IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 13, 2019, accepted August 28, 2019, date of publication September 24, 2019, date of current version October 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943446

Beyond Cumulative Sum Charting in Non-Stationarity Detection and Estimation

FELIX ZHAN[1], ANTHONY MARTINEZ[1], NILAB RAI[1], RICHARD MCCONNELL[1], MATTHEW SWAN[1], MOINAK BHADURI[2], JUSTIN ZHAN[3], LAXMI GEWALI[1], AND PAUL OH[1]

[1]University of Nevada, Las Vegas, NV 89154, USA
[2]Bentley University, Waltham, MA 02452, USA
[3]University of Arkansas, Fayetteville, AR 72701, USA

Corresponding author: Justin Zhan (justinzzhan@gmail.com)

This work was supported in part by the United States Department of Defense under the Army Educational Outreach Program (AEOP) under Grant W911NF1910482, Grant W911NF-17-1-0088, and Grant W911NF-16-1-0416, and in part by the National Science Foundation under Grant 1625677 and Grant 1710716.

ABSTRACT In computer science, stochastic processes, and industrial engineering, stationarity is often taken to imply a stable, predictable flow of events and non-stationarity, consequently, a departure from such a flow. Efficient detection and accurate estimation of non-stationarity are crucial in understanding the evolution of the governing dynamics. Pragmatic considerations include protecting human lives and property in the context of devastating processes such as earthquakes or hurricanes. Cumulative Sum (CUSUM) charting, the prevalent technique to weed out such non-stationarities, suffers from assumptions on a priori knowledge of the pre and post-change process parameters and constructs such as time discretization. In this paper, we have proposed two new ways in which non-stationarity may enter an evolving system - an easily detectable way, which we term strong corruption, where the post-change probability distribution is deterministically governed, and an imperceptible way which we term hard detection, where the post-change distribution is a probabilistic mixture of several densities. In addition, by combining the ordinary and switched trend of incoming observations, we develop a new trend ratio statistic in order to detect whether a stationary environment has changed. Surveying a variety of distance metrics, we examine several parametric and non-parametric options in addition to the established CUSUM and find that the trend ratio statistic performs better under the especially difficult scenarios of hard detection. Simulations (both from deterministic and mixed inter-event time densities), sensitivity-specificity type analyses, and estimated time of change distributions enable us to track the ideal detection candidate under various non-stationarities. Applications on two real data sets sampled from volcanology and weather science demonstrate how the estimated change points are in agreement with those obtained in some of our previous works, using different methods. Incidentally, this study sheds light on the inverse nature of dependence between the Hawaiian volcanoes Kilauea and Mauna Loa and demonstrates how inhabitants of the now-restless Kilauea may be relocated to Mauna Loa to minimize the loss of lives and moving costs.

INDEX TERMS Non-stationarity, classification, CUSUM chart, change points, strong corruption, weak corruption, trend reversal, distribution-free methods, mixture densities.

I. INTRODUCTION

Future decisions and plans are made based on the assumptions of an underlying model. When the parameters of the underlying statistical model change, enacting the future plan can have significant consequences. In 1913, there was strong statistical evidence suggesting that the rate of mining accidents in the UK was increasing. However, because no action was taken, this resulted in a mining accident that took the lives of over 400 people [1]. Detecting and estimating the changes

The associate editor coordinating the review of this manuscript and approving it for publication was Lei Wu.

in parameters of an underlying model is known formally as the change point detection and estimation problem.

The change point detection and estimation problem is concerned with detecting and estimating points in time in which the model for a random process changes. Formally speaking, let $T = \{\tau_i | i \in N_0\}$ be an increasing enumeration of non negative integers with $\tau_0 = 0$. For $i \in N_0$ let F_i be an arbitrary cdf, and define the random variable X_i by

$$X_i \sim F_i \text{ if } \tau_i < i \leq \tau_{i+1} \tag{1}$$

where the $F_i's$ are not necessarily distinct. At a given time t with $\tau_i < t \leq \tau_{i+1}$ samples from X_i are observed. Under the above scenario the change point detection problem is to determine if there exist an i such that $F_i \neq F_{i+1}$ while the change point estimation problem is concerned with constructed the set $T' := \{\tau_i | i > 0, F_i \neq F_{i+1}\}$. The problem is often looked though the lens of a single change detection, since if a method can detect a single change, then applying it multiply times provides a solution to the general problem.

Due to the general setting in which change point detection and estimation is defined, it has found uses in many fields. In the field of financial security, Bolton and Hand [2] used change point detection to detected fraudulent credit card purchases. Climatologists use change point detection for the analysis of global climate and temperature series [3], [4]. It has also found use in image analysis [6], speech recognition [5], and human activity analysis [7]. Problem-specific peculiarities originate from sources where reliable detection with small sample sizes is necessary. Tan *et al.* [11] and Ho *et al.* [10] review such cases with strong sandstorms and bank failures. Efficient estimation naturally improves other performances that depend on it - drift detection (Bhaduri *et al.* [15]), weak estimation (Bhaduri *et al.* [14], Zhan *et al.* [17]), time series clustering (Bhaduri and Zhan [49]), to name a few.

The rest of the paper is formatted as follows. Section two is devoted a review of previous methods and the establishing nine new methods for the change point detection problem. Section three is devoted to analyzing the proposed methods on simulated data using two new ways of introducing non-stationarity into an evolving environment, while section four concerns analyzing real data sets. Finally, in section five we summarize our conclusions along with future directions.

II. THEORY AND METHODS

Various change detection algorithms, from mathematical statistics, computer science, physics are in constant use. Xie *et al.* [55], Sugiyama and Kawanabe [56], and Imani *et al.* [57] offer excellent exposure to the latest advancements, through tools such as linear discriminant-based Kalman smoothers and the Bayesian paradigm-inspired Markov Decision Processes. Methods are being developed to tackle specific data types such as panel data (Bardwell *et al.* [58]). This section lays out the fundamentals necessary to understand the formulation and surveys several relevant candidates. The previous section has briefly touched upon the problem and one should, in the spirit of Ross [31], start by being aware of two distinct approaches used to tackle the analysis. A detailed description of the literature review to follow may be found in Bhaduri [16].

A. BATCH DETECTION SCENARIO

With F_is representing cumulative distribution functions of inter-event times, change detection under this framework amounts to choosing one of the following competing hypotheses:

$$H_0 : X_i \sim F_0(x; \theta_0), \quad i = 1, 2, .., n \tag{2}$$

$$H_1 : X_i \sim \begin{cases} F_0(x; \theta_0), & i = 1, 2, \ldots, k \\ F_1(x; \theta_1), & i = k+1, k+2, ..n. \end{cases} \tag{3}$$

Here, it is assumed that there exists only one change point, which occurs immediately after the kth observation in a fixed sample of size n. This point might be triggered, under a parametric setup, by changing θ_0 to θ_1. Depending on the change point, the variables are independently and identically distributed according to some F_i ($i = 0$ before the change, and $i = 1$ after it).

Assuming normal-like conditions of the probability distributions of the inter-event times, two sample t or F tests are often used, in case the θs represent location or scale parameters. In the absence of such knowledge, Mann-Whitney or Mood tests can be employed to detect possible location or scale updates, and other non-parametric options such as Lepage, Kolmogorov-Smirnov or Cramer Von-Misses, to unearth more intricate structural changes. These statistics have been elaborated in section (2.3). In theory, a two sample statistic $D_{k,n}$ is devised, and subsequently studied for signs of large values. This, due to its construction described later, signals dissimilarity between the pre-change sample (those before the kth observation) and the post-change sample (those after the kth observation), and consequently, a drift in the underlying model. A tolerance level $h_{k,n}$ helps to quantify alarmingly large or small values. For ready implementation, a working statistic D_n is created by choosing the largest of the $D_{k,n}$'s:

$$D_n = \max_{k=2,3,..,n-1} D_{k,n} = \max_{k=2,3,..,n-1} \left| \frac{\hat{D}_{k,n} - \mu_{\hat{D}_{k,n}}}{\sigma_{\hat{D}_{k,n}}} \right| \tag{4}$$

since the true value of k is unknown, and in fact, almost our target. A natural estimate for the true change point τ will thus, be

$$\tau = argmax_{k=2,3,..,n-1} D_{k,n} \tag{5}$$

Here $\mu_{\hat{D}_{k,n}}$ and $\sigma_{\hat{D}_{k,n}}$ represent the average and the standard deviations of $\hat{D}_{k,n}$'s. A general formal test runs thus:

$$\phi(D_n) = \begin{cases} 1 & \text{if } D_n > h_n \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where h_n is chosen to satisfy the level α condition, typically as the upper α point of the null density of D_n.

In this distribution, the probability density under the null assumption of no change (i.e. stationarity) is often intractable for several popular choices of the two sample statistics $D_{k,n}$s. Hawkins [32] and Pettitt [33] provide large sample approximations for the t and Mann-Whitney choice of the $D_{k,n}$s. Asymptotic bounds for a class of other choices may be had from Worsley [34]. The CPM framework detailed later, exploits numerical simulations to estimate the null densities for small sample sizes.

B. SEQUENTIAL DETECTION SCENARIO

The sequential framework, unlike the batch model, assumes a continually increasing sample size, with observations pouring in completely at random on the time axis. The two sample approach described in the previous subsection can be generalized here as follows: as the tth observation x_t arrives, one may treat $\{x_1, x_2, .., x_t\}$ as a t-length set, and compare D_t to h_t with $D_t > h_t$ signaling significant change. The attractive computational efficiency with updating D_t to D_{t+1} may be estimated from Hawkins *et al.* [25] and Ross *et al.* [28]. The thresholds h_t's are not straightforward to calculate unlike the batch scenario, since the Type-I error gets inflated due to a series of dependent tests. A common choice is to make this error time-homogeneous, i.e.

$$P(D_1 > h_1) = \alpha \tag{7}$$

$$P(D_t > h_t | D_{t-1} > h_{t-1}, .., D_1 > h_1) = \alpha, \quad t > 1 \tag{8}$$

for a fixed α. The average run length (ARL) is defined as the mean number of observations scanned before sounding a false alarm. Under the null assumption, this can be proved to be $1/\alpha$, while under the alternate assumption of a change, this quantity should preferably, be high. Summarized versions of the above conditional distributions (created using Monte Carlo simulations) are stored as lookup tables within the CPM package in R.

C. CHANGE-POINT MODEL (CPM) FRAMEWORK

In addition to the sample size type (i.e., whether fixed or not), it is imperative to have some insights into the nature of model parameters. Hawkins *et al.* [25] introduces this CPM framework with a normal choice for F, i.e. with

$$X_i \sim \begin{cases} N(\mu_1; \sigma_1^2), & i = 1, 2, \ldots, \tau \\ N(\mu_2; \sigma_2^2), & i = \tau + 1, \tau + 2, ..n. \end{cases} \tag{9}$$

This work identifies the following scenarios:

i) Complete knowledge about process parameters

Here, μ_1, μ_2 and $\sigma = \sigma_1 = \sigma_2$ are assumed to be known completely. The only unknown quantity to estimate is the change-point location τ. The Cumulative Sum (CUSUM) chart, constructed using:

$$S_0 = 0 \tag{10}$$

$$S_i = max(0, S_{i-1} + X_i - k) \tag{11}$$

with $k = \frac{\mu_1 + \mu_2}{2}$ is the most widely used method to detect a drift under such knowledge. A shift in mean from μ_1 to μ_2 with $\mu_2 > \mu_1$ is indicated if $S_i > h$ where h (constructed using known parameters μ and σ) is created to fix ARL_0 at some predefined level. Two typical CUSUM charts are graphed below in Figs (1, 2), one indicating a stable stationary flow (with all the data points contained by the two bounds), and the other, a non-stationary process (with at least one point going beyond the threshold).

More about this technique and its theoretical attractiveness can be found in Lai [35] and Hawkins and Olwell [36].

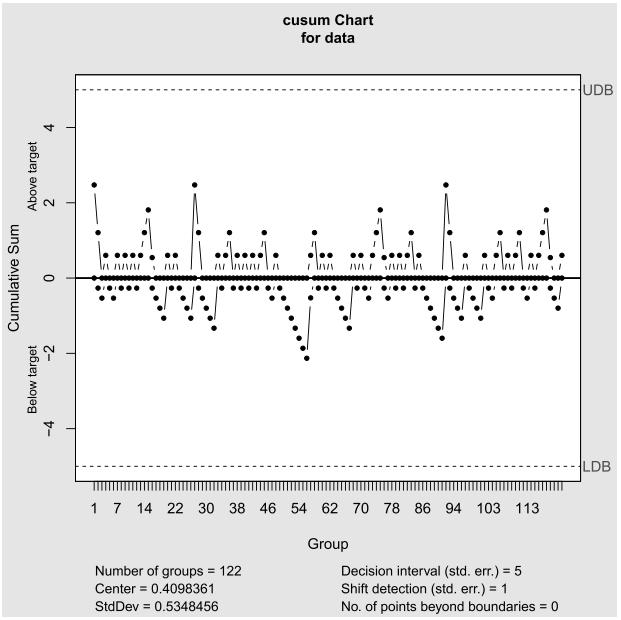

FIGURE 1. CUSUM chart depicting stationarity.

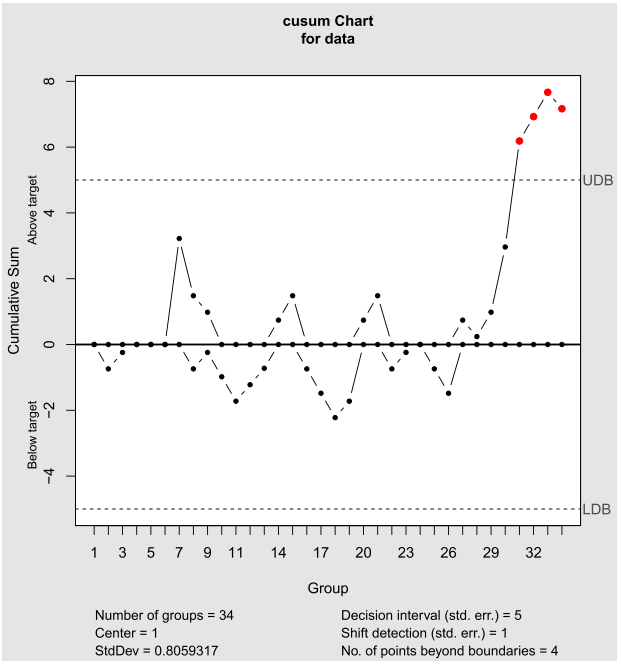

FIGURE 2. CUSUM chart depicting non-stationarity.

The Exponentially Weighted Moving Average (EWMA) procedure is closely related and similarly depends on complete knowledge of process parameters.

ii) Partial knowledge about process parameters

As an improvement, Lai [35] allowed the post change process mean μ_2 be unknown. Generalized likelihood ratio tests were conducted using the maximum likelihood estimates of μ_2 and τ to test the assumption of a change point against the one of a "clean" data set. Pignatiello and Samuel [37] have worked with an identical framework and Gombay [38] generalized it even further by allowing the nuisance parameter σ be unknown.

iii) Complete ignorance about process parameters

Hawkins *et al.* [25] developed their CPM formulation under this scenario of complete darkness, arguably, the most realistic of all. They prescribe conducting another generalized likelihood ratio test as follows:

Defining $\bar{X_{jn}} = \frac{\sum_{i=1}^{j} X_i}{j}$ as the pre-change mean, $\bar{X_{jn}}^* = \frac{\sum_{i=j+1}^{n} X_i}{n-j}$ as the post-change mean, and $V_{JN} = \sum_{i=1}^{j}(X_i - \bar{X_{jn}})^2 + \sum_{i=j+1}^{n}(X_i - \bar{X_{jn}}^*)^2$ as the error sum of squares, a traditional two-sample t-statistic for comparing the two means would be

$$T_{jn} = \sqrt{\frac{j(n-j)}{n}}\frac{\bar{X_{jn}} - \bar{X_{jn}}^*}{\hat{\sigma_{jn}}} \tag{12}$$

with $\hat{\sigma_{jn}^2} = \frac{V_{jn}}{n-2}$. Under the null assumption of stationarity, $T_{jn} \sim t_{n-2}$. The MLE of the true change-point is thus

$$\hat{\tau} = argmax_{1\le j\le n-1}|T_{jn}| \tag{13}$$

and a change is signaled if

$$max_{1\le j\le n-1}|T_{jn}| > h_n \tag{14}$$

The thresholds $\{h_n\}$'s are constructed using Bonferroni type bounds. For creating the competitors to follow, we will remove the normality assumption and hence the t statistic (which, for this example can be taken as the D_n candidate), and will instead survey more recent choices.

The problem of checking whether a Poisson process (a special type of a point process, where the number of events in a given interval follows a Poisson density) is stationary or trend-infected, is well studied. In recent times, the problem has attracted attention both from the theoretic (Brodsky [39]) and pragmatic (Chen and Gupta [40]) viewpoints. Antoch and Jaruskova [41], and Lindqvist [42] may be consulted for a review. It is time for us to survey different possible candidates for the two sample statistic $D_{k,n}$ mentioned previously and check their performance against the established CUSUM technique.

1) CPM-EXP TEST

Using $Exponential(\lambda_0)$ and $Exponential(\lambda_1)$ choices for F_0 and F_1 in (2), (3), Ross [26] has constructed expressions for the generalized likelihood ratio statistic

$$M_{k,n} = -2log(\frac{L_0}{L_1}) \tag{15}$$

where L_0 and L_1 denote the likelihoods under the null and the alternate hypothesis, respectively. $M_{k,n}$ can then be taken as $D_{k,n}$ under the general CPM framework.

2) CPM-ADJUSTED EXP TEST

As n explodes, it can be shown (Ross [26]) that the average of $M_{k,n}$ defined previously, approaches $-2k\{\psi(k) - log(k)\}$, which need not be 1, the expectation of a chi-square variable with one degree of freedom. Here $\psi(k) = \frac{\Gamma(k)}{\Gamma'(k)}$ is the usual digamma function. To rectify this, Ross [26] scaled $M_{k,n}$ down as

$$M_{k,n}^c = \frac{M_{k,n}}{E(M_{k,n})} \tag{16}$$

which makes the mean hover around 1. $D_{k,n}$ in the original CPM framework may thus, now be played by $M_{k,n}^c$.

3) CPM-MANN-WHITNEY TEST

This relies on Pettitt's [33] proposal of a U statistic based on the Mann-Whitney two-sample test:

$$U_{k,n} = \sum_{i=1}^{k}\sum_{j=k+1}^{n} P_{ij} \quad 1 \le k \le n-1 \tag{17}$$

where

$$P_{ij} = sgn(X_i - X_j) = \begin{cases} 1 & \text{if } X_i > X_j \\ 0 & \text{if } X_i = X_j \\ -1 & \text{if } X_i < X_j \end{cases} \tag{18}$$

Conover [43] relates $U_{k,n}$ to the rank of X_i, i.e. R_i as

$$U_{k,n} = 2\sum_{i=1}^{k} R_i - k(n+1) \tag{19}$$

implying

$$E(U_{k,n}) = 0, \quad Var(U_{k,n}) = \frac{k(n-k)(n+1)}{3} \tag{20}$$

Thus, $D_{k,n}$ in the CPM framework may be taken as

$$D_{k,n} = \frac{U_{k,n}}{\sqrt{k(n-k)(n+1)/3}} \tag{21}$$

More on this technique can be found in Hawkins and Deng [27].

4) CPM-MOOD TEST

The Mood test developed by Mood [44] is efficient in detecting scale parameter shifts, with reasonable power performance, observed by Duran [45]. Defining the rank of the ith observation as

$$r(X_i) = \sum_{i\ne j}^{n} I(X_i \ge X_j) \tag{22}$$

the statistic quantifies the amount of discrepancy between the rank of a point and its average

$$M' = \sum_{X_i}(r(X_i) - \frac{n+1}{2})^2. \tag{23}$$

M, a standardized version of M' can then be taken as D_n in the CPM framework. Details about the standardization can be had from Ross *et al.* [28].

5) CPM-LEPAGE TEST

With the Mann-Whitney test designed to detect location changes and the Mood test to detect scale shifts, a need is often felt to combine the two and create a test efficient for both aspects. Lepage-type tests offers (see [59]) an alternative by using

$$L = U^2 + M^2 \tag{24}$$

with U and M defined previously. L can then be incorporated into the CPM framework. More on this test can be found in Ross *et al.* [28].

6) CPM-KOLMOGOROV-SMIRNOV TEST

This exploits the comparison between the empirical distribution functions of the pre-change and the post-change sample defined as:

$$\hat{F}_{S_1}(x) = \frac{1}{k}\sum_{i=1}^{k} I(X_i \leq x) \tag{25}$$

$$\hat{F}_{S_2}(x) = \frac{1}{n-k}\sum_{i=k+1}^{n} I(X_i \leq x) \tag{26}$$

and $D_{k,n}$ in the CPM framework is taken as:

$$D_{k,n} = sup_x|\hat{F}_{S_1}(x) - \hat{F}_{S_2}(x)| \tag{27}$$

Techniques for standardization can be had from Ross and Adams [29].

7) CPM-CRAMER-VON-MISES TEST

This uses the square of the average distance to quantify discrepancy between the two empirical functions. $D_{k,n}$ in the CPM framework is now:

$$D_{k,n} = \int_{-\infty}^{\infty} |\hat{F}_{S_1} - \hat{F}_{S_2}| dF_t(x) \tag{28}$$

with $F_t(.)$ standing for the empirical c.d.f. for the pooled sample. For implementation purposes, one may use:

$$D_{k,n} = \sum_{i=1}^{n} |\hat{F}_{S_1}(X_i) - \hat{F}_{S_2}(X_i)|^2 \tag{29}$$

Ross and Adams [29] may be consulted for the necessary standardization.

8) E-DIVERGENCE TEST

A technique originally developed by Matteson and James [30] to detect any number of change points in multivariate time series observations, it is distribution free and is capable of detecting changes of several kinds. A priori knowledge on the number of change points is not required, however, the observations must be independent, and have finite αth absolute moments with $\alpha \in (0, 2]$. It uses Szekely and Rizzo (2005, 2010) [50], [51]'s divergence measure to check whether two vectors $X, Y \in R^d$ with characteristic functions $\phi_X(t)$ and $\phi_Y(t)$ are identically distributed. Using Matteson and James (2013)'s [30] proposal of a specific weight function, the measure takes the form

$$D(X, Y; \alpha) = \int_{R^d} |\phi_X(t) - \phi_Y(t)|^2 \left(\frac{2\pi^{d/2}\Gamma(1-\alpha/2)}{\alpha 2^\alpha \Gamma((d+\alpha)/2)}|t|^{d+\alpha}\right)^{-1} dt \tag{30}$$

The null assumption of similarity is rejected for exceedingly high values of this divergence. James and Matteson (2013, 2014) [30], [52] introduce a binary tree based bisection algorithm called "E-divisive" for hierarchical divisive change point estimation. The significance of an estimated change point and its corresponding p-value is found through permuting the observation collected thus far.

D. TREND RATIO STATISTIC

If we denote the times of the incoming observations by $t_1, t_2, \ldots, t_n$, the *forward* statistic Z defined through

$$Z = -2\sum_{i=1}^{n} log(t_i/t_n) \tag{31}$$

has been examined by Rigdon and Basu [53], among others. Ho [54] made a slight modification to Z, introducing a *backward* version of it, as

$$Z_B = -2\sum_{i=1}^{n-1} log(1 - t_i/t_n). \tag{32}$$

Thus, if departure from stationarity leads to process deterioration, i.e., if events seem to occur more and more frequently in recent times, the t_i/t_n values will tend to cluster around 1, inflating the value of Z_B and deflating the value of Z. If the process improves, Z will be high while Z_B will be low. Ho [54] examined these versions in the context of a Poisson process setting and found situations under which the use of Z_B in detecting non-stationarity could be advantageous. One such instance is if the intensity of the underlying stochastic process happens to be a series of gradually increasing steps. But there are other situations, notably when the process improves (through maybe a gradually decreasing step intensity) when Z_B's statistical power drops in comparison to Z's.

Noting both versions perform well under certain disjoint conditions and that the goal of the present paper is to detect non-stationarity *broadly* defined (as opposed to only deteriorating or improving systems), we combine Z and Z_B to propose a new trend ratio statistic defined by

$$Z_{TR} = \frac{Z_B}{Z} = \frac{\sum_{i=1}^{n-1} log(1 - t_i/t_n)}{\sum_{i=1}^{n} log(t_i/t_n)}. \tag{33}$$

It is easy to see that Z_{TR} is bounded by 0 and 1. When the process is stationary (or homogeneous, in Poisson process literature) the elements in the $\{t_i\}$ sequence are roughly

equidistant from each other leading to $Z = Z_B$, and consequently, $Z_{TR} = 1$. The probability distribution of this random quantity Z_{TR} therefore, will be centered around 1, under the assumption of stationarity. Statistically significant departures to the right of 1 will signify deterioration (with $Z_B > Z$) while significant departures to the left will imply improvement (with $Z_B < Z$). We determine the thresholds differentiating stationarity from non-stationarity using the lower and upper α points of the null probability distribution centered around 1. Here α represents the usual probability of Type-I error, set to 0.05 in this study.

We shall next implement each of these change detection options along with an array of realistic synthetic scenarios to discover the best choice under every situation.

III. SIMULATION STUDIES

Prior to analyzing real data sets, in this section, we endeavor to examine the suitability of the several change detection methods described in the previous section using simulated data sets. Inter-event times $\{X\}$s were generated using the parametric densities mentioned below on the statistical software R. These gap times were pasted together to generate the global times $\{T\}$s seen in (31) and (32) (and indeed to generate the entire process for any of the tools described in the previous section to be applied) through

$$T_k = \sum_{i=1}^{k} X_i, \quad k = 1, 2, .., n. \tag{34}$$

In the simulation results to follow, 10^6 such processes were generated and each was classified by the ten competing tools (described in the previous section) into one of two types - stationary or not. To make our simulations as realistic as possible, we next propose two new ways in which a stable and stationary process might get corrupted.

A. STRONG CORRUPTION

We have mentioned in the introductory section that one of the ways in which a departure from stationarity occurs is if at some point during the natural progression of the process, the probability distribution driving it changes. By strong corruption, we indicate a change in which the modified probability distribution is deterministically known. To illustrate, in the simulations to follow, we sample inter-event times (Xs in (34)) from a *Weibull*(k, λ) density of the form

$$f_W(x) = \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1} \exp(-x/\lambda)^k, \quad x > 0, \tag{35}$$

and let the shape parameter λ vary to represent different non-stationarities. For visual convenience and to graphically emphasize the deviation from a stable flow, we also sample inter-event times from a *Gamma*(α, β), represented by the density

$$f_G(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \exp(-x/\beta)x^{\alpha-1}, \quad x > 0. \tag{36}$$

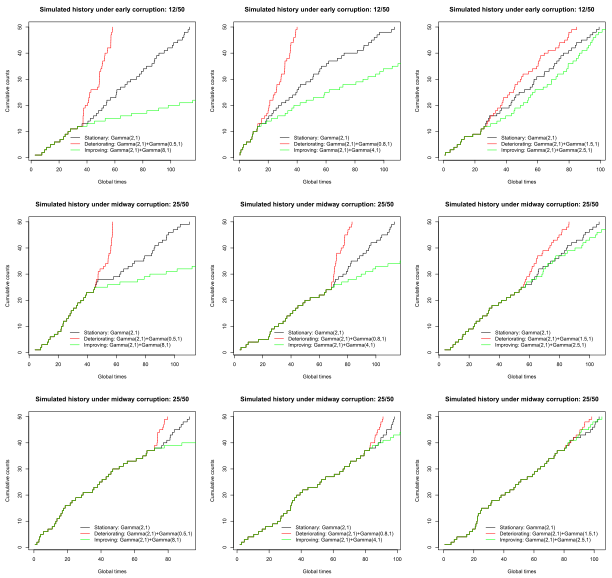

FIGURE 3. Non-stationarity of different kinds explained through a Gamma example.

Figure 3 above describes strong corruption graphically. Each step represents the occurrence of an event and the distance between steps denote the inter-event times X_is in (34). Consequently, the occurrence times on the horizontal axis are the global times T_is in (34). The first panel, for instance, tells us the stationary process represented by the black curve was driven by a *Gamma*$(2, 1)$ density generated inter-event times. In case the post-change density *deterministically* switches to a *Gamma*$(0.5, 1)$, the shocks would occur more frequently, represented by the red curve and the process will be deteriorating. Similarly, if the post-change density *deterministically* switches to a *Gamma*$(5, 1)$, the shocks would occur less frequently, represented by the green curve and the process will be improving. On the diagrams in the first row, the jump from a stationary flow to a non-stationary one happens early in the process and as we move to the right, the differences between the pre and post change parameters reduce, thereby making the change detection harder. Along the second row, the change happens midway into the process, and along the third row, it happens late into the process. Strong corruption is thus, generally characterized by improvement or deterioration lines clearly showing, never deviating to the other side. From a practical viewpoint, strongly corrupted processes would typically have one cause, or more than one that acts in the same direction, that make(s) them non-stationary. The appointment of a new CEO who wants to push sales, for instance, might make the stochastic process of making sales deteriorate (i.e., sales would occur more frequently). Both the Weibull and Gamma densities are reasonable models to work with since they can account for a wide array of skew scenarios and the inter-event times are non-negative, by definition.

For our simulations, thus, at some predetermined time, the parameters either improve or deteriorate deterministically

TABLE 1. Detection performance of classifiers under late change (non.stationary shape = 0.6).

Option	Sensitivity (%)	Specificity (%)
Exp	97	96
Exp-Adj	95	98
Mann	91	89
Mood	67	92
Lepage	92	91
K-S	95	94
CVM	95	97
Div	83	95
CUSUM	61	40
Z_{TR}	97	95

TABLE 2. Detection performance of classifiers under late change (non.stationary shape = 1.8).

Option	Sensitivity (%)	Specificity (%)
Exp	83	81
Exp-Adj	84	86
Mann	81	78
Mood	64	88
Lepage	80	75
K-S	83	84
CVM	79	81
Div	81	87
CUSUM	49	38
Z_{TR}	93	94

TABLE 3. Detection performance of classifiers under late change (non.stationary shape = 2.2).

Option	Sensitivity (%)	Specificity (%)
Exp	78	75
Exp-Adj	76	72
Mann	73	69
Mood	61	67
Lepage	73	62
K-S	81	78
CVM	68	71
Div	74	79
CUSUM	43	41
Z_{TR}	92	93

TABLE 4. Detection performance of classifiers under late change (non.stationary shape = 2.6).

Option	Sensitivity (%)	Specificity (%)
Exp	81	83
Exp-Adj	84	81
Mann	87	90
Mood	73	91
Lepage	88	89
K-S	92	91
CVM	91	93
Div	89	92
CUSUM	65	53
Z_{TR}	98	96

(as opposed to randomly, described in the next subsection) and stays that way throughout the rest of the process. This corruption may happen at different points in the history and becomes harder to detect the later it occurs. With this, in the spirit of Figure (3), we categorize them into early, midway, and late corruptions. Early and midway corruptions are easy to identify because deviation from stationarity occurs very soon after the process starts, and has an algorithm has a large portion of the it to pick up a change. Late corruptions, by contrast, are harder to detect, as is evidenced in Fig (3) above.

As our first exercise, we have looked at the detection problem, i.e., checking whether a change from stationarity has occurred at all. It it worthwhile to recall that Sensitivity and Specificity are two established performance indicators of almost any binary classifier. In our context, Sensitivity is defined as the probability (estimated through a relative frequency) of correctly identifying a non-stationary process as non-stationary while Specificity is taken as the probability of correctly classifying a stationary process as stationary. Higher these probabilities, better is the classifier. The shape parameter for the stationary Weibull density was held at 2 and the tables below describe the different non-stationarity parameters. Since late changes are the most difficult ones to identify, through Tables (1) - (4) below, we have recorded performance measures under this condition. We, however, offer summary tables below that takes into account the other possibilities too.

Table (1) above represents a situation when the difference between the stationary parameter (2) and the non-stationary parameter (0.6) is noticeable. Thus, although the change corrupts the process late into its evolution, the difference in parameter values makes it easily detectable, evidenced by the high sensitivity values for most of the competitors. Table (2) below makes the non-stationary parameter (1.8) closer to the stationary one and records similar results.

The sensitivity values have dropped considerably owing to the proximity of the non-stationary and stationary parameters, however, similar to its performance in Table (1) above, our new proposal Z_{TR} offers the best classification accuracy even under this difficult scenario. While the two tables shown above investigate departures to the left of the stationary parameter (suggesting process deterioration), the two to follow will consider deviations to the right (hence process improvements).

Table (3) above represents improvement detections that are harder to detect owing to the facts that the change appeared late into the process and in addition, the non-stationary parameter was extremely close to the stationary one, while table (4) represents improvement detections that are relatively easier due to the large difference between the non-stationary and stationary parameters.

In the four tables above, we have observed how in separating stable flows from unstable ones, our new proposal Z_{TR} outperforms both its established competitor CUSUM, and a host of other recent ones. Certain options (such as the Divergence-based measure) do not detect deteriorations as efficiency as they do improvements. The new candidate Z_{TR} however doesn't suffer seriously from such asymmetries. This is expected since we constructed it in (33) combining

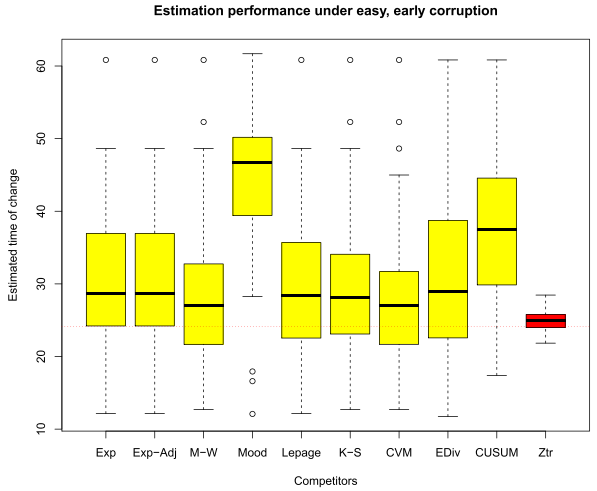

FIGURE 4. Estimation accuracy through time-of-change distributions (easy detection).

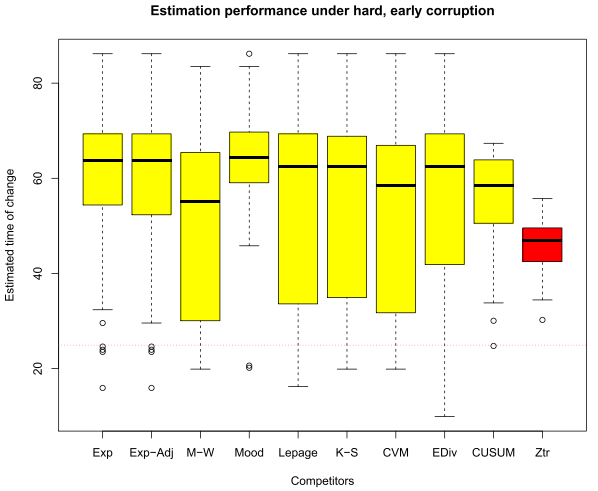

FIGURE 5. Estimation accuracy through time-of-change distributions (hard detection).

and borrowing strength from components designed to work best in one specific direction. Most notably, Z_{TR} should be the preferred choice in case the change detection seems hard, either in the sense of a late occurrence or in the sense of close similarity of the post-change probability density to the pre-change one, or both.

Next, we turn our attention to the case of estimation. This is important since detecting a change (what we have done in the previous part of this section) is only the first hurdle a classifier has to overcome. Once a change gets detected (if it gets detected) the next step is to inquire *when* did it occur. Such estimation and comparing those estimates with user-defined change points form the purpose of the last portion of this subsection on strong corruption.

The red broken line in Figure (4) above represent the true time of change and the boxplots represent the estimated change time distributions from the different competitors. The jump was placed early into the process (after the 12th observation out of 50) and the pre (2) and post (0.6) change parameters were held markedly different. These imply an easy detection scenario, which help explain the good performance from almost every change-detection choice - the medians from most of the ten time-of-change distributions hover around the true change time. The distribution from our new proposal Z_{TR}, however, is the closest to the truth, along with the tightest spread, promising accurate and consistent inferences.

Keeping the change position unaltered (i.e., early into the process), we next make the detection harder by moving the non-stationary parameter (1.8) closer to the stationary one (2). Figure (5) records those results. The difficult situation gets reflected both through the widening of the box-lengths and the farther drifting of the medians away from the true change dotted line. Nevertheless, Z_{TR} still retains its superiority.

We next generalize the above into a large-scale study combining both types of non-stationarities (i.e., deterioration and improvement), two types of parameter proximities (close, suggesting hard detection and wide, suggesting easy detection), and three types of knot-placements (early, midway, and late into the process), covering thereby $2 \times 2 \times 3 = 12$ possibilities. We have tracked down the best option under each scenario, taking into account both aspects of change analysis - detection and estimation. These have been summarized in the following recommendation tables (Tables (5) and (6)) below.

TABLE 5. Recommendation table (deterioration).

	Early shock	Midway shock	Late shock
Easy	CVM or Z_{TR}	K-S	Z_{TR}
Hard	M-W or Z_{TR}	M-W or Z_{TR}	Z_{TR}

TABLE 6. Recommendation table (improvement).

	Early shock	Midway shock	Late shock
Easy	M-W or Z_{TR}	K-S	Z_{TR}
Hard	Lepage or Z_{TR}	E-Div or Z_{TR}	Z_{TR}

These tables have been constructed to store the strongest candidates, defined by detection accuracies (evidenced by high sensitivity and specificity values) and closest average proximity to the red broken line, and may be used as follows: if one wants to guard against a deteriorating process (suggesting events are happening more and more frequently in recent times, which, in turn, could spell disaster, especially with examples studied in the next section) and suspects that the change happened midway into the process with parameters hugely different from the stable one (implying an easy detection), the best two sample statistic $D_{k,n}$ to use in the CPM framework would be the Kolmogorov-Smirnov one. Other conditions remaining the same, if the change is suspected to happen in more recent times (implying a late change), the best option to implement would be the new proposal Z_{TR}. It is interesting to seen that this trend ratio option occupies most of these twelve cells.

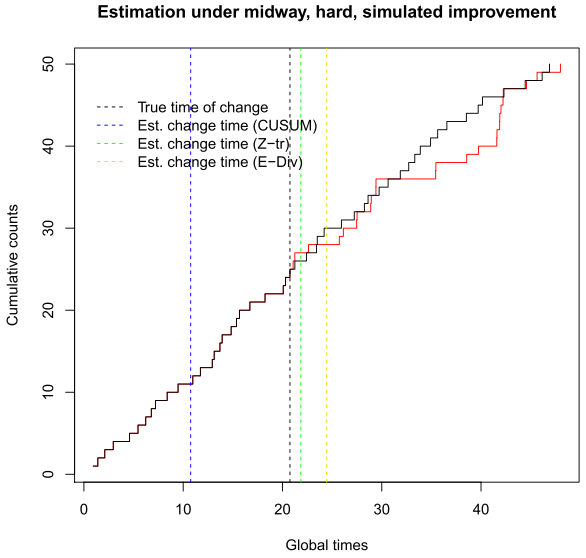

FIGURE 6. Change detection and estimation under hard, midway change.

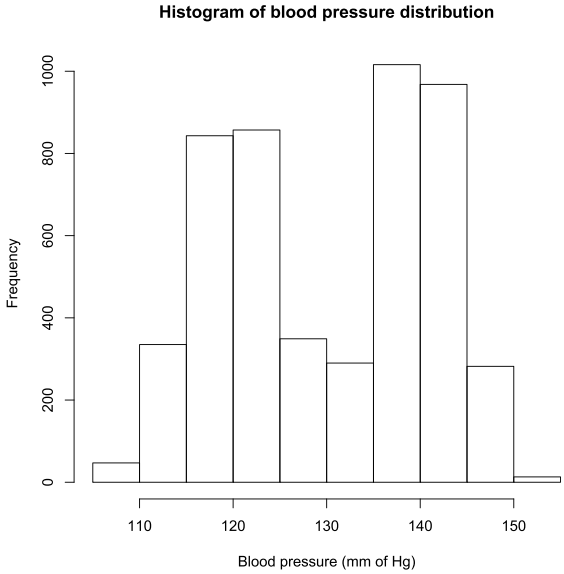

FIGURE 7. Hypothetical blood pressure distribution from 5000 patients.

We also implemented one of these recommendations with one isolated simulation shown in Figure (6) representing a hard, midway improvement. Table (6) prescribes either the Energy-divergence or the Z_{TR} option. The black steps represent a stationary and stable flow, while the red departure signals an improvement somewhat midway into its evolution. In addition to the evident closeness of the newer candidates (Z_{TR} and E-div) to the true time of change and our latest proposal Z_{TR}'s better performance, another observation merits mention: we have found in almost all the cases possible, Z_{TR} and the CPM-based candidates suggest estimated times of change that are *after* the true time of change. This is not always true for the established CUSUM, as Figure (6) shows, suggesting its extreme sensitiveness and tendency to sound too many false alarms.

B. WEAK CORRUPTION

The preceding subsection assumes that the type of departure from stationarity is known, which requires considerable mastery and field knowledge about the process unfolding. If, by any chance, we are uncertain of the nature of an alteration, we can combine improvement and deterioration setups probabilistically to generate a mixed distribution for the inter-event times. By way of illustration, let us consider a case where a certain set of individuals buy cigarettes in a smoke shop over a period of time. If we gather data on how much cigarettes was bought per day and graph it through a histogram, in all likelihood, it will exhibit several peaks. The reason being, there exists an inherent categorization among smokers - heavy, moderate, occasional, etc. and the buying habits vary considerably across these groups. In such cases, it is useful to introduce the notion of a mixture density.

Formally, if $\{f_i(.)\}, i = 1, 2, .., k$ represents a set of valid probability densities, then

$$f(x) = \sum_{i=1}^{k} p_i f_i(x) \tag{37}$$

is also a valid density, termed a k-mixture, as long as $\sum_{i=1}^{k} p_i = 1$. Except for artificially constructed scenarios, when some or all of the p_i's vanish, mixture densities of the type (37) exhibit multimodalities due to uncertainty about the "sub-population" that generated a specific value. To demonstrate, let us consider another example of patients pouring into a clinic with different complaints. If our variable of interest is the observed blood pressure of a randomly chosen patient, a histogram of 5000 patients sampled might reveal the following bimodality (Fig (7)). This is expected, since the very definition of the study variable forces a dichotomy among the patient population - a *sub*-population who came in for high blood pressure treatment, and another who came in for some other reason (perhaps a broken arm) but with perfectly normal blood pressure. Assuming in the first group, blood pressures vary according to a normal model with mean 140 mm of Hg and standard deviation 4 mm and in the second, they vary similarly with mean 120 mm and standard deviation 5 mm, a random draw of 5000 patients generates Fig(7) above, assuming fairness between groups. Thus in (37), we have used $k = 2, p_1 = p_2 = 0.5, f_1(x) = Norm(140, 4), f_2(x) = Norm(120, 5)$ to get the combined or generalized blood pressure distribution $f(x)$, graphed in Fig (7). Mixture densities have a tendency of cropping up in unexpected places, such as the marginal distribution for a Hidden Markov Chain, used for speech recognition, among others. A detailed account of their statistical peculiarities may be garnered from Zucchini and MacDonald [18].

This process is known as weak corruption in our specific context, where we generally introduce varying degrees of both improvement and deterioration. To be concrete, the inter-event time distributions in this subsection will not be deterministically governed (like in the previous on strong corruption), but will be generated according to (37) with some appropriate choice of the mixing probabilities and

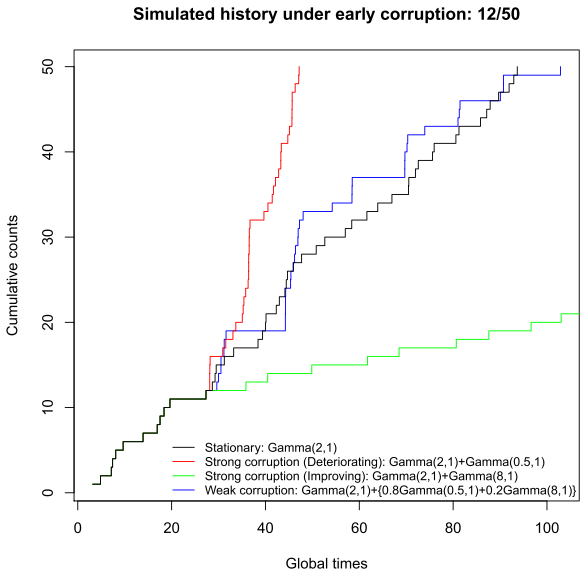

FIGURE 8. Weak corruption explained through Gamma mixing.

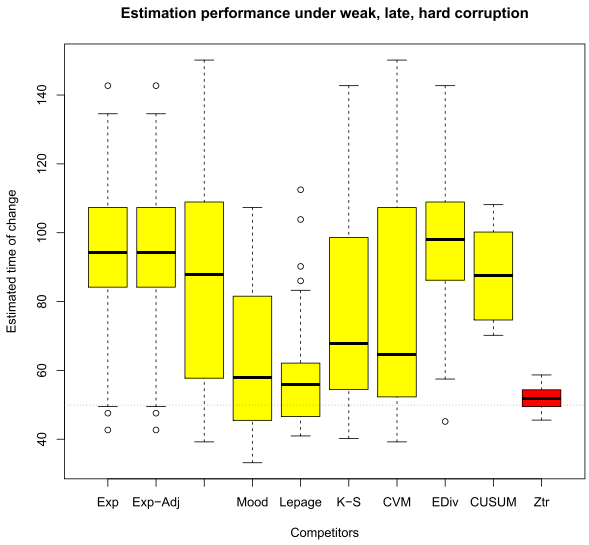

FIGURE 9. Estimation accuracy through time-of-change distributions (weak, late, hard detection).

contributing densities. This will enable the process to follow an alternating deterioration-improvement pattern and incorporate our possible lack of complete knowledge on the nature of its history. The mixing was carried out using the "UnivarMixingDistribution" function within the "distr" package in R and the random generation of inter-event times for our simulation studies was done using functional programming.

Figure (8) portrays both strong (i.e., deterministic, introduced in the last subsection) and weak (i.e., random, introduced here) corruption simultaneously. The black line represents a stable, stationary process. The red represents a (deterministically) deteriorating case where, at a certain point of time, admittedly, early in the process, the process transitioned from a stationary to a non-stationary environment. The green line represents an (deterministically) improving case where, also, the line transitioned early in the process. The blue process, however, represents a mixed case in which we combined (20%) improving and (80%) deteriorating cases to produce a weak non-stationarity. Gamma densities have been used as the contributing distributions for both improvement and deterioration. As shown in the illustration, the deterministic changes from stationarity to non-stationarity is very noticeable, as both the red and green curves demonstrate. However, as a glance at the black and the blue steps shows, such differentiation gets tough if the update through weak non-stationarity, gets random.

Analyses similar to the ones carried out in the previous subsection may be conducted here too, with the simplification that the separate cases of improvement and deterioration goes away, since weak corruption, by its very definition, takes both sides into account. We now, therefore, have $2 \times 3 = 6$ possibilities - two from the similarity of parameters and three from the placement of change points. Figure (9) above represents the most difficult of such cases, when in addition to the corruption being weak, the parameters are similar

TABLE 7. Recommendation table (weak corruption).

	Early shock	Midway shock	Late shock
Easy	M-W	K-S	Z_{TR}
Hard	E-Div or Z_{TR}	E-Div or Z_{TR}	Z_{TR}

(i.e., hard) and the change happened late into the process. Table (7) below offers the recommendations, combining all cases. It is interesting to note that Z_{TR} detects departures from stability more accurately than the rest, even under this difficult situation of weak, random corruption, especially if the difference between the pre-change and post-change process parameters is minimal (suggesting a harder classification) and if the change is late in the history, signifying a more recent deviation. This accuracy under late changes is particularly appealing since from a practical viewpoint, we are concerned about changes in the immediate future as opposed to the distant past.

IV. REAL DATA ANALYSES

In a simulation scenario, experimenters have control over the time at which they want to corrupt a stationary process and, subsequently, measure the average difference between the estimates furnished by the competitors and the true change point. In a real data example, however, the "truth" is unknown. In this section, therefore, we shall investigate two processes and confirm the change points found by the alternatives proposed here with those generated using different techniques in some of our previous works.

A. VOLCANISM

Volcanic eruptions can be disastrous situations in terms of the safety of the population, the environment and the local economy. The current situation in Hawaii with the continued eruption of the Kilauea volcano proves this point. At this

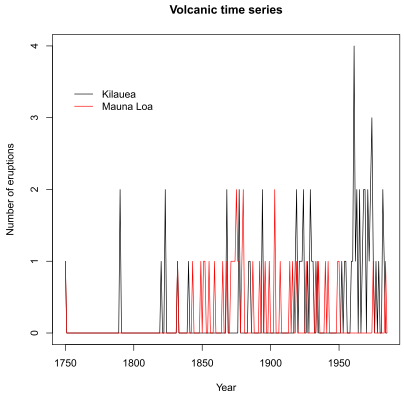

FIGURE 10. Annual time series storing eruption counts for Kilauea and Mauna Loa.

writing over 100 acres and 35 structures have been destroyed. The eruption has displaced hundreds of residents. One explosion sent gas and ash over 30,000 feet in the air. Using the 1955 Kilauea eruption to estimate the cost in today's dollars, we are looking at about 20 million dollars.

Kilauea is unique in that there is a second volcano, Mauna Loa, only 20.6 miles away. A handful of studies in geology, including the ones from notable scholar Frank Trusdell, hints at a possible inverse dependence between these two, in the sense that one's restlessness goes hand in hand with the other's dormancy. The most notable non-mathematical analysis to this end is probably Lipman's [22]. Klein [21] studied statistical tests for non-randomness on the repose times between one's eruptions to conclude that they appear to be associated with the increased activity of the other. The study adopted a random action model, assuming that the probability of eruption in an interval of time is constant, regardless of the interval's actual location. King [20] voiced similar conclusions and suggested generalized non-linear regression models to predict eruption times and lava volume forecasts. Miklius and Cervelli [23] used continuous deformation monitoring to track the eruption correlation over short time scales. Gonnermann and Houghton [19] considers the issue of coupling in a geophysical framework and also surveys some crustal-level interactions between the two. Most of these conjectures rely on a possible physical connection and lack rigorous statistical backing. As a remedy, Ho and Bhaduri [12] tackled the problem through their recently proposed statistic, bootstrapped Empirical Recurrence Rates Ratio, constructed out of discretized time series counts. For a more extensive literature review on the topic, we direct interested readers to Ho and Bhaduri [12] or Rhodes and Hart [24]. The present section exploits the change point detection strategies detailed in the last two, conducted in continuous time, to offer a fresh perspective on the problem, through the lens of non-stationarity estimation.

Eruption data related to dates and lava volumes were collected from the United States Geological Survey (USGS) and the Smithsonian Institute's Global Volcanism Program, spanning the period from 1750 till the present day. Figure (10)

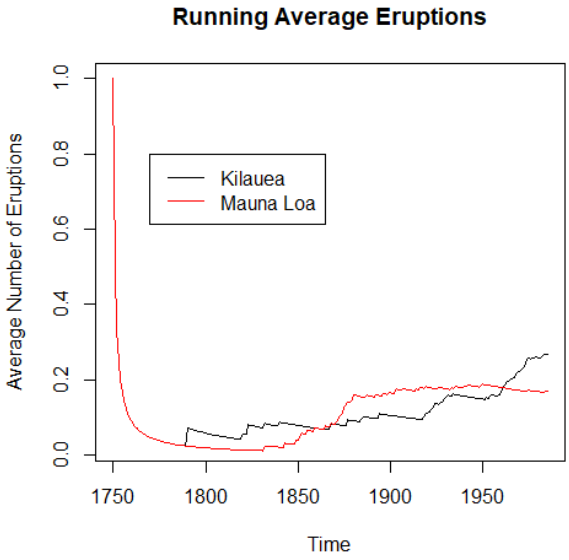

FIGURE 11. Running average counts for Kilauea and Mauna Loa.

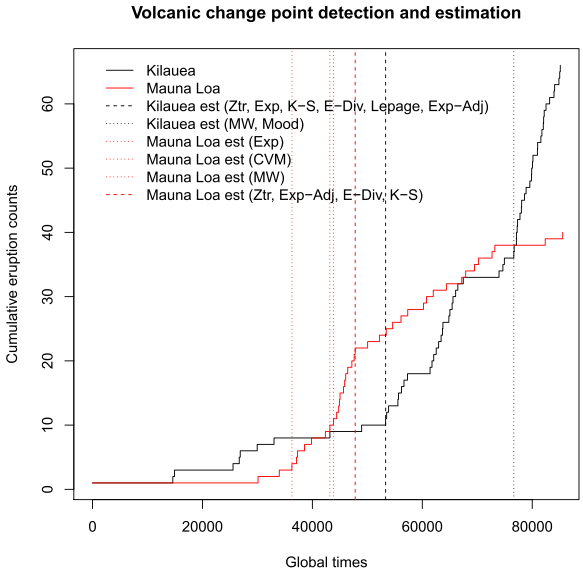

FIGURE 12. Change estimation in Kilauea and Mauna Loa eruption counts.

shows the annual eruption count time series for these volcanoes while figure (11) keeps a record of the cumulative running averages for these yearly counts. The change detection results are shown in Fig (12) above. The step graphs track the volcanoes cumulatively - one additional eruption leads to one additional step. The global times on the horizontal axis have been recorded in days. The vertical lines represent different change points obtained through different tools - the blacks for changes in Kilauea, while the reds for those in Mauna Loa. One might notice that two of these vertical lines are heavy while the rest are faded. This aspect represents the strength of our confidence in them. For instance, through the simulation analyses carried out previously, we found that our latest proposal Z_{TR} performs the best under the

widest array of scenarios. Thus if a change point is estimated through this tool, it gets coded through a heavy vertical line. In this instance, however, several good competitors agree on the same change point, which increases our confidence in it even further. It is interesting to note the proximity of the heaviest change points of Mauna Loa and Kilauea. One feels due to such closeness, that a downward slope (implying process improvement, rarer eruptions) from Mauna Loa almost *causes* an upward slope (implying process improvement, increased eruptions) for Kilauea. This inverse interaction has been hinted in our previous work Ho and Bhaduri (2017) [12] through a different function termed Empirical Recurrence Rates Ratio and was explained through the possibility of a connected magma reservoir. Thus, the volcanoes fight for the same resource causing ones restlessness imply the other's dormancy. Additionally, it is our conjecture that the change points bands will alternate between ''red-black'' and ''black-red''. To clarify, as one traverses from the left to the right on the horizontal axis, there will be a chunk of period where the Mauna Loa change points will appear first, closely followed by those for Kilauea. The next clustering will start with Kilauea change points, followed by Mauna Loa. We see initial glimpses of this phenomenon already and it will be more apparent as future eruption data come in, or if forecasts are extracted in spirits similar to Ho and Bhaduri (2017) [12] or Ho and Bhaduri (2015) [9]. This pattern would not have happened had the two processes been independent of each other in which case the red and the black vertical lines would have been randomly mixed.

Government officials may exploit this dependence in several ways. Relocating inhabitants seems the most pressing. Moving people from dangerous zones remains crucial. As does doing so in a cost-efficient way. Shifting people from the now-active Kilauea to a great distance from it might save lives but will be forbidding financially. The inverse dependence established through this change point research might suggest a prudent alternative - shifting them close to Mauna Loa. It is significantly dormant currently and simultaneously, not a great distance away from Kilauea, either, suggesting minimal moving costs.

B. WEATHER SCIENCE

The National Oceanic and Atmospheric Association (NOAA) classifies hurricanes based on their maximum wind speeds attained (described in Table (8) below) and following is an analysis on whether the frequencies of such storms have changed in recent times. Following the recommendations of leading weather scientists like Emanuel (2003, 2006, 2007) [46]–[48] and also because of the damage they inflict, we have agreed to define hurricane categories 3 through 5 as the strong group and 1 and 2 as the weak group. Data on their past/projected tracks, origination dates, etc. are freely available from the NOAA web page. The West Atlantic basin, bordering the US east coast, is one of the most well studied oceanic regions and our conversations with NOAA experts about a reliable time frame led us to choose 1923-2013.

TABLE 8. NOAA hurricane classification.

Category	Maximum wind speed
Hurricane 5	$>$ 135 knots
Hurricane 4	114-135 knots
Hurricane 3	96 - 113 knots
Hurricane 2	83 - 95 knots
Hurricane 1	64 - 82 knots
Trop/Subtropical storms	34 - 63 knots

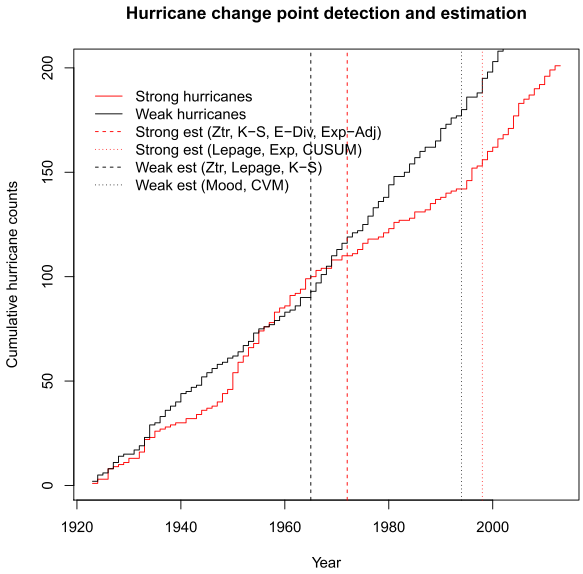

FIGURE 13. Change estimation in strong and weak hurricane counts, West Atlantic basin.

Over this period and over this region, we have found 32 H5 storms, 84 H4 storms, 87 H3 storms, 93 H2 storms, and 150 H1 storms. The cumulative counts, through steps, have been plotted in Figure (13) below. Scientists ([46]–[48]) believe that with a continually warming climate, the likelihood of a strong hurricane drops due to a decline in moisture deficit, however, if it gets started somehow, it has the potential to become deadly. Thus, although the number of hurricanes might increase in recent times, the proportion of those that are strong should fall.

We have implemented the ten options described previously on both the strong and weak series with the estimates collected in Figure (13). While some candidates (especially CUSUM on the weak series) fail to identify any changes at all, our low confidence in those (evidenced through their simulation performances) leads us to turn to the more reliable ones mentioned in the recommendation tables (such as Z_{TR}, E-Div, etc). We find that for the strong series, most of the candidates (including the newly proposed Z_{TR}) identify a change point around 1972 (the heavy dotted vertical red line) while for the weak series, most (again, including the newly proposed Z_{TR}) identify a change around 1965. It is interesting to note that using entirely different modeling techniques (through a statistic termed Empirical Recurrence Rates Ratio), in one of our previous works (Bhaduri and Ho (2018) [49]) we

found a change in the *interaction* pattern between these two series around 1970. The advent of global warming around this time was noted in this work and offered as a possible reason. In the present context, however, the proximity of these dates 1972, 1965 (obtained from the current study) and 1970 (obtained from the previous), derived using different strategies, strengthens the conclusion that changes did occur both in the ways the strong and weak hurricane count series propagate individually, and hence in the pattern of their interactions.

V. CONCLUSION

The issue of non-stationarity intrigues several decision making processes on a daily basis, with consequences both profound and mild. Examples originate from myriad sources such as industrial engineering, where a production manager might be curious about detecting a change in the quality of the items manufactured, the financial sector, where one might be concerned with the rate of bank failures, or health care, where the prevalence of a rare disease might be the quantity of interest. The current state of art relies on some variant of the Cumulative Sum (CUSUM) charting technique, where departure from an insipid flow is usually flagged by a chain of points, documenting the flow, venturing beyond a specified threshold. While the principle sounds alluring and works reasonably well in theory, there exist several strong, often untenable assumptions, perhaps the most striking one being complete knowledge of the pre and post-change parameters, necessary to come up with the bounds. In addition, the time discretization method involved invariably loses information that could offer insights into the time and nature of a change.

Noting such drawbacks, this work takes stock of a host of statistics, both parametric and non-parametric, and examines their applicability under a wide array of change scenarios. We have proposed two new ways in which non-stationarity might corrupt a stable flow - a strong way, in which the post change density is deterministically governed and a weak way, in which it is a random mixture of deterministic densities. We have explained the advantages of each and formulated a new trend ratio statistic Z_{TR} combining ideas of ordinary trend (when time flows from the left to the right) and reversed trend (when it is switched to the right to the left). To replicate non-homogeneous environments, we have conducted extensive simulations, both under the deteriorating and improving framework, with the changes placed early, midway, and late into the process. To invite further intricacies, within each category, we have generated easily identifiable non-stationarities, when the corrupted parameters are extremely different from the stable parameters, and difficult-to-identify non-stationarities, when they are not. The established CUSUM, its eight prevalent competitors and our proposal were subsequently employed to address two crucial aspects of unstability research: detection - investigating whether a change happened at all, and estimation - guessing the true time of change, in case it did. This rewarding exercise revealed that in every possible corrupted scenario, at least one of the newer proposals outperforms CUSUM with regards to both aspects mentioned previously. High sensitivity and specificity values imply efficient detection, and distributions for the estimated change times hovering around the true change, suggest accurate estimation. We have constructed recommendation tables, summarizing the apt proposals, for the ease of applicability and the benefit of practitioners. Weak corruptions were then considered, with the non-stationary piece construed as a probabilistic mixture of an improving and a deteriorating system, and the supremacy over CUSUM was observed to be retained. Our newest proposal Z_{TR} was found to occupy most of the cells in these recommendation tables, with some shared by other non-CUSUM options. Interestingly, however, Z_{TR} was found to be the only undisputed choice in case a suspected change happened late into the process, arguably the hardest detection scenario. Such classification power holds practical benefits too since in general, we are more concerned about changes in the immediate future as opposed to ones in the distant past. The rationale behind Z_{TR}'s accuracy has also been explained.

Attention was then turned to harrowing ordeal inhabitants of the Hawaiian volcano, Kilauea are going through, and the confidence gained through the simulation studies was channelized both to fuel and resolve an ongoing geological debate - whether Kilauea and its close neighbor Mauna Loa are inversely related. The proximity of the estimated change points of these neighbor provides an answer in the affirmative. Such statistical validation should aid government authorities to relocate Kilauea inhabitants close to Mauna Loa. Lives may be saved and moving costs may be minimized, owing to the geographical closeness. Interactions between strong and weak hurricanes originating in the West Atlantic basin, striking the US east coast, was then considered and the change detection results found through this study was observed to be in agreement with some of our other studies published recently, using different interaction ideas.

This research is still in its infancy, but with its promising prospects, the road ahead looks enticing. We are especially excited about ways of putting Bayesian priors on the number of mixture components involved, their biasing intensities, and corruption locations. The sixth author's Ph.D. Dissertation (forthcoming) (Bhaduri (2018) [16]) offers the relevant intricacies. Average Run Length calculations may be carried out as well and multiple changes may be detected simply by restarting the algorithm. Our proposals are not tethered to the suffocating CUSUM assumptions brought to the fore and provide better change estimates without paying a hefty price in terms of model complexity. Their computational simplicity and intuitive appeals should also help them find a way into every modeler's arsenal.

REFERENCES

[1] (Jun. 6, 2019). *Senghenydd Colliery Disaster*. Accessed: Aug. 3, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Senghenyddcollierydisaster
[2] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," *Stat. Sci.*, vol. 17, no. 3, pp. 235–255, 2002.

[3] J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, "A review and comparison of changepoint detection techniques for climate data," *J. Appl. Meteorol. Climatol.*, vol. 46, no. 6, pp. 900–915, 2007.

[4] J. F. Ducré-Robitaille, L. A. Vincent, and G. Boulet, "Comparison of techniques for detection of discontinuities in temperature series," *Int. J. Climatol., J. Roy. Meteorol. Soc.*, vol. 23, no. 9, pp. 1087–1101, 2003.

[5] F. R. Chowdhury, "A soft computing approach for on-line automatic speech recognition in highly non-stationary acoustic environments," Ph.D. dissertation, Inst. Nat. Recherche Scientifique, Univ. Québec, Québec City, QC, Canada, 2012.

[6] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[7] L. Wei and E. Keogh, "Semi-supervised time series classification," in *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2006, pp. 748–753.

[8] L. J. Bain and M. Engelhardt, "Inferences on the parameters and current system reliability for a time truncated Weibull process," *Technometrics*, vol. 22, no. 3, pp. 421–426, 1980.

[9] C.-H. Ho and M. Bhaduri, "On a novel approach to forecast sparse rare events: Applications to Parkfield earthquake prediction," *Natural Hazards*, vol. 78, no. 1, pp. 669–679, 2015.

[10] C.-H. Ho, G. Zhong, F. Cui, and M. Bhaduri, "Modeling interaction between bank failure and size," *J. Finance Bank Manage.*, vol. 4, no. 1, pp. 15–33, 2016.

[11] S. Tan, M. Bhaduri, and C.-H. Ho, "A statistical model for long-term forecasts of strong sand dust storms," *J. Geosci. Environ. Protection*, vol. 2, pp. 16–26, Jun. 2014.

[12] C.-H. Ho and M. Bhaduri, "A quantitative insight into the dependence dynamics of the Kilauea and Mauna Loa volcanoes, Hawaii," *Math. Geosci.*, vol. 49, pp. 893–911, Oct. 2017.

[13] M. Bhaduri and J. Zhan, "Using empirical recurrence rates ratio for time series data similarity," *IEEE Access*, vol. 6, pp. 30855–30864, 2018.

[14] M. Bhaduri, J. Zhan, and C. Chiu, "A novel weak estimator for dynamic systems," *IEEE Access*, vol. 5, pp. 27354–27365, 2017.

[15] M. Bhaduri, J. Zhan, C. Chiu, and F. Zhan, "A novel online and non-parametric approach for drift detection in big data," *IEEE Access*, vol. 5, pp. 15883–15892, 2017.

[16] M. Bhaduri, "Bi-directional testing for change point detection in Poisson processes," Ph.D. dissertation, Dept. Math. Sci., Univ. Nevada, Las Vegas, Las Vegas, NV, USA, 2018.

[17] J. Zhan, B. J. Oommen, and J. Crisostomo, "Anomaly detection in dynamic systems using weak estimators," *ACM Trans. Internet Technol.*, vol. 11, no. 1, 2011, Art. no. 3.

[18] W. Zucchini and I. L. MacDonald, *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton, FL, USA: CRC Press, 2009.

[19] H. M. Gonnermann and B. F. Houghton, "Magma degassing during the Plinian eruption of Novarupta, Alaska, 1912," *Geochem. Geophys. Geosyst.*, vol. 13, Oct. 2012, Art. no. Q10009. doi: 10.1029/2012GC004273.

[20] C.-Y. King, "Volume predictability of historical eruptions at Kilauea and Mauna Loa volcanoes," *J. Volcanol. Geothermal Res.*, vol. 37, pp. 281–285, Sep. 1989.

[21] F. W. Klein, "Patterns of historical eruptions at Hawaiian volcanoes," *J. Volcanol. Geothermal Res.*, vol. 12, pp. 1–35, Mar. 1982.

[22] P. W. Lipman, "The southwest rift zone of Mauna Loa: Implications for structural evolution of Hawaiian volcanoes," *Amer. J. Sci.*, vol. 280, no. 2, pp. 752–776, 1980.

[23] A. Miklius and P. Cervelli, "Interaction between Kilauea and Mauna Loa," *Nature*, vol. 421, p. 229, Jan. 2003.

[24] J. M. Rhodes and S. R. Hart, "Episodic trace element and isotopic variations in historical Mauna Loa lavas: Implications for magma and plume dynamics," in *Mauna Loa Revealed: Structure, Composition, History, and Hazards*, J. M. Rhodes and J. P. Lockwood, Eds. Washington, DC, USA: American Geophysical Union, 1995. doi: 10.1029/GM092p0263.

[25] D. M. Hawkins, P. Qiu, and C. W. Kang, "The changepoint model for statistical process control," *J. Qual. Technol.*, vol. 35, no. 4, pp. 355–366, 2003.

[26] G. J. Ross, "Sequential change detection in the presence of unknown parameters," *Statist. Comput.*, vol. 24, no. 6, pp. 1017–1030, 2014.

[27] D. M. Hawkins and Q. Deng, "A nonparametric change-point control chart," *J. Qual. Technol.*, vol. 42, no. 2, pp. 165–173, 2010.

[28] G. J. Ross, D. K. Tasoulis, and N. M. Adams, "Nonparametric monitoring of data streams for changes in location and scale," *Technometrics*, vol. 53, no. 4, pp. 379–389, 2011.

[29] G. J. Ross and N. M. Adams, "Two nonparametric control charts for detecting arbitrary distribution changes," *J. Qual. Technol.*, vol. 44, no. 2, pp. 102–116, 2012.

[30] D. S. Matteson and N. A. James, "A nonparametric approach for multiple change point analysis of multivariate data," *J. Amer. Stat. Assoc.*, vol. 109, no. 505, pp. 334–345, 2013.

[31] G. J. Ross, "Parametric and nonparametric sequential change detection in R: The CPM package," *J. Stat. Softw.*, vol. 66, no. 3, pp. 1–20, 2015.

[32] D. M. Hawkins, "Testing a sequence of observations for a shift in location," *J. Amer. Stat. Assoc.*, vol. 72, no. 357, pp. 180–186, 1977.

[33] A. N. Pettitt, "A non-parametric approach to the change-point problem," *J. Roy. Stat. Soc. C*, vol. 28, no. 2, pp. 126–135, 1979.

[34] K. J. Worsley, "An improved Bonferroni inequality and applications," *Biometrika*, vol. 69, no. 2, pp. 297–302, 1982.

[35] T. L. Lai, "Sequential analysis: Some classical problems and new challenges," *Statist. Sinica*, vol. 11, no. 2, pp. 303–350, 2001.

[36] D. M. Hawkins and D. H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*. New York, NY, USA: Springer-Verlag, 1998.

[37] J. J. Pignatiello, Jr., and T. R. Samuel, "Estimation of the change point of a normal process mean in SPC applications," *J. Qual. Technol.*, vol. 33, no. 1, pp. 82–95, 2001.

[38] E. Gombay, "Sequential change-point detection with likelihood ratios," *Statist. Probab. Lett.*, vol. 49, pp. 195–204, Aug. 2000.

[39] B. Brodsky, *Change-Point Analysis in Nonstationary Stochastic Models*. Boca Raton, FL, USA: CRC Press, 2017.

[40] J. Chen and A. K. Gupta, *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*, 2nd ed. Cambridge, MA, USA: Birkhaüser, 2011.

[41] J. Antoch and D. Jarušková, "Testing a homogeneity of stochastic processes," *Kybernetika*, vol. 43, no. 4, pp. 415–430, 2007.

[42] B. H. Lindqvist, "On the statistical modeling and analysis of repairable systems," *Stat. Sci.*, vol. 21, no. 4, pp. 532–551, 2006.

[43] W. J. Conover, *Practical Nonparametric Statistics*. New York, NY, USA: Wiley, 1999.

[44] A. Mood, "On the asymptotic efficiency of certain nonparametric two-sample tests," *Ann. Math. Statist.*, vol. 25, no. 3, pp. 514–522, 1954.

[45] B. Duran, "A survey of nonparametric tests for scale," *Commun. Statist.-Theory Methods*, vol. 5, no. 14, pp. 1287–1312, 1976.

[46] K. Emanuel, "Tropical cyclones," *Annu. Rev. Earth Planet. Sci.*, vol. 31, pp. 75–104, Feb. 2003.

[47] K. Emanuel, "Hurricanes: Tempests in a greenhouse," *Phys. Today*, vol. 59, no. 8, pp. 74–75, 2006.

[48] K. Emanuel, "Environmental factors affecting tropical cyclone power dissipation," *J. Climate*, vol. 20, no. 22, pp. 5497–5509, 2007.

[49] M. Bhaduri and C.-H. Ho, "On a temporal investigation of hurricane strength and frequency," *J. Environ. Model. Assessment*, pp. 1–13, Oct. 2018. doi: 10.1007/s10666-018-9644-0.

[50] G. J. Szekely and M. L. Rizzo, "Hierarchical clustering via joint between-within distances: Extending Ward's minimum variance method," *J. Classification*, vol. 22, no. 2, pp. 151–183, 2005.

[51] M. L. Rizzo and G. J. Székely, "Disco analysis: A nonparametric extension of analysis of variance," *Ann. Appl. Statist.*, vol. 4, no. 2, pp. 1034–1055, 2010.

[52] N. A. James and D. S. Matteson, "ecp: An R package for nonparametric multiple change point analysis of multivariate data," *J. Stat. Softw.*, vol. 62, no. 7, pp. 1–25, 2014.

[53] S. E. Rigdon and A. P. Basu, *Statistical Methods for the Reliability of Repairable Systems* (Wiley Series in Probability and Statistics). Hoboken, NJ, USA: Wiley, 2000.

[54] C.-H. Ho, "Forward and backward tests for an abrupt change in the intensity of a Poisson process," *J. Stat. Comput. Simul.*, vol. 48, nos. 3–4, pp. 245–252, 1993.

[55] S. Xie, M. Imani, E. R. Dougherty, and U. M. Braga-Neto, "Nonstationary linear discriminant analysis," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, 2018, pp. 161–165.

[56] M. Sugiyama and M. Kawanabe, *Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation*. Cambridge, MA, USA: MIT Press, 2012.

[57] M. Imani, S. F. Ghoreishi, and U. M. Braga-Neto, "Bayesian control of large MDPs with unknown dynamics in data-poor environments," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst. (NIPS)*, 2018, pp. 8157–8167.

[58] L. Bardwell, P. Fearnhead, I. A. Eckley, S. Smith, and M. Spott, "Most recent changepoint detection in panel data," *Technometrics*, vol. 61, p. 1, pp. 88–98, 2019.
[59] Y. Lepage, "A combination of Wilcoxon's and Ansari–Bradley's statistics," *Biometrika*, vol. 58, no. 1, pp. 213–217, 1971. doi: 10.2307/2334333.

FELIX ZHAN is a member of AEOP/RET program with the University of Nevada Las Vegas. His research interests include computer and data science.

ANTHONY MARTINEZ is a member of AEOP/RET program with the University of Nevada Las Vegas. His research interests include computer and data science.

NILAB RAI is a member of AEOP/RET program with the University of Nevada Las Vegas. His research interest includes computer and data science.

RICHARD MCCONNELL is a member of AEOP/RET program with the University of Nevada Las Vegas. His research interests include computer and data science.

MATTHEW SWAN is currently pursuing a degree with the University of Nevada Las Vegas. His research interests include computer and data science.

MOINAK BHADURI is currently a tenure-track Assistant Professor with the Department of Mathematical Sciences, Bentley University, Massachusetts. He studies spatio-temporal Poisson processes and others like the self-exciting Hawkes or log-Gaussian Cox processes, which are natural generalizations. Recently, while completing his doctoral work, he was a Research Member of the Big Data Hub within the Department of Computer Science, University of Nevada, Las Vegas. His current research interests include change detection in non-stationary systems, especially through trend permutations.

JUSTIN ZHAN is currently an ARA Scholar and a Professor of data science with the Department of Computer Science and Computer Engineering, University of Arkansas. He is also an Adjunct Professor with the Department of Biomedical Informatics, University of Arkansas for Medical Sciences. He has been the Director of the Big Data Hub and a Professor with the Department of Computer Science, College of Engineering, University of Nevada, Las Vegas. He has published 230 articles in peer-reviewed journals and conferences and delivered more than 30 keynote speeches and invited talks. He has been involved more than 45 projects as a Principal Investigator (PI) or a Co-PI, which were funded by the National Science Foundation, Department of Defense, National Institute of Health, and so on. His research interests include data science, biomedical informatics, artificial intelligence, information assurance, and social computing. He was a Steering Chair of the IEEE International Conference on Social Computing (SocialCom), IEEE International Conference on Privacy, Security, Risk, and Trust (PASSAT), and the IEEE International Conference on BioMedical Computing (BioMedCom). He has served as a Conference General Chair, a Program Chair, a Publicity Chair, a Workshop Chair, and a Program Committee Member for 150 international conferences; he has also served as an Editor-in-Chief, an Editor, an Associate Editor, a Guest Editor, an Editorial Advisory Board Member, and an Editorial Board Member for 30 journals. He has been an Editor-in-Chief of *International Journal of Privacy, Security and Integrity* and *International Journal of Social Computing and Cyber-Physical Systems*.

LAXMI GEWALI is currently a Professor with the Department of Computer Science, University of Nevada Las Vegas. His research interests include computer and data science.

PAUL OH is currently a Professor with the Department of Mechanical Engineering, University of Nevada Las Vegas. From 2000 to 2014, he was a Professor of mechanical engineering with Drexel University, Philadelphia, and he founded and directed the Drexel Autonomous Systems Laboratory (DASL). He was the former Program Director for robotics with the National Science Foundation, where he managed a portfolio that supported almost all robotics research in American universities. He has worked for Boeing, the Office of Naval Research, and the NASA Caltech/Jet Propulsion Lab. He is also the Lincy Professor of unmanned aerial systems with the Mechanical Engineering Department, Howard R. Hughes College of Engineering. He is establishing an unmanned autonomous systems laboratory at UNLV, complete with a fleet of drones and several humanoid robots. His research interests include robotics and data science.