We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Exponentially Weighted Moving Averages of Counting Processes When the Time between Events Is Weibull Distributed

Ross Stewart Sparks and Hossein Hazrati-Marangaloo

Abstract

There are control charts for Poisson counts, zero-inflated Poisson counts, and over dispersed Poisson counts (negative binomial counts) but nothing on counting processes when the time between events (TBEs) is Weibull distributed. In our experience the in-control distribution for time between events is often Weibull distributed in applications. Counting processes are not Poisson distributed or negative binomial distributed when the time between events is Weibull distributed. This is a gap in the literature meaning that there is no help for practitioners when this is the case. This book chapter is designed to close this gap and provide an approach that could be helpful to those applying control charts in such cases.

Keywords: average run length, counts, monitoring, time between events

1. Introduction

Statistical process control and monitoring (SPCM) methods originally arose in the context of industrial/manufacturing applications, developed during and after World War II. Since then, it has become a popular way of monitoring all processes. Today, large volumes of data are often available from a variety of sources, in a variety of environments that need to be monitored. This means one needs to make sense of these data and then be able to make efficient monitoring decisions based on them. While constructing and applying monitoring tools, a fundamental assumption, necessary to justify the end results, involves the assumption about the distribution that the data have been generated from. This is the heart of outbreak detection of events particularly for estimating the in-control false discovery rates.

Selecting an appropriate probability distribution for the data is one of the most important and challenging aspects of data analysis. The estimates of outbreak false discovery rates often hinge on this crucial selection. We focus on the distribution for the time between events (TBEs) because there are often many of these events as compared to counts. Therefore, it is easier to fit an appropriate distribution using these many TBE values. The most commonly assumed distribution in the application of TBE is the Weibull distribution which is asymmetric and sometimes severely skewed. However, depending on the context, other distributions may also be used, such as the exponential distribution (which is a special case of a Weibull

distribution) or the gamma distribution. If the TBE distribution is exponentially distributed, then the related counts are Poisson distributed. This book chapter focusses on counting processes when the distribution of TBE values is known to be Weibull distributed.

The challenge of making and meeting the distributional assumption is faced by all practitioners and data analysts. In many monitoring settings, event data are collected in a nearly continuous stream, and it is often more meaningful to monitor the individual TBE data [1–4] when outbreaks are of large magnitude. This individual event data are aggregated over fixed time intervals (e.g., daily) to form counts. In this chapter, these counts are monitored to detect outbreaks resulting from small changes in the incident of events. The focus is on the steady-state situation because this is the most common situations in event monitoring. Note that we cannot stop the process and investigate the out-of-control situation because often in nonmanufacturing settings it is not under our control. Events may include warranty claims of a product, health presentation at emergency departments, sales of an online products, etc. Here the term "quality of the process" is used in a general sense, which is context dependent. In the case of sales, an outbreak would represent an increased sales opportunity, provided the inventory stock can support this outbreak and the products are not sold out before the next order arrives. However, for warrantee claims, this would represent an undesirable outbreak of increased claims which may require a failure mode effect analysis [5]. Monitoring of in-control nonhomogeneous counting processes has traditionally been carried out using either Poisson or the negative binomial distributions for the counts. Many statistical tools used in SPCM for counts data are documented in Sparks et al. [1–3], Sparks et al. [6], Weiß [7, 8], Weiß and Testik [9, 10], Yontay et al. [11], and Albarracin et al. [12]. These control charts are perhaps the most well-known count monitoring methodologies. The control chart graphic is a time series plot of a signal-to-noise ratio designed for the user to make decisions about the outbreak of events.

In any case, before defining an "in-control" process, we need the information about the probability distribution of the events being monitored. When this information is available, it is possible to calculate the probabilities (or the chance) from the event in-control distribution, which could be defined by the TBEs or the counting of events within a fixed time interval. Deciding on whether event distribution constitutes an event outbreak is based on whether the event distribution is extreme compared to what is usual, i.e., counts are higher than expected. This is usually gauged by some upper threshold for the signal-to-noise ratios.

In a vast majority of SPCM applications, it is common to assume that the underlying probability distribution is of a (given) known form. In this chapter we assume that the TBEs are distributed as a Weibull distribution. However, we explore approaches to monitoring the counts of these events over fixed periods of varying length to find the period width that leads to earlier detection of outbreaks in terms of the average time to signal (ATS). The distribution of these counts when the TBEs are Weibull distributed is neither Poisson nor negative binomial distributed. Therefore, this chapter offers a different approach compared to others in the literature. In addition, the appropriate period of aggregation for the counts is explored in terms of how it influences early detection of outbreak events.

In practice event data are often collected in a continuous stream defined by the TBEs. Besides, wherever outbreaks are of large magnitude, it is more meaningful to monitor these individual TBE values [4]. In such situations, one also needs to deal with the issue of autocorrelation which may be thought of as the effect of time, as data values in close proximity with respect to time and space are likely dependent. This violates one of the basic assumptions in process monitoring, and a common way to deal with this issue is by fitting a time series model and monitoring the

standardized residuals. The assumption of a distribution is also an important part of this analysis. In this chapter, our focus is on the Weibull distributional assumption for the TBE values. However, rather than monitoring the TBEs, we monitor the counts over a fixed time interval, because this improves the early detection of smaller outbreaks (Sparks et al., [4]).

2. Monitoring homogeneous counts

Considering a fixed distribution for TBEs during the monitoring period, we assume that the time of the day and date stamp of all events are available. For a series comprised of n events, let the day numbers for events be denoted as

$$d_1, d_2, \ldots, d_n$$

with the event times within days which are defined as

$$\tau_1, \tau_2, \ldots, \tau_n.$$

Note that these times are measured in fractions of a day, e.g., $0 \leq \tau_i < 1$. Daily counts are given by counting the number d_i values that are the same and then adding the days with zero counts for those days that are missing from the d_i values. Then TBEs are given by

$$w_1 = d_2 - d_1 + \tau_2 - \tau_1, \ldots, w_{n-1} = d_n - d_{n-1} + \tau_n - \tau_{n-1}$$

where w_i represents the time between $(i+1)$th and ith events. We flag an outbreak wherever w_i for any $i = 1, \ldots, n-1$ are consistently lower than an expected value. These TBEs are assumed to be Weibull distributed with fixed scale and shape parameters. Using R code, we define the counting process, for say daily counts, as

$$x < - rweibull(n = number\ of\ TBE, shape = shape, scale = scale)$$

$$counts < - table(floor(cumsum(x)))$$

$$counts < - counts[-length(counts)]$$

Denote these counts as c_i for the ith day. We define the exponentially weighted moving average (EWMA) statistic for these daily homogeneous counts as follows:

$$e_i = \max(0, \alpha c_i + (1 - \alpha)e_{i-1}) \tag{1}$$

where $e_0 = E(c_i)$ is assumed constant when in-control for days $i = 1, 2, \ldots$. The smoothing parameter, $\alpha(0 < \alpha < 1)$, determines the level of memory of past observations included in e_i. Smaller values of α retain the more memory of past counts than do the larger values. Therefore, smaller values of α are more efficient at detecting smaller changes in mean counts, while larger values of α are more efficient at detecting larger changes. EWMA statistic, e_i, has a minimum value of zero, and we do not allow it to fall below zero because we are only interested in outbreaks, i.e., counts that are higher than expected. This EWMA statistic controls the worst-case scenarios, whereas the traditional EWMA statistic can wonder well below zero before going out of control causing the EWMA statistic to take a long time to signal the outbreak. An outbreak is flagged when e_i exceeds a predetermined threshold, $h_c(ARL_0, scale, shape, \alpha)$. For a given pair of shape and scale parameters,

h_c is determined so that a desired in-control average run length (ARL_0) is achieved. Appendix A provides models for establishing the thresholds for various values for $0.2 \leq scale \leq 0.46$, $0.6 \leq shape \leq 1.4$, $\alpha = 0.1$ and $ARL_0 = 100, 200, 300$ or 400. Given a h_c and the process is out of control, we want the ARL to be as low as possible. Monitoring daily counts, the *ARL* is defined as the number of days before a signaled outbreak. Later we will use the ATS to assess the relative performance of control chart plans, because the ARL may vary according to the aggregation period we select for the counts. Note that a false outbreak is flagged by this EWMA statistic being significantly larger than expected given it is in-control.

3. Monitoring in-control nonhomogeneous counts

For most of the real-world cases, TBE distribution may vary due to different circumstances while not experiencing an outbreak. For instance, if we aggregate the TBEs in daily intervals, occurring intervals for the events may vary based on the day of the week, and even working and non-working days may affect the distribution of TBEs. As a result, we often face nonhomogeneous count processes. Hence, we define an adaptive exponentially weighted move average (AEWMA) statistic for nonhomogeneous daily counts as

$$ae_i = \max\left(0, \alpha c_i / h_c(ARL_0, scale, shape, \alpha) + (1-\alpha)ae_{i-1}\right) \quad (2)$$

where ae_i is the AEWMA statistic at time t, and $ae_0 = E(c_i)/h_c(ARL_0, scale, shape, \alpha)$ for days $i = 1, 2, \ldots$. The other notations are as defined earlier. To control the false discovery rate, we set $h_c(ARL_0 = c, scale, shape, \alpha)$ so that a desiredARL_0 is achieved. An outbreak is flagged approximately whenever $ae_i > 1$.

The $\alpha(0 < \alpha < 1)$ in Eq. (1) determines the level of memory of past observations in this average ae_i. Smaller α values retain more memory of past counts; therefore, small α values are efficient at retaining enough memory of past counts to have the power to flag smaller outbreaks. However, larger values of α are needed when there is a larger size outbreak, because a shorter range of memory is adequate to build enough power to detect the outbreak. In addition, the shorter memory for the EWMA average with larger values of α is less inclined to be influenced by too many past in-control counts.

4. Simulation results

To assess the validity and applicability of our proposed adaptive method, we employ simulations studies. We restrict our attention to plans that have $10 < E(c_i) < 50$ which limits the volume of simulations that are required to make a realistic judgment. This is also the range where the EWMA counts become competitive for small changes in the scale parameter (see [4] and Sparks et al., 2020). To do so, we consider four counting processes with daily aggregations. The shape parameters for the TBE distributions when the processes are in-control are equal to 0.85, 0.95, 1.15, and 1.25. The in-control scale parameter for each process is equal to 0.02, 0.025, 0.03, and 0.035, respectively. We intentionally imposed outbreaks in the simulated data to assess if the proposed method is capable of detecting them. In the simulation study, we assume that the outbreaks result in a decrease in the scale parameter. Various surveillance plans are devised to monitor each process. The ATS_0 for the plans associated with the first process (with shape and scale

parameters of 0.85 and 0.02, respectively) is set to be equal to 400. The ATS_0 for the plans associated with other processes are equal to 300, 200, and 100, respectively. 100,000 events are employed to estimate the count mean for each process. In addition, 500 events are also considered for the burn-in period of the simulations. The performance of the devised plans is presented in Tables 1–4. The lowest ATS values are colored in the tables below in black bold text to make it easier to see which plans are more efficient in certain situations.

Table 1 shows the performance results for the plans employed to monitor counting process where the in-control data are Weibull with scale and shape parameters of 0.035 and 1.25, respectively. ATS_0 (in days) for these plans are set approximately equal to 100. As shown in Table 1, for the outbreaks of larger magnitude, plans with larger smoothing parameter are superior. On the contrary,

	EWMA counts (α), shape = 1.25							
Threshold	31.835	32.162	32.524	32.807	33.207	33.470	33.803	34.108
Scale					α			
	0.04	0.06	0.08	0.10	0.125	0.15	0.175	0.20
0.035	100.05	101.16	100.41	101.95	100.31	100.02	100.55	101.21
0.034	41.952	40.255	**39.997**	41.552	40.954	38.936	40.372	43.733
0.033	19.116	18.757	**18.275**	19.098	19.564	19.106	19.538	20.302
0.032	13.054	11.924	11.721	**11.259**	11.359	10.733	11.370	11.843
0.031	8.598	8.231	8.156	7.860	7.847	**7.491**	7.515	7.566
0.030	6.774	5.937	5.943	5.749	5.628	**5.393**	5.508	5.437
0.029	5.824	4.715	4.724	4.696	4.426	4.211	**4.161**	4.183
0.028	4.408	4.149	3.912	3.739	3.660	3.424	3.446	**3.283**
0.027	3.699	3.358	3.386	3.025	3.002	2.886	2.832	**2.788**

Table 1.
Performance of plans when the in-control TBEs are Weibull distributed with scale = 0.035 and shape = 1.25.

	EWMA counts (α), shape = 1.15							
Threshold	36.20	37.058	37.4917	37.904	38.315	38.725	39.1025	39.515
Scale					α			
	0.04	0.06	0.08	0.10	0.125	0.15	0.175	0.20
0.03	205.09	200.54	201.68	201.21	200.23	200.82	200.96	200.51
0.029	50.792	**47.665**	54.813	53.764	52.916	54.227	56.616	61.987
0.028	22.027	21.034	20.747	**21.067**	21.624	22.292	23.361	23.265
0.027	12.791	11.790	11.953	11.655	**11.265**	11.344	11.397	12.319
0.026	8.678	8.414	7.811	7.414	7.211	**7.230**	7.333	7.497
0.025	6.718	6.163	5.718	5.508	5.444	5.243	**5.159**	5.162
0.024	5.078	4.650	4.462	4.283	4.228	3.921	3.920	**3.877**
0.023	4.482	4.010	3.742	3.558	3.465	3.190	3.155	**3.072**

Table 2.
Performance of plans when the in-control TBEs are Weibull distributed with scale = 0.03 and shape = 1.15.

EWMA counts (α), shape = 0.95								
Threshold	**41.35**	**41.9028**	**42.4899**	**42.9932**	**43.5669**	**44.1448**	**44.7079**	**45.1527**
Scale					α			
	0.04	**0.06**	**0.08**	**0.10**	**0.125**	**0.15**	**0.175**	**0.20**
0.025	302.19	301.69	302.03	301.51	302.29	301.85	301.29	301.83
0.024	**51.937**	56.110	59.550	65.442	63.928	66.311	78.925	74.448
0.023	23.359	**20.844**	22.686	22.031	22.855	23.860	25.744	26.173
0.022	12.681	11.357	11.807	**11.345**	11.412	11.493	12.104	12.056
0.021	8.784	8.253	7.723	7.394	**7.150**	7.326	7.196	7.231
0.0205	7.474	6.866	6.855	6.410	5.997	**5.962**	6.113	5.977
0.020	6.788	5.962	5.416	5.368	5.075	5.038	5.136	**4.976**
0.019	4.840	4.707	4.415	4.032	3.942	3.844	3.805	**3.757**

Table 3.
Performance of plans when the in-control TBEs are Weibull distributed with scale = 0.025 and shape = 0.95.

EWMA counts (α), shape = 0.85								
Threshold	**48.7256**	**49.5976**	**50.3141**	**50.9389**	**51.6532**	**52.35**	**53.0104**	**53.5827**
Scale					α			
	0.04	**0.06**	**0.08**	**0.10**	**0.125**	**0.15**	**0.175**	**0.20**
0.02	401.92	401.28	399.92	400.89	399.89	401.09	401.08	399.93
0.0195	**113.23**	121.88	123.74	135.49	144.80	145.77	145.82	156.46
0.019	51.948	**48.144**	51.531	57.694	57.722	63.461	66.637	69.652
0.018	19.344	18.374	17.494	**17.316**	17.663	18.892	19.627	19.487
0.0175	12.846	12.711	12.370	12.263	**12.260**	12.272	12.651	12.765
0.017	10.966	10.696	9.583	9.032	9.052	**8.752**	9.179	9.074
0.0165	8.261	8.510	7.599	7.301	6.946	6.948	**6.868**	6.954
0.016	7.191	6.386	6.198	5.803	5.666	5.429	5.456	**5.277**
0.015	5.474	4.868	4.354	4.119	4.031	4.047	3.862	**3.589**

Table 4.
Performance of plans when the in-control TBE are Weibull distributed with scale = 0.02 and shape = 0.85.

for the early detection of the small outbreaks, smaller values for the smoothing parameter are preferred.

Table 2 shows the performance results of the plans when the shape and the scale parameters of the Weibull distribution are equal to 1.15 and 0.03, respectively. The ATS_0 for the plans employed to monitor this counting process is set to be approximately 200. In the results indicated in **Table 1**, plans with smaller smoothing parameter work better in early detection of small outbreaks than those with larger smoothing parameter. For the larger outbreaks, the detection power of the plans increases as the smoothing parameters increases.

Table 3 shows the performance results of the plans when the shape and the scale parameters of the Weibull distribution are equal to 0.95 and 0.025, respectively. The ATS_0 for the plans employed to monitor this counting process is set to be

approximately 300. Conclusions similar to those regarding **Tables 1** and **2** can be drawn from **Table 3**. It is clearly observed that the larger the magnitude of the outbreak, the larger the smoothing parameter should be. On the other hand, for the detection of small outbreaks, plans with larger values of ATS_0 need even smaller values for smoothing parameters than do the plans with smaller ATS_0. As shown in **Table 1**, smoothing parameters equal to or larger than 0.08 work better for outbreak detection. As the ATS_0 increases, analogous to results presented in **Tables 2** and **3**, even smaller values for α are needed to devise a plan of larger detection power. Last but not the least, similar results can be driven from **Table 4**, which presents the performance of the monitoring plans applied to a counting process with underlying Weibull distribution for TBEs. The shape and the scale parameters of the aforementioned distribution are equal to 0.85 and 0.028, respectively.

5. Real-world example

In this section, we apply our proposed method to a real-world example. The counting process to monitor is the number of presentations at Gold Coast University Hospital emergency department for a broad definition of influenza. The events
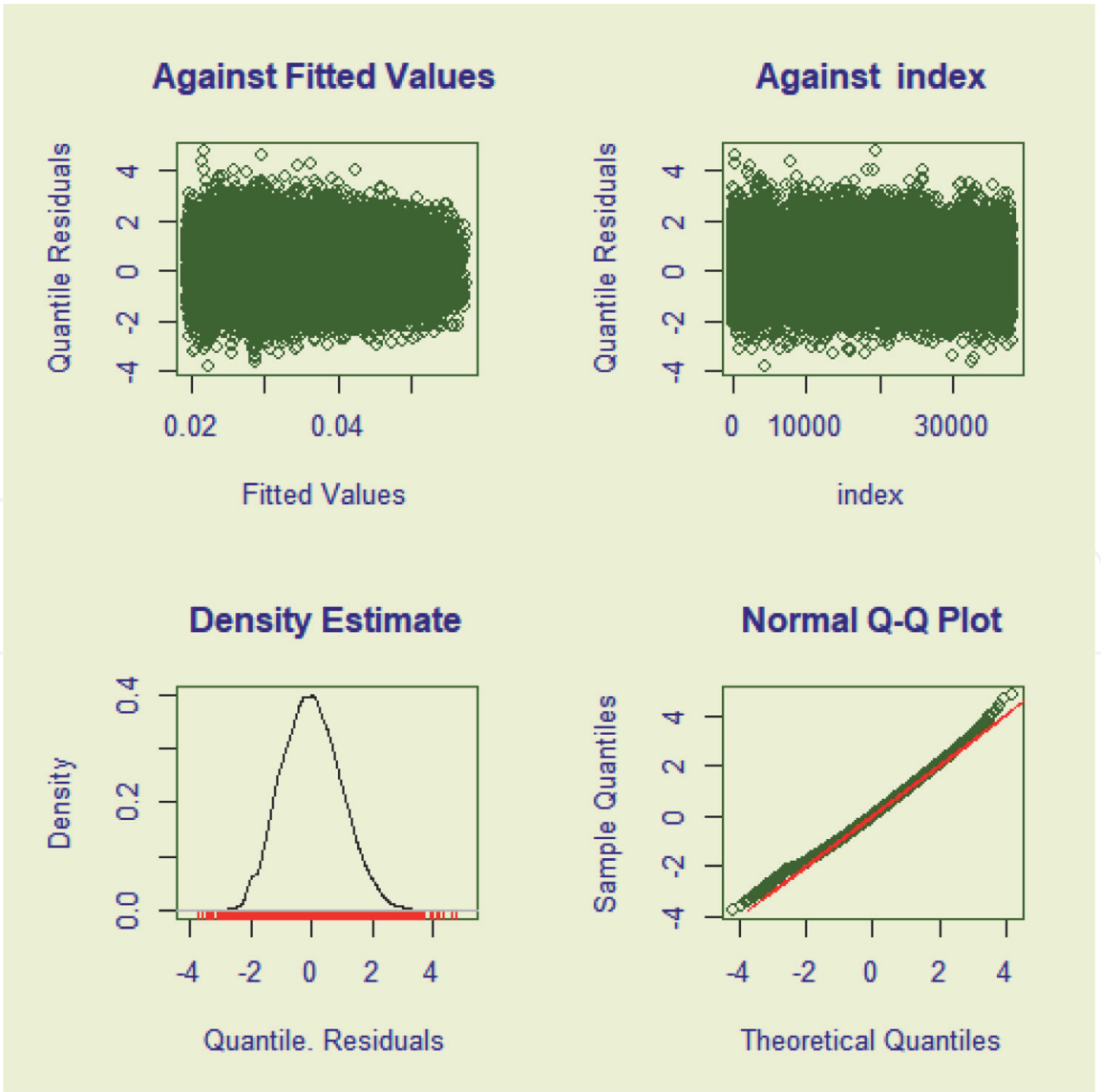

Figure 1.
Residual analysis for the fitted model.

are presentations at Gold Coast University Hospital with flu symptoms. Data is gathered for four consecutive years starting from January 2015. We first check if the TBEs are Weibull distributed. We fitted a Weibull regression model to data, and all the parameters of the conditional distribution of the response variable are modeled using explanatory variables as the hour of the day harmonics, seasonal harmonics, and day of the week. To do so, we employ "gamlss" [13] R package for statistical modeling. This package includes functions for fitting the generalized additive models for location, scale, and shape introduced by Rigby and Stasinopoulos [14]. The R procedure for the model fitting is as follows:

```
gamlss(formula = TBE ~ wd + wd=="Monday") + (wd=="Thursday"))*
(cos(2*pi*hr/24) + sin(2*pi*hr/24)) + cos(2*pi*nday/365.25)
+ sin(2*pi*nday/365.25)sigma.formula = ~(cos2*pi*hr/24) + sin
(2*pi*hr/24))+cos(2*pi*nday/365.25)+sin(2*pi*day/365.25)
family=WEI()data=data)
```

where wd, nday, and hr. are week day, number of the day in a year, and the time of day (0–24), respectively. **Figure 1** summarizes the analysis of the residuals for the fitted model. As shown in **Figure 1**, since residuals do not represent any particular pattern in data, and the model describes the response variables quite well, then we conclude that TBEs are Weibull distributed. Details for the analyzing the model adequacy is presented in Appendix B.

As mentioned earlier, the threshold, h_c, is determined according to parameters of the underlying Weibull distribution for the TBEs. For any timestamp during the monitoring period, the fitted model is used to predict the parameters of the Weibull distribution. Then these parameters along with the smoothing parameter, α, and the in-control ARL are substituted in models in Appendix A to establish the h_c for Weibull-distributed counts. The parameter for the day of the count is taken as the average of the estimated parameters for the day of events. Considering $\alpha = 0.1$ and
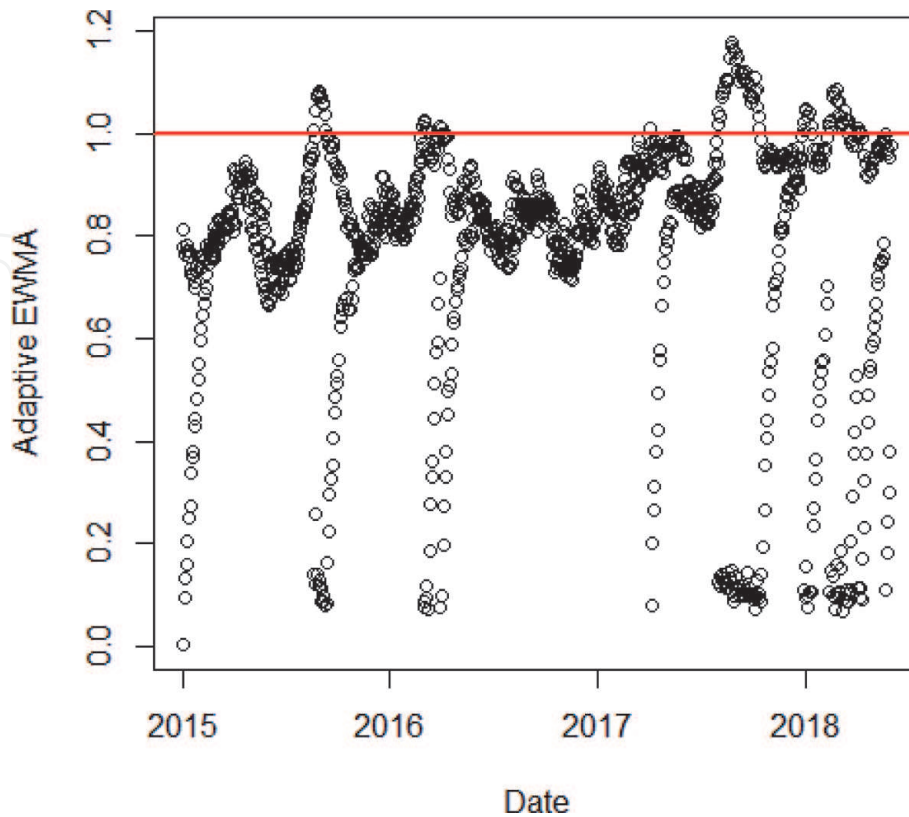

Figure 2.
AEWMA chart using $\alpha = 0.1$ for influenza presentations at Gold Coast University Hospital.
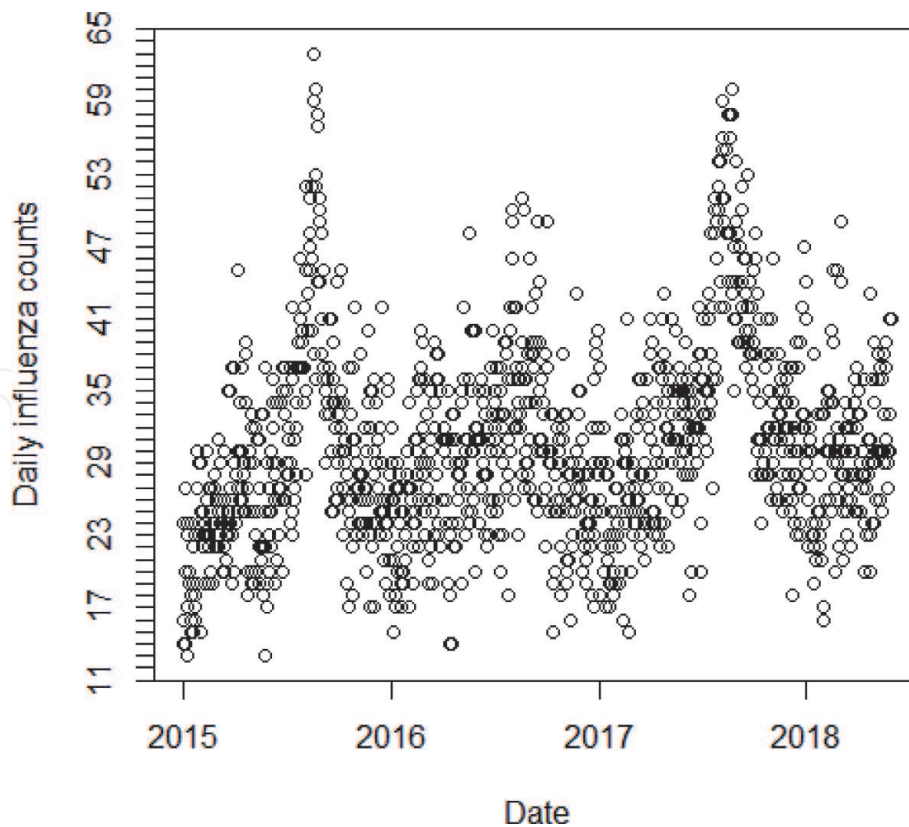
Figure 3.
Event counts vs. time.

ARL_0 for the plan to be 400, the estimated parameters are used to establish the AEWMA plan for monitoring the daily counts of flu presentations at the Gold Coast University Hospital. **Figure 2** shows the time series of the monitoring statistic over the study period, and **Figure 3** is the time series of the daily flu counts. The proposed plan indicates six potential flu outbreaks during 4 years as the monitoring statistic falls beyond 1. Note that influenza counts in 2015 were bigger than usual, but not unusual in 2016 (apart from early in that year). However, in 2017 this was flagged as a very unusual influenza outbreak and this seems to persist into early 2018.

6. Concluding remarks

In this chapter, we proposed an adaptive EWMA surveillance plan to monitor a counting process of which the time between its events is Weibull distributed. The proposed method can be applied to both homogeneous and nonhomogeneous processes. To implement the proposed surveillance plan, the scale and shape parameters for the underlying distribution of the TBEs are estimated using a distributional regression approach [15]. Then the threshold for the counts is established using the estimated parameters and the desired ARL. The proposed plan is applied to both simulated and real data. Simulation results indicate that the proposed method is applicable for detecting outbreaks of any magnitude and also signals them in a reasonable time after their incidence. In addition, simulations revealed that for the detection of the large outbreak, plans with larger smoothing parameter are superior. However, for the early detection of small outbreaks, we need to employ smaller smoothing weights. Applying the proposed surveillance method to real data, we conclude that the proposed method is capable of detecting outbreaks in nonhomogeneous counting processes.

A. Appendix

The thresholds for the linear model fitted

$$\text{sqrt(hc)} \sim (\text{scale} + \text{shape} + \log{(\text{scale})} + \log{(\text{shape})})^2 + \text{sqrt(shape)} * \text{scale}$$

The multiply R-squared for all these models is equal to 1 corrected to the sixth decimal place. The square of the fitted values for these models provides an starting estimate of the threshold for the respectively in-control ARL of 100, 200, 300, or 400. This can be further revised using simulation (**Table A1**).

In-control ARL	100	200	300	400
(Intercept)	55.27965623580790	14.50063556867970	35.4751774678155	1.62174314752660
scale	−626.65874427126700	−750.65355682164500	−853.0616160220110	−666.97214806267300
shape	54.99181051902600	6.91623162316955	26.9351743566771	−2.52567948128196
log(scale)	−8.64992403297571	−8.66215508213318	−8.6812026091707	−8.66460977364899
log(shape)	16.34706116414410	6.99770664515964	12.4345299969720	3.32335746029235
sqrt (shape)	−132.57740079984300	−43.82444799901730	−84.7132341475612	−21.35094544552350
scale:shape	−224.25112834314800	−263.89677049308800	−301.0663396807370	−236.64136324217100
scale:log (scale)	−69.61113437236530	−70.49847595504570	−69.8492890948244	−69.64522997112020
scale:log (shape)	−160.56446815370100	−201.15404356846600	−235.8054471494210	−174.62069894929300
shape:log (scale)	2.19016501325625	2.17689719954047	2.2105503473413	2.20073389465297
shape:log (shape)	−7.97657987867648	4.93365063336234	0.1464213924397	6.88699339675467
log(scale): log(shape)	−3.19688371718715	−3.19951671299892	−3.1780225086612	−3.20431683519426
scale:sqrt (shape)	795.53717813564600	957.71276755290300	1098.4086506150800	848.24798318601300

Table A1.
Fitted models for predicting the thresholds when the scale parameter is between 0.02 and 0.046 (inclusive) and the shape parameter ranges from 0.6 to 1.4 (inclusive) for the EWMA statistic used in this chapter with smoothing constant α = 0.1.

B. Appendix

```
*************************************************************
Family:  c("WEI", "Weibull")

Call:  gamlss(formula = TBE ~ wd + (wd == "Thursday") * (cos(2 *
    pi * hr/24) + sin(2 * pi * hr/24)) + cos(2 * pi *
    day/365.25) + sin(2 * pi * day/365.25), sigma.formula = ~(cos(2 *
    pi * hr/24) + sin(2 * pi * hr/24)) + cos(2 * pi *
    day/365.25) + sin(2 * pi * day/365.25), family = WEI(),      data = temp)

Fitting method: RS()
```

```
     -----------------------------------------------------------------
     Mu link function:  log
     Mu Coefficients:
                                            Estimate Std. Error  t value Pr(>|t|)
     (Intercept)                           -3.369757   0.014082 -239.291  < 2e-16 ***
     wdMonday                              -0.078713   0.019199   -4.100 4.14e-05 ***
     wdSaturday                             0.002747   0.019667    0.140   0.8889
     wdSunday                              -0.077758   0.019225   -4.045 5.25e-05 ***
     wdThursday                            -0.021966   0.019891   -1.104   0.2695
     wdTuesday                             -0.037717   0.019439   -1.940   0.0524 .
      wdWednesday                          -0.031362   0.019439   -1.613   0.1067
     cos(2 * pi * hr/24)                    0.170919   0.008217   20.802  < 2e-16 ***
     sin(2 * pi * hr/24)                    0.244262   0.008684   28.127  < 2e-16 ***
     cos(2 * pi * day/365.25)               0.143229   0.007876   18.186  < 2e-16 ***
     sin(2 * pi * day/365.25)               0.130402   0.007559   17.252  < 2e-16 ***
     wd == "Thursday"TRUE:cos(2 * pi * hr/24) 0.035362 0.020530   1.722   0.0850 .
     wd == "Thursday"TRUE:sin(2 * pi * hr/24) 0.030439 0.021941   1.387   0.1654
     ---
     Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     -----------------------------------------------------------------
     Sigma link function:  log
     Sigma Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
     (Intercept)             -0.007815   0.004079  -1.916  0.05535 .
     cos(2 * pi * hr/24)     -0.028340   0.005434  -5.216 1.84e-07 ***
     sin(2 * pi * hr/24)     -0.008343   0.006060  -1.377  0.16859
     cos(2 * pi * day/365.25) -0.009082  0.005691  -1.596  0.11058
     sin(2 * pi * day/365.25)  0.015188  0.005507   2.758  0.00581 **
     ---
     Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     -----------------------------------------------------------------
     No. of observations in the fit:  38143
     Degrees of Freedom for the fit:  18
           Residual Deg. of Freedom:  38125
                           at cycle:  5

      Global Deviance:     -187355.7
                  AIC:     -187319.7
                  SBC:     -187165.8
```

Author details

Ross Stewart Sparks* and Hossein Hazrati-Marangaloo
Data61, CSIRO, Sydney, NSW, Australia

*Address all correspondence to: ross.sparks@data61.csiro.au

IntechOpen

References

[1] Sparks R, Keighley T, Muscatello D. Exponentially weighted moving average plans for detecting unusual negative binomial counts. IIE Transactions. 2010;**42**(10):721-733

[2] Sparks R, Carter C, Graham P, Muscatello D, Churches T, Kaldor J, et al. Understanding sources of variation in syndromic surveillance for early warning of natural or intentional disease outbreaks. IIE Transactions. 2010;**42**(9): 613-631

[3] Sparks RS, Keighley T, Muscatello D. Early warning CUSUM plans for surveillance of negative binomial daily disease counts. Journal of Applied Statistics. 2010;**37**(11):1911-1929

[4] Sparks R, Jin B, Karimi S, Paris C, MacIntyre CR. Real-time monitoring of events applied to syndromic surveillance. Quality Engineering. 2019; **31**(1):73-90

[5] Liu HC, Liu L, Liu N. Risk evaluation approaches in failure mode and effects analysis: A literature review. Expert Systems with Applications. 2013;**40**(2): 828-838

[6] Sparks RS, Keighley T, Muscatello D. Optimal exponentially weighted moving average (EWMA) plans for detecting seasonal epidemics when faced with non-homogeneous negative binomial counts. Journal of Applied Statistics. 2011;**38**(10):2165-2181

[7] Weiß CH. EWMA monitoring of correlated processes of Poisson counts. Quality Technology and Quantitative Management. 2009;**6**(2):137-153

[8] Weiß CH. Detecting mean increases in Poisson INAR (1) processes with EWMA control charts. Journal of Applied Statistics. 2011;**38**(2):383-398

[9] Weiss CH, Testik MC. CUSUM monitoring of first-order integer-valued autoregressive processes of Poisson counts. Journal of Quality Technology. 2009;**41**(4):389-400

[10] Weiß CH, Testik MC. The Poisson INAR (1) CUSUM chart under overdispersion and estimation error. IIE Transactions. 2011;**43**(11):805-818

[11] Yontay P, Weiß CH, Testik MC, Pelin Bayindir Z. A two-sided cumulative sum chart for first-order integer-valued autoregressive processes of poisson counts. Quality and Reliability Engineering International. 2013;**29**(1):33-42

[12] Albarracin OYE, Alencar AP, Lee Ho L. CUSUM chart to monitor autocorrelated counts using negative binomial GARMA model. Statistical Methods in Medical Research. 2018; **27**(9):2859-2871

[13] gamlss: Generalised Additive Models for Location Scale and Shape. Available from: https://cran.r-project. org/web/packages/gamlss

[14] Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. JRSS: Series C (Applied Statistics). 2005;**54**:507-554. DOI: 10.1111/j.1467-9876.2005.00510.x

[15] Sparks RS. CUSUM charts for signalling varying location shifts. Journal of Quality Technology. 2000; **32**(2):157-171