Scientific Reports

https://doi.org/10.1038/s41598-025-30205-z

Article in Press

Inhomogeneous spatio-temporal epidemic-type aftershock sequence model incorporating seismicity-triggering slow slip events

Received: 31 March 2025

Accepted: 21 November 2025

Published online: 29 November 2025

Cite this article as: Bañales I., Nishikawa T., Ito Y. *et al.* Inhomogeneous spatio-temporal epidemic-type aftershock sequence model incorporating seismicity-triggering slow slip events. *Sci Rep* (2025). https://doi.org/10.1038/s41598-025-30205-z

Isaías Bañales, Tomoaki Nishikawa, Yoshihiro Ito, Vladimir Kostoglodov, Ekaterina Kazachkina & José Santiago

Inhomogeneous Spatio-Temporal Epidemic-Type Aftershock Sequence Model Incorporating Seismicity-Triggering Slow Slip Events

Isaías Bañales[1,*], Tomoaki Nishikawa[1], Yoshihiro Ito[1], Vladimir Kostoglodov[2], Ekaterina Kazachkina[2], and José Santiago[2]

[1]Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, 611-0011, Japan
[2]Instituto de Geofísica, Universidad Nacional Autónoma de México, Mexico City, 04510, Mexico
[*]isaias@ciencias.unam.mx

ABSTRACT

Clarifying the relationship between regular earthquakes and slow fault slip is essential for understanding the mechanisms behind seismic activity. We hypothesize that the background seismic activity is partially triggered by interplate slow-slip events (SSEs). Consequently, we present an extension of the spatio-temporal epidemic-type aftershock sequence (ETAS) model, which incorporates background seismicity as a piecewise constant function over time based on recent advances in the inference of space–time inhomogeneous point processes. In this study, Global Navigation Satellite System (GNSS) data is employed to identify the occurrence periods of SSEs, thereby delineating the intervals during which changes in background seismicity may occur. Due to the technical complexity of performing inference with an inhomogeneous ETAS model, this work employs a maximum likelihood inference method using the Expectation-Maximization (EM) algorithm. This approach also enables the inference of the branching process for aftershocks, allowing for the estimation of earthquake genealogy. This study elucidates how the background seismicity increases during the periods of SSEs in Guerrero, Mexico and Boso Peninsula, Japan, which allows for a more comprehensive understanding of seismic activity and the relationship between slow and fast earthquakes.

1 Introduction

At the boundaries between tectonic plates, two types of spontaneous and episodic fault slip phenomena occur: fast (regular) earthquakes and slow earthquakes. These fault slip phenomena are closely related. Among slow earthquakes, those with a relatively large magnitude (approximately Mw 5 or greater) that can be detected geodetically are referred to as slow slip events (SSEs)[1]. It has been observed that SSEs often trigger small to moderate earthquakes in the Sagami Trough subduction zone in Japan[2]. Furthermore, SSEs have preceded and possibly triggered megathrust earthquakes at several subduction plate boundaries[3,4].

Global Navigation Satellite System (GNSS) networks provide detailed daily information on crustal deformation and allow to detect and describe SSEs[5–8]. SSEs have also been discovered and studied in various regions worldwide, including, Middle America Trench[9], the Japan Trench[10,11], the Hikurangi subduction zone in New Zealand[12] , and Peru[13]. In these regions, SSEs occurring at the plate boundary are thought to significantly impact seismic activity, and therefore quantifying the impact of SSEs is essential for improving the accuracy of earthquake forecasts.

The modeling and forecasting of fast-earthquake activity in a stochastic context has been widely accepted by the seismological community since the presentation of the seminal article by Ogata[14], in which the epidemic-type aftershock sequence (ETAS) model is defined using the Hawkes process. Moreover, Zhuang et al.[15] made an important extension to the ETAS model by the inclusion of a spatio-temporal component in the intensity function (i.e., seismicity rate). In addition, Li et al.[16] have studied in detail the inference of the background intensity function (i.e., the background seismicity rate in seismological terms) in spatio-temporally inhomogeneous point processes, this work was a pillar in the development of the model presented in section 2.

The detection of aseismic transients and their relationship to seismicity have been extensively studied. There have been several studies of seismicity changes due to aseismic transients considering only the time domain[17–21]. For example, Okutani & Ide[22] investigated the impacts of SSEs on seismic activity using the temporal ETAS model. They proposed a model called the boxcar model, in which the background seismicity rate increases in a boxcar-like manner during the slow slip period estimated from geodetic observations. Their approach is similar to that one presented by Mattews & Reasenberg[23], who investigated the quiescence of microearthquakes through a temporally inhomogeneous Poisson process, using a piecewise constant function.

Nishikawa & Nishimura[24] presented a variant of the ETAS model that explicitly links an increase in background seismicity to detected SSE. Although their research has made significant advances in the modeling and forecasting of seismic activity associated with SSEs, it does not account for the spatio-temporal changes that may occur in background seismic activity during SSE periods.

As described above, extensive research has been conducted on time-domain analysis. In this work, however, the focus is on spatio-temporal seismicity modeling.

Llenos & McGuire[25] proposed a complex model that combined the ETAS model and the rate- and state-dependent friction seismicity model[26] to detect seismicity rate changes induced by aseismic transients. They adopted a tricky approach to subtract the coseismically triggered seismicity rate estimated by the conventional spatio-temporal ETAS model from the total seismicity rate and related the residual seismicity rate to the rate- and state-dependent friction seismicity model.

It is important to highlight the contributions of Marsan et al.[27] and Reverso et al.[28] to the spatio-temporal modeling of seismicity associated with aseismic transients. Both papers propose a way to model the evolution of the seismicity using a mesh over space and time. They used the conventional spatio-temporal ETAS model throughout the entire period as a null model, and for each earthquake occurring at time t_i in the location (x_i, y_i), they fit a locally elevated background intensity using earthquake records where (t_j, x_j, y_j) satisfy

$$
\begin{aligned}
|t_j - t_i| &< \frac{\tau}{2}, \\
|x_j - x_i| &< \frac{\mathscr{L}}{2}, \\
|y_j - y_i| &< \frac{\mathscr{L}}{2},
\end{aligned}
$$

for all n, where τ and $\mathscr{L}$ are parameters that control the size of the spatio-temporal window. If the locally estimated background intensity significantly differs from that of the null model, as determined by the criterion specified in each study, they replace the background intensity of the null model by the locally estimated value within the vicinity. Additionally, Reverso et al.[29] have presented a pioneering work relating the ETAS model with SSE, following the ideas presented in[28].

Possible improvements to the methods of Marsan et al.[27] and Reverso et al.[28] include the utilization of geodetic observations. In their methods, τ and $\mathscr{L}$ are subjectively chosen (e.g., τ is set to 1 day, 40 days, or 100 days). However, particularly for τ, the duration of an aseismic transient can sometimes be estimated based on geodetic data, which can then be used as τ.

Furthermore, they use the small spatio-temporal windows for each earthquake one by one to estimate the local background intensity via maximum likelihood estimation, sequentially comparing it to the null model. However, this is an approach adopted for simplicity. Ideally, the background intensities of multiple windows should be varied simultaneously to estimate the set of background intensities that maximize the likelihood.

In light of the aforementioned studies, this research proposes a new modification of the spatio-temporal ETAS model that incorporates the impact of SSEs on seismic activity (Section 2.2). Our model determines the periods in which the background intensity changes based on GNSS observations and simultaneously estimates the spatial distribution of the background seismicity rate for each period. This model is mathematically grounded by Li et al.[16].

We apply the new model to Mexican earthquakes in the Middle America subduction zone and elucidate the impact of SSEs on the background seismicity within this subduction zone, which is a topic that has not been addressed from the perspective of statistical modeling to date. In addition, the model was applied to the Boso Peninsula, located in the Sagami Trough subduction zone in Japan, a region that has been previously studied and has exhibited substantial changes in the seismic activity during SSE[22,30]. Our new model will be a useful tool in the future for elucidating the characteristics of seismic activity associated with SSEs worldwide.

2 Methodology

2.1 Introduction to Spatio-Temporal ETAS Model

The spatio-temporal ETAS model is a marked branching point process for earthquake occurrences, and its behavior can be completely defined through its conditional intensity function given by

$$
\begin{aligned}
&\mathbb{P}(\text{an event in } [t,t+dt]\times[x,x+dx]\times[y,y+dy] \\
&\qquad \times [M,M+dM]|\mathscr{H}_t) = \lambda(t,x,y,M|\mathscr{H}_t)dtdxdydM + o(dtdxdydM),
\end{aligned} \tag{1}
$$

where M is the magnitude, (x,y) denotes the spatial coordinates, t represents the elapsed time, and $\mathscr{H}_t$ denotes the space-time and magnitude occurrence history of the earthquakes up to time t[15]. In particular, based on its assumptions, the spatio-temporal

ETAS model has

$$\lambda(t,x,y|\mathscr{H}_t) = \mu(x,y) + \sum_{\{k:t_k<t\}} k(M_k)g(t-t_k)f(x-x_k,y-y_k|M_k), \tag{2}$$

where the subindex k refers to the values of the $k-$th event considered, with $k = 1,2,...,n$, (x_k,y_k) are its spatial coordinates, t_k is the time of occurrence, and M_k is its magnitude. Without loss of generality, in this work it is assumed that t_k, $k = 1,2,...,n$ are ordered in increasing order. Furthermore,

$$k(M) = Ae^{\alpha(M-M_0)} \tag{3}$$

$$g(t) = (p-1)c^{p-1}(t+c)^{-p}\mathbb{1}(t>0) \tag{4}$$

$$f(x,y|M) = \frac{1}{2\pi\sqrt{d_1 d_2}e^{\alpha(M-M_0)}}\exp\Big\{-\frac{1}{2}\frac{1}{e^{\alpha(M-M_0)}}\Big(\frac{x^2}{d_1}+\frac{y^2}{d_2}\Big)\Big\}, \tag{5}$$

where M_0 denotes a reference magnitude, typically the minimum observed value, although in some cases a higher cutoff magnitude is adopted to reduce computational load[31], and

$$\mathbb{1}(t>0) = \begin{cases} 1, & \text{if } t>0 \\ 0, & \text{otherwise} \end{cases}.$$

Different authors[15,32] have used the Gaussian probability density function in equation (5). In this work, we assume that $d_1 = d_2$ for the sake of parsimony, since the aftershock classification does not show significant differences compared to the case $d_1 \neq d_2$ for the Mexican data, as can be seen in Figures 6 and 15a.

Another popular option is to use a Pareto distribution[33–35], given by

$$f_{\text{Pareto}}(x,y|M) = \frac{q-1}{\pi\sqrt{d_1 d_2}e^{\gamma(M-M_0)}}\Big(1+\frac{1}{e^{\gamma(M-M_0)}}\Big(\frac{x^2}{d_1}+\frac{y^2}{d_2}\Big)\Big)^{-q},$$

which can also be seen as a particular case of the bivariate t-distribution. The advantage of using the Pareto distribution is to avoid overestimate the background seismicity function. Nevertheless, for our Mexican data, the heavy tails of the Pareto distrubtion classify as aftershocks earthquakes that are unrealistically far from their respective mainshocks as it can be seen in Figure 15b in Appendix App. 1.

By defining $\mu(x,y) = \nu u(x,y)$ and assuming stationarity, Zhuang et al.[15] propose the estimator of μ as

$$\hat{\mu}(x,y) = \frac{1}{T}\sum_j (1-\rho_j)\frac{1}{2\pi d_j^2}\exp\Big\{-\frac{x^2+y^2}{2d_j^2}\Big\}. \tag{6}$$

where d_j is a bandwidth that depends on how many earthquakes are close to the event j, and $1-\rho_j$ is the probability that the $j-$event is an immigrant (i.e., background event).

$\hat{\mu}$ is fitted in an iterative two-step procedure, in the first iteration the vector $\eta = (\nu,A,\alpha,c,p,d)$ is fitted, in the second step the vector η is taken as known and $\hat{\mu}$ is updated until the convergence of the log-likelihood

$$\ell(\eta) := \sum_{k=1}^{n}\log(\lambda_\eta(t_k,x_k,y_k|H_{t_k})) - \int_0^T\iint_S \lambda_\eta(t,x,y|H_t)dxdydt, \tag{7}$$

is reached, where the analysis time is $[0,T]$, and S is the analysis region.

It is worth mentioning the novel nonparametric Bayesian approaches to model μ. Ross & Kolev[32] also assume that μ fulfills $\mu(x,y) = \nu u(x,y)$, where u is a probability density function and ν is a positive real number. This allows the use of a mixture of Dirichlet processes (MDP)[36] as the prior for $u(x,y)$. On the other hand, Molkenthin et al.[35] assume that μ can be written as

$$\mu = \frac{\nu}{1+e^{-w(x,y)}},$$

where a Gaussian Process (GP) prior is used for w. The main advantage of using the MDP approach is that the function u always integrates 1, since it is a probability density function. In the case of GP, the integral can not be solved analytically; nevertheless, using the GP approach avoids the need for a finite approximation of the infinite mixture required in the MDP case.

It is important to mention that Veen & Schoenberg[37] discuss in detail the numerical stability problems of maximizing the likelihood of the spatio-temporal ETAS model directly. They proposed using the Expectation-Maximization (EM) algorithm to improve the inference performance, this idea was one of our motivations to develop our model (Section 2.2).

Extending the ideas presented by Veen & Schoenberg[37], Fox et al.[38] modeled the background intensity rate as a piecewise constant function, which allows a numerically efficient way to realize the approach proposed by Veen & Schoenberg[37] with spatial inhomogeneities. The modeling of the background seismicity as a piecewise function is also used by different authors[39,40] to infer the intensity function. Our approach (Section 2.2) intends to advance the method proposed by Fox et al.[38] by considering temporal inhomogeneities that will act as triggering effects due to SSEs.

Another advantage in following the approach presented by Veen & Schoenberg[37] and Fox et al.[38] is that the structure of the branching process is inferred in addition to estimating the intensity function. Therefore, we can easily distinguish which earthquakes are background events and which are aftershocks, as shown in the results obtained in section 4.

Since the ETAS model was introduced by[14], the completeness of catalogs remains crucial for the correct estimation of the model. This is the reason that authors such as Seif et al.[33] have made significant efforts to analyze the consequences of missing recorded aftershocks on the biases of the estimators. In their work, they argue that a possible solution to avoid incompleteness in aftershock sequences after a strong earthquake is to use a large cutoff magnitude. However, this can result in some seismic activity as foreshocks are not being observed and it is important to consider that the inclusion of small earthquake reduce the bias in the ETAS model estimators.

Short-term aftershock incompleteness has been extensively studied in sismology[41,42].This is a problem of blindness, whereby small earthquakes are undetectable when the signal is saturated by a strong earthquake. Empirical relationships have been propose to model the STAI effect[43]. In the context of the ETAS algorithm, in recent years Hainzl[31] has proposed the ETASI model, which incorporates the STAI phenomena into the ETAS model by adding a temporary dependence on the number of expected events. The ETASI model has been recently extended by Asayesh et al.[44], where the spatial kernel of the ETASI model was modified to incorporate information from the stress scalars.

Although STAI is an important phenomenon to take into account, the cutoff magnitudes used in this study are not low enough to appreciate this phenomenon as it can be seen in Figure 13 in Appendix App. 1. Therefore, we chose to keep b-values and cutoff magnitudes constant throughout the entire period.

We also wish to point out that, in order to model spatial variation in aftershock activity parameters, authors such as Ogata[45] and Ueda et al.[46] extended the spatio-temporal ETAS model[15], allowing parameters such as p and the productivity K to become spatially varying functions. However, the issue of identifiability in their approach required the imposition of a smoothness penalization on the functions p, and K. Nevertheless, extending μ, K, and p to functions of (x,y,t) leads to identifiability problems, since changes in seismicity induced by strong earthquakes can occur abruptly in both space and time, rendering smoothness penalization ineffective. For this reason, our work only considers temporal inhomogeneity through μ.

2.2 Model

Following the idea of Nishikawa and Nishimura[24], we model the inhomogeneities in the ETAS model through changes in the background seismic activity. The reason for this is that μ is usually related to tectonic loading and the relative velocity of plate motion[47]. If we regard slow slip as an increase in tectonic loading or interplate slip rate, it is natural to model it through changes in the background seismic activity. However, we cannot deny the impact of SSEs on aftershock activity. Addressing this issue is an important direction for future work.

Since the aim of this work is to model the triggering effect of SSEs, the assumption of stationarity may not be realistic. In addition, Veen & Schoenberg[37] have discussed that the approach introduced by Zhuang et al.[15] is numerically expensive and unstable because it is necessary to optimize (7).

To describe the triggering effect due to SSEs, this work proposes to use

$$\lambda(t,x,y|\mathscr{H}_t) = \mu(x,y,t) + \sum_{\{k:t_k<t\}} k(M_k)g(t-t_k)f(x-x_k,y-y_k|M_k),$$

which is an extension of (2) because it allows a time dependent μ function. However, due to the complexity of working with an arbitrary form of μ, in this study, it is defined as

$$\mu(x,y,t) = \sum_{i=0}^{m} \mu^i(x,y)\mathbb{1}(t \in S_i,E_i), \tag{8}$$

where m is the total number of SSEs and S_i,E_i are the start and end times of the $i-$th SSE, with $i = 1,\ldots,m$. In the case of $i = 0$, $S_i = 0, E_i = T$, i.e. μ^0 represents the background seismicity over the entire period, regardless of the presence of an SSE, and the remaining μ^i with i=1,2,...,m represent the increase of the background seismicity with respect to μ^0. Thus, the intensity is given by

$$\lambda(t,x,y|\mathscr{H}_t) = \sum_{i=0}^{m} \mu^i(x,y)\mathbb{1}(t \in S_i,E_i) + \sum_{\{k:t_k<t\}} k(M_k)g(t-t_k)f(x-x_k,y-y_k|M_k). \tag{9}$$

As presented by Veen & Schoenberg[37] and discussed in detail by others[48,49], the Hawkes process could be defined as a marked Poisson cluster process where there are two kind of events, immigrants (background events) and offspring (aftershocks), which allows the use of the EM algorithm to maximize the likelihood.

McLachlan & Krishnan[50] discussed the EM algorithm in detail and showed that it is useful when there is missing information which, if known, makes the maximization of the complete likelihood easier. The EM algorithm is an iterative algorithm that maximizes the likelihood using a two-step procedure, the first step (Expectation) replaces the unknown information by the expected one with the help of an initial value for all the parameters, and the second step (Maximization) uses the expected values obtained in the previous step to optimize the likelihood defined by the augmented information, these two steps are repeated until the convergence of the augmented likelihood.

The development of the model in this section is mathematically based on that proposed by Li et al.[16], but the definition of $\lambda(t,x,y|\mathscr{H}_t)$ by them differs from that presented in (9). In their work they assumed

$$\mu(x,y,t) = \alpha u(x,y)v(t),$$

where α is in $\mathbb{R}^+$, and u and v are positive functions. Since SSEs are spontaneous large fault slip events with variable slip evolution, the spatial and temporal contributions to background seismicity cannot be assumed to form a product over the entire period. Therefore, we preferred to use the expression in (8).

It is also important to mention that they are not working with a marked point process and that their intensity function does not incorporate information about event magnitudes. However, they use the kernel of a gamma distribution for the decay in the aftershock activity over the time with parameter $\alpha = 1$:

$$g^L(t) = \beta e^{-\beta t}$$

where β is in $\mathbb{R}^+$. Their expression differs from (4), proposed by Zhuang et al.[15], which is based on Omori's law[51] and is also adopted in our study.

From the observed data, it is not known whether an earthquake is a background event or an aftershock, and if it is a background event, it is also not known whether it comes from the process with intensity μ^0 or from μ^i when the $i-$th SSE occurs. Therefore, the following random variables are defined:

$$\chi_{ii}^s = \begin{cases} 1, & \text{if earthquake } i \text{ is a background event, and it is produced by } \mu^s \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

$$\chi_{ij} = \begin{cases} 1, & \text{if earthquake } i \text{ is an aftershock of } j \\ 0, & \text{otherwise} \end{cases}, \tag{11}$$

where $s = 0,\ldots,m$.

The random variables defined in (10) and (11) are an extension of those presented in the supplementary material by Fox et al.[38], where only one χ_{ii} is introduced for the background intensity over the entire period, while here the triggering effect of multiple SSEs is modeled in a similar way to the model by Li et al.[16]. In the present work, $\mu^i(x,y)$ is defined as a piecewise constant function for all i:

$$\mu^i(x,y) = \sum_{u=1}^{n_y}\sum_{v=1}^{n_x} \mu_{uv}^i \mathbb{1}((x,y) \in D_{uv}), \tag{12}$$

where $D_{kl} = ((u-1)\Delta x, u\Delta x) \times ((v-1)\Delta y, v\Delta y)$, Δx and Δy are the step size in the partition of the x and y axes, also n_x and n_y are the number of grids for each axis. The advantages of defining $\mu^i(x,y)$ in this way is to facilitate the integral over the space in (17) to recover a closed expression in the maximization step, as it can be seen in (19). The grid applied in this study is shown in figure 1.

The total amount of parameters to be estimated by (12) are $n_y n_x$ for each μ^i, with $i \in 0,1,\ldots m$. Furthermore, because of (3) to (5), we have the five aftershock parameters (A,α,c,p,d) to be estimated. Additionally, we must estimate the whole branch structure given by $n(n-1)/2$ elements in χ_{ij} and $\sum_{i=1}^m n_i$ elements in χ_{ii}^s, where n_i is the total number of earthquakes that occurred during (S_i,E_i).

In this work, the array that contains all the parameters of the previous paragraph is denoted by θ, i.e.

$$\theta = (\{\chi_{ii}\}_{i\in I}, \{\chi_{ij}\}_{ij\in I\times I}, A, \alpha, c, p, d).$$

where $I = \{1, 2, ..., n\}$ and $\times$ denotes the cartesian product. It is important to note that the enormous amount of parameters will produce a not well-posed problem, the optimization of the likelihood has been constrained, following[52] according to the expression (20). This study assumes that the parameters controlling the aftershock intensity (A, α, c, p, d) do not change over time.

If the branch structure were known, the log-likelihood of the complete information (ℓ_c) would be defined as

$$\ell_c(\theta) = \ell_{\mathrm{O}}^*(\theta) + \ell_{\mathrm{I}}^*(\theta),$$

where $\ell_{\mathrm{O}}(\theta)$ is the log-likelihood due to the offspring, and $\ell_{\mathrm{I}}(\theta)$ is the log-likelihood due to immigrants. They are given by

$$\ell_{\mathrm{I}}^*(\theta) = \sum_{s=0}^{m} \Big[\sum_{i=1}^{n} \chi_{ii}^s \log(\mu_\theta^s(x_i, y_i)) \mathbb{1}(t_i \in (S_s, E_s)) - \int_0^T \iint_S \mu_\theta^s(x_i, y_i) \mathbb{1}(t \in (S_s, E_s)) dx dy dt \Big] \tag{13}$$

$$\ell_{\mathrm{O}}^*(\theta) = \sum_{j=1}^{n} \Big[\sum_{i>j} \chi_{ij} \log\Big(k_\theta(M_j) g_\theta(t_i - t_j) f_\theta(x_i - x_j, y_i - y_j | M_j) \Big) - \int_{t_j}^T \iint_S k_\theta(M_j) g_\theta(t - t_j) f_\theta(x - x_j, y - y_j | M_j) dx dy dt \Big], \tag{14}$$

where the subindex θ in $k_\theta, g_\theta, f_\theta$ and μ_θ^s means that the functions k, g and f are defined using the parameters (A, α, c, p, d) given by the array θ where the log-likelihood is evaluated.

It is important to note that (10) and (11) are not independent random variables, the sum of all entries in the vector $(\chi_{ii}^0, ..., \chi_{ii}^m, \chi_{i,1}, ..., \chi_{i,i-1})$ is always 1 for all $i = 1, ..., n$, i.e., it is a vector with a multinomial distribution. This means that if we have a set of parameters θ_r, the expected value of the entries for all i in $1, 2, ..., n$ is given by

$$\mathbb{E}(\chi_{ii}^s | \theta_r) = \hat{p}_{ii}^s = \frac{\mu^s(x_i, y_i)}{\lambda_{\theta_0}(t_i, x_i, y_i | H_{t_i})} \mathbb{1}(t_i \in S_s, E_s) \quad \text{with } s = 0, 1, ..., m \tag{15}$$

$$\mathbb{E}(\chi_{ij} | \theta_r) = \hat{p}_{ij} = \frac{k(M_j) g(t_i - t_j) f(x_i - x_j, y_i - y_j | M_j)}{\lambda_{\theta_0}(t_i, x_i, y_i | H_{t_i})} \quad \text{with } j = 1, ..., i-1. \tag{16}$$

The equations (15) and (16) correspond to the expectation step in the EM algorithm. The next step is to replace the χ values in (13) and (14) by their expected values. Using θ_r as the initial parameter value,

$$\ell_{\mathrm{I}}(\theta_r) = \sum_{s=0}^{m} \Big[\sum_{i=1}^{n} \hat{p}_{ii}^s \log(\mu_{\theta_r}^s(x_i, y_i)) \mathbb{1}(t_i \in (S_s, E_s)) - \int_0^T \iint_S \mu_{\theta_r}^s(x_i, y_i) \mathbb{1}(t \in (S_s, E_s)) dx dy dt \Big], \tag{17}$$

$$\ell_{\mathrm{O}}(\theta_r) = \sum_{j=1}^{n} \Big[\sum_{i>j} \hat{p}_{ij} \log\Big(k_{\theta_r}(M_j) g_{\theta_r}(t_i - t_j) f_{\theta_r}(x_i - x_j, y_i - y_j | M_j) \Big) - \int_{t_j}^T \iint_S k_{\theta_r}(M_j) g_{\theta_r}(t - t_j) f_{\theta_r}(x - x_j, y - y_j | M_j) dx dy dt \Big]. \tag{18}$$

It is important to note that since the real branch structure (genealogy) of the earthquakes is unobservable, the equations (13) and (14) cannot be evaluated. However, when the real values are replaced by their expected values as in (17) and (18), they can be computed.

Another advantage of the EM approach is that all parameters concerning μ^i are only in (17), while the parameters (A, α, c, p, d) are only in (18), therefore the optimization can be done separately. To solve (18) it is important to note that

$$\int_{t_j}^T \iint_S k_{\theta_r}(M_j) g_{\theta_r}(t - t_j) f_{\theta_r}(x - x_j, y - y_j | M_j) dx dy dt$$

can be rewritten as

$$k_{\theta_r}(M_j) \int_{t_j}^T g_{\theta_r}(t - t_j) dt \iint_S f_{\theta_r}(x - x_j, y - y_j | M_j) dx dy.$$

The utility of taking the expressions in (4) and (5) is that the integrals can be solved easily, since

$$\int_{t_j}^{T} g_{\theta_r}(t-t_j)dt = 1 - c^{p-1}(T-t_j+c)^{-p+1}$$

and the spatial integral

$$\iint_S f_{\theta_r}(x-x_j, y-y_j|M_j)dxdy,$$

which corresponds to the double integral of a bivariate Gaussian distribution, which can be solved efficiently using Monte Carlo methods. As an additional note,[53] suggests the following approximation with different k, g and f functions

$$\int_{t_j}^{T}\iint_S k_{\theta_r}(M_j)g_{\theta_r}(t-t_j)f_{\theta_r}(x-x_j, y-y_j|M_j)dxdydt \approx k_{\theta_r}(M_j),$$

this approximation holds if almost all aftershock activity occurs in the observed space and time, which is a very easy assumption to satisfy and it can lead to produce important reductions in computing time. We mention this approach because it has been used by different authors[32,38], nevertheless, the implementation of algorithm 1 in this work does not use it because, as mentioned in[54], it can generate inaccuracies in the estimation process.

Regarding (17), it is easy to find that the solution of

$$\frac{\partial}{\partial \mu_{uv}^j}\sum_{s=0}^{m}\Bigg[\sum_{k=1}^{n}\hat{p}_{ii}^s \log(\mu_{\theta_r}^s(x_k,y_k))\mathbb{1}(t_k \in (S_s,E_s)) - \int_0^T\iint_S \mu_{\theta_r}^s(x_k,y_k)\mathbb{1}(t \in (S_s,E_s))dxdydt\Bigg] = 0$$

is

$$\hat{\mu}_{uv}^j = \frac{\sum_{k=1}^{n}\hat{p}_{ii}^j \mathbb{1}((x_k,y_k)\in D_{uv})\mathbb{1}(t_k \in (S_j,E_j))}{(E_j-S_j)\Delta x\Delta y}. \tag{19}$$

The main advantage of using the piecewise expression of μ^i is to retrieve a closed expression in (19), which enables a fast updating of the parameters. However, the main weakness of this approach is that only the earthquakes inside of D_{kl} update the value of μ_{kl}^i, and a smooth estimation of μ^i is not achieved.

To address the not well-conditioned problem we restrict the optimization problem, the parameter A must be less than 1, i.e. we are assuming that the smallest earthquake have less than one expected aftershock. Regarding the α value we followed the idea of Ross & Kolev[32] of use the Helmstetter et al.[52] inequalities.

Assuming the Gutenberg-Richter law[55], the random variable $(M-M_0)$ is distributed as an exponential random variable[56] with rate $b\log(10)$, where b is the $b-$value of the Gutenberg-Richter law. Then, we can calculate the average number of offspring created per event, which is defined by

$$r := \int_{M_0}^{\infty} Ae^{\alpha(M-M_0)}b\log(10)e^{-b\log(10)(M-m_0)}dM.$$

In order to guarantee that r is finite and it is less than one, the following conditions must be fulfilled[52]

$$\alpha < b\log(10), \quad \frac{Ab\log(10)}{b\log(10)-\alpha} < 1. \tag{20}$$

Consequently, in this work we allow only parameters of A and α that satisfy the inequalities in (20).

During this study, we assumed a constant b-value over time and across the entire study region. This value was estimated using the maximum likelihood method[57]. Tests were also performed in Mexico using the b-positive estimator[42] and the novel estimator presented by Lippiello & Petrillo[58] called b-more-positive; however, the ETAS model estimators did not change significantly. Furthermore, Table 1 in Appendix App. 1 reports estimations assuming b-value lower and higher than those obtained via MLE, b-positive method and b-more-positive and the parameters did not show any important change with respect to those presented in section 4.1.

Although the choice of the MLE estimator or the family of b-positive estimators does not generate significant changes in our ETAS model estimators, we note that Nandan et al.[59] introduced a variant of the ETAS model that allows different b-values

to be considered. They speculate that the magnitude of the mainshock may affect the b-value for its triggered earthquakes. Also, Ito & Kaneko[60], who, using continuum models of fully dynamic earthquake cycles with fault frictional heterogeneities, have observed that the b-value of foreshocks decreases with time prior to the mainshock.

Furthermore, as noted by[61], the b-value also exhibits a spatiotemporal distribution, which we plan to analyze in detail in future work.

For c and p parameters, we restrict the acceptable values to be between $[0,5]$ and $[1,2]$. We are being very flexible in our boundaries, since Holschneider et al.[62] have observed that c and p are weakly identifiable.

In the case of d, with units deg^2, we restrict the optimization algorithm to be in (0,1), the idea of take 1 as the upper limit is be noninformative.

The maximization step in the EM algorithm consists in solving

$$\theta_{r+1} = \underset{\theta}{\operatorname{argmax}} \quad \ell_{\mathrm{I}}(\theta_r) + \ell_{\mathrm{O}}(\theta_r) \tag{21}$$

Finally, by replacing r by $r+1$ the Expectation and Maximization steps must be iterated until $|\theta_r - \theta_{r+1}| < \varepsilon$, where $\varepsilon > 0$ is the convergence criterion. In this work, the optimization was done using the implementation of the *L-BFGS-B* algorithm in *python* using the implementation in *scipy*[63].

As a summary of the EM algorithm that is used in section 4 to detect the earthquake triggering effect of SSEs, the pseudo code is presented in the algorithm 1.

Algorithm 1 Expectation Maximization Algorithm of Inhomogeneous Spatio-Temporal ETAS

1: Input values: Initial values of θ, tolerance level ε
2: Start $r \leftarrow 0$, define $\theta_r = \theta$, $\varepsilon_r = \varepsilon$
3: **while** $\varepsilon_r \geq \varepsilon$ **do**
4: Calculate $\hat{\mu}^i_{kl}$ for i in 0,1,...,m, with k in $1,2,...,n_x$ and l in $1,2,...,n_y$ as in (19)
5: Find (A,α,c,p,d) that maximize $\ell_{\mathrm{O}}(\theta_r)$
6: Calculate $\hat{p}^s_{ii}$ and $\hat{p}_{ij}$ as it was presented in (15) and (16)
7: Define θ_{r+1} using $\hat{\mu}^i_{kl}$ from line 4, (A,α,c,p,d) from line 5, and $\hat{p}^s_{ii}$ and $\hat{p}_{ij}$ from line 6
8: $\varepsilon_r \leftarrow |\theta_r - \theta_{r+1}|$
9: $r \leftarrow r+1$
10: **end while**
11: Return θ_r

The initial values of $\hat{p}^s_{ii}$ and $\hat{p}_{ij}$ used in Algorithm 1 are defined by the following matrices

$$P^O = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & \dots & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \dots & 0 \end{pmatrix},$$

$$P^I = \begin{pmatrix} \frac{\mathbb{1}(t_1\in(S_0,E_0))}{\sum_{i=0}^m \mathbb{1}(t_1\in(S_i,E_i))} & \frac{\mathbb{1}(t_1\in(S_1,E_1))}{\sum_{i=0}^m \mathbb{1}(t_1\in(S_i,E_i))} & \cdots & \frac{\mathbb{1}(t_1\in(S_m,E_m))}{\sum_{i=0}^m \mathbb{1}(t_1\in(S_i,E_i))} \\ \frac{\mathbb{1}(t_2\in(S_0,E_0))}{2\sum_{i=0}^m \mathbb{1}(t_2\in(S_i,E_i))} & \frac{\mathbb{1}(t_2\in(S_2,E_2))}{2\sum_{i=0}^m \mathbb{1}(t_1\in(S_i,E_i))} & \cdots & \frac{\mathbb{1}(t_2\in(S_m,E_m))}{2\sum_{i=0}^m \mathbb{1}(t_2\in(S_i,E_i))} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathbb{1}(t_n\in(S_0,E_0))}{n\sum_{i=0}^m \mathbb{1}(t_n\in(S_i,E_i))} & \frac{\mathbb{1}(t_n\in(S_2,E_2))}{n\sum_{i=0}^m \mathbb{1}(t_n\in(S_i,E_i))} & \cdots & \frac{\mathbb{1}(t_n\in(S_m,E_m))}{n\sum_{i=0}^m \mathbb{1}(t_n\in(S_i,E_i))} \end{pmatrix},$$

taking $\hat{p}^s_{ii}$ as the entry (i,s) of P^I and $\hat{p}_{ij}$ as the entry (i,j) of P^O. The idea of the above matrices is to try to reflect the unknowns about the branching structure, starting with uniform distributions between being an aftershock or a background event, and uniformity among μ^s that could generate the event if it is a background event.

All the codes used to fit the above model and generate the images in this work are available in https://github.com/isaiasmanuel/ETAS. The code used to perform Algorithm 1 is available in *EM2.py*. Also, the figures in this study were produced using the code *GNSSETAS_Figures.py*.

2.3 Hypothesis Testing

To verify if there is an improvement of our model, i.e., by adding μ^i for $i = 1,2,...,m$ with respect to the model which only has μ_0, a hypothesis test must be performed to analyze the significance of the results. Due to the complexity of the model presented in Algorithm 1, it is not straightforward to derive the theoretical joint distribution of the estimators and determine the rejection regions. For this reason, we propose a Likelihood Ratio Test (LRT) based on parametric bootstrapping[64].

Since bootstrap is a resampling approach that requires simulating synthetic earthquake catalogs, we will simulate the synthetic catalog in a similar way to Ross & Kolev[32], based on the cluster process representation of the Hawkes process[65], the advantage of this approach is avoiding the methodology of thining as in Fox et al.[38] that requires more computational time. The simulation of synthetic catalogs is as follows

1. We simulate our synthetic mainshocks for each square D_{uv} defined in the equation (12), and for each interval $i = 1,2,...,m$, from a Poisson distribution with mean $\mu^i_{uv}(E_i - S_i)\Delta x\Delta y$, the occurrence time of this event is simulated from a uniform distribution in (S_i, E_i) and the magnitude is simulated from the Gutenberg-Richter law with the $b-$value estimated via maximum likelihood from the real catalog[57].

2. For each simulated earthquake j that it is direct offspring have not been calculated previously, we simulate the number of direct offspring from a Poisson distribution with mean $Ke^{\alpha(M_j - M_0)}$, and the offspring locations and times are simulated directly from random variables with densities given by equations (4) using $t = t_j$ and (5) using (x_j, y_j). If the new earthquakes have times larger than E_0 or if they are outside of our study area, they are discarded.

3. Repeat the previous step until no new offspring is generated.

The code to simulate synthetic earthquake catalogs is available in *Hypothesis_testing.py*.

The hypothesis to test in the LRT are

- H_0: θ has μ^i for $i = 1,2,..$ equal to 0
- H_1: θ has at least one μ^i for $i = 1,2,..$ different to 0

From the real catalog $\{x_i, y_i, t_i, m_i\}_{i=1}^n$ we can obtain the maximum likelihood estimators (MLE) of θ using the complete model and the model that only has μ^0, which will be denoted by $\hat{\theta}$ and $\tilde{\theta}$ respectively, then our test statistic is defined by

$$T_{LR} = \frac{L_1(\hat{\theta})}{L_0(\tilde{\theta})},$$

where L_1 and L_0 denote the likelihoods of the complete and reduced model. In order to define the rejection region for our hypothesis testing, we will generate B synthetic catalogs $\{x_i, y_i, t_i, m_i\}_{i=1}^{n_b}$, with $b = 1,...,B$, using $\tilde{\theta}$ from the reduced model. Once the synthetic catalogs are obtained, we estimate the MLE for both models to each synthetic catalog to obtain a pair $\hat{\theta}_b$ and $\tilde{\theta}_b$, which allows to define

$$T_{LR,b} = \frac{L_{1,b}(\hat{\theta}_b)}{L_{0,b}(\tilde{\theta}_b)},$$

where $L_{0,b}, L_{1,b}$ denote the Likelihood functions of the reduced and complete model using the synthetic catalogue b. It is important to note that the idea of the bootstrap methodology is to obtain an approximation of the distribution of our estimators through empirical distributions obtained by resampling[64].

Finally we will reject H_0 in favor of H_1 at a confidence level γ if the following expression is fulfilled

$$\frac{\#\{T_{LR,b} : T_{LR,b} > T_{LR}\}}{B} \leq \gamma,$$

where # denotes the cardinality of the set.

While the LRT allows us to determine whether there is evidence that the full model is preferable to the reduced model, it is not informative in terms of seeing which μ^i_{uv} are contributing to that conclusion. Taking this into account and the fact that we already calculate $\hat{\theta}_b$ for all synthetic catalogs, we can examine marginally which μ^i_{uv} from $\hat{\theta}$, denoted by $\hat{\mu}^i_{uv}$, are significantly greater than the corresponding μ^i_{uv} from $\hat{\theta}_b$, denoted by $\mu^{i,b}_{uv}$. i.e., we can define

$$T_{uv} = \frac{\#\{\mu^{i,b}_{uv} : \mu^{i,b}_{uv} \geq \hat{\mu}^i_{uv}\}}{B},$$

and conclude that $\hat{\mu}^i_{uv}$ is significantly higher than μ^i_{uv} under the reduced model at level γ if $T_{uv} \leq \gamma$.

Note that the value μ_{uv}^{i} for $i = 1, 2, \ldots$ can be 0 if no earthquake was observed in the square D_{uv} during the interval (S_i, E_i) or if the observed earthquakes were classified as offspring of an earthquake that ocurred prior to (S_i, E_i).

In the section 4 this hypothesis testing methodology is applied for records in Guerrero, Mexico and the Boso Peninsula, Japan.

Additionally, since we are obtaining synthetic catalogs with a sample size n_b, we select the size of our mesh in equation 12 as the $n_x = n_y$ that minimize

$$\frac{1}{B}\sum_{b=1}^{B}(n - n_b)^2.$$

In other words, the mesh size is selected through minimize the mean squared error of the number of earthquakes. It is worthy to mention that Asim et al.[66] have proposed the idea of use a multiresolution mesh to improve the computational time, nevertheless due the size of our region a non multiresolution mesh is still computationally affordable.

3 Data

3.1 Mexico

In figure 1, all the epicenters in the study are shown in blue, on a mesh with $n_x = n_y = 14$, following the notation in equation (12). The earthquake catalog was obtained from the Advanced National Seismic System (ANSS)[67], with dates from 2000/01/01 to 2016/12/31, and magnitudes greater than or equal to 4.3. Only the earthquakes related to plate subduction (i.e., on the landward side of the trench) were considered in this study. The magnitude completeness of the ANSS catalog in the study region was assessed using the Maximum Curvature Method[68], and the magnitude of completeness (Mc) was found to be M4.0. However, as this method is known to systematically underestimate Mc by a fraction of a unit[69], M4.3 was used as the threshold M_0 in equations (3) and (5) in the present work.

On September 8, 2017, an earthquake of M8.2 occurred in Mexico and some studies, such as[70], have suggested that it may have broken the entire subducted Cocos lithosphere and significantly altered the seismicity in the subduction system. For this reason, we limit our study period to the end of 2016.

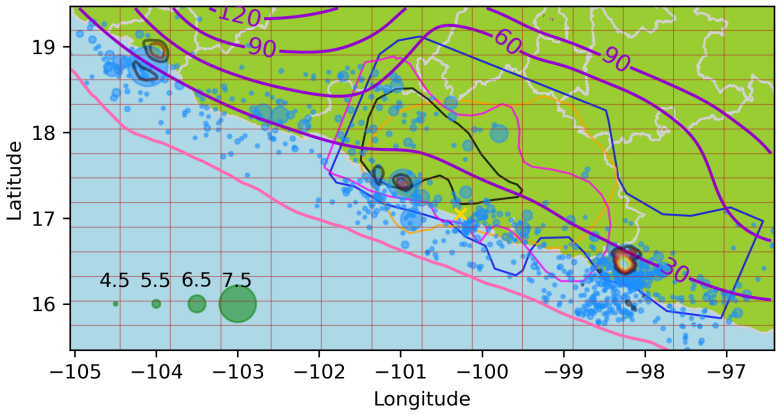


Figure 1. Seismicity data from Guerrero, Mexico. Earthquakes epicenters[67] are shown in blue. Green circles are magnitude scale. The red squares define the grid used in Algoritm 1. The pink line denotes the Middle America Trench[71]. The gray lines show the political division of Mexico[72], the blue, orange, fuchsia contours are the 4 cm slip contours curves digitalized from[73] of the 2001-2002, 2006 and 2009-2010 SSE respectively, the black curve is the 15 cm slip contour of the 2014 SSE digitalized from[4]. The yellow cross correspond to the location of Cayaco GNSS station. The contour lines in purple are the slab model from[74]. We also present the slip distribution contours at 1-meter intervals from 1 to 6 meters for the earthquakes with $M > 7$, using data from[67]. These earthquakes ocurred on 2003/01/22, 2012/03/20, and 2014/04/18, with epicenters -104.1040E° 18.770N°, -98.2310E° 16.493N°, and -100.9723E° 17.397N° and magnitudes 7.6, 7.4 and 7.2, respectively.

In figure 2, the North-South component of the GNSS time series of the Cayaco (CAYA) continuous GPS monitoring station is presented. CAYA is located in Guerrero state, Mexico, at -100.2672E°,17.0485N°. The duration of SSEs can be defined as the periods of southward displacements at this station[9], and the SSEs appear to have a periodicity of approximately four years. This GNSS data is available in https://github.com/isaiasmanuel/ETAS as *caya_ns*.

To estimate the start and end times (S_i and E_i) of the SSEs, b-splines[75] are fitted to the raw data. The duration of the SSEs is defined by the periods when the spline has a negative slope, as shown in figure 2. The data for this work was provided by the Servicio Sismológico Nacional (SSN) of Mexico and by the Department of Seismology, IGEF UNAM.

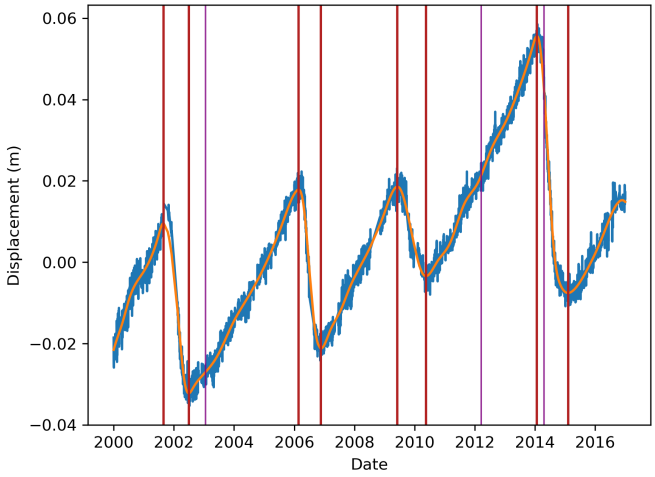


Figure 2. Global navigation satellite system data at Cayaco station. North-to-south component at Cayaco station (blue). Fitted b-spline polynomial (orange). The periods when the data slope is negative is indicated by vertical red lines, and the purple vertical lines are when an earthquake with $M > 7$ ocurred.

The SSE slip contour curves with 4 cm of slip spacing (from 2 cm to 18 cm) of the first 3 SSEs (the 2001-2002, 2006, and 2009-2010 SSEs) have been digitalized from Radiguet et al.[73]. The 15 cm slip contour of the 2014 SSE was digitalized from[4] and is presented in Figure 1. The expected result of the algorithm 1 is that if an earthquake triggering effect of SSEs exists, μ^i with $i = 1,2,3,4$ should have an increase near the SSE slip contours.

3.2 Japan

To test the flexibility and capability of our model, we additionally explored the Boso Peninsula, located in the Sagami Trough subduction zone in Japan. The seismicity-triggering effect of SSEs and swarm activity produced during SSE periods in the Boso Peninsula have been previously studied by Okutani & Ide[22] and Fukuda[30]. For this analysis, we used data from the Japanese Meteorological Data[76] from 2001/01/01 to 2009/01/01 and $M_0 = 3$. Accoding to the geodetic analysis by Fukuda[30], during this period, two SSE ocurred from 2002/10/01 to 2002/10/19 and from 2007/08/12 to 2008/08/25. These periods were used to define μ^1 and μ^2, respectively.

Figure 3 shows the seismicity during our study period, and the mesh used in Algorithm 1, in this example $n_x = n_y = 6$. The daily slip rate distributions of $4\frac{\text{m}}{\text{year}}$ for the 2002 and 2007 SSEs are also presented. The digitized data correspond to the slip rates on 2022/10/08 and 2007/08/16 from[30], which represent the days with the highest slip rates.

It is important to note that, while the SSEs in Mexico occurred over several months, those in the Boso Peninsula lasted only several weeks. This substantial difference in time scale is a key reason for analyzing the performance of our model in both regions.

4 Results

4.1 Mexico

Using μ^1, we observe a substantial increase in the background seismicity of $0.18\frac{\text{Events}}{\text{deg}^2\text{day}}$, compared to the stationary background rate of 0.01 given by μ^0, during the 2001-2002 SSE in the square with a vertex at -100.11E°, 17.18N° (Figure 4). It is important to note that this area is located close to the regions with the largest slip gradient in the slip contours of figure 4.The significant influence of μ^1 on the background seismicity is further supported by Figure 5, which highlights the squares with a significant T^i_{uv} at the 90% confidence level using $B = 100$. Additionally, the LRT rejects the null hypothesis (p-value = 0.09) in favor of our model.

Since SSE slips release stored elastic energy, the regions with the largest SSE slips have experienced a reduction in cumulative stress. In contrast, stress increases just outside the margins of the SSE patches, where an increase in background seismicity can be expected due to the triggering effects. This is consistent with our estimation in Figure 4.

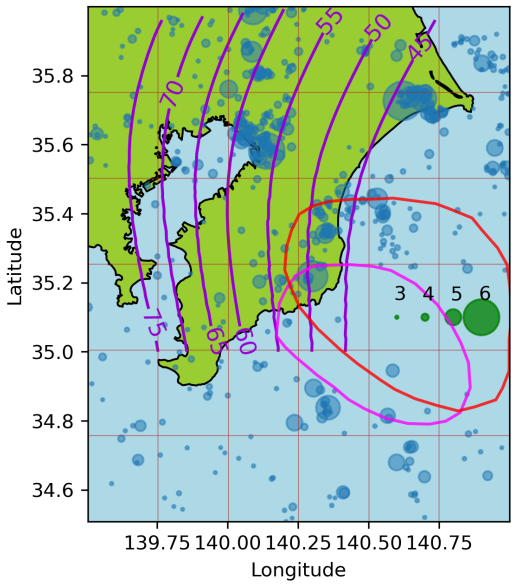


Figure 3. Seismicity data from the Boso Peninsula, Japan. The colors are the same as in Figure 1, except for the SSEs. The fuchsia and red contours denote the daily slip rate distributions of $4\frac{\text{m}}{\text{year}}$ for the 2002 and 2007 SSEs, respectively, digitized from[30].

Regarding μ^2 and μ^3, we observe values greater than zero next to the epicenter of the 2003/01/22 earthquake and in the margins of the 2006 and 2009-2010 SSEs. Nevertheless, these values are too small and dispersed to conclude that a triggering effect could be visible. Furthermore, as shown in Figure 5, the observed values in nearly all the squares are not statistically significant.

Radiguet et al.[4] have discussed that the 2014/04/18 M7.3 Papanoa earthquake was triggered by the 2014 SSE. Consistently, our model shows a significant increase in the background seismicity rate during the SSE period in the square that contains the epicenter, as well as in the square with vertex at -100.73E°, 16.89N° where, according to[4], the SSE caused a slip of approximately 0.05 m.

We also found a significant value of μ^4 in squares with vertices at -98.88E°, 16.03N° and -98.26E°, 16.89N° which are next to the epicenter of the 2012/03/20 earthquake. Although we have focused this work on discussing changes in seismicity during periods of SSEs, it is important in future work to consider the background seismicity variations around large earthquakes to account for possible tectonic stress rearrangements.

While μ^1 reached the highest values among all μ^i, μ^2 and μ^3 exhibited much lower levels of activity. This may be related to differences in the source characteristics of each SSE (i.e., moment magnitude, duration, and slip distribution).

In terms of moment magnitude, the 2001–2002 SSE was the largest (Mw 7.65), while the 2006, 2009–2010, and 2014 SSEs had smaller magnitudes of Mw 7.49, Mw 7.54, and Mw 7.60, respectively, according to Radiguet et al.[4,73]. The duration of the 2001–2002 SSE was 436 days, while the other three SSEs lasted 314, 510, and 378 days, respectively.

In particular, the 2009-2010 SSE exhibited relatively small moment magnitude and longest duration, indicating a lower moment rate. This may explain our observation that this SSE did not produce a detectable triggering effect for earthquakes of M4.3 or larger.

To show the validity of the estimation presented in this work, the genealogy of earthquakes defined by the estimators obtained using Algorithm 1, is presented in Figure 6. In this figure, all the earthquakes in our dataset are shown with arrows pointing to their aftershocks, in the figure only arrows where $P_{ij}^O \geq 0.5$ are presented.

An anomalous arrow was observed originating from the earthquake with coordinates -102.540E°, 18.026N°. This may be due to the fact that the time between both earthquakes was only 6.86 hours, and the fact the second earthquake is in a square where only 3 events were recorded throughout the entire study period. This arrow may be an unreliable result.

To clarify the results in Figure 6, a zoomed-in view of indices 183 to 192 is presented in the Figure 7a. This window was selected because it contains the largest earthquake (event 183) in our data set, with a magnitude of 7.6. As it can be seen in the Figure 7a, earthquake 190, with an epicenter at -104.37E°,18.54N°, could be an offspring of the earthquakes 183,184, 188 or 189. Therefore, the 4 arrows pointing to it are lighter in color than the arrow between earthquakes 183 and 184, which have an associated probability close to 1.

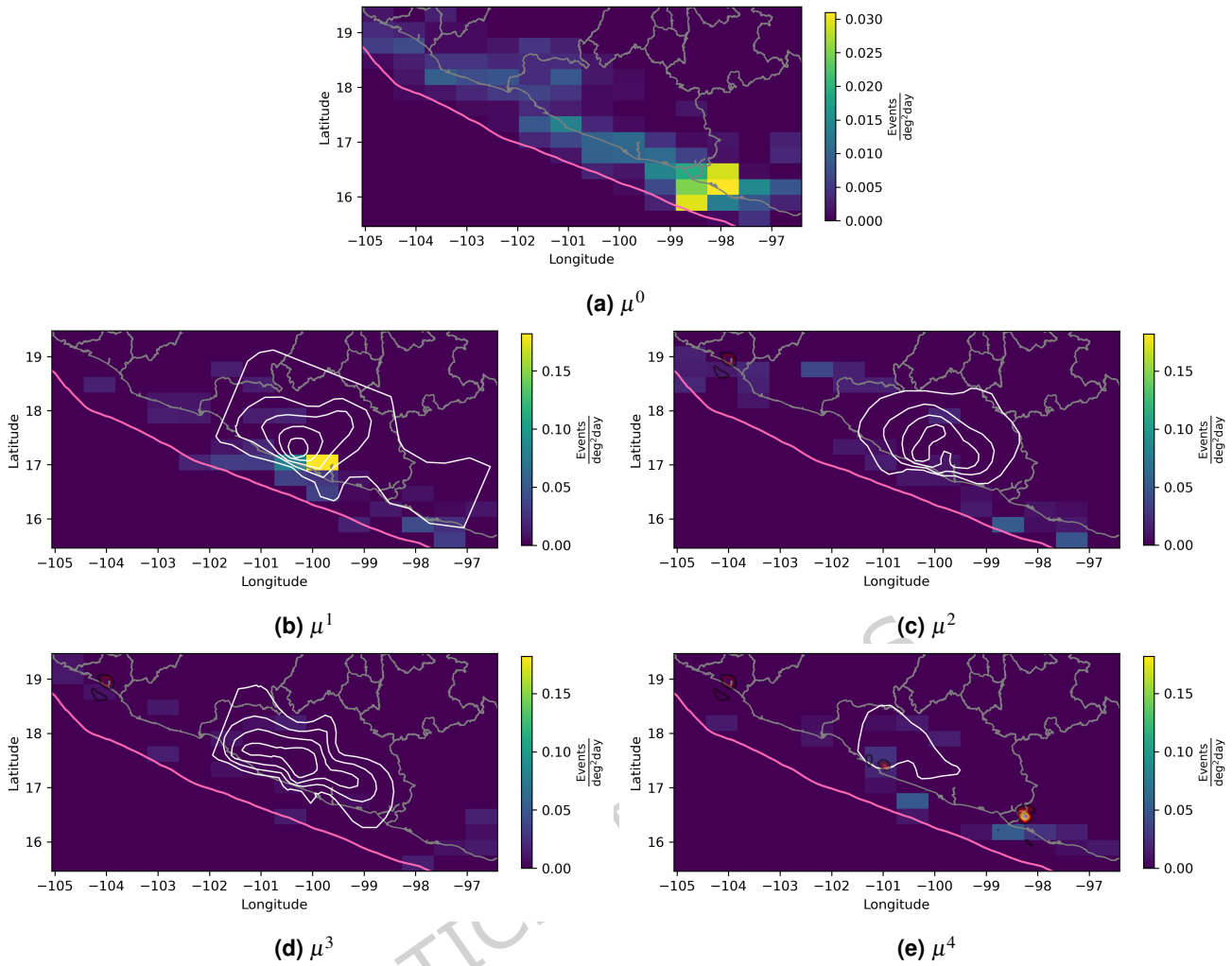


Figure 4. Estimated background seismicity for theMexican dataset. μ^i with i=0,1,2,3,4 are estimated using Algorithm 1. The white contours in μ^1, μ^2 and μ^3 represent the slip distributions with 4-cm slip intervals for the 2001-2002, 2006, and 2009-2010 SSEs, respectively, digitized from[73]. In μ^4, the slip contour of the 2014 SSE is shown.

The figure in 7b allows not only to see which earthquakes are background events and which are offsprings, but also to recover an estimated entire genealogy—an idea that has been previously explored in the works of Rasmussen[48] and Fox et al.[38].

The b-value for our dataset is 1.254 and the estimated vector of aftershock parameters (A, α, c, p, d) is (0.118 events, 1.112, 0.021 days, 1.363, 0.0048 deg^2). The expected total value of earthquakes using the data in this section and Algorithm 1 is

$$\int_0^T \iint_S \lambda(t,x,y|\mathscr{H}_t) = 774.15,$$

using our sample of 794 earthquakes. Figure 8a shows a histogram of the values (i.e., the probability of each earthquake being a background event):

$$(1-\rho_j) = \sum_{s=0}^{4} \hat{p}_{ii}^s.$$

In this figure, we observe a bimodal density concentration near 0 and 1, which is similar to the one presented by Zhuang et al.[15]. The probability of being an aftershock is considerably lower. This could be due to the use of $M_0 = 4.3$, because, as shown in figure 8b, for the Boso Peninsula in Japan ($M_0 = 3$), there are more aftershocks in proportion than in the Mexican case.

4.2 Japan

According to[30], the magnitudes of the observed SSEs are 6.67 and 6.65 for the 2002 and 2007 events, respectively. Although they had shorter durations than the SSEs in Guerrero, Mexico, the SSEs in the Boso Peninsula, Japan, have exhibited notable

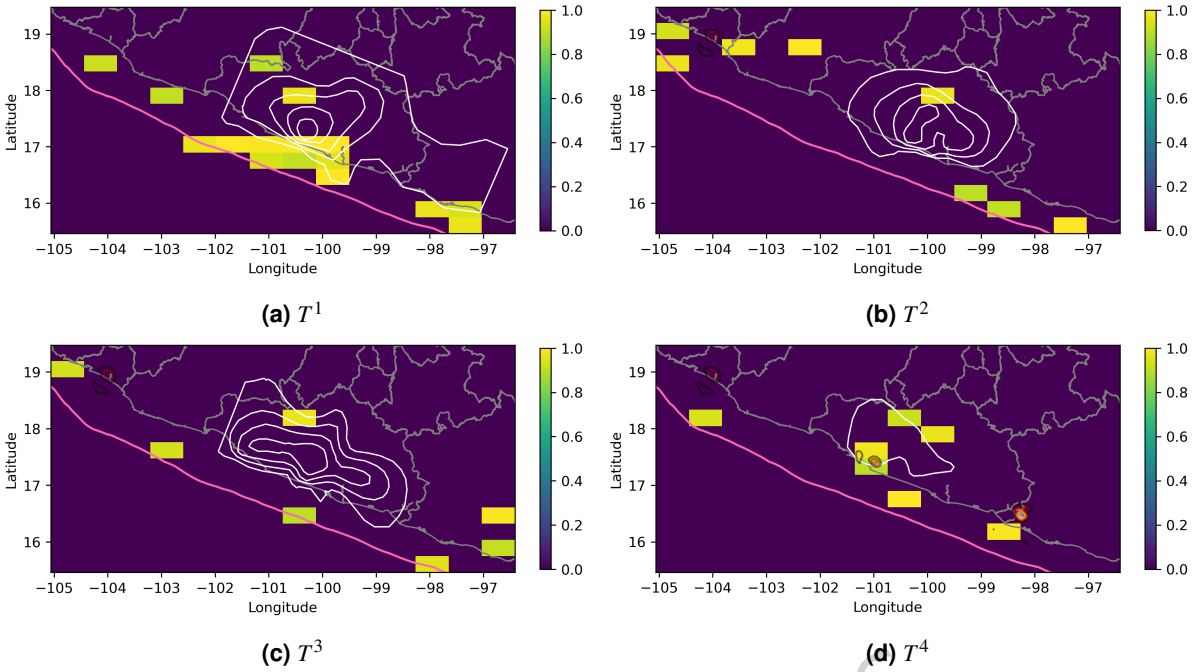


Figure 5. Significant T^i values for the Mexican dataset. The T^i values with $i = 1,2,3,4$ that are significantly different from zero are shown.

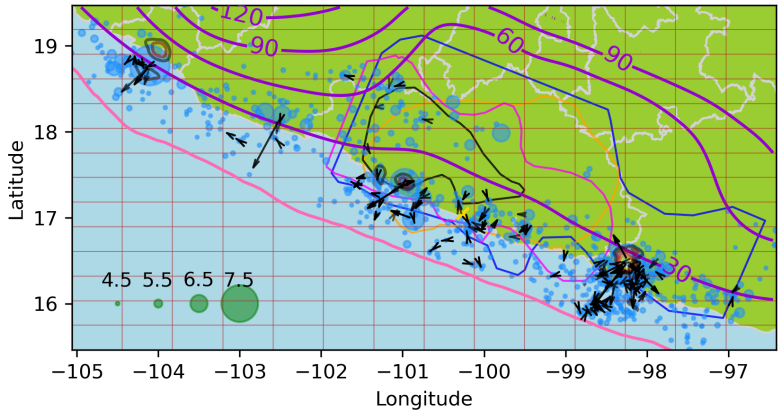


Figure 6. Offspring estimation for the Mexican dataset. Map of earthquakes with arrows pointing from events to their offspring according to P^O.

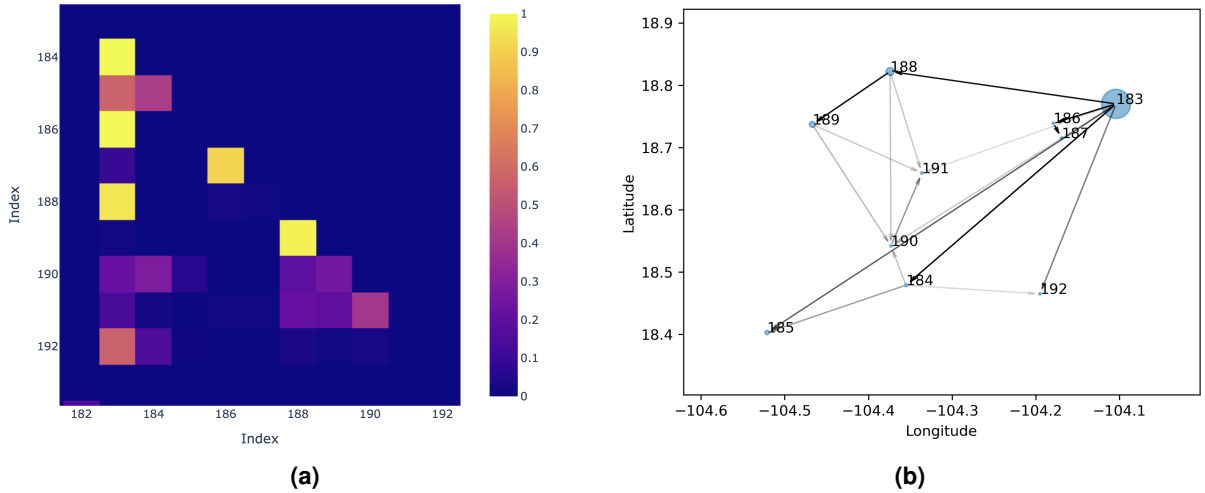
increases of the background seismicity rate, as it can be observed in figure 9. In both cases, the square with the highest value has a vertex 140.5E°, 35.75N$^\circ$, where μ^1 and μ^2 reach values of 2.65 and 3.02, respectively, while μ^0 for the same square was 0.03.

For both SSEs, T^i for the squares associated with the maximum value of μ^1 and μ^2 was statistically significant, as shown in Figure 10, and the the LRT rejects the null hypothesis (p-value = 0.01) in favor of our model. Also, as expected, the increase in the background seismicity rate was observed in the margins of the SSE patches. This is consistent with the Mexican results

The b-value for this data is 0.709, and the estimated vector of aftershock parameters (A, α, c, p, d) is (0.365 events, 1.03, 0.002 days, 1.029, 4.101e-05 deg^2). And we obtain

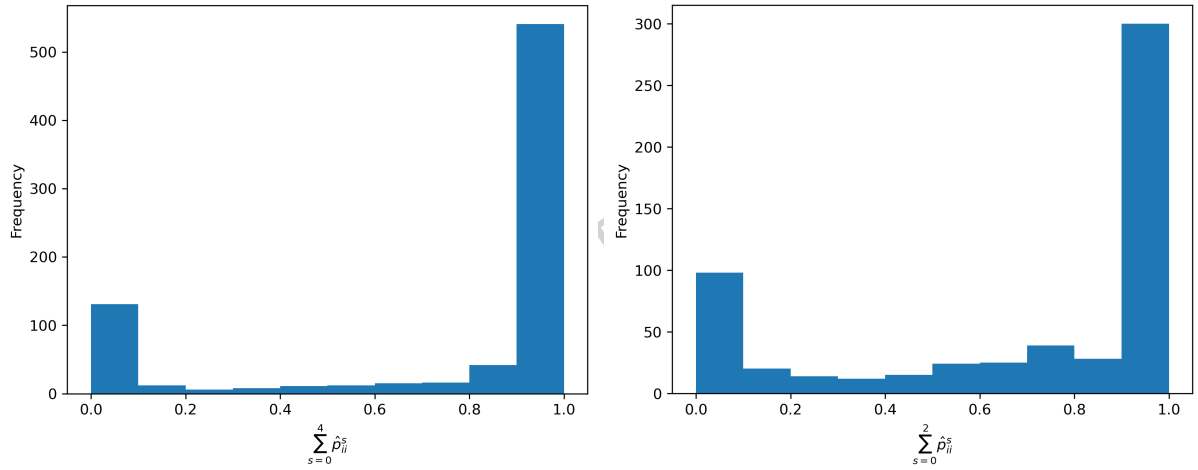
$$\int_0^T \iint_S \lambda(t,x,y|\mathscr{H}_t) = 554.10,$$

for our sample of 575 earthquakes. As in section 4.1, Figure 12 presents the estimated genealogy.



(a) (b)

Figure 7. Genealogy of events 183-192. In Figure 7a, the matrix P^O for indices 183 to 192 is shown. 7b presents the same events as a genealogy.



(a) (b)

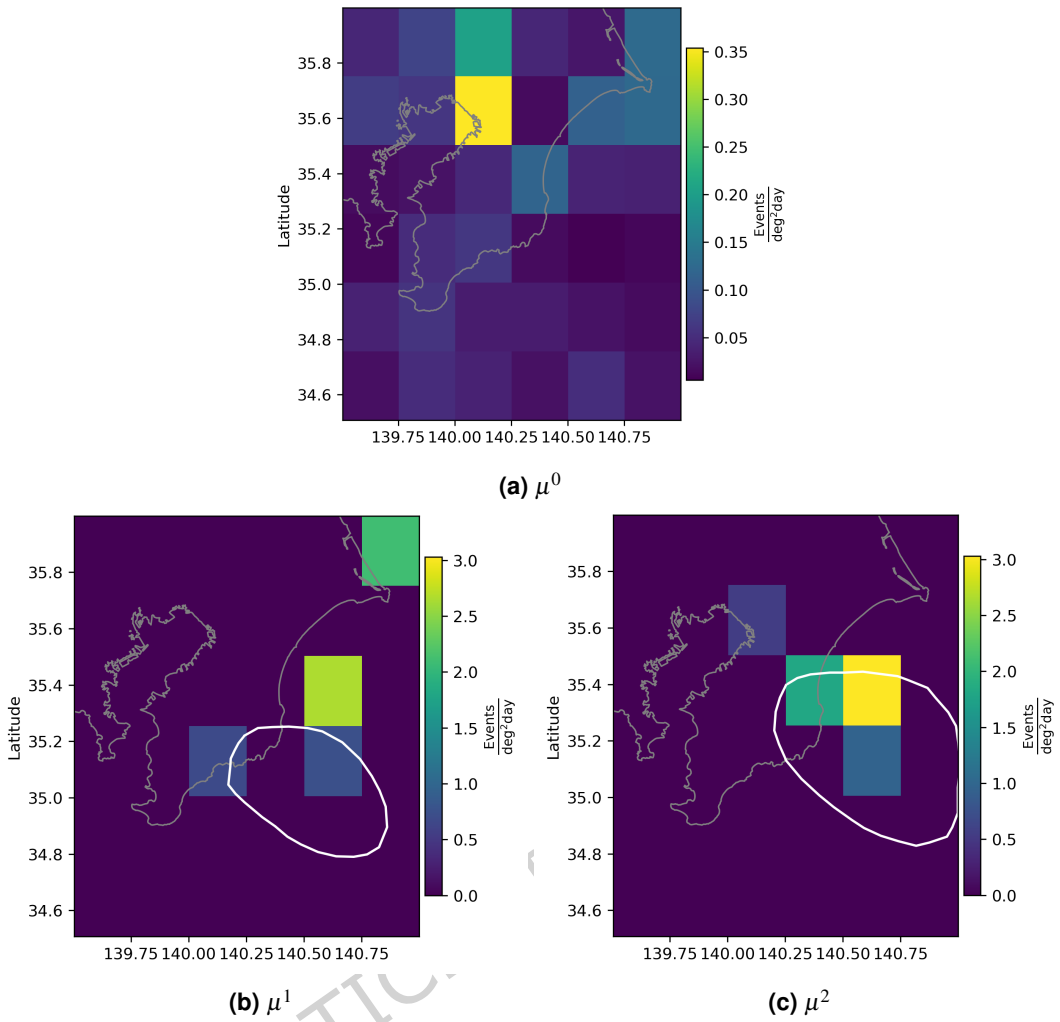
Figure 8. Background probability histogram. Histograms of probabilities of being a background event for all earthquakes in our dataset are presented. **(a)** Mexican dataset **(b)** Japanese dataset.

5 Discussion and Conclusions

We have successfully developed an inhomogeneous spatio-temporal ETAS model. Our approach extends the ideas of Fox et al.[38] by introducing a piecewise constant spatial background seismicity over time, which contrasts with their non-time-dependent approach. Furthermore, this approach can be regarded as the spatio-temporal extension of the model proposed by Okutani & Ide[22], which considers only the temporal components.

We also extended the work of Li et al.[16] by employing a marked point process. Furthermore, we did not assume that spatial and time contributions of the background seismicity interact solely through multiplication, and we accounted for differences in the spatial distribution of increased background seismicity for each SSE.

Unlike Marsan et al.[27] and Reverso et al.[28], which rely on subjectively chosen local spatio-temporal windows, we used the estimation of the SSE durations from GNSS data and utilized them in our seismicity modeling. Moreover, rather than using an approximate approach that sequentially optimizes changes in the background seismicity rate across numerous small spatio-temporal windows, our method simultaneously performs a maximum likelihood estimation of spatio-temporal variations in the background seismicity rate over the entire analysis region and period, which represents a key technical advance of our model.



(a) μ^0

(b) μ^1

(c) μ^2

Figure 9. Estimated background seismicity for the Japanese dataset. μ^i with i=0,1,2 estimated using Algorithm 1 are shown. The white contours represent the daily slip rates presented in figure 3 for the 2002 and 2007 SSE.

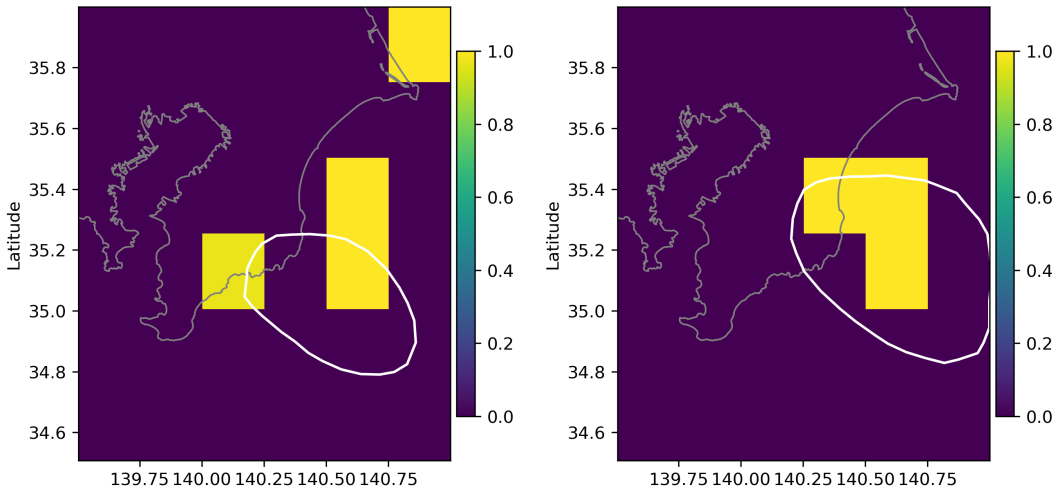
Furthermore, for the first time, we quantitatively evaluated the relationship between seismic activity and SSEs in the Mexican subduction zone using a statistical seismicity model that accounts for not only regular earthquakes but also SSEs. Our results from Mexico demonstrate statistically significant increases in the background seismicity accompanying Guerrero SSEs, highlighting the seismological novelty of this study.

Our model also successfully detected statistically significant increases in background seismicity associated with SSEs in the offshore Boso Peninsula region of Japan. Despite the substantial differences in the time scales of SSEs between the Boso Peninsula and Guerrero, Mexico, consistent results were obtained—namely, an increase in background seismicity near the margins of the SSE patches. This demonstrates the adaptability and broad applicability of our model to different datasets.

The assumptions made for our model were flexible enough to be applied in different regions, and we expect to continue improving the model and exploring other plate boundaries where SSEs occur, such as in New Zealand and Peru[12,13]. Applying our model to these regions will enable a deeper understanding of how SSEs trigger fast earthquakes.

In future work, we plan to extend our algorithm to allow dependencies between mainshock and aftershock magnitudes. Currently, we model magnitudes independently. However, it is important to explore possible dependencies between magnitudes, as Nandan et al.[59] found evidence of such dependencies in California, USA.

We also intend to improve our algorithm by replacing the piecewise estimators of μ^i with smooth functions in the space domain. In the context of the ETAS model, interesting ideas have recently been explored for this purpose, such as the use of Dirichlet processes[32] (DP), or Gaussian processes[35] (GP). An additional advantage of using DP or GP is that they allow inference to be performed within a Bayesian framework.



(a) T^1 **(b)** T^2

Figure 10. Significant T^i values for the Japanese dataset. The T^i with $i = 1, 2$ that are significantly different from zero are presented.

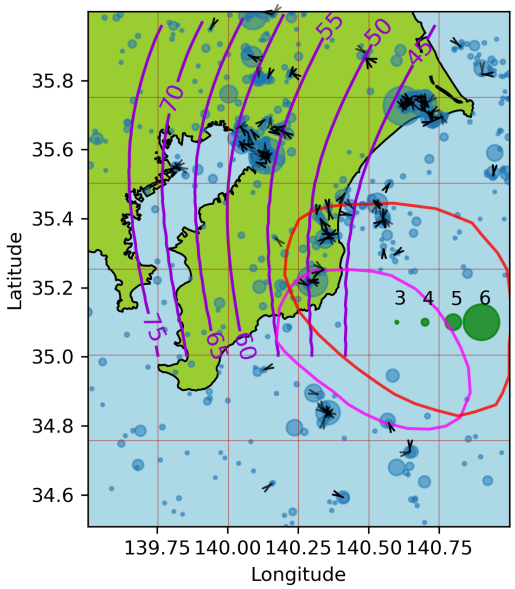


Figure 11

Figure 12. Offspring estimation for the Japanese dataset. Map of earthquakes with arrows pointing from events to their offspring according to P^0.

6 Data availability

Data is provided within the manuscript.

7 Acknowledgments

This work was supported by JSPS KAKENHI (Grant 21H05200 and 24K17144) and JST SATREPS project (Grant JP-MJSA2310).

8 Funding

This work was sponsored by grants received by Isaías Bañales from the Japan Society for the Promotion of Science (Grant 21H05200 and 24K17144) and the Science and Technology Research Partnership for Sustainable Development (Grant JP-MJSA2310).

9 Author Contributions Statement

Isaías Bañales wrote the main manuscript text under the supervision of Tomoaki Nishikawa and Yoshihiro Ito. The data was provided by Vladimir Kostoglodov, Ekaterina Kazachkina and José Santiago. All authors reviewed the manuscript.

References

1. Ide, S. & Beroza, G. C. Slow earthquake scaling reconsidered as a boundary between distinct modes of rupture propagation. *Proc. Natl. Acad. Sci.* **120**, e2222102120 (2023).
2. Ozawa, S. *et al.* Characteristic silent earthquakes in the eastern part of the boso peninsula, central japan. *Geophys. Res. Lett.* **30** (2003).
3. Ito, Y. *et al.* Episodic slow slip events in the japan subduction zone before the 2011 tohoku-oki earthquake. *Tectonophysics* **600**, 14–26 (2013).
4. Radiguet, M. *et al.* Triggering of the 2014 m w 7.3 papanoa earthquake by a slow slip event in guerrero, mexico. *Nat. Geosci.* **9**, 829–833 (2016).
5. Dragert, H., Wang, K. & James, T. S. A silent slip event on the deeper cascadia subduction interface. *Science* **292**, 1525–1528 (2001).
6. Kostoglodov, V. *et al.* A large silent earthquake in the guerrero seismic gap, mexico. *Geophys. Res. Lett.* **30** (2003).
7. Schwartz, S. Y. & Rokosky, J. M. Slow slip events and seismic tremor at circum-pacific subduction zones. *Rev. Geophys.* **45** (2007).
8. Nishimura, T., Matsuzawa, T. & Obara, K. Detection of short-term slow slip events along the nankai trough, southwest japan, using gnss data. *J. Geophys. Res. Solid Earth* **118**, 3112–3125 (2013).
9. Cruz-Atienza, V. M. *et al.* Short-term interaction between silent and devastating earthquakes in mexico. *Nat. communications* **12**, 2171 (2021).
10. Kawasaki, I. *et al.* The 1992 sanriku-oki, japan, ultra-slow earthquake. *J. Phys. Earth* **43**, 105–116 (1995).
11. Nishikawa, T., Ide, S. & Nishimura, T. A review on slow earthquakes in the japan trench. *Prog. Earth Planet. Sci.* **10**, 1–51 (2023).
12. Delahaye, E., Townend, J., Reyners, M. & Rogers, G. Microseismicity but no tremor accompanying slow slip in the hikurangi subduction zone, new zealand. *Earth Planet. Sci. Lett.* **277**, 21–28 (2009).
13. Villegas-Lanza, J. C. *et al.* A mixed seismic–aseismic stress release episode in the andean subduction zone. *Nat. Geosci.* **9**, 150–154 (2016).
14. Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. association* **83**, 9–27 (1988).
15. Zhuang, J., Ogata, Y. & Vere-Jones, D. Stochastic declustering of space-time earthquake occurrences. *J. Am. Stat. Assoc.* **97**, 369–380 (2002).
16. Li, C., Song, Z. & Wang, W. Space–time inhomogeneous background intensity estimators for semi-parametric space–time self-exciting point process models. *Annals Inst. Stat. Math.* **72**, 945–967 (2020).
17. Hainzl, S. & Ogata, Y. Detecting fluid signals in seismicity data through statistical earthquake modeling. *J. Geophys. Res. Solid Earth* **110** (2005).
18. Hainzl, S., Zakharova, O. & Marsan, D. Impact of aseismic transients on the estimation of aftershock productivity parameters. *Bull. Seismol. Soc. Am.* **103**, 1723–1732 (2013).
19. Marsan, D., Prono, E. & Helmstetter, A. Monitoring aseismic forcing in fault zones using earthquake time series. *Bull. Seismol. Soc. Am.* **103**, 169–179 (2013).
20. Kumazawa, T. & Ogata, Y. Quantitative description of induced seismic activity before and after the 2011 tohoku-oki earthquake by nonstationary etas models. *J. Geophys. Res. Solid Earth* **118**, 6165–6182 (2013).

21. Passarelli, L. *et al.* Aseismic transient driving the swarm-like seismic sequence in the pollino range, southern italy. *Geophys. J. Int.* **201**, 1553–1567 (2015).

22. Okutani, T. & Ide, S. Statistic analysis of swarm activities around the boso peninsula, japan: Slow slip events beneath tokyo bay? *Earth, planets space* **63**, 419–426 (2011).

23. Matthews, M. V. & Reasenberg, P. A. Statistical methods for investigating quiescence and other temporal seismicity patterns. *Pure Appl. Geophys.* **126**, 357–372 (1988).

24. Nishikawa, T. & Nishimura, T. Development of an epidemic-type aftershock-sequence model explicitly incorporating the seismicity-triggering effects of slow slip events. *J. Geophys. Res. Solid Earth* **128**, e2023JB026457 (2023).

25. Llenos, A. L. & McGuire, J. J. Detecting aseismic strain transients from seismicity data. *J. Geophys. Res. Solid Earth* **116** (2011).

26. Dieterich, J. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res. Solid Earth* **99**, 2601–2618 (1994).

27. Marsan, D., Reverso, T., Helmstetter, A. & Enescu, B. Slow slip and aseismic deformation episodes associated with the subducting pacific plate offshore japan, revealed by changes in seismicity. *J. Geophys. Res. Solid Earth* **118**, 4900–4909 (2013).

28. Reverso, T., Marsan, D. & Helmstetter, A. Detection and characterization of transient forcing episodes affecting earthquake activity in the aleutian arc system. *Earth Planet. Sci. Lett.* **412**, 25–34 (2015).

29. Reverso, T., Marsan, D., Helmstetter, A. & Enescu, B. Background seismicity in boso peninsula, japan: Long-term acceleration, and relationship with slow slip events. *Geophys. Res. Lett.* **43**, 5671–5679 (2016).

30. Fukuda, J. Variability of the space-time evolution of slow slip events off the boso peninsula, central japan, from 1996 to 2014. *J. Geophys. Res. Solid Earth* **123**, 732–760 (2018).

31. Hainzl, S. Etas-approach accounting for short-term incompleteness of earthquake catalogs. *Bull. Seismol. Soc. Am.* **112**, 494–507 (2022).

32. Ross, G. J. & Kolev, A. A. Semiparametric bayesian forecasting of spatiotemporal earthquake occurrences. *The Annals Appl. Stat.* **16**, 2083–2100 (2022).

33. Seif, S., Mignan, A., Zechar, J. D., Werner, M. J. & Wiemer, S. Estimating etas: The effects of truncation, missing data, and model assumptions. *J. Geophys. Res. Solid Earth* **122**, 449–469 (2017).

34. Guo, Y., Zhuang, J. & Ogata, Y. Modeling and forecasting aftershocks can be improved by incorporating rupture geometry in the etas model. *Geophys. Res. Lett.* **46**, 12881–12889 (2019).

35. Molkenthin, C. *et al.* Gp-etas: semiparametric bayesian inference for the spatio-temporal epidemic type aftershock sequence model. *Stat. computing* **32**, 29 (2022).

36. Antoniak, C. E. Mixtures of dirichlet processes with applications to bayesian nonparametric problems. *The annals statistics* 1152–1174 (1974).

37. Veen, A. & Schoenberg, F. P. Estimation of space–time branching process models in seismology using an em–type algorithm. *J. Am. Stat. Assoc.* **103**, 614–624 (2008).

38. Fox, E. W., Schoenberg, F. P. & Gordon, J. S. Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences. *The Annals Appl. Stat.* **10**, 1725 – 1756 (2016).

39. Heikkinen, J. & Arjas, E. Non-parametric bayesian estimation of a spatial poisson intensity. *Scand. J. Stat.* **25**, 435–450 (1998).

40. Geng, J., Shi, W. & Hu, G. Bayesian nonparametric nonhomogeneous poisson process with applications to usgs earthquake data. *Spatial Stat.* **41**, 100495 (2021).

41. Hainzl, S. Apparent triggering function of aftershocks resulting from rate-dependent incompleteness of earthquake catalogs. *J. Geophys. Res. Solid Earth* **121**, 6499–6509 (2016).

42. van der Elst, N. J. B-positive: A robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *J. Geophys. Res. Solid Earth* **126**, e2020JB021027 (2021).

43. Helmstetter, A., Kagan, Y. Y. & Jackson, D. D. Comparison of short-term and time-independent earthquake forecast models for southern california. *Bull. Seismol. Soc. Am.* **96**, 90–106 (2006).

44. Asayesh, B. M., Hainzl, S. & Zöller, G. Improved aftershock forecasts using mainshock information in the framework of the etas model. *J. Geophys. Res. Solid Earth* **130**, e2024JB030287 (2025).

45. Ogata, Y. Significant improvements of the space-time etas model for forecasting of accurate baseline seismicity. *Earth, planets space* **63**, 217–229 (2011).

46. Ueda, T., Kato, A., Ogata, Y. & Yamaya, L. Spatial variations in seismicity characteristics in and around the source region of the 2019 yamagata-oki earthquake, japan. *Earth, Planets Space* **73**, 40 (2021).

47. Ide, S. The proportionality between relative plate velocity and seismicity in subduction zones. *Nat. Geosci.* **6**, 780–784 (2013).

48. Rasmussen, J. G. Bayesian inference for hawkes processes. *Methodol. Comput. Appl. Probab.* **15**, 623–642 (2013).

49. Daley, D. J. & Vere-Jones, D. *An introduction to the theory of point processes: volume I: elementary theory and methods* (Springer, 2003).

50. McLachlan, G. J. & Krishnan, T. *The EM algorithm and extensions* (John Wiley & Sons, 2007).

51. Utsu, T., Ogata, Y. *et al.* The centenary of the omori formula for a decay law of aftershock activity. *J. Phys. Earth* **43**, 1–33 (1995).

52. Helmstetter, A., Sornette, D. & Grasso, J.-R. Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. *J. Geophys. Res. Solid Earth* **108** (2003).

53. Schoenberg, F. P. Facilitated estimation of etas. *Bull. seismological Soc. Am.* **103**, 601–605 (2013).

54. Kolev, A. A. & Ross, G. J. Inference for etas models with non-poissonian mainshock arrival times. *Stat. Comput.* **29**, 915–931 (2019).

55. Gutenberg, B. & Richter, C. *Seismicity of the earth and associated phenomena* (Princeton University Press, 1949).

56. Lombardi, A. M. The maximum likelihood estimator of b-value for mainshocks. *Bull. Seismol. Soc. Am.* **93**, 2082–2088 (2003).

57. Aki, K. Maximum likelihood estimate of b in the formula log n= a-bm and its confidence limits. *Bull. Earthq. Res. Inst., Tokyo Univ.* **43**, 237–239 (1965).

58. Lippiello, E. & Petrillo, G. b-more-incomplete and b-more-positive: Insights on a robust estimator of magnitude distribution. *J. Geophys. Res. Solid Earth* **129**, e2023JB027849 (2024).

59. Nandan, S., Ouillon, G. & Sornette, D. Magnitude of earthquakes controls the size distribution of their triggered events. *J. Geophys. Res. Solid Earth* **124**, 2762–2780 (2019).

60. Ito, R. & Kaneko, Y. Physical mechanism for a temporal decrease of the gutenberg-richter b-value prior to a large earthquake. *J. Geophys. Res. Solid Earth* **128**, e2023JB027413 (2023).

61. Herrmann, M., Piegari, E. & Marzocchi, W. Revealing the spatiotemporal complexity of the magnitude distribution and b-value during an earthquake sequence. *Nat. Commun.* **13**, 5087 (2022).

62. Holschneider, M., Narteau, C., Shebalin, P., Peng, Z. & Schorlemmer, D. Bayesian analysis of the modified omori law. *J. Geophys. Res. Solid Earth* **117** (2012).

63. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).

64. Shao, J. & Tu, D. *The jackknife and bootstrap* (Springer Science & Business Media, 2012).

65. Hawkes, A. G. & Oakes, D. A cluster process representation of a self-exciting process. *J. applied probability* **11**, 493–503 (1974).

66. Asim, K. M. *et al.* Multi-resolution grids in earthquake forecasting: The quadtree approach. *Bull. Seismol. Soc. Am.* **113**, 333–347 (2023).

67. U.S. Geological Survey. Advanced national seismic system (anss) comprehensive catalog. http://earthquake.usgs.gov/earthquakes/search/ (2017).

68. Wiemer, S. & Wyss, M. Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bull. Seismol. Soc. Am.* **90**, 859–869 (2000).

69. Woessner, J. & Wiemer, S. Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty. *Bull. Seismol. Soc. Am.* **95**, 684–698 (2005).

70. Melgar, D. *et al.* Deep embrittlement and complete rupture of the lithosphere during the m w 8.2 tehuantepec earthquake. *Nat. Geosci.* **11**, 955–960 (2018).

71. The General Bathymetric Chart of the Oceans. Undersea feature names gazetteer. https://www.ngdc.noaa.gov/gazetteer/ (2015).

72. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Portal de información geoespacial. http://www.conabio.gob.mx/informacion/gis (2024).

73. Radiguet, M. *et al.* Slow slip events and strain accumulation in the guerrero gap, mexico. *J. Geophys. Res. Solid Earth* **117** (2012).

74. Hayes, G. P. *et al.* Slab2, a comprehensive subduction zone geometry model. *Science* **362**, 58–61 (2018).

75. De Boor, C. On calculating with b-splines. *J. Approx. theory* **6**, 50–62 (1972).

76. Japan Meteorological Agency. Epicenter data catalog. https://www.data.jma.go.jp/eqev/data/bulletin/hypo.html (2025).

77. Ogata, Y. & Katsura, K. Immediate and updated forecasting of aftershock hazard. *Geophys. research letters* **33** (2006).

App. 1 Additional Figures

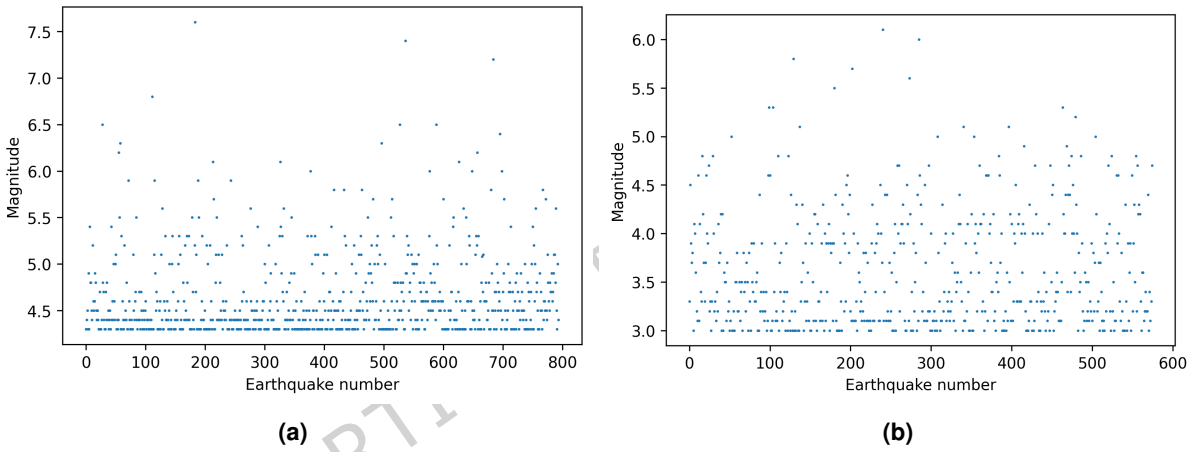


Figure 13. Ranked time vs magnitude. **(a)** Mexican dataset **(b)** Japanese dataset.

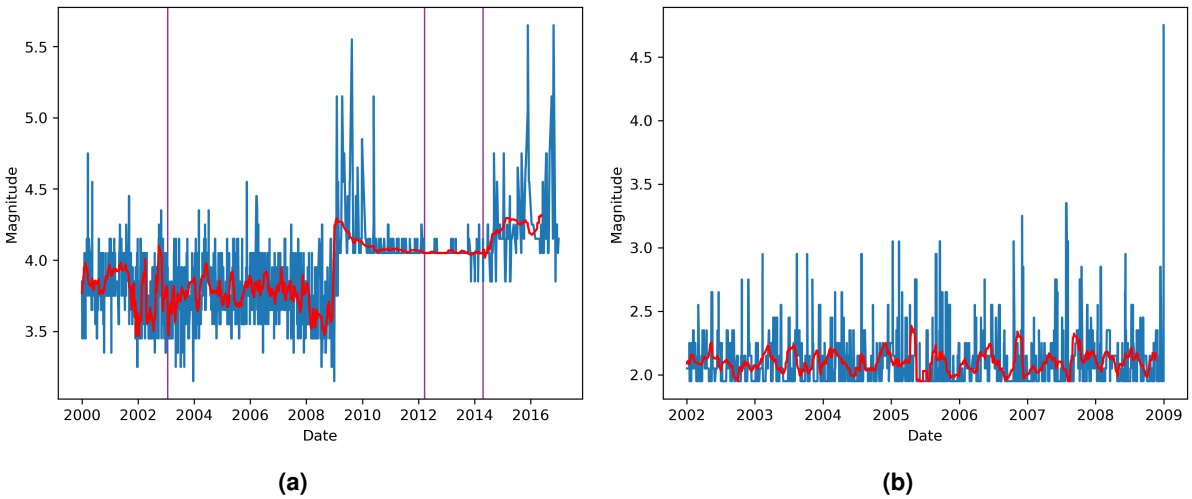


Figure 14. Maximum curvature magnitudes. **(a)** Mexican dataset **(b)** Japanese dataset. In purple the date of earthquake with magnitude greather than 7, and in red a moving average taking 50 points.

As mentioned in Section 2.1, if a catalog exhibits STAI after strong earthquakes, a reduction in the occurrence of small earthquakes would be expected, which is not evident in Figure 13. We also estimate the maximum curvature magnitude using 10-day bins for our catalog, as shown in Figure 14. For the Mexican case, we set the cutoff magnitude at 4.3, primarily to reduce the bias introduced by the change in the mean in 2009, while for the Japanese case, we set it at 3 to reduce the computational load. We also do not see any substantial changes associated with large earthquakes in this figure.

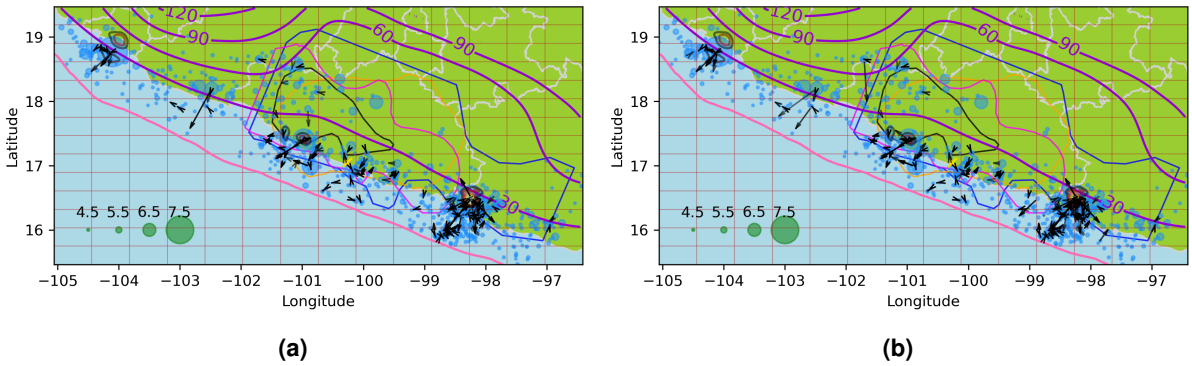


Figure 15. Mexican genealogy using alternative kernels. **(a)** Bivariate Gaussian **(b)** Bivariate t kernel.

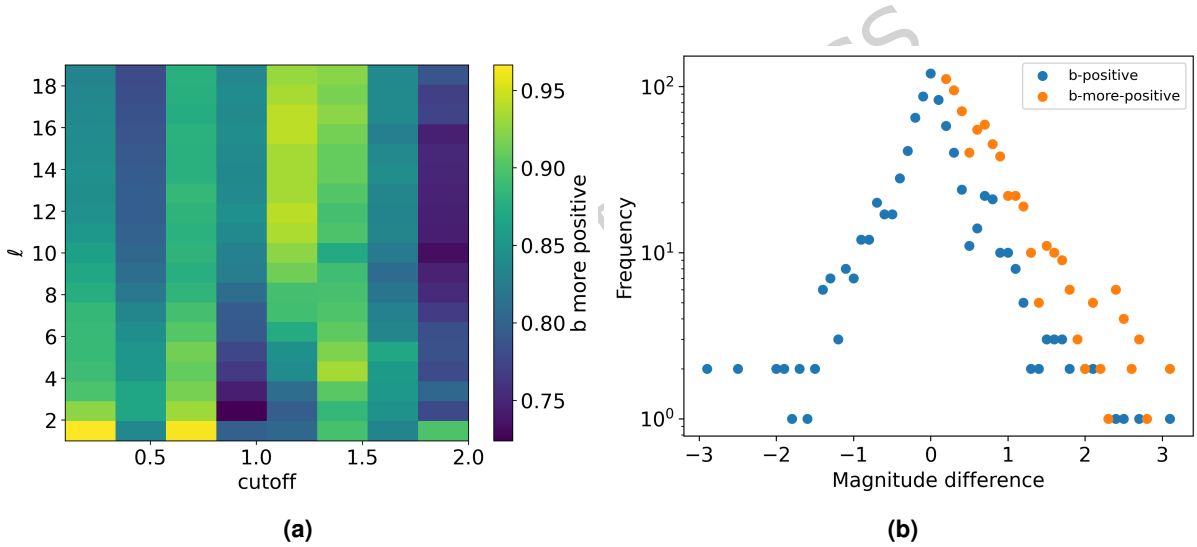


Figure 16. b-more-positive estimator for the Mexican dataset. **(a)** b-more-positive estimator for different cutoffs and lags, **(b)** Magnitude differences used for each estimator, with $\ell = 10$.

b-value	A	α	c	p	d
0.6	1.178e−01	1.121e+00	2.121e−02	1.358e+00	4.724e−03
2.5	1.180e−01	1.106e+00	2.173e−02	1.372e+00	4.979e−03

Table 1. Parameter values for different b-values.

With the aim of obtaining more robust b-value estimators, van der Elst[42] proposed the b-positive estimator. Instead of work with a sample of magnitudes m_i, $i = 1,2,...,m$, which are assumed to follow the Guttenberg Richter law[55], that is, independent exponential random variables with rate $b\log(10)$ shifted by M_c, he proposed using the auxiliary variables $Z_i = |m_{i+1} - m_i|$ with $i = 1,2,...,m-1$. He defines the b-positive estimator as

$$b_+(M_c') = \left(\frac{\sum_{i=1}^{m-1} Z_i \mathbb{1}(Z_i \geq M'c)}{\sum_{i=1}^{m-1} \mathbb{1}(Z_i \geq M'c)} - M'c \right)^{-1} \log(e).$$

In his work, he exhibited through simulations that the estimator can improve the estimation of the b-value, particularly in the presence of incompleteness as modeled by Ogata & Katsura model[77].

One of the main concerns regarding the b-positive estimator is that if we only use the Z_i where $\mathbb{1}(Z_i \geq M'c)$, the sample size decreases significantly, in order to solve this problem, Lippiello & Petrillo[58] have proposed the b-more-positive estimator, they added an extra hyperparameter l to define

$$\vartheta_i^l = \min\{i : |m_{i+j+1} - m_{i+j}| > M'c, j = 0, .., l-1\},$$

for the non empty sets, which it allows to define

$$Z_i' = |m_{i+\vartheta_i^j+1} - m_{i+\vartheta_i^j}|.$$

Then the b-more-positive estimator is defined as

$$b_{+l}(M_c') = \left(\frac{1}{|\mathscr{I}|}\sum_{\mathscr{I}} Z_i - M'c\right)^{-1} \log(e).$$

where $\mathscr{I}$ is the set of indexes which $\{i : |m_{i+j+1} - m_{i+j}| > M'c, j = 0, .., l-1\}$ is non empty. As can be seen in Figure 16b, this allows an increase in the sample size for our estimation, nevertheless, it is important to note that this increases the dependence between the Z_i' variables compared to the Z_i in the b-positive estimator.

Lippiello & Petrillo studied the performance of the b-more-positive estimator through numerical simulations, in their example the b-more-positive showed an improvement with respect to the b-positive approach, in part because the b-positive estimator is a particular case of the b-more-positive one if we take $l = 0$.

Although estimators have proven useful, new areas of research on this new type of estimator emerges, mainly due to the lack of independence of the Z_i' variables and how to obtain maximum likelihood estimators since the current b-positive and b-more-positive estimators rely on the pseudolikelihood. Although this facilitates the calculation of estimators, it remains an approximation of the likelihood.