



HAL
open science

A distribution-free EWMA control chart for monitoring time-between-events-and-amplitude data

Shu Wu, Philippe Castagliola, Giovanni Celano

► **To cite this version:**

Shu Wu, Philippe Castagliola, Giovanni Celano. A distribution-free EWMA control chart for monitoring time-between-events-and-amplitude data. *Journal of Applied Statistics*, 2021, 48 (3), pp.434-454. 10.1080/02664763.2020.1729347 . hal-03123025

HAL Id: hal-03123025

<https://hal.science/hal-03123025>

Submitted on 12 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Distribution-Free EWMA Control Chart for Monitoring Time-Between-Events-and-Amplitude Data

Shu Wu^a and Philippe Castagliola^b and Giovanni Celano^c

^a School of Logistics Engineering, Wuhan University of Technology, Wuhan, China, ^b Université de Nantes & LS2N UMR CNRS 6004, Nantes, France; ^c Università di Catania, Catania, Italy

ARTICLE HISTORY

Compiled July 12, 2021

ABSTRACT

During the recent years, many control charts have been developed for the simultaneous monitoring of the time interval T between successive occurrences of an event E and its magnitude X . All these TBEA (Time Between Events and Amplitude) control charts assume a known distribution for the random variables T and X (like, for instance, the normal, lognormal, gamma or Weibull distributions). But, in practice, as it is rather difficult to know the actual distribution of the time T or the amplitude X , proposing a distribution-free approach could be a way to overcome this “distribution choice” dilemma. For this reason, we propose in this paper a distribution-free upper-sided EWMA (Exponentially Weighted Moving Average) type control chart, for simultaneously monitoring the time interval T and the magnitude X of an event. In order to investigate the performance of this control chart and obtain its run length properties, we also develop a specific method called “continuousify” which, coupled with a classical Markov chain technique, allows to obtain reliable and replicable results. A numerical comparison shows that our distribution-free EWMA TBEA chart performs as the parametric Shewhart TBEA chart, but without the need to pre-specify any distribution. An illustrative example obtained from a French forest fire database is also provided to show the implementation of the proposed distribution-free EWMA TBEA control chart.

KEYWORDS

Distribution-Free; Markov Chain; Statistical Process Monitoring; Time Between Events and Amplitude

1. Introduction

Control charts are undeniably the most powerful tools in SPM (Statistical Process Monitoring) for improving the quality and the productivity of productions. But, during the recent years, several techniques based on control charts have also been developed to monitor processes in non-manufacturing sectors, such as in the health-care sector (like diseases, see Woodall [30]), the meteorological sector (like extreme weather or climate events, see Vučijak et al. [29]) or the geological sector (like earthquakes or volcanic eruptions, see Santiago and Smith [25], Ho [18]). In general, when a particular negative event E is of interest, two important characteristics should be recorded: the time T between two consecutive occurrences of this event and its amplitude X . In

most situations, a decrease in T and/or an increase in X will result in a negative, hazardous or disastrous consequence and, therefore, it has to be monitored with dedicated control charts called TBEA (Time Between Events and Amplitude) control charts.

It is worth recalling that Calvin [9] firstly proposed to monitor the cumulative number of conforming items between two nonconforming ones for improving the performance of traditional attribute control charts for monitoring high-quality processes. Lucas [19] and Vardeman and Ray [28] were at the origin of TBE (Time Between Events) control charts, as they were the first to propose new control charts with the idea of monitoring TBEs data for monitoring high-quality processes. Since then, several TBE control charts (both for phases I and II) have been proposed in the literature. For instance, the TBE exponential chart has been studied by Chan et al. [14] and Xie et al. [34]. Bourke [7] developed a geometric CUSUM chart for monitoring TBE data. Gan [17], Borrór et al. [6] and Shafae et al. [26] investigated an exponential TBE CUSUM (cumulative sum) control chart. A design procedure for TBE control charts with runs rules has been proposed by Cheng and Chen [15]. Qu et al. [22] studied some TBE control charts for sampling inspection. Readers can also refer to Shamsuzzaman et al. [27], Zhang et al. [36], Zhang et al. [35], Fang et al. [16].

In some applications, it is clearly important to monitor the time between events but also the amplitudes associated with these events. Recent enhancements to the statistical monitoring of an event E , not only quantified by its time T between two consecutive events but also by its amplitude X , have been introduced in the literature and they have been called TBEA (Time Between Events and Amplitude) control charts. Wu et al. [31] were the first to propose a combined T/X control chart based on a T chart to monitor the time interval and on a X chart to monitor the amplitude. From that moment, several other TBEA control charts have been developed, see Wu et al. [32], Wu et al. [33], Qu et al. [21], Qu et al. [23], Ali and Pievatolo [5], Rahali et al. [24].

All existing methods mentioned above for monitoring TBEA data are *parametric*, i.e. they assume that the distributions of both T and X are perfectly known. However, in many practical situations, the distributions of these random variables are unknown or their parameters cannot be correctly estimated by means of a Phase I retrospective study. This has been studied by Qiu [20], who has shown that using parametric control charts is not a reliable solution when the validity of the distribution is in question. In the specific case of monitoring TBEA data, Rahali et al. [24] investigated the use of *parametric* approaches in order to monitor “fires in forests”. They experienced difficulties in selecting the most suitable distribution for the time T and the amplitude X due the limited number of Phase I data and the small differences in terms of the statistic (Kolmogorov-Smirnov) measuring the quality of the fit (making several distributions being a possible candidate, such as the gamma, lognormal, normal and Weibull distributions). It is well known that a fitting error on the distribution of observations can result in a poor in-control performance of the control chart. To overcome this problem, *distribution-free* control charts have been investigated in the literature. Among the most recent ones, we can cite Celano et al. [11] who investigated the statistical performance of a Shewhart sign control chart in a process with a finite production horizon, Abid et al. [1] who proposed a nonparametric EWMA control chart based on the sign test using RSS (Ranked Set Sampling), Castagliola et al. [10] who proposed a new Phase II EWMA-type chart for count data based on

the sign statistic, Abid et al. [2] who introduced a nonparametric EWMA control chart based on the Wilcoxon signed-rank statistic using ranked set sampling, Abid et al. [3, 4] who suggested nonparametric CUSUM sign and Wilcoxon signed-rank control charts for monitoring and detecting possible deviations from the process mean using ranked set sampling. Interested readers can find a comprehensive discussion in the very recent review of Chakraborti and Graham [13] who discuss many Phase I and Phase II distribution-free control charts and give some suggestions for next research directions. Practical guidelines for the distribution-free control charts implementation can be found in the books of Qiu [20] and Chakraborti and Graham [12].

But, as far as we know, no research has been conducted so far on proposing a distribution-free control chart for monitoring TBEA data. This will be the main and first goal of this paper where a new upper-sided distribution-free EWMA control chart for monitoring TBEA data will be introduced. As evaluating Run Length related values (ARL, SDR, ...) for an EWMA scheme based on *discrete data* is a challenging problem, the second goal of this paper will consist in proposing a dedicated method, called “continuousify”, which allows reliable and replicable results to be obtained.

The structure of this paper is as follows. In Section 2, the distribution-free statistic to be monitored as well as the “continuousify” technique used for computing the Run Length properties are both introduced. Then, in Section 3 the optimal design of the proposed distribution-free chart is presented. In Section 4 a comparison with parametric Shewhart TBEA charts proposed in Rahali et al. [24] is performed. A real case example implementing the new upper-sided distribution-free EWMA TBEA control chart is presented in Section 5 and conclusions and future researches are given in Section 6.

2. A distribution-free EWMA TBEA control chart

Let $D_0 = 0, D_1, D_2, \dots$ be the dates of occurrence of a specific negative event E, let $T_1 = D_1 - D_0, T_2 = D_2 - D_1, \dots$ be the time intervals between two consecutive occurrences of the event E and let X_1, X_2, \dots be the corresponding magnitudes of this event occurring at times D_1, D_2, \dots and assumed to be independent of T_1, T_2, \dots (see Figure 1). It must be noted that $D_0 = 0$ is the date of a “virtual” event which has no amplitude associated with.

Let $F_T(t|\theta_T)$ and $F_X(x|\theta_X)$ be the *unknown* continuous c.d.f. (cumulative distribution functions) of T_i and $X_i, i = 1, 2, \dots$, where θ_T and θ_X are *known* α -quantiles, respectively. More precisely, when the process is in-control, we have $\theta_T = \theta_{T_0}, \theta_X = \theta_{X_0}$ and, when the process is out-of-control, we have $\theta_T = \theta_{T_1}, \theta_X = \theta_{X_1}$. Without loss of generality, we will consider in this paper that θ_T and θ_X are the median values (i.e. the 0.5-quantiles) of T_i and X_i , respectively. Other α -quantiles can be considered based on the investigated event’s severity of consequences.

Let $p_T = P(T_i > \theta_{T_0}|\theta_T) = 1 - F_T(\theta_{T_0}|\theta_T)$ and $p_X = P(X_i > \theta_{X_0}|\theta_X) = 1 - F_X(\theta_{X_0}|\theta_X), i = 1, 2, \dots$, be the probabilities that T_i and X_i are larger than θ_{T_0} and θ_{X_0} assuming that the actual median values are θ_T and θ_X , respectively. If the process is in-control, we have $p_T = p_{T_0} = 1 - F_T(\theta_{T_0}|\theta_{T_0}) = 0.5, p_X = p_{X_0} = 1 - F_X(\theta_{X_0}|\theta_{X_0}) = 0.5$ and, when the process is out-of-control, we have

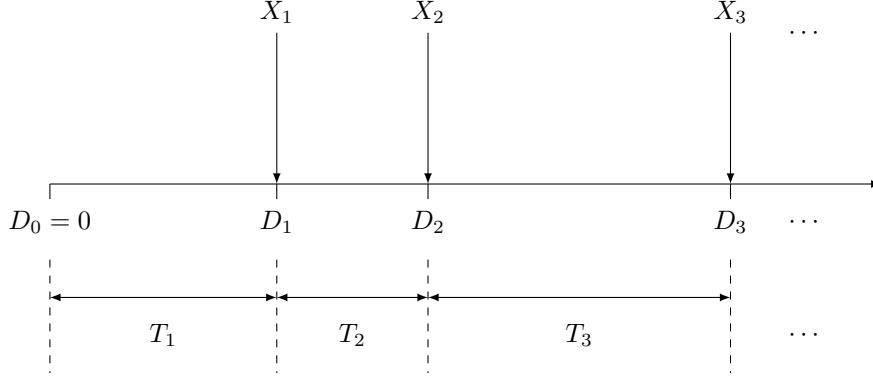


Figure 1. Times of occurrence D_i , time intervals T_i and amplitudes X_i of a negative event E

$$p_T = p_{T_1} = 1 - F_T(\theta_{T_0} | \theta_{T_1}), \quad p_X = p_{X_1} = 1 - F_X(\theta_{X_0} | \theta_{X_1}).$$

Let us define the statistics ST_i and SX_i , for $i = 1, 2, \dots$ as

$$\begin{aligned} ST_i &= \text{sign}(T_i - \theta_{T_0}), \\ SX_i &= \text{sign}(X_i - \theta_{X_0}), \end{aligned}$$

where $\text{sign}(x) = -1$ if $x < 0$ and $\text{sign}(x) = +1$ if $x > 0$. Because T_i and X_i are assumed to be continuous random variables, the unlikely case $x = 0$ will not be considered in the definition of the charting statistic. Nevertheless, as sometimes ties may occur in practice (for example, as a consequence of the selected time unit for T_i), we will explain in the “Illustrative example” section how this situation can be handled by practitioners without significantly affecting the final result. In order to simultaneously monitor, *in a distribution-free way*, the time interval T_i between consecutive occurrences of the event E and its magnitude X_i , we suggest to define the statistic S_i , for $i = 1, 2, \dots$ as

$$S_i = \frac{SX_i - ST_i}{2}.$$

By definition, we have $S_i \in \{-1, 0, +1\}$ and, more precisely, we have:

- $S_i = -1$ when the process is in an *acceptable* situation, i.e. when T_i increases ($ST_i = +1$) and, at the same time, X_i decreases ($SX_i = -1$).
- $S_i = +1$ when the process is in an *unacceptable* situation, i.e. when T_i decreases ($ST_i = -1$) and, at the same time, X_i increases ($SX_i = +1$).
- $S_i = 0$ when the process is in an *intermediate* situation, i.e. when both T_i and X_i increase or when both T_i and X_i decrease.

It is easy to prove that the p.m.f. (probability mass function) $f_{S_i}(s | p_T, p_X) = P(S_i = s | p_T, p_X)$ of S_i is equal to

$$f_{S_i}(s | p_T, p_X) = \begin{cases} p_T q_X & \text{if } s = -1 \\ p_T p_X + q_T q_X & \text{if } s = 0 \\ q_T p_X & \text{if } s = +1 \\ 0 & \text{if } s \notin \{-1, 0, 1\} \end{cases},$$

where $q_T = 1 - p_T$, $q_X = 1 - p_X$ and its c.d.f. $F_{S_i}(s|p_T, p_X) = P(S_i \leq s|p_T, p_X)$ is equal to

$$F_{S_i}(s|p_T, p_X) = \begin{cases} 0 & \text{if } s \in (-\infty, -1) \\ p_T q_X & \text{if } s \in [-1, 0) \\ p_T + q_T q_X & \text{if } s \in [0, 1) \\ 1 & \text{if } s \in [1, +\infty) \end{cases}.$$

In practice, it is actually *always* possible to define and implement an EWMA TBEA type control chart directly monitoring the statistic S_i using an equation like $Z_i = \lambda S_i + (1 - \lambda)Z_{i-1}$, where $\lambda \in [0, 1]$ is some smoothing parameter to be fixed and $Z_0 = 0$. The problem of this approach is that, because of the discrete nature of the random variable S_i , it is impossible to *accurately* compute (using Markov chain or integral equation methods, for instance) the run length properties (average run length ARL and standard deviation of the run length SDRL) of such a control chart and, therefore, it is impossible to tune the chart parameters in order to obtain a predefined in-control performance. If, for instance, the Markov chain approach, (as detailed hereafter), is used in order to compute the ARL or the SDRL, the results will i) *heavily* fluctuate depending on the value of the selected number m of subintervals and ii) not exhibit any monotonic convergence when m increases, making useless such an approach. This point will be highlighted at the end of this section. Of course, it is always possible to obtain these values using simulations but, even in this case, if it is quite easy to compute *small* ARL or SDRL values with some precision, it becomes just impossible to obtain reliable results when these values become very large.

Since the Markov chain and integral equation methods give good results in the case of *continuous* random variables, (and more particularly in the case of the normal distribution, which is an unbounded one), we therefore suggest to transform the discrete random variable S_i into a new continuous one, denoted as S_i^* , (say that we “continuousify” the random variable S_i), and to monitor it using a traditional EWMA scheme. We suggest to define the statistic S_i^* as a mixture of 3 normal random variables $Y_{i,-1} \sim \text{Nor}(-1, \sigma)$, $Y_{i,0} \sim \text{Nor}(0, \sigma)$ and $Y_{i,+1} \sim \text{Nor}(+1, \sigma)$, with weights $w_{-1} = p_T q_X$, $w_0 = p_T p_X + q_T q_X$ and $w_{+1} = q_T p_X$ (corresponding to the probabilities $f_{S_i}(s|p_T, p_X)$, $s \in \{-1, 0, +1\}$), respectively, i.e.

$$S_i^* = \begin{cases} Y_{i,-1} & \text{if } S_i = -1, \\ Y_{i,0} & \text{if } S_i = 0, \\ Y_{i,+1} & \text{if } S_i = +1. \end{cases}$$

Concretely speaking, this means that if, at $i = 1, 2, \dots$, we have $S_i = s \in \{-1, 0, +1\}$ then, in order to obtain S_i^* , we just have to generate a $\text{Nor}(s, \sigma)$ random number to “continuousify” the random variable S_i . The fact that random numbers have to be generated does not imply that the Run Length properties (ARL, SDRL, ...) are obtained using simulations. As shown below, the Run Length properties of the upper-sided distribution-free EWMA TBEA control chart are obtained with an exact Markov chain based method. But, in order to use this approach, it is necessary to assume that the discrete random variables S_i have been transformed into continuous ones. This is why random numbers are generated. The parameter $\sigma > 0$ has to be fixed and, as it will be shown later, its value does not significantly affect the performance of the control chart as long it is neither too small nor too large. Since S_i^* is defined as a

mixture of normal distributions, its c.d.f. $F_{S_i^*}(s|p_T, p_X) = P(S_i^* \leq s|p_T, p_X)$ is equal to

$$F_{S_i^*}(s|p_T, p_X) = p_T q_X F_{\text{Nor}}(s| -1, \sigma) + (p_T p_X + q_T q_X) F_{\text{Nor}}(s|0, \sigma) + q_T p_X F_{\text{Nor}}(s| +1, \sigma) \quad (1)$$

where $F_{\text{Nor}}(s|\mu, \sigma)$ is the c.d.f. of the normal $\text{Nor}(\mu, \sigma)$ distribution. As an example, we plotted in Figure 2 the p.d.f. of S_i^* when the process is in-control, (i.e. the weights are $w_{-1} = 0.25$, $w_0 = 0.5$ and $w_{+1} = 0.25$), for $\sigma \in \{0.1, 0.125, 0.15, 0.2\}$. We suggest to only investigate this range of values as it seems that when $\sigma < 0.1$ the “peaks” around $\{-1, 0, +1\}$ become too sharp and when $\sigma > 0.2$ these “peaks” become too smooth. It is not difficult to demonstrate that the expectance $E(S_i^*)$ and variance $V(S_i^*)$ of S_i^* are equal to

$$\begin{aligned} E(S_i^*) &= p_X - p_T, \\ V(S_i^*) &= \sigma^2 + p_T q_T + p_X q_X. \end{aligned}$$

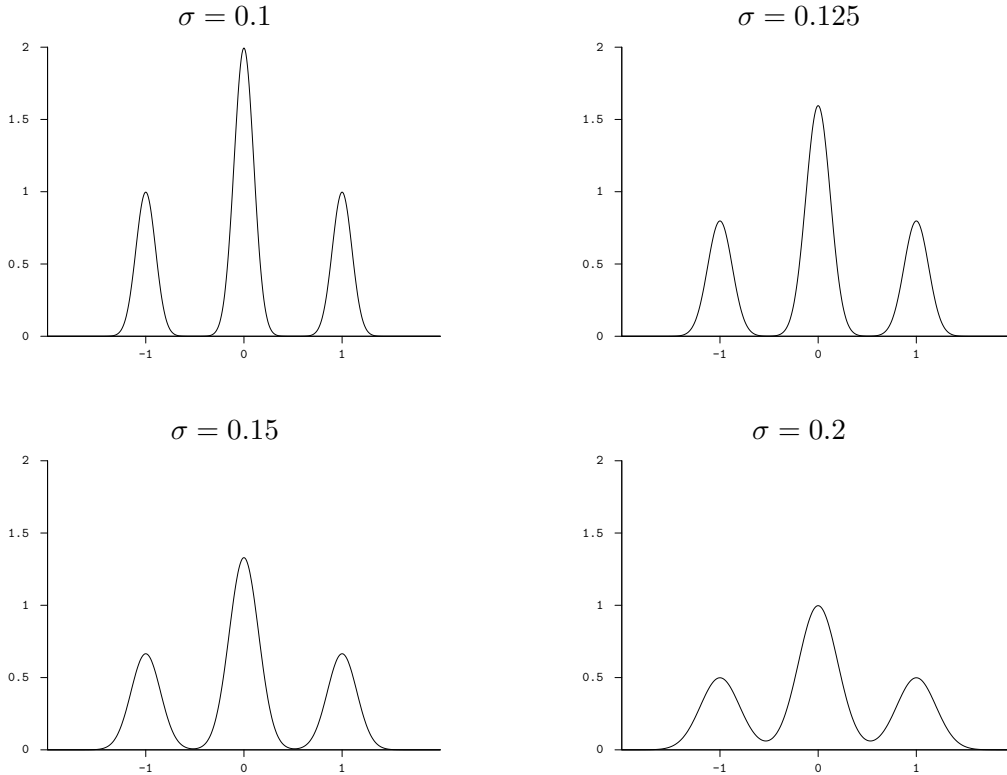


Figure 2. p.d.f. of S_i^* when the process is in-control for $\sigma \in \{0.1, 0.125, 0.15, 0.2\}$

The in- and out-of-control c.d.f., expectance and variance of S_i^* can be simply obtained by replacing, in the previous equations, p_T and p_X by either p_{T_0} and p_{X_0} or p_{T_1} and p_{X_1} , respectively. In particular, if the process is in-control, we have $p_{T_0} = q_{T_0} = 0.5$, $p_{X_0} = q_{X_0} = 0.5$ and the expectance and variance of S_i^* simplify to $E(S_i^*) = 0$ and $V(S_i^*) = \sigma^2 + 0.5$.

As it is more important to detect an increase in S_i or S_i^* (in order to avoid more damages or injuries/costs, for instance) rather than a decrease, we suggest to define the following upper-sided EWMA TBEA control chart based on the statistic

$$Z_i^* = \max(0, \lambda S_i^* + (1 - \lambda)Z_{i-1}^*), \quad (2)$$

with the following upper asymptotic control limit UCL defined as

$$\text{UCL} = \underbrace{\text{E}(S_i^*)}_{=0} + K \sqrt{\frac{\lambda}{2 - \lambda}} \times \underbrace{\sqrt{\text{V}(S_i^*)}}_{=\sqrt{\sigma^2 + 0.5}} = K \sqrt{\frac{\lambda(\sigma^2 + 0.5)}{2 - \lambda}}, \quad (3)$$

where $\lambda \in [0, 1]$ and $K > 0$ are the control chart parameters to be fixed and the initial value $Z_0^* = 0$.

In order to obtain the zero-state ARL and SDRL of the proposed distribution-free upper-sided EWMA TBEA control chart, we suggest to use the standard approach proposed by Brook and Evans [8], which assumes that the behavior of this control chart can be well represented by a discrete-time Markov chain with $m + 2$ states, where states $i = 0, 1, \dots, m$ are transient and state $m + 1$ is an absorbing one. The transition probability matrix \mathbf{P} of this discrete-time Markov chain is

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{r} \\ \mathbf{0}^\top & 1 \end{pmatrix} = \begin{pmatrix} Q_{0,0} & Q_{0,1} & \cdots & Q_{0,m} & r_0 \\ Q_{1,0} & Q_{1,1} & \cdots & Q_{1,m} & r_1 \\ \vdots & \vdots & & \vdots & \vdots \\ Q_{m,0} & Q_{m,1} & \cdots & Q_{m,m} & r_m \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

where \mathbf{Q} is the $(m + 1, m + 1)$ matrix of transient probabilities, where $\mathbf{0} = (0, 0, \dots, 0)^\top$ and where the $(m + 1, 1)$ vector \mathbf{r} satisfies $\mathbf{r} = \mathbf{1} - \mathbf{Q}\mathbf{1}$ (i.e. row probabilities must sum to 1) with $\mathbf{1} = (1, 1, \dots, 1)^\top$. The transient states $i = 1, \dots, m$ are obtained by dividing the interval $[0, \text{UCL}]$ into m subintervals of width 2Δ , where $\Delta = \frac{\text{UCL}}{2m}$. By definition, the midpoint of the i -th subinterval (representing state i) is equal to $H_i = (2i - 1)\Delta$. The transient state $i = 0$ corresponds to the “restart state” feature of our chart (due to the presence of the $\max(\dots)$ in (2)). This state is represented by the value $H_0 = 0$. Concerning the proposed upper-sided EWMA TBEA control chart, it can be easily proven that the generic element $Q_{i,j}$, $i = 0, 1, \dots, m$, of the matrix \mathbf{Q} is equal to

- if $j = 0$,

$$Q_{i,0} = F_{S_i^*} \left(-\frac{(1 - \lambda)H_i}{\lambda} \middle| p_T, p_X \right), \quad (4)$$

- if $j = 1, 2, \dots, m$,

$$Q_{i,j} = F_{S_i^*} \left(\frac{H_j + \Delta - (1 - \lambda)H_i}{\lambda} \middle| p_T, p_X \right) - F_{S_i^*} \left(\frac{H_j - \Delta - (1 - \lambda)H_i}{\lambda} \middle| p_T, p_X \right) \quad (5)$$

Let $\mathbf{q} = (q_0, q_1, \dots, q_m)^\top$ be the $(m + 1, 1)$ vector of initial probabilities associated with the $m + 1$ transient states. In our case, we assume $\mathbf{q} = (1, 0, \dots, 0)^\top$, i.e. the

initial state corresponds to the “restart state”. When the number m of subintervals is sufficiently large (say $m = 300$), this finite approach provides an effective method that allows the ARL and SDRL to be accurately evaluated using the following classical formulas

$$\text{ARL}(p_X, p_T) = \mathbf{q}^\top (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}, \quad (6)$$

$$\text{SDRL}(p_X, p_T) = \sqrt{2\mathbf{q}^\top (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q} \mathbf{1} + \text{ARL}(1 - \text{ARL})}. \quad (7)$$

In order to clearly illustrate, for the proposed upper-sided EWMA TBEA control chart, the difference between using or not the suggested “continuousify” technique, we present in Table 1 the ARL values obtained for several combinations of (p_X, p_T) , $m \in \{100, 120, \dots, 400\}$ and $\sigma = 0.125$. In Table 1, we also provide ARL values obtained by simulations (last row of Table 1). Based on Table 1, the following conclusions can be drawn:

- when the “continuousify” technique is not used, (see the left side of Table 1, denoted as “without continuousify”), the ARL values obtained using the Markov chain method have a large variability with m ; furthermore, they do not show any visible monotonic convergence when m increases. The worst case is for $(p_X = 0.7, p_T = 0.4)$ for which some ARL values are even negative! This phenomenon is known to happen even in the case of continuous random variables when the smoothing parameter λ is too small. In this case, the Markov chain approach does not converge and provide meaningless (i.e. either negative or too large) ARL values. The fact that the random variables are discrete makes this phenomenon even stronger due to the fact that the probabilities in (4) and (5) are not necessarily continuous / smooth. For the remaining combinations (p_X, p_T) the fluctuation is noticeable with a particularity for $m = 260$ which gives (for some unclear reason) larger ARL values, if compared to the others.
- when the “continuousify” technique is used, (see the right side of Table 1, denoted as “with continuousify”), the ARL values obtained using the Markov chain method are very stable, even for small values of $m \in \{100, \dots, 150\}$. The ARL values obtained with this technique are a bit larger than those obtained by simulation “without continuousify” (for instance compare the values 87.24, 26.08, 12.23, 27.88 with the results 84.46, 24.71, 11.66, 26.46 obtained for $m = 400$). This is logical as the control limits “with continuousify” are a bit larger than those “without continuousify” due to the extra term $\sigma > 0$ in (3).

3. Numerical analysis

The goal of this section is twofold:

- (1) obtaining optimal values (λ^*, K^*) for the upper-sided EWMA TBEA control chart parameters (λ, K) as to minimize the out-of-control $\text{ARL}(\lambda^*, K^*, \sigma, p_T, p_X)$ for $p_T \neq 0.5$ and $p_X \neq 0.5$ under the constraint $\text{ARL}(\lambda^*, K^*, \sigma, 0.5, 0.5) = \text{ARL}_0$, where ARL_0 is a predefined value for the in-control ARL;
- (2) demonstrating that the choice of the parameter σ does not significantly impact the out-of-control performance of this chart as long as this value is not too small nor too large.

The optimal values for (λ^*, K^*) are listed in Table 2 with the corresponding out-of-

Table 1. ARL for the distribution-free EWMA TBEA chart computed with and without the “continuousify” technique.

m	“without continuousify”				“with continuousify” ($\sigma = 0.125$)			
	$p_T = 0.4$ $p_X = 0.7$	$p_T = 0.3$ $p_X = 0.8$	$p_T = 0.2$ $p_X = 0.9$	$p_T = 0.1$ $p_X = 0.6$	$p_T = 0.4$ $p_X = 0.7$	$p_T = 0.3$ $p_X = 0.8$	$p_T = 0.2$ $p_X = 0.9$	$p_T = 0.1$ $p_X = 0.6$
100	-10629.23	32.96	12.18	40.16	87.22	26.08	12.23	27.87
120	35.68	18.57	10.89	18.55	87.23	26.08	12.23	27.87
140	579.56	28.88	11.91	31.56	87.23	26.08	12.23	27.87
160	42.43	20.31	11.08	20.81	87.23	26.08	12.23	27.87
180	286.68	28.36	11.82	31.36	87.24	26.08	12.23	27.87
200	77.24	24.47	11.55	26.42	87.24	26.08	12.23	27.87
220	33.76	17.97	10.05	18.39	87.24	26.08	12.23	27.87
240	108.67	27.98	11.88	31.36	87.24	26.08	12.23	27.87
260	-174.46	57.68	16.41	77.81	87.24	26.08	12.23	27.87
280	53.94	21.17	11.41	21.75	87.24	26.08	12.23	27.87
300	122.35	26.75	11.75	29.94	87.24	26.08	12.23	27.88
320	45.08	21.69	11.43	22.43	87.24	26.08	12.23	27.88
340	179.06	26.07	11.93	27.39	87.24	26.08	12.23	27.88
360	29.81	16.68	10.43	16.5	87.24	26.08	12.23	27.88
380	1630.32	29.2	13.09	29.83	87.24	26.08	12.23	27.88
400	53.38	20.33	11.02	20.7	87.24	26.08	12.23	27.88
Simu	84.46	24.71	11.66	26.46	87.23	26.09	12.23	27.87

control values of (ARL, SDRL) for $p_T \in \{0.1, 0.2, \dots, 0.4\}$ (as we are only interested in a decrease in T), $p_X \in \{0.5, 0.6, \dots, 0.9\}$ (as we are only interested in an increase in X), for four possible choices for $\sigma \in \{0.1, 0.125, 0.15, 0.2\}$ and assuming $ARL_0 = 370.4$. For instance, in Table 2, when $\sigma = 0.125$, $p_T = 0.4$ and $p_X = 0.6$ the optimal chart parameters are $(\lambda^*, K^*) = (0.025, 2.174)$ and the corresponding values for the out-of-control (ARL, SDRL) are $ARL = 51.11$ and $SDRL = 32.63$. From Table 2 we can draw the following conclusions:

- No matter the value of σ , when $p_T = p_X = 0.5$ we *exactly* obtain $ARL = ARL_0 = 370.4$ (as expected). In this case, it exists an infinite number of couples (λ^*, K^*) *exactly* satisfying the constraint $ARL = ARL_0 = 370.4$. These couples are denoted with “(-,-)”. It has to be noted that without the “continuousify” technique used in this paper, it would have been impossible to exactly obtain $ARL = ARL_0 = 370.4$ due to the discrete nature of the random variable S_i .
- No matter the value of σ , the out-of-control ARL values monotonically decrease when the values of p_T decrease and/or the values of p_X increase. Due to the symmetry of ST_i and SX_i in the definition of the random variable S_i , the performance of the distribution-free upper-sided EWMA TBEA chart is the same for any combination of $(p_T = \alpha_T, p_X = \alpha_X)$ or $(p_T = 1 - \alpha_X, p_X = 1 - \alpha_T)$ where α_T and α_X are two probabilities in $[0, 1]$. For this reason, only the lower side of each table is presented, being the upper side immediately be derived by symmetry. For example, if $\sigma = 0.125$, the optimal parameters (λ^*, K^*) and corresponding out-of-control ARL and SDRL for $p_T = 0.4$ and $p_X = 0.7$ are the same as the ones for $p_T = 0.3$ and $p_X = 0.6$, i.e. $(\lambda^* = 0.045, K^* = 2.387)$, $ARL = 30.79$ and $SDRL = 18.25$.
- As long as $\sigma \in \{0.1, 0.125, 0.15, 0.2\}$, the optimal design parameters (λ^*, K^*) and the out-of-control ARL and SDRL values are almost the same. For instance, if $p_T = 0.3$ and $p_X = 0.6$, then the optimal parameters are $(\lambda^* = 0.045, K^* = 2.387)$ (irrespective of the value of σ) and the out-of-control ARL and SDRL values are $(ARL = 30.55, SDRL = 18.04)$, $(ARL = 30.79, SDRL = 18.25)$, $(ARL = 31.08, SDRL = 18.51)$ and $(ARL = 31.81, SDRL = 19.15)$ when

Table 2. Optimal values for (λ^*, K^*) with the corresponding out-of-control values of (ARL, SDRL) for $p_T \in \{0.1, 0.2, \dots, 0.4\}$, $p_X \in \{0.5, 0.6, \dots, 0.9\}$ and $\sigma \in \{0.1, 0.125, 0.15, 0.2\}$

$\sigma = 0.1$					
p_T	0.5	0.6	p_X 0.7	0.8	0.9
0.5	(-, -) (370.40, -)				
0.4	(0.010, 1.773) (105.66, 74.04)	(0.025, 2.174) (50.77, 32.32)			
0.3	(0.025, 2.174) (51.54, 32.55)	(0.045, 2.387) (30.55, 18.04)	(0.070, 2.515) (20.50, 11.38)		
0.2	(0.040, 2.348) (31.30, 17.51)	(0.070, 2.515) (20.74, 11.40)	(0.100, 2.591) (14.85, 7.67)	(0.145, 2.639) (11.19, 5.55)	
0.1	(0.060, 2.474) (21.40, 10.76)	(0.090, 2.571) (15.16, 7.37)	(0.135, 2.634) (11.32, 5.40)	(0.180, 2.645) (8.76, 3.84)	(0.240, 2.627) (6.99, 2.74)

$\sigma = 0.125$					
p_T	0.5	0.6	p_X 0.7	0.8	0.9
0.5	(-, -) (370.40, -)				
0.4	(0.010, 1.774) (106.19, 74.55)	(0.025, 2.174) (51.11, 32.63)			
0.3	(0.025, 2.174) (51.88, 32.87)	(0.045, 2.387) (30.79, 18.25)	(0.070, 2.515) (20.68, 11.53)		
0.2	(0.040, 2.348) (31.53, 17.72)	(0.065, 2.496) (20.91, 11.27)	(0.100, 2.592) (14.99, 7.80)	(0.140, 2.638) (11.32, 5.57)	
0.1	(0.060, 2.474) (21.57, 10.92)	(0.090, 2.572) (15.30, 7.50)	(0.135, 2.634) (11.44, 5.49)	(0.175, 2.648) (8.88, 3.89)	(0.225, 2.639) (7.10, 2.75)

$\sigma = 0.15$					
p_T	0.5	0.6	p_X 0.7	0.8	0.9
0.5	(-, -) (370.40, -)				
0.4	(0.010, 1.775) (106.83, 75.16)	(0.025, 2.175) (51.53, 33.01)			
0.3	(0.025, 2.175) (52.30, 33.26)	(0.045, 2.387) (31.08, 18.51)	(0.070, 2.515) (20.90, 11.71)		
0.2	(0.040, 2.348) (31.82, 17.97)	(0.065, 2.496) (21.13, 11.45)	(0.095, 2.584) (15.17, 7.81)	(0.135, 2.636) (11.47, 5.61)	
0.1	(0.055, 2.449) (21.79, 10.76)	(0.090, 2.573) (15.47, 7.64)	(0.130, 2.632) (11.59, 5.53)	(0.170, 2.651) (9.02, 3.96)	(0.215, 2.646) (7.23, 2.80)

$\sigma = 0.2$					
p_T	0.5	0.6	p_X 0.7	0.8	0.9
0.5	(-, -) (370.40, -)				
0.4	(0.010, 1.777) (108.43, 76.68)	(0.025, 2.176) (52.57, 33.96)			
0.3	(0.020, 2.085) (53.33, 32.51)	(0.045, 2.387) (31.81, 19.15)	(0.065, 2.496) (21.44, 11.90)		
0.2	(0.040, 2.348) (32.53, 18.61)	(0.065, 2.496) (21.66, 11.92)	(0.090, 2.574) (15.60, 8.01)	(0.125, 2.630) (11.84, 5.74)	
0.1	(0.055, 2.449) (22.31, 11.21)	(0.085, 2.562) (15.89, 7.83)	(0.120, 2.624) (11.95, 5.64)	(0.155, 2.652) (9.34, 4.06)	(0.195, 2.658) (7.53, 2.91)

$\sigma \in \{0.1, 0.125, 0.15, 0.2\}$, respectively.

4. Comparative studies

The goal of this section is to compare the proposed upper-sided distribution-free EWMA TBFA chart with the three *parametric* Shewhart type control charts introduced in Rahali et al. [24] based on statistics Z_1 , Z_2 and Z_3 . It is important to note that these statistics depend on standardized versions $X' = \frac{X}{\mu_{X_0}}$ and $T' = \frac{T}{\mu_{T_0}}$ of X and T , respectively, where μ_{X_0} and μ_{T_0} are the in-control mean values for X and T . The 2-parameters distributions considered in Rahali et al. [24] were i) the gamma, lognormal, normal and Weibull distributions for the amplitude X and ii) the gamma, lognormal and Weibull distributions for the time between events T , leading to a combination of 11 scenarios. For more details concerning the definition of statistics Z_1 , Z_2 and Z_3 and the parametrization of these distributions, do refer to Rahali et al. [24]. In this paper, we will only investigate two scenarios:

- Scenario #1: a Normal distribution for X with in-control mean $\mu_{X_0} = 10$ and standard-deviation $\sigma_{X_0} = 1$ and a gamma distribution for T with in-control mean $\mu_{T_0} = 10$ and standard-deviation $\sigma_{T_0} = 2$, i.e. $X \sim \text{Nor}(10, 1)$ and $T \sim \text{Gam}(25, 0.4)$.
- Scenario #2: a Normal distribution for X with in-control mean $\mu_{X_0} = 10$ and standard-deviation $\sigma_{X_0} = 2$ and a Weibull distribution for T with in-control mean $\mu_{T_0} = 10$ and standard-deviation $\sigma_{T_0} = 1$, i.e. $X \sim \text{Nor}(10, 2)$ and $T \sim \text{Wei}(12.1534, 10.4304)$.

In this parametric framework, when an upper shift is occurring, it can be due to: i) either a mean shift in the amplitude X from μ_{X_0} to $\mu_{X_1} = \delta_X \mu_{X_0}$ where $\delta_X \geq 1$ is the parameter quantifying the change in the amplitude, ii) or a mean shift in the time T from μ_{T_0} to $\mu_{T_1} = \delta_T \mu_{T_0}$ where $\delta_T \leq 1$ is the parameter quantifying the change in the time, ii) or also a change in *both* the amplitude X from μ_{X_0} to $\mu_{X_1} = \delta_X \mu_{X_0}$ and the time T from μ_{T_0} to $\mu_{T_1} = \delta_T \mu_{T_0}$.

As it is usually very difficult to know the actual values of δ_X and δ_T , we will use the Expected Average Run Length (EARL) criterion and, more particularly:

- the EARL_X for X (assuming $\delta_T = 1$) defined as:

$$\text{EARL}_X = \sum_{\delta_X \in \Omega_X} f_{\delta_X}(\delta_X) \text{ARL}(\delta_X, 1)$$

- the EARL_T for T (assuming $\delta_X = 1$) defined as:

$$\text{EARL}_T = \sum_{\delta_T \in \Omega_T} f_{\delta_T}(\delta_T) \text{ARL}(1, \delta_T)$$

- the EARL_{XT} for both of X and T defined as:

$$\text{EARL}_{XT} = \sum_{\delta_X \in \Omega_X} \sum_{\delta_T \in \Omega_T} f_{\delta_X}(\delta_X) f_{\delta_T}(\delta_T) \text{ARL}(\delta_X, \delta_T)$$

where Ω_X and Ω_T are the “range of possible shifts” for δ_X and δ_T , respectively, $f_{\delta_X}(\delta_X)$ and $f_{\delta_T}(\delta_T)$ are the p.m.f. (probability mass functions) of the shifts δ_X and δ_T over Ω_X and Ω_T , respectively. Since the goal of a TBEA control chart is to detect an increase in the amplitude X and/or a decrease in the time between events T , we suggest to define $\Omega_X = \{1.05, 1.1, 1.15, 1.2, 1.25, 1.3\}$ and $\Omega_T = \{0.7, 0.75, 0.8, 0.85, 0.9, 0.95\}$. As the distributions of δ_X and δ_T are unknown, we assume that $f_{\delta_X}(\delta_X)$ and $f_{\delta_T}(\delta_T)$ are the p.m.f. of discrete uniform distributions on Ω_X and Ω_T , respectively.

In the parametric framework, the values EARL_X , EARL_T and EARL_{XT} depend on the values $\text{ARL}(\delta_X, 1)$, $\text{ARL}(1, \delta_T)$ and $\text{ARL}(\delta_X, \delta_T)$ that can be computed using formulas presented in Rahali et al. [24]. Concerning the distribution-free upper-sided EWMA TBEA chart, the same formulas for EARL_X , EARL_T and EARL_{XT} can be used with the difference that, for each parametric scenario, the values of δ_X and δ_T have to be transformed into equivalent probabilities p_X and p_T and the values $\text{ARL}(\delta_X, 1)$, $\text{ARL}(1, \delta_T)$ and $\text{ARL}(\delta_X, \delta_T)$ have to be replaced by $\text{ARL}(p_X, 0.5)$, $\text{ARL}(0.5, p_T)$ and $\text{ARL}(p_X, p_T)$, that can be computed using (6). This allows a direct comparison between the parametric methods proposed in Rahali et al. [24] and the distribution-free method proposed in this paper.

The results of EARL_X , EARL_T and EARL_{XT} for the distribution-free upper-sided EWMA TBEA control chart are in Table 3 (see the values in bold) for both scenarios #1 and #2, i.e.

- $\text{EARL}_X = 24.91$, $\text{EARL}_T = 45.08$ and $\text{EARL}_{XT} = 10.49$ for scenario #1,
- $\text{EARL}_X = 44.30$, $\text{EARL}_T = 23.54$ and $\text{EARL}_{XT} = 9.93$ for scenario #2.

In Table 3, we also have the value of p_X and p_T (which depend on the scenario) corresponding to the values $\delta_X \in \Omega_X = \{1.05, 1.1, 1.15, 1.2, 1.25, 1.3\}$ and $\delta_T \in \Omega_T = \{0.7, 0.75, 0.8, 0.85, 0.9, 0.95\}$, respectively. For each of these combinations (δ_X, δ_T) or (p_X, p_T) we have the optimal parameters (λ^*, K^*) with the corresponding values for $(\text{ARL}, \text{SDRL})$. For example, in scenario #1, the combination $(\delta_X = 1.2, \delta_T = 0.9)$ corresponds to $(p_X = 0.9773, p_T = 0.3105)$ and the optimal parameters for the distribution-free upper-sided EWMA TBEA control chart are $(\lambda^* = 0.15, K^* = 2.643)$ with $(\text{ARL} = 9.78, \text{SDRL} = 4.23)$.

The values of EARL_X , EARL_T and EARL_{XT} for the 3 parametric Shewhart control charts proposed in Rahali et al. [24] (based on statistics Z_1 , Z_2 and Z_3) are presented in Table 4 for scenarios #1 and #2. The upper control limits used for each case have also been recorded in this table. A comparison between Tables 3 and 4 immediately shows that, no matter the scenario or the statistic considered, the values of EARL_X , EARL_T and EARL_{XT} for the distribution-free upper-sided EWMA TBEA control chart are always smaller than the ones obtained for the parametric Shewhart control charts proposed in Rahali et al. [24], thus showing the advantage of using the proposed distribution-free control chart in situations where the distributions for T and X are unknown.

Table 3. Values EARL_X, EARL_T and EARL_{XT} of the distribution-free upper-sided EWMA TBEA control chart for scenarios #1 and #2

Scenario #1									
	δ_X	1	1.05	1.1	1.15	1.2	1.25	1.3	
	p_X	0.5	0.6918	0.8416	0.9333	0.9773	0.9938	0.9987	
		(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	
δ_T	p_T	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	EARL _X
1	0.5	(—, —) (370.40, —)	(0.020, 2.084) (54.45, 33.49)	(0.045, 2.387) (26.63, 14.05)	(0.065, 2.496) (19.35, 9.30)	(0.075, 2.532) (16.93, 7.71)	(0.075, 2.532) (16.15, 7.04)	(0.080, 2.547) (15.93, 7.05)	24.91
0.95	0.4007	(0.010, 1.774) (106.86, 75.17)	(0.045, 2.387) (32.03, 19.35)	(0.075, 2.532) (18.27, 9.50)	(0.100, 2.592) (13.99, 6.65)	(0.110, 2.608) (12.48, 5.58)	(0.115, 2.615) (11.98, 5.25)	(0.115, 2.615) (11.84, 5.13)	16.77
0.90	0.3105	(0.020, 2.084) (55.20, 34.16)	(0.065, 2.496) (22.13, 12.48)	(0.110, 2.608) (13.74, 6.97)	(0.140, 2.638) (10.84, 5.02)	(0.150, 2.643) (9.78, 4.23)	(0.155, 2.644) (9.42, 3.97)	(0.160, 2.646) (9.32, 3.94)	12.53
0.85	0.2333	(0.035, 2.301) (36.62, 21.45)	(0.085, 2.560) (17.01, 9.04)	(0.140, 2.638) (11.10, 5.36)	(0.175, 2.648) (8.92, 3.89)	(0.190, 2.648) (8.11, 3.30)	(0.195, 2.647) (7.83, 3.09)	(0.200, 2.647) (7.75, 3.06)	10.12
0.80	0.1706	(0.045, 2.387) (27.93, 15.17)	(0.105, 2.601) (14.12, 7.18)	(0.165, 2.647) (9.48, 4.32)	(0.205, 2.645) (7.72, 3.14)	(0.225, 2.639) (7.04, 2.66)	(0.235, 2.634) (6.81, 2.51)	(0.235, 2.634) (6.75, 2.45)	8.65
0.75	0.1220	(0.055, 2.449) (23.28, 12.04)	(0.120, 2.621) (12.38, 6.02)	(0.185, 2.648) (8.47, 3.65)	(0.235, 2.634) (6.94, 2.66)	(0.255, 2.622) (6.35, 2.22)	(0.260, 2.618) (6.15, 2.05)	(0.265, 2.614) (6.09, 2.01)	7.73
0.70	0.0858	(0.060, 2.474) (20.58, 10.06)	(0.135, 2.634) (11.29, 5.35)	(0.205, 2.645) (7.82, 3.24)	(0.250, 2.625) (6.44, 2.29)	(0.275, 2.606) (5.90, 1.88)	(0.280, 2.602) (5.72, 1.72)	(0.280, 2.602) (5.66, 1.66)	7.14
	EARL _T	45.08	18.16	11.48	9.14	8.28	7.99	7.90	10.49

Scenario #2									
	δ_X	1	1.05	1.1	1.15	1.2	1.25	1.3	
	p_X	0.5	0.5985	0.6913	0.7732	0.8412	0.8943	0.9331	
		(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	(λ, K)	
δ_T	p_T	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	(ARL, SDRL)	EARL _X
1	0.5	(—, —) (370.40, —)	(0.010, 1.774) (107.63, 75.90)	(0.020, 2.084) (54.62, 33.64)	(0.035, 2.301) (35.52, 20.48)	(0.045, 2.387) (26.68, 14.09)	(0.055, 2.449) (22.00, 10.94)	(0.065, 2.496) (19.36, 9.31)	44.30
0.95	0.300027	(0.025, 2.174) (51.89, 32.88)	(0.045, 2.387) (31.00, 18.43)	(0.070, 2.515) (21.34, 12.09)	(0.090, 2.572) (16.25, 8.57)	(0.115, 2.615) (13.35, 6.76)	(0.130, 2.631) (11.61, 5.55)	(0.145, 2.640) (10.55, 4.87)	17.35
0.90	0.129897	(0.050, 2.420) (23.95, 12.21)	(0.085, 2.560) (16.78, 8.60)	(0.120, 2.621) (12.66, 6.26)	(0.155, 2.644) (10.17, 4.77)	(0.180, 2.648) (8.63, 3.74)	(0.210, 2.644) (7.67, 3.15)	(0.230, 2.636) (7.06, 2.74)	10.50
0.85	0.034429	(0.070, 2.515) (17.53, 8.00)	(0.105, 2.601) (12.89, 5.81)	(0.150, 2.643) (10.00, 4.43)	(0.190, 2.648) (8.19, 3.38)	(0.230, 2.636) (7.03, 2.69)	(0.255, 2.622) (6.28, 2.16)	(0.275, 2.606) (5.81, 1.80)	8.37
0.80	0.004493	(0.080, 2.547) (16.08, 7.17)	(0.115, 2.615) (11.96, 5.23)	(0.160, 2.646) (9.35, 3.96)	(0.200, 2.647) (7.69, 3.01)	(0.240, 2.631) (6.62, 2.37)	(0.270, 2.610) (5.93, 1.88)	(0.290, 2.593) (5.49, 1.53)	7.84
0.75	0.000222	(0.080, 2.547) (15.89, 7.01)	(0.115, 2.615) (11.83, 5.12)	(0.160, 2.646) (9.26, 3.89)	(0.205, 2.645) (7.62, 2.99)	(0.245, 2.628) (6.56, 2.34)	(0.275, 2.606) (5.88, 1.86)	(0.290, 2.593) (5.45, 1.48)	7.77
0.70	0.000003	(0.060, 2.474) (15.88, 7.00)	(0.135, 2.634) (11.83, 5.12)	(0.205, 2.645) (9.25, 3.88)	(0.250, 2.625) (7.62, 2.99)	(0.275, 2.606) (6.56, 2.34)	(0.280, 2.602) (5.88, 1.85)	(0.280, 2.602) (5.44, 1.48)	7.76
	EARL _T	23.54	16.05	11.98	9.59	8.13	7.21	6.63	9.93

Table 4. Values EARL_X, EARL_T and EARL_{XT} of the 3 parametric Shewhart control charts proposed in Rahali et al. [24] for scenarios #1 and #2

	Scenario #1			Scenario #2		
	Z ₁	Z ₂	Z ₃	Z ₁	Z ₂	Z ₃
UCL	0.5550	1.9692	2.9115	0.5470	1.6742	2.6171
EARL _X	52.4749	93.1968	134.3478	85.9319	121.0851	110.8255
EARL _T	55.3093	44.5776	41.6612	69.5180	57.4231	57.5436
EARL _{XT}	10.9375	14.8817	18.9865	18.8108	20.8753	20.5259

5. Illustrative example

We consider the same illustrative example as in Rahali et al. [24], which is based on a real data set concerning the time T_i (in days) between fires in forests of the french region “Provence - Alpes - Côte D’Azur” and their amplitudes X_i (burned surface in $ha = 10000m^2$, where only surfaces larger than $1ha$ have been included). This data set reports a total of 92 fires that have been divided into two subsets:

- 47 fires, from October 2016 to approximately mid-June 2017. This subset, corresponding to the “low season” for fires, is used here as a Phase I data set.
- 45 fires, from approximately mid-June 2017 to the end of September 2017. This subset, corresponding to the “high season” for fires, is used here as a Phase II data set.

The dates D_i (from October 1st 2016, in days), the times between fires T_i as well as their amplitudes X_i have been recorded in Table 5. The values of T_i and X_i are also plotted in Figure 3 (top and middle): it is evident that the Phase II values of $T_i(X_i)$ are shorter(larger) than those observed during Phase I.

In order to compute the upper control-limit UCL of the distribution-free EWMA TBFA chart, the following values have been fixed: $p_T = 0.3$, $p_X = 0.7$, $\sigma = 0.125$ and $ARL_0 = 370.4$. Using the results in Table 2 we have $\lambda^* = 0.07$ and $K^* = 2.515$ and we get

$$UCL = 2.515 \times \sqrt{\frac{0.07 \times (0.125^2 + 0.5)}{2 - 0.07}} = 0.344.$$

From the Phase I data set, the following in-control median values have been estimated $\theta_{T_0} = 3$ and $\theta_{X_0} = 5.3$. These values are used to compute the values ST_i , SX_i , S_i and S_i^* in Table 5. As it can be noticed, for some dates we have $T_i = \theta_{T_0} = 3$: for these values, we have $T_i - \theta_{T_0} = 0$. This is not supposed to happen as T_i is supposed to be a continuous random variable but, due to the measurement scale (days), this situation actually may occur. In this case, we have decided to keep the corresponding values and to assign $ST_i = 0$ (instead of -1 or $+1$). For this reason, some values of $S_i = s = \pm 0.5$ and the corresponding values for S_i^* are obtained by randomly generating a $Nor(s, \sigma)$ random variable, as it is already the case for values $s \in \{-1, 0, +1\}$. For instance, in Table 5, when $D_i = 70$ we have $S_i = 0.5$ and the corresponding value for S_i^* has been randomly generated from a $Nor(0.5, 0.125)$ distribution ($S_i^* = 0.552$). When $D_i = 140$ we have $S_i = -0.5$ and the corresponding value for S_i^* has been randomly generated from a $Nor(-0.5, 0.125)$ distribution ($S_i^* = -0.507$).

The values Z_i^* have been computed using (2) for both Phase I and II data sets. They are recorded in Table 5 and plotted in Figure 3 (bottom) along with the distribution-free EWMA TBFA upper control limit $UCL = 0.344$:

- For the Phase I data set, all the values Z_i^* are below the upper control limit $UCL = 0.344$; therefore, this data set is in-control and the estimated median values $\theta_{T_0} = 3$ and $\theta_{X_0} = 5.3$ can be used for the Phase II monitoring.
- For the Phase II data set, the distribution-free EWMA TBFA detects several out-of-control situations during the period mid-June 2017 – end of September

2017, (see also the bold values in Table 5), confirming that a decrease in the time between fires occurred with a concurrent increase in the amplitude of these fires. Similar conclusions have been obtained using the parametric approaches in Rahali et al. [24] and assuming a lognormal distribution for both T_i and X_i .

- It is interesting to investigate the robustness of the Z_i^* after the generation of new normal random values for S_i^* . In this case, are the new Z_i^* totally different from the others? Is it possible to detect the same out-of-control situations? How robust is the “continuousify” method if it is replicated several times? In order to answer these questions, we plotted, in Figure 4, 10 replicated sequences of Z_i^* corresponding to the same Phase I and II fires example. The D_i , T_i , X_i , ST_i , SX_i and S_i are exactly the same as in Table 5: only the S_i^* have been randomly generated and the Z_i^* recomputed. As it can be seen, the 10 trajectories are quite similar and none of them significantly diverges from the others. During the Phase I they are all below the UCL (confirming that the process is actually in-control) and during the Phase II, they all exhibit out-of-control situations more or less at the same moments (the “peaks” occur almost at the same time). Therefore, we can conclude that the “continuousify” method is robust vs. the random generation of the S_i^* values.

6. Conclusions

In this paper we have investigated a distribution-free EWMA control chart based on sign statistics for monitoring time between events and amplitude data. Implementing a distribution-free control chart allows the problem of estimating the in-control distributions of time between events and amplitudes to be overcome. Only a couple of selected quantiles from the in-control distributions need to be estimated to start the implementation of the EWMA TBEA control chart. Using the proposed distribution-free EWMA TBEA control chart allows to efficiently monitor TBEA data with an out-of-control performance significantly better than any parametric Shewhart TBEA control charts currently available in literature. Being the EWMA TBEA control chart based on a discrete sample statistic, we have defined a technique, called “continuousify”, which allows to compute the ARL values for the EWMA scheme by using Markov chains in a reliable and replicable way.

Future research in the same area will consider the case of possible ties in the data, i.e. situations where $ST_i = 0$ or $SX_i = 0$ (like it happens in the Illustrative example). In this case, the distribution of S_i and S_i^* are no longer the same and this new situation is worth to be investigated. Furthermore, this technique can be extended to any kind of discrete distribution like the Poisson or the binomial distribution and zero-inflated versions of these distributions, for instance. It could also be adapted to work with the multivariate version of the EWMA for multivariate discrete data like multivariate Poisson data.

References

- [1] M. Abid, H.Z. Nazir, M. Riaz, and Z. Lin. Use of Ranked Set Sampling in Nonparametric Control Charts. *Journal of the Chinese Institute of Engineers*, 39(5):627–636, 2016.
- [2] M. Abid, H.Z. Nazir, M. Riaz, and Z. Lin. An Efficient Nonparametric EWMA Wilcoxon

Table 5. Phase I and II values of D_i , T_i , X_i , ST_i , SX_i , S_i , S_i^* and Z_i^* for the forest fires example.

Phase 1								Phase 2							
D_i	T_i	X_i	ST_i	SX_i	S_i	S_i^*	Z_i^*	D_i	T_i	X_i	ST_i	SX_i	S_i	S_i^*	Z_i^*
9	9	3.68	1	-1	-1.0	-0.917	0.000	258	1	1.00	-1	-1	0.0	-0.078	0.000
26	17	1.99	1	-1	-1.0	-0.802	0.000	260	2	3.70	-1	-1	0.0	0.119	0.008
60	34	6.00	1	1	0.0	-0.081	0.000	262	2	3.17	-1	-1	0.0	-0.063	0.003
67	7	1.19	1	-1	-1.0	-0.901	0.000	265	3	18.40	0	1	0.5	0.333	0.026
70	3	135.80	0	1	0.5	0.552	0.039	268	3	1.00	0	-1	-0.5	-0.145	0.014
72	2	14.37	-1	1	1.0	1.113	0.114	269	1	2.22	-1	-1	0.0	0.208	0.028
86	14	8.10	1	1	0.0	-0.104	0.099	271	2	19.09	-1	1	1.0	1.001	0.096
88	2	32.31	-1	1	1.0	0.892	0.154	272	1	2.00	-1	-1	0.0	0.027	0.091
94	6	3.07	1	-1	-1.0	-1.056	0.069	274	2	34.28	-1	1	1.0	1.086	0.161
95	1	10.03	-1	1	1.0	0.867	0.125	276	2	3.00	-1	-1	0.0	0.070	0.154
96	1	7.93	-1	1	1.0	1.033	0.189	277	1	6.63	-1	1	1.0	0.955	0.210
97	1	1.50	-1	-1	0.0	0.409	0.204	278	1	4.47	-1	-1	0.0	-0.097	0.189
103	6	23.30	1	1	0.0	-0.116	0.182	285	7	8.24	1	1	0.0	0.160	0.187
106	3	3.73	0	-1	-0.5	-0.708	0.120	286	1	769.45	-1	1	1.0	1.024	0.246
109	3	4.73	0	-1	-0.5	-0.677	0.064	287	1	4.37	-1	-1	0.0	-0.144	0.218
111	2	3.19	-1	-1	0.0	0.179	0.072	288	1	90.70	-1	1	1.0	0.961	0.270
113	2	6.25	-1	1	1.0	1.032	0.139	289	1	11.49	-1	1	1.0	1.044	0.324
114	1	3.60	-1	-1	0.0	-0.155	0.118	295	6	3590.78	1	1	0.0	0.033	0.304
115	1	6.12	-1	1	1.0	1.112	0.188	296	1	1427.92	-1	1	1.0	0.949	0.349
118	3	1.50	0	-1	-0.5	-0.740	0.123	297	1	255.96	-1	1	1.0	1.054	0.399
122	4	1.33	1	-1	-1.0	-1.009	0.044	298	1	1.00	-1	-1	0.0	-0.051	0.367
134	12	1.42	1	-1	-1.0	-1.037	0.000	302	4	13.88	1	1	0.0	-0.074	0.336
137	3	5.75	0	1	0.5	0.629	0.044	303	1	138.28	-1	1	1.0	1.117	0.391
140	3	3.47	0	-1	-0.5	-0.507	0.005	305	2	8.90	-1	1	1.0	1.153	0.444
142	2	13.31	-1	1	1.0	1.217	0.090	308	3	1.50	0	-1	-0.5	-0.342	0.389
143	1	26.31	-1	1	1.0	1.041	0.157	312	4	34.63	1	1	0.0	-0.217	0.347
144	1	18.54	-1	1	1.0	0.923	0.210	313	1	82.56	-1	1	1.0	0.811	0.379
146	2	66.17	-1	1	1.0	1.124	0.274	314	1	2.00	-1	-1	0.0	-0.019	0.351
147	1	9.90	-1	1	1.0	0.916	0.319	315	1	162.08	-1	1	1.0	1.071	0.402
150	3	4.22	0	-1	-0.5	-0.534	0.260	319	4	3.26	1	-1	-1.0	-1.056	0.300
157	7	34.28	1	1	0.0	-0.110	0.234	321	2	285.91	-1	1	1.0	0.729	0.330
161	4	2.23	1	-1	-1.0	-1.102	0.140	322	1	2.00	-1	-1	0.0	-0.283	0.287
162	1	1.84	-1	-1	0.0	0.152	0.141	325	3	11.57	0	1	0.5	0.347	0.291
163	1	2.88	-1	-1	0.0	-0.018	0.130	334	9	34.70	1	1	0.0	0.068	0.275
164	1	21.46	-1	1	1.0	1.087	0.197	335	1	431.00	-1	1	1.0	1.150	0.337
165	1	4.46	-1	-1	0.0	-0.001	0.183	336	1	10.89	-1	1	1.0	1.003	0.383
166	1	58.27	-1	1	1.0	1.034	0.243	340	4	1.00	1	-1	-1.0	-1.004	0.286
167	1	8.84	-1	1	1.0	0.863	0.286	346	6	1.50	1	-1	-1.0	-0.921	0.202
180	13	1.03	1	-1	-1.0	-0.905	0.203	347	1	1.17	-1	-1	0.0	0.100	0.195
187	7	16.57	1	1	0.0	0.156	0.199	349	2	1.27	-1	-1	0.0	0.129	0.190
201	14	4.96	1	-1	-1.0	-1.084	0.110	350	1	26.25	-1	1	1.0	1.098	0.254
202	1	1.37	-1	-1	0.0	-0.087	0.096	353	3	11.66	0	1	0.5	0.332	0.259
205	3	23.39	0	1	0.5	0.498	0.124	354	1	3.03	-1	-1	0.0	0.127	0.250
225	20	1.70	1	-1	-1.0	-1.032	0.043	355	1	12.00	-1	1	1.0	1.130	0.311
247	22	5.30	1	0	-0.5	-0.727	0.000	356	1	1.00	-1	-1	0.0	-0.206	0.275
248	1	15.64	-1	1	1.0	1.161	0.081								
257	9	5.14	1	-1	-1.0	-0.921	0.011								

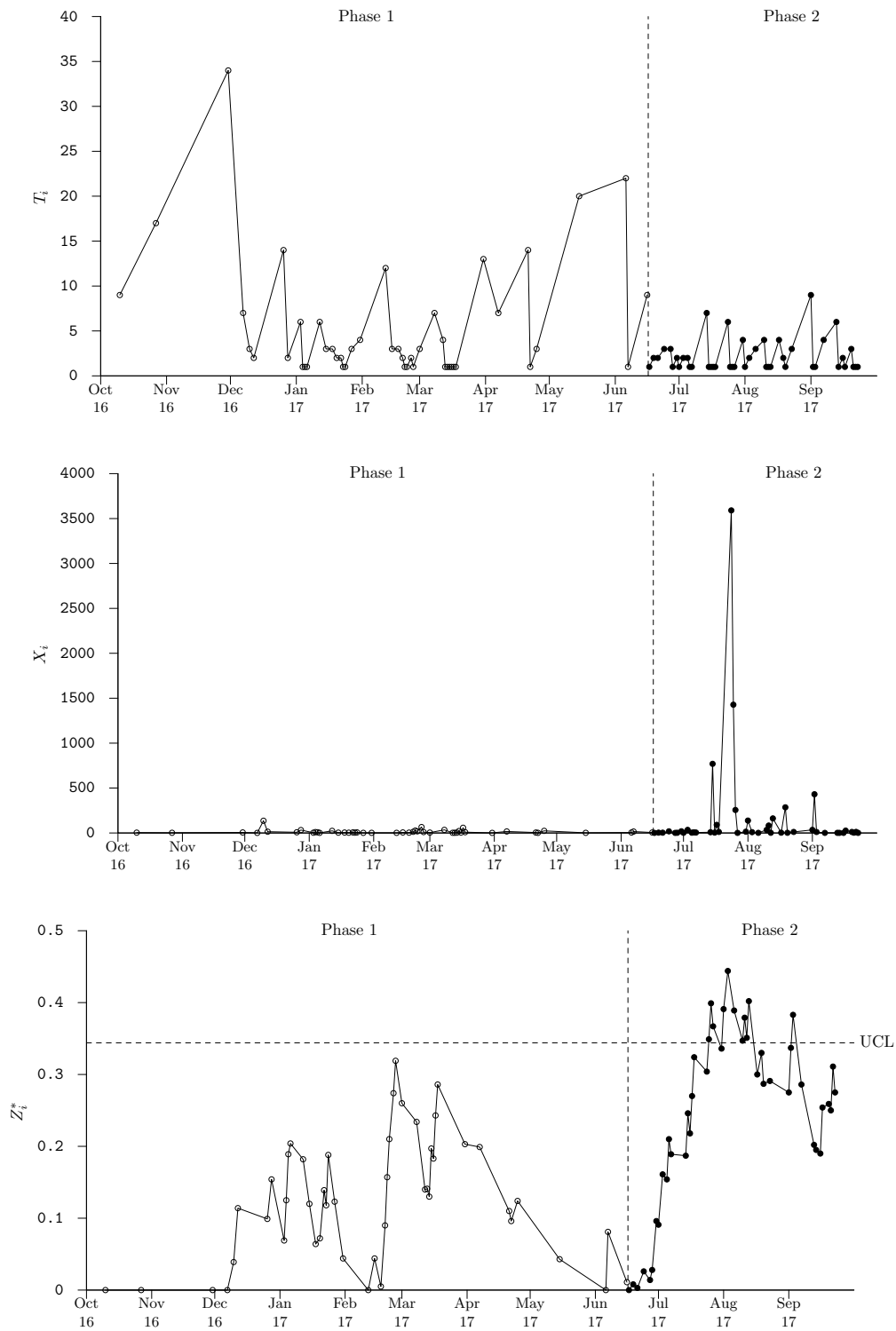


Figure 3. Time T_i in days between fires (top), amplitudes X_i as the burned surface in ha (middle) and the distribution-free EWMA TBEA chart with statistic Z_i^* (bottom) corresponding to the data set in Table 5

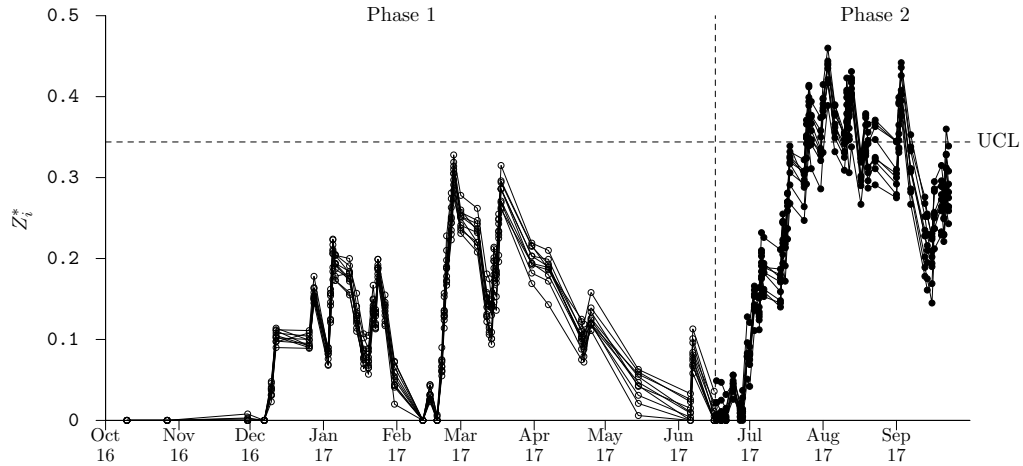


Figure 4. 10 new trajectories for the Z_i^* based on randomly regenerated S_i^* values corresponding to the data set in Table 5

- Signed-Rank Chart for Monitoring Location. *Quality and Reliability Engineering International*, 33(3):669–685, 2017.
- [3] M. Abid, H.Z. Nazir, M. Riaz, and Z. Lin. Investigating the Impact of Ranked Set Sampling in Nonparametric CUSUM Control Charts. *Quality and Reliability Engineering International*, 33(1):203–214, 2017.
 - [4] M. Abid, H.Z. Nazir, M. Tahir, and M. Riaz. On Designing a New Cumulative Sum Wilcoxon Signed Rank Chart for Monitoring Process Location. *PloS one*, 13(4):e0195762, 2018.
 - [5] S. Ali and A. Pievatolo. Time and Magnitude Monitoring based on the Renewal Reward Process. *Reliability Engineering & System Safety*, 179:97–107, 2018.
 - [6] C.M. Borrer, J.B. Keats, and D.C. Montgomery. Robustness of the Time Between Events CUSUM. *International Journal of Production Research*, 41(15):3435–3444, 2003.
 - [7] P.D. Bourke. The Geometric CUSUM Chart with Sampling Inspection for Monitoring Fraction Defective. *Journal of Applied Statistics*, 28(8):951–972, 2001.
 - [8] D. Brook and D.A. Evans. An Approach to the Probability Distribution of CUSUM Run Length. *Biometrika*, 59(3):539–549, 1972.
 - [9] T. Calvin. Quality Control Techniques for “Zero Defect”. *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, 6(3):323–328, 1983.
 - [10] P. Castagliola, K.P. Tran, G. Celano, A.C. Rakitzis, and P.E. Maravelakis. An EWMA-Type Sign Chart with Exact Run Length Properties. *Journal of Quality Technology*, 51(1):51–63, 2019.
 - [11] G. Celano, P. Castagliola, S. Chakraborti, and G. Nenes. The Performance of the Shewhart Sign Control Chart for Finite Horizon Processes. *International Journal of Advanced Manufacturing Technology*, 84(5):1497–1512, 2016.
 - [12] S. Chakraborti and M.A. Graham. *Nonparametric Statistical Process Control*. Wiley, 2019.
 - [13] S. Chakraborti and M.A. Graham. Nonparametric (Distribution-free) Control Charts: an Updated Overview and Some Results. *Quality Engineering*, 31(4):523–544, 2019.
 - [14] L.Y. Chan, M. Xie, and T.N. Goh. Cumulative Quantity Control Charts for Monitoring Production Processes. *International Journal of Production Research*, 38(2):397–408, 2000.
 - [15] C.S. Cheng and P.W. Chen. An ARL-Unbiased Design of Time-Between-Events Control Charts with Runs Rules. *Journal of Statistical Computation and Simulation*, 81(7):857–871, 2011.
 - [16] Y.Y. Fang, M.B.C Khoo, S.Y. Teh, and M. Xie. Monitoring of Time-Between-Events with

- a Generalized Group Runs Control Chart. *Quality and Reliability Engineering International*, 32(3):767–781, 2016.
- [17] F.F. Gan. Exact Run Length Distributions for One-Sided Exponential CUSUM Schemes. *Statistica Sinica*, 2(1):297–312, 1992.
- [18] C.H. Ho. Statistical Control Chart for Regime Identification in Volcanic Time Series. *Mathematical Geology*, 24(7):775–787, 1992.
- [19] J.M. Lucas. Counted Data CUSUM's. *Technometrics*, 27(2):129–144, 1985.
- [20] P. Qiu. *Introduction to Statistical Process Control*. Chapman and Hall, Boca Raton(USA), 2014.
- [21] L. Qu, Z. Wu, M.B.C. Khoo, and P. Castagliola. A Cusum Scheme for Event Monitoring. *International Journal of Production Economics*, 145(1):268–280, 2013.
- [22] L. Qu, Z. Wu, M.B.C. Khoo, and A. Rahim. Time-Between-Event Control Charts for Sampling Inspection. *Technometrics*, 56(3):336–346, 2014.
- [23] L. Qu, SG. He, M.B.C. Khoo, and P. Castagliola. A Cusum Chart for Detecting the Intensity Ratio of Negative Events. *International Journal of Production Research*, 56(19):6553–6567, 2018.
- [24] D. Rahali, P. Castagliola, H. Taleb, and M.B.C. Khoo. Evaluation of Shewhart Time-Between-Events-and-Amplitude Control Charts for Several Distributions. *Quality Engineering*, 31(2):240–254, 2019.
- [25] E. Santiago and J. Smith. Control Charts based on the Exponential Distribution: Adapting Runs Rules for the t Chart. *Quality Engineering*, 25(2):85–96, 2013.
- [26] M.S. Shafae, R.M. Dickinson, W.H. Woodall, and J.A. Camelio. Cumulative Sum Control Charts for Monitoring Weibull-Distributed Time Between Events. *Quality and Reliability Engineering International*, 31(5):839–849, 2015.
- [27] M. Shamsuzzaman, M. Xie, T.N. Goh, and H.Y. Zhang. Integrated Control Chart System for Time-Between-Events Monitoring in a Multistage Manufacturing System. *The International Journal of Advanced Manufacturing Technology*, 40(3):373–381, 2009.
- [28] S. Vardeman and D.O. Ray. Average Run Lengths for CUSUM Schemes when Observations are Exponentially Distributed. *Technometrics*, 27(2):145–150, 1985.
- [29] B. Vučijak, T. Kupusović, S. Midžić-Kurtagić, and A. Čerić. Climate Change Assessment using Statistical Process Control Methods. In *Climate-Smart Technologies*, pages 113–125. Springer, 2013.
- [30] W.H. Woodall. The Use of Control Charts in Health-Care and Public-Health Surveillance. *Journal of Quality Technology*, 38(2):89–104, 2006.
- [31] Z. Wu, J. Jiao, and Z. He. A Control Scheme for Monitoring the Frequency and Magnitude of an Event. *International Journal of Production Research*, 47(11):2887–2902, 2009.
- [32] Z. Wu, J. Jiao, and Z. He. A Single Control Chart for Monitoring the Frequency and Magnitude of an Event. *International Journal of Production Economics*, 119(1):24–33, 2009.
- [33] Z. Wu, Y. Liu, Z. He, and M.B.C. Khoo. A Cumulative Sum Scheme for Monitoring Frequency and Size of an Event. *Quality and Reliability Engineering International*, 26(6):541–554, 2010.
- [34] M. Xie, T.N. Goh, and P. Ranjan. Some Effective Control Chart Procedures for Reliability Monitoring. *Reliability Engineering & System Safety*, 77(2):143–150, 2002.
- [35] H.Y. Zhang, M. Shamsuzzaman, M. Xie, and T.N. Goh. Design and Application of Exponential Chart for Monitoring Time-Between-Events Data under Random Process Shift. *The International Journal of Advanced Manufacturing Technology*, 57(9):849–857, 2011.
- [36] H.Y. Zhang, M. Xie, T.N. Goh, and M. Shamsuzzaman. Economic Design of Time-Between-Events Control Chart System. *Computers and Industrial Engineering*, 60(4):485–492, 2011.