Computers & Industrial Engineering 154 (2021) 107100

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Optimal monitoring of Poisson data with known and unknown shifts

Junjie Wang [a,*], Zhi Lin Chong [b], Peihua Qiu [c]

[a] School of Business Administration, Zhongnan University of Economics and Law, Wuhan, China
[b] School of Mathematical Sciences, Universiti Sains Malaysia (USM), Penang, Malaysia
[c] Department of Biostatistics, College of Public Health & Health Professions, University of Florida, Gainesville, FL, United States

ARTICLE INFO

Keywords:
Count data
Expected average run length
Fibonacci search algorithm
Poisson EWMA chart
Statistical process control

ABSTRACT

The number of event occurrences, called counts are prevalent in many fields such as manufacturing industry and public health. Control charts have been widely employed to monitor such count data for quality improvement of products or medical service by assuming the data follows the Poisson distribution. However, the shift information of Poisson mean has not been well considered in current design of the exponentially weighted moving average (EWMA) control chart. This article studies the optimal design of the Poisson EWMA chart with known and unknown shift sizes integrated respectively in order to bridge the research gap. We simplify these two optimization problems to searching for a unique smoothing parameter in minimizing the out-of-control (OC) average run length (ARL) and OC expected ARL (EARL) over random shifts respectively. Due to the intractability of obtaining a closed-form solution, the Fibonacci search algorithm is proposed to find out the optimal smoothing parameter in a short time. The satisfactory performance of proposed optimal design method is demonstrated by numerous simulation results and two real datasets from manufacturing industry and public health.

1. Introduction

In recent years, the number of event occurrences has attracted plenty of attention in various fields. For example, recording the monthly number of patients suffering from polio disease can help forecast and prevent potential epidemic situations in public health (Wang & Qiu, 2018) whereas nonconforming products are counted to reflect the variation of a manufacturing process (Li, Wang, & Zhu, 2019). Statistical process control (SPC) has been widely employed to analyze such count data for either public health surveillance or quality improvement. In SPC applications, it is usual to assume that these count data follow a Poisson distribution, allowing for the establishments of reasonable benchmarks such as control limits to determine whether an anomaly occurs given a specific magnitude of Type I error. Due to its widespread applications, monitoring Poisson counts has become an important research field, to which, this article is trying to make a contribution.

In the literature, an early representative method should be c-chart, the detailed introduction of which can be found in Montgomery (2013) and Qiu (2014). Since only current observations are used, c-chart is of Shewhart type and is more suitable for detecting large shifts. In order to detect moderate and small shifts effectively, Lucas (1985) and Gan (1990) proposed a cumulative sum (CUSUM) chart and an EWMA chart respectively, both of which can combine current and past sample information. Borror, Champ, and Rigdon (1998) calculated the ARL of the Poisson EWMA chart by Markov chain approximation. Zhang, Govindaraju, Lai, and Bebbington (2003) improved the detection capability of Poisson EWMA chart by calculating the EWMA charting statistic values with two smoothing parameters. Fast initial response (FIR) features were further incorporated by Chiu and Sheu (2008) into a general EWMA chart initiated by Sheu and Chiu (2007). Shu, Jiang, and Wu (2012) transformed Poisson counts into normally distributed observations and monitored the process by an upper-sided CUSUM control chart. In recent years, some sampling methods have been integrated into the traditional Poisson EWMA chart to improve its performance such as the ranked set sampling (Abujiya, Abbasi, & Riaz, 2016) and progressive sampling (Abbasi, 2017). Particularly, Alevizakos and Koukouvinos (2020) compared the performance of the aforementioned representative control charts.

Other features of Poisson counts have also been considered in existing research. Testik (2007) evaluated the performance of Poisson CUSUM chart with an estimated in-control (IC) parameter. The problem to monitor Poisson counts with varying sample size was tackled by Ryan and Woodall (2010), Jiang, Shu, and Tsui (2011), Zhou, Zou, Wang, and Jiang (2012) and Shen, Zou, Jiang, and Tsung (2013) respectively. All

* Corresponding author.
E-mail address: junjwang3-c@my.cityu.edu.hk (J. Wang).

https://doi.org/10.1016/j.cie.2021.107100
Received 23 May 2020; Received in revised form 17 November 2020; Accepted 4 January 2021
Available online 9 January 2021

the above-mentioned charting techniques assume that the mean and variance of Poisson counts are equal. However, the variance of counts may be either larger than their mean (over-dispersion) or smaller than the mean (under-dispersion). Saghir and Lin (2015) provided a detailed review of the methods on count data with under-dispersion or over-dispersion. The latest work goes to a novel CUSUM scheme proposed by Yu, Wu, Wang, and Tsung (2018), which integrated weighted likelihood ratio to improve the method's robustness to over-dispersed counts. Wang, Li, and Xue (2018) proposed a general approach to monitor both types of counts based on the Conway–Maxwell–Poisson model.

Moreover, researchers introduced zero-inflated Poisson models for counts with excessive zeros and developed all kinds of monitoring schemes, which have been reviewed well by Li et al. (2019), Mahmood and Xie (2019). To alleviate the limits of assuming independence among the sequential observations, researchers exploited time series models such as Poisson integer-values aggressive models to characterize the serial dependence or autocorrelation of Poisson counts. Then novel EWMA and CUSUM charting techniques can be established to monitor the parameters or residuals of the model. Such work includes Weiß (2011), Weiß and Testik (2012, 2015), Sales, Pinho, Vivacqua, and Ho (2020). To be different, Li et al. (2019) regarded an autocorrelated Poisson variable as multiple dependent Poisson variables and proposed a multivariate control chart for online monitoring. Bourguignon, Rodrigues, and Santos-Neto (2019) further considered the dispersion in monitoring autocorrelated counting processes. Apart from these parametric approaches, Wang and Qiu (2018) and Qiu, He, and Wang (2019) developed nonparametric methods to monitor univariate and multivariate count data with unknown probability distributions.

Despite the efficiency of existing monitoring schemes for Poisson data, priori shift information is seldom considered in the EWMA chart to improve the power of change detection. Such shift information may be obtained according to the domain knowledges or engineers' past experiences. The information about how many process parameters may become abnormal has been utilized by Wang and Jiang (2009) in high-dimensional processes, by Li, Tsung, and Zou (2012) in multivariate categorical processes and Li, Liu, and Xian (2017) in causal networks for the improvement of monitoring performance. Another type of priori information is acquired by estimating the shift size in advance. With regard to a normally distributed variable, Capizzi and Masarotto (2003) proposed an adaptive EWMA chart, which can adjust the smoothing parameter automatically according to the score functions that consider estimated shift information. Moreover, Shu, Huang, and Jiang (2014) proposed a novel charting design method with estimated shift sizes, which obtains the optimal smoothing parameter and control limit by minimizing OC ARL subject to given IC ARL. By analogy, Huang, Shu, and Jiang (2016) redesigned the CUSUM chart by assuming the shift size is unknown but its underlying distribution can be determined. The two works were further extended by Huang, Shu, and Jiang (2018) to find optimal parameters for multivariate normal EWMA charts. In monitoring Poisson data, CUSUM and adaptive CUSUM charts can integrate priori shift information although they are one-sided (He, Shu, & Tsui, 2014). However, the traditional Poisson EWMA chart proposed by Borror et al. (1998) can only select appropriate smoothing parameter according to the rule found by Lucas and Saccucci (1990) that smaller smoothing parameter leads to quicker detection of smaller shift while the larger one is more suitable for detecting larger shifts. In general, this rule fails to mathematically characterize the relationship between smoothing parameter and predetermined priori shift information, which may cause subjective parameter selection.

This article discusses the optimal design of Poisson EWMA chart in two cases to bridge the research gap mentioned above. When the magnitude of shift can be estimated in advance, we prove that there exist an optimal smoothing parameter and two-sided control limits such that OC ARL is minimized. When only the probability distribution of shift is known, we suggested minimizing the expected ARL (EARL) to search for the chart's optimal parameters, which definitely exist under a few mild conditions. Note that the ARL is calculated based on Markov chain approximation in line with Borror et al. (1998), which avoids a large computational burden of the simulation method. We can further approximately calculate EARL by averaging the ARLs associated with all possible mean shifts. Since it is intractable to obtain the closed form of optimal parameters, we make another contribution to introducing a heuristic algorithm, i.e., Fibonacci search algorithm (Hassin & Sarid, 2018), which provides a good solution in comparison with the conventional binary search algorithm. Two real datasets respectively from a manufacturing process and public medical health are employed to illustrate the implementation and effectiveness of the proposed methods.

The remainder of this article is organized as follows. Section 2 introduces the traditional Poisson EWMA chart, which is efficiently optimized in Section 3 with proposed Markov approximation method and the Fibonacci search algorithm. The performance of the optimal monitoring scheme is evaluated by simulation in Section 4 and two real examples are implemented in Section 5 to show the effectiveness of the proposed scheme. Finally, Section 6 provides some concluding remarks.

2. Poisson EWMA control chart

In this article, we follow the tradition in SPC field that a Poisson process is said to be in IC state when only common cause variation is present. In contrast, the process becomes out of control once extra variation occurs due to some assignable causes (Qiu, 2014). Suppose X_t denotes the observation of a Poisson variable at time point t and all IC observations are independently and identically distributed (i.i.d.). For any $t > \tau$, Poisson mean μ may change from IC μ_0 to some unknown μ_1, which indicates that the process transmits into the OC state. All IC and OC observations can be presumably collected according to the following change-point model

$$X_t \sim \begin{cases} \text{Poisson}(\mu_0), & \text{for } t = 1, 2, \cdots, \tau \\ \text{Poisson}(\mu_1), & \text{for } t = \tau + 1, \cdots \end{cases}$$

Since τ is unknown in advance, we should conduct the following hypothesis test at each time point:

$$H_0 : \mu = \mu_0 \qquad H_1 : \mu = \mu_1$$

where "H_0" and "H_1" corresponds to the null hypothesis and alternative hypothesis respectively. According to Borror et al. (1998), observation X_t can be regarded as the monitoring statistic. To integrate current and historical samples in surveillance, the plotting statistic of Poisson EWMA chart is computed by the following formula:

$$Z_t = (1 - \lambda)Z_{t-1} + \lambda X_t$$

where $0 < \lambda \leq 1$ is smoothing parameter. When the process is in IC state, it is easy to know that

$$E(Z_t) = \mu_0 \text{ and } Var(Z_t) = \frac{\lambda}{2 - \lambda}\left[1 - (1 - \lambda)^{2t}\right]\mu_0 \approx \frac{\lambda \mu_0}{2 - \lambda} \text{ with } t \to +\infty$$

Then the upper and lower control limits denoted by h_U and h_L respectively can be obtained based on the asymptotic variance of Z_t as below:

$$h_U = \mu_0 + A_U\sqrt{\frac{\lambda \mu_0}{2 - \lambda}}$$

$$h_L = \mu_0 - A_L\sqrt{\frac{\lambda \mu_0}{2 - \lambda}}$$

where A_U and A_L are two constants determined based on given Type I error. According to Borror et al. (1998), $A_U = A_L = A$ is often used to establish symmetric control charts although asymmetric control charts

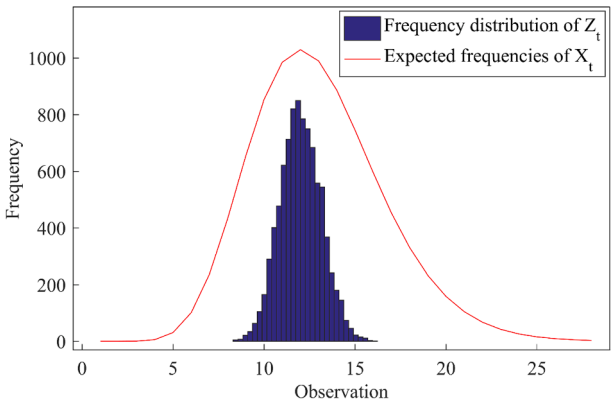

Fig. 1. Comparison between frequency distribution of Z_t and expected frequencies of X_t with $t \in [1001, 10000]$, $\mu_1 = 12$ and $\lambda = 0.2$.

may be preferred sometimes. In this work, we intend to design a symmetric Poisson EWMA chart with control limits set to:

$$h_U = \mu_0 + A\sqrt{\frac{\lambda\mu_0}{2-\lambda}}$$

$$h_L = \mu_0 - A\sqrt{\frac{\lambda\mu_0}{2-\lambda}}$$

Note that if resultant $h_L < 0$, we should set h_L to 0. Once $Z_t > h_U$ or $Z_t < h_L$, we should reject the null hypothesis and conclude that the process has gone out of control.

When a process is in control, we usually define the number of samples taken before the release of a false OC signal as run length. Thus, the run length for an EWMA chart approximately follows a geometric distribution with the success probability equal to Type I error α so that the IC ARL can be computed as $1/\alpha$ according to Lucas and Saccucci (1990). Note that all samples are collected at equally spaced time points throughout this work. Correspondingly, the OC run length can be defined as the number of samples collected from the time of shift occurrence to the time of correct OC signal. Simulation results from Lucas and Saccucci (1990) show that the OC run length of an EWMA chart will follow some distribution other than the geometric distribution but ARL remains to be effective to reflect the chart's performance. When the Poisson process is in steady state i.e., $t \to +\infty$ with mean shift δ, it is easy to know that the statistic Z_t will follow a stable distribution, the shape of which is similar to Poisson distribution according to our simulation results (see Fig. 1 as an example). Therefore, the probability of failing to trigger the OC signal denoted by $\beta = Pr\{h_L \le Z_t \le h_U\}$ is fixed given mean shift δ and all other parameters. Moreover, the larger β is, the larger OC ARL denoted by $\mathrm{ARL}(\lambda, A, \delta)$ will be. We express their positively correlated relationship as below:

$$\mathrm{ARL}(\lambda, A, \delta) \propto \beta = Pr\left\{ \left| \frac{z-\mu_0}{\sqrt{\frac{\lambda\mu_0}{2-\lambda}}} \right| \le A \right\} \tag{1}$$

where the observation of random variable z is $Z_t (t = 1, 2, \cdots)$. According to the Central Limit Theorem, when $\delta = 0$ and sample size approaches to infinity, random variable $x = (z-\mu_0)/\sqrt{\frac{\lambda\mu_0}{2-\lambda}}$ will asymptotically follow the standard normal distribution with probability density function $\phi(x)$ and cumulative distribution function $\Phi(x)$. If $\delta \neq 0$, the Poisson mean would become $\mu_1 = \mu_0 + \delta$ and $y = (z-\mu_1)/\sqrt{\frac{\lambda\mu_1}{2-\lambda}}$ would asymptotically follow the standard normal distribution with sample size approaching to infinity.

3. Optimal design with Markov chain approximation

It can be seen that the performance metric $\mathrm{ARL}(\lambda, A, \delta)$ of Poisson EWMA chart is determined by three parameters λ, A and δ. According to Qiu (2014), for two-sided EWMA chart, smoothing parameter λ is usually determined beforehand while parameter A is then specified to reach a predefined IC ARL, i.e., ARL_0. In other words, any value combinations of λ and A are feasible as long as they satisfy $\mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$. Lucas and Saccucci (1990) proposed one general rule to guide the selection of λ that larger λ can help the chart detect larger shifts while the smaller one promotes quicker detection of smaller shifts. Instead of determining whether the most probable shift is small or large in a subjective way, this work assumes that we can estimate Poisson mean shift δ in advance based on engineering knowledges or practitioners' past experiences in line with many existing articles mentioned in our Introduction.

Once δ is specified, numerous values of $\mathrm{ARL}(\lambda, A, \delta)$ can be obtained according to numerous feasible value combinations of λ and A satisfying $\mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$. The idea is natural that we may choose the combination of λ and A that induces the minimum OC ARL. The resultant control chart should detect the shift most quickly than the charts with other parameter combinations when estimated shift δ really occurs. To this end, the design of Poisson EWMA chart becomes searching a combination of parameters A, λ to minimize the OC ARL, which induces the following optimization problem:

$$\min_{\lambda, A} \ \mathrm{ARL}(\lambda, A, \delta)$$

$$\text{subject to} \quad \mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$$

where $\mathrm{ARL}_0 = 1/\alpha$ represents IC ARL. Actually, Shu et al. (2014) and Huang et al. (2018) have solved this problem for univariate and multivariate normal variables respectively. They used the integral function proposed by Crowder, S. V. (1987) to calculate the ARLs and obtained the derivatives of integral function with respect to λ and A respectively. A gradient approach was established to search for the optimal combination of λ and A that minimize the OC ARL. Since a Poisson variable is discrete, neither the integral function nor the gradient approach is applicable in this article. To make an extension, we calculate the ARLs via Markov chain approximation and then introduce a new method for optimal design of EWMA control chart.

Since it may be computationally intensive to search for λ and A simultaneously, this work makes another difference that the two-parameter optimization problem is transformed into one parameter search. Notice that the constraint $\mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$ exhibits an implicit relationship between λ and A, which enables us to calculate one parameter given the other. Another issue arises that how we can assure there is solution to the transformed optimization problem. In terms of this, it is necessary to investigate the properties of function $\mathrm{ARL}(\lambda, A, \delta)$ with respect to its parameters λ and A. All these are based on the assumptions that all observations are independent and identically distributed (i.i.d.) and samples are collected at equally spaced time periods. In this article, we would like to express A as a function of λ according to the constraint $\mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$ so that the relationship between these two parameters will be discussed first. It is straightforward to see from the image of normal probability distribution that $Pr\left\{ \left| \frac{z-\mu_0}{\sqrt{\frac{\lambda\mu_0}{2-\lambda}}} \right| \le A \right\}$ exhibits a positive correlation relationship with A, i.e., $\frac{\partial \mathrm{ARL}(\lambda, A, 0)}{\partial A} > 0$. Then one property can be clarified as the following proposition:

Proposition 1:. *Given* $\delta = 0$, *we have the results: a.* $\frac{dA}{d\lambda} > 0$; *b.*$\frac{\partial ARL(\lambda, A, 0)}{\partial \lambda} < 0$.

Proof.. *By analogy with formula* (1), *there exists* $\mathrm{ARL}(\lambda, A, 0) \propto$

$Pr\left\{\left|\frac{z-\mu_0}{\sqrt{\frac{\lambda\mu_0}{2-\lambda}}}\right| \le A\right\}$. *When* λ *increases, the standard deviation of* z *denoted by* $Std(z)=\sigma_0=\sqrt{\frac{\lambda\mu_0}{2-\lambda}}$ *will also increase, which means that observations of* z *become more scattered. It is easy to see from the image of normal probability distribution that* $Pr\left\{\left|\frac{z-\mu_0}{\sqrt{\frac{\lambda\mu_0}{2-\lambda}}}\right| \le A\right\}$ *shall decrease. To maintain* $\mathrm{ARL}(\lambda,A,0)=\mathrm{ARL}_0$, *we have to let* A *increase. Therefore, when the constraint* $\mathrm{ARL}(\lambda,A,0)=\mathrm{ARL}_0$ *is imposed, parameter* A *shall increase with increasing* λ, *i.e.,* $\frac{dA}{d\lambda}>0$. *The first-order derivative of* $\mathrm{ARL}(\lambda,A,0)=\mathrm{ARL}_0$ *with regard to* λ *can be described as*

$$\frac{\partial \mathrm{ARL}(\lambda,A,0)}{\partial\lambda}+\frac{\partial \mathrm{ARL}(\lambda,A,0)}{\partial A}\times\frac{dA}{d\lambda}=0.$$

According to $\frac{\partial \mathrm{ARL}(\lambda,A,0)}{\partial A}>0$ and $\frac{dA}{d\lambda}>0$, we can obtain $\frac{\partial \mathrm{ARL}(\lambda,A,0)}{\partial\lambda}<0$, which is the second half statement of Proposition 1.

This proposition tells us that given any value of smoothing parameter $\lambda\in(0,1)$, there exist a unique A such that prespecified IC ARL is satisfied. It sets a base for searching the optimal λ instead of searching λ and A simultaneously to minimize OC ARL. These properties of Poisson EWMA chart have been validated by Borror et al. (1998) by various simulation results. However, Borror et al. (1998) failed to discuss the relationships among λ, A and $\mathrm{ARL}(\lambda,A,\delta)$ when the process is in OC state. Actually, it is complicated to analyze these relationships with mathematical formulas. In monitoring a normal variable with a constant shift, Crowder (1987) provided numerous simulation results to show that OC ARL is convex with $\lambda\in(0,1)$ and fixed control limit. Shu et al. (2014) found that a unique combination of λ and A exists to minimize OC ARL subject to a prespecified IC ARL when a constant shift occurs with the univariate normal variable. Similar results can also be found in designing multivariate EWMA chart (Huang et al., 2018). Besides, Qiu (2014) suggested a general guideline that λ can be chosen such that the OC ARL for detecting the target shift is minimized. Despite the assumption of monitoring normal variables, these results may also be true for general EWMA charts including Poisson EWMA chart. In Section 5, we will prove this with numerous simulation results.

Next, the Markov chain approach will be applied to approximate the ARL of the proposed Poisson EWMA control chart and optimal λ, A can be further obtained. It is well known that Markov chain focuses on characterizing the transition of a process from one state to another according to a transition probability matrix. As mentioned above, the process is in IC state when $h_L\le Z_t\le h_U$ and will be regarded as going out of control otherwise. To make it further, we divide the IC interval $[h_L, h_U]$ into m subintervals, the width of each should be $\frac{L}{m}$. When $Z_t\in(\frac{(i-1)L}{m}, \frac{iL}{m}]$, the process can be said in the ith transient state. Note that $Z_t>h_U$ or $Z_t<h_L$ represents the absorbing state or OC state since the recurrent hypothesis test will stop in this state. In this way, the process will transmit in $m+1$ states with m transient states and one absorbing state, which can be naturally characterized by a Markov chain.

Denote p_{ij} as the probability of moving from ith state to jth state and let the midpoint $d_i=\frac{(i-0.5)L}{m}$ represent the ith subinterval, $c_j=\frac{jL}{m}$ as the jth cutting point, we can have

$$p_{ij}=Pr\left(c_{j-1}<Z_t\le c_j|Z_{t-1}=d_i\right)$$

$$=Pr\left(c_{j-1}<(1-\lambda)Z_{t-1}+\lambda X_t\le c_j|Z_{t-1}=d_i\right)$$

$$=Pr\left(c_{j-1}<(1-\lambda)d_i+\lambda X_t\le c_j\right)$$

$$=Pr\left(\frac{c_{j-1}-(1-\lambda)d_i}{\lambda}<X_t\le\frac{c_j-(1-\lambda)d_i}{\lambda}\right)$$

Since X_t follows Poisson distribution with mean μ_0 in the IC state, it is straightforward to compute p_{ij}. Once all $p_{ij}(i,j=1,2,\cdots,m)$ are figured out, p_{ij} can form a $m\times m$ matrix denoted by $\mathbf{R}$ with p_{ij} as its entry in ith row and jth column. Furthermore, by adding $(m+1)$th row and column to consider the transition to absorbing state, we can formulate the complete transition probability matrix

$$\mathbf{P}=\begin{bmatrix}\mathbf{R} & (\mathbf{I}-\mathbf{R})\mathbf{1}\\ \mathbf{0}^{\mathbf{T}} & 1\end{bmatrix},$$

where $\mathbf{I}$ is the $m\times m$ identity matrix and $\mathbf{1}$ represents a $m\times 1$ column vector with all elements equal to one. According to Lucas and Saccucci (1990), the average run length (ARL) can be calculated by

$$\mathrm{ARL}=\mathbf{p}^T(\mathbf{I}-\mathbf{R})^{-1}\mathbf{1},$$

where $\mathbf{p}$ represents the initial state of the process. Here we only consider the zero-state case, where the process transits from the first state, i.e., $Z_0=\mu_0$. In this way, we can let m be an odd number and $\mathbf{p}=[0,\cdots,1,\cdots,0]^T$ with only the $\frac{m+1}{2}$th element as 1.

To this end, when λ, A, δ are known, ARL can be computed easily according to the Markov chain approximation. Note that $\delta=0$ induces the IC ARL while $\delta\neq 0$ corresponds to the OC ARL. Since IC ARL, i.e., ARL_0 is determined in advance and shall not change throughout the monitoring process, we can express A as the function of λ denoted by $A(\lambda)$ according to the constraint $\mathrm{ARL}(\lambda,A,0)=\mathrm{ARL}_0$. The consequent optimization problem becomes

$$\min_{\lambda}\mathrm{ARL}(\lambda,A(\lambda),\delta)$$

The optimization problem has been transformed into searching for a λ between $[\lambda_{min},\lambda_{max}]$ leading to the smallest OC ARL. It is intuitive to conduct the grid search algorithm to complete this task. However, the grid search algorithm requires much computational cost since it considers each kind of situation. Hassin and Sarid (2018) reviewed existing dichotomous search methods in operations research, which are available to us. In this work, we exploit two fast and efficient algorithms, the Fibonacci search algorithm and the bisection search algorithm. The latter algorithm is introduced in Appendix while the former involves the following steps:

a. Generate Fibonacci numbers according to $F(\gamma)=F(\gamma-1)+F(\gamma-2)$ with γ as a positive integer and $F(1)=0, F(2)=1$. In line with most research, we consider 20 Fibonacci numbers so that $\gamma_{max}=20$;
b. Set $\lambda_1=\lambda_{max}-(\lambda_{max}-\lambda_{min})\times F(\gamma_{max}-1)/F(\gamma_{max})$, $\lambda_2=\lambda_{min}+(\lambda_{max}-\lambda_{min})\times F(\gamma_{max}-1)/F(\gamma_{max})$ and calculate $\mathrm{ARL}(\lambda_1, A(\lambda_1),\delta)$ denoted by ARL_1 and $\mathrm{ARL}(\lambda_2,A(\lambda_2),\delta)$ denoted by ARL_2;
c. For the ith step, conduct the following procedure:

if $\mathrm{ARL}_1<\mathrm{ARL}_2$, let $\lambda_{max}=\lambda_2$, $\lambda_2=\lambda_1$, $\lambda_1=\lambda_{max}-(\lambda_{max}-\lambda_{min})\times F(\gamma_{max}-i-1)/F(\gamma_{max}-i)$, $\mathrm{ARL}_2=\mathrm{ARL}_1$ and then calculate new ARL_1;
if $\mathrm{ARL}_1>\mathrm{ARL}_2$, let $\lambda_{min}=\lambda_1$, $\lambda_1=\lambda_2$, $\lambda_2=\lambda_{min}+(\lambda_{max}-\lambda_{min})\times F(\gamma_{max}-i-1)/F(\gamma_{max}-i)$, $\mathrm{ARL}_1=\mathrm{ARL}_2$ and then calculate new ARL_2;

d. Repeat Step c until $i=\gamma_{max}-2$ or $|\mathrm{ARL}_1-\mathrm{ARL}_2|<T$ with T as a given threshold and optimal $\lambda=(\lambda_{max}+\lambda_{min})/2$.

4. Extensions of Poisson EWMA chart with unknown shift sizes

The preceding contents assume that the shift size of a Poisson process is known in advance, which may not be the usual case in practice. Sometimes we can only obtain a range of possible shift sizes that follow certain probability distribution based on engineering knowledges or practitioners' past experiences (Huang et al., 2016). As an extension to the aforementioned study, this section discusses how to design a Poisson EWMA control chart when the shift size follows a certain probability distribution. Since there is one-to-one correspondence between δ and

$\mathrm{ARL}(\lambda, A(\lambda), \delta)$, it is suitable to minimize expected average run length (EARL) instead of OC ARL in optimal design. Note that Huang et al. (2016) introduced a weight function to indicate the importance of the shift magnitude in computing EARL for the CUSUM chart. However, the weight function is omitted here for the ease of exposition. Assume the most possible range of δ is $[a, b]$ and the optimization problem can be expressed as

$$\min_{\lambda} \mathrm{EARL}(\lambda, A(\lambda), \delta) = \int_a^b f(\delta)\mathrm{ARL}(\lambda, A(\lambda), \delta)d\delta$$

$$\text{s.t} \quad \mathrm{ARL}(\lambda, A, 0) = \mathrm{ARL}_0$$

where $f(\delta)$ represents the probability density function of δ and other parameters are defined as above. Since it is intractable to obtain the explicit function of $\mathrm{EARL}(\lambda, A(\lambda), \delta)$, a gradient approach is proposed by Huang et al. (2016) to search for optimal parameters. Apparently, the closed-form derivative of EARL with respect to all parameters can hardly be obtained in this article due to the exploitation of the Markov chain approximation to calculate ARLs. To exhibit the design of Poisson EWMA chart easily, we assume the shift δ follows a uniform distribution $U(a, b)$ with a and b as parameters. In this way, our task is to figure out the optimal smoothing parameter λ and the corresponding constant A given δ $U(a, b)$. According to integral properties, we can divide the uniform interval $[a, b]$ into q subintervals and make an approximation:

$$\mathrm{EARL}(\lambda, A(\lambda)) \approx \frac{1}{q}\sum_{i=1}^{q} \mathrm{ARL}(\lambda, A(\lambda), \delta_i)$$

with $\delta_i = a + \frac{i}{q}(b - a)$. Note that similar approximations of integral calculation can be made when other probability distributions are involved.

Numerous researches have shown that OC ARL shall decline with increasing magnitude of process shift δ. In particular, when δ ranges from a to b with $a < 0$ and $b > 0$, $\mathrm{ARL}(\lambda, A(\lambda), \delta)$ will first increase and then decrease. In this case, it will be difficult to determine how the selection of λ will influence the average of $\mathrm{ARL}(\lambda, A(\lambda), \delta_i)(i = 1, \cdots, q)$, i.e., $\mathrm{EARL}(\lambda, A(\lambda))$ although Proposition 1 indicates that $\mathrm{ARL}(\lambda, A(\lambda), \delta)$ is a convex function of λ given δ. However, when δ ranges from a to b with $ab > 0$, we will see $\mathrm{ARL}(\lambda, A(\lambda), \delta)$ either decrease $(a, b < 0)$ or increase $(a, b > 0)$. Due to the convex relationship of $\mathrm{ARL}(\lambda, A(\lambda), \delta_i)$ with λ given δ_i, $\mathrm{EARL}(\lambda, A(\lambda))$ is also convex with respect to λ. To assure the existence of a unique λ to minimize EARL with δ $U(a, b)$, we have to let $ab > 0$. The proposed Fibonacci search algorithm is also applicable in searching the optimal smoothing parameter λ. It can be seen that the proposed design of Poisson EWMA chart requires the forecast of process shift direction and magnitude. In SPC field, this is quite common before the implementation of one-sided control charts such as CUSUM chart and can be completed based on engineering knowledges and practitioners' past experiences.

5. Performance study

This section evaluates performance of the proposed optimal design of Poisson EWMA chart based on simulation study. Specifically, this section is divided into six parts. The first part shows how to determine the number of subintervals for Markov chain approximation and evaluates the performance of ARL approximation whereas the second part validates the existence of a unique solution to the optimization problem mentioned in this work. In the third subsection, the Fibonacci search algorithm is compared with the bisection search algorithm and some optimal combinations of parameters are provided in. The sensitivity analysis of IC and OC ARL is conducted in the next part with varying A, which is followed by the performance study with unknown shift sizes in the fifth subsection. Finally, we compare the optimally designed Poisson EWMA chart with its non-optimally designed version.

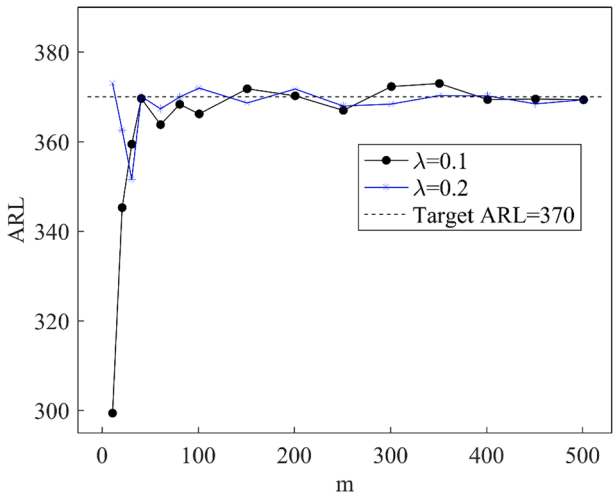

Fig. 2. The effect of m on actual ARL.

Table 1
IC ARL comparison between Markov approximation method and simulation method.

μ_0	ARL_o=	100	200	370	500	800
1	IC_A	99.28	201.6	373.7	493.3	799.9
	IC_S	103.7	205.2	371.8	522.4	792.8
2	IC_A	101.6	210.3	368.1	501.3	798.3
	IC_S	102.5	207.4	381.2	511.2	813.6
3	IC_A	98.83	198.2	367.6	498.4	793.4
	IC_S	97.74	198.5	375.7	501.5	818.1
4	IC_A	99.49	199.4	363.5	493.4	800.0
	IC_S	98.77	199.4	372.7	511.1	806.6
6	IC_A	101.7	201.2	378.4	500.3	792.8
	IC_S	103.9	207.8	380.2	503.9	834.8
8	IC_A	99.17	198.5	371.2	498.7	791.7
	IC_S	99.60	204.5	371.2	485.8	782.6
10	IC_A	99.01	198.6	371.6	507.7	790.1
	IC_S	100.9	203.3	366.0	500.5	793.0

5.1. Performance of Markov chain approximation

As mentioned in Section 3, ARL calculation with Markov chain approximation is highly affected by m, the number of subintervals between the upper and lower control limits. To figure out the appropriate m, we assume the Poisson mean shift δ amounts to 0 and set target IC ARL to 370. With varying m, different actual ARLs can be calculated based on Markov approximation and formulate Fig. 2. To display the

Table 2
OC ARL comparison between Markov approximation method and simulation method with IC ARL = 370

μ_0	δ =	0.5	1.0	2.0	4.0	6.0	8.0
1	OC_A	28.31	–	–	–	–	–
	OC_S	28.69	–	–	–	–	–
2	OC_A	47.90	16.12	–	–	–	–
	OC_S	48.01	16.37	–	–	–	–
3	OC_A	66.39	22.69	7.672	–	–	–
	OC_S	65.40	22.82	7.775	–	–	–
4	OC_A	82.72	29.06	9.417	–	–	–
	OC_S	83.65	29.15	9.592	–	–	–
6	OC_A	114.2	42.62	13.19	4.767	–	–
	OC_S	112.3	42.13	13.13	4.777	–	–
8	OC_A	136.8	54.64	16.83	5.738	3.452	–
	OC_S	137.9	54.87	16.84	5.755	3.476	–
10	OC_A	156.7	66.32	20.59	6.706	3.923	2.840
	OC_S	153.8	65.03	20.27	6.730	3.912	2.830

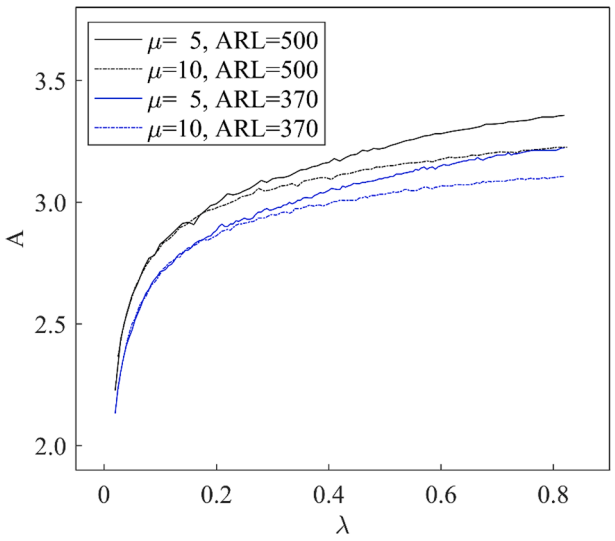

Fig. 3. Relationship between A and λ.

effect of smoothing parameter, ARLs with $\lambda = 0.1$ and $\lambda = 0.2$ are plotted. It can be seen that actual ARL approaches to its target value 370 when m gets larger despite the value of λ. Good ARL approximation requires that m should exceed 100. In this sense, all simulation results in our work are obtained with $m = 101$. Additional simulation results are available upon request from the authors.

It is also necessary to evaluate how much close the ARLs given by Markov chain approximation are to the actual ones given m. We conduct a simulation study to calculate actual ARLs with the optimal parameters obtained via our proposed method. Table 1 shows various IC ARLs from approximation (denoted by IC_A) and those from simulation (denoted by IC_S) when the IC Poisson mean is set to different values. It can be seen that the two types of IC ARLs are close to each other in most cases and all of them are within ±10% of their nominal values. Similar results can be seen from Table 2, which displays the OC ARL comparison between Markov chain approximation method and simulation method with IC ARL = 370. It seems that the magnitude of IC mean will not affect the performances of Markov chain approximation method and the optimally designed Poisson EWMA chart.

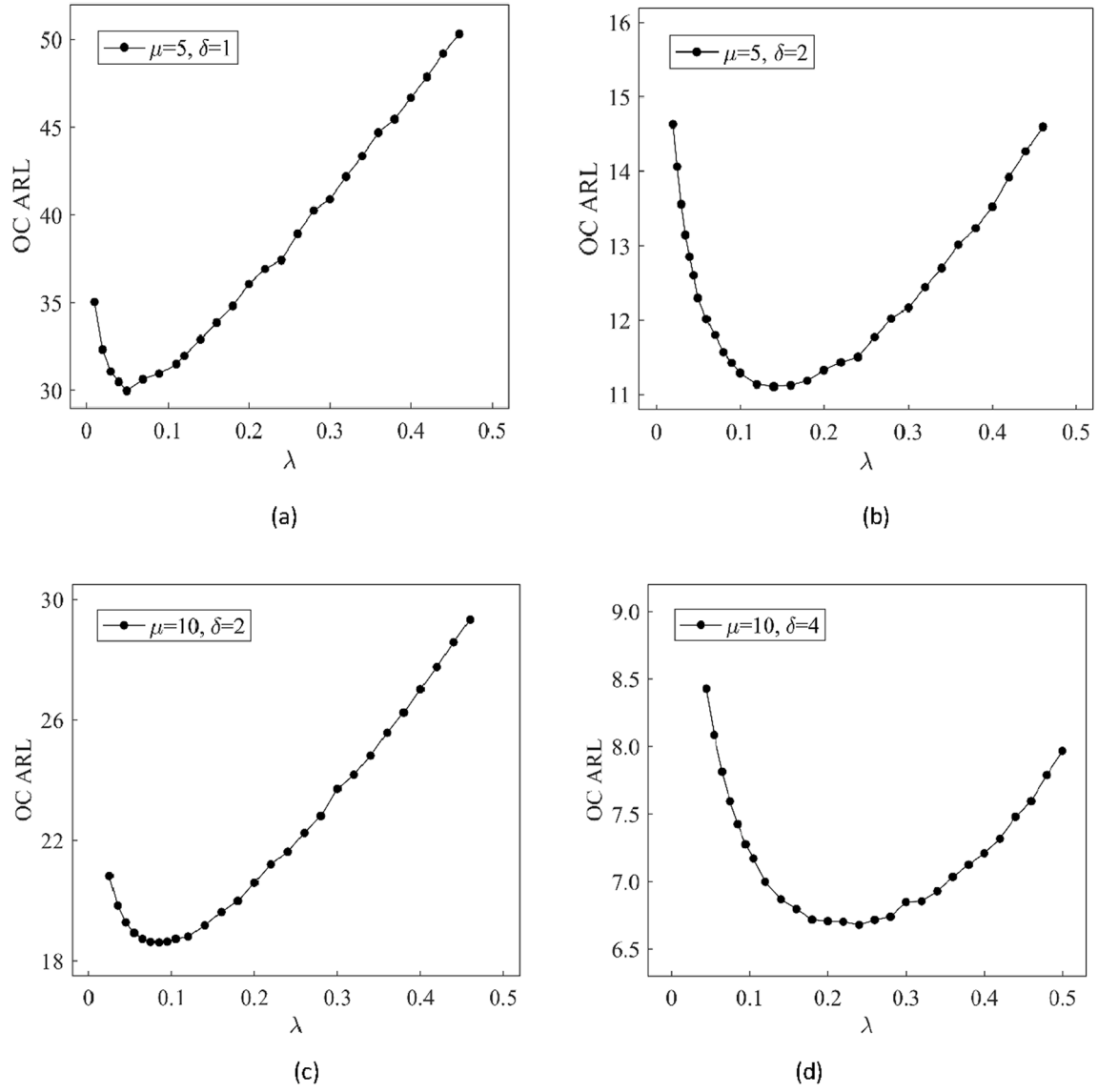

Fig. 4. The effect of λ on OC ARL.

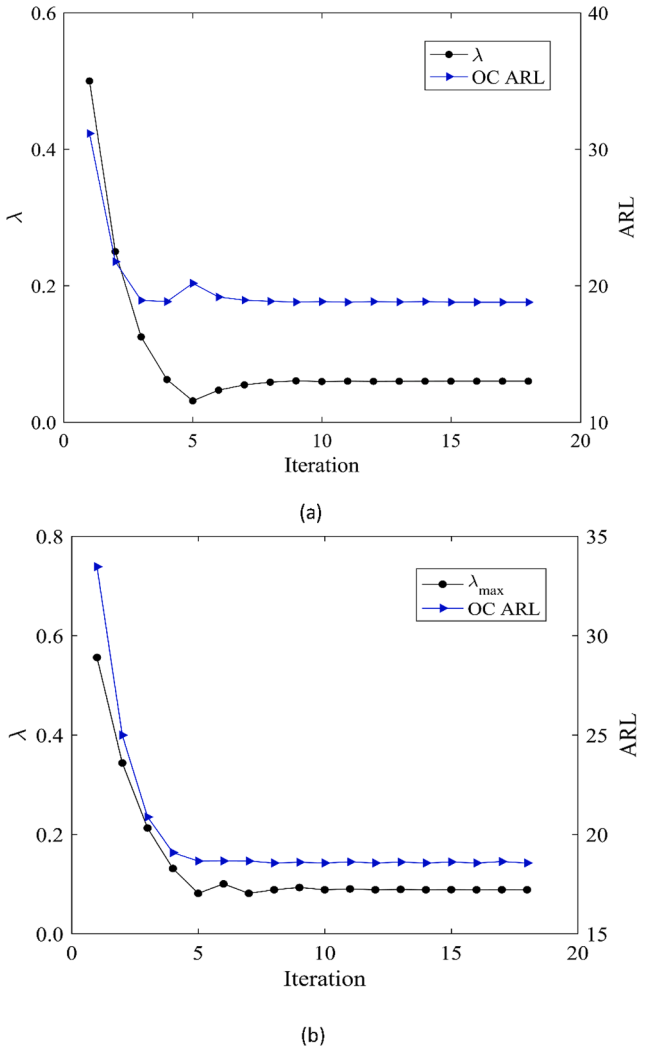

Fig. 5. Searching process with $ARL_0 = 370$ and $\delta = 2$. (a) Bisection search; (b) Fibonacci search.

5.2. *Validating the existence of optimal λ*

Actually, part of the simulation results (Figs. 2–9) provided by Borror et al. (1998) supported $\frac{\partial \mathrm{ARL}(\lambda,A,0)}{\partial \lambda} < 0$ and $\frac{\partial \mathrm{ARL}(\lambda,A,\delta)}{\partial A} > 0$ mentioned in our Proposition 1. Here we investigate the relationship between λ and A subject to a fixed IC ARL. In the simulation, the IC ARL is set to 370 and 500 respectively while $\mu_0 = 5$ and $\mu_0 = 10$ are chosen separately as IC Poisson mean. With varying λ, different A can be attained based on bisection search algorithm under the four scenarios. According to Fig. 3, given μ and IC ARL, constant A is monotonically increasing with respect to smoothing parameter λ, which validates the statement $\frac{dA}{d\lambda} > 0$ in Proposition 1.

In the sequel, simulation results will be provided to validate the uniqueness of a smoothing parameter that minimizes OC ARL. Here IC ARL is set to 370 and $\mu_0 = 5$ and $\mu_0 = 10$ are considered respectively. When $\mu_0 = 5$, the mean shift is set to $\delta = 1$, $\delta = 2$, which are 20% and 40% of the mean separately. It should be notified that the results with $\delta < 0$ are similar and hence they are not listed here due to page limits. According to Fig. 4a-4b, OC ARL is smaller than IC ARL 370 and exhibits a convex relationship with smoothing parameter λ. Furthermore, the optimal λ that minimizes OC ARL lies in $(0, 0.1)$ when $\delta = 1$ and in $(0.1, 0.2)$ when $\delta = 2$. In the scenario $\mu_0 = 10$, δ is chosen as 2 and 4 to maintain the 10% and 20% of the mean respectively. Similar conclusions can be made except for the case where the OC ARL is minimized

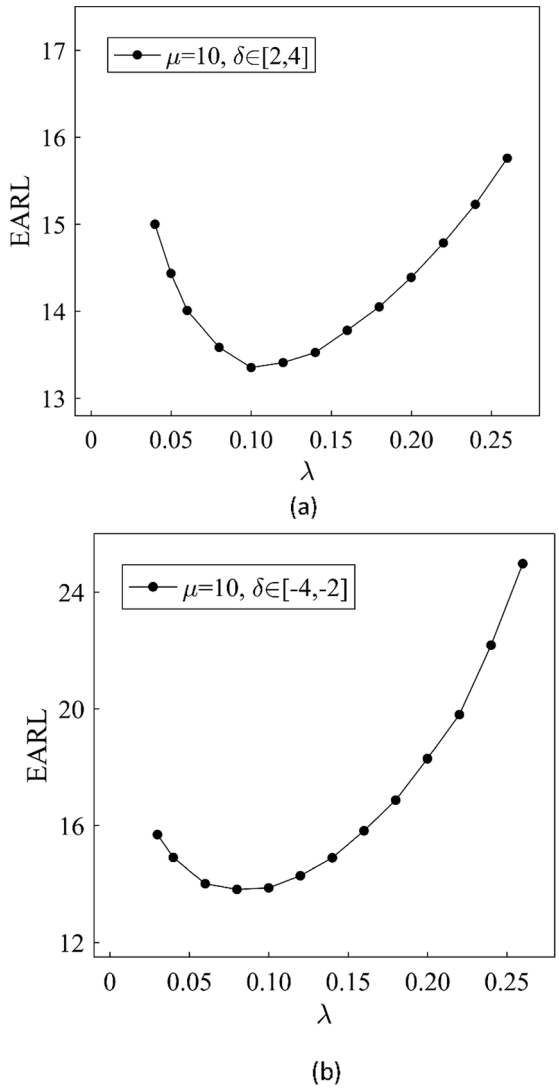

Fig. 6. The effect of λ on EARL.

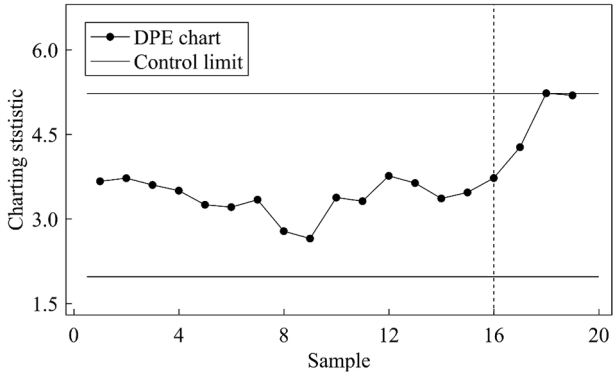

Fig. 7. Optimal monitoring of PCB packaging processes.

with $\lambda \in (0.2, 0.3)$ when $\delta = 4$. It seems that optimal λ will increase when the Poisson mean gets larger and the proportion of shift is maintained. From these and many other simulation results, we can validate the existence of optimal smoothing parameter.

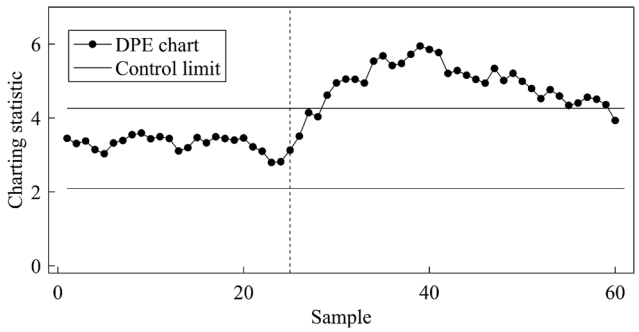

Fig. 8. Optimal monitoring of monthly deaths caused by hepatitis C.

Table 3
Performance comparison between grid search, Fibonacci search and bisection search algorithm with IC ARL = 370.

		Grid search		Fibonacci search		Bisection search	
	δ	λ_{opt}	ARL_{min}	λ_{opt}	ARL_{min}	λ_{opt}	ARL_{min}
$\mu_0 = 5$	1	0.055	30.22	0.054	30.18	0.046	30.29
	2	0.140	11.11	0.150	11.09	0.120	11.14
	3	0.240	6.142	0.233	6.154	0.245	6.150
$\mu_0 = 10$	1	0.026	48.86	0.025	48.85	0.024	48.93
	2	0.085	18.61	0.085	18.56	0.060	18.83
	3	0.140	10.23	0.147	10.23	0.124	10.28
	4	0.240	6.679	0.212	6.671	0.250	6.694
	5	0.280	4.811	0.293	4.820	0.250	4.836

5.3. Searching for the optimal λ

As mentioned before, it is intractable to obtain the closed-form derivative of $ARL(\lambda, A(\lambda), \delta)$ with regard to λ. Hence, it is impossible to search for the optimal λ based on gradient approach proposed by Huang et al. (2018). We recommend Fibonacci search algorithm be employed to complete this task and it will be compared with famous bisection search algorithm to demonstrate its good performances. Detailed introduction of Fibonacci search algorithm is provided in Section 3 and that of bisection search algorithm used here is provided in Appendix to save the space of context. In numerical study, Poisson mean is set to $\mu_0 = 5, 10$ with shifts 1, 2, 3 and 1, 2, 3, 4, 5 respectively. Both of mentioned algorithms are implemented in Fortran languages to obtain optimal smoothing parameter denoted by λ_{opt} and corresponding minimized OC ARL denoted by ARL_{min} with IC ARL = 370. We also list the optimal λ_{opt} and ARL_{min} obtained from grid search algorithm for comparison. Among the three search algorithms, grid search algorithm may provide the parameters that are closest to the real optimal parameters since it searches nearly all feasible values of parameters. Therefore, we use the results from grid search algorithm as benchmarks to compare the performance of the other algorithms. Notice that Table 3 only lists results with $\delta > 0$ as similar results can be obtained with $\delta < 0$.

It can be seen from Table 3 that Fibonacci search algorithm and the bisection search algorithm can be said to perform equally well due to their close OC ARLs to those provided by grid search algorithm regardless of the magnitude of δ. It should be notified that Fibonacci search algorithm provides closer results to grid search than bisection search and their ARL_{min} are also smaller than those from bisection search in most cases. To further investigate performances of the two search algorithms, we display their searching details in Fig. 5 with IC ARL = 370 and $\delta = 2$. It should be notified that the two algorithms will behave similarly under other settings. Each subfigure is plotted with left vertical axis representing smoothing parameter λ and the right one representing OC ARL. In Fig. 5(b), λ_{min} is not provided for Fibonacci search algorithm. Fig. 5 shows that it nearly takes five steps before both algorithms reach to the optimal smoothing parameters and corresponding OC ARLs. The convergence speed of Fibonacci search algorithm seems to be no different from that of bisection search algorithm.

Table 4
Optimal λ and A of proposed Poisson EWMA chart with $\mu_0 = 10$.

δ	ARL_0=	100	200	370	500	800	1000
1	λ_{opt}	0.050	0.031	0.031	0.031	0.031	0.031
	A_{opt}	1.875	2.021	2.314	2.451	2.659	2.746
	ARL_{min}	29.26	38.94	48.87	54.12	62.89	67.32
2	λ_{opt}	0.134	0.093	0.088	0.079	0.068	0.063
	A_{opt}	2.246	2.422	2.668	2.751	2.895	2.969
	ARL_{min}	12.65	15.67	18.56	20.06	22.44	23.59
3	λ_{opt}	0.181	0.181	0.148	0.141	0.122	0.120
	A_{opt}	2.324	2.616	2.808	2.905	3.032	3.111
	ARL_{min}	7.343	8.820	10.22	10.92	12.04	12.59
4	λ_{opt}	0.297	0.254	0.212	0.193	0.184	0.187
	A_{opt}	2.441	2.686	2.876	2.969	3.125	3.205
	ARL_{min}	4.932	5.821	6.670	7.094	7.733	8.052
5	λ_{opt}	0.555	0.335	0.294	0.293	0.263	0.253
	A_{opt}	2.549	2.734	2.944	3.049	3.193	3.258
	ARL_{min}	3.686	4.254	4.811	5.086	5.510	5.734

Without generality, we obtain the optimal smoothing parameter, λ_{opt} with corresponding A_{opt}, ARL_{min} in different cases and display them in Table 4. The results show that when δ is fixed and IC ARL gets larger, both of A_{opt} and ARL_{min} increase whereas λ_{opt} declines except for that with $\delta = 1$. This result coincides with Proposition 1. It should be clarified that approximation errors in display make λ_{opt} stay the same as 0.031 when $ARL_0 > 200$. Besides, once IC ARL is fixed, increasing δ generates decreasing ARL_{min} while λ_{opt} increases in this situation, which satisfies the usual idea that the Poisson EWMA control chart with a large smoothing parameter is suitable for detecting large process shifts and that with smaller smoothing parameter will be more sensitive to smaller shifts.

5.4. Sensitivity analysis

After the optimal design of Poisson EWMA chart, we need further investigate how the parameters will affect performance of the control chart, i.e., sensitivity analysis. Here we first obtain the optimal parameters A and λ with Fibonacci search algorithm under the settings $\delta^* = 1, 3$, $\mu_0 = 10$ and IC ARL = 500. Here we use δ^* to represent the predetermined shift and δ to indicate the practical shift. Next, shifts are added to constant A (denoted by $\Delta A(\%)$) and smoothing parameter (denoted by $\Delta\lambda(\%)$). Third, we compute the varying percentage of corresponding IC ARL (with $\delta = 0$) and OC ARL ($\delta = 1,2,3,4$). It can be seen from Table 5 that both of IC ARL and OC ARL are positively correlated with A while they exhibit a negative correlation with smoothing parameter, which are in line with statements in Proposition 1. It is interesting that IC ARL is more sensitive to the shifts in A and smoothing parameter than OC ARL. For example, when control limit is encountered with -5% change and $\delta = 2$, IC and OC ARL undergo the changes of 30.4% and 6.2% respectively while the changes of IC and OC ARL are -10.6% and -4.4% respectively with smoothing parameter varying 15% and $\delta = 2$. Furthermore, both IC and OC ARL exhibit higher sensitivity to A than to smoothing parameter, which can be validated by the variation of IC and OC ARL at the 3% change of A and the smoothing parameter. This sensitivity analysis indicates that constant A should be carefully handled since it can strongly affect target IC ARL and designed minimal OC ARL.

5.5. Validating optimal design with unknown shift sizes

In this section, we further investigate the performance of proposed design method under the assumption that the specific magnitude of shift is unknown in advance. Note that simulation results are similar to those

Table 5
Sensitivity analysis of proposed Poisson EWMA chart with IC ARL = 500 and δ^*,δ representing the respective designed, actual shifts in mean $\mu_0 = 10$.

		ΔARL(%)									
		$\delta^* = 1, A = 2.451, \lambda = 0.031$					$\delta^* = 3, A = 2.905, \lambda = 0.141$				
	%	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$
ΔA	1.0	3.4	1.0	0.9	0.8	0.7	10.1	5.0	2.9	2.1	1.6
	2.0	10.7	3.4	2.4	2.1	1.9	17.4	8.6	5.0	3.5	2.8
	3.0	16.1	4.9	3.5	3.1	2.9	27.7	13.5	7.8	5.5	4.4
	4.0	24.8	7.4	5.1	4.4	4.1	41.3	19.6	11.1	7.7	6.1
	5.0	30.9	9.1	6.2	5.4	5.1	52.9	24.3	13.7	9.5	7.6
	−1.0	−6.3	−2.2	−1.4	−1.2	−1.1	−8.0	−4.0	−2.4	−1.7	−1.4
	−2.0	−10.3	−3.6	−2.6	−2.2	−2.1	−13.6	−7.1	−4.3	−3.0	−2.4
	−3.0	−16.9	−6.0	−4.2	−3.7	−3.5	−20.9	−11.2	−6.9	−5.0	−4.1
	−4.0	−19.9	−7.1	−5.1	−4.4	−4.2	−26.5	−14.4	−8.9	−6.5	−5.4
	−5.0	−24.1	−8.8	−6.3	−5.6	−5.3	–32.1	−17.8	−11.1	−8.2	−6.8
$\Delta\lambda$	3.0	−2.1	−0.7	−0.9	−1.1	−1.2	0.1	1.1	0.7	0.1	−0.2
	6.0	−5.0	−1.6	−2.0	−2.3	−2.4	−3.8	−0.1	0.0	−0.7	−1.2
	9.0	−6.6	−2.1	−2.7	−3.1	−3.3	−5.4	−0.1	−0.1	−1.1	−1.8
	12.0	−9.1	−2.9	−3.8	−4.2	−4.5	−7.3	−0.3	−0.2	−1.5	−2.5
	15.0	−10.6	−3.3	−4.4	−5.0	−5.2	−8.0	0.2	0.1	−1.6	−2.9
	−3.0	2.0	0.5	0.8	0.9	1.0	1.5	−0.1	0.0	0.3	0.6
	−6.0	4.0	1.2	1.8	2.1	2.3	4.9	0.6	0.4	1.0	1.5
	−9.0	8.6	2.5	3.2	3.5	3.7	8.1	1.1	0.9	1.8	2.5
	−12.0	11.7	3.4	4.4	4.8	5.0	9.4	0.7	0.7	2.1	3.0
	−15.0	16.6	4.9	6.2	6.7	7.0	10.7	0.1	0.6	2.5	3.8

Table 6
OC ARL comparison with δ^*,δ representing the respective designed and actual shifts in Poisson mean

	DPE chart			TPE chart		
δ	$\delta^* = 2$ $\lambda_{opt} = 0.088$	$\delta^* = 4$ $\lambda_{opt} = 0.212$	$\delta^* \in [2,4]$ $\lambda_{opt} = 0.139$	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.3$
1	54.3 (0.64)	66.4 (0.91)	58.8 (0.20)	54.8 (0.66)	64.8 (0.87)	75.6 (1.06)
2	18.6 (0.17)	20.5 (0.23)	19.1 (0.07)	18.9 (0.17)	20.3 (0.22)	23.8 (0.28)
3	10.6 (0.08)	10.4 (0.10)	10.2 (0.04)	10.4 (0.08)	10.4 (0.10)	11.2 (0.12)
4	7.37 (0.05)	6.72 (0.06)	6.90 (0.04)	7.21 (0.05)	6.65 (0.05)	6.82 (0.06)
5	5.70 (0.03)	4.93 (0.04)	5.20 (0.02)	5.51 (0.03)	4.98 (0.03)	4.82 (0.04)
−1	64.5 (0.77)	132 (1.78)	86.0 (0.75)	67.6 (0.84)	122 (1.64)	218 (2.96)
−2	19.5 (0.15)	28.9 (0.31)	21.5 (0.19)	19.9 (0.16)	27.3 (0.29)	42.3 (0.53)
−3	10.6 (0.06)	11.7 (0.10)	10.5 (0.08)	10.5 (0.06)	11.5 (0.09)	14.7 (0.15)
−4	7.19 (0.03)	6.85 (0.04)	6.85 (0.05)	7.04 (0.03)	6.81 (0.04)	7.55 (0.06)
−5	5.50 (0.02)	4.80 (0.02)	5.03 (0.03)	5.31 (0.02)	4.79 (0.02)	4.86 (0.03)

Note: Standard errors are in parentheses.

under situations with known shift sizes. Due to the space constraint, we only provide the results that verify the existence of optimal smoothing parameter such that EARL is minimized. In simulation study, the IC mean of Poisson distribution is set to $\mu = 10$ and the shift respectively follows $U(2,4)$ and $U(-4,-2)$ to formulate two OC cases. Note that we assume the estimated shift should be either positive or negative, which means the predefined uniform distribution $U(a,b)$ should satisfy the condition $ab \geq 0$. As mentioned in Section 4, optimal smoothing parameter may not exist when $ab < 0$. Fig. 6 shows that there is convex relationship between EARL and smoothing parameter λ in both cases. The optimal λ can be found through proposed Fibonacci search algorithm or grid search algorithm. Similar results can be obtained when shift size is assumed to follow other uniform distributions. It is feasible to minimize EARL for optimal design of Poisson EWMA chart. More simulation results are available upon request.

5.6. Monitoring performance comparison with existing control charts

This section conducts the comparison between the proposed optimally designed Poisson EWMA (DPE) chart and the traditional Poisson EWMA (TPE) chart. Due to page limits, only partial simulation results are listed in Table 6. Note that IC ARL is set to 370 and all OC ARLs are computed based on 5000 replications. As stated before, the optimal design of Poisson EWMA chart is to find a unique smoothing parameter in accordance with a predetermined magnitude or range of mean shift. With proposed Fibonacci search algorithm, we obtain three optimal smoothing parameters in three separate cases, i.e., $\delta^* = 2$, $\delta^* = 4$ and $\delta^* \in [2,4]$ for the DPE chart. Accordingly, $\lambda = 0.1, 0.2, 0.3$ are selected respectively for TPE chart. Table 6 shows that the monitoring performance of control charts are largely affected by the magnitude of smoothing parameter. Specifically, smaller smoothing parameter can help the charts detect smaller shifts more quickly whereas a larger smoothing parameter promotes quicker detection of larger shifts, which are basically in line with the conclusion from Lucas and Saccucci (1990). This conclusion usually guides us to choose appropriate λ for the TPE chart, which is quite rough and subjective. In contrast, the proposed DPE chart is advantageous due to its optimal design of λ according to the predetermined magnitude or range of mean shift.

6. Two illustrative examples

6.1. Monitoring packaging process of the printed circuit board (PCB)

In this section, the proposed DPE chart is illustrated in a packaging process of printed circuit board (PCB). According to Wang, Li, and Zhou (2017), in the plating stage of packaging process, some factors such as high temperature and humidity may separate the material layers of PCB to cause delamination, the number of which at different positions follow Poisson distributions independently in approximation. Hence, it is feasible to equip one Poisson EWMA control chart for the delamination monitoring at each position. Here we choose to monitor the number of delamination in MR lead position for illustrating the optimal design of Poisson EWMA chart. Wang et al. (2017) pointed out that the IC and OC mean can be separately estimated as $\mu_0 = 3.6$ and $\mu_1 = 5.5$. In this case, the process shift can be determined as $\delta = 1.9$ for the optimal design of

Poisson EWMA chart. Then Fibonacci search algorithm provides optimal smoothing parameter $\lambda = 0.167$ together with constant $A = 2.837$, which leads to an upper control limit (UCL) 5.225 and a lower control limit (LCL) 1.975. According to Wang et al. (2017), the shift $\delta = 1.9$ occur at the 16th sample while Fig. 7 shows that the OC signal is released at the 18th signal. The optimally designed Poisson EWMA chart is able to detect the process shift efficiently.

6.2. Monitoring the deaths caused by hepatitis C in Australia

SPC techniques have been widely utilized for public health surveillance. In this section, the optimally designed Poisson chart will be applied to monitor the deaths caused by Hepatitis C in South Australia. Hepatitis C is an infectious disease that may harm the livers of people to death. According to World Health Organization (2002), the infection of this disease is mainly through exposure to small quantities of blood. It is important to surveil its transmissions so that instant actions can be taken before an epidemic situation occurs. We use the monthly deaths caused by Hepatitis C in South Australia between the year of 2010 and 2014 provided by Pascual and Akhundjanov (2020). As a tradition in SPC, the first 24 observations are regarded as IC samples, which follow Poisson distribution with mean $\mu_0 = 3.167$ at the 5% significance level according to a Kolmogorov-Smirnov goodness-of-fit test.

The following data in 36 successive months are collected to formulate Phase II samples. Different from the first case, the mean shift of deaths is unknown in advance here. We may design the Poisson EWMA chart according to the shift size that should be avoided in practice. In this case, the shift is assumed to follow uniform distribution $U(0.9, 1.9)$, which indicates that we consider the shift exceeding 30% of the IC Poisson mean to be severe and it is unlikely to see Poisson mean increase by 60%. With the proposed procedure of minimizing EARL, the optimal smoothing parameter $\lambda = 0.098$ together with constant $A = 2.695$, which leads to UCL = 4.259 and LCL = 2.081. In line with Pascual and Akhundjanov (2020), we plot Fig. 8 with all the data in 60 months so that Phase II monitoring starts with the 25th sample. It shows that the designed Poisson EWMA chart released an OC signal in the 29th month. In other words, the deaths had increased significantly since May 2012, to which public health authorities should pay attention. Our findings are basically in accordance with the results from Pascual and Akhundjanov (2020).

7. Conclusions

This article focuses on the optimal design of Poisson EWMA control chart with shift information integrated to improve the change detection power. When the mean shift of Poisson distribution can be prespecified, the minimization of OC ARL is chosen as a criterion for searching for the optimal parameters of the control chart. The unique solution to the minimization problem is proved to exist by numerous simulation results and can be obtained based on proposed Fibonacci search algorithm. When the shift is unknown and randomly distributed, it will be no longer suitable to use ARL as a criterion and thus we choose to minimize the expectation of ARL for optimal design of the chart. Both of ARL and EARL are calculated based on the Markov chain approximation. The simulation results and two real monitoring examples illustrate the effectiveness of proposed optimal design methods. Despite this, there are several issues worthy of future efforts. Fibonacci search algorithm is compared with only bisection and grid search algorithm in obtaining the optimal parameters. Some other algorithms may be used to solve the optimization problem and make comparisons among them. This paper limits the discussion to optimal monitoring of a single Poisson variable, which may be extended to the multivariate case. In addition, the optimal design of a Poisson CUSUM chart is worthy of further investigation under the assumption that the shift is randomly distributed. Finally, one future research topic is to develop an adaptive EWMA chart for monitoring count data, in which the shift size is estimated from the currently available data and then the weighting parameter of the EWMA chart is adjusted by using the proposed method for choosing the optimal parameter. However, the computation of this problem would be intensive, since the optimal parameter needs to be updated at every observation time point. It is interesting for future research to develop the recursive computing formula for updating the optimal parameters to ease the computation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the Editor and three anonymous referees for their insightful comments that have helped improve the article significantly. This work was partly supported by the grant from National Natural Science Foundation of China under a Youth Project (No. 72002220) and Universiti Sains Malaysia (USM), Short Term Grant (No. 304/PMATHS/6315485).

Appendix

Here we introduce the bisection search algorithm to solve the following optimization problem:

$$\min_{\lambda} \mathrm{ARL}(\lambda, A(\lambda), \delta),$$

where $\lambda \in [\lambda_{min}, \lambda_{max}]$ and δ is prespecified according to domain knowledges or engineers' experiences. The algorithm procedures can be illustrated as below:

a. Set $\lambda_{min} = 0$, $\lambda_{max} = 1$, ARL_1=300, ARL_2=400 and$\mathrm{ARL}_3 = 370$;
b. Repeat the following steps for iteration $i = 1, 2, \cdots$ until $|\mathrm{ARL}_3 - \mathrm{ARL}_2| < 10^{-5}$:
(1) Let $\mathrm{ARL}_3 = \mathrm{ARL}_2$ and $\lambda = \frac{\lambda_{max}+\lambda_{min}}{2}$;
(2) Obtain the OC ARL, i.e., ARL_2 corresponding to the combination of δ and λ based on proposed Markov chain approximation;
(3) Let $\lambda_{max} = \lambda$, $\mathrm{ARL}_1 = \mathrm{ARL}_2$ if $\mathrm{ARL}_2 < \mathrm{ARL}_1$ and set $\lambda_{min} = \lambda$ otherwise.

Note that the practitioners can choose different stopping criterion and initial values of ARL_1, ARL_2 and ARL_3 to conduct the bisection algorithm.

References

Abbasi, S. A. (2017). Poisson progressive mean control chart. *Quality and Reliability Engineering International, 33*(8), 1855–1859.

Abujiya, M. A. R., Abbasi, S. A., & Riaz, M. (2016). A new EWMA control chart for monitoring Poisson observations. *Quality and Reliability Engineering International, 32*(8), 3023–3033.

Alevizakos, V., & Koukouvinos, C. (2020). A comparative study on Poisson control charts. *Quality Technology & Quantitative Management, 17*(3), 354–382.

Borror, C. M., Champ, C. W., & Rigdon, S. E. (1998). Poisson EWMA control charts. *Journal of Quality Technology, 30*(4), 352–361.

Bourguignon, M., Rodrigues, J., & Santos-Neto, M. (2019). Extended Poisson INAR (1) processes with equidispersion, underdispersion and overdispersion. *Journal of Applied Statistics, 46*(1), 101–118.

Capizzi, G., & Masarotto, G. (2003). An adaptive exponentially weighted moving average control chart. *Technometrics, 45*(3), 199–207.

Chiu, W. C., & Sheu, S. H. (2008). Fast initial response features for Poisson GWMA control charts. *Communications in Statistics—Simulation and Computation, 37*(7), 1422–1439.

Crowder, S. V. (1987). A simple method for studying run–length distributions of exponentially weighted moving average charts. *Technometrics, 29*(4), 401–407.

Gan, F. F. (1990). Monitoring Poisson observations using modified exponentially weighted moving average control charts. *Communications in Statistics-Simulation and Computation, 19*(1), 103–124.

Hassin, R., & Sarid, A. (2018). Operations research applications of dichotomous search. *European Journal of Operational Research, 265*(3), 795–812.

He, F., Shu, L. J., & Tsui, K. (2014). Adaptive CUSUM charts for monitoring linear drifts in Poisson rates. *International Journal of Production Economics, 148*, 14–20.

Huang, W., Shu, L., & Jiang, W. (2016). A gradient approach to the optimal design of CUSUM charts under unknown mean-shift sizes. *Journal of Quality Technology, 48*(1), 68–83.

Huang, W., Shu, L., & Jiang, W. (2018). A gradient approach to efficient design and analysis of multivariate EWMA control charts. *Journal of Statistical Computation and Simulation, 88*(14), 2707–2725.

Jiang, W., Shu, L., & Tsui, K. L. (2011). Weighted CUSUM control charts for monitoring Poisson processes with varying sample sizes. *Journal of Quality Technology, 43*(4), 346–362.

Li, C., Wang, D., & Zhu, F. (2019). Detecting mean increases in zero truncated INAR (1) processes. *International Journal of Production Research, 57*(17), 5589–5603.

Li, J., Liu, K., & Xian, X. (2017). Causation-based process monitoring and diagnosis for multivariate categorical processes. *IISE Transactions, 49*(3), 332–343.

Li, J., Tsung, F., & Zou, C. (2012). Directional control schemes for multivariate categorical processes. *Journal of Quality Technology, 44*(2), 136–154.

Lucas, J. M. (1985). Counted data CUSUM's. *Technometrics, 27*(2), 129–144.

Lucas, J. M., & Saccucci, M. S. (1990). Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics, 32*(1), 1–12.

Mahmood, T., & Xie, M. (2019). Models and monitoring of zero-inflated processes: The past and current trends. *Quality and Reliability Engineering International, 35*(8), 2540–2557.

Montgomery, D. C. (2013). *Introduction to statistical quality control* (7th ed.). New York, NY: Wiley.

Pascual, F., & Akhundjanov, S. B. (2020). Copula-based control charts for monitoring multivariate Poisson processes with application to hepatitis C counts. *Journal of Quality Technology, 52*(2), 128–144.

Qiu, P., He, Z., & Wang, Z. (2019). Nonparametric monitoring of multiple count data. *IISE Transactions, 51*(9), 972–984.

Qiu, P. (2014). *Introduction to Statistical Process Control*. Boca Raton, FL: Chapman & Hall/CRC.

Ryan, A. G., & Woodall, W. H. (2010). Control charts for Poisson count data with varying sample sizes. *Journal of Quality Technology, 42*(3), 260–275.

Saghir, A., & Lin, Z. (2015). Control charts for dispersed count data: An overview. *Quality and Reliability Engineering International, 31*(5), 725–739.

Sales, L. O., Pinho, A. L., Vivacqua, C. A., & Ho, L. L. (2020). Shewhart control chart for monitoring the mean of Poisson mixed integer autoregressive processes via Monte Carlo simulation. *Computers & Industrial Engineering, 140*, Article 106245.

Shen, X., Zou, C., Jiang, W., & 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.

Sheu, S. H., & Chiu, W. C. (2007). Poisson GWMA control chart. *Communications in Statistics—Simulation and Computation, 36*(5), 1099–1114.

Shu, L., Huang, W., & Jiang, W. (2014). A Novel Gradient Approach for Optimal Design and Sensitivity Analysis of EWMA Control Charts. *Naval Research Logistics, 61*(3), 223–237.

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

Testik, M. C. (2007). Conditional and marginal performance of the Poisson CUSUM control chart with parameter estimation. *International Journal of Production Research, 45*(23), 5621–5638.

Wang, K., & Jiang, W. (2009). High-dimensional process monitoring and fault isolation via variable selection. *Journal of Quality Technology, 41*(3), 247–258.

Wang, R., Li, J., & Xue, L. (2018). Joint monitoring of mean and dispersion of count data. *Journal of Industrial and Production Engineering, 35*(4), 269–276.

Wang, Z., Li, Y., & Zhou, X. (2017). A Statistical Control Chart for Monitoring High-dimensional Poisson Data Streams. *Quality and Reliability Engineering International, 33*(2), 307–321.

Wang, Z., & Qiu, P. (2018). Count data monitoring: Parametric or nonparametric? *Quality and Reliability Engineering International, 34*(8), 1763–1774.

Weiß, C. H. (2011). Detecting mean increases in Poisson INAR (1) processes with EWMA control charts. *Journal of Applied Statistics, 38*(2), 383–398.

Weiß, C. H., & Testik, M. C. (2012). Detection of abrupt changes in count data time series: Cumulative sum derivations for INARCH (1) models. *Journal of quality technology, 44*(3), 249–264.

Weiß, C. H., & Testik, M. C. (2015). Residuals-based CUSUM charts for Poisson INAR (1) processes. *Journal of Quality Technology, 47*(1), 30–42.

World Health Organization (WHO). (2002). Hepatitis C accessed February 22, 2020 https://www.who.int/en/news-room/fact-sheets/detail/hepatitis-C.

Yu, M., Wu, C., Wang, Z., & Tsung, F. (2018). A robust CUSUM scheme with a weighted likelihood ratio to monitor an overdispersed counting process. *Computers & Industrial Engineering, 126*, 165–174.

Zhang, L., Govindaraju, K., Lai, C. D., & Bebbington, M. S. (2003). Poisson DEWMA control chart. *Communications in Statistics-Simulation and Computation, 32*(4), 1265–1283.

Zhou, Q., Zou, C., Wang, Z., & 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.