JOURNAL OF APPLIED STATISTICS
2024, VOL. 51, NO. 5, 913–934
https://doi.org/10.1080/02664763.2023.2170336

Taylor & Francis
Taylor & Francis Group

Check for updates

Investigating zero-state and steady-state performance of MEWMA-CoDa control chart using variable sampling interval

Muhammad Imran [a], Jinsheng Sun[a], Xuelong Hu [b], Fatima Sehar Zaidi [c] and Anan Tang [b]

[a]Nanjing University of Science and Technology, Nanjing, People's Republic of China; [b]Nanjing University of Posts and Telecommunications, Nanjing, People's Republic of China; [c]Guangzhou University, Guangzhou, People's Republic of China

ABSTRACT

Traditional process monitoring control charts (CCs) focused on sampling methods using fixed sampling intervals (FSIs). The variable sampling intervals (VSIs) scheme is receiving increasing attention, in which the sampling interval (SI) length varies according to the process monitoring statistics. A shorter SI is considered when the process quality indicates the possibility of an out-of-control (OOC) situation; otherwise, a longer SI is preferred. The VSI multivariate exponentially moving average for compositional data (VSI-MEWMA CoDa) CC based on a coordinate representation using isometric log-ratio (ilr) transformation is proposed in this study. A methodology is proposed to obtain the optimal parameters by considering the zero-state (ZS) average time to signal (ZATS) and the steady-state (*SS*) average time to signal (SATS). The statistical performance of the proposed CC is evaluated based on a continuous-time Markov chain (CTMC) method for both cases, the ZS and the *SS* using a fixed value of in-control (IC) ATS_0. Simulation results demonstrate that the VSI-MEWMA CoDa CC has significantly decreased the OOC average time to signal (ATS) than the FSI MEWMA CoDa CC. Moreover, it is found that the number of variables (d) has a negative impact on the ATS of the VSI-MEWMA CoDa CC, and the subgroup size (n) has a mildly positive impact on the ATS of the VSI-MEWMA CoDa CC. At the same time, the SATS of the VSI-MEWMA CoDa CC is less than the ZATS of the VSI-MEWMA CoDa CC for all the values of n and d. The proposed VSI-MEWMA CoDa CC under steady-State performs effectively compared to its competitors, such as the FSI-MEWMA CoDa CC, the VSI-T^2 CoDa CC and the FSI-T^2 CoDa CC. An example of an industrial problem from a plant in Europe is also given to study the statistical significance of the VSI-MEWMA CoDa CC.

ARTICLE HISTORY
Received 20 July 2022
Accepted 15 January 2023

KEYWORDS
Average time to signal; compositional data; steady-state; variable sampling interval; zero-state

1. Introduction

Monitoring manufacturing processes has become increasingly difficult due to sophisticated consumer demands for quality products. Statistical process monitoring (SPM) is a commonly used statistical approach for quality control in the industrial scenario. Control

CONTACT Jinsheng Sun jssun67@163.com

charts (CCs) are the most widely used tool in SPM. W.A. Shewhart introduced the concept of CCs in 1924. Conventional CCs are simple and sensitive to large process variations but have poor sensitivity to small process variations. Several measurements are required to improve the detection speed for small process shift.

Aitchison defines Compositional data (CoDa) analysis as an adequate geometry model for the transformation of CoDa [1]. Since Karl Pearson first emphasized problems in the analysis of CoDa in 1897. CoDa has unique numerical characteristics that have significant implications for statistical analysis studied by many researchers (cf. [3,11,37]). Because CoDa represents parts of a larger whole, they have unique properties. The standard statistical techniques designed for probabilistic random variables that cannot analyze CoDa in raw form are studied (cf. [15]). In recent years, researchers have started focusing on CCs for CoDa. The First CC for CoDa was a Chi-square CC, assuming CoDa follows the properties of Dirichlet distribution. After the $d = 3$-part CoDa was analyzed using Hotelling T^2 CC to interpret the out-of-control (OOC) signals [59]. The Hotelling T^2 CC can also be applied on CoDa after deleting one component from the CoDa vector or after applying the isometric log-ratio (ilr) transformation, but the one with transformed values outperforms the other [58]. As these methods deal with $d = 3$-parts CoDa, a method to deal with high dimensional CoDa is introduced [60]. After the advancement of Hotelling T^2 CC for CoDa, multivariate exponentially moving average (MEWMA) CoDa CC using ilr transformation [56] and the effect of measurement error on Hotelling T^2 CC [62] and MEWMA [63] have been evaluated. The multivariate cumulative sum (MCUSUM) CC for CoDa has been studied with parameter estimation [17]. Recently, MEWMA CC for CoDa using variable sampling interval (VSI) has been studied using zero-state (ZS) average time to signal for ilr transformed $d = 3$-part CoDa using $n = 1$ subgroup size [35].

A VSI strategy reduces the detecting time in CCs. A small sampling interval (SI) is used if there is any signal that the process has changed; if there is no signal, a longer SI is used. The fixed sampling interval (FSI) CC is used when the SI length stays the same through all the samples (please see [64]). The multivariate CC to monitor the mean vector and variance-covariance matrix with VSI was investigated by Reynolds and Cho [43]. The MEWMA and MEWMA-type CCs were combined to get the best performance of the CC. The variable sampling rate (VSR) scheme has been used to study the increase and decrease in process dispersion in inverse normal transformation [50]. Further, the VSI CC to monitor the coefficient of variation has been introduced [7]. The CCs with double warning lines are faster at detecting small shifts in the mean vector [22]. The CC with VSI and variable sample size (VSS) was used to monitor the variance-covariance matrix of a multivariate normally distributed process [23].

Many researchers [4,25,41,61] examined the VSI and the FSI features for univariate and multivariate CCs for process monitoring. Simulation is used to investigate the average run length (ARL) properties of the exponentially weighted moving average (EWMA) CC to effectively detect the small changes in the process's desired value [52]. More recently, the ARL performance of EWMA techniques based on the VSI for the monitoring logistic profiles has been proposed [31]. The CUSUM CCs are found to be an effective method for monitoring changes in aquatic toxicity [16]. The robust measures of the location were applied to improve exponential-cum-ratio estimators [14]. The multivariate EWMA (MEWMA) CCs using unequal sample sizes were studied [20]. Improving the multivariate CUSUM and EWMA CCs for monitoring purposes has focused most research on quality control proposed by Jarrett [18]. Further, MCUSUM CCs using VSI were used to monitor the ratio when more than two mixture components were considered [33].

The performance of the MEWMA CC was evaluated using a continuous-time Markov chain (CTMC) method [39]. Several researchers have attempted to improve the MEWMA CC's efficiency in identifying shift patterns in the process mean vector through various methods. For example, the MEWMA CC using sequential sampling [20] and the MEWMA CC using unequal sample sizes [44]. Changing the SI value in response to process data is a frequently used technique for increasing the efficiency of CCs [57]. The VSI scheme has been amalgamated to study the performance of $\bar{X}$ CCs [9], the CUSUM CCs for monitoring process mean [42], double EWMA CCs [49], Hotelling T^2 CC for exponentially distributed random variables [12], multivariate Shewhart and MEWMA-type CCs for simultaneously monitoring vectors of means and standard deviations matrix [13]. Further, two SIs were used for designing the optimal process for Taguchi's online monitor and control method with and without misclassification errors [5]. The performance of the VSI MEWMA CC was investigated using a proposed CTMC approach by Sabahno et al. [48]. To study the benefits of using VSI scheme, the VSI and the FSI MEWMA CC's performance has been compared [34].

In SPM literature, two different types of performance are typically considered: ZS and steady-state (*SS*). The term '*SS* performance' shows the time required for the CC to identify a process shift for control statistics to reach a static distribution. Some processes are initially uncontrollable; the procedure is initiated under control in most realistic scenarios and then changes randomly [40]. When the average number of samples is taken from the start of signal monitoring in an OOC situation, then the CC's performance is evaluated using ZS ARL (cf. [8,10]). The comparison between the zero-state average time to signal (ZATS) symmetric and asymmetric distributions to the steady-state average time to signal (SATS) using the CTMC showed the SATS performed better in terms of ARL [19]. The CTMC method is also used to determine the *SS* ARL of the CC [21]. To detect changes in the process mean, the *SS* properties of synthetic CCs have been examined [10]. The CCs give more significant results by using *SS* ARL to create a CC with m-of-m run rules [24]. Numerous researchers have distinguished between the *SS* optimal and ZS optimal VSI schemes (cf. [53,54]). The CUSUM CC for two possible SIs and probability ratio tests were used to study the *SS*-optimized VSI methods [55]. The VSI-based CC scheme is superior to the FSI-based CC scheme in terms of average time to signal (ATS) performance [29].

As discussed earlier, many researchers are currently working on CC for CoDa, but all the above-mentioned studies deal with the Markov chain model with ZS ATS to study the CC performance for CoDa. Also, most of the research on CoDa deals with $d = 3$-part CoDa. As far as the author knows, till now, the *SS* ATS performance has not been used for monitoring CoDa. The literature shows the *SS* ATS performs better than the ZS ATS (cf. [6,30,51]). Hence to fill this gap, this paper makes an attempt to take *SS* ATS performance into account. The VSI-MEWMA CoDa CC has been proposed using ilr transformation to investigate the ATS using the different number of variables p (i.e. $d = 3, 5$), subgroup sizes n (i.e. $n = 1, 3$) and the VSI h (i.e. h_1, h_2). The ZS ATS has also been computed to study the difference between ZS and *SS* ATS performances.

The rest of this paper is as follows: Section 2 discusses brief details about how to model and manipulate the CoDa. In Section 3, the model for VSI-MEWMA CC for CoDa has been presented. Section 4 presents the CTMC for both ZS and *SS* for the VSI and the

FSI. Section 5 gives the CCs performance and compares the VSI-MEWMA CoDa and the FSI-MEWMA CoDa CC. Finally, an illustrative example and conclusions are presented in Sections 6 and 7.

2. Compositional data

A row vector is defined as a CoDa vector if it belongs to simplex space $\mathcal{S}^d$,

$$\mathcal{S}^d = \left\{ \mathbf{y} = (y_1, y_2, \ldots, y_d) | y_i > 0,\ i = 1, 2, \ldots, d \text{ such that } \sum_{i=1}^{d} y_i = \kappa \right\}, \tag{1}$$

where κ is a constant sum of the CoDa vector. Because of the constraint of constant sum, Euclidean geometry is unsuitable for CoDa. To overcome this problem, J. Aitchison proposed a specific geometry known as Aitchison's geometry [2]. In which advanced operators for sum and multiplications have been defined,

- the *perturbation* operator for the sum of CoDa vectors,

$$\mathbf{y} \oplus \mathbf{z} = \mathcal{C}(y_1 z_1, y_2 z_2, \ldots, y_p z_d), \tag{2}$$

- the *powering* operator for multiplication of CoDa vector with a constant,

$$c \odot \mathbf{y} = \mathcal{C}(y_1^c, y_2^c, \ldots, y_d^c). \tag{3}$$

To overcome the constant sum constraints, CoDa can be transformed from simplex sample space $\mathcal{S}^d$ to real space $\mathbb{R}^{d-1}$ using the predefined transformations,

- *Centered log-ratio* transformation,

$$\text{clr}(\mathbf{y}) = \left(\ln \frac{y_1}{\bar{y}_G}, \ln \frac{y_2}{\bar{y}_G}, \ldots, \ln \frac{y_d}{\bar{y}_G} \right), \tag{4}$$

  where $\bar{y}_G$ is the component-wise geometric mean of $\mathbf{y}$, i.e.

$$\bar{y}_G = \left(\prod_{i=1}^{p} y_i \right)^{\frac{1}{d}} = \exp\left(\frac{1}{d} \sum_{i=1}^{d} \ln y_i \right). \tag{5}$$

- *Isometric log-ratio*

$$\text{ilr}(\mathbf{y}) = \mathbf{y}^* = \text{clr}(\mathbf{y})\mathbf{B}^\mathsf{T}, \tag{6}$$

  where

$$B_{i,j} = \begin{cases} \sqrt{\dfrac{1}{(d-i)(d-i+1)}} & j \leq d-i \\ -\sqrt{\dfrac{d-i}{d-i+1}} & j = d-i+1 \\ 0 & j > d-i+1 \end{cases}.$$

To transform the vector from real space to simplex space, we use inverse isometric log-ratio,

$$\mathrm{ilr}^{-1}(\mathbf{y}^*) = \mathbf{y} = \mathcal{C}(\exp(\mathbf{y}^*\mathbf{B})). \tag{7}$$

There are two ways to deal with CoDa, one is to use CoDa as it is by using powering and perturbation operator, and the second way is to transform CoDa into real space by using the above-mentioned log-ratio transformations so that the classical methods can be applied to CoDa after making some important amendments. For more details about CoDa, the readers can refer to [17].

3. The VSI-MEWMA CoDa chart

Let us assume that there are have n measures $\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,n}$ of the quality characteristic $\mathbf{y}_i$ at the time $i = 1, 2, \ldots$. $\mathbf{y}_i \sim \mathrm{MNOR}_{\mathcal{S}^d}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$ when the process is IC, where $\boldsymbol{\mu}_0$ is the IC mean vector and $\mathbf{y}_i \sim \mathrm{MNOR}_{\mathcal{S}^d}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ when the process is OOC, where $\boldsymbol{\mu}_1$ is the OOC mean vector and $\boldsymbol{\Sigma}$ remain unchanged in both cases. Let $\bar{\mathbf{x}}_i^* = \mathrm{ilr}(\bar{\mathbf{x}}_i)$ and $\mathbf{y}_i^* \sim \mathrm{MNOR}_{\mathbb{R}^{d-1}}(\boldsymbol{\mu}_0^*, \boldsymbol{\Sigma}^*)$, where $\boldsymbol{\mu}_0^* = \mathrm{ilr}(\boldsymbol{\mu}_0)$ and $\boldsymbol{\Sigma}^* = \mathrm{ilr}(\boldsymbol{\Sigma})$. According to [36], x follows a multivariate normal distribution on S^d if the vector of random orthonormal coordinates, $x^* = ilr(x)$, follows a multivariate normal distribution on $\mathbb{R}^{D-1}$. This paper is an extension of [35] $\mathrm{VSI} - MEWMA$ CoDa CC, considering the SS ATS performance analysis. The VSI-MEWMA CoDa CC statistic is,

$$\mathbf{Q}_i = \mathbf{w}_i \boldsymbol{\Sigma}_{\mathbf{w}_i}^{-1} \mathbf{w}_i^{\mathsf{T}}, \tag{8}$$

with

$$\mathbf{w}_i = r(\bar{\mathbf{x}}_i^* - \mu_0^*) + (1 - r)\mathbf{w}_{i-1}, \tag{9}$$

where $\mathbf{w}_0 = \mathbf{0}$ and $r \in (0, 1]$ are smoothing parameters. The VSI-MEWMA CoDa CC shows a signal when

$$\mathbf{Q}_i = \mathbf{w}_i \boldsymbol{\Sigma}_{\mathbf{w}_i}^{-1} \mathbf{w}_i^{\mathsf{T}} > H, \tag{10}$$

where H is the upper control limit (UCL), and $\boldsymbol{\Sigma}_{\mathbf{w}_i}$ is the variance-covariance matrix of $\mathbf{w}_i$. Here the author used the asymptotic variance-covariance matrix proposed by Lowry et al. [27], i.e.

$$\boldsymbol{\Sigma}_{\mathbf{w}_i} = \frac{r}{n(2 - r)} \boldsymbol{\Sigma}^*. \tag{11}$$

Due to the directional invariant property, the MEWMA CC's ATS depend on the non-centrality parameter δ [26]. Where the value of δ is,

$$\delta = (\boldsymbol{\mu}_1^* - \boldsymbol{\mu}_0^*)(\boldsymbol{\Sigma}^*)^{-1}(\boldsymbol{\mu}_1^* - \boldsymbol{\mu}_0^*)^{\mathsf{T}}. \tag{12}$$

When the SI is fixed, the SI of FSI-MEWMA CoDa CC is denoted by h_0. But for VSI-MEWMA CoDa CC, the selection of SI is based on the charting statistics $\mathbf{Q}_i$. The time interval between the sample x_i and $x_i + 1$ can vary. Using two sampling intervals is reasonable to limit the VSI-MEWMA CoDa CC's complexity and achieve the proposed chart's efficacy [46]. Hence following [46], two SIs are used in this paper, h_1 and h_2, where h_2

denotes the small SI and h_1 denotes the long one. For the VSI-MEWMA CoDa CC, the UCL $= H$ is the same as that of FSI-MEWMA CoDa CC. A warning limit (L) is introduced such that $0 < L < H$ and $h_2 < h_0 < h_1$. The switch between the small and the long SI depends on the value of the CC parameter $\mathbf{Q}_i$. If the CC parameter $\mathbf{Q}_i$ lies within the L, a long SI h_1 will be used, and if the value of $\mathbf{Q}_i$ lies between the L and the H, then the small SI h_2 should be used.

4. The average time to signal

Prabhu and Runger [47] suggested calculating the statistics $q_i = \| Y_i \|_2$ as the standardized form of $Q_i = a \| Y_i \|_2^2$ with $a = \frac{2-r}{r}$ to determine the IC and OOC ATS of the MEWMA CC using CTMC models. For the IC case, one-dimensional CTMC (i.e. $[0, \mathrm{UCL}']$) is used to approximate the ATS of q_i, where $\mathrm{UCL}' = (H/a)^{1/2}$ have $m+1$ sub-interval with the length of the first sub-interval $g/2$ and others g and the width of sub-interval $g = 2\mathrm{UCL}/2m+1$. Concerning the VSI-MEWMA CoDa CCs $\mathrm{WL} = (L/a)^{1/2}$ is also added to the one-dimensional CTMC (i.e. $[0, \mathrm{WL}, \mathrm{UCL}']$). The IC one-dimensional CTMC is also shown in Figure 1.

The probability of transition for i to j state is,

$$\mathbf{P}_1(i,j) = \begin{cases} Pr\left(\chi^2(d-1,c) < \left(\frac{g}{2r}\right)^2\right) & \text{for } j = 0 \\ Pr\left(\left(\frac{(j-0.5)g}{r}\right)^2 < \chi^2(d-1,c) < \left(\frac{(j+0.5)g}{r}\right)^2\right) & \text{for } j = 1,2,\ldots,m \end{cases} \tag{13}$$

where $\mathbf{P}_1(i,j)$ follows a non-central chi-square distribution $\chi^2(d-1,c)$ with a non-centrality parameter $c = (\frac{(1-r)ig}{r})^2$ having $d-1$ degree of freedom.

The ATS of the VSI-MEWMA CoDa CCs for IC case is as follows,

$$\mathrm{ATS} = \mathrm{s}^{\mathsf{T}}(\mathbf{I}_{m+1} - \mathbf{P}_1)^{-1}\mathbf{h}_{m+1}, \tag{14}$$

with $\mathbf{I}_{m+1}$ is the identity matrix of size $m+1$ and $\mathrm{s} = (1,0,0,\ldots,0)^{\mathsf{T}}$ is the initial probability vector and $\mathbf{h}_{m+1}$ is vector of SI. The SI average for the proposed CC can be written as

$$\bar{h} = \frac{\mathrm{s}^{\mathsf{T}}(\mathbf{I}_{m+1} - \mathbf{P}_1)^{-1}\mathbf{h}_{m+1}}{\mathrm{s}^{\mathsf{T}}(\mathbf{I}_{m+1} - \mathbf{P}_1)^{-1}\mathbf{1}_{m+1}}. \tag{15}$$

For the OOC case, two-dimensional CTMC is used to approximate the ATS of q_i with the partition of $\boldsymbol{Y}_i \in \mathbb{R}^{d-1}$ into $Y_{i1} \in \mathbb{R}$ and $\boldsymbol{Y}_{i2} \in \mathbb{R}^{d-2}$ with δ and zero mean, respectively and $||\boldsymbol{Y}_i||_2 = (Y_{i1}^2 + \boldsymbol{Y}_{i2}^{\mathsf{T}}\boldsymbol{Y}_{i2})^{\frac{1}{2}}$. A two-dimensional CTMC can also be used for the MEWMA CoDa CC. The approach used to approximate the component Y_{i1}^2, and for $||\boldsymbol{Y}_{i2}||_2$ is given in [28]; the same method for IC CTMC is used where $d-1$ is replaced by $d-2$. For Y_{i1}, the OOC component is analyzed using the transition probability $u(i_1, j_1)$

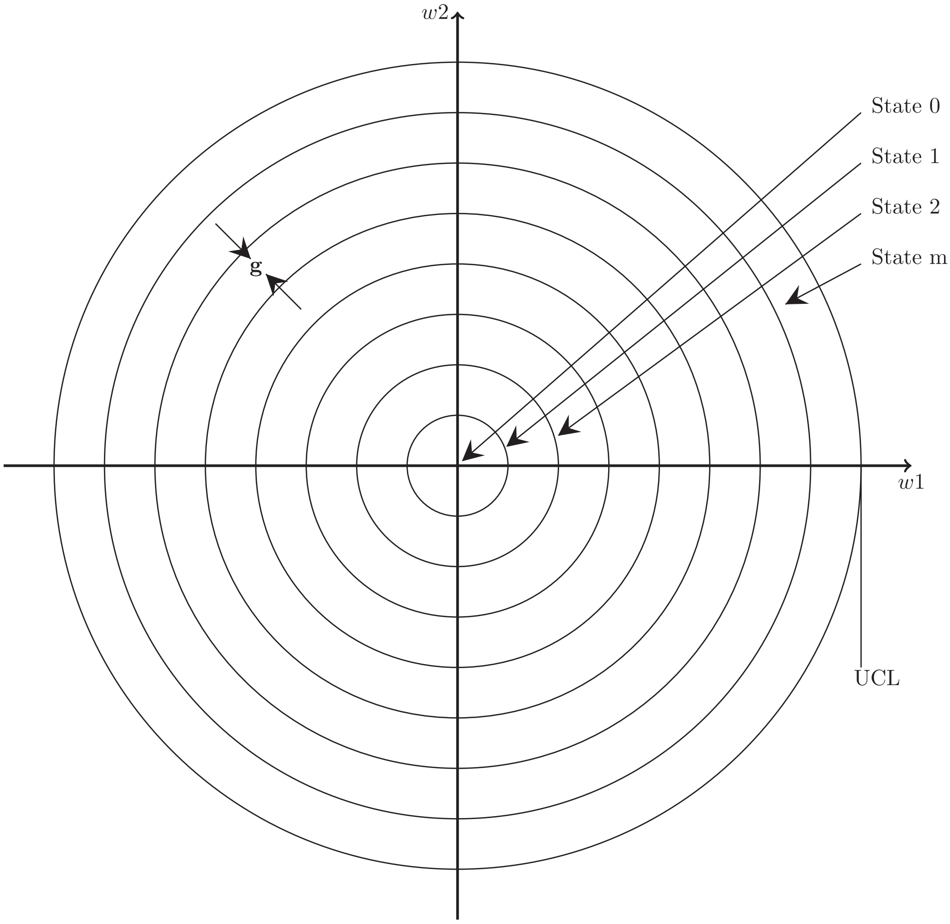

Figure 1. IC CTMC Distribution for the VSI MEWMA CoDa CC.

from state i_1 to state j_1 with $2m_1 + 1$ states,i.e.

$$u(i_1, j_1) = \Phi\left(\frac{-\mathrm{UCL}' + j_1 g_1 - (1-r)c_i}{r} - \delta\right) - \Phi\left(\frac{-\mathrm{UCL}' + (j_1 - 1)g_1 - (1-r)c_i}{r} - \delta\right), \tag{16}$$

where Φ is the cumulative standard normal distribution function with $c_i = -\mathrm{UCL}' + (i - 0.5)g_1$ being the midpoint of the state i and $g_1 = \frac{2\mathrm{UCL}'}{2m_1+1}$ be the width of each state. The OOC two-dimensional CTMC is also shown in Figure 2.

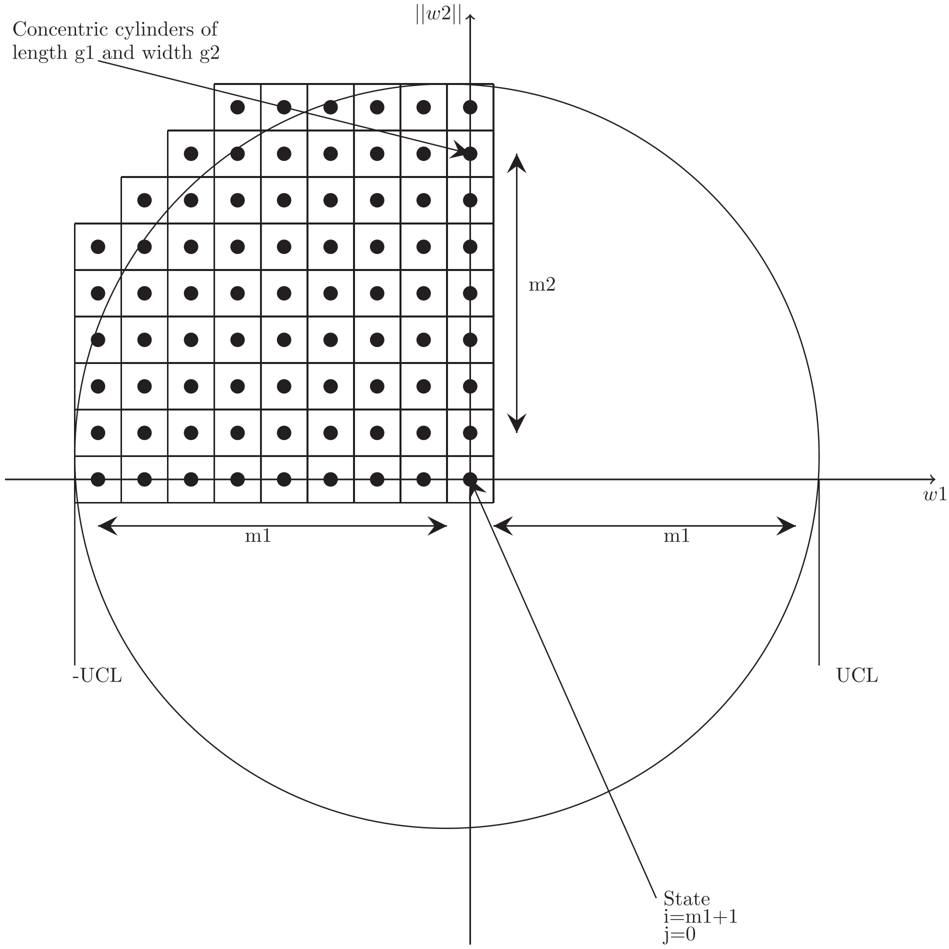

Figure 2. OOC CTMC Distribution for the VSI MEWMA CoDa CC.

For $\boldsymbol{Y}_{i2}$, the IC component is analyzed using the transition probability $v(i_2, j_2)$ from state i_2 to state j_2 with $m_2 + 1$ states. i.e.

$$v(i_2, j_2) = \left\{ \begin{array}{l} Pr\left(\chi^2(d-2, c) < \left(\frac{g_2}{2r}\right)^2\right) \quad \text{for } j_2 = 0 \\ Pr\left(\left(\frac{(j_2 - 0.5)g_2}{r}\right)^2 < \chi^2(d-2, c) < \left(\frac{(j_2 + 0.5)g_2}{r}\right)^2\right) \\ \qquad \text{for } j_2 = 1, 2, \ldots, m_2 \end{array} \right\}, \tag{17}$$

where $c = (\frac{(1-r)ig_2}{r})^2$ with width of states $g_2 = \frac{2\text{UCL}'}{2m_2+1}$. All the transient states of CTMC can be summarized in a transition probability matrix $\boldsymbol{Pr}$. Then,

$$\boldsymbol{Pr} = \boldsymbol{T}(i_1, i_2) \circledast \boldsymbol{P}_2, \tag{18}$$

where the symbol $\circledast$ is used for element-wise matrices multiplication, $\boldsymbol{T}$ is the $(2m_1+1)\times(m_2+1)$ dimensional matrix defined as

$$\boldsymbol{T}(i_1,i_2)=\begin{cases}1 & \text{if state } (\alpha,\beta) \text{ is a transient state}\\ 0 & \text{otherwise}\end{cases} \tag{19}$$

and $\boldsymbol{P}_2$ denotes the transition probability matrix of two-dimensional CTMC, $P_2=\boldsymbol{U}\otimes\boldsymbol{V}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the transitional probability matrices of Y_{i1} and $||\boldsymbol{Y}_{i2}||_2$ respectively and $\otimes$ is the Kronecker's matrices product. The ZS OOC ATS of the VSI-MEWMA CoDa CC is defined as,

$$\text{ZATS}=\text{s}^{\mathsf{T}}(\mathbf{I}_{m+1}-\boldsymbol{Pr})^{-1}\mathbf{h}_{m+1}. \tag{20}$$

with $\mathbf{I}_{m+1}$ is the identity matrix of size $m+1$, and s is the initial probability vector with all components equal to zero except the component corresponding to the state $(\alpha,\beta)=(m_1+1,0)$ which is equal to one and $\mathbf{h}_{m+1}$ is the vector of SI.

The *SS* OOC ATS of the VSI-MEWMA CoDa CC is defined as,

$$\text{SATS}=\text{w}^{\mathsf{T}}(\boldsymbol{I}-\boldsymbol{Pr})^{-1}\boldsymbol{s}. \tag{21}$$

where w is a $(2m+1,m+1)$ *SS* vector with $\text{w}_i=\frac{\boldsymbol{s}_i\boldsymbol{b}_i}{\boldsymbol{sb}}$ and $\boldsymbol{s}$ is $(2m+1,m+1)$ vector of SI with elements h_1 if $\sqrt{i_1-(m+1))^2 * g^2+i_2^2g^2}<\text{WL}$, h_2 if $\text{WL}<\sqrt{i_1-(m+1))^2 * g^2+i_2^2g^2}<\text{UCL}$ and zero when the process is OOC. Where b is a *SS* probability vector obtained by solving the following equation: $\boldsymbol{b}=\text{P}_1^{\mathsf{T}}\boldsymbol{b}$ subject to $1^{\mathsf{T}}\boldsymbol{b}=1$. Where P_1 is the transition probability matrix when the process is IC, i.e. $\delta=0$.

The average of SI for the VSI-MEWMA CoDa CC can be written as,

$$\bar{h}=p_1h_1+(1-p_1)h_2. \tag{22}$$

where p_1 is the proportion of time to signal. If $h_0=h_1=h_2$ the CC will be the standard MEWMA CoDa CC. The number of states greatly impacts the ATS of the CC. (see [32]), but due to limited resources and time, the author cannot use a large number of m. Following literature reviews, hence $m_1=m_2=30$ will be used. (see [32] or [56]).

5. Comparative analysis of the VSI-MEWMA CoDa chart

This section presents an optimization approach for statistical designing the VSI-MEWMA CoDa CC. An optimal VSI CC can be achieved using two different SIs, with the small SI h_2 taken as small as possible and the long SI h_1 dependent on δ and h_2, where the CC is best for tracking shifts δ. Similar to [38,45], the practitioners need to set the h_2 fixed for the minimum interval hmin. The VSI-MEWMA CoDa CC is designed by determining the CC parameters (i.e. r, h_1, W, and H) that minimize the OOC ATS aspect to the target value specified for constraints h, $h_2 = hmin$, and ATS_0, for the provided values of n, d, and δ. The value of H will be the same for the FSI and the VSI-MEWMA-CoDa CCs for given values of r, n, h, and ATS_0. The following is the optimized statistical layout process for the VSI-MEWMA CoDa CC:

- Specify n, ATS_0, d, $h2$, $\bar{h}$ and δ.

Table 1. ZS optimum charting parameters for VSI-MEWMA CoDa and FSI-MEWMA CoDa CC.

		$n = 1$		$n = 3$	
		VSI-MEWMA CoDa	FSI-MEWMA CoDa	VSI-MEWMA CoDa	FSI-MEWMA CoDa
δ	p	$(r, H, W, h1, h2)$	(r, H)	$(r, H, W, h1, h2)$	(r, H)
0.25	3	(0.05, 7.35, 1.74, 1.62, 0.10)	(0.05, 7.35)	(0.05, 7.35, 0.82, 2.70, 0.10)	(0.05, 7.35)
	5	(0.05, 9.46, 2.27, 2.10, 0.10)	(0.05, 9.46)	(0.05, 9.46, 2.30, 2.10, 0.10)	(0.05, 9.46)
0.50	3	(0.05, 7.50, 1.72, 1.65, 0.10)	(0.06, 7.50)	(0.05, 7.50, 1.24, 2.50, 0.10)	(0.06, 7.50)
	5	(0.05, 9.50, 1.73, 2.28, 0.10)	(0.06, 9.50)	(0.05, 9.50, 1.78, 2.28, 0.10)	(0.06, 9.50)
0.75	3	(0.09, 8.51, 1.64, 1.73, 0.10)	(0.09, 8.51)	(0.10, 8.51, 1.48, 2.20, 0.10)	(0.10, 8.51)
	5	(0.08, 10.27, 1.69, 2.09, 0.10)	(0.08, 10.27)	(0.09, 10.27, 1.76, 2.08, 0.10)	(0.08, 10.27)
1.00	3	(0.14, 9.19, 2.90, 1.32, 0.10)	(0.14, 9.19)	(0.15, 9.41, 1.88, 1.89, 0.10)	(0.15, 9.41)
	5	(0.13, 10.71, 1.95, 1.83, 0.10)	(0.13, 10.71)	(0.14, 10.93, 2.05, 1.82, 0.10)	(0.13, 10.93)
1.25	3	(0.19, 9.61, 1.65, 1.82, 0.10)	(0.20, 9.61)	(0.21, 9.77, 2.45, 1.69, 0.10)	(0.20, 9.77)
	5	(0.18, 10.97, 2.45, 1.65, 0.10)	(0.19, 10.97)	(0.20, 11.13, 2.58, 1.64, 0.10)	(0.19, 11.13)
1.50	3	(0.25, 9.91, 3.51, 1.22, 0.10)	(0.25, 9.91)	(0.27, 10.05, 2.86, 1.49, 0.10)	(0.25, 10.05)
	5	(0.24, 11.11, 2.81, 1.46, 0.10)	(0.25, 11.11)	(0.26, 11.25, 2.96, 1.45, 0.10)	(0.25, 11.25)
1.75	3	(0.31, 10.12, 3.72, 1.24, 0.10)	(0.31, 10.12)	(0.34, 10.23, 4.01, 1.28, 0.10)	(0.31, 10.23)
	5	(0.31, 11.16, 3.92, 1.26, 0.10)	(0.31, 11.16)	(0.33, 11.27, 4.09, 1.25, 0.10)	(0.31, 11.27)
2.00	3	(0.37, 10.26, 3.62, 1.25, 0.10)	(0.38, 10.26)	(0.41, 10.36, 4.70, 1.10, 0.10)	(0.38, 10.36)
	5	(0.37, 11.14, 4.57, 1.10, 0.10)	(0.37, 11.14)	(0.40, 11.24, 4.77, 1.10, 0.10)	(0.37, 11.24)

- Initialize r as 0.05.
- Initialize h_1 as $\bar{h} + 0.1$. Set $h_2 = 0.1$, then find the value of H for the fixed value of IC ATS_0, where $h_2 < \bar{h} < h_1$. Given δ is calculated, the OOC ZATS and SATS. Increasing r with a step size of 0.005, iterate Steps 3 to 5.
- The r, h_1, W, and H values are used to determine the minimum OOC ZATS and SATS for the optimal VSI-MEWMA CoDa CC parameters.

For comparison of the VSI with the FSI CC, the average SI $\bar{h}$ of the VSI-MEWMA CoDa CC is assumed to be the same as the h_0 of the FSI-MEWMA CoDa CC, i.e. when the process is IC when $\bar{h} = 1$. In other words, SI for the VSI-MEWMA CoDa CC is chosen to have a similar IC ATS as the FSI-MEWMA CoDa CC; in the specific context, the VSI-MEWMA CoDa CC's false alarm rate (i.e. $\text{ATS}_0 \approx 200$) is the same as the FSI-MEWMA CoDa CC.

5.1. ZATS of the VSI-MEWMA CoDa control chart

The values of optimal couples of the VSI and the FSI MEWMA CoDa CCs under ZS are presented in Table 1. The OOC ATS values of the VSI and the FSI MEWMA CoDa CCs for ZS ATS are given in Table 2 when $d \in \{3, 5\}$ and $n \in \{1, 3\}$. The OOC ATS values of the VSI and the FSI T^2 CoDa CCs for the ZS are also given in Table 2.

5.1.1. Impact of sampling interval h

Based on Table 2, it can be seen that the ZATS of the VSI CC is less than the ZATS of the FSI CC. When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$ and $W = 1.78$, the ZATS for the FSI-T^2 CoDa CC is ZATS $= 41.916$, while for the VSI-T^2 CoDa CC is ZATS $=$ 22.986. Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$, $r = 0.14$, $H = 9.19$ and $W = 1.78$, the ZATS for the FSI-MEWMA CoDa CC is ZATS $= 6.98$, while for the VSI-MEWMA CoDa CC is ZATS $= 9.90$. Hence it is summarized that the VSI CCs have a greater degree of efficacy than the FSI CCs for CoDa.

Table 2. OOC ZATS of the VSI-MEWMA CoDa CC.

		$n = 1$				$n = 3$			
		MEWMA CoDa		T^2 CoDa		MEWMA CoDa		T^2 CoDa	
δ	p	VSI	FSI	VSI	FSI	VSI	FSI	VSI	FSI
0.25	3	56.842	64.6	115.528	88.553	48.346	53.980	112.877	85.902
	5	64.362	70.27	120.928	99.909	53.746	59.654	118.277	97.258
0.50	3	19.935	26.4	76.86	51.885	17.731	22.440	75.900	50.925
	5	26.695	31.65	81.86	59.076	22.731	27.686	80.900	58.116
0.75	3	10.441	15.1	55.323	33.367	9.095	12.900	54.784	32.828
	5	15.69	19.72	59.723	38.901	13.495	17.525	59.184	38.362
1.00	3	6.913	9.9	41.916	22.986	5.489	8.440	41.559	22.629
	5	10.748	13.89	45.716	27.482	9.289	12.431	45.359	27.125
1.25	3	4.934	7.1	32.942	16.039	3.885	6.050	32.687	15.784
	5	8.333	10.67	36.342	19.932	7.285	9.623	36.087	19.677
1.50	3	3.728	5.4	26.618	11.744	3.185	4.610	26.425	11.551
	5	6.978	8.55	29.618	15.142	6.185	7.757	29.425	14.948
1.75	3	3.008	4.3	21.984	9.14	2.852	3.680	21.832	8.989
	5	6.075	7.03	24.584	12.088	5.452	6.407	24.432	11.936
2.00	3	2.419	3.5	18.484	7.484	2.246	3.000	18.362	7.362
	5	4.95	5.81	20.684	9.684	4.446	5.306	20.562	9.562

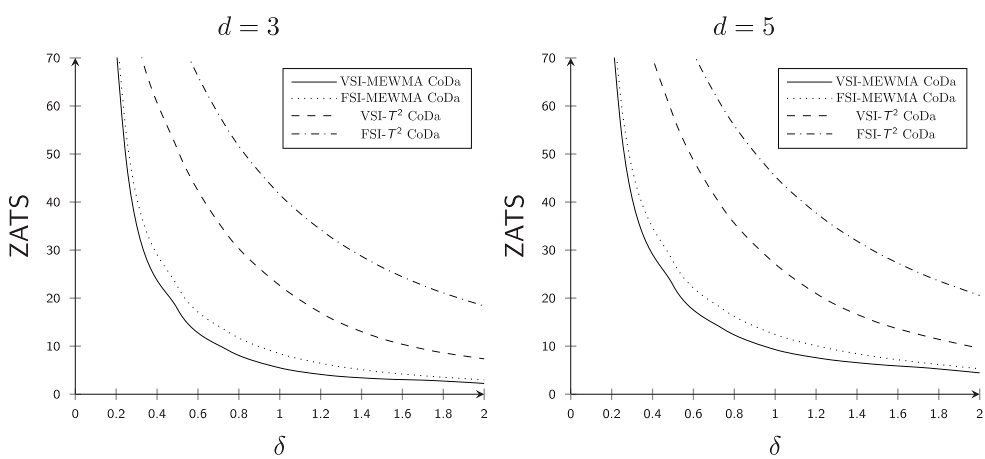

Figure 3. ZATS Curves for $n = 3$ and $d \in \{3, 5\}$.

5.1.2. Impact of number of the variables d

Based on Table 2 and Figure 3, it can be seen that d has a negative effect on the *ZS* ATS of the CC for CoDa; that is, the OOC ZATS values increase with an increase in the value of d.

When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$ and $W = 1.78$, the ZATS for the FSI-T^2 CoDa CC is ZATS $= 41.916$ and for the VSI-T^2 CoDa CC is ZATS $= 22.986$. But when the value of d increases to $d = 5$, the ZATS for the FSI-T^2 CoDa CC increases to ZATS $=$ 45.716 and the VSI-T^2 CoDa CC increases to ZATS $= 27.482$.

Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$, $r = 0.14$, $H = 9.19$ and $W = 1.78$, the ZATS for the FSI-MEWMA CoDa CC is ZATS $= 6.98$, and for the VSI-MEWMA CoDa CC is ZATS $= 9.90$. But, when the value of d increases to $d = 5$, the ZATS for the FSI-MEWMA CoDa CC increases to ZATS $= 10.748$ and the VSI-MEWMA CoDa

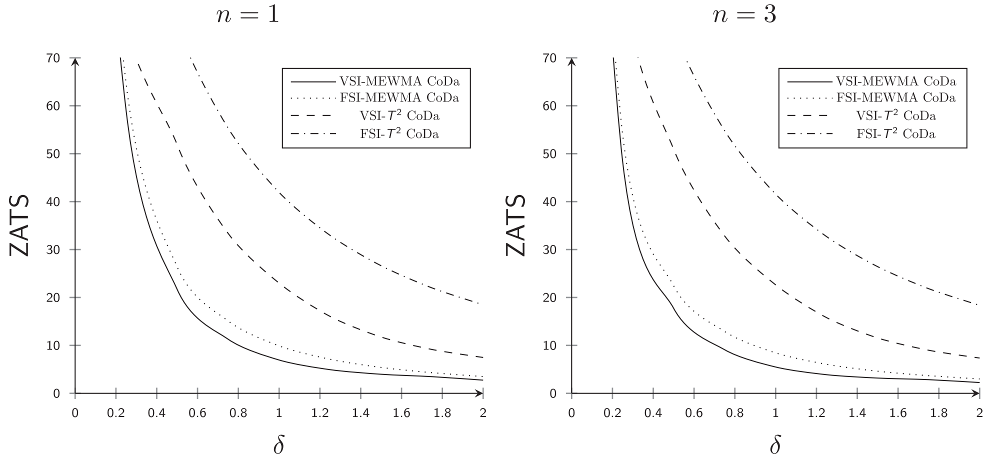

Figure 4. ZATS Curves for $d = 3$ and $n \in \{1, 3\}$.

CC increases to ZATS = 13.89. Figure 3 also shows the impact of the number of variables d on ZATS of the T^2 CoDa and the MEWMA CoDa CC for both the VSI and the FSI situations.

Where a solid black line shows the ZATS of the VSI-MEWMA CoDa CC, the ZATS of the FSI-MEWMA CoDa CC is shown by the dotted line, and the dashed line shows the ZATS of the VSI-T^2 CoDa CC, the dashed-dotted line shows the ZATS of the FSI-T^2 CoDa CC. From Figure 3, it is also clearly visible that d has a negative effect on the ATS of the VSI and the FSI CC for CoDa.

5.1.3. Impact of subgroup size n

Based on Table 2, it can be seen that n has a mild positive effect on the ATS of the CC for CoDa; that is, the OOC ZATS values decrease with an increase in the value of n.

When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$ and $W = 1.78$, the ZATS for the FSI-T^2 CoDa CC is ZATS = 41.916, and for the VSI-T^2 CoDa CC is ZATS = 22.986. But when the value of n increases to $n = 3$, the ZATS for the FSI-T^2 CoDa CC decreases to ZATS = 41.559, and the VSI-T^2 CoDa CC decreases to ZATS = 22.629.

Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.90$, $h_2 = 0.1$, $r = 0.14$, $H = 9.19$ and $W = 1.78$, the ZATS for the FSI-MEWMA CoDa CC is ZATS = 6.98 and for the VSI-MEWMA CoDa CC is ZATS = 9.90. But when the value of n increases to $n = 3$, the ZATS for the FSI-MEWMA CoDa CC decreases to ZATS = 5.489, and the VSI-MEWMA CoDa CC decreases to ZATS = 8.44. Figure 4 also shows the impact of the subgroup size n on ZATS of the Hotelling T^2 CoDa and the MEWMA CoDa CC for both the VSI and the FSI situations. From Figure 4, it is also clearly visible that n has a positive effect on the ATS of the VSI and the FSI CC for CoDa.

5.2. SATS of the VSI-MEWMA CoDa control chart

The values of optimal couples of the VSI and the FSI MEWMA CoDa CCs under *SS* are presented in Table 3. The OOC ATS values of the VSI and the FSI MEWMA CoDa CCs

Table 3. *SS* optimum charting parameters for VSI-MEWMA CoDa and FSI-MEWMA CoDa CC.

		$n = 1$		$n = 3$	
		VSI-MEWMA CoDa	FSI-MEWMA CoDa	VSI-MEWMA CoDa	FSI-MEWMA CoDa
δ	p	$(r, H, W, h1, h2)$	(r, H)	$(r, H, W, h1, h2)$	(r, H)
0.25	3	(0.05, 7.38, 0.80, 2.50, 0.10)	(0.05, 7.38)	(0.05, 7.38, 0.83, 2.50, 0.10)	(0.05, 7.38)
	5	(0.05, 9.33, 2.25, 1.92, 0.10)	(0.05, 9.33)	(0.05, 9.33, 2.30, 1.91, 0.10)	(0.05, 9.33)
0.50	3	(0.05, 7.53, 1.19, 2.30, 0.10)	(0.05, 7.53)	(0.05, 7.53, 1.24, 2.30, 0.10)	(0.05, 7.53)
	5	(0.05, 9.37, 1.71, 2.09, 0.10)	(0.05, 9.37)	(0.05, 9.37, 1.78, 2.09, 0.10)	(0.05, 9.37)
0.75	3	(0.09, 8.54, 1.42, 2.00, 0.10)	(0.10, 8.54)	(0.10, 8.54, 1.49, 2.00, 0.10)	(0.10, 8.54)
	5	(0.08, 10.14, 1.68, 1.89, 0.10)	(0.09, 10.14)	(0.09, 10.14, 1.77, 1.89, 0.10)	(0.08, 10.14)
1.00	3	(0.14, 9.22, 1.80, 1.70, 0.10)	(0.15, 9.22)	(0.15, 9.44, 1.90, 1.69, 0.10)	(0.15, 9.44)
	5	(0.13, 10.58, 1.95, 1.64, 0.10)	(0.14, 10.58)	(0.14, 10.80, 2.07, 1.63, 0.10)	(0.14, 10.80)
1.25	3	(0.19, 9.64, 2.35, 1.50, 0.10)	(0.20, 9.64)	(0.21, 9.80, 2.47, 1.49, 0.10)	(0.20, 9.80)
	5	(0.19, 10.84, 2.45, 1.46, 0.10)	(0.19, 10.84)	(0.20, 11.00, 2.59, 1.45, 0.10)	(0.19, 11.00)
1.50	3	(0.25, 9.94, 2.74, 1.30, 0.10)	(0.26, 9.94)	(0.27, 10.08, 2.89, 1.29, 0.10)	(0.26, 10.08)
	5	(0.25, 10.98, 2.82, 1.27, 0.10)	(0.25, 10.98)	(0.27, 11.12, 2.99, 1.25, 0.10)	(0.25, 11.12)
1.75	3	(0.31, 10.15, 3.86, 1.10, 0.10)	(0.32, 10.15)	(0.34, 10.26, 4.04, 1.10, 0.10)	(0.32, 10.26)
	5	(0.31, 11.03, 3.94, 1.07, 0.10)	(0.31, 11.03)	(0.33, 11.14, 4.12, 1.10, 0.10)	(0.31, 11.14)
2.00	3	(0.38, 10.29, 4.54, 1.10, 0.10)	(0.38, 10.29)	(0.41, 10.39, 4.74, 1.10, 0.10)	(0.38, 10.39)
	5	(0.37, 11.17, 4.60, 1.08, 0.10)	(0.37, 11.17)	(0.40, 11.27, 4.81, 1.10, 0.10)	(0.37, 11.27)

Table 4. OOC SATS of the VSI-MEWMA CoDa CC.

		$n = 1$				$n = 3$			
		MEWMA CoDa		T^2 CoDa		MEWMA CoDa		T^2 CoDa	
δ	p	VSI	FSI	VSI	FSI	VSI	FSI	VSI	FSI
0.25	3	57.882	63.350	114.028	87.804	47.267	52.740	111.378	85.155
	5	62.882	68.600	119.028	99.900	52.905	58.698	117.129	97.044
0.50	3	20.695	25.250	75.480	51.356	16.732	21.290	74.521	50.398
	5	25.295	30.080	80.080	59.195	21.819	26.642	79.647	57.718
0.75	3	10.410	14.100	54.123	33.169	8.216	11.910	53.586	32.633
	5	14.410	18.300	58.123	39.213	12.662	16.579	58.049	38.194
1.00	3	6.188	9.050	40.896	23.119	4.732	7.590	40.543	22.766
	5	9.588	12.620	44.296	27.986	8.560	11.616	44.380	27.255
1.25	3	4.253	6.350	32.042	16.394	3.209	5.310	31.792	16.144
	5	7.253	9.500	35.042	20.564	6.629	8.899	35.218	20.016
1.50	3	3.378	4.750	25.838	12.320	2.590	3.960	25.652	12.135
	5	5.978	7.480	28.438	15.901	5.606	7.131	28.672	15.509
1.75	3	2.955	3.750	21.324	9.774	2.338	3.130	21.172	9.622
	5	5.155	6.060	23.524	12.524	4.952	5.857	23.772	12.222
2.00	3	2.310	2.950	17.824	6.274	1.806	2.450	17.702	6.152
	5	4.510	5.260	20.024	9.024	4.006	4.756	19.902	8.352

for *SS* ATS are given in Table 4 when $d \in \{3, 5\}$ and $n \in \{1, 3\}$. The OOC ATS values of the VSI and the FSI T^2 CoDa CCs for the *SS* are also given in Table 4.

5.2.1. Impact of sampling interval h

Based on Table 4, it can be seen that the SATS of the VSI CC is less than the SATS of the FSI CC. When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$ and $W = 1.80$, the SATS for the FSI-T^2 CoDa CC is SATS $= 40.896$, while for the VSI-T^2 CoDa CC is SATS $=$ 23.119. Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$, $r = 0.14$, $H = 9.22$ and

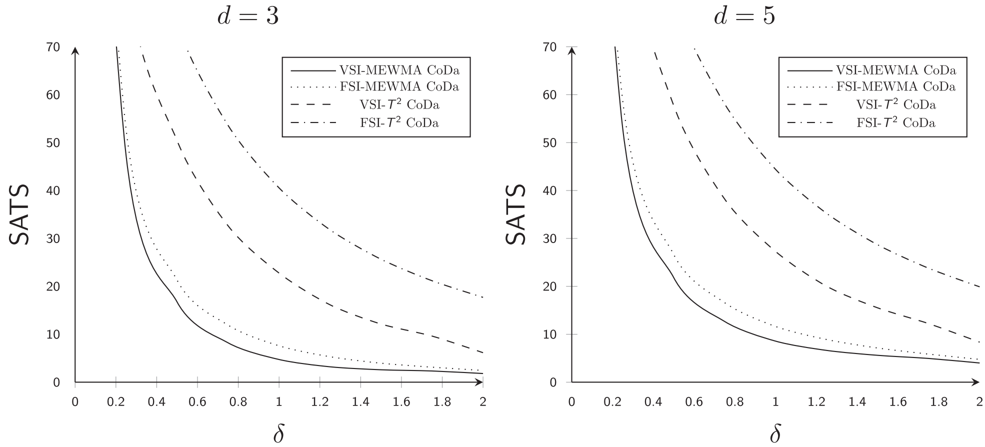

Figure 5. SATS Curves for $n = 3$ and $d \in \{3, 5\}$.

$W = 1.80$, the SATS for the FSI MEWMA CoDa CC is SATS $= 9.05$, while for the VSI-MEWMA CoDa CC is SATS $= 6.188$. Hence it is summarized that the VSI CCs have a greater degree of efficacy than the FSI CCs for CoDa.

5.2.2. Impact of number of the variables d

Based on Table 4, it can be seen that d has a negative effect on the ATS of the CC for CoDa; that is, the OOC SATS values increase with an increase in the value of d.

When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$ and $W = 1.80$, the SATS for the FSI-T^2 CoDa CC is SATS $= 40.896$ and for the VSI-T^2 CoDa CC is SATS $= 23.119$. But when the value of d increases to $d = 5$, the SATS for the FSI-T^2 CoDa CC increases to SATS $=$ 44.296 and the VSI-T^2 CoDa CC increases to SATS $= 27.986$.

Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$, $r = 0.14$, $H = 9.22$ and $W = 1.80$, the SATS for the FSI-MEWMA CoDa CC is SATS $= 9.05$ and for the VSI-MEWMA CoDa CC is SATS $= 6.188$. But when the value of d increases to $d = 5$, the SATS for the FSI-MEWMA CoDa CC increases to SATS $= 12.62$, and the VSI-MEWMA CoDa CC increases to SATS $= 9.58$. Figure 5 also shows the impact of the number of variables d on SATS of the MEWMA CoDa CC. From Figure 5, it is also clearly visible that d has a negative effect on the ATS of the VSI and the FSI CC for CoDa.

5.2.3. Impact of subgroup size n

Based on Table 4, it can be seen that n has a mild positive effect on the ATS of the CC for CoDa; that is, the OOC SATS values decrease with an increase in the value of n.

When $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$ and $W = 1.80$, the SATS for the FSI-T^2 CoDa CC is SATS $= 40.896$ and for the VSI-T^2 CoDa CC is SATS $= 23.119$. But when the value of n increases to $n = 3$, the SATS for the FSI-T^2 CoDa CC decreases to SATS $=$ 40.543, and the VSI-T^2 CoDa CC decreases to SATS $= 22.766$.

Similarly, when $\delta = 1$, $n = 1$, $d = 3$, $h_1 = 1.70$, $h_2 = 0.1$, $r = 0.14$, $H = 9.22$ and $W = 1.80$, the SATS for the FSI-MEWMA CoDa CC is SATS $= 9.05$ and for the VSI-MEWMA CoDa CC is SATS $= 6.188$. But when the value of n increases to $n = 3$, the SATS for the FSI-MEWMA CoDa CC decreases to SATS $= 7.59$, and the VSI-MEWMA CoDa

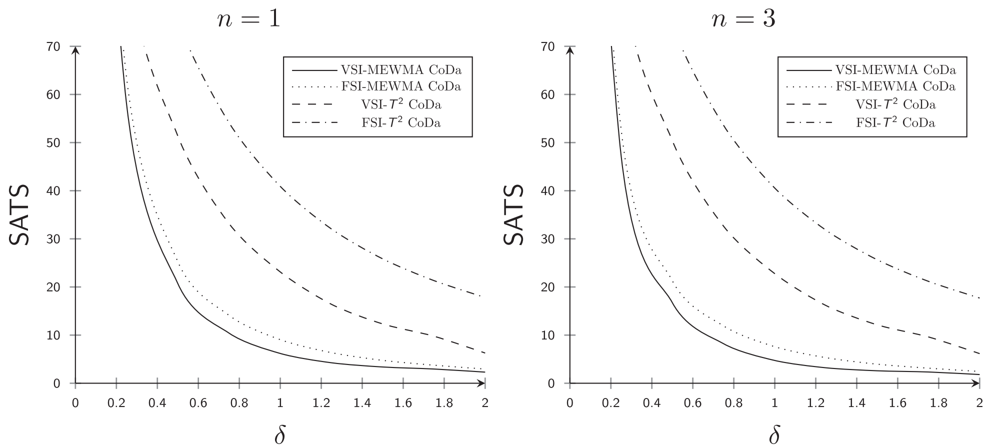

Figure 6. SATS Curves for $d = 3$ and $n \in \{1, 3\}$.

Table 5. OOC ZATS and SATS of the FSI and VSI-MEWMA CoDa CC.

	$d = 3$				$d = 5$			
	$n = 1$		$n = 3$		$n = 1$		$n = 3$	
δ	VSI	FSI	VSI	FSI	VSI	FSI	VSI	FSI
				ZATS				
0.25	56.842	64.6	48.346	53.98	64.362	70.27	53.746	59.654
0.5	19.935	26.4	17.731	22.44	26.695	31.65	22.731	27.686
0.75	10.441	15.1	9.095	12.9	15.69	19.72	13.495	17.525
1	6.913	9.9	5.489	8.44	10.748	13.89	9.289	12.431
1.25	4.934	7.1	3.885	6.05	8.333	10.67	7.285	9.623
1.5	3.728	5.4	3.185	4.61	6.978	8.55	6.185	7.757
1.75	3.008	4.3	2.852	3.68	6.075	7.03	5.452	6.407
2	2.419	3.5	2.246	3	4.95	5.81	4.446	5.306
				SATS				
0.25	55.762	63.35	47.267	52.74	62.882	68.6	52.905	58.698
0.5	18.935	25.25	16.732	21.29	25.295	30.08	21.819	26.642
0.75	9.561	14.1	8.216	11.91	14.41	18.3	12.662	16.579
1	6.153	9.05	4.732	7.59	9.588	12.62	8.56	11.616
1.25	4.254	6.35	3.209	5.31	7.253	9.5	6.629	8.899
1.5	3.128	4.75	2.59	3.96	5.978	7.48	5.606	7.131
1.75	2.488	3.75	2.338	3.13	5.155	6.06	4.952	5.857
2	1.979	2.95	1.806	2.45	4.51	5.26	4.006	4.756

CC decreases to SATS $= 4.732$. Figure 5 also shows the impact of the subgroup size n on the SATS of the MEWMA CoDa CC. From Figure 6, it is also clearly visible that n has a positive effect on the ATS of the VSI and FSI CC for CoDa.

5.3. Comparison of ZATS and SATS of the VSI-MEWMA CoDa control chart

To compare the ZS and the *SS* performance of the VSI-MEWMA CoDa CC, all the ZATS and the SATS values for different combinations of the involved variables are given in Table 5. It can be seen from Table 5 that the SATS for both the FSI and the VSI MEWMA CoDa CC are less than the ZATS of both the FSI and the VSI MEWMA CoDa CC.

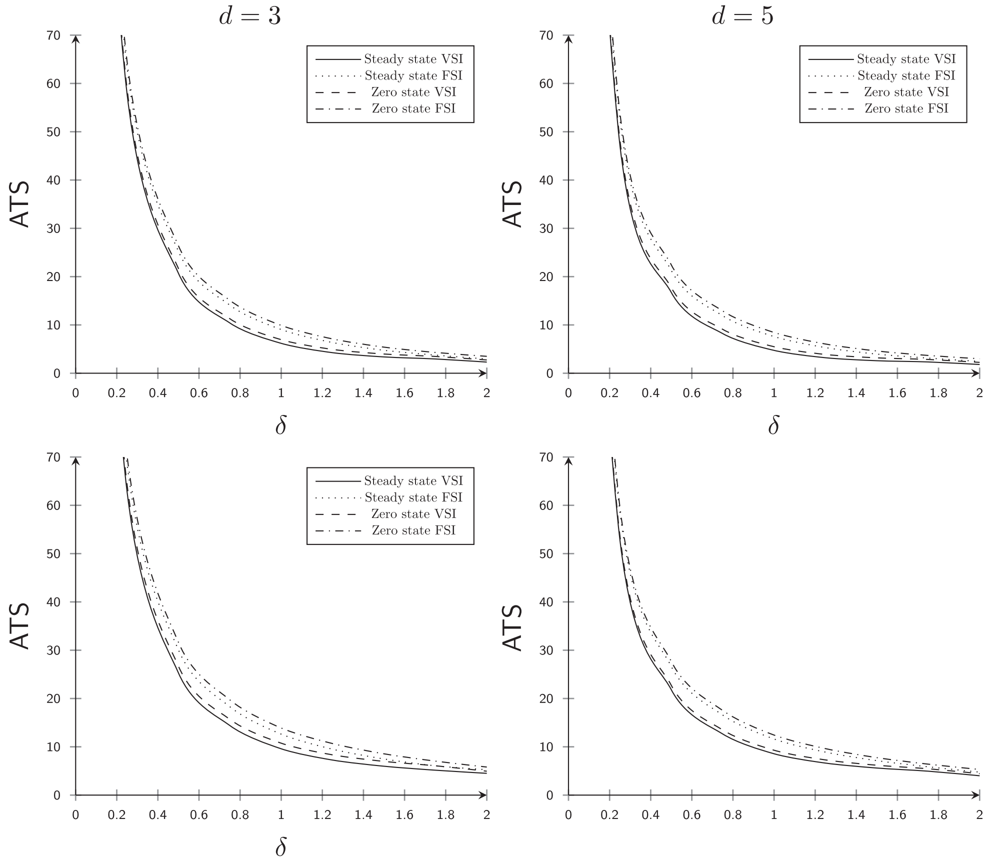

Figure 7. SATS and ZATS Curves.

When $\delta = 1$, $n = 1$, $d = 3$, the ZATS for the FSI and the VSI MEWMA CoDa CC are ZATS = 9.90 and ZATS = 6.948 respectively. While the SATS for the FSI and the VSI MEWMA CoDa CC are SATS = 9.05 and SATS = 6.188, respectively, are less than ZATS for both the FSI and the VSI MEWMA CoDa CC.

Figure 7 also shows the ZATS and the SATS curves for both the FSI and the VSI MEWMA CoDa CCs. A solid line shows the SATS of VSI in all the figures, the SATS of the FSI is shown by a dotted line in all the figures, the ZATS of the VSI is shown by a dashed line in all the figures, while a dashed-dotted line shows the ZATS of the FSI in all the figures. From Figure 7, we can see that the SATS for both the FSI and the VSI MEWMA CoDa CC are less than the ZATS of both the FSI and the VSI MEWMA CoDa CC.

Also the out-of-control performances of VSI MEWMA CoDa CC under ZS and *SS* can be compared using percentage improvement indicator as,

$$\Delta = \frac{100(\mathrm{ATS}_{ZS} - \mathrm{ATS}_{SS})}{\mathrm{ATS}_{ZS}}$$

Table 6 presents the percentage improvement in terms of out-of-control ATSs under the ZS and *SS* of the VSI-MEWMA CoDa CC for $n = 1$ and $p = 3$. The *SATS* of VSI-MEWMA CoDa CC is always smaller than the *ZATS* of the VSI-MEWMA CoDa CC.

Table 6. Comparison in terms of out-of-control ATSs under the ZS and *SS* of the VSI-MEWMA CoDa CC for $n = 1$ and $p = 3$.

δ	ZATS	SATS	Δ
0.25	56.842	55.762	1.90
0.50	19.935	18.935	5.02
0.75	10.441	9.561	8.43
1.00	6.913	6.153	10.99
1.25	4.934	4.254	13.78
1.50	3.728	3.128	16.09
1.75	3.008	2.488	17.29
2.00	2.419	1.979	18.19

In terms of their percentage improvement indicators it can be seen that, depending on the level of shift δ, when $n = 1$ and $p = 3$, the VSI-MEWMA CoDa CC under *SS* proposed in this paper is between 2% to 18% more efficient than the VSI-MEWMA CoDa CC under ZS presented in [35].

6. Illustrative example

Similar to [56,58], the example of the particle-size distribution for a plant in Europe is used in this study. According to [58], there were four OOC points in the data (i.e. (#1, #26, #45, #52)). Following [56], the author removed all the four OOC points described by Vives et al. [58] to get the IC phase I data set. Assume that the author would like to use the VSI-MEWMA CoDa CC with $r = 0.05$ and $H = 7.35$ to control a process of $d = 3$-part CoDa. After removing the OOC point, the IC phase-I dataset is given in Table 7. The estimates for the parameters of the multivariate normal distribution of the ilr transformed mean vector and variance-covariance matrix are given by

$$\boldsymbol{\mu}_0^* = \begin{pmatrix} 1.962 \\ 1.184 \end{pmatrix},$$

and

$$\boldsymbol{\Sigma}^* = \begin{pmatrix} 0.099 & -0.022 \\ -0.022 & 0.088 \end{pmatrix}.$$

while the mean of original CoDa can be written as

$$\boldsymbol{\mu}_0 = \begin{pmatrix} 0.892 \\ 0.056 \\ 0.052 \end{pmatrix},$$

For the phase II dataset, using simulation, 20 samples of size $n = 3$ have been generated using $\boldsymbol{\mu}_0^*$. The process is IC up to sample 10, after sample 10, a shift with the assignable cause in the mean vector has been introduced, and the next 10 samples are generated using $\boldsymbol{\mu}_1^*$.

Table 7. The Phase I dataset from [56].

i	*M*	*L*	*S*	*i*	*M*	*L*	*S*	*i*	*M*	*L*	*S*	*i*	*M*	*L*	*S*
1	92.6	3.2	4.2	14	94.5	2.6	2.9	27	83.6	7.4	9	40	84.5	6.9	8.6
2	91.7	5.2	3.1	15	94.5	2.7	2.8	28	84.8	6.8	8.4	41	84.4	7.4	8.2
3	86.9	3.5	9.6	16	88.7	7.9	3.4	29	87.1	6.3	6.6	42	84.3	8.9	6.8
4	90.4	2.9	6.7	17	84.6	6.6	8.8	30	87.2	6.1	6.7	43	89.8	8.2	2
5	92.1	4.6	3.3	18	90.7	4	5.3	31	87.3	6.6	6.1	44	90.4	6.7	2.9
6	91.5	4.4	4.1	19	90.2	2.5	7.3	32	84.8	6.2	9	45	90.1	5.9	4
7	90.3	5	4.7	20	92.7	3.8	3.5	33	87.4	6.5	6.1	46	83.6	8.7	7.7
8	85.1	8.4	6.5	21	91.5	2.8	5.7	34	86.8	6	7.2	47	88	6.4	5.6
9	89.7	4.2	6.1	22	91.8	2.9	5.3	35	88.8	4.8	6.4	48	84.7	8.4	6.9
10	92.5	3.8	3.7	23	90.6	3.3	6.1	36	89.8	4.9	5.3	49	93	5.1	1.9
11	91.8	4.3	3.9	24	87.3	7.2	5.5	37	86.9	5.8	7.3	50	91.4	5	3.6
12	91.7	3.7	4.6	25	82.6	7	10.4	38	83.8	7.2	9	51	86.2	5	8.8
13	90.3	3.8	5.9	26	83.5	6	10.5	39	89.2	5.6	5.2	52	87.2	5.9	6.9

Hence the mean vector shifted from

$$\boldsymbol{\mu}_0^* = \begin{pmatrix} 1.962 \\ 1.184 \end{pmatrix},$$

to

$$\boldsymbol{\mu}_1^* = \begin{pmatrix} 2.070 \\ 1.15 \end{pmatrix},$$

or it can be written in original CoDa as

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 0.901 \\ 0.048 \\ 0.051 \end{pmatrix}.$$

Where the shift from $\boldsymbol{\mu}_0^*$ to $\boldsymbol{\mu}_1^*$ equals $\delta = 0.34$. But here, the author has used a shift of size $\delta = 0.25$ in the mean vector. $\delta = 0.25$ is considered enough to detect the shift in μ^* quickly, as it is interpreted as a signal that something is not right in the process. For this reason, $\delta = 0.25$ is used to implement the VSI-MEWMA CoDa CC. For $n = 1$ and $\delta = 0.25$, the optimal parameters for the VSI-MEWMA CoDa CC are $r = 0.05$ and $H = 7.35$ (see Table 2).

Here the author has taken the subgroup of size $n = 3$ and the IC $\text{ATS}_0 = 200$. Using $h_2 = 0.1$ and $h_1 = 2.5$, the author gets a WL of the VSI-MEWMA CoDa CC $W = 0.8$ with the *SS* OOC $SATS = 57.882$. The next SI depends on the position of the VSI-MEWMA CoDa CC; if the CC lies below W, the SI will be h_1, while if the CC lies between W and H, the SI will be h_2. The VSI-MEWMA CoDa CC has a greater degree of efficacy than the FSI-MEWMA CoDa CC; as for the same values of r, H, d and n, the FSI-MEWMA CoDa CC have the *SS* OOC $SATS = 63.35$. For the sake of comparison, the author has used a percentage improvement indicator,

$$\Delta = \frac{100(\text{ATS}_{\text{FSI}} - \text{ATS}_{\text{VSI}})}{\text{ATS}_{\text{FSI}}}.$$

The percentage improvement indicator shows that the VSI-MEWMA CoDa CC has almost **8.63%** greater degree of efficacy than the FSI-MEWMA CoDa CC in terms of *SS* OOC

SATS. The same is the case with ZS OOC ZATS, and the VSI-MEWMA CoDa CC have an almost **8.73%** greater degree of efficacy than the FSI-MEWMA CoDa CC in terms of *ZATS*.

7. Conclusions

This article has investigated the performance of the VSI function in the MEWMA CoDa CCs using a normal random vector described as the inverse log-ratio of a d-part CoDa to monitor the mean vector. This article focuses on the VSIs instead of using the FSI-based charting schemes to monitor the shift in the process mean vector. In the VSI CC, the length of the SI depends on the charting statistic. A WL was introduced, and the SI length was divided into two values, h_1 for the large SI and h_2 for the small SI. By fixing the small SI to 0.1, the author can find the values of optimal parameters considering a fixed value of the IC ATS_0 for a wide range of shifts in the process mean. If the monitored statistics lie below the WL, then a large SI has been used. But, when the monitored statistics lie between the warning and the UCL, the small SI has been used. The proposed study has investigated the VSI MEWMA CoDa CC's performance under zero-state and steady-state properties of the run length using the CTMC method. The process mean and the variance-covariance matrix is supposed to be known for the ATS comparative study. Different values of the number of variables d and subgroup size n have been used to investigate the OOC ATS using a fixed value of IC ATS. The main conclusions of this article are (i). The ZATS and the SATS of the FSI-MEWMA CoDa CC are greater than the ZATS and the SATS of the VSI-MEWMA CoDa CC; (ii). The ATS of the VSI-MEWMA CoDa CC increases with an increase in the d; (iii). The ATS of the VSI-MEWMA CoDa CC decreases with an increase in the n; (iv). The SATS of the proposed CC is less than the ZATS for all the different combinations of n and d. A comparison of the VSI-MEWMA CoDa CC with the FSI-MEWMA CoDa CC, the VSI and the FSI Hotelling T^2 CoDa CC has also been made to study the statistical sensitivity of the proposed CC. For future research, MCUSUM-CoDa, Hotelling T^2 CoDa, for the location mean vector and dispersion matrix, can be designed using VSI.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by National Natural Science Foundation of China [Grant number: 71802110]; Foundation of Nanjing University of Posts and Telecommunications [Grant number: NY222176]; The Excellent Innovation Teams of Philosophy and Social Science in Jiangsu Province [Grant number: 2017ZSTD022]; Key Research Base of Philosophy and Social Sciences in jiangsu Information Industry Integration Innovation and Emergency Management Research Center [Grant number: None]; Humanity and Social Science Foundation of the Ministry of Education of China [Grant number: 19YJA630061].

ORCID

Muhammad Imran http://orcid.org/0000-0003-0592-4635
Xuelong Hu http://orcid.org/0000-0002-5556-2612

Fatima Sehar Zaidi http://orcid.org/0000-0003-0265-4051
Anan Tang http://orcid.org/0000-0002-3933-1523

References

[1] J. Aitchison, *The Statistical Analysis of Compositional Data*, Monographs on Statistics and Applied Probability, Reprinted 2003 with additional material by Blackburn Press. Chapman and Hall Ltd., London, 1986.

[2] J. Aitchison, *The Statistical Analysis of Compositional Data*, Monographs on Statistics and Applied Probability. Springer Dordrecht Netherlands, 1986.

[3] J. Aitchison and J.J. Egozcue, *Compositional data analysis: Where are we and where should we be heading?*, Math. Geol. 37 (2005), pp. 829–850.

[4] S. Asghar, F. Alireza, H. Cédric, S. Erwin, and M.B. Moghadam, *A modified economic-statistical design of the T^2 control chart with variable sample sizes and control limits*, J. Appl. Stat. 38 (2011), pp. 2459–2469.

[5] L. Bessegato, R. Quinino, L.L. Ho, and L. Duczmal, *Variable interval sampling in economical designs for online process control of attributes with misclassification errors*, J. Oper. Res. Soc. 62 (2011), pp. 1365–1375.

[6] P.D. Bourke, *The geometric CUSUM chart with sampling inspection for monitoring fraction defective*, J. Appl. Stat. 28 (2001), pp. 951–972.

[7] P. Castagliola, A. Achouri, H. Taleb, G. Celano, and S. Psarakis, *Monitoring the coefficient of variation using a variable sampling interval control chart*, Qual. Reliab. Eng. Int. 29 (2013), pp. 1135–1149.

[8] C.W. Champ, *Steady-state run length analysis of a shewhart quality control chart with supplementary runs rules*, Commun. Stat. Theory Methods 21 (1992), pp. 765–777.

[9] R.Q. Cui and M.R. Reynolds Jr, *Chart with runs and variable sampling intervals*, Commun. Stat. Simul. Comput. 17 (1988), pp. 1073–1093.

[10] R.B. Davis and W.H. Woodall, *Evaluating and improving the synthetic control chart*, J. Qual. Technol. 34 (2002), pp. 200–208.

[11] J.J. Egozcue and V. Pawlowsky-Glahn, *Compositional data: The sample space and its structure*, Test Off. J. Span. Soc. Stat. Oper. Res. 28 (2019), pp. 599–638. https://doi.org/10.1007/s11749-019-00670-6.

[12] A.R. Faraz and E. Saniga, *Economic statistical design of A T2 control chart with double warning lines*, Qual. Reliab. Eng. Int. 27 (2011), pp. 125–139.

[13] R. Ghanaatiyan, A. Amiri, and F. Sogandi, *Multi-Objective economic-Statistical design of VSSI-MEWMA-DWL control chart with multiple assignable causes*, J. Ind. Syst. Eng. 10 (2017), pp. 34–58.

[14] M.A. Gulzar, W. Latif, M. Abid, H.Z. Nazir, and M. Riaz, *On enhanced exponential-cum-ratio estimators using robust measures of location*, Concurr. Comput. Pract. Exp. 34 (2022), Article ID e6763. e6763 CPE-21-1421.R1.

[15] K. Hron, M. Templ, and P. Filzmoser, *Imputation of missing values for compositional data using classical and robust methods*, Comput. Stat. Data Anal. 54 (2010), pp. 3095–3107.

[16] S. Hussain, M. Sun, T. Mahmood, M. Riaz, and M. Abid, *IQR CUSUM charts: An efficient approach for monitoring variations in aquatic toxicity*, J. Chemom. 35 (2021), Article ID e3336.

[17] M. Imran, J.S. Sun, F.S. Zaidi, Z. Abbas, and H.Z. Nazir, *Multivariate cumulative sum control chart for compositional data with known and estimated process parameters*, Qual. Reliab. Eng. Int. 38 (2022), pp. 2691–2714.

[18] J.E. Jarrett, *Total quality management (TQM) movement in public health*, Int. J. Qual. Reliab. Manag. 33 (2015), pp. 25–41. https://doi.org/10.1108/IJQRM-12-2013-0193.

[19] S.K. Khilare and D.T. Shirke, *Nonparametric synthetic control charts for process variation*, Qual. Reliab. Eng. Int. 28 (2012), pp. 193–202.

[20] K. Kim, R. Marion, and J. Reynolds, *Multivariate monitoring using an MEWMA control chart with unequal sample sizes*, J. Qual. Technol. 37 (2005), pp. 267–281.

[21] S. Knoth, *Steady-state average run length(s): Methodology, formulas, and numerics*, Seq. Anal. 40 (2021), pp. 405–426.

[22] M.H. Lee, *Variable sampling rate multivariate exponentially weighted moving average control chart with double warning lines*, Qual. Technol. Quant. Manag. 10 (2013), pp. 353–368.

[23] M.H. Lee and M.B.C. Khoo, *Double sampling |s| control chart with variable sample size and variable sampling interval*, Commun. Stat. Simul. Comput. 47 (2018), pp. 615–628.

[24] T.J. Lim and M. Cho, *Design of control charts with M-of-M runs rules*, Qual. Reliab. Eng. Int. 25 (2009), pp. 1085–1101.

[25] Y.C. Lin and C.Y. Chou, *Robustness of the EWMA and the combined $\bar{X}$-EWMA control charts with variable sampling intervals to non-normality*, J. Appl. Stat. 38 (2011), pp. 553–570.

[26] K.W. Linna, W.H. Woodall, and K.L. Busby, *The performance of multivariate control charts in the presence of measurement error*, J. Qual. Technol. 33 (2001), pp. 349–355.

[27] C.A. Lowry, W.H. Woodall, C.W. Champ, and S.E. Rigdon, *A multivariate exponentially weighted moving average control chart*, Technometrics 34 (1992), pp. 46–53.

[28] J.M. Lucas and M.S. Saccucci, *Exponentially weighted moving average control schemes: Properties and enhancements*, Technometrics 32 (1990), pp. 1–12.

[29] S.B. Mahadik and D.T. Shirke, *On the superiority of a variable sampling interval control chart*, J. Appl. Stat. 34 (2007), pp. 443–458.

[30] M.A. Mahmoud, W.H. Woodall, and R.E. Davis, *Performance comparison of some likelihood ratio-based statistical surveillance methods*, J. Appl. Stat. 35 (2008), pp. 783–798.

[31] M. Mohammadzadeh, A. Yeganeh, and A.R. Shadman, *Monitoring logistic profiles using variable sample interval approach*, Comput. Ind. Eng. 158 (2021), Article ID 107438.

[32] W.E. Molnau, G.C. Runger, D.C. Montgomery, K.R. Skinner, E.N. Loredo, and S.S. Prabhu, *A program of ARL calculation for multivariate EWMA charts*, J. Qual. Technol. 33 (2001), pp. 515–521.

[33] H.D. Nguyen, K.P. Tran, and H.L. Heuchenne, *Cusum control charts with variable sampling interval for monitoring the ratio of two normal variables*, Qual. Reliab. Eng. Int. 36 (2020), pp. 474–497.

[34] Q.T. Nguyen, V. Giner-Bosch, K.D. Tran, C. Heuchenne, and K.P. Tran, *One-sided variable sampling interval EWMA control charts for monitoring the multivariate coefficient of variation in the presence of measurement errors*, Int. J. Adv. Manuf. Technol. 115 (2021), pp. 1821–1851.

[35] T.V. Nguyen, C. Heuchenne, and K.P. Tran, *Anomaly detection for compositional data using VSI MEWMA control chart*, in *Scientific Congresses and Symposiums*, Elsevier, June 2022

[36] V. Pawlowsky-Glahn, J.J. Egozcue, and R. Tolosana-Delgado, *Modeling and Analysis of Compositional Data*, John Wiley & Sons, Hoboken, New Jersey, US, 2015.

[37] V. Pawlowsky-Glahn and J.J. Egozcue, *Compositional data in geostatistics: A log-ratio based framework to analyze regionalized compositions*, Math. Geosci. 52 (2020), pp. 1067–1084.

[38] S.S. Prabhu, D.C. Montgomery, and G.C. Runger, *A combined adaptive sample size and sampling interval $\bar{X}$ control scheme*, J. Qual. Technol. 26 (1994), pp. 164–176.

[39] S.S. Prabhu and G.C. Runger, *Designing a multivariate EWMA control chart*, J. Qual. Technol. 29 (1997), pp. 8–15.

[40] P. Qiu and D. Hawkins, *A rank-based multivariate CUSUM procedure*, Technometrics 43 (2001), pp. 120–132.

[41] M.F. Ramalhoto and M. Morais, *Shewhart control charts for the scale parameter of a weibull control variable with fixed and variable sampling intervals*, J. Appl. Stat. 26 (1999), pp. 129–160.

[42] M.R. Reynolds, R.W. Amin, and J.C. Arnold, *CUSUM charts with variable sampling intervals*, Technometrics 32 (1990), pp. 371–384.

[43] M.R. Reynolds and G.Y. Cho, *Multivariate control charts for monitoring the mean vector and covariance matrix with variable sampling intervals*, Seq. Anal. 30 (2011), pp. 1–40.

[44] M. R. Reynolds and K. Kim, *Multivariate control charts for monitoring the process mean and variability using sequential sampling*, Seq. Anal. 26 (2007), pp. 283–315.

[45] G.C. Runger and D.C. Montgomery, *Adaptive sampling enhancements for shewhart control charts*, IIE Trans. 25 (1993), pp. 41–51.

[46] G.C. Runger and J.J. Pignatiello, *Adaptive sampling for process control*, J. Qual. Technol. 23 (1991), pp. 135–155.
[47] G.C. Runger and S.S. Prabhu, *A Markov chain model for the multivariate exponentially weighted moving averages control chart*, J. Am. Stat. Assoc. 91 (1996), pp. 1701–1706.
[48] H. Sabahno, A. Amiri, and P. Castagliola, *A new adaptive control chart for the simultaneous monitoring of the mean and variability of multivariate normal processes*, Comput. Ind. Eng. 151 (2021), Article ID 106524.
[49] S.E. Shamma, R.W. Amin, and A.K. Shamma, *A double exponentially weigiited moving average control procedure with variable sampling intervals*, Commun. Stat. Simul. Comput. 20 (1991), pp. 511–528.
[50] L. Shi, C. Zou, Z. Wang, and K.C. Kapur, *A new variable sampling control scheme at fixed times for monitoring the process dispersion*, Qual. Reliab. Eng. Int. 25 (2009), pp. 961–972.
[51] S.C. Shongwe, J. Malela-Majika, and P. Castagliola, *A combined mixed-s-skip sampling strategy to reduce the effect of autocorrelation on the $\bar{X}$ scheme with and without measurement errors*, J. Appl. Stat. 48 (2021), pp. 1243–1268.
[52] M.S. Srivastava and Y. Wu, *Comparison of EWMA, CUSUM and Shiryayev–Roberts procedures for detecting a shift in the mean*, Ann. Stat. 21 (1993), pp. 645–670. https://doi.org/10.1214/aos/1176349142.
[53] Z.G. Stoumbos, M.R. Reynolds, T.P. Ryan, and W.H. Woodall, *The state of statistical process control as we proceed into the 21st century*, J. Am. Stat. Assoc. 95 (2000), pp. 992–998.
[54] Z.G. Stoumbos, J. Mittenthal, and G.C. Runger, *Steady-state-optimal adaptive control charts based on variable sampling intervals*, Stoch. Anal. Appl. 19 (2001), pp. 1025–1057.
[55] Z.G. Stoumbos and M.R. Reynolds, *Control charts applying a general sequential test at each sampling point*, Seq. Anal. 15 (1996), pp. 159–183.
[56] K.P. Tran, P. Castagliola, G. Celano, and M.B.C. Khoo, *Monitoring compositional data using multivariate exponentially weighted moving average scheme*, Qual. Reliab. Eng. Int. 34 (2018), pp. 391–402.
[57] P.H. Tran and C. Heuchenne, *Monitoring the coefficient of variation using variable sampling interval CUSUM control charts*, J. Stat. Comput. Simul. 91 (2021), pp. 501–521.
[58] M. Vives-Mestres, J. Daunis-I-Estadella, and J.A. Martin-Fernandez, *Individual T^2 control chart for compositional data*, J. Qual. Technol. 46 (2014), pp. 127–139.
[59] M. Vives-Mestres, J. Daunis-I-Estadella, and J.A. Martín-Fernández, *Out-of-control signals in three-part compositional T^2 control chart*, Qual. Reliab. Eng. Int. 30 (2014), pp. 337–346.
[60] M. Vives-Mestres, J. Daunis i Estadella, and J. Martín-Fernández, *Signal interpretation in hotelling's T^2 control chart for compositional data*, IIE Trans. 48 (2016), pp. 661–672.
[61] Z. Wu, Y. Tian, and S. Zhang, *Adjusted-loss-function charts with variable sample sizes and sampling intervals*, J. Appl. Stat. 32 (2005), pp. 221–242.
[62] F.S. Zaidi, P. Castagliola, K.P. Tran, and M.B.C. Khoo, *Performance of the hotelling T^2 control chart for compositional data in the presence of measurement errors*, J. Appl. Stat. 46 (2019), pp. 2583–2602.
[63] F.S. Zaidi, P. Castagliola, K.P. Tran, and M.B.C. Khoo, *Performance of the MEWMA-CoDa control chart in the presence of measurement errors*, Qual. Reliab. Eng. Int. 36 (2020), pp. 2411–2440.
[64] C. Zou, F. Tsung, and Z. Wang, *Monitoring general linear profiles using multivariate exponentially weighted moving average schemes*, Technometrics 49 (2007), pp. 395–408.