A new multivariate extended homogeneously weighted moving average monitoring scheme incorporated with a support vector machine

Luke Pieters, Jean-claude Malela-Majika, Schalk Human, Philippe Castagliola

A New Multivariate Extended Homogeneously Weighted Moving Average Monitoring Scheme Incorporated with a Support Vector Machine

Luke Pieters[1], Jean-Claude Malela-Majika[1], Schalk Human[1], and Philippe Castagliola[2]

[1]Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria, Hatfield, Pretoria, South Africa

[2]Nantes Université & LS2N UMR CNRS 6004, Nantes, France

September 7, 2023

Abstract

Statistical process monitoring (SPM) provides a way, by means of monitoring schemes, to be alerted when a process significantly changes, and to initiate investigations for potential causes of variation. However, these schemes can only determine if the process is stable or unstable and they cannot provide any additional information. Several researchers have recommended the integration of machine learning (ML) approaches into SPM as a solution for the shortfalls of the techniques used to construct traditional monitoring schemes. This paper introduces a new multivariate extended homogeneously weighted moving average (MEHWMA) monitoring scheme and investigates its performance in terms of the run-length distribution using simulation. In addition, the proposed MEHWMA scheme is integrated with a support vector machine to allow for the classification of the out-of-control events which facilitates the identification of the root causes of variation in the process. A numerical illustrative example is provided using real-life data.

Keywords: Enhanced monitoring scheme; Extended HWMA; Machine learning; Performance metrics; Multivariate process; Support vector machine.

1 Introduction

In today's competitive markets, the quality of a product or a service is the key that gives a significant advantage over other competitors; see for example Montgomery (2020). To keep a business ahead of the competition, it is required to continuously monitor the production, manufacturing and business processes. This can be accomplished by using statistical process monitoring (SPM) tools.The most popular tool used in SPM is the control chart (or monitoring scheme) with examples being the Shewhart, cumulative sum (CUSUM), exponentially weighted moving average (EWMA) and homogeneously weighted moving average (HWMA) schemes (Roberts (2000), Abbas (2018) and Montgomery (2020)). A modern monitoring scheme is a graphical representation that has a charting statistic on the vertical axis and the sample number or time on the horizontal axis. In addition, it has two main thresholds, namely the lower and upper control limits (LCL and UCL), that divide the vertical axis into two regions: the in-control (IC) region (within the control limits) and the out-of-control (OOC) region (on or beyond the control limits).

Monitoring schemes can be classified based on the types of data, charting statistic, distributional assumptions, monitoring strategies, and number of quality characteristics. Depending on the number of quality characteristics, monitoring schemes are divided into either univariate or multivariate schemes. A univariate monitoring scheme is used to monitor a single quality characteristic, while a multivariate scheme monitors two or more quality characteristics. More details on univariate and multivariate monitoring schemes can be found in Maravelakis et al. (2002), Bersimis et al. (2007) and Malela-Majika et al. (2022a,b). A monitoring scheme can additionally be grouped as either a memoryless or a memory-type scheme. The former uses only the most recent information to decide whether the process is IC or OOC. The latter, on the other hand, incorporates current as well as past information to decide whether the process is IC or not; see for example Adegoke et al. (2019). Memoryless-type schemes are more efficient in detecting large shifts in the process parameter, while memory-type schemes are better at detecting small and moderate shifts. Among the existing memory-type schemes, the cumulative sum (CUSUM) and the EWMA schemes are the most popular in SPM (see Page (1961) and Roberts (2000)). Naveed et al. (2018) proposed the extended EWMA (EEWMA) scheme for monitoring the process mean. They reported that the EEWMA scheme outperforms both the EWMA and Shewhart schemes in detecting up- or downwards step shifts in the mean. More recently, Abbas (2018) introduced the HWMA scheme to improve the detection ability of the EWMA scheme in monitoring small shifts. The HWMA scheme assigns a certain weight to the most recent observation and the rest of the weight is assigned homogeneously, i.e. equally, between all previous observations. Therefore, the HWMA scheme has been shown to be superior to the Shewhart and EWMA schemes in detecting small sustained shifts. It is generally more beneficial to use more efficient schemes such as memory-type schemes as they allow faster detection of OOC shifts in the process parameters and they give the chance to get the process back IC before the quality of the products or services suffers too heavily. For a recent account on the improvement of memory-type schemes, readers are referred to Chan et al. (2021), Anwar et al. (2022) and Anwer et al. (2022).

The aforementioned schemes allow for the detection of any abnormality in the process by giving a signal that initiates an investigation of the causes of variation. However, these schemes cannot provide further information on what caused the OOC signal, the size of the shift or if there are any special patterns present in the data. These shortfalls become very apparent in the multivariate process case as multiple quality variables of a process are combined into a single charting statistic. Multivariate monitoring schemes are also not able to identify which specific quality characteristic moved OOC; see Demircioglu Diren et al. (2020). To overcome this problem, one would suggest to combine a multivariate scheme with two or more univariate schemes. This, however, would increase the false alarm rate and increase the cost of inspection. Therefore, a more reliable solution would be to integrate a multivariate monitoring scheme with machine learning (ML) tools or techniques. Some of the ML techniques used recently in SPM include artificial neural networks (ANN), multivariate adaptive regression splines (MARS), and support vector machines (SVM); see Chowdhury and Janan (2020) and Yeganeh et al. (2021). Based on the works by Apsemidis and Psarakis (2020) as well as Chowdhury and Janan (2020), SVMs have been shown to be useful in SPM as they provide good results even when compared to other ML techniques such as the ANN, and thus SVMs will be used and integrated with the new monitoring scheme.

In this paper we introduce new univariate and multivariate extended homogeneously weighted moving average (EHWMA and MEHWMA) schemes for monitoring the location parameter of a process. In addition, since traditional multivariate monitoring schemes cannot identify which one of the variables (or quality characteristics) caused the OOC situation, the MEHWMA scheme is integrated with a SVM trained to determine which quality characteristic went OOC.

The remainder of this paper is structured as follows: In Section 2, a brief background is given on the existing EWMA, EEWMA , HWMA, and the multivariate HWMA (MHWMA) schemes. In Section 3, the EHWMA monitoring scheme is introduced along with its properties. The zero-state performance of the EHWMA scheme is evaluated for both the univariate and multivariate cases in Section 4. Section 5 investigates the zero and steady-state performances of the proposed schemes using the conditional expected delay. Section 6 provides details on how a SVM can be used in SPM. Moreover, this section provides an illustrative example based on real-life data where the MEHWMA scheme is integrated with a SVM. Lastly, Section 7 presents the concluding remarks.

2 Mathematical background

2.1 The EWMA scheme

The EWMA monitoring scheme was first introduced as the geometrically weighted moving average scheme and only later became known as the EWMA scheme; see Roberts (2000). It was designed to incorporate current information from the process as well as that from the previous charting statistic. Let X_t (for $t = 1, 2, 3...$) be a sequence of independently and identically distributed (i.i.d.) normal random variables from a process with IC mean μ_0 and standard deviation σ_0. The EWMA statistic at time t is given by:

$$Z_t = \phi X_t + (1 - \phi) Z_{t-1}, \tag{1}$$

where ϕ is the smoothing parameter with $\phi \in (0, 1]$ and a starting value of $Z_0 = \mu_0$. The expected value and variance of the EWMA statistic defined in (1) are given by:

$$E\left(Z_t\right) = \mu_0 \tag{2}$$

and

$$V(Z_t) = \sigma_0^2 \left(\frac{\phi}{2 - \phi}\right) \left(1 - (1 - \phi)^{2t}\right), \tag{3}$$

respectively. The EWMA control limits are defined by:

$$\mathrm{UCL}_{Z_t}/\mathrm{LCL}_{Z_t} = E(Z_t) \pm L_Z \sqrt{V(Z_t)}, \tag{4}$$

where L_Z is the control limit constant chosen such that the attained IC average run-length (ARL) is equal to a pre-specified IC ARL (denoted by ARL_0). The EWMA scheme gives a signal if the charting statistic defined in (1) plots on or beyond the control limits defined in (4).

2.2 The extended EWMA scheme

More recently, Naveed et al. (2018) introduced the EEWMA scheme by extending the EWMA with an additional weighting parameter allowing the charting statistic to give a positive weight to the most recent data and negative weights to the oldest ones. The addition of the new parameter resulted in a lower variance than that of the conventional EWMA statistic. Thus, the charting statistic of the EEWMA scheme at the t^{th} sampling time is defined as follows:

$$\mathrm{EZ}_t = \phi_1 X_t - \phi_2 X_{t-1} + (1 - \phi_1 + \phi_2)\mathrm{EZ}_{t-1}, \tag{5}$$

where $\phi_1 \in (0,1]$ and $\phi_2 \in [0,\phi_1)$ are the smoothing parameters, $X_0 = \text{EZ}_0 = \mu_0$, and the expected value and variance of the EEWMA statistic are equal to:

$$E\left(\text{EZ}_t\right) = \mu_0 \tag{6}$$

and

$$\begin{aligned} V(\text{EZ}_t) = \sigma_0^2 \Bigg[& (\phi_1^2+\phi_2^2)\left(\frac{1-(1-\phi_1+\phi_2)^{2t}}{2(\phi_1-\phi_2)-(\phi_1-\phi_2)^2}\right) \\ & -2\phi_1\phi_2(1-\phi_1+\phi_2)\left(\frac{1-(1-\phi_1+\phi_2)^{2(t-1)}}{2(\phi_1-\phi_2)-(\phi_1-\phi_2)^2}\right)\Bigg], \end{aligned} \tag{7}$$

respectively. If $\phi_2 = 0$, then the EEWMA monitoring scheme reduces to the EWMA scheme and therefore $\text{EZ}_t = Z_t$. The control limits of the EEWMA scheme are defined by:

$$\text{UCL}_{\text{EZ}_t}/\text{LCL}_{\text{EZ}_t} = E(\text{EZ}_t) \pm L_{\text{EZ}}\sqrt{V(\text{EZ}_t)}, \tag{8}$$

where L_{EZ} represents the control limit constant of the EEWMA scheme. Thus, the EEWMA scheme gives a signal if the charting statistic defined in (5) plots on or beyond the control limits defined in (8).

2.3 The HWMA scheme

The HWMA monitoring scheme was proposed in order to improve the sensitivity of the EWMA scheme in detecting very small shifts; see Abbas (2018). Thus, the design of the HWMA scheme is similar to that of the EWMA scheme, except that instead of using the previous value of the charting statistic, the average of all the previous process observations is used. It also uses a similar weighting structure to the one of the EWMA scheme. Therefore, the HWMA statistic is defined as:

$$H_t = \phi X_t + (1-\phi)\bar{X}_{t-1}, \tag{9}$$

where $\phi \in (0,1]$ is the smoothing parameter, $\bar{X}_{t-1}$ is the average of the previous $t-1$ observations (i.e. $\bar{X}_{t-1} = \frac{1}{t-1}\sum_{k=1}^{t-1} X_k$) and the starting value is $\bar{X}_0 = \mu_0$. The mean and variance of the HWMA statistic are equal to:

$$E\left(H_t\right) = \mu_0 \tag{10}$$

and

$$V(H_t) = \begin{cases} \phi^2\sigma_0^2, & \text{for } t = 1 \\ \left(\phi^2 + \frac{(1-\phi)^2}{t-1}\right)\sigma_0^2, & \text{for } t > 1 \end{cases}, \tag{11}$$

respectively. Thus, the HWMA control limits are defined by:

$$\mathrm{UCL}_{H_t}/\mathrm{LCL}_{H_t} = E(H_t) \pm L_H\sqrt{V(H_t)}, \tag{12}$$

where L_H is the control limit constant chosen such that the attained IC ARL value is equal to the desired ARL_0 value. Therefore, the HWMA scheme gives a signal if the charting statistic defined in (9) plots beyond the control limits defined in (12).

2.4 The multivariate HWMA scheme

The HWMA statistic was later expanded to the multivariate case and denoted as MHWMA (Adegoke et al. (2019)). Let $\boldsymbol{y}_t$ $(t = 1, 2, ...)$ be a p-variate random vector of size n from a multivariate population with known mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$. When $n = 1$, the MHWMA statistic can be defined as:

$$\boldsymbol{H}_t = \phi\boldsymbol{y}_t + (1-\phi)\bar{\boldsymbol{y}}_{t-1}, \tag{13}$$

where $\bar{\boldsymbol{y}}_{t-1}$ is a $p \times 1$ vector where the elements are the averages of the previous $t-1$ observations of each p-variate, $\boldsymbol{H}_t$ is a $p \times 1$ vector and $\bar{\boldsymbol{y}}_0 = \boldsymbol{\mu}_0$. The mean vector and covariance matrix of $\boldsymbol{H}_t$ are given by:

$$E\left(\boldsymbol{H}_t\right) = \boldsymbol{\mu}_0 \tag{14}$$

and

$$\boldsymbol{\Sigma}_{\boldsymbol{H}_t} = \begin{cases} \phi^2\boldsymbol{\Sigma}_0, & \text{for } t = 1 \\ \left(\phi^2 + \frac{(1-\phi)^2}{t-1}\right)\boldsymbol{\Sigma}_0, & \text{for } t > 1 \end{cases}. \tag{15}$$

In the case where $p = 1$, the MHWMA statistic reduces to the HWMA statistic. If $p = 1$ and $\phi = 1$ it also reduces to the Shewhart $\bar{X}$ scheme. However, if $p > 1$ and $\phi = 1$, it reduces to the χ^2 monitoring scheme in the case where the parameters are assumed known. If the parameters are unknown, it reduces to the T^2 monitoring scheme.

For the three univariate monitoring schemes mentioned above, their UCL and LCL are calculated using the formula: $E\left(C\right) \pm L\sqrt{V(C)}$, where C is the charting statistic of the respective scheme and L is the corresponding control limit constant. However, for the MHWMA statistic the approach is different. Firstly, the T^2 value has to be calculated for $\boldsymbol{H}_t$ as

$$T_t^2 = (\boldsymbol{H}_t - \boldsymbol{\mu}_0)^{\intercal}\boldsymbol{\Sigma}_{H_t}^{-1}(\boldsymbol{H}_t - \boldsymbol{\mu}_0), \tag{16}$$

allowing the multivariate statistic to be converted to a scalar value that can easily be plotted and then the process is considered OOC if

$$T_t^2 > h,$$

where h is the control limit chosen such that the attained IC $ARL = ARL_0$.

3 Proposed univariate and multivariate EHWMA monitoring schemes

3.1 The univariate EHWMA scheme

The proposed EHWMA monitoring scheme is designed following the same approach as the one for the EEWMA scheme described in Section 2.2. An additional smoothing parameter will be added to the HWMA scheme to allow the resulting statistic (i.e. EHWMA statistic) to give negative weights to previous process observations and a positive weight to the most recent observations in order to potentially decrease the variance. If the second smoothing parameter is set to 0 then the statistic becomes equivalent to that of the HWMA statistic. The EHWMA statistic will also be expanded to the multivariate case in a similar way as described in Section 2.4.

For the univariate case, let X_t (for $t = 1, 2, 3...$) be a sequence of i.i.d. random observations from a normal population with IC process mean μ_0 and standard deviation σ_0.

Then, the EHWMA charting statistic at time t (denoted as EH_t) is defined as:

$$\text{EH}_t = \phi_1 X_t - \phi_2 X_{t-1} + (1 - \phi_1 + \phi_2)\bar{X}_{t-1}, \tag{17}$$

where ϕ_1, ϕ_2, X_0 and $\bar{X}_0$ are defined as in Section 2. When $\phi_2 = 0$, the EHWMA scheme is reduced to the HWMA scheme with smoothing parameter ϕ_1. When $\phi_1 = 1$ and $\phi_2 = 0$, the EHWMA scheme is reduced to the Shewhart scheme.

The expected value and variance of the proposed EHWMA statistic are given by:

$$E(\text{EH}_t) = \mu_0 \tag{18}$$

and

$$V(\text{EH}_t) = \begin{cases} \phi_1^2 \sigma_0^2, & \text{for } t = 1 \\ \left(\phi_1^2 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2 (t-2)\right)\sigma_0^2, & \text{for } t > 1 \end{cases}, \tag{19}$$

respectively. For more details on the derivations of the above properties of the EHWMA statistic, see Appendix A. The lower and upper control limits of the proposed EHWMA scheme are given by:

$$\text{UCL}_{\text{EH}_t}/\text{LCL}_{\text{EH}_t} = E(\text{EH}_t) \pm L_{\text{EH}}\sqrt{V(\text{EH}_t)} \tag{20}$$

respectively, where L_{EH} is the control limit constant selected such that the attained IC ARL is equal to some pre-specified ARL_0. The EHWMA scheme gives a signal at the sampling time t if the charting statistic defined in (17) plots beyond the control limits defined in (20).

3.2 The multivariate EHWMA scheme

For the multivariate case, let $\boldsymbol{y}_t : p \times 1$ $(t = 1, 2, ...)$ be a p-variate normal random vectors from a multivariate population with mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$.
Then, the MEHWMA statistic at time t (denoted as MEH_t) can be defined as:

$$\mathbf{MEH}_t = \phi_1 \boldsymbol{y}_t - \phi_2 \boldsymbol{y}_{t-1} + (1 - \phi_1 + \phi_2)\bar{\boldsymbol{y}}_{t-1}, \tag{21}$$

where $\boldsymbol{y}_0 = \bar{\boldsymbol{y}}_0 = \boldsymbol{\mu}_0$. The expected mean vector and covariance matrix of the MEHWMA statistic are mathematically defined by:

$$E\left(\mathbf{MEH}_t\right) = \boldsymbol{\mu}_0 \tag{22}$$

and

$$\boldsymbol{\Sigma}_{\text{MEH}_t} = \begin{cases} \phi_1^2 \boldsymbol{\Sigma}_0, & \text{for } t = 1 \\ \left(\phi_1^2 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2 (t-2)\right) \boldsymbol{\Sigma}_0, & \text{for } t > 1 \end{cases}, \tag{23}$$

respectively. The MEHWMA monitoring scheme will give an OOC signal if

$$T_t^2 = (\mathbf{MEH}_t - \boldsymbol{\mu}_0)^\intercal \boldsymbol{\Sigma}_{\text{MEH}_t}^{-1} (\mathbf{MEH}_t - \boldsymbol{\mu}_0) > h_{\text{MEH}}$$

where h_{MEH} is the control limit of the MEHWMA scheme selected such that the IC $ARL = ARL_0$.

The expected value and variance of the MEHWMA statistic are derived similarly to the ones of the EHWMA statistic. The optimal values of ϕ_1, ϕ_2, h_{MEH} and L_{EH} will be chosen to achieve a high desired ARL_0 value and a small OOC ARL (ARL_1) value.

4 Zero-state performance of the EHWMA and MEHWMA monitoring schemes

In this section, the zero-state performance and robustness of the EHWMA and MEHWMA schemes are investigated using extensive simulations. The simulations for the univariate cases are conducted as follows:

1. Set the mean and variance of the process to μ_0 and 1, respectively. When the process is IC, $\mu_0 = 0$. However, for the OOC case, the mean is set to $\mu_1 = \mu_0 + \delta = \delta$ where $\delta \neq 0$.

2. Set the value of the control limit constant.

3. Generate a sample of size n from a normal distribution with mean μ_1 and variance σ_0^2, and compute the sample mean, the charting statistic and control limits of the proposed scheme.

4. Compare the charting statistic to the control limits computed in Step 3.

5. Repeat steps 3 and 4 until the scheme detects an OOC event for the first time and record the number of samples needed before the OOC signal is triggered.

6. Repeat Steps 3 to 5 for a predetermined number of iterations, and calculate the ARL as the average number of samples taken before an OOC signal is triggered.

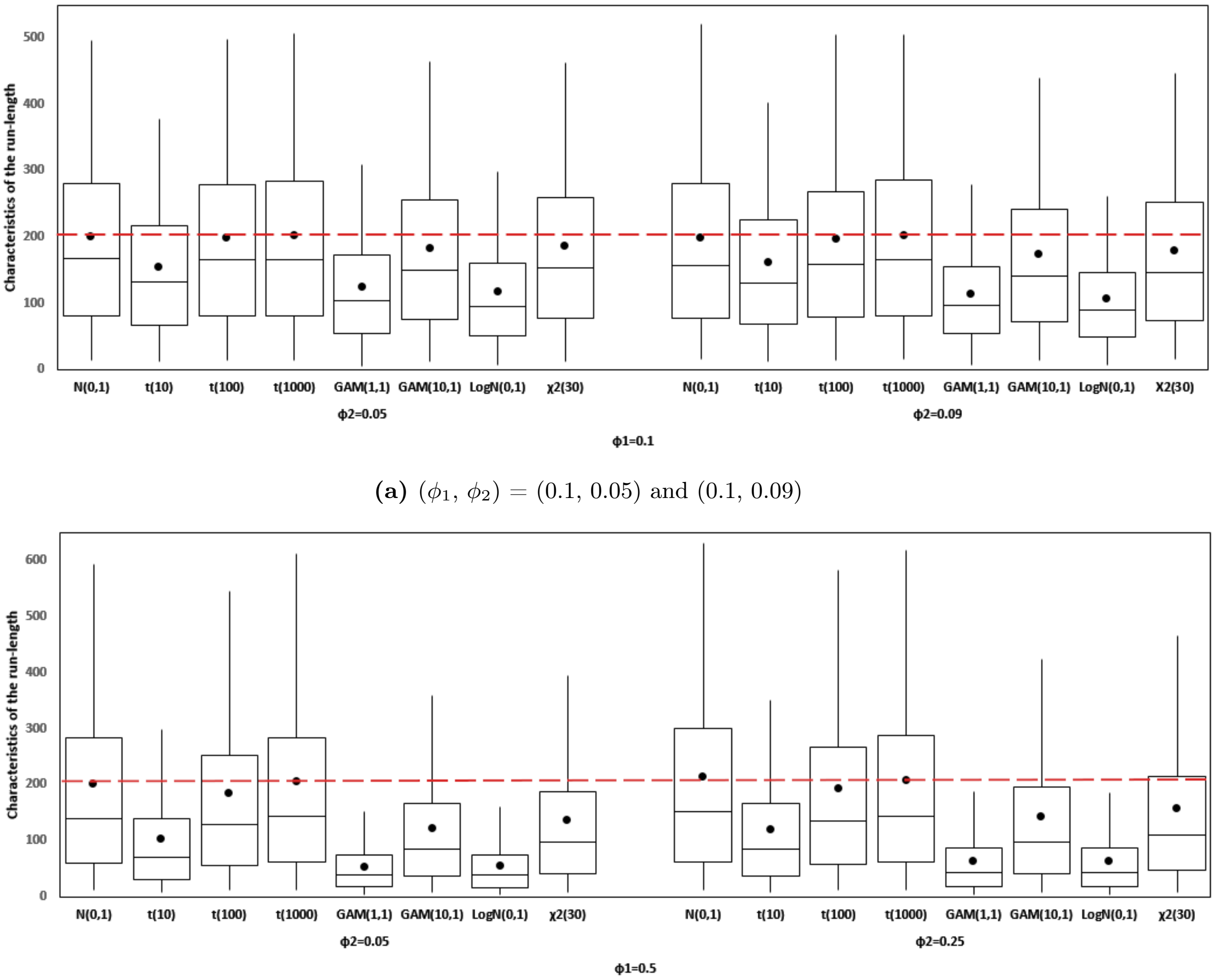
In this paper, we used 20000 iterations. The simulations for the multivariate cases are constructed in a similar fashion, where the samples are taken from a p-variate distribution instead of a univariate one, and the OOC mean is set to $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_0 \boldsymbol{\Delta}_p$, where $\boldsymbol{\mu}_0$ is a $p \times 1$ vector of zeros, $\boldsymbol{\mu}_1$ is a $p \times 1$ vector, $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_{p\times p}$ identity matrix, and $\boldsymbol{\Delta}_p$ is a $p \times 1$ vector with elements all equal to $\sqrt{\frac{\delta^2}{p}}$. For all of the following results and comparisons, the values of L_{EH} and h_{MEH} have been chosen to achieve an ARL_0 value of 200.

4.1 Results and discussion

4.1.1 Robustness of the proposed EHWMA and MEHWMA schemes

The proposed EHWMA and MEHWMA schemes are designed based on the assumption of normality. Recall that a monitoring scheme is said to be IC robust if the characteristics of the IC run-length, such as the ARL_0 values are the same (or almost the same) across all continuous probability distributions. In this paper, we considered the Student's t distribution with 10, 100 and 1000 degrees of freedom (denoted as $t(10)$, $t(100)$ and $t(1000)$, respectively), the gamma distribution with shape parameter α= 1 and 10, and scale parameter β= 1 (denoted as $GAM(1,1)$ and $GAM(10,1)$), the lognormal distribution with parameters $\mu = 0$ and $\sigma = 1$ (denoted as $LogN(0,1)$), and the Chi-squared distribution with 30 degrees of freedom (denoted as $\chi^2(30)$). The characteristics of the IC run-length distribution of the proposed EHWMA and MEHWMA schemes are displayed in Figures 1 and 2, respectively, using box-plots for different pairs of parameters ϕ_1 and ϕ_2. The value of L_{EH} has been chosen to achieve an ARL_0 value close to 200 for the normal case and thereafter used for the other distributions. Note that for each box-plot, the extreme values represent the 5th and 95th percentile of the run-length (PRL) while the lower, middle and upper horizontal lines represent the 25th, 50th (i.e. median) and 75th PRL values, respectively. The attained ARL_0 value is represented by a dot. From Figure 1, it can be seen that the attained ARL_0 value under the $t(10)$ distribution is much smaller than the desired ARL_0 for all pairs of ϕ_1 and ϕ_2 considered, but as the degrees of freedom increase, the ARL_0 value gets closer to that of the normal. This is expected as the t distribution tends to a normal distribution as the degrees of freedom increase. For the gamma and the lognormal distributions, the ARL_0 values are also much smaller than the one for the normal distribution. This is due to the heavier tails of these distributions compared to the normal distribution. The attained ARL_0 values for the Chi-squared distribution are also significantly smaller than the ARL_0 value of 200, but not as small as for the gamma and the lognormal distributions. Similar findings can be observed when examining the IC PRL (PRL_0) profiles, i.e. the PRL values for each percentile are not the same across all continuous probability distributions considered in this paper. When ϕ_2 is kept fixed, the proposed EHWMA scheme loses its IC robustness as ϕ_1 increases. However, when ϕ_1 is kept fixed, the attained ARL_0 of EHWMA scheme increases as ϕ_2 increases.

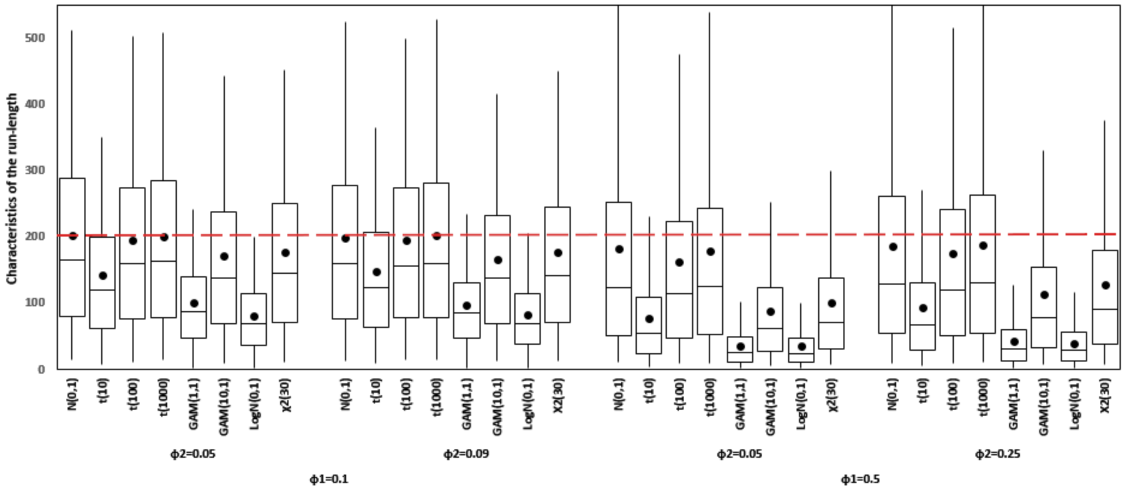
The results of the robustness of the MEHWMA scheme are given in Figures 2. The simulations for these cases were built such that the p-variate characteristics are sampled from p independent distributions. The distributions used were the same as the ones for the univariate cases. Here similar findings as the ones of the univariate case are observed. Again, the value of h_{MEH} was chosen to achieve a ARL_0 value of 200 for the normal case.



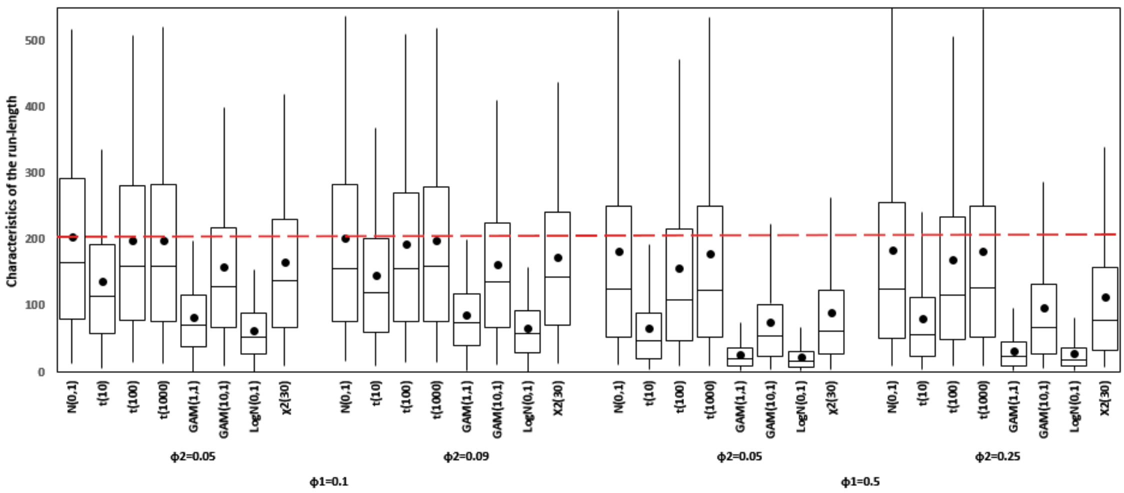
(a) $(\phi_1, \phi_2) = (0.1, 0.05)$ and $(0.1, 0.09)$

(b) $(\phi_1, \phi_2) = (0.5, 0.05)$ and $(0.5, 0.25)$

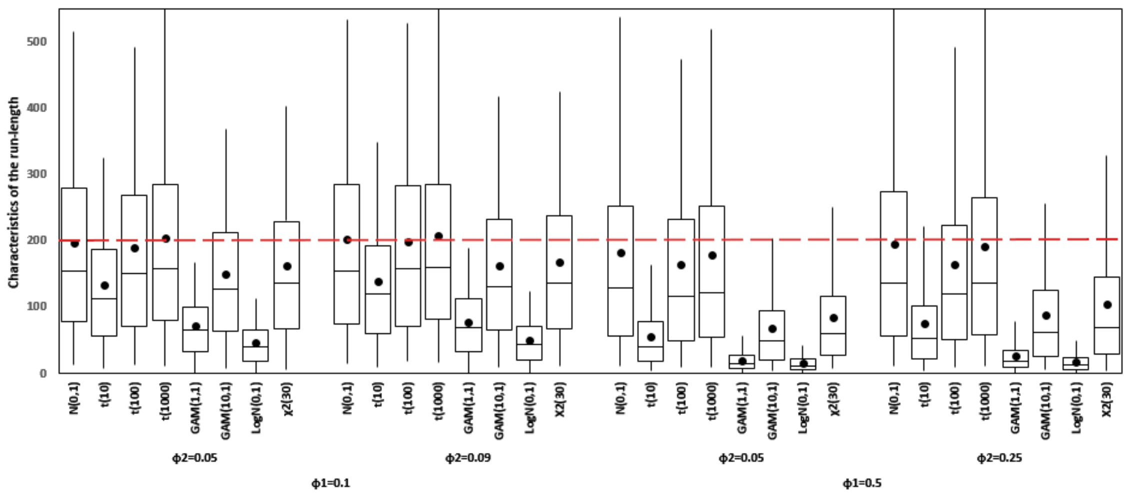
Figure 1: Box-plots of the run-length distributions for different distributions of the univariate case



(a) $p = 2$



(b) $p = 4$



(c) $p = 10$

Figure 2: Box-plots of the run-length distributions for different distributions of the multivariate case

4.1.2 OOC performances of the proposed EHWMA and MEHWMA schemes

The performance evaluation of the proposed EHWMA scheme is given in Table 1 where the ARL values are computed for different shifts in the process mean of the quality characteristic being monitored for different values of ϕ_1 and ϕ_2 when $n = 1$ for a nominal $ARL_0 = 200$. Here the quality characteristic is assumed to have a normal distribution. The shift values represent the number of standard deviation units by which the mean was shifted from 0. This means, if the shift value is 0.5, the mean shifted from $\mu_0 = 0$ to $\mu_1 = 0 + 0.5(1) = 0.5$. From Tables 1 and 2, it can be seen that when the shift $\delta = 0$, the ARL is close to 200 for each pair of ϕ_1 and ϕ_2 and then it starts dropping rapidly as the shift size increases. In addition to the ARL, the expected ARL ($EARL$) is also computed to measure the overall performance of the proposed schemes where $EARL(\delta_1, \delta_2)$ is the average ARL over the shift sizes $\delta \in (\delta_1, \delta_2]$. For example, $EARL(1, 2)$ for $\phi_1 = 0.1$ and $\phi_2 = 0.01$ is calculated as

$$EARL(1, 2) = \frac{(5.3 + 4.0 + 3.3 + 2.7)}{4} = 3.8.$$

The $EARL(0, 1)$, $EARL(1, 2)$, $EARL(2, 3)$, $EARL(0, 2)$, $EARL(1, 3)$ and $EARL(0, 3)$ are used to evaluate the overall performances of the proposed schemes for small, moderate, large, small-to-moderate, moderate-to-large and small-to-large shifts, respectively. In Table 1, it can be seen that the EHWMA scheme performs better for smaller values of ϕ_1 and ϕ_2 regardless of the size of the mean shift. The larger the value of ϕ_1 and/or ϕ_2, the larger the $EARL$ profile. The scheme seems to trigger OOC signals the fastest on average for larger shifts in the process mean.

Table 2 presents the performance of the MEHWMA scheme when $n = 1$, $\phi_1 \in \{0.1, 0.25\}$, $\phi_2 \in \{0.01, 0.05, 0.09, 0.1, 0.2\}$ for different values of p. Tables containing the results of additional pairs of ϕ_1 and ϕ_2 are given in the Github link provided in appendix B. Similar trends to the ones of the univariate case are observed. The $EARL$ increases as the values of ϕ_1 and ϕ_2 increase. Though smaller pairs of ϕ_1 and ϕ_2 seem to trigger an OOC signal faster, depending on the process being monitored, other combinations could result in better results (this is shown in Section 6.3). Moreover, it can also be observed that as p increases, the performance of the MEHWMA scheme decreases. In other words, the smaller the p value, the better the performance of the proposed MEHWMA scheme. Additional tables for the standard deviation run length ($SDRL$), median run length (MRL), expected $SDRL$ ($ESDRL$) and expected MRL ($EMRL$) are also given in the Github link provided in appendix B for both the EHWMA and MEHWMA schemes where the findings are summarised as follows (These tables are also available from the corresponding author on request):

- In terms of the $SDRL$ and $ESDRL(\delta_1, \delta_2)$ profiles, similar findings to the ones of the ARL and $EARL(\delta_1, \delta_2)$ profiles are observed, respectively.
- In terms of MRL and $EMRL(\delta_1, \delta_2)$ profiles, when ϕ_1 is kept fixed for small shifts in the process mean, the proposed EHWMA and MEHWMA monitoring schemes are more likely to give signals the quickest for small values of ϕ_2. However, when ϕ_2 is kept fixed, the proposed schemes are likely

to give OOC signals for small values of ϕ_1 for both small and moderate shifts in the process mean. Note though that for large shifts in the process mean, when p and ϕ_1 are kept fixed, the proposed schemes are equally likely to give OOC signal as ϕ_2 increases.

- The smaller the value of p, the worse the *SDRL* and *MRL* profiles.
- The IC *SDRL* profile is larger for large values of p. In other words, the proposed schemes is likely to give more false alarms for large values of p.

Table 1: *ARL* and *EARL* results for the EHWMA scheme when $n = 1$ for a nominal ARL_0 value of 200.

	ϕ_1	**0.1**			**0.25**			**0.5**			**0.9**		
	ϕ_2	**0.01**	**0.05**	**0.09**	**0.05**	**0.1**	**0.2**	**0.05**	**0.1**	**0.25**	**0.05**	**0.1**	**0.25**
	L_{EH}	2.516	2.540	2.600	2.772	2.763	2.762	2.804	2.803	2.794	2.804	2.809	2.801
	0.00	200.9	200.9	200.5	200.3	200.8	200.5	200.7	200.4	200.1	200.9	200.1	200.3
	0.25	55.7	58.4	61.3	70.1	72.4	81.4	103.5	105.5	114.0	146.4	147.7	151.9
	0.50	21.4	22.1	23.5	25.3	26.3	29.4	38.8	40.6	47.1	79.8	81.7	84.0
	0.75	11.4	11.8	12.4	12.8	13.4	15.1	18.3	18.7	21.8	41.7	42.6	44.5
	1.00	7.3	7.6	8.0	8.1	8.3	9.0	10.1	10.3	11.7	23.4	23.0	25.0
	1.25	5.3	5.4	5.7	5.6	5.7	6.2	6.4	6.5	7.2	13.6	13.6	14.3
Shifts	**1.50**	4.0	4.1	4.3	4.2	4.3	4.5	4.5	4.6	5.0	8.4	8.5	8.9
	1.75	3.3	3.3	3.4	3.3	3.4	3.5	3.4	3.5	3.6	5.6	5.7	6.0
	2.00	2.7	2.8	2.9	2.8	2.8	2.9	2.7	2.7	2.8	4.0	4.1	4.2
	2.25	2.3	2.4	2.4	2.4	2.4	2.4	2.2	2.3	2.3	3.0	3.0	3.1
	2.50	2.0	2.0	2.1	2.0	2.0	2.0	1.9	1.9	1.9	2.3	2.3	2.4
	2.75	1.7	1.8	1.8	1.8	1.8	1.8	1.7	1.7	1.7	1.9	2.0	1.9
	3.00	1.5	1.6	1.6	1.6	1.6	1.6	1.5	1.5	1.5	1.6	1.6	1.7
	***EARL*(0,1)**	24.0	25.0	26.3	29.1	30.1	33.7	42.7	43.8	48.7	72.8	73.8	76.3
	***EARL*(1,2)**	3.8	3.9	4.1	4.0	4.0	4.3	4.3	4.3	4.7	7.9	8.0	8.4
	***EARL*(2,3)**	1.9	1.9	2.0	1.9	1.9	2.0	1.8	1.8	1.9	2.2	2.2	2.3
	***EARL*(0,2)**	13.9	14.4	15.2	16.5	17.1	19.0	23.5	24.1	26.7	40.4	40.9	42.4
	***EARL*(1,3)**	2.9	2.9	3.0	3.0	3.0	3.1	3.1	3.1	3.3	5.1	5.1	5.3
	***EARL*(0,3)**	9.9	10.3	10.8	11.7	12.0	13.3	16.3	16.7	18.4	27.7	28.0	29.0

Table 2: *ARL* and *EARL* results for the MEHWMA scheme when $n = 1$ for a nominal ARL_0 value of 200.

	ϕ_1=0.1									ϕ_1=0.25								
	$p=2$			$p=3$			$p=4$			$p=2$			$p=3$			$p=4$		
ϕ_2	**0.01**	**0.05**	**0.09**	**0.01**	**0.05**	**0.09**	**0.01**	**0.05**	**0.09**	**0.05**	**0.1**	**0.2**	**0.05**	**0.1**	**0.2**	**0.05**	**0.1**	**0.2**
h_{MEH}	9	9.15	9.44	11.09	11.32	11.73	13.1	13.29	13.62	10.34	10.37	10.38	12.7	12.7	12.62	14.6	14.62	14.68
0	201.7	205.4	204.1	201.0	203.8	198.5	205.7	206.7	194.3	201.2	202.7	197.9	203.0	204.9	203.9	199.3	206.1	205.1
0.25	62.8	68.3	70.6	68.8	72.4	78.9	74.9	80.9	81.2	83.5	90.2	97.8	95.5	97.9	104.5	100.7	99.0	119.6
0.5	25.1	26.5	28.0	27.2	29.0	32.2	30.1	31.4	32.7	30.8	33.3	37.6	35.5	36.5	42.7	38.5	38.6	47.2
0.75	13.5	14.1	14.8	14.9	15.7	16.7	16.0	16.9	17.9	15.2	16.1	18.4	17.7	18.6	21.1	18.7	19.9	23.7
1	8.7	8.9	9.3	9.4	9.8	10.4	10.2	10.6	11.2	9.5	10.0	11.0	10.9	10.9	12.5	11.7	12.0	13.6
1.25	6.2	6.5	6.5	6.8	7.0	7.4	7.3	7.5	7.8	6.5	6.7	7.3	7.3	7.7	8.3	7.7	7.9	8.9
Shifts 1.5	4.7	4.8	5.1	5.1	5.2	5.5	5.6	5.6	6.1	4.9	5.0	5.4	5.4	5.5	6.0	5.7	5.8	6.4
1.75	3.8	3.8	4.0	4.0	4.2	4.4	4.4	4.5	4.6	3.8	3.9	4.2	4.3	4.3	4.5	4.5	4.6	4.9
2	3.2	3.2	3.3	3.4	3.5	3.6	3.7	3.7	3.8	3.2	3.2	3.3	3.5	3.5	3.7	3.7	3.8	3.9
2.25	2.7	2.7	2.8	2.9	3.0	3.0	3.1	3.2	3.3	2.7	2.7	2.8	3.0	2.9	3.0	3.1	3.1	3.2
2.5	2.3	2.4	2.4	2.5	2.6	2.7	2.7	2.8	2.8	2.3	2.3	2.3	2.5	2.6	2.6	2.7	2.7	2.8
2.75	2.1	2.0	2.1	2.2	2.3	2.3	2.4	2.4	2.5	2.0	2.1	2.1	2.2	2.2	2.2	2.3	2.4	2.4
3	1.8	1.8	1.9	1.9	2.0	2.1	2.1	2.1	2.2	1.8	1.8	1.8	2.0	2.0	1.9	2.1	2.1	2.1
***EARL*(0,1)**	27.5	29.4	30.7	30.1	31.7	34.5	32.8	35.0	35.8	34.8	37.4	41.2	39.9	41.0	45.2	42.4	42.4	51.0
***EARL*(1,2)**	4.5	4.6	4.7	4.8	5.0	5.2	5.2	5.3	5.6	4.6	4.7	5.0	5.1	5.3	5.6	5.4	5.5	6.0
***EARL*(2,3)**	2.2	2.2	2.3	2.4	2.5	2.5	2.6	2.6	2.7	2.2	2.2	2.2	2.4	2.4	2.4	2.5	2.6	2.6
***EARL*(0,2)**	16.0	17.0	17.7	17.4	18.3	19.9	19.0	20.1	20.7	19.7	21.1	23.1	22.5	23.1	25.4	23.9	23.9	28.5
***EARL*(1,3)**	3.4	3.4	3.5	3.6	3.7	3.9	3.9	4.0	4.1	3.4	3.5	3.6	3.8	3.8	4.0	4.0	4.0	4.3
***EARL*(0,3)**	11.4	12.1	12.6	12.4	13.1	14.1	13.6	14.3	14.7	13.9	14.8	16.2	15.8	16.2	17.8	16.8	16.8	19.9

4.2 Performance comparison of the proposed schemes with the existing schemes in terms of the zero-state ARL values

For comparison purposes, the performances of the existing HWMA, EWMA and EEWMA schemes are investigated in a similar way to that of the new EHWMA scheme described in Section 4. The HWMA and EWMA schemes are constructed using $\phi \in \{0.1, 0.25, 0.5\}$, while the EEWMA scheme is constructed using $\phi_1 \in \{0.1, 0.25, 0.5\}$ and $\phi_2 \in \{0.01, 0.05\}$. Figure 3 provides a visual comparison of the competing schemes in terms of the OOC ARL values. Figures 3 (a) to (c) compare the ARL values of the EHWMA to that of the EWMA and HWMA, while Figures 3 (d) to (f) compare the performance of the EHWMA scheme to that of the EEWMA scheme. The values of L_Z, L_{EZ} and L_H, as explained in Section 2, are chosen to achieve an ARL_0 value of 200 for each scheme. The ARL values are computed for the shifts of sizes 0.25 to 1 since for shifts larger than 1 the competing schemes perform similarly. From Figure 3, it can be seen that the new EHWMA scheme performs better than the EWMA and EEWMA schemes in almost all cases. The new scheme has smaller ARL values for different shifts meaning that it will detect a small shift faster than the EWMA and EEWMA schemes. Comparing the new scheme to the HWMA scheme, both seem to perform similarly in terms of the ARL values.

Figure 4 presents a visual comparison of the performances of the competing schemes for multivariate cases. The MEHWMA scheme is compared to the multivariate versions of the HWMA, EWMA and EEWMA schemes in terms of the ARL values using the same values for ϕ_1 and ϕ_2. The competing schemes are compared when $p \in \{2, 3, 4\}$. In Figure 4, the first four graphs are for $p = 2$ characteristics, the second four are for $p = 3$ characteristics, and the last four are for $p = 4$ characteristics. For each value of p, Figures (i) and (ii) compare the ARL profile of the MEHWMA scheme to those of the multivariate EWMA and HWMA schemes for different values of ϕ_1 and ϕ_2, while Figures (iii) and (iv) compare the ARL profile of the MEHWMA scheme to those of the multivariate EEWMA scheme for different values of ϕ_1 and ϕ_2 when $\delta \in \{0.25, 0.5, 0.75\}$. The MEHWMA scheme can be seen to have smaller ARL values compared to the multivariate EWMA and EEWMA schemes in most cases. This means that the new MEHWMA scheme will detect small shifts faster than the existing multivariate EWMA and EEWMA schemes. The MEHWMA scheme also has similar ARL values to that of the multivariate HWMA scheme and seemingly it will detect small shifts at a similar rate.

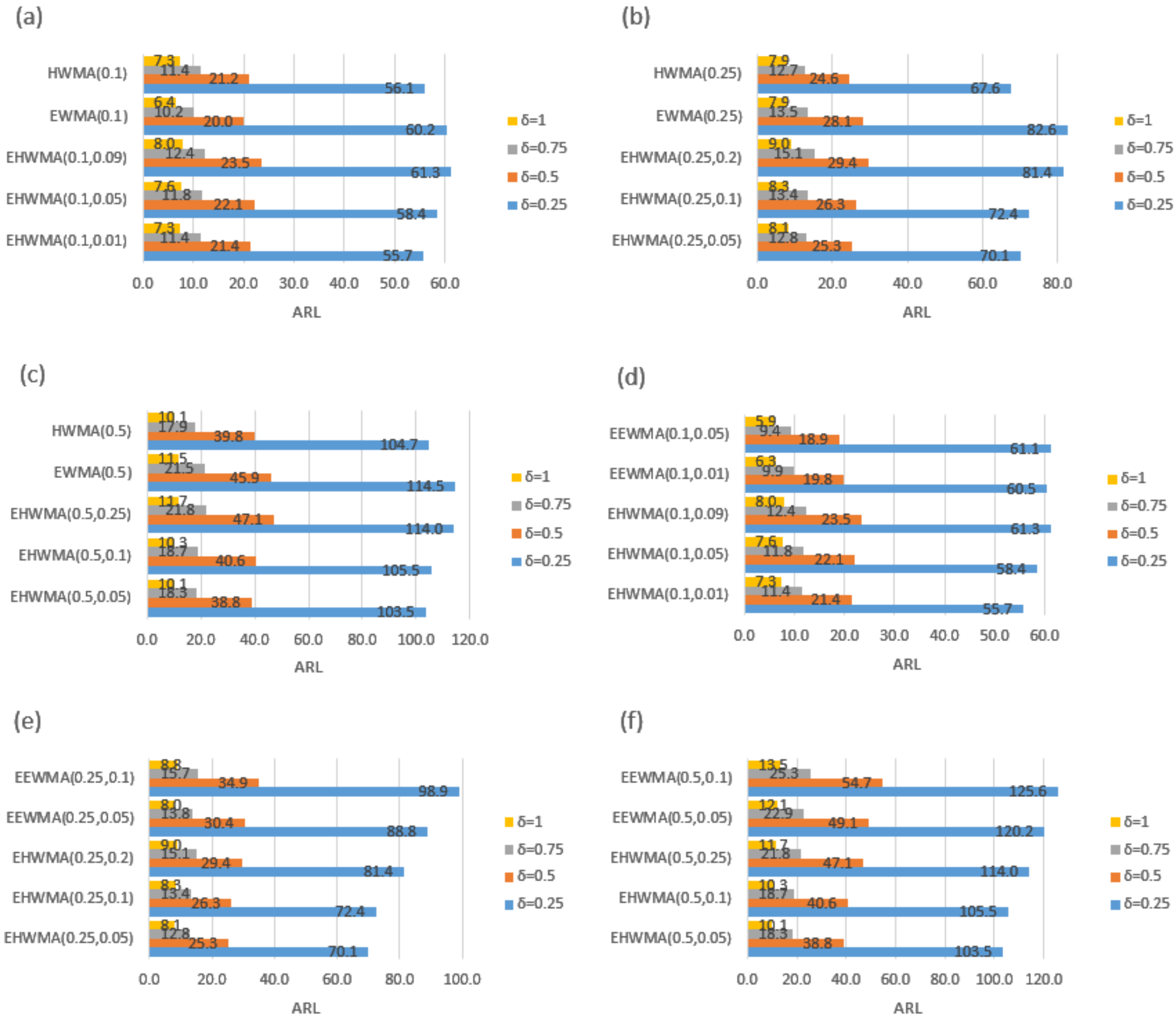
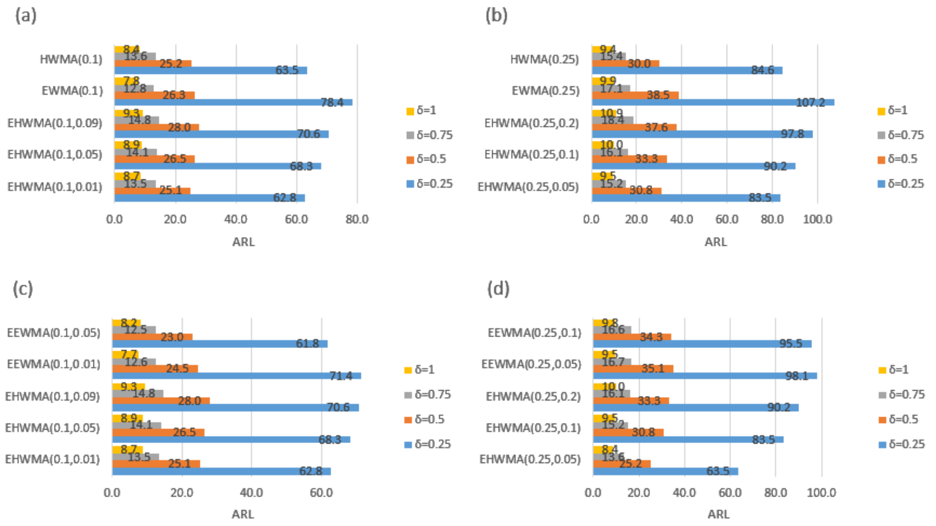
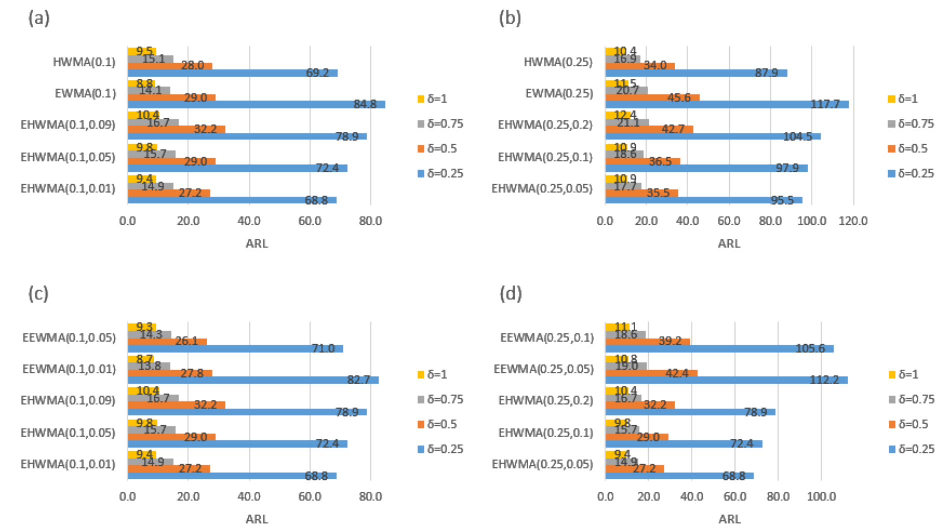


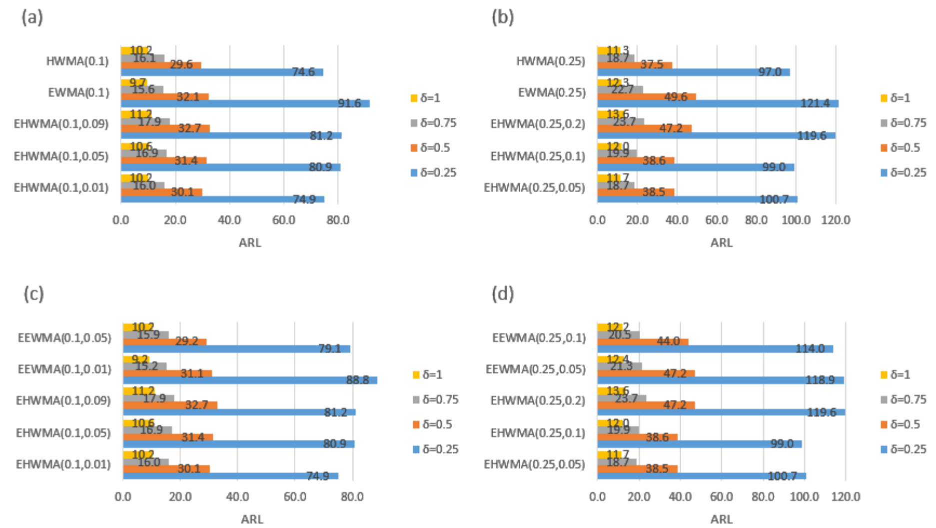
Figure 3: *ARL* performance comparison of the EHWMA, EWMA, HWMA, and EEWMA schemes for different shift sizes δ.



Pane 1: $p = 2$



Pane 2: $p = 3$



Pane 3: $p = 4$

Figure 4: *ARL* performance comparison of the multivariate EHWMA, HWMA, EWMA and EEWMA schemes for different shift sizes δ.

5 Performance comparison in terms of the conditional expected delay

5.1 The conditional expected delay criterion

A number of researchers have argued that when the IC process distribution is unknown or unspecified, then the OOC ARL metric is not a good criterion for the performance comparison of different schemes; see for instance, Kenett and Pollak (2012) and Knoth (2022). In such case, they recommend the use of the conditional expected delay (CED) criterion. Kenett and Pollak (2012) defined the CED as "the delay from the first opportunity to detect a change and not from the time of change itself". Thus, the CED, denoted as D_τ, is mathematically defined by:

$$D_\tau = E_\tau\left(R - \tau + 1 | R \geq \tau\right), \, \tau = 1, 2, ..., \tag{24}$$

where R represents the number of samples until an alarm is raised and τ is the change point applied to the following model:

$$\mu = \begin{cases} \mu_0 = 0, & \text{for } t < \tau \\ \mu_1 = \delta, & \text{for } t \geq \tau \end{cases}, \tag{25}$$

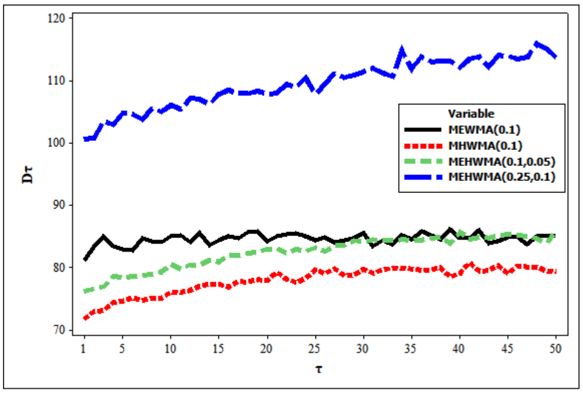
where μ_0 and μ_1 represent the IC and OOC process means, respectively.

When $\tau = 1$, then the CED is equivalent to the zero-state ARL, and when $\tau > 1$ the CED is equivalent to the steady-state ARL.

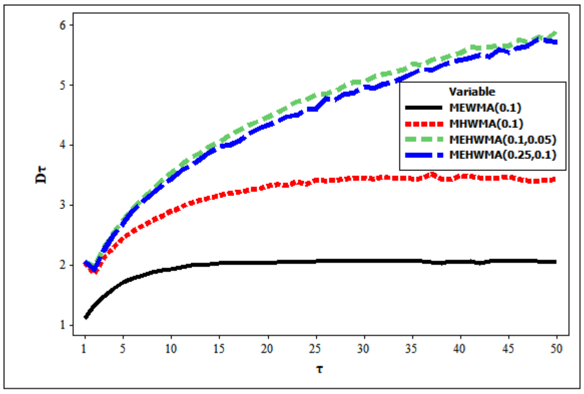
5.2 Performance comparison of the competing schemes in terms of the CED

In this section, the MEHWMA scheme is compared to the existing MEWMA and MHWMA schemes in terms of the CED criterion (i.e. in terms of the steady-state ARL values) when $n = 1$, $p = 3$ and $\delta \in \{0.25, 3\}$ for a nominal ARL_0 value of 200. The MEWMA and MHWMA schemes are designed using smoothing parameters $\phi \in \{0.1, 0.25\}$. For instance, MEWMA (0.1) indicates that the MEWMA scheme is designed with $\phi = 0.1$. The MEHWMA scheme is designed using (ϕ_1, ϕ_2)= (0.1, 0.05), (0.25,0.05) and (0.25,0.1). Figure 5 displays the performance comparison of the MEWMA, MHWMA and MEHWMA schemes in terms of the D_τ when $\tau = 1, 2, ..., 50$. From Figure 5, it can be seen that for small smoothing parameters, the MHWMA scheme outperforms both the MEWMA and MEHWMA schemes for small shifts in the process mean and the MEHWMA (0.1,0.05) scheme performs better than the MEWMA (0.1) scheme when $\tau \leq 30$; however, for $\tau > 30$, the MEHWMA and MEWMA schemes are almost similar in performance (see Figure 5(a)). For large shifts in the process mean, the MEWMA scheme performs better followed by the MHWMA scheme (see Figure 5(b)). For moderate smoothing parameters, the

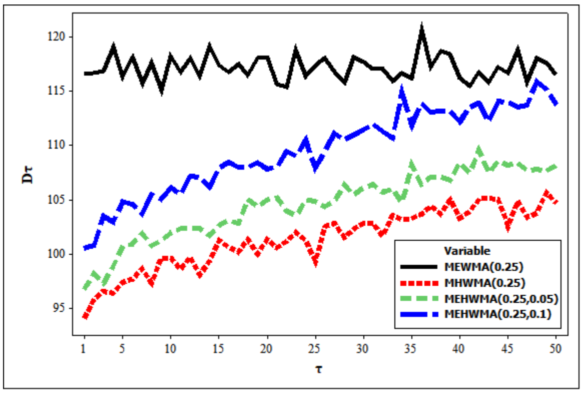
MHWMA scheme performs better for small shifts followed by the MEHWMA scheme for small shifts in the process mean (see Figure 5(c)). However, for large shifts, the MEWMA scheme outperforms both the MHWMA and MEHWMA schemes (see Figure 5(d)). This comparison reveals that in steady-state, the MEHWMA scheme is inferior to the competing MHWMA and MEWMA schemes in many cases.



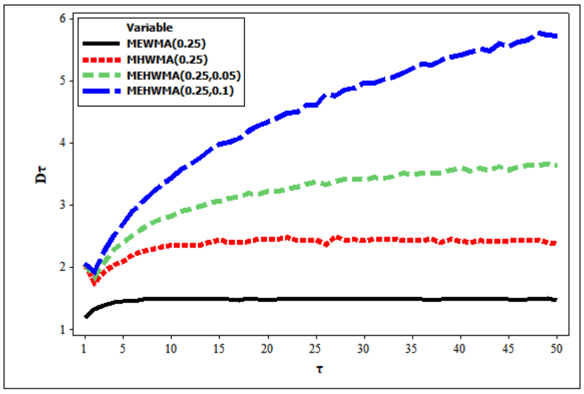
(a) $\delta = 0.25$



(b) $\delta = 3$



(c) $\delta = 0.25$



(d) $\delta = 3$

Figure 5: Performance comparison of the MEWMA, MHWMA and MEHWMA monitoring scheme in terms of the D_τ values when and $n = 1$ and $\delta = 0.25$ and 3 for a nominal $ARL_0 = 200$

6 Enhancement of the MEHWMA scheme using a support vector machine

6.1 Support vector machine in SPM

A SVM is a supervised learning technique that separates classes through the construction of a hyperplane. Before obtaining this hyperplane, the data have to be mapped into a higher dimensional space through the use of a kernel function, such as the Gaussian radial basis or a polynomial function. This is done to allow the hyperplane to be constructed linearly in this space if it cannot be done in the original space; see Noble (2006). Knowing up to how many more dimensions the data need to be mapped into is, unfortunately, a process of trial and error, and mapping the data to a too-high dimension can cause overfitting. SVMs are designed to mainly deal with binary classification problems, but they can also be applied to multi-classification problems by constructing multiple binary SVM in either the one-against-all or one-against-one method. The one-against-all method constructs a SVM for each class that separates the data as being either of that class or not. The one-against-one method, on the other hand, constructs SVMs for each class against every other class where the final classification is then decided by which class has the largest number of positive classifications. SVMs have shown to be useful in both classification and pattern recognition problems and scales relatively well with dimensionality. SVMs do, however, require parameter estimation and thus need to be combined with estimation algorithms (Cuentas et al. (2017)).

The most popular way SVMs are implemented into SPM is to solve the problems of pattern recognition; see Apsemidis and Psarakis (2020). In SPM, there are different types of patterns that can be present in the data. Some of these patterns include: an upward or downward trend, an upward or downward shift, or a cyclical pattern. These patterns can occur while the charting statistic is still plotting in the IC region and might not cause the monitoring scheme to consider the process as OOC. The effective use of SVMs to determine patterns in SPM has recently been demonstrated by Shao and Hu (2020) and Chowdhury and Janan (2020).

6.1.1 SVM classifier

Let $\{(\boldsymbol{x}_i, y_i), i = 1, 2, ..., N\}$ be a set where $\boldsymbol{x}_i \in \mathbb{R}^p$ are the training vectors and $y_i \in \{-1, 1\}$ are the indicator values for a two-class case. Thus, the optimal hyperplane can be found by minimising the following quadratic optimisation problem:

$$\min_{\boldsymbol{w},b,\boldsymbol{\xi}} \left\{ \frac{1}{2}\boldsymbol{w}^{\intercal}\boldsymbol{w} + C\sum_{i=1}^{N} \xi_i \right\}, \tag{26}$$

subject to

$$y_i \left(\boldsymbol{w}^{\intercal}\phi(\boldsymbol{x}_i) + b\right) \geq 1 - \xi_i, \text{ where } \xi_i \geq 0, \tag{27}$$

where b is the bias term, C is the regularization or cost parameter, $\boldsymbol{w}$ is the vector of the hyperplane, and $\phi(\boldsymbol{x}_i)$ maps $\boldsymbol{x}_i$ into a higher dimension. Note that the slack variables, denoted ξ_i $(i = 1, 2, ..., N)$, are needed to relax the separability constraint defined in (27).
The decision function is given by:

$$\operatorname{sgn}\left(\sum_{i=1}^{N} y_i\alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b\right), \tag{28}$$

where $\alpha_i \in [0, C]$, and $K(\boldsymbol{x}_i, \boldsymbol{x})$ is a kernel function. If the decision function is not linear then a specific kernel function is needed in order to apply linear classifiers. This is discussed in the next section.

6.1.2 Kernel functions

Different kernel functions, denoted $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$, are documented in the literature. The most commonly used ones include the polynomial, radial basis, and sigmoid (i.e. neural network activation) kernel functions which are mathematically defined by:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\mathbf{x}_i^{\intercal}\mathbf{x}_j)^p, \tag{29}$$

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\gamma||\boldsymbol{x}_i - \boldsymbol{x}_j||^2\right), \ \gamma > 0, \tag{30}$$

and

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh\left(k\mathbf{x}_i^{\intercal}\mathbf{x}_j + \delta\right), \ k > 0 \text{ and } \delta > 0, \tag{31}$$

where p is a tunable parameter that is specified a priori by users.

6.2 The MEHWMA scheme incorporated with SVM

In this paper, a SVM will be implemented along with a MEHWMA scheme using the radial basis kernel defined in (30). Note that the radial basis kernel was selected because its simplicity, high accuracy and extended use in SPM. Once the MEHWMA scheme gives an OOC signal, a SVM will then determine the variable (or quality characteristic) that caused the OOC signal. A SVM with a radial basis kernel will be fitted using the `svm` function using the e1071 R package. The γ parameter of the kernel function as well as the cost parameter C both need to be determined through additional steps. A dataset consisting of samples that caused OOC signals and which variable was responsible for each signal will be used to train the SVM. A Cross-validation procedure will then be used to train the model and determine the parameters γ and C that optimize the models prediction accuracy.

6.3 MEHWMA monitoring scheme integrated with a SVM: An illustrative example using real-life data

This illustrative example is based on bivariate data from Chen et al. (2005) which represents the production of springs and contains samples of two variables namely, the spring inner diameter (X_1) and the spring elasticity (X_2). Twelve samples of size 5 for X_1 and X_2 are given in Table 3. From Chen et al. (2005), the IC mean vector and covariance matrix for the process are given by

$$\boldsymbol{\mu}_0 = \begin{pmatrix} 28.29 \\ 45.85 \end{pmatrix} \text{ and } \boldsymbol{\Sigma}_0 = \begin{pmatrix} 0.0035 & -0.0046 \\ -0.0046 & 0.0226 \end{pmatrix}, \tag{32}$$

respectively. To detect which sample is OOC, the bivariate EHWMA (BEHWMA) scheme with smoothing parameters $\phi_1 = 0.25$ and $\phi_2 = 0.05$ will be used. Then, from Table 2, in order to have an ARL_0 value close to 200, the value of h_{MEH} needs to be set to 10.34. Once the OOC samples have been determined by the scheme, a SVM is used to identify which characteristic caused the signal for each OOC sample. The SVM will use the sample averages of X_1 and X_2 as input.

A radial SVM was trained using data that were simulated from a bivariate normal distribution with the same IC mean and variance given in (32) where one variable was forced to be OOC by shifting its mean. Since the data is bivariate, the radial function of the SVM first maps the data into 3 dimensions and then fits a linear hyperplane separating the observations in this space. The contour of linear hyperplane can then be plotted in 2 dimensions which will then provide the prediction regions for the model. The sigma parameter of the radial function was estimated by determining the value that maximises the prediction accuracy of the model. The final model was fitted with a γ parameter of 1.69 and a cost parameter of 1. This model achieves a 94.7% prediction accuracy when using 5-fold cross-validation, and the prediction regions of the fitted SVM are given in Figure 6.

Table 3: Spring manufacturing process data

Sample	**Spring manufacturing process data**									
	X1					**X2**				
1	28.1	28.3	28.3	28.2	28.3	46.3	45.8	45.9	45.9	45.8
2	28.5	28.4	28.3	28.3	28.2	45.9	45.9	45.8	45.9	45.9
3	28.3	28.3	28.3	28.4	28.3	45.8	45.8	45.8	45.5	45.6
4	28.2	28.3	28.3	28.3	28.3	45.8	46.0	45.8	46.0	45.9
5	28.3	28.4	28.3	28.3	28.3	45.8	45.9	46.0	45.8	45.7
6	28.3	28.3	28.3	28.3	28.2	45.8	45.9	45.8	45.9	46.0
7	28.2	28.3	28.3	28.4	28.4	45.9	45.8	45.7	45.8	45.7
8	28.2	28.4	28.3	28.3	28.3	45.8	45.9	45.7	45.8	45.7
9	28.3	28.4	28.3	28.4	28.3	45.6	46.1	45.9	45.6	45.9
10	28.3	28.3	28.3	28.4	28.3	45.7	45.8	45.8	45.9	45.9
11	28.4	28.4	28.4	28.5	28.4	45.8	45.4	45.8	45.8	45.9
12	28.2	28.2	28.3	28.1	28.4	45.3	45.3	45.7	45.8	45.9

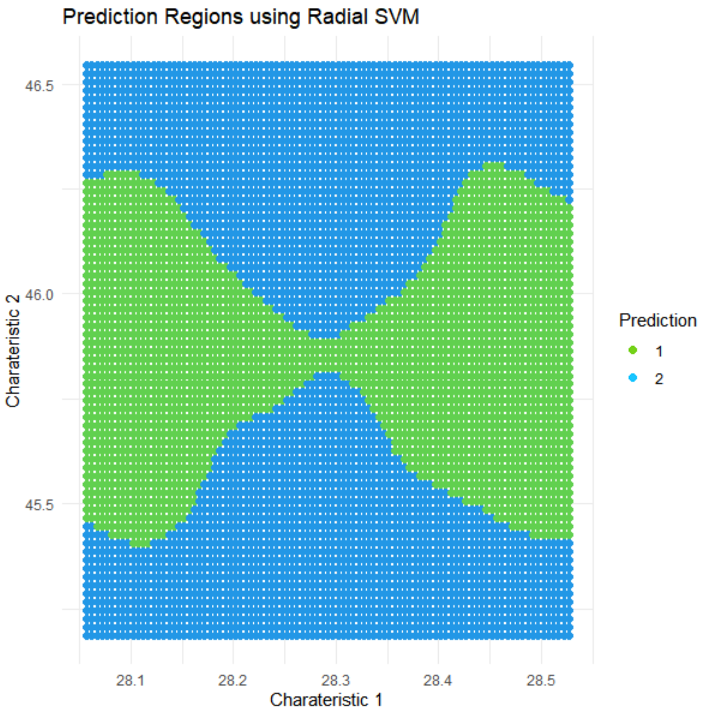


Figure 6: Prediction regions of radial SVM

Applying the BEHWMA scheme to the data, the plot of T^2 statistics is shown in Figure 7, where the red dashed line represents the control limit h_{MEH}. It can be seen that the BEHWMA scheme gives an OOC signal on the 11th and 12th samples as their T^2 values plot above h_{MEH}. Using the proposed SVM algorithm, it is determined that there was a significant change in X_1 on the 11th sample and another significant change in X_2 on the 12th sample. Operators would then be recommended to make correctional actions to restore the process. Since these results are in line with what was expected (see Chen et al. (2005)), the BEHWMA and proposed SVM are shown to perform accurately in this case. The final results of this example are summarized in Table 4.

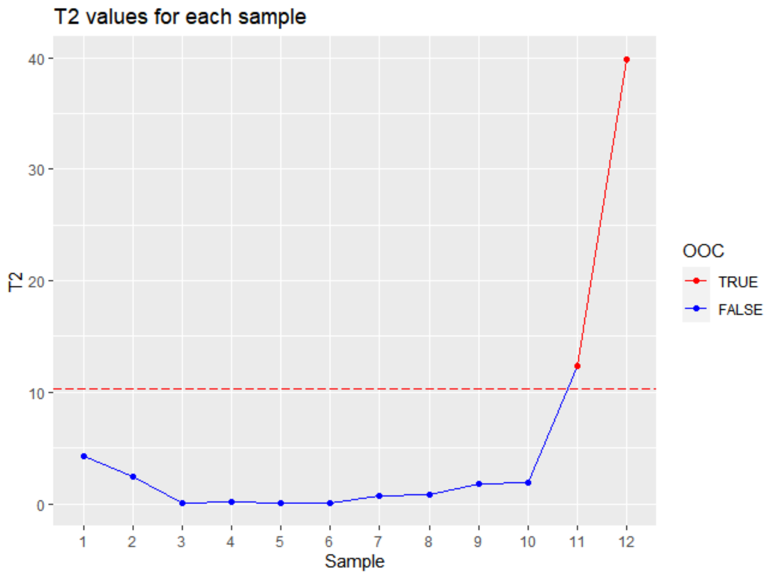


Figure 7: Bivariate EHWMA chart for spring manufacturing samples

Table 4: Final results for spring manufacturing process data

Sample	**Spring manufacturing process data**										**MEHWMA**		**SVM**
	X1					**X2**					**T2**	**OOC?**	**Cause**
1	28.1	28.3	28.3	28.2	28.3	46.3	45.8	45.9	45.9	45.8	4.22	NO	-
2	28.5	28.4	28.3	28.3	28.2	45.9	45.9	45.8	45.9	45.9	2.48	NO	-
3	28.3	28.3	28.3	28.4	28.3	45.8	45.8	45.8	45.5	45.6	0.01	NO	-
4	28.2	28.3	28.3	28.3	28.3	45.8	46.0	45.8	46.0	45.9	0.20	NO	-
5	28.3	28.4	28.3	28.3	28.3	45.8	45.9	46.0	45.8	45.7	0.06	NO	-
6	28.3	28.3	28.3	28.3	28.2	45.8	45.9	45.8	45.9	46.0	0.08	NO	-
7	28.2	28.3	28.3	28.4	28.4	45.9	45.8	45.7	45.8	45.7	0.75	NO	-
8	28.2	28.4	28.3	28.3	28.3	45.8	45.9	45.7	45.8	45.7	0.82	NO	-
9	28.3	28.4	28.3	28.4	28.3	45.6	46.1	45.9	45.6	45.9	1.76	NO	-
10	28.3	28.3	28.3	28.4	28.3	45.7	45.8	45.8	45.9	45.9	1.94	NO	-
11	28.4	28.4	28.4	28.5	28.4	45.8	45.4	45.8	45.8	45.9	12.38	YES	X1
12	28.2	28.2	28.3	28.1	28.4	45.3	45.3	45.7	45.8	45.9	39.89	YES	X2

7 Concluding remarks

In this paper, new univariate and multivariate EHWMA schemes for monitoring the process mean and mean vector have been introduced and their performances and IC robustness have also been evaluated using intensive simulations. It has been found that the new schemes are not IC robust but perform better for small smoothing parameters ϕ_1 and ϕ_2. Therefore, practitioners in the industries are recommended to use small smoothing parameters. The new schemes have also been compared to existing monitoring schemes and the results showed that they outperformed them in many cases in zero-state. However, in steady-state, the proposed MEHWMA scheme is outperformed by the MHWMA and MEWMA schemes. A BEHWMA scheme has been integrated with a SVM to enable it to detect the variable that caused the OOC signal using real-life data from a spring manufacturing process. The integrated SVM was shown to perform accurately in correctly classifying which quality characteristic went OOC with 94.7% prediction accuracy. Note that the proposed schemes are designed under the assumption of normality. In the case where the underlying process distribution departs from normality, the properties and results provided in this paper should be revisited. If there is doubt concerning the nature and shape of the underlying process distribution, practitioners in the industries are advised to use non-parametric versions of the proposed schemes (this will be investigated in future). Future research opportunities could include:

- Integrating other ML techniques such as the ANN, MARS, etc. to the MEHWMA scheme.
- Designing new SVMs with higher sets of hyperplanes to facilitate the integration of SVM to multivariate schemes when p is considerably larger.
- Integrating ML tools to non-parametric multivariate monitoring schemes to provide more robust and efficient schemes when the assumption of normality is violated.

References

Abbas, N. (2018). Homogeneously weighted moving average control chart with an application in substrate manufacturing process. *Computers & Industrial Engineering*, 120:460–470.

Adegoke, N., Abbasi, S., Smith, A., Anderson, M., and Pawley, M. (2019). A multivariate homogeneously weighted moving average control chart. *IEEE Access*, 7:9586–9597.

Anwar, S., Aslam, M., Zaman, B., and Riaz, M. (2022). An enhanced double homogeneously weighted moving average control chart to monitor process location with application in automobile field. *Quality and Reliability Engineering International*, 38(1):174–194.

Anwer, F., Sanaullah, A., Ahmad, A., and Asghar, A. (2022). An improved mixed-homogeneously weighted moving average-cusum control chart for efficient monitoring of a process mean. *Journal of Statistical Computation and Simulation*, pages 1–23.

Apsemidis, A. and Psarakis, S. (2020). Support vector machines: A review and applications in statistical process monitoring. *Data Analysis and Applications 3: Computational, Classification, Financial, Statistical and Stochastic Methods*, 5:123–144.

Bersimis, S., Psarakis, S., and Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability engineering international*, 23(5):517–543.

Chan, K., Mukherjee, A., Chong, Z., and Lee, H. (2021). Distribution-free double exponentially and homogeneously weighted moving average lepage schemes with an application in monitoring exit rate. *Computers & Industrial Engineering*, 161:107370.

Chen, G., Cheng, S., and Xie, H. (2005). A new multivariate control chart for monitoring both location and dispersion. *Communications in Statistics Simulation and Computation*, 34(1):203–217.

Chowdhury, N. and Janan, F. (2020). Control chart pattern recognition: A comparison between statistical correlation measure and support vector machine. In *Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management Detroit*, pages 10–14.

Cuentas, S., Peñabaena-Niebles, R., and Garcia, E. (2017). Support vector machine in statistical process monitoring: a methodological and analytical review. *The International Journal of Advanced Manufacturing Technology*, 91(1):485–500.

Demircioglu Diren, D., Boran, S., and Cil, I. (2020). Integration of machine learning techniques and control charts in multivariate processes. *Scientia Iranica*, 27(6):3233–3241.

Kenett, R. and Pollak, M. (2012). On assessing the performance of sequential procedures for detecting a change. *Quality and Reliability Engineering International*, 28(5):500–507.

Knoth, S. (2022). An expanded case against synthetic-type control charts. *Quality and Reliability Engineering International*.

Malela-Majika, J., Shongwe, S., Chatterjee, K., and Koukouvinos, C. (2022a). Monitoring univariate and multivariate profiles using the triple exponentially weighted moving average scheme with fixed and random explanatory variables. *Computers & Industrial Engineering*, 163:107846.

Malela-Majika, J.-C., Chatterjee, K., and Koukouvinos, C. (2022b). Univariate and multivariate linear profiles using max-type extended exponentially weighted moving average schemes. *IEEE Access*, 10:6126–6146.

Maravelakis, P., Panaretos, J., and Psarakis, S. (2002). Effect of estimation of the process parameters on the control limits of the univariate control charts for process dispersion. *Communications in Statistics-Simulation and Computation*, 31(3):443–461.

Montgomery, D. (2020). *Introduction to statistical quality control. 8th ed.* John Wiley & Sons.

Naveed, M., Azam, M., Khan, N., and Aslam, M. (2018). Design of a control chart using extended ewma statistic. *Technologies*, 6(4):108.

Noble, W. (2006). What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567.

Page, E. (1961). Cumulative sum charts. *Technometrics*, 3(1):1–9.

Roberts, S. (2000). Control chart tests based on geometric moving averages. *Technometrics*, 42(1):97–101.

Shao, Y. and Hu, Y. (2020). Using machine learning classifiers to recognize the mixture control chart patterns for a multiple-input multiple-output process. *Mathematics*, 8(1):102.

Yeganeh, A., Abbasi, S., and Shongwe, S. (2021). A novel simulation-based adaptive mewma approach for monitoring linear and logistic profiles. *IEEE Access*, 9:124268–124280.

Appendices

Appendix A:

Derivations of the properties of the EHWMA statistic

To derive the properties of the univariate EHWMA statistic, (17) can be rewritten for the cases $t = 1$ and $t > 1$.

For $t = 1$, we have:

$$
\begin{aligned}
\mathrm{EH}_1 &= \phi_1 X_1 - \phi_2 X_0 + (1-\phi_1+\phi_2)\bar{X}_0 \\
&= \phi_1 X_1 - \phi_2 \mu_0 + (1-\phi_1+\phi_2)\mu_0 \\
&= \phi_1 X_1 + (1-\phi_1)\mu_0, \qquad \text{(A1)}
\end{aligned}
$$

And for $t > 1$,

$$
\begin{aligned}
\mathrm{EH}_t &= \phi_1 X_t - \phi_2 X_{t-1} + (1-\phi_1+\phi_2)\bar{X}_{t-1} \\
&= \phi_1 X_t - \phi_2 X_{t-1} + (1-\phi_1+\phi_2)\left(\frac{1}{t-1}\sum_{i=1}^{t-1} X_i\right) \\
&= \phi_1 X_t - \phi_2 X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i + X_{t-1}\right) \\
&= \phi_1 X_t + \left(\frac{1-\phi_1+\phi_2-(t-1)\phi_2}{t-1}\right) X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i\right) \\
&= \phi_1 X_t + \left(\frac{1-\phi_1+(1-(t-1))\phi_2}{t-1}\right) X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i\right) \\
&= \phi_1 X_t + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right) X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i\right). \qquad \text{(A2)}
\end{aligned}
$$

Now, to derive the expected value of EH_t we start with the case where $t = 1$. Using (A1) and the properties of expected value, we get:

$$
\begin{aligned}
E(\mathrm{EH}_1) &= E\left[\phi_1 X_1 - \phi_2 \mu_0 + (1-\phi_1+\phi_2)\mu_0\right] \\
&= \phi_1 E\left[X_1\right] - \phi_2 \mu_0 + (1-\phi_1+\phi_2)\mu_0 \\
&= \phi_1 \mu_0 - \phi_2 \mu_0 + (1-\phi_1+\phi_2)\mu_0 \\
&= \mu_0.
\end{aligned}
$$

Next, for the case where $t > 1$, we use (A2) and get:

$$\begin{aligned}
E\left(\mathrm{EH}_t\right) &= E\left(\phi_1 X_t + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right) X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i\right)\right) \\
&= \phi_1 E\left(X_t\right) + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right) E\left(X_{t-1}\right) + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} E\left(X_i\right)\right) \\
&= \phi_1\mu_0 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)\mu_0 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)(t-2)\mu_0 \\
&= \left(\frac{(t-1)\phi_1 + 1 - \phi_1 - (t-2)\phi_2 + (1-\phi_1+\phi_2)(t-2)}{t-1}\right)\mu_0 \\
&= \left(\frac{t\phi_1 - \phi_1 + 1 - \phi_1 + (1-\phi_1+\phi_2-\phi_2)(t-2)}{t-1}\right)\mu_0 \\
&= \left(\frac{t\phi_1 - 2\phi_1 + 1 + (1-\phi_1)(t-2)}{t-1}\right)\mu_0 \\
&= \left(\frac{t\phi_1 - 2\phi_1 + 1 + t - t\phi_1 - 2 + 2\phi_1}{t-1}\right)\mu_0 \\
&= \left(\frac{t-1}{t-1}\right)\mu_0 \\
&= \mu_0,
\end{aligned}$$

and thus,

$$E\left(\mathrm{EH}_t\right) = \mu_0.$$

To derive the variance of EH_t, we will again start with the case where $t = 1$ and use (A1), which results in:

$$\begin{aligned}
V\left(\mathrm{EH}_1\right) &= V\left(\phi_1 X_1 - \phi_2\mu_0 + (1-\phi_1+\phi_2)\mu_0\right) \\
&= \phi_1^2 V\left(X_1\right) + 0 + 0 \\
&= \phi_1^2\sigma_0^2. \qquad \text{(A3)}
\end{aligned}$$

Then, for the case where $t > 1$, using (A2) we get:

$$\begin{aligned}
V\left(\mathrm{EH}_t\right) &= V\left(\phi_1 X_t + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right) X_{t-1} + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)\left(\sum_{i=1}^{t-2} X_i\right)\right) \\
&= \phi_1^2 V\left(X_t\right) + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2 Var\left[X_{t-1}\right] + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2\left(\sum_{i=1}^{t-2} V\left(X_i\right)\right) \\
&= \phi_1^2\sigma_0^2 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2\sigma_0^2 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2(t-2)\sigma_0^2 \\
&= \left(\phi_1^2 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2(t-2)\right)\sigma_0^2,
\end{aligned}$$

and thus, the variance of EH_t is given by:

$$V(\mathrm{EH}_t) = \begin{cases} \phi_1^2\sigma_0^2, & \text{for } t = 1 \\ \left(\phi_1^2 + \left(\frac{1-\phi_1-(t-2)\phi_2}{t-1}\right)^2 + \left(\frac{1-\phi_1+\phi_2}{t-1}\right)^2 (t-2)\right)\sigma_0^2, & \text{for } t > 1. \end{cases} \tag{A4}$$

Lastly, the control limits of the univariate EHWMA statistic are found by using the formula

$$\mathrm{UCL}_{\mathrm{EH}_t}/\mathrm{LCL}_{\mathrm{EH}_t} = E\left(\mathrm{EH}_t\right) \pm L_{\mathrm{EH}}\sqrt{V(\mathrm{EH}_t)}, \tag{A5}$$

where L_{EH} is the control limit constant.

Appendix B:

R code used and additional tables

`https://github.com/Luke-Pieters/EHWMA_Scheme_2021`