Self-exciting point process in modeling earthquake occurrences

H. Pratiwi¹, I. Slamet², D. R. S. Saputro³, Respatiwulan⁴
¹,²,⁴Statistics Study Program, Universitas Sebelas Maret, Jl. Ir. Sutami 36A, Surakarta 57126, Indonesia
³Mathematics Study Program, Universitas Sebelas Maret, Jl. Ir. Sutami 36A, Surakarta 57126, Indonesia
E-mail: hpratiwi@mipa.uns.ac.id, isnandarlamet@staff.uns.ac.id, dewiretnoss@staff.uns.ac.id, respatiwulan@staff.uns.ac.id

Abstract. In this paper, we present a procedure for modeling earthquake based on spatial-temporal point process. The magnitude distribution is expressed as truncated exponential and the event frequency is modeled with a spatial-temporal point process that is characterized uniquely by its associated conditional intensity process. The earthquakes can be regarded as point patterns that have a temporal clustering feature so we use self-exciting point process for modeling the conditional intensity function. The choice of main shocks is conducted via window algorithm by Gardner and Knopoff and the model can be fitted by maximum likelihood method for three random variables.

1. Introduction
A spatio-temporal point process is a random collection of points, where each point represents the time and location of an event. Examples of events include incidence of disease, sightings or births of a species, or the occurrences of fires, earthquakes, lightning strikes, tsunamis, or volcanic eruptions. Typically the spatial locations are recorded in three spatial coordinates, e.g. longitude, latitude, and height or depth, though sometimes only one or two spatial coordinates are available or of interest. One particularly useful approach to modeling spatio-temporal point processes is through the conditional intensity function which depend on the location and times of past events occurring in the time interval \((0, t)\) [9]. In this paper we analyze earthquake model in Java island using a spatial-temporal point process approach. We present self-exciting point process where the earthquake magnitude as the mark.

Point process and its notation are defined in second section. Third section defines the truncated exponential distribution and fourth section defines the spatio-temporal point precess and the self-exciting point process. We consider the distributions of earthquake magnitudes, mainshocks, and aftershocks in estimating the intensity function. Finally we apply the model to the earthquake data in Java island in the last section.
2. Point process
Consider a sequence \((X_n)\) of random vectors in the state space \(E\) and define, for \(A \subset E\),
\[
N(A) = \#\{i \geq 1 : X_i \in A\},
\]
i.e., \(N(A)\) counts the number of \(X_i\)’s falling into \(A\). The state space \(E\), where the points live, is a Borel subset of a finite-dimensional Euclidean space and \(E\) is equipped with the \(\sigma\)-field \(\mathcal{E} = \mathcal{B}(E)\) of the Borel sets generated by the open sets of \(E\). It is convenient to write a point process using the Dirac measure \(\delta_x\) at \(x \in E\):
\[
\delta_x(A) = I_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A, \end{cases} \quad A \in \mathcal{E}
\]
For a given sequence \((x_i)_{i \geq 1}\) in \(E\),
\[
m(A) = \sum_{i=1}^{\infty} \delta_{x_i}(A) = \#\{i \geq 1 : x_i \in A\}, \quad A \in \mathcal{E}
\]
define a counting measure on \(\mathcal{E}\) which is called point measure if \(m(K) < \infty\) for all compact sets \(K \subset E\). This means that any compact set \(K\) must not contain infinitely many points \(x_i\).

Let \(M_p(E)\) be the space of all point measures on \(E\) equipped with the smallest \(\sigma\)-field \(M_p(E)\) which contains all sets of the form
\[
\{m \in M_p(E) : m(A) \in B\}
\]
for any \(A \in \mathcal{E}\) and any Borel set \(B \subset [0, \infty]\), i.e., it is the smallest \(\sigma\)-field making the maps \(m \rightarrow m(A)\) measurable for all \(A \in \mathcal{E}\) [6].

**Definition 1.** A point process \(N\) with state space \(E\) is a measurable map from the underlying outcome space \(\Omega\) equipped with a \(\sigma\)-field \(\mathcal{F}\) to \((M_p(E), M_p(E))\).

In other words, a point process \(N\) is random element or a random function which assumes point measures as values: for every \(\omega \in \Omega\) the value \(m(\cdot) = N(\cdot, \omega)\) is a point measure. In particular, \(N(K) < \infty\) for compact sets \(K \subset E\). The next result gives a justification of the fact that we may interpret a point process \(N\) as a collection \((N(A))_{A \in \mathcal{E}}\) of the random variables \(N(A)\) with values in \(\{0, 1, \ldots, \infty\}\).

**Lemma 1.** The mapping \(N\) from \((\Omega, \mathcal{F})\) to \((M_p(E), M_p(E))\) is a point process on \(E\) if and only if, for every \(A \in \mathcal{E}\), \(N(A)\) is a random variable with values in \(\{0, 1, \ldots, \infty\}\) such that \(N(A) < \infty\) for compact \(A \subset E\).

3. Truncated Exponential Distribution
The truncated exponential model is based on Gutenberg-Richter magnitude recurrence relation. The Gutenberg-Richter recurrence relation is expressed as:
\[
\log N(M) = a - bM
\]
where \(N(M)\) is the mean number of earthquakes per unit volume and per unit time having magnitude greater than \(M\) [4]. The \(a\)-value represents the number of events per unit time with magnitudes greater than zero, which describes the seismic activity rate in the area of interest. The \(b\)-value defines the ratio of small and large magnitude events and typically takes a value between 0.8-1.0 [3], [5]. Equation (1) describes the frequency of earthquakes as a power of
magnitude. Therefore, exponential probability distribution would describe the probability of occurrence rate of earthquake.

Although the standard Gutenberg-Richter recurrence relation can be applied to an infinite range of magnitudes, it is common to apply bounds at minimum and maximum magnitude values. This is because seismic sources are usually associated with a capacity to produce some maximum magnitude and for engineering purposes earthquakes of very small magnitudes that do not cause damage to structures are not of interest. If parameter of magnitude distribution is $\beta$ then the corresponding probability density function using these minimum and maximum values is expressed below in its bounded form [3]:

$$f(m) = \frac{\beta e^{-\beta m}}{e^{-\beta M_{\text{min}}}}$$

$M_{\text{min}}$ is the smallest magnitude considered in the model (here $M_{\text{min}} = 5$) and $M_{\text{max}}$ is the largest magnitude event which can be expected in the region. Different $M_{\text{max}}$ values may be assigned in different regions.

4. Self Exciting Point Process

A spatio-temporal point process $N$ is mathematically defined as a random measure on a region $S \subseteq \mathbb{R} \times \mathbb{R}^3$ of space and time, taking values in the non-negative integers $\mathbb{Z}^+$. In this framework the measure $N(A)$ represents the number of points falling in the subset $A$ of $S$. Any analytical spatiotemporal point process is characterized uniquely by its associated conditional intensity process $\lambda[6]$. $\lambda(t, x, y|H_t)$ may be thought of as the frequency with which events are expected to occur around a particular location $(t, x, y)$ in spacetime, conditional on the prior history, $H_t$, of the point process up to time $t$ [11].

A self exciting point processes can be used for modeling events that are clustered together in time and space. Consider events $\{(t_i, x_i, y_i, m_i) : i = 1 \ldots , N\}$, where $(x_i, y_i) \in D$ is location of $i$-th earthquake, $D \subset \mathbb{R}^2$, $t_i \in [0, T]$ is time of $i$-th earthquake, $m_i > m_0$ is magnitude of $i$-th earthquake, $m_0$ is cut off magnitude. The self-exciting point process captures the tendency for earthquakes to spawn aftershocks, clustered about their parents [10]. The conditional intensity function is

$$\lambda(t, x, y|H_t) = \mu(x, y) + \sum_{i : t_i < t} k(m_i)g(t - t_i)v(x - x_i, y - y_i|m_i)$$

where

(i) the function $\mu(x, y)$ represents the deterministic background rate and is assumed to take the semi-parametric form

$$\mu(x, y) = \mu u(x, y)$$

where $\mu$ is an unknown constant and $u(x, y)$ is an unknown function

(ii) mainshocks of magnitude $m$ produce a number of aftershock with distribution $k(m)$

(iii) aftershocks are spatially distributed about their parents according to the bivariate density function $v(\cdot)$

(iv) aftershocks are temporally distributed about their parents according to a nonnegative function $g(\cdot)$.

In application we make use of the specific function

$$k(m) = Ae^{\alpha(m-m_0)}$$

(4)
Table 1. Window Parameters

| $M$  | $\Delta R$ (km) | $\Delta T$ (days) |
|------|-----------------|-------------------|
| 2.5  | 19.5            | 6                 |
| 3.0  | 22.5            | 11.5              |
| 3.5  | 26              | 22                |
| 4.0  | 30              | 42                |
| 4.5  | 35              | 83                |
| 5.0  | 40              | 155               |
| 5.5  | 47              | 290               |
| 6.0  | 54              | 510               |
| 6.5  | 61              | 790               |
| 7.0  | 70              | 915               |
| 7.5  | 81              | 960               |
| 8.0  | 94              | 985               |

for number of aftershocks. The temporal aftershocks follow Omori’s Law: the frequency of aftershocks at time $t$ after the main shock decays hyperbolically [12]:

$$g(t) = (p - 1)e^{p-1}(t + c)^{-p}$$

and the spatial distribution of aftershock is determined by the Gaussian density function [7]:

$$v(x, y|m) = \frac{1}{2\pi d\alpha(m-m_0)}exp\left\{-\frac{1}{2d\gamma(m-m_0)}\right\}$$

For determining the intensity function (3) the declustering algorithm is applied. We focus on the window algorithms, where magnitude-dependent windows in space and time are used to identify earthquakes of the same cluster. Here we use the window method by Gardner and Knopoff to identify the mainshocks (Table 1) [1], [2]. Furthermore we should find the background intensity $u(x, y)$ first to estimate the conditional intensity function. An estimator of the background intensity can be estimated using thinning procedure and the model parameters $\theta = (\mu, A, c, \alpha, p, d, q, \gamma)$ are estimated simultaneously using an iterative approach [13], [14], [15]. The parameter vector $\theta$ is estimated by maximizing the log-likelihood function

$$l(\theta) = \sum_i \lambda(t_i, x_i, y_i|H_i) - \int \lambda(t, x, y|H)dx dy dt.$$

5. Earthquake modeling

In this research we use the earthquake data in Java Island obtained from U.S. Geological Survey. The data was restricted to events with time period January 1973 - December 2010 from the rectangular area 5.46° - 11°S and 104.98° - 114.87°E. The data are shorted in ascending magnitude order and the cumulative number of events $N$, of magnitude $M$ or greater is found. A straight line fit to the plot of log $N$ versus $M$ yields Gutenberg-Richter recurrence model with $R^2 = 99\%$ and $p$-value = 0. Figure 1 describes the frequency of earthquakes as a power of magnitude. Therefore, exponential probability distribution would describe the probability of occurrence rate of earthquakes. We take $M_{min} = 5$ and $M_{max} = 7.9$ and estimate $\beta$ in equation (2) using maximum likelihood method. Hence we have $\beta = 3.235$ and we obtain the magnitude distribution. Furthermore we apply Gardner-Knopoff window to identify mainshocks.
and aftershocks. Here we consider shallow deep (depth ≤ 70 km)\[^8\] with magnitude ≥ 5 for mainshocks.

**Algorithm 1**

(i) Identify an event with magnitude ≥ 5 and depth ≤ 70 km. Record this event by \(i\) with magnitude \(m_i\), location \((x_i, y_i)\) and occurrence time \(t_i\), \(i = 1, 2, \ldots\).

(ii) For every event after \(i\), say \(ij\), \(j = 1, 2, \ldots\) calculate

(a) \(\Delta t_{ij} = |t_{ij}| - |t_i|\)
(b) \(\Delta x_{ij} = |x_{ij}| - |x_i| \times 111.3195\)
(c) \(\Delta y_{ij} = |y_{ij}| - |y_i| \times 111.3195\)
(d) \(\Delta r_{ij} = \sqrt{\Delta x_{ij}^2 + \Delta y_{ij}^2}\)

Calculation is done for events \(ij\) having magnitude ≤ \(m_i\) and satisfy window parameters (Table 1), i.e. \(\Delta r_{ij} \leq \Delta R\) and \(\Delta t_{ij} \leq \Delta T\).

(iii) Stop step (ii) if Gardner-Knopoff criteria is not satisfied, go to step (i).

**Algorithm 2**

(i) Input the length of the time intervals, number of events, starting points of time, the magnitude threshold, the starting point of time target interval, the number of points for the polygon region, number of grids for background.

(ii) Input an initial \(u_0(x, y)\).

(iii) Input initial parameters.

(iv) Fit the conditional intensity function \(\lambda(t, x, y | H_t)\).

(v) Calculate the background rate at the position of all events \(\hat{\mu}(x, y)\).

(vi) Calculate probability of the \(j\)th earthquake is trigger by \(i\)th earthquake \(\rho_{i,j}\).

(vii) Calculate probability of the \(j\)th event belongs to the background \(\phi_j\).

(viii) Steps (iv) - (vii) are repeated enough times such that the results converge.
Figure 2. Epicentral locations of mainshocks

Table 2. Maximum Likelihood Estimates for the Java earthquake

| parameter | estimate       |
|-----------|---------------|
| $\mu$     | 0.822674974   |
| $A$       | 0.325757138   |
| $c$       | 0.026135075   |
| $\alpha$  | 1.408691215   |
| $p$       | 1.325195590   |
| $d$       | 0.004627986   |
| $q$       | 1.501570185   |
| $\gamma$  | 1.020769545   |

By using Algorithm 1 we can identify 45 mainshocks from 491 events including three great earthquakes, the 2006 Yogyakarta earthquake, the 2006 Ciamis-Cilacap earthquake, and the 2009 Tasikmalaya earthquake. The epicentral locations of mainshocks is presented in Figure 2 while Table 2 shows the maximum likelihood estimates for parameters of the self-exciting point process obtained from Algorithm 2. This estimates have log likelihood = -1589.4533 and AIC = 3194.9. From the probability $\rho_j$ we can see that most of the events are either independent events or aftershocks. About 36.46% of the events have probability of being an aftershock. Figure 3 shows that high values of conditional intensity function are located in 112.019°E, 6.969°S (Java Sea, north of Tuban, East Java), 113.247°E, 7.15°S (Sampang, Madura Island), 114.521°E, 7.27°S (Madura Strait), 105.756°E, 10.429°S (Hindia Ocean, south of Banten Province).

Although we obtain maximum likelihood estimates for the parameters of the model, there remains some room for further improvement. In this model, information on shocks less than 5 in magnitude was not available. Modifications of (4) - (6) can be made to model variability in mean aftershock number and the spatio-temporal clustering of aftershocks. However we should be care to avoid over parametrization of model.

6. Conclusion
A probabilistic approach for earthquake data in this research is self-exciting point process. This model consist of three variables and eight parameters to be estimated simultaneously. Four high values of the conditional intensity function in Java Island are located in the north of Tuban regency, Sampang regency, Madura strait, and the south of Banten Province. From the
algorithm, we also obtain the probability of the event being an aftershock.

7. Acknowledgments
We are grateful to Universitas Sebelas Maret for financial support through Grant of Maintenance Research Group 2015. We also thank referees for their critical and helpful comments.

8. References
[1] Christophersen, A., M.C. Gerstenberger, D.A. Rhoades, and M.W. Stirling, Quantifying the Effect of Declustering on Probabilistic Seismic Hazard, *Proceedings of the Ninth Pacific Conference on Earthquake Engineering Building on Earthquake-Resilient Society*, Auckland, New Zealand, 2011.
[2] Gardner, J.K. and L. Knopoff, Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian?, *Bulletin of the Seismological Society of America*, vol. 64, No. 5, pp. 1363-1367, 1974.
[3] Godinho, J., Probabilistic Seismic Hazard Analysis - An Introduction to Theoretical Basis and Applied Methodology, *Dissertation*, University of Patras, Greece, 2007.
[4] Knopoff, L., The Magnitude Distribution of Declustered Earthquakes in Southern California, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 97, No. 22, pp. 11880-11884, 2003.
[5] Marzocchi, W. and L. Sandri, A Review and New Insights on the Estimation of the b-value and Its Uncertainty, *Annals of Geophysics*, Vol. 46, No. 6, pp. 1271-1282, 2003.
[6] Mikosch, T., *Non-Life Insurance Mathematics*, 2nd ed., Berlin: Springer-Verlag, 2009.
[7] Musmeci, F. and D. Vere-Jones, A Space-Time Clustering Model for Historical Earthquakes, *Annals of the Institute of Statistical Mathematics*, Vol. 44, No. 1, pp. 1-11, 1992.
[8] NEIC, *M7.5 Northern Peru Earthquake of 26 September 2005*. National Earthquake Information Center, U.S. Geological Survey, October 17, 2005.
[9] Pratiwi, H., Subanar, and Danardono, Analysis of Earthquake Data Based on Space-Time Epidemic Type Aftershocks Sequence Model, *International Journal of Applied Mathematics and Statistics*, vol. 52, No. 5, 2014.
[10] Rathbun, S. L., Asymptotic Properties of the Maximum Likelihood Estimator for Spatio-Temporal Point Process, *Journal of Statistical Planning and Inference*, vol. 51, pp. 55-74, 1996.
[11] Schoenberg, F. P., D. R. Brillinger, and P. Guttorp, Point Processes, Spatial-Temporal, *Encyclopedia of Environmetrics*, Vol. 3, pp. 1573-1577, 2002.
[12] Vere-Jones, D., Seismology - A Statistical Vignette, *Journal of the American Statistical Association*, Vol. 95, No. 451, pp. 975-978, 2000.
[13] Zhuang, J., Second-order Residual Analysis of Spatio-temporal Point Processes and Applications in Model Evaluation, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 4, No. 68, pp. 635-653, 2006.
[14] Zhuang, J., Y. Ogata, and D. Vere-Jones, Analyzing Earthquake Clustering Features by Using Stochastic Reconstruction, *Journal of Geophysical Research*, Vol. 109, No. B5, B05301, 2004.
[15] Zhuang, J., Y. Ogata, and D. Vere-Jones, Stochastic Declustering of Space-Time Earthquake Occurrences, *Journal of the American Statistical Association*, Vol. 97, No. 458, pp. 369-380, 2002.