Design of EWMA-type control chart

H W You¹, M A Shahrin² and Z Mustafa²

¹ Pusat GENIUS@Pintar Negara, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
² Pusat Pengajian Sains Matematik, Fakulti Sains dan Teknologi, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
Corresponding author: hwyou@ukm.edu.my

Abstract. A control chart is typically examined based on the overall performance. This is crucial, as it will influence the final decision on the selection of a control chart. For this reason, the application of average run length (ARL) has been identified as the standard performance metric. In order to calculate ARL, practitioners must determine the size of the process shift in advance. Nevertheless, in practice, the process shift size cannot be predetermined, as practitioners may not have the requisite information associated with the process. Hence, in this work, the expected average run length (EARL) is employed as a performance metric for the unknown shift size in a process. This work evaluates the performance of a memory-type control chart, i.e. exponentially weighted moving average median (EWMA median) chart, with respect to ARL and EARL. The results indicate that, as long as the deterministic shift size is within the range of the process shift size, the chart’s performance is nearly the same, regardless of the use of the optimal pair based on minimizing ARL or EARL.

1. Introduction
In the business world, the primary objective is to produce high quality products or services. In view of this, a collection of statistical tools with the objective of reducing the variability in the process has been suggested, namely, Statistical Process Control (SPC). A control chart is a useful tool in SPC [1], and has been widely adopted in business settings. This has motivated researchers to develop novel control charts that are more efficient with respect to the detection ability. For example, memory-type control charts have been introduced for better detection ability in a process. Some recent works by [2]-[4], to name a few, investigated memory-type control charts extensively.

The exponentially weighted moving average (EWMA) mean chart is the first EWMA chart introduced by [5]. The EWMA mean chart performs well in detecting small-to-moderate shifts in a process, under normality assumption. However, [6] showed that when the process contains contaminated normal data, the EWMA mean chart is insensitive to detect process shifts. In view of this, the EWMA median chart was introduced by [2] under such settings. The key benefit of the EWMA median chart is that it is robust against outliers.

After evaluation of the performance of a control chart, average run length (ARL) was considered as a standardized performance metric. The ARL is defined as the expected number of samples taken before the process is signalled as out-of-control [7]. The employment of ARL requires the practitioner to specify the magnitude of a shift. In real-world scenarios, the practitioners have limited knowledge to predefined the process shift size [8], [9]. This eventually leads to the introduction of the expected average
run length (EARL), which is used to investigate the performance of the EWMA median chart. This is because the computation of EARL does not demand practitioners to define the process shift size. Moreover, research dealing with EARL as a performance measure in EWMA median charts is not yet available in the related literature.

This work aims to investigate the overall performance evaluation of the EWMA median chart using ARL and EARL. The run length profiles of the EWMA median chart are demonstrated in Section 2. Section 3 presents the comparison performance measure using ARL and EARL. Finally, some conclusions are drawn in the final section.

2. EWMA median chart

The EWMA median chart was pioneered by [2]. In this particular median chart, the sample median is used to monitor the process, as follows:

\[
\tilde{T}_j = \begin{cases} 
T_{i,(n+1)/2} & \text{if } n \text{ is odd} \\
\frac{T_{i,(n/2)} + T_{i,(n/2+1)}}{2} & \text{if } n \text{ is even}
\end{cases}
\]  

(1)

where \( n \) denotes the sample size, and we heuristically make the assumption that \( n \) is an odd number. With such an assumption, the sample median \( \tilde{T}_i \) can be determined more quickly and easily. Work by [6] shows that the sample median is an ideal alternative to monitor if the process contains contaminated normal data.

Assuming that the process is distributed independently and identically (i.i.d) and has a normal distribution with the in-control mean and in-control standard deviation, i.e. \( \mu_0 \) and \( \sigma_0 \), respectively, the statistic of the EWMA median chart is defined as follows:

\[
U_j = (1 - a)U_{j-1} + a\tilde{T}_j,
\]

(2)

for \( i \in \{1, 2, \ldots\} \). Note that \( U_0 = \mu_0 \) and \( a \) represent a smoothing constant, i.e. \( a \in (0, 1] \). The control limits, i.e. \( UCL \) and \( LCL \), which represents the upper control limit and lower control limit of the EWMA median control chart, respectively, are defined as follows:

\[
UCL / LCL = \mu_0 \pm b\sigma_0
\]

(3)

where \( b \) is the design constant.

Here, the Markov chain method is applied to obtain the run length profiles of the EWMA median chart. The transition probability matrix (tpm) for the transient states is denoted as \( P \), as follows:

\[
P = \begin{bmatrix} P_{i,j} \end{bmatrix}_{i \times j},
\]

(4)

where \( i, j = 1, 2, \ldots, l \). The interval between the control limits is divided into subintervals, \( l = 2g+1 \), each with a width of \( 2w \). Let \( K_j \) be the midpoint of the \( j^{th} \) subinterval, for \( j = 1, 2, \ldots, 2g+1 \). The \( P_{i,j} \) corresponding to the matrix \( P \) are as follows:

\[
P_{i,j} = F_{\tilde{T}} \left( \frac{K_j + w - (1-a)K_i}{a} \right) - F_{\tilde{T}} \left( \frac{K_j - w - (1-a)K_i}{a} \right)
\]

(5)
for \(i, j = 1, 2, \ldots, 2g+1\). Here, \(F_{T_i}(...|n)\) is the cumulative distribution function (cdf) of \(T_i\), with \(i = 1, 2, \ldots\).

Finally, the ARL is computed as follows:

\[
\text{ARL} = v_i, \tag{6}
\]

where

\[
v_i = p^T (I - P)^{-1} \cdot \tag{7}
\]

Note that \(p\) is the vector of initial probabilities and \(I\) is the identity matrix. The computation of the ARL requires the practitioners to pre-determine the size of the process shift. In practical application, the practitioners may not know the precise size of the process shift in advance, as they do not have past knowledge of this process. Accordingly, EARL criterion is used as the EWMA median chart’s performance measure, and is defined as follows:

\[
\text{EARL} = \int_{\delta_{\text{low}}}^{\delta_{\text{up}}} f_\delta (\delta) \text{ARL} \, d\delta. \tag{8}
\]

Note that \(f_\delta (\delta)\) is the probability density function of the magnitude of the process shift, i.e. \(\delta\).

### 3. Performance of the EWMA median chart with unknown shift sizes

The efficiency of a control chart is generally assessed by the ARL. Here, in-control and out-of-control ARL are denoted as \(\text{ARL}_0\) and \(\text{ARL}_1\), respectively. The \(\text{ARL}_1\) for various combinations of sample sizes, \(n \in \{3, 5, 7, 9\}\) and the process shift sizes, \(\delta \in \{0.25, 0.5, 0.75, 0.95, 1.25, 1.5, 1.75, 2.0\}\) are presented in Table 1.

| \(n\) | \(\delta\) | \((a, b)\) | \(\text{ARL}_1\) | \(n\) | \(\delta\) | \((a, b)\) | \(\text{ARL}_1\) |
|---|---|---|---|---|---|---|---|
| 3 | 0.25 | \(0.1000, 0.4155\) | 46.53 | 0.25 | \(0.1000, 0.3322\) | 31.78 |
| 0.50 | \(0.1000, 0.4155\) | 14.87 | 0.50 | \(0.1265, 0.3847\) | 10.64 |
| 0.75 | \(0.1648, 0.5669\) | 8.09 | 0.75 | \(0.2273, 0.5538\) | 5.76 |
| 0.95 | \(0.2311, 0.7000\) | 5.66 | 0.95 | \(0.3163, 0.6825\) | 4.03 |
| 1.25 | \(0.3393, 0.8932\) | 3.74 | 1.25 | \(0.4768, 0.8952\) | 2.67 |
| 1.50 | \(0.4457, 1.0696\) | 2.84 | 1.50 | \(0.6294, 1.0918\) | 2.01 |
| 1.75 | \(0.5684, 1.2675\) | 2.24 | 1.75 | \(0.7571, 1.2613\) | 1.57 |
| 2.00 | \(0.6834, 1.4546\) | 1.81 | 2.00 | \(0.8521, 1.3940\) | 1.29 |
| \(7\) | 0.25 | \(0.1000, 0.2845\) | 24.51 | 0.25 | \(0.1000, 0.2528\) | 20.20 |
| 0.50 | \(0.1589, 0.3795\) | 8.43 | 0.50 | \(0.1889, 0.3753\) | 7.04 |
| 0.75 | \(0.2830, 0.5444\) | 4.56 | 0.75 | \(0.3348, 0.5383\) | 3.81 |
| 0.95 | \(0.3968, 0.6772\) | 3.19 | 0.95 | \(0.4788, 0.6823\) | 2.67 |
| 1.25 | \(0.6072, 0.9098\) | 2.09 | 1.25 | \(0.7089, 0.9088\) | 1.73 |
| 1.50 | \(0.7588, 1.0813\) | 1.56 | 1.50 | \(0.8443, 1.0503\) | 1.31 |
| 1.75 | \(0.8663, 1.2103\) | 1.25 | 1.75 | \(0.9300, 1.1461\) | 1.11 |
| 2.00 | \(0.9362, 1.2995\) | 1.10 | 2.00 | \(0.9761, 1.2005\) | 1.03 |

The optimal pair \((a, b)\) is obtained in order to attain the intended \(\text{ARL}_0 = 370\) for the corresponding \(n\) and \(\delta\). For example, in the case \(n = 5\) and \(\delta = 0.95\), the optimal pair, \((a, b) = (0.3163, 0.6825),\)
yielding the lowest \( ARL_1 \) value, i.e. \( ARL_1 = 4.03 \). Meanwhile, these optimal pairs achieve the desired \( ARL_0 = 370 \).

According to Table 1, for a fixed \( n \), the \( ARL_1 \) value decreases as \( \delta \) increases. For instance, for \( n = 3 \), when \( \delta = 0.25 \), the corresponding \( ARL_1 = 46.53 \) with the optimal pair, \( (a, b) = (0.1000, 0.4155) \), as opposed to when \( \delta = 1.50 \), the \( ARL_1 = 2.84 \) is attained using the optimal pair, \( (a, b) = (0.4457, 1.0696) \).

Nevertheless, the calculation of the ARL required the practitioners to define the process shift. In real-world application, the practitioners may not have the requisite judgement on the process. Therefore, if a practitioner considers a particular shift size, \( \delta \), and uses the corresponding optimal pair, the performance of the EWMA median chart is affected if a different shift size is detected in the process. In light of this, the EARL is proposed to monitor the process when the size of the shift is unknown. Here, two EARLs are usually of interest, namely, the in-control EARL, \( EARL_0 \), and the out-of-control EARL, \( EARL_1 \).

Table 2 reveals the optimal pair \((a, b)\) that is obtained by minimizing the \( EARL_1 \) of the EWMA median chart, based on the corresponding \((\delta_{\text{min}}, \delta_{\text{max}}) = (0.25, 2.0)\). Note that the \( EARL_0 \) is set at 370. For comparison purposes, the similar \( n \) value is shown in Table 2.

| \(n\) | 3     | 5     | 7     | 9     |
|-------|-------|-------|-------|-------|
| \((a, b)\) | (0.1033, 0.4240) | (0.1366, 0.4035) | (0.1677, 0.3925) | (0.2032, 0.3928) |
| \(EARL_1\) | 8.47  | 6.15  | 4.94  | 4.18  |

| \(\delta\) | 0.25  | 0.50  | 0.75  | 0.95  | 1.25  | 1.50  | 1.75  | 2.00  |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|
| \(\delta_{\text{min}}\) | 47.02 | 14.89 | 8.35  | 6.18  | 4.48  | 3.68  | 3.14  | 2.75  |
| \(\delta_{\text{max}}\) | 34.94 | 10.65 | 6.00  | 4.47  | 3.28  | 2.72  | 2.34  | 2.10  |
| \(0.25\) | 28.38 | 8.43  | 4.77  | 3.57  | 2.65  | 2.23  | 1.98  | 1.78  |
| \(0.50\) | 24.52 | 7.05  | 3.98  | 3.00  | 2.25  | 1.93  | 1.68  | 1.44  |

According to Table 2, when \( n = 3 \), \( \delta_{\text{min}} = 0.25 \) and \( \delta_{\text{max}} = 2.0 \), the optimal pair is \((a, b) = (0.1033, 0.4240)\), yielding the smallest \( EARL_1 \), i.e. \( EARL_1 = 8.47 \) while attaining the \( EARL_0 = 370 \). By considering \( \delta = 0.5 \) (i.e. \( \delta \in (\delta_{\text{min}}, \delta_{\text{max}}) \)) for the same \( n \) value, the \( ARL_1 = 14.89 \) is obtained using the optimal pair \((a, b) = (0.1033, 0.4240)\) from Table 2. Meanwhile, for the same combination \((n, \delta)\), the \( ARL_1 = 14.87 \) using the optimal pair \((a, b) = (0.1000, 0.4155)\). This demonstrates that the performance of the EWMA median chart is nearly similar in the case of using the optimal pair based on minimizing \( ARL_1 \) and \( EARL_1 \). In view of this, when the practitioner is unable to determine the precise process shift size, the optimal pair evaluated by minimizing the \( EARL_1 \) can be employed.

4. Conclusion
This paper explicitly demonstrated that the EARL can be applied as an alternative performance metric for the EWMA median chart in the case the process shift size is unknown in advance. Moreover, the
The proposed optimal pair \((a, b)\) based on minimizing the EEARL\(_1\) can be implemented, as long as the considered shift size is within \((\delta_{\text{min}}, \delta_{\text{max}})\). This can overcome the inaccuracy of the process when monitoring using EWMA median chart, in the case a different shift size is detected. This study is based on the process parameters, i.e. the in-control mean and standard deviation are known. In real-world scenarios, the process parameters may not be known. Therefore, further research to examine the performance of the EWMA median chart with exact and random process shifts, when process parameters are unknown, is needed. The findings of this research are to be discussed in our future work.

5. References

[1] Montgomery DC 2019 Introduction to Statistical Quality Control. New York: John Wiley & Sons, Inc.

[2] Castagliola P 2001 An \(\bar{X}/R\)–EWMA control chart for monitoring the process sample median Int. J. Reliab. Qual. Saf. Eng. 8 123-135.

[3] Raza SMM and Siddiqi AF 2017 EWMA and DEWAM control charts for Poisson-Exponential Distribution: Conditional median approach for censored data Qual. Reliab. Eng. Int. 33 387-399.

[4] Mitra A, Kok LK and Chakraborti S 2019 An adaptive exponentially weighted moving average-type control chart to monitor the process mean Eur. J. Oper. Res. 279 902-911.

[5] Roberts SW 1959 Control chart tests based on geometric moving averages Technometrics 1 239-250.

[6] Human SW, Kritzinger PE and Chakraborti S 2011 On the robustness of the EWMA chart J. Appl. Stat. 38 2071-2087.

[7] Tang A, Castagliola P, Sun J and Hu X 2019 Optimal design of the adaptive EWMA chart for the mean based on median run length and expected median run length Qual. Technol. Quant. Manag. 16 439-458.

[8] You HW, Khoo MBC, Castagliola P and Qu L 2016 Optimal exponentially weighted moving average charts with estimated parameters based on median run length and expected median run length Int. J. Prod. Res. 54 5073-5094.

[9] Chong ZL, Khoo MBC, Teoh WL, You HW and Castagliola P 2019 Optimal design of the side-sensitive modified group runs (SSMGR) chart when process parameters are estimated Qual. Reliab. Eng. Int. DOI: 10.1002/qre.2395.

Acknowledgement

This research is supported by the Universiti Kebangsaan Malaysia, Geran Galakan Penyelidik Muda, GGPM-2017-062.