IMPROVED ADAPTIVE THRESHOLDING LASSO CHART FOR MONITORING DISPERSION OF HIGH-DIMENSIONAL PROCESSES USING GENERALIZED MULTIPLE DEPENDENT STATE SAMPLING

Mehdrad Hajiesmaeili[1], Mohammad Reza Maleki[2*], Ali Salmasnia[3]

•

[1]Department of Industrial Engineering, University of Eyvanekey, Semnan, Iran
esmaili4802@gmail.com
[2]Industrial Engineering Group, Golpayegan College of Engineering, Isfahan University of Technology, Golpayegan 87717-67498, Iran
m.maleki@iut.ac.ir
[3]Department of Industrial Engineering, Faculty of Engineering, University of Qom, Iran
a.salmasnia@qom.ac.ir

Abstract

In many applications of multivariate statistical quality control, it is commonly observed that the number of quality characteristics exceeds the sample size. This poses significant challenges in monitoring high-dimensional data. In such conditions, it is challenging to detect sparse changes where an assignable cause leads to the deviation of only a few elements in the covariance matrix. On the other hand, the utilization of the multiple dependent state (MDS) sampling technique to enhance the sensitivity of control charts has recently attracted the attention of researchers. However, to the best of the authors' knowledge, no previous research has been conducted on equipping multivariate dispersion control charting methods with the MDS technique under high dimensionality. Therefore, this article integrates the adaptive thresholding Lasso statistic with the MDS and generalized MDS techniques to track all types of disturbances in the covariance matrix of high-dimensional processes, including diagonal, off-diagonal, and joint diagonal/off-diagonal deviations. The performance of the proposed control charts will be compared through a numerical example under seven out-of-control patterns in terms of three metrics: average, standard deviation, and median of run length. The results clearly indicate that the use of both sampling techniques significantly improves the run length properties of the adaptive thresholding Lasso chart.

Keywords: Adaptive thresholding Lasso control chart; High dimensional process; Multiple dependent state sampling; Run length; Covariance matrix.

I. Introduction

In recent years, increasing customer expectations and technological advancements have resulted in the development of new production processes that require monitoring numerous quality characteristics. In this context, we can mention imaging processes and multi-stage production processes. Conventional multivariate control charts lose their sensitivity to process changes as the number of quality characteristics under study increases. This challenge becomes more significant when only a limited number of variables are affected by assignable causes [1]. Therefore, in recent years, significant efforts have been devoted to monitoring high-dimensional data, where the number of quality characteristics exceeds the sample size. One effective strategy for monitoring high-dimensional processes involves implementing control charts based on variable selection approach. This method involves identifying a small subset of variables that may deviate from their nominal value and subsequently conducting the monitoring process using this subset. In this connection, Abdella et al. [2] developed a multivariate cumulative sum (MCUSUM) chart using a variable selection approach for the quick detection of small mean disturbances in high-dimensional process monitoring. Sangahn [3] integrated the deviance residual-based multivariate exponentially weighted moving average statistic with a variable selection procedure for Phase II monitoring of high-dimensional multistage processes. Zhang et al. [4] proposed a sensitized variable selection-based control charting method that utilizes a classification algorithm for Phase II monitoring of high-dimensional industrial process.

Although variable selection-based control charts have significant advantages for monitoring high-dimensional processes under sparse conditions, they face challenges in identifying suspicious variables. The efficacy of this control chart is significantly reduced when the suspicious variables are not accurately identified. Specifically, when the shift magnitude in process parameters is small, these monitoring schemes lose their effectiveness because they fail to accurately identify all the variables responsible for the deviation in the selected set. As one of the most effective alternative methods for monitoring the dispersion of high-dimensional processes, the adaptive thresholding Lasso control chart has gained significant attention from researchers in recent years. In this monitoring procedure, the first step involves calculating the difference matrix through the subtraction of the sample covariance from its corresponding target matrix. Then, according to the sparsity assumption, only the values of the difference matrix that exceed a certain threshold value are taken into consideration. In other words, any elements in the difference matrix that are smaller than the threshold value are considered to be equal to zero. Interested readers can refer to important references, such as [5], [6], [7], [8], and [9], for more comprehensive details regarding the adaptive thresholding LASSO control chart.

On the other hand, the performance of control charts significantly depends on the sampling strategy used to collect process observations. This issue becomes even more crucial in situations where any delay in identifying process disturbances, regardless of shift magnitude, results in substantial costs for the production system. In this context, researchers have recently been focusing on new sampling strategies, such as double sampling, ranked set sampling [10], repetitive sampling [11], and multiple dependent state sampling [12], to enhance the power of various control charts. Recently, the multiple dependent state (MDS) sampling strategy has been successfully utilized to enhance the diagnostic capability of control charts. In this regard, Arshad et al. [13] equipped a variability control charting mechanism based on sample variance statistic to MDS sampling strategy in order to improve the detection of increased variance shifts. Aiming to achieve rapid detection of small mean deviations, Naveed et al. [14] have developed a modified version of the exponentially weighted moving average (EWMA) control chart that utilizes MDS strategy. Aslam et al. [15] designed an enhanced version of the $\bar{X}$ control chart based on the generalized multiple dependent state (GMDS) sampling for rapid detection of small mean deviations. Rao et al. [16] studied the efficiency of the GMDS sampling technique in improving the detectability of the coefficient of variation (CV) control chart. Please refer to [17] and [18] for additional information on MDS-based control charts.

The high dispersion of the quality characteristics under investigation will lead to an increase in the production rate of non-conforming products, thereby resulting in an increase in quality loss costs. In many manufacturing environments, product fluctuations often increase due to various factors, including changes in raw materials, machinery depreciation, operator mistakes, environmental factors, and insufficient calibration of measuring equipment calibration. If these factors are not promptly identified and eliminated, the costs imposed on the system due to the production of non-conforming products will increase greatly. According to the mentioned cases, monitoring the covariance matrix in multivariable production processes is of great importance for maintaining product quality in both industrial and service environments. In recent years, researchers such as [19] and [20] have dedicated their attention to tracing the deviations of the covariance matrix. Most covariance matrix monitoring control charts are commonly designed with the assumption that the sample size exceeds the problem dimension or the number of quality characteristics. However, this assumption does not hold true in numerous statistical quality control applications. Fortunately, there have been recent efforts to monitor the scatter matrix of high-dimensional data. In this context, Kim et al. [21] developed a covariance matrix monitoring scheme based on a ridge penalized likelihood ratio statistic for the identification of general disturbances, without the need for sparsity assumption. A monitoring scheme for Phase I monitoring of high-dimensional variability matrices, utilizing the sparse-leading-eigenvalue statistic was developed by [22]. Safikhani et al. [23] employed an additive covariate model to examine how imprecise observations affect the effectiveness of the ridge penalized likelihood ratio chart in Phase II monitoring of high-dimensional process dispersion. As far as we know, the MDS sampling technique has mainly been used in control charts to monitor univariate processes. Therefore, due to the appropriate performance of the MDS method in improving the sensitivity of control charts. Therefore, this paper aims to utilize the MDS and GMDS sampling methods for Phase II monitoring of high-dimensional process dispersion, as they have been shown to improve the sensitivity of control charts. Specifically, in this article two adaptive thresholding Lasso-based control charting schemes named MDS-ATL and GMDS-ATL are presented to identify the covariance matrix disturbances under high-dimensionality.

The structure of this article is organized as: Section 2 describes the adaptive thresholding Lasso control chart, which is used to detect covariance matrix disturbances in high-dimensional processes. Two variability control charts are developed by integrating the adaptive thresholding Lasso statistic with the multiple dependent state and generalized multiple dependent state sampling strategies in Section 3. In Section 4, firstly, seven out-of-control scenarios for the covariance matrix are defined, based on three patterns: diagonal, off-diagonal, and joint diagonal/off-diagonal patterns. We then conduct extensive simulation studies to demonstrate the effectiveness of the MDS and GMDS sampling strategies in enhancing the sensitivity of the adaptive thresholding Lasso control chart. Section 5 is devoted to the conclusion and recommendations for future research.

II. Adaptive thresholding Lasso monitoring scheme

In this section, we discuss the adaptive thresholding Lasso control chart as an efficient approach in monitoring the dispersion of high-dimensional processes. This control chart has the ability to detect covariance matrix disturbances across all three categories: diagonal, non-diagonal, and diagonal-non-diagonal deviations. In this regard, Table 1 presents the indices, distribution parameters, sample parameters, and chart parameters employed in the development of the adaptive thresholding Lasso chart along with its improved versions based on MDS and GMDS sampling strategies.

Table 1. *Notations*

Indices	Description
t	Index of sample
i	Index of observation
j,k	Indices of quality characteristics
	Distribution parameters
y_j	Quality characteristic j
$\mathbf{Y}_t$	The matrix of observations obtained in the t^{th} sample
$\mathbf{y}_{ti}$	The column vector of quality characteristics in the i^{th} observation of the t^{th}
y_{tij}	The value of the j^{th} quality characteristic in the i^{th} observation of the t^{th} sample
p	Process dimension
$\boldsymbol{\mu}_{ic}$	Target mean vector
μ_j	Target mean parameter of the j^{th} quality characteristic
$\boldsymbol{\Sigma}_{ic}$	Target covariance matrix
σ_j^2	Target variance of the j^{th} quality characteristic
σ_{jk}	Target covariance of quality characteristics j and k
γ_{jk}	Threshold value for the covariance between quality characteristics j and k
θ_{jk}	Variance of $\hat{d}_{t,jk}$
$\boldsymbol{\Sigma}_{oc}$	Off-target covariance matrix
$\boldsymbol{\Psi}$	The matrix of shift magnitudes within the covariance matrix
	Sample parameters
$\bar{\mathbf{y}}_t$	Sample mean vector at the t^{th} sample.
$\mathbf{S}_t$	Sample covariance matrix at the t^{th} sample
S_{tj}^2	Sample variance of quality characteristic j at the t^{th} sample
$S_{t,jk}$	Sample covariance of quality characteristics j and k at the t^{th} sample
$\mathbf{D}_t$	Difference matrix at the t^{th} sample
$\hat{\mathbf{D}}_t$	Shrinkage difference matrix at the t^{th} sample
$d_{t,jk}$	The component of matrix $\mathbf{D}_t$ at row j and column k
$\hat{d}_{t,jk}$	The component of matrix $\hat{\mathbf{D}}_t$ at row j and column k
ATL_t	Adaptive thresholding LASSO statistic at the t^{th} sample
	Chart/Other parameters
α	Probability of type I error
n	Sample size
φ	Regularization parameter
η	Shrinkage parameter
r	Number of additional samples taken
q	Minimum number of additional samples
UCL	Control limit of adaptive thresholding LASSO chart
UCL_{inner}	Inner control limit of the proposed MDS- based adaptive thresholding LASSO charts
UCL_{outer}	Outer control limit of the proposed MDS- based adaptive thresholding LASSO charts
ARL_{ic}	In-control average run length
τ	The time of occurring an assignable cause of deviation
T	The time of alerting an out-of-control signal by the control chart
$(a)_+$	Operator for converting negative values to zero

Consider a process in which a product manufacturer or service provider is interested in concurrently monitoring p correlated quality characteristics $y_1, y_2, ..., y_p$ which all follow a normal distribution. In this case, the observations obtained in the t^{th} sample are described by the matrix $\mathbf{Y}_t = (\mathbf{y}_{t1}, \mathbf{y}_{t2}, ..., \mathbf{y}_{tn_t})_{p\times n_t}$, where $\mathbf{y}_{ti} = \left(y_{ti1}, y_{ti2}, ..., y_{tip}\right)^T ; t = 1, ..., T, i = 1, ..., n_t$. In addition, n_t denotes the number of observations at the t^{th} sample, while T indicates the point at which the out-of-control signal is issued. In this thesis, it is assumed, in accordance with the convention of statistical quality control, that the number of observations in each subgroup remains constant and equal to n between consecutive samples. Furthermore, it is presumed in this study that the observations within a sample are independent from each other. In other words, the columns of the observation matrix $\mathbf{Y}_t$, denoted as $\mathbf{y}_{t1}, \mathbf{y}_{t2}, ..., \mathbf{y}_{tn}$, are considered to be independent. The mean vector and covariance matrix of the study quality characteristics are defined by Equations (1) and (2), respectively, when the process is statistically in-control.

$$\boldsymbol{\mu}_{ic} = \left(\mu_1, \mu_2, ..., \mu_p\right)^T \tag{1}$$

$$\boldsymbol{\Sigma}_{ic} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{pmatrix} \tag{2}$$

Since the purpose of adaptive thresholding Lasso control chart is to identify variations in the covariance matrix in high-dimensional processes, the mean vector is considered to remain constant throughout each production cycle, and the assignable cause solely affects the elements of the covariance matrix. In other words, when an assignable cause occurs at time τ, the covariance matrix of random vectors $\mathbf{y}_{ti}; t = \tau+1, \tau+2, ..., T; i = 1, ..., n$ changes from $\boldsymbol{\Sigma}_{ic}$ to $\boldsymbol{\Sigma}_{oc} = \boldsymbol{\Sigma}_{ic} + \boldsymbol{\Psi}$, while the mean vector remains constant at $\boldsymbol{\mu}_{ic}$. The adaptive thresholding Lasso control chart is designed to compare the sample covariance matrix in each sample with its target matrix. In other words, this monitoring procedure conducts the following hypothesis test:

$$\begin{aligned} H_0 &: \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_{ic} = \mathbf{0}_{p\times p} \\ H_1 &: \boldsymbol{\Sigma} - \boldsymbol{\Sigma}_{ic} \neq \mathbf{0}_{p\times p} \end{aligned} \tag{3}$$

To test hypothesis (3) using the adaptive thresholding Lasso method, the difference matrix is calculated based on the difference between the sample and target covariance matrices as:

$$\mathbf{D}_t = \mathbf{S}_t - \boldsymbol{\Sigma}_{ic} \tag{4}$$

By calculating the sample mean vector of the t^{th} subgroup as $\bar{\mathbf{y}}_t = \frac{1}{n}\sum_{i=1}^{n} \mathbf{y}_{ti}$, we can obtain the sample covariance matrix as follows:

$$\mathbf{S}_t = \begin{pmatrix} S_{t1}^2 & S_{t,12} & \cdots & S_{t,1p} \\ S_{t,21} & S_{t2}^2 & \cdots & S_{t,2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{t,p1} & S_{t,p2} & \cdots & S_{tp}^2 \end{pmatrix} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{y}_{ti} - \bar{\mathbf{y}}_t\right)\left(\mathbf{y}_{ti} - \bar{\mathbf{y}}_t\right)^T \tag{5}$$

It is evident that the probability of rejecting the null hypothesis increases as the distance between the elements of matrix $\mathbf{D}_t$ increases from zero. Since $\mathbf{Y}_t$ is a random matrix, the components of the difference matrix exhibit slight deviations from their nominal values even when the covariance matrix remains stable. Therefore, according to [5], $\hat{\mathbf{D}}_t$ is defined by setting a threshold value based on Equation (6) such that the respective values of the in-control components are equal to zero.

$$\hat{d}_{t,jk} = d_{t,jk}\left(1-\left|\gamma_{jk}/d_{t,jk}\right|^{\eta}\right)_{+};\ j,k=1,\ldots,p \tag{6}$$

In Equation (6), the output of the operator $(.)_+$ is equal to $1-\left|\frac{\gamma_{jk}}{d_{jk}}\right|^{\eta}$ when $\left|\gamma_{jk}/d_{jk}\right|^{\eta}$ is less than 1; otherwise the output will be zero. Moreover, η is a predetermined constant parameter that determines the amount of shrinkage, while γ_{jk} represents an specified threshold which is calculated as follows.

$$\gamma_{jk} = \varphi\sqrt{\frac{\theta_{jk}\log p}{n}} \tag{7}$$

In Equation (7), φ is a non-negative constant known as the regularization parameter, while θ_{jk} represents the variance of $\hat{d}_{t,jk}$ which is calculated in terms of the components of $\mathbf{\Sigma}_{ic}$ as:

$$Var(\hat{d}_{t,jk}) = \sigma_j^2\sigma_k^2 + \sigma_{jk}^2 \tag{8}$$

In Equation (8), σ_j^2 and σ_k^2 denote the variances of variables j and k, respectively, whereas σ_{jk} represents the covariance between them. Accordingly, the value of the component in the j^{th} row and k^{th} column of the shrinkage difference matrix at the t^{th} sample can be rewritten as:

$$\hat{d}_{t,jk} = d_{t,jk}\left(1-\left|\frac{\varphi\sqrt{\left(\sigma_j^2\sigma_k^2+\sigma_{jk}^2\right)n^{-1}\log p}}{d_{t,jk}}\right|^{\eta}\right)_{+};\ j,k=1,\ldots,p \tag{9}$$

Finally, the adaptive thresholding Lasso chart statistic for the t^{th} sample based on $\hat{d}_{t,jk}$ and θ_{jk} is calculated according to Equation (10).

$$ATL_t = n^{-1}\sum_{k=1}^{p}\sum_{l=1}^{p}\left(\frac{\hat{d}_{t,jk}^2}{\theta_{jk}}\right) \tag{10}$$

The ideal value of ATL_t is zero, and as the sample covariance matrix deviates further from $\mathbf{\Sigma}_{ic}$, the chart statistic increases. As can be observed, each component in Equation (10) is a positive value, and therefore the adaptive thresholding Lasso control chart only has an upper control limit, denoted by UCL. In this study, the value of UCL is determined through 10,000 iterations of Monte Carlo simulation to ensure that the in-control average run length (ARL_{ic}) of the chart is equal to a predetermined value of $\frac{1}{\alpha}$.

III. Proposed MDS-ATL and GMDS-ATL control charting methods

One of the key factors that significantly affects the control chart's ability to detect various shifts in the distribution parameters is the sampling methodology that is employed. This section presents two monitoring approaches, namely MDS-ATL and GMDS-ATL to enhance the detectability of the adaptive thresholding Lasso control chart for high-dimensional process monitoring based on MDS and GMDS sampling strategies. In the proposed control charting methods, the control limit interval is partitioned into three zones of safety, warning, and rejection as follows:

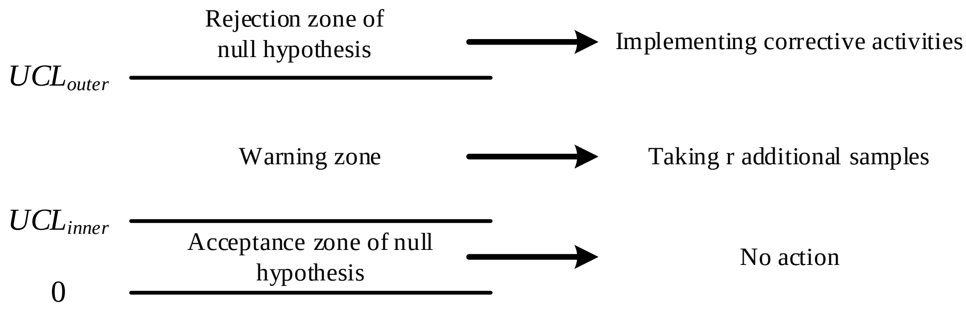

Figure 1. *Safety, warning, and rejection zones*

In the MDS-ATL chart, the covariance matrix is declared to be in-control when the adaptive thresholding Lasso statistic is less than or equal to the internal control limit i.e. $MDS-ATL_t \in [0, UCL_{inner}]$. On the other hand, if the chart statistic exceeds the outer control limit, or equivalently $MDS-ATL_t \in [UCL_{outer}, \infty)$, an out-of-control situation is triggered. Then, the production process is stopped to implement necessary corrective actions. If neither of the two mentioned situations occurs, i.e. $MDS-ATL_t \in (UCL_{inner}, UCL_{outer})$, the final decision regarding the process dispersion is postponed until the additional r samples are taken. In this case, the process variability is deemed to be in-control if the adaptive thresholding Lasso statistic associated with each additional *r* samples fall within the acceptance zone. In other words, if the chart statistic for any of the additional r samples exceeds UCL_{inner}, the process is out-of-control.

The MDS-ATL chart employs a rigorous approach to declare the covariance matrix as being in control when the chart statistic falls within the warning region. To address this issue and enhance chart's flexibility, we propose using a control chart called the GMDS-ATL. In this chart, we define an additional parameter q, as the generalization parameter, alongside the four previous parameters *n*, *r*, UCL_{inner}, and UCL_{outer}. The difference between the GMDS chart and MDS-ATL lies in the condition where the adaptive thresholding Lasso statistic falls within the warning area. In such cases, the process is considered in-control if both of the following conditions are fulfilled: (1) The chart statistic corresponding to at least q samples from r additional samples is less than or equal to UCL_{inner}; (2) the chart statistic for r additional samples are all smaller than UCL_{outer}.

IV. Performance evaluation

In this section, the sensitivity of the proposed MDS-ATL and GMDS-ATL control charts to the covariance matrix deviations of high-dimensional processes is compared with the adaptive thresholding Lasso chart in terms of *ARL*, *SDRL* and *MRL* metrics. To illustrate this, a numerical example based on simulation is presented, wherein the quality of the product or service is assessed through $p=15$ normally distributed variables. When the process is statistically in-control, the mean vector and covariance matrix of the study variables are considered to be equal to $\boldsymbol{\mu}_{ic} = \mathbf{0}_{15\times1}$ and $\boldsymbol{\Sigma}_{ic} = \mathbf{I}_{15}$, respectively. In all ATL, MDS-ATL and GMDS-ATL control charts, the regularization and shrinkage parameters of adaptive thresholding Lasso statistic are set to $\varphi=1$ and $\eta=4$, respectively, and samples of size $n=10$ are used to monitor the process dispersion. Furthermore, the repetition parameter for the MDS-RPLR and GMDS-RPLR charts is set to $r=3$, and two values of $q \in \{1,2\}$ are utilized for the generalization parameter. To ensure a fair comparison, the UCL, UCL_{inner} and UCL_{outer} values of the ATL-based control charting methods are determined through 20,000 iterations of Monte Carlo simulation in such a way that the in-control ARL of all charts is equal to the predefined value of $ARL_{ic} = 200$. The upper control limits for the competing control charting schemes to achieve $ARL_{ic} = 200$ are reported in Table 2.

Table 2. *UCL values*

Chart		UCL_{inner}	UCL_{outer}
ATL		3.6710	3.6710
MDS-ATL		1.7325	3.8002
GMDS-ATL	$q=1$	1.3346	3.9495
	$q=2$	1.0126	3.9910

Next, we will introduce seven out-of-control scenarios for the components of the covariance matrix in order to compare the sensitivity of the competing variability charts in response to process disturbances. It is important to note that the defined out-of-control scenarios are divided into three categories: diagonal, off-diagonal, and diagonal/off-diagonal. The first two scenarios are of the diagonal type, while the third and fourth scenarios belong to the off-diagonal category. Finally, the remaining three scenarios are categorized as diagonal/off-diagonal.

Scenario 1. In this scenario according to Equation (11), the assignable cause has no effect on the covariance components, while the variance of each of the 15 quality characteristics under study deviates equally from their target values.

$$\mathbf{\Sigma}_{oc_1} = \begin{bmatrix} 1+\Delta & 0 & \dots & 0 \\ 0 & 1+\Delta & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1+\Delta \end{bmatrix}_{15\times15} \tag{11}$$

Scenario 2. In this scenario, according to Equation (12), the variance of the first quality characteristic is increased by Δ units from its target value. As a result, scenario 2 will lead to a sparse deviation in the process covariance matrix.

$$\mathbf{\Sigma}_{oc_2} = \begin{bmatrix} 1+\Delta & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{15\times15} \tag{12}$$

Scenario 3. In this off-diagonal scenario, according to Equation (13), the covariance values associated with 70% of the quality characteristics increase from 1 to $1+\Delta$.

$$\mathbf{\Sigma}_{oc_3} = \begin{bmatrix} 1 & \Delta & \dots & \Delta & 0 & \cdots & 0 \\ \Delta & 1 & \dots & \Delta & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Delta & \Delta & \dots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \dots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \cdots & 1 \end{bmatrix}_{15\times15} \tag{13}$$

Scenario 4. The main difference between this scenario and scenario 3 is that only the covariance of the first three variables deviate from their nominal value. Given that only 2.66% of all the elements of the covariance matrix change under this scenario, it can be characterized as a sparse pattern.

$$\mathbf{\Sigma}_{oc_4} = \begin{bmatrix} 1 & \Delta & \Delta & 0 & \cdots & 0 \\ \Delta & 1 & \Delta & 0 & \cdots & 0 \\ \Delta & \Delta & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{15\times15} \tag{14}$$

Scenario 5. As indicated in Equation (15), the fifth scenario represents a diagonal/off-diagonal out-of-control pattern wherein the variance and covariance components associated with 30% of quality characteristics deviate from their nominal values by Δ units.

$$\mathbf{\Sigma}_{oc_5} = \begin{bmatrix} 1+\Delta & \Delta & \Delta & \Delta & 0 & \cdots & 0 \\ \Delta & 1+\Delta & \Delta & \Delta & 0 & \cdots & 0 \\ \Delta & \Delta & 1+\Delta & \Delta & \vdots & \ddots & \vdots \\ \Delta & \Delta & \Delta & 1+\Delta & 0 & \cdots & 0 \\ 0 & 0 & \dots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \cdots & 1 \end{bmatrix}_{15\times15} \tag{15}$$

Scenario 6. In this scenario, the first three variables are impacted by the source of deviation, so that each of the variance and covariance components increase by Δ and Δ^2 units, respectively.

$$\mathbf{\Sigma}_{oc_6} = \begin{bmatrix} 1+\Delta^2 & \Delta & \Delta & 0 & \cdots & 0 \\ \Delta & 1+\Delta^2 & \Delta & 0 & \cdots & 0 \\ \Delta & \Delta & 1+\Delta^2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{15\times15} \tag{16}$$

Scenario 7. Based on Equation (17), the variance and covariance components corresponding to the first three variables, and the last three variables change by the amount of Δ units. However, the components related to the 4th to 12th quality characteristics have remained unchanged.

$$\mathbf{\Sigma}_{oc_7} = \begin{bmatrix} 1+\Delta & \Delta & \Delta & 0 & \dots & 0 & 0 & 0 & 0 \\ \Delta & 1+\Delta & \Delta & 0 & \dots & 0 & 0 & 0 & 0 \\ \Delta & \Delta & 1+\Delta & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1+\Delta & \Delta & \Delta \\ 0 & 0 & 0 & 0 & \dots & 0 & \Delta & 1+\Delta & \Delta \\ 0 & 0 & 0 & 0 & \dots & 0 & \Delta & \Delta & 1+\Delta \end{bmatrix} \tag{17}$$

The *ARL*, *SDRL*, and *MRL* values for the ATL, competing charts under $\Delta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75\}$ are summarized in Tables 3-9. It is evident that in all scenarios and shift magnitudes, the MDS-ATL and GMDS-ATL charts show significantly better performance than the ATL in timely detecting the covariance matrix disturbances. That is to say, incorporating MDS and GMDS techniques into the ATL chart significantly enhances its performance in terms of *ARL*, *SDRL*, and *MRL* metrics. In particular, the values of *ARL*, *SDRL*, and *MRL* indices under diagonal changes based on 20,000 replicates are reported in Tables 3 and 4. Table 3 shows that the

GMDS-ATL scheme has superior performance compared to the ATL and MDS-ATL charts. As seen, when Δ is equal to 0.1, 0.2, 0.3, 0.4, 0.5, and 0.75, equipping the ATL chart with the MDS technique leads to improvements in its power by 36.30%, 60.57%, 62.70%, 49.92%, 31.84%, and 6.16%, respectively. This means that in the first scenario, the average sensitivity improvement of the ATL chart in terms of the *ARL* index, when using the MDS technique, is 41.25% across different values of Δ. The findings from Table 3 demonstrate that implementing the MDS sampling technique in the design of the ATL control chart enhances its *SDRL* index by 35.63%, 61.57%, 75.82%, 84.03%, 89.11%, and 100% when Δ is equal to 0.1, 0.2, 0.3, 0.4, 0.5, and 0.75, respectively.

Table 3. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 1*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	31.3154	8.7568	3.6496	2.1015	1.4789	1.0657
		SDRL	201.2460	30.8788	8.1388	3.1113	1.5094	0.8413	0.2640
		MRL	138	22	6	3	2	1	1
MDS-ATL		*ARL*	199.1255	19.9484	3.4527	1.3611	1.0525	1.0080	1.0000
		SDRL	198.4491	19.8767	3.1274	0.7522	0.2410	0.0916	0
		MRL	138	14	2	1	1	1	1
GMDS-ATL	$q=1$	*ARL*	199.9991	17.8049	2.7309	1.0772	1.0441	1.0000	1.0000
		SDRL	197.4807	17.3989	2.3829	0.5795	0.1678	0.0638	0
		MRL	138	12	2	1	1	1	1
	$q=2$	*ARL*	198.6688	17.6665	2.9177	1.1110	1.0518	1.0000	1.0000
		SDRL	199.5893	17.3772	2.4720	0.5972	0.1628	0.0619	0
		MRL	138	12	2	1	1	1	1

Table 4 shows that in the second scenario, equipping the ATL chart with the MDS technique under $\Delta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.75$ leads to improvements of 3.64, 6.11, 8.59, 11.07, 14.80 and 23.45 percent in the *ARL*. The improvement percentages by employing the GMDS technique, with $q=1$, are 10.86%, 15.15%, 28.51%, 34.12%, 36.89%, and 43.27%, respectively. Hence, as the shift magnitude increases, the percentage of improvement in the ATL scheme from the application of both MDS and GMDS techniques also increases. Furthermore, both the *SDRL* and *MRL* indices demonstrate a similar trend as the *ARL*. The main finding from comparing Tables 3 and 4 is that using MDS and GMDS techniques for detecting non-sparse diagonal shifts (scenario 1) has a considerably greater influence on the performance of the ATL chart compared to sparse diagonal shifts (scenario 2).

Table 4. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 2*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	173.9385	147.6309	124.5177	104.2381	88.0053	55.6005
		SDRL	201.2460	175.1907	145.4316	123.1942	104.4754	88.9015	55.0208
		MRL	138	121	103	87	72	60	39
MDS-ATL		*ARL*	199.1255	167.5999	138.6120	113.8206	92.6929	74.9781	42.5623
		SDRL	198.4491	168.7403	139.5496	114.3189	95.0044	75.6889	42.7268
		MRL	138	115	95	79	64	52	30
GMDS-ATL	$q=1$	*ARL*	199.9991	155.0499	125.2636	89.0154	68.6665	55.5435	31.5425
		SDRL	197.4807	147.4571	128.7383	89.7331	65.1677	51.9184	26.8687
		MRL	138	109	88	65	51	41	24
	$q=2$	*ARL*	198.6688	154.9197	125.2593	91.2727	69.2921	56.0495	30.1794
		SDRL	199.5893	157.0294	124.2078	93.0953	66.7335	53.1658	23.6370
		MRL	138	109	88	67	52	42	23

Table 5 indicates that in terms of *ARL* index, the MDS-ATL chart outperforms the ATL at various values of Δ (0.1, 0.2, 0.3, 0.4, 0.5, and 0.75), with percentages of 7.58, 18.45, 30.48, 37.38, 41.21, and 39.89. Additionally, when the GMDS technique is implemented with $q=1$, the performance of the ATL chart improves by 7.92, 21.16, 34.86, 43.39, 46.86, and 55.76 percent, respectively, under the mentioned values of Δ. As can be seen, the positive effect of using the MDS and GMDS techniques on the power of the ATL chart becomes more noticeable with the increase of Δ. A similar trend can be observed in *SDRL* and *MRL* indices produced by the MDS and GMDS charts.

Table 5. *RL* features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 3

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	120.7972	40.3205	15.9755	8.2071	5.1151	2.4923
		SDRL	201.2460	120.7832	39.4855	15.2993	7.6347	4.5666	1.9078
		MRL	138	83	28	11	6	4	2
MDS-ATL		*ARL*	199.1255	111.6419	32.8822	11.1054	5.1396	3.0072	1.4980
		SDRL	198.4491	111.2441	33.2822	11.2720	5.0651	2.7762	1.0197
		MRL	138	78	22	8	3	2	1
GMDS-ATL	$q=1$	*ARL*	199.9991	111.2271	31.7886	10.4066	4.6458	2.7183	1.1026
		SDRL	197.4807	111.1097	31.4771	10.5509	4.6317	2.5387	0.7109
		MRL	138	77	21	8	3	2	1
	$q=2$	*ARL*	198.6688	111.2013	32.0921	10.2317	4.4474	2.6022	1.0711
		SDRL	199.5893	111.2748	31.1282	10.4600	4.3239	2.3699	0.7083
		MRL	138	77	21	7	3	2	1

Table 6 indicates that the GMDS-ATL scheme demonstrates superior performance compared to the MDS-ATL and ATL charts in terms of all *ARL*, *SDRL*, and *MRL* metrics when the assignable cause leads to off-diagonal disturbances according to the fourth out-of-control scenario. Similar to the previous scenario, the performance improvement of the ATL chart through the utilization of MDS and GMDS strategies will be more significant as the shift magnitude increases. For example, when $\Delta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.75$, the MDS-ATL chart demonstrates superiority over ATL in terms of the *ARL* index by 1.15%, 5.47%, 7.92%, 13.93%, 19.36%, and 30.17% respectively. The the *ARL* improvement achieved by implementing the GMDS strategy in the ATL chart for the mentioned shifts, is 2.72%, 17.30%, 24.02%, 47.44%, 50.75%, and 36.34% when $q=1$ while in the case of $q=2$, the *ARL* improvements are 4.16%, 16.92%, 28.19%, 54.93%, 57.78%, and 38.37%. In other words, under this off-diagonal pattern, increasing q from 1 to 2 will lead to improved detectability of the GMDS-ATL chart. By comparing the results in Tables 5 and 6, it is observed that the improvement percentage of all three *RL*-based indices is more tangible when ATL chart is equipped with MDS and GMDS sampling approaches in detection of an off-diagonal non-sparse pattern (as in the third scenario) compared to the off-diagonal sparse pattern (as in the fourth scenario).

Table 6. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 4*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	187.9099	155.3939	111.0139	75.9599	51.0764	20.6715
		SDRL	201.2460	186.4131	154.5071	111.4313	75.0029	50.9071	20.2622
		MRL	138	131	108	77	53	35	14
MDS-ATL		*ARL*	199.1255	185.7435	146.8887	102.2206	65.3805	41.1865	14.4354
		SDRL	198.4491	186.5648	147.2405	102.0745	65.2222	41.3985	14.5201
		MRL	138	129	102	70	45	29	10
GMDS-ATL	$q = 1$	*ARL*	199.9991	182.8032	128.5112	84.3483	39.9277	25.1525	13.1602
		SDRL	197.4807	185.3521	129.5276	85.2083	37.6359	23.8887	8.4463
		MRL	138	127	90	58	28	18	9
	$q = 2$	*ARL*	198.6688	180.0868	129.0974	79.7237	34.2348	21.5663	12.7393
		SDRL	199.5893	176.4070	127.7548	79.4708	32.3020	20.5031	5.7574
		MRL	138	125	90	55	25	16	9

Tables 7 to 9 present the run length properties of the ATL, MDS-ATL, and GMDS-ATL control charting methods for joint diagonal/off-diagonal disturbances in high-dimensional covariance matrices. The results confirm that using the MDS and GMDS control charts, instead of the ATL, will significantly lead to faster detection of co-diagonal/off-diagonal changes in the covariance matrix components. The results presented in Tables 7 to 9 reveal that as the shift magnitudes increase, the superiority of MDS-ATL and GMDS-ATL compared to the ATL control chart becomes more evident.

Table 7. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 5*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	98.9288	44.2593	20.4434	11.0304	6.8975	3.1627
		SDRL	201.2460	98.1537	43.3755	20.2059	10.4759	6.3302	2.5983
		MRL	138	69	31	14	8	5	2
MDS-ATL		*ARL*	199.1255	86.9386	32.4176	12.7290	5.8847	3.2767	1.4862
		SDRL	198.4491	86.4191	32.6108	12.6924	5.6984	3.0323	0.9624
		MRL	138	61	22	9	4	2	1
GMDS-ATL	$q = 1$	*ARL*	199.9991	86.2533	31.0648	12.5609	5.4242	3.0203	1.2224
		SDRL	197.4807	84.1714	30.9066	12.4237	5.2813	2.8104	0.7855
		MRL	138	61	21	9	4	2	1
	$q = 2$	*ARL*	198.6688	86.7379	32.3513	12.8048	5.6068	3.1220	1.2655
		SDRL	199.5893	88.8912	32.1598	12.7473	5.3554	2.8498	0.8444
		MRL	138	61	22	9	4	2	1

Table 8. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 6*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	178.3261	128.3747	78.7263	43.6322	24.2064	7.2093
		SDRL	201.2460	178.1352	128.3269	76.9652	43.3895	23.6536	6.7377
		MRL	138	124	89	55	30	17	5
MDS-ATL		*ARL*	199.1255	176.7739	120.3104	66.4980	33.5460	16.5317	3.6357
		SDRL	198.4491	177.6378	120.3849	66.5117	33.6048	16.5764	3.3766
		MRL	138	123	84	46	23	11	2
GMDS-ATL	$q = 1$	*ARL*	199.9991	168.6586	108.5676	57.8567	27.8740	13.7365	2.8737
		SDRL	197.4807	169.3584	105.8655	57.3664	29.0077	14.3088	2.5912
		MRL	138	118	76	40	19	9	2
	$q = 2$	*ARL*	198.6688	176.2301	114.1713	59.7670	29.2099	14.3948	3.1825
		SDRL	199.5893	166.7157	114.6325	60.3128	29.7811	14.6903	2.7198
		MRL	138	121	79	41	19	9	2

Table 9. *RL features of ATL, MDS-ATL, and GMDS-ATL charts under scenario 7*

Chart		Index	Δ						
			0	0.1	0.2	0.3	0.4	0.5	0.75
ATL		*ARL*	200.6506	78.8354	30.6467	13.7325	7.1715	4.4884	2.1488
		SDRL	201.2460	77.8354	30.0446	13.3228	6.7644	3.9699	1.5554
		MRL	138	55	21	10	5	3	2
MDS-ATL		*ARL*	199.1255	65.0101	20.0844	6.9717	3.1433	1.8793	1.1167
		SDRL	198.4491	64.9520	20.0714	6.8471	2.8302	1.4152	0.3920
		MRL	138	45	14	5	2	1	1
GMDS-ATL	$q = 1$	*ARL*	199.9991	64.3264	20.0155	6.7645	3.1169	1.8635	1.0000
		SDRL	197.4807	63.8176	20.2910	6.8467	2.8459	1.4227	0.3619
		MRL	138	44	14	5	2	1	1
	$q = 2$	*ARL*	198.6688	65.6486	19.8908	7.1158	3.1635	1.8914	1.0241
		SDRL	199.5893	64.4291	20.7047	6.9845	2.7734	1.3868	0.3694
		MRL	138	45	13	5	2	1	1

V. Conclusion remarks

Manufacturers often prefer using small sample sizes to reduce production costs, while also needing to consider a wide range of quality characteristics in order to enhance their market share. Therefore, in today's competitive world, encountering high-dimensional data where the sample size is smaller than the number of quality characteristics, has become a significant challenge in various industrial and service applications. Due to the singularity of the sample covariance matrix, conventional multivariate statistical methods cannot be used to monitor the process dispersion under high-dimensionality. On the other hand, in recent years, multiple dependent state sampling has been effectively utilized to enhance the effectiveness of control charts, primarily for the purpose of monitoring the mean parameter of univariate processes. As a result, in this paper, we extended two adaptive thresholding Lasso control charts for the rapid detection of covariance matrix disturbances in high-dimensional process monitoring using multiple dependent state and generalized multiple dependent state sampling techniques. The performance of the improved MDS-ATL and GMDS-ATL control charting methods was compared to the conventional adaptive thresholding Lasso through Monte Carlo simulation, using three metrics of *ARL*, *SDRL* and *MRL*. The simulation results demonstrate that the utilization of MDS and GMDS methods in design of the ATL control chart effectively enhances the detection of covariance matrix disturbances in high-dimensional processes. Particularly when confronted with assignable causes that result in significant deviations in the components of the covariance matrix, the incorporation of MDS and GMDS methods in the design of the ATL chart becomes even more crucial. One assumption of this article is that the assignable cause solely impacts the components of the covariance matrix. However, there are instances where the deviation source can also result in a simultaneous change in the mean vector and the covariance matrix of high-dimensional processes. To overcome this limitation, it is recommended to employ MDS and GMDS techniques to monitor the coefficient of variation in high-dimensional processes.

References

[1] Wang, K. and Jiang, W. (2009). High-dimensional process monitoring and fault isolation via variable selection. Journal of Quality Technology, 41(3): 247-258.

[2] Abdella, G. M., Al-Khalifa, K. N., Kim, S., Jeong, M. K., Elsayed, E. A. and Hamouda, A. M. (2017). Variable selection-based multivariate cumulative sum control chart. Quality and Reliability Engineering International, 33(3):565-578.

[3] Sangahn, K. I. M. (2019). Variable selection-based SPC procedures for high-dimensional multistage processes. Journal of Systems Engineering and Electronics, 30(1):144-153.

[4] Zhang, S., Xue, L., He, Z., Liu, Y. and Xin, Z. (2023). A sensitized variable selection control chart based on a classification algorithm for monitoring high-dimensional processes. Quality and Reliability Engineering International, 39(7):2837-2850.

[5] Abdella, G. M., Kim, J., Kim, S., Al-Khalifa, K. N., Jeong, M. K., Hamouda, A. M. and Elsayed, E. A. (2019). An adaptive thresholding-based process variability monitoring. Journal of Quality Technology, 51(3):242-256.

[6] Abdella, G. M., Maleki, M. R., Kim, S., Al-Khalifa, K. N. and Hamouda, A. M. S. (2020). Phase-I monitoring of high-dimensional covariance matrix using an adaptive thresholding LASSO rule. Computers & Industrial Engineering, 144, 106465.

[7] Jafari, M., Maleki, M. R. and Salmasnia, A. (2023). A high-dimensional control chart for monitoring process variability under gauge imprecision effect. Production Engineering, 17(3):547-564.

[8] Jalilibal, Z., Karavigh, M. H. A., Maleki, M. R. and Amiri, A. (2024). Control charting methods for monitoring high dimensional data streams: A conceptual classification scheme. Computers & Industrial Engineering, 191, 110141.

[9] Salmasnia, A., Maleki, M. R. and Mirzaei, M. (2025). Double Sampling Adaptive Thresholding LASSO Variability Chart for Phase II Monitoring of High-Dimensional Data Streams. Journal of Industrial Integration and Management, In Press, Doi: 10.1142/S242486222350001X.

[10] Salmasnia, A., Maleki, M. R. and Niaki, S. T. A. (2018). Remedial measures to lessen the effect of imprecise measurement with linearly increasing variance on the performance of the MAX-EWMAMS scheme. Arabian Journal for Science and Engineering, 43:3151-3162.

[11] Shaheen, U., Azam, M. and Aslam, M. (2020). A control chart for monitoring the lognormal process variation using repetitive sampling. Quality and Reliability Engineering International, 36(3):1028-1047.

[12] Saemian, M., Salmasnia, A. and Maleki, M. R. (2025). A generalized multiple dependent state sampling chart based on ridge penalized likelihood ratio for high-dimensional covariance

matrix monitoring. Scientia Iranica, In Press, Doi: 10.24200/SCI.2022.60169.6640.

[13] Arshad, A., Azam, M., Aslam, M. and Jun, C. H. (2018). A control chart for monitoring process variation using multiple dependent state sampling. Communications in Statistics-Simulation and Computation, 47(8):2216-2233.

[14] Naveed, M., Azam, M., Khan, N. and Aslam, M. (2020). Designing a control chart of extended EWMA statistic based on multiple dependent state sampling. Journal of Applied Statistics, 47(8):1482-1492.

[15] Aslam, M., Raza, M. A. and Jun, C. H. (2020). A new variable control chart under generalized multiple dependent state sampling. Communications in Statistics-Simulation and Computation, 49(9):2321-2332.

[16] Rao, G. S., Aslam, M., Alamri, F. S. and Jun, C. H. (2024). Comparing the efficacy of coefficient of variation control charts using generalized multiple dependent state sampling with various run-rule control charts. Scientific Reports, 14(1):2726.

[17] García-Bustos, S., León, J. and Pastuizaca, M. N. (2020). Hotelling T^2 chart using the generalized multiple dependent state sampling scheme. International Journal of Quality & Reliability Management, 38(6):1265-1277.

[18] Khan, N., Ahmad, L. and Aslam, M. (2022). Monitoring using X-bar control chart using neutrosophic-based generalized multiple dependent state sampling with application. International Journal of Computational Intelligence Systems, 15(1):73.

[19] Machado, M. A., Lee Ho, L., Quinino, R. C. and Celano, G. (2022). Monitoring the covariance matrix of bivariate processes with the DVMAX control charts. Applied Stochastic Models in Business and Industry, 38(1):116-132.

[20] Maleki, M. R., Salmasnia, A. and Yousefi, S. (2023). Multivariate ELR control chart with estimated mean vector and covariance matrix. Communications in Statistics-Theory and Methods, 52(24):8814-8827.

[21] Kim, J., Abdella, G. M., Kim, S., Al-Khalifa, K. N. and Hamouda, A. M. (2019). Control charts for variability monitoring in high-dimensional processes. *Computers & Industrial Engineering*, 130:309-316.

[22] Fan, J., Shu, L., Yang, A. and Li, Y. (2021). Phase I analysis of high-dimensional covariance matrices based on sparse leading eigenvalues. Journal of Quality Technology, 53(4):333-346.

[23] Safikhani, E., Salmasnia, A. and Maleki, M. R. (2023). A ridge penalized likelihood ratio chart for Phase II monitoring of high-dimensional process dispersion under measurement system inaccuracy. International Journal of Industrial Engineering, 34(2):1-17.