Optimal one- and two-sided adaptive EWMA scheme for monitoring Poisson count data

Anan Tang, Philippe Castagliola, Xuelong Hu, Xiaojian Zhou

Optimal one- and two-sided adaptive EWMA scheme for monitoring Poisson count data

Anan Tang[a], Philippe Castagliola[b], Xuelong Hu[a,*], Xiaojian Zhou[a]

[a] *Nanjing University of Posts and Telecommunications, Nanjing, China*
[b] *Université de Nantes & LS2N UMR CNRS 6004, Nantes, France*

Abstract

The Poisson distribution assumption often arises in several industrial applications for modeling defects or non-conformities. In this work, we investigate the one- and two-sided performance of a new adaptive EWMA (Exponentially Weighted Moving Average)-type chart for monitoring Poisson count data. An appropriate discrete-state Markov chain technique is provided to compute the *exact* ARL (Average Run Length) properties. Moreover, comparative studies are conducted to demonstrate the higher sensitivity of the proposed chart in the detection of shifts with various magnitudes. Advices on how to select the appropriate chart parameters are provided and an illustrative numerical example is proposed.

Keywords: Adaptive EWMA chart, Poisson Count Data, Markov Chain

1. INTRODUCTION

Control charts have become a key means of ensuring the quality of processes in the field of SPM (Statistical Process Monitoring). When we deal with different kinds of processes, data can usually be classified as either continuous or discrete. The corresponding control charts are called as variable or attribute charts, respectively.

*Corresponding author
Email address: hxl0419@njupt.edu.cn (Xuelong Hu)

A variable control chart is commonly used for the monitoring of quality characteristics which are measured on a continuous scale. For example, the Shewhart $\bar{X}$ and the R/S charts are widely employed to detect shifts in the process mean and variance, respectively. More advanced control charts, like the EWMA and the CUSUM (Cumulative Sum) charts, that combine the information from the beginning up to the current state of the process are expected to be more sensitive to small shifts than Shewhart-type charts. Specifically, Capizzi and Masarotto [5] have proposed an adaptive EWMA (AEWMA) chart in which the smoothing parameter is adjusted based on the difference between the last AEWMA statistic and the new observation. This adaptive strategy has proven to be effective in providing additional protection for small and large shifts simultaneously as well as in compensating the negative effect of inertia, see Woodall and Mahmoud [28], Reynolds Jr and Stoumbos [12]. Due to the remarkable adaptability of this AEWMA scheme, it has been investigated by several researchers, see, for instance, Shu [17], Su et al. [19], Aly et al. [2], Zaman et al. [31], Haq [9], Mitra et al. [10], Tang et al. [20, 21, 22].

On the other hand, in several applications, the quality characteristics is not a continuous variable. Such data are widely prevalent in the manufacturing industry, the public health surveillance and the network trafficking, see, for example, Woodall [29], Weiß [25], Yu et al. [30]. In that case, inspection by attribute is used due to its simple implementation and low testing cost. Considering the EWMA-type scheme for monitoring Poisson processes, Gan [7] presented a modified EWMA scheme, which rounds the EWMA statistic to a whole integer-value at each step. Furthermore, Borror et al. [3] proposed a new Poisson EWMA (PEWMA) chart having a superior ARL performance than the Shewhart c and the Gan's EWMA charts. For the detection of shifts on a specific direction, Shu et al. [18] investigated a one-sided Poisson EWMA chart which focuses on the detection of an increasing rate. An extensive amount of researches has also been performed on Poisson processes with time varying sampling rate, see Ryan and Woodall [13], Zhou et al. [32], Shen et al. [16]. A detailed state-of-the-art con-

cerning control charts for attribute data is included in Woodall [27], Topalidou and Psarakis [24], Saghir and Lin [15].

It has to be noted that, the use of the classical EWMA scheme (with a real-valued smoothing parameter λ) or the classical AEWMA scheme (with a real-valued score function φ) for Poisson count data (defined in an integer-valued domain) will lead to statistics that are no longer integer-valued. In this case, the calculation of the ARL properties is not reliable since it is not converging to a stable value based neither on pure simulation techniques nor on a Markov chain method. If the former approach is clearly dependent on the number of simulation cycles generated, the second one, unfortunately, is also heavily affected by the level of discretization of the control limit interval. Weiß [26], Rakitzis et al. [11] and Castagliola et al. [6] have shown that, the ARL results of the Markov chain method strongly fluctuate because the discrete nature of this statistic depends on the number of states selected. For this reason, Rakitzis et al. [11] proposed a discrete EWMA-type chart with exact run length properties, denoted as the CEWMA chart, for monitoring Poisson count data. This CEWMA chart uses two integer-valued parameters (γ_X, γ_Z), instead of the classical real smoothing parameter λ, to allocate the weight of past observations.

The main motivation of this work is that, there are few adaptive EWMA-type attribute charts for monitoring discrete data (compared to adaptive EWMA-type variable charts). In this paper, we study the one- and two-sided statistical performance of a discrete AEWMA chart for monitoring Poisson count data. An appropriate Markov chain technique is also developed to guarantee *exact* run length properties for this scheme. The rest of the paper is organized as follows. In Section 2, the CEWMA chart is reviewed and our proposed CAEWMA chart is introduced. Section 3 introduces an appropriate discrete-state Markov chain methodology and Section 4 presents a detailed procedure for the optimal design of the CAEWMA scheme. Section 5 evaluates the performance of the CEWMA and the CAEWMA charts for detecting a wide range of shifts in the process

mean. A numerical example is presented in Section 6 to illustrate the application of this new chart. Concluding remarks are drawn in Section 7.

2. THE CAEWMA CONTROL CHART

Suppose that $\{X_t, t = 1, 2, \ldots\}$ are the number of non-conformities in a unit which are assumed to follow a Poisson distribution with mean θ. By definition, the p.m.f. (probability mass function) of X_t is equal to

$$f_p(x|\theta) = \frac{e^{-\theta}\theta^x}{x!}, x = 0, 1, 2, \ldots. \tag{1}$$

The process is said to be in-control if $\theta = \theta_0$ and it is out-of-control if $\theta = \theta_1 \neq \theta_0$. When $\theta_1 > \theta_0$, it corresponds to an upward shift in the process mean. Of course, we need to detect this kind of shift as soon as possible and make adjustments to return the system to its normal state. On the other hand, a downward shift ($0 < \theta_1 < \theta_0$) reflects the decrease in the number of non-conformities. Modern SPM methods also emphasize the monitoring of this kind of shift as soon as possible since any beneficial preventive interventions and strategies to the process may be worth of interest.

The classical continuous EWMA chart is defined as $Z_t = (1-\lambda)Z_{t-1} + \lambda X_t$, where $\lambda \in (0, 1]$ is a real-valued smoothing parameter. However, note that, when applying this EWMA chart to monitor discrete count data, the ARL properties, obtained based on the Markov Chain approach presented in Brook and Evans [4], are known to not strictly monotonically converge with the increase in the number of Markov chain states. The same fact has also been emphasized in Weiß [26]. At this point, Rakitzis et al. [11] proposed a CEWMA monitoring technique for count data, in which not only the observations but also the smoothing constants (γ_X, γ_Z) and the EWMA statistic Z_t are all integers. The CEWMA statistics are defined as follows:

$$Z_t = \left\lfloor \frac{\gamma_X X_t + \gamma_Z Z_{t-1} + R_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \tag{2}$$

$$R_t = \gamma_X X_t + \gamma_Z Z_{t-1} + R_{t-1} - (\gamma_X + \gamma_Z) Z_t, \tag{3}$$

where γ_X and γ_Z are two positive integer-valued parameters, $\lfloor \ldots \rfloor$ denotes a rounded down integer (e.g. $\lfloor 1.5 \rfloor = 1$). By definition, Z_t corresponds to the quotient of the Euclidean division and $R_t \in \{0, 1, \ldots, \gamma_X + \gamma_Z - 1\}$ is the remainder of this division. Please note that these $R_1, R_2, \ldots$ values are only recorded but will not be used for monitoring. The monitored statistics are only the values $Z_1, Z_2, \ldots$, and (2) and (3) can be rewritten as

$$(\gamma_X + \gamma_Z)Z_t + R_t = \gamma_X X_t + \gamma_Z Z_{t-1} + R_{t-1}. \tag{4}$$

This CEWMA chart is optimally designed to detect a *specific* shift magnitude by adjusting the relative value of γ_X and γ_Z. More weight can be given to past observations when the ratio $\gamma_X/(\gamma_X + \gamma_Z)$ gets smaller, thus allowing a better sensitivity to the detection of small shifts. Otherwise, large values of $\gamma_X/(\gamma_X + \gamma_Z)$ must be used to quickly detect large shifts. However, it is quite difficult to pre-determine the exact magnitude of shifts occurring in practice. A natural option is to find a potential tendency presented in the data and then to dynamically adjust the weight coefficient according to the predicted difference. Therefore, we suggest the following adaptive EWMA (denoted as CAEWMA) scheme proposed in Tang et al. [23] for count data:

$$Z_t = \left\lfloor \frac{\varphi(X_t - Z_{t-1}) + (\gamma_X + \gamma_Z)Z_{t-1} + R_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \tag{5}$$

$$R_t = \varphi(X_t - Z_{t-1}) + (\gamma_X + \gamma_Z)Z_{t-1} + R_{t-1} - (\gamma_X + \gamma_Z)Z_t, \tag{6}$$

where $\varphi(e_t) = \varphi(X_t - Z_{t-1})$ is a score function. If we define $C_t \stackrel{\text{def}}{=} (\gamma_X + \gamma_Z)Z_t + R_t$ then (5) and (6) can be rewritten as

$$\underbrace{(\gamma_X + \gamma_Z)Z_t + R_t}_{C_t} = \varphi(e_t) + \underbrace{(\gamma_X + \gamma_Z)Z_{t-1} + R_{t-1}}_{C_{t-1}}, \tag{7}$$

where the initial values for Z_t and R_t are fixed as $Z_0 = z_0$ and $R_0 = r_0$. So, at time t, when the X_t, Z_{t-1}, R_{t-1} values as well as the score function $\varphi(e_t)$ are fixed, we can uniquely get the values of Z_t and R_t. It goes without saying that the error terms e_t between X_t and Z_{t-1} have to be also integers, so a new score

function is defined as

$$\varphi(e) = \begin{cases} e(\gamma_X + \gamma_Z) + \gamma_Z k & \text{if } e < -k \\ e\gamma_X & \text{if } |e| \leq k \\ e(\gamma_X + \gamma_Z) - \gamma_Z k & \text{if } e > k \end{cases}, \tag{8}$$

where k is also a positive integer-valued parameter. As mentioned in Capizzi and Masarotto [5], when we construct the score function, the following conditions need to be met:

i) $\varphi(e)$ must be monotone increasing;

ii) if $k \to \infty$ and we have $\varphi(e) = e\gamma_X$ and (7) becomes $(\gamma_X + \gamma_Z)Z_t + R_t = \gamma_X X_t + \gamma_Z Z_{t-1} + R_{t-1}$, which agrees with the CEWMA scheme in Rakitzis et al. [11];

iii) if $k = 0$, we have $\varphi(e) = e(\gamma_X + \gamma_Z)$ and (7) reduces to a Shewhart-type scheme.

The process is said to be out-of-control whenever $Z_t < h_L$ or $Z_t > h_U$, where $h_L \in \{1, 2, \dots\}$ and $h_U \in \{h_L, h_L + 1, \dots\}$ are the integer-valued lower and upper control limits, respectively.

3. RUN LENGTH PROPERTIES

The RL (Run length) properties are the most widely used indicator for evaluating control charts. It represents the number of sample points drawn before the control chart signals, i.e. RL $= \inf\{t \geq 1 | Z_t < h_L \text{ or } Z_t > h_U\}$. In this paper, we proposed a discrete-state Markov chain method to exactly calculate the RL properties of the CAEWMA chart. At time $t-1$, the set of Markov states are defined as $c_{t-1} = (\gamma_X + \gamma_Z)z_{t-1} + r_{t-1}$. As $z_{t-1} \in \{h_L, h_L + 1, \dots, h_U\}$ and $r_{t-1} \in \{0, 1, \dots, \gamma_X + \gamma_Z - 1\}$, the minimum and the maximum values of c_{t-1} are

$$c_{\min} = (\gamma_X + \gamma_Z)h_L, \tag{9}$$

$$c_{\max} = (\gamma_X + \gamma_Z)(h_U + 1) - 1, \tag{10}$$

respectively. So, at time $t-1$, all the discrete states are $c_{t-1} \in \{c_{\min}, c_{\min}+1, \ldots, c_{\max}\}$, and the total number of states is $m+1 = c_{\max} - c_{\min} + 1$. When the state c_{t-1} moves to the state $c_t = (\gamma_X + \gamma_Z)z_t + r_t$, the score $\varphi(e_t)$ needs to satisfy $c_t = \varphi(e_t) + c_{t-1} \in \{c_{\min}, c_{\min}+1, \ldots, c_{\max}\}$. Since $\varphi(e_t)$ is monotone increasing in e_t, the allowable values of $x_t \geq 0$ must simultaneously satisfy

$$\begin{cases} (\gamma_X + \gamma_Z)(x_t - \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor) + k\gamma_Z + c_{t-1} \geq c_{\min} & \text{if } e_t < -k, \\ \gamma_X(x_t - \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor) + c_{t-1} \geq c_{\min} & \text{if } |e| \leq k, \end{cases} \tag{11}$$

and

$$\begin{cases} (\gamma_X + \gamma_Z)(x_t - \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor) - k\gamma_Z + c_{t-1} \leq c_{\max} & \text{if } e_t > k, \\ \gamma_X(x_t - \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor) + c_{t-1} \leq c_{\max} & \text{if } |e| \leq k, \end{cases} \tag{12}$$

from which we obtain the minimum $x_{\min}$ and the maximum $x_{\max}$ values as

$$x_{\min} = \max\left(0, \left\lceil \frac{c_{\min} - c_{t-1} - k\gamma_Z}{\gamma_X + \gamma_Z} \right\rceil + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \left\lceil \frac{c_{\min} - c_{t-1}}{\gamma_X} \right\rceil + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor\right), \tag{13}$$

$$x_{\max} = \min\left(\left\lfloor \frac{c_{\max} - c_{t-1} + k\gamma_Z}{\gamma_X + \gamma_Z} \right\rfloor + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \left\lfloor \frac{c_{\max} - c_{t-1}}{\gamma_X} \right\rfloor + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor\right), \tag{14}$$

where $\lceil \ldots \rceil$ denotes a rounded up integer.

Now, the transition probabilities $q_{i,j}$ of going from one state to another can be calculated using the following steps:

Given n, h_L, h_U, γ_X, γ_Z, k

$c_{\min} \leftarrow (\gamma_X + \gamma_Z)h_L$

$c_{\max} \leftarrow (\gamma_X + \gamma_Z)(h_U + 1) - 1$

$m \leftarrow c_{\max} - c_{\min}$

For $i = 0, 1, \ldots, m$

$c_{t-1} \leftarrow i + c_{\min}$

$z_{t-1} \leftarrow \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor$

$x_{\min} \leftarrow \max\left(0, \left\lceil \frac{c_{\min} - c_{t-1} - k\gamma_Z}{\gamma_X + \gamma_Z} \right\rceil + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \left\lceil \frac{c_{\min} - c_{t-1}}{\gamma_X} \right\rceil + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor\right)$

$x_{\max} \leftarrow \min\left(\left\lfloor \frac{c_{\max} - c_{t-1} + k\gamma_Z}{\gamma_X + \gamma_Z} \right\rfloor + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor, \left\lfloor \frac{c_{\max} - c_{t-1}}{\gamma_X} \right\rfloor + \left\lfloor \frac{c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor\right)$

For $x_t = x_{\min}, x_{\min} + 1, \ldots, x_{\max}$

$e_t \leftarrow x_t - z_{t-1}$

If $e_t < -k$

$\varphi(e_t) \leftarrow (\gamma_X + \gamma_Z)e_t + k\gamma_Z$

Else if $e_t > k$

$\varphi(e_t) \leftarrow (\gamma_X + \gamma_Z)e_t - k\gamma_Z$

Else $-k \leq e_t \leq k$

$\varphi(e_t) \leftarrow \gamma_X e_t$

End If

$z_t \leftarrow \left\lfloor \frac{\varphi(e_t) + c_{t-1}}{\gamma_X + \gamma_Z} \right\rfloor$

$r_t \leftarrow \varphi(e_t) + c_{t-1} - (\gamma_X + \gamma_Z)z_t$

$j \leftarrow (\gamma_X + \gamma_Z)z_t + r_t - c_{\min}$

$q_{i,j} \leftarrow f_p(x_t|\theta)$

End For

End For

Let $\mathbf{Q}$ be the $(m+1, m+1)$ transition matrix containing all of the probabilities $q_{i,j}$, and let $\mathbf{s}$ be the initial probability vector represented by a $(m+1, 1)$ vector $\mathbf{s} = (s_0, s_1, \ldots, s_m)$, where $\mathbf{s}$ contains the probabilities that the Markov chain starts in a given state. We will set 1 in the entry corresponding to $(\gamma_X + \gamma_Z)Z_0 + r_0$ and zeros in the remaining ones. For example, If $z_0 = \lfloor \theta_0 \rfloor$ and $r_0 = 0$, the entry $s_{(\gamma_X + \gamma_Z)\lfloor \theta_0 \rfloor + r_0} = 1$ and $s_i = 0$ for $i \neq (\gamma_X + \gamma_Z)\lfloor \theta_0 \rfloor + r_0$.

The mean value $\mathrm{ARL} = E(\mathrm{RL})$, as well as the standard deviation $\mathrm{SDRL} = \sigma(\mathrm{RL})$ can be easily obtained as

$$\mathrm{ARL} = \mathbf{s}^{\intercal}(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}, \tag{15}$$

$$\mathrm{SDRL} = \sqrt{2\mathbf{s}^{\intercal}(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{Q}\mathbf{1} - \mathrm{ARL}^2 + \mathrm{ARL}}. \tag{16}$$

where $\mathbf{I}$ is the $(m+1, m+1)$ identity matrix and $\mathbf{1}$ is a $(m+1, 1)$ vector of 1's.

4. DESIGN OF THE CAEWMA SCHEME

When there is an objective of detecting a shift in a particular direction (either upward or downward), one-sided schemes could be straightforwardly adapted from the two-sided situation. For the upper-sided chart, $h_L = 0$ is set, and it results in $c_{\min} = 0$ and $c_{\max} = m$. Analogously, for the lower-sided case, an upper restriction $h_U = A$ will be picked to prevent the statistic Z_t from shifting to infinity.

Next, we provide a procedure for the design of the proposed CAEWMA scheme, in which the statistical performances are evaluated in terms of the ARL. Usually, the same ARL_0 value is taken when the process is in-control, and when the process is out-of-control, the smaller the out-of-control ARL_1 value, the better the performance of control charts. Due to the discreteness of the monitoring statistic Z_t, $\mathrm{ARL}(h_L, h_U, \gamma_X, \gamma_Z, k, \theta_0) = \mathrm{ARL}_0$ cannot be exactly attained. Therefore, to satisfy $\mathrm{ARL}(h_L, h_U, \gamma_X, \gamma_Z, k, \theta_0) \approx \mathrm{ARL}_0$, only the combinations of $(h_L, h_U, \gamma_X, \gamma_Z, k, \theta_0)$ with a ARL value close to the design one, i.e. $\left|\frac{\mathrm{ARL}(h_L, h_U, \gamma_X, \gamma_Z, k, \theta_0) - \mathrm{ARL}_0}{\mathrm{ARL}_0}\right| \leq \zeta$, will be considered, where ζ is a pre-determined threshold.

- For the upper-sided CAEWMA chart, we suggest to find parameters $(h_U, \gamma_X, \gamma_Z, k)$

giving an optimal performance for detecting a range of shifts:

$$\text{Min EARL}(h_U, \gamma_X, \gamma_Z, k, \Omega), \tag{17}$$

$$\text{s.t. ARL}(h_U, \gamma_X, \gamma_Z, k, \theta_0) \approx \text{ARL}_0, \tag{18}$$

where

$$\text{EARL} = \int_{\Omega} f(\theta_1)\text{ARL}(h_U, \gamma_X, \gamma_Z, k, \theta_1)\mathrm{d}\theta_1, \tag{19}$$

and Ω is assumed to be a uniform distribution for the mean shift $U[\theta_{1\min}, \theta_{1\max}]$ with $f(\theta_1) = \frac{1}{\theta_{1\max} - \theta_{1\min}}$.

- For the two-sided CAEWMA chart, one important problem that must be taken into account is that the ARL_0 value is not always larger than the corresponding ARL_1 values. Inspired by Su et al. [19], a two-steps searching procedure is recommended:

  1. For a desired ARL_0 value, find the optimal combination of (γ_X, γ_Z) for the CEWMA chart, which gives a minimum ARL_1 value at $\theta_{1\min}$. Some of these combinations could be found in Rakitzis et al. [11].
  2. Based on the same (γ_X, γ_Z) value obtained in Step 2, for the CAEWMA chart, find an optimal k value having the minimum ARL_1 value at $\theta_{1\max}$ subject to $\text{ARL}(h_L, h_U, \gamma_X, \gamma_Z, k, \theta_0) \approx \text{ARL}_0$.

5. PERFORMANCE COMPARISONS

Since it is quite difficult to predict the magnitude and direction of shifts occurring in practice, Han and Tsung [8] recommended using the RMI (Relative Mean Index) to evaluate the global performance of control charts, which is defined as follows:

$$\text{RMI}(I) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\text{ARL}(I, \theta_1^{(i)}) - \text{ARL}^*(\theta_1^{(i)})}{\text{ARL}^*(\theta_1^{(i)})}\right), \tag{20}$$

where $\text{ARL}(I, \delta^{(i)})$ is the ARL value for a particular chart I for the shift $\theta_1(i)$, and among these comparative charts $I = 1, 2, \ldots$, $\text{ARL}^*(\theta_1^{(i)})$ denotes the the

minimum ARL value when detecting $\theta_1^{(i)}$. Here, a discrete uniform distribution for the shift θ_1 is assumed, for example, when $\theta_0 = 8$, consider upward shifts $\theta_1 \in \{9, 10, \ldots, 20\}$, $\theta_1^{(i)}$ denotes the i^{th} shift, i.e. $\theta_1^{(1)} = 9, \theta_1^{(2)} = 10, \ldots, \theta_1^{(12)} = 20$), and the total number of shifts in this example is $N = 12$. Obviously, the smaller the value of RMI(I), the better the performance of the proposed chart.

We have compared the CAEWMA chart with the CEWMA chart proposed in Rakitzis et al. [11]. From Tables 1 to 6, we can draw the following conclusions:

- It can be seen that, when designing the CEWMA charts, a small 'smoothing ratio' $\gamma_X/(\gamma_X + \gamma_Z)$ value is effective for detecting small shifts, and a large one is effective for larger shifts. Therefore, for a single CEWMA control chart, it is difficult to balance the overall performance for different shift sizes, at the same time. Taking the upper-sided case as an example, in Table 1, if the CEWMA chart, with $(\gamma_X, \gamma_Z) = (1, 83)$, can be efficient in the detection of small shifts (e.g. $\theta_1 = 9$), it is unfortunately insensitive to moderate or large ones (e.g. $\theta_1 > 12$).

- For a specified in-control $\text{ARL}_0 = 1000$, Table 1 and 2 compare the performance for upper-sided shifts in the Poisson mean $\theta_0 \in \{8, 12\}$. Note that, according to Rakitzis et al. [11], the optimal combination $(\gamma_X, \gamma_Z) = (1, 83)$ of the CEWMA chart could help in detecting small shifts $\theta_1 \in \{9, 10\}$ (when $\theta_0 = 8$) faster. However, based on the results shown in Table 1, the CAEWMA chart always has the minimum ARL_1 for $\theta_1 \in \{9, 10, 11, 12\}$. The same case can also be observed in Table 2, when $\theta_0 = 12$, the CAEWMA chart has the minimum $\text{ARL}_1 = 2.1$ value for $\theta_1 = 24$ compared with the optimal $\text{ARL}_1 = 2.2$ obtained with the CEWMA chart.

- For the lower-sided case in Table 3, when $\theta_0 = 16$, note that the CEWMA chart with $(\gamma_X, \gamma_Z) = (3, 28)$ is the only optimally recommended combination in Rakitzis et al. [11] for all shift sizes. However, the comparison reveals that the CAEWMA chart has a higher detection probability for

decreasing shifts $\theta_1 \in \{15, 14, 13, 12\}$ and $\theta_1 \in \{5, 4, 3\}$ while, for shifts $\theta_1 \in \{10, 9, 8, 7\}$, the differences between these two schemes are negligible. The advantage of the CAEWMA chart in detecting different shift sizes is also fully verified in terms of the RMI performance. For example, when $\theta_0 = 20$ in Table 4, the CAEWMA chart gives the minimum RMI $= 0.206$ value compared with all CEWMA charts.

- For the two-sided case in Tables 5 and 6, it is interesting to note when similar values of γ_X and γ_Z are selected, the CAEWMA chart guards against small and large shifts more effectively than the CEWMA chart. For example, in Table 5, when compared with the CEWMA chart with $\gamma_X = 5$ and $\gamma_Z = 37$, the CAEWMA chart with $\gamma_X = 5$ and $\gamma_Z = 38$ provides a slightly worse performance for detecting increasing shifts $\theta_1 \in \{22, 24, 26\}$, but better performance for detecting decreasing shifts $\theta_1 \in \{18, 16, 14, 12\}$ and increasing shifts $\theta_1 \in \{30, 35, 40\}$. In Table 6, when $\gamma_X = 1$ and $\gamma_Z = 15$, there is no loss in the efficiency for the ARL_1 in detecting decreasing shifts $\theta_1 \in \{18, 16, \ldots, 6\}$, however, the CAEWMA chart performs better in the detection of increasing shifts $\theta_1 \in \{22, 24, \ldots, 40\}$ compared to the CEWMA chart.

To sum up, it is shown that the CAEWMA chart outperforms the CEWMA chart in terms of the out-of-control ARL_1 and RMI to efficiently detect both small and large shifts. As the practitioner has usually no clear idea about the actual magnitude of potential shifts, we recommend the use of the CAEWMA chart for practical purposes.

(Please Insert Tables 1 to 6 Here)

6. AN ILLUSTRATIVE EXAMPLE

The implementation of the proposed CAEWMA chart is illustrated with an example from the traffic flow monitoring. Based on historical data, the number X_t of vehicles passing every minute in a particular section of the road follows a

Poisson distribution with in-control mean $\theta_0 = 12$. We actually have two data sets recording the number of vehicles passing in a certain section of the road

- in the morning period $8:00-9:00$ (Table 7),
- in the afternoon period $17:00-18:00$ (Table 8).

For comparison, two upper-sided CEWMA charts are also implemented for the same date sets. The desired $\text{ARL}_0 = 1000$ value is fixed, and based on Rakitzis et al. [11], the optimal combinations $(h_U, \gamma_X, \gamma_Z) = (13, 1, 19)$ and $(16, 2, 5)$ of the CEWMA charts (denoted as CEWMA-1 and CEWMA-2, respectively) are considered here. Concerning the CAEWMA chart, the unique optimal combination is obtained to be $(h_U, \gamma_X, \gamma_Z, k) = (15, 3, 14, 12)$.

All results are presented in Figures 1 and 2. For the morning case (Table 7), the CAEWMA and CEWMA-1 charts give an out-of-control signal at the 38^{th} observation. It indicates that the traffic congestion caused by the morning peak occurred around $8:38$. While, no point exceeds the control limit in the CEWMA-2 chart, so no alarm is issued. For the afternoon case (Table 8), both the CAEWMA and CEWMA-2 charts find anomalies at the 23^{th} observation and the CEWMA-1 chart signals at the 26^{th} observation. It indicates that the traffic congestion caused by the afternoon peak occurred around $17:23$. In short, the CAEWMA chart is not only as effective as the CEWMA-1 chart when the shift size is small, but it is also as effective as the CEWMA-2 chart for large shifts.

(Please Insert Figures 1 and 2 Here)

7. CONCLUSIONS

In this work, we have extended the AEWMA scheme to the monitoring of Poisson count data. An appropriate Markov chain technique has been developed to guarantee exact ARL results for discrete statistics, without approximation. Particularly, to enhance its detection capability for asymmetric distribution,

an in-depth study for the one- and two-sided cases have been performed. The one-sided CAEWMA scheme is recommended when shifts toward a particular side are definitely more significant than toward the opposite one. The two-sided CAEWMA scheme should be used when shifts toward either side are equally important. The comparative results support the fact that the CAEWMA chart is able to provide a more balanced detection property against shifts of different magnitudes than the CEWMA chart in both one- and two-sided conditions. Therefore, this scheme is worth to be considered when the exact size of the shift cannot be pre-determined.

References

[1] Aly AA, Hamed RM and Mahmoud MA. 2017. "Optimal design of the adaptive exponentially weighted moving average control chart over a range of mean shifts." *Communications in Statistics-Simulation and Computation* 46 (2):890–902.

[2] Aly AA, Saleh NA, Mahmoud MA and Woodall WH. 2015. "A reevaluation of the adaptive exponentially weighted moving average control chart when parameters are estimated." *Quality and Reliability Engineering International* 31(8):1611–1622.

[3] Borror CM, Champ CW and Rigdon SE. 1998. "Poisson EWMA control charts." *Journal of Quality Technology* 30(4):352–361.

[4] Brook DAED and Evans DA. 1972. "An approach to the probability distribution of CUSUM run length." *Biometrika* 59(3):539–549.

[5] Capizzi G and Masarotto G. 2003. "An adaptive exponentially weighted moving average control chart." *Technometrics* 45(3):199–207.

[6] Castagliola P, Tran KP, Celano G, Rakitzis AC and Maravelakis PE. 2019. "An EWMA-Type Sign Chart with Exact Run Length Properties." *Journal of Quality Technology* 51(1):51–63.

[7] Gan FF. 1990b. "Monitoring Poisson observations using modified exponentially weighted moving average control charts." *Communications in Statistics-Simulation and Computation* 19(1):103–124.

[8] Han D and Tsung F. 2006. "A reference-free cuscore chart for dynamic mean change detection and a unified framework for charting performance comparison." *Journal of the American Statistical Association* 101(473):368–386.

[9] Haq A. 2018. "A New Adaptive EWMA Control Chart for Monitoring the Process Dispersion." *Quality and Reliability Engineering International* 34(5): 846-C857.

[10] Mitra A, Kang BL and Subhabrata C. 2019. "An Adaptive Exponentially Weighted Moving Average-Type Control Chart to Monitor the Process Mean." *European Journal of Operational Research* 279(3):902–911.

[11] Rakitzis AC, Castagliola P and Maravelakis PE. 2015. "A new memory-type monitoring technique for count data." *Computers & Industrial Engineering* 85:235–247.

[12] Reynolds Jr MR and Stoumbos ZG. 2006. "Comparisons of some exponentially weighted moving average control charts for monitoring the process mean and variance." *Technometrics* 48(4):550–567.

[13] Ryan AG and Woodall WH. 2010. "Control charts for Poisson count data with varying sample sizes." *Journal of Quality Technology* 42(3):260–275.

[14] Ryu JH, Wan G and Kim S. 2010. "Optimal design of a CUSUM chart for a mean shift of unknown size." *Journal of Quality Technology* 42(3):311–326.

[15] Saghir A and Lin Z. 2015. "Control charts for dispersed count data: an overview." *Quality and Reliability Engineering International* 31(5):725–739.

[16] Shen X, Zou C, Jiang W and Tsung F. 2013. "Monitoring Poisson count data with probability control limits when sample sizes are time varying." *Naval Research Logistics (NRL)* 60(8):625–636.

[17] Shu L. 2008. "An adaptive exponentially weighted moving average control chart for monitoring process variances." *Journal of Statistical Computation and Simulation* 78(4):367–384.

[18] Shu L, Jiang W and Wu Z. 2012. "Exponentially weighted moving average control charts for monitoring increases in Poisson rate." *IIE Transactions* 44 (9):711–723.

[19] Su Y, Shu L and Tsui KL. 2011. "Adaptive EWMA procedures for monitoring processes subject to linear drifts." *Computational Statistics & Data Analysis* 55(10):2819–2829.

[20] Tang A, Castagliola P, Sun J and Hu X. 2017. "An adaptive exponentially weighted moving average chart for the mean with variable sampling intervals." *Quality and Reliability Engineering International* 33(8):2023–2034.

[21] Tang A, Castagliola P, Sun J and Hu X. 2018. "The effect of measurement errors on the adaptive EWMA $\bar{X}$ chart." *Quality and Reliability Engineering International* 34(4):609–630.

[22] 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." *Quality Technology & Quantitative Management* 16(4): 439-458.

[23] Tang A, Sun J, Hu X and Castagliola P. 2019. "A new nonparametric adaptive EWMA control chart with exact run length properties." *Computers & Industrial Engineering*, 130:404–419.

[24] Topalidou E and Psarakis S. 2009. "Review of multinomial and multiattribute quality control charts." *Quality and Reliability Engineering International* 25(7):773–804.

[25] Weiß CH. 2007. "Controlling correlated processes of Poisson counts." *Quality and reliability engineering international* 23(6):741–754.

[26] Weiß CH. 2009. "EWMA monitoring of correlated processes of Poisson counts." *Quality Technology & Quantitative Management* 6(2):137–153.

[27] Woodall WH, 1997. "Control charts based on attribute data: bibliography and review." *Journal of quality technology* 29(2):172–183.

[28] Woodall WH and Mahmoud MA. 2005. "The inertial properties of quality control charts." *Technometrics* 47(4):425–436.

[29] Woodall WH, 2006. "The use of control charts in health-care and public-health surveillance." *Journal of quality technology* 38(2):89–104.

[30] Yu FJ, Yang YY., Wang MJ and Wu Z. 2011. "Using EWMA control schemes for monitoring wafer quality in negative binomial process." *Microelectronics Reliability* 51(2):400–405.

[31] Zaman B, Muhammad HL, Muhammad Riaz and Muazu RA. 2017. "An Adaptive EWMA Scheme-Based CUSUM Accumulation Error for Efficient Monitoring of Process Location" *Quality and Reliability Engineering International* 33(8):2463–2482.

[32] Zhou Q, Zou C, Wang Z and Jiang W. 2012. "Likelihood-based EWMA charts for monitoring Poisson count data with time-varying sample sizes." *Journal of the American Statistical Association* 107(499):1049–1062.

Table 1: Comparisons of the CEWMA and CAEWMA charts, upper-sided case, $\text{ARL}_0 = 1000$ and $\theta_0 = 8$

θ	CEWMA			CAEWMA
8	995.1	987.8	1011.6	1008.8
9	83.7	125.7	154.5	75.8
10	42.7	32.0	40.8	21.4
11	28.7	13.4	16.3	11.1
12	21.7	7.7	8.7	7.4
13	17.4	5.2	5.6	5.6
14	14.6	4.0	4.1	4.5
15	12.6	3.2	3.2	3.7
16	11.1	2.7	2.7	3.2
17	9.9	2.4	2.3	2.8
18	8.9	2.1	2.0	2.5
19	8.2	1.9	1.8	2.2
20	7.5	1.8	1.7	1.9
21	7.0	1.6	1.5	1.8
RMI	2.542	0.129	0.207	0.119
h_L	0	0	0	0
h_U	8	12	13	10
γ_X	1	4	6	8
γ_Y	83	7	7	43
k				13

Table 2: Comparisons of the CEWMA and CAEWMA charts, upper-sided case, $\text{ARL}_0 = 1000$ and $\theta_0 = 12$

θ	CEWMA			CAEWMA
12	1016.0	1032.0	1008.3	1009.3
13	74.5	92.0	151.5	135.0
14	28.1	27.7	41.2	35.8
15	17.0	14.7	17.6	16.3
16	12.2	9.9	10.0	9.9
17	9.6	7.4	6.8	7.0
18	7.9	6.0	5.1	5.4
19	6.7	5.1	4.1	4.4
20	5.9	4.4	3.5	3.7
21	5.2	3.9	3.0	3.1
22	4.7	3.5	2.7	2.7
23	4.3	3.2	2.4	2.4
24	4.0	2.9	2.2	2.1
25	3.7	2.7	2.0	1.9
RMI	0.523	0.210	0.141	0.113
h_L	0	0	0	0
h_U	13	14	16	15
γ_X	1	1	2	3
γ_Y	19	8	5	14
k				12

Table 3: Comparisons of the CEWMA and CAEWMA charts, lower-sided case, $\text{ARL}_0 = 1000$ and $\theta_0 = 16$

θ	CEWMA			CAEWMA
16	879.0	1015.0	934.8	1010.2
15	131.4	100.2	66.6	68.6
14	35.3	27.6	21.6	21.8
13	15.1	13.4	12.1	12.1
12	8.7	8.6	8.3	8.3
11	6.0	6.3	6.4	6.3
10	4.5	5.0	5.2	5.1
9	3.7	4.1	4.4	4.3
8	3.1	3.5	3.8	3.7
7	2.7	3.1	3.4	3.2
6	2.4	2.8	3.1	2.8
5	2.2	2.5	2.8	2.4
4	2.0	2.2	2.6	2.0
3	2.0	2.1	2.3	1.5
RMI	0.172	0.175	0.177	0.079
h_L	13	14	15	15
h_U	30	30	30	30
γ_X	3	3	3	5
γ_Y	13	28	68	114
k				12

Table 4: Comparisons of the CEWMA and CAEWMA charts, lower-sided case, $\text{ARL}_0 = 1000$ and $\theta_0 = 20$

θ	CEWMA			CAEWMA
20	993.9	1207.2	1005.1	1002.8
19	285.4	123.1	78.2	79.1
18	95.1	34.3	26.0	26.1
17	38.0	16.7	14.6	14.6
16	18.1	10.7	10.0	10.0
15	10.2	7.8	7.7	7.7
14	6.6	6.1	6.2	6.2
13	4.7	5.1	5.2	5.2
12	3.6	4.4	4.5	4.5
11	2.9	3.8	4.0	4.0
10	2.5	3.4	3.6	3.5
9	2.2	3.1	3.3	3.2
8	1.9	2.9	3.1	2.9
7	1.7	2.6	2.9	2.6
RMI	0.625	0.274	0.233	0.206
h_L	14	18	19	19
h_U	30	30	30	30
γ_X	3	3	2	1
γ_Y	4	35	55	27
k				15

Table 5: Comparisons of the CEWMA and CAEWMA charts, two-sided case, $\text{ARL}_0 = 1000$ and $\theta_0 = 20$

θ	CEWMA			CAEWMA
6	2.4	2.0	1.9	2.4
8	2.9	2.2	2.1	2.9
10	3.6	2.8	2.6	3.6
12	4.7	3.9	3.6	4.6
14	6.9	6.6	5.9	6.8
16	13.5	16.9	14.0	13.2
18	62.4	100.8	70.9	58.2
20	1017.2	1000.3	982.8	1000.0
22	50.2	110.4	335.9	53.2
24	14.7	21.1	42.2	15.2
26	8.3	8.8	13.3	8.4
28	5.8	5.4	7.0	5.8
30	4.5	3.9	4.7	4.3
35	3.0	2.4	2.7	2.5
40	2.3	1.8	2.0	1.9
RMI	0.168	0.220	0.656	0.131
h_L	17	15	15	17
h_U	23	26	27	23
γ_X	5	7	3	5
γ_Y	37	17	7	38
k				17

Table 6: Comparisons of the CEWMA and CAEWMA charts, two-sided case, $\text{ARL}_0 = 1500$ and $\theta_0 = 20$

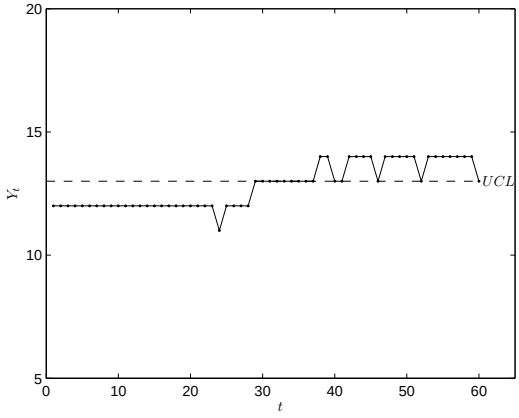
θ	CEWMA			CAEWMA
6	3.0	2.1	2.2	3.0
8	3.5	2.4	2.7	3.5
10	4.2	3.0	3.4	4.2
12	5.4	4.1	4.7	5.4
14	7.7	6.3	8.1	7.7
16	13.5	13.7	22.4	13.5
18	47.0	70.7	191.8	47.0
20	1520.4	1495.6	1517.1	1504.0
22	50.1	142.9	85.0	50.0
24	17.2	23.2	17.9	17.2
26	10.4	10.0	8.3	10.3
28	7.5	6.3	5.3	7.4
30	5.9	4.6	3.9	5.7
35	3.9	2.9	2.5	3.5
40	3.6	2.2	1.9	2.2
RMI	0.319	0.258	0.373	0.249
h_L	18	16	15	18
h_U	22	25	25	22
γ_X	1	6	7	1
γ_Y	15	25	23	15
k				21

Table 7: Data sets for the morning period $8:00-9:00$

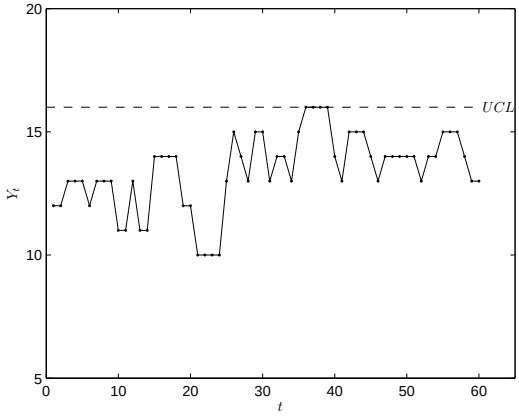
t	X_t	t	X_t	t	X_t
1	13	21	2	41	10
2	15	22	11	42	19
3	15	23	10	43	16
4	12	24	10	44	14
5	12	25	20	45	14
6	9	26	22	46	8
7	16	27	10	47	19
8	15	28	12	48	14
9	11	29	19	49	13
10	8	30	15	50	15
11	10	31	9	51	13
12	17	32	14	52	11
13	6	33	15	53	15
14	12	34	10	54	15
15	20	35	22	55	16
16	17	36	17	56	15
17	14	37	16	57	15
18	11	38	17	58	12
19	8	39	15	59	12
20	14	40	10	60	11

Table 8: Data sets for the afternoon period $17:00-18:00$

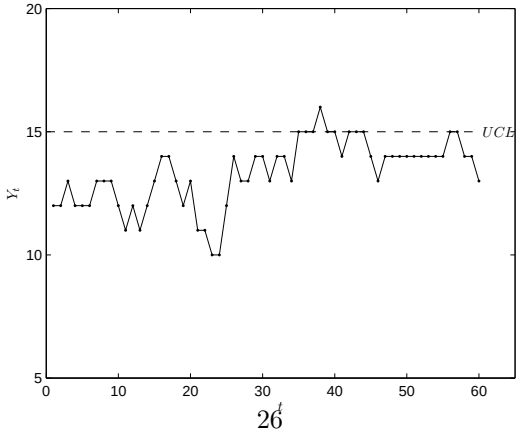
t	X_t	t	X_t	t	X_t
1	17	21	14	41	19
2	7	22	19	42	20
3	10	23	26	43	20
4	10	24	18	44	26
5	10	25	21	45	20
6	12	26	28	46	22
7	16	27	31	47	23
8	10	28	20	48	21
9	16	29	16	49	24
10	6	30	18	50	25
11	15	31	12	51	16
12	5	32	20	52	25
13	14	33	21	53	21
14	13	34	11	54	16
15	13	35	24	55	20
16	16	36	16	56	23
17	8	37	25	57	22
18	14	38	17	58	18
19	13	39	19	59	23
20	7	40	23	60	24



(a) CEWMA-1

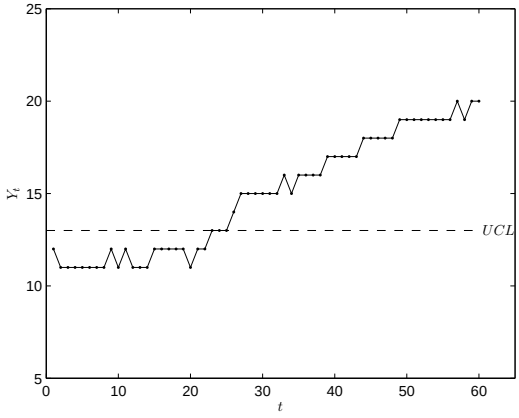


(b) CEWMA-2

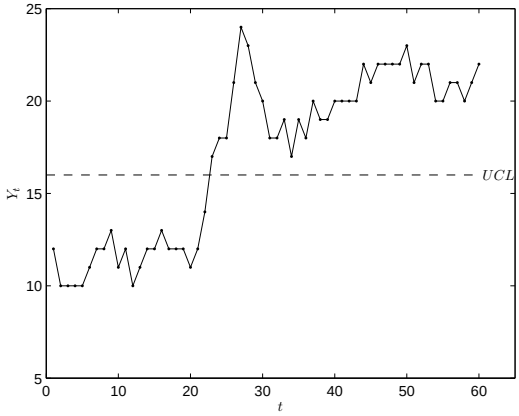


(c) CAEWMA

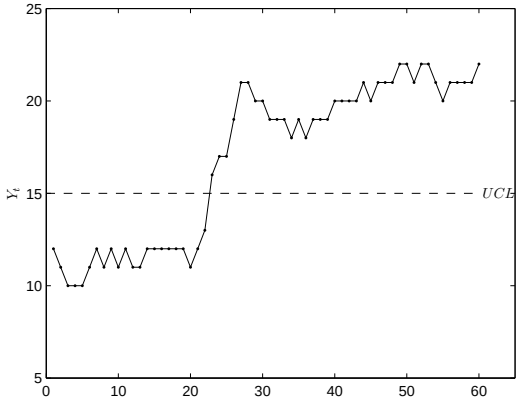
Figure 1: Control charts applied to the morning period $8:00-9:00$



(a) CEWMA-1



(b) CEWMA-2



(c) CAEWMA

Figure 2: Control charts applied to the afternoon period $17:00-18:00$