Efficient Homogeneously Weighted Moving Average Chart for Monitoring Process Mean Using an Auxiliary Variable

NURUDEEN A. ADEGOKE1, ADAM N. H. SMITH1, MARTI J. ANDERSON2, RIDWAN ADEYEMI SANUSI3, AND MATTHEW D. M. PAWLEY1

1School of Natural and Computational Sciences, Massey University, Auckland 0632, New Zealand
2New Zealand Institute for Advanced Study, Massey University, Auckland 0632, New Zealand
3Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong

Corresponding author: Nurudeen A. Adegoke (nurudeen.adegoke@yahoo.com)

ABSTRACT In this paper, we propose an efficient control chart for monitoring small shifts in a process mean for scenarios where the process variable is observed with a correlated auxiliary variable. The proposed chart, called an auxiliary homogeneously weighted moving average (AHWMA) chart, is a homogeneously weighted moving average type control chart that uses both the process and auxiliary variables in the form of a regression estimator to provide an efficient and unbiased estimate of the mean of the process variable. We provide the design structure of the chart and examine its performance in terms of its run length properties. Using a simulation study, we compare its run length performance with several existing methods for detecting a small shift in the process mean. Our simulation results show that the proposed chart is more efficient in detecting a small shift in the process mean than its competitors. We provide a detailed study of the chart’s robustness to non-normal distributions and show that the chart may also be designed to be less sensitive to non-normality. We give some recommendations on the application of the chart when the process parameters are unknown and provide an example to show the implementation of the proposed new technique.

INDEX TERMS Auxiliary variable, average run length, control chart, parameter estimation, robustness.

I. INTRODUCTION Monitoring programs are designed to detect unnatural changes in process variables for a wide variety of applications, particularly in industrial and manufacturing settings. Control charts are popular tools for tracking processes of interest, ensuring they are kept in control by monitoring essential quality characteristics [1]. To date, several univariate control charts have been proposed in statistical process control (SPC) literature; they are classified into (i) memory-less control charts and (ii) memory-type control charts for monitoring large and small-to-moderate shifts in the process, respectively. For example, the Shewhart chart is a memory-less control chart that uses only the current process information and not the past behavior of the process. It is very effective for detecting large shifts in the process mean (i.e., $\delta \geq 2$, where $\delta$ is the size of the shift in standard deviation units [2]). The homogeneously weighted moving average (HWMA) control chart by [3] is a memory-type chart proposed for efficient monitoring of small (i.e., $\delta \leq 0.5$) to moderate (i.e., $0.5 < \delta < 2$) shifts in the process mean. Other memory-type charts include the EWMA chart by [4], the CUSUM chart by [5], and the mixed EWMA-CUSUM chart proposed by [6].

These univariate classical charts are widely used in most of today’s industries; their attractiveness is motivated by the simplicity of their construction, implementation, and interpretation, as well as their prompt detection of small, moderate, or large shifts in a process mean. These techniques have been implemented by [7] to monitor the quality of garments produced on the sewing floor, by [8] to monitor and control steam boiler generation for vacuum degassing processes, and by [9] to evaluate critical control point hygiene data. Also, see [10]–[13] for some other industrial applications of these classical charts.

Several applications of classical charts focus on monitoring the process in situations where the process variable is independent of other variables; however, in some cases,
the process variable may be observed along with another correlated auxiliary variable. The concept of using supplemental information to provide an efficient estimate of a population parameter is popular in the field of survey sampling [14]. Several researchers have studied and recommended the use of auxiliary variables in the monitoring of a process variable of interest, and have proposed a variety of different control chart tools for this purpose.

For example, [15] proposed a regression control chart, while [16] proposed a cause-selecting control chart. Recently, [17] proposed a Shewhart-type chart in the form of a regression-based estimator, called a \( V \) chart, for monitoring process variability. He compared the proposed \( V \) chart with some other existing charts (specifically, \( R \), \( S \) and \( S^2 \) charts for the same purpose), and showed that the \( V \) chart was effective in detecting moderate to large shifts in the process variability under certain conditions on the correlation between the process variable and auxiliary variable. Similarly, a Shewhart-type control chart using a regression-based estimator (\( M_r \) chart) for monitoring a process mean (proposed by [18]) was shown to be more powerful at detecting shifts in the process mean. This work was later extended to an EWMA chart for detecting small-to-moderate changes in the process mean under different correlation structures between the process and auxiliary variables (see [19]–[24]).

Here, we propose a more efficient control chart for monitoring the process mean when the process variable is observed along with an auxiliary variable. The proposed chart, called an auxiliary homogeneously weighted moving average (AHWMA) chart, is an HWMA-type control chart that uses both the deviation of the process mean from its target value (known apriori or estimated from historical reference samples), as well as a regression estimator for the process mean provided through its relationship (or estimated relationship) with an auxiliary variable with which it is known to be correlated.

The rest of the article is organized as follows: in Section II, we outline the structure of the chart. Section III compares the AHWMA chart (run length) performance in detecting a small shift in the process mean with several other existing charts. Section IV gives a detailed study of the chart’s robustness to non-normality. We give recommendations regarding the application of the chart when the process parameters are unknown in Section V. Section VI also provides an example to demonstrate practical implementation of the chart, followed by a conclusion and discussion in Section VII.

II. THE AHWMA CONTROL CHART

Consider a control chart based on observations \( z_{ij} \) of the quality characteristics \( Z_{ij} \), for each of \( i = 1, \ldots, m \) time-points and \( j = 1, \ldots, n \) sampling units per time-point (i.e., \( n \) is the sample size). Assume that these quality characteristics \( (Z_{ij}) \) are identically distributed as normal random variables with a known in-control mean \( (\mu_Z) \) and standard deviation \( (\sigma_Z) \), i.e., \( Z_{ij} \sim N(\mu_Z, \sigma_Z^2) \) and represents the main process variable. The HWMA statistic, \( H_i \) (in Equation (1)), at time-point \( i \), gives a specific weight to the current sample and the remaining weight is equally distributed among the previous samples, and is given by:

\[
H_i = w\bar{z}'_i + (1-w)\bar{z}'_{i-1}
\]

where \( \bar{z}'_i \) is the sample average for the \( ith \) sample, and \( w \) is a smoothing constant (also called the sensitivity parameter) selected such that \( 0 < w \leq 1 \). The HWMA structure becomes the Shewhart plotting structure whenever \( w = 1 \). \( \bar{z}'_{i-1} \) is the average of the sample means of all of the previous samples (i.e., up to and including the \( (i-1)th \) sample), and is given by \( \bar{z}'_{i-1} = \frac{1}{n} \sum_{k=1}^{i-1} \bar{z}_k \). The mean and variance of the HWMA statistic in Equation (1) are given by \( \mu_H = \mu_Z \), and

\[
\sigma_H^2 = \begin{cases} 
\frac{1}{n}w^2\sigma_Z^2 & \text{if } i = 1 \\
\frac{1}{n}((1-w)^2\sigma_Z^2) & \text{if } i > 1 
\end{cases}
\]

where \( \mu_H = \mu_Z \) and \( \sigma_Z^2 \) are the mean and variance of the normally distributed random variable \( Z \) [3].

Let an auxiliary variable, \( Y_{ij} \), be correlated with the main variable of interest, \( Z_{ij} \), with correlation \( \rho \). We assume the observations of \( Z_{ij} \) and \( Y_{ij} \) are observed in pairs from a bivariate normal distribution, given by \((Z_{ij}, Y_{ij}) \sim N_2(\mu_Z, \mu_Y, \sigma_Z^2, \sigma_Y^2, \rho)\), where \( N_2 \) is the bivariate normal distribution, and \( \mu_Y \) and \( \sigma_Y^2 \) are the population mean and variance of \( Y \), respectively. We assume the linear relationship between the variables can be modeled using linear least squares obtained by adjusting the process mean at time \( i \), \( z_i \), to reflect its known relationship with the auxiliary variable. This yields the regression-informed estimator (i.e., \( R_i \)) for the process mean given by:

\[
R_i = \bar{z}_i + b(\mu_Y - \bar{y}_i)
\]

where \( b \) (given as \( b = \frac{\sigma_Z}{\sigma_Y} \)) is the slope of the regression line; given as the change in the process variable, \( Z \), due to a unit change in the auxiliary variable, \( Y \) [14]. The mean and variance of \( R \) are given as:

\[
\mu_R = \mu_Z \text{ and } \sigma_R^2 = \frac{\sigma_Z^2}{n}(1 - \rho^2).
\]

respectively.

Using Equation (3), the plotting statistic \( (T_i) \) of the AHWMA chart is given as:

\[
T_i = wR_i + (1-w)\bar{R}_{i-1}
\]

where \( w \) is the smoothing parameter of the chart (selected such that \( 0 \leq w \leq 1 \)), \( R_i \) is the regression-informed estimate of the process variable, given in Equation (3) for the \( ith \) sample, and \( \bar{R}_{i-1} \) is the average of the sample means of all of the previous samples (i.e., up to and including the \( (i-1)th \) sample) of the plotting statistic, and is given as \( \bar{R}_{i-1} = \frac{1}{n} \sum_{k=1}^{i-1} R_k \). The mean and variance of the plotting
The ARL statistic in Equation (5) are given as $\mu_H = \mu_Z$ (also called the center line of the AHWMA chart), and

$$\sigma^2_T = \begin{cases} \frac{(1 - \rho^2)}{n} w^2 \sigma^2_Z & \text{if } i = 1 \\ \frac{(1 - \rho^2)}{n} \left( \frac{w^2 \sigma^2_Z + (1 - w)^2 \sigma^2_Z}{i - 1} \right) & \text{if } i > 1 \end{cases} \quad (6)$$

respectively. The time varying lower ($L_i$) and upper ($U_i$) control chart limits of the plotting statistic given in Equation (5) are given as:

$$L_i = \begin{cases} \mu_Z - C \sigma_Z \sqrt{\frac{w^2}{n} (1 - \rho^2)} & \text{if } i = 1 \\ \mu_Z - C \sigma_Z \sqrt{\frac{w^2}{n} + \frac{(1 - w)^2}{n(i - 1)}} (1 - \rho^2) & \text{if } i > 1 \end{cases} \quad (7)$$

and

$$U_i = \begin{cases} \mu_Z + C \sigma_Z \sqrt{\frac{w^2}{n} (1 - \rho^2)} & \text{if } i = 1 \\ \mu_Z + C \sigma_Z \sqrt{\frac{w^2}{n} + \frac{(1 - w)^2}{n(i - 1)}} (1 - \rho^2) & \text{if } i > 1 \end{cases} \quad (8)$$

respectively, where, $C$ determines the width of the control limits; the values of $C$ and $w$ are chosen to achieve a desired in-control average run length (ARL) for the chart. ARL is the average number of plotted samples on the control chart before a shift is detected. We provide R-code [25] (in the supplementary material) which practitioners can use to obtain the value of $C$, given $w$, that fix the in-control ARL of the chart to a desired value. We adopted the ARL numerics algorithm for the EWMA chart [26]; implemented in the spc ($R$) package [27], to obtain an arbitrary start value ($C_{\text{start}}$) of the AHWMA chart limit, and used a binary search algorithm to determine $C$ for the chart.

### III. PERFORMANCE ASSESSMENTS AND COMPARISONS

#### A. PERFORMANCE ASSESSMENTS

Here, we provide a comprehensive assessment of the AHWMA chart in detecting a shift in the process mean in terms of the chart’s ARL and standard deviation of run length (SDRL). The value of the ARL when a process is in control is denoted by $ARL_0$, while $ARL_1$ denotes the value of the ARL when the process is out of control. SDRL is used to determine the variation of the run length distribution for a given value of shift. Similarly, $SDRL_0$ and $SDRL_1$ can be defined as the SDRL for the in-control and out-of-control process, respectively. When comparing two charts, the $ARL_0$ is fixed to a specific value, and a chart having a smaller value of $ARL_1$ than another is said to be more efficient in detecting the shift in the process [28]–[31].

To ensure a fair comparison of the AHWMA chart with existing charts of the same $ARL_0$, we examined the performance of the chart with $w \in \{0.03, 0.05, 0.10, 0.25, 0.5, 0.75\}$, and the corresponding values of $C$ that fix $ARL_0$ to 500 are used, the R-code provided in the supplementary material finds the value of $C$ (for each value of $w$), that fixes $ARL_0$ to 500. We examined the ARL performance of the chart under different correlation values between the process and the auxiliary variables. Specifically, we considered $\rho \in \{0.05, 0.25, 0.5, 0.75, 0.95\}$. The $ARL$ values of the AHWMA chart are given in Tables 1 - 5. In these tables, $\delta$ is the size of shifts, and is calculated as $\delta = \frac{n^{1/2} |\mu_Z - \mu_1|}{\sigma_Z}$, where $n$ is the sample size at each time $i$ (here, we assume $n = 1$ across $i$), and $\mu_Z$ and $\mu_1$ are the in-control and out-of-control mean, respectively.

The main findings of the AHWMA chart (cf. Tables 1 - 5) are:

- For fixed values of $\delta$ and $\rho$, the chart is more efficient for smaller values of $w$. For example, where $\rho = 0.05$ (Table 1), when $\delta = 0.5$, the values of the $ARL_1$ when $w = 0.03$ and $0.75$ were 20.05 and 132.08, respectively. Thus, the chart detects a shift in the process mean faster when a small value of $w$ is used.
- For fixed values of $\delta$, $w$ and $C$, the chart is more efficient when large values of $\rho$ are used. For example, when $w = 0.03$, $L = 2.272$, and $\delta = 0.5$, $ARL_1$ values were 20.05 and 3.43 in (Tables 1 and 5) for $\rho = 0.05$ and $\rho = 0.95$, respectively. Thus, increases in the correlation structure between the process variable and the auxiliary variable lead to an increase in the chart’s ability to detect a shift.
- The chart is $ARL$ unbiased. That is, the $ARL_1$ values never exceed the corresponding $ARL_0$ for any choice of $\delta$ examined.
- As $\delta$ increases, the $ARL_1$ and $SDRL_1$ values approach 1 and 0, respectively, especially for large values of $\rho$; that is, the charts detect large shifts promptly.

#### B. COMPARISONS

We provide detailed comparisons of the proposed AHWMA chart with some existing control charts: the classical HWMA chart by [3], the classical EWMA chart by [4], the classical CUSUM chart by [4], the auxiliary-based EWMA chart (i.e., $M_x$EWMA) by [19], and the auxiliary-based CUSUM chart (i.e., $A_u$CUSUM$_2$ by [22]), in terms of their ARL values. The auxiliary-based EWMA and CUSUM charts are also based on a regression estimator; they provide efficient applications of the classical EWMA and CUSUM charts, respectively, in those situations where the process variable is observed along with a variable. For comparison with the $M_x$EWMA and $A_u$CUSUM$_2$ charts, we considered three different values of $\rho$: namely, $\rho \in \{0.05, 0.5, 0.95\}$. In all cases, the charts’ parameters were set to values that fix $ARL_0$ at 500. We provide the charts’ $ARL$ results that optimized $\delta$ at $w \in \{0.05, 0.1, 0.2\}$.

The results of the comparisons are provided in Table 6. As shown on the table, the AHWMA chart outperformed the classical CUSUM, EWMA and HWMA charts in detecting shifts in the mean, especially when $\rho > 0.05$. For fixed values of $w$ and $\rho$, the AHWMA chart was more efficient than the $A_u$CUSUM$_2$ chart, especially for small-to-moderate values...
TABLE 1. ARL and SDRL values of the AHWMA chart when the correlation between the variables is $\rho = 0.05$. The values of $C$ are chosen to fix the chart’s ARL to 500 for each chosen value of $w$.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\omega$</th>
<th>$w$</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.000</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.050</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.050</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.075</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.075</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.100</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.100</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.125</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.125</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.150</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.150</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.175</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.175</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.200</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.200</td>
<td>0.05</td>
<td>0.1</td>
</tr>
</tbody>
</table>

TABLE 2. ARL and SDRL values of the AHWMA chart when the correlation between the variables is $\rho = 0.25$. The values of $C$ are chosen to fix the chart’s ARL to 500 for each chosen value of $w$.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\omega$</th>
<th>$w$</th>
<th>0.03</th>
<th>0.05</th>
<th>0.1</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.000</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.050</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.050</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.075</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.075</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.100</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.100</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.125</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.125</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.150</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.150</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.175</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.175</td>
<td>0.05</td>
<td>0.1</td>
</tr>
<tr>
<td>0.200</td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
<td>0.5</td>
<td>0.75</td>
<td>0.200</td>
<td>0.05</td>
<td>0.1</td>
</tr>
</tbody>
</table>

of $\delta$ (i.e., $\delta < 2$). For fixed values of $w$ and $\rho$, the AHWMA chart was less efficient than the $M_2$EWMA chart in detecting moderate-to-large shifts (i.e., $\delta > 0.5$) in the process mean. However, the AHWMA chart shows greater efficiency than the $M_2$EWMA chart in detecting small shifts (i.e., $\delta \leq 0.5$) in the mean.

IV. ROBUSTNESS TO NON-NORMALITY OF THE CHART

The AHWMA chart described in Section II relies on the assumption that the process variable and the auxiliary variable are bivariate normally distributed. In practice, this assumption does not always hold. Non-normality is not a major concern with a large sample size because the central limit theorem warrants that the sample mean will be approximately normally distributed for any continuous variables [32]. When $n = 1$, however, it is important to check the sensitivities of control charts to departures from normality [2]. We refer readers to [33]–[35], and [36] for detailed studies on the robustness of the EWMA control chart to non-normality.

Here, we investigate the robustness of the AHWMA chart to non-normality. As mentioned by [33] “a control chart is robust if its in-control run-length distribution remains stable (unchanged or nearly unchanged) when the underlying distributional assumption(s) (e.g. normality) are violated”. Following previous investigators [33]–[37], we considered a heavy-tailed bivariate distribution (the bivariate Student’s $t$-distribution), and a skewed distribution (the bivariate gamma distribution). We denote the bivariate $t$-distribution with $v$ degrees of freedom by $t_2(v)$. The probability density
The function of a bivariate $t$-distribution is given by

$$f(x) = \frac{\Gamma(v + 2)/2}{\Gamma(v/2)v\pi|\Sigma|^{1/2}}\left[1 + \frac{1}{v}(x - \mu)^T\Sigma^{-1}(x - \mu)\right]^{-(v+2)/2}$$

(9)

where $x \in R^2$, $\mu = [\mu_1, \mu_2]^T$ is the $2 \times 1$ vector of location parameters, $\Sigma$ is a $2 \times 2$ positive-definite (or covariance) matrix, $v$ is the number of the degrees of freedom, and $\Gamma(n) = (n - 1)!$ for $n = 1, 2, \ldots$. The mean vector and covariance matrix are given as $\mu$ (if $v > 1$, else, undefined), and $\frac{v}{\sqrt{v}}\Sigma$ (when $v > 2$, else, undefined), respectively.

We denote the bivariate gamma distribution with shape parameter, $\alpha_1p$, and scale parameter, $\beta_1p$, by $G_2(\alpha_1p, \beta_1p, \Sigma)$, where $1_p$ is a column vector of ones of size $p = 2$. The probability density function of the bivariate gamma distribution is given as:

$$f(x) = \frac{|\Sigma|^{-\alpha}}{\beta^{\alpha_1p}\Gamma_2(\alpha)}|x|^{\alpha_1p-\alpha/2}\exp\left(\frac{1}{\beta}\Gamma(\alpha)\Gamma_2(\alpha/2)\right)$$

(10)

where $\alpha > 0$ is the scale parameter, $\beta > 0$ is the shape parameter, and $\Gamma_2(\alpha) = \pi^{1/2}/\Gamma(\alpha)(\alpha - 1/2)$. See [38] and [39] for detailed information on the bivariate gamma distribution and its properties.

We studied the chart’s robustness under a large range of degrees of freedom ($v$) for the bivariate $t$-distribution; namely, $v \in \{4, 6, 8, 10, 15, 20, 30, 40, 50, 100, 1000\}$. For the bivariate gamma distribution, without loss of generality, we considered the scale parameter, $\beta = 1$, and a range of values of the shape parameter,
i.e., $\alpha \in \{1, 2, 3, 4, 5, 10, 50, 100, 1000\}$. Hence, we denote the bivariate gamma distribution as $G_2(\alpha)$, for short. The $ARL_0$ values of the chart for $\rho \in \{0.25, 0.5, 0.95\}$, and $w \in \{0.03, 0.05, 0.25, 0.75\}$ for the bivariate $t$ and bivariate gamma distributions are given in Tables 7 and 8, respectively. The $ARL_0$ results in Tables 7 and 8 are summarized below:

- For a fixed value of $w$, the $ARL_0$ values of the bivariate $t$-distributions are the same for all the correlation values (i.e., $\rho_{ZY} = 0.25, 0.5, \text{or} 0.95$) examined. This result is due to the symmetry of the $t$-distribution.
- However, for a fixed value of $w$, the $ARL_0$ of the bivariate gamma distributions differ across all the values of $\rho_{ZY}$ examined. Here, the chart appears to be more robust to non-normality only for smaller values of $\rho_{ZY}$.
- For both non-normal distributions, as expected, the $ARL_0$ value increases, and tends to converge to values obtained for bivariate normal data (i.e., $\rho = 0.5$), especially when $w$ is large (i.e., $w = 0.75$), and other parameters are unchanged (i.e., $v = 50, \beta = 50$, and $\delta = 0.5$), the $ARL_1$ for the AHWMA under normal, $t$ and gamma distributions were 125.44, 118.63, and 68.10, respectively; the percentage deviation of these $ARL_1$ values from obtained under the normal distribution were $-5.43\%$ and $-45.71\%$ for the $t$ and gamma distributions, respectively.

V. STEP-BY-STEP ALGORITHM FOR CONSTRUCTING THE AHWMA CHART WHEN PARAMETERS ARE UNKNOWN

The AHWMA chart in Section II was formulated based on the assumption that the parameters associated with the process variable and auxiliary variable are all known. However, these parameters are generally unknown in practice and need to be estimated. In this case, the regression model in Equation (3) would be based on estimated parameters, and is given as:

$$\tilde{R}_i = \tilde{Z}_i + \hat{b}(\mu_Y - \hat{\mu}_Y)$$

where $\hat{b}$ is the estimated slope of the regression line; given as the estimated change in the process variable $Z$ due to a unit change in the auxiliary variable $Y$ [14], and $\hat{\mu}_Y$ ($\hat{\mu}_Y = \frac{1}{m} \sum_{i=1}^{m} \tilde{Z}_i$) is the unbiased estimate of the mean of the auxiliary variable (i.e., $\mu_Y$). The estimated mean and variance of $R$ are given as $\tilde{R} = \tilde{\mu}_Z$, and $S_R^2 = \frac{\hat{\sigma}_Z^2}{n} (1 - r^2)$, where $r$ is the estimated value of the correlation size between the variables, $\hat{\mu}_Z$ and $\hat{\sigma}_Z^2$ are the unbiased estimates of $\mu_Z$ and $\sigma_Z^2$, respectively. The $\hat{\mu}_Z$ and $\hat{\sigma}_Z^2$ are calculated from a specified set of sample values measured when the process was known to be in control. They are given as $\hat{\mu}_Z = \frac{1}{m} \sum_{i=1}^{m} \tilde{Z}_i$ and $\hat{\sigma}_Z^2 = \frac{S_p^2}{c_{4,m}}$, where $S_p$ is $\left( \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n} (Z_{ij} - \bar{Z}_i)^2 \right)^{1/2}$.
TABLE 6. ARL comparisons of the charts.

<table>
<thead>
<tr>
<th>ω</th>
<th>CUSUM</th>
<th>HWMA</th>
<th>Classical chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.035</td>
<td>0.4673</td>
<td>929.7</td>
<td>492.71</td>
</tr>
<tr>
<td>0.075</td>
<td>0.3924</td>
<td>684.8</td>
<td>387.97</td>
</tr>
<tr>
<td>0.15</td>
<td>0.3242</td>
<td>538.1</td>
<td>359.37</td>
</tr>
<tr>
<td>0.25</td>
<td>0.2723</td>
<td>448.32</td>
<td>279.48</td>
</tr>
<tr>
<td>0.50</td>
<td>0.2176</td>
<td>322.34</td>
<td>193.97</td>
</tr>
<tr>
<td>0.75</td>
<td>0.1651</td>
<td>234.23</td>
<td>135.69</td>
</tr>
</tbody>
</table>

and $c_{4,m} = \frac{1}{2} \left( \frac{m(n-1)}{m(n-1)+1} \right)^{1/2} \left( \frac{n-1}{m} \right)^{1/2}$ is an unbiassing constant [3], [40].

Using Equation (11), the plotting statistic for the AHWMA control chart based on estimated parameters is given as:

$$\hat{T}_i = w\hat{R}_i + (1 - w)\hat{R}_{i-1}$$
The estimated mean and variance of the plotting statistic in Equation (12) are given by \( \hat{\mu}_T = \hat{\mu}_Z \), and

\[
\hat{\sigma}^2_{\hat{Z}} = \begin{cases} 
\frac{(1 - r^2)}{n} \hat{\sigma}^2_Z & \text{if } i = 1 \\
\frac{(1 - r^2)}{i - 1} \left( \frac{w^2 \hat{\sigma}^2_Z}{n} + (1 - w) \hat{\sigma}^2_{Z} \right) & \text{if } i > 1
\end{cases}
\]

(13)

The upper and lower control limits for the (plotting statistic given in Equation (12)) estimated time varying control chart are given as:

\[
\hat{L}_i = \begin{cases} 
\hat{\mu}_Z - C' \hat{\sigma}_Z \sqrt{\frac{w^2}{n} \left( 1 - r^2 \right)} & \text{if } i = 1 \\
\hat{\mu}_Z - C' \hat{\sigma}_Z \sqrt{\frac{w^2}{n} \left( 1 - r^2 \right)} \left( 1 - \frac{w}{n(i - 1)} \right) & \text{if } i > 1
\end{cases}
\]

(14)

\[
\hat{U}_i = \begin{cases} 
\hat{\mu}_Z + C' \hat{\sigma}_Z \sqrt{\frac{w^2}{n} \left( 1 - r^2 \right)} & \text{if } i = 1 \\
\hat{\mu}_Z + C' \hat{\sigma}_Z \sqrt{\frac{w^2}{n} \left( 1 - r^2 \right)} \left( 1 - \frac{w}{n(i - 1)} \right) & \text{if } i > 1
\end{cases}
\]

(15)
where $C'$ determines the width of the estimated control limits. Also, the estimated center line (CL) of the AHWMA chart is given by:

$$\hat{CL} = \hat{\mu}_Z$$

When the chart is based on estimated parameters, implementation occurs in two phases. In phase I (retrospective phase), a historical reference sample is studied to establish the in-control state and to evaluate the stability of the process [41], [42]. Once the in-control reference sample is characterized, the process parameters are estimated from phase I and control chart limits are obtained for use in phase II. Phase II involves regular monitoring of the process. If successive observed values obtained at the beginning of Phase II fall within the in-control limits calculated from Phase I, the process is considered to be in control. In contrast, any observed values during Phase II that fall outside the control limits indicate that the process may be out of control, and remedial responses are then required [43], [44]. A shift in a process parameter needs to be detected quickly so that corrective actions can be taken as early as possible.

We give below a step-by-step algorithm to implement the chart in phase I and phase II [3], [45].

- **Phase I**
  
  1. Simulate $m$ bivariate samples $(Z, Y)$ each of size $n$ from the in-control bivariate (normally distributed) process.
  2. Calculate the sample means, $(\bar{Z}_i, \bar{Y}_i)$, and the sample variances ($S^2_{Zi}$, $S^2_{Yi}$), where $S^2_{Zi} = \frac{\sum_{j=1}^{n}(Z_{ij} - \bar{Z}_i)^2}{(n-1)}$, and $S^2_{Yi} = \frac{\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_i)^2}{(n-1)}$, for each sample $i = 1, 2, \ldots, m$.  

**FIGURE 1. Application of the EWMA, HWMA, M$\mu$EWMA, and AHWMA charts.**
TABLE 10. Calculation of the AHWMA chart statistic and its limits.

<table>
<thead>
<tr>
<th>i</th>
<th>Z_i</th>
<th>Y_i</th>
<th>R_i</th>
<th>T_i</th>
<th>LCL_t</th>
<th>UCL_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.39</td>
<td>-0.865</td>
<td>0.8255</td>
<td>0.0247</td>
<td>-0.059</td>
<td>0.059</td>
</tr>
<tr>
<td>2</td>
<td>-0.242</td>
<td>-1.686</td>
<td>0.601</td>
<td>0.8159</td>
<td>-1.9095</td>
<td>1.9095</td>
</tr>
<tr>
<td>3</td>
<td>-0.919</td>
<td>-1.046</td>
<td>-0.396</td>
<td>0.6785</td>
<td>-1.3509</td>
<td>1.3509</td>
</tr>
<tr>
<td>4</td>
<td>-1.22</td>
<td>-1.366</td>
<td>-0.537</td>
<td>0.3161</td>
<td>-1.1035</td>
<td>1.1035</td>
</tr>
<tr>
<td>5</td>
<td>2.01</td>
<td>0.574</td>
<td>1.723</td>
<td>0.1706</td>
<td>-0.9561</td>
<td>0.9561</td>
</tr>
<tr>
<td>6</td>
<td>1.395</td>
<td>1.61</td>
<td>0.59</td>
<td>0.4471</td>
<td>-0.8556</td>
<td>0.8556</td>
</tr>
<tr>
<td>7</td>
<td>1.66</td>
<td>1.542</td>
<td>0.889</td>
<td>0.4799</td>
<td>-0.7814</td>
<td>0.7814</td>
</tr>
<tr>
<td>8</td>
<td>-0.514</td>
<td>0.816</td>
<td>-0.922</td>
<td>0.484</td>
<td>-0.7238</td>
<td>0.7238</td>
</tr>
<tr>
<td>9</td>
<td>-0.213</td>
<td>-0.907</td>
<td>0.2405</td>
<td>0.3431</td>
<td>-0.6774</td>
<td>0.6774</td>
</tr>
<tr>
<td>10</td>
<td>-0.588</td>
<td>-1.923</td>
<td>0.3735</td>
<td>0.3237</td>
<td>-0.6389</td>
<td>0.6389</td>
</tr>
<tr>
<td>11</td>
<td>0.074</td>
<td>0.132</td>
<td>0.008</td>
<td>0.3285</td>
<td>-0.6064</td>
<td>0.6064</td>
</tr>
<tr>
<td>12</td>
<td>1.673</td>
<td>1.64</td>
<td>0.853</td>
<td>0.3247</td>
<td>-0.5785</td>
<td>0.5785</td>
</tr>
<tr>
<td>13</td>
<td>1.765</td>
<td>0.575</td>
<td>1.4775</td>
<td>0.3875</td>
<td>-0.5541</td>
<td>0.5541</td>
</tr>
<tr>
<td>14</td>
<td>0.061</td>
<td>-0.008</td>
<td>0.065</td>
<td>0.429</td>
<td>-0.5326</td>
<td>0.5326</td>
</tr>
<tr>
<td>15</td>
<td>1.537</td>
<td>-1.084</td>
<td>2.079</td>
<td>0.4636</td>
<td>-0.5135</td>
<td>0.5135</td>
</tr>
<tr>
<td>16</td>
<td>-0.519</td>
<td>-0.52</td>
<td>-0.259</td>
<td>0.5010*</td>
<td>-0.4963</td>
<td>0.4963</td>
</tr>
<tr>
<td>17</td>
<td>1.198</td>
<td>-0.246</td>
<td>1.321</td>
<td>0.509</td>
<td>-0.4808</td>
<td>0.4808</td>
</tr>
<tr>
<td>18</td>
<td>1.853</td>
<td>0.028</td>
<td>1.839</td>
<td>0.5646*</td>
<td>-0.4666</td>
<td>0.4666</td>
</tr>
<tr>
<td>19</td>
<td>0.733</td>
<td>1.715</td>
<td>-0.1245</td>
<td>0.5765</td>
<td>-0.4537</td>
<td>0.4537</td>
</tr>
<tr>
<td>20</td>
<td>0.108</td>
<td>-0.6</td>
<td>0.408</td>
<td>0.5556*</td>
<td>-0.4418</td>
<td>0.4418</td>
</tr>
</tbody>
</table>

* Out-of-control signal

VI. ILLUSTRATIVE EXAMPLE

In this section, we provide an example to illustrate the implementation of the AHWMA chart, using the simulated dataset provided in [19]. The data were obtained by simulating $m = 20$ bivariate samples, each of size $n = 1$, from $(Z, Y) \sim N_2(\mu_Z + \delta Z, \mu_Y, \sigma_Z^2, \sigma_Y^2, \rho)$. The values of the parameters used for the simulation were: $\mu_Z = 0, \mu_Y = 0, \sigma_Z^2 = 1, \sigma_Y^2 = 0.5$, and $\delta = 0.5$, where $\delta$ is the size of the shift applied to the in-control mean, $\mu_Z$, of the process variable of interest, and $\rho$ is the correlation between the process variable and the auxiliary variable. The bivariate dataset is given in the first two columns of Table 10. We examined the ability of the AHWMA chart to detect a shift in the process variable and compared this to the $M_X$ EWMA, the classical EWMA and the HWMA charts. In all cases, the chart parameters: $w$ and $C$, were chosen to fix the $ARL_0$ to 500. The parameters for the classical EWMA and $M_X$ EWMA were $w = 0.03$ and $C = 2.483$ (see Table 6). For the classical HWMA and AHWMA charts, we used $w = 0.03$ and $C = 2.272$ (see Table 6). We give the calculations for the AHWMA chart in Table 10, and the results for all the control charts are shown graphically in Figure 1.

The AHWMA chart detected the shift in the process mean faster than any of the other methods. In particular, it detected the shift after the 14th sample, whereas the $M_X$ EWMA chart detected the shift after the 15th sample, and the classical EWMA, and HWMA charts both detected the shift only after the 18th sample.

VII. CONCLUSION AND DISCUSSION

We propose here a new efficient control-chart method, for monitoring small shifts in the process mean where the process variable of interest is correlated with and observed alongside an auxiliary variable. Based on the homogeneously weighted moving average, the proposed chart uses both the process and auxiliary variable to form a regression estimator that yields an efficient and unbiased estimate of the mean of the process variable. We provided the design structure of the chart and examined its performance in terms of its run length properties. Our simulation results showed that the chart detects a shift in the process mean more rapidly than other methods. Also, the $ARL$ comparisons showed that the proposed chart is more efficient than existing control charts used for the same purpose, especially when interest lies in detecting a small shift in the process mean. We provided a detailed study of the chart’s robustness to non-normality. The chart’s $ARL$ values showed that the chart is more robust to non-normality when a smaller value of $w$ is used. In particular, when a small value is chosen for the chart’s smoothing parameter (for example $w \leq 0.05$), the proposed chart can be designed to have an in-control $ARL$ that is reasonably close to the $ARL$ for the chart under a normally distributed process. We gave some recommendations on the application of the chart when the process parameters are unknown, and provided a step-by-step algorithm to construct the chart for phase I and phase II of SPC. Also, we applied the chart to a simulated dataset and showed that it detected a small shift in the process mean faster than other examined charts including EWMA, HWMA, $M_X$ EWMA, and $A_X$ CUSUM$_2$ methods. We consider that the effect of estimating parameters during phase I of the process on subsequent performance of the AHWMA chart warrants further study.

REFERENCES


MARTI J. ANDERSON is currently a Distinguished Professor with the New Zealand Institute for Advanced Study (NZIAS), Massey University, Auckland, New Zealand. Her current research interests include multivariate analysis, community ecology, biodiversity, ecological monitoring, experimental design, and resampling methods.

RIDWAN ADEYEMI SANUSI received the B.Sc. degree in statistics from the University of Ibadan, Ibadan, Nigeria, and the M.Sc. degree in applied statistics from the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. He is currently pursuing the Ph.D. degree with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong. His research interests include statistical process monitoring and applied statistics.

MATTHEW D. M. PAWLEY is currently a Senior Lecturer in statistics with the School of Natural and Computational Sciences, Massey University, Auckland, New Zealand. His research interests include high-dimensional data monitoring, statistical process control, data mining and analytics, and applied statistics.

* * *