ is a smoothing parameter that determines the falling speed of the weights of historical data in forecasting a new mean vector. The above modeling procedure is analogous to the multivariate EWMA procedure for monitoring multivariate applications. The limiting form of the variance of $\boldsymbol{\mu}_t$ is given as

$$\boldsymbol{\Sigma}_\mu = \frac{\lambda}{2-\lambda}\boldsymbol{\Sigma} \tag{25}$$

The tuning parameter, λ, provides a lever to adjust how quickly a dynamic process is followed by the forecasted sequence. If a larger λ is used, more weight is assigned to the latest observations. Therefore, a sudden process shift will be captured faster. However, the forecasting is easily misled by process fluctuation and noise. In contrast, a smaller λ provides more stable forecasting, while its response to large sudden shifts may be slow.

The trajectories of two PI-controlled processes are shown in Figure 1. The corresponding EWMA forecasts are also plotted. As is seen from Figure 1(a), in the transient stage, when the process undergoes strong oscillations, the EWMA estimation approaches the mean of the process gradually. In the steady-state stage, the process is stabilized; the estimated mean almost overlaps with the true process mean. In Figure 1(b), the process decays smoothly when a sudden shift occurs. The EWMA procedure provides an even better forecasting performance.

5. PERFORMANCE STUDY OF THE ADR SCHEME

The performance of quality-control charts is commonly measured by their ARL. To explore this, we carry out simulation studies in this section to compare the proposed method with existing schemes.

Several existing schemes are chosen for comparison. The first chart is the general T^2 chart (denoted as GT2 hereafter) proposed by Tsung *et al.*[9], which involves a joint monitoring of e_t and x_t based on the vector $\mathbf{V}_t = [e_t, x_t]^T$. The second and the third charts to be considered are the U_0 chart and the U_∞ chart proposed by Jiang[11]. The fourth and the fifth competitors are two multi-charts. By combining multiple Shewhart charts, the resulting multi-chart signals if any individual chart signals[38]. As all projected statistics derived in (18) and (21) are possibly useful for fault detection, a collection of Shewhart charts can be set up for the statistics. We denote the multi-chart derived from (18) and (21) as Multi-1 and Multi-2, respectively.

The process being investigated follows model (1) with $\phi = 0.8$ and $\theta = 0.3$. The optimal PI controller suggested by Tsung *et al.*[19] has $k_P = -0.125$ and $k_I = -0.427$. Process shifts modeled by (4) are added to the process. All shift magnitudes, s, satisfy $0 \leq s \leq 5$, which covers both small and large shift ranges. For a fair comparison, the in-control ARL of all charts is forced to be 200, and each ARL is computed using at least 100 000 replicates.

We first review the performance of the competing charts. The ARLs of these charts are shown in Table II. The results clearly show that for small mean shifts, the x_t chart and the U_∞ chart have shorter ARLs than other charts. This finding is consistent with the conclusions drawn by Jiang[11]. In the presence of large mean shifts, the GT2 chart and the Multi-1 and Multi-2 charts outperform all others. The optimality of the GT2 chart in a large shift range is explained by its Shewhart-type property. In contrast to a cumulative-sum- or EWMA-type chart, a Shewhart-type chart is not influenced by historical observations and can respond quickly to sudden process changes. With regard to the Multi-1 and the Multi-2 charts, all informative variables are monitored, and these variables contribute to the fast detection of large mean shifts.

Table III shows the charting performance of the proposed ADR charts. Different smoothing parameters are employed by each chart, as shown in the second row of Table III. A careful examination of Table III

Table II. Performance comparison of competing charts

s	GT2	U_0	U_∞	e_t	x_t	Multi-1	Multi-2
0.0	199.97	199.94	200.04	200.02	200.05	200.47	200.10
0.5	149.88	177.89	124.44	198.68	126.25	153.51	149.15
1.0	80.26	132.13	56.48	194.02	57.74	82.88	79.16
1.5	40.92	89.86	27.89	181.68	28.50	41.17	40.28
2.0	21.20	59.13	15.36	158.37	15.56	20.68	21.22
2.5	11.20	37.53	9.28	123.38	9.22	10.64	11.59
3.0	5.87	22.47	6.08	83.21	5.84	5.49	6.38
3.5	3.16	12.41	4.32	47.34	3.93	2.94	3.57
4.0	1.85	6.37	3.32	21.72	2.79	1.75	2.12
4.5	1.28	3.17	2.72	8.44	2.07	1.24	1.42
5.0	1.08	1.74	2.36	2.96	1.60	1.07	1.14

Table III. ARL of the ADR-1 and the ADR-2 charts

s	ADR-1			ADR-2		
	$\lambda=0.01$	$\lambda=0.1$	$\lambda=0.5$	$\lambda=0.01$	$\lambda=0.1$	$\lambda=0.5$
0.0	199.17	200.66	200.35	199.93	199.82	199.91
0.5	123.32	139.95	151.09	123.77	134.97	148.30
1.0	54.87	67.83	80.05	55.98	64.30	77.89
1.5	25.75	32.12	39.39	27.00	31.26	39.06
2.0	12.74	15.80	19.74	13.97	16.14	20.33
2.5	6.50	8.02	10.12	7.52	8.69	11.01
3.0	3.43	4.16	5.22	4.14	4.78	6.02
3.5	1.98	2.30	2.81	2.41	2.74	3.37
4.0	1.34	1.47	1.69	1.56	1.72	2.01
4.5	1.10	1.15	1.22	1.19	1.25	1.38
5.0	1.02	1.04	1.06	1.06	1.08	1.12

shows that ADR-1 and the ADR-2 charts exhibit comparable performance over the whole shift range. For $\lambda=0.5$ and $\lambda=0.1$, ADR-2 is favored over ADR-1 for small shifts; whereas for $\lambda=0.01$, the performance of ADR-1 is slightly better than that of ADR-2 for all shift magnitudes. However, the large shift performance of both charts has not improved when λ increases. Therefore, a smaller smoothing parameter is always recommended for practitioners. Extensive simulations show that the value of 0.01 for λ is a reasonably good choice.

Before further comparisons of the ADR schemes and the competing charts in Table II, we conduct a thorough study of the way in which the ADR schemes work. As is seen from (18) and (21), both schemes consist of multiple projection matrices. At each step, only one of the projected statistics is chosen to function. Therefore, isolating effective statistics in each scheme can help us to identify the contribution of each statistic to the overall performance.

Figure 2 demonstrates the switching patterns in the ADR-1 chart. Two hundred out-of-control signals are collected. The height of each bin represents the number of alarms released by the corresponding projection matrix, which is given in (18).

As is seen from Figure 2, for small shifts, $s=0.5$ or 1.0, the alarms are mostly accumulated on scale $L(2)$. As the second projected statistic is x_t, this result suggests that x_t plays an instructive role in detecting small shifts. Small shifts usually require long runs before they are detected. During the undetected out-of-control period, as the trajectories in Figure 1 suggest, the output, e_t, approaches zero gradually. However, a sustained change is observed in x_t. Therefore, it is x_t that contributes most to the identification of small shifts. When the shift magnitude increases to 2.0 or 3.0, the possibility that the first projected statistic, e_t, be chosen
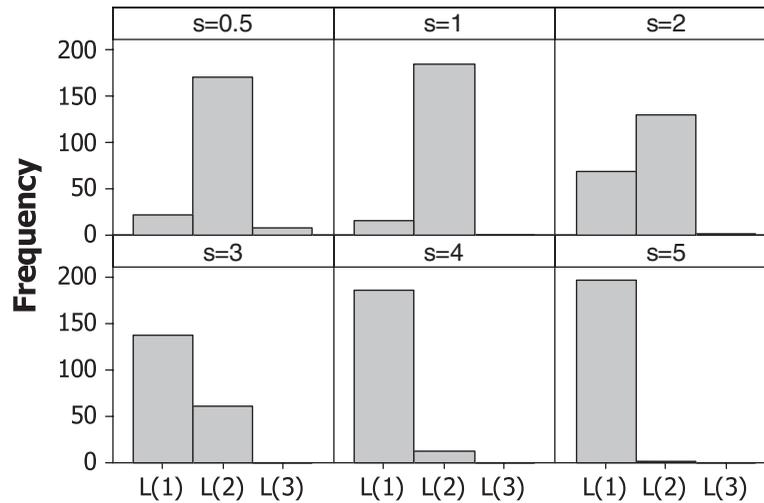
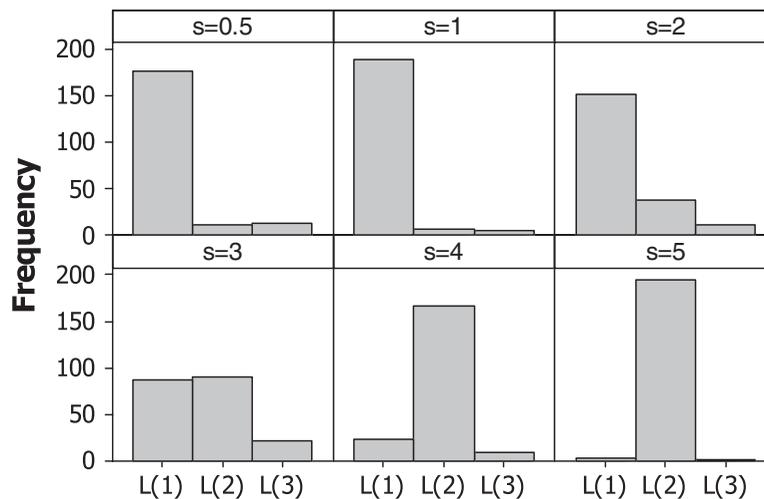Figure 2. Switching patterns of the ADR-1 chart

Figure 3. Switching patterns of ADR-2 chart

increases. Starting from $s = 3.0$, e_t is more frequently chosen than x_t, which suggests that e_t becomes more informative in detecting large mean shifts. This is consistent with the conclusions drawn in Jiang[11] and Tsung and Tsui[8], who concluded that monitoring output e_t is more efficient for large shift detection.

Similar switching patterns are found in Figure 3, in which the run-time status of the ADR-2 chart is recorded. For small shifts, the out-of-control alarms are frequently reported by projection matrix 1, which corresponds to the direct monitoring of x_t. As the magnitude of the process shifts increases, projection 2 is more frequently chosen. The patterns agree with the findings by Jiang[11], which suggested that the U_0 is sensitive to large mean shifts.

Now we compare the ADR-1 and the ADR-2 schemes with the charts in Table II. First, we focus on large shifts. Compared with the GT2 chart, which is claimed to be sensitive to large shifts, significant improvement is observed with the ADR-1 and the ADR-2 charts. This suggests that the ADR schemes are always favored

if large shifts are of interest. For small shifts, both the ADR-1 and the ADR-2 charts with small λ outperform the U_∞ chart.

Overall, the benefit of applying the ADR charts is obvious. Conventional monitoring schemes usually stick to a particular statistic. The ADR scheme, however, switches automatically among the projected statistics and chooses the most efficient one for functioning at each step. The dimension reduction principle guarantees the optimality of each statistic, and the ADR procedure successfully takes advantage of each statistic to produce a method with overall good performance.

6. CONCLUSIONS

In order to improve the performance of a multivariate control chart, it is always desirable to reduce dimensions and preserve only informative variables for monitoring. However, in a feedback-controlled process, most variables have time-varying shifts and their contribution to fault detection also changes over time, which makes it difficult to choose important variables statically. A dynamic and ADR scheme is therefore of great need.

In this paper, we first proposed an MSN. The MSN maps the detection power of multivariate charts to a common scale and compares their sensitivity in detecting a particular shift. Multivariate charts with diverse dimensions and false alarm rates can be compared on the basis of MSN.

An ADR scheme has been proposed and two-dimensional reduction procedures are considered to project the original vector into a low-dimensional space. Based on MSN, the ADR scheme makes decisions on choosing the most informative variables or components at each step to maximize the detection power. Simulation results have demonstrated that the ADR scheme substantially improves both large and small shift performance, and the two charts, ADR-1 and ADR-2, have demonstrated almost equally good performance.

Although developed for a feedback-control process, the ADR charts can be easily extended to a general application with multiple process variables. The ADR-1 chart may suffer from huge demand for computational power, while the searching algorithm of the ADR-2 chart has been optimized to minimize the demand for computational power. Therefore, the ADR-2 chart is recommended for application to a large number of variables.

Furthermore, the ADR scheme is a flexible framework and can be easily extended. New dimension reduction techniques can be fitted to the unified framework in Equation (14) without significant modification, which is a topic that deserves additional research effort.

REFERENCES

1. Capilla C, Ferrer A, Romero R, Hualda A. Integration of statistical and engineering process control in a continuous polymerization process. *Technometrics* 1999; **41**(1):14–28.
2. Tucker WT, Faltin FW, VanderWiel SA. Algorithmic statistical process-control—An elaboration. *Technometrics* 1993; **35**(4):363–375.
3. Montgomery DC, Keats JB, Runger GC, Messina WS. Integrating statistical process control and engineering process control. *Journal of Quality Technology* 1994; **26**(2):79–87.
4. Woodall WH, Montgomery DC. Research issues and ideas in statistical process control. *Journal of Quality Technology* 1999; **31**(4):376–386.

5. Wang K, Tsung F. Monitoring feedback-controlled processes using adaptive T2 schemes. *International Journal of Production Research* 2007; **45**(23):5601–5619.
6. Faltin FW, Tucker WT. On-line quality control for the factory of the 1990s and beyond. *Statistical Process Control in Manufacturing*, Keats JB, Montgomery DC (eds.). Marcel Dekker: New York, 1991.
7. Box GEP, Kramer T. Statistical process monitoring and feedback adjustment—A discussion. *Technometrics* 1992; **34**(3):251–267.
8. Tsung F, Tsui KL. A mean-shift pattern study on integration of SPC and APC for process monitoring. *IIE Transactions* 2003; **35**(3):231–242.
9. Tsung F, Shi J, Wu CFJ. Joint monitoring of PID-controlled processes. *Journal of Quality Technology* 1999; **31**(3): 275–285.
10. Tsung F, Apley DW. The dynamic T-2 chart for monitoring feedback-controlled processes. *IIE Transactions* 2002; **34**(12):1043–1053.
11. Jiang W. A joint monitoring scheme for automatically controlled processes. *IIE Transactions* 2004; **36**(12):1201–1210.
12. Runger GC, Alt FB. Choosing principal components for multivariate statistical process control. *Communications in Statistics—Theory and Methods* 1996; **25**(5):909–922.
13. Murphy BJ. Selecting out of control variables with the T2 multivariate quality control procedure. *The Statistician* 1987; **36**:571–583.
14. Mastrangelo CM, Runger GC, Montgomery DC. Statistical process monitoring with principle components. *Quality and Reliability Engineering International* 1996; **12**:203–210.
15. Jackson JE. *A User's Guide to Principle Components*. Wiley: New York, 1991.
16. Ingolfsson A, Sachs E. Stability and sensitivity of an EWMA controller. *Journal of Quality Technology* 1993; **25**(4): 271–287.
17. Sachs E, Hu A, Ingolfsson A. Run by run process control—Combining SPC and feedback-control. *IEEE Transactions on Semiconductor Manufacturing* 1995; **8**(1):26–43.
18. DelCastillo E, Hurwitz AM. Run-to-run process control: Literature review and extensions. *Journal of Quality Technology* 1997; **29**(2):184–196.
19. Tsung F, Wu HQ, Nair VN. On the efficiency and robustness of discrete proportional-integral control schemes. *Technometrics* 1998; **40**(3):214–222.
20. Del Castillo E. *Statistical Process Adjustment for Quality Control*. Wiley: New York, 2002.
21. Nembhard HB, Kao MS. Adaptive forecast-based monitoring for dynamic systems. *Technometrics* 2003; **45**(3): 208–219.
22. Runger GC. Assignable causes and auto correlation: Control charts for observations or residuals? *Journal of Quality Technology* 2002; **34**(2):165–170.
23. Tsung F, Shi JJ. Integrated design of run-to-run PID controller and SPC monitoring for process disturbance rejection. *IIE Transactions* 1999; **31**(6):517–527.
24. Jiang W, Tsui KL, Woodall WH. A new SPC monitoring method: The ARMA chart. *Technometrics* 2000; **42**(4): 399–410.
25. Jiang W, Tsui KL. SPC monitoring of MMSE- and PI-controlled processes. *Journal of Quality Technology* 2002; **34**(4):384–398.
26. Jiang W. Multivariate control charts for monitoring autocorrelated processes. *Journal of Quality Technology* 2004; **36**(4):367–379.
27. Lowry CA, Montgomery DC. A review of multivariate control charts. *IIE Transactions* 1995; **27**(6):800–810.
28. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*: *Data Mining*, *Inference*, *and Prediction*. Springer: New York, 2001.
29. Doganaksoy N, Faltin FW, Tucker WT. Identification of out of control quality characteristics in a multivariate manufacturing environment. *Communications in Statistics—Theory and Methods* 1991; **20**(9):2775–2790.
30. Mason RL, Tracy ND, Young JC. Decomposition of T-2 for multivariate control chart interpretation. *Journal of Quality Technology* 1995; **27**(2):99–108.
31. Mason RL, Young JC. Improving the sensitivity of the T-2 statistic in multivariate process control. *Journal of Quality Technology* 1999; **31**(2):155–165.
32. Weisberg S. *Applied Linear Regression* (2nd edn). Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics. Wiley: New York, 1985.
33. Hawkins DM. Regression adjustment for variables in multivariate quality control. *Journal of Quality Technology* 1993; **25**(3):170–182.
34. Pandit SM, Wu SM. *Time Series and System Analysis with Applications*. Krieger: Malabar, FL, 1993.

35. Apley DW, Tsung F. The autoregressive T-2 chart for monitoring univariate autocorrelated processes. *Journal of Quality Technology* 2002; **34**(1):80–96.

36. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis*: *Forecasting and Control*. Prentice-Hall: Englewood Cliffs, NJ, 1994.

37. Alwan LC, Roberts HV. Time series modeling for statistical process control. *Journal of Business & Economic Statistics* 1989; **6**(1):87–95.

38. Sparks RS. CUSUM charts for signaling varying location shifts. *Journal of Quality Technology* 2000; **32**(2):157–171.

Authors' biographies

Dr Kaibo Wang is an Assistant Professor in the Department of Industrial Engineering at Tsinghua University. He received his BS and MS degrees in Mecha-tronics from Xi'an Jiaotong University, Xi'an, People's Republic of China, and PhD in Industrial Engineering and Engineering Management from Hong Kong University of Science and Technology (HKUST), Hong Kong. He is an ASQ Certified Six Sigma Black Belt. He is on the Editorial Board for Journal of the Chinese Institute of Industrial Engineers (JCIIE). He has published papers on Journal of Quality Technology, Quality and Reliability Engineering International, International Journal of Production Research and others. His research interests include quality management, statistical quality control, statistical process control, and run-to-run process control.

Prof. Fugee Tsung is Director of the Quality Lab and Professor in the Department of Industrial Engineering and Logistics Management at the Hong Kong University of Science & Technology. He received both his MS and PhD from the University of Michigan, Ann Arbor. He is currently an Associate Editor of Technometrics, a Department Editor of the IIE Transactions, and on the Editorial Boards for Quality and Reliability Engineering International (QREI), the International Journal of Reliability, Quality and Safety Engineering (IJRQSE) and the International Journal of Six Sigma and Competitive Advantage (IJSSCA). He is an ASQ Certified Six Sigma Black Belt, ASQ authorized Six Sigma Master Black Belt Trainer, and former Chair of the Quality, Statistics, and Reliability (QSR) Section at INFORMS. He is also the winner of the Best Paper Award for the IIE Transactions in 2003. His research interests include quality engineering and management, statistical process control, monitoring and diagnosis.