0.4pt=0pt

Detection and replenishment of missing data in marked point processes

[1)]Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

[2)]Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand

[3)]Organization of Advanced Science and Technology, Kobe University

1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

Supplementary Material

This online Supplementary Material includes the following topics:

(1) a proof of the existence of a solution to the equation system in (21) to (24);

(2) the asymptotic properties of the solution;

(3) additional simulations for the case in which the missing region is wrongly specified;

(4) list of the history record of the Hakone volcano; and

(5) comments on the Wenchuan aftershock sequence.

S1 Convergency of the iterative algorithm

The existence of a solution to equations (21) to (24) is ensured by the

Schauder fixed-point theorem: Every continuous mapping from a convex compact subset K of a Banach space to K itself has a fixed point. First, for the given observed point process $N_{\rm obs}$ in $[0,T]\times M$, consider the space consisting of the following stepwise constant function pairs:

$$f(t) = \begin{cases} \tilde{f}_0, & t < t_1; \\ \tilde{f}_i, & t_i \le t < t_{i+1}, \quad i = 1, 2, \cdots, n; \\ \tilde{f}_n, & t_n \le t < T; \end{cases} \tag{S1.1}$$

$$g(m) = \begin{cases} \tilde{g}_0, & m < m_{(1)}; \\ \tilde{g}_i, & m_{(i)} \le m < m_{(i+1)}, \quad i = 1, 2, \cdots, n; \ , \\ \tilde{g}_n, & m \ge m_{(n)}; \end{cases} \tag{S1.2}$$

where $\tilde{f}_0, \cdots, \tilde{f}_n, \tilde{g}_0, \cdots, \tilde{g}_n$ are constants, and $m_{(i)}$ are the ith-order statistics for $\{m_1, m_2, \cdots, m_n\}$. When equipped with a metric

$$d[(f_1, g_1), (f_2, g_2)] = \sup_{x\in[0,T]} |f_1(x) - f_2(x)| + \sup_{x\in M} |g_1(x) - g_2(x)|,$$

such a function space is complete and equivalent to a finite dimensional Euclidean space. The function pairs that satisfy

$$0 = \tilde{f}_0 \le \tilde{f}_1 \le \tilde{f}_2 \le \cdots \le \tilde{f}_n = 1, \quad 0 = \tilde{g}_0 \le \tilde{g}_1 \le \tilde{g}_2 \le \cdots \le \tilde{g}_n = 1$$

form a convex compact subset X of this space, because X is complete and totally bounded. When the missing area S is fixed, we have the following mapping:

$$\Gamma(f,g) \equiv \left(\frac{\sum_{j=1}^{n} w_1(t_j,m_j,S)\mathbf{1}\left(t_j < t\right)}{\sum_{j=1}^{n} w_1(t_j,m_j,S)}, \frac{\sum_{j=1}^{n} w_2(t_j,m_j,S)\mathbf{1}\left(m_j < m\right)}{\sum_{j=1}^{n} w_2(t_j,m_j,S)} \right), \tag{S1.3}$$

where

$$w_1(t,m,S) \equiv \frac{\mathbf{1}\left((t,m) \notin S\right)}{\int_M \mathbf{1}\left((t,m') \notin S\right) \mathrm{d}g(m')} \tag{S1.4}$$

$$w_2(t,m,S) \equiv \frac{\mathbf{1}\left((t,m) \notin S\right)}{\int_0^T \mathbf{1}\left((t',m) \notin S\right) \mathrm{d}f(t')} \tag{S1.5}$$

is a continuous mapping from X to X. By the Schauder fixed-point theorem, the solution to equations (21) to (24) exists. Denote this by $(F^*(t), G^*(m))$, which is in fact a mapping from $[0,T] \times M$ to $[0,1] \times [0,1]$.

S2 Asymptotic properties of the solution

Consider the expectation of F^* and G^* with respect to all the possibilities of the point process N,

$$\mathbf{E}\left[F^*(t)\right] = \mathbf{E}\left[\frac{\sum_{j=1}^{n} w_1(t_j,m_j,S)\mathbf{1}\left(t_j < t\right)}{\sum_{j=1}^{n} w_1(t_j,m_j,S)} \right], \tag{S2.6}$$

$$\mathbf{E}\left[G^*(m)\right] = \mathbf{E}\left[\frac{\sum_{j=1}^{n} w_2(t_j,m_j,S)\mathbf{1}\left(m_j < m\right)}{\sum_{j=1}^{n} w_2(t_j,m_j,S)} \right], \tag{S2.7}$$

and set

$$U(t) = \sum_{j=1}^{n} w_1(t_j,m_j,S)\mathbf{1}\left(t_j < t\right) \tag{S2.8}$$

and

$$V(m) = \sum_{j=1}^{n} w_2(t_j, m_j, S)\mathbf{1}\,(m_j < m). \tag{S2.9}$$

By the delta method, asymptotically,

$$\begin{pmatrix} F^*(t) \\ \\ G^*(m) \end{pmatrix} \xrightarrow{D} \mathbf{N}\left(\begin{array}{c} \dfrac{\mathbf{E}[U(t)]}{\mathbf{E}[U(T)]} \\ \dfrac{\mathbf{E}[V(m)]}{\mathbf{E}[V(\infty)]} \end{array}, \; J^T\, W\, J \right) \tag{S2.10}$$

where

$$J = \begin{pmatrix} \dfrac{\partial F^*(t)}{\partial U(t)} & 0 \\ \dfrac{\partial F^*(t)}{\partial U(T)} & 0 \\ 0 & \dfrac{\partial G^*(m)}{\partial V(m)} \\ 0 & \dfrac{\partial G^*(m)}{\partial V(\infty)} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{U(T)} & 0 \\ -\dfrac{U(t)}{[U(T)]^2} & 0 \\ 0 & \dfrac{1}{V(\infty)} \\ 0 & -\dfrac{V(m)}{[V(\infty)]^2} \end{pmatrix} \tag{S2.11}$$

and

$$W = \begin{pmatrix} \mathbf{Var}[U(t)] & \mathbf{Cov}[U(t), U(T)] & 0 & 0 \\ \mathbf{Cov}[U(t), U(T)] & \mathbf{Var}[U(T)] & 0 & 0 \\ 0 & 0 & \mathbf{Var}[V(m)] & \mathbf{Cov}[V(m), V(\infty)] \\ 0 & 0 & \mathbf{Cov}[V(m), V(\infty)] & \mathbf{Var}[V(\infty)] \end{pmatrix}. \tag{S2.12}$$

Denote as $\mu(t, m)$ and $\mu_2(t, m, t', m')$ the first- and second-order moment intensities of the complete process, respectively. By the separability of the mark distribution, the first- and second-order moment intensities can be written as $\mu = \mu_g(t)g(m)$ and $\mu_2(t, m, t', m') = \mu_{2g}(t, t')g(m)g(m')$.

We also denote $A(t) \equiv \int_0^t \mu_g(t')\,\mathrm{d}t'$, $c(t,t') \equiv \mu_{2g}(t,t') - \mu_g(t)\mu_g(t')$, and $C(t,t') \equiv \int_0^{t'}\int_0^t c(u,u')\,\mathrm{d}u\mathrm{d}u'$ (the separability of the mark distribution in the conditional intensity implies its separability in moment intensities). By the Campbell theorem (e.g., Møller and Waagepetersen, 2003; Daley and Vere-Jones, 2008),

$$\begin{aligned}
\mathbf{E}[U(t)] &= \mathbf{E}\left[\int_{[0,T]\times M} \frac{\mathbf{1}\left((t',m') \notin S\right)\mathbf{1}\left(t' < t\right)}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\,\mathrm{d}G^*(m'')} N(dt' \times \mathrm{d}m')\right] \\
&= \mathbf{E}\left[\int_{[0,t)\times M} \frac{\mathbf{1}\left((t',m') \notin S\right)\,\mu_g(t')\,g(m')dt'\mathrm{d}m'}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\mathrm{d}G^*(m'')}\right] \\
&= \mathbf{E}\left[\int_0^t \mu_g(t')\,\frac{\int_M \mathbf{1}\left((t',m') \notin S\right)\,\mathrm{d}G(m')}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\mathrm{d}G^*(m'')} dt'\right]. \qquad \text{(S2.13)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbf{E}[V(m)] &= \mathbf{E}\left[\int_{[0,T]\times M} \frac{\mathbf{1}\left((t',m') \notin S\right)\mathbf{1}\left(m' < m\right)N(dt' \times \mathrm{d}m')}{\int_0^T \mathbf{1}\left((t'',m') \notin S\right)\mathrm{d}F^*(t'')}\right] \\
&= \mathbf{E}\left[\int_{[0,T]\times[0,m)} \frac{\mathbf{1}\left((t',m') \notin S\right)\,\mu_g(t')\,g(m')\,dt'\mathrm{d}m'}{\int_M \mathbf{1}\left((t'',m') \notin S\right)\mathrm{d}F^*(t'')}\right] \\
&= \mathbf{E}\left[\int_0^m g(m')\,\frac{\int_0^T \mathbf{1}\left((t',m') \notin S\right)\,\mu_g(t')\,dt'}{\int_0^T \mathbf{1}\left((t'',m') \notin S\right)\,dF^*(t'')}\,\mathrm{d}m'\right]. \qquad \text{(S2.14)}
\end{aligned}$$

Substitute (S2.13) and (S2.14) into (S2.10). Then, we have

$$\mathbf{E}\left[F^*(t)\right] \to \frac{\mathbf{E}\left[\int_0^t \mu_g(t')\,\frac{\int_M \mathbf{1}\left((t',m') \notin S\right)\,\mathrm{d}G(m')}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\mathrm{d}G^*(m'')}\mathrm{d}t'\right]}{\mathbf{E}\left[\int_0^T \mu_g(t')\,\frac{\int_M \mathbf{1}\left((t',m') \notin S\right)\,\mathrm{d}G(m')}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\mathrm{d}G^*(m'')}\mathrm{d}t'\right]}, \qquad \text{(S2.15)}$$

$$\mathbf{E}\left[G^*(m)\right] \to \frac{\mathbf{E}\left[\int_0^m g(m')\,\frac{\int_0^T \mathbf{1}\left((t',m') \notin S\right)\,\mu_g(t')\,\mathrm{d}t'}{\int_0^T \mathbf{1}\left((t'',m') \notin S\right)\,\mathrm{d}F^*(t'')}\mathrm{d}m'\right]}{\mathbf{E}\left[\int_M g(m')\,\frac{\int_0^T \mathbf{1}\left((t',m') \notin S\right)\,\mu_g(t')\,\mathrm{d}t'}{\int_0^T \mathbf{1}\left((t'',m') \notin S\right)\,\mathrm{d}F^*(t'')}\mathrm{d}m'\right]}, \qquad \text{(S2.16)}$$

where $\to$ represents "converging to in probability." From these equations, we have that

$$\mathbf{E}\left[F^*(t)\right] \to \frac{\int_0^t \mu_g(t')\,\mathrm{d}t'}{\int_0^T \mu_g(t')\,\mathrm{d}t'} = \frac{A(t)}{A(T)}, \tag{S2.17}$$

$$\mathbf{E}\left[G^*(m)\right] \to \int_0^m g(m')\,\mathrm{d}m' = G(m), \tag{S2.18}$$

which yield

$$\mathbf{E}[U(t)] \to \int_0^t \mu_g(t')\,\mathrm{d}t' = A(t) \tag{S2.19}$$

$$\begin{aligned}\mathbf{E}[V(m)] &\to \mathbf{E}[U(T)]\int_0^m g(m')\,\mathrm{d}m' \\ &= A(T)\,G(m). \end{aligned} \tag{S2.20}$$

Now, consider the variance. By the Campbell theorem for higher-order moment intensities, if $t \leq s$, then

$$\begin{aligned}
& \mathbf{E}\left[U(t)U(s)\right] \\
= \;& \mathbf{E}\Bigg[\left(\int_{[0,T]\times M} \frac{\mathbf{1}\left((t',m') \notin S\right)\mathbf{1}\left(t' < t\right)}{\int_M \mathbf{1}\left((t',m''') \notin S\right)\, \mathrm{d}G^*(m''')} N(\mathrm{d}t' \times \mathrm{d}m')\right) \\
& \quad \times \left(\int_{[0,T]\times M} \frac{\mathbf{1}\left((t'',m'') \notin S\right)\mathbf{1}\left(t'' < s\right)}{\int_M \mathbf{1}\left((t'',m''') \notin S\right)\, \mathrm{d}G^*(m''')} N(\mathrm{d}t'' \times \mathrm{d}m'')\right)\Bigg] \\
= \;& \mathbf{E}\left[\int_M \int_0^T \left(\frac{\mathbf{1}\left((t',m') \notin S\right)\mathbf{1}\left(t' < t\right)}{\int_M \mathbf{1}\left((t',m''') \notin S\right)\, \mathrm{d}G^*(m''')}\right)^2 \mu_g(t')\, g(m')\, \mathrm{d}t'\mathrm{d}m'\right] \\
& + \mathbf{E}\Bigg[\int_M \int_0^T \int_M \int_0^T \frac{\mathbf{1}\left((t',m') \notin S\right)\mathbf{1}\left(t' < t\right)}{\int_M \mathbf{1}\left((t',m''') \notin S\right)\, \mathrm{d}G^*(m''')} \frac{\mathbf{1}\left((t'',m'') \notin S\right)\mathbf{1}\left(t'' < s\right)}{\int_M \mathbf{1}\left((t'',m''') \notin S\right)\, \mathrm{d}G^*(m''')} \\
& \qquad \times \mu_{2g}(t',t'')\, g(m)\, g(m'')\mathrm{d}t'\mathrm{d}m'\mathrm{d}t''\mathrm{d}m''\Bigg] \\
= \;& \mathbf{E}\left[\int_0^t \mu_g(t') \frac{\int_M \mathbf{1}\left((t',m') \notin S\right)\, \mathrm{d}G(m')}{\left(\int_M \mathbf{1}\left((t',m''') \notin S\right)\, \mathrm{d}G^*(m''')\right)^2}\, \mathrm{d}t'\right] \\
& + \mathbf{E}\left[\int_0^t \int_0^s \mu_{2g}(t',t'') \frac{\int_M \mathbf{1}\left((t',m') \notin S\right)\, \mathrm{d}G(m')}{\int_M \mathbf{1}\left((t',m''') \notin S\right)\, \mathrm{d}G^*(m''')} \frac{\int_M \mathbf{1}\left((t'',m'') \notin S\right)\, \mathrm{d}G(m'')}{\int_M \mathbf{1}\left((t'',m''') \notin S\right)\, \mathrm{d}G^*(m''')}\, \mathrm{d}t'\, \mathrm{d}t''.\right]
\end{aligned} \tag{S2.21}$$

Again, by applying the delta method to the expectation part only, we have

$$\mathbf{E}\left[U(t)U(s)\right] \to \int_0^t \frac{\mu_g(t')\, \mathrm{d}t'}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\, \mathrm{d}G(m'')} + \int_0^t \int_0^s \mu_{2g}(t',t'')\, \mathrm{d}t'\, \mathrm{d}t''. \tag{S2.22}$$

Set

$$Q(t) \equiv \int_0^t \frac{\mu_g(t')\, \mathrm{d}t'}{\int_M \mathbf{1}\left((t',m'') \notin S\right)\, \mathrm{d}G(m'')}.$$

Then, the above equation gives

$$\begin{aligned}
\mathbf{Var}\left[U(t)\right] &= \mathbf{E}\left[U(t)U(t)\right] - \mathbf{E}\left[U(t)\right]\mathbf{E}\left[U(t)\right] \\
&\to \int_0^t \frac{\mu_g(t')\,\mathrm{d}t'}{\int_M \mathbf{1}\left((t',m'')\notin S\right)\,\mathrm{d}G(m'')} + \int_0^t \int_0^t \left[\mu_{2g}(t',t'') - \mu_g(t')\,\mu_g(t'')\right]\,\mathrm{d}t'\,\mathrm{d}t'', \\
&= \int_0^t \frac{\mu_g(t')\,\mathrm{d}t'}{\int_M \mathbf{1}\left((t',m'')\notin S\right)\,\mathrm{d}G(m'')} + \int_0^t \int_0^t c(t',t'')\,\mathrm{d}t'\,\mathrm{d}t'' \\
&= Q(t) + C(t,t), \qquad \text{(S2.23)}
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{Cov}\left[U(t),U(T)\right] &= \mathbf{Var}\left[U(t)\right] + \mathbf{Cov}\left[U(t),U(T)-U(t)\right] \\
&\to \int_0^t \int_0^T c(t',t'')\,\mathrm{d}t'\,\mathrm{d}t'' + \int_0^t \frac{\mu_g(t')\,\mathrm{d}t'}{\int_M \mathbf{1}\left((t',m'')\notin S\right)\,\mathrm{d}G(m'')} \\
&= Q(t) + C(t,T). \qquad \text{(S2.24)}
\end{aligned}$$

Similarly, for high-order moments related to $V(\cdot)$, we can obtain the following for $m \le u$:

$$\begin{aligned}
\mathbf{Cov}\left[V(m),V(u)\right] &\to G(m)\,G(u)\left(\int_0^T \mu_g(t')\,\mathrm{d}t'\right)^2 \int_0^T\!\!\int_0^T c(t',t'')\,\mathrm{d}t'\,\mathrm{d}t'' \\
&\quad + \left(\int_0^T \mu_g(t')\,\mathrm{d}t'\right)^2 \int_0^m \frac{g(m')\,\mathrm{d}m'}{\int_0^T \mathbf{1}\left((t'',m')\notin S\right)\,\mu_g(t'')\mathrm{d}t''} \\
&= A^2(T)\left[G(m)\,G(u)\,C(T,T) + \int_0^m \frac{g(m')\,\mathrm{d}m'}{\int_0^T \mathbf{1}\left((t'',m')\notin S\right)\,\mu_g(t'')\mathrm{d}t''}\right] \\
&= A^2(T)\,G(m)\,G(u)\,C(T,T) + A(T)\,P(m), \qquad \text{(S2.25)}
\end{aligned}$$

where

$$P(m) \equiv A(T)\int_0^m \frac{g(m')\,\mathrm{d}m'}{\int_0^T \mathbf{1}\left((t'',m')\notin S\right)\,\mu_g(t'')\mathrm{d}t''}.$$

Combining (S2.11), (S2.12), (S2.19), (S2.23), and (S2.24), we have

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}, \tag{S2.26}$$

where

$$\begin{aligned} \sigma_1^2 &= \frac{C(t,t)}{A^2(T)} - 2\frac{C(T,t)\,A(t)}{A^3(T)} + \frac{A^2(t)\,C(T,T)}{A^4(T)} \\ &\quad + \frac{A(T) - 2A(t)}{A^3(T)}\,Q(t) + \frac{A^2(t)\,Q(T)}{A^4(T)} \end{aligned} \tag{S2.27}$$

and

$$\sigma_2^2 = \frac{[1 - 2G(m)]\,P(m) + G^2(m)P(\infty)}{A(T)}. \tag{S2.28}$$

From the above results, and adding the conditions that guarantee the existence of related integral and division operations, we obtain the following asymptotic theorem for the estimates in (21) to (24).

Theorem 1. *Consider a point process N with continuous marks that satisfies the following conditions:*

1. *The mark distribution is separable (i.e., $\mu(t,m) = \mu_g(t)\,g(m)$);*

2. *The ground intensity is bounded (i.e., there exist two positive numbers K_1 and K_2, such that $K_1 \le \lambda_g(t) \le K_2$, for all t;*

3. *The covariance density $c(t,t') = \mu_{2g}(t,t') - \mu_g(t)\,\mu_g(t')$ satisfies*

$$\int_0^T \int_0^T |c(t,t')|\,\mathrm{d}t\,\mathrm{d}t' < \infty,$$

and

$$\int_0^\infty |c(t,t')|\,\mathrm{d}t' < K$$

for each t and some constant K;

4. *The missing area S satisfies*

$$\int_M \mathbf{1}\left((t,m) \notin S\right)\,\mathrm{d}G(m) > 0$$

for all $t \in [0,T]$ and

$$\int_0^T \mathbf{1}\left((t,m) \notin S\right)\,\mu_g(t)\,\mathrm{d}t > 0$$

for all $m \in M$.

Then, the solutions to equations (21) to (24), $F^*(t)$ and $G^*(m)$, satisfy

$$\sqrt{A(T)}\begin{pmatrix} F^*(t) - \frac{A(t)}{A(T)} \\ \\ G^*(m) - G(m) \end{pmatrix} \xrightarrow{D} \mathbf{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_f^2 & 0 \\ \\ 0 & \sigma_g^2 \end{pmatrix}\right) \tag{S2.29}$$

for fixed $(t,m) \in [0,T] \times M$, where

$$\begin{aligned} \sigma_f^2 \;=\; & \frac{C(t,t)}{A(T)} - 2\frac{C(T,t)\,A(t)}{A^2(T)} + \frac{A^2(t)\,C(T,T)}{A^3(T)} \\ & + \frac{A(T) - 2A(t)}{A^2(T)}\,Q(t) + \frac{A^2(t)}{A^3(T)}\,Q(T) \end{aligned} \tag{S2.30}$$

and

$$\sigma_g^2 \;=\; [1 - 2G(m)]\,P(m) + G^2(m)\,P(\infty). \tag{S2.31}$$

In other words, both processes of $\xi_1(t) = \sqrt{A(T)}[F^*(t) - A(t)/A(T)]$ *and* $\xi_2(m) = \sqrt{A(T)}[G^*(m) - G(m)]$ *converge in distribution to some Brownian bridges with covariance functions*

$$\begin{aligned} & \mathbf{Cov}\left\{\xi_1(t), \xi_1(s)\right\} \\ = \; & \frac{A^2(s)[Q(t)+C(t,t)] + 2A(s)A(t)[Q(t)+C(t,s)] + A^2(t)[Q(s)+C(s,T)]}{A^3(T)} \\ & - \frac{4A(t)A^2(s)[Q(t)+C(t,T)] + 4A^2(t)A(s)[Q(s)+C(s,T)]}{A^4(T)} \\ & + \frac{4A^2(t)A^2(s)[Q(T)+C(T,T)]}{A^5(T)}, \end{aligned} \tag{S2.32}$$

for $t \leq s$, *and*

$$\begin{aligned} & \mathbf{Cov}\left\{\xi_2(m), \xi_2(u)\right\} \\ = \; & G(m)\,G(u)\left\{P(m)\,[3 - 4G(u)] + P(u)\,[1 - 4G(m)] + 4P(\infty)\,G(u)\,G(m)\right\}, \end{aligned} \tag{S2.33}$$

for $m \leq u$.

Proof: Equalities (S2.29) to (S2.31) can be proved from the earlier discussions. Here, we only prove (S2.32) and (S2.33).

Denoting $Z(t,s) \equiv F^*(t)\,F^*(s)$, and observing

$$Z(t,s) = \frac{U(t)\,U(s)}{U^2(T)} \tag{S2.34}$$

and using the delta method, for fixed t and s,

$$Z(t,s) \xrightarrow{D} \mathbf{N}\left(\frac{A(t)A(s)}{A^2(T)}, K^T \Omega K\right), \tag{S2.35}$$

where

$$K = \begin{pmatrix} \dfrac{\partial Z(t,s)}{\partial U(t)} \\ \dfrac{\partial Z(t,s)}{\partial U(s)} \\ \dfrac{\partial Z(t,s)}{\partial U(T)} \end{pmatrix} = \begin{pmatrix} \dfrac{U(s)}{U^2(T)} \\ \dfrac{U(t)}{U^2(T)} \\ -\dfrac{2\,U(t)\,U(s)}{U^3(T)} \end{pmatrix} \tag{S2.36}$$

and

$$\Omega = \begin{pmatrix} \mathbf{Var}[U(t)] & \mathbf{Cov}[U(t),U(s)] & \mathbf{Cov}[U(t),U(T)] \\ \mathbf{Cov}[U(t),U(s)] & \mathbf{Var}[U(s)] & \mathbf{Cov}[U(s),U(T)] \\ \mathbf{Cov}[U(t),U(T)] & \mathbf{Cov}[U(s),U(T)] & \mathbf{Var}[U(T)] \end{pmatrix}. \tag{S2.37}$$

Substituting (S2.23) and (S2.24) into the above equation and considering the definitions of ξ_1, we obtain (S2.32). Equality (S2.33) can be proved in a similar way.

S3 How does the misspecification of the missing region S influence the replenishment results?

In this replenishment algorithm, the key point is the specification of the region S that contains the missing events. To check whether S is appropriately specified we can inspect whether the output point process N^*_{obs} in

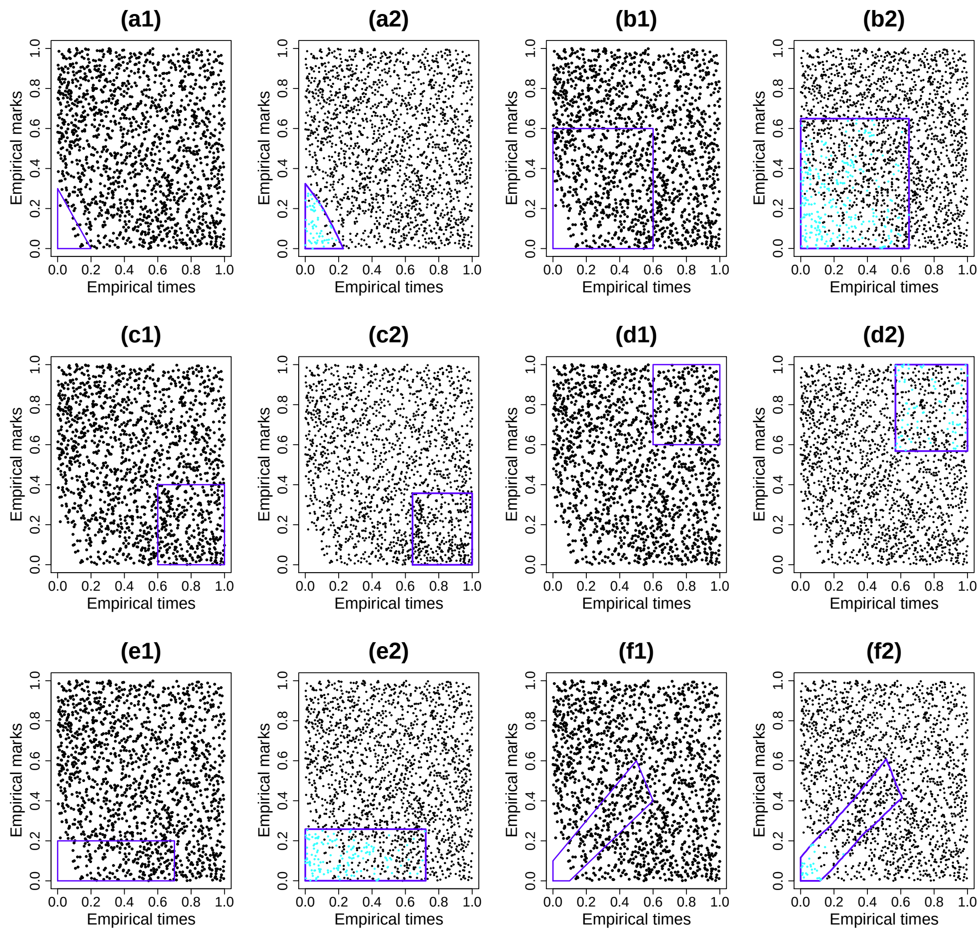

Figure S1: Outputs of $S^{(1)}$ and S^* in examinations of the replenishment algorithm when the missing data region is misspecified. (a) when S is too small. (b) when S is too large. (c) and (d) when S is completely misspecified. (e) and (f) when S is partially misspecified. In each pair of panels, the first one displays $S^{(1)}$ and the transformation of N_{obs} under $\Gamma_{N_{\text{obs}}}$, and the second displays S^* and N^*_{obs}. The blue dots mark the empirical time-mark locations of the simulated events. See text in Section 2.3 for detail of notations.

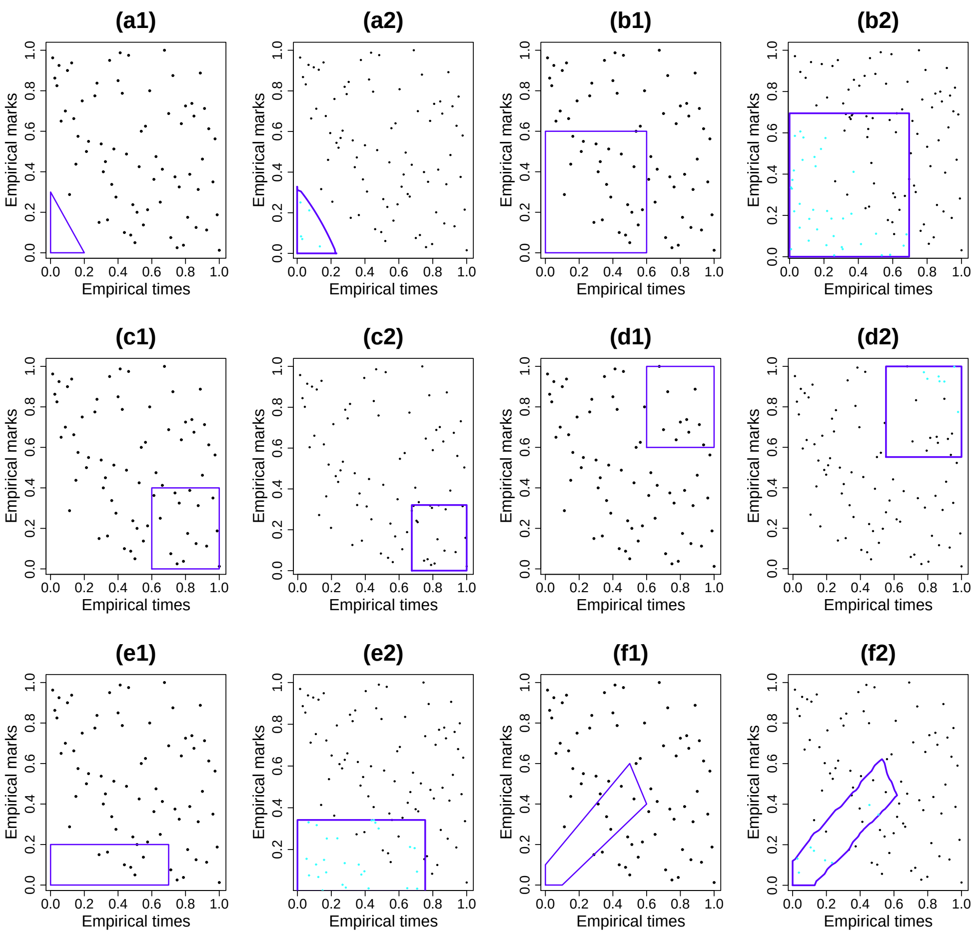

Figure S2: Same as Figure S2 but a small sample size.

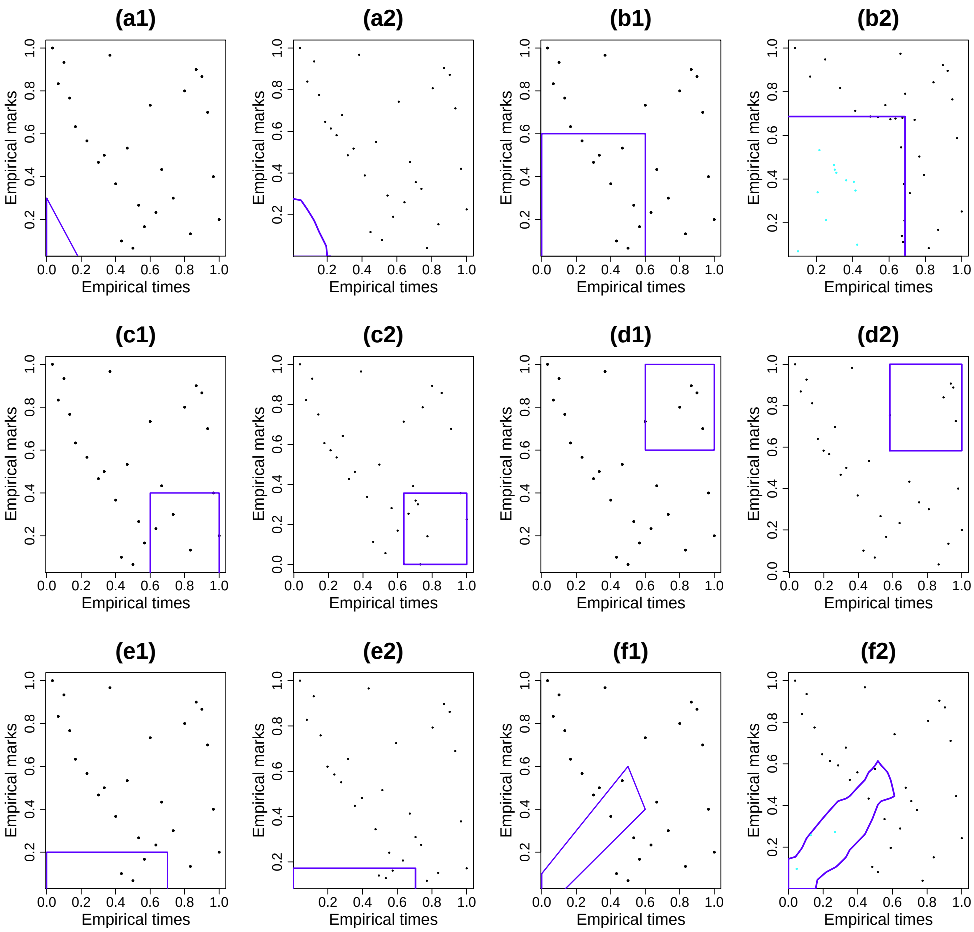

Figure S3: Same as Figure S2 but with a sample size even smaller than in Figure S2.

Step 2 is homogenously distributed outside of S^*. We test four cases of the misspecification of S: (a) too small, (b) too large, (c) completely misspecified, and (d) partially misspecified. Figures S1 to S3 give several examples where S is misspecified. We can see that, when S is too small (Panel a in Figures S1 to S3), does not contains the missing data (Panels c and d in Figures S1 to S3), or only partially covers the missing data (Panels e and f in Figures S1 to S3), N^*_{obs} is not homogeneous outside of S. Such inhomogeneity can be detected by the R or D statistics defined in (6). When the specified region $S^{(1)}$ is larger than necessary but does not break Condition 1 (Figure S1b), N^*_{obs} is still homogeneous outside of S. In this case, the replenishment algorithm can also give unbiased replenishment results but with higher uncertainty. For the case in Figure S1c, the algorithm halts at Step 5 because there are always more events in S^* from N^*_{obs} than from N^*_{rep}, i.e., not enough points to remove.

S4 Hakone Vocalno data

The estimated volcano eruption indices (VEI) and ages for historical eruptions of the Hakone Volcano are listed Table S1.

S5 More on Wenchuan earthquake aftershock sequence

An epicenter map of the observed earthquakes in the catalog is shown in Figure S4.

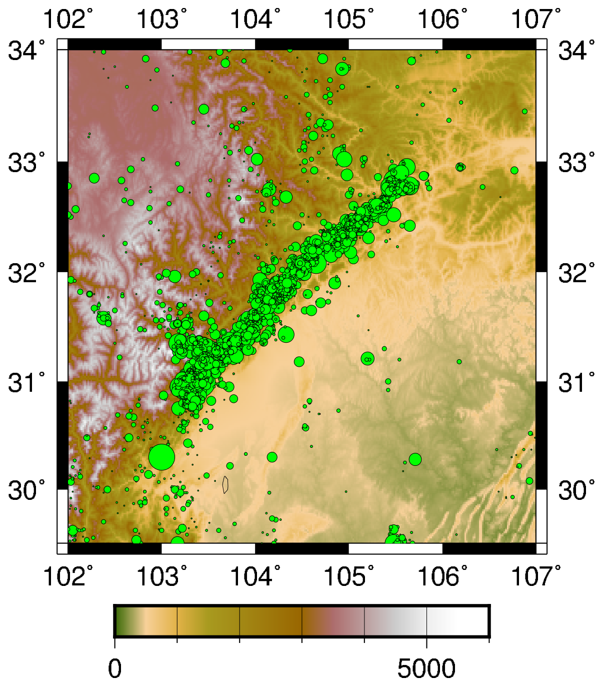

Figure S4: Epicenter map of earthquakes in the Wenchuan area.

We also fit the Omori formula to the original dataset, but only consider earthquakes that occurred at least 54 days after the mainshock. In this case,$c \approx 0$, implies that the selected data, which starts from 54 days after the mainshock, cannot be used to estimate the the temporal pattern of seismicity at beginning stage of the aftershock sequence when the aftershock occurrence rate decays away, whereas by (25), when $t = 0$, $\lambda(t) = K/c^{-p} \approx$

∞. Moreover the parameter p becomes lower than 1, which might be caused by a secondary triggering effect by large aftershocks. However, this does not change much when the magnitude threshold changes from 2.95 to 4.15 (Table S2).

Bibliography

Daley, D. D. and Vere-Jones, D. (2008). *An Introduction to Theory of Point Processes – Volume II: General Theory and Structure (2nd Edition)*. Springer, New York, NY.

Møller, J. and Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC.

Table S1: Eruption records of the Hakone volcano . TUN: Tephra unit name

TUN	Age (ka)	EPI	TUN	Age (ka)	VEI
Kamiyama 1	45	4	CC4	52.5	4.8
CC3	53	4	CC2	56	4
CC1	58	5.1	Hk-S	64	5.1
Hk-T, Hk-TP	65.98	6.1	Hk-MP	71	5
Hk-AP	76	5	Hk-OP	88	6
Da5	89	5.2	KmP12	96.4	4
KmP11	97	4	Kmp10	975	5
Hk-Da4	98.5	6	KmP9	98.5	4
KmP8	99	4	KmP7	101	5.4
KmP6	102	5	KmP5	107	4
KmP4	107	4	KmP3	108	5
KmP2	109	5	KmP1	111	5.5
KIP13	117	5.4	KIP11	124	5
KIP10	125	5	KIP-9(Hk-Da1)	126	6
KIP8	127	5.4	KIP7	128	5.8
KIP6	129	5	KIP5	131	5
KIP3	134	5	KIP2	137	4
KIP1	138	4	TAu-12	150	5.4
TAu2	160	5	TAm-7	165	5
TAm-6	170	5	TAm-5	178	6.1
TAm-4	181	6	TAm-1	185	5.9
TB13	190	6	TB-7	228	5
TB-1	240	5.7	TCu-1	255	5.8
TCl4	310	5			

Magnitude threshold	Replenished dataset $[t_{\text{main}}, T]$			Orig. dataset $[t_{\text{main}}, T]$			Orig. dataset $[t_{\text{main}} + 54 \text{ days}, T]$		
	$\hat{K}$	$\hat{c}$	$\hat{p}$	$\hat{K}$	$\hat{c}$	$\hat{p}$	$\hat{K}$	$\hat{c}$	$\hat{p}$
2.95	804.4	.1140	1.003	82.29	.0553	.6205	567.9	.0000	.9374
3.05	639.2	.1131	1.003	80.31	.0596	.6547	461.9	.0000	.9410
3.15	511.5	.1134	1.001	79.25	.0660	.6872	384.0	.0000	.9447
3.25	412.9	.1110	.9965	79.04	.0737	.7185	311.4	.6961	.9407
3.35	327.3	.1067	.9926	78.80	.0825	.7555	228.2	.0009	.9254
3.45	260.3	.1141	.9925	80.67	.0991	.7986	182.6	.0026	.9271
3.55	213.8	.1142	.9953	83.33	.1177	.8407	150.5	.0000	.9309
3.65	171.6	.1135	.9907	85.73	.1360	.8799	105.5	.0000	.9048
3.75	135.9	.1132	.9911	90.18	.1642	.9278	84.24	.0003	.9058
3.85	111.2	.1029	.9941	95.17	.1935	.9708	67.65	.0000	.9033
3.95	100.0	.1241	1.015	103.2	.2383	1.023	70.62	.0000	.9498
4.05	74.12	.1082	1.013	79.20	.1938	1.027	58.65	.0002	.9663
4.15	60.65	.1266	1.026	62.92	.1690	1.034	47.95	.0003	.9761

Table S2: Results from fitting the Omori-Utsu formula to the original and the replenished datasets of earthquakes from Southwest China, with different magnitude thresholds and different fitting time intervals. t_{main}: occurrence time of the mainshock; T: end of the time interval.