



Likelihood based inference for partially observed renewal processes



M.N.M. van Lieshout*

CWI, P.O. Box 94079, NL-1090 GB Amsterdam, The Netherlands

Department of Applied Mathematics, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:

Received 26 January 2016

Received in revised form 28 June 2016

Accepted 4 July 2016

Available online 8 July 2016

Keywords:

Markov chain Monte Carlo

Renewal process

Sequential point process

State estimation

ABSTRACT

This paper is concerned with inference for renewal processes on the real line that are observed in a broken interval. For such processes, the classic history-based approach cannot be used. Instead, we adapt tools from sequential spatial point process theory to propose a Monte Carlo maximum likelihood estimator that takes into account the missing data. Its efficacy is assessed by means of a simulation study and the missing data reconstruction is illustrated on real data.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Inference for point processes on the real line has been dominated by a dynamic approach based on the stochastic intensity (Brémaud, 1972; Karr, 1991; Last and Brandt, 1995) which relates the likelihood of a point at any given time to the history of the process. Such an approach is quite natural in that it is the mathematical translation of the intuitive idea that information becomes available as time passes. Furthermore, the approach allows the utilization of powerful tools from martingale theory, and, since the stochastic intensity is closely related to the hazard functions of the inter-arrival times distributions conditional on the past, a likelihood is immediately available (Daley and Vere-Jones, 2003).

Censoring in the sense of truncation at a random time independently of the point process can be dealt with (Andersen et al., 1993). However, as shown in Section 3, the dynamic approach does not seem capable of dealing with situations in which the flow of time is interrupted. In such cases, combined state estimation techniques are needed that are able to simultaneously carry out inference and reconstruct the missing points.

The aim of this paper is to apply ideas from sequential point process theory (Lieshout, 2006a,b), in particular the Papangelou conditional intensity which describes the probability of finding a point at a particular time conditional on the remainder of the process. Thus, the concept is related to the stochastic intensity, except that the future is taken into account as well as the past. It is this last feature that allows the incorporation of missing data. Moreover, for hereditary point processes at least, the sequential Papangelou conditional intensity defines a likelihood.

We focus on renewal processes on the positive half line and begin by recalling their definition in Section 2. Section 3 shows that a substantial bias may be incurred by the classic approach to parameter estimation when the process is only partially observed. An alternative method based on Markov chain Monte Carlo maximum likelihood (Geyer, 1999) and the sequential Papangelou conditional intensity (Lieshout, 2006b) is proposed in Section 4; Section 5 presents a simulation study

* Correspondence to: Department of Applied Mathematics, University of Twente, P.O. Box 217, NL-7500 AE Enschede, The Netherlands.
E-mail address: M.N.M.van.Lieshout@cwi.nl.

to quantify the bias reduction. It should be noted that the method applies equally to other partially observed point processes. Section 6 illustrates this point by analysing data about calls to a medical helpline.

2. Renewal processes

A renewal process on $[0, T]$, $T > 0$, is defined as follows (Karr, 1991, Chapter 8). Starting at time 0, let U_i be a sequence of (non-defective) independent and identically distributed inter-arrival times with probability density function π , cumulative distribution function F and hazard function h . Set $S_0 = 0$ and $S_i = S_{i-1} + U_i$ for $i \in \mathbb{N}$. Then those S_i , $i \geq 1$, that fall in $(0, T]$ form a simple sequential point process Y (Lieshout, 2006b). For simplicity, we assume that the process starts at time 0, but other initial distributions such as the forward recurrent time may also be accommodated.

Due to the independence assumptions in the model, the stochastic intensity $h^*(\cdot)$ of Y is particularly appealing. Write V_t for the backward recurrence time at t , that is, the difference between t and the last event falling before or at time t . Then, by Karr (1991, Prop. 8.10), $h^*(t) = h(V_{t-})$ and Y admits a density f with respect to $\nu_{[0,T]}$, the distribution of a unit rate Poisson process on $[0, T]$, that can be written as

$$f(t_1, \dots, t_n) = e^T \exp \left[- \int_0^T h^*(t) dt \right] \prod_{i=1}^n h^*(t_i) = e^T (1 - F(T - t_n)) \prod_{i=1}^n \pi(t_i - t_{i-1}) \tag{1}$$

for $(t_1, \dots, t_n) \in H_n([0, T]) = \{(t_1, \dots, t_n) \in [0, T]^n : t_1 < \dots < t_n\}$, cf. Karr (1991, Thm 8.17), under the conventions that an empty product is set to one and that $t_0 = 0$.

By definition, the stochastic intensity h^* is a function of the past of the process. A more versatile concept of conditional intensity is the sequential Papangelou conditional intensity (Lieshout, 2006b) for inserting t at position $k \in \{1, \dots, n + 1\}$ into the vector (t_1, \dots, t_n) . It is defined by

$$\lambda_k(t|t_1, \dots, t_n) = \frac{f(t_1, \dots, t_{k-1}, t, t_k, \dots, t_n)}{f(t_1, \dots, t_n)}$$

whenever both $t_{k-1} < t < t_k$ (with $t_0 = 0$ and $t_{n+1} = \infty$ by convention) and $f(t_1, \dots, t_n) > 0$; it is set to zero otherwise. Note that λ_k depends on both the past and the future. Conversely, provided $f(\cdot)$ is hereditary in the sense that $f(t_1, \dots, t_n) > 0$ implies that $f(s_1, \dots, s_m) > 0$ for all sub-sequences (s_1, \dots, s_m) of (t_1, \dots, t_n) , the factorization

$$f(t_1, \dots, t_n) = f(\emptyset) \prod_{i=1}^n \lambda_i(t_i|t_1, \dots, t_{i-1})$$

holds.

Specializing to renewal processes, note that (1) is not necessarily hereditary, for example when π has small bounded support. Therefore we will assume that $\pi > 0$ and set $t_0 = 0$. Then (1) is hereditary and the sequential Papangelou conditional intensity $\lambda_k(t|t_1, \dots, t_n)$ for $t_1 < \dots < t_n$ reads

$$\lambda_k(t|t_1, \dots, t_n) = \begin{cases} \pi(t - t_{k-1}) \frac{\pi(t_k - t)}{\pi(t_k - t_{k-1})} & \text{if } 1 \leq k \leq n \text{ and } t_{k-1} < t < t_k \\ \pi(t - t_{k-1}) \frac{1 - F(T - t)}{1 - F(T - t_{k-1})} & \text{if } k = n + 1 \text{ and } t_n < t \leq T \end{cases} \tag{2}$$

and zero otherwise. Clearly, (2) depends on the vector (t_1, \dots, t_n) only through the two immediate neighbours t_{k-1} and t_k of t , provided such neighbours exist.

Next, consider the conditional distribution of Y on (T_1, T_2) , $0 < T_1 < T_2 < T$, given $Y \cap [0, T_1] = \vec{\mathbf{r}}$ and $Y \cap [T_2, T] = \vec{\mathbf{s}}$. Then, for $t \in (T_1, T_2)$ and $(t_1, \dots, t_n) \in H_n((T_1, T_2))$,

$$\lambda_k(t|t_1, \dots, t_n; \vec{\mathbf{r}}, \vec{\mathbf{s}}) = \begin{cases} \pi(t - t_{k-1}) \frac{\pi(t_k - t)}{\pi(t_k - t_{k-1})} & \text{if } 1 \leq k \leq n \text{ and } \max(T_1, t_{k-1}) < t < t_k \\ \pi(t - t_{k-1}) \frac{\pi(s_1 - t)}{\pi(s_1 - t_{k-1})} & \text{if } k = n + 1, \vec{\mathbf{s}} \neq \emptyset \text{ and } t_n < t < T_2 \\ \pi(t - t_{k-1}) \frac{1 - F(T - t)}{1 - F(T - t_{k-1})} & \text{if } k = n + 1, \vec{\mathbf{s}} = \emptyset \text{ and } t_n < t < T_2 \end{cases} \tag{3}$$

and zero otherwise. Now, we use the convention that $t_0 = \max(\vec{\mathbf{r}})$ if $\vec{\mathbf{r}} \neq \emptyset$ and $t_0 = 0$ otherwise. As (2), (3) depends on the vector (t_1, \dots, t_n) only through the two immediate neighbours t_{k-1} and t_k of t , provided such neighbours exist. Moreover, assuming that $\vec{\mathbf{s}} \neq \emptyset$, the density of the conditional distribution of Y on (T_1, T_2) with respect to a unit rate Poisson process on (T_1, T_2) is, for $n \in \mathbb{N}$ and $(t_1, \dots, t_n) \in H_n((T_1, T_2))$,

$$f(t_1, \dots, t_n | \vec{\mathbf{r}}, \vec{\mathbf{s}}) = f(\emptyset | \vec{\mathbf{r}}, \vec{\mathbf{s}}) \frac{\pi(t_1 - \max(\vec{\mathbf{r}})) \pi(s_1 - t_n)}{\pi(s_1 - \max(\vec{\mathbf{r}}))} \prod_{i=2}^n \pi(t_i - t_{i-1}) \tag{4}$$

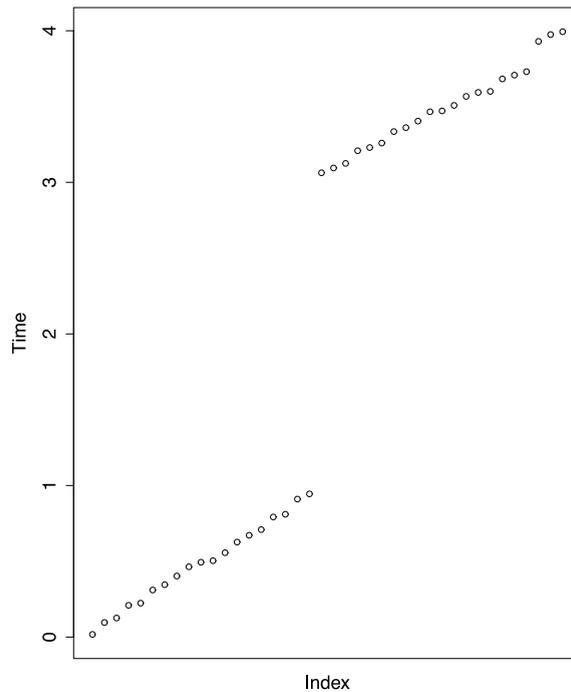


Fig. 1. Sample from a renewal process with Erlang(2)-distributed inter-arrival times with rate parameter $\lambda = 40$ observed in $[0, 1] \cup [3, 4]$. Time is plotted against index number.

under the convention that $\max(\emptyset) = 0$. If $\vec{s} = \emptyset$, the ratio $\pi(s_1 - t_n)/\pi(s_1 - \max(\vec{r}))$ should be replaced by $(1 - F(T - t_n))/(1 - F(T - \max(\vec{r})))$.

A few remarks are in order. Firstly, (3) and (4) depend on \vec{r} and \vec{s} only through $\max(\vec{r})$ and $\min(\vec{s})$, the points adjacent to (T_1, T_2) . Secondly, except when π is the density of an exponential distribution, or, in other words, when Y is a Poisson process, there seems to be no closed form expression for the normalization constant $f(\emptyset|\vec{r}, \vec{s})^{-1}$, that is, for

$$e^{-(T_2-T_1)} \left[1 + \sum_{n=1}^{\infty} \int_{H_n((T_1, T_2))} \frac{\pi(t_1 - \max(\vec{r}))\pi(s_1 - t_n)}{\pi(s_1 - \max(\vec{r}))} \prod_{i=2}^n \pi(t_i - t_{i-1}) dt_1 \cdots dt_n \right].$$

Nevertheless, $e^{-(T_2-T_1)} f(\emptyset|\vec{r}, \vec{s})$ can be interpreted as the conditional probability of finding no points in the interval (T_1, T_2) given the realizations \vec{r} and \vec{s} on $[0, T_1]$, respectively, $[T_2, T]$.

3. Inference based on fully observed inter-arrival times

This section considers the Erlang probability density function

$$\pi(x) = \frac{\lambda^\alpha}{(\alpha - 1)!} x^{\alpha-1} e^{-\lambda x}, \quad x \geq 0,$$

with parameters $\lambda > 0$ and $\alpha \in \mathbb{N}$. Fig. 1 shows a realization of a renewal process on $[0, 4]$ with inter-arrival density π for $\lambda = 40$ and $\alpha = 2$ that is observed within $[0, 1] \cup [3, 4]$.

For renewal processes observed in an unbroken interval $[0, T]$, $T > 0$, Karr (1991, Chapter 8) surveys two approaches to estimate the parameters of π . The first method is to treat the fully observed inter-arrival times $U_i, i = 1, \dots, N(T)$, as a random sample from π and apply maximum likelihood or the method of moments to obtain parameter estimates. The second approach is to use the explicit representation (1) of the likelihood in terms of the stochastic intensity h^* and apply maximum likelihood estimation directly. Note that the first approach applies equally to broken observation intervals but that the second one does not.

Under the Erlang model, if the lengths of the fully observed inter-arrival intervals would constitute a valid random sample, for fixed α we would have an exponential family with sufficient statistic $\sum_i U_i$ and maximum likelihood estimator $\hat{\lambda} = n\alpha / \sum_i U_i$. The parameter α could be estimated by profile likelihood. The problem, though, is that the sample size $n = N(T)$ is random. Nevertheless, Karr (1991) suggests to proceed as if the U_i were a random sample and claims this causes relatively little loss of information.

As an example, the realization consisting of 40 points depicted in Fig. 1 has 38 observable inter-arrival times. Based on this sample, $\hat{\lambda} = 40.91$, whilst the profile likelihood method yields $\hat{\alpha} = 2$. To assess the bias and variance of $\hat{\lambda}$, we generated

Table 1

Classic estimation of λ based on 100 partially observed renewal processes on $[0, 4]$ with Erlang(2)-distributed inter-arrival times with rate parameter $\lambda = 40$. The first column lists the observation window, the second and third columns the empirical mean and standard deviations of $\hat{\lambda}$.

Observation window	Mean	Standard error
$[0, 1] \cup [3, 4]$	42.1	5.0
$[0, 0.25] \cup [0.75, 1]$	47.4	11.7
$[0, 0.0625] \cup [0.1875, 0.25]$	128.1	100.6

100 data patterns on $[0, 4]$ for the parameter values $\lambda = 40$ and $\alpha = 2$. We iteratively excluded the middle of the left-most interval and estimated λ . The results are summarized in Table 1.

We conclude that the bias and variance increase as the observation window contains less complete inter-arrival intervals. The bias occurs as smaller intervals are more likely to fall in the observation window, a phenomenon known as length bias (Karr, 1991). For the smallest observation window, the bias is severe. Indeed, of the samples on the union of two intervals of length $1/16$ each, less than half had observable inter-arrival times at all. Moreover, $1/16$ is only slightly larger than the expected inter-arrival time of $1/20$.

4. Monte Carlo maximum likelihood with missing data

The approach of Section 3 does not make full use of the available data, as it completely ignores partially observed inter-arrival intervals. Here, we adapt the method of Geyer (1999) for dealing with missing data to better account for such intervals.

As before, suppose Y is a renewal process on $[0, T]$ that is observed in $[0, T_1] \cup [T_2, T]$. Write $\pi = \pi_\theta$ for the density of the inter-arrival time distribution, which depends on a parameter θ . Furthermore, denote by \vec{r} the vector of points in $[0, T_1]$ and let \vec{s} contain those in $[T_2, T]$. Writing U for the vector of un-observed points of Y , that is, those that fall in (T_1, T_2) , and Z for the vector of points that are observed, the log likelihood ratio with respect to a given reference parameter θ_0 (Geyer, 1999; Møller and Waagepetersen, 2004) is

$$L(\theta) = \log \mathbb{E}_{\theta_0} \left[\frac{f_\theta(\vec{r}, U, \vec{s})}{f_{\theta_0}(\vec{r}, U, \vec{s})} \mid Z = (\vec{r}, \vec{s}) \right]$$

where f_θ is of the form (1).

In general, the expectation cannot be calculated explicitly and is replaced by an average over a sample from the conditional distribution. Doing so, we obtain the Monte Carlo log likelihood ratio

$$L_N(\theta) = \log \left[\frac{1}{N} \sum_{i=1}^N \frac{f_\theta(\vec{r}, U_i^*, \vec{s})}{f_{\theta_0}(\vec{r}, U_i^*, \vec{s})} \right]. \tag{5}$$

Specifically, the U_i^* form a sample from the conditional distribution of U on (T_1, T_2) given the observed vector $Z = (\vec{r}, \vec{s})$ of points in $[0, T_1] \cup [T_2, T]$ under the reference parameter θ_0 . By taking derivatives, we get the Monte Carlo score and Fisher information. Note that even for exponential families, in the missing data case there may not be a unique maximum likelihood estimator.

To generate a sample $U_i^*, i = 1, \dots, N$, we assume that the sequential Papangelou conditional intensity (3) is uniformly bounded by some $\beta > 0$ and use the Metropolis–Hastings approach (Brooks et al., 2011; Geyer and Møller, 1994; Green, 1995; Lieshout, 2006a; Møller and Waagepetersen, 2004). By Geyer (1999, Prop. 3.3), the algorithm is geometrically ergodic.

An alternative is to use birth-and-death processes (Lieshout, 2006a; Møller, 1989; Preston, 1977; Ripley, 1977), but special care is needed to avoid explosion while ensuring sufficient mixing at the same time. Indeed, the classic birth-and-death process (Preston, 1977) is often implemented by means of a thinning of a birth process whose rate is equal to a global bound on the sequential Papangelou conditional intensity. In our context, since the latter may fluctuate quite a lot, we use adaptive thinning based on local bounds

$$\lambda_k(u|t_1, \dots, t_n; \vec{r}, \vec{s}) \leq g(u|t_1, \dots, t_n; \vec{r}, \vec{s}) \leq \beta$$

for some integrable function g that is constant on $(t_{k-1}, t_k), k = 1, \dots, n + 1$, with an appropriate convention for $k = 1$ and $k = n + 1$. Conditions for geometrically fast convergence can be found in Møller (1989) and Møller and Waagepetersen (2004).

Asymptotic results, including a central limit theorem for the Monte Carlo error, are presented in Geyer (1994). In practice, though, one examines trace plots of carefully chosen statistics (for example, the sufficient statistics in an exponential family model) to assess convergence (Møller and Waagepetersen, 2004).

Table 2

Missing data Monte Carlo maximum likelihood estimation of λ based on 100 partially observed renewal processes on $[0, 4]$ with Erlang(2)-distributed inter-arrival times with rate parameter $\lambda = 40$. The first column lists the observation window, the second and third columns the empirical mean and standard deviations of $\hat{\lambda}$.

Observation window	Mean	Standard error
$[0, 1] \cup [3, 4]$	40.9	4.8
$[0, 0.25] \cup [0.75, 1]$	40.8	8.8
$[0, 0.0625] \cup [0.1875, 0.25]$	40.8	21.4

5. Results

Let us return to the model of Section 3 and the data (\vec{r}, \vec{s}) plotted in Fig. 1. For this model, the parameter θ is the rate λ of the Erlang density function and, for $1 = T_1 < t_1 < \dots < t_n < T_2 = 3$, since $\vec{s} \neq \emptyset$,

$$\frac{f_{\lambda}(\vec{r}, t_1, \dots, t_n, \vec{s})}{f_{\lambda_0}(\vec{r}, t_1, \dots, t_n, \vec{s})} = \left(\frac{\lambda}{\lambda_0}\right)^{(n+n(\vec{r})+n(\vec{s}))\alpha} \frac{e^{-\lambda T} \sum_{i=0}^{\alpha-1} \lambda^i (T - \max(\vec{s}))^i / i!}{e^{-\lambda_0 T} \sum_{i=0}^{\alpha-1} \lambda_0^i (T - \max(\vec{s}))^i / i!}.$$

Hence, to calculate (5) it is sufficient to keep track of $n(U_i^*) + n(\vec{r}) + n(\vec{s})$, the total number of points, during the Monte Carlo simulations. For the data shown in Fig. 1, \vec{r} consists of 19 points, \vec{s} of 21.

In order to implement the Monte Carlo samplers discussed in the previous section, local and global bounds on the sequential Papangelou conditional intensity are needed. For the local bound, recall (3) and note that, for $a < \xi < b$,

$$\frac{\pi(\xi - a)\pi(b - \xi)}{\pi(b - a)} \leq \frac{\lambda^\alpha}{(\alpha - 1)!} \left(\frac{b - a}{4}\right)^{\alpha-1}.$$

Furthermore,

$$\frac{\pi(\xi - a)(1 - F(b - \xi))}{1 - F(b - a)} = \frac{\lambda^\alpha}{(\alpha - 1)!} (\xi - a)^{\alpha-1} \frac{\sum_{i=0}^{\alpha-1} \lambda^i (b - \xi)^i / i!}{\sum_{i=0}^{\alpha-1} \lambda^i (b - a)^i / i!}$$

is bounded from above by

$$\frac{\lambda^\alpha (b - a)^{\alpha-1}}{(\alpha - 1)! \sum_{i=0}^{\alpha-1} \lambda^i (b - a)^i / i!} \sum_{i=0}^{\alpha-1} \left[\frac{\lambda(b - a)}{4}\right]^i / i!.$$

As an aside, the bound can be improved upon in some cases by noting that the function $\phi : \xi \rightarrow \pi(\xi - a)(1 - F(b - \xi))$ increases on (a, b) if $\lambda - (\alpha - 1)/(b - a) < 0$. Finally, since the sub-interval length $b - a \leq T$ is bounded, a global bound and hence local stability follows.

To find a suitable reference parameter λ_0 for $L_N(\lambda)$, we use the Newton–Raphson method (Geyer, 1999). For the data of Fig. 1, this method gives $\lambda_0 = 40.531$. Next, we compute the Monte Carlo log likelihood ratio (5) using $N = 10,000$ samples from a long Metropolis–Hastings run sub-sampled every thousand steps after a thousand steps burn-in. It has a unique maximum at $\hat{\lambda} = 40.36$; the Monte Carlo inverse Fisher information is 20.3.

To assess the bias and variance of the Monte Carlo maximum likelihood estimator, we considered a hundred data patterns as in Section 3. For each pattern, after ten Newton–Raphson steps from the true value $\lambda_0 = 40.0$, $N = 1000$ Monte Carlo samples were obtained by sub-sampling in a long run every thousand steps after a burn-in of a thousand steps to yield a Monte Carlo log likelihood ratio $L_N(\lambda)$, which, in turn, was optimized over the parameter λ . The results are summarized in Table 2; similar results were obtained by the birth-and-death approach.

Compared to the naive approach (cf. Table 1), the bias is much reduced by correctly taking into account missing data and does not increase noticeably for smaller observation intervals. The variance is also reduced in comparison to that of the naive approach but does increase when more data is missing.

6. Extensions and application

The state estimation approach developed in this paper can be applied to point process models other than renewal processes as well. Consider, for example, daily records of calls to NHS Direct, a phone service operated by the National

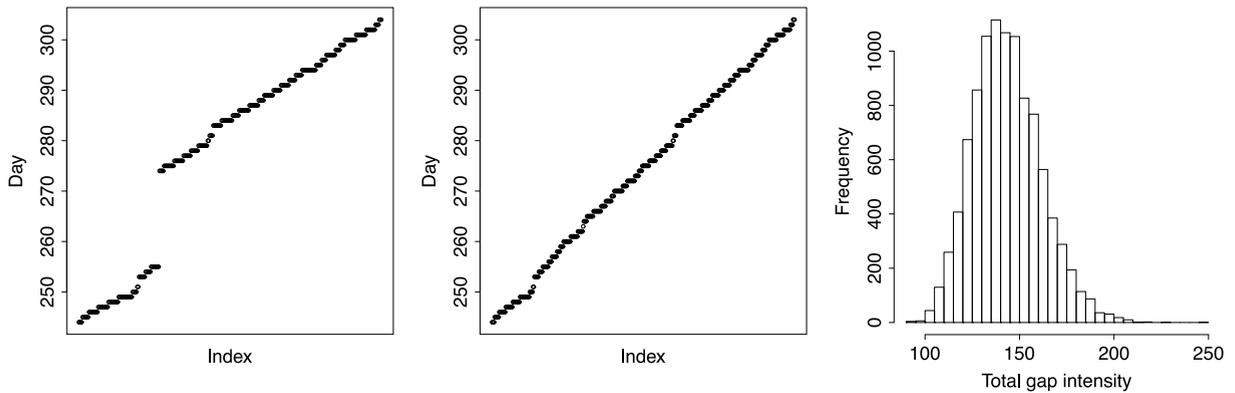


Fig. 2. Dates of calls to NHS Direct in days starting January 1st, 2001, plotted against index number (left panel) with state estimation (middle panel). Right: histogram of the conditional total intensity of calls during September 13–30, 2001, given the observed calls.

Health Service in Britain which people could call 24 h a day to get medical advice.¹ The 327 calls shown in Fig. 2 are those recorded during September and October 2001 that reported acute gastroenteric complaints in the county of Hampshire as downloaded from www.lancaster.ac.uk/staff/diggle/pointpatternbook/datasets/AEGISS.

A salient feature is that no calls were registered during the period September 13–30 with recording resuming on October 1st, 2001. For further details, see Diggle et al. (2005).

As in Brix and Diggle (2001) and Diggle et al. (2005), we assume that calls are made according to a log-Gaussian Cox process (Møller et al., 1998) and approximate the rate of calls during the i th day of the two-month period by $\mu_0(i) \exp[S_T(i)]$, $i = 0, \dots, T = 60$, for S_T that satisfy the Markov recursion

$$S_T(i) = -\frac{\sigma^2}{2} (1 - e^{-\beta}) + e^{-\beta} S_T(i-1) + \sigma \Gamma(i)$$

with initial state $S_T(0) = -\frac{\sigma^2}{2} + \sigma \Gamma(0)$. Here, the $\Gamma(i)$ are independent zero-mean normally distributed random variables having variance $1 - e^{-2\beta}$ for $i > 0$ and 1 for $\Gamma(0)$. The function $\mu_0(\cdot)$ corrects for the increased intensity in calls during weekends and is assumed known (Diggle et al., 2005).

To fill the gap in Fig. 2, note that on a logarithmic scale, the conditional probability density function of the random measure $\Gamma(\cdot)$, up to a constant $c(\beta, \sigma^2, \mu_0, n(i)_i)$, is given by

$$-\frac{1}{2} \gamma(0)^2 - \sum_{i=1}^T \frac{\gamma(i)^2}{2(1 - e^{-2\beta})} + \sum_{i \in \{T_1-6, \dots, T_1, T_2, \dots, T_2+6\}} [n(i) S_\gamma(i) - \mu_0(i) e^{S_\gamma(i)}] \quad (6)$$

where we condition on the counts $n(i)$ for a week on either side of the gap. Since $c(\beta, \sigma^2, \mu_0, n(i)_i)$ cannot be evaluated explicitly, we use the missing data Monte Carlo approach of Section 4 with a Metropolis–Hastings algorithm as in Brix and Diggle (2001) to sample from (6), which, in turn, can be used to sample the missing counts $N(i)$, $i \in \{T_1 + 1, \dots, T_2 - 1\}$, during September 13–30 as independent Poisson random variables with rate $\mu_0(i) \exp[S_\gamma(i)]$. We ran the chain for 10,000 steps and obtained the result shown in the central panel of Fig. 2. The parameter estimates based on the completed data were $\hat{\sigma}^2 = 0.11$ and $\hat{\beta} = 0.93$ which should be compared to the estimates $\hat{\sigma}^2 = 0.12$ and $\hat{\beta} = 0.55$ obtained by a naive approach. Finally, we extended the run, sub-sampling every 1000 steps, to obtain 10,000 realizations and calculated the histogram of the total intensity in the gap period shown in the right-most panel of Fig. 2. In conclusion, around 150 calls may have been missed.

References

- Andersen, P., Borgan, Ø., Gill, R., Keiding, N., 1993. *Statistical Models based on Counting Processes*. Springer, New York.
- Brémaud, P., 1972. *A martingale approach to point processes* (Ph.D. dissertation), University of Berkeley.
- Brix, A., Diggle, P.J., 2001. Spatiotemporal prediction for log-Gaussian Cox processes. *J. R. Stat. Soc. Ser. B* 63, 823–841.
- Brooks, S., Gelman, A., Jones, G., Meng, X.-L., 2011. *Handbook of Markov Chain Monte Carlo*. In: *Handbooks of Modern Statistical Methods*, Chapman & Hall/CRC Press, Boca Raton.
- Daley, D.J., Vere-Jones, D., 2003. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, second ed.. Springer, New York.

¹ Data were provided by Professor Peter Diggle, Lancaster University, and Dr Peter Hawtin, Health Protection Agency, Southampton Laboratory. The data were derived from anonymized data-sets collected during project AEGISS, a collaborative surveillance project based in the south of England. The project was financially supported by the Food Standards Agency.

- Diggle, P.J., Rowlingson, B., Su, T.-L., 2005. Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* 16, 423–434.
- Geyer, C.J., 1994. On the convergence of Monte Carlo maximum likelihood calculations. *J. R. Stat. Soc. Ser. B* 56, 261–274.
- Geyer, C., 1999. Likelihood inference for spatial point processes. In: Barndorff-Nielsen, O.E., Kendall, W.S., Lieshout, M.N.M. van (Eds.), *Stochastic Geometry. In: Monographs in Statistics and Applied Probability*, Vol. 80. Chapman & Hall/CRC Press, Boca Raton, pp. 79–140.
- Geyer, C.J., Møller, J., 1994. Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Stat.* 21, 359–373.
- Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82, 711–732.
- Karr, A.F., 1991. *Point Processes and their Statistical Inference*, second ed.. Marcel Dekker, New York.
- Last, G., Brandt, A., 1995. *Marked Point Processes on the Real Line. The Dynamic Approach*. Springer, New York.
- Lieshout, M.N.M. van, 2006a. Markovianity in space and time. In: Denteneer, D., Den Hollander, F., Verbitskiy, E. (Eds.), *Dynamics & Stochastics: Festschrift in Honour of M.S. Keane*. In: *Lecture Notes–Monograph Series*, Vol. 48. Institute for Mathematical Statistics, Beachwood, pp. 154–168.
- Lieshout, M.N.M. van, 2006b. Campbell and moment measures for finite sequential spatial processes. In: Hušková, M., Janžura, M. (Eds.), *Proceedings Prague Stochastics 2006*. Matfyzpress, Prague, pp. 215–224.
- Møller, J., 1989. On the rate of convergence of spatial birth-and-death processes. *Ann. Inst. Statist. Math.* 3, 565–581.
- Møller, J., Syversveen, A.R., Waagepetersen, R.P., 1998. Log Gaussian Cox processes. *Scand. J. Stat.* 25, 451–482.
- Møller, J., Waagepetersen, R.P., 2004. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton.
- Preston, C.J., 1977. Spatial birth-and-death processes. *Bull. Inst. Internat. Statist.* 46, 371–391.
- Ripley, B.D., 1977. Modelling spatial patterns. *J. R. Stat. Soc. Ser. B* 39, 172–212.