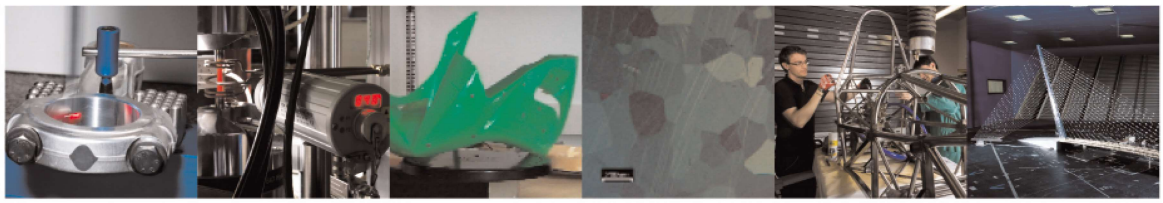

Sequential detection framework for real-time biosurveillance based on Shiryaev-Roberts procedure with illustrations using COVID-19 incidence data

Zamba, K. D.; Tsiamyrtzis, P.

This is an Accepted Manuscript of an article published by Taylor & Francis in SEQUENTIAL ANALYSIS on 10 May 2021, available online: https://doi.org/10.1080/07474946.2021.1912503

Sequential Detection Framework for Real-time Biosurveillance Based on Shiryaev–Roberts Procedure with Illustrations Using COVID-19 Incidence Data

K. D. Zamba[1] **and Panagiotis Tsiamyrtzis**[2]
[1]Department of Biostatistics, University of Iowa, Iowa City, USA
[2]Department of Mechanical Engineering, Politecnico di Milano, Milan, Italy

Abstract: This manuscript develops a detection framework using Bayesian philosophy by adaptation of Shiryaev (1963) and Roberts (1966) methods. We propose two unifying versions directly applicable in industrial process control; and easily extendable to public health infectious disease surveillance *via* some data detrending and/or demodulation. The root idea uses the sum of likelihood ratios upon which an optimal stopping criterion is based. It sets a prior on the epoch of a change, allows the flexibility to elicit a prior distribution on other process parameters, and attempts to minimize an expected loss function. A sensitivity analysis is conducted for validation and performance assessment, and analytical formulas derived. The methods are successfully applied to the European Union Centre for Disease Control (ECDC) open source global COVID-19 incidence data. We further lay out scenarios where interest may switch to the detection of separate outbreaks with similar syndromes during an already evolving epidemics. We view our approach as a toolkit with a potential to augment early reports to sentinels in syndromic surveillance and in biosurveillance.

Keywords: Bayesian sequential update; Biosurveillance; Change point; COVID-19; Dynamic processing, Shiryaev–Roberts procedure, Syndromic surveillance.

Subject Classifications: 62L10; 60G15; 60G40; 62F15.

1. INTRODUCTION

Biosurveillance, like any other sequential aberration detection field, relies on *watchdog*-type of statistical procedure for early detection of unanticipated change in stochastic measurements that occur in real-time or near real-time. In countries around the world, medical data are being used to assess data-driven evidence of natural epidemics or intentional release of biological agents with operations similar to natural epidemics. The real-time *watchdog*-type of platform created in biosurveillance has allowed a shift in disease surveillance from the classical retrospective disease chart

Address correspondence to K. D. Zamba, Department of Biostatistics, The University of Iowa, Iowa City, IA 52242 , USA; E-mail: gideon-zamba@uiowa.edu

review to a prospective and early disease detection, where medical characteristics become investigational tools used to augment early reports to sentinels. Examples of these medical characteristics and investigational parameters are emergency room visits, over-the-counter sales, veterinary data and medical and public health information. See for example Arnon et al. (2001); Dennis et al. (2001); Henderson et al. (1999); Inglesby et al. (1999); Inglesby et al. (2000); Buehler et al. (2003); Green and Kaufman (2002) and Pavlin (2003) for relevant publications.

The optimality of surveillance procedures is usually assessed through minimizing a desired detection delay penalty function, subject to a constraint on a preset false alarm rate. These known facts of standard control theory have long been a central part of classical disease surveillance, laboratory-based outbreak detection algorithms, so that tools that are usually employed to reach optimality have incorporated false alarm rates, and reluctance to act in the presence of contaminants, as part of their operating characteristic. In the US for example, the Centers for Disease Control and Prevention (CDC) has routinely applied cumulative sum techniques (Cusum) to laboratory-based data for outbreak detection; see Hutwagner et al. (1997) and Stern and Lightfoot (1999). As most industrial control tools rely on strict statistical underpinnings, and on a relative knowledge of distributional parameters, their transition to medical data application has fallen short of optimality. As an example in biosuveillance, the main incidence parameters may be unknown, or too raw for a steady-state inference. Sometimes the main incidence parameter may be unstable or depict some unusual structural breaks. In the case of infectious diseases with seasonal trend such as influenza, for example, it has been recognized that their year-to-year incidence parameter behaves like a constantly moving target, thereby generating new data processes to be monitored each year *de novo*. In cases such as these, statistical optimality becomes clouded by the background of variability and structural changes inherent to the data. For some in-depth discussion on biosurveillance and control methods in medical surveillance, see Woodall (2006); Zamba et al. (2008); Zamba et al. (2013), Shmueli and Burkom (2010). Stoto et al. (2004) was quick to point out that as background variability and structural changes affect these data, the monitoring scheme used for signal detection yields more false signals than anticipated; thus resulting in sub-par optimality and sometimes coin toss decisions.

The contextual platform on which biosurveillance operates usually demands individual data monitoring. Cusums, exponentially weighted moving averages (Ewma), and change point methods might seem to be first-blush natural candidates. These tools are limited when faced up with health-related data–as their performance tends to be governed by statistical assumptions such as independent readings, normal distribution with known parameters and a requirement of stable and steady-state historical data set (HDS); which clearly lack in biosurveillance; see Hawkins and Olwell (1997). For example the statistical underpinnings of Cusums make them lesser suitable candidates for monitoring real-time biosurveillance data. In addition, Cusums are better known for responding to step (not gradual) changes. Ewma requires advance knowledge and stability of the process parameters. Change point methods, although less demanding in their calibration needs, and better suited for handling individual reading and start-up processes, rely on independent and identically distributed normal assumption along with the assumption of change from a constant set of distributional parameters to another constant set; see for example Hawkins et al. (2003); Hawkins and Zamba (2005a); and Hawkins and Zamba (2005b). Other methods usually applied to autocorrelated data can be found in Zhang (1997) and Winkel and Zhang (2007).

The current manuscript presents a surveillance method based on Bayesian philosophy. The Bayesian approach to SPC was spearheaded by the pioneering work of Shiryaev (1963) and

Roberts (1966), which considered change point types of modeling with Bayesian thinking in view. Hodley (1981), for example, offered a Bayesian alternative to c-chart, while Woodward and Naylor (1993) applied Bayesian approaches to handling short runs in normal data. Recently, Tsiamyrtzis and Hawkins (2005), Tsiamyrtzis and Hawkins (2010) and Tsiamyrtzis and Hawkins (2019) provide Bayesian change point approaches using a mixture of distributions in modeling Gaussian or Poisson phase I data, and Apley (2012) introduced posterior distribution plots for phase II monitoring.

The approach we propose herein is based on the Shiryaev–Roberts change point detection method such as outlined in Shiryaev (1963) and Roberts (1966). It uses a sequential scheme to build an optimal stopping criterion based on the sum of likelihood ratios while not necessarily requiring independent observations. The approach sets a prior distribution for the epoch of aberration τ, and attempts to minimize an expected loss function. An optimal decision is reached based on the partial sum of likelihood ratios exceeding a threshold directly related to a preset false alarm rate. The Shiryaev–Roberts approach can be generalized to accommodate settings where the pre- and post-change densities are unknown, in accordance with Lorden and Pollack (2005), and can be adapted to autocorrelated series *via* adequate transformations. Historically, Bayesian methods have been slow to percolate into control theory and biosurveillance due in part to their theoretical and computational intimidation and also to problems surrounding prior density specifications. Regardless, they appear to be quite appealing in modeling the type of data gathered in medical settings where short runs and unknown parameters are prevalent. Our penchant for Bayesian thinking in this context relates to their ease of adaptation to individual monitoring, their ease of handling unknown parameters, added to the luxury that comes with deriving posterior predictive distributions–key ingredients for error management.

The manuscript is organized as follows: Section 2 defines the proposed Shiryaev and Shiryaev-Roberts Bayesian frameworks and the algebraic formulas entailed. In Section 3, we study the operating characteristics of the methods using simulations, and apply the methods to ECDC COVID-19 global incidence data. Section 4 outlines a sensitivity analysis, documenting performance for overlapping epidemics. In Section 5, we close our manuscript with discussion and conclusion.

2. DATA AND BAYESIAN MODEL

2.1. The Data Structure

Biosurveillance data can be likened to readings from a propagated outbreak in which incidence flattens at first, then slowly increases to an apex, after which it slowly decreases as a function of decreasing transfer rate associated with herd immunity, public health measures and/or other eradicative measures. As these infective states are likely related to influenza-like-illnesses, their incidence follows similar epic curves as those observed in diseases of human vector and those spread by contacts through bodily fluids or droplets. Regardless of their etiology, statistical observations suggest an underlying latent serial process fluctuating across a constant value (corresponding to a baseline *noise*, or a *walk* segment) before epic surges are observed. Immediately following this segment is a *ramp model*, or a *rise* segment where the incidence data display positive association with time scale, which continues until the apex, a turning point from which recordings have a negative trend (a *decay* segment), which later transitions into a constant *noise* segment. The slope indicating the *ramp model* and the *decay* segment may be epidemic-by-year specific; and so is the

beginning of the *ramp model*, the peak infective time, and the end of the *decay* segment. See for example the top panel of figure 1 for an illustration of the serial evolution of COVID-19 incidence data. Firstly, we focus our attention to the baseline and rise sections of the data with the goal to capture the beginning of the inflection (*i.e.* the epoch τ) while optimizing the false alarm probability. Next, by using an adequate demodulation we switch attention to the entire epidemic curve, on which spikes of additional outbreaks co-occur during an epidemic. For the latter of these *foci*, a good argument can be made for surveillance of COVID-19 during a regular flu season or vice versa.

2.2. The Statistical Model

Assume X_i to be time-indexed normally distributed readings, or approximately so, observed with the goal to identify an epoch τ of a persistent increase in their mean structure. Relating back to the epidemic curve, this epoch τ can be viewed as the juncture between the *walk* and the *ramp model* segments. In that case, one is led to consider:
Baseline data: for $i = 1, 2, \ldots, \tau - 1$

$$X_i \overset{iid}{\sim} N\left(\mu, \ \ \sigma^2\right)$$

Rise data: for $j = \tau, \tau + 1, \ldots, n$

$$X_j \overset{iid}{\sim} N\left(\mu + (j - \tau + 1)\delta, \ \ \sigma^2\right).$$

Assuming demodulated data and/or a first difference detrending filter Y_i as follows:
for $1 \leq i < \tau$ we have

$$Y_i = X_i - X_{i-1} \sim N\left(0, \ \ 2\sigma^2\right),$$

while for $j \geq \tau$ one has

$$Y_j = X_j - X_{j-1} \sim N\left(\delta, \ \ 2\sigma^2\right).$$

The Y_i process defines a *walk* upon which there is a *shock* starting time τ.

The mixture of likelihood functions that encapsulate the process Y_i may be summarized as:

$$Y_i|\sigma^2, \delta, \tau \sim \left\{ \begin{array}{ll} f_o \equiv N\left(0, \ \ 2\sigma^2\right) & \text{if } 1 \leq i < \tau \\ f_1 \equiv N\left(\delta, \ \ 2\sigma^2\right) & \text{if } \tau \leq i \leq n \end{array} \right\}. \tag{2.1}$$

Note that the model in (2.1) is similar to a family of change point models with random change point τ. In order to find the epoch τ of change in Y_i process, we resort to methods from Bayesian Statistical Control, namely those of Shiryaev and Shiryaev-Roberts; see Shiryaev (1963) and Roberts (1966). The methods are documented for their optimality properties in detecting persistent shifts. One is inclined to consider the *baseline segment* as the statistical in-control state, and the *ramp* portion as an out-of-control state. In the classical implementation of the Shiryaev and the Shiryaev-Roberts methods, it is necessary to pre-specify both the in-and out-of-control states, and to target the one-parameter change point problem as dictated by the random variable τ; along the same vein as known parameter control schemes. Our current development differs from the classical approach in that we allow the parameters in the likelihood function (2.1) to follow a prior

distribution while keeping our gaze on the change point τ. By so doing, we relax the known-parameter distributional assumption by allowing the remaining parameters (δ, σ^2) to be nuisance in (2.1), thereby creating a generalized version of the classical application of the Shiryaev and the Shiryaev-Roberts methods. Within the Bayesian paradigm, the nuisance parameters are handled as random variables for which one can elicit a prior distribution before any data are recorded. The hierarchical structure adopted follows:

$$\begin{aligned} \delta|\sigma^2 &\sim N(\delta_0, \ k\sigma^2) && (2.2) \\ \sigma^2 &\sim \text{Inv.Gamma}(\alpha, \ \beta); && (2.3) \end{aligned}$$

where the four hyper-parameters δ_0, k, α and β remain to be exposed. This structure assumes $(\delta, \sigma^2) \perp \tau$, and will require the hyper-parameters to operate. This is not an over-parametrization of the detection scheme, but rather an attempt to generalize our schema to account for a background of variability such as those observed in syndromic surveillance settings. To elaborate further, δ_0 refers to the size of the *shock* one wishes to detect. This *shock* is intuitively expected to be greater than the maximum possible fluctuation one can anticipate in the *baseline segment*; barring false signal. The hyper-parameter k reflects the uncertainty around the mean choice of δ_0. The suggestion is to keep $k \approx 1$, or $k > 1$ if one wishes to be as less informative as possible. Finally, α and β will handle the prior regarding σ^2; with small values associated with more vague priors, which are usually connected to robustness due to finite dimensionality of the parameter space; see for example Gelman et al. (2014).

Theorem 2.1. Using the likelihood in (2.1) and the prior settings in (2.2) & (2.3), the marginal distribution of the data point is as follows:

$$\begin{aligned} Y_i|\tau \sim f_o(y_i) &= \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)} \frac{1}{\sqrt{4\pi\beta}} \left[1 + \frac{y_i^2}{4\beta}\right]^{-\alpha-\frac{1}{2}} \quad \text{if } 1 \leq i < \tau \\ Y_i|\tau \sim f_1(y_i) &= \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)} \frac{1}{\sqrt{2(k+2)\pi\beta}} \left[1 + \frac{(y_i-\delta_0)^2}{2(k+2)\beta}\right]^{-\alpha-\frac{1}{2}} \quad \text{if } \tau \leq i \leq n. \end{aligned}$$

For proof, please see the technical appendix section.

2.3. The Shiryaev–Roberts Approach to Biosurveillance

The sequential change detection criterion, under the Shiryaev and Shiryaev-Roberts approaches, makes use of the likelihood ratio (see technical appendix for proof):

$$\frac{f_1(y_i)}{f_0(y_i)} = \sqrt{\frac{2}{k+2}} \left[\frac{1 + \frac{y_i^2}{4\beta}}{1 + \frac{(y_i-\delta_0)^2}{2(k+2)\beta}}\right]^{\alpha+\frac{1}{2}}.$$

Under the assumption that τ follows a geometric prior distribution, *i.e.*

$$\tau \sim Geometric(p),$$

and in accordance with Shiryayev (1963), we base our stopping criterion on the posterior probabilities

$$P(\tau \leq n|y_1, y_2, \ldots, y_n) = P_n = \frac{R_n^S}{R_n^S + 1};$$

where

$$R_n^S = \frac{p}{(1-p)^{n+1}} \sum_{k=1}^{n} (1-p)^k \prod_{j=k}^{n} \frac{f_1(y_j)}{f_0(y_j)};$$

and stopping time:

$$\inf\{n : P_n \geq p^\star\}, \tag{2.4}$$

with $p^\star$ predetermined by optimizing a required false alarm rate.
Roberts (1966) set $p \to 0$ and derived the recursive formula known as Shiryaev-Roberts approach:

$$R_n^{SR} = \sum_{k=1}^{n} \prod_{j=k}^{n} \frac{f_1(y_j)}{f_0(y_j)} = \left(1 + R_{n-1}^{SR}\right) \frac{f_1(y_n)}{f_0(y_n)};$$

with initial value $R_0^{SR} = 0$ and stopping time:

$$\inf\{n : R_n^{SR} \geq \gamma\}, \tag{2.5}$$

where γ is predetermined, just like in the Shiryayev case, according to the required false alarm rate. Direct mathematical optimization of these formulas so as to obtain $(p^\star, \gamma)$ is a daunting task; thus, we resort to numerical optimization to approximate the solutions.

3. OPERATING CHARACTERISTICS AND APPLICATION

A solution to the stopping criteria (2.4) and (2.5) is not amenable to direct mathematical formulas; points that have been extensively discussed by Pollack (1987) and Yakir (1995). Instead of direct mathematical optimization, we resort to a simulation study in order to 1.) establish $p^\star$ and γ, and 2.) assess the operating characteristics under both normal and contaminated persistent regimes. Note that some key factors may influence the performance of any sequential detection algorithm. Among these one can identify the size of a shift (δ) and its epoch–*i.e.* the position of the shift relative to the series length and relative to the volume of the *walk* prior to the *shock* segment. As an example, shifts that occur after short *walk* series may rarely enjoy the same detection rate as those occurring after long and steady-state series. To calibrate $p^\star$ and γ, we conduct a Monte Carlo simulation to define empirical thresholds, on the basis of false alarm probability, when both systems are tuned to detect $\kappa \times \sigma$-shift in standard Gaussian series. We target series of various lengths; although a window of size 365 would be realistic for the sake of the current application–in which real-time daily incidence is measured all year around. One hundred thousands (100,000) Monte Carlo simulations were used for tuning purposes. After establishing the stopping rules, we then spike the series with contaminations after $\tau =$ 10, 20, 30, 40, 50 steady-state observations, then subsequently by step 50 till 300. We also include shifts at mid-series; after 182 observations. We only report a few tables as an exposition of our methods in this manuscript. The `R` functions used to generate these thresholds are appended to the manuscript. Users can modify these functions as they have firsthand knowledge or some educated estimates of their process parameters. These

functions serve as means to simplify practitioners' lives so they can proceed straight to charting after these thresholds are generated. The entries in table 1 are optimal thresholds for 1/2- and 1σ-shift in a data window of 365 *walk* series. The limits are derived from the quantiles of the maximum Shiryaev and maximum Shiryaev-Roberts statistics. We use the `R` software version `R.4.0.0` on `LINUX` platform with generator seed `123321`.

Table 1. Threshold as function of False Alarm Probability

	$\sigma/2$				σ			
	0.050	0.025	0.010	0.001	0.050	0.025	0.010	0.001
Shiryaev ($p^\star$)	.0378	.0436	.0521	.0807	.0543	.0673	.0881	.1686
Shiryaev-Roberts (γ)	38.84	45.08	54.19	86.51	56.98	71.55	95.65	200.03

Table 2. Achieved False Alarm Probability

Shiryaev	0.044
Shiryaev-Roberts	0.044
DI Cusum-K	0.045
DI Cusum-U	0.043

3.1. Operating Characteristics

A combination of methods are used to measure the operating characteristic of the Shiryaev sequential Bayesian methods, and compare them to other monitoring schemes. These characteristics are:

- NS: The proportion of series that have gone degenerate; *i.e.* the series failed to signal in the presence of contamination
- FS: The proportion of signals prior to contamination time τ
- DS: The average length of time from contamination time τ and detection time $\iota \geq \tau$.

The shifts considered are $\{0.25, 0.50, 0.75, 1, 2\}$-$\sigma$; which are introduced after steady state observations in a window of size 365. The idea behind these simulation settings is to be able to document the systems' reaction to early or late shifts during the monitoring period. We used the false alarm probability (FAP) approach to design the optimal schemes at an in-control nominal FAP of 0.05. The FAP concept finds its niche in sequential multiple hypotheses testing procedure in which we ensure the overall family wise error rate is set at some predetermined level, which is 5% in our case. To maintain an overall low false alarm probability we adopt the Sidák correction, Sidák (1967), which accounts for conservativeness or anti-conservativeness associated with dependent tests. The FAP, based on joint probability assessment, is defined as the probability of obtaining at least one false alarm during the monitoring period. This approach is preferred in low volume

Table 3. Operating characteristic measured by delay to signal (**DS**)

						τ					
	δ	10	20	30	40	50	100	150	182	200	300
	.50	67.4	68.8	69.1	64.0	67.3	66.5	62.2	58.3	56.4	34.7
SH	.75	26.9	25.8	25.9	24.8	25.6	25.6	25.8	24.9	25.3	23.0
	1.0	15.6	14.7	14.6	14.5	14.6	14.6	13.9	14.5	14.5	14.4
	2.0	6.0	5.6	5.4	5.5	5.6	5.5	5.5	5.5	5.5	5.5
	.50	67.7	69.4	70.1	64.4	67.5	67.1	62.3	58.6	56.7	34.8
SHR	.75	26.9	25.8	26.0	24.8	25.6	25.6	25.9	25.0	25.4	23.0
	1.0	15.6	14.7	14.6	14.5	14.7	14.6	13.9	14.5	14.5	14.5
	2.0	6.0	5.6	5.4	5.5	5.6	5.5	5.5	5.5	5.5	5.5
	.50	43.9	43.6	43.8	43.3	44.6	43.5	43.0	42.6	42.4	35.8
Cusum-K	.75	23.6	23.2	23.2	22.3	23.2	22.9	22.9	23.0	23.0	22.4
	1.0	16.1	15.8	15.8	15.7	15.7	15.6	15.3	15.7	15.6	15.6
	2.0	7.2	7.0	6.9	7.0	7.1	7.0	6.9	6.9	6.9	7.0
	.50	137.7	89.2	79.7	67.6	74.8	63.0	56.7	52.4	50.0	36.5
Cusum-U	.75	84.7	60.8	49.7	44.2	41.1	30.1	27.4	26.9	26.1	23.9
	1.0	61.5	41.0	30.8	26.0	22.9	18.7	17.1	17.2	18.7	16.4
	2.0	23.9	12.1	10.1	9.3	8.7	7.8	7.5	7.5	7.3	7.3

startup productions because it takes into account issues pertaining to multiplicity of testing criterion, simultaneity and dependence. For more on FAP, see King (1954), Jones and Champ (2002), Chakraborti et al. (2008). The FAP probability is different from the usual false alarm rate (FAR), which is the probability of a false alarm at every sampling stage–computation of which is based on the marginal distribution of the charting statistics in state of in-control. The FAR, used as an operating characteristic, does not take into account multiplicity and dependence among tests. Table 2 shows the achieved FAP when thresholds are chosen at a nominal FAP value of 0.05; with the corresponding $(p^\star, \gamma) = (0.03784, 38.8434)$ respectively for Shiryaev and Shiryaev-Roberts. These limits were obtained when $n = 365$, $\delta_0 = 0.5$, $k = \alpha = \beta = 1$; while $p = 0.001$ for Shiryaev; and Shiryaev-Roberts is free of the parameter p. We next compare the Bayesian Shiryaev and Shiryaev-Roberts schemas to two decision interval cusums. The first decision interval cusum is a known parameter cusum (Cusum-K), while the second is a cusum with sequentially estimated parameters (Cusum-U). For Cusum-U, the mean and the variance at time k are running mean and running variance up till time *k-1*. Needless to say that by design, Cusum-K, the known parameter cusum, would be optimal for small and persistent shifts. The setting of Cusum-K would be the best case scenario if a long historical dataset is available for parameter estimation. In our current study though, this safety net is unavailable and the parameters have to be estimated on the fly–thus, the necessity of Cusum-U.

Results: All four monitoring methods show a bit of conservativeness, as their achieved FAP fall

a little below the nominal value of 0.05 (see table 2). Their performance being similar in the FAP department, we can compare them on the basis of other operating characteristics. That Cusum-K outperforms all other methods in detecting persistent small shifts is not a surprise (see table 3). This is the hallmark of known parameter DI cusums. But this advantage is quickly offset by the Shiryaev and the Shiryaev-Roberts nearly equal, if not better, performance in moderate to large shifts. What has been abundantly clear is the poor performance of the estimated parameter cusum (Cusum-U) compared to all other methods–specifically when shifts occur after a small sample of in-control observations. This is not too difficult to see. As parameters are estimated to build Cusum-U, this scheme behaves like a self-starting cusum. When contaminated values are fed into the running mean, they are also fed into both the estimated means and variances. The combined effect is that the cusum now sees these contaminations as a normal part of its process; and even if it started out rising initially, it will progressively adapt to the readings and end up falling below the decision threshold. Consequently, it will take a very long time for Cusum-U to signal. As the in-control gathering gains in size, we will eventually find ourselves in the Law of Large Numbers territory; as a result its performance will progressively compare to that of Cusum-K. This finding is not new. It simply enhances the work by Hawkins and Olwell (1997) on unknown parameter cusums. What transpires from this early prescription is, when the monitoring parameters are unknown, the Shiryaev and Shiryaev-Roberts methods are respectable candidates for post-shift operating characteristics. Given that the previous simulation study is designed for sustained and persistent post-shift distribution, what remains to be exposed is whether the system could function in the presence of isolated temporary causes of variability. This latter inquisition reflects the hallmark of epidemics and the nature of data gathered in public health arena. On a separate note, it has been observed that for a system designed for 0.5σ shift in the mean of a standard Gaussian process, there were no degeneracy by the moment contaminations of size 1σ have been introduced. All such series have signaled from 1σ onwards.

3.2. Application to EU COVID-19 Data

The European Centre for Disease Prevention and Control (ECDC) was created in 2004 with a mission to strengthen Europe's defense against infectious diseases. In 2019, a strain of the severe acute respiratory syndrome (SARS) species, SARS-CoV-2, was discovered and linked to COVID-19, a disease that brought a global COVID-19 pandemic. During the COVID pandemic, ECDC publishes real-time data related to the global outbreak as well as the number of people affected in the EU. The data are in public domain and available at `https://www.ecdc.europa.eu/en`. We provide a pictorial representation of the global COVID cases from ECDC after applying a seven-day smoothing filter to the real-time observations starting December 31, 2019. We also provide a second graph in which we zoom in on the early real-time data gathering on COVID-19 infections. This early part of the data is the object of our illustrative example (see figure 1). Our goal throughout this prescription is to assess how fast our sequential algorithm reacts to real-time data gathering in order to sound and report alarms to sentinels. Three methods, tuned at the same FAP, were compared with respect to their ability to detect early surge in reported cases of the pandemic. These are the Shiryaev, the Shiryaev-Roberts and the Cusum-U. We sequentially compute their statistics as the data stream evolves and compared them to their individual threshold. The Shiryaev and the Shiryaev-Roberts statistics after crossing their thresholds suggest unusual surge as far back as observation `20`, while the Cusum-U took a two-day delay to signal at `22`.

These correspond to January 19 and January 21 respectively. Parts of the sequence of actual counts leading up to those dates are `(0, 0, 1, 0, 1, 0, 5, 17, 136, 20, 153)`, and the first biggest global daily surge was observed on January 19, 2020. The Shiryaev and the Shiryaev-Roberts jumped on the first biggest surge, while the cusum focussed on the second.

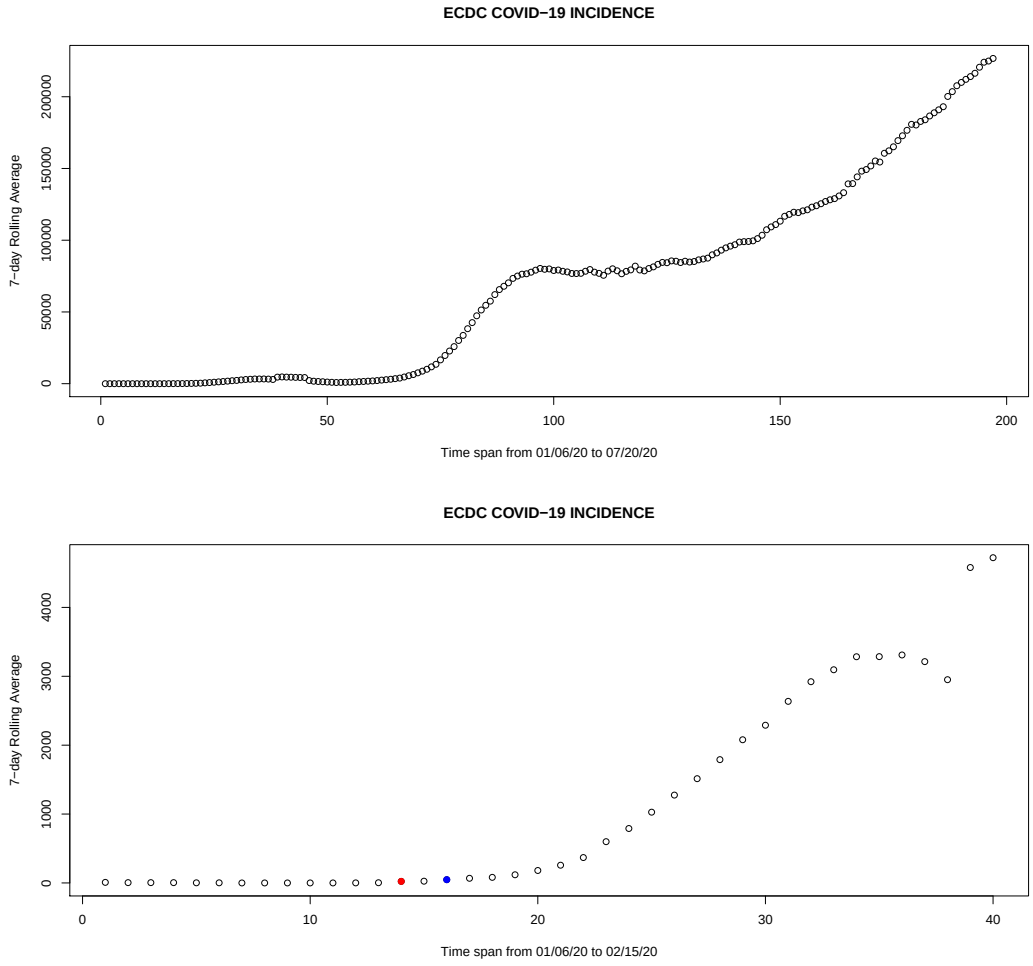

Figure 1. EU global COVID-19 surveillance data. The top panel is the daily cases spanning December 31, 2019 to July 20, 2020. The bottom panel highlights the early forty-day observations. In red, Shiryaev and Shiryaev-Roberts detections, in blue the unknown parameter Cusum.

4. SENSITIVITY ANALYSIS FOR OVERLAPPING EPIDEMICS

Now we switch our attention to co-occurring outbreaks during an epidemics. A primary problem of interest to researchers and policy makers is the ability for a surveillance tool to detect an epidemic in general, provide a backup data based evidence to medical diagnosis, and augment reports to sentinels. This concern can be put to rest by devising a surveillance algorithm rooted in timeliness between activity and detection while optimizing the statistical type I and type II errors, to a reasonable degree, as it is traditionally the case. A different set of more challenging problems

create itself when two different epidemics with similar syndromes overlap. For example, one may be interested in the usual influenza operation when a SARS-CoV-2 epidemics is underway or another biological agent is released. This scenario presents one of the most challenging problems in the area of syndromic surveillance especially if one or both infectious diseases under investigation have a paucity of historical data while sharing similar syndromes. Medical timely screening and diagnosis would be paramount in those settings, while most surveillance tools would rather be reactionary, taking a back sit to firsthand medical expertise. Our hope is, at best, one of the two overlapping epidemics would have some respectable history available before the operation of the subsequent one; and these serial readings, or a demodulated version can be used in error management fashion. In what follows we adopt some of the techniques similar to Zamba et al. (2013) to investigate this issue. Zamba et al. (2013) simulated outbreaks in different ways that a pathogen could alter syndromes daily counts. Simulated extra cases are superposed on another infectious disease serial evolution and the data run through a surveillance scheme for signal detection. Within this paradigm, we simulate outbreaks of one and of two weeks and use them to spike the COVID-19 data stream. Care has been taken so that the simulated outbreaks represent multiple ways pathogens could spread through a community. Among the pathways considered, a flat outbreak corresponds to point-source infections such as Bacillus Anthracis. Linear, exponential and sigmoid outbreaks may relate to infectious diseases that are highly infectious, such as smallpox, SARS and the seasonal flu variants. These infections tend to have common flu-like symptoms such as high temperature, fever, respiratory illness, shortness of breath and cough–no wonder they are labeled in the literature influenza-like-illnesses. The belief is that these epidemics in turn would reflect on any COVID daily counts and potentially create confounding. Thus, a strategy consists of overlaying these additional counts on the detrended COVID-19 stream while mimicking a known pattern and shape that reflects the hallmark of an infective curve, according to a seasonal trend and a varying error profile. Since it is assumed that the COVID data stream has considerably gained in size, its demodulated and detrended error profile seems a reasonable baseline candidate for the simulated outbreaks. To illustrate, assuming one is interested in a 2σ-shift. If at day d the count is 100 cases and the running standard deviation up to d is 20, for the subsequent 7 (or 14) days, each day can generate between 0 to 40 extra cases. Thus, one can easily generate the extra cases and readjust their volumes to fit the distributional pathway under study. As the error profile varies between summer, fall, winter and spring, one can also implement seasonal moving windows to get a handle on various error profiles. In our simulation various error profiles are used as we space outbreaks by 10, 20, and 50 days throughout the support of the data $(\mathcal{S})$. We clean the system between outbreaks so that response to one outbreak does not impede on the other. For an outbreak of duration d (7 days, 14 days), pathway q (flat, linear, exponential, sigmoid), and yearly hit frequency f (once, twice), on can proceed as follows:

- Choose $n_1 \in \{1, \ldots, \mathcal{S} - d\}$ for the beginning of the outbreak with duration d
- Overlay the outbreak on the original series from $n_1, \ldots, n_1 + d$ according to pathway q
- Run the simulated data through the system for signal detection
- If the hit frequency f is twice, space both outbreaks either by 10, 20, or 50 observations
- Assess signal detection.

This algorithm has been followed for 10,000 Monte Carlo each combination. The flat outbreak can be likened to a temporary sustained shift after an in-control state. However, this sustained shift lasts only for 7 or 14 days, after which the readings return to normal. This may be viewed as a snapshot of an isolated cause of variability over a span of 7 or 14 days in a data window of 365 observations. Consequently, the systems are at competitive distributional disadvantage from the get go–as shifts were just occasional spikes; not continually sustained as we may wish in the perfect scenario. The linear, exponential and sigmoid are rather viewed as temporary gradual shifts that increase in size toward the end of the outbreak where they may turn out to resemble outliers around day 7 or 14. However, these also quickly fade away, making ways to the normal series to pursue its course. We define the following operating characteristics:

- NS: The series have gone degenerate as they failed to signal
- TS: The series signal during the epidemics; *i.e.* within d days from n_1
- FS: The series signal before n_1
- DS: The series signal after $n_1 + d$.

A final point that remains to be exposed is detrending the series. This may be done heuristically through first difference recursively Gaussian transformed, more elegantly using a Kalman filtering, Kalman (1960), or an adaptation of the Hodrick-Prescott (HP) one-sided filter with the Ravn-Uhlig adjustment to sensitivity (see Hodrick and Prescot (1997); Ravn and Uhlig (2002)). For more on dynamic time series detrending, see Shumway and Stoffer (2000); Mills (2003) or Enders (2010).

Results: We report the results for the 14-day outbreak. We observe an overall high FS rate with the Cusum when shifts are introduced after a long history of steady observations. These FS increase to around 50% in some cases while TS have plateaued at a maximum of 65%. In most combinations, the Cusums yield the lowest number of degenerate series compared to the Bayesian methods. The Shiryaev and the Shiryaev-Roberts have displayed consistent behaviors across combination of epidemics and pathways. In those cases the rate of degeneracy is relatively high (ranging from 4% to 23%); but the silver lining is their ability to stay true to the FS rate while maintaining a respectable proportion of true signal. The FS rate ranges between 0 and 8%; the NS rate (ranging from 4% to 23%). In the TS department, they performed reasonably well above and beyond a coin toss (ranging from 75% to 90%). The DS ranges between 3 to 31%. What has been abundantly clear in this comparison is the ability of the Cusum to be easily confused when the in-control series gain in volume. Having started to gain some confidence in the process parameter estmation, the Cusum tends to regret not having previously identified sources of variability that crept into the system under low volume. The Cusum tries to correct these sources but triggers them too early as they become false signals. This tendency has led the Cusum to interpret most early flukes as signals; although the true signals have yet to be observed. Several authors have raised the red flag on the confusion that may ensue when using the unknown parameter Cusum; see Hawkins and Olwell (1997). Note too that similar performances were obtained in disease surveillance by Zamba et al. (2013) and Stoto et al. (2004) to name a few.

5. DISCUSSION AND CONCLUSION

As pointed out by Stoto et al. (2004), traditional tools for monitoring data pertaining to biosurveillance fall short of capturing all the important structural changes depicted on these data and also lose in performance due to the inherent background variability in the readings. Our demonstration also embeds a lesson that care is needed when traditional tools are being employed, especially the ones that are sensitive to distributional assumptions or specifically designed for some type of characteristics. In this paper, we provide generalized Shiryaev and Shiryaev-Roberts approaches to control theory. The technicality allows us to elicit prior distributions on process parameters and chart in the presence of unknowns. The proposed methods perform as well as, if not better than, the known parameter Cusum for sustained post-shift series, and outperformed the self-starting Cusum for sustained changes. Their adaptation to Biosurveillance and syndromic surveillance data suggests these tools outperform Cusums over the range of *flat, linear, exponential* and *sigmoid* outbreaks. The application to COVID-19 data gathered in the European Union by the ECDC, and its positive results give us added incentive that our methods will find applicability across industrial processing, health quality monitoring, and disease surveillance.

Like any Bayesian methodology, our proposal relies on prior parameter elicitation; and this needs some extra caution. While one can alleviate the need for prior derivation by using objective (non-informative) prior distributions, we rather recommend subjective priors for the following two reasons: 1.) relevant information does typically exist and could be wasted if one opts for the objective approach, and 2.) the subjective prior will offer a head-start in the monitoring approach, enhancing performance, especially when shifts can be identified soon after the beginning of the process at the lower tail of the observation window. For subjective priors, highly informative ones, *i.e.* those with "very small" variability, should be avoided as they tend to cause high false alarm rates if mis-specified. Thus, the safest choice would be to select hyper-parameter values that yield low information priors, *i.e.* with some non-negligible variance, so that we do not impact false alarm probability, but increase detection power at startup of the process and finally allow the prior settings to wash out quickly in the recursive updating stage as data accrues.

To provide assistance and succor practitioners less familiar with Bayesian methods and computing, we document our codes and make the main functions available. The ultimate goal is to create an open source `R` menu program to generate thresholds needed for any parameter combination specific to their setting. Given the dynamic nature of the detection problem, we recommend HP filters with Ravn-Uhlig adjustment at detrending stage. In medical settings, we believe the current proposal will augment early reports of natural epidemics and intentional operations to sentinels. Though we view our approach as self-reliant, we also recognize that our proposal is not a 'one size fits all'. In the health science, Biosurveillance data may differ from one location to another even within the same country. As one switches from one geographic proximity to another, or from one syndromic variable to the next, the general philosophy behind our proposal holds; but parameters and their prior distributions will require some tweaking.

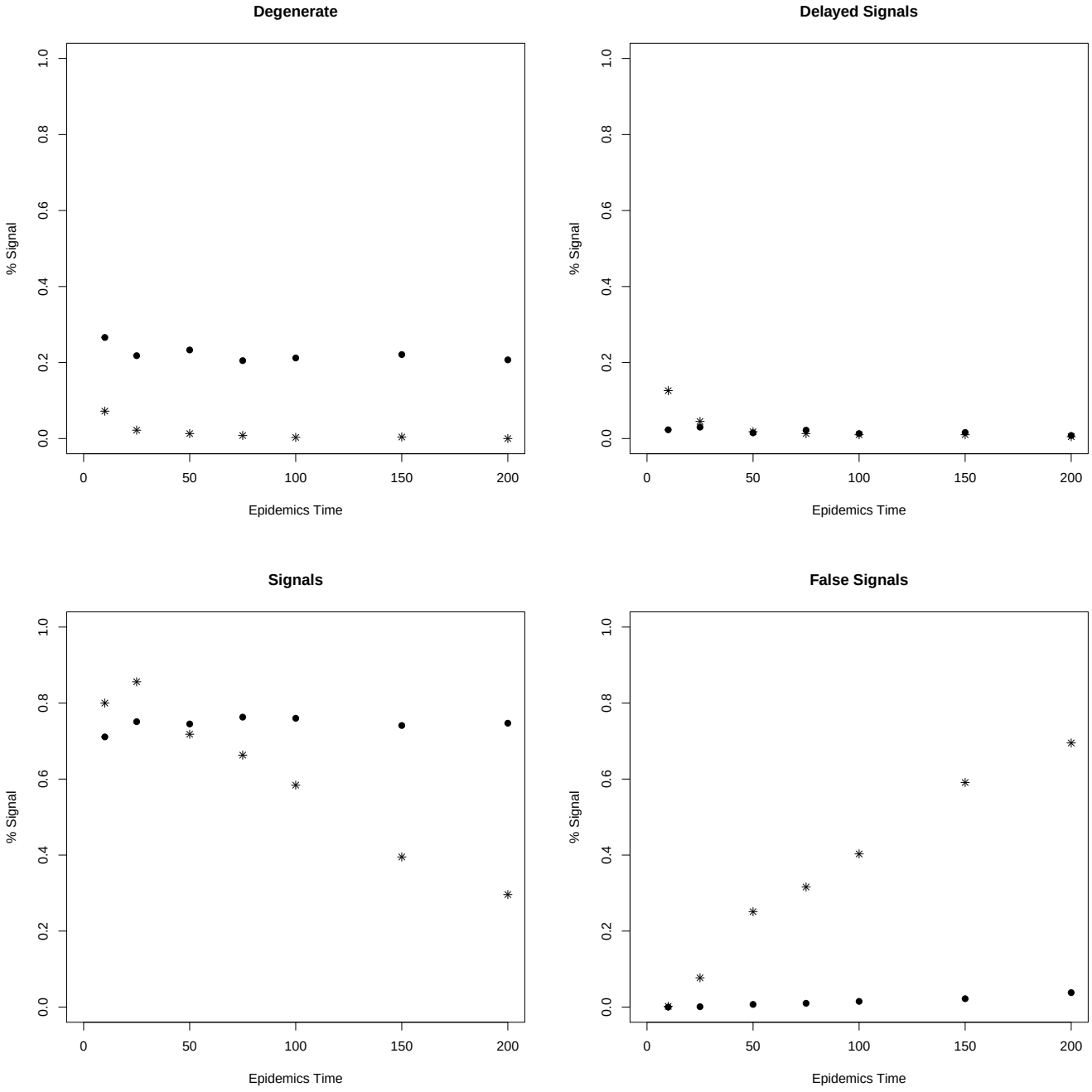

Figure 2. Flat outbreak operating characteristics with self-starting cusum ($\star$) and Shiryaev ($\bullet$)

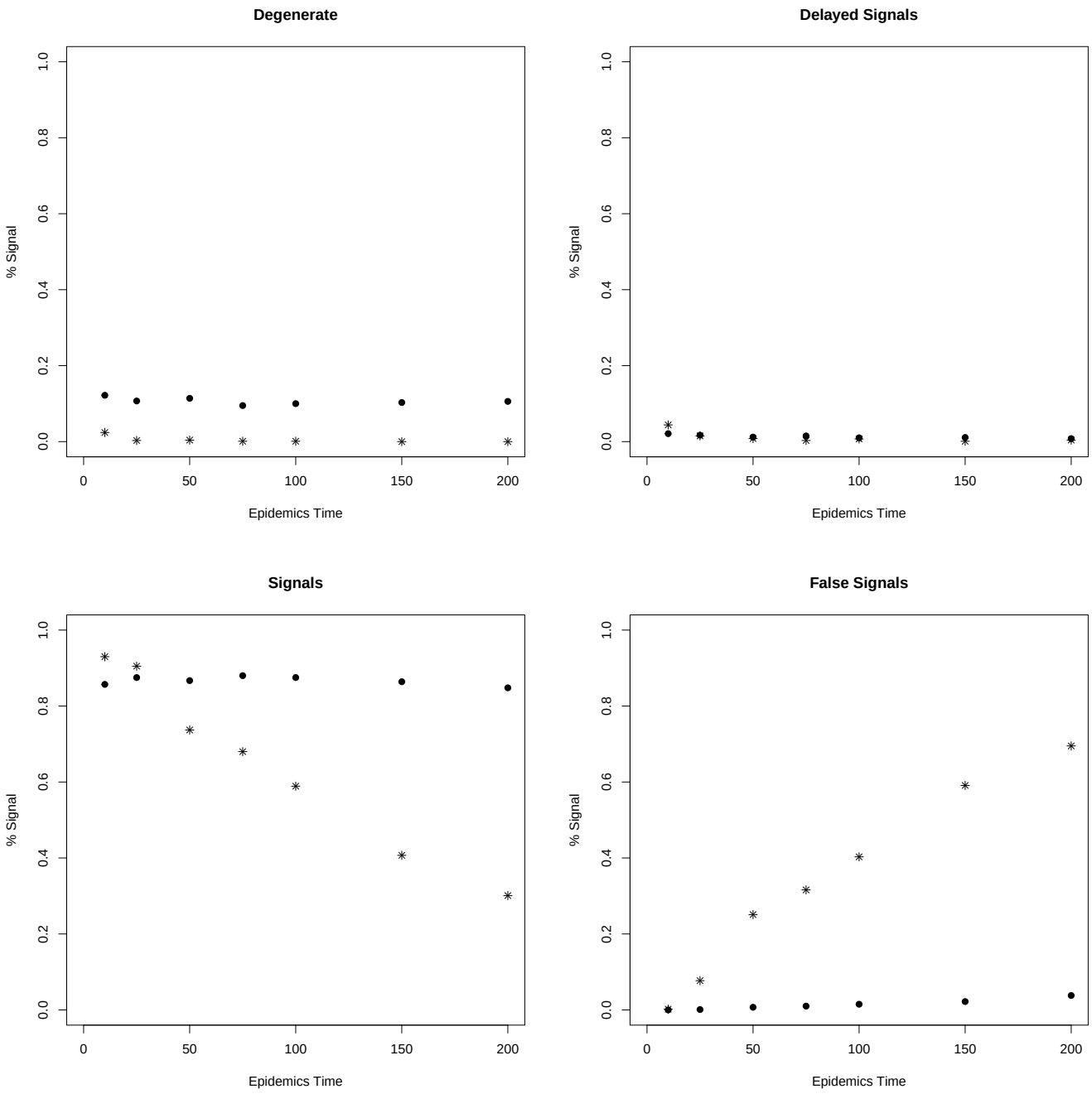

Figure 3. Linear outbreak operating characteristics with self-starting cusum ($\star$) and Shiryaev ($\bullet$)

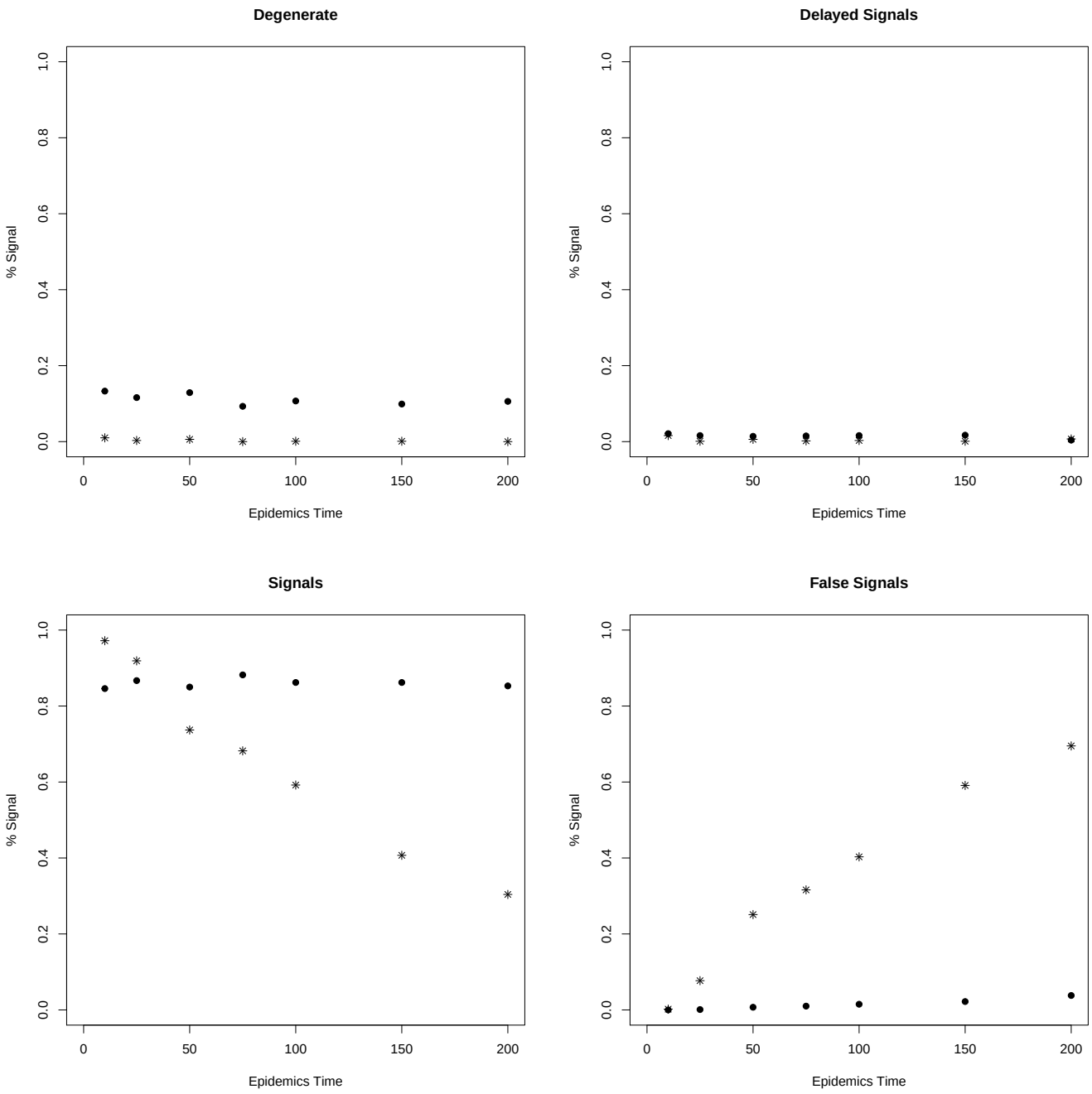

Figure 4. Exponential outbreak operating characteristics with self-starting cusum ($\star$) and Shiryaev ($\bullet$)

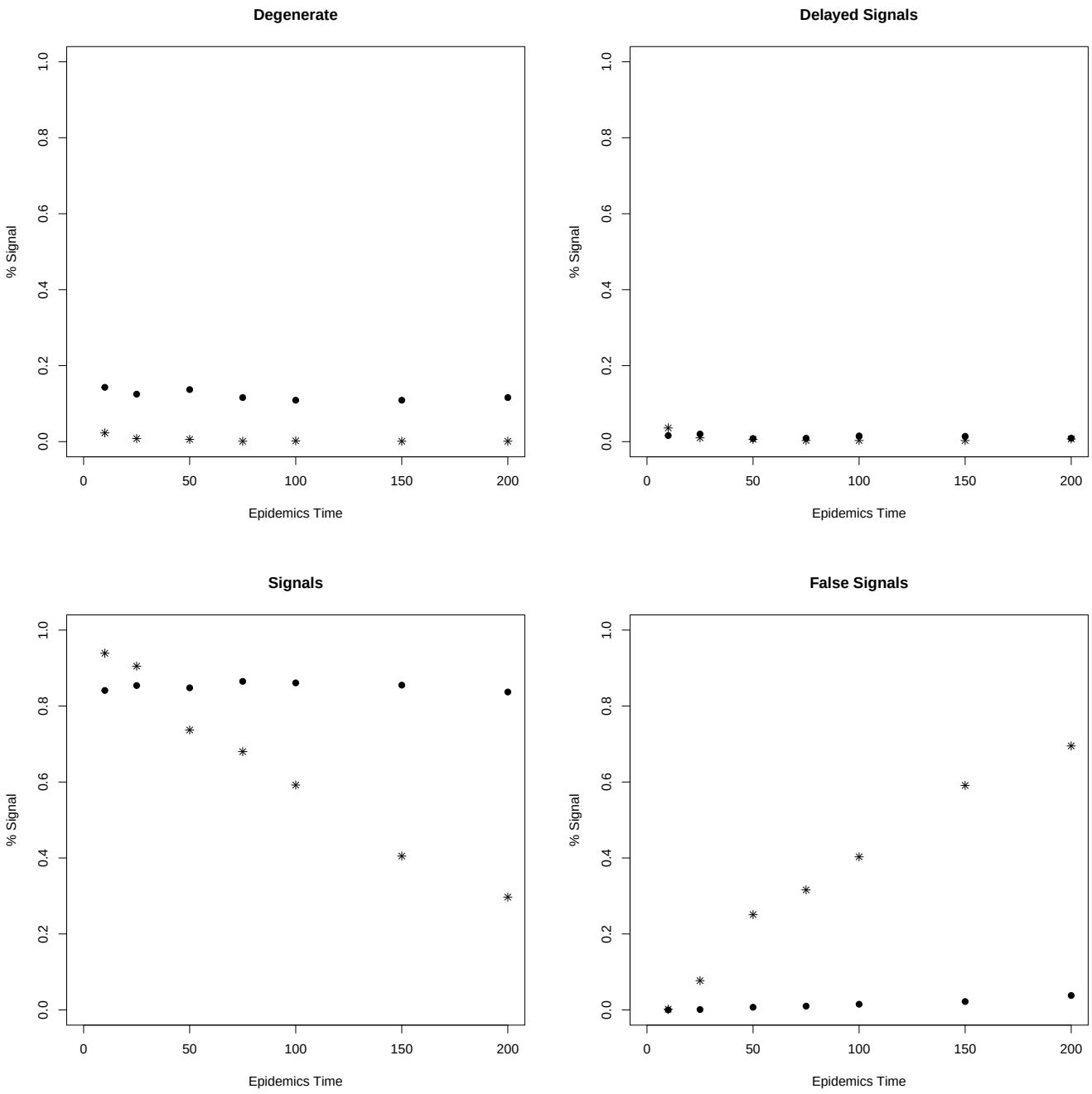

Figure 5. Sigmoid outbreak operating characteristics with self-starting cusum ($\star$) and Shiryaev ($\bullet$)

6. TECHNICAL APPENDIX

6.1. Theoretical Formulas

The likelihood is given by:

$$Y_i|\sigma^2,\delta,\tau \sim \left\{ \begin{array}{ll} f_o \equiv N\left(0, \ \ 2\sigma^2\right) & \text{if } 1 \leq i < \tau \\ f_1 \equiv N\left(\delta, \ \ 2\sigma^2\right) & \text{if } \tau \leq i \leq n \end{array} \right\}$$

while for the prior setting we have:

$$\begin{array}{rcl} \delta|\sigma^2 & \sim & N(\delta_0, \ \ k\sigma^2) \\ \sigma^2 & \sim & Inv.Gamma(\alpha, \ \ \beta) \ . \end{array}$$

We will consider the two components of the likelihood separately.

Case I: $1 \leq i < \tau$

$$\begin{array}{rcl} f_0(y_i|\tau) & = & \displaystyle\int_0^{+\infty} f_0(y_i,\sigma^2|\tau)d\sigma^2 = \int_0^{+\infty} f_0(y_i|\sigma^2,\tau)\pi(\sigma^2|\tau)d\sigma^2 \\ & = & \displaystyle\int_0^{+\infty} \frac{1}{\sqrt{2\pi 2\sigma^2}} \exp\left\{-\frac{y_i^2}{2(2\sigma^2)}\right\} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\sigma^2\right)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\} d\sigma^2 \\ & = & \displaystyle\frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\sqrt{4\pi}} \int_0^{+\infty} \left(\sigma^2\right)^{-\left(\alpha+\frac{1}{2}\right)-1} \exp\left\{-\frac{1}{\sigma^2}\left[\beta + \frac{y_i^2}{4}\right]\right\} d\sigma^2 \\ & = & \displaystyle\frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\sqrt{4\pi}} \frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\left[\beta+\frac{y_i^2}{4}\right]^{\alpha+\frac{1}{2}}} = \frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha)} \frac{1}{\sqrt{4\pi\beta}} \frac{1}{\left[\beta+\frac{y_i^2}{4}\right]^{\alpha+\frac{1}{2}}} \ . \end{array}$$

Case II: $\tau \leq i \leq n$

$$\begin{array}{rcl} f_1(y_i|\tau) & = & \displaystyle\int_0^{+\infty} \left[\int_{-\infty}^{+\infty} f_1(y_i,\delta,\sigma^2|\tau)d\delta\right] d\sigma^2 \\ & = & \displaystyle\int_0^{+\infty} \left[\int_{-\infty}^{+\infty} f_1(y_i|\delta,\sigma^2,\tau)\pi(\delta,\sigma^2|\tau)d\delta\right] d\sigma^2 \\ & = & \displaystyle\int_0^{+\infty} \left[\int_{-\infty}^{+\infty} f_1(y_i|\delta,\sigma^2,\tau)\pi(\delta|\sigma^2,\tau)d\delta\right] \pi(\sigma^2|\tau)d\sigma^2 \\ & = & \displaystyle\int_0^{+\infty} [I_1]\, \pi(\sigma^2)d\sigma^2 = \textbf{(*)} \ ; \end{array}$$

but:

$$
\begin{aligned}
I_1 &= \int_{-\infty}^{+\infty} f_1(y_i|\delta,\sigma^2)\pi(\delta|\sigma^2,\tau)d\delta \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi 2\sigma^2}} \exp\left\{-\frac{(y_i-\delta)^2}{2(2\sigma^2)}\right\} \frac{1}{\sqrt{2\pi k\sigma^2}} \exp\left\{-\frac{(\delta-\delta_0)^2}{2(k\sigma^2)}\right\} d\delta \\
&= \frac{1}{\sqrt{8\pi^2 k\sigma^4}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{4k\sigma^2}\left[ky_i^2 - 2ky_i^2\delta + k\delta^2 + 2\delta^2 - 4\delta_0\delta + 2\delta_0^2\right]\right\} d\delta \\
&= \frac{1}{\sqrt{8\pi^2 k\sigma^4}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{4k\sigma^2}\left[(k+2)\delta^2 - 2(ky_i+2\delta_0)\delta\right]\right\} d\delta \times \exp\left\{-\frac{ky_i^2+2\delta_0^2}{4k\sigma^2}\right\} \\
&= \frac{1}{\sqrt{8\pi^2 k\sigma^4}} \times \exp\left\{-\frac{ky_i^2+2\delta_0^2}{4k\sigma^2}\right\} \times \exp\left\{\frac{\frac{(ky_i+2\delta_0)^2}{(k+2)^2}}{\frac{4k\sigma^2}{k+2}}\right\} \times \\
&\quad \times \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\left(\frac{2k}{k+2}\right)\sigma^2}\left[\delta^2 - 2\left(\frac{ky_i+2\delta_0}{k+2}\right)\delta + \left(\frac{ky_i+2\delta_0}{k+2}\right)^2\right]\right\} d\delta \\
&= \frac{1}{\sqrt{8\pi^2 k\sigma^4}} \times \exp\left\{-\frac{ky_i^2+2\delta_0^2}{4k\sigma^2}\right\} \times \exp\left\{\frac{\frac{(ky_i+2\delta_0)^2}{(k+2)^2}}{\frac{4k\sigma^2}{k+2}}\right\} \times \sqrt{\frac{4k\pi}{k+2}\sigma^2} \\
&= \frac{\sqrt{\frac{4k\pi}{k+2}\sigma^2}}{\sqrt{8\pi^2 k\sigma^4}} \times \exp\left\{-\frac{1}{2[2k(k+2)]\sigma^2}\left[k(k+2)y_i^2 + 2(k+2)\delta_0^2 - (ky_i+2\delta_0)^2\right]\right\} \\
&= \frac{1}{\sqrt{2(k+2)\pi\sigma^2}} \exp\left\{\frac{k^2y_i^2 + 2ky_i^2 + 2k\delta_0^2 + 4\delta_0^2 - k^2y_i^2 - 4ky_i\delta_0 - 4\delta_0^2}{2[2k(k+2)]\sigma^2}\right\} \\
&= \frac{1}{\sqrt{2(k+2)\pi\sigma^2}} \exp\left\{\frac{2k}{2(2k)(k+2)\sigma^2}\left(y_i^2 - 2y_i\delta_0 + \delta_0^2\right)\right\} \\
&= \frac{1}{\sqrt{2(k+2)\pi\sigma^2}} \exp\left\{\frac{(y_i-\delta_0)^2}{2(k+2)\sigma^2}\right\} .
\end{aligned}
$$

Then:

$$
\begin{aligned}
\textbf{(*)} &= \int_0^{+\infty} \frac{1}{\sqrt{2(k+2)\pi\sigma^2}} \exp\left\{\frac{(y_i-\delta_0)^2}{2(k+2)\sigma^2}\right\} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\sigma^2\right)^{-\alpha-1} \exp\left\{-\frac{\beta}{\sigma^2}\right\} d\sigma^2 \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)\sqrt{2(k+2)\pi}} \int_0^{+\infty} \left(\sigma^2\right)^{-\alpha-1-\frac{1}{2}} \exp\left\{-\frac{1}{\sigma^2}\left[\beta + \frac{(y_i-\delta_0)^2}{2(k+2)}\right]\right\} d\sigma^2 \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\sqrt{2(k+2)\pi}} \frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\left[\beta + \frac{(y_i-\delta_0)^2}{2(k+2)}\right]^{\alpha+\frac{1}{2}}} \\
&= \frac{\Gamma\left(\alpha+\frac{1}{2}\right)}{\Gamma(\alpha)} \frac{1}{\sqrt{2(k+2)\pi\beta}} \frac{1}{\left[1 + \frac{(y_i-\delta_0)^2}{2(k+2)\beta}\right]^{\alpha+\frac{1}{2}}} .
\end{aligned}
$$

Thus:

$$\begin{aligned} Y_i|\tau \sim f_o(y_i) &= \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}\frac{1}{\sqrt{4\pi\beta}}\left[1+\frac{y_i^2}{4\beta}\right]^{-\alpha-\frac{1}{2}} \quad \text{if } 1<i<\tau \\ Y_i|\tau \sim f_1(y_i) &= \frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}\frac{1}{\sqrt{2(k+2)\pi\beta}}\left[1+\frac{(y_i-\delta_0)^2}{2(k+2)\beta}\right]^{-\alpha-\frac{1}{2}} \quad \text{if } \tau\le i\le n; \end{aligned}$$

and therefore we have:

$$\frac{f_1(y_i)}{f_0(y_i)} = \frac{\frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}\frac{1}{\sqrt{2(k+2)\pi\beta}}\left[1+\frac{(y_i-\delta_0)^2}{2(k+2)\beta}\right]^{-\alpha-\frac{1}{2}}}{\frac{\Gamma(\alpha+\frac{1}{2})}{\Gamma(\alpha)}\frac{1}{\sqrt{4\pi\beta}}\left[1+\frac{y_i^2}{4\beta}\right]^{-\alpha-\frac{1}{2}}} = \sqrt{\frac{2}{k+2}}\left[\frac{1+\frac{y_i^2}{4\beta}}{1+\frac{(y_i-\delta_0)^2}{2(k+2)\beta}}\right]^{\alpha+\frac{1}{2}}.$$

6.2. Useful Functions for Recursive Implementation

In the following functions the argument `y` is the data, `d` is the shift δ_0, `k` is the multiplier of σ^2 for the variance term in the prior distribution of the jump parameter δ_0, `a` is the hyperparameter α of the Inverse Gamma prior distribution on σ^2, and `b` is the hyperparameter β. For Shiryaev, `p` is the hyperparameter for the assumed geometric prior distribution.

6.2.1. The Likelihood Ratio Function

```
LR <- function(y, d, k, a, b)
 (sqrt(2/(k+2))*((1+y^2/(4*b))/(1+(y-d)^2/(2*(k+2)*b)))^(a+1/2))
```

6.2.2. The Shiryaev Function

```
S<-function(y,p,d,k,a,b)
{
          n<- length(y)
       PLRn<- 1
         Sn<- 0
for (j in n:1){
            PLRn<- LR(y[j],d,k,a,b)*PLRn
              Sn<- ((1-p)^j)*PLRn + Sn
               }
Sn<- p/((1-p)^(n+1))*Sn
Sn<- Sn/(Sn+1)
Sn
}
```

6.2.3. The Shiryaev-Roberts Function

```
SR<-function(y,d,k,a,b)
{
```

```
            n<- length(y)
          PLRn<- 1
           SRn<- 0
for (j in 1:n){
            SRn<- (1+SRn)*LR(y[j],d,k,a,b)
              }
SRn
}
```

REFERENCES

Apley, D. W. (2012). Posterior Distribution Charts: A Bayesian Approach for Graphically Exploring a Process Mean, *Technometrics* 54: 296–310.

Arnon, S. S., Schechter, R., and Inglesby, T. V. (2001). Tulinum toxin as a biological weapon: medical and public health management, *Journal of American Medical Association* 285: 1059–1070.

Buehler, J. W., Berkelman, R. L., Hartley, D. M., and Peters, C. J. (2003). Syndromic surveillance and bioterrorism-related epidemics, *Emerging Infectious Diseases* 9: 1197–1204.

Chakraborti, S., Human, S. W., and Graham, M. A. (2008). Phase I statistical process control charts: An overview and some results, *Quality Engineering* 21: 52–62.

Dennis, D. T., Inglesby, T. V., and Henderson, D. A. (2001). Tularemia as a biological weapon: medical and public health management, *Journal of American Medical Association* 285: 2763–2773.

Enders, W. (2010). *Applied Econometric Time Series*, Wiley: New York.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B. (2014). *Bayesian Data Analysis (Third ed.)*, Taylor and Francis: Boca Raton.

Green, M. S., and Kaufman, Z. (2002). Surveillance for early detection and monitoring of infection disease outbreaks associated with bioterrorism, *Israel Medical Association Journal* 4 (7): 503–506.

Hawkins, D. M., and D. H. Olwell (1997). *Cumulative Sum Charts and Charting for Quality Improvement*, Springer Verlag: New York.

Hawkins, D. M., Qiu, P., and Kang, C. W. (2003). The change point model for statistical process control, *Journal of Quality Technology* 35: 355–365.

Hawkins, D. M., and Zamba, K. D. (2005a). Change point model for a shift in variance, *Journal of Quality Technology* 37: 21–31.

Hawkins, D. M., and Zamba, K. D. (2005b). A change point model for statistical process control with shift in mean or variance, *Technometrics* 47: 164–173.

Henderson, D. A., Inglesby, T. V., and Bartlett, J. G. (1999). Smallpox as a biological weapon: medical and public health management, *Journal of American Medical Association* 281: 2127–2137.

Hoadley B. (1981). The quality measurement plan (QMP), *Bell System Technical Journal* 60: 215-274.

Hodrick, R., and Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation, *Journal of Money, Credit, and Banking* 29: 116.

Hutwagner, L. C., Maloney, E. K., Bean, N. H., Slutsker, L., and Martin, S. M. (1997). Using laboratory-based surveillance data for prevention: An algorithm to detect salmonella outbreaks, *Emerging Infectious disease* 3: 395–400.

Inglesby, T. V., Dennis, D. T., and Henderson, D. A. (2000). Plague as a biological weapon: medical and public health management, *Journal of American Medical Association* 283: 2281–2290.

Inglesby, T. V., Henderson, D. A., and Bartlett, J. G. (1999). Anthrax as a biological weapon: medical and public health management, *Journal of American Medical Association* 281: 1735.

Jones, L. A., and Champ C. W. (2002). Phase I control charts for times between events, *Quality and Reliability Engineering International* 18: 479–488.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering* 82: 3545.

King, E. P. (1954). Probability limit for the average chart when process standards are unspecified, *Industrial Quality Control* 10: 62–64.

Lorden, G., and Pollak, M. (2005). Nonanticipating estimation applied to sequential analysis and changepoint detection, *Annals of Statistics* 33: 1422–1454.

Mills, T. C. (2003). *Modelling Trends and Cycles in Economic Time Series*, Palgrave Macmillan: New York.

Pavlin, J. A. (2003). Investigation of disease outbreaks detected by syndromic surveillance systems, *Journal of Urban Health* 80: 107–114.

Pollack, M. (1987). Average run lengths of an optimal method of detecting a change in distribution, *Annals of Statistics* 15: 749–779.

Ravn, M. O., and Uhlig H. (2002). On Adjusting the Hodrick-Prescott Filter for Frequency of Observations. *Review of Economics and Statistics*, 84: 371380.

Roberts, S. W. (1966). A comparison of some control chart procedures, *Technometrics* 8: 411–430.

Shiryaev, A. N. (1963). On optimum methods in quickest detection problems, *Theory and Probability Applications* 8: 22–46.

Shmueli, G., and Burkom, H. S. (2010). Statistical challenges facing early outbreak detection in biosurveillance, *Technometrics* 52: 39–51.

Shumway, R. H., and Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*, Springer: New York.

Sidák, Z. K. (1967). Rectangular confidence regions for the means of multivariate normal distributions, *Journal of American Statistical Association* 62: 626–633.

Stern, L., and D. Lightfoot (1999). Automated outbreak detection: A quantitative retrospective analysis, *Epidemiology of Infectious Disease* 122: 103–110.

Stoto, M., Schonlau, M., and Mariano, L. T. (2004). Syndromic surveillance: Is it worth the effort? *Chance* 17: 19–24.

Tsiamyrtzis P., and Hawkins D. M. (2005). A Bayesian scheme to detect changes in the mean of a short run process, *Technometrics* 47: 446-456.

Tsiamyrtzis P., and Hawkins D. M. (2010). Bayesian Start up Phase Mean Monitoring of an Autocorrelated Process that is Subject to Random Sized Jumps, *Technometrics* 52: 438-452.

Tsiamyrtzis P., and Hawkins D. M. (2019) Statistical process control for Phase I count type data, *Applied Stochastic Models in Business and Industry* 35: 766-787.

Winkel, P., and Zhang, N. F. (2007). *Statistical Development of Quality in Medicine*, John Wiley & Sons Ltd: England.

Woodall, W. H. (2006). The use of control charts in health-care and public-health surveillance, *Journal of Quality Technology* 38: 89–104.

Woodward, P. W., and Naylor, J. C. (1993). An Application of Bayesian Methods in SPC, *Statistician* 42: 461-469.

Yakir, B. (1995). A note on the run length to false alarm of a change-point detection policy, *Annals of Statistics* 23: 272–281.

Zamba, K. D., Tsiamyrtzis, P., and Hawkins, D. M. (2008). A sequential bayesian control model for influenza-like illnesses and early detection of intentional outbreaks, *Quality Engineering (Health Edition)* 20: 495–507.

Zamba, K. D., P. Tsiamyrtzis, and D. M. Hawkins (2013). A three-state Recursive Sequential Bayesian Algorithm for Biosurveillance, *Computational Statistics and Data Analysis* 58: 82–97.

Zhang, N. F. (1997). Detection capability of residual control chart for stationary process data, *Journal of Applied Statistics* 24: 475–492.