Scientia Iranica E (2016) 23(1), 316–329

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
www.scientiairanica.com

Bayesian multiple change-point estimation of Poisson rates in control charts

H. Assareh[a,*], R. Noorossana[b], M. Mohammadi[c] and K. Mengersen[d]

a. *Simpson Centre for Health Services Research, Australian Institute of Health Innovation, Faculty of Medicine, University of New South Wales, Sydney, NSW 2052, Australia.*
b. *School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran.*
c. *Department of Industrial Engineering, Islamic Azad University-Science and Research Branch, Tehran, Iran.*
d. *Faculty of Science and Engineering, Queensland University of Technology, Brisbane, QLD 4001, Australia.*

Received 22 November 2013; received in revised form 24 June 2014; accepted 16 February 2015

KEYWORDS
Bayesian hierarchical model;
Multiple change-point;
Control charts;
Reversible jump Markov chain Monte Carlo;
Poisson process.

Abstract. Effectiveness of root cause analysis efforts, following a control chart signal, will be enhanced if there exists more accurate information about the true time of change in the process. In this study, we consider a Poisson process experiencing an unknown multiple number of step changes in the Poisson rate. We formulate the multiple change-point scenario using Bayesian hierarchical models. We compute posterior distributions of the change-point parameters including number, location, and magnitude of changes and also corresponding probabilistic intervals and inferences through Reversible Jump Markov Chain Monte Carlo methods. The performance of the Bayesian estimator is investigated over several simulated change-point scenarios. Results show that when the proposed Bayesian estimator is used in conjunction with the c-chart, it can provide precise estimates about the underlying change-point scenario (number, timing, direction, and size of step changes). In comparison with alternatives, including Poisson EWMA and CUSUM built-in estimators and a maximum likelihood estimator, our estimator performs satisfactorily over consecutive monotonic and non-monotonic changes. The proposed Bayesian model and computation framework also benefit from probability quantification as well as flexibility, which allow us to formulate other process types and change scenarios.

1. Introduction

Control charts are used to distinguish if a process works in presence of common and known variations versus unpredictable variations. Control charts signal when an assignable cause occurs. Following a signal, we initiate a search to identify potential causes of the change and, accordingly, conduct corrective or preventive actions [1]. Knowing when the process truly began to change (change-point) enables us to conduct root cause analysis more efficiently, since a tighter time-frame prior to the signal in the control charts is investigated [2,3].

Several statistical methods have been employed in development of change-point estimators for a broad range of processes and change types [4,5]. Maximum Likelihood (ML)-based estimators were found superior in estimation of step changes [6,7] and linear trends [8] in Poisson rates compared to early built-in estimators [9-11] of cumulative sum (CUSUM) [10,11] and Exponentially Weighted Moving Average (EWMA) [12] control charts. Perry et al. [13] further relaxed the underlying assumption of knowing the form of change types and developed an ML estimator for non-

*. *Corresponding author. Tel.: 0061 2 93858695;*
E-mail addresses: hassan.assareh@health.nsw.gov.au (H. Assareh); rassoul@iust.ac.ir (R. Noorossana); majid_mohamadi44@yahoo.com (M. Mohammadi); k.mengersen@qut.edu.au (K. Mengersen)

decreasing multiple step change points (unknown number of consecutive changes) using isotonic regression models. The ML framework has also been applied for different change scenarios in correlated Poisson observations [14-16] and other attributes [17-22]. A Bayesian modeling and computation framework has recently been proposed as an alternative platform for both attribute [3,23] and variable characteristics [24-26]. They were found highly flexible in incorporation of process complexity (e.g. case mix [27] and censoring time [28]) while providing precise estimates and informative probabilistic inferences for various sudden step [29] and linear disturbances [30].

In practice, it is not uncommon to experience non-monotonic unknown consecutive changes in a Poisson process that may occur as a result of one influential process input variable changing several times or several influential process input variables changing at different times. Indeed, these changes could influence the process mean in any direction and lead to multiple change-points in the Poisson mean, which are not necessarily monotonic. None of the built-in estimators of CUSUM and EWMA charts can describe such multiple change-point scenarios and provide specific estimates for more than one change. Moreover, the recent ML-based and Bayesian estimators of multiple change-point still remain dependent on *a priori* knowledge about the shifts, such as monotonic form [13] or the number of changes [23].

In this study, we model a multiple change-point scenario in a Bayesian framework assuming that no *a priori* knowledge about the form of consecutive changes exists. In this setting, we relax two constrains:

1. Knowing the number of change points described in Assareh et al. [23];
2. Monotonically increasing or decreasing the consecutive changes form previously assumed in Perry et al. [13] and elsewhere [17,22].

Hence, the number of change points is incorporated into the proposed Bayesian multiple change-point model in order to be directly estimated; whereas the model that was recently developed by Assareh et al. [23] focused on identifying the change point in the case of knowing the number of changes. We employ Reversible Jump Markov Chain Monte Carlo (RJMCMC) methods [31] to obtain posterior distributions of all multiple change-point model parameters (number, time, and size). First, we describe the Bayesian model and RJMCMC components. The application of the model is demonstrated through an illustrative simulation of a scenario with one change point. Then, we investigate performance of the model over a wide range of consecutive changes and also false alarms. The model is explained in Section 2 and implemented and analyzed in Section 3. Then, we compare performance of the estimator with those of alternatives in Section 4 and summarize the study and obtained results in Section 5.

2. Bayesian Poisson process multiple change-point model and RJMCMC method

2.1. Model

We employed Bayesian Hierarchical Models (BHM) to formulate a multiple change-point scenario in a Poisson process. Consider a process X_t, $t = 1, \cdots, T$, that is initially in-control and then k change points with unknown locations and magnitude occur in the Poisson rate. Thus, at k unknown points in time $\tau_{k,1}, \tau_{k,2}, \cdots, \tau_{k,k}$, the rate parameter changes from its known in-control state of $\lambda_{k,0}$ to $\lambda_{k,l}$, $\lambda_{k,l} = \lambda_{k,0} + \delta_{k,l}$ and $\lambda_{k,l} \neq \lambda_{k,0}$ for $l = 1, \cdots, k$. The Poisson process multiple change-point model can be parameterized as $x_t \sim \text{Poisson}(\lambda_{k,i})$, $t = \tau_{k,i} + 1, \cdots, \tau_{k,i+1}$ for $i = 0, \cdots, k$, where $\tau_{k,0} = 0$ and $\tau_{k,k+1} = T$. That is:

$$p(x_t|\lambda_t) = \begin{cases} \exp(-\lambda_{k,0})\lambda_{k,0}^{x_t}/x_t! \\ \quad \text{if} \quad t = 1, 2, \cdots, \tau_{k,1} \\ \\ \exp(-\lambda_{k,1})\lambda_{k,1}^{x_t}/x_t! \\ \quad \text{if} \quad t = \tau_{k,1} + 1, \cdots, \tau_{k,2} \\ \vdots \\ \exp(-\lambda_{k,k})\lambda_{k,k}^{x_t}/x_t! \\ \quad \text{if} \quad t = \tau_{k,k} + 1, \cdots, T. \end{cases} \tag{1}$$

Thus, the quantities of interest are the number, the time, and the magnitude of the changes.

Let the maximum number of change points be $K - 1 > 0$, so that there exist K models, m_k, $k = 0, 1, \cdots, K - 1$, where k is the number of changes in the Poisson process. We assign a discrete distribution for k; for example, in the following simulation study, a non-informative flat uniform distribution is imposed; however, K is set to 7 based on the problem context, so that $f(m = k) = 1/7$, $k = 0, \cdots, 6$. This setting lets the model test seven hypotheses including occurrence of a false alarm (no change point), one change point, two change points, and up to six consecutive change points. In other contexts, other distributions, such as truncated Poisson or Gamma, might also be of interest (see [32] for more details on the selection of prior distributions).

We place a Gamma distribution as a prior for the mean of the Poisson process, so that $\lambda_{k,i} \sim \Gamma(\alpha_{k,i}, \beta_{k,i})$, $i = 0, \cdots, k$. For example, in the simulation study described below, since no other information on which to base the choice of the hyperparameters exists, we follow Carlin and Louis [33] and set all $\alpha_{k,i}$, $\beta_{k,i}$ for $k = 0, \cdots, K - 1$ and $i = 0, \cdots, k$ to be

$$\alpha = \min\left(1, \frac{f(x|m',\theta'_{m'})f(\theta'_{m'}|m')f(m')j(m',m)q(u'|\theta_{m'},m',m)}{f(x|m,\theta_m)f(\theta_m|m)f(m)j(m,m')q(u|\theta_m,m,m')}\left|\frac{\partial g(\theta_m,u)}{\partial(\theta_m,u)}\right|\right). \quad (2)$$

Box I

equal and use empirical Bayes methods to estimate α and β. Thus, we let the prior have a mean (α/β) of 20, equal to the in-control rate $\lambda_{k,0}$ and a variance (α/β^2) of at least $6 \times \sqrt{\lambda_{k,0}}$, approximately. This is a reasonably informative prior for the magnitude of the change in an in-control Poisson rate as the control chart is sensitive enough to detect very large shifts and estimate associated change points. We thus set $\alpha = 10$ and $\beta = 0.5$. A non-informative uniform distribution was also considered to be a prior distribution for the time of change points in the model. Specifications of this prior were modified within each iteration of the posterior computation algorithm (see Section 2.2).

In the above setting, minimum knowledge was incorporated into the model through the priors. It only included the general design criteria in the SPC context and the in-control Poisson rate. We intentionally let the model not be highly informative in order to run a balanced performance comparison with non-Bayesian methods in this study. In practice, for a specific process, knowledge about possible change patterns, timing, and magnitude of the changes as well as design characteristics, such as desired sensitivity, can also contribute to the prior setting which may lead to more accurate estimates.

2.2. Parameter estimation

To obtain posterior estimates of the parameters of interest, we apply the RJMCMC method developed by Green [31], which has been extensively studied and used in complex change point and model selection problems [34-36].

RJMCMC provides a general framework for Markov Chain Monte Carlo (MCMC) simulation in which the dimension of the parameter space can vary between iterations of the Markov chain. Thus, the dimension of the space, here the number of change points as well as the time and magnitude of the changes given the dimension are considered to be stochastic variables. In this view, the reversible jump sampler can be seen as an extension of the standard Metropolis-Hastings algorithm into more general state spaces that jumps between models with parameter spaces of different dimensions.

Let θ_m denote the parameter vector corresponding to model m, where θ_m has dimension d_m. If the current state of the Markov chain is (m, θ_m), then a general version of the algorithm is the following:

1. Propose a new model m' with probability $j(m, m')$;
2. Generate u from a specified proposal density q $(u|\theta_m, m, m')$;
3. Propose a new vector of parameters $\theta'_{m'}$ by setting $(\theta'_{m'}, u') = g_{m,m'}(\theta_m, u)$ where $g_{m,m'}$ is a specified invertible function;
4. Accept model m' with the probability as obtained in Box I.
5. Return to step 1 until the required number of iterations is reached.

The portion of times that a model m is accepted in the simulation represents the posterior probability of the model and the samples from each iteration within the model m are drawn from the posterior distributions of the parameter set of θ_m.

Important elements of the algorithm are the proposal distributions $q(u'|\theta_{m'}, m', m)$ and the matching function $g_{m,m'}$. The vectors u and u' are used to make the dimensions of the parameter spaces of m and m' equal.

The corresponding proposal distributions are usually constructed by single MCMC runs within each model, while the matching function $g_{m,m'}$ is constructed by considering the structural properties of each model and their possible association.

We adopt the approach taken by Zhao and Chu [36]. In this approach, proposing a new model is restricted to adjacent models which have one change point more or less. This is known as birth or death of a change point, respectively. We have outlined the algorithm components in Appendix A: *RJMCMC Components*.

3. Performance analysis

3.1. Simulation method

We used Monte Carlo simulation to study the performance of the constructed BHM in multiple change detection following a signal from a c-chart.

To demonstrate the generalizability of the proposed model, processes with one, two, and three change points were considered, $k = \{1, 2, 3\}$. We generated 25 observations of a Poisson process with an in-control rate of $\lambda_{k,0} = 20$. Then, we induced step changes until the c-chart [37,38] signalled. Because we knew that the process was in-control, if an out-of-control observation was generated in the simulation of the first 25 in-control observations, similar to the approach applied by Perry et al. [13] and others, it was taken as a false alarm and the simulation was restarted. However, in practice, a false alarm may lead the process to stop and search for root causes. When no root cause is found, the process would continue without adjustment. We, separately, investigated the performance of the Bayesian model on false alarms. The simulation was also repeated for rate parameters of 5 and 10 over equivalent step changes; since the results are similar to those obtained for $\lambda_{k,0} = 20$, $k = \{1, 2, 3\}$, they are not reported here.

To investigate the behavior of the Bayesian estimator of the number, k, magnitude, $\delta_{k,1}, \ldots, \delta_{k,k}$, and location, $\tau_{k,1}, \cdots, \tau_{k,k}$, of changes, we replicated the simulation method explained above 100 times.

The number of replication studies, indeed, is a compromise between excessive computational time, considering RJMCMC iterations, and sufficiency of the achievable distributions even for tails. Since scenarios with more than one change point prior to the control chart's signal were also of interest, simulated datasets that did not meet the desired processes were excluded.

The multiple changes and control charts were simulated in MATLAB. For each change-point scenario, we modified and used RJMCMC algorithm made available in MATLAB by Zhao and Chu [36] to generate 100,000 samples with the first 20000 samples ignored as burn-in.

3.2. An illustrative example

Two processes were simulated in which step change of sizes $\delta_{1,1} = +5$ and $\delta_{1,1} = -5$ were induced at $\tau_{1,1} = 25$. The posterior distributions for the number, time, and magnitude of step changes for the two processes are presented in Figure 1. For both change sizes, a larger mass of the probability function concentrates on the model with one change point; see Figure 1(a) and (b). Acceptance of the model with one change point, m_1, leads to posterior distributions of the time, $\tau_{1,1}$, and the magnitude, $\delta_{1,1}$, of the change. As seen in Figure 1(c) and (d), the posteriors for time concentrate on the 25th sample, which is the real change point. The posteriors for magnitude approximately concentrate on the true size of shifts, $\delta_{1,1} = \pm 5$, as shown in Figure 1(e) and (f). Since the posteriors tend to be asymmetric, in particular for time, the mode of the posteriors is used as an estimator for the change-point model parameter.

Table 1 shows the posterior estimates for the induced change sizes, $\delta_{1,1} = \pm 5$, in the process mean. The Bayesian estimator suggests that one change point is more probable, $p(m_1) = 0.43$, prior to signal of the c-chart where a change of size $\delta_{1,1} = -5$, around one standard deviation, was induced. Where the c-chart detects a fall after 51 samples, the mode of the posterior distribution of $\tau_{1,1}$ reports the 25th sample accurately as the change point. For an increase in size

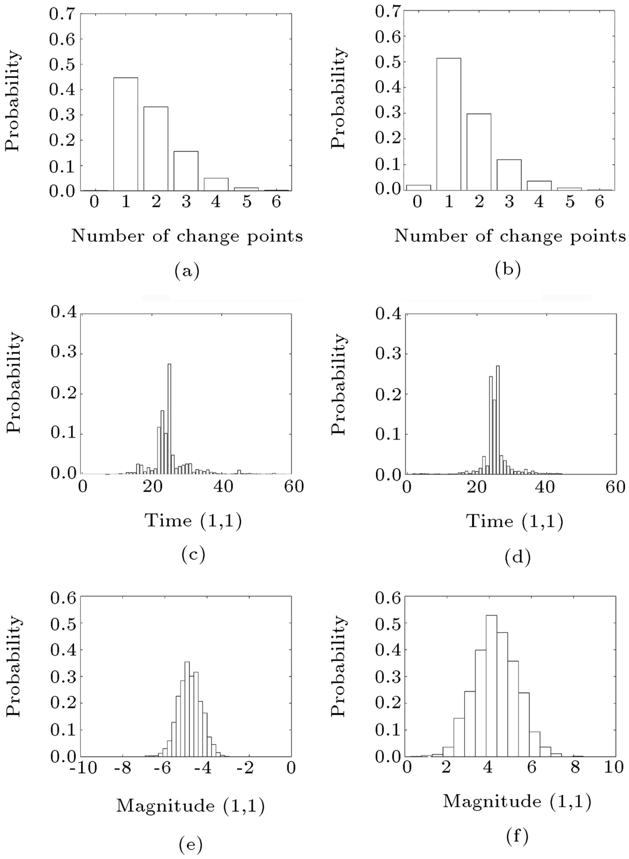

Figure 1. Posterior distributions of the number k, time $\tau_{1,1}$, and magnitude $\delta_{1,1}$ of a step change of sizes: (a,c,e) $\delta_{1,1} = -5$, and (b,d,f) $\delta_{1,1} = +5$ following signals from c-chart where $\lambda_{1,0} = 20$ and $\tau_{1,1} = 25$.

$\delta_{1,1} = +5$, although the posterior mode underestimates the time of the change, $\tau_{1,1} = 24$, it still provides a reasonable accurate estimation with a bias of one sample. Posterior estimates of the magnitude of the change tend to be accurate.

Applying the Bayesian framework enables us to construct probability-based intervals around estimated parameters. A Credible Interval (CI) is a posterior probability-based interval which involves those values of highest probability in the posterior density of the parameter of interest. Table 1 also presents 80% credible intervals for the estimated time and magnitude of step changes. As expected, the CIs are affected by the dispersion and higher order behavior of the posterior distributions. As shown in Table 1 and discussed above, the magnitude of the changes is not estimated as precisely as the time of the change.

The probability of having a specified number of changes is presented in Table 1. The probability of having more than three changes prior to signalling of the chart is 0.08 when an increase of size $\delta_{1,1} = -5$ has occurred. It is less unlikely, 0.06, for a change of size $\delta_{1,1} = +5$. An interesting inference can also be made on the falseness of the signal using $p(m_0)$,

Table 1. Posterior distributions (mode, standard deviation (sd)) of multiple change-point model parameters, m_k and $\theta_{m_1} = (\tau_{1,1}, \delta_{1,1})$, following signals ($RL$) from c-chart where $\lambda_{1,0} = 20$ and $\tau_{1,1} = 25$. Standard deviations and 80% credible intervals are shown in parentheses and brackets, respectively.

$\delta_{1,1}$	RL	$p(m_k)$				θ_{m_1}	
		m_0	m_1	m_2	m_3	$\hat{\tau}_{1,1}$	$\hat{\delta}_{1,1}$
-5	76	0.001	0.43	0.34	0.15	25 (5.46) [24.47,25.55]	-4.97 (0.51) [-5.01,-4.95]
+5	44	0.019	0.51	0.29	0.12	24 (3.65) [23.94,24.28]	4.76 (1.02) [4.66,4.81]

Table 2. Average of posterior estimates ($E(mode), E(sd.)$) of multiple change-point model parameters, m_k and $\theta_{m_1} = (\tau_{1,1}, \delta_{1,1})$, following signals ($RL$) from c-chart where $\lambda_{1,0} = 20$ and a step change was induced at $\tau_{1,1} = 25$. Standard deviations are shown in parentheses.

$\delta_{1,1}$	$E(RL)$	$E(p(m_k))$				θ_{m_1}		
		m_0	m_1	m_2	m_3	$E(\hat{\tau}_{1,1})$	$E(\hat{\sigma}_{\hat{\tau}_{1,1}})$	$E(\hat{\delta}_{1,1})$
-15	26.30 (0.64)	0.001 (0.05)	0.36 (0.04)	0.34 (0.02)	0.18 (0.02)	25.65 (0.71)	2.08 (1.54)	-6.45 (2.29)
-10	32.83 (7.23)	0.006 (0.01)	0.68 (0.08)	0.24 (0.05)	0.05 (0.02)	26.00 (0.93)	2.54 (3.84)	-9.52 (1.41)
-5	148.92 (126.32)	0.10 (0.08)	0.36 (0.05)	0.27 (0.03)	0.15 (0.03)	28.25 (2.81)	127.92 (55.34)	-1.49 (0.70)
-3	293.51 (257.68)	0.24 (0.09)	0.35 (0.07)	0.22 (0.02)	0.10 (0.02)	33.35 (13.12)	529.13 (180.05)	-0.31 (0.28)
+3	100.39 (58.25)	0.22 (0.06)	0.32 (0.03)	0.24 (0.02)	0.13 (0.01)	28.14 (12.38)	88.24 (24.77)	1.09 (1.04)
+5	45.10 (20.21)	0.05 (0.07)	0.40 (0.05)	0.30 (0.02)	0.15 (0.02)	27.72 (9.25)	15.95 (10.80)	3.25 (1.36)
+10	29.18 (3.34)	0.002 (0.01)	0.58 (0.08)	0.28 (0.04)	0.09 (0.03)	26.01 (1.35)	1.53 (1.98)	8.38 (2.15)
+15	26.63 (1.06)	0.008 (0.03)	0.51 (0.12)	0.28 (0.05)	0.12 (0.04)	25.86 (0.91)	1.16 (1.35)	11.3 (3.71)

the probability of having no step change point. Note that only an upper threshold ($K = 7$) was imposed on the number of change points in the model, letting the model examine a no change-point scenario.

We can also construct other probabilistic inferences using the posterior distributions of parameters. As an example, the probability that a change point occurs in the last 10, 20, and 40 observed samples prior to signalling in the control charts can be obtained. For a step change of size $\delta_{1,1} = -5$, since the c-chart signals very late (see Table 1), it is unlikely that the change point has occurred in the last 10, 20, and even 40 samples with probabilities 0.0, 0.0 and 0.03, respectively, whereas for a change of size $\delta = +5$, it is very probable that the change has occurred in the last 40 samples with probability 0.99, and with probability of 0.58, it is between the last 10 and 20 samples. Such probability computations and inferences can be extended to the magnitude of the changes.

3.3. Simulation results

We examined the behavior of the proposed Bayesian model in estimation of the number, location, and magnitude of changes using 100 replications of the simulated datasets outlined in Section 3.1. Table 2 shows the average of the estimated parameters obtained from the replicated datasets where there existed one change point prior to signal of the control chart.

For all change sizes, the model with one change point, m_1, has the highest posterior probability; however, the strength of this comparison varies over different change sizes. As seen, this probability, $p(m_1)$, almost doubles when the magnitude of the shift increases from $\delta_{1,1} = \pm 3$, approximately half a standard deviation, to $\delta_{1,1} = \pm 10$, two standard deviations. As the magnitude of the change increases, the posterior probability of the model with no change point, $p(m_0)$, decreases in favor of the model with one change point. This implies that the model with no change point, m_0, closely competes with the model with one change point, m_1, over small shifts, whereas the model with two change points, m_2, is the runner-up over medium to large shifts. For a large shift, $\delta_{1,1} = \pm 15$ around three standard deviations, the probability of a model with one change point significantly drops, particularly where there exists a drop in the Poisson mean. This is due to the early detection of such shifts by the c-chart that leads to a very short run of samples after the change, which then compresses the data and hence informs the RJMCMC algorithm.

For changes of size of less than one standard deviation ($\delta_{1,1} = \pm 3, \pm 5$) in the Poisson rate, the average of the posterior modes, denote here by $E(\hat{\tau})$, reports 33rd sample (at latest) as the change point. The corresponding c-charts detect the changes with delays greater than 75 samples. The superiority of using Bayesian estimator in conjunction with c-chart persists where a medium shift of size $\delta = \pm 10$ has occurred in the process mean. In this scenario, the bias of the Bayesian estimator does not exceed one observation, whereas the minimum delay is four samples in detection of the fall. As expected, for large shift sizes ($\delta_{1,1} = \pm 15$) around three standard deviations, the c-chart performs well, yet the expected values of modes report more accurate estimations with a delay of less than one observation.

Table 2 reveals that the variation of the Bayesian estimates for time tends to reduce when the magnitude of shift in the process mean increases. However, for drops, the observed variation is almost less than those obtained in detection of jumps. This can be explained by the nature of mean and variance parameters in a Poisson distribution, so that a drop in the mean leads to less dispersed observations. The mean of the standard deviation of the posterior estimates of time, $E(\hat{\sigma}_{\hat{\tau}_{1,1}})$, also decreases dramatically when moving from small shift sizes to medium and large shift sizes.

The average of the Bayesian estimates of the magnitude of the change, $E(\hat{\delta}_{1,1})$, shows that the modes of posteriors for change sizes do not perform as well as the corresponding posterior modes of the time across different shift sizes; however, promising results are obtained where a medium shift, $\delta_{1,1} = \pm 10$, has occurred in the process mean. This estimator tends to underestimate the sizes. Having said that, Bayesian estimates of the magnitude of the change must be studied in conjunction with their corresponding standard deviations. In this manner, analysis of credible intervals is effective.

We extended the simulation study to investigate the performance of the proposed model on scenarios with two and three consecutive change points, $k = 2, 3$, prior to the signal of the control chart. A series of monotonic and non-monotonic changes were considered in this setting. Average of Bayesian estimates, mode of posterior distributions, are summarized and depicted in Figures 2 and 3. See Appendix B: *Simulation results* for details of simulation results for two and three change-points scenarios.

As shown in Figure 2(a), in all two change scenarios (monotonic and non-monotonic), the posterior probability of the model with two change points, m_2, is the highest; however, the strength of this varies over different change sizes. The estimator tends to distinguish non-monotonic scenarios better than monotonic cases (e.g. $p(m_2(+5, -10)) = 0.59$ versus $p(m_2(+5, +10)) = 0.40$). With an increase in the absolute difference between two consecutive changes, the likelihood of the model with two changes increases (e.g. $p(m_2(+5, -10)) = 0.59$ versus $p(m_2(+5, -5)) = 0.35$). For monotonic changes, the model with one change point, m_1, competes with the true model, m_2, whereas for small non-monotonic changes, the model with no change, m_0, also contends with the true model.

Figure 2(b) shows that the proposed model reasonably well estimates change parameters of two change-point scenarios; however, the performance is

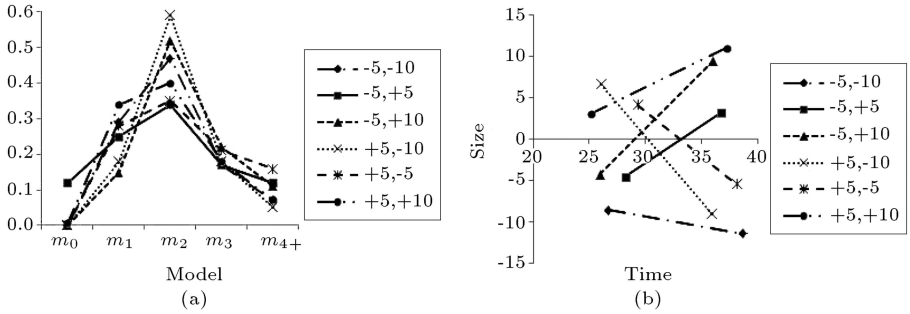

Figure 2. Average of posterior estimates for (a) the probability of models with k change points, m_k, and (b) the time, $\tau_{2,1}$ and $\tau_{2,2}$, and the magnitude, $\delta_{2,1}$ and $\delta_{2,2}$, of changes in scenarios with two consecutive change points following signals from c-chart where $\lambda_{2,0} = 20$.

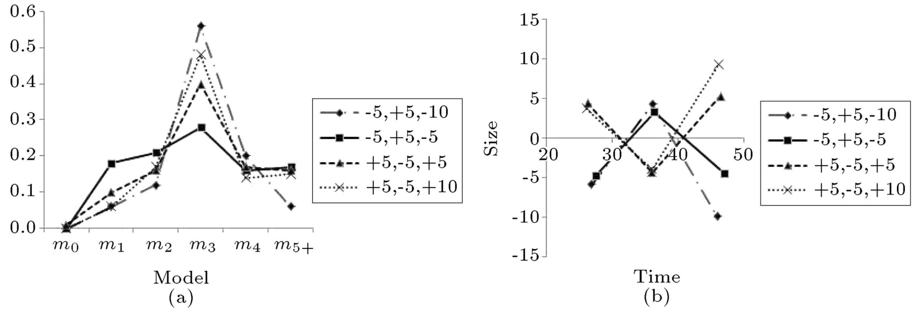

Figure 3. Average of posterior estimates for (a) the probability of models with k change points, m_k, and (b) the time, $\tau_{3,1}$, $\tau_{3,2}$ and $\tau_{3,3}$, and the magnitude, $\delta_{3,1}$, $\delta_{3,2}$ and $\delta_{3,3}$, of changes in scenarios with three consecutive change points following signals from c-chart where $\lambda_{3,0} = 20$.

affected by the direction of changes and the size of their absolute differences. For time of changes, $\hat{\tau}_{2,1}$ and $\hat{\tau}_{2,2}$, more accurate estimates were obtained where there existed non-monotonic changes with larger absolute differences. The average of the Bayesian estimates of the magnitude of the changes, $\hat{\delta}_{2,1}$ and $\hat{\delta}_{2,2}$, shows that while the point estimates slightly deviate from the true values, there is no consistent pattern in these deviations and the true values are typically encompassed in the corresponding 80% CIs.

Similar to the two change-point cases, Figure 3(a) shows that in all scenarios with three changes, the true model, m_3 model with three change points, has the highest probability. However, the strength of this probability varies over different change sizes. It is seen that when the magnitude of the third change increases from one standard deviation to two standard deviations approximately, the Bayesian estimator distinguishes the true model more strongly. Accordingly, when the magnitude of the third change increases, the Bayesian estimator tends to be more accurate in estimation of time, $\hat{\tau}_{3,1}$, $\hat{\tau}_{3,2}$ and $\hat{\tau}_{3,3}$ (see Figure 3(b)). The accuracy and the direction of bias of Bayesian estimates for the magnitude of the changes, $\hat{\delta}_{3,1}$, $\hat{\delta}_{3,2}$ and $\hat{\delta}_{3,3}$, are not consistent across different scenarios. However, there exist some gains in studying the estimated sizes and directions, particularly when the obtained standard deviations are also considered.

3.4. False alarms

Incorporation of the number of change points within the proposed Bayesian hierarchical model and allocation of a prior distribution, which includes a no change-point hypothesis, enable us to evaluate the possibility of occurrence of a false alarm within the process along other multiple change-point hypotheses. To investigate the performance of the proposed model in detection of false alarms against true change points, 1000 datasets of Poisson processes with an in-control rate of $\lambda_{k,0} = 20$ were generated. Upon c-chart's false signals, the change-point model was employed and the posterior estimate for the number of change points was obtained and then compared with simulated scenarios in which a genuine small shift had occurred ($\delta_{1,1} = \pm 3$ at 25th observation). We also analyzed the performance of the proposed estimator over false alarms in presence of over-dispersed observations (over-dispersion index: $\sigma_x^2/\text{mix} > 1$). We generated over-dispersed datasets using a Gamma-Poisson mixture parameterization, $\Gamma(\alpha = 1/d, \beta = \lambda_0 \times d)$, letting an over-dispersion index of $\sigma_x^2/\lambda_0 = 1 + d \times \lambda_0$ with d as over-dispersion parameter. To simulate a range of small to large over-dispersed Poisson observations, three magnitudes of $d = (0.025, 0.01, 0.1)$ were considered. For homogeneity of comparison, datasets with false signals prior the 25th observation were excluded.

Likelihood averages of alternative models for all scenarios are summarized in Table 3. In presence of a false alarm for equip-dispersed observations, although the model with no change point is marginally the superior model compared to the model with one change point with a slightly higher likelihood ($p(m_0) = 0.315$ vs. $p(m_1) = 0.313$), over replications and taking Precision of estimates into account, the observed difference between the two alternatives is not statistically significant. Having said that, for scenarios with genuine small shifts of size $\delta_{1,1} = \pm 3$, the estimated likelihood of the model with one change point was significantly larger than the associated probability for the model with no change point (32% vs. 25% for a decrease and 33% vs. 22% for an increase, respectively, as shown in Table 3). The observed interval estimates over replications for likelihood of the two alternative hypotheses of false alarm and a change point between the three simulated scenarios demonstrate some capability of the proposed Bayesian change-point model in differentiation of random and genuine shifts in the Poisson process as long as underlying assumptions are met. As shown in Table 3, by increase in the magnitude of variance versus mean of Poisson-based observations, i.e. in presence of over-dispersion, the c-chart falsely signals earlier. Our simulation study reveals that if the process exhibits any degree of over-dispersion, not only insignificant marginal superiority of the model with no change point

Table 3. Average of posterior estimates for the probability of models with k change points, m_k, in presence of false signals, over-dispersion and small step changes in a c-chart with $\lambda_{1,0} = 20$. 95% confidence intervals of estimates over replications are shown in parentheses.

Scenario	$E(RL)$	$E(p(m_k))$			
		m_0	m_1	m_2	m_3
False signal	325.77	0.315	0.313	0.207	0.104
equi-dispersed ($d = 0$)	(310.27-341.27)	(0.311-0.319)	(0.311-0.315)	(0.205-0.208)	(0.103-0.105)
False signal	169.45	0.301	0.312	0.212	0.109
small over-dispersed ($d = 0.025$)	(160.15-178.75)	(0.300-0.301)	(0.311-0.312)	(0.212-0.213)	(0.109-0.110)
False signal	94.98	0.284	0.306	0.219	0.117
medium over-dispersed ($d = 0.01$)	(90.82-99.14)	(0.283-0.284)	(0.306-0.307)	(0.219-0.220)	(0.117-0.118)
False signal	38.79	0.215	0.291	0.241	0.146
large over-dispersed ($d = 0.1$)	(37.74-39.85)	(0.215-0.216)	(0.291-0.292)	(0.241-0.242)	(0.146-0.147)
True signal following	284.82	0.256	0.329	0.221	0.115
a small decrease ($\delta_{1,1} = -3$)	(269.33-300.31)	(0.255-0.257)	(0.329-0.330)	(0.221-0.222)	(0.115-0.116)
True signal following	95.25	0.226	0.338	0.234	0.126
a small increase ($\delta_{1,1} = +3$)	(91.55-98.95)	(0.221-0.231)	(0.336-0.341)	(0.232-0.236)	(0.125-0.127)

is no longer persistent, but also the likelihood of the model with one change point is significantly greater over replications. With an increased over-dispersion, the likelihood of no change point markedly decreases in favor of the model with one change point, suggesting a true change in the process whereas no change has genuinely occurred.

Further investigation is required to fully formulate and evaluate false alarm detection capability of the proposed model. Sensitivity and comparative analyses across a broader range of desired change size scenarios, setting up decision threshold and use of more informative or truncated priors, are of interest. Furthermore, caution should be taken in employment of the proposed model in distinguishing false signals from true shifts if underlying Poisson assumptions are not satisfied. Within this setting, appropriate selection of control charting methods and, accordingly, modification of change-point models are recommended [39,40].

4. Comparison of Bayesian estimator with other methods

To study the performance of the proposed Bayesian estimator in comparison with alternatives, we considered Poisson EWMA and CUSUM charts and associated built-in estimators [9,10] and the proposed ML estimator for a step change in Poisson processes [6] within replications discussed in Section 3.1.

As expected, since both EWMA and CUSUM charts are very sensitive to shifts, simulation of more than one change point before signalling is unlikely. However, we considered the application of these charts in contexts in which the monitoring process and charts were not terminated when the chart had signalled. Woodall [41] highlighted this circumstance as a significant characteristic of monitoring in a clinical setting, where an out-of-control process may not be able to stop and root cause analysis procedures are conducted simultaneously or with a delay. We chose the ML estimator for step change proposed by Samuel and Pignatiello [6], because it is the only one proposed ML method that can be applied over different change scenarios, as the developed ML estimators for linear trend [8] and multiple change [13] in a Poisson mean are restricted to increasing trends and monotonic changes.

To construct control charts, we applied the procedures of Brook and Evans [42] and Trevanich and Bourke [43] for Poisson CUSUM and Poisson EWMA control charts, respectively. A Poisson CUSUM accumulates the difference between an observed value and a reference value k through $S_i^+ = \max\{0, x_i - k^+ + S_{i-1}^+\}$ and $S_i^- = \max\{0, k^- - x_i + S_{i-1}^-\}$ where $k^+ = (\lambda_1^+ - \lambda_0)/(\ln(\lambda_1^+) - \ln(\lambda_0))$ and $k^- = (\lambda_0 - \lambda_1^-)/(\ln(\lambda_0) - \ln(\lambda_1^-))$. If $S_i^\pm$ exceeds a specified decision interval $h^\pm$, then the control chart signals that an increase (a decrease) in the Poisson rate has occurred. We calibrated the charts to detect a 25% shift in Poisson rates and have an in-control average run length ($A\hat{R}L_0$) of 370, approximately, close to standard c-chart (see Woodall and Adams [44]). The resultant Poisson CUSUM charts had $(k^+, h^+) = (22.4, 22)$ and $(k^-, h^-) = (17.4, 14)$. For simplicity, the values were rounded to one decimal place.

In a Poisson EWMA, cumulative values of observations are obtained through $Z_i = r \times x_i + (r-1) \times Z_{i-1}$, where $Z_0 = \lambda_0$, and plotted in a chart with $UCL = \lambda_0 + A^+\sqrt{\text{Var}Z_i}$ and $LCL = \lambda_0 - A^-\sqrt{\text{Var}Z_i}$. We let

Table 4. Average of change point estimates obtained through the built-in EWMA (τ_{ewma}) and CUSUM (τ_{cusum}), ML (τ_{mle}) and Bayesian (τ_b, time of the first change) estimators following signals from Poisson EWMA (RL_{ewma}), Poisson CUSUM (RL_{cusum}), and c-chart (RL_c), where $\lambda_{k,0} = 20$ and $\tau_{k,1} = 25$. Standard deviations are shown in parentheses.

δ	c-chart		Poisson EWMA		Poisson CUSUM		τ_b
	$E(RL_c)$	$E(\hat{\tau}_{mle})$	$E(RL_{ewma})$	$E(\hat{\tau}_{ewma})$	$E(RL_{\text{cusum}})$	$E(\hat{\tau}_{\text{cusum}})$	
-10	32.83	25.05	27.76	21.82	28.15	23.40	26.00
	(7.23)	(0.92)	(0.95)	(5.23)	(0.79)	(2.59)	(0.93)
-5	148.92	25.13	32.31	22.32	33.18	24.27	28.25
	(126.32)	(3.74)	(3.80)	(5.53)	(3.80)	(3.69)	(2.81)
+5	45.10	26.08	32.14	23.67	33.14	25.23	27.72
	(20.21)	(4.02)	(4.19)	(4.76)	(4.52)	(3.50)	(9.25)
+10	29.18	24.99	28.04	22.35	28.35	23.40	26.01
	(3.34)	(1.64)	(1.31)	(4.69)	(1.32)	(3.31)	(1.35)
-5,-10	42.28	28.39	31.71	22.12	32.62	24.01	26.71
	(6.73)	(4.51)	(3.01)	(5.53)	(2.94)	(3.49)	(2.75)
-5,+10	38.64	33.69	32.16	24.00	32.84	25.87	25.96
	(2.97)	(4.73)	(3.65)	(7.24)	(3.42)	(5.45)	(1.94)
+5,-10	42.30	31.03	31.96	24.87	32.70	26.33	26.07
	(6.42)	(8.69)	(5.39)	(3.93)	(6.84)	(3.74)	(2.21)
+5,+10	39.30	26.62	31.98	22.76	32.61	24.55	25.20
	(2.90)	(4.78)	(3.28)	(5.30)	(3.33)	(3.71)	(1.29)
-5,+5,-10	52.88	39.68	33.02	23.85	34.12	26.71	26.90
	(6.49)	(7.23)	(5.48)	(8.12)	(5.42)	(6.91)	(2.63)
+5,-5,+10	49.23	37.91	33.18	25.77	34.39	27.50	26.15
	(3.59)	(9.10)	(5.90)	(7.93)	(6.00)	(6.75)	(2.27)

$r = 0.1$ and $A^{\pm} = 2.67$ to build a chart with an ARL_0 of 370, close to a standard c-chart.

Table 4 shows the expected values of the Bayesian estimates and detected change points provided by built-in estimators of EWMA [9] and CUSUM [10,11] charts and the ML estimator [6] for a step change in a Poisson process. For scenarios of more than one change, the posterior estimates for the time of the first change were considered.

Where there exists a step change, the Bayesian estimator, $\hat{\tau}_b$, only outperforms the built-in estimators of EWMA, $\hat{\tau}_{ewma}$, and CUSUM, $\hat{\tau}_{\text{cusum}}$, charts over large shifts of $\delta = \pm 10$. It is outperformed by them over medium size changes, $\delta = \pm 5$; however, a larger variation is associated with the estimates obtained by the alternatives. The ML estimator, $\hat{\tau}_{mle}$, also outperforms the Bayesian estimator over medium and large step changes.

Where there exist two consecutive changes that are monotonic, the Bayesian estimator, $\hat{\tau}_b$, outperforms all alternatives, except the CUSUM built-in estimator for $\delta_{1,2} = (-5, -10)$ reporting slightly more accurate, 24.01 vs. 26.71, but a less precise, 3.49 vs. 2.75, estimation of the location of the first change point. If there exist two non-monotonic changes, the Bayesian estimator, $\hat{\tau}_b$, provides reasonably accurate and precise estimates against the alternatives. For three non-monotonic step changes scenarios, it also competes very well with the alternatives taking into account the obtained standard deviations over replications.

In this comparison analysis, we applied all estimators on the same datasets generated at each iteration; thus, in CUSUM and EWMA procedures, we let charts remain in out-of-control state following the signals of chart until the c-chart signals. Considering accuracy and precision of estimates over different change scenarios, the Bayesian estimator exhibits a reasonable performance compared to the alternatives. Poisson EWMA and CUSUM are known to detect small to medium shifts, quickly [1], as also observed in Table 4. However, if the process is not terminated in following the signals of chart, as it is a case within a clinical context [41], the Bayesian estimator can provide more insight about the underlying changes compared to the built-in estimators for root cause analysis purposes. Here, we limited the comparative analysis to runs of Poisson processes up to the signal of c-chart. Certainly, within contexts with late or no stopping points following the signals of CUSUM and EWMA chart, other termination points can be considered. The performance of our Bayesian estimator over different termination points in conjunction with

CUSUM and EWMA charts remains a subject of further research. The superiority of ML estimator over one step change scenarios is not surprising since no limitation is made for the Bayesian estimator; whereas the ML estimator was specifically designed to detect one step change.

By relaxing some of the previously incorporated assumptions in the design of multiple change-point estimators, the form [13] and the number of changes [23], we let post-signal change-point estimation be almost purely data driven. In addition to accuracy and precision criteria used for the comparison study, the obtainable posterior distributions for the number, the time, and the magnitude of changes enable us to construct probabilistic intervals around estimates and probabilistic inferences as discussed in Section 3.2. This is a significant advantage of the proposed Bayesian approach compared to other methods. Although similar results may be obtained when resampling in conjunction with ML methods, the inferential capabilities of this approach are more limited (see Bernardo and Smith [45] for more details).

5. Conclusion

Knowing the true time when a process has changed enhances efficiency of root cause analysis efforts by restricting the search to a tighter window of observations and related variables. In monitoring a quality characteristic, it is likely to experience consecutive changes prior to signalling of quality control procedures (or termination of the process). In this paper, we modeled a multiple change-point problem for a Poisson process in a Bayesian framework. The change point was developed using Bayesian hierarchical models and a RJMCMC computation method was employed to obtain posterior distributions of the model parameters.

We considered three scenarios of changes, a step change, and two and three consecutive changes where they are monotonic and non-monotonic. The performance of the proposed model was investigated through simulation in which the estimator was activated following the signal of a c-chart. Over different scenarios, the Bayesian estimator provided reasonably accurate estimates for underlying change patterns including the number, time, and magnitude of the changes. Incapability of the proposed model in differentiation of genuine and false signals and sensitiveness to the Poisson assumption were also addressed; further investigation is subject to a broader design and analyses. The abilities of the Bayesian model in probability quantification through credible intervals and construction of probabilistic inferences for all parameters, including the location of the change, were illustrated. Then, we compared the Bayesian estimator with built-in estimators of EWMA, CUSUM, and an ML-based estimator. The Bayesian estimator performed reasonably well and remained a strong alternative.

In this study, to fairly compare our Bayesian estimator with alternatives, we incorporated a limited knowledge about the underlying process into the model through the priors. In practice, there may exist some knowledge about the underlying change-point model, as addressed by Assareh et al. [23]; thus, it is worthwhile to better tune the priors accordingly. In this setting more gain in accuracy and precision is expected from the Bayesian model compared to non-Bayesian methods. Further research is required to investigate the performance of the model using well informative priors. Such setting could, in particular, enhance the performance of the Bayesian model in identification of false alarms. Considering extreme Poisson rate, very large or very small, and extension of the model to other processes and change characteristics, such as non-constant inspection unit in a Poisson process or linear trends in rates, are of interest. We employed an algorithm developed by Zhao and Chu [36] based on RJMCMC methods [31] to derive posteriors. There exist other Bayesian formulation and computation methods for multiple change-point problems, e.g. product partition models [46], stochastic approximation in Monte Carlo algorithms [47] and others [48,49]. Comparison of RJMCMC methods with alternatives in control charting is also of interest for further research.

Our results showed that the Bayesian estimator can provide an integrated and comprehensive view of the number, location, magnitude, and direction of changes where no *a priori* knowledge about the change model exists. Compared to the alternatives, it becomes the superior estimator when considering other characteristics including relief of setting assumptions, ability to incorporate any available knowledge through informative priors, probabilistic quantification and inferences features, flexibility of the model, ease of extension to more complicated change scenarios such as combination of steps and linear and nonlinear trends, and relief of analytic calculation of likelihood function, particularly for non-tractable likelihood functions.

Acknowledgments

The authors gratefully acknowledge financial support from the Queensland University of Technology and St Andrews Medical Institute through an ARC Linkage Project.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

Hassan Assareh contributed to the conception, design, and implementation of statistical analysis and to writing and modification of the manuscript. Rassoul Noorossana, Majid Mohammadi, and Kerrie Mengersen contributed to the conception, design, and modification of the manuscript. All the authors read and approved the final manuscript.

References

1. Montgomery, D.C., *Introduction to Statistical Quality Control*, Wiley (2008).
2. Samuel, T. and Pignatiello, J. "Identifying the time of a step change in the process fraction nonconforming", *Quality Engineering*, **13**(3), pp. 357-365 (2001).
3. Assareh, H., Smith, I. and Mengersen, K. "Bayesian change point detection in monitoring cardiac surgery outcomes", *Quality Management in Health Care*, **20**(3), pp. 207-222 (2011).
4. Amiri, A. and Allahyari, S. "Change point estimation methods for control chart postsignal diagnostics: A literature review", *Quality and Reliability Engineering International*, **28**(7), pp. 673-685 (2012).
5. Atashgar, K. "Identification of the change point: an overview", *The International Journal of Advanced Manufacturing Technology*, **64**(9-12), pp. 1663-1683 (2013).
6. Samuel, T. and Pignatiello, J. "Identifying the time of a change in a Poisson rate parameter", *Quality Engineering*, **10**(4), pp. 673-681 (1998).
7. Perry, M.B. "Robust change detection and change point estimation for Poisson count processes", Ph.D. Thesis, Florida State University, USA (2004).
8. Perry, M.B., Pignatiello, J.J. and Simpson, J. "Estimating the change point of a Poisson rate parameter with a linear trend disturbance", *Quality and Reliability Engineering International*, **22**(4), pp. 371-384 (2006).
9. Nishina, K. "A comparison of control charts from the viewpoint of change-point estimation", *Quality and Reliability Engineering International*, **8**(6), pp. 537-541 (1992).
10. Page, E.S. "Continuous inspection schemes", *Biometrika*, **41**(1/2), pp. 100-115 (1954).
11. Page, E.S. "Cumulative sum charts", *Technometrics*, **3**(1), pp. 1-9 (1961).
12. Roberts, S. "Control chart tests based on geometric moving averages", *Technometrics*, **1**, pp. 239-250 (1959).
13. Perry, M.B., Pignatiello, J.J. and Simpson, J. "Change point estimation for monotonically changing Poisson rates in SPC", *International Journal of Production Research*, **45**, pp. 1791-1813 (2007).
14. Niaki, S.T.A. and Khedmati, M. "Detecting and estimating the time of a step-change in multivariate Poisson processes", *Scientia Iranica*, **19**(3), pp. 862-871 (2012).
15. Niaki, S.T.A. and Khedmati, M. "Estimating the change point of the parameter vector of multivariate Poisson processes monitored by a multi-attribute t 2 control chart", *The International Journal of Advanced Manufacturing Technology*, **64**(9-12), pp. 1625-1642 (2013).
16. Sharafi, A., Aminnayeri, M. and Amiri, A. "An MLE approach for estimating the time of step changes in Poisson regression profiles", *Scientia Iranica*, **20**(3), pp. 855-860 (2013).
17. Perry, M.B., Pignatiello, J.J. and Simpson, J.R. "Estimating the change point of the process fraction nonconforming with a monotonic change disturbance in SPC", *Quality and Reliability Engineering International*, **23**(3), pp. 327-339 (2007).
18. Noorossana, R., Saghaei, A., Paynabar, K. and Abdi, S. "Identifying the period of a step change in high-yield processes", *Quality and Reliability Engineering International*, **25**(7), pp. 875-883 (2009).
19. Amiri, A., Koosha, M. and Azhdari, A. "Profile monitoring for Poisson responses", In *Industrial Engineering and Engineering Management (IEEM), 2011 IEEE International Conference on*, pp. 1481-1484, IEEE (2011).
20. Hou, C.-D., Shao, Y.E. and Huang, S. "A combined MLE and generalized P chart approach to estimate the change point of a multinomial process", *Applied Mathematics and Information Sciences*, **7**(4), pp. 1487-1493 (2013).
21. Niaki, S.T.A. and Khedmati, M. "Identifying the change time of multivariate binomial processes for step changes and drifts", *Journal of Industrial Engineering International*, **9**(1), pp. 1-11 (2013).
22. Niaki, S.T.A. and Khedmati, M. "Change point estimation of high-yield processes experiencing monotonic disturbances", *Computers & Industrial Engineering*, **67**, pp. 82-92 (2014).
23. Assareh, H., Noorossana, R. and Mengersen, K. "Bayesian change point estimation in Poisson based control charts", *Journal of Industrial Engineering International*, **9**(1), pp. 32-44 (2013).
24. Pan, R. and Rigdon, S.E. "A Bayesian approach to change point estimation in multivariate SPC", *Journal of Quality Technology*, **44**(3), pp. 231-248 (2012).
25. Tan, M.H. and Shi, J. "A Bayesian approach for interpreting mean shifts in multivariate quality control", *Technometrics*, **54**(3), pp. 294-307 (2012).
26. Shariatmadari, M. "Change point detection in multivariate normal process by Bayesian approach", Master's Thesis, Islamic Azad University, South Tehran Branch, Iran (2013).
27. Assareh, H., Smith, I. and Mengersen, K. "Change

point detection in risk adjusted control charts", *Statistical Methods in Medical Research* (2011).

28. Assareh, H. and Mengersen, K. "Change point estimation in monitoring survival time", *PLoS ONE*, **7**(3), pp. e33630 (2012).

29. Assareh, H. and Mengersen, K. "Bayesian estimation of the time of a decrease in risk-adjusted survival time control charts", *IAENG International Journal of Applied Mathematics*, **41**(4), pp. 360-366 (2011).

30. Assareh, H., Smith, I. and Mengersen, K. "Bayesian estimation of the time of a linear trend in risk-adjusted control charts", *IAENG International Journal of Computer Science*, **38**(4), pp. 409-417 (2011).

31. Green, P. "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination", *Biometrika*, **82**(4), pp. 711-732 (1995).

32. Gelman, A., Carlin, J., Stern, H. and Rubin, D., *Bayesian Data Analysis*, Chapman & Hall/CRC (2004).

33. Carlin, B. and Louis, T. "Empirical bayes: Past, present and future", *Journal of the American Statistical Association*, **95**(452), pp. 1286-1289 (2000).

34. Brooks, S.P. "Markov chain Monte Carlo method and its application", *Journal of the Royal Statistical Society. Series D (The Statistician)*, **47**(1), pp. 69-100 (1998).

35. Lavielle, M. and Lebarbier, E.S.P. "An application of MCMC methods for the multiple change-points problem", *Signal Processing*, **81**(1), pp. 39-53 (2000).

36. Zhao, X. and Chu, P.S. "Bayesian change-point analysis for extreme events (typhoons, heavy rainfall, and heat waves): a RJMCMC approach", *Journal of Climate*, **23**(5), pp. 1034-1046 (2010).

37. Shewhart, W. "Quality control charts", *Bell System Technical Journal*, **5**, pp. 593-602 (1926).

38. Shewhart, W. "Quality control", *Bell System Technical Journal*, **6**, pp. 722-735 (1927).

39. Saghir, A. and Lin, Z. "Control charts for dispersed count data: an overview", *Quality and Reliability Engineering International*, **31**(5), pp. 725-739 (2015). Doi:10.1002/qre.1642.

40. Wilken, B.A. "Change-point methods for overdispersed count data", *Tech. Rep.*, Department of Air Force, Air University, Ohio, United States (2007).

41. Woodall, D.H. "The use of control charts in health-care and public-health surveillance", *Journal of Quality Technology*, **38**(2), pp. 89-104 (2006).

42. Brook, D. and Evans, D. "An approach to the probability distribution of CUSUM run length", *Biometrika*, **59**(3), pp. 539-549 (1972).

43. Trevanich, A. and Bourke, P. "EWMA control charts using attributes data", *The Statistician*, **42**(3), p. 215 (1993).

44. Woodall, W.H. and Adams, B.M. "The statistical design of CUSUM charts", *Quality Engineering*, **5**(4), pp. 559-570 (1993).

45. Bernardo, J.M. and Smith, A.F.M., *Bayesian Theory*, Wiley (1994).

46. Barry, D. and Hartigan, J.A. "A Bayesian analysis for change point problems", *Journal of the American Statistical Association*, **88**(421), pp. 309-319 (1993).

47. Liang, F., Liu, C. and Carroll, R.J. "Stochastic approximation in Monte Carlo computation", *Journal of the American Statistical Association*, **102**(477), pp. 305-320 (2007).

48. Lai, T.L. and Xing, H. "A simple Bayesian approach to multiple change-points", *Statistica Sinica*, **21**(2), p. 539 (2011).

49. Chib, S. "Estimation and comparison of multiple change-point models", *Journal of Econometrics*, **86**(2), pp. 221-241 (1998).

Appendix A

RJMCMC components

Birth and death of a change point

In Step 1 of the RJMCMC algorithm, a model, m_k, is randomly proposed and can be limited to adjacent models, m_{k-1} and m_{k+1}, say of the last iteration. We set the probability of transition to adjacent models, the so-called birth and death of a change point, $j(m_k, m_{k+1}) = j(m_k, m_{k-1}) = 0.5$ for $0 < k < K-1$ and $j(m_0, m_1) = j(m_{K-1}, m_{K-2}) = 1$ where there exists only one adjacent model.

In the followings, for ease of expositions, subscripts of new parameters obtained through birth and death moves are dropped. In the birth of a new change point, τ, in a move from m_k to m_{k+1}, all existing change points and most Poisson rates remain untouched. A non-informative prior for τ is $p(\tau) = 1/(n-k-1)$ as the birth cannot occur on x_t, $t = (1; \tau_{k1}, \cdots, \tau_{kk})$. Assume that τ occurs within $(\tau_{k,j}, \tau_{k,j+1})$ and splits this epoch into two parts. In this circumstance, the old $\lambda_{k,j}$ is replaced by two new rates λ_1 and λ_2, where under the competing model, m_{k+1}, their conditional posteriors are:

$$\lambda_1 | x, \theta_{m_k}, \tau \sim \Gamma\left(\alpha + \sum_{t=\tau_{k,j}}^{\tau-1} x_t, \beta + \tau - \tau_{k,j}\right),$$

$$\lambda_2 | x, \theta_{m_k}, \tau \sim \Gamma\left(\alpha + \sum_{t=\tau}^{\tau_{k,j+1}-1} x_t, \beta + \tau_{k,j+1} - \tau\right). \quad \text{(A.1)}$$

In contrast, for the death of a change point, in a move from m_k to m_{k-1}, two epochs are merged and the two rates are replaced by one rate. The conditional posterior of the merged rate is:

$$\lambda|x,\theta_{m_{k+1}} \sim \Gamma\left(\alpha+\sum_{t=\tau_{k+1,j}}^{\tau_{k+1,j+2}-1} x_t, \beta+\tau_{k+1,j+2}-\tau_{k+1,j}\right). \quad (A.2)$$

Proposal distributions

Finding appropriate proposal densities for moves, $q_{m_k,m_{k+1}}(u'|\theta_{m_{k+1}}, m_{k+1}, m_k)$ for birth and q_{m_{k+1},m_k} $(u|\theta_{m_k}, m_k, m_{k+1})$ for death are critical in the RJMCMC algorithm.

For a birth move, the vector u' includes three parameters τ, λ_1 and λ_2. We let the proposal density be:

$$q(\tau,\lambda_1,\lambda_2|\theta_{m_k}) = p(\lambda_1|\theta_{m_k},\tau) \times p(\lambda_2|\theta_{m_k},\tau) \times p(\tau|\theta_{m_k}), \quad (A.3)$$

where $p(\lambda_1|\theta_{m_k},\tau)$ and $p(\lambda_2|\theta_{m_k},\tau)$ are the posteriors obtained in Eq. (A.1) and $p(\tau|\theta_{m_k})$ is set to the posterior of the new change point calculated as follows (see [36] for derivation details):

$$p(\tau|\theta_{m_k},\lambda_1,\lambda_2,m_{k+1},x) \propto e^{(\tau-\tau_{k,j})(\lambda_1-\lambda_2)}(\lambda_1/\lambda_2)^{\sum_{t=\tau_{k,j}}^{\tau-1} x_t}, \quad (A.4)$$

where λ_1 and λ_2 are replaced by the mean of the posteriors obtained in Eq. (A.1).

For a death move, the vector u includes one parameter λ. We need to propose a new rate for the period $[\tau_{k,j}, \tau_{k,j+1}-1]$ under m_k. Here, the proposal is set in a straightforward manner by applying the posterior of λ as follows:

$$p(\lambda|\theta_{m_{k+1}}) \sim \Gamma\left(\alpha+\sum_{t=\tau_{k+1,j}}^{\tau_{k+1,j+2}-1} x_t, \beta + \tau_{k+1,j+2}-\tau_{k+1,j}\right). \quad (A.5)$$

All priors and proposals introduced in Section 2.1 and the Appendix are then replaced in the acceptance ratio defined in Eq. (2) for birth and death moves, appropriately (see [36] for more details).

Appendix B

Simulation results

To investigate the performance of the proposed model in scenarios with two and three consecutive change points prior to the c-charts signal, we followed the method described in Section 3.1. A series of monotonic and non-monotonic changes were considered in this setting. Average of Bayesian estimates, mode of posterior distributions, over 100 replications, are summarized in Tables B.1 and B.2.

Table B.1. Average of posterior estimates ($E(mode), E(sd.)$) of multiple change-point model parameters m_k and $\theta_{m_2} = (\tau_{2,i}, \delta_{2,i})$, $i = 1, 2$, following signals (RL) from c-chart where $\lambda_{2,0} = 20$, and two consecutive step changes were induced at $\tau_{2,1} = 25$ and $\tau_{2,2} = 35$, respectively. Standard deviations are shown in parentheses.

$\delta_{2,1}, \delta_{2,2}$	$E(RL)$	$E(p(m_k))$				$\theta_{m_2,i=1}$			$\theta_{m_2,i=2}$		
		m_0	m_1	m_2	m_3	$E(\hat\tau_{2,1})$	$E(\hat\sigma_{\hat\tau_{2,1}})$	$E(\hat\delta_{2,1})$	$E(\hat\tau_{2,2})$	$E(\hat\sigma_{\hat\tau_{2,2}})$	$E(\hat\delta_{2,2})$
-5,-10	42.28	0.00	0.29	0.47	0.17	26.71	4.47	-8.53	38.71	6.76	-11.38
	(6.73)	(0.00)	(0.08)	(0.05)	(0.03)	(2.75)	(1.53)	(2.38)	(3.89)	(4.94)	(1.06)
-5,+5	68.25	0.12	0.25	0.34	0.17	28.21	12.38	-4.55	36.75	20.97	3.28
	(28.02)	(0.09)	(0.08)	(0.03)	(0.01)	(4.69)	(7.17)	(1.72)	(1.70)	(16.79)	(1.48)
-5,+10	38.64	0.00	0.15	0.52	0.22	25.96	3.84	-4.26	36.00	1.37	9.45
	(2.97)	(0.01)	(0.12)	(0.08)	(0.05)	(1.94)	(1.61)	(4.18)	(0.42)	(1.16)	(2.09)
-3,+10	38.50	0.006	0.22	0.45	0.21	24.64	5.11	-4.02	35.91	2.26	8.21
	(2.97)	(0.03)	(0.11)	(0.08)	(0.05)	(3.48)	(1.75)	(4.04)	(0.58)	(1.14)	(2.42)
+3,-10	42.10	0.00	0.27	0.54	0.14	24.21	4.12	2.28	35.96	3.46	-9.79
	(4.61)	(0.00)	(0.11)	(0.08)	(0.05)	(3.70)	(1.50)	(4.49)	(0.18)	(1.73)	(1.56)
+5,-10	42.30	0.00	0.18	0.59	0.18	26.07	3.41	6.63	35.95	4.16	-9.03
	(6.42)	(0.00)	(0.14)	(0.10)	(0.05)	(2.21)	(1.28)	(1.77)	(0.31)	(3.88)	(1.28)
+5,-5	156.19	0.001	0.28	0.35	0.21	29.32	14.17	4.12	38.14	22.08	-5.45
	(120.58)	(0.05)	(0.04)	(0.02)	(0.02)	(5.71)	(8.54)	(3.32)	(2.39)	(11.34)	(3.29)
+5,+10	39.30	0.007	0.34	0.40	0.18	25.20	7.11	3.12	37.21	5.14	11.03
	(2.90)	(0.01)	(0.04)	(0.03)	(0.02)	(1.39)	(1.65)	(3.04)	(4.42)	(1.19)	(2.15)

Table B.2. Average of posterior estimates ($E(mode), E(sd.)$) of multiple change-point model parameters m_k and $\theta_{m_3} = (\tau_{3,i}, \delta_{3,i})$, $i = 1, 2, 3$, following signals (RL) from c-chart where $\lambda_{3,0} = 20$, and three consecutive step changes were induced at $\tau_{3,1} = 25$, $\tau_{3,2} = 35$ and $\tau_{3,3} = 45$, respectively. Standard deviations are shown in parentheses.

$\delta_{3,1}$, $\delta_{3,2}$, $\delta_{3,3}$	$E(RL)$	$E(p(m_k))$				$\theta_{m_3,i=1}$			$\theta_{m_3,i=2}$			$\theta_{m_3,i=3}$		
		m_0	m_1	m_2	m_3	$E(\hat{\tau}_{3,1})$	$E(\hat{\sigma}_{\hat{\tau}_{3,1}})$	$E(\hat{\delta}_{3,1})$	$E(\hat{\tau}_{3,2})$	$E(\hat{\sigma}_{\hat{\tau}_{3,2}})$	$E(\hat{\delta}_{3,2})$	$E(\hat{\tau}_{3,3})$	$E(\hat{\sigma}_{\hat{\tau}_{3,3}})$	$E(\hat{\delta}_{3,3})$
-5,+5,-10	52.88	0.00	0.06	0.12	0.56	26.90	2.75	-5.87	36.15	1.71	4.30	45.95	2.56	-9.87
	(6.49)	(0.00)	(0.08)	(0.11)	(0.11)	(2.63)	(1.40)	(2.69)	(1.36)	(0.89)	(2.40)	(0.39)	(1.63)	(1.45)
-5,+5,-5	187.33	0.00	0.18	0.21	0.28	27.61	7.34	-4.65	36.35	6.32	3.31	46.95	8.91	-4.37
	(133.80)	(0.00)	(0.03)	(0.03)	(0.02)	(3.45)	(1.92)	(3.87)	(2.01)	(2.22)	(3.31)	(4.11)	(5.32)	(3.31)
+5,-5,+5	60.05	0.01	0.10	0.16	0.40	26.29	5.27	4.41	36.01	5.61	-4.32	46.41	11.14	5.33
	(11.08)	(0.02)	(0.07)	(0.09)	(0.07)	(2.80)	(1.65)	(4.53)	(0.61)	(3.28)	(1.68)	(3.89)	(4.94)	(1.06)
+5,-5,+10	49.23	0.00	0.06	0.17	0.48	26.15	4.07	3.73	36.03	1.80	-3.90	46.06	1.37	9.32
	(3.59)	(0.01)	(0.09)	(0.12)	(0.10)	(2.27)	(1.43)	(4.66)	(0.84)	(0.82)	(1.61)	(0.43)	(0.71)	(1.96)

Biographies

Hassan Assareh is a Senior Statistician (PhD of Statistics) at Western Sydney Local Health District and a Conjoint Lecturer at University of New South Wales. His primary research interests include Bayesian methods and engineering statistics and its applications in statistical quality control broadly and, in particular, within a health sector.

Rassoul Noorossana is Professor of Industrial Engineering at Iran University of Science and Technology. His primary research interests include statistical process control, process optimization, Six Sigma, and statistical analysis.

Majid Mohammadi is a Research Fellow (PhD of Industrial Engineering) at Islamic Azad University. His primary research interests include Bayesian methods, statistical quality control, and monitoring.

Kerrie Mengersen is Professor of Statistics at Queensland University of Technology. Her research interests are Bayesian statistics, mixture models, hierarchical modelling, and meta-analysis.