Bulletin of the Seismological Society of America, Vol. 103, No. 1, pp. 169–179, February 2013, doi: 10.1785/0120110304

# Monitoring Aseismic Forcing in Fault Zones Using Earthquake Time Series

by David Marsan, Elodie Prono, and Agnès Helmstetter

**Abstract** The dynamics of earthquake occurrences is controlled both by fault interaction processes and by long-term, tectonic loading of the faults. In addition, transient loading can be caused by aseismic deformation episodes, for example during crustal fluid migration or slow slip events. These forcing transients are best revealed by geodetic measurements. However, this type of instrumentation is not always available, or is not always sensitive enough to detect significant anomalies. In such cases, one is better off exploiting the seismicity signature of these transients in order to characterize them. We here explore different ways to do so. Interearthquake time statistics are found to be prone to damping out fluctuations in forcing rate. A more accurate method is developed by comparing the data with a triggering model that accounts for earthquake interactions. The changes in fault loading rates are then well recovered, both in duration and in intensity.

## Introduction

Earthquakes are caused by stress build-up on the rupturing fault. This loading can stem from two distinct contributions: (1) aseismic forcing such as tectonic loading, more rapidly time-varying processes such as fluid or magma intrusions, or aseismic transients such as slow earthquakes; and (2) elastic loading by stress transfer from previous earthquakes. These two contributions differ by the fact that (1) is aseismic. It can vary quickly in time, in which case its intensity is difficult to evaluate, most notably when no geodetic information is available.

If we define the earthquake population as the system of interest, then contribution (1) is external, while (2) is internal. When studying large tectonic areas, the main external forcing is tectonic loading. Separating the two contributions then amounts to finding so-called mainshocks and their related aftershocks, which is equivalent to declustering the earthquake catalog (Gardner and Knopoff, 1974; Reasenberg, 1985; Zhuang *et al.*, 2002). However, the rate of mainshock occurrences, hereafter called the forcing (or background) rate, can exhibit large fluctuations in time. This is more particularly the case when investigating small zones characterized by bursts of activity, for example, in the case of volcanic or fluid-induced earthquake swarms. Then the forcing rate typically evolves through rapid, transient changes, forcing the system to switch from low to high seismic activities. Such swarms generally have no clear mainshock–aftershock sequences. Moreover, they are sometimes characterized by hypocenter migration or diffusion, thought to be a signature of fluid movement (Aoyama *et al.*, 2002; Hainzl, 2004; Vidale and Shearer, 2006; Jenatton *et al.*, 2007; Pacchiani and Lyon-Caen, 2010; Daniel *et al.*, 2011). Other examples include swarms related to slow earthquakes (Wolfe *et al.*, 2007) that can also exhibit hypocenter migration (Lohman and McGuire, 2007).

The goal of this paper is to discuss and propose statistical methods for estimating time-varying forcing rates due to fluid/magma intrusions or aseismic deformation transients, by only using earthquake time series in the absence of complementary deformation monitoring based on geodetic measurements (Global Positioning System [GPS], strain or tiltmeters, Interferometric Synthetic Aperture Radar [InSAR]). While the latter can directly reveal and allow characterizing aseismic transients, they either have limited resolution (GPS, InSAR) or are difficult to interpret due to their sensitivity to other, unrelated, possibly very local phenomena (strain and tiltmeters). Moreover, seismicity is by far the most abundant type of data to constrain changes in crustal stress, and it must therefore be accounted for when searching for transient deformation episodes.

We postulate that the rate of earthquakes $\lambda(t) = \mu(t) + \nu(t)$ is the sum of the two contributions mentioned earlier: the aseismic forcing $\mu(t)$ and the forcing $\nu(t)$ due to interactions between earthquakes. To separate between $\mu$ and $\nu$ is a difficult problem: An earthquake time series can always be modeled by assuming there are no interactions between earthquakes, that is, $\nu(t) = 0$, as suggested by Traversa and Grasso (2010) for seismicity swarms triggered by dyke intrusions, so that it results solely from a forcing $\mu(t)$ that exactly mimics the earthquake rate $\lambda(t)$.

The case of aftershock sequences serves here as a good illustration. While most models invoke coseismic stress transfer from the mainshock to the aftershock foci in order

to explain the occurrences of the latter (e.g., Dieterich, 1994), so that the term $\nu(t)$ dominates during these sequences, other models instead assume that aftershocks are caused by afterslip, hence a dominant $\mu(t)$ exists (Perfettini and Avouac, 2007). How to distinguish between these two mechanisms of aftershock triggering is still a matter of debate (Helmstetter and Shaw, 2009).

Earlier works have used the epidemic-type aftershock sequence (ETAS) model (Kagan and Knopoff, 1981; Ogata, 1988) for estimating the time-varying forcing rates $\nu(t)$ during swarms. In all these studies, the swarm is divided into separate time intervals, and the forcing $\nu(t)$ is then estimated individually for these intervals. The forcing is generally assumed to be constant over each individual time interval (Hainzl and Ogata, 2005; Lombardi, Marzocchi, and Selva, 2006; Lombardi, Cocco, and Marzocchi, 2010) although more sophisticated forms have been tested by Matsu'ura and Karakama (2005) to model the Matsushiro swarm. Because a model with $\nu(t) = 0$ is always a possible candidate, the tendency of these earlier studies to find anomalously low interaction terms, most notably low productivity $\alpha$-values (see equation 5 for the definition of $\alpha$) raise the question as to their robustness: these methods are possibly not efficient in determining $\mu(t)$.

Recently, studies on the distribution of waiting times between consecutive earthquakes have shown that the forcing rate could be simply estimated with nonparametric methods. Corral (2004) proposed that the probability density function of interearthquake times $\delta t$ (here not normalized by the mean) can be expressed as

$$f(\delta t) = Ce^{-\mu\delta t}\delta t^{-n}, \tag{1}$$

with $C$ a normalization constant, $\mu$ the forcing rate, and $n$ a model parameter. Hainzl *et al.* (2006), based on an argument by Molchan (2005), further showed that $\mu$ can then be estimated from the first two sample moments of $\delta t$ as

$$\mu = \frac{E\{\delta t\}}{\mathrm{var}\delta t}, \tag{2}$$

where $E\{.\}$ denotes the mean and var is the variance. Furthermore, the exponent $n$ is given by

$$n = 1 - \mu E\{\delta t\}. \tag{3}$$

The $n$ value then theoretically corresponds to the fraction of triggered events among all seismic events and is called the branching ratio (Helmstetter and Sornette, 2003), while $\mu$ is the forcing rate (Hainzl *et al.*, 2006). However, the estimated $\mu$ of equation (2) can be shown to overestimate the real rate, sometimes by several orders of magnitude, when the analyzed time period is dominated by aftershocks (see the Synthetic Datasets section). Moreover, ensemble-averaged (rather than sample-averaged) distributions with different branching ratios $n$ were shown by Sornette *et al.* (2008) to be very similar, questioning the use of interevent times for performing the estimation of $\mu(t)$. We address these issues in this paper.

We first describe how we generate synthetic datasets used to test the methods. We then detail the limitations of exploiting interearthquake time statistics as defined in equation (2), for example, to deduce the time-varying forcing. An improved method that attempts to avoid periods dominated by aftershocks is shown to generally damp the real fluctuations of the forcing rate. A proper account of aftershock triggering is thus required, allowing for a robust and accurate estimation of the contribution of aseismic processes in earthquake time series.

## Synthetic Datasets

### Seismicity Model

In our simulations, earthquakes are characterized by their time of occurrence $t$ and their magnitude $m$. The latter is constrained to verify the Gutenberg–Richter law (Gutenberg and Richter, 1944), which gives the probability density $f(m)$ of magnitude $m$. We only consider positive magnitudes, so that $f(m) = \beta e^{-\beta m}$. There is no upper cutoff in magnitude. The $b$-value of the Gutenberg–Richter law is $b = \beta/\log 10$.

The time series of earthquake occurrences is modeled as an inhomogeneous Poisson process with rate $\lambda(t)$ depending on time. We use ETAS models for generating our synthetic datasets (Kagan and Knopoff, 1981; Ogata, 1988). In these models, the rate is a function of the past history of earthquake occurrences, to simulate mechanical interactions between faults that are responsible for correlations in seismicity dynamics. Namely, the rate is

$$\lambda(t) = \mu(t) + \nu(t), \tag{4}$$

with

$$\nu(t) = \sum_{t_i<t} Ae^{\alpha m_i}(t + c - t_i)^{-p}, \tag{5}$$

where the sum is on all earthquakes $i$ of magnitudes $m_i$ that occurred before $t$, that is, at times $t_i < t$. The first term on the right side of equation (4) is the forcing rate, which gives the rate of background earthquakes. In the absence of any interactions between earthquakes, for $A = 0$, only background earthquakes occur and $\lambda(t) = \mu(t)$. We describe in the following text how $\mu(t)$ is chosen in our simulations.

We take $A = 0.0059$, $\alpha = 2$, $p = 1.2$, $c = 10^{-3}$, and a $b$-value equal to 1. These parameters give a branching ratio $n$ of 0.9; the branching ratio is the average number of direct aftershocks triggered by a mainshock, when averaging over all mainshock magnitudes. It also controls the proportion $1 - n$ of background earthquakes in the overall earthquake population (Helmstetter and Sornette, 2003; Hainzl *et al.*, 2006; Sornette *et al.*, 2008), in the limit of infinitely long time series. These values are in the typical range that has been found when estimating ETAS parameters on real data including seismic swarms (Hainzl and Ogata, 2005; Lombardi, Marzocchi, and Selva, 2006; Daniel *et al.*, 2011).

### Synthetic Dataset $\mathcal{D}_1$: Control Catalog

The first synthetic catalog is a control catalog so that $\mu(t)$ is constant over the whole duration $T$ of the dataset. A constant forcing rate is meant to represent the simple case of a fault system subject to a constant tectonic loading rate. We here take $\mu(t) = 0.4$ and simulate earthquakes from $t = 0$ up to $T = 1000$, assuming there was no earthquake for $t < 0$. Catalog $\mathcal{D}_1$ contains $N = 1034$ earthquakes with magnitudes ranging from just above 0 to 3.64 (see Fig. 1).

### Synthetic Dataset $\mathcal{D}_2$: Swarmlike Seismicity with Sudden Onset

Earthquake swarms are often characterized by a rapid transition from very low to very high activity (Zollo *et al.*, 2002; Jenatton *et al.*, 2007; Lohman and McGuire, 2007; Pacchiani and Lyon-Caen, 2010). For example, fluid- or magma-induced activity in volcanoes, geothermal areas, or fault systems dominated by crustal extension occurs in sudden bursts of earthquakes that can last several days to months. This rapid onset of activity is generally attributed to aseismic processes (e.g., fluid-pulse migration or slow earthquakes), because there is no large magnitude shock at the start of the sequence that could explain it.

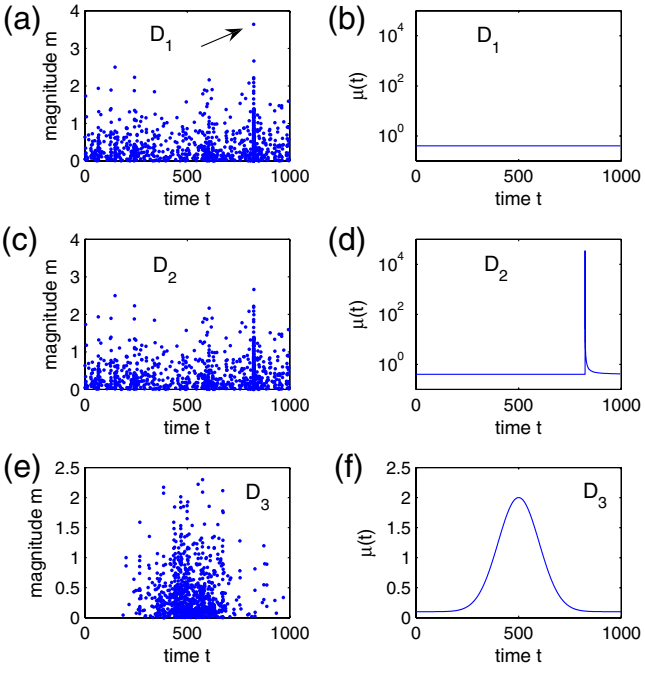


**Figure 1.** The three synthetic catalogs $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$, from top to bottom (a,c,e): magnitude versus time of occurrence. The only difference between catalogs $\mathcal{D}_1$ and $\mathcal{D}_2$ is the magnitude value of the earthquake pointed out in the graph (a): this shock has the largest magnitude of all in $\mathcal{D}_1$ ($m = 3.64$), and has only magnitude $m = 0.43$ in $\mathcal{D}_2$. (b,d,f) forcing rate $\mu(t)$. In $\mathcal{D}_2$, this rate jumps from 0.4 to $3.40 \times 10^4$ per unit time, before decaying. The color version of this figure is available only in the electronic edition.

This type of activity is here modeled by a sharply increasing and then decaying $\mu(t)$. We simulate it from the control catalog $\mathcal{D}_1$ by simply changing the magnitude of the largest shock: the six hundred and fourth earthquake of $\mathcal{D}_1$, with magnitude $m_0$ 3.64, occurring at time $t_0 = 823.83$, in $\mathcal{D}_2$ is given the new magnitude value $m_0' = \frac{1}{\log 10} = 0.43$, which corresponds to the mean magnitude for the Gutenberg–Richter law with a $b$-value of 1.

This change implies that the forcing rate of $\mathcal{D}_2$ is $\mu(t) = 0.4$ for $t < t_0$ and becomes $\mu(t) = 0.4 + A\{e^{\alpha m_0} - e^{\alpha m_0'}\}(t + c - t_0)^{-p}$ for $t > t_0$, see Figure 1.

Catalogs for $\mathcal{D}_1$ and $\mathcal{D}_2$ only differ by one magnitude value. The remaining 1034 occurrence times and 1033 magnitudes are the same for both catalogs. Despite this strong similarity, the forcing rates are very different, for $t > t_0$. This shows that the methods developed here cannot solely rely on the earthquake occurrence times, but they must also make good use of the magnitude information.

### Synthetic Dataset $\mathcal{D}_3$: Slow Transient Deformation with Gradual Onset

The last synthetic catalog has a smoother forcing rate compared with $\mathcal{D}_2$. It is thought of as simulating a seismic swarm that would originate from a slow transient deformation. As an example, the swarms related to silent slip on the south flank of the Kilauea volcano do not have the very sudden, sharp increase in activity and the Omori-type relaxation as modeled with $\mathcal{D}_2$ (Montgomery-Brown *et al.*, 2009). Here we take $\mu(t) = \mu_0 + (\mu_1 - \mu_0)e^{-(t-t_0)^2/2\sigma^2}$ with $\mu_0 = 0.1$, $\mu_1 = 2$, $t_0 = 500$, and $\sigma = 100$. This gives a pulse of activity with the shape of a Gauss curve culminating in a 20-fold increase in the forcing rate (see Fig. 1). We obtain $N = 879$ earthquakes, with a maximum magnitude of 2.29 over the total duration $T = 1000$ of the catalog.

## Using Interearthquake Times to Estimate the Forcing Rate

### Background Rate Estimators

*Estimators for Gamma-Distributed Interearthquake Times.* The possibility of using the distribution of interearthquake times to estimate the forcing rate was proposed and tested by Hainzl *et al.* (2006). The forcing rate was then computed as in equation (2), which uses the method of moments estimators and the first two moments to estimate $\mu$. As is well known in statistical mathematics, this estimate of $\mu$ for a Gamma law as expressed in equation (1) is less efficient (Fisher, 1922) than the maximum likelihood estimate (MLE), denoted $\hat{\mu}_{\mathrm{MLE}}$, which amounts to maximizing

$$\ell(\mu, n) = (1 - n)\log\mu - \log\Gamma(1 - n) - \mu E\{\delta t\} - nE\{\log\delta t\}. \quad (6)$$

This maximization implies that $\hat{\mu}_{\mathrm{MLE}} = \frac{1-n}{E\{\delta t\}}$, and that the function

$$\ell(n) = \bar{n}\log\bar{n} - \bar{n} - \log\Gamma(\bar{n}) \tag{7}$$

be maximized, where $\bar{n} = 1 - n$. The $n$ parameter that maximizes $\ell(n)$ is numerically easy to find because $0 < n < 1$. In the following sections, we will, however, show that $\hat{\mu}_H$ of Hainzl *et al.* (2006) is better suited for the current application of finding the background rate out of an earthquake time series.

*Bias toward Small Rates.* The estimators $\hat{\mu}_H$ of Hainzl *et al.* (2006) and $\hat{\mu}_{\mathrm{MLE}}$ in most but not all cases, for example, during aftershock sequences, correctly find the forcing rate, because it is strongly constrained by the minimum instantaneous rate over the investigated time period. We recall that the earthquake rate $\lambda(t)$ is the sum of the background rate $\mu$, plus triggered rates $\nu$ caused by previous earthquakes. Thus, $\lambda(t) \geq \mu$, but $\lambda(t) \simeq \mu(t)$ during periods devoid of large shocks. This suggests that the estimation of $\mu$ is equivalent to searching for the minimum value $\lambda(t)$ can take.

We now consider the oversimplistic but illustrative example where a time series is made of a Poisson process with constant rate $\lambda_1$ for $0 < t < 1$, followed by another Poisson process with constant rate $\lambda_2 \geq \lambda_1$ for $1 < t < 1 + \frac{\lambda_1}{\lambda_2}$. By construction, both time windows contain on average the same number of earthquakes. We run the estimator $\hat{\mu}_H$ of Hainzl *et al.* (2006) on the whole time period $0 < t < 1 + \frac{\lambda_1}{\lambda_2}$ at once, that is, considering the two time intervals $0 < t < 1$ and $1 < t < 1 + \frac{\lambda_1}{\lambda_2}$ together rather than separately. This mimics an oversimplifed case when one tries to estimate the minimum rate $\min_t \lambda(t)$ in a nonhomogeneous Poisson time series with a time-varying rate $\lambda(t)$. On ensemble average,

$$\hat{\mu}_H = \frac{2\lambda_1\lambda_2(\lambda_1 + \lambda_2)}{(\lambda_2 - \lambda_1)^2 + 2(\lambda_1^2 + \lambda_2^2)}. \tag{8}$$

For $\lambda_1 = \lambda_2$, this gives $\hat{\mu}_H = \lambda_1$, as expected. For $\lambda_2 \gg \lambda_1$, it yields $\hat{\mu}_H = \frac{2}{3}\lambda_1$; hence, the very high rate $\lambda_2$ is somewhat ignored, and the estimator instead finds a value very close to the minimum rate $\lambda_1$. Similarly, $\hat{\mu}_{\mathrm{MLE}}$ is mostly controlled by $\lambda_1$ when $\lambda_2 \gg \lambda_1$ (see Fig. 2a).

We can make the model slightly more complex by changing the duration of the second time window that now extends from 1 to $1 + \frac{x}{y}$, with $y = \frac{\lambda_2}{\lambda_1}$, so that $x$ is the ratio of the average number of earthquakes in the second time window to the average number of earthquakes in the first time window ($x = 1$ was assumed so far). We then find that, on average,

$$\hat{\mu}_H = \lambda_1 \frac{(1+x)\times\left(1+\frac{x}{y}\right)}{\left(1-\frac{x}{y}\right)^2 + 2x\left(1+\frac{1}{y^2}\right)}. \tag{9}$$

As an example, the curve $\hat{\mu}_H$ versus $y = \frac{\lambda_2}{\lambda_1}$ is shown in Figure 2 for both $x = 1$ and $x = 10^3$. In the latter case, although the second time window has $10^3$ times more earthquakes, the estimator $\hat{\mu}_H$ still has a mean value close to the minimum rate $\lambda_1$, when $\lambda_2 \gg \lambda_1$.

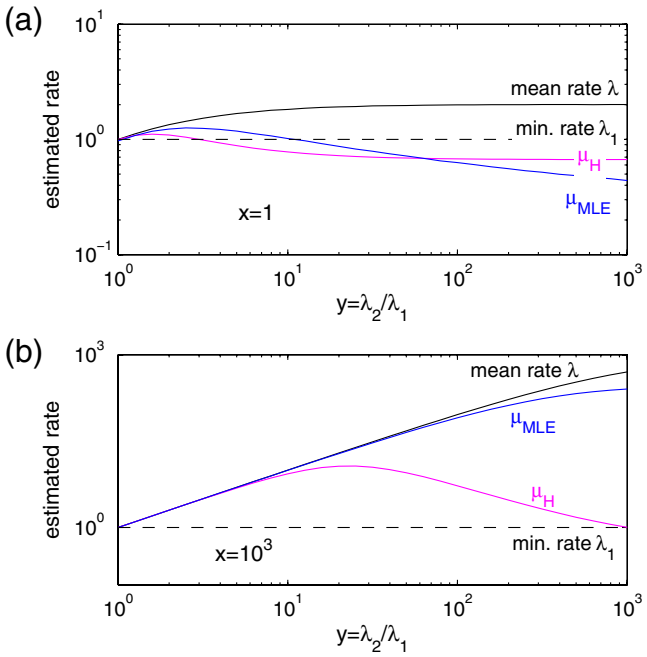


**Figure 2.** Estimated rate versus $y = \frac{\lambda_2}{\lambda_1}$ for our simple model with two nonoverlapping homogeneous Poisson processes (see the Bias toward Small Rates section). The rate $\lambda_1$, that is, for the first time window, is equal to 1. We compare the estimator $\hat{\mu}_H$ of Hainzl *et al.* (2006), see equation (9), to the MLE estimator $\hat{\mu}_{\mathrm{MLE}}$ for Gamma-distributed interearthquake times. The mean rate of equation (10) is also shown. All three rates are computed by taking the two time windows as a whole. (a) $x = 1$, hence the two time windows have on average the same number of earthquakes. (b) $x = 10^3$, so there are on average 1000 times more earthquakes in the second window. The estimator of Hainzl *et al.* (2006) approximates $\lambda_1$ even when $\lambda_2 \gg \lambda_1$ so that there is only a small portion of the earthquakes generated with rate $\lambda_1$. The color version of this figure is available only in the electronic edition.

The fact that $\hat{\mu}_H$ is strongly biased toward the minimum rate $\lambda_1$ in this simple model is indeed a desirable feature: the estimator is well designed to find the forcing (hence, minimum) rate out of an inhomogeneous Poisson process.

As a comparison, the simple statistic $\hat{\lambda} = 1/E\{\delta t\}$, which is the mean rate, is on average, for our model

$$\hat{\lambda} = \lambda_1 \frac{1+x}{1+\frac{x}{y}}. \tag{10}$$

Figure 2 shows that this $\hat{\lambda}$ rapidly departs from $\lambda_1$ when $x \gg 1$ and $y \gg 1$, unlike $\hat{\mu}_H$. Also, the same departure is observed for $\hat{\mu}_{\mathrm{MLE}}$ when assuming Gamma-distributed interearthquake times: this mixture of two very different Poisson processes (at $x \gg 1$ and $y \gg 1$) causes the actual distribution of $\delta t$ to strongly differ from a Gamma law, hence this departure.

Moving from this academic example to one directly related to earthquake time series, we now consider a synthetic dataset $\mathcal{D}$ obtained by running the ETAS model with

parameters as for $\mathcal{D}_1$, except that the background rate is here set to $\mu(t) = 1$. We zoom in on its largest shock of magnitude $m = 4.8$. We select $n_1$ (varying between 0 and 60) earthquakes before, and $n_2 = 1000$ earthquakes after this mainshock. We then apply both the estimators $\hat{\mu}_H$ and $\hat{\mu}_{\mathrm{MLE}}$ to these subsets.

As shown in Figure 3, the estimated rate is very high ($> 10^4$) for $n_1 = 0$ for both estimators. The first 1000 earthquakes after the large shock are mostly aftershocks, occurring at a high rate that masks the background rate. As a matter of fact, these 1000 earthquakes occur within a time equal to 0.0215 after the mainshock, implying that the probability of having at least one background earthquake among those 1000 events is very small (2.15%). However, as soon as we add one earthquake prior to the mainshock ($n_1 = 1$) to our subset, $\hat{\mu}_H$ drops to less than 10. It is remarkable that such an estimate is found, because the subset is made of 1000 earthquakes occurring at an average rate of $4.6 \times 10^4$ events per unit time, plus only 2 (the earthquake prior to, and the mainshock itself) events occurring at the background rate of 1. The background rate is eventually correctly estimated when $n_1 > 50$. A similar behavior, although less pronounced, is also observed with $\hat{\mu}_{\mathrm{MLE}}$.

We have shown that the two estimators $\hat{\mu}_H$ and $\hat{\mu}_{\mathrm{MLE}}$, when applied to a time window with fluctuating instantaneous rate $\lambda(t)$, yield estimated rates close to the minimum of $\lambda(t)$ over this time window. Thus, if the time window contains background earthquakes, then the estimated rate is the background rate. However, if the time window only contains aftershocks of previous earthquakes, then both $\hat{\mu}_H$ and $\hat{\mu}_{\mathrm{MLE}}$ will overestimate the background rate, the more so as the rate of aftershocks is high. These features can explain the negative results of Sornette *et al.* (2008) regarding the use of $\hat{\mu}_H$ for estimating the forcing rate. Thus, it is necessary to first evaluate whether the analyzed time window could be made of only aftershocks, in which case this time window must be made longer in order to also contain background earthquakes. We now describe an attempt at achieving so.

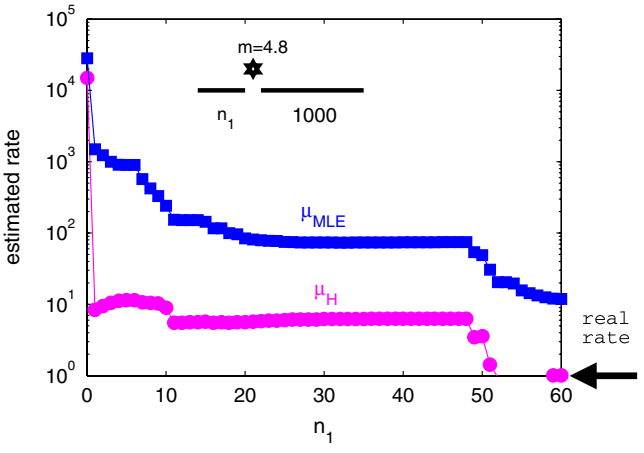


**Figure 3.** Estimated rate for a time window containing $n_1$ earthquakes before and 1000 earthquakes after the largest shock of dataset $\mathcal{D}$ using the MLE estimator $\mu_{\mathrm{MLE}}$ and the estimator $\mu_H$. The color version of this figure is available only in the electronic edition.

### Description of the Method

The rationale is to find, for any time $t$, the smallest time window centered on $t$ that is not made only of aftershocks. We first compute $\hat{\mu}_H$ for sliding windows. The estimator $\hat{\mu}_H$ is here preferred to $\hat{\mu}_{\mathrm{MLE}}$ for its better capacity to estimate the minimum rate (see the Bias toward Small Rates section. For each earthquake $i$, $\hat{\mu}_H^{(1)}$ is estimated for the window extending from event $i - \rho$ to event $i$, and $\hat{\mu}_H^{(2)}$ from $i$ to $i + \rho$. Both windows thus contain $\rho$ interevent times. A measure of the relative change in rate is given by

$$\Delta\mu = (\hat{\mu}_H^{(2)} - \hat{\mu}_H^{(1)}) \times \left( \frac{1}{\hat{\mu}_H^{(1)}} + \frac{1}{\hat{\mu}_H^{(2)}} \right). \qquad (11)$$

When the variation is greater in absolute value than a threshold parameter $\xi$, we examine whether this effectively amounts to a change in the forcing rate. We consider the set of all times $T' = \{t'_1, t'_2, \ldots\}$ when $|\Delta\mu| > \xi$, so that $[t'_n, t'_{n+1}]$ contains at least $\rho$ events. We recompute $\hat{\mu}_H$ for the windows extending between two consecutive $t'$. Such windows correspond to time intervals with no strong changes in estimated rates. A change at $t'_n$ is assumed to be either due to the occurrence of a large shock and the triggering of its aftershocks or to a real change in the forcing rate.

If no large shock is found within a window $[t'_n, t'_{n+1}]$, that is, all the earthquakes in this window have magnitudes less than a threshold magnitude $M_S$, we consider that the variation is really caused by a change in the forcing rate, and we keep this window and corresponding rate $\hat{\mu}_H$ as they are. On the contrary, if there is a large shock with magnitude $m > M_S$, then the time window is likely to contain only aftershocks. In that case, we remove $t'_{n+1}$ from the subset of times $T'$ and merge this time window with the next one, hence defining an extended, new window $[t'_n, t'_{n+2}]$. We subsequently go on merging this new time window $[t'_n, t'_{n+2}]$ with the next ones $[t'_{n+2}, t'_{n+3}, \ldots]$, as long as $\hat{\mu}_H$ obtained for $[t'_{n+2}, t'_{n+3}, \ldots]$ is greater than $\hat{\mu}_H$ obtained just prior to the large shock, hence for $[t'_{n-1}, t'_n]$. The rationale is simply to select a time window longer than the duration of the aftershock sequence of a large shock.

Finally, we remove the time windows that contain too few events, as natural randomness of the statistics $\hat{\mu}_H$ can give birth to short time windows even in the absence of variation of the forcing rate. Namely, if the number of events in a window is lower than a threshold parameter $\rho_2$, we remove $t'_n$ from the subset of time $T'$, and this window is merged with the previous one. To do this correction, we start with the window characterized by the highest forcing rate, and then proceed to the second largest and so on, rather than in chronological order as we did previously. If a time window before a real variation in the forcing rate is too small, it will merge

with the previous window rather than with the next one, and we thus preserve the onset of the change in forcing.

### Parameterization

Four parameters are used in this algorithm. First, $\rho$ is the number of events in the initial time windows used to compute $\hat{\mu}_H^{(1)}$ and $\hat{\mu}_H^{(2)}$. Second, $\rho_2$ is the minimal number of events in the final time windows. The time windows used to compute $\hat{\mu}_H^{(1)}$ and $\hat{\mu}_H^{(2)}$ must contain enough events to be statistically significant, but must be short enough to show variations in the forcing rate. As a trade off, we limit the total number of windows for the entire dataset to 50, that is, $\rho_2 = N/50$ events, with $N$ the total number of events, and take $\rho_2 = 2\rho$ with a minimum of 10 interevent times in order to reach a statistically significant estimate.

Third, $\xi$ is the threshold parameter to detect significant changes in the forcing rate, when $|\Delta\mu| > \xi$. Considering a Poisson process with a constant rate $\beta$, the time $t$ to observe $\rho$ successive events has the probability density function following a Gamma law $f(t) = \beta^\rho \frac{t^{\rho-1}}{(\rho-1)!} e^{-\beta t}$. The sample rate $\hat{\mu} = \frac{\rho}{t}$ thus has the probability density function

$$f(\hat{\mu}) = \frac{1}{\hat{\mu}(\rho-1)!}\left(\frac{\beta\rho}{\hat{\mu}}\right)^\rho e^{-\beta\rho/\hat{\mu}}. \qquad (12)$$

For $\rho \gg 1$, $\hat{\mu}$ tends to a normal law $\mathcal{N}(\beta, \beta/\sqrt{\rho})$ with mean $\beta$ and standard deviation $\frac{\beta}{\sqrt{\rho}}$. Then $\hat{\mu}^{(1)} - \hat{\mu}^{(2)}$ tends to $\mathcal{N}(0, \beta\sqrt{\frac{2}{\rho}})$ and $\Delta\mu$ to $\mathcal{N}(0, 2\sqrt{\frac{2}{\rho}})$. The statistics $\Delta\mu$ does not depend on the initial rate $\beta$ anymore. The probability of $\Delta\mu$ being lower than the threshold parameter $\xi$ is therefore

$$P(\Delta\mu \le \xi) = \frac{1}{2}\left[1 + \mathrm{er}\,f\left(\frac{\xi\sqrt{\rho}}{4}\right)\right], \qquad (13)$$

where er $f$ is the error function. Thus, the probability $P$ that $|\Delta\mu|$ is greater than $\xi$ is

$$P = P(|\Delta\mu| > \xi) = 1 - \mathrm{er}\,f\left(\frac{\xi\sqrt{\rho}}{4}\right). \qquad (14)$$

Given a value of $P$, the threshold parameter $\xi$ is chosen as $\xi = \frac{4}{\sqrt{\rho}}\mathrm{er}\,f^{-1}(1-P)$. In the following, we take $P = 10\%$.

Finally, $M_S$ is the threshold magnitude over which an event is considered as a large shock. To constrain this parameter, we use a Monte Carlo test with 64 synthetic datasets based on the background rate $\mu(t)$ as in the synthetic $\mathcal{D}_2$, that is, we run 64 independent realizations of the same ETAS model. Parameters $\rho$, $\rho_2$, and $\xi$ are fixed as described previously. We run our algorithm with $M_S$ varying from $M_{\max} - 0$ to $M_{\max} - 1.0$, with $M_{\max}$ representing the highest magnitude. We measure the quality of the estimated background rate by computing the error

$$\mathcal{E} = E\{[\log_{10}(\mu) - \log_{10}(\hat{\mu}_H)]^2\}, \qquad (15)$$

that is, the departure of the estimated $\hat{\mu}_H(t)$ from the true $\mu(t)$ in logarithmic scale. The mean and median of $\mathcal{E}$ for each tested value of $M_S$ are shown in Figure 4. Our favored value of $M_S$ is $M_{\max} - 0.5$, although the quality of the estimate does not vary much as long as $M_S$ is greater than $M_{\max} - 0.9$.

We performed the same test with datasets $\mathcal{D}_1$ and $\mathcal{D}_3$ (see Fig. 5). For $\mathcal{D}_1$, the result is the same for all tested values of $M_S$. For $\mathcal{D}_3$, the best result is obtained for parameter $M_S$ between $M_{\max} - 0.4$ and $M_{\max} - 0.6$. Thus, we conclude that $M_S$ is optimally taken close to $M_{\max} - 0.5 \pm 0.1$. In the following, we fix $M_S$ as $M_S = M_{\max} - 0.4$.

### Results

The results for datasets $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$ are shown in Figures 4 and 5. For all three datasets, the variations of $\hat{\mu}_H(t)$ reproduce well the actual variations of $\mu(t)$, with relative departure typically within a ratio of $1/2$ to 2. However, the estimate is clearly too low for several realizations of the model $\mathcal{D}_2$ (see Fig. 6). Indeed, our estimator is conservative because it cannot account for variations of $\mu(t)$ during an aftershock sequence. This happens for several realizations of $\mathcal{D}_2$, for which the algorithm defines a unique time

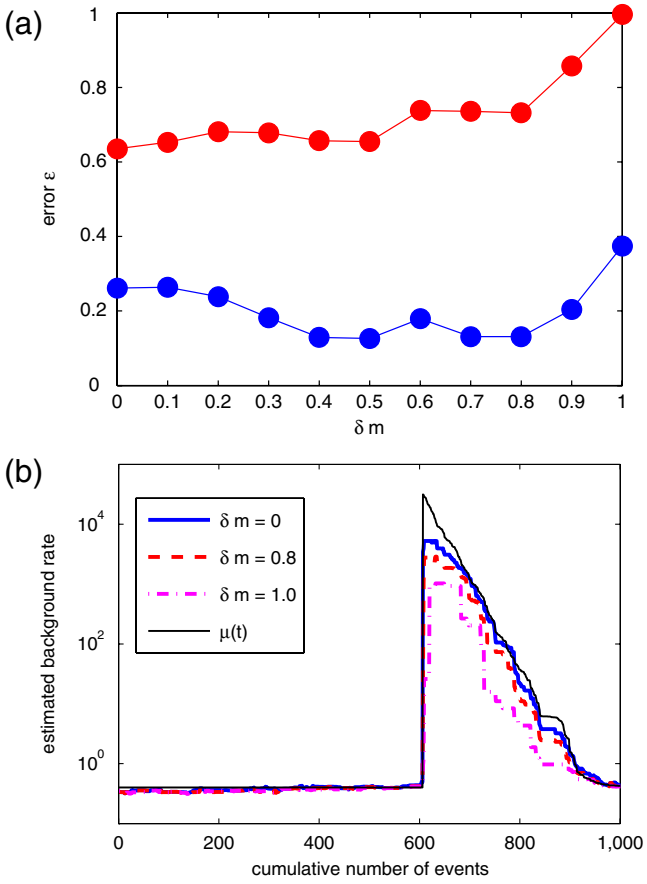


**Figure 4.** (a) Mean (upper curve) and median (lower curve) value of the error $\mathcal{E}$ defined by (15) for 64 realizations of $\mathcal{D}_2$ for $M_S = M_{\max} - \delta m$ with $\delta m$ varying from 0 to 1.0. (b) Corresponding estimated forcing rates, only shown for several values of $\delta m$. The curves show the median estimates for $M_S = M_{\max} - \delta m$. The color version of this figure is available only in the electronic edition.

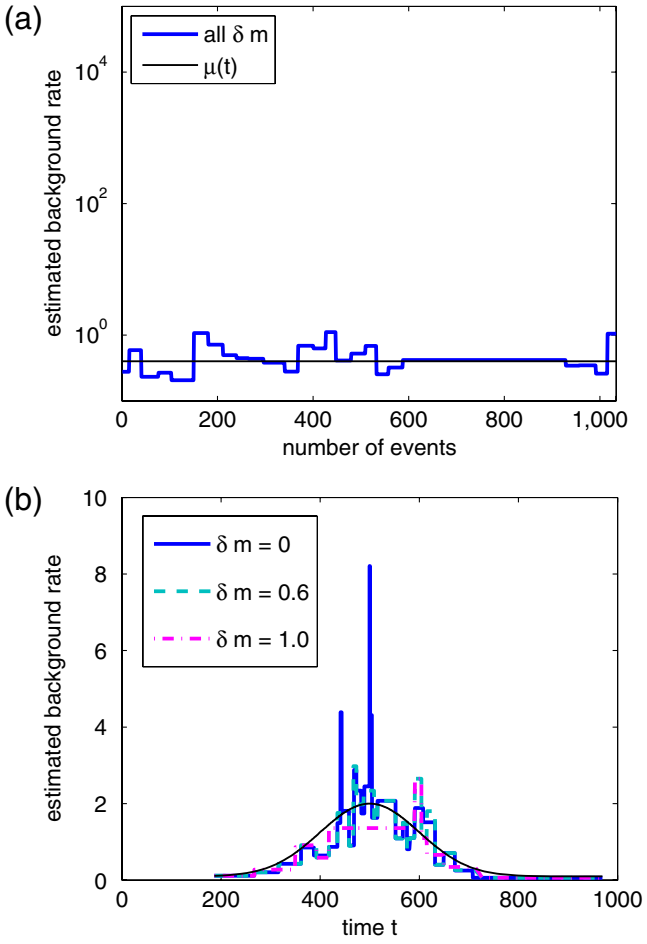


**Figure 5.** (a) Same as Figure 4 for the synthetic datasets $\mathcal{D}_1$ and (b) $\mathcal{D}_3$. Note that we plot the rate as a function of time in graph (b), while it is function of the cumulative number of earthquakes in graph (a) and in Figure 4b. The color version of this figure is available only in the electronic edition.

window extending from the jump in the forcing rate until the end of the dataset.

Thus, we conclude that interearthquake time statistics can be used to estimate the time-varying forcing rate, although it is necessary to adjust the lengths of the time windows to avoid probing only aftershocks. When high rates of aftershocks dominate the time series, these lengths must be increased, and potentially important fluctuations in the forcing rate are then damped. This clearly limits the use of these statistics for detecting short-lived forcing transients, especially if such transients occur in conjunction with sufficiently large earthquakes. The aftershocks triggered by the latter then are likely to mask the forcing transients, because the use of the interearthquake time statistics then require time windows potentially too long compared with the durations of the transients.

## Method Based on Triggering Models

We now develop a more sophisticated algorithm, inspired from Zhuang *et al.* (2002), but here recast it for a purely temporal dataset. The rationale is to estimate both contributions $\mu(t)$ and $\nu(t)$ to the observed rate $\lambda(t)$. The interaction term $\nu(t)$ is here modeled using ETAS, although other models could also be exploited. Its parameters $\{A, \alpha, p, c\}$ (equation 5), are inverted, and the forcing rate $\mu(t)$ is then deduced. We assume that only $\mu$ can vary with time, while the parameters $A$, $\alpha$, $p$, and $c$ are constant. This amount neglects possible changes in the processes involved in earthquake nucleation and triggering that govern earthquake interactions, in comparison with a time-varying stress loading by aseismic processes that can freely, and possibly wildly, fluctuate in time. More specifically, the method consists of:

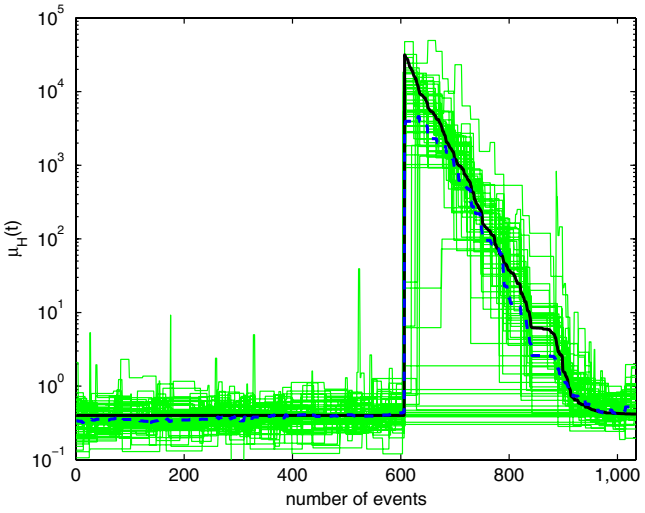


**Figure 6.** Estimated forcing rate $\hat{\mu}_H$ (thin lines) for 65 realizations ($\mathcal{D}_2$ and the 64 synthetics based on $\mu(t)$ from $\mathcal{D}_2$) and their median (dashed line). The real $\mu$ is shown as the thick line. All rates are computed at the times $t_i$ of occurrence of the 1034 events in the synthetic $\mathcal{D}_2$. The color version of this figure is available only in the electronic edition.

1. initially assuming a constant forcing rate $\mu(t) = \mu$, using either estimators $\hat{\mu}_H$ or $\hat{\mu}_{\mathrm{MLE}}$, this gives the first estimate $\mu^{(1)}(t)$;
2. computing the MLE ETAS parameters $\theta^{(1)} = \{A^{(1)}, \alpha^{(1)}, p^{(1)}, c^{(1)}\}$ knowing $\mu^{(1)}(t)$;
3. updating the estimate of the forcing rate based on these parameters, this yields $\mu^{(2)}(t)$;
4. computing $\theta^{(k)}$ knowing $\mu^{(k)}(t)$, and computing $\mu^{(k+1)}(t)$ knowing $\theta^{(k)}$, for $k \geq 2$, until convergence of both the ETAS parameters and the forcing rate is reached.

The second step consists of maximizing the (log-)likelihood $\ell$ of the ETAS model, given the data and the current estimate $\mu(t)$:

$$\ell(\theta|\mu(t)) = -\int_0^T dt\lambda(t) + \sum_{i=1}^{N} \ln \lambda(t_i), \qquad (16)$$

where the integral is computed on the duration $0 \leq t \leq T$ of the dataset, and the sum is performed on all $N$ earthquakes $i$ in this dataset. Here we use equations (4) and (5) that define $\lambda(t)$.

The third step first requires the computation of the probability $\omega_i$ that earthquake $i$ is a background earthquake, for all $i$. This probability is defined as

$$\omega_i = \frac{\mu(t_i)}{\mu(t_i) + \nu(t_i)}. \tag{17}$$

The forcing rate is then obtained by smoothing these probabilities over time:

$$\mu(t) = \frac{1}{t_b - t_a} \sum_{i=a}^{b} \omega_i, \tag{18}$$

where indices $a$ and $b$ are the indices of the earthquakes such that there are $n_e/2$ earthquakes in between $t - t_a$ and $t$ and $n_e/2$ earthquakes in between $t$ and $t + t_b$. This defines a window of $n_e$ events centered on $t$.

The resulting $\mu(t)$ obtained after convergence strongly depends on the smoothing parameter $n_e$. This is illustrated in Figure 7, where the estimated $\mu(t)$ is shown for dataset $\mathcal{D}_1$ and $n_e = 4$, 10, and 1000, and in Figure 8 for dataset $\mathcal{D}_2$ and $n_e = 4$, 40, and 350. For small values of $n_e$, the estimated $\mu(t)$ is allowed to vary quickly, and the solution where $\nu \ll \mu$ is obtained: the observed rate of earthquakes is explained by an *ad hoc* time-varying forcing rate and very little interactions between earthquakes. The vanishingly small ETAS parameter $A$ at $n_e \to 1$ indeed causes the interaction term $\nu$ to be negligible compared with $\mu$. On the contrary, for large $n_e$, the estimated $\mu(t)$ tends toward a constant rate, with a value very close to $\hat{\mu}_H$ and $\hat{\mu}_{\mathrm{MLE}}$ computed using the whole dataset.

Zhuang *et al.* (2002) suggest to choose the best $n_e$ as the parameter that minimizes the discrepancy between the numbers $N_1$ and $N_2$ of triggered earthquakes defined as:

$$N_1 = \sum_{i=1}^{N} 1 - \omega_i, \tag{19}$$

and

$$N_2 = \sum_{i=1}^{N} A e^{\alpha m_i} \frac{(T + c - t_i)^{1-p} - c^{1-p}}{1 - p}. \tag{20}$$

However, in our method, both numbers always remain very close to each other, whatever $n_e$, as shown in Figure 9. This is a consequence of updating $\mu(t)$ with equation (18), which forces the total number of background earthquakes to equal the time integral of the forcing rate.

To select the appropriate smoothing parameter $n_e$, we instead propose to search for the minimum of the Akaike Information Criterion (AIC):

$$\mathrm{AIC}(n_e) = -\ell + \frac{N}{n_e}, \tag{21}$$

where $\ell$ is the log-likelihood of equation (16), and $N$ is the number of earthquakes in the dataset. The second term is a proxy for the number of free parameters in $\mu(t)$ and penalizes models with low $n_e$ values that allow $\mu(t)$ to vary too quickly. We then obtain the minimum AIC for $n_e = 1000$ and $n_e = 40$ for catalogs $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively (Figs. 7 and 8). The selected $\mu(t)$ is seen to well approximate the real forcing rate, shown with the black curve.

For dataset $\mathcal{D}_2$, we compute the Kullback–Leibler discrepancy $\Delta_{\mathrm{KL}}$, that measures how close the solution obtained for a given $n_e$ is from the true $\mu(t)$. Here we can compute this discrepancy because we know the true forcing rate by construction. We use the same 64 independent realizations of the model that generated $\mathcal{D}_2$, and each time run our inversion of

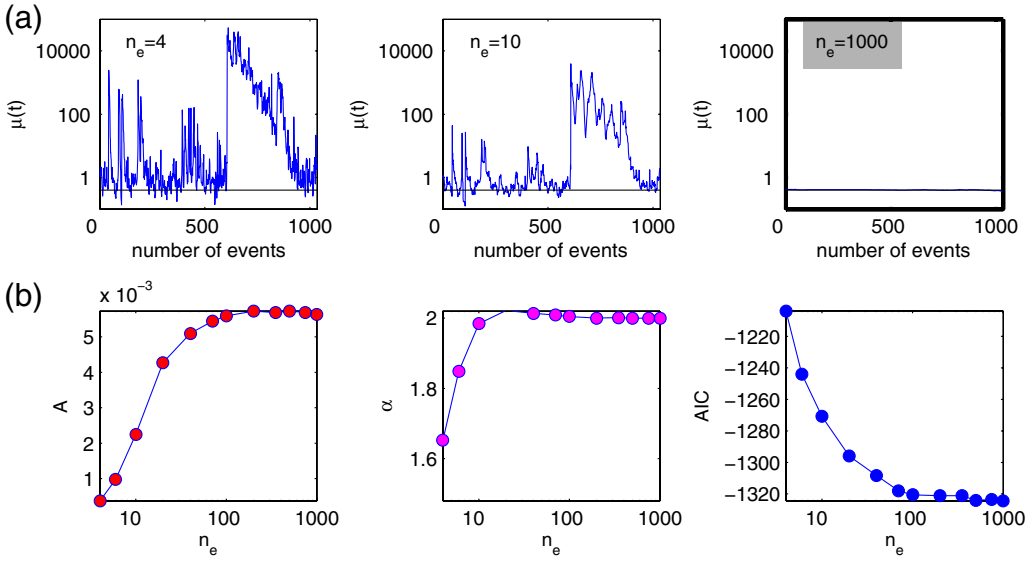


**Figure 7.** (a) Estimated forcing rate $\mu(t)$ for the synthetic dataset $\mathcal{D}_1$ and three values of the smoothing parameter $n_e$ as labeled. The real $\mu(t)$ is the horizontal line $\mu(t) = 1$. (b) ETAS parameters $A$, $\alpha$, and AIC, function of $n_e$. We recall that the dataset was generated with $A = 0.0059$ and $\alpha = 2$. The minimum AIC is obtained for $n_e = 1000$. The color version of this figure is available only in the electronic edition.

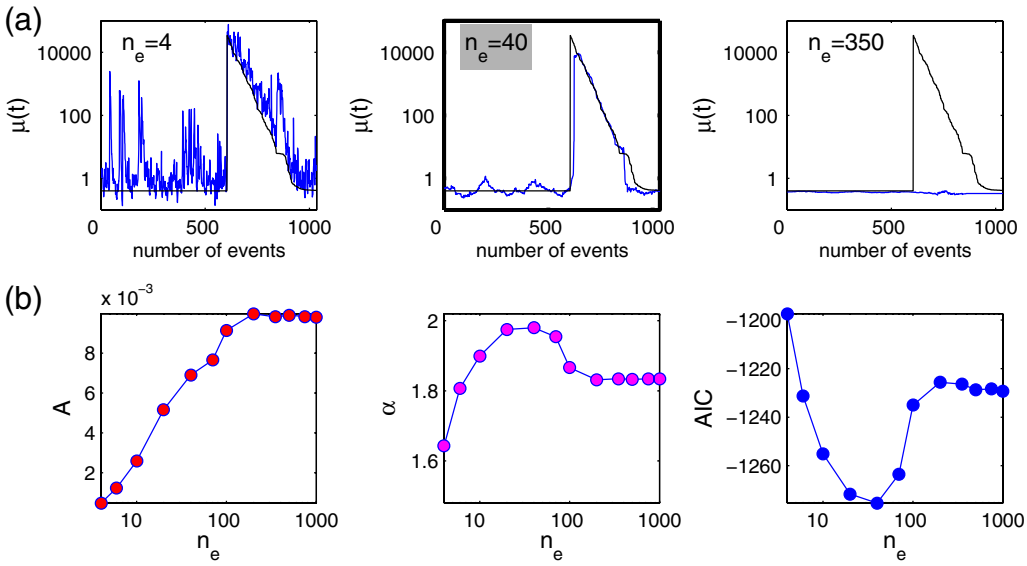


**Figure 8.** Same as Figure 7, but for the synthetic catalog $\mathcal{D}_2$. The minimum AIC is here obtained for $n_e = 40$. The color version of this figure is available only in the electronic edition.

$\theta$ and $\mu(t)$ as explained previously, for a set of values for the smoothing parameter $n_e$. Then all 64 solutions are tested against all 64 synthetic datasets: namely, we compute the log-likelihood $\ell_{i,j}$ of the ETAS parameters and the forcing rate inverted for the $i$th dataset, given the data of the $j$th dataset. The Kullback–Leibler discrepancy is finally given by $\Delta_{\mathrm{KL}} = -E\{\ell\}$, the average being taken over the $64^2$ pair combinations $(i, j)$.

Figure 10 shows that the minimum is found for $n_e = 30$, hence close to the $n_e = 40$ value found with the AIC (see Fig. 7). This shows that the AIC is indeed a good criterion for selecting the smoothing parameter $n_e$.

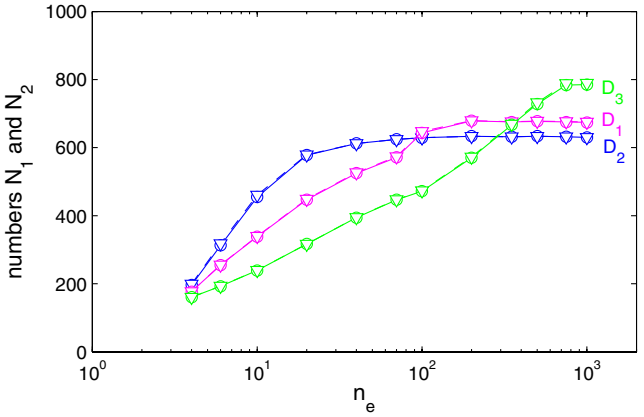


**Figure 9.** Total number of triggered earthquakes $N_1$ (circles) and $N_2$ (triangles) computed as in equations (19) and (20) for the three synthetic datasets $\mathcal{D}_1$, $\mathcal{D}_2$, and $\mathcal{D}_3$, functions of the smoothing parameter $n_e$. As expected, these numbers drop with decreasing $n_e$. There is no visual departure between $N_1$ and $N_2$, preventing their use for selecting the best $n_e$. The color version of this figure is available only in the electronic edition.

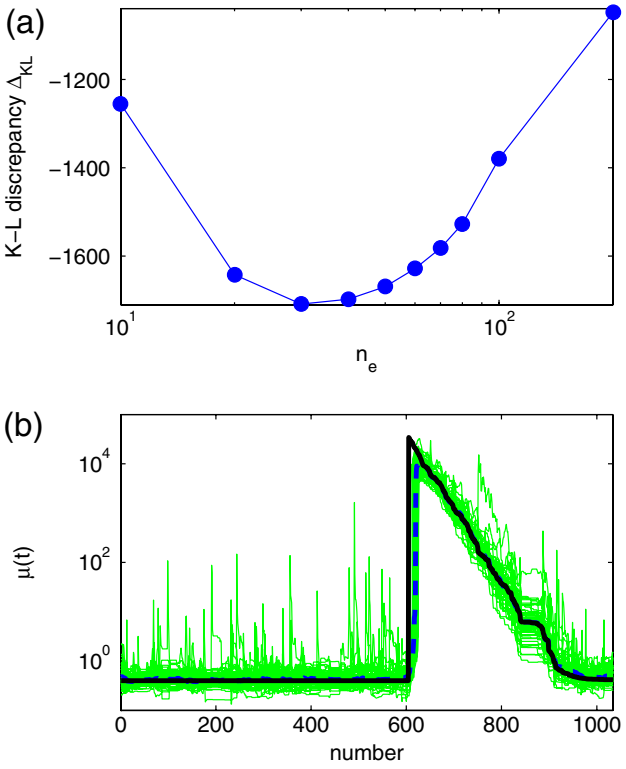


**Figure 10.** (a) Kullback–Leibler discrepancy defined as $\Delta_{\mathrm{KL}} = -E\{\ell\}$ where the average of the log-likelihood is performed on $64^2$ pair combinations of synthetic realizations and inversions. The minimum is obtained for $n_e = 30$. (b) True background rate (thick line) compared with all 64 estimated rates (thin lines) and their median (dashed line) for $n_e = 30$. All rates are computed at the times $t_i$ of occurrence of the 1034 earthquakes in the synthetic $\mathcal{D}_2$ and shown as a function of the number $i$. The median is found to be a very good approximation of the true rate. The color version of this figure is available only in the electronic edition.

We also run this method on the synthetic $\mathcal{D}_3$ (Fig. 11). Again, a good approximation of the real $\mu(t)$ is found, here for $n_e = 70$. We thus conclude that this method is able to correctly estimate the real forcing rate in all situations: constant rate, rapidly varying rate, and slowly varying rate. It provides a better estimate than methods based on interearthquake time statistics, at the cost of a more involved computation, in particular during the second step, which requires inverting the ETAS parameters.

Previous methods for estimating $\mu(t)$ based on the ETAS model failed to consider the dependence of the estimated model parameters ($A$, $\alpha$, $p$, and $c$) on the estimated $\mu(t)$. For example, Lombardi, Marzocchi, and Selva (2006) and Lombardi, Cocco, and Marzocchi (2010) estimate these parameters by imposing a constant $\mu$, and then update $\mu(t)$ with a piecewise uniform background model. However, the optimized ETAS parameters are different if using this updated $\mu(t)$, so that an iterative procedure as proposed here must be run. Similarly, Llenos and McGuire (2011) assume *a priori* ETAS parameter values without considering the possible dependence of their estimates on $\mu(t)$.

The division in independent time intervals, for which the ETAS parameters and the (locally constant) background rate are inverted (Hainzl and Ogata, 2005) can also yield misleading estimates if the durations of these time intervals are not optimized. In particular, low $\alpha$-values are typically found for swarm seismicity with these methods. For example, Hainzl and Ogata (2005) found $\alpha = 0.73$ for the Vogtland swarm and mention that this value "is similar to previous findings for Japanese earthquake swarm activity, where the $\alpha$-value has been found to scatter in the range [0.35,0.85], in contrast to nonswarm activity which is characterized by higher values, namely, $\alpha \in [1.2, 3.1]$". As shown in Figure 8, underestimation of $\alpha$ is indeed a signature of a badly optimized smoothing parameter: actual fluctuations of $\mu(t)$ cannot be explained by the model unless too strong a weight is given to the triggering by small earthquakes, hence a low $\alpha$-value.

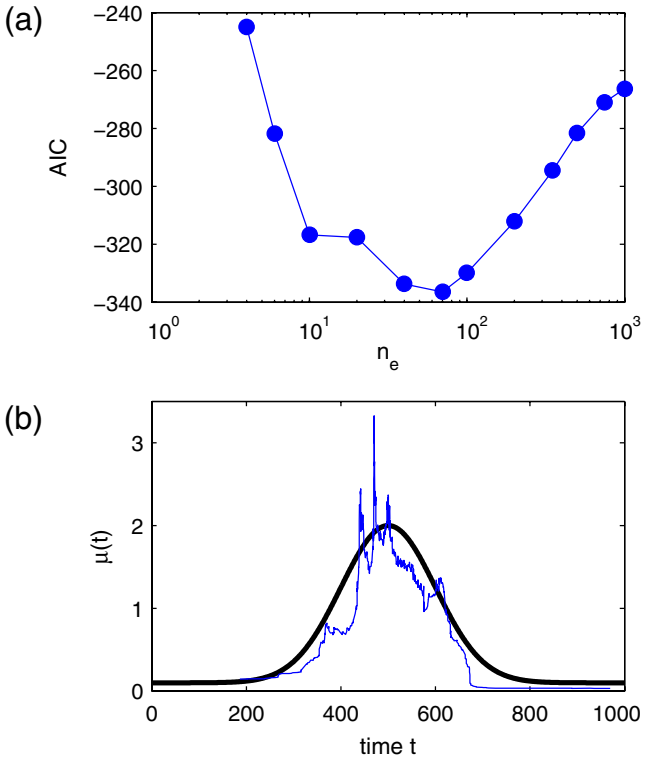


**Figure 11.** (a) AIC function of $n_e$ for synthetic catalog $\mathcal{D}_3$. (b) Estimated (thin line) and real (thick line) background rate $\mu(t)$. The color version of this figure is available only in the electronic edition.

## Conclusions

Determining the changes in the forcing rate that best characterize the evolution of seismicity is a delicate issue. A simple way to model any earthquake time series is to assume the absence of triggering between earthquakes, so that the time series is only explained by *ad hoc* changes in the forcing rate. To explore other possible, more realistic models, some smoothing of the rates has to be performed, and, more importantly, this smoothing has to be optimized. We have shown that this smoothing corresponds to adequately choosing the duration of the time windows when using interearthquake time statistics. Unfortunately, this can damp real and important changes in the forcing rate. Such fluctuations are well estimated with the methods based on triggering models, for which the smoothing parameter is selected according to an information criterion.

The use of these methods for anaylzing earthquake datasets is particularly appropriate when considering swarms, either in extensional or in volcanic zones, but also to detect and characterize aseismic forcing transients that impact the seismicity dynamics, for example, in subduction zones. Other applications that will be further explored in subsequent works relate to dynamic triggering: seismic activity can be first initiated by the passage of waves radiated by remote, large mainshocks; this activity then continues past the wave train because of local earthquake interactions (Brodsky, 2006). For example, this type of pattern was proposed for the Yalova region in Turkey following the 1999 Düzce earthquake (Daniel *et al.*, 2006). The present methods then allow us to characterize the transient initial forcing, both in duration and in intensity. Finally, the specific triggering caused by a specific earthquake can be extracted from the earthquake time series thanks to such methods: as we did with our synthetic datasets $\mathcal{D}_1$ and $\mathcal{D}_2$, the specific event to be analyzed can be artifically removed from the catalog. Any change in the estimated forcing, as compared with that obtained with the full catalog, can then be attributed to this specific event. This type of analysis can be used to explore the departures of the triggering caused by individual mainshocks from the mean-field laws provided by stochastic triggering models.

## Data and Resources

No data were used in this paper.

## Acknowledgments

We thank Sebastian Hainzl for stimulating discussions about this work, and attempts at testing other methods. Constructive comments by two anonymous reviewers helped improving this paper. This work was supported by the ANR ASEISMIC project and by ECFP7 REAKT Project.

## References

Aoyama, H., M. Takeo, and S. Ide (2002). Evolution mechanisms of an earthquake swarm under the Hida mountains, central Japan, in 1998, *J. Geophys. Res.* **107,** no. B8, 2174, doi: 10.1029/2001JB000540.

Brodsky, E. E. (2006). Long-range triggered earthquakes that continue after the wave train passes, *Geophys. Res. Lett.* **33,** L15313, doi: 10.1029/2006GL026605.

Corral, A. (2004). Universal local versus unified global scaling laws in the statistics of seismicity, *Physica A* **340,** 590–597.

Daniel, G., D. Marsan, and M. Bouchon (2006). Perturbation of the Izmit earthquake aftershock decaying activity following the 1999 $M_w$ 7.2 Düzce, Turkey, earthquake, *J. Geophys. Res.* **111,** B05310, doi: 10.1029/2005JB003978.

Daniel, G., E. Prono, F. Renard, F. Thouvenot, S. Hainzl, D. Marsan, A. Helmstetter, P. Traversa, J. L. Got, L. Jenatton, and R. Guiguet (2011). Changes in effective stress during the 2003–2004 Ubaye seismic swarm, France, *J. Geophys. Res.* **116,** B01309, doi: 10.1029/2010JB007551.

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.* **99,** no. B2, 2601–2618.

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics, *Phil. Trans. R. Soc. Lond.* **222,** 309–368.

Gardner, J. K., and L. Knopoff (1974). Is the sequence of earthquakes in southern California, with aftershocks removed, poissonian? *Bull. Seismol. Soc. Am.* **64,** no. 5, 1363–1367.

Gutenberg, B., and C. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34,** no. 4, 185–188.

Hainzl, S. (2004). Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering, *Geophys. J. Int.* **159,** 1090–1096.

Hainzl, S., and Y. Ogata (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.* **110,** 5. B05S07, doi: 10.1029/2004JB003247.

Hainzl, S., F. Scherbaum, and C. Beauval (2006). Estimating background activity based on interevent-time distribution, *Bull. Seismol. Soc. Am.* **96,** no. 1, 313–320.

Helmstetter, A., and B. Shaw (2009). Afterslip and aftershocks in the rate-and-state friction law, *J. Geophys. Res.* **114,** B01308, doi: 10.1029/2007JB005077.

Helmstetter, A., and D. Sornette (2003). Importance of direct and indirect triggered seismicity in the ETAS model of seismicity, *Geophys. Res. Lett.* **30,** no. 11, doi: 10.1029/2003GL017670.

Jenatton, L., R. Guiguet, F. Thouvenot, and N. Daix (2007). The 16,000-event 2003–2004 earthquake swarm in Ubaye (French Alps), *J. Geophys. Res.* **112,** B11304, doi: 10.1029/2006JB004878.

Kagan, Y., and L. Knopoff (1981). Stochastic synthesis of earthquake catalogs, *J. Geophys. Res.* **86,** no. B4, 2853–2862.

Llenos, A. L., and J. J. McGuire (2011). Detecting aseismic strain transients from seismicity data, *J. Geophys. Res.* **116,** B06305, doi: 10.1029/2010JB007537.

Lohman, R., and J. McGuire (2007). Earthquake swarms driven by aseismic creep in the Salton Trough, California, *J. Geophys. Res.* **112,** B04405, doi: 10.1029/2006JB004596.

Lombardi, A., W. Marzocchi, and J. Selva (2006). Exploring the evolution of a volcanic seismic swarm: The case of the 2000 Izu Islands swarm, *Geophys. Res. Lett.* **33,** L07310, doi: 10.1029/2005GL025157.

Lombardi, A. M., M. Cocco, and W. Marzocchi (2010). On the increase of background seismicity rate during the 1997–1998 Umbria–Marche, central Italy, sequence: Apparent variation or fluid-driven triggering? *Bull. Seismol. Soc. Am.* **100,** 1138–1152.

Matsu'ura, R. S., and I. Karakama (2005). A point-process analysis of the Matsushiro earthquake swarm sequence: The effect of water on earthquake occurrence, *Pure Appl. Geophys.* **162,** 1319–1345.

Molchan, G. (2005). Interevent time distribution in seismicity: A theoretical approach, *Pure Appl. Geophys.* **162,** 1135–1150.

Montgomery-Brown, E. K., P. Segall, and A. Miklius (2009). Kilauea slow slip events: Identification, source inversions, and relation to seismicity, *J. Geophys. Res.* **114,** B00A03, doi: 10.1029/2008JB006074.

Ogata, Y. (1988). Statistical models for earthquake occurences and residual analysis for point processes, *J. Am. Stat. Assoc.* **83,** no. 401, 9–27.

Pacchiani, F., and H. Lyon-Caen (2010). Geometry and spatio-temporal evolution of the 2001 Agios Ioanis earthquake swarm (Corinth rift, Greece), *Geophys. J. Int.* **180,** 59–72.

Perfettini, H., and J.-P. Avouac (2007). Modeling afterslip and aftershocks following the 1992 Landers earthquake, *J. Geophys. Res.* **112,** B07409, doi: 10.1029/2006JB004399.

Reasenberg, P. (1985). Second-order moment of central California seismicity, 1969–1982, *J. Geophys. Res.* **90,** no. B7, 5479–5495.

Sornette, D., S. Utkin, and A. Saichev (2008). Solution of the nonlinear theory and tests of earthquake recurrence times, *Phys. Rev. E* **77,** 066109, doi: 10.1103/PhysRevE.77.066109.

Traversa, P., and J.-R. Grasso (2010). How is volcano seismicity different from tectonic seismicity? *Bull. Seismol. Soc. Am.* **100,** no. 4, 1755–1769.

Vidale, J., and P. Shearer (2006). A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers, *J. Geophys. Res.* **111,** B05312, doi: 10.1029/2005JB004034.

Wolfe, C., B. Brooks, J. Foster, and P. Okubo (2007). Microearthquake streaks and seismicity triggered by slow earthquakes on the mobile south flank of Kilauea volcano, Hawai'i, *Geophys. Res. Lett.* **34,** L23306, doi: 10.1029/2007GL031625.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.* **97,** no. 458, 369–380.

Zollo, A., W. Marzocchi, P. Capuano, A. Lomax, and G. Innaccone (2002). Space and time behavior of seismic activity at Mt. Vesuvius volcano, southern Italy, *Bull. Seismol. Soc. Am.* **92,** no. 2, 625–640.

ISTerre, CNRS
Université de Savoie
Campus Scientifique
73376 Le Bourget du Lac, France
(D.M., E.P.)

ISTerre, CNRS
Université Joseph Fourier
38041 Grenoble, France
(A.H.)

Manuscript received 31 October 2011