



**HAL**  
open science

# An EWMA-type chart based on signed ranks with exact run length properties

T. Perdikis, S. Psarakis, P. Castagliola, G. Celano

## ► To cite this version:

T. Perdikis, S. Psarakis, P. Castagliola, G. Celano. An EWMA-type chart based on signed ranks with exact run length properties. *Journal of Statistical Computation and Simulation*, 2021, 91 (4), pp.732-751. 10.1080/00949655.2020.1828415 . hal-03149541

**HAL Id: hal-03149541**

**<https://hal.science/hal-03149541v1>**

Submitted on 19 Aug 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.

# An EWMA-type Chart based on Signed Ranks with Exact Run Length Properties

T. Perdikis<sup>a</sup> and S. Psarakis<sup>a</sup> and P. Castagliola<sup>b</sup> and G. Celano<sup>c</sup>

<sup>a</sup>Department of Statistics & Laboratory of Statistical Methodology, Athens University of Economics and Business, Athens, Greece; <sup>b</sup>Université de Nantes & LS2N UMR CNRS 6004, Nantes, France; <sup>c</sup>Università di Catania, Catania, Italy

## ARTICLE HISTORY

Compiled August 19, 2021

## ABSTRACT

In this work, a new distribution-free Phase II EWMA-type chart based on the Wilcoxon signed rank statistic is introduced and its *exact* Run Length properties are discussed. A proper discrete Markov-chain approach is used for the determination of the Average Run Length in the in-control and out-of-control process operating conditions and its exact performance is derived without any knowledge of the distribution of sample observations. In addition, an algorithm for the determination of the chart's optimal design has been developed. Moreover, the performance of our proposed chart is compared with several existing nonparametric schemes available from literature. Finally, two illustrative examples are provided to show the practical implementation of our proposed chart.

## KEYWORDS

nonparametric control chart; Wilcoxon signed rank statistic; Phase II control chart

## 1. Introduction

Quality improvement is of high importance for the manufacturing industries and, as a result, statistical methods play a vital role in this direction. SPM (Statistical Process Monitoring) techniques allow the on-line monitoring of the product's characteristic to be performed during production. Control charts are the most important tool to perform statistical process monitoring. They are an on-line process monitoring technique whose purpose is to detect shifts in the process as quickly as possible. One of the most commonly used control chart is the Shewhart-type control chart, see [1]. Shewhart control charts are easy to design and able to detect large shifts in the parameter of interest during the process monitoring. On the other hand, in cases where small shifts of process parameters are likely to occur, memory-type control charts like the EWMA (Exponentially Weighted Moving Average, see [2]) or the CUSUM (Cumulative Sum, see [3]) charts are preferable.

It should be noted that the design and operations of control charts are commonly based on the assumption that the observations, collected over time, are random variables following the Normal distribution or at least some continuous distribution.

However, in many cases this assumption is violated since, in practice, the actual distribution of the characteristic to be monitored is unknown. As a result, a new class of distribution-free (nonparametric) control schemes have been introduced into the literature. Chakraborti et al. presented in [4] an extensive literature review, illustrating the design and operations of several univariate nonparametric control charts. Recently, Chakraborti and Graham presented in [5] an updated version of the former review by adding recent publications related to univariate and multivariate nonparametric charting techniques in SPM. Regarding the use of nonparametric techniques in EWMA-type control charts, Amin and Searcy introduced in [6] a nonparametric EWMA procedure for monitoring the location of a measurement based on the Wilcoxon signed rank statistic. Li et al. suggested in [7] two nonparametric extensions of the CUSUM and EWMA control charts using the Wilcoxon rank-sum test. Graham et al. investigated in [8] a distribution-free EWMA sign chart for location based on individual measurements and they presented in [9] a nonparametric EWMA chart based on the Wilcoxon signed rank statistic. Yang et al. introduced in [10] a modified EWMA sign chart using the arcsine transformation. Graham et al. proposed in [11] a modified distribution-free binomial-type EWMA chart called the EWMA-EX chart, in which exceedance statistics are used. Extensions of nonparametric EWMA-type schemes based on the sign statistic can be found in [12–14]. Chakraborty et al. proposed in [15] a distribution-free Generally Weighted Moving Average (GWMA) control chart based on the Wilcoxon signed rank statistic. Abid et al. investigated in [16] the performance of an EWMA signed rank chart under ranked set sampling. Haq presented in [17] a new nonparametric EWMA sign chart using an auxiliary-information-based (AIB) estimator of the process mean and, in [18], he proposed a new synthetic EWMA (SynEWMA) sign chart based on the arcsine transformation for monitoring the process mean. Recently, Raza et al. presented in [19] a Double Exponentially Weighted Moving Average (DEWMA) control chart based on the Wilcoxon signed rank statistic. Finally, Abbas et al. proposed in [20] two DEWMA schemes based on the Wilcoxon signed rank statistic in which simple random and rank set sampling techniques are investigated.

It is worth stretching that, regarding the design phase of an EWMA-type control chart a proper computation of its RL (Run length) properties is crucial. Specifically, for the determination of the design parameters  $\lambda$  (smoothing parameter) and  $K$  (control limit parameter) of an EWMA-type chart, a search algorithm needs to be run, for a particular shift in the process, in order to find the optimal pair  $(\lambda^*, K^*)$  which minimizes the out-of-control ARL, under the constraint  $ARL = ARL_0$  where  $ARL_0$  is a predefined value of the in-control ARL. In conventional EWMA-type schemes the RL properties are often obtained by using the Markov Chain approach of Brook and Evans presented in [21], which is based on a discretization of the control limits interval into  $m$  subintervals. In the case of measurement data (usually assuming normality) as the number of subintervals,  $m$ , increases the method of Brook and Evans tends to give reliable approximations of the chart's RL properties. Generally, the nonparametric statistics (such as sign, Wilcoxon signed rank, etc.) used by the distribution-free schemes mentioned above, are defined on a discrete domain. When the standard continuous EWMA chart is applied to discrete data, a common technique in order to compute its RL properties is to use simulation-based techniques. The main disadvantage of this approach is that it leads to approximated results which depend on the number of runs. Additionally, Weiß in [22] stated that the approach presented in [21], when applied to an EWMA control chart for count data, leads

to an approximation of the chart's actual statistical performance, which does not necessarily converge as the number of the Markov chain states increases. In a recent work of Wu et al. in [23] where a distribution-free EWMA-TBEA (Time Between Events and Amplitude) control chart based on the sign statistic is considered, they showed that the approach of Brook and Evans yields to unreliable results since it is seriously affected by the number of subintervals,  $m$ . In order to overcome this problem, Castagliola et al. in [24] introduced a new class of EWMA-type nonparametric scheme using the sign statistic called as the CEWMA SN chart. The design of the CEWMA SN chart is based on the approach of Rakitzis et al. who introduced in [25] a new memory type EWMA chart (denoted as the CEWMA control chart) for monitoring count data from a Poisson distribution. This chart is able to monitor the process location when the observations have an unknown distribution and, at the same time, it guarantees exact results for their run length distribution without any approximation. Tang et al. in [26] extended the approach of [24] and proposed an adaptive distribution-free EWMA sign chart using an adaptive feature in the smoothing parameter. As Castagliola et al. pointed out in [24], the design of the CEWMA chart can be utilised for any other nonparametric test statistics besides the sign test statistic. As far as we are concerned, in EWMA-type schemes based on the Wilcoxon signed rank statistic presented in the literature, their out-of-control RL properties are derived based on a specified distribution and using simulation-based techniques. In this current work, we aim to present a new nonparametric EWMA chart based on the recursive formula proposed in [25] in which the Wilcoxon signed rank statistic will be utilised. We also aim to provide a methodology for the exact determination of the RL properties of a EWMA-type chart based on signed ranks not only for the in-control but also for the out-of-control cases.

The paper is organized as follows. In section 2, the CEWMA SN chart is reviewed. Section 3, consists of a brief review of the statistical properties and operations of the Wilcoxon signed rank statistic and an exact formula for the computation of its general distribution is provided. In sections 4 and 5, our proposed nonparametric scheme based on the Wilcoxon signed rank statistic is introduced along with a specified algorithm for the computation of the exact in- and out-of-control ARL values based on a proper discrete Markov Chain model. In Section 6, the optimal design parameters are provided for different shifts of the process location. Moreover, our proposed chart is numerically compared with several nonparametric control charts proposed in the SPM. In Section 7, two illustrating examples are discussed to show a practical implementation of the operation of our proposed chart. Finally, in Section 8, some concluding remarks and suggestions for future work are discussed.

## 2. The CEWMA SN control chart

As it was noted before, when the standard EWMA chart is applied to discrete data, the Markov chain method of Brook and Evans (see [21]) leads to an approximation of its actual statistical performance. Using the approach presented in [25] and the Shewhart sign control chart proposed in [27], Castagliola et al. introduced in [24] a new nonparametric EWMA-type chart based on the sign statistic (CEWMA SN chart) for monitoring shifts in the location parameter providing exact results for its Run Length distribution.

Let  $X$  be a quality characteristic following an unknown continuous distribution with c.d.f. (cumulative distribution function)  $F_X(x|\theta)$  where  $\theta$  is the location parameter to be monitored. With reference to the Phase II implementation of the control chart, when the process is in-control we have  $\theta = \theta_0$ , (assumed as known). On the other hand, when the process is out-of-control, we have  $\theta = \theta_1$ . Suppose that at each sampling point  $t = 1, 2, \dots$  a subgroup  $\{X_{t,1}, X_{t,2}, \dots, X_{t,n}\}$  of size  $n$  is collected. The plotting statistic,  $Y_t$ , for the CEWMA SN control chart is obtained through the following formula:

$$(\gamma_x + \gamma_y)Y_t + R_t = \gamma_x \text{SN}_t + \underbrace{\gamma_y Y_{t-1} + R_{t-1}}_{B_{t-1}},$$

$$\text{SN}_t = \sum_{j=1}^n \text{sign}(X_{t,j} - \theta_0),$$

where  $\text{sign}(x) = -1, 0$  or  $+1$  if  $x < 0, x = 0$  or  $x > 0$ , respectively. Note that, the condition  $\text{sign}(x) = 0$  corresponds to occasional events occurring in practice due to rounding-off errors of the measurement system. Additionally,  $(\gamma_x, \gamma_y) \in \mathbb{N}^2$  are two fixed positive *integer-valued* parameters,  $B_{t-1} \stackrel{\text{def}}{=} \gamma_y Y_{t-1} + R_{t-1}$ , and  $Y_t$  is the quotient of the Euclidean division

$$\left\lfloor \frac{\gamma_x \text{SN}_t + B_{t-1}}{\gamma_x + \gamma_y} \right\rfloor,$$

where  $\lfloor \dots \rfloor$  denotes the rounded towards zero integer and  $R_t$  is the remainder of this Euclidean division defined as:

$$R_t = \gamma_x \text{SN}_t + B_{t-1} - (\gamma_x + \gamma_y)Y_t.$$

As stated in [24], when the initial values  $Y_0 = y_0, R_0 = r_0$  and the current values  $\text{SN}_t, Y_{t-1}, R_{t-1}$  are fixed, both  $Y_t$  and  $R_t$  are uniquely defined. The initial values  $y_0$  and  $r_0$ , are set equal to  $y_0 = r_0 = 0$ . However, if a head-start feature is desired, any choice of  $y_0 \neq 0$  or  $r_0 \neq 0$  can be considered. For a better understanding of the operation and design of the CEWMA SN plotting statistic the reader is advised to refer to [24].

The Run length properties of the CEWMA SN chart are obtained based on the fact that  $\text{SN}_t$  is defined on  $\{-n, -n+2, \dots, n-2, n\}$  and its distribution can be obtained based on the relationship  $\text{SN}_t = 2D_t - n$ , where  $D_t$  is the number of observations  $\{X_{t,1}, X_{t,2}, \dots, X_{t,n}\}$  larger than  $\theta_0$ . If the process is in-control,  $P(X_{t,j} > \theta_0 | \theta = \theta_0) = P(X_{t,j} < \theta_0 | \theta = \theta_0) = p_0 = 0.5$ . As a result,  $D_t$  is a binomial random variable with parameters  $n$  and  $p_0 = 0.5$ . On the other hand, let  $p_1 = P(X_{t,j} > \theta_0 | \theta = \theta_1) = 1 - F_X(\theta_0 | \theta_1)$  be the probability of having an observation larger than the in-control median  $\theta_0$  when the process runs out-of-control with median  $\theta = \theta_1$ . Assuming that the median is the location parameter of interest, values of  $p_1$  close to  $p_0 = 0.5$  correspond to small shifts, while values of  $p_1$  close to 0 or 1 correspond to large shifts from  $\theta_0$  to  $\theta_1$ . In addition, since the in-control distribution of  $\text{SN}_t$  is symmetric about 0, the

control limits (UCL, LCL) and the center line (CL) are equal to  $LCL = -K$ ,  $CL = 0$  and  $UCL = K$  where  $K = \{2, 4, \dots, n\}$  (if  $n$  is an even integer) or  $K = \{1, 3, \dots, n\}$  (otherwise). The process is declared to be in-control if  $-K < Y_t < K$  and out-of-control otherwise. Regarding the computation of the optimal design parameters of the CEWMA SN chart a discrete Markov chain approach is used which will be also utilised in our proposed scheme.

### 3. The Wilcoxon signed rank (WSR) statistic

Suppose that at each sampling point a subgroup  $\{X_{t,1}, X_{t,2}, \dots, X_{t,n}\}$  of size  $n$ , following a continuous *symmetric* distribution, is collected at time  $t = 1, 2, \dots$ . Let  $L_{t,j} \in \{1, 2, \dots, n\}$  denote the rank of the absolute value of the differences  $|X_{t,j} - \theta_0|$ ,  $j = 1, 2, \dots, n$  for subgroup  $t = 1, 2, \dots$ . By definition, the Wilcoxon signed rank statistic  $SR_t$  is equal to

$$SR_t = \sum_{j=1}^n \text{sign}(X_{t,j} - \theta_0) L_{t,j},$$

and it is defined on  $\{-\frac{n(n+1)}{2}, -\frac{n(n+1)}{2} + 2, \dots, \frac{n(n+1)}{2} - 2, \frac{n(n+1)}{2}\}$ . Therefore,  $SR_t$  is the sum of the signed ranks. It should be noted that the domain on which  $SR_t$  is defined contains zero only if  $\frac{n(n+1)}{2}$  is an even integer. The statistic  $SR_t$  can also be expressed as:

$$SR_t = 2SR_t^+ - \frac{n(n+1)}{2},$$

where  $SR_t^+$  is the sum of the *positive* ranks only. Under the *null* hypothesis (i.e. for  $p = 0.5$ ), the p.m.f. (probability mass function)  $f_{SR_t^+}(s|n)$  of  $SR_t^+$  can be easily and exactly obtained (see [28], for instance) by recursively evaluating the number  $N_{SR_t^+}(s|n)$  of subsets of integers in  $\{1, \dots, n\}$  having a sum equal to  $s \in \{0, 1, \dots, \frac{n(n+1)}{2}\}$ , i.e.

$$N_{SR_t^+}(s|n) = N_{SR_t^+}(s|n-1) + N_{SR_t^+}(s-n|n-1),$$

and by computing

$$f_{SR_t^+}(s|n) = \frac{N_{SR_t^+}(s|n)}{2^n}.$$

Computing the p.m.f.  $f_{SR_t^+}(s|n, p)$  of  $SR_t^+$  (exactly, without any approximation) under the *alternative* hypothesis (i.e.  $p \neq 0.5$ ) is more tricky. The solution we opted in this paper consists in evaluating firstly the p.g.f. (probability generating function)  $G_{SR_t^+}(\omega)$  of  $SR_t^+$  as (see [29])

$$G_{SR_t^+}(\omega) = \prod_{i=1}^n (p\omega^i + q),$$

where  $q = 1 - p$  and to obtain the p.m.f.  $f_{SR_t^+}(s|n, p)$  of  $SR_t^+$  by differentiating

$G_{\text{SR}_t^+}(\omega)$ ,  $s$  times, for  $\omega = 0$ , using the formula

$$f_{\text{SR}_t^+}(s|n, p) = \frac{1}{s!} G_{\text{SR}_t^+}^{(s)}(\omega) \Big|_{\omega=0},$$

where  $G_{\text{SR}_t^+}^{(s)}(\omega)$  is the  $s^{\text{th}}$  derivative of  $G_{\text{SR}_t^+}(\omega)$ . As  $G_{\text{SR}_t^+}(\omega)$  is a polynomial of degree  $\frac{n(n+1)}{2}$  then, for  $s \in \{0, 1, \dots, \frac{n(n+1)}{2}\}$ ,  $G_{\text{SR}_t^+}^{(s)}(\omega)$  is a polynomial of degree  $\frac{n(n+1)}{2} - s$  and we can write

$$\frac{1}{s!} G_{\text{SR}_t^+}^{(s)}(\omega) = \sum_{j=0}^{\frac{n(n+1)}{2}-s} c_{s,j} \omega^j,$$

where  $c_{s,j}$  is the coefficient of degree  $j$  corresponding to the polynomial  $\frac{1}{s!} G_{\text{SR}_t^+}^{(s)}(\omega)$ . For a better understanding of this approach let us consider the following example with  $n = 4$  and  $p = 0.2$ . In this case  $\text{SR}_t^+ \in \{0, 1, \dots, 10\}$  and we have

- $\frac{1}{0!} G_{\text{SR}_t^+}(\omega) = 0.0016\omega^{10} + 0.0064\omega^9 + 0.0064\omega^8 + 0.032\omega^7 + 0.032\omega^6 + 0.0512\omega^5 + 0.128\omega^4 + 0.128\omega^3 + 0.1024\omega^2 + 0.1024\omega + 0.4096$ . The coefficients  $c_{0,j}$  are listed in the first row (for  $s = 0$ ) of Table 1. Replacing  $\omega = 0$  in this polynomial gives  $f_{\text{SR}_t^+}(0|4, 0.2) = 0.4096$ .
- Now, if we evaluate the first derivative, we have  $\frac{1}{1!} G_{\text{SR}_t^+}^{(1)}(\omega) = 0.016\omega^9 + 0.0576\omega^8 + 0.0512\omega^7 + 0.224\omega^6 + 0.192\omega^5 + 0.256\omega^4 + 0.512\omega^3 + 0.384\omega^2 + 0.2048\omega + 0.1024$ . The coefficients  $c_{1,j}$  are listed in the second row (for  $s = 1$ ) of Table 1. Replacing  $\omega = 0$  in this polynomial gives  $f_{\text{SR}_t^+}(1|4, 0.2) = 0.1024$ .
- $\vdots$
- If we evaluate the 9<sup>th</sup> derivative, we have  $\frac{1}{9!} G_{\text{SR}_t^+}^{(9)}(\omega) = 0.016\omega + 0.0064$ . The coefficients  $c_{9,j}$  are listed in the row corresponding for  $s = 9$  of Table 1. Replacing  $\omega = 0$  in this polynomial gives  $f_{\text{SR}_t^+}(9|4, 0.2) = 0.0064$ . A final derivative gives  $\frac{1}{10!} G_{\text{SR}_t^+}^{(10)}(\omega) = 0.0016$  and we have  $f_{\text{SR}_t^+}(10|4, 0.2) = 0.0016$ .

Because polynomials can be efficiently coded as vectors of coefficients, fast arithmetic operations (addition, multiplication and power) and derivation can be efficiently implemented (as in Matlab for instance) and the evaluation of  $f_{\text{SR}_t^+}(s|n, p)$  can be obtained in a very fast way. It should be noted that this method can also be applied when  $p = 0.5$ , i.e. for the null hypothesis case.

#### 4. Design of the CEWMA WSR control chart

As for the design of the CEWMA SN control chart, at each sampling point, the plotting statistic  $Y_t$  for the CEWMA WSR control chart is obtained through the following formula:

$$(\gamma_x + \gamma_y)Y_t + R_t = \gamma_x \text{SR}_t + \underbrace{\gamma_y Y_{t-1} + R_{t-1}}_{B_{t-1}}. \quad (1)$$

**Table 1.** Computation of the p.m.f. of  $SR_t^+$  for  $n = 4$  and  $p = 0.2$ 

$s$	$c_{s,10}$	$c_{s,9}$	$c_{s,8}$	$c_{s,7}$	$c_{s,6}$	$c_{s,5}$	$c_{s,4}$	$c_{s,3}$	$c_{s,2}$	$c_{s,1}$	$c_{s,0}$	$f_{SR_t^+}(s 4, 0.2)$
0	0.0016	0.0064	0.0064	0.0320	0.0320	0.0512	0.1280	0.1280	0.1024	0.1024	0.4096	0.4096
1	-	0.0160	0.0576	0.0512	0.2240	0.1920	0.2560	0.5120	0.3840	0.2048	0.1024	0.1024
2	-	-	0.0720	0.2304	0.1792	0.6720	0.4800	0.5120	0.7680	0.3840	0.1024	0.1024
3	-	-	-	0.1920	0.5376	0.3584	1.1200	0.6400	0.5120	0.5120	0.1280	0.1280
4	-	-	-	-	0.3360	0.8064	0.4480	1.1200	0.4800	0.2560	0.1280	0.1280
5	-	-	-	-	-	0.4032	0.8064	0.3584	0.6720	0.1920	0.0512	0.0512
6	-	-	-	-	-	-	0.3360	0.5376	0.1792	0.2240	0.0320	0.0320
7	-	-	-	-	-	-	-	0.1920	0.2304	0.0512	0.0320	0.0320
8	-	-	-	-	-	-	-	-	0.0720	0.0576	0.0064	0.0064
9	-	-	-	-	-	-	-	-	-	0.0160	0.0064	0.0064
10	-	-	-	-	-	-	-	-	-	-	0.0016	0.0016

**Table 2.** An example of calculation of  $Y_t$  and  $R_t$  given  $SR_t$ 

$t$	$SR_t$	$Y_t$	$R_t$
0	-	0	0
1	-17	-2	-5
2	15	0	0
3	21	3	3
4	-7	1	5
5	-15	0	-5
6	-13	-3	0
7	-31	-7	-4
8	-9	-8	0
9	37	0	-3
10	47	7	2
11	25	10	2
12	13	10	5
13	27	13	4
14	21	15	0
15	47	20	2

Concerning the initial values  $y_0$  and  $r_0$  for  $Y_0$  and  $R_0$  they are both set to  $y_0 = r_0 = 0$ . The process is declared to be in-control if  $-K < Y_t < K$  and out-of-control otherwise where  $K \in \{2, \dots, \frac{n(n+1)}{2}\}$ . Let us consider an example presented in Table 2 in order to clarify the design and operations of the charting statistic  $Y_t$  of our proposed scheme. A simulated dataset containing  $m = 15$  subgroups of size  $n = 10$  have been simulated where the first 10 subgroups are considered as in-control samples ( $p_0 = 0.5$ ) and the other 5 subgroups as out-of-control samples ( $p_1 = 0.8$ ). For illustrative purposes the design parameters are  $\gamma_x = 1$  and  $\gamma_y = 5$ , respectively and no head-start feature is used (i.e.  $Y_0 = 0$  and  $R_0 = 0$ ). The values of  $Y_t$  are computed as follows:

- For  $t = 1$  we have  $Y_0 = 0$ ,  $R_0 = 0$ ,  $SR_1 = -17$  and the equation to be solved is  $6 \times Y_1 + R_1 = 1 \times (-17) + 5 \times 0 + 0 = -17$ . The unique solution of this equation (as an Euclidean division) is  $Y_1 = -2$  and  $R_1 = -5$ .
- For  $t = 2$  we have  $Y_1 = -2$ ,  $R_1 = -5$ ,  $SR_2 = 15$  and the equation to be solved is  $6 \times Y_2 + R_2 = 1 \times (15) + 5 \times (-2) - 5 = 0$ . The unique solution of this equation is  $Y_2 = 0$  and  $R_2 = 0$ .
- $\vdots$
- For  $t = 15$  we have  $Y_{14} = 15$ ,  $R_{14} = 0$ ,  $SR_{15} = 47$  and the equation to be solved is  $6 \times Y_{15} + R_{15} = 1 \times (47) + 5 \times (15) + 0 = 122$ . The unique solution of this equation is  $Y_{15} = 20$  and  $R_{15} = 2$ .



It is interesting to note that the values of the charting statistic,  $Y_t$  react similarly as for a classical EWMA chart, i.e. they are randomly distributed around zero when the process is in-control (samples 1 – 10) and they start to increase when a shift occurs (samples 11 – 15). This fact has also been mentioned in [24].

Also note that, our proposed scheme can both be extended to monitor any particular quantiles (provided that the sample size  $n$  is large enough) by changing the value  $p_0 = 0.5$  to any other value of interest.

It is well acknowledged into the literature that in a conventional EWMA chart varying the value of the smoothing parameter,  $\lambda$ , allows more weight to be assigned to the past, or to the current observations. As pointed out in [25], if  $\gamma_x < \gamma_y$  then the CEWMA scheme assigns more weight to the past observations while, if  $\gamma_x > \gamma_y$ , this scheme assigns more weight to the most recent ones. Additionally, the corresponding limiting conditions for the CEWMA WSR chart in the case where  $\lambda = 1$  or  $\lambda = 0$  can be expressed respectively as:

- for  $\gamma_x > 0$  and  $\gamma_y = 0$  we have  $Y_t = SR_t$  and  $R_t = r_0, t = 1, 2, \dots$ . Consequently, in this situation, the CEWMA WSR chart coincides with the nonparametric Shewhart-type control chart based on signed ranks proposed in [30].
- for  $\gamma_x = 0$  and  $\gamma_y > 0$  we have  $Y_t = y_0$  and  $R_t = r_0, t = 1, 2, \dots$

## 5. Run length properties of the CEWMA WSR chart

In order to obtain the exact RL properties of our proposed scheme, the discrete-time Markov chain approach presented for the CEWMA SN control chart is used. Specifically, the transition probability matrix  $\mathbf{P}$  is defined as:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{r} \\ \mathbf{0}^\top & 1 \end{pmatrix} = \begin{pmatrix} q_{-b,-b} & q_{-b,-b+1} & \dots & q_{-b,b-1} & q_{-b,b} & r_{-b} \\ q_{-b+1,-b} & q_{-b+1,-b+1} & \dots & q_{-b+1,b-1} & q_{-b+1,b} & r_{-b+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ q_{-b+1,-b} & q_{b-1,-b+1} & \dots & q_{b-1,b-1} & q_{b-1,b} & r_{b-1} \\ 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

where  $\mathbf{Q}$  is the  $(2b+1, 2b+1)$  matrix of transient probabilities  $q_{i,j}$ ,  $\mathbf{0}^\top = (0, 0, \dots, 0)$  and  $\mathbf{r} = \mathbf{1} - \mathbf{Q}\mathbf{1}$ . Similar to the approach used for the CEWMA and CEWMA SN charts, at time  $t-1$  the transient states of the discrete-time Markov chain will be defined as the integers  $b_{t-1} \in \{-b, -b+1, \dots, +b\}$  where  $b = \gamma_x + K\gamma_y - 1$ . More specifically, assuming that at time  $t-1$ , as long as  $y_{t-1} \in \{-K+1, \dots, K-1\}$  the process is declared to be in-control, the transient states of the discrete-time Markov chain are defined as the integers  $b_{t-1} \in \{\gamma_y \times y_{t-1} + r_{t-1}\}$  where  $r_{t-1} \in \{-\gamma_x - \gamma_y + 1, \dots, \gamma_x + \gamma_y - 1\}$ . As a consequence, the minimum,  $b_{\min}$ , and the maximum,  $b_{\max}$ , numbers of states are defined as  $b_{\min} = \gamma_y(-K+1) - \gamma_x - \gamma_y + 1 = -\gamma_x - K\gamma_y + 1$  and  $b_{\max} = \gamma_y(K-1) + \gamma_x + \gamma_y - 1 = \gamma_x + K\gamma_y - 1$  respectively where  $b = b_{\max} = -b_{\min}$ . As a result, the total number of transient states is  $2b+1 = 2(\gamma_x + K\gamma_y - 1) + 1$ . Finally, the transient probabilities  $q_{i,j}$  will be computed as detailed in Algorithm 1.

The RL properties of the CEWMA WSR chart will be obtained using the fact that RL is a Discrete Phase-type (DPH) random variable of parameters  $(\mathbf{Q}, \mathbf{q})$  (see

---

**Algorithm 1** Computation of transient probabilities  $q_{i,j}$ 

---

Define  $n, K, \gamma_x, \gamma_y$  and  $p_1$ .  
 $b \leftarrow \gamma_x + K\gamma_y - 1$ .  
**for**  $i = -b, -b + 1, \dots, b - 1, b$  **do**  
  **for**  $\text{SR} = -\frac{n(n+1)}{2}, -\frac{n(n+1)}{2} + 2, \dots, \frac{n(n+1)}{2} - 2, \frac{n(n+1)}{2}$  **do**  
     $y \leftarrow \lfloor \frac{\gamma_x \times \text{SR} + i}{\gamma_x + \gamma_y} \rfloor$ .  
    **if**  $-K < y < K$  **then**  
       $r \leftarrow \gamma_x \times \text{SR} + i - (\gamma_x + \gamma_y) \times y$ .  
       $j \leftarrow \gamma_y \times y + r$ .  
       $q_{i,j} \leftarrow q_{i,j} + f_{\text{SR}^+} \left( \frac{\text{SR} + \frac{n(n+1)}{2}}{2} \mid n, p_1 \right)$ .  
    **end if**  
  **end for**  
**end for**

---

[31,32]). Specifically, the p.m.f.  $f_{\text{RL}}(t|\mathbf{Q}, \mathbf{q})$  and the c.d.f.  $F_{\text{RL}}(t|\mathbf{Q}, \mathbf{q})$  of RL are defined for  $t = 1, 2, \dots$  and they are respectively, equal to:

$$f_{\text{RL}}(t|\mathbf{Q}, \mathbf{q}) = \mathbf{q}^\top \mathbf{Q}^{t-1} \mathbf{r},$$
$$F_{\text{RL}}(t|\mathbf{Q}, \mathbf{q}) = 1 - \mathbf{q}^\top \mathbf{Q}^t \mathbf{1}.$$

So, the Average Run Length  $\text{ARL} = E(\text{RL})$  and the Standard Deviation Run Length  $\text{SDRL} = \sigma(\text{RL})$  will be computed as:

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

where  $\mathbf{q} = (q_{-b}, q_{-b+1}, \dots, q_{b-1}, q_b)^\top$  is the vector of initial probabilities associated with the  $2b + 1$  transient states. Assuming that  $y_0 = r_0 = 0$  then  $q_i = 0$  for  $i \in \{-b, -b + 1, \dots, -1, 1, \dots, b\}$  and  $q_0 = 1$ .

## 6. Performance Comparisons

In this work, the parameter  $\theta_0$  will be considered as the median of the process distribution, so we assume  $p_0 = 0.5$ . For fixed values of the sample size  $n$  and  $p_1 = P(X_{t,j} > \theta_0 | \theta = \theta_1)$ , in order to obtain the optimal combination of the design parameters  $(K^*, \gamma_x^*, \gamma_y^*)$  which will provide the minimum corresponding out-of-control ARL, a specific search algorithm is used aiming to select the optimal values of  $(K, \gamma_x, \gamma_y)$  which guarantee the following conditions:

$$(K^*, \gamma_x^*, \gamma_y^*) = \underset{(K, \gamma_x, \gamma_y)}{\text{argmin}} \text{ARL}(K, \gamma_x, \gamma_y, n, p_1)$$

$$\text{ARL}(K^*, \gamma_x^*, \gamma_y^*, n, p_0) = \text{ARL}_0$$

where  $ARL_0 = 370.4$  is the target in-control ARL value. Due to the fact that the charting statistic,  $Y_t$ , is a discrete random variable, it is not possible to find an optimal combination  $(K^*, \gamma_x^*, \gamma_y^*)$  that *exactly* meets the constraint  $ARL(K^*, \gamma_x^*, \gamma_y^*, n, p_0) = ARL_0$ . Similarly to the design of the CEWMA SN chart, we suggest to accept, as tentative design parameters, all those combinations of parameters  $(K^*, \gamma_x^*, \gamma_y^*)$  satisfying the following condition:

$$D(K^*, \gamma_x^*, \gamma_y^*, n) = \frac{|\text{ARL}(K^*, \gamma_x^*, \gamma_y^*, n, p_0) - \text{ARL}_0|}{\text{ARL}_0} \leq \delta \quad (2)$$

where  $\delta$  is a predefined constant. For the determination of the optimal parameters for  $n = \{10, \dots, 25\}$  in the CEWMA SN chart, Castagliola et al. set  $\delta = 0.05$  (see [24]), to ensure that the corresponding optimal combination of  $(K^*, \gamma_x^*, \gamma_y^*)$  will be obtained successfully. Based on our findings, we saw that in general the CEWMA WSR chart guarantees a smaller error than the CEWMA SN chart with respect to the constraint on the nominal in-control performance. As a consequence, for the determination of the design parameters of the CEWMA WSR control chart we suggest to set  $\delta = 0.01$ . Additionally, since  $ARL(K, \gamma_x, \gamma_y, n, p_1) = ARL(K, \gamma_x, \gamma_y, n, 1 - p_1)$  we will only focus on shifts  $p_1 \in (0, 0.5)$  as the same results would have been obtained for  $p_1 \in (0.5, 1)$ . As a consequence, for given values of  $n$  and  $p_1$ , the optimal combination of the design parameters  $(K^*, \gamma_x^*, \gamma_y^*)$  will be obtained through the searching Algorithm 2.

---

**Algorithm 2** Computation of optimal combination  $(K^*, \gamma_x^*, \gamma_y^*)$

---

```

Define  $n$  and  $p_1 \in (0, 0.5)$ .
 $ARL_0 \leftarrow 370.4$ .
 $ARL^* \leftarrow \infty$ .
for  $K = 1, \dots, \frac{n(n+1)}{2}$  do
  for  $\gamma_x = 1, 2, \dots$  do
     $\gamma_y \leftarrow 0$ .
    repeat
       $\gamma_y \leftarrow \gamma_y + 1$ 
    until  $D(K, \gamma_x, \gamma_y, n) \leq 0.01$ 
    if  $ARL(K, \gamma_x, \gamma_y, n, p_1) < ARL^*$  then
       $ARL^* \leftarrow ARL(K, \gamma_x, \gamma_y, n, p_1)$ ,
       $K^* \leftarrow K, \gamma_x^* \leftarrow \gamma_x$  and  $\gamma_y^* \leftarrow \gamma_y$ .
    end if
  end for
end for

```

---

In the following tables some results regarding the performance of our proposed chart are provided, including comparisons with the CEWMA SN chart. In Table 3, the optimal combinations  $(K^*, \gamma_x^*, \gamma_y^*)$  for our proposed chart are given, along with the corresponding in-control and out-of-control ARLs for  $n = \{10, \dots, 25\}$ . Based on the results presented in Table 3 it can be concluded that when  $p_1 < 0.5$ , for large values of the sample size or large shifts ( $p_1 \simeq 0$ ) the corresponding optimal  $ARL_1$  values tend to decrease. In addition, for fixed values in the sample size, as the value of  $p_1 < 0.5$  becomes closer to 0.5 the “smoothing ratio”  $\frac{\gamma_x}{\gamma_x + \gamma_y}$  and the value of the control limit  $K$  of our proposed chart also tend to decrease. For example for  $n = 20$  and  $p_1 \in \{0.05, \dots, 0.2\}$ , we have  $\frac{\gamma_x}{\gamma_x + \gamma_y} = \frac{10}{10+3} = 0.77$  and the control limit equals to

$K = 117$ . On the other hand for  $p_1 = 0.4$  the corresponding smoothing ratio equals to  $\frac{\gamma_x}{\gamma_x + \gamma_y} = \frac{4}{4+27} = 0.13$  and  $K = 39$ . It is worth mentioning that these findings are consistent with the ones of the CEWMA (see [25]) and CEWMA SN (see [24]) charts.

In Table 4 the performance of our proposed chart is compared to the CEWMA SN chart for  $n \in \{5, \dots, 25\}$  and  $p_1 \in \{0.05, 0.1, \dots, 0.45\}$ . In order to make fair comparisons, we obtained the optimal design parameters of the CEWMA SN with corresponding  $D(K^*, \gamma_x^*, \gamma_y^*, n) \leq 0.01$  (except from some values of  $n$  where the bound 0.01 in (2) has been replaced by 0.02), for  $n \in \{5, 6, \dots, 25\}$  and  $p_1 \in \{0.05, 0.1, \dots, 0.45\}$ . The  $\Delta$  values are defined as:

$$\Delta = 100 \times \frac{\text{ARL}_{\text{SR}} - \text{ARL}_{\text{SN}}}{\text{ARL}_{\text{SN}}},$$

where  $\text{ARL}_{\text{SN}}$  and  $\text{ARL}_{\text{SR}}$  are the ARL values of the CEWMA SN and CEWMA WSR charts respectively. A negative value for  $\Delta$  corresponds to an outperformance of the CEWMA WSR scheme versus the CEWMA SN chart. According to our findings, it can be concluded that when  $5 \leq n \leq 15$ , for large shifts ( $p_1 \leq 0.2$ ) in 55% of the cases the CEWMA WSR scheme has a better performance. In addition for small to moderate shifts ( $p_1 > 0.2$ ) in 60% of the cases the CEWMA WSR chart also outperforms the CEWMA SN chart. On the other hand, for  $n > 15$ , when  $p_1 \leq 0.2$  or  $p_1 > 0.2$  the CEWMA WSR chart performs better only in 38% of the cases. As a consequence, in most cases, our proposed scheme has an overall better performance when  $n < 15$  and as the sample size increases the CEWMA SN chart performs better.

In Tables 5 (for  $n = 10$ ) and 6 (for  $n = 20$ ) the CEWMA WSR chart is compared with other nonparametric schemes presented in the literature. More specifically, we examined the performance of the CEWMA WSR chart against the CEWMA SN chart, the two EWMA sign charts proposed by Yang et al. in [10], denoted as standard sign EWMA (EWMA) and arcsine sign EWMA (A-EWMA), the distribution free cumulative sum mean chart (CUSUM) of Yang et al. in [33] and the modified sign EWMA chart proposed by Lu in [14] (S-GWMA). For the computation of the in- and out-of-control ARL values of the competitors of the CEWMA WSR chart, a Monte Carlo simulation was performed, except for the CEWMA SN chart where the exact Markov chain approach was used. Regarding the design parameters of each chart, for the CEWMA WSR and CEWMA SN charts  $(K, \gamma_x, \gamma_y)$  have been selected to get the optimal performance for each chart; for the EWMA charts proposed in [10], the pair  $(\lambda, W)$  corresponds to the optimal values of the smoothing parameter  $\lambda$  and the distance  $W$ . Moreover, the pair  $(k, h)$  corresponds to the values of the reference value  $k$  and the decision interval  $h$  of the CUSUM chart. Finally,  $(q, \alpha, W)$  correspond to the values of the design parameters for the S-GWMA chart. For more information regarding the design and operations of the above schemes the reader is advised to see [10,14,33].

Based on the results presented in Tables 5 and 6 it can be concluded that the CEWMA WSR chart outperforms its competitors for moderate to large shifts ( $p_1 < 0.3$ ). On the other hand, for small shifts, i.e. when  $p_1$  tends to be close to 0.5 the standard or the arcsine transformed EWMA chart have better performance. Finally, in cases where  $0.25 < p_1 < 0.3$  the use of the S-GWMA chart or the SN EWMA can be considered.

**Table 3.** Optimal combinations  $(K^*, \gamma_x^*, \gamma_y^*)$  (first line of each block) for the CEWMA WSR chart along with the corresponding in-control ARL's (second line) and the out-of-control (ARL, SDRL) (third line) for  $n \in \{10, \dots, 25\}$  and  $p_1 \in \{0.05, 0.1, \dots, 0.45\}$

$n$	$p_1$								
	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45
10	(26,8,15)	(26,8,15)	(26,8,15)	(26,8,15)	(22,10,28)	(19,7,27)	(15,9,57)	(9,1,16)	(5,6,249)
	369	369	369	369	369.4	368.4	369.3	367.7	367.8
	(2.2,0.4)	(2.6,0.7)	(3.2,1.1)	(4.1,1.8)	(5.5,2.5)	(7.8,4)	(12.3,6.7)	(22.4,12.1)	(57.6,35.4)
11	(47,5,2)	(47,5,2)	(36,10,11)	(27,5,12)	(25,7,20)	(20,8,37)	(16,8,58)	(12,9,110)	(6,3,115)
	373.9	373.9	369.3	367.4	374	368.9	371.6	370.3	366.9
	(1.5,0.7)	(2.1,1.1)	(2.9,1.1)	(3.8,1.5)	(5.1,2.3)	(7.3,3.4)	(11.5,5.9)	(21,12)	(54.2,33.4)
12	(55,9,3)	(53,5,2)	(34,6,11)	(34,6,11)	(29,10,27)	(25,10,38)	(21,7,38)	(13,7,92)	(7,7,255)
	367.3	368.1	370.4	370.4	370.5	371.2	369.9	366.7	370.3
	(1.5,0.7)	(2,1)	(2.8,0.9)	(3.6,1.5)	(4.8,2.1)	(6.9,3.3)	(10.8,5.9)	(19.7,10.8)	(51.5,31.6)
13	(43,3,4)	(43,3,4)	(43,9,12)	(43,9,12)	(32,6,17)	(30,8,26)	(22,6,37)	(15,7,87)	(7,1,43)
	368.4	368.4	368.4	368.4	373.2	368.2	368.7	368.1	368.2
	(2,0,2)	(2.2,0.5)	(2.7,0.9)	(3.4,1.5)	(4.6,1.9)	(6.5,3.2)	(10.1,5.2)	(18.6,10.2)	(49.1,28.5)
14	(73,8,2)	(73,8,2)	(57,9,7)	(49,4,5)	(36,9,25)	(34,7,22)	(24,5,32)	(16,3,40)	(9,3,103)
	369.1	369.1	367.6	368	369.9	367.3	371.1	369.1	369.1
	(1.4,0.6)	(1.9,0.9)	(2.4,1)	(3.2,1.4)	(4.3,1.8)	(6.1,3)	(9.6,4.7)	(17.8,9.3)	(46.7,28.2)
15	(81,4,1)	(81,4,1)	(71,4,2)	(57,10,11)	(49,9,15)	(34,7,27)	(29,8,43)	(21,6,59)	(11,4,117)
	366.8	366.8	371.5	373.6	368.8	367.2	366.8	368.9	372.6
	(1.3,0.5)	(1.7,0.9)	(2.3,1.1)	(3,1.3)	(4.1,1.9)	(5.8,2.6)	(9.1,4.6)	(16.9,9.5)	(44.8,27.7)
16	(91,9,2)	(91,9,2)	(67,9,8)	(55,7,11)	(45,5,13)	(47,9,21)	(31,3,17)	(23,6,59)	(12,3,88)
	367	367	370.7	366.8	371.3	368.6	372.8	368.3	366.7
	(1.3,0.5)	(1.7,0.8)	(2.2,0.8)	(2.9,1.1)	(3.9,1.6)	(5.6,2.9)	(8.7,4.3)	(16.1,8.9)	(42.8,26.1)
17	(96,7,2)	(96,7,2)	(83,7,4)	(74,7,6)	(60,5,8)	(45,7,22)	(37,4,19)	(23,2,23)	(14,3,79)
	369.2	369.2	372.4	367.4	371.3	369.8	371.2	368.3	367.8
	(1.2,0.4)	(1.6,0.7)	(2.1,0.9)	(2.8,1.3)	(3.7,1.7)	(5.3,2.4)	(8.3,4.2)	(15.4,8)	(41.2,25.4)
18	(117,8,1)	(117,8,1)	(98,5,2)	(77,10,10)	(69,8,11)	(51,8,23)	(40,5,24)	(30,7,58)	(14,1,30)
	367.1	367.1	372.5	371.7	368.6	370.6	368.3	367.3	369.6
	(1.1,0.4)	(1.5,0.8)	(2,1)	(2.7,1.1)	(3.6,1.7)	(5.1,2.3)	(8,4)	(14.8,8.3)	(39.7,23.4)
19	(108,8,3)	(108,8,3)	(96,8,5)	(96,8,5)	(71,7,11)	(60,8,19)	(48,7,27)	(29,6,61)	(16,2,55)
	370.5	370.5	367.5	367.5	370.5	370	367.8	368.4	370.3
	(1.1,0.4)	(1.4,0.6)	(1.9,0.8)	(2.5,1.3)	(3.4,1.5)	(4.9,2.4)	(7.7,4)	(14.2,7.4)	(38.3,22.8)
20	(117,8,3)	(117,8,3)	(117,8,3)	(102,9,6)	(75,3,5)	(57,7,22)	(51,8,32)	(39,4,27)	(17,2,56)
	372.7	372.7	372.7	373.6	370.5	369.5	371.8	368.2	368
	(1.1,0.3)	(1.4,0.6)	(1.8,0.9)	(2.5,1.2)	(3.3,1.4)	(4.7,2)	(7.4,3.8)	(13.8,8)	(37,21.7)
21	(131,10,3)	(131,10,3)	(131,10,3)	(118,10,5)	(85,9,13)	(70,3,7)	(52,7,31)	(37,2,17)	(23,3,58)
	373.1	373.1	373.1	370.8	367.3	371.3	368.8	368.9	370.6
	(1.1,0.3)	(1.3,0.5)	(1.8,0.9)	(2.4,1.2)	(3.2,1.4)	(4.5,2.1)	(7.1,3.4)	(13.2,7)	(36.1,23)
22	(156,7,1)	(156,7,1)	(148,9,2)	(107,10,9)	(91,9,13)	(80,7,14)	(63,9,31)	(44,4,28)	(23,2,43)
	367.8	367.8	368.6	367.5	370.4	367.3	367.9	366.9	367.1
	(1.1,0.2)	(1.3,0.5)	(1.7,0.9)	(2.3,0.9)	(3.1,1.3)	(4.4,2.1)	(6.9,3.6)	(12.8,7.1)	(34.7,21.3)
23	(164,6,1)	(164,6,1)	(139,9,4)	(123,7,5)	(107,9,10)	(92,9,15)	(69,4,13)	(41,1,9)	(28,3,52)
	368.8	368.8	373.2	370.2	369.2	373.1	371.8	370.8	368.8
	(1.1,0.2)	(1.2,0.5)	(1.6,0.7)	(2.2,1)	(3,1.4)	(4.3,2.2)	(6.7,3.5)	(12.4,6.3)	(34,21.9)
24	(157,6,2)	(148,9,4)	(157,9,3)	(131,7,5)	(109,8,10)	(88,6,13)	(71,4,14)	(49,4,29)	(17,1,42)
	371.3	367.6	371.3	373.4	368	366.8	372.5	368.5	372
	(1,0,2)	(1.2,0.5)	(1.6,0.7)	(2.1,0.9)	(2.9,1.2)	(4.1,1.9)	(6.5,3.2)	(12,6.4)	(33.7,17.1)
25	(163,8,3)	(148,7,4)	(163,8,3)	(148,7,4)	(104,3,5)	(102,8,14)	(78,4,13)	(56,3,19)	(29,1,20)
	373	372.1	373	372.1	369.6	366.9	368.7	370.6	369.2
	(1,0,2)	(1.2,0.5)	(1.5,0.7)	(2,0.9)	(2.8,1.1)	(4,1.9)	(6.2,3.2)	(11.7,6.5)	(31.9,19.3)

**Table 4.** Comparison between the CEWMA SN and CEWMA WSR chart for  $n \in \{5, \dots, 25\}$  and  $p_1 \in \{0.05, 0.1, \dots, 0.45\}$

$n$	0.05	0.1	0.15	0.2	$p_1$ 0.25	0.3	0.35	0.4	0.45
5	<b>-37.21</b>	<b>-22.45</b>	<b>-10.53</b>	<b>-2.90</b>	3.45	11.21	18.75	6.91	<b>-11.55</b>
6	37.50	26.67	12.20	0.00	<b>-14.89</b>	<b>-31.93</b>	<b>-32.31</b>	<b>-18.97</b>	1.42
7	<b>-22.58</b>	<b>-8.82</b>	0.00	8.16	12.70	15.91	9.72	<b>-6.23</b>	<b>-2.05</b>
8	9.09	11.54	15.62	16.67	8.33	12.20	12.50	1.16	<b>-3.06</b>
9	57.14	35.00	21.43	<b>-2.22</b>	<b>-24.36</b>	<b>-44.44</b>	<b>-38.32</b>	<b>-25.00</b>	<b>-3.61</b>
10	<b>-69.01</b>	<b>-67.09</b>	<b>-64.44</b>	<b>-60.58</b>	<b>-55.65</b>	<b>-49.68</b>	<b>-40.00</b>	<b>-27.04</b>	<b>-5.73</b>
11	<b>-28.57</b>	<b>-12.50</b>	3.57	8.57	15.91	17.74	16.16	<b>-1.41</b>	<b>-6.39</b>
12	0.00	11.11	21.74	16.13	6.67	<b>-8.00</b>	<b>-28.95</b>	<b>-50.50</b>	<b>-63.14</b>
13	81.82	57.14	42.11	21.43	2.22	<b>-24.42</b>	<b>-44.20</b>	<b>-31.11</b>	<b>-8.74</b>
14	16.67	26.67	20.00	10.34	<b>-6.52</b>	<b>-25.61</b>	<b>-44.19</b>	<b>-30.74</b>	<b>-8.61</b>
15	<b>-35.00</b>	<b>-26.09</b>	<b>-14.81</b>	<b>-6.25</b>	5.13	11.54	18.18	19.01	5.66
16	30.00	41.67	37.50	31.82	11.43	16.67	19.18	20.15	9.18
17	<b>-77.78</b>	<b>-74.19</b>	<b>-70.00</b>	<b>-65.43</b>	<b>-61.46</b>	<b>-55.46</b>	<b>-47.47</b>	<b>-34.47</b>	<b>-11.59</b>
18	<b>-63.33</b>	<b>-57.14</b>	<b>-50.00</b>	<b>-42.55</b>	<b>-35.71</b>	<b>-27.14</b>	<b>-14.89</b>	0.68	16.08
19	0.00	7.69	11.76	13.64	6.25	<b>-10.91</b>	<b>-35.29</b>	<b>-36.61</b>	<b>-13.93</b>
20	0.00	7.69	5.88	19.05	17.86	14.63	15.62	21.05	15.62
21	<b>-78.00</b>	<b>-76.79</b>	<b>-71.43</b>	<b>-67.12</b>	<b>-63.22</b>	<b>-58.33</b>	<b>-50.35</b>	<b>-38.03</b>	<b>-14.45</b>
22	<b>-35.29</b>	<b>-31.58</b>	<b>-15.00</b>	0.00	10.71	18.92	21.05	17.43	<b>-11.93</b>
23	10.00	0.00	6.67	15.79	20.00	19.44	3.08	<b>-6.06</b>	14.48
24	0.00	9.09	23.08	23.53	20.83	13.89	12.07	20.00	11.22
25	0.00	9.09	15.38	17.65	21.74	14.29	19.23	20.62	19.03

**Table 5.** ARL values of the CEWMA WSR chart, the CEWMA SN chart, the standard and the arcsine transformed EWMA chart, the CUSUM chart and the S-GWMA chart ( $n = 10$ ,  $p_0 = 0.50$ ,  $ARL_0 \approx 370$ )

$p$	CEWMA WSR	CEWMA SN	EWMA	A-EWMA	CUSUM	S-GWMA
0.50	369.0	371.4	366.3	367.6	370.0	370.6
0.45	131.4	61.1	<b>51.4</b>	51.4	63.2	58.0
0.40	38.0	30.9	<b>19.0</b>	19.0	20.2	19.0
0.30	8.6	15.5	8.1	7.9	7.9	<b>7.2</b>
0.25	5.6	12.4	6.3	6.1	6.0	<b>5.4</b>
0.20	<b>4.0</b>	10.4	5.2	4.9	4.9	4.3
0.15	<b>3.2</b>	9	4.4	4.1	4.2	3.6
0.10	<b>2.6</b>	7.0	3.9	3.4	3.6	3.1
0.05	<b>2.2</b>	7.0	3.4	2.9	3.2	2.7
$\gamma_X$	8	1	-	-	-	-
$\gamma_Y$	15	1	-	-	-	-
$K$	26	58	-	-	-	-
$\lambda$	-	-	0.05	0.05	-	-
$W$	-	-	2.487	2.487	-	2.698
$k$	-	-	-	-	0.5	-
$h$	-	-	-	-	10.65	-
$q$	-	-	-	-	-	0.9
$\alpha$	-	-	-	-	-	0.9

**Table 6.** ARL values of the CEWMA WSR chart, the CEWMA SN chart, the standard and the arcsine transformed EWMA chart, the CUSUM chart and the S-GWMA chart ( $n = 20$ ,  $p_0 = 0.50$ ,  $ARL_0 \approx 370$ )

$p$	CEWMA WSR	CEWMA SN	EWMA	A-EWMA	CUSUM	S-GWMA
0.50	370.6	370.2	370.0	367.6	373.7	370.9
0.45	84.7	37.3	31.0	<b>30.9</b>	40.3	33.0
0.40	20.2	<b>11.4</b>	12.3	12.2	11.7	11.5
0.30	4.9	<b>4.5</b>	5.6	5.4	4.5	4.7
0.25	<b>3.3</b>	3.5	4.4	4.2	3.5	3.7
0.20	<b>2.6</b>	2.9	3.7	3.5	2.9	3.0
0.15	<b>2.1</b>	2.4	3.2	2.9	2.4	2.5
0.10	<b>1.8</b>	2.1	2.9	2.5	2.1	2.2
0.05	<b>1.5</b>	2.0	2.6	2.1	2.0	2.0
$\gamma_X$	3	3	-	-	-	-
$\gamma_Y$	5	16	-	-	-	-
$K$	75	4	-	-	-	-
$\lambda$	-	-	0.05	0.05	-	-
$W$	-	-	2.487	2.487	-	2.709
$k$	-	-	-	-	1.0	-
$h$	-	-	-	-	11.62	-
$q$	-	-	-	-	-	0.9
$\alpha$	-	-	-	-	-	0.9

## 7. Illustrative Examples

Motivated by existing works related to applications of nonparametric schemes for monitoring the characteristic of interest in a process, we present two examples, originally discussed by Celano et al. in [34] and [35] respectively, to show a practical Phase II implementation of the design and operations of our proposed chart. It should be noted that the distributions of the observations are unknown and so no control chart based on normality assumption can be implemented. In the first one (entitled as the *Radial example*), the quality characteristic to be monitored is the radial error, which is defined as “a quality characteristic frequently monitored in hole drilling processes of mechanical parts and assembly processes of printed circuit boards”. In the second one (entitled as the *Beverage example*), the quality characteristic of interest is the quantity of CO<sub>2</sub> dissolved in soft drink PET (polyethylene terephthalate) bottles.

### 7.1. Radial example

In this example, at each sampling point  $t$ , a subgroup of size  $n = 20$  is collected in order to detect a shift in the median of the quality of interest such that  $p_0 = 0.5$  shifts to  $p_1 = 0.3$ . As shown in Table 3, the optimal parameters to be used are  $K^* = 57$ ,  $\gamma_x^* = 7$ , and  $\gamma_y^* = 22$ . In addition, the in-control value of the median for the radial error is  $\theta_0 = 0.338$ . In Table 7 the differences between  $X_{t,j}$  and  $\theta_0$  for  $t \in \{1, 2, \dots, 10\}$  are provided and the values of  $SR_t$ ,  $Y_t$  and  $R_t$  are also reported. Using the recursive formula presented in (1) the values of the charting statistic  $Y_t$  are computed as follows:

- For  $t = 1$  we have  $Y_0 = 0$ ,  $R_0 = 0$ ,  $SR_1 = 45$  and the equation to be solved is  $29 \times Y_1 + R_1 = 7 \times 45 + 22 \times 0 + 0 = 315$ . The unique solution of this equation (as an Euclidean division) is  $Y_1 = 10$  and  $R_1 = 25$ .
- For  $t = 2$  we have  $Y_1 = 10$ ,  $R_1 = 25$ ,  $SR_2 = 27$  and the equation to be solved is  $29 \times Y_2 + R_2 = 7 \times (27) + 22 \times (10) + 25 = 434$ . The unique solution of this equation is  $Y_2 = 14$  and  $R_2 = 28$ .
- 
-

**Table 7.** Radial error example: Phase II dataset of  $t = 1, \dots, 10$  subgroups of size  $n = 20$  and the corresponding values for  $SR_t$ ,  $Y_t$  and  $R_t$

$t$	$X_{t,j}$									
	1	2	3	4	5	6	7	8	9	10
1	0.289	0.380	0.483	0.288	0.544	0.390	0.567	0.512	0.433	0.168
2	0.447	0.599	0.207	0.317	0.256	0.433	0.218	0.329	0.432	0.674
3	0.081	0.368	0.435	0.216	0.246	0.229	0.623	0.455	0.394	0.616
4	0.954	0.537	0.621	0.513	1.540	0.609	0.801	1.080	1.069	0.954
5	0.316	0.237	0.286	0.879	0.190	0.104	0.570	0.448	0.269	0.746
6	0.342	0.378	0.287	0.328	0.589	0.233	0.255	0.119	0.284	0.499
7	0.370	0.391	0.525	0.459	1.280	0.470	0.482	0.032	0.525	0.628
8	0.352	0.264	0.759	0.154	0.256	0.426	0.363	0.310	0.303	0.316
9	0.305	0.352	0.468	0.224	0.739	0.234	0.171	0.250	0.308	0.431
10	0.603	0.363	0.628	0.314	0.029	0.436	0.207	0.553	0.645	0.122

$t$	$X_{t,j}$										$SR_t$	$Y_t$	$R_t$
	11	12	13	14	15	16	17	18	19	20			
1	0.128	0.428	0.081	0.575	0.396	0.574	0.730	0.407	0.367	0.452	45	10	25
2	0.233	0.570	0.748	0.364	0.372	0.798	0.218	0.405	0.060	0.632	27	14	28
3	0.116	0.611	0.666	0.262	0.410	0.234	0.692	0.719	1.033	0.376	44	22	6
4	0.852	0.425	1.389	0.794	1.081	0.900	0.521	0.576	0.761	0.535	210	67	17
5	0.344	0.191	0.366	0.315	0.408	0.522	0.598	0.232	0.671	0.448	0	51	12
6	0.410	0.668	0.385	0.594	0.390	0.265	0.409	0.434	0.628	0.316	-11	36	13
7	0.686	0.584	0.300	0.245	0.555	0.113	0.194	0.932	0.597	0.523	84	48	1
8	0.807	0.235	0.173	0.183	1.105	0.068	0.368	0.736	0.097	0.060	-54	23	12
9	0.092	0.326	0.455	0.569	0.354	0.475	0.530	0.312	0.102	0.651	-31	10	11
10	0.759	0.296	0.691	0.425	0.441	0.323	0.287	0.310	0.194	0.582	18	12	9

- For  $t = 10$  we have  $Y_9 = 10$ ,  $R_9 = 11$ ,  $SR_{10} = 18$  and the equation to be solved is  $29 \times Y_{10} + R_{10} = 7 \times (18) + 22 \times (10) + 11 = 357$ . The unique solution of this equation is  $Y_{10} = 12$  and  $R_{10} = 9$ .

In Figure 1 the differences  $X_{t,j} - \theta_0$  for  $t \in \{1, 2, \dots, 10\}$  are plotted, where at each sampling point  $t$ , some values are tied together. The values of the charting statistic  $Y_t$  are plotted in Figure 2. It can be seen that at the 4<sup>th</sup> sampling point ( $t = 4$ ) an out-of-control signal is given stating that the process median has changed.

## 7.2. Beverage example

In this example, at each sampling point  $t$ , a subgroup of size  $n = 7$  is collected in order to detect a shift in the median of the quality of interest such that  $p_0 = 0.5$  shifts to  $p_1 = 0.4$ . Based on the results obtained by the optimization method presented in Section 6 (Algorithm 3) the optimal parameters are equal to  $K^* = 6$ ,  $\gamma_x^* = 10$ , and  $\gamma_y^* = 140$ . In Table 8 the deviations between  $X_{t,j}$  and  $\theta_0$  for  $t \in \{1, 2, \dots, 10\}$  are illustrated, (Figure 3), along with the corresponding values of  $SR_t$ ,  $Y_t$  and  $R_t$ , (for confidentiality reasons, the in-control / target median value  $\theta_0$  of the quality characteristic is not provided). Regarding the computation of the ranks,  $L_{t,j}$ , in presence of ties, we have used the average rank method: i.e., for the rank allocated to a group of ties / ex-aequos we have assigned the average of the ranks that would have been allocated to these observations if they would have been all different. Consequently, similarly to the above example, using the recursive formula presented in (1) the values of the charting statistic  $Y_t$  are computed as:

- For  $t = 1$  we have  $Y_0 = 0$ ,  $R_0 = 0$ ,  $SR_1 = 2$  and the equation to be solved is  $150 \times Y_1 + R_1 = 10 \times 2 + 140 \times 0 + 0 = 20$ . The unique solution of this equation



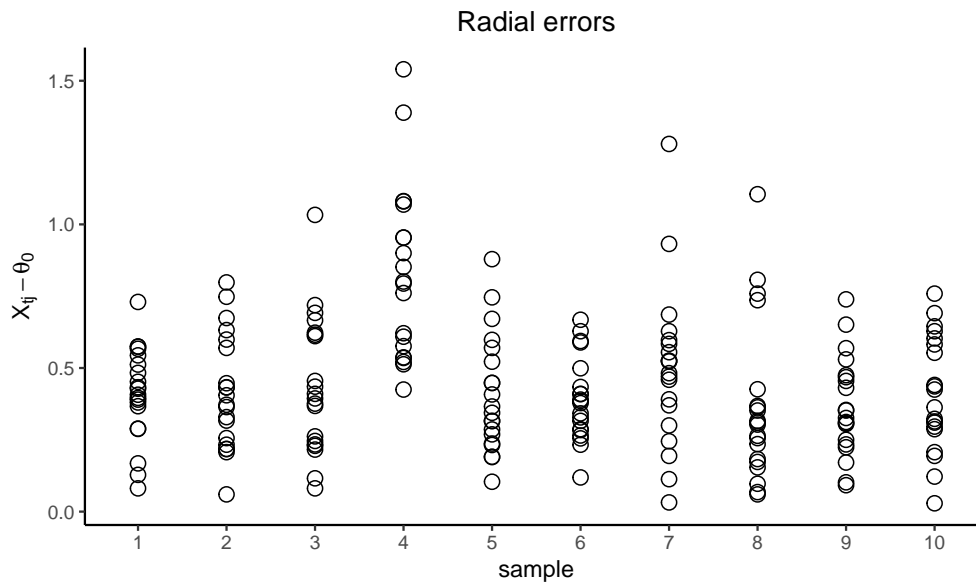


Figure 1. Radial error example: individual value plot of the observations

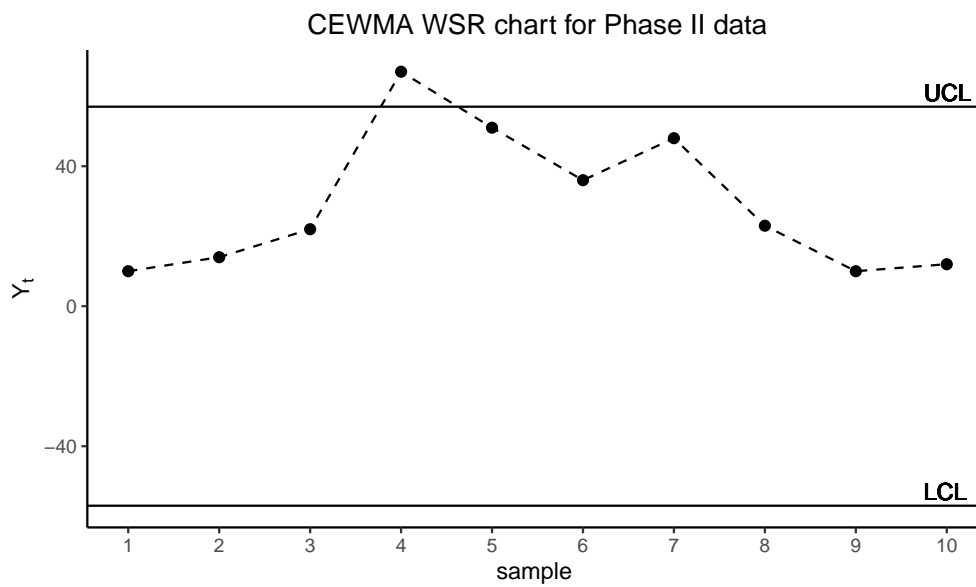
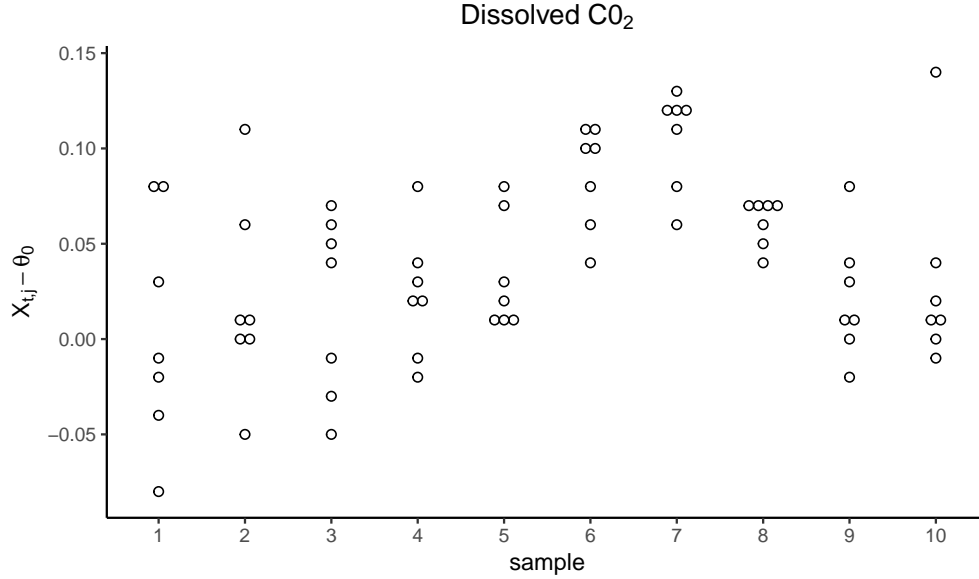


Figure 2. Radial error example: the CEWMA WSR chart for the Phase II sample presented in Table 7

**Table 8.** Beverage example: Phase II dataset of  $t = 1, \dots, 10$  subgroups of size  $n = 7$  and the corresponding values for  $SR_t$ ,  $Y_t$  and  $R_t$

$t$	$X_{t,j} - \theta_0$							$SR_t$	$Y_t$	$R_t$
	1	2	3	4	5	6	7			
1	-0.01	-0.04	0.08	-0.08	0.03	-0.02	0.08	2	0	20
2	-0.05	0.01	0.01	0.06	0.00	0.00	0.11	15	1	20
3	0.06	-0.05	-0.01	-0.03	0.04	0.05	0.07	13	1	140
4	0.03	-0.02	0.02	0.04	0.02	-0.01	0.08	20	3	30
5	0.02	0.07	0.01	0.03	0.01	0.01	0.08	28	4	130
6	0.11	0.10	0.08	0.10	0.11	0.04	0.06	28	6	70
7	0.08	0.11	0.12	0.12	0.13	0.12	0.06	28	7	146
8	0.07	0.04	0.06	0.07	0.07	0.05	0.07	28	9	50
9	0.04	0.01	0.01	0.00	-0.02	0.03	0.08	19	10	0
10	-0.01	0.04	0.02	0.00	0.01	0.01	0.14	21	10	110



**Figure 3.** Beverage example: individual value plot of the observations

- (as an Euclidean division) is  $Y_1 = 0$  and  $R_1 = 20$ .
- For  $t = 2$  we have  $Y_1 = 0$ ,  $R_1 = 20$ ,  $SR_2 = 15$  and the equation to be solved is  $150 \times Y_2 + R_2 = 10 \times 15 + 140 \times (0) + 20 = 170$ . The unique solution of this equation is  $Y_2 = 1$  and  $R_2 = 20$ .
  - $\vdots$
  - For  $t = 10$  we have  $Y_9 = 10$ ,  $R_9 = 0$ ,  $SR_{10} = 21$  and the equation to be solved is  $150 \times Y_{10} + R_{10} = 10 \times 21 + 140 \times (10) + 0 = 1610$ . The unique solution of this equation is  $Y_{10} = 10$  and  $R_{10} = 110$ .

In Figure 4 the values of the charting statistic  $Y_t$  are plotted. It can be seen that the control chart triggers a signal at sample #6. Anyway, the evidence of an upward trend in the EWMA statistics suggests to look for factors affecting the occurrence of assignable causes since sample #4.

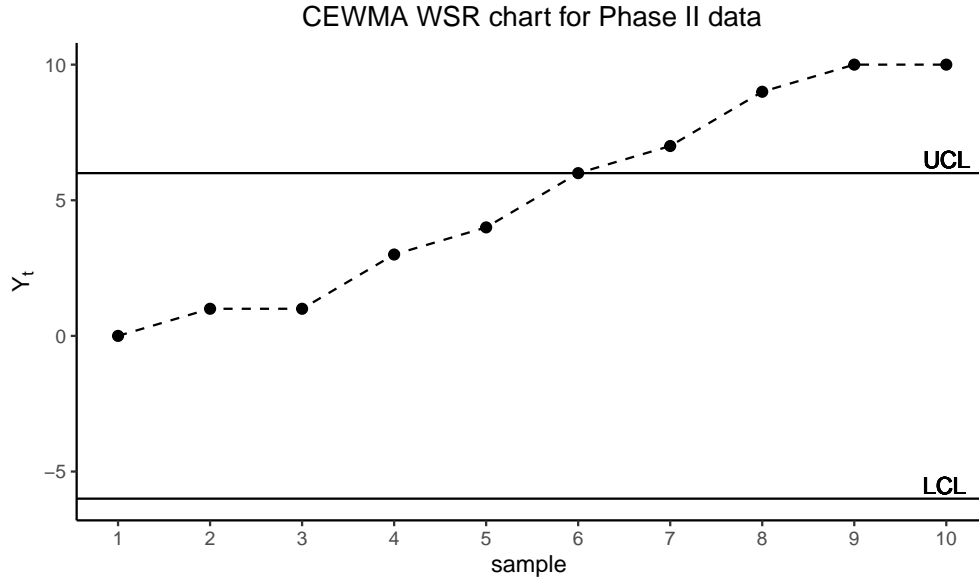


Figure 4. Beverage example: the CEWMA WSR chart for the Phase II sample presented in Table 8

## 8. Concluding Remarks

In this paper, we designed a new EWMA distribution-free control chart based on signed ranks using a new memory-type EWMA technique, introduced by Rakitzis et al. in [25] and extended by Castagliola et al. in [24] to the nonparametric case. As it was explained in the previous sections, the use of the conventional approach of Brook and Evans [21] in nonparametric EWMA charts yields to an unreliable determination of their exact RL properties. The proposed CEWMA WSR scheme is designed with guaranteed exact Run Length properties obtained by a specific finite state Markov chain approach without any approximations. In addition, for almost all the nonparametric schemes based on the Wilcoxon signed rank statistic, their out-of-control performances are examined via Monte-Carlo simulation-based techniques for a given distribution. On the other hand, for our proposed chart the run length properties are examined and computed exactly for both the in-control and the out-of-control conditions regardless of the observations' underlying distribution. Moreover, the performance of the CEWMA WSR chart was examined and compared to several existing distribution-free control charts available from the literature. Our results show that the CEWMA WSR chart can be an efficient means in detecting moderate to large shifts in the process.

Many future work directions are worth of interest in this field of research. For instance, it would be interesting to improve our chart's overall shift detection properties by using the adaptive feature in the smoothing parameter, as it was proposed in [26], where the sign statistic was used. In addition, it would be challenging to extend the CEWMA WSR chart to the multivariate case. Finally, the use of the proposed EWMA control chart can be extended to other non-parametric statistics as, for instance, the Mann-Whitney, and the Ansari-Bradley statistics.

## References

- [1] Shewhart W. Statistical Method from the Viewpoint of quality control, graduate school department of agriculture. Washington, DC. 1939;.
- [2] Roberts S. Properties of Control Chart Zone Tests. *Bell System Technical Journal*. 1958; 37(1):83–114.
- [3] Page E. Continuous Inspection Schemes. *Biometrika*. 1954;41(1/2):100–115.
- [4] Chakraborti S, Van der Laan P, Bakir S. Nonparametric Control Charts: an Overview and Some Results. *Journal of Quality Technology*. 2001;33(3):304–315.
- [5] Chakraborti S, Graham M. Nonparametric (Distribution-Free) Control Charts: An Updated Overview and Some Results. *Quality Engineering*. 2019;31(4):523–544.
- [6] Amin R, Searcy A. A Nonparametric Exponentially Weighted Moving Average Control Scheme. *Communications in Statistics-Simulation and Computation*. 1991;20(4):1049–1072.
- [7] Li S, Tang L, Ng S. Nonparametric CUSUM and EWMA Control Charts for Detecting Mean Shifts. *Journal of Quality Technology*. 2010;42(2):209–226.
- [8] Graham M, Chakraborti S, Human S. A Nonparametric EWMA Sign Chart for Location Based on Individual Measurements. *Quality Engineering*. 2011;23(3):227–241.
- [9] Graham M, Chakraborti S, Human S. A Nonparametric Exponentially Weighted Moving Average Signed-Rank Chart for Monitoring Location. *Computational Statistics & Data Analysis*. 2011;55(8):2490–2503.
- [10] Yang SF, Lin JS, Cheng S. A New Nonparametric EWMA Sign Control Chart. *Expert Systems with Applications*. 2011;38(5):6239–6243.
- [11] Graham M, Mukherjee A, Chakraborti S. Distribution-Free Exponentially Weighted Moving Average Control Charts for Monitoring Unknown Location. *Computational Statistics & Data Analysis*. 2012;56(8):2539–2561.
- [12] Aslam M, Azam M, Jun C. A New Exponentially Weighted Moving Average Sign Chart using Repetitive Sampling. *Journal of Process Control*. 2014;24(7):1149–1153.
- [13] Riaz M. A Sensitive Non-Parametric EWMA Control Chart. *Journal of the Chinese Institute of Engineers*. 2015;38(2):208–219.
- [14] Lu S. An Extended Nonparametric Exponentially Weighted Moving Average Sign Control Chart. *Quality and Reliability Engineering International*. 2015;31(1):3–13.
- [15] Chakraborty S Nand Chakraborti, Human S, Balakrishnan N. A Generally Weighted Moving Average Signed-Sank Control Chart. *Quality and Reliability Engineering International*. 2016;32(8):2835–2845.
- [16] Abid H Mand Nazir, Riaz M, Lin Z. An Efficient Nonparametric EWMA Wilcoxon Signed-Rank Chart for Monitoring Location. *Quality and Reliability Engineering International*. 2017;33(3):669–685.
- [17] Haq A. A conparametric EWMA Chart with Auxiliary Information for Process Mean. *Communications in Statistics-Theory and Methods*. 2018;49(5):1–16.
- [18] Haq A. A New Nonparametric Synthetic EWMA Control Chart for Monitoring Process Mean. *Communications in Statistics-Simulation and Computation*. 2019;48(6):1665–1676.
- [19] Raza M.A, Nawaz T, Aslam M, Bhatti S.H, Sherwani R.A.K. A New Nonparametric Double Exponentially Weighted Moving Average Control Chart. *Quality and Reliability Engineering International*. 2020;36(1):68–87.
- [20] Abbas Z, Nazir H.Z, Abid M, Akhtar N, Riaz M. Enhanced Nonparametric Control Charts Under Simple and Ranked Set Sampling Schemes. *Transactions of the Institute of Measurement and Control*. 2020;42(14):2744–2759.
- [21] Brook D, Evans D. An Approach to the Probability Distribution of CUSUM Run Length. *Biometrika*. 1972;59(3):539–549.
- [22] Weiß C. EWMA Monitoring of Correlated Processes of Poisson Counts. *Quality Technology & Quantitative Management*. 2009;6(2):137–153.
- [23] Wu S, Castagliola P, Celano G. A Distribution-free EWMA Control Chart for Monitoring Time-Between-Events-and-Amplitude Data. *Journal of Applied Statistics*.2020;1–21. doi:

10.1080/02664763.2020.1729347

- [24] Castagliola P, Tran K, Celano G, et al. An EWMA-Type Sign Chart with Exact Run Length Properties. *Journal of Quality Technology*. 2019;51(1):51–63.
- [25] Rakitzis A, Castagliola P, Maravelakis P. A New Memory-type Monitoring Technique for Count Data. *Computers & Industrial Engineering*. 2015;85:235–247.
- [26] Tang A, Sun J, Hu X, et al. A New Nonparametric Adaptive EWMA Control Chart with Exact Run Length Properties. *Computers & Industrial Engineering*. 2019;130:404–419.
- [27] Amin R, Reynolds Jr M, Bakir S. Nonparametric Quality Control Charts based on The Sign Statistic. *Communications in Statistics-Theory and Methods*. 1995;24(6):1597–1623.
- [28] McCornack R. Extended Tables of the Wilcoxon Matched Pair Signed Rank Statistic. *Journal of the American Statistical Association*. 1965;60(311):864–871.
- [29] Bennett B. On the Non-Null Distribution of Wilcoxon’s Signed Rank Test. *Metrika*. 1972; 19(1):36–38.
- [30] Bakir S. A Distribution-Free Shewhart Quality Control Chart based on Signed-Ranks. *Quality Engineering*. 2004;16(4):613–623.
- [31] Neuts M. *Matrix-Geometric Solutions in Stochastic Models—An Algorithmic Approach*. Johns Hopkins University Press. Baltimore, Md. 1981;.
- [32] Latouche G, Ramaswami V. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Vol. 5. SIAM; 1999.
- [33] Yang SF, Cheng S. A New Non-Parametric CUSUM Mean Chart. *Quality and Reliability Engineering International*. 2011;27(7):867–875.
- [34] Celano G, Castagliola P, Chakraborti S, et al. On the Implementation of the Shewhart Sign Control Chart for Low-volume Production. *International Journal of Production Research*. 2016;54(19):5886–5900.
- [35] Celano G, Castagliola P, Chakraborti S, et al. The Performance of the Shewhart Sign Control Chart for Finite Horizon Processes. *The International Journal of Advanced Manufacturing Technology*. 2016;84(5-8):1497–1512.