

**SUPPLEMENT TO “SPATIALLY INHOMOGENEOUS BACKGROUND
RATE ESTIMATORS AND UNCERTAINTY QUANTIFICATION FOR
NONPARAMETRIC HAWKES POINT PROCESS MODELS OF
EARTHQUAKE OCCURRENCES”**

BY ERIC WARREN FOX, FREDERIC PAIK SCHOENBERG AND JOSHUA SETH GORDON
University of California, Los Angeles

APPENDIX A: DERIVATION OF EM HISTOGRAM ESTIMATORS

In this appendix we derive the histogram estimators given in Algorithm 1 in the main article using an expectation-maximization (EM) framework.

The log-likelihood function for the space-time Hawkes process model, given in Equation (7) in the main article, is often analytically intractable for direct maximization (Veen and Schoenberg, 2008). For the EM algorithm, we instead work with the complete data log-likelihood, which incorporates information about the branching structure, and conveniently decomposes into separate components for the background and triggering processes.

Suppose the latent branching structure for a sequence of earthquakes $\{(t_i, x_i, y_i, m_i) : i = 1, \dots, N\}$ is given by the following random variables:

$$\chi_{ii} = \begin{cases} 1, & \text{if earthquake } i \text{ is a mainshock} \\ 0, & \text{otherwise;} \end{cases}$$

$$\chi_{ij} = \begin{cases} 1, & \text{if earthquake } i \text{ is an aftershock of } j \text{ and } i > j; \\ 0, & \text{otherwise.} \end{cases}$$

Then the complete data log-likelihood is given by $l_c(\Theta) = l_\mu(\Theta) + l_\nu(\Theta)$, where

$$(S1) \quad l_\mu(\Theta) = \sum_{i=1}^N \chi_{ii} \log(\mu(x_i, y_i)) - \int_0^T \int \int_S \mu(x, y) dx dy dt,$$

$$(S2) \quad l_\nu(\Theta) = \sum_{j=1}^N \left[\sum_{i>j} \chi_{ij} \log(\nu(t_i - t_j, x_i - x_j, y_i - y_j; m_j)) \right. \\ \left. - \int_{t_j}^T \int \int_S \nu(t_i - t_j, x_i - x_j, y_i - y_j; m_j) dx dy dt \right],$$

where Θ represents the set of parameters (i.e., all the piecewise constant rates for the background and triggering components). We can further simplify (S2) by applying the

integral approximation of [Schoenberg \(2013\)](#):

$$\begin{aligned} & \int_{t_j}^T \int \int_S \nu(t_i - t_j, x_i - x_j, y_i - y_j; m_j) dx dy dt \\ &= \kappa(m_j) \int_{t_j}^T g(t - t_j) dt \int \int_S f(x - x_i, y - y_i) dx dy \\ &\approx \kappa(m_j) \int_{t_j}^{\infty} g(t - t_j) dt \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - x_i, y - y_i) dx dy = \kappa(m_j) \end{aligned}$$

The approximation is close provided that most of the aftershock activity is contained within the observation region.

E-Step

Since the latent variables representing the mainshock-aftershock labels are unobservable for real earthquake catalogs, we take their expectations, given a current estimate of the model:

$$(S3) \quad E[\chi_{ii} | \Theta^{(v)}] = p_{ii}^{(v)} = \frac{\mu^{(v)}(x_i, y_i)}{\lambda^{(v)}(t_i, x_i, y_i | H_{t_i})},$$

$$(S4) \quad E[\chi_{ij} | \Theta^{(v)}] = p_{ij}^{(v)} = \frac{\nu^{(v)}(t_i - t_j, x_i - x_j, y_i - y_j; m_j)}{\lambda^{(v)}(t_i, x_i, y_i | H_{t_i})}.$$

Here $\Theta^{(v)}$ represents the current estimate of the model's background and triggering rates. Since $\mu^{(v)}$ and $\lambda^{(v)}$ in (S3) are estimates of the mainshock and total seismicity rates at (t_i, x_i, y_i) , respectively, the ratio is the estimated probability that earthquake i is a mainshock. Similarly, since $\nu^{(v)}$ in (S4) is an estimate of the rate of aftershocks triggered by previously occurring earthquake j , the ratio is the estimated probability that earthquake i is an aftershock of j . Note, (S3) and (S4) are the update equations in step 4 of Algorithm 1.

Substituting $p_{ii}^{(v)}$ for χ_{ii} in (S1) and $p_{ij}^{(v)}$ for χ_{ij} in (S2) gives the expected complete data log-likelihood, $E[l_c(\Theta) | \Theta^{(v)}]$, which we maximize in the M-step.

M-Step

Now, assume that the true background rate μ , magnitude productivity κ , and triggering densities g and h are piecewise constant over a space-time region $S \times [0, T]$ and can be

expressed as:

$$(S5) \quad \mu(x, y) = \sum_{l=1}^{n_y} \sum_{k=1}^{n_x} \mu_{kl} \mathbf{1}_{D_{kl}}(x, y), \quad D_{kl} = ((k-1)\Delta x, k\Delta x] \times ((l-1)\Delta y, l\Delta y];$$

$$(S6) \quad \kappa(m) = \sum_{k=1}^{n_m} \kappa_k \mathbf{1}_{A_k}(m), \quad A_k = (\delta m_k, \delta m_{k+1}];$$

$$(S7) \quad g(t) = \sum_{k=1}^{n_t} g_k \mathbf{1}_{B_k}(t), \quad B_k = (\delta t_k, \delta t_{k+1}];$$

$$(S8) \quad h(r) = \sum_{k=1}^{n_r} h_k \mathbf{1}_{C_k}(r), \quad C_k = (\delta r_k, \delta r_{k+1}];$$

where $\mathbf{1}$ denotes the indicator function. Note, the notation has slightly different meaning than in Section 3.1 of the main article, where the sets referred to the indices of observed points that fall within each interval (bin); here, the sets refer to the intervals themselves.

The remainder of this appendix will be divided into two parts. First, we show that if $\mu(x, y)$ is assumed piecewise constant, then the estimator given in step 2 of Algorithm 1 is the maximum likelihood estimate of $E[l_\mu(\Theta)|\Theta^{(v)}]$. Second, we show that if κ , g , and h are all assumed piecewise constant then the estimators in step 3 of Algorithm 1 are the maximum likelihood estimates of $E[l_\nu(\Theta)|\Theta^{(v)}]$ when using the integral approximation.

Part 1: Assume that $\mu(x, y)$ is piecewise constant as expressed by (S5). Then plugging (S5) into the expected complete data log-likelihood for the background process and solving the integral term gives:

$$E[l_\mu(\Theta)|\Theta^{(v)}] = \sum_{i=1}^N p_{ii}^{(v)} \log \left(\sum_{l=1}^{n_y} \sum_{k=1}^{n_x} \mu_{kl} \mathbf{1}_{D_{kl}}(x_i, y_i) \right) - T \sum_{l=1}^{n_y} \sum_{k=1}^{n_x} \mu_{kl} \Delta x \Delta y.$$

Take the partial derivative with respect to the rate μ_{kl} in cell D_{kl} :

$$(S9) \quad \frac{\partial E[l_\mu(\Theta)|\Theta^{(v)}]}{\partial \mu_{kl}} = \sum_{i=1}^N \left(\frac{p_{ii}^{(v)} \mathbf{1}_{D_{kl}}(x_i, y_i)}{\sum_{l=1}^{n_y} \sum_{k=1}^{n_x} \mu_{kl} \mathbf{1}_{D_{kl}}(x_i, y_i)} \right) - T \Delta x \Delta y$$

$$(S10) \quad = \sum_{i=1}^N \left(\frac{p_{ii}^{(v)} \mathbf{1}_{D_{kl}}(x_i, y_i)}{\mu_{kl}} \right) - T \Delta x \Delta y.$$

Equation (S10) follows by realizing that when evaluating the sum in (S9) the fraction is equal to $p_{ii}^{(v)}/\mu_{kl}$ when $(x_i, y_i) \in D_{kl}$, and 0 when $(x_i, y_i) \notin D_{kl}$. Setting (S10) equal to 0 and solving gives

$$\mu_{kl}^{(v+1)} = \frac{\sum_{i=1}^N p_{ii}^{(v)} \mathbf{1}_{D_{kl}}(x_i, y_i)}{T \Delta x \Delta y},$$

which is the background rate histogram estimator in step 2 of Algorithm 1.

Part 2: Assume that $\kappa(m)$, $g(t)$, and $h(r)$ are all piecewise constant as expressed by Equations (S6)–(S8). If we apply the integral approximation, then the expected complete data log-likelihood for the triggering process is given by:

$$\begin{aligned} E[l_\nu(\Theta)|\Theta^{(v)}] = & \sum_{j=1}^N \left[\sum_{i>j} p_{ij}^{(v)} \left(\log \left(\sum_{k=1}^{n_m} \kappa_k \mathbf{1}_{A_k}(m_j) \right) + \log \left(\sum_{k=1}^{n_t} g_k \mathbf{1}_{B_k}(t_i - t_j) \right) + \log \left(\sum_{k=1}^{n_r} h_k \mathbf{1}_{C_k}(r_{ij}) \right) - \log(2\pi r_{ij}) \right) \right] \\ & - \sum_{j=1}^N \sum_{k=1}^{n_m} \kappa_k \mathbf{1}_{A_k}(m_j) - c_1 \left(\sum_{k=1}^{n_t} g_k \Delta t_k - 1 \right) - c_2 \left(\sum_{k=1}^{n_r} h_k \Delta r_k - 1 \right), \end{aligned}$$

where c_1 and c_2 are Lagrange multipliers for the constraints that g and h integrate to 1. Also note that $f(r_{ij}) = h(r_{ij})/(2\pi r_{ij})$, where r_{ij} is the distance between events i and j .

First, take the partial derivative with respect to κ_k :

$$\frac{\partial E[l_\nu(\Theta)|\Theta^{(v)}]}{\partial \kappa_k} = \sum_{j=1}^N \sum_{i>j} \left(\frac{p_{ij}^{(v)} \mathbf{1}_{A_k}(m_j)}{\kappa_k} \right) - \sum_{j=1}^N \mathbf{1}_{A_k}(m_j).$$

Setting the partial derivative equal to 0 and solving gives

$$\kappa_k^{(v+1)} = \frac{\sum_{j=1}^N \sum_{i>j} p_{ij}^{(v)} \mathbf{1}_{A_k}(m_j)}{\sum_{j=1}^N \mathbf{1}_{A_k}(m_j)},$$

which is the histogram estimator for the magnitude productivity in step 3 of Algorithm 1. Next, set the partial derivatives with respect to g_k and c_1 equal to 0:

$$\begin{aligned} \frac{\partial E[l_\nu(\Theta)|\Theta^{(v)}]}{\partial g_k} &= \sum_{j=1}^N \sum_{i>j} \left(\frac{p_{ij}^{(v)} \mathbf{1}_{B_k}(t_i - t_j)}{g_k} \right) - c_1 \Delta t_k = 0, \\ \frac{\partial E[l_\nu(\Theta)|\Theta^{(v)}]}{\partial c_1} &= 1 - \sum_{k=1}^{n_t} g_k \Delta t_k = 0. \end{aligned}$$

Note the second line gives the histogram density estimator constraint. It is straightforward to show that solving for g_k and c_1 gives

$$g_k^{(v+1)} = \frac{\sum_{j=1}^N \sum_{i>j} p_{ij}^{(v)} \mathbf{1}_{B_k}(t_i - t_j)}{\Delta t_k \sum_{j=1}^N \sum_{i>j} p_{ij}^{(v)}},$$

which is the estimator of $g(t)$ in Algorithm 1. Similarly, solving $\partial E[l_\nu(\Theta)|\Theta^{(v)}]/\partial h_k = 0$ and $\partial E[l_\nu(\Theta)|\Theta^{(v)}]/\partial c_2 = 0$ gives the histogram density estimator of $h(r)$ in Algorithm 1.

APPENDIX B: BOOTSTRAP SIMULATION ALGORITHM

This appendix provides an algorithm for simulating from a nonparametric Hawkes process model $\hat{\lambda}(t, x, y|H_t)$ which is estimated using Algorithm 1 or 2. This is the procedure we use in step 1 of Algorithm 3 to generate the bootstrap datasets used to compute the error bars for the histogram estimates of the triggering function.

Algorithm S1 explains how to simulate from the estimated background rate $\hat{\mu}(x, y)$. Here we use the thinning procedure of [Lewis and Shedler \(1979\)](#). Algorithm S2 explains how to simulate aftershocks from the histogram estimates $\hat{\kappa}(m)$, $\hat{g}(t)$, and $\hat{h}(t)$ of the triggering function components.

Algorithm S1: Simulating the Background Process

1. Let μ^{max} be the maximum of the background rate estimate $\hat{\mu}(x, y)$ over S .
2. Draw N^{max} from $Pois(\mu^{max} \cdot |S| \cdot T)$; this is an upper bound on the simulated number of background events.
3. Generate N^{max} event locations and times independently and uniformly over the observation region $S \times [0, T]$.
4. Retain each event generated in step 3 with probability $\hat{\mu}(x, y)/\mu^{max}$, otherwise remove it from our background set. Let $N(0)$ be the number of retained events.
5. Draw a bootstrap sample of size $N(0)$ from the magnitudes of events $\{m_i : i = 1, \dots, N\}$ in the earthquake catalog used to estimate the model.
6. Let $G(0) = \{(t_i^{(0)}, x_i^{(0)}, y_i^{(0)}, m_i^{(0)}) : i = 1, \dots, N(0)\}$ be the set of times, locations, and magnitudes of the mainshocks selected in steps 4 and 5, which we refer to as generation 0.

If the background rate $\hat{\mu}(x, y)$ was estimated using the histogram method (Algorithm 1) we can generate the event times and locations by simulating a homogeneous Poisson process in each cell.

After laying down the mainshocks in $G(0)$, we can simulate subsequent generations of aftershocks from the nonparametric triggering function and then form the bootstrap dataset with the following algorithm:

Algorithm S2: Simulating the Triggering Process

1. Set generation index $v = 1$.
2. For each event $i \in \{1, \dots, N(v - 1)\}$ draw n_i^{aft} from $Pois(\hat{\kappa}(m_i^{(v-1)}))$; this is the number of aftershocks triggered by each event i in the previous generation $G(v - 1)$.
3. Let $N(v) = \sum_{i=1}^{N(v-1)} n_i^{aft}$ be the number of aftershocks in the current generation v . If $N(v) = 0$, there are no more aftershocks, set $v = v^{end}$ and go to step 8.
4. For each event $i \in \{1, \dots, N(v - 1)\}$, generate n_i^{aft} aftershocks times and distances by sampling from the histogram density estimates $\hat{g}(t)$ and $\hat{h}(r)$.

5. Draw a bootstrap sample of size $N(v)$ from the magnitudes of events $\{m_i : i = 1, \dots, N\}$ in the earthquake catalog used to estimate the model.
6. Let $G(v) = \{(t_i^{(v)}, x_i^{(v)}, y_i^{(v)}, m_i^{(v)}) : i = 1, \dots, N(v)\}$ be the set of times, locations, and magnitudes of aftershocks generated in steps 4 and 5, which we refer to as generation v .
7. Update $v \leftarrow v + 1$. Repeat steps 2 – 6.
8. Form the bootstrap dataset b^* by taking the union of events generated over all generations. That is, $b^* = \bigcup_{v=0}^{v^{end}} G(v) = \{(t_i^*, x_i^*, y_i^*, m_i^*) : i = 1, \dots, N^*\}$ where $N^* = \sum_{v=0}^{v^{end}} N(v)$.

To sample from the histogram density estimates $\hat{g}(t)$ and $\hat{h}(r)$ in step 4 we use the `sample()` function in [R Core Team \(2014\)](#), setting the probability of selecting time bin $(\delta t_k, \delta t_{k+1}]$ to $\Delta t_k \hat{g}_k$ and distance bin $(\delta r_k, \delta r_{k+1}]$ to $\Delta r_k \hat{h}_k$. Here \hat{g}_k is the density estimate for bin $k \in \{1, \dots, n_t^{bins}\}$, and \hat{h}_k is the density estimate for bin $k \in \{1, \dots, n_r^{bins}\}$. Once a bin is selected, the corresponding time or distance is drawn uniformly over that bin's interval.

All events generated by Algorithm S2 are retained in the bootstrap dataset b^* , including aftershocks that occur outside the observation region $S \times [0, T]$. That is, each time a bootstrap dataset is generated from a fitted nonparametric model, aftershocks that occur outside of the observation region are retained in that bootstrap dataset, and used to re-estimate the nonparametric model and form the replicate. Not keeping these aftershock events may result in further boundary effects on the bootstrap intervals, and error bars which are not centered around the histogram estimates of the triggering function.

APPENDIX C: ADDITIONAL FIGURES AND TABLES FOR SECTION 4

In this appendix we provide the additional tables and figures mentioned in Section 4. These show a realization of the bootstrap intervals (Section 3.3) for each of the 3 simulation studies mentioned in Sections 4.1, 4.2, and 4.3.

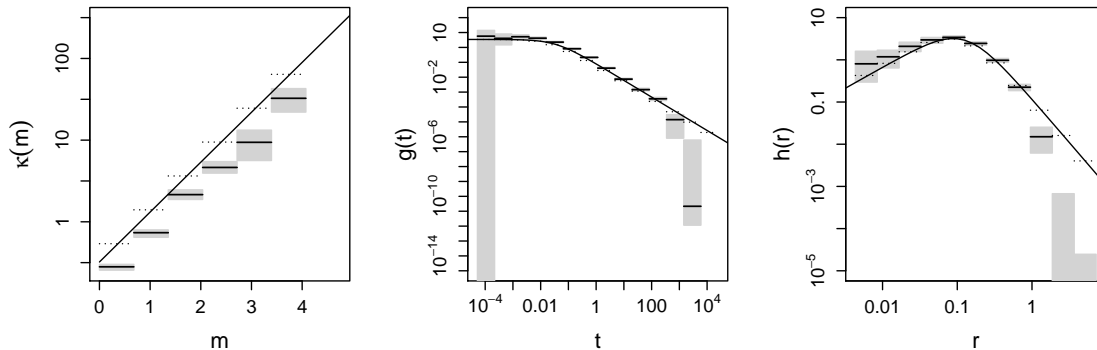


FIG S1. Estimate of the magnitude, temporal, and distance components of the triggering function from one realization of the ETAS model specified in Section 4.1. The solid horizontal lines (black) are the estimates in each bin, with 95% bootstrap error bars (gray) formed from 200 replicates of the estimated nonparametric model. The black curves are the true triggering components governing the simulation, and the dashed horizontal lines are the binned average values of these functions.

TABLE S1

Estimate of the background rate from one realization of the ETAS model specified in Section 4.1. Approximate 95% bootstrap confidence intervals (CI), which quantify the sampling variability of the background rate estimator for each cell, were formed from 200 replicates of the estimated nonparametric model (i.e., the intervals are the 0.025 and 0.975 percentiles of the replicates of the mainshock rate in each cell). Note that the CIs are upwardly biased, since the estimates are upwardly biased due to the boundary issues discussed in Section 4. Cell numbers correspond to Figure 1 in the main article.

Cell	Truth	Estimate	CI
1	0.001	0.0015	(0.0014, 0.0018)
2	0.005	0.0059	(0.0055, 0.0063)
3	0.005	0.0060	(0.0056, 0.0065)
4	0.001	0.0016	(0.0014, 0.0018)

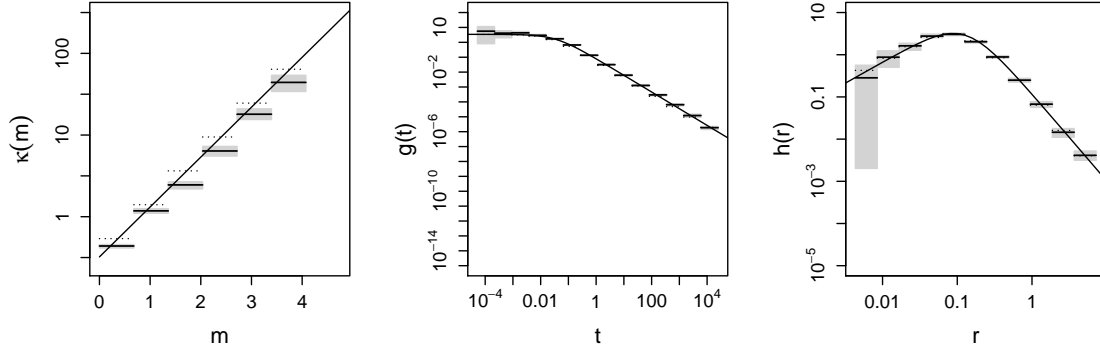


FIG S2. Estimate of the magnitude, temporal, and distance components of the triggering function from one realization of the ETAS model specified in Section 4.1, with boundary correction (Section 4.2) for aftershock activity $\epsilon_r = 1000$ and $\epsilon_t = 10^6$. The solid horizontal lines (black) are the estimates in each bin, with 95% bootstrap error bars (gray) formed from 200 replicates of the estimated nonparametric model. The black curves are the true triggering components governing the simulation, and the dashed horizontal lines are the binned average values of these functions.

TABLE S2

Estimate of the background rate from one realization of the ETAS model specified in Section 4.1, with boundary correction (Section 4.2) for aftershock activity $\epsilon_r = 1000$ and $\epsilon_t = 10^6$. Approximate 95% bootstrap confidence intervals (CI), which quantify the sampling variability of the background rate estimator for each cell, were formed from 200 replicates of the estimated nonparametric model (i.e., the intervals are the 0.025 and 0.975 percentiles of the replicates of the mainshock rate in each cell). The boundary correction substantially reduces the bias in the estimates and their respective CIs. Note, the boundary correction used here can only be applied to simulated earthquake data, and thus only serves to illustrate the extent to which the estimates and CIs are affected by boundary issues. The cell numbers correspond to Figure 1 in the main article.

Cell	Truth	Estimate	CI
1	0.001	0.0010	(0.0008, 0.0012)
2	0.005	0.0051	(0.0047, 0.0056)
3	0.005	0.0048	(0.0044, 0.0053)
4	0.001	0.0010	(0.0008, 0.0012)

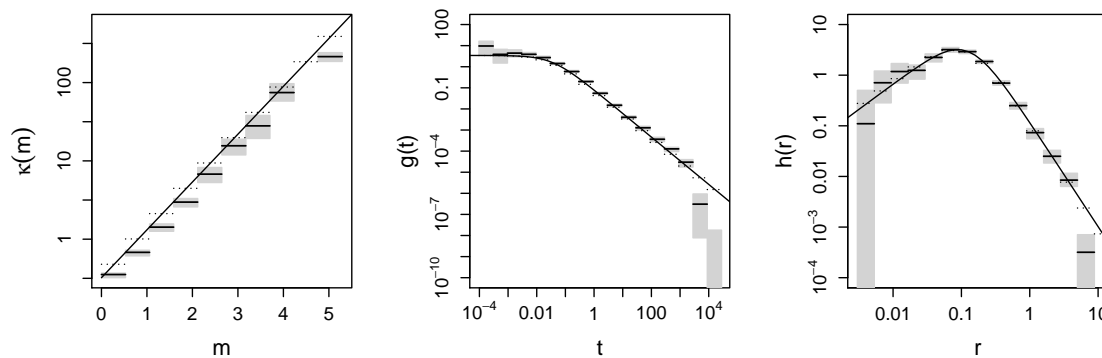


FIG S3. Estimate of the magnitude, temporal, and distance components of the triggering function from one realization of the ETAS model specified in Section 4.3. The solid horizontal lines (black) are the estimates in each bin, with 95% bootstrap error bars (gray) formed from 200 replicates of the estimated nonparametric model. The black curves are the true triggering components governing the simulation, and the dashed horizontal lines are the binned average values of these functions. The gap in the histogram estimate of the productivity function is due to there being no simulated events (for this particular realization) with magnitudes in the bin (4.2, 4.8].

TABLE S3

Estimate of total mainshock rate from one realization of the ETAS model specified in Section 4.3. An approximate 95% bootstrap confidence interval (CI), which quantifies the sampling variability in the estimation of the total mainshock rate, was formed from 200 replicates of the estimated nonparametric model (i.e., the intervals are the 0.025 and 0.975 percentiles of the replicates of the total mainshock rate).

Truth	Estimate	CI
2000	2012.4	(1981.8, 2251.8)

APPENDIX D: ADDITIONAL SIMULATION STUDIES

D.1. Mainshocks Occurring Outside the Observation Region. In this appendix we conduct a simulation study to assess the impact that mainshocks occurring outside the observation region have on the nonparametric estimates. We simulate earthquakes on a $[0, 8] \times [0, 12] \times [0, 25000]$ space-time region. The observation region for the estimation is a subset of this larger region given by $S \times [0, T] = [2, 6] \times [3, 9] \times [0, 25000]$. The specification of the ETAS model used for this simulation study is the same as in Section 4.1, except we allow for mainshock events to occur outside the observation region. Figure S4 (left) shows the background rate with the subset observation region denoted by the dashed rectangle. Figure S4 (right) shows ones simulated realization of the ETAS model over this region.

The average, bias, standard deviation, and RMSE of the background rate estimates in each cell from 200 simulations of ETAS are shown in Table S4. Note that the estimates are only from events occurring within the observation region. The main result is that the background estimates are more positively biased for the simulations with mainshocks occurring outside the observation region than the simulations with mainshocks constrained within the observation region. This can be seen by comparing the bias values in Table S4 with Table 1 in the main article. This increase in bias is perhaps attributable to mainshocks outside the observation region triggering aftershocks inside the observation region which are then misclassified by the EM algorithm.

The histogram estimates of the triggering function when allowing mainshock events outside the observation region are not substantially different than the estimates when constraining mainshocks within the observation region. For instance, the bias, standard deviation, and RMSE of the triggering estimates for 200 simulations of ETAS in Tables S5–S7 are very similar to the values in Tables 2–4 in the main article. Also, the shape of triggering function estimates in Figure S5 appears very similar to Figure 3 in the main article.

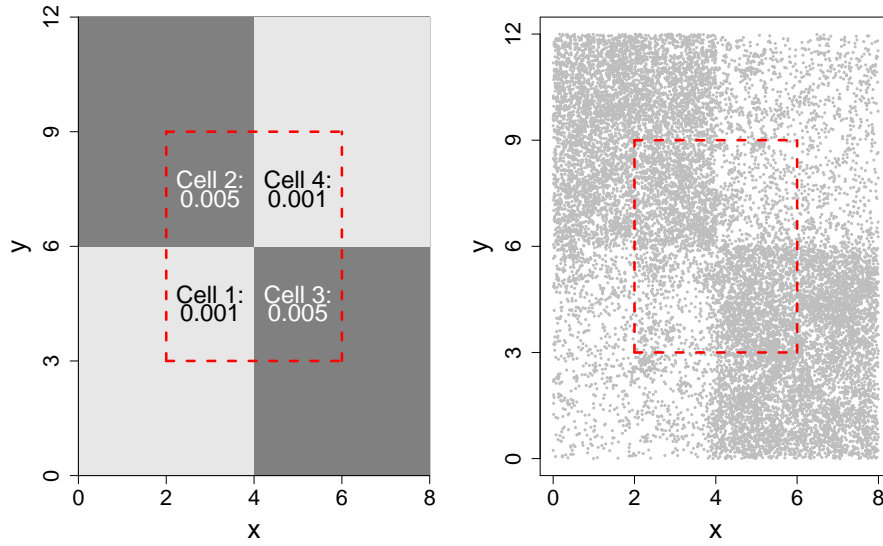


FIG S4. True background rate used in simulation study with mainshocks allowed to occur outside the observation region (left). Epicenters of earthquakes from one simulated realization of ETAS model (right). Spatial observation region is denoted by the red dashed box.

TABLE S4

Average, bias, standard deviation, and RMSE of background rate estimates in each cell of the observation region from 200 simulations of ETAS with mainshocks allowed to occur outside observation window. Cell numbers correspond to Figure S4 (left).

Cell	Truth	Avg	Bias	SD	RMSE
1	0.001	0.00167	0.00067	0.00014	0.00068
2	0.005	0.00669	0.00169	0.00030	0.00172
3	0.005	0.00669	0.00169	0.00032	0.00172
4	0.001	0.00166	0.00066	0.00017	0.00068

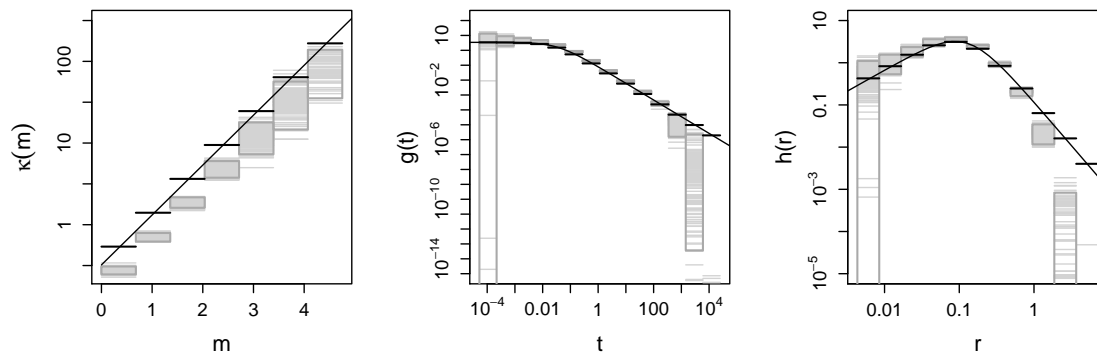


FIG S5. Triggering function estimates (gray horizontal lines) from 200 simulations of ETAS when allowing mainshocks to occur outside the observation region. The black curves are the true triggering components governing the simulation, and the black horizontal lines are the binned average values of these functions.

TABLE S5

Average, bias, standard deviation, and RMSE for the histogram estimates of $\kappa(m)$ for simulation study with mainshocks occurring outside the observation region. Bias and RMSE are evaluated using the binned average values of $\kappa(m)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0, 0.679]	0.539	0.277	-0.262	0.0163	0.263
(0.679, 1.36]	1.4	0.71	-0.691	0.0441	0.693
(1.36, 2.04]	3.64	1.85	-1.79	0.149	1.79
(2.04, 2.72]	9.47	4.86	-4.61	0.542	4.64
(2.72, 3.39]	24.6	12.3	-12.3	2.51	12.5
(3.39, 4.07]	64	32.6	-31.4	10.6	33.1
(4.07, 4.75]	166	86.2	-80	28.9	85.1

TABLE S6

Average, bias, standard deviation, and RMSE for the histogram estimates of $g(t)$ for simulation study with mainshocks occurring outside the observation region. Bias and RMSE are evaluated using the binned average values of $g(t)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0.000052, 0.00022]	3.41	5.3	1.89	3.77	4.22
(0.00022, 0.00091]	3.37	5.25	1.88	1.85	2.64
(0.00091, 0.0038]	3.19	4.77	1.58	0.915	1.83
(0.0038, 0.016]	2.62	3.97	1.35	0.406	1.41
(0.016, 0.066]	1.51	2.31	0.792	0.149	0.806
(0.066, 0.27]	0.534	0.807	0.273	0.0465	0.277
(0.27, 1.1]	0.132	0.197	0.0649	0.0116	0.0659
(1.1, 4.8]	0.0282	0.0414	0.0132	0.0027	0.0135
(4.8, 20]	0.00577	0.00825	0.00248	0.000668	0.00257
(20, 83]	0.00117	0.0016	0.000426	0.000153	0.000453
(83, 345]	0.000236	0.000281	4.48E-05	4.47E-05	6.33E-05
(345, 1439]	4.77E-05	2.61E-05	-2.16E-05	1.5E-05	2.63E-05
(1439, 5998]	9.63E-06	1.98E-07	-9.43E-06	6.67E-07	9.45E-06
(5998, 25000]	1.94E-06	9.58E-18	-1.94E-06	6.67E-17	1.94E-06

TABLE S7

Average, bias, standard deviation, and RMSE for the histogram estimates of $h(r)$ for simulation study with mainshocks occurring outside the observation region. Bias and RMSE are evaluated using the binned average values of $h(r)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0.0043, 0.0085]	0.426	0.511	0.0853	0.264	0.278
(0.0085, 0.017]	0.826	1.01	0.185	0.263	0.321
(0.017, 0.033]	1.55	1.87	0.324	0.259	0.415
(0.033, 0.064]	2.58	3.12	0.539	0.249	0.593
(0.064, 0.13]	3.11	3.68	0.566	0.176	0.593
(0.13, 0.25]	2.14	2.46	0.322	0.104	0.338
(0.25, 0.49]	0.846	0.909	0.0629	0.0524	0.0819
(0.49, 0.95]	0.246	0.206	-0.0397	0.0207	0.0447
(0.95, 1.9]	0.064	0.0225	-0.0414	0.00581	0.0418
(1.9, 3.7]	0.0161	0.000118	-0.016	0.000271	0.016
(3.7, 7.2]	0.00402	2.42E-07	-0.00402	3.42E-06	0.00402

D.2. Using a Larger Observation Region. In this appendix we conduct a simulation study on a larger observation region than used for the simulation study in Section 4.1. Specifically, we simulate from an ETAS model over $S \times T = [0, 6] \times [0, 9] \times [0, 30000]$ with background rate shown in Figure S6, and the same triggering function as specified in Section 4.1 of the main article. We simulate and re-estimate the ETAS model 200 times using Algorithm 1 and no boundary correction (i.e., only events occurring within the observation region are used for estimation). The results are shown in Figure S7 and Tables S8–S11.

Table S8 reveals that the positive bias in the background rate estimates do not change much when enlarging the observation region. This can be seen by comparison with Table 1 in the main article. Although, there is a reduction in the standard deviation of the background estimates due to the larger sample size.

Overall, there does appear to be some improvement in the bias, standard deviation, and RMSE of the estimates of the triggering function on the larger observation region. This can be seen by comparison with the results in Section 4.1 of the main article. However, boundary effects on the right-tail ends of the triggering function are still quite noticeable on the larger region.

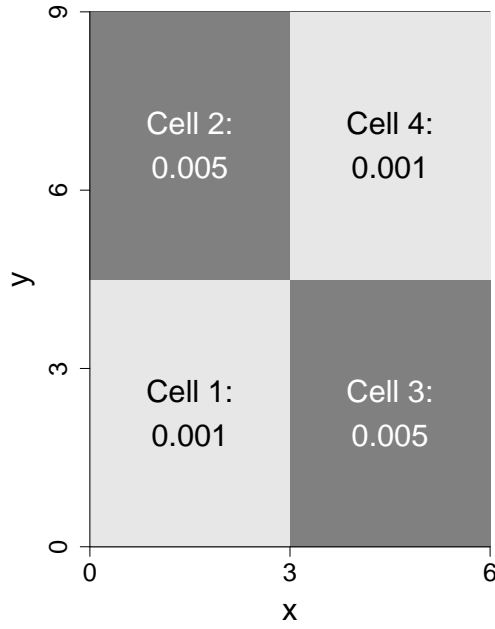


FIG S6. True background rate used in simulation study with larger observation region.

TABLE S8

Average, bias, standard deviation, and RMSE of background rate estimates in each cell from 200 simulations of ETAS with larger observation region. Cell numbers correspond to Figure S6.

Cell	Truth	Avg	Bias	SD	RMSE
1	0.001	0.00136	0.00036	0.00008	0.00037
2	0.005	0.00600	0.00100	0.00019	0.00102
3	0.005	0.00598	0.00098	0.00018	0.00100
4	0.001	0.00135	0.00035	0.00008	0.00036

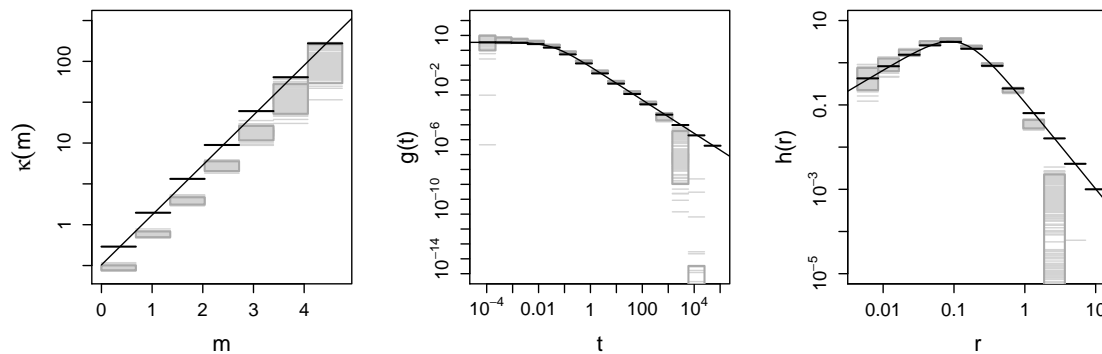


FIG S7. Triggering function estimates (gray horizontal lines) from 200 simulations of ETAS with larger observation region. The black curves are the true triggering components governing the simulation, and the black horizontal lines are the binned average values of these functions.

TABLE S9

Average, bias, standard deviation, and RMSE for the histogram estimates of $\kappa(m)$ for simulation study with larger observation region. Bias and RMSE are evaluated using the binned average values of $\kappa(m)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0, 0.679]	0.539	0.296	-0.243	0.0102	0.243
(0.679, 1.36]	1.4	0.764	-0.637	0.032	0.638
(1.36, 2.04]	3.64	1.98	-1.67	0.102	1.67
(2.04, 2.72]	9.47	5.13	-4.33	0.369	4.35
(2.72, 3.39]	24.6	13.4	-11.2	1.44	11.3
(3.39, 4.07]	64	35.4	-28.6	7.5	29.6
(4.07, 4.75]	166	96.9	-69.3	29	75.2

TABLE S10

Average, bias, standard deviation, and RMSE for the histogram estimates of $g(t)$ for simulation study with larger observation region. Bias and RMSE are evaluated using the binned average values of $g(t)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0.000052, 0.00022]	3.41	4.88	1.47	2.1	2.56
(0.00022, 0.00091]	3.37	5.07	1.7	1.07	2.01
(0.00091, 0.0038]	3.19	4.69	1.5	0.522	1.58
(0.0038, 0.016]	2.62	3.87	1.25	0.267	1.28
(0.016, 0.066]	1.51	2.21	0.699	0.0898	0.705
(0.066, 0.27]	0.534	0.781	0.247	0.0287	0.249
(0.27, 1.1]	0.132	0.191	0.059	0.0072	0.0594
(1.1, 4.8]	0.0282	0.04	0.0118	0.00163	0.012
(4.8, 20]	0.00577	0.00811	0.00234	0.000381	0.00237
(20, 83]	0.00117	0.00161	0.000441	9.64E-05	0.000451
(83, 345]	0.000236	0.000301	6.52E-05	2.72E-05	7.06E-05
(345, 1439]	4.77E-05	4.12E-05	-6.46E-06	1.04E-05	1.22E-05
(1439, 5998]	9.63E-06	4.18E-07	-9.21E-06	1.11E-06	9.28E-06
(5998, 25000]	1.94E-06	1.31E-11	-1.94E-06	1.65E-10	1.94E-06
(25000, 30000]	3.92E-07	8.52E-58	-3.92E-07	1.2E-56	3.92E-07

TABLE S11

Average, bias, standard deviation, and RMSE for the histogram estimates of $h(r)$ for simulation study with larger observation region. Bias and RMSE are evaluated using the binned average values of $h(r)$, which are provided in the column labeled Truth.

Bin	Truth	Avg	Bias	SD	RMSE
(0.0043, 0.0085]	0.426	0.49	0.064	0.148	0.162
(0.0085, 0.017]	0.826	0.939	0.112	0.162	0.197
(0.017, 0.033]	1.55	1.75	0.197	0.148	0.246
(0.033, 0.064]	2.58	2.95	0.373	0.14	0.398
(0.064, 0.13]	3.11	3.55	0.435	0.109	0.449
(0.13, 0.25]	2.14	2.41	0.27	0.061	0.277
(0.25, 0.49]	0.846	0.909	0.0628	0.0311	0.0701
(0.49, 0.95]	0.246	0.227	-0.0189	0.0126	0.0227
(0.95, 1.9]	0.064	0.0354	-0.0285	0.00447	0.0289
(1.9, 3.7]	0.0161	0.000571	-0.0155	0.000687	0.0156
(3.7, 7.2]	0.00402	3.31E-07	-0.00402	4.38E-06	0.00402
(7.2, 10.8]	0.001	2.36E-31	-0.001	3.3E-30	0.001

APPENDIX E: USING DISTANCE TO THE FAULT

In this appendix we apply estimation Algorithm 2 (Section 3.2) to the Japan earthquake dataset using distance to the fault instead of epicentral distance. To estimate a fault we use the location of all nearby events, both before and after, and perform a weighted least squares regression. The details for estimating mainshock faults are outlined below.

E.1. Fault Approximation. Using the following steps we can approximate a fault for each mainshock:

1. Let earthquake j be called the mainshock. Using the entire catalog of events, both before and after the mainshock, identify all $m_i \geq c_1$ events within some area surrounding a $m_j \geq c_2$ event, where $c_1 \leq c_2$. Here c is a magnitude cut-off and m is the magnitude of an earthquake.
2. Calculate the Euclidean distance between a m_j earthquakes epicenter and all other m_i events denoted r_{ij} .

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

3. Fit a weighted ordinary least squares regression for each event j constrained to pass through the m_j epicenter with weights $w_j = 1/r_{ij}$.
4. Store the coefficient of the weighted regression β_j as the estimated fault line for each m_j event.

If any event is in the same exact location as the mainshock as a result of a location error, we remove the event since it adds no information to the fault estimate.

E.2. Distance to Fault. Distance to fault was calculated by computing the perpendicular distance for each pair (i, j) of earthquakes, where $t_i > t_j$, in the catalog using the fault approximated by the method described above. Since the regression line must pass through the mainshock location we can rewrite the regression equation as $-\beta_j x + y = 0$ and apply the formula:

$$d_{ij} = \frac{|-\beta_j x_i + y_i|}{\sqrt{\beta_j^2 + 1^2}}.$$

E.3. Results. We apply Algorithm 2 to estimate the triggering function and background rate for the Japan earthquake dataset (Section 5) using distance to the fault, d_{ij} , instead of epicentral distance. Fault orientation was estimated by setting the magnitude cut-offs to $c_1 = c_2 = 4.0$ and a surrounding area of .045 degrees by .045 degrees (roughly 100km²) for each event in the catalog. All input parameters were set identically to the epicentral approach (i.e., number of bins and smoothing bandwidth).

The method estimated 137 mainshocks as opposed to the 801 estimated using epicentral distance. Figure S8 shows the estimate of the underlying spatial inhomogeneous Poisson

processes $\mu(x, y)$ for mainshock activity. Although the estimated background rate is much lower, its spatial distribution is somewhat similar to the background rate described in Section 5 (Figure 12). Histogram estimates of the components of the triggering function are shown Figure S9. Compared to the estimate derived using epicentral distance in Figure 13, the density estimate of $g(t)$ and the magnitude productivity function $\kappa(m)$ are quite similar; however, the estimates for the spatial distribution $h(r)$ shows a much higher rate of triggering at shorter distances.

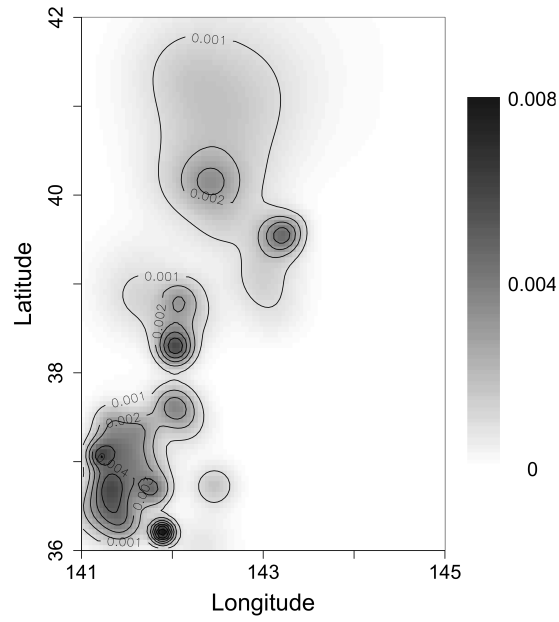


FIG S8. Estimate of background rate (Algorithm 2, step 2) for Japan earthquake dataset using distance to fault. Rate values are in events/day/degree².

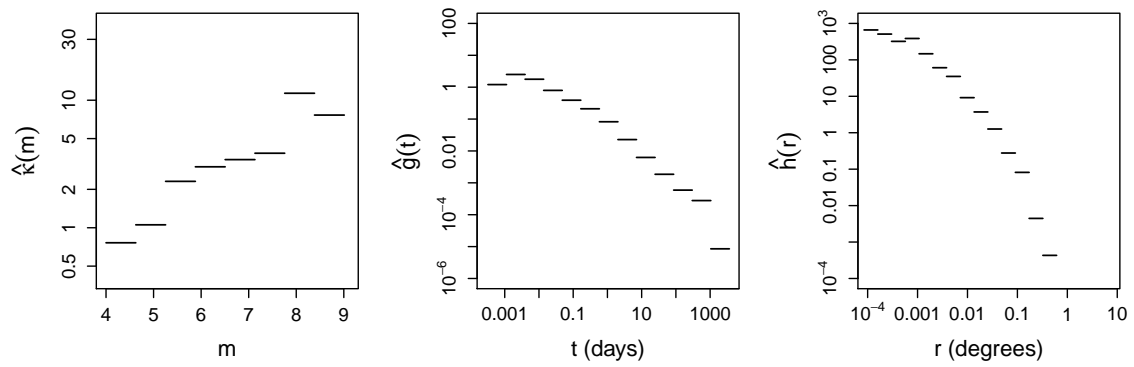


FIG S9. Magnitude, temporal, and distance components for the triggering function estimated from the Japan earthquake dataset using distance to fault.

REFERENCES

- LEWIS, P. A. and SHEDLER, G. S. (1979). Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistics Quarterly* **26** 403–413.
- SCHOENBERG, F. P. (2013). Facilitated estimation of ETAS. *Bulletin of the Seismological Society of America* **103** 601–605.
- R CORE TEAM (2014). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.
- VEEN, A. and SCHOENBERG, F. P. (2008). Estimation of space–time branching process models in seismology using an EM–type algorithm. *Journal of the American Statistical Association* **103** 614–624.

E. W. FOX
F. P. SCHOENBERG
J. S. GORDON
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
8125 MATH SCIENCES BUILDING
LOS ANGELES, CALIFORNIA 90095-1554
E-MAIL: eric.fox@stat.ucla.edu
frederic@stat.ucla.edu
joshuagordon@g.ucla.edu