The performance evaluation of the bivariate EWMAControl chart using CARL distribution and EPC

Selly Acita\(^1\), Muhammad Mashuri\(^1\), and Dedy Dwi Prastyo\(^1\)

\(^1\)Department of Statistics, Faculty of Mathematics, Computing and Data Science, Institut Teknologi Sepuluh Nopember (ITS) Surabaya 60111, Indonesia

E-mail: m_mashuri@statistika.its.ac.id

Abstract. In general, control charts are developed with the assumption that the critical quality of a production process is normally distributed with known parameters. However, in practice, critical quality is not always normally distributed and process parameters are typically unknown. In such a case, it is necessary to estimate the parameters. One of the measures to evaluate the performance control charts is Average Run Length (ARL). When the process parameters are estimated, the run-length follows its conditional distribution, so-called Conditional Average Run Length (CARL). In this study, the performance of the Bivariate Exponentially Weighted Moving Average (BEWMA) control chart will be evaluated by considering the practitioner to practitioner variability using the CARL distribution and the Exceedance Probability Criterion (EPC). The value of CARL is calculated using the Markov Chain method. The EPC is used to evaluate practitioner to practitioner variability that is closely related to parameter estimation. The results show that to guarantee the in-control performance Phase II chart based on EPC, the large size of observations in Phase I data is needed. However, in practice, it is difficult to collect such a huge size of data in Phase I. Therefore, to produce the best in-control performance with available Phase I, the control limits are adjusted.

1. Introduction

Control charts are important tools used in quality control of production to detect the out-of-control signal. Quality control of production usually involves more than one quality characteristic. One of the control charts to monitor more than one quality characteristic is Multivariate Exponentially Weighted Moving Average (MEWMA) control chart. The MEWMA control chart was introduced by [1] to detect small to medium shifts in the mean process.

The control charts are generally assumed to be normally distributed with known parameters. However, in the real case, the distribution of data is often unknown. So, it needs to be estimated. The author’s [2] and [3] explained that research on the effect of parameter estimation on the performance of control charts mostly focused on the unconditional run-length distribution. However, when the process parameters are estimated from the in-control reference sample, the run length follows the conditional distribution by giving a parameter estimate so-called CARL. Some studies suggest that the unconditional in-control ARL distribution does not represent the actual performance of the chart. The reason is that the distribution of in-control run length is conditioned on the estimated parameter, which is a random variable whose realization varies from practitioner to practitioner and will not be the same as the unconditional run length. This is because the CARL distribution is a function of parameter estimation obtained from data in Phase I that can vary significantly from data set to data set. Author [2] also make recommendations about the minimum sample size in Phase I needed to guarantee a high probability of \(1 - \beta\) as a function of Phase I is \(Mn\) (\(M\) is the number of subgroups and \(n\) is the size of subgroups), that CFAR will not exceed the specified nominal CFAR value. This is the Exceedance Probability Criterion (EPC) introduced by [4].
The author [5] examined the EWMA control chart using CARL and EPC. The results of their study indicate that for a small value of smoothing constant ($\lambda$), a larger number subgroups in Phase I is needed to guarantee the performance of an in-control Phase II chart. Based on this research, the performance evaluation of the Bivariate Exponentially Weighted Moving (BEWMA) control chart is carried out using CARL and EPC and guarantee with several available Phase I subgroups that the BEWMA Phase II performance charts remain in control. In this research, CARL values were estimated using the Markov chain method based on research [7] and [8].

2. Material and Methodology

In this section, we will explain the formulas and procedures used in this study. This study uses a BEWMA control chart with estimated parameters. The BEWMA control chart in this study will be evaluated using the CARL distribution.

2.1 BEWMA Chart with Estimated Parameters

Suppose that two variables $X_1$ and $X_2$ represent the quality characteristics that are monitored. So that $X_{it} = (X_{i1}, X_{i2})^T$ represents the observation at the time $i$ that follows the bivariate normal distribution, i.e. $X_{it} \sim N_2(\mu_{it}, \Sigma_{ii})$[8]. Statistics of BEWMA control chart with the number of subgroups $M$ and subgroup size $n$ are formulated as follows

$$Z_i = \lambda(\hat{X}_i - \mu) + (1 - \lambda)Z_{i-1},$$

with value $Z_0 = 0$. The BEWMA chart gives the out-of-control signal when

$$Q_i = Z_i^T \Sigma^{-1} Z_i > h,$$

with $h(>0)$ is selected to achieve certain ARL in-control.

In the condition of parameters $\mu$ and $\Sigma$ are unknown, these two parameters are often replaced by $\hat{\mu} = \bar{X} = \frac{1}{M} \sum_{i=1}^{M} X_i$ and $\hat{\Sigma} = \bar{S} = \frac{1}{M} \sum_{i=1}^{M} S_i$. References[9] and [10] showed that if a random sample $X_1, X_2, \ldots, X_n$ is taken from $N_r(\mu, \Sigma)$ at sampling state $i$, then the sample means vector $ar{X}_i \sim N_r(\mu, \frac{1}{n} \Sigma)$ and $(n-1) S_i \sim \text{Wishart}_r(\Sigma, (n-1))$, where Wishart$(\Sigma, (n-1))$ is the $r$-variate Wishart distribution with parameter $\Sigma$ and $n-1$ degrees of freedom. Therefore, the in-control estimator of mean vector $\bar{X} \sim N_r(\mu, \frac{1}{Mn} \Sigma)$ and $M(n-1) \bar{S} \sim \text{Wishart}_r(\Sigma, M(n-1))$.

In the Phase II chart, $Q_i$ is designed with the assumption that $2 \times 1$ random vector $X$ has a normal bivariate distribution with parameters $\mu$ and $\Sigma$. The process is said to be in-control when $\mu = \mu_0$ and $\Sigma = \Sigma_0$. Assume $\Sigma$ is positive definite and note that there is exist a non-singular matrix $V$ and $V_0$, so that $\Sigma = VV^T$ dan $\Sigma_0 = V_0V_0^T$.

Based on the same procedure in research [5], the statistical control chart of equation (1) can be written as follows

$$Y_i = \lambda W_i + (1 - \lambda)Y_{i-1}$$

with $W_i = \sqrt{n} \left( \bar{X}_i - \hat{\mu}_0 \right), \hat{\mu}_0 = \bar{X}, \bar{X}_i$ is the average of Phase II samples. The canonical form $W_i$ can be written as follows

$$W_i = \sqrt{n} \left( \bar{X}_i - \hat{\mu}_0 \right), \hat{\mu}_0 = \bar{X}, \bar{X}_i$$
\[ W_i = \Lambda T_i + \sqrt{nd} - \frac{1}{\sqrt{M}} T_0 \]

where \( T_i = \sqrt{n} \left( \frac{X_i - \mu}{V} \right), \) \( \Lambda = V_0^{-1} V, \) \( d = \left( \frac{\mu - \mu_0}{V_0} \right), \) \( T_0 = \sqrt{Mn} \left( \frac{X - \mu_0}{V_0} \right). \) So, equation (2) can be written

\[ Q_i = Y_i^T \hat{\Sigma}_Y^{-1} Y_i = \left( \Lambda T_i + \sqrt{nd} - \frac{1}{\sqrt{M}} T_0 \right) + (1 - \lambda) Y_{i-1} \]

where

\[ \hat{\Sigma}_Y = O = \begin{bmatrix} \hat{\sigma}_1^2 & \rho \hat{\sigma}_1 \hat{\sigma}_2 \\ \rho \hat{\sigma}_1 \hat{\sigma}_2 & \hat{\sigma}_2^2 \end{bmatrix} \]

Covariance matrix \( Y_i \)

\[ \hat{\Sigma}_Y = \begin{bmatrix} \frac{\lambda_1 \hat{\sigma}_1^2 (1 - (1 - \lambda_1)^2)}{2 - \lambda_1} & \frac{\lambda_1 \lambda_2 \rho \hat{\sigma}_1 \hat{\sigma}_2 (1 - (1 - \lambda_1) (1 - \lambda_2))}{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2} \\ \frac{\lambda_1 \lambda_2 \rho \hat{\sigma}_1 \hat{\sigma}_2 (1 - (1 - \lambda_1) (1 - \lambda_2))}{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2} & \frac{\lambda_2 \hat{\sigma}_2^2 (1 - (1 - \lambda_2)^2)}{2 - \lambda_2} \end{bmatrix} \]

then the asymptotic covariance matrix \( Y_i \) (when \( i \to \infty \)) is

\[ \hat{\Sigma}_Y = \begin{bmatrix} \frac{\lambda_1 \hat{\sigma}_1^2}{2 - \lambda_1} & \frac{\lambda_1 \lambda_2 \rho \hat{\sigma}_1 \hat{\sigma}_2}{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2} \\ \frac{\lambda_1 \lambda_2 \rho \hat{\sigma}_1 \hat{\sigma}_2}{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2} & \frac{\lambda_2 \hat{\sigma}_2^2}{2 - \lambda_2} \end{bmatrix} \]

The inverse of the asymptotic covariance matrix \( \hat{\Sigma}_Y^{-1} \) is

\[ \hat{\Sigma}_Y^{-1} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \]

where

\[ a = \frac{(2 - \lambda_1)(\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)^2}{\lambda_1 \hat{\sigma}_1^2 (\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)^2 - \lambda_1^2 \lambda_2 \rho^2 \hat{\sigma}_1^2 (2 - \lambda_1)(2 - \lambda_2)} \]

\[ b = \frac{\rho (2 - \lambda_1)(2 - \lambda_2)(\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)}{\lambda_1 \lambda_2 \rho \hat{\sigma}_1 \hat{\sigma}_2 (2 - \lambda_1)(2 - \lambda_2) - \hat{\sigma}_1 \hat{\sigma}_2 (\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)^2} \]

\[ c = \frac{(2 - \lambda_2)(\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)^2}{\lambda_2 \hat{\sigma}_2^2 (\lambda_1 + \lambda_2 - \lambda_1 \lambda_2)^2 - \lambda_1^2 \lambda_2 \rho^2 \hat{\sigma}_2^2 (2 - \lambda_1)(2 - \lambda_2)} \]

Note that the random variables \( T_i \) and \( T_0 \) are bivariate normal independent standard variables that are independent of each other. In the case of BEWMA it is known that
\[ X \sim N_2 \left( \mu, \frac{1}{Mn} \Sigma \right) \]

\[ M(n-1)S \sim \text{Wishart}_2(\Sigma, M(n-1)) \]

\[ T_n \sim N_2(0, I) \]

\[ T_0 = \sqrt{MnV_0^{-1}(\bar{X} - \mu_0)} \sim N_2(0, I) \]

\[ O = V_0^{-1}S(V_0^{-1})^T \]

\[ M(n-1)O \sim \text{Wishart}_2(I, M(n-1)) \]

2.2 CARL of BEWMA Chart

The control chart performance is evaluated by using a run-length distribution, especially the mean which is denoted as ARL. When parameters are estimated, run-length distribution follows the conditional distribution [2]. In this research, conditional run length distribution is a run-length distribution which is calculated based on \( T_0 \) and \( O \) values that is gained from Phase I analysis. Values from these distributions are denoted by CARL. The conditional run length distribution and CARL from the BEWMA chart can be calculated using the Markov chain method. The Markov chain for the BEWMA control chart is based on a conditional distribution.

\[ Q_i = Y_i^T \tilde{\Sigma}_i^{-1} Y_i > h \]

Equation (3) forms an oblique ellipse (figure 1) on \( Y_1 \) and \( Y_2 \) with the center \((0,0)\). The \( Y_1 \) control limit is \([-UCL_1,UCL_1]\) that is divided into \(2m_1+1\) subinterval and \( Y_2 \) control limit is \([-UCL_2,UCL_2]\) that is divided into \(2m_2+1\) subinterval. If \( g_1 \) and \( g_2 \) indicate subinterval length, then

\[ UCL_i = (2m_i + 1) \frac{g_i}{2}; \ i = 1, 2 \]

by distinguishing equation (3) with respect to \( Y_1 \) and \( Y_2 \) results in subinterval length

\[ g_1 = \frac{2\sqrt{hc}}{(2m_1 + 1)\sqrt{(ac-b^2)}} \quad \text{and} \quad g_2 = \frac{2\sqrt{ha}}{(2m_2 + 1)\sqrt{(ac-b^2)}} \]

Suppose that the subinterval axis, \( I_1, \ldots, I_{2m_1+1} \) for \( Y_1 \) and \( J_1, \ldots, J_{2m_2+1} \) for \( Y_2 \). Induction of \((2m_1+1) \times (2m_2+1)\) subrectangle \( I_i \times J_j, \ i = 1, \ldots, 2m_1 + 1 \) and \( j = 1, \ldots, 2m_2 + 1 \). Subinterval endpoints are notated with \( I_i = [A_i, B_i] = [-UCL_1 + (i-1)g_1, -UCL_1 + ig_1] \) and \( J_j = [C_j, D_j] = [-UCL_2 + (j-1)g_2, -UCL_2 + jg_2] \). The center \( I_i \times J_j \) is \((\alpha_i, \beta_j) = (-UCL_1 + (i-0.5)g_1, -UCL_2 + (j-0.5)g_2) \). The transient state of the bivariate Markov chain consists of all \( I_i \times J_j \) sub rectangles whose centers \((\alpha_i, \beta_j)\) are inside the ellipse, that is, in the control region[8].
Figure 1. State in the Markov chain. The dots represent the midpoint of the state. Adapted from [8]

Suppose the transient states $i_j$ and $k_l$ with $i,k=1,2,\cdots,2m_i+1$ and $j,l=1,2,\cdots,2m_j+1$.

This shows that the probability of $Y_i$ transition from $i_j \times J_j$ to $i_j \times J_j$ state is

$$P(Y_i \in I_i \times J_j | Y_{i,l-1} \in I_k \times J_j)$$

$$= P\left(\frac{Y_{i,l} + Y_{i,l-1} - \bar{Y}_{i,l} - \bar{Y}_{i,l-1}}{\sqrt{\text{Var}(Y_{i,l})}} \in \mathbb{R}^2, \frac{Y_{k,l} + Y_{k,l-1} - \bar{Y}_{k,l} - \bar{Y}_{k,l-1}}{\sqrt{\text{Var}(Y_{k,l})}} \in \mathbb{R}^2\right)$$

$$= P\left(\frac{\hat{Y}_{i,l} + \hat{Y}_{i,l-1} - \bar{Y}_{i,l} - \bar{Y}_{i,l-1}}{\sqrt{\text{Var}(Y_{i,l})}} \in \mathbb{R}^2, \frac{\hat{Y}_{k,l} + \hat{Y}_{k,l-1} - \bar{Y}_{k,l} - \bar{Y}_{k,l-1}}{\sqrt{\text{Var}(Y_{k,l})}} \in \mathbb{R}^2\right)$$

where $\alpha_k = -UCL_1 + (k-0.5)g_1$ and $\beta_l = -UCL_2 + (l-0.5)g_2$

$$P\left(\frac{\hat{Y}_{i,l} + \hat{Y}_{i,l-1} - \bar{Y}_{i,l} - \bar{Y}_{i,l-1}}{\sqrt{\text{Var}(Y_{i,l})}} \in \mathbb{R}^2\right)$$

Assume that there is no change in the covariance matrix. Thus $\Lambda = I$.  

\[ P \left( A < \lambda_1 \left( T_i + \sqrt{n}d_i - \frac{T_{t0}}{\sqrt{M}} \right) + (1 - \lambda_1) \alpha_k \leq B_i \right) \times P \left( C_i < \lambda_2 \left( T_{i2} + \sqrt{n}d_{i2} - \frac{T_{t20}}{\sqrt{M}} \right) + (1 - \lambda_2) \beta_i \leq D_i \right) \times \left( 1 - \lambda_1 \right) \alpha_k = \alpha_i \times \left( 1 - \lambda_2 \right) \beta_i = \beta_i \times \frac{A_i - (1 - \lambda_1) \alpha_k}{\lambda_1} - \sqrt{n}d_i + \frac{T_{t10}}{\sqrt{M}} \leq \frac{B_i - (1 - \lambda_1) \alpha_k}{\lambda_1} - \sqrt{n}d_i + \frac{T_{t10}}{\sqrt{M}} \right) \times P \left( C_i - (1 - \lambda_2) \beta_1 \leq \frac{D_i - (1 - \lambda_2) \beta_1}{\lambda_2} - \sqrt{n}d_{i2} + \frac{T_{t20}}{\sqrt{M}} \right) \left( 1 - \lambda_1 \right) \alpha_k = \alpha_i \times \frac{A_i - (1 - \lambda_1) \alpha_k}{\lambda_1} - \sqrt{n}d_i + \frac{T_{t10}}{\sqrt{M}} \right). \]

The application of the Markov chain method conditionally on \( O \) and \( T_0 \) from the BEWMA Phase II chart can be written

\[ \text{CARL} = u' \left( I - P \right)^{-1} a, \]

with \( u' \) is a row vector \( 1 \times \left( (2m_1 + 1)(2m_2 + 1) \right) \) with one in the center position and null for the other entries, \( a \) is a column vector one with the size \( \left( (2m_1 + 1)(2m_2 + 1) \right) \times 1 \), \( I \) is the identity matrix, \( P = \left[ p_{(k-1)(2m_1+1)l, (l-1)(2m_1+1)} \right] \) is a conditional transition probability matrix of size \( \left( (2m_1 + 1)(2m_2 + 1) \right) \times \left( (2m_1 + 1)(2m_2 + 1) \right) \), \( i,k = 1, \ldots, 2m_1 + 1 \) and \( j,l = 1, \ldots, 2m_2 + 1 \) on condition \( a(\alpha - (m_1 + 1))^2 + 2b(\alpha - (m_1 + 1))(\beta - (m_2 + 1))g_1g_2 + c(\beta - (m_2 + 1))^2 g_2^2 < h \), with \( \alpha = 1, \ldots, 2m_1 + 1 \) and \( \beta = 1, \ldots, 2m_2 + 1 \).

\[ p_{(k-1)(2m_1+1)l, (l-1)(2m_1+1)} = \left[ \Phi \left( \frac{B_i - (1 - \lambda_1) \alpha_k}{\lambda_1} - \sqrt{n}d_i + \frac{T_{t10}}{\sqrt{M}} \right) - \Phi \left( \frac{A_i - (1 - \lambda_1) \alpha_k}{\lambda_1} - \sqrt{n}d_i + \frac{T_{t10}}{\sqrt{M}} \right) \right] \times \left[ \Phi \left( \frac{D_i - (1 - \lambda_2) \beta_1}{\lambda_2} - \sqrt{n}d_{i2} + \frac{T_{t20}}{\sqrt{M}} \right) - \Phi \left( \frac{C_i - (1 - \lambda_2) \beta_1}{\lambda_2} - \sqrt{n}d_{i2} + \frac{T_{t20}}{\sqrt{M}} \right) \right] \]

where \( \Phi \) is the cdf function of standard normal distribution.

The CARL values depend on the random variables \( O \) and \( T_0 \). So, we can say that the CARL is a random variable. In the study of [5], the CARL value varies to nominal ARL. When parameters are estimated using limits for known parameter case to design a Phase II BEWMA chart is risky, because it can produce a small in-control CARL value. We can increase the number of subgroup \( M \) to reduce risk and get a large in-control CARL value. However, the required \( M \) value can be very large. Some literature design control charts with estimated parameters with

\[ P(\text{CARL}_n > \text{ARL}_0) = 1 - p. \]

Therefore, ARL \( _0 \) is the 100\( p \)th percentile value of in-control CARL. This is an EPC approach that is used to evaluate and design the BEWMA chart.

2.3 Procedure

1. Using R package “spc”, we estimate the control limits \( h \) needed to design the ARL 100, 200, 370, and 500 in-control values when different values of the smoothing constant (\( \lambda \)) are used to design a BEWMA control chart with known parameters.
2. Simulating an observation of a standard normal bivariate distribution \( T_0 \).
3. Simulating a matrix $Y$ of the Wishart bivariate distribution with $M(n-1)$ degrees of freedom and the covariance matrix is identity, i.e. $\text{Wishart}_p \left( I, M(n-1) \right)$. Calculate $O = \frac{1}{M(n-1)} Y$.

4. Constructing the transition probability matrix $P = [p_{(k+1)(2m+1)\ell+1}, (i+1)(2m+1)\ell+1]$ using simulation data in step 2-3 and calculate $\text{CARL}_{IN}$ (CARL value when zero shift ($d = 0$)).

5. Repeating steps 1-4 many times (say 1000 times). After obtaining 1000 $\text{CARL}_{IN}$ values, calculate the value of the percentile of $\text{CARL}_{IN}$ in this study using the 10th percentile.

6. Evaluating the performance of a standard Phase II BEWMA chart using EPC based on the $\text{CARL}_{IN}$ value.

7. Adjusting the control limits to achieve the EPC criteria, which is guaranteed with a high probability that the CARL value is greater than the specified $\text{ARL}_0$ value.

3. Result and Discussion

The following are the results of the study using CARL and EPC criteria used to evaluate the performance of BEWMA Phase II charts.

3.1 Performance of A Standard Phase II BEWMA Chart using EPC

EPC is used to evaluate the BEWMA chart because the EPC approach takes into account variability in-control CARL with

$$P(\text{CARL}_{IN} \leq \text{ARL}_0) = p.$$  

So for $p \in (0,1)$, we sought $100p$th percentile that symbolized $\text{CARL}_{IN,p}$. $\text{CARL}_{IN,p}$ value compared with $\text{ARL}_0$, which is a theoretical value that must be exceeded with probability $1-p$. Comparison between $\text{CARL}_{IN,p}$ and $\text{ARL}_0$ is calculated based on the percentage different (PD) formulated as follows $PD = \frac{\text{CARL}_{IN,p} - \text{ARL}_0}{\text{ARL}_0} \times 100$, see [5] and [6] as the reference.

**Table 1.** The $5^{th}$ and $10^{th}$ percentile of $\text{CARL}_{IN}$ values for $\lambda = 0.5, 0.1$, $n = 5$ and $\text{ARL}_0 = 100, 200, 370, 500$

| $\lambda$ | M | $\text{ARL}_0=100$ (h = 6.980) | $\text{ARL}_0=200$ (h = 8.641) | $\text{ARL}_0=370$ (h = 10.091) | $\text{ARL}_0=500$ (h = 10.791) |
|----------|---|-----------------|-----------------|-----------------|-----------------|
| 0.1      | 30 | p = 0.05        | p = 0.10        | p = 0.05        | p = 0.10        |
|          |    | (-71.36%)       | (-64.74%)       | (-85.05%)       | (-81.23%)       |
|          |    | (-89.31%)       | (-86.85%)       | (-90.95%)       | (-88.35%)       |
|          | 50 | (-67.01%)       | (-61.32%)       | (-77.14%)       | (-72.55%)       |
|          |    | (-82.24%)       | (-78.73%)       | (-83.29%)       | (-80.64%)       |
|          | 100| (-50.24%)       | (-40.94%)       | (-63.73%)       | (-57.54%)       |
|          |    | (-64.86%)       | (-58.77%)       | (-72.66%)       | (-63.80%)       |
|          | 500| (-20.80%)       | (-19.27%)       | (-24.87%)       | (-20.85%)       |
|          |    | (-32.31%)       | (-28.52%)       | (-35.59%)       | (-31.59%)       |
|          | 1000| (-11.87%)       | (-10.33%)       | (-19.02%)       | (-13.16%)       |
|          |    | (-20.37%)       | (-18.26%)       | (-21.88%)       | (-19.38%)       |
| 1500     |    | (-11.20%)       | (-8.17%)        | (-12.52%)       | (-9.79%)        |
|          |    | (-14.18%)       | (-12.56%)       | (-18.00%)       | (-14.38%)       |
| 2000     |    | (-8.61%)        | (-8.17%)        | (-10.10%)       | (-8.83%)        |
|          |    | (-12.35%)       | (-11.12%)       | (-15.45%)       | (-12.34%)       |
| 4000     |    | (-5.71%)        | (-4.89%)        | (-7.37%)        | (-6.53%)        |
|          |    | (-7.56%)        | (-6.93%)        | (-8.10%)        | (-7.23%)        |
In Table 1, the negative values in parentheses are PD values, which indicates that the CARL_{IN,p} less than ARL_0. PD value will be positive if the CARL_{IN,p} value is higher than ARL_0. Based on table 1, it can be seen that the CARL_{IN,p} value from the Phase II BEWMA chart with the determined ARL_0 value has a high value (in absolute value) for a small M value. For example, when smoothing constant λ = 0.1, h=10.091, p =0.05 and the parameter estimated using M = 30 Phase I samples, the CARL_{IN,p} value is 39.56 which is 89.31% below the value of ARL_0 = 370. So, we expect that CARL_{IN} from the chart at least 39.56 with 95% probability, or conversely. In other words, we want the chart to provide at least a large CARL_{IN} value with 95% certainty. CARL_{IN} value increases when M increases and convergence faster when value λ = 0.5. A large p also increases CARL_{IN} value. Therefore, when parameters are estimated, small CARL_{IN} values occur more frequently than high CARL_{IN} values (CARL_{IN}> ARL_0).

| λ  | M     | ARL_p=100  | ARL_p=200  | ARL_p=370  | ARL_p=500  |
|----|-------|------------|------------|------------|------------|
| p  | 0.05  | 0.10       | 0.05       | 0.10       | 0.05       | 0.10       |
| 0.1| 95.65 | 95.97      | 189.79     | 190.38     | 347.65     | 348.47     | 465.82     | 470.93     |
|    | (-4.35%) | (-4.03%) | (-5.11%) | (-4.81%) | (-6.04%) | (-5.82%) | (-6.837%) | (-5.81%) |
| 0.5| 30    | 34.03      | 36.95      | 55.63      | 60.40      | 93.41      | 10.64      | 98.54      | 131.49     |
|    |      | (-65.97%) | (-63.05%) | (-72.19%) | (-69.80%) | (-74.75%) | (-70.64%) | (-80.29%) | (73.70%)   |
| 50 | 52.98 | 57.15      | 80.79      | 91.24      | 125.65     | 143.28     | 153.87     | 167.75     |
|    |      | (-47.02%) | (-42.85%) | (-59.61%) | (-54.38%) | (-66.04%) | (-61.28%) | (-69.22%) | (-66.45%)  |
| 100| 66.48 | 69.82      | 98.20      | 121.05     | 192.76     | 209.16     | 228.90     | 280.54     |
|    |      | (-33.52%) | (-30.18%) | (-50.90%) | (-39.47%) | (-47.90%) | (-43.47%) | (-55.80%) | (-43.89%)  |
| 500| 84.91 | 87.55      | 155.33     | 165.74     | 286.37     | 292.11     | 380.99     | 391.97     |
|    |      | (-15.09%) | (-12.45%) | (-22.33%) | (-17.13%) | (-22.60%) | (-21.05%) | (-23.80%) | (-21.61%)  |
| 1000| 85.97| 89.59     | 164.42     | 172.73     | 302.53     | 309.04     | 415.56     | 424.64     |
| 1500| 90.82| 91.33     | 174.64     | 179.55     | 318.62     | 328.23     | 433.78     | 439.34     |
|    |      | (-9.17%) | (-8.67%) | (-12.68%) | (-10.22%) | (-13.89%) | (-11.29%) | (-13.24%) | (-12.13%)  |
| 2000| 90.71| 91.49     | 176.85     | 181.38     | 328.64     | 336.16     | 435.20     | 443.19     |
|    |      | (-9.29%) | (-8.05%) | (-11.58%) | (-9.31%) | (-11.18%) | (-9.14%) | (-12.96%) | (-11.36%)  |
| 4000| 92.44| 93.48     | 181.35     | 183.43     | 336.09     | 341.35     | 450.08     | 455.37     |
|    |      | (-7.56%) | (-6.52%) | (-9.33%) | (-8.28%) | (-9.16%) | (-7.74%) | (-9.98%) | (-8.92%)   |
| 10000| 94.85| 95.34    | 186.82     | 187.81     | 353.86     | 339.59     | 461.40     | 464.21     |
|    |      | (-5.15%) | (-4.66%) | (-6.56%) | (-6.09%) | (-9.23%) | (-8.22%) | (-7.72%) | (-7.16%)   |

Table 2. The minimum M required CARL_{IN,p} to be ε = 0%, 10%, 20% for λ = 0.5, 0.1, n = 5 and ARL_0 = 100,200,370,500

| λ  | ε    | ARL_p=100  | ARL_p=200  | ARL_p=370  | ARL_p=500  |
|----|------|------------|------------|------------|------------|
| p  | 0.05 | 0.10       | 0.05       | 0.10       | 0.05       | 0.10       |
| 0.1| 0%   | >10000     | >10000     | >10000     | >10000     | >10000     | >10000     |
|    | 10%  | 1500       | 1000       | 2000       | 1500       | 2000       | 4000       |
|    | 20%  | 500        | 500        | 1000       | 500        | 1000       | 1000       |
| 0.5| 0%   | >10000     | >10000     | >10000     | >10000     | >10000     | >10000     |
|    | 10%  | 1500       | 1000       | 4000       | 1500       | 4000       | 4000       |
|    | 20%  | 500        | 500        | 500        | 500        | 500        | 500        |

Table 2 shows the number of subgroups M needed to guarantee the CARL_{IN,p} value is higher than ARL_0 with (1-p) probability formulated with
\[ P\left( \text{CARL}_{\text{IN}} > \text{CARL}_{\text{IN},p} \right) \geq 1 - p \]
\[ P\left( \text{CARL}_{\text{IN}} > \text{ARL}_0 \left( 1 - \varepsilon \right) \right) \geq 1 - p. \]

\( \varepsilon \) is the specified PD value. The number of \( M \) needed to achieve EPC in-control performance with a value of \( \varepsilon = 0\% \) for all value combinations and is higher than 10000 subgroups. If the value of \( \varepsilon \) gets larger, the number of Phases I subgroups needed will be smaller.

### 3.2 Adjustment of BEWMA Chart Limits

Based on the evaluation results, we need a very large number of Phase I subgroups to achieve adequate Phase II EPC performance. However, in real cases, large Phase I subgroup samples are not available or difficult to obtain. When we use a few phases I subgroups to evaluate the BEWMA control chart with parameter estimates, it will produce inferior chart performance. Therefore, to overcome the poor chart performance, it is necessary to adjust the limit to ensure that the chart remains in-control.

| \( \lambda \) | M   | 100 | 200 | 370 | 500 |
|-----------|-----|-----|-----|-----|-----|
| 0.1       | 30  | 12.68| 16.04| 18.79| 19.29|
| 50        |     | 10.43| 12.09| 15.64| 15.99|
| 100       |     | 9.08 | 11.44| 13.09| 14.29|
| 500       |     | 7.43 | 9.29 | 10.94| 11.79|
| K         |     | 6.980| 8.641| 10.091| 10.791|

\*K is the case when the parameter is known

The value \( \lambda = 0.1 \) is a value commonly used as the default to calculate control statistics based on [11], so the adjustment of the control limits in table 3 uses \( \lambda = 0.1 \). Table 3 shows that when control limits are adjusted to a certain Phase I sample size, the control limit values used to design the Phase II control chart will have wider control limits. In other words, if the number of Phases I subgroup samples to design a Phase II chart gets smaller, the control limits obtained will be wider.

### 4. Conclusion

Based on the results of the study, to achieve the \( \text{CARL}_{\text{IN}} \) value based on the predetermined \( \text{ARL}_0 \) with \( (1 - p) \) probability, a very large number of Phase I subgroups (\( M > 10000 \)) are needed. Due to the difficulty of obtaining the large \( M \) at Phase I, the control limits to design Phase II BEWMA chart are adjusted for the number of available Phases I subgroups. The new control limit value \( h \) has a higher value than is commonly used for the case \( K \) (i.e., the control limits from the condition when the parameters are known). Thus, the BEWMA control chart that was built using the new limit value \( h \) has a wider limit, especially when the number of subgroups \( M \) used is small. In this research, we use the same smoothing constant value for each variable. In further research, this research can be developed by using different smoothing constant values on each variable. This research can also be developed using the copula function. So, when the production process is not normally distributed, we can use the production process without ignoring the assumption of non-normal distribution.

### Acknowledgments

The authors thank to the Department of Statistics, Faculty of Mathematics, Computing and Data Science, Institut Teknologi Sepuluh Nopember (ITS).
References

[1] Lowry C A, Woodall W H, Champ C W, and Ridgon S E 1992 A Multivariate Exponentially Weighted Moving Average Control Chart Technometrics 34(1)46-53

[2] Epprecht E K, Loureiro L D, and Chakraborti S 2015 Effect of the Amount of Phase I Data on the Phase II Performance of $S^2$ and $\bar{S}$ Control Charts Journal of Quality Technology 47(2) 139-155

[3] Kumar R and Chakraborti S 2016 Phase II Shewhart-type Control Chart for Monitoring Times Between Events and Effects of Parameter Estimation Quality and Reliability Engineering International 32(1) 315-328

[4] Albers W, Kallenberg W C M, and Nurdiati S 2005 Exceedance Probabilities for Parametric Control Charts Statistics 39(5) 429-443

[5] Diko M D, Chakraborti S, and Does R J M M 2019 Guaranteed In-Control Performance of the EWMA Chart for Monitoring the Mean Quality and Reliability Engineering International

[6] Diko M D, Chakraborti S, and Does R J M M 2019 An Alternative design of the two-side CUSUM chart for monitoring the mean parameters are estimated Computers & Industrial Engineering

[7] Runger G C and Prabu S S 1996 A Markov Chain Model for the Multivariate Exponentially Weighted Moving Averages Control Chart. Journal of the American Statistical Association 91(436) 1701-1706

[8] Huh I, Viveros-Aguilera R, and Balakrishnan N 2017 Differential Smoothing in the Bivariate Exponentially Weighted Moving Average Chart Journal of Quality Technology 45(4) 377-393

[9] Anderson T W 2013 An Introduction to Multivariate Statistical Analysis Third Edition, New Jersey, John Wiley & Sons

[10] Mahmoud A, Maravelakis P E 2010 The Performance of the MEWMA Control Chart when Parameters are Estimated Communications in Statistics – Simulation and Computation 39(9) 1803-1817

[11] Montgomery D C 2013 Introduction to Statistical Quality Control 7th Edition, New Jersey, John Wiley & Sons