STATISTICAL MONITORING PERFORMANCE FOR STARTUP OPERATIONS IN A FEEDBACK CONTROL SYSTEM

HARRIET BLACK NEMBHARD1*, CHRISTINA M. MASTRANGELO2 AND MING SHU KAO1
1 Department of Industrial Engineering, University of Wisconsin-Madison, 1513 University Ave., Madison WI 53706, USA
2 Department of Systems Engineering, University of Virginia, Thorton Hall, Charlottesville VA 22903, USA

SUMMARY
Previously, it has been held that statistical process control (SPC) and engineering process control (EPC) were two distinct domains for process improvement. However, we specifically consider the impact for integrating the two approaches on a first-order dynamic system with ARIMA disturbances. We show how to model and analyze this system over a range of practical conditions. Our work results in a set of response surfaces that characterize the performance of the integrated design. We also compare these results to the case where the SPC and EPC policies are applied separately. In general, we find that the EPC approach performs best in terms of minimizing error, but that we can reduce the number and magnitude of adjustments using the integrated monitoring and control approach. This work also further supports our earlier findings that the integrated design is effective on complex dynamic systems during the initial transient or startup period. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: statistical process control (SPC); engineering process control (EPC); dynamic processes; time series; simulation

1. INTRODUCTION
With globalization of many aspects of our culture and technology, manufacturing environments are facing higher demands with the introduction of new products and shorter production cycles of existing products. Such changes inevitably invoke a transient or startup period in the basic system operations. Product grade change, recipe change and raw material change are examples of these operations. An immediate consequence of these operational changes is that a process tends to move from the existing level to another level. Further, the process response to the changes is often sluggish resulting in a substantial startup period of transient output. A key question for quality engineering is: what action(s) should be taken during the transition period? This is a very complex question with many facets. To focus the question, we first consider a motivating case from the plastics industry and discuss an appropriate monitoring and control philosophy.

1.1 Motivation
In order to convey the nature of a startup period, we draw an example using an extrusion process that produces plastic parts. The operation involves a color changeover which is routine in the plastics industry. Parts are usually made in a variety of colors to appeal to a wide range of consumers. Because of the dynamic nature of the extrusion system, a color change operation will produce a startup period of mixed-color (or undesirable) parts that do not conform to specifications. That is, the next targeted color parts will appear after a period of mixed-color parts.

To illustrate, consider Figure 1 which depicts observations on color intensity measured from plastic samples of a color changing operation. The old color had a target intensity level of 0.43. At sample \( t = 18 \), a transition began to a new color with a target intensity level of 0.85. These plastic samples were collected at approximately 30 s apart. Observe that the operation had reached a steady level around the old color intensity target. Then sample color intensity moves from that target to a new target over a period with fluctuating response. Notice that no monitoring or control has been applied in this example so the data

*Correspondence to: Dr. H. B. Nembhard, Department of Industrial Engineering, University of Wisconsin-Madison, 1513 University Ave., Madison WI 53706, USA. Email: hbnem@engr.wisc.edu

Copyright © 2001 John Wiley & Sons, Ltd.

Received 28 February 2001
Revised 2 July 2001
represent the natural response of the process. Ideally, we would like the change to occur instantaneously because the output during the transition is unusable.

1.2. Philosophy and contribution

Our intent is to show how the transient period of a first-order system performance can be improved using what we have termed integrated process control (IPC). An IPC policy uses engineering process control (EPC) to make adjustments to the process and statistical process control (SPC) to monitor the process. We initially reported on using IPC for startup operations in Nembhard and Mastrangelo [1]. That earlier work introduced the idea of IPC for a system with stationary disturbance.

In this paper, there are four main contributions:

- an analysis of the performance of the IPC design for startup operations with a nonstationary autoregressive integrated moving average (ARIMA) disturbance model which mimics the changeover process shown in Figure 1;
- an investigation of the range of responses that may occur in a dynamic system;
- the development of a set of response surfaces for the key parameters that help to identify robust conditions of the IPC design; and
- a comparison of the IPC strategy to its component EPC and SPC policies—that is where the adjustment and monitoring policy is implemented separately—to determine the circumstances under which IPC provides an advantage.

A part of the motivation behind our IPC designs is to use basic tools that are most relevant to practitioners. The IPC designs discussed here use a proportional-integral (PI) controller to provide the EPC component. To provide the SPC component, we use a moving center-line exponentially weighted moving average chart (MCEWMA) for individual observations. The MCEWMA chart is based on the familiar EWMA chart that is also standard in the literature; however, it adapts the EWMA for the autocorrelated data given by the ARIMA disturbance model (Montgomery and Mastrangelo [2]).

The main measure of improvement is to reduce the sum of squared errors (SSE) from target, where the target changes when the transition begins. So, for example, in the color intensity case the SSE would be measured against 0.43 for $0 < t < 18$ and against 0.85 for $18 < t < 60$. We also consider the number of adjustments and the magnitude of adjustments because, in many operations, a decrease in these metrics is desirable even at the expense of an increase in output variation (Box and Luceño [3]). We found that while the IPC policy has a larger mean square error than EPC, it is a good control strategy when adjustment costs are non-zero. The mean square error of the IPC strategy may even be improved with knowledge of the noise model that may come from using identification procedures on observational data. If the noise process is highly non-stationary, the IPC strategy is comparable with an EPC policy.

1.3. Recent related work

There has been a significant amount of discussion in the literature on the differences of SPC and EPC in terms of their applications context. Traditional SPC emphasizes that any observed difference between a controlled process variable and its desired value
may be due to inherent common cause variation (thus no system change is necessary) or systematic special cause variation (where a system change may be necessary). The goal is to monitor the process so that there is a good chance of detecting the special cause so that it can be removed. However, several authors have illustrated that the traditional control charts will not do a good job of detection in the case of autocorrelated data. Alwan [4], Harris and Ross [5], and Wardell et al. [6] investigate the robustness of traditional control charts to autocorrelated data that can be represented by an ARMA(1, 1) process. They show that even moderate levels of autocorrelation have a significant impact on a control chart and will give false ARL performance.

While performing the monitoring task, the underlying assumption is that there is a disincentive to make an adjustment after every observation period because they are costly (e.g. see Box and Kramer [7], MacGregor and Harris [8], and Janakiram and Keats [9]). On the other hand, traditional EPC assumes that the process input variables can be adjusted as frequently as desired, that the adjustments are virtually cost free, and that the process output variables can be measured on-line without any associated uncertainty (Ogunnaike and Ray [10]). The amount of adjustment is often determined using some variation of a proportional-integral-derivative (PID) controller. This type of controller is a very robust tool for process adjustments and enjoys a high esteem in the process-control literature (e.g. see Seborg et al. [11], Ogunnaike and Ray [10], and Dorf and Bishop [12]). In process-control applications, PI control is responsible for more than 95% of the system control loops (Aström and Hagglund [13]).

Recently, we have seen an increasing research enthusiasm in combining (or integrating) SPC and EPC techniques despite their differences. MacGregor [14] points out the utility of control charts on controlled systems to enhance system performance. Process control by means of SPC may be designed to accommodate for the costs associated with adjustment and sampling (i.e. adjusting the process at each opportunity is not economical). To address the cost constraint problem, Box and Kramer [7] suggest a bounded feedback adjustment scheme that makes adjustment only when an EWMA forecast moves outside the upper and lower boundaries or dead band. The boundaries are like control-chart limits but cost based. Part of our work is related to the Box and Kramer approach. In the IPC model, we make PI controller adjustments only when observations move outside the limits of the MCEWMA chart. The design of our controller is somewhat different, however, because of the transient period dynamics that are not considered in the Box and Kramer work.

The algorithmic statistical process control (ASPC) methodology proposed by Vander Wiel et al. [15] is an excellent work on the concurrent use of SPC scheme (for removing sources of variability) and EPC technique (for compensating the process deviation from target) for improving product and process quality. They demonstrated the integration of EPC (minimum mean square error (MMSE) control) and SPC (CUSUM) on a pure gain polymerization process. Several other works that advocate the concurrent implementation of SPC and EPC include for example, Box and Luceño [1], Janakiram and Keats [9], Capilla et al. [16], Montgomery et al. [17], and Tsung et al. [18].

1.4. Organization

The remainder of the paper is organized as follows. Section 2 shows the make-up and behavior of a first-order system and its state-space representation. Section 3 shows how to model ARIMA disturbances using the same state-space form. Section 4 gives a description of the equations that represent the EPC and SPC components of the IPC design. Section 5 shows a graphical simulation model that represents the first-order system with disturbance and also the integrated control mechanism applied to the system. In Section 6, we present a set of experiments conducted with the simulation model and analyze the results. Finally, in Section 7 we summarize the work and make some concluding remarks.

2. FIRST-ORDER DYNAMICS IN THE STARTUP PROBLEM

Under the Shewhart philosophy, a steady-state process is monitored for a sign of a special cause (e.g. a leak in a mixing-tank pipe). The underlying assumption is that the process will operate at a given level unless an unexpected event occurs. This type of system has been well investigated. By contrast, we focus on the startup problem where a process is operating at one level for some time (and has presumably reached a steady-state condition) and then has to achieve a new operating level. Due to the dynamics of the system, the new level will not be achieved instantaneously.

We consider the discrete-time first-order system, which is commonly used in control engineering. Our specific interest in this system is what happens to the output $y$ in the time periods immediately following a
step change in the input $u$. The step input is given by

$$u(k) = \begin{cases} 
0 & \text{if } k < c, \\
M & \text{otherwise}
\end{cases}$$

where $c$ is the time at which the step change of magnitude $M$ occurs. The discrete-time first-order process can be described by the following difference equation:

$$y(k) = ay(k - 1) + bu(k - 1) \quad (2)$$

where $y(k)$ is $k$th sample of the output sequence or time $kT$, where $T$ is the sampling interval, $u(k)$ is the input of the system, $a$ measures the degree of inertia for the process dynamics and $0 < a < 1$, and $b$ is some constant. In general, a discrete-time dynamic process of order $n$ can be represented by the $n$th-order difference equation

$$y(k) = \sum_{i=1}^{n} a_i y(k-i) + \sum_{j=0}^{l} b_j u(k-i) \quad (3)$$

where $y(k)$ and $u(k)$ are the systems output and input and $a_i$ ($i = 1, 2, \ldots, n$) and $b_j$ ($j = 0, 1, 2, \ldots, l$) are the corresponding coefficients (e.g. see Ogata [19] and Ogunnaike and Ray [10]).

The state-space representation of process dynamics is often preferred because the vector format allows for easy updating of the equations in each time period as a new observation becomes available. For this reason, we use the state-space representation here to facilitate the simulation modeling in Section 5. (The interested reader may refer to Seborg, Edgar, and Mellichamp [11] for an introduction to state-space equations for dynamic systems.) The state-space form for the discrete-time dynamic process described by equation (3) has the form

$$x(k + 1) = Ax(k) + Bu(k)$$

$$y(k) = Cx(k) + Du(k)$$

where $A$ is the state matrix, $B$ the input matrix, $C$ the output matrix, $D$ the direct transition matrix, and $x$ is the relational matrix that links $u$ and $y$. The discrete-time sequence, $\ldots, k-1, k, k+1$, is such that $t_k = kT$. The state equation and observation equation of the above state-space representation of the discrete-time dynamic process are

$$\begin{bmatrix}
    x_1(k+1) \\
    x_2(k+1) \\
    x_3(k+1) \\
    \vdots \\
    x_n(k+1)
\end{bmatrix} = \begin{bmatrix}
    a_1 & a_2 & a_{n-1} & a_n \\
    1 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    \vdots & \vdots & \vdots & \vdots \\
    0 & 0 & 1 & 0
\end{bmatrix} \begin{bmatrix}
    x_1(k) \\
    x_2(k) \\
    x_3(k) \\
    \vdots \\
    x_n(k)
\end{bmatrix} + \begin{bmatrix}
    0 \\
    0 \\
    0 \\
    \vdots \\
    0
\end{bmatrix} u(k) \quad (4)$$
\[ y(k) = [b_1 + a_1b_0 \ b_2 + a_2b_0 \ \cdots \ b_n + a_nb_0] \]
\[ \times \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix} + [b_0]u(k). \]  
\hspace{5cm} (5)

As an example, consider a startup operation using equation (1) with \( c = 10, M = 30 \) and equation (2) with \( a = 0.78, b = 0.22 \). A model with these parameters gives the discrete-time first-order response curve shown in Figure 2. We can observe that the output response has a transient period that starts at time 10 with a level of zero and ends approximately at time 32 with a new level of 30.

3. REPRESENTING DISTURBANCES USING THE ARIMA PROCESS

In practice, a disturbance is typically associated with the output response \( y_k \). This disturbance is represented here using a nonstationary stochastic ARIMA process. The ARIMA process is a widely used mathematical model that is a special case of the linear filter of white noise. As with the process dynamics model, the ARIMA model can be represented in several forms. Here we will consider both the rational polynomial form which is often used in quality engineering and the state-space form which will be used in the simulation modeling. (The interested reader may refer to Box et al. [23] for an introduction to state-space equations for ARIMA models.)

The ARIMA process of order \( (p, d, q) \) can be represented as

\[ \Phi_p(B)V^d y_k = \Theta_q(B)e_k \]  
\hspace{5cm} (6)

where \( B \) is the backward shift operator defined by \( B^m y_k = y_{k-m} \), \( V \) is the difference operator defined by \( V = 1 - B \), \( \Phi_p \) is the autoregressive operator defined by \( \Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \), \( \Theta_q \) is the moving average operator defined by \( \Theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \), and \( e_k \) is a white-noise process with mean zero and variance \( \sigma^2_e \) (Box et al. [23]). A discrete state-space representation of the form

\[
\begin{align*}
  x(k+1) &= \hat{A}x(k) + \hat{B}e(k) \\
  y(k) &= \hat{C}x(k) + \hat{D}e(k)
\end{align*}
\]

for the ARIMA\((p, d, q)\) process is given by the following two equations:

\[
\begin{bmatrix}
  x_1(k+1) \\
  x_2(k+1) \\
  \vdots \\
  x_r(k+1)
\end{bmatrix}
= \begin{bmatrix}
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 0 \\
  \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  \phi_r \\
  \phi_{r-1} \\
  \vdots \\
  \phi_1
\end{bmatrix}
\]

\[
\begin{bmatrix}
  x_1(k) \\
  x_2(k) \\
  x_3(k) \\
  \vdots \\
  x_r(k)
\end{bmatrix}
+ \begin{bmatrix}
  1 \\
  \psi_1 \\
  \vdots \\
  \psi_{r-1}
\end{bmatrix}e(k) 
\]  
\hspace{5cm} (7)

\[ y(k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix}
  x_1(k) \\
  x_2(k) \\
  x_3(k) \\
  \vdots \\
  x_r(k)
\end{bmatrix} + [N]e(k) \]  
\hspace{5cm} (8)

where \( r = \max(p+d, q+1) \) and \( N \) is 0 or 1 depending on whether the process is subject to an additional white-noise term (in our examples, we assume \( N = 0 \)). The \( \psi \) weights can be found from the \( \phi \) and \( \theta \) weights by equating coefficients of like powers in the expansion

\[
(\psi_0 + \psi_1 B + \cdots)(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d
\]

\[ = 1 - \theta_1 B - \cdots - \theta_q B^q \]

Equation (7) is known as a state or transition equation and equation (8) is known as an observation equation.

As an example, consider an ARIMA\((0, 1, 1)\) disturbance using \( \theta = 0.6 \) in equation (6). A model with these parameters—used in combination with the step change and first-order dynamics in Figure 2—gives the output response curve shown in Figure 3.

4. CONTROL AND MONITORING OF THE STARTUP PROCESS

In a system with transient periods, a key issue is making process adjustments in the startup phase. In integrating EPC and SPC to make adjustments in this situation, we compare the ability to activate both monitoring and control simultaneously. In other words, the monitoring tool regulates the controller to start and stop adjustments to the input. The IPC mechanisms studied here combine a PI controller with a MCEWMA control chart for the case where the output observations during the transient period of a first-order process are autocorrelated.

An important question is: What represents an improvement in control? One answer is that the benefit of the combination stems from reducing the mean squared error. By integrating basic EPC and SPC policies, it is possible to reduce the output variance in some systems beyond the capabilities of either tool individually (Vander Wiel et al. [15] and MacGregor [14]). Montgomery et al. [18] demonstrate this using a modified funnel experiment in combination with an SPC chart monitoring deviation from target. Another aspect is that the benefit is associated with having to make an adjustment less often. This may be particularly relevant for systems where there is a distinct cost associated with making each adjustment or where the adjustment is done manually.

4.1. EPC policies

In EPC, proportional-integral-derivative (PID) controllers are a standard. Several texts cover the basics of PID controller design (see e.g. Seborg et al. [11] or Dorf and Bishop [12]). Although these tools are common for making adjustments to a process, they are usually designed for steady-state behavior and do not perform effectively during the startup. Manual adjustments are often used for this reason.

A typical control objective is MMSE control, or to minimize the variance of the output deviations from a target or setpoint. One criterion in developing an adjustment policy under this objective is cost. If the cost of a process adjustment is zero, then an adjustment is made at every sampling interval. If the cost of adjustment is nonzero, then an adjustment is made only if a significant deviation occurs in relation to cost of being off target. There are many optimal adjustment policies for key underlying process representations, disturbance models, and adjustment costs for MMSE control. For example, integral control is an optimal policy for a steady-state process with a well-characterized ARIMA(0, 1, 1) disturbance and zero adjustment costs (Box et al. [20]). For the same process and disturbance model with nonzero adjustment costs, the EWMA chart with a dead band is the recommended control strategy (e.g. see MacGregor [14] and Box and Luceño [1]).

For the case of a change in the process mean, Hunter [21] uses an empirical equation to arrive at MMSE control,

$$\hat{y}_{t+1} = \hat{y}_t + \lambda_1 e_t + \lambda_2 \sum e_t + \lambda_3 V e_t$$

where $\hat{y}_{t+1}$ is the predicted value at time $t+1$; $\hat{y}_t$ is the predicted value at time $t$; $\lambda_1$, $\lambda_2$, and $\lambda_3$ weight the historical data; $e_t$ is the observed error from target at time $t$; and $V$ is the first-difference operator. This constitutes the PID control equation with one quantity that is proportional to $e_t$, a second that is a function of the sum (integral) of $e_t$, and a third that is a function of the difference (derivative) $e_t$. In control engineering, $\lambda_1$, $\lambda_2$, and $\lambda_3$ are often referred to as $k_p$, $k_i$, and $k_d$ constants. We will apply equation (9) in the IPC model discussed in Section 5.

As an illustration, we consider a first-order dynamic system represented by equation (2) with $a = \delta$ and $b = K(1 - \delta)$ and an IMA(1, 1) disturbance. Using equation (6), the nonstationary IMA(1, 1) disturbance process is given by

$$(1 - B)\gamma_k = (1 - \theta B)\varepsilon_k$$
It is well-known that the PI controller minimizes the variance of output deviation from target for such a stochastic system setup. From equation (9), the corresponding weights for $\lambda_1$ and $\lambda_2$ ($\lambda_3$ is zero) are,

$$\lambda_1 = \frac{(1 - \theta)\delta}{K(1 - \delta)}$$

and

$$\lambda_2 = \frac{1 - \theta}{K}$$

where $\delta = e^{-\tau/\tau}$ with $\tau$ as the time constant of the continuous first-order dynamic system representation and $K$ is the long-term gain.

4.2. SPC policies

Montgomery and Mastrangelo [3] introduce a control chart methodology for application to autocorrelated data. The MCEWMA is a procedure that approximates the true underlying time series model with the familiar EWMA. The EWMA is typically used with independent data for either subgroup or individual measurements, but can be applied to correlated data if some modifications are made to the procedure. The EWMA for independent data is defined as

$$Z_t = \lambda Y_t + (1 - \lambda)Z_{t-1} \quad (11)$$

By replacing $\lambda$ with $1 - \theta$ and rearranging, the EWMA is similar to the IMA$(1, 1)$ model (Box et al. [20]). The EWMA with $\lambda = 1 - \theta$ is the optimal one-step-ahead forecast for the IMA$(1, 1)$ process.

The MCEWMA uses the $Z_t$ of equation (11) as the center-line for period $t + 1$. The upper and lower limits are given by $Z_t \pm L\sigma_t$ where $L$ can be a prescribed constant and $\sigma_t$ is the standard deviation of the single-period forecast errors. To test for statistical control, the output observation is compared to these control limits. The value of $\sigma_t$ can be estimated by several methods. We use the smoothed variance approach:

$$\hat{\sigma}_t^2 = \alpha e_t^2 + (1 - \alpha)\hat{\sigma}_{t-1}^2$$

where $e_t$ is the one step-ahead prediction error and $\alpha$ is a smoothing constant similar in function to $\lambda$ in the EMWA. Note that in using the smoothed variance method, the estimate of $\sigma_t$ is updated each period, which in turn, results in the varying control limits. In essence, this approach bases the estimate of $\sigma_t$ on recent process history. The initial values for $Z_t$ and $\hat{\sigma}_t^2$ are the mean and the variance of the noise model, respectively.

5. SIMULATION MODEL OF THE STARTUP SYSTEM

We developed a simulation model using SIMULINK [22] to represent the IPC policies executed on the first-order dynamic system with ARIMA disturbance. A schematic of the model is shown in Figure 4. In this section, we will discuss the operation and block components of this model. It will be used in Section 6 to experiment with the IPC policies.

Selecting any of the simulation model blocks reveals an input environment for the model parameters. Primarily, the state-space representation (i.e. the $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ matrices) of the ARIMA filter and the first-order process must be provided as input.
Table 1. Model components and simulation parameters for simulation experiments

<table>
<thead>
<tr>
<th>Model components</th>
<th>Simulation settings and parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Step function, $u(k)$, at $c = 10$ with $M = 10, 20, 30$</td>
</tr>
<tr>
<td>Dynamic process</td>
<td>First-Order with $\delta = 0.78$ and $K = 1$</td>
</tr>
<tr>
<td>Disturbance model</td>
<td>ARIMA(0, 1, 1) with $\theta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ and $\sigma^2 = 1$</td>
</tr>
<tr>
<td>Feedback controller</td>
<td>PI controller with $\lambda_1 = 2$ and $\lambda_2 = 0.5$</td>
</tr>
<tr>
<td>Monitoring scheme</td>
<td>MCEWMA with $\lambda = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ and $\alpha = 0.05$ and $L = 3$</td>
</tr>
</tbody>
</table>

The white noise block provides random shocks from a normal distribution with any value for the mean and variance. Replications of the simulation are achieved by specifying a new random-number stream in this block. The ARIMA filter block provides the discrete state-space form of the linear filter process.

The step input block provides a shift change in the input at a specified time. The first order process block represents the dynamics of the system. The noise process and the plant process are combined using the sum block to model the overall process. The statistical process control computation block contains the code for the MCEWMA algorithm.

6. SIMULATION EXPERIMENTS

In these simulation experiments, we consider three control strategies, EPC, SPC, and IPC, for a system consisting of a first-order dynamic process and an ARIMA(0, 1, 1) disturbance process. We explore a range of values for three parameters: the step size, $M$; the coefficient of the noise model, $\theta$; and the MCEWMA parameter, $\lambda$. The step sizes, $M = 10, 20, 30$, are used to represent small, moderate, and large input changes. We varied both $\lambda$ and $\theta$ from 0.1 to 0.9 to assess their effect on IPC. The complete set of model components and simulation parameters are summarized in Table 1.

We employ a model that represents the first-order dynamic process. By equations (4) and (5), the matrices for the state-space representation of this process are

$$
A = \begin{bmatrix} 0.78 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0.22 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \end{bmatrix}
$$

In each case, the step change is made at time 10 as given by equation (1). The new target level that the first-order process will eventually reach is 10, 20, and 30 respectively.

The ARIMA(0, 1, 1) process representing the disturbance to the system is given by equation (10) with $\sigma^2 = 1$. By equations (7) and (8), the matrices for the state-space representation of the disturbance are

$$
\tilde{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 1 \\ 1 - \theta \end{bmatrix}, \\
\tilde{C} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} 0 \end{bmatrix}
$$

The state-space matrices are then used in the simulation model discussed in Section 5. The results from this modeling and experimentation are discussed below.

6.1. Simulation results and discussion

We use three metrics to compare the IPC, EPC, and SPC control policies: (1) SSE; (2) the number of adjustments; and (3) the magnitude of the adjustments. Our analysis based on these metrics is discussed below.

6.1.1. Sum of squared errors. The SSE is defined as the sum of squared deviations between the target and the observed value; the final value is then averaged over all simulation runs. The EPC policy results in smaller overall squared error when compared to the SPC and IPC policies. This result is true for any system because EPC policies are designed to minimize the sum of squared errors. If this is the only criterion, EPC will always give the best performance (e.g. see MacGregor [14] or Box and Lucas [3]).

Figure 5 shows the results for a step size of 30, and we see that there is little difference in the behavior of the IPC and SPC policies. For IPC the SSE range...
is about 1800–2600 while for SPC the SSE range is about 2400–3300. Empirically, we found that the behavior shown in Figure 5 between EPC, IPC, and SPC held true over the step sizes of 10, 20, and 30 that were considered in this study.

The SSE is not only lower for EPC but it is also very flat over a range of \( \theta \). Clearly, for the cases of IPC and SPC we can say that the noise model does have an impact. If the process is highly non-stationary, then using a large value of \( \theta \) will reduce the SSE.

6.1.2. Number of adjustments. While the EPC policy has small error rates, the other two policies outperform it in terms of the average number of adjustments made to the process and the average magnitude of adjustments. Again, this relationship holds across step sizes. Figure 6 shows the number of adjustments for a step size of 30. Note that by design the EPC policy makes an adjustment in every period. The IPC policy makes an adjustment only when the observation is outside the limits of the MCEWMA. The SPC policy does not make any adjustment; however, in the figure we show the number of alarms that would have been signaled to indicate that an adjustment should be made. The downward sloping behavior on the IPC and SPC policies is expected because the MCEWMA with small \( \lambda \) values is slower to react to the input change than larger values. Consequently, there are more alarms on the monitoring chart and hence more adjustments for larger \( \lambda \) values. However, for the small step size of 10, the number of average number of adjustments is very small (around 2.0) because the control chart often does not detect the small process change.

6.1.3. Magnitude of adjustments. While the number of adjustments is small for the IPC policy, so is the size of the adjustments. This fact is illustrated in Figure 7 which shows the average of the cumulative absolute value of the magnitude of the adjustments. In other words, for a single simulation run, the absolute value of the adjustment is summed over the 50 time periods. This value is then averaged over all simulation runs. Note that in this figure, the size of adjustment for the SPC policy is zero because it is a monitoring policy and adjustment is not taking place.

7. CONCLUSIONS

Dynamic systems often have a transient period that occurs after an operating shift in the process. We demonstrated this effect with an example from polymer processing where a change in color of a plastic...
Figure 6. The number of adjustments for the EPC policy, the number of alarms and adjustments for the IPC policy, and the number of alarms (only) for SPC policy, respectively. The input change has a step size of 30.

Figure 7. The average size of adjustment for the EPC, IPC, and SPC policies, respectively. The input change has a step size of 30.
extrudate was desired. Other examples of transient periods include production startup, shutdown, and grade changes. There is a significant need for better monitoring and control of the transient period in dynamic systems. For instance, manufacturers in the paper, polymer, and pharmaceutical industries report that a large portion of their out-of-specification products is attributable to this period.

In this paper, we have extended our investigation into using an integrated approach that combines EPC and SPC for a noisy dynamic system. We have shown how to model and analyze this system over a range of practical conditions. Specifically, we analyzed the performance of the integrated design for transient period operations with a first-order dynamic process and a nonstationary ARIMA disturbance model. This type of model closely represents the industry systems where transient periods present a concern. Our work resulted in a set of response surfaces that characterize the performance of the integrated design. We also compared these results to the case where the SPC and EPC policies are applied separately.

This work confirms our earlier findings as well as identifies several process improvement opportunities. In general, we determined that the EPC approach performs best in terms of minimizing error, but that we can reduce the number and magnitude of adjustments using the integrated monitoring and control approach. There is a modest improvement in the IPC policy, that is, a reduction in the variability as measured by the SSE, if the disturbance process is known to a certain degree (through identification procedures on observational data). We also found that the SSE is smaller for EPC; however, the number of adjustments and size of adjustments is smaller for IPC. This implies that for non-zero adjustment costs, IPC is a better alternative. Furthermore, for a small input change, our results parallel the SPC research that indicates that a small shift is difficult to detect.

Overall, we believe that the integration of process monitoring and control is a promising alternative.

Our future plans involve applying these policies to startup operations that may not be perfectly described by the first-order and nonstationary models in order to develop an understanding of the robustness of this approach. An economic-statistical model that addresses the level of costs that would be justified using IPC instead of EPC or SPC alone is another exciting opportunity.

ACKNOWLEDGEMENTS
The authors would like to acknowledge the advice and input of an anonymous referee.

Copyright © 2001 John Wiley & Sons, Ltd.
Authors’ biographies:

Harriet Black Nembhard is an Assistant Professor of Industrial Engineering at the University of Wisconsin-Madison. She has previously held manufacturing positions with Dow Chemical, General Mills, and Pepsi-Cola. Her PhD degree is in Industrial and Operations Engineering from The University of Michigan. She is a member of ASQ, IIE, and INFORMS.

Christina M. Mastrangelo is an Associate Professor of Systems Engineering at the University of Virginia. Her PhD degree is in Industrial Engineering from Arizona State University. She is a member of ASQ, IIE, and INFORMS.

Ming Shu Kao is a PhD candidate in Industrial Engineering at the University of Wisconsin-Madison. He is a member of ASQ.