A NONPARAMETRIC CUMULATIVE SUM SCHEME BASED ON SEQUENTIAL RANKS AND ADAPTIVE CONTROL LIMITS

Michael Lang and Abdelhak M. Zoubir

Technische Universität Darmstadt, Germany
Signal Processing Group, 64283 Darmstadt, and
Graduate School of Computational Engineering, 64293 Darmstadt
{mlang , zoubir}@spg.tu-darmstadt.de

ABSTRACT

We consider the problem of quickest detection, i.e. we sequentially monitor a data sequence to detect a shift in the sampling distribution which may occur at an unknown time instance. Conventional quickest detection procedures typically require a-priori knowledge of the underlying pre- and post-change distributions of the process. Such knowledge may not be available in practice or be flawed, e.g. because the distributional assumptions itself or the respective parameter estimates are inadequate. In this paper we propose a distribution-free cumulative sum (CUSUM) procedure based on sequential ranks and adaptive control limits. The presented procedure does not require a historical set of training data and is therefore especially suited for initial monitoring phases.

Index Terms— Quickest Detection, Nonparametric, Sequential Ranks, Cumulative Sums

1. INTRODUCTION

Originating from an industrial statistics background, statistical process control (SPC) techniques such as Page's Cumulative Sum charts (CUSUM) [1] have found their way into numerous application areas other than quality monitoring in manufacturing lines. In fact, the need to quickly detect changes in the sampling distribution is encountered in virtually every discipline that relies on data processing, well known examples being engineering, economics, medicine and science. Essential to the usefulness of such techniques in practice is often that they be robust to flaws in the underlying distributional assumptions or inaccurate model parameter estimates.

In this paper the focus is on distribution-free CUSUM control charts. We propose a distribution-free CUSUM based on a sequential ranks test statistic and a sequence of adaptive control limits. Since no historical training data is required our procedure is especially useful in initial monitoring phases, where one may collect observations in order to construct a more powerful detection scheme to then use in the monitoring of following observations.

The remainder of this paper is organized as follows. In Section 2 the univariate sequential change detection problem and the conventional CUSUM are introduced, followed by a discussion of two nonparametric approaches in Section 3. Our proposed method is described in Section 4 while simulation and real data results are presented in Section 5 and Section 6, respectively. Section 7 completes this paper with a discussion.

This work was supported by the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt.

2. PROBLEM FORMULATION AND CUSUM CONTROL CHARTS

Consider an observed sequence $\{x(n), n \geq 1\}$ of independent random variables such that $\{x(1), \ldots, x(\tau-1)\} \sim F$ and $\{x(\tau), x(\tau+1), \ldots\} \sim G$, i.e. a distributional shift $F \to G$ occurs at time instance τ. In this paper we will monitor $\{x(n)\}$ for a shift from F to a stochastically larger distribution G.

If we were to assume both F and G to be normally distributed with known parameters, Page's CUSUM [1] can be regarded as the standard change detection technique and can be computed sequentially as

$$C_{\mathrm{C}}(0) = 0, \quad C_{\mathrm{C}}(n) = \max\{0, C_{\mathrm{C}}(n-1) + x(n) - k_{\mathrm{C}}\}, \quad n \geq 1 \tag{1}$$

A distributional shift is declared if $C_{\mathrm{C}}(n) > h_{\mathrm{C}}$, with h_{C} and k_{C} being the pre-specified control limit and reference constant, respectively. Hereby, h_{C} and k_{C} are chosen such that a nominal in-control average run length (ARL) ARL_0 is attained, cf. [2]. The in-control ARL is defined as the expected time until a change is signaled under F, i.e.

$$ARL = E_F \inf\{n > 0 : C_{\mathrm{C}}(n) > h_{\mathrm{C}}\}. \tag{2}$$

Intuitively this can be thought of as the equivalent of setting a nominal type-I error level in hypothesis testing. The closeness of the actual in-control ARL to ARL_0 can be regarded as an indicator of the CUSUM chart's robustness [3]. It is well known that, under some regularity conditions, choosing $k_{\mathrm{C}} = \delta/2$, with δ being the shift in the transition $F \to G$, is optimal [4].

3. NONPARAMETRIC APPROACHES

Optimality of the conventional CUSUM comes at the price of a considerable required a-priori knowledge concerning the process distribution. CUSUM's sensitivity to deviations from the assumption that both F and G are normally distributed with known parameters is well documented [5, 3]. Various nonparametric change-point procedures have been proposed in the literature, yet the topic appears to be less well explored than parametric counterparts. Nonparametric adaptations of Page's CUSUM usually involve sign or rank statistics. For a thorough overview of the existing literature we refer to [6, 7]. Without claiming to have provided a complete and exhaustive representation of the field we would like to point to relevant contributions by Gordon and Pollak. In [8] the said authors proposed a Shiryaev-Roberts (SR) procedure wherein a finite sequence of likelihood ratios (LR) of sequential vectors of signs and ranks takes the place of the LRs in the parametric SR procedure.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

In the following subsections we introduce two nonparametric procedures: a Bootstrap-Aided CUSUM (BAC) using a sequence of control limits instead of a fixed threshold, as introduced by Chatterjee and Qiu [3], and McDonald's Sequential Ranks CUSUM [9] (SRC).

3.1. Sequential Ranks CUSUM (SRC)

The sequential rank of $x(n)$ is defined as

$$R(n) = 1 + \sum_{j=1}^{n-1} (x(n) - x(j))^{+} \quad (3)$$

where $(x)^{+}$ is 1 for $x > 0$ and 0 otherwise. The SRC is then

$$C_{\text{SRC}}(n) = \max\{0, C_{\text{SRC}}(n-1) + \frac{R(n)}{n+1} - k_{\text{SRC}}\}, n \geq 1 \quad (4)$$

with $C_{\text{SRC}}(0) = 0$ and k_{SRC} again being the reference constant. The SRC signals a change if $C_{\text{SRC}}(n)$ exceeds a predetermined control limit h_{SRC}.

It can be shown [9] that, given the observed process is in-control, the quantities $\frac{R(n)}{n+1}$ are independent and discrete uniform on

$$\{\frac{1}{n+1}, \frac{2}{n+1}, \cdots, \frac{n}{n+1}\}$$

Herein lies a key advantage of the SRC. In fact, for a fixed reference value k_{SRC} we can obtain the control limit h_{SRC} and thus construct the sequential ranks CUSUM as follows without any historical training data being required. Set a constant N_{SRC}, then construct the set of random variables $\{U(n)\}_{n=1}^{N_{\text{SRC}}}$ as discrete uniform on $\{\frac{1}{n+1}, \frac{2}{n+1}, \cdots, \frac{n}{n+1}\}$ and construct

$$\tilde{C}_{\text{SRC}}(n) = \max\{0, \tilde{C}_{\text{SRC}}(n-1) + U(n) - k_{\text{SRC}}\} \quad (5)$$

Extract the maximum value of $\{\tilde{C}_{\text{SRC}}(n)\}_{n=1}^{N_{\text{SRC}}}$ and repeat the above steps B times. Finally, set the control limit h_{SRC} as the $B \cdot (1 - ARL_0{}^{-1})$ ordered extracted maximum value.

3.2. Bootstrap-Aided CUSUM (BAC)

Chatterjee and Qiu [3] proposed a nonparametric CUSUM where instead of a fixed decision interval a sequence of control limits, determined by the bootstrap estimate of the conditional distribution of the CUSUM test statistic given the last time it was zero, is used.

Starting from the conventional CUSUM formulation as in Eq. (1)

$$C_{\text{BAC}}(n) = \max\{0, C_{\text{BAC}}(n-1) + x(n) - k_{\text{BAC}}\}, n \geq 1 \quad (6)$$

with $C_{\text{BAC}}(0) = 0$ we define the sprint length $T_{\text{BAC}}(n)$, to express the time elapsed since $C_{\text{BAC}}(n)$ was last zero.

$$\begin{aligned} T_{\text{BAC}}(n) &= 0 \text{ if } C_{\text{BAC}}(n) = 0 \\ T_{\text{BAC}}(n) &= j \text{ if } C_{\text{BAC}}(n) \neq 0, \ldots, C_{\text{BAC}}(n-j+1) \neq 0, \\ & \qquad C_{\text{BAC}}(n-j) = 0; j = 1, 2, \ldots, n \end{aligned}$$

Furthermore, let us denote by Y_{BAC_j} a random variable with distribution $[C_{\text{BAC}}(n) | T_{\text{BAC}}(n) = j]$,

$$Y_{\text{BAC}_j} \sim [C_{\text{BAC}}(n) | T_{\text{BAC}}(n) = j] \quad (7)$$

It can be shown that, as opposed to the unconditional distribution of $C_{\text{BAC}}(n)$, the conditional distributions $[C_{\text{BAC}}(n) | T_{\text{BAC}}(n) = j]$ are easier to handle and under some regularity conditions depend only on j and F, but not on n [3]. Then, for any positive integer $j_{\max} \leq n$, the distribution of $C_{\text{BAC}}(n)$ can be expressed as

$$C_{\text{BAC}}(n) \sim \sum_{j=1}^{j_{\max}} Y_{\text{BAC}_j} I_{T_{\text{BAC}}(n)=j} + Y^{*}_{\text{BAC}} I_{T_{\text{BAC}}(n)>j_{\max}} \quad (8)$$

$$Y^{*}_{\text{BAC}} \sim C_{\text{BAC}}(n) | T_{\text{BAC}}(n) > j_{\max} \quad (9)$$

Chatterjee and Qiu argue that, instead of a single control limit h_{BAC}, one may construct a sequence of control limits $\{h_{\text{BAC}_j}\}_{j=1}^{j_{\max}}$ based on the distribution of Y_{BAC_j}. To limit computational complexity, control limits are calculated up to a reasonably small value $j_{\max}$ after which the control limit is kept fixed at its last value $h_{\text{BAC}_{j_{\max}}}$.

Prior to determining the sequence of control limits $\{h_{\text{BAC}_j}\}_{j=1}^{j_{\max}}$ we have to set k_{BAC} and $j_{\max}$. The latter is essentially chosen according to the available computational power. With no knowledge regarding the shift δ (cf. Section 2), k_{BAC} should be calibrated according to the average sprint length $E\ T_{\text{BAC}}$ (cf. [3] for a detailed discussion). Note that an increase of k_{BAC} increases the probability that $C_{\text{BAC}}(n) = 0$, thus yielding a smaller $E\ T_{\text{BAC}}$. Vice versa, a smaller k_{BAC} diminishes the chances of the test statistic to hit the lower bound 0, thereby increasing the average sprint length. Clearly the agility of the detection procedure can be influenced by appropriately calibrating k_{BAC} as follows: set an initial value k_{BAC_0} and draw a number B_1 of bootstrap samples from the historical training data. Then run the BAC on each of these bootstrap samples and save the time index when C_{BAC} first bounces back to zero. Take the mean of the so obtained B_1 sprint lengths as the average sprint length and repeat the above steps in a binary search algorithm until the average sprint length is close enough to the desired value, e.g. $\lfloor \frac{3 j_{\max}}{4} \rfloor$.

Having set k_{BAC} and $j_{\max}$, $\{h_{\text{BAC}_j}\}_{j=1}^{j_{\max}}$ is determined as follows

- for each $j \in \{1, \ldots, j_{\max}\}$
  - **Initialization**: $C^{*}_{\text{old}} = 0, T^{*}_{\text{old}} = 0, b = 0$
  - **Step 0**: set $b = b + 1$;
  - **Step 1**: draw a bootstrap sample X^{*} from X.
  - **Step 2**: update $C^{*}_{\text{new}} = \max\{C^{*}_{\text{old}} + X^{*} - k, 0\}$ and set $T^{*}_{\text{new}} = T^{*}_{\text{old}} + 1 \quad if \quad C^{*}_{\text{new}} > 0$ and $T^{*}_{\text{new}} = 0 \quad if \quad C^{*}_{\text{new}} = 0$.
  - **Step 3**: if $T_{\text{new}} = j$ record $Y_{\text{BAC}_{j:b}} = C^{*}_{\text{new}}$ and return to step 0 if $b < B$ else update C^{*} and T^{*} and return to step 1
- once the above steps are performed we end up with B numbers $Y_{\text{BAC}_{j:1}}, Y_{\text{BAC}_{j:2}}, \ldots, Y_{\text{BAC}_{j:B}}$
- take the $B(1 - ARL_0{}^{-1})$ ordered value of $Y_{\text{BAC}_{j:1}}, Y_{\text{BAC}_{j:2}}, \ldots, Y_{\text{BAC}_{j:B}}$ as control limit h_{BAC_j}
- set $h_{\text{BAC}_j} = h_{\text{BAC}_{j_{\max}}}$ for $j > j_{\max}$

Then, the BAC chart signals if

$$\begin{aligned} T_{\text{BAC}}(n) &= j \text{ and } C_{\text{BAC}}(n) > h_{\text{BAC}_j} \text{ for } 1 \leq j \leq j_{\max} \\ T_{\text{BAC}}(n) &> j_{\max} \text{ and } C_{\text{BAC}}(n) > h_{\text{BAC}_{j_{\max}}} \end{aligned}$$

As a final step, we subject the obtained control limits $\{h_{\text{BAC}_j}\}_{j=1}^{j_{\max}}$ to a further fine-tuning using a binary search algorithm, wherein similarly to the empirical calibration of k_{BAC} we run the procedure on B_1 bootstrap resamples of the training data and fine tune our sequence of control limits such that the actual ARL is close enough to the nominal counterpart ARL_0.

The advantage of having an adaptive threshold depending on the elapsed time since $C_{\text{BAC}}(n)$ last bounced back to 0 is illustrated in Figure 1 where we chose $k_{\text{BAC}} = k_{\text{C}}$, i.e. we did not empirically calibrate k_{BAC}, yielding the special case $C_{\text{BAC}}(n) = C_{\text{C}}(n)$. The tracking statistic of both CUSUMs is represented by the solid line ($C_{\{\text{BAC,C}\}}$) whereas the respective control limit(s) are depicted by dashed (h_{BAC}) and dash dotted (h_{C}) line. In this example, a positive shift in the mean occurs at time instance 50.

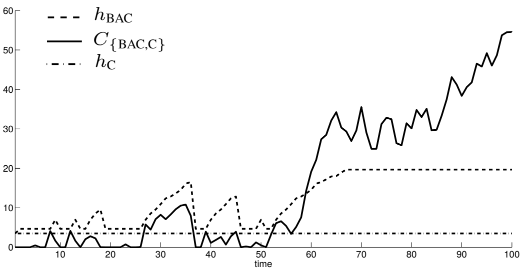

Fig. 1. adaptive (h_{BAC}) vs. fixed (h_{C}) control limit

4. THE PROPOSED METHOD (AC-SRC)

We propose merging the SRC and BAC and - before formally introducing it - would like give the reasoning behind it. It turns out that the BAC is a powerful and and highly tunable procedure with the caveat however that this is only true if it has been well adjusted to the particular problem at hand. The latter requires a reasonably large set of training data and computationally expensive refinements. This is not the case for the SRC. In fact, setting up the SRC requires neither training data nor is the algorithm to get h_{SRC} for an arbitrarily fixed k_{SRC} computationally expensive. This leads us to strongly recommend the SRC as a control chart for startup use.

A simple concatenation of SRC and BAC, i.e. use the SRC while training data is gathered (thereby checking that the training data we accumulate is indeed in-control) and the BAC is being calibrated, seems both appealing and intuitive. While such a combination is feasible and makes sense (the alternative would be to just require by definition that the historical data collected be in-control, as is often done and as a matter of fact has been done by Chatterjee as well), problems arise if very early in the process of collecting the allegedly in-control training data there were an abrupt change.

Whilst we recommend use of the SRC in a start-up phase, we argue that it first needs to be revised in order to reduce (at least to some degree) its inadequate performance given that a change occurs soon after SRC monitoring begins. The aim of the detection scheme we present in the following is meant to address exactly this issue.

We propose a combination of SRC and BAC wherein we use control limits $\{h_{\text{AC-SRC}_j}\}_{j=1}^{j_{\max}}$ obtained from

$$Y_{\text{AC-SRC}_j} \sim \left[\tilde{C}_{\text{AC-SRC}}(n) \,|T_{\text{AC-SRC}}(n) = j\right] \tag{10}$$

similar to the algorithm described in Section 3.2 (in that here we repeatedly generate $\{U(n)\}_{n=1}^{N_{\text{AC-SRC}}}$ instead of bootstrapping historical data). $\tilde{C}_{\text{AC-SRC}}(n)$ is defined analogously to Eq. (5).

As we suggest this CUSUM scheme to be used while the actual BAC is collecting training data and being computed, we will design it with a relatively short ARL_0 (i.e. smaller than the amount of training data we want to collect for the BAC) and set $k_{\text{AC-SRC}}$ such that the average sprint lengths will be small too (cf. Section 3.2) in order to obtain a more agile detection scheme. Keep in mind that, even with perfectly in-control data, every CUSUM chart will eventually signal a change. Therefore, since we design with a relatively small ARL_0 to begin with, it's crucial that when collecting long training data sets we restart our procedure periodically (if no change was signaled), e.g. after having collected observations in the range of say $\lfloor \frac{ARL_0}{2} \rfloor$ to $\lfloor \frac{3 \cdot ARL_0}{4} \rfloor$.

5. SIMULATION RESULTS

In this section we present results of simulations we carried out in order to assess the detection delay (DD) and the ARL. We interpret the closeness of the actual ARL to ARL_0 as a measure of robustness. We applied the above mentioned CUSUM schemes to $N_{\text{ph2}} = 500$ samples of a process, where at time $\tau = 20$ the distribution shifts from $F \sim N(0,1) \longmapsto G \sim N(1,1)$. For the conventional CUSUM we chose $k_{\text{C}} = 0.5$ and $h_{\text{C}} = 3.5$. For the SRC we used $k_{\text{SRC}} = 0.642$ and $h_{\text{SRC}} = 1$, as recommended by McDonald [9] if $ARL_0 = 200$. All presented CUSUM charts were designed with a nominal $ARL_0 = 200$ in mind. Results were averaged over 1000 Monte-Carlo runs. Furthermore, for the BAC and AC-SRC we set $j_{\max} = 8$ and empirically determined the respective k such that $ET_n = 6$. $N = 1000$ training data samples were assumed available, similarly we set $N_{\text{AC-SRC}} = N_{\text{SRC}} = 1000$, $B = 1000$, $B_1 = 100$.

To examine robustness we contaminated our observations using a two-term Gaussian Mixture Model. For Table 2 we contaminated our data using a two-term Gaussian mixture model, i.e. our in-control distribution can be expressed as

$$f = (1-\eta)\mathcal{N}(0,1) + \eta\mathcal{N}(0,\kappa) \tag{11}$$

with $0 \leq \eta \leq 1$ being the mixture probability. We used a mixing probability of $\eta = 0.2$ and $\kappa = 100$.

We start by focusing on Table 1 which displays DD and the actual in-control ARL for the conventional parametric CUSUM (C), McDonald's sequential ranks CUSUM (SRC), Chatterjee's bootstrap-based CUSUM (BAC) and our proposed scheme (AC-SRC).

	C	**SRC**	**AC-SRC**	**BAC**
DD	5.8266	29.7385	14.9470	6.8212
ARL	196.3780	268.6330	246.6070	224.9200

Table 1. no contamination

As was to be expected, under the ideal scenario $F \sim N(0,1) \mapsto G \sim N(1,1)$ with $k_{\text{C}} = 0.5$, C clearly outperforms its competitors yielding the smallest detection delay as well as an ARL that is closest to the nominal $ARL_0 = 200$. Note that the BAC yielded a detection delay that compares very favorably to the one attained by C. As for the ARL, the average run length attained by the BAC is close enough to ARL_0 and consistent with the margins of error we allowed for it in the fine-tuning of $\{h_{\text{BAC}_j}\}_{j=1}^{j_{\max}}$ (cf. Section 3.2). Focusing on the SRC's DD its aforementioned poor performance for small τ (here

$\tau = 20$) becomes evident. Comparing the SRC with our proposed method AC-SRC we point out that AC-SRC performed significantly better with respect to both DD and ARL.

	C	SRC	AC-SRC	BAC
DD	3.8678	112.1189	67.5657	21.6258
ARL	11.6310	243.0950	219.5240	225.4710

Table 2. 20% contamination

We now focus on Table 2 to discuss the impact of deliberately injected noise. Obviously the closer a respective result is to its counterpart in Table 1 the better. The impact the deviation from its distributional assumptions has for C is evident and devastating - C breaks down. This however comes as no surprise. It's worth re-emphasizing that the ARL expresses the average time elapsed before the monitoring scheme to yields a false-positive. Note that C in the scenario considered in Table 2 on average signals after about 12 observations as opposed to the aspired $ARL_0 = 200$. As for BAC, SRC and AC-SRC we note that qualitatively the assertions with regard to Table 1. hold for Table 2. as well, a finding that is consistent with the distribution-free nature of BAC, SRC, and AC-SRC.

6. APPLICATION EXAMPLE: C. TRACHOMATIS INFECTIONS IN THE U.K.

Chlamydia trachomatis infections have been recognized as the most prevalent sexually transmitted infection (STI) both in industrialized and developing countries and are a common cause of urethritis, cervicitis, conjunctivitis etc., cf. [10] and references therein for a detailed (medical) discussion of the topic.

Figure 2 shows the annual incidence rate of Chlamydia infections per 100,000 population for the U.K. in the years 1988-2012 (no data for 2011) [11].

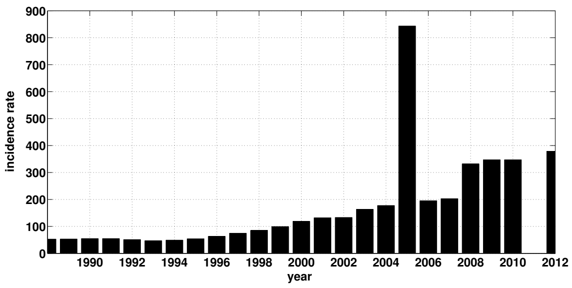

Fig. 2. incidence rate per 100,000 population

SPC techniques are widely used for infectious disease monitoring [12]. In the following we will apply the SRC and AC-SRC to the incidence data as illustrated in Figure 2. Note that for the particular real-world problem at hand we assume our knowledge is limited to the sequential observation of the incidences in Fig. 2. An immediate implication then is that, since we stipulate to have no knowledge of and make no distributional assumptions about the specific infectious disease being monitored and have no historical disease data either, it follows that C and BAC as they were presented here are not suitable to the problem under investigation.

We assume an expert epidemiologist to have determined (retrospectively, e.g. to validate our findings) through visual inspection that a signal should be raised at the 10th sample, i.e. 1997. Note that the question of what a (communicable disease) outbreak is and when exactly an alarm should be raised is not well settled and retrospective visual analysis of the data by an expert epidemiologist is common practice [13]. We acknowledge that our hypothetical expert's choice is somewhat arbitrary (there is however a clear trend starting around 1996), however necessary for the sake of algorithm comparison. We apply the SRC and the AC-SRC to the given data. We set $ARL_0 = 30$, $N_{\text{SRC}} = 100$, $k_{\text{SRC}} = 0.5$, $B = 1000$, $B_1 = 100$, $j_{\max} = 6$, $ET_n = 4$ and determine h_{SRC} and $\{h_{\text{AC-SRC}_j}\}_{j=1}^{j_{\max}}$ as outlined in Section 3.1 and Section 4, averaging over 1000 Monte-Carlo runs. Results are shown in Table 3.

	SRC	AC-SRC
DD	13	1
ARL	396.4871	30.2510

Table 3. SRC vs. AC-SRC for the Chlamydia data

This application example, firstly, makes a strong and vivid case for the importance of methods that fit well into challenging real-world scenarios and are able to cope with respective implications adequately rather than by imposing various modeling assumptions or requirements that tend to contradict reality. Secondly, following up on our discussion in Section 5, we point out the remarkable improvement our proposed method AC-SRC is able to attain compared to the SRC. In fact, on average the AC-SRC signals at year 1998, i.e. with a detection delay of 1 observations as opposed to an average detection delay of 13 for the SRC. Besides the improvement in performance, it should be noted that AC-SRC, as opposed to SRC, allows for better and easier control of performance characteristics. In fact, while AC-SRC enables a relatively easy and precise control of the actual ARL, this is not true for the SRC (cf. [9]).

7. DISCUSSION

We proposed a distribution-free CUSUM procedure based on sequential ranks and adaptive control limits, which does not require training data and is therefore especially suited for initial monitoring phases. In our simulations the BAC outperformed nonparametric competitors. This was to be expected since the BAC is known to perform well, given that a reasonably large amount of training data is available and that it has been well calibrated. The SRC does not require historical data and is computationally inexpensive. Performance of the SRC is however inadequate if a change occurs soon after we start monitoring. Furthermore, the SRC exhibits substantial discrepancies between the actual ARL and ARL_0. Results from a numerical simulation study as well as a real-world application indicate that our proposed procedure is able to successfully tackle these shortcomings, significantly reducing the detection delay compared to the SRC and yielding average run lengths close to ARL_0. Consistent results were obtained for other values of $\tau, j_{\max}, ET_n$, given that they were relatively small.

8. REFERENCES

[1] E. S. Page, “Continuous inspection schemes,” *Biometrika*, vol. 41, no. 1/2, pp. pp. 100–115, 1954.

[2] J. M. Lucas, “The design and use of cumulative sum control schemes,” *Technometrics*, , no. 14, pp. 51–59, 1976.

[3] S. Chatterjee and P. Qiu., “Distribution-free cumulative sum control charts using bootstrap-based control limits,” *Ann. Appl. Stat*, vol. 3, no. 1, pp. 349–369, 2009.

[4] M. R. JR. Reynolds, “Approximations to the average run length in cumulative sum control charts,” *Technometrics*, , no. 17, pp. 65–71, 1975.

[5] A. Gandy and J. T. Kvaloy, “Guaranteed conditional performance of control charts via bootstrap methods,” *Scandinavian Journal of Statistics*, vol. 40, no. 4, pp. 647–668, December 2013.

[6] B. E. Brodsky and B. S. Darkhovsky, *Nonparametric methods in change-point problems*, Springer, 1993.

[7] P. van der Laan S. Chakraborti and S. T. Bakir, “Nonparametric control charts: An overview and some results.,” *J. Quality Technology*, vol. 33, no. 3, 2001.

[8] L. Gordon and M. Pollak., “An efficient sequential nonparametric scheme for detecting a change of distribution,” *Ann. Statist.*, vol. 22, no. 2, pp. 763–804, 1994.

[9] D.R. McDonald, *A Cusum Procedure Based on Sequential Ranks*, Laboratory for Research in Statistics and Probability, Carleton University, 1985.

[10] K. K. Holmes, P. F. Sparling, W. E. Stamm, P. Piot, J. N. Wasserheit, L. Corey, M. S. Cohen and D. H. Watts, *Sexually Transmitted Diseases*, McGraw-Hill, 4th edition, 2007.

[11] “Centralized Information System for Infectious Diseases, Communicable Diseases, Surveillance and Response. WHO Regional Office for Europe,” http://data.euro.who.int/cisid/, Online; accessed September-2014.

[12] R. Brookmeyer and D. F. Stroup, *Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance*, Oxford University Press, 1st edition, 2003.

[13] R. E. Watkins, S. Eagleson, B. Veenendaal, G. Wright and A. J. Plant, “Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia,” *BMC Medical Informatics and Decision Making*, vol. 8, no. 37, 2008.