Scientia Iranica E (2022) 29(4), 2210–2229

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
http://scientiairanica.sharif.edu

Monitoring process mean using a second-order filter: Signal and system approach

O. Ahmadi* and H. Shahriari

Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran.

Received 4 October 2019; received in revised form 6 August 2020; accepted 14 September 2020

KEYWORDS
Control charts;
Linear filters;
LTI systems;
Markov chain;
Statistical process control.

Abstract. In statistical process control, one objective is to control the stability of a process. A process is stable when its mean is in control and variance bounded. Different control charts were introduced for monitoring the mean and variance of a process by plotting suitable test statistics on the chart. In this research, design of a system which converts the sample mean into test statistics was proposed. The second-order filter, a special class of the Linear Time Invariant (LTI) systems, was used to design the converting system. It was shown that design of a low-pass filter was better for detecting a level (mean) change in the process. Markov chain approach was also employed to construct an appropriate control chart and to estimate its control limits. Simulated data under normality assumption for different scenarios were used to compare the proposed control chart with Shewhart and Exponentially Weighted Moving Average charts by means of Average Run Length (ARL) and Percentage of False Signals (PFS) criteria. Existing data from the Central Bank of Iran was also applied to evaluate the suggested method. The signal-to-noise ratio was used to assess the performance of this method at different stages. Results indicate that the proposed method detects shifts more rapidly.

© 2022 Sharif University of Technology. All rights reserved.

1. Introduction

Statistical process control techniques are tools used to reduce process dispersion and to improve its stability. In statistical process control, it is usually assumed that the output of the process is normally distributed with in-control mean μ_0 and in-control standard deviation σ_0. Control charts are one of the quality control tools which are frequently used to investigate the process stability. In the simplest case where the quality characteristic is normally distributed, it is of interest to investigate if the process mean and standard deviation remain at in-control levels. According to Montgomery [1], one may use Shewhart control charts to monitor the process parameters. Construction and implementation of Shewhart charts require introduction of some test statistics for monitoring the stability of the process. Then, the probability distribution of the test statistics under the assumption of process being in control must be determined. Using this probability distribution, the values of test statistics which may be generated when the process is in control are determined. It is expected that when the process is in control, test statistics fall inside a specific interval. The upper and lower bounds of this interval are called Upper Control Limit (UCL) and Lower Control Limit (LCL), respectively. Determination of the control limits requires a desired value for Average Run Length

*. *Corresponding author.*
E-mail addresses: oroarm@yahoo.com (O. Ahmadi); hshahriari@kntu.ac.ir (H. Shahriari)

doi: 10.24200/sci.2020.54204.3647

(ARL). ARL is the mean number of samples needed to be taken from the process until the test statistics exceed its control limits. Control limits are defined to obtain a predetermined in-control ARL, such as 370. In this research, it is assumed that the process standard deviation remains in control when monitoring the process mean. One of the most popular control charts for monitoring the process mean is the $\bar{X}$ chart. For implementing this chart, samples of size n are taken from the process and the sample mean $\bar{X}$ is computed for each one of them. As long as $\bar{X} \in (\mu_0 \pm 3\sigma_0/\sqrt{n})$, the process mean is judged to be in-control. According to Montgomery [1], Shewhart $\bar{X}$ chart is suitable to detect large shifts and is less sensitive to small shifts. Some authors including Jiang [2], Wu and Spedding [3], Zhang [4], Chang and Aw [5], and Harris and Ross [6] proposed other charts for monitoring the process mean which are sensitive to small and moderate changes. Two of such charts are Exponentially Weighted Moving Average (EWMA) and Cumulative Sum (CUSUM). In these two charts, other test statistics rather sample mean is plotted on the chart. It was shown that EWMA control chart was highly robust against non-normality of the process distribution. Thus, this chart has been adopted to construct the proposed control chart. Saif [7] proposed a framework to integrate the control charts with automation process control.

Monitoring the process mean may be categorized into two approaches. In the first approach, monitoring procedure is directly based on plotting the sample mean on the appropriate chart. The second approach consists of monitoring some appropriate functions of the sample mean. These functions must be chosen so that the corresponding charts would find small to moderate changes rapidly.

These two approaches could be considered as a unified approach. In this unified approach, one may design a function which is called system. This system takes $\bar{X}_t$, $t = 1, 2, \cdots s$ as input and transforms them into output shown by Y_t which is the charting test statistics. The system along with its input and output is shown in Figure 1.

Using this unified approach, one may study the behavior of monitoring procedure more precisely. In other words, selecting appropriate test statistics to monitor the process will be reduced to selecting an appropriate system. Process monitoring scheme is usually designed only based on statistical considerations. In this strategy, the behavior of processes in the time domain is usually assessed. However, some important features of the processes may not be discovered if processes are considered only in the time domain. Accordingly, any change in the process mean may not be detected effectively.

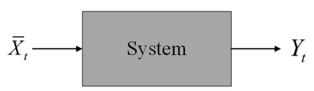


Figure 1. The unified controlling system.

On the other hand, considering statistical process monitoring from the system viewpoint enables practitioners to better understand the processes. The reason is that using the system approach one could study the process in the frequency domain. This domain contains information about the process which may not be evident in the time domain. Thus, it is possible to monitor complicated processes more precisely using system approach. Choosing an appropriate system, shown in Figure 1, is very important in constructing a control scheme. Designing such a system requires that the process be studied in the frequency domain, as well. In this research, the system approach is followed to construct a unified scheme to monitor processes.

Various authors, including Rabyk and Schmid [8] Shokrizadeh et al. [9], Yang and Arnold [10], Chen et al. [11], Apley and Shi [12] and Lu and Reynolds [13], proposed different methods for monitoring process mean. However, signal and system approach was not explicitly used to define a monitoring scheme. Previous studies have not tried the hybrid method discussed in this study which adopts the second-order filter and the Markov chain approach as a means of detecting changes in the process mean. In this context, the signal and system approach is applied to define control scheme with suitable properties. The proposed monitoring procedure must be designed so that the resulting chart can have desirable properties.

The rest of this paper is organized as follows. In Section 2, some preliminaries about system theory are provided. In Section 3, LTI systems are introduced. The proposed monitoring procedure is presented in Section 4. In the proposed method, the test statistics is obtained by using a suitable LTI system. ARL of the proposed procedure is obtained using Markov chain approach in Section 5. In Section 6, the performance of the proposed method is evaluated via simulation. Results of monitoring a real stochastic process using the proposed control scheme are provided in Section 7. Finally, discussions and conclusion are made. Symbols and notations used in this research are provided in Nomenclature.

2. Theory and system design

The test statistics for EWMA chart is given in Eq. (1):

$$Z_t = \lambda \bar{X}_t + (1-\lambda) Z_{t-1}; \qquad Z_0 = \mu_0; \quad t = 1, 2, \cdots. \tag{1}$$

The upper and lower control limits for this chart are as follows:

$$UCL = \mu_0 + L\sigma_0\sqrt{\frac{\lambda}{(2-\lambda)n}},$$

$$LCL = \mu_0 - L\sigma_0\sqrt{\frac{\lambda}{(2-\lambda)n}}, \tag{2}$$

where the charting parameters $0 \leq \lambda \leq 1$ and $L > 0$ are designed by the analyst for desired properties such as having a predetermined in-control ARL. As long as $Z_t \in (LCL, UCL)$, the process is considered to be in control.

According to Zhang et al. [14], CUSUM chart is based on plotting the following statistics on the chart:

$$S_0^+ = 0; \qquad S_t^+ = \max\left\{0, S_{t-1}^+ + \frac{\sqrt{n}\left(\bar{X}_t - \mu_0\right)}{\sigma_0}\right\}; \quad t = 1, 2, \cdots, \tag{3}$$

$$S_0^- = 0; \qquad S_t^- = \min\left\{0, S_{t-1}^- + \frac{\sqrt{n}\left(\bar{X}_t - \mu_0\right)}{\sigma_0}\right\}; \quad t = 1, 2, \cdots. \tag{4}$$

For a given $h > 0$, the process mean is said to be in control at time t if $S_t^+ < h$ and $S_t^- > -h$. The parameter $h > 0$ is the charting parameter chosen to have a predetermined in-control ARL.

Considering EWMA and CUSUM control charts, one may conclude that it is possible to use a system that transforms $\bar{X}_t$ values to generate some other test statistics. These test statistics then may be used to monitor the process mean rather than $\bar{X}_t$ values.

A class of systems is called the Linear Time Invariant (LTI) which is important and suitable for defining test statistics, Y_t. In this research, input and output of a LTI system at time t are shown by $x[t]$ and $y[t]$, respectively. In system theory, $x[t]$ and $y[t]$ are called input and output signals, respectively. For the sake of simplicity, the relation between $x[t]$ and $y[t]$ is shown by $x[t] \to y[t]$. The LTI systems, as stated by Oppenheim et al. [15], are the systems that transform input to output using two important properties. These properties include linearity and time invariance. For defining linearity, suppose that $x_1[t]$ and $x_2[t]$ are two arbitrary input signals of a system with corresponding output signals $y_1[t]$ and $y_2[t]$, respectively, i.e., $x_1[t] \to y_1[t]$ and $x_2[t] \to y_2[t]$. The system is said to be linear if for every a_1 and a_2, $a_1x_1[t] + a_2x_2[t] \to a_1y_1[t] + a_2y_2[t]$. A system intuitively is said to be time invariant if its behavior does not change with time. More formally, $x[t] \to y[t]$ represents a system; this system is time invariant if for every $t_0 \in \mathbb{Z}$, $x[t-t_0] \to y[t-t_0]$. In this study, the LTI systems are used to define suitable test statistics for monitoring the process mean.

3. Statistical monitoring procedure: The LTI system

As mentioned in Section 2, LTI systems have two important properties: linearity and time invariance. These properties make it easy to obtain a closed form for output signal of a system in terms of input signal. Outputs of a system may be called system responses. To show how to obtain LTI system response, it is required to define two important input signals. These two input signals are unit impulse, $\delta[t]$, and unit step, $u[t]$, which are defined as follows:

$$\delta[t] = \begin{cases} 1 & t = 0 \\ 0 & t \in \mathbb{Z} - \{0\} \end{cases} \tag{5}$$

$$u[t] = \begin{cases} 1 & t \in \mathbb{Z}^+ \cup \{0\} \\ 0 & t \in \mathbb{Z}^- \end{cases} \tag{6}$$

Let $h[t]$; $t \in \mathbb{Z}$ be the system response to the unit impulse, i.e., $\delta[t] \to h[t]$. In the literature on signals and systems, $h[t]$ is called impulse response. According to Oppenheim et al. [15], the output of a LTI system to any input signal $x[t]$ can be expressed as follows:

$$y[t] = \sum_{k=-\infty}^{+\infty} h[k]x[t-k] = h[t]^*x[t]; \qquad t, k \in \mathbb{Z}, \tag{7}$$

where $y[t]$ is the system response to $x[t]$ and (*) is the convolution operator.

Based on Eq. (7), having $h[t]$, one can find a system response to any input signal. Eq. (7) represents the output of a system in the time domain. In Figure 2, the relation between input and output of a LTI system in the time domain is shown.

It is possible to represent a system in the frequency domain, as well. To this end, one may take Fourier transformation of both sides of Eq. (7). It is shown in Appendix A that taking the Fourier Transformation of Eq. (7), one may conclude that:

$$Y(e^{jw}) = H(e^{jw})X(e^{jw}), \tag{8}$$

where $j = \sqrt{-1}$, w is the frequency variable, and $Y(e^{jw})$, $H(e^{jw})$, and $X(e^{jw})$ are the Fourier transformation of $y[t]$, $h[t]$, and $x[t]$, respectively, which are shown below:

$$Y(e^{jw}) = \sum_{t=-\infty}^{+\infty} y[t]e^{-jwt}, \tag{9}$$

$$H(e^{jw}) = \sum_{t=-\infty}^{+\infty} h[t]e^{-jwt}, \tag{10}$$

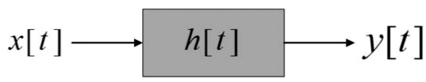


Figure 2. Time representation of an LTI system.

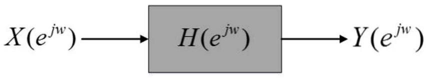


Figure 3. Frequency representation of an LTI system.

$$X(e^{jw}) = \sum_{t=-\infty}^{+\infty} x[t]e^{-jwt}. \tag{11}$$

The function $H(e^{jw})$ is usually called the frequency response of system. According to Eq. (8), one may multiply $H(e^{jw})$ by $X(e^{jw})$ to obtain $Y(e^{jw})$. Then, taking inverse Fourier transformation of $Y(e^{jw})$ yields $y[t]$. Representation of an LTI system in the frequency domain is illustrated in Figure 3.

LTI systems are usually called filters. In designing a filter, it is very important to define $h[t]$ or its Fourier transformation $H(e^{jw})$ properly. Generally, three types of filters, including low-pass, high-pass, and medium-pass filters could be defined. Low-pass filters only permit low frequency pass through the filter. In high-pass filters, only high frequencies pass the filters and low frequencies are vanished. Median-pass filters only send moderate frequencies out.

One type of frequently-used LTI systems or filters are those represented by linear difference equations with constant coefficients. In a more formal term, let $x[t]$ and $y[t]$ be input and output signals of a system which are related to each other according to Eq. (12):

$$\sum_{k=0}^{M} b_k y[t-k] = \sum_{k=0}^{N} a_k x[t-k]. \tag{12}$$

Eq. (12) is a linear difference equation with constant coefficients. Solving Eq. (12) needs some auxiliary conditions. It is rational to assume that if $x[t] = 0$ for $t < \tau$, then $y[t] = 0$ for $t < \tau$. This auxiliary condition is called initial rest. Based on Oppenheim et al. [15], any system that is represented by Eq. (12) with initial rest condition specifies a LTI system, which is a linear filter in turn. Designing linear filters has been studied by various authors, especially in the field of communication systems. Wang et al. [16] proposed linear feedback control loops to detect sensor faults. Chen et al. [17] studied the problem of detecting faults in linear stochastic dynamic systems. Zuo et al. [18] designed a linear filter to estimate missing measurements. Liu and Shi [19] studied optimal linear filtering for correlated data. Liu et al. [20] investigated the optimal design of linear filters. Eijnden et al. [21] proposed a hybrid low-pass filter to control nonlinear motion.

Two frequently used filters include first-order and second-order filters. In this paper, the second-order filter is used to monitor the process mean. The first- and second-order filters are reviewed briefly in Subsection 3.1.

3.1. The first and the second order filters

The simplest linear filter used to design a system is a first-order system with initial rest conditions:

$$y[t] = \phi_1 y[t-1] + x[t]; \qquad t \in \mathbb{Z}, \tag{13}$$

where ϕ_1 is a constant. The block diagram of the first-order system is shown in Figure 4.

In Figure 4, D represents the difference operator. From Eq. (13), it is evident that for a given input signal $x[t]$, the output signal $y[t]$ depends on the value of $y[t-1]$. Accordingly, the corresponding system is called a first-order filter.

In designing filters, usually first-order systems are connected serially or in parallel. Deciding which types of connection to use depends on the application of the resulting filter.

Another frequently used filter is the second-order filter that is represented by Eq. (14):

$$y[t] = \phi_1 y[t-1] + \phi_2 y[t-2] + x[t]; \qquad t \in \mathbb{Z}, \tag{14}$$

where ϕ_1 and ϕ_2 are some constants. Block diagram of the second-order filter is shown in Figure 5.

According to Eq. (14), for a given input signal $x[t]$, the output signal $y[t]$ depends on the values of $y[t-1]$ and $y[t-2]$. Because of this dependency, the system which is represented by Eq. (14) is called the second-order filter.

As stated before, the behavior of each LTI system including the first-order and second-order systems in the time domain is determined by impulse response, $h[t]$. In the frequency domain, the behavior of a filter is modeled by its frequency response, $H(e^{jw})$. In this study, the authors attempted to use the second-order filter to design a unified control system, as shown in Figure 1, to transform $\bar{X}_t$ into Y_t. Designing this

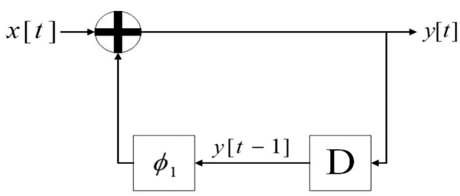


Figure 4. Block diagram of first-order system.

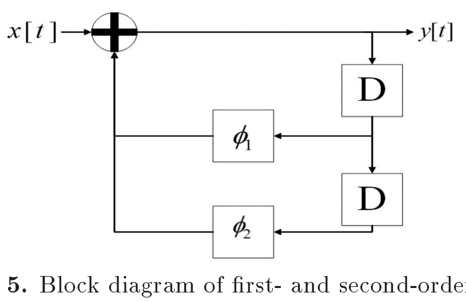


Figure 5. Block diagram of first- and second-order systems.

second-order filter requires determination of the values of ϕ_1 and ϕ_2 in Eq. (14). These coefficients should be determined so that the resulting test statistics, Y_t, becomes sensitive to a certain type of change in the process mean. In this study, it is assumed that the process mean is μ_0 by time t_0. At time t_0, the process mean changes to an out-of-control level shown by $\mu_0+\gamma\sigma_0$, where γ is the magnitude of shift in process mean in terms of process standard deviation. The unified controlling system should be designed in order to detect this change rapidly. In other words, the values of ϕ_1 and ϕ_2 in Eq. (14) must be determined so that the change in process mean is magnified in the sequence of Y_t's. Note that the mean and variance of the filter output, Y_t, will not be, in general, equal to those of $\bar{X}_t$. One may define Signal to Noise Ratio (SNR) of $\bar{X}_t$ and Y_t by $SNR_{\bar{X}_t} = \left(\frac{E^2[\bar{X}_t]}{\mathrm{var}(\bar{X}_t)}\right)$ and $SNR_{Y_t} = \left(\frac{E^2[Y_t]}{\mathrm{var}(Y_t)}\right)$, respectively. If $SNR_{Y_t} > SNR_{\bar{X}_t}$, changes in process mean will be unfolded more prominently using Y_t. Then, out-of-control conditions will be detected more rapidly. Thus, it is better to use a filter that increases SNR_{Y_t} compared to $SNR_{\bar{X}_t}$. In this study, a second-order low-pass filter is proposed to transform $\bar{X}_t$ into Y_t. Use of low-pass filters is logical since these filters remove high frequency components and smooth out input signals. Using the ratio $\frac{SNR_{Y_t}}{SNR_{\bar{X}_t}}$, one may compare levels of dispersion and mean before and after filtering. Thus, the ability of filter to remove noises may be studied properly. It is important to study the behavior of a second-order system in time and frequency domains for designing the filter properly. In the next subsection, the behavior of the second-order system in terms of $h[t]$ and $H(e^{jw})$ is investigated. This investigation illustrates designing of an appropriate unified control system.

3.2. Behavior of the second order filters

To study the behavior of a second-order system, its impulse response, $h[t]$, and frequency response, $H(e^{jw})$, must be determined. It is better to specify $H(e^{jw})$ first. According to Oppenheim et al. [15], in order to have a frequency response, the system must be stable. Intuitively speaking, a system is stable if its responses to bounded inputs remain finite, i.e., $|x[t]| < \infty \Rightarrow |y[t]| < \infty$.

A second-order system shown by Eq. (14) is stable if the following conditions are met:

$$\phi_1 + \phi_2 < 1; \qquad \phi_2 - \phi_1 < 1; \qquad |\phi_2| < 1. \tag{15}$$

Thus, if a second-order filter is used to transform $\bar{X}_t$ into Y_t as long as conditions of Eq. (15) are satisfied and $|\bar{X}_t| < \infty$, the output signal Y_t will be bounded. Using Markov inequality, one may conclude that the sufficient condition for $|\bar{X}_t| < \infty$ is that $E[|\bar{X}_t|] < \infty$. On the other hand, $E[|\bar{X}_t|] \leq E[|X|]$. Thus, in the stability region of a second-order filter, if $E[|X|] < \infty$, then $|Y_t| < \infty$.

Suppose that a second-order system defined by Eq. (14) is stable. For determining $H(e^{jw})$, one may take Fourier transformation of both sides of Eq. (14) which results in the following:

$$Y(e^{jw}) = \phi_1 e^{-jw} Y(e^{jw}) + \phi_2 e^{-2jw} Y(e^{jw}) + X(e^{jw}). \tag{16}$$

Rearranging Eq. (16) yields Eq. (17):

$$\frac{Y(e^{jw})}{X(e^{jw})} = \frac{1}{1 - \phi_1 e^{-jw} - \phi_2 e^{-2jw}}. \tag{17}$$

According to Eq. (8), Eq. (17) may be written as:

$$H(e^{jw}) = \frac{1}{1 - \phi_1 e^{-jw} - \phi_2 e^{-2jw}}. \tag{18}$$

Eq. (18) is called the frequency response of system. Taking the inverse Fourier transformation of both sides of Eq. (18), one may obtain the impulse response of system. The impulse response of the second-order filter is obtained in Appendix B.

In Appendix B, it is shown that the behavior of $h[t]$ depends on the sign of $\phi_1^2 + 4\phi_2$. Thus, two cases are possible for computing $h[t]$ which are explained in the following.

Case I: $\phi_1^2 + 4\phi_2 > 0$. If $\phi_1^2 + 4\phi_2 > 0$, the impulse response of the system is equal to:

$$h[t] = \frac{x_2}{x_2 - x_1}\left(\frac{1}{x_1}\right)^t u[t] + \frac{x_1}{x_1 - x_2}\left(\frac{1}{x_2}\right)^t u[t]; \quad t \in \mathbb{Z}, \tag{19}$$

where:

$$x_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2},$$

$$x_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}, \tag{20}$$

where $u[t]$ is a unit step function defined in Eq. (6). The stability conditions of the system require $|x_1| > 1$ and $|x_2| > 1$. Conversely, given x_1 and x_2, one could compute ϕ_1 and ϕ_2 as follows:

$$\phi_1 = \frac{1}{x_1} + \frac{1}{x_2},$$

$$\phi_2 = -\frac{1}{x_1 x_2}. \tag{21}$$

Note that x_1^{-1} and x_2^{-1} are called poles of system. If $\phi_1^2 + 4\phi_2 > 0$, the system has two real poles. The behavior of the second-order filter is determined by

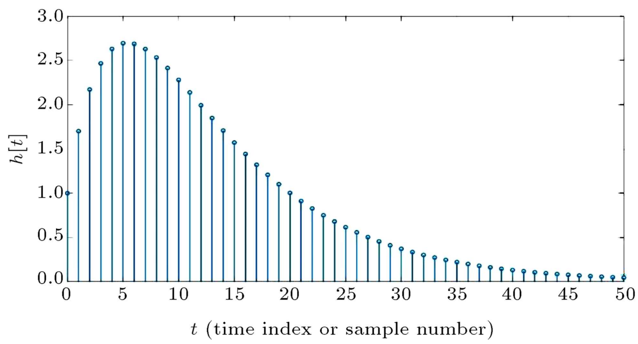


Figure 6. Impulse response with $\phi_1 = 1.7$ and $\phi_2 = -0.72$.

location of these poles. According to Oppenheim et al. [15], when poles move toward 1, the corresponding system behaves as a low-pass filter. If poles move toward -1, the system will be a high-pass filter.

It could be seen that for case $\phi_1^2 + 4\phi_2 > 0$, the impulse response of the system damps out without oscillation as $t \to +\infty$. Thus, to design a second-order system for this case, it is important to define the values of ϕ_1 and ϕ_2. For example, consider a second-order filter with poles $x_1^{-1} = 0.9$ and $x_2^{-1} = 0.8$. Using Eq. (21), the values of ϕ_1 and ϕ_2 were computed as 1.7 and -0.72, respectively. The impulse response of this system is shown in Figure 6. Note that these values for ϕ_1 and ϕ_2 are in a non-oscillatory region;

Case II: $\phi_1^2 + 4\phi_2 < 0$. In case $\phi_1^2 + 4\phi_2 < 0$, the impulse response of a system becomes:

$$h[t] = \frac{\sin((t+1)\theta)}{\sin(\theta)}\rho^t u[t]; \qquad t \in \mathbb{Z}, \tag{22}$$

where:

$$2\rho\cos\theta = \phi_1; \qquad -\rho^2 = \phi_2. \tag{23}$$

For case $\phi_1^2 + 4\phi_2 < 0$, the impulse response of the system damps with oscillation as $t \to +\infty$. Let x_1 and x_2 be the following:

$$x_1 = \frac{-\phi_1 + j\sqrt{-\phi_1^2 - 4\phi_2}}{2\phi_2},$$

$$x_2 = \frac{-\phi_1 - j\sqrt{-\phi_1^2 - 4\phi_2}}{2\phi_2}. \tag{24}$$

According to Oppenheim et al. [15], in this case, x_1^{-1} and x_2^{-1} are called poles of system. In order for the system to be stable, it is required that $|x_1| > 1$ and $|x_2| > 1$. Again, given x_1 and x_2, the values of ϕ_1 and ϕ_2 could be obtained using Eq. (21). According to Oppenheim et al. [15], if poles move toward 1, the system will be a low-pass filter. If they move toward -1, the system will be a high-pass filter. As $x_1 \to j$ and $x_2 \to -j$, the resulting system is a medium-pass filter. For example, consider a system with poles $x_1^{-1} = 0.9 + 0.1j$ and $x_2^{-1} = 0.9 - 0.1j$. Using Eqs. (21) and (23), the following values were obtained: $\phi_1 = 1.8$, $\phi_2 = -0.82$, $\rho = \sqrt{-\phi_2} = 0.91$, and $\theta = \cos^{-1}\left(\frac{\phi_1}{2\rho}\right) = 0.11$. The impulse response of this system is shown in Figure 7.

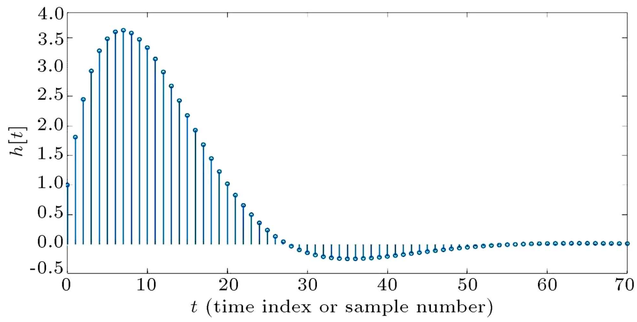


Figure 7. Impulse response with $\rho = 0.91$ and $\theta = 0.11$.

In summary, for designing a second-order system, the values of its poles must be chosen. These values are set such that the resulting filter poses some suitable features. Based on poles' values and using Eq. (21), ϕ_1 and ϕ_2 values are determined. When monitoring the process mean, a desired feature that the filter is better to have is sensitivity to step changes. Based on this desired feature, a suitable second-order filter can be designed. The designed filter could be used to transform sample means $\bar{X}_t$ into test statistics, Y_t's, as shown in Figure 1. Proper design of filter makes Y_t sensitive to step changes. Based on this guideline, a second-order filter is used to define a unified control system here. The proposed monitoring method using a second-order filter is introduced in Section 4.

4. Proposed monitoring method

4.1. Underlying logic

In this section, a unified control system is designed to monitor the process mean. More formally, suppose that when the process is in-control, the quality characteristic of interest is normally distributed with mean μ_0 and variance σ_0^2. It is assumed that when the process is out of control, the process mean jumps to a new level $\mu_0 + \gamma\sigma_0$. It is important to detect this change as soon as possible. Let t_0 be the change point of the process mean. Let $\bar{X}_t$ be the sample mean of a random sample of size n taken at time t. $\bar{X}_t$ can be represented as follows:

$$\bar{X}_t = \begin{cases} \mu_0 + \frac{\sigma_0}{\sqrt{n}}\varepsilon_t & t = 0, 1, \cdots, t_0 - 1 \\ \mu_0 + \gamma\sigma_0 + \frac{\sigma_0}{\sqrt{n}}\varepsilon_t & t = t_0, t_0 + 1, \cdots \end{cases} \tag{25}$$

where ε_t's are Normally Independently Distributed

Table 1. Response of system to level shift.

	$t < t_0$	$t \geq t_0$
$\phi_1^2 + 4\phi_2 > 0$	$\lambda[t] = \frac{\mu_0}{1-\phi_1-\phi_2}$	$\lambda[t] = (\mu_0 + \gamma\sigma_0)\left\{A_1 \frac{x_1 - \left(\frac{1}{x_1}\right)^{t-t_0}}{x_1 - 1} + A_2 \frac{x_2 - \left(\frac{1}{x_2}\right)^{t-t_0}}{x_2 - 1}\right\}$
$\phi_1^2 + 4\phi_2 < 0$	$\lambda[t] = \frac{\mu_0}{1-\phi_1-\phi_2}$	$\lambda[t] = \frac{\mu_0+\gamma\sigma_0}{\sin(\theta)} \sum_{k=0}^{t-t_0} \sin\left((k+1)\theta\right) \rho^k$

[NID] with mean 0 and variance 1, i.e., $\varepsilon_t \sim NID(0,1)$. Expected value of $\bar{X}_t$ in Eq. (25) could be represented as:

$$E\left[\bar{X}_t\right] = v[t] = \mu_0 + (\gamma\sigma_0)u[t-t_0]. \qquad (26)$$

Suppose that the sequence of $\bar{X}_t$'s are added to a stable second-order system. Thus, we have:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \bar{X}_t; \qquad t = 1, 2, \cdots;$$

$$Y_0 = Y_{-1} = \frac{\mu_0}{1-\phi_1-\phi_2}. \qquad (27)$$

Note that it is required to have the initial values for Y_0, and Y_{-1} to compute Y_t for $t > 0$. Reasonable initial values for Y_0 and Y_{-1} are the expected values of the stochastic process Y_t which is $\frac{\mu_0}{1-\phi_1-\phi_2}$. In Eq. (27), this initial value was used. According to Box et al. [22], when $\bar{X}_t$ is normally distributed, Y_t follows normal distribution.

Eq. (27) defines a second-order autoregressive process, called AR(2) in time series analysis. It is suitable to determine the effect of step change on the mean value of Y_t. Let $E[Y_t] = \lambda[t]$. Taking expectation from both sides of Eq. (27) yields the following:

$$\lambda[t] = \phi_1\lambda[t-1] + \phi_2\lambda[t-2] + v[t]. \qquad (28)$$

Eq. (28) defines an LTI system with input signal $v[t]$ and response signal $\lambda[t]$. Let $h[t]$ be the impulse response of the system. When $\phi_1^2 + 4\phi_2 > 0$, $h[t]$ is given by Eq. (19). For $\phi_1^2 + 4\phi_2 < 0$, $h[t]$ is obtained according to Eq. (22). Based on Eq. (7), the response of the system to the input signal $v[t]$ is given by:

$$\lambda[t] = h[t]^* v[t] = \sum_{k=-\infty}^{+\infty} h[k]v[t-k]. \qquad (29)$$

Eq. (29) was provided in Appendix C. Solution of Eq. (29) is summarized in Table 1.

A numerical example is here provided to assess the response of the system to the level shift.

4.2. Test evaluation of process parameters

Suppose that the quality characteristic of interest is normally distributed with in-control mean $\mu_0 = 0$ and variance $\sigma_0^2 = 1$. At time $t_0 = 15$, a positive shift of one standard deviation ($\gamma = 1$) occurs, hence $\mu_0 + \gamma\sigma_0 = 1$. A second-order filter with two poles $x_1^{-1} = 0.65$ and $x_2^{-1} = -0.15$ is considered. In this pattern of poles, x_1^{-1} is closer to 1 than x_2^{-1} to -1. Thus, the system behaves as a low-pass filter. Using Eq. (21), values of ϕ_1 and ϕ_2 are computed as 0.5 and 0.1, respectively. This filter is used to convert X_t into Y_t. In this case, $\phi_1^2 + 4\phi_2 = 0.65 > 0$. Thus, the response of the filter does not oscillate. It is expected that when $t \to \infty$, $\lambda[t] \to \frac{\mu_0+\gamma\sigma_0}{1-\phi_1-\phi_2} = 2.5$. Based on Table 1, $\lambda[t]$ is determined and shown in Figure 8(a), along with $v[t]$. From Figure 8(a), it is obvious that before the change

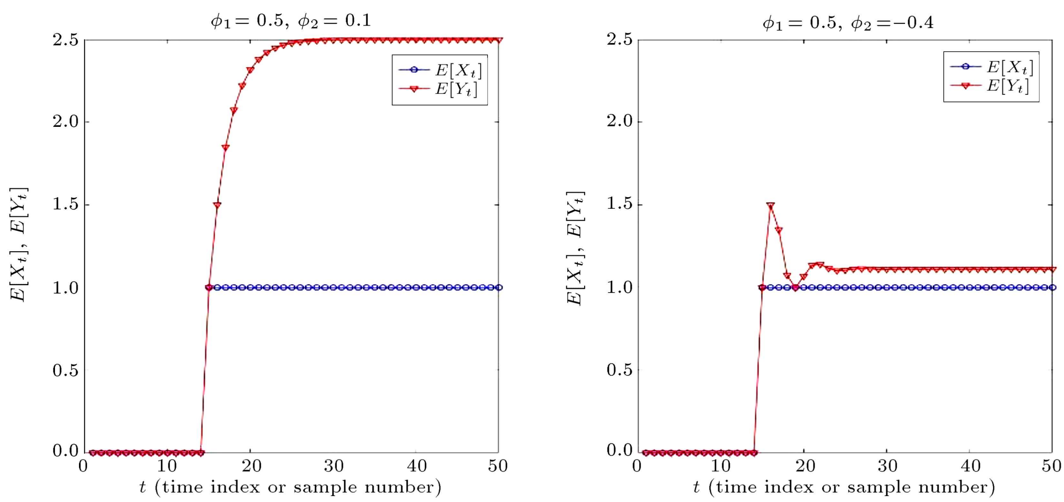


Figure 8. Expected value of X_t and Y_t: (a) $\phi_1 = 0.5$ and $\phi_2 = 0.1$ and (b) $\phi_1 = 0.5$ and $\phi_2 = -0.4$.

Table 2. Characteristics of the proposed method and Shewhart control chart.

	Test statistics	Control limits	In-control region
Proposed method	$Y_t - \frac{\mu_0}{1-\phi_1-\phi_2}$	$\pm 2.9\sigma_Y$	$\left\lvert Y_t - \frac{\mu_0}{1-\phi_1-\phi_2}\right\rvert < 2.9\sigma_Y$
Shewhart	$\bar{X}_t - \mu_0$	$\pm\frac{3\sigma_0}{\sqrt{n}}$	$\left\lvert \bar{X}_t - \mu_0\right\rvert < \frac{3\sigma_0}{\sqrt{n}}$

point, $E[X_t]$ and $E[Y_t]$ remains at their constant levels. At change point, $E[X_t]$ and $E[Y_t]$ jump to the new level 1. After change point, $E[X_t]$ remains at 1. While $E[Y_t]$ tends to 2.5. This behavior makes it easy to detect a level change in $E[X_t]$. In other words, using a proper second-order filter magnifies the level change and it would be detected sooner.

Now, consider a case in which $x_1^{-1}, x_2^{-1} = 0.25 \pm 0.58j$. These values of poles are chosen arbitrarily in stability region to have an oscillatory impulse response. Using Eq. (21), $\phi_1 = 0.5$, $\phi_2 = -0.4$, and $\phi_1^2 + 4\phi_2 = -1.35 < 0$. In case when $t \to \infty$, $\lambda[t] \to \frac{\mu_0+\gamma\sigma_0}{1-\phi_1-\phi_2} = 1.1$. According to Table 1, $\lambda[t]$ is computed and shown in Figure 8(b) along with $v[t]$.

Figure 8 reveals that use of a proper second-order filter magnifies the change in the process level. ■

4.3. Proposed control chart

Considering the example presented in Section 3.2, it is logical to use Y_t instead of only $\bar{X}_t$ for monitoring the process mean. Hence, in the proposed method, a second-order filter with proper parameters, ϕ_1 and ϕ_2, must be defined. Then, using this system, $\bar{X}_t$'s are filtered to obtain test statistics Y_t's. The centered test statistics $Y_t - \frac{\mu_0}{1-\phi_1-\phi_2}$ are then compared with control limits $\pm L\sigma_Y$ in which L is the coefficient of the proposed control chart and σ_Y is the in-control standard deviation of Y_t. Note that according to Eq. (27), the sequence of Y_t's generates a stationary AR(2) process. Therefore, var(Y_t) is the variance of an AR(2) process. According to Box et al. [22], when the process is in-control, the variance for an AR(2) process is as follows:

$$\sigma_Y^2 = \frac{(1-\phi_2)\sigma_0^2}{n(1+\phi_2)(1-\phi_1-\phi_2)(1-\phi_2+\phi_1)}. \tag{30}$$

The value of L in $\pm L\sigma_Y$ is determined so that the in-control ARL equals a pre-specified value. The approach to determining L is discussed in the next section. As an illustration of the proposed method application, the mentioned example in Section 3.2, Case I, is considered. The proposed method with $\phi_1 = 1.7$ and $\phi_2 = -0.72$ as well as the Shewhart control chart is used to detect any change in the process mean. The control limits for the two monitoring methods are set in a way that the in-control ARL becomes 370. Note that in the proposed method, the test statistics are not independent. Thus, the distribution of the run length is not geometric and the traditional method of computing ARL does not work anymore. In Section 4, a Markov chain approach is applied to compute the in-control ARL. In Table 2, characteristics of the two monitoring methods are presented.

For comparing the performance of the two methods, 50 observations were generated from the following model:

$$X_t = \begin{cases} \varepsilon_t & t < 15 \\ 1 + \varepsilon_t & t \geq 15 \end{cases} \tag{31}$$

where ε_t's were independent standard normal random variables. Note that here $n = 1$, hence $\bar{X}_t = X_t$. Figure 9 shows the sequence of $X_t - \mu_0$ along with the corresponding Shewhart control limits.

As could be seen from Figure 9, the Shewhart control chart does not issue any out-of-control signal.

Figure 10 shows the sequence of $Y_t - \mu_0/(1-\phi_1-\phi_2)$ along with the proposed control limits.

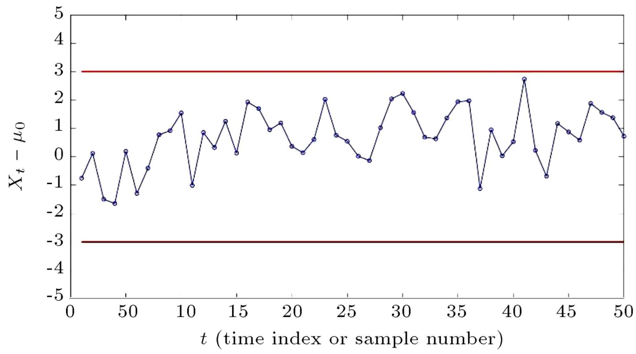


Figure 9. Shewhart control chart.

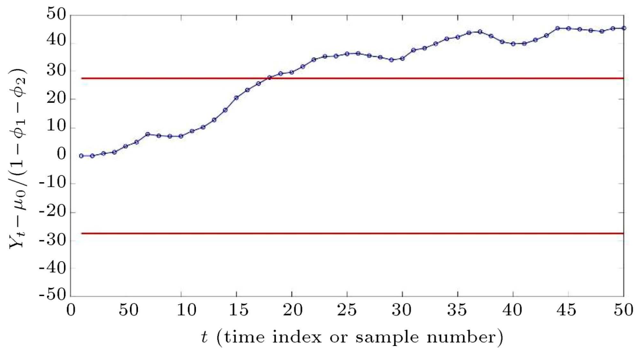


Figure 10. Proposed control chart.

Figure 10 reveals that the proposed method signals out of control in the 18th sample. Because of using a second-order filter, a level shift is magnified and the chart signals rapidly.

It is useful to compare SNR for two control schemes in this example. When $t \to \infty$, $E[Y_t] \to \frac{E[X_t]}{1-\phi_1-\phi_2} = 50$. According to Eq. (30), variance of Y_t equals 89.8. Thus:

$$SNR_{Y_t} = \left(\frac{E^2[Y_t]}{\mathrm{var}(Y_t)}\right) = 27.84.$$

For Shewhart control chart as $t \to \infty$, $E[X_t] = 1$ according to Eq. (29). It is obvious that variance of X_t equals 1. Thus:

$$SNR_{X_t} = \left(\frac{E^2[X_t]}{\mathrm{var}(X_t)}\right) = 1 \quad \text{for} \quad t \to \infty.$$

In other words, before using the proposed filter, SNR equals 1, while SNR raised to the value of 27.84 after using this filter. By removing noise and smoothing the input signal, the proposed second-order filter increases SNR in comparison to Shewhart control chart. This property helps detect changes in the process mean more rapidly.

Using the proposed method for monitoring the process mean requires determination of the values of ϕ_1 and ϕ_2. The precise determination of the values of ϕ_1 and ϕ_2 is left as further research. However, some guidelines are provided in the next subsection.

4.3.1. Determination of ϕ_1 and ϕ_2

ϕ_1 and ϕ_2 are two real numbers that characterize the behavior of the filter. These two parameters must be chosen so that the conditions in 15 are satisfied. These conditions are stability conditions for a second-order system or stationary conditions for an AR(2) process. These conditions could be plotted in ϕ_1 and ϕ_2 space, as shown in Figure 11.

If one chooses a point in the non-oscillatory region, according to Eq. (19), the impulse response of the system becomes non-oscillatory and decays exponentially. While if a point in the oscillatory region is chosen, the corresponding impulse response, provided in Eq. (22), oscillates and damps. According to Eq. (29), for choosing the values of ϕ_1 and ϕ_2, the response of the system to $\nu\ [t] = \mu_0 + (\gamma\sigma_0)u[t - t_0]$ must be determined. This response could be found in the time domain using Eq. (29). However, it is better to study the behavior of the second-order system in the frequency domain. Without loss of generality, assuming that $\mu_0 = 0$, it is convenient to study the behavior of level shift $\nu\ [t]$ in the frequency domain. The Fourier transformation of $\nu[t]$ equals to:

$$V(e^{jw}) = (\gamma\sigma_0)\frac{e^{-jwt_0}}{1 - e^{-jw}}. \tag{32}$$

The function shown in Eq. (32) is complex valued. The magnitude of this function is:

$$\left|V(e^{jw})\right| = \frac{|\gamma|\sigma_0}{\sqrt{2(1 - \cos(w))}}. \tag{33}$$

Note that $|V(e^{jw})|$ is a periodic function of w. One period of this function is shown in Figure 12 for $w \in [-\pi, \pi)$ when $|\gamma|\sigma_0 = 1$.

Figure 12 shows that the low frequencies near $w = 0$ make a large contribution in $v[t]$. Thus, for magnifying level shift, a low-pass filter must be designed. For studying the effect of using a second-order system for magnifying the level shift, it is more suitable to investigate system in the frequency domain. Taking Fourier transformation of both sides of Eq. (29) results in the following:

$$\Lambda(e^{jw}) = H(e^{jw})V(e^{jw}), \tag{34}$$

where $\Lambda(e^{jw})$ is the Fourier transformation of $\lambda[t]$ and $H(j\omega)$ is the frequency response of the system shown in Eq. (18). Substituting Eq. (32) into Eq. (34) and taking an absolute value yields:

$$\left|\Lambda(e^{jw})\right| = \frac{|\gamma|\sigma_0}{|1-\phi_1 e^{-jw} - \phi_2 e^{-j2w}|\sqrt{2(1-\cos(w))}}. \tag{35}$$

A low-pass filter requires large absolute values for

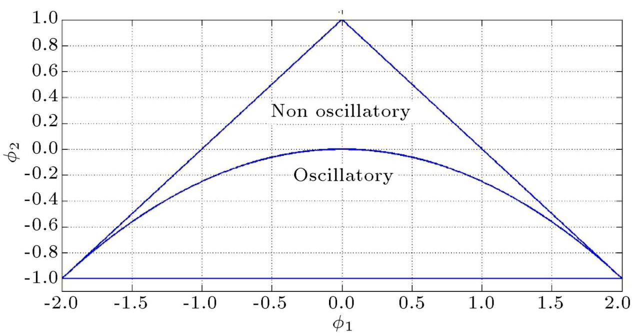


Figure 11. Stability region of the second-order filter.

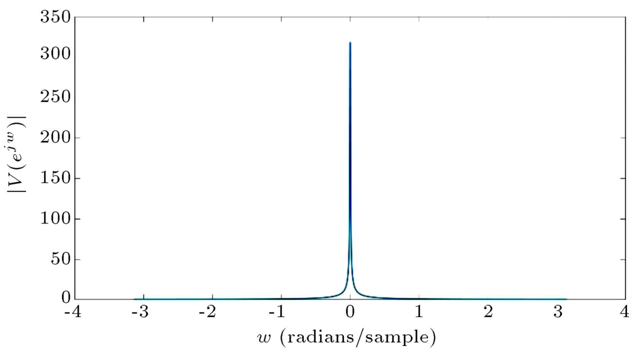


Figure 12. Magnitude of Fourier transformation of level shift.

$w \simeq 0$ in order to magnify low frequencies. Thus, according to Eq. (35), for $w \simeq 0$, it is appropriate to minimize $|1 - \phi_1 - \phi_2|$. This value is minimized when $\phi_1 + \phi_2 = 1$. However, according to the stability conditions for a system, $\phi_1 + \phi_2$ must be less than 1. Thus, for determining the values of ϕ_1 and ϕ_2, it is better to choose these values in stability region near the border of $\phi_1 + \phi_2 = 1$. This choice makes the system sensitive to level changes. In the next section, Markov chain approach is used to determine the in-control ARL of the proposed method. Determination of the in-control ARL is necessary for specifying the control limit.

Note that according to Box et al. [22], the main purpose of process control is to cancel out the entropy and disorganization of processes. Process control techniques are categorized into two classes. The first class is Statistical Process Control (SPC). The main objective of SPC is to detect when the process is out of control by process monitoring and remove assignable causes of variations. The second class of statistical process control methods is Automatic Process Control (APC). In APC, it is attempted to compensate disturbances by process adjustment. The main goal of process adjustment is to maintain the output of a process at a desired level by manipulating a control variable. This manipulation is usually performed through feedback and feedforward control loops. Note that the proposed method, which is a combination of second-order filters and control charts, falls into the SPC category. More formally, the purpose of the proposed method is to monitor the process rather than to adjust the process using feedback or feedforward control loops. However, according to English and Case [23], SPC methods could be combined with APC tools to control the processes more effectively. For example, the proposed method may be placed within a feedback control loop to adjust the process better. The block diagram of this structure is shown in Figure 13.

In Figure 13, R_t is the desired level of process. Disturbances are added to the output of process to make X_t. Using the proposed second-order filter, X_t is transformed into Y_t. Statistical stability of Y_t is then monitored by means of the proposed control chart. If the control chart shows that the process is in-control, the control signal C_t will be R_t. Otherwise, for out-of-control situations, $C_t = Y_t$. The control error, $E_t = R_t - C_t$, is then computed. Afterwards, a Proportional-Integral-Derivative (PID) controller adjusts the process based on E_t value. For more details about PID controllers and feedback control loops, one may refer to Smith and Corripio [24]. Thus, it is possible to use the proposed method within feedback control loops.

5. Markov chain approach to determining the in-control ARL

In a statistical process, control performances of control charts are usually evaluated using ARL criterion. Control charts are designed to have a large in-control ARL. However, the control charts are desired to have a small ARL when the process is out of control. In traditional Shewhart control charts, test statistics used to monitor the process are independent. Thus, the distribution of the Run Length (RL) is geometric and computation of ARL is straightforward. However, in some cases including EWMA chart, CUSUM chart, and the proposed control chart in this study, the test statistics are correlated. Computation of the ARL for these control charts is somehow complicated. Three different approaches may be used to compute ARL of a control chart. These include analytical, simulation, and Markov chains. The ARLs of different control charts such as EWMA and CUSUM charts in Chang and Wu [25], Fu and Spring [26], and Bohm and Hackl [27] and ARMA chart in Jiang et al. [28] were computed. In this research, Markov chain approach was employed to compute the in-control ARL. Suppose that two

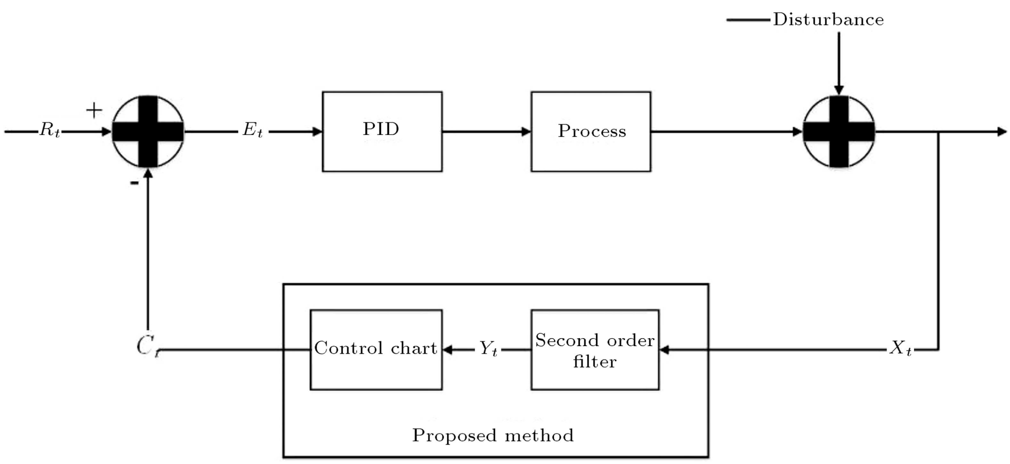


Figure 13. Combination of the proposed method with feedback control loop.

parameters of the second-order system, namely ϕ_1, ϕ_2, and the parameter of control limits, L, are determined. Let $\mathbf{z}_t = (Y_t, Y_{t-1})'$ be expressed as follows:

$$\mathbf{z}_t = \begin{pmatrix} Y_t \\ Y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \bar{X}_t$$

$$= \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \mathbf{z}_{t-1} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \bar{X}_t. \quad (36)$$

According to Eq. (36), the conditional distribution of $\mathbf{z}_t$ given $\mathbf{z}_{t-1}$ only depends on the distribution of $\bar{X}_t$. Thus, $\mathbf{z}_t$ has Markov property and generates a Markov process with a continuous state space. This Markov process may be approximated by a Markov chain with a finite state space.

For a given $L > 0$, each side of the control chart is segmented into $m + 1$ states where one state is used to show the values exceeding the control limits. The value of m could be determined by an analyst. Note that at a fixed value of L, the larger the value of m, the more precise the approximation of ARL. The total number of states is then equal to $2m+3$. For each state in the control chart, a strip with width $\Delta = \frac{2L\sigma_Y}{2m+1}$ is then defined. Total number of strips inside the control chart is then $2m + 1$. The reason that the number of strips is 2 units less than the number of states is that 2 out of $2m + 3$ states are used to show the situations in which the test statistics exceed the control limits.

The segmentation of the proposed control chart along with the strips for $m = 3$ is shown in Figure 14.

One may discretize the value of Y_t based on the segment in which Y_t falls. According to the notation used by Chang and Wu [25], let $D_0(Y_t)$ be a discretized value of Y_t defined as follows:

$$D_0(Y_t) = \begin{cases} i & (i - 0.5)\Delta \le Y_t \le (i + 0.5)\Delta \\ & \text{for} \quad i = -m, -m+1, \cdots, m \\ m + 1 & Y_t > (m + 0.5)\Delta \\ -m - 1 & Y_t < -(m + 0.5)\Delta \end{cases} \quad (37)$$

By discretizing Y_t, the vector $\mathbf{z}_t$ is discretized and shown by:

$$D_0(\mathbf{z}_t) = (D_0(Y_t), D_0(Y_{t-1}))', \quad (38)$$

where $D_0(Y_t)$ is defined in Eq. (37). Since the $\mathbf{z}_t$'s are characterized by the Markov property, the vector $D_0(\mathbf{z}_t)$ constructs a Markov chain with a finite state space. The state space of this Markov chain is as:

$$S = \{(S_1, S_2)', \alpha\}, \quad (39)$$

where $S_1, S_2 \in \{-m, -m + 1, \cdots, m\}$ and α is used to show the state in which the process is out of control. This state is an absorbing state which is:

$$\alpha = \left\{ \begin{pmatrix} m+1 \\ S_1 \end{pmatrix}, \begin{pmatrix} -m-1 \\ S_1 \end{pmatrix}, \begin{pmatrix} m+1 \\ m+1 \end{pmatrix}, \begin{pmatrix} -m-1 \\ -m-1 \end{pmatrix} \right\};$$

$$S_1 \in \{-m, -m + 1, \cdots, m\}. \quad (40)$$

If the $\mathbf{z}_t$ process moves to each element of set α, it will remain in this set. Thus, it is an absorbing state.

Let $\mathbf{P} = [p_{ab}]$ be the transition matrix of $D_0(\mathbf{z}_t)$, where p_{ab} is the probability of $D_0(\mathbf{z}_t)$ moving from state a to state b. The method of computing p_{ab}'s is provided in Appendix D. After constructing $\mathbf{P}$, one could compute the in-control ARL of the control chart using the formula provided in Appendix E. For different values of ϕ_1, ϕ_2 and the in-control ARL, the values of L for the proposed control chart are provided in Table 3. These values are used in the next section to assess the performance of the proposed control chart.

As stated in Section 3, values of system poles, x_1^{-1} and x_2^{-1}, determine the system to be low-pass, high-pass, or medium-pass filter. Thus, depending on the values of system poles, the proposed second-order filter could be low pass, high pass, or medium pass. The system relation shown in Eq. (36) is valid in all of these cases. As a result, the conditional distribution of $\mathbf{z}_t$ given $\mathbf{z}_{t-1}$ only depends on the distribution of $\bar{X}_t$, as long as $|x_1| > 1$ and $|x_2| > 1$. Thus, the stochastic process $\mathbf{z}_t$ is a Markov process, whether the proposed second order filter is low pass, high pass, or medium pass. So, the hybrid method discussed in this study could be used even in cases where the proposed filter is high pass or medium pass. Note that the main purpose of this study is to detect shifts in process mean. To reduce noises, it is recommended that system poles be selected close to 1 and a low pass filter be designed.

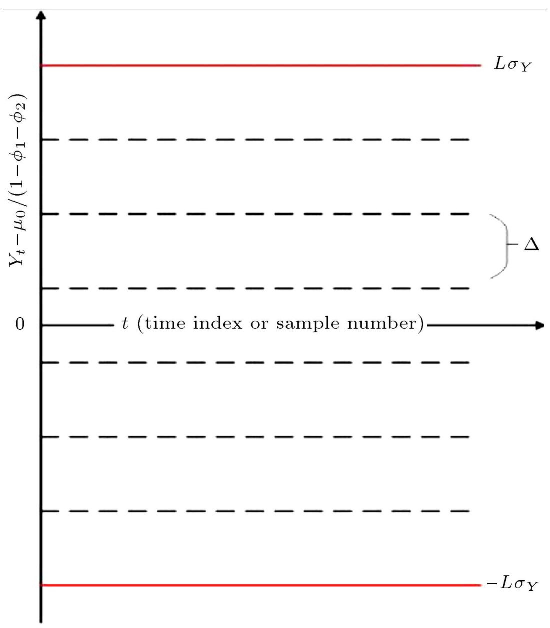


Figure 14. Segmentation of the proposed control chart.

Table 3. Values of L for the proposed control chart ($m = 10$).

ARL	$\phi_1^2 + 4\phi_2 > 0$	$\phi_1^2 + 4\phi_2 < 0$
	$\phi_1 = 0.2$ $\phi_2 = 0.79$	$\phi_1 = 1.8$ $\phi_2 = -0.85$
200	1.39	2.63
250	1.50	2.79
300	1.61	2.89
370	1.71	3.05
500	1.88	3.25
ARL	$\phi_1 = 0.85$ $\phi_2 = 0.14$	$\phi_1 = 1.5$ $\phi_2 = -0.6$
200	1.51	2.54
250	1.62	2.63
300	1.73	2.70
370	1.86	2.79
500	2.05	2.92

6. Simulation studies

In this section, the performance of the proposed control chart to detect level shifts in the process mean is evaluated by means of simulation in MATLAB R2018b software. The proposed control chart is compared with two frequently used control charts, including Shewhart $\bar{X}$ chart and EWMA chart. Shewhart $\bar{X}$ and EWMA control charts were described in Section 1. The Markov chain approach was used to design the control charts proposed by Chang and Wu [25] and Bohm and Hackl [27]. However, these charts are used to monitor autocorrelated processes. The proposed method was designed to monitor an independent sequence of observations. Thus, control charts to monitor autocorrelated observations were excluded from comparison.

The proposed control chart to be discussed here is shown by AR(2). Two criteria were used in this section to compare the control charts. The first one is ARL, and the second criterion is the percentage of times when a chart signals before change occurs, i.e., Percentage of False Signals (PFS). The smaller the value of the latter criterion, the better a control chart. The control limits of the charts were specified so that the in-control ARL be equal to 370. In Eq. (2), using the Markov chain approach, $L = 2.085$ and $\lambda = 0.15$ lead to an in-control ARL as being equal to 370. An EWMA chart with $\lambda = 0.15$ is a special case of AR(2) model with $\phi_1 = 0.85$ and $\phi_2 = 0$. The introduction of the second parameter ϕ_2 to a first-order system yields a second-order model, AR(2). Let $\phi_1 = 0.85$ and $\phi_2 = 0.14$ for the simulation study. For these values of parameters, $L = 1.86$ was selected from Table 3. According to Eq. (20), by these values for $\phi_1$and ϕ_2, $x_1 = 1.0088$ and $x_2 = -7.08$ were obtained. Thus, poles of the system are equal to $x_1^{-1} = 0.99$ and $x_2^{-1} = -0.14$. As could be seen, one pole is very close to 1. Thus, the corresponding system is a low-pass filter. In each simulation run, the data were generated using the following model:

$$X_t = \begin{cases} \varepsilon_t & t < t_0 \\ \gamma + \varepsilon_t & t \geq t_0 \end{cases} \tag{41}$$

where ε_t's are independent standard normal random variables, t_0 is the time instance at which level shift in process mean occurs, and γ is the value of the level shift. Note that individual observations in this section are used to control the process, i.e., $n = 1$. As stated in Section 3.1, it is possible to use $\frac{SNR_{Y_t}}{SNR_{X_t}}$ to study the effect of the proposed filter on input signal. Note that $E[X_t] = \gamma$, $E[Y_t] = \frac{\gamma}{1-\phi_1-\phi_2}$ for $t \geq t_0$. Using Eq. (30), one may compute var(Y_t) for the proposed filter which is equal to 27.96. It is possible to compute SNR values for different values of γ. For example, if $\gamma = 1$, then $SNR_{X_t} = 1$ and $SNR_{Y_t} = 3.58$ for $t \geq t_0$. Thus, after the occurrence of shifts, the proposed low-pass filter smooths out input signals. This will help the control scheme detect shifts rapidly.

After generating each observation, test statistics for each control chart was computed. For each control chart, generating a new observation continued until an observation would exceed its control limits and then, the run length was recorded. For each value of $t_0 = 20, 50, 75$ and $\gamma = -2(0.1)2$, this procedure was repeated 10000 times. The mean and standard deviation of the run lengths for each control chart as well as the proportion of the false signals were then computed. Results for ARL criterion are shown in Figure 15.

In Figure 15, the 95% confidence intervals of ARL values are also shown using vertical error bars. These intervals are defined as $ARL \pm 2\frac{SDRL}{\sqrt{10000}}$ in which

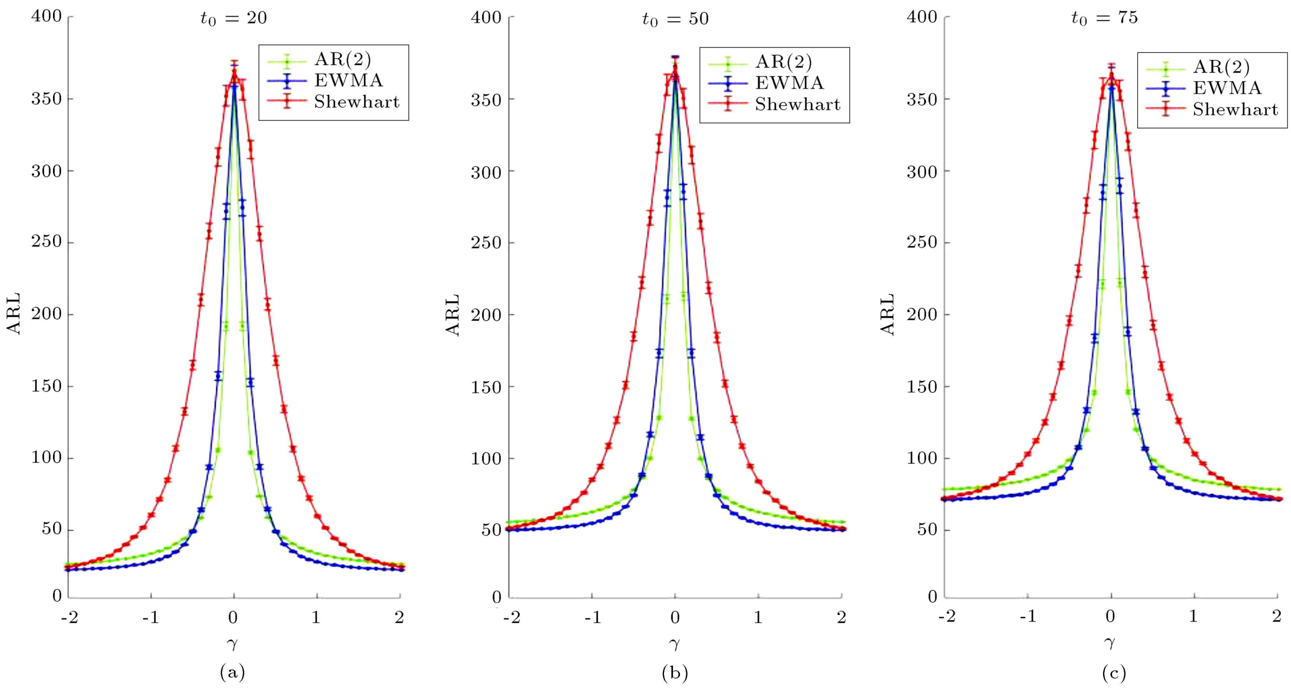


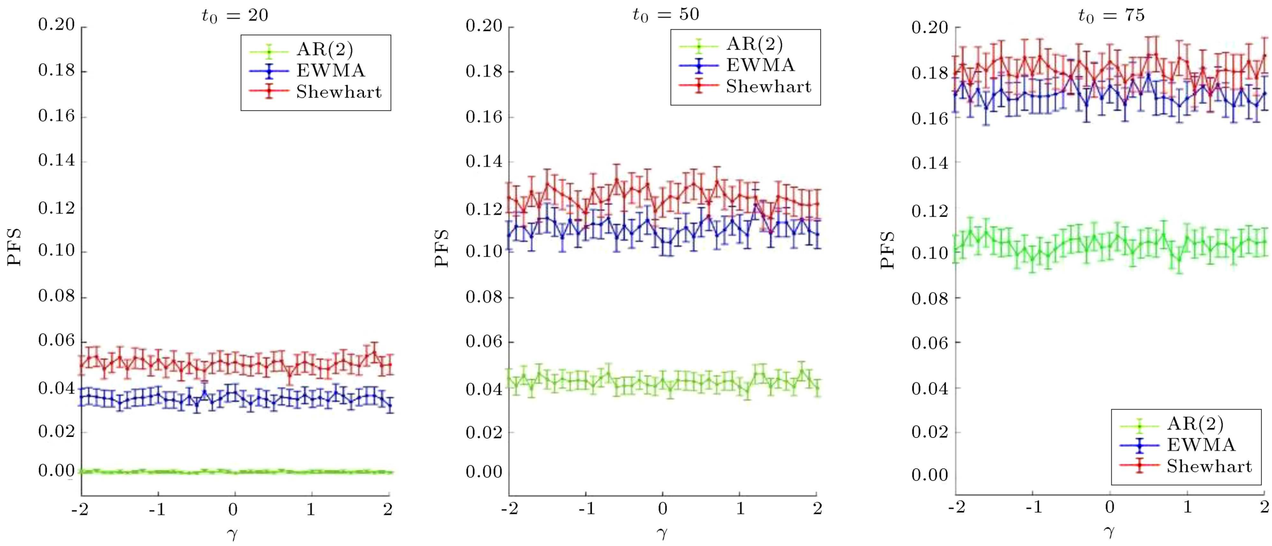
Figure 15. ARL of control charts: (a) $t_0 = 20$, (b) $t_0 = 50$, and (c) $t_0 = 75$.

SDRL is the estimated standard deviation of run length. Figure 15 shows that the ARL of the three control charts reduces when $|\gamma|$ increases. However, the reduction rate of ARL for AR(2) and EWMA charts is steeper than that for the Shewhart chart. It is shown that EWMA and AR(2) detect the level change more rapidly than the Shewhart chart. Comparing the AR(2) and the EWMA charts, one could find that for small changes, the AR(2) chart on average detects the level shifts before the EWMA charts. For large changes, the EWMA chart performs slightly better than AR(2). AR(2) and EWMA charts outperform Shewhart chart in terms of ARL. From Figure 15, it is obvious that as $|\gamma|$ increases, the SDRL decreases for all control charts. This indicates that as shifts become larger, the standard deviation of run length decreases and shifts are detected more precisely. Note that the confidence level is the same for estimating ARL of the three control charts. However, changes in SDRL values result in changes in width of confidence intervals. Note that the performance of the proposed method in detecting shifts in the process could be enhanced by assigning proper values to ϕ_1 and ϕ_2. The proper values of these parameters may be determined by a number of optimization methods. By using this approach, the behavior of the proposed method could be modified to accommodate larger mean changes. Optimizing the performance of the proposed method is left as an area for further research. Results for PFS criterion are shown in Figure 16.

In Figure 16, the 95% confidence intervals of PFS values are also shown. These intervals are defined as $PFS \pm 2\frac{SDPFS}{\sqrt{10000}}$ in which SDPFS is the estimated standard deviation of PFS. Considering PFS curves shown in Figure 16 reveals that PFS of the three charts is relatively constant. However, the probability of false alarms for AR(2) is much less than that for the other two charts. Considering the two criteria shows that the AR(2) chart detects changes more rapidly and precisely. The AR(2) chart outperforms EWMA in terms of PFS. In terms of PFS, the proposed AR(2) chart is much better than the other two charts. In summary, the results of simulation studies indicate that the level shifts in process mean are detected more rapidly and precisely when the proposed AR(2) chart is used.

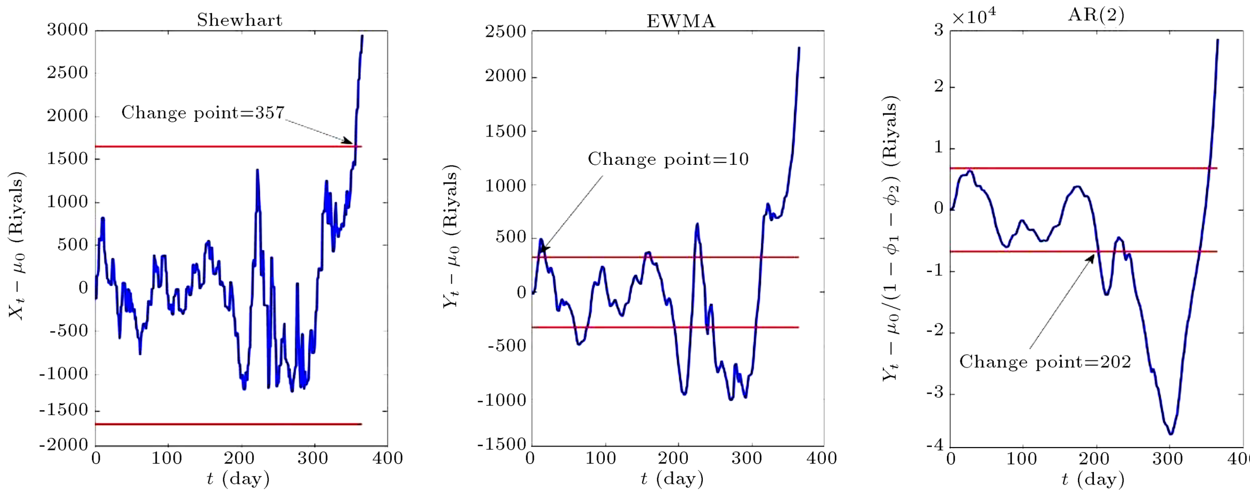
7. Case study – Existing data

In this section, the application of the proposed control scheme to monitor a real process is shown. The stochastic process which was monitored is Euro exchange rate to Iranian Riyal. Time series data of this stochastic process may be found on the website of Islamic Republic of Iran Central Bank: https://cbi.ir/exratesadv/exratesadv_en.aspx. Euro exchange rates, from 01/08/2019 to 30/07/2020, a total of 365 observations, were used as a process to be monitored. The main purpose of monitoring this process was to detect if the mean of Euro exchange rates shifted. Three control charts including Shewhart, EWMA, and AR(2) were employed in this section to monitor the process. Parameters of these charts are the same as those in Section 6. To construct control



(a) (b) (c)

Figure 16. PFS of control charts: (a) $t_0 = 20$, (b) $t_0 = 50$, and (c) $t_0 = 75$.



(a) (b) (c)

Figure 17. Control charts for monitoring exchange rate: (a) Shewhart, (b) EWMA, and (c) AR(2).

charts, it is required to compute in-control process mean and standard deviation. To remove the effects of outliers, robust estimators of process mean and standard deviation were applied to estimate the values of these parameters. Process mean was estimated using the sample median of 365 observations being equal to 46525 Riyals. To estimate the process standard deviation, Median Absolute Deviation (MAD) from sample median was used. For this data set, $MAD =$ 549.6 Riyals. Using these estimated parameters and parameters introduced in Section 6, control limits of each control chart were computed. These charts along with the corresponding sequences of test statistics are shown in Figure 17.

Considering Figure 17(a), it is obvious that Shewhart control chart detects changes in the process mean at point 357. Whereas some abrupt changes are evident before this point. In fact, the presence of heavy noise in observations has made Shewhart control chart insensitive to detecting changes in process mean. In EWMA control chart, Figure 17(b), fluctuations were removed to some extent. In this chart, a few initial observations fall outside the control limits. While it seems that the process mean is initially in-control and these observations fall outside control limits due to noises.

Considering AR(2) control chart, Figure 17(c), it is evident that noises were removed. In fact, by using a low pass filter, input signal was smoothed out. As a result, changes in process mean were magnified and were detected using proposed AR(2) control chart.

8. Results and discussion

In this study, a unified approach based on signal and system theory was proposed to monitor the process mean. Most of the approaches to monitoring processes are designed only based on statistical properties of processes in the time domain. Some important features of processes may be overlooked if the processes are only studied in the time domain. Thus, this research study proposed studying the process in the frequency domain, as well. In this research, the test statistics for monitoring the process were considered as an output of the system. Based on the signal theory, the behavior of this transforming system was studied in the frequency domain. In this domain, some features of the process are unfolded. Thus, one could design the control scheme more efficiently.

The problem of monitoring process mean was addressed in this paper. An attempt was made to design a control chart that can be sensitive to level shifts in process mean. The design of such a control chart is based on the signal and system approach. More accurately, a unified control scheme based on the LTI systems was developed. This control scheme is called a filter. It was suggested that a second-order system be used to filter out the sample means and then, to control the process via the proposed control chart. This filter has two parameters that must be predetermined in order to detect the level shifts rapidly. To determine the values of these parameters, the behavior of the second-order system was investigated in the frequency domain. It was shown that the level shift only has low frequency components. Thus, it is required to design a filter that magnifies the low frequencies. Based on this fact, the values of the system parameters are then chosen. After designing the filter, control limits of the proposed control chart must be determined. To adjust the control limits, it is recommended that the Markov chain approach be used. The proposed control chart was compared with EWMA and the Shewhart control charts through simulation. It must be noted that the EWMA chart is a special case of the proposed AR(2) chart with $\phi_1 = 1 - \lambda$ and $\phi_2 = 0$. The Shewhart chart could be considered as a special case of AR(2) chart when $\phi_1 = \phi_2 = 0$. The simulation results showed that the proposed control chart detected the level shifts more rapidly and precisely. In addition to the second-order filters, the other filters may be used to monitor the process. The proposed method is able to detect shifts in process mean rapidly with low frequency components. If out of control conditions with high frequency components need to be detected, it is suggested to use other filters that magnify high frequencies.

9. Conclusion

In this paper, monitoring process mean was considered. A second-order low-pass filter was proposed to remove noises from sequence of sample means. This filter is a Linear Time Invariant (LTI) system that is used to transform sample mean into a test statistics. It was shown that by using this filter, any changes in process mean were magnified, while at the same time, noises were filtered out. As a consequence, the SNR value increased which resulted in detecting shifts rapidly. By choosing poles of the proposed system appropriately, low-pass, high-pass, or medium-pass filters could be obtained. All of these second-order filters could be used to remove nuisance components of signals. As a result, changes in process could be detected rapidly. The proposed second-order filter has two parameters that must be determined. Performance of the proposed control chart depends on the values of these parameters. To identify appropriate values of these parameters, some optimization methods, such as gradient based search or metaheuristic algorithms, should be used. This is an open area that could be studied to design the proposed control chart better. The filtering approach proposed in this research may be used to control correlated observations. These areas are left for further research. The proposed monitoring scheme may be used to monitor various stochastic processes, including chemical, financial, and electrical processes.

Acknowledgements

Financial supports granted by Niroo Research Institute (NRI), under the contract number 95/20031/01, are gratefully acknowledged. The authors also would like to thank the anonymous reviewer and the editor whose comments and suggestions improved greatly the presentation of the paper.

Nomenclature

μ_0	In-control process mean
σ_0	In-control process standard deviation
n	Sample size to compute sample mean
$\mathbb{Z}$	Set of integer numbers
$t \in \mathbb{Z}$	Time index
$\bar{X}_t$	Sample mean at time t
$\lambda \in [0, 1]$	Parameter of EWMA control chart
L	Coefficient of control chart
$j = \sqrt{-1}$	Imaginary unit
$x[t]$	Input signal of LTI system
$y[t]$	Output signal of LTI system
$\delta[t]$	Unit impulse function
$u[t]$	Unit step function
$\mathbb{Z}^+ \subset \mathbb{Z}$	Set of positive integer numbers

$\mathbb{Z}^- \subset \mathbb{Z}$	Set of negative integer numbers
Y_t	Proposed test statistic to be monitored
$h[t]$	Impulse response of LTI system
$t', k \in \mathbb{Z}$	Shifts in time index
w	Frequency variable
$H(e^{jw})$	Fourier transform of $h[t]$
$X(e^{jw})$	Fourier transform of $x[t]$
$Y(e^{jw})$	Fourier transform of $y[t]$
ϕ_1	Coefficient of $y[t-1]$
ϕ_2	Coefficient of $y[t-2]$
t_0	Change point in process mean
γ	Magnitude of shift in process mean
$E[.]$	Expectation operator
x_1, x_2	Inverse of second order filter poles
$\nu[t]$	Expected value of $\bar{X}_t$
p_{ab}	One step transition probability from state a to state b
$\lambda[t]$	Expected value of Y_t
$V(e^{jw})$	Fourier transform of $\nu[t]$
$\Lambda(e^{jw})$	Fourier transform of $\lambda[t]$
m	Parameter used to determine states of proposed control chart
Δ	Width of control chart strips
$D_0(.)$	Discretizing function
S	Markov chain state space
$\mathbf{P}$	Transition probability matrix
r	Magnitude of x_1, x_2
ρ	Inverse of r
$\theta = \measuredangle x_1$	Phase of x_1
$P(.\vert.)$	Conditional probability operator
$\mathbf{I}$	Identity matrix
α	Absorbing state

References

1. Montgomery, D.C., *Introduction to Statistical Quality Control*, John Wiley and Sons, New York (2013).
2. Jiang, W. "Average run length computation of ARMA charts for stationary processes", *Communications in Statistics-Simulation and Computation*, **30**(3), pp. 699–716 (2007).
3. Wu, Z. and Spedding, T.A. "A synthetic control chart for detecting small shifts in the process mean", *Journal of Quality Technology*, **32**(1), pp. 32–38 (2000).
4. Zhang, N.F. "A statistical control chart for stationary process data", *Technometrics*, **40**(1), pp. 24–38 (1998).
5. Chang, S.I. and Aw, C.A. "A neural fuzzy control chart for detecting and classifying process mean shifts", *International Journal of Production Research*, **34**(8), pp. 2265–2278 (1996).
6. Harris, T.J. and Ross, W.H. "Statistical process control procedures for correlated observations", *The Canadian Journal of Chemical Engineering*, **69**, pp. 48–57 (1991).
7. Saif, A.W. "A frame work for the integration of statistical process control and engineering process control", *Industrial & Systems Engineering Conference (ISEC)*, Jeddah, Saudi Arabia (2019).
8. Rabyk, L. and Schmid, W. "EWMA control charts for detecting changes in the mean of a long-memory process", *Metrika*, **79**(3), pp. 267–301 (2016).
9. Shokrizadeh, R., Saghaei, A., and Yaquninejad, Y. "An evaluation of an adaptive generalized likelihood ratio charts for monitoring the process mean", *International Journal of Applied Operational Research*, **6**(3), pp. 59–68 (2016).
10. Yang, S.F. and Arnold, B.C. "A simple approach for monitoring process mean and variance simultaneously", *Frontiers in Statistical Quality Control*, **11**, Springer International Publishing, pp. 135–149 (2015).
11. Chen, G., Cheng, S.W., and Xie, H. "Monitoring process mean and variability with one EWMA chart", *Journal of Quality Technology*, **33**(2), pp. 223–233 (2001).
12. Apley, D.W. and Shi, J. "The GLRT for statistical process control of autocorrelated processes", *IIE Transactions*, **31**, pp. 1123–1134 (1999).
13. Lu, C.W. and Reynolds, M.R. "EWMA control charts for monitoring the mean of autocorrelated processes", *Journal of Quality Technology*, **31**(2), pp. 166–188 (1999).
14. Zhang, W., Jiao, J., Yang, M., et al. "An enhanced adoptive CUSUM control chart", *IIE Transactions*, **41**(7), pp. 642–653 (2009).
15. Oppenheim, A.V., Wilsky, A.S., and Hamid, S., *Signal and Systems*, Pearson New International Edition (2014).
16. Wang, K., Chen, J., and Song, Z. "Data-driven sensor fault diagnosis systems for linear feedback control loops", *Journal of Process Control*, **54**, pp. 152–171(2017).
17. Chen, M., Xu, G., Yan, R., et al. "Detecting scalar intermittent faults in linear stochastic dynamic systems", *International Journal of System Science*, **46**(8), pp. 1337–1348 (2015).
18. Zuo, C., Song, X., and Park, J.H. "Linear estimation for system with unknown measurement input and missing measurement", *Chinese Automation Congress (CAC)*, Jinan, China (2017).
19. Liu, W. and Shi, P. "Optimal linear filtering for networked control systems with time-correlated fading channels", *Automatica*, **101**, pp. 345–353 (2018).
20. Liu, W., Zhang, H., Yu, K., et al. "Optimal linear filtering for networked systems with communication constraints, fading measurements, and multiplicative noises", *International Journal of Adaptive Control and Signal Processing*, **31**(7), pp. 1019–1039 (2016).

21. Eijnden, V.D., Knops, Y., and Heertjes, M.F. "A hybrid integrator-gain based low-pass filter for nonlinear motion control", *Conference on Control Technology and Applications (CCTA)*, Copenhagen, Denmark (2018).
22. Box, G.E.P., Jenkins, G.M., and Reinsel, G.C., *Time Series Analysis, Forecasting and Control*, Prentice-Hall International, New Jersey (1994).
23. English, J.R. and Case, K.E. "Control charts applied as filtering devices within feedback control loop", *IIE Transactions*, **22**(3), pp. 255–269 (1990).
24. Smith, C.A. and Corripio, A.B., *Principals and Practice of Automatic Process Control*, John Wiley and Sons, New York (1997).
25. Chang, Y.M. and Wu., T.L. "On the average run length of control charts for autocorrelated processes", *Methodology and Computing in Applied Probability*, **13**(2), pp. 419–431 (2011).
26. Fu, J.C. and Spring, F. "On the average run length of quality control schemes using a Markov chain approach", *Statistics & Probability Letters*, **56**, pp. 369–380 (2002).
27. Bohm, W. and Hackl, P. "The effect of serial correlation on the in-control average run length of cumulative score charts", *Journal of Statistical Planning and Inferences*, **54**, pp. 15–30 (1996).
28. Jiang, W., Tsui, K.L., and Woodall, W.H. "A new SPC monitoring method: The ARMA chart", *Technometrics*, **42**(4), pp. 399–410 (2000).

Appendix A

Fourier transformation of LTI system response

Substituting Eq. (7) into Eq. (9) yields the following:

$$Y(e^{jw}) = \sum_{t=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} h[k]x[t-k]e^{-jwt}. \tag{A.1}$$

Let $t - k = t' \in \mathbb{Z}$, thus:

$$\begin{aligned} Y(e^{jw}) &= \sum_{t'=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} h[k]x[t']e^{-jw(k+t')} \\ &= \sum_{t'=-\infty}^{+\infty} x[t']e^{-jwt'} \sum_{k=-\infty}^{+\infty} h[k]e^{-jwk}. \end{aligned} \tag{A.2}$$

Using Eqs. (10) and (11), we have:

$$Y(e^{jw}) = X(e^{jw})H(e^{jw}). \tag{A.3}$$

Appendix B

Impulse response of second-order filter

To obtain the inverse Fourier transformation of $H(e^{jw})$, it is better to first factorize $H(e^{jw})$. Let $x = e^{-jw}$; thus, the denominator of $H(e^{jw})$ becomes $D(x) = 1 - \phi_1 x - \phi_2 x^2$. Suppose x_1 and x_2 denote the roots of $D(x)$. It is easy to verify that these roots are:

$$x_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2},$$

$$x_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}. \tag{B.1}$$

Note that $D(x)$ may be written as follows:

$$\begin{aligned} D(x) &= 1 - \phi_1 x - \phi_2 x^2 = -\phi_2 \left(x^2 + \frac{\phi_1}{\phi_2}x - \frac{1}{\phi_2}\right) \\ &= -\phi_2(x - x_1)(x - x_2) \\ &= -\phi_2 x_1 x_2 \left(1 - \frac{x}{x_1}\right)\left(1 - \frac{x}{x_2}\right). \end{aligned} \tag{B.2}$$

According to Eq. (B.1), $\phi_2 x_1 x_2 = -1$; thus, $D(x)$ could be written as follows:

$$D(x) = \left(1 - \frac{x}{x_1}\right)\left(1 - \frac{x}{x_2}\right). \tag{B.3}$$

If $\phi_1^2 + 4\phi_2 \neq 0$, two roots are distinct. For the sake of simplicity, suppose that this is the case. Now, $H(e^{jw})$ could be written as follows:

$$H(e^{jw}) = \frac{1}{(1 - x_1^{-1}e^{-jw})(1 - x_2^{-1}e^{-jw})}. \tag{B.4}$$

Then, $H(e^{jw})$can be factorized in the following manner:

$$H(e^{jw}) = \frac{\frac{1}{1 - x_2^{-1}x_1}}{1 - x_1^{-1}e^{-jw}} + \frac{\frac{1}{1 - x_1^{-1}x_2}}{1 - x_2^{-1}e^{-jw}}. \tag{B.5}$$

Taking inverse Fourier transformation of both sides of Eq. (B.5), one may obtain the following:

$$h[t] = \frac{x_2}{x_2 - x_1}\left(\frac{1}{x_1}\right)^t u[t] + \frac{x_1}{x_1 - x_2}\left(\frac{1}{x_2}\right)^t u[t], \tag{B.6}$$

where $u[t]$ is a unit step function defined in Eq. (6). According to system stability, $|x_1| > 1$ and $|x_2| > 1$.

Based on the sign of $\phi_1^2 + 4\phi_2$ in Eq. (B.1), two cases could be considered. If $\phi_1^2 + 4\phi_2 > 0$, two roots x_1 and x_2 are real numbers and $h[t]$ will tend to zero without oscillation as $t \to +\infty$. If $\phi_1^2 + 4\phi_2 < 0$, the two roots are complex conjugate numbers. In this case, x_1 and x_2 could be written as:

$$x_1 = \frac{-\phi_1 + j\sqrt{-\phi_1^2 - 4\phi_2}}{2\phi_2} = re^{j\theta} = r\cos\theta + jr\sin\theta,$$

$$x_2 = \frac{-\phi_1 - j\sqrt{-\phi_1^2 - 4\phi_2}}{2\phi_2} = re^{-j\theta} = r\cos\theta - jr\sin\theta, \tag{B.7}$$

where $r = |x_1| = |x_2|$ and $\theta = \measuredangle x_1$. It could be shown that:

$$x_1 + x_2 = -\frac{\phi_1}{\phi_2} = 2r\cos\theta, \tag{B.8}$$

$$x_1 - x_2 = j\frac{\sqrt{-\phi_1^2 - 4\phi_2}}{\phi_2} = 2jr\sin\theta. \tag{B.9}$$

Let $r^{-1} = \rho$. Solving Eqs. (B.8) and (B.9) for ϕ_1 and ϕ_2 results in the following:

$$\phi_1 = 2\rho\cos\theta, \qquad \phi_2 = -\rho^2. \tag{B.10}$$

Substituting Eqs. (B.7) and (B.9) in (B.6) results in:

$$h[t] = \left(\frac{-r^{1-t}e^{-j\theta(t+1)}}{2jr\sin\theta} + \frac{r^{1-t}e^{j\theta(t+1)}}{2jr\sin\theta}\right)u[t]. \tag{B.11}$$

Note that:

$$e^{j\theta(t+1)} - e^{-j\theta(t+1)} = 2j\sin(\theta(t+1)). \tag{B.12}$$

Substituting (B1.2) in (B.11) results in the following:

$$h[t] = \frac{\sin((t+1)\theta)}{\sin(\theta)}\rho^t u[t]. \tag{B.13}$$

Thus, if $\phi_1^2 + 4\phi_2 < 0$, the impulse response of the second-order system oscillates and damps as $t \to +\infty$.

Appendix C

Response of unified controlling system to level shift

Define A_1 and A_2 as in the following:

$$A_1 = \frac{x_2}{x_2 - x_1}, \qquad A_2 = \frac{x_1}{x_1 - x_2}. \tag{C.1}$$

Case I: $\phi_1^2 + 4\phi_2 > 0$. Let $\phi_1^2 + 4\phi_2 > 0$. Substituting Eqs. (19) and (26) in Eq. (29), we have:

$$\lambda[t] = \sum_{k=0}^{+\infty}\left[\left(A_1\left(\frac{1}{x_1}\right)^k + A_2\left(\frac{1}{x_2}\right)^k\right)\right.$$
$$\left.(\mu_0 + (\gamma\sigma_0)u[t - t_0 - k])\right]. \tag{C.2}$$

For $t < t_0$:

$$\lambda[t] = \mu_0\sum_{k=0}^{+\infty}\left[\left(A_1\left(\frac{1}{x_1}\right)^k + A_2\left(\frac{1}{x_2}\right)^k\right)\right]. \tag{C.3}$$

After some algebraic computation of Eq. (C.3), one may obtain:

$$\lambda[t] = \frac{\mu_0}{1 - \phi_1 - \phi_2}; \qquad t < t_0. \tag{C.4}$$

For $t \geq t_0$, Eq. (C.2) becomes:

$$\lambda[t] = \sum_{k=0}^{t-t_0}\left[\left(A_1\left(\frac{1}{x_1}\right)^k + A_2\left(\frac{1}{x_2}\right)^k\right)(\mu_0 + \gamma\sigma_0)\right]. \tag{C.5}$$

Simplifying Eq. (C.5) yields:

$$\lambda[t] = (\mu_0 + \gamma\sigma_0)\left\{A_1\frac{x_1 - \left(\frac{1}{x_1}\right)^{t-t_0}}{x_1 - 1}\right.$$
$$\left. + A_2\frac{x_2 - \left(\frac{1}{x_2}\right)^{t-t_0}}{x_2 - 1}\right\}; \qquad t \geq t_0. \tag{C.6}$$

Note that if $t \to +\infty$, the control system reveals steady state behavior. This behavior is modeled by $\lim_{t\to+\infty}\lambda[t]$ according to the following:

$$\lim_{t\to+\infty}\lambda[t] = \frac{\mu_0 + \gamma\sigma_0}{1 - \phi_1 - \phi_2}. \tag{C.7}$$

Case II: $\phi_1^2 + 4\phi_2 < 0$. Now, suppose $\phi_1^2 + 4\phi_2 < 0$. Substituting Eqs. (22) and (26) in Eq. (29) results in:

$$\lambda[t] = \sum_{k=0}^{+\infty}\frac{\sin((k+1)\theta)}{\sin(\theta)}\rho^k(\mu_0 + (\gamma\sigma_0)u[t - t_0 - k]). \tag{C.8}$$

For $t < t_0$, Eq. (C.8) becomes:

$$\lambda[t] = \frac{\mu_0}{\sin(\theta)}\sum_{k=0}^{+\infty}\sin((k+1)\theta)\rho^k; \qquad t < t_0. \tag{C.9}$$

Using Euler formula, Eq. (C.9) is simplified into the following:

$$\lambda[t] = \frac{\mu_0}{1 - 2\rho\cos\theta + \rho^2} = \frac{\mu_0}{1 - \phi_1 - \phi_2}; \qquad t < t_0. \tag{C.10}$$

For $t \geq t_0$, Eq. (C.8) becomes:

$$\lambda\ [t] = \frac{\mu_0 + \gamma\sigma_0}{\sin(\theta)}\sum_{k=0}^{t-t_0}\sin((k+1)\theta)\rho^k. \qquad t \geq t_0. \tag{C.11}$$

Appendix D

Computing in control p_{ab}

For computing p_{ab}, let $a = (S_{1a}, S_{2a})'$ and $b = (S_{1b}, S_{2b})'$ be two states in the state space. Thus, p_{ab} is equal to:

$$p_{ab} = \Pr(D_0(\mathbf{z}_t) = b | D_0(\mathbf{z}_{t-1}) = a)$$
$$= \Pr(D_0(Y_t) = S_{1b}, D_0(Y_{t-1})$$
$$= S_{2b} | D_0(Y_{t-1}) = S_{1a}, D_0(Y_{t-2}) = S_{2a}). \tag{D.1}$$

If $S_{1a} \neq S_{2b}$, then $p_{ab} = 0$. Due to this fact, only some transitions are possible in the intended Markov chain. These transitions are shown below:

$$a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow b = \begin{pmatrix} S_{1b} \\ S_{1a} \end{pmatrix};$$

$$S_{1a}, S_{2a}, S_{1b} \in \{-m, -m+1, \cdots, m\}, \quad \text{(D.2)}$$

$$a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow b$$

$$\in \left\{ \begin{pmatrix} m+1 \\ S_{1a} \end{pmatrix}, \begin{pmatrix} -m-1 \\ S_{1a} \end{pmatrix}, \begin{pmatrix} m+1 \\ m+1 \end{pmatrix}, \begin{pmatrix} -m-1 \\ -m-1 \end{pmatrix} \right\},$$

$$S_{1a}, S_{2a} \in \{-m, -m+1, \cdots, m\}, \quad \text{(D.3)}$$

$$a = \alpha \rightarrow b = \alpha. \quad \text{(D.4)}$$

Probability of transition shown in Eq. (D.2) is equal to:

$$\Pr\left(D_0(Y_t) = S_{1b}, D_0(Y_{t-1}) = S_{1a} | D_0(Y_{t-1}) = S_{1a}, D_0(Y_{t-2}) = S_{2a}\right). \quad \text{(D.5)}$$

For small values of Δ, Eq. (D.5) becomes:

$$\begin{aligned} p_{ab} &= \Pr\left((S_{1b} - 0.5)\Delta \leq Y_t \leq (S_{1b} + 0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right) = \Pr\big((S_{1b} - 0.5)\Delta \\ &\leq \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \bar{X}_t \leq (S_{1b} + 0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\big) = \Pr\big((S_{1b} - 0.5)\Delta \\ &\leq \phi_1 S_{1a}\Delta + \phi_2 S_{2a}\Delta + \bar{X}_t \leq (S_{1b} + 0.5)\Delta\big) \\ &= \Pr\left(\frac{(S_{1b} - 0.5 - \phi_1 S_{1a} - \phi_2 S_{2a})\Delta\sqrt{n}}{\sigma_0} \leq Z \leq \frac{(S_{1b} + 0.5 - \phi_1 S_{1a} - \phi_2 S_{2a})\Delta\sqrt{n}}{\sigma_0}\right), \end{aligned} \quad \text{(D.6)}$$

where Z is standard normally distributed.

To compute the probability of transitions shown in Eq. (D.3), note that:

$$\Pr\left\{a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow \begin{pmatrix} m+1 \\ m+1 \end{pmatrix}\right\} = \Pr\left\{a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow \begin{pmatrix} -m-1 \\ -m-1 \end{pmatrix}\right\} = 0. \quad \text{(D.7)}$$

for $S_{1a}, S_{2a} \in \{-m, -m+1, \cdots, m\}$.

Other transition probabilities of Eq. (D.3) for small Δ are computed as follows:

$$\begin{aligned} &\Pr\left\{a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow \begin{pmatrix} m+1 \\ S_{1a} \end{pmatrix}\right\} \\ &= \Pr\left\{D_0(Y_t) = m+1, D_0(Y_{t-1}) = S_{1a} | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} = \Pr\left\{Y_t > (m+0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} \\ &= \Pr\left\{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \bar{X}_t > (m+0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} \\ &= \Pr\left\{\phi_1 S_{1a}\Delta + \phi_2 S_{2a}\Delta + \bar{X}_t > (m+0.5)\Delta\right\} \\ &= \Pr\left\{Z > \frac{(m + 0.5 - \phi_1 S_{1a} - \phi_2 S_{2a})\Delta\sqrt{n}}{\sigma_0}\right\}, \end{aligned} \quad \text{(D.8)}$$

and:

$$\begin{aligned} &\Pr\left\{a = \begin{pmatrix} S_{1a} \\ S_{2a} \end{pmatrix} \rightarrow \begin{pmatrix} -m-1 \\ S_{1a} \end{pmatrix}\right\} \\ &= \Pr\left\{D_0(Y_t) = -m-1, D_0(Y_{t-1}) = S_{1a} | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} = \Pr\left\{Y_t < (-m-0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} \\ &= \Pr\left\{\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \bar{X}_t < (-m-0.5)\Delta | Y_{t-1} = S_{1a}\Delta, Y_{t-2} = S_{2a}\Delta\right\} \\ &= \Pr\left\{\phi_1 S_{1a}\Delta + \phi_2 S_{2a}\Delta + \bar{X}_t < (-m-0.5)\Delta\right\} \\ &= \Pr\left\{Z < \frac{(-m - 0.5 - \phi_1 S_{1a} - \phi_2 S_{2a})\Delta\sqrt{n}}{\sigma_0}\right\}. \end{aligned} \quad \text{(D.9)}$$

The probability of transition of Eq. (D.4) is equal to 1.

Other transition probabilities except those shown in Eqs. (D.2) to (D.5) are equal to zero.

Appendix E

Computing in control ARL

Let $\mathbf{P}$ be the transition matrix whose computation procedure was explained in Appendix D. According to Chang and Wu [25], $\mathbf{P}$ could be written as:

$$\mathbf{P}_{(2m+2)(2m+2)} = \begin{pmatrix} \mathbf{A}_{(2m+1)\times(2m+1)} & \mathbf{B}_{(2m+1)\times 1} \\ \mathbf{0}'_{(2m+1)\times 1} & \mathbf{1} \end{pmatrix}, \quad \text{(E.1)}$$

where $\mathbf{A}_{(2m+1)(2m+1)}$ is the transition matrix of transient states, $\mathbf{B}_{(2m+1)\times 1}$ is a vector whose entries are the probability of process moving to absorbing state,

α. Let RL be the in-control run length of the proposed chart defined as follows:

$$RL = \inf\left\{t : D_0(\mathbf{z}_t) = \alpha\right\}. \tag{E.2}$$

Thus, the in-control ARL of the model can be defined as follows:

$$ARL = E[RL] = \sum_{t=0}^{+\infty} \Pr(RL > t). \tag{E.3}$$

According to Fu and Spring [26], Eq. (E.3) could be written as follows:

$$ARL = \boldsymbol{\pi}_0'(\mathbf{I} - \mathbf{A})^{-1}\mathbf{1}, \tag{E.4}$$

where $\boldsymbol{\pi}_0$ is a vector whose elements are the probability of $\mathbf{z}_0$ being in each state, $\mathbf{I}$ is the identity matrix, and $\mathbf{1}$ is the vector whose elements are all equal to one.

Biographies

Orod Ahmadi is a post-doctoral researcher at K. N. Toosi University of Technology. He presently is an Assistant Professor of Industrial Engineering at Kharazmi University. He received his PhD in Industrial Engineering from K. N. Toosi University of Technology in Tehran, Iran in 2015. His research interests include statistical quality control, applied multivariate statistics, robust statistics, time series analysis, signal processing, and data analysis.

Hamid Shahriari is a Professor at the Department of Industrial Engineering at K. N. Toosi University of Technology in Tehran, Iran. He received his PhD in Industrial Engineering at the Arizona State University in 1992. His research interests include quality control, applied multivariate statistics, engineering statistics, and data analysis. He is a member of IIIE, IITC, ISA, and IPA and a former member of ASA, SQC, and IIE.