Distribution related to all samples and extreme events in the ETAS cluster

Ilaria Spassiani[1], Giuseppe Petrillo[2], Jiancang Zhuang[2]

[1]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Via di Vigna Murata 605, 00143 Rome, Italy
[2]The Institute of Statistical Mathematics, Tokyo

*

Key Points:

- Earthquake forecasting is a challenging goal for statistical seismologists and geophysical modelers
- Epidemic Type Aftershock Sequence (ETAS) model for seismic hazard assessment
- Derivation of the probability for the largest event in any ETAS cluster to occur in (t, x, y, m)

*All the authors contributed equally to this work

Corresponding author: Ilaria Spassiani, ilaria.spassiani@ingv.it

Abstract

Earthquake forecasting is a challenging goal for seismologists and geophysicists due to the complex nature of the earthquake phenomenon. Delivering reliable forecasts is crucial for public safety. This study derives the probability of extreme events in any seismic cluster generated by the Epidemic Type Aftershock Sequence (ETAS) model. This probability is obtained as a function of time, space and magnitude. The results contribute to understanding the distinguishing features between mainshocks and foreshocks and provide insights into earthquake prediction and the probability assessment of extreme events within seismic clusters.

Plain Language Summary

Forecasting extreme events in a seismic sequence is a challenging but crucial goal for statistical seismologists to prevent catastrophic disasters. To date, the most used model in earthquake forecasting applications is the Epidemic-Type-Aftershock-Sequence (ETAS) model. It accounts for the most widely recognized feature of seismicity, that is, events' spatiotemporal clustering. In this work, we explicitly derive the probability for the largest event, in any seismic ETAS cluster, to occur in the time-space-magnitude location (t, x, y, m). To do that, we use some advanced mathematical tools, such as Laplace transforms, Tauberian theorems, Bessel/Struve functions. In the pure-temporal case, depending on the specific stability condition considered (explosion or not of the process), we derive the probability of the largest event occurring in (t, m) as the product between an only-time function and an only-magnitude function. Numerical results show that an ETAS synthetic catalogue fits well the theoretical prediction. When including the spatial component in the analysis, due to the high-complexity of the functions involved, we cannot analytically derive the probability of the largest event to occur in (t, x, y, m); we derive instead its Laplace transform, which is shown to decrease with time and increase with space. A numerical procedure is suggested to obtain the inverse transform.

1 Introduction

Earthquake forecasting represents one of the most important and challenging goal for statistical seismologists and seismic-modeling researchers. The earthquake phenomenon is very complex, as it represents an example of self-organized criticality (Bak & Tang, 1989; Sornette & Sornette, 1989; Jagla et al., 2014; Lippiello et al., 2019; Petrillo et al., 2020; Lippiello et al., 2021), but at the same time it is concretely very often tangible and the need to deliver reliable forecasts is essential for public safety.

Seismic hazard assessment is based on several scaling laws, which have been found to represent a good approximation for earthquake sequences, such as the constitutive laws in the Epidemic Type Aftershock Sequence (ETAS) model, briefly illustrated in Section 2, and representing the most used model for operational forecasting issues (Marzocchi, 2008; Marzocchi & Lombardi, 2009; Jordan et al., 2011; Marzocchi et al., 2012; Jordan et al., 2014; Marzocchi et al., 2014; Page et al., 2016; Marzocchi et al., 2017; Llenos & Michael, 2017; Omi et al., 2018; Spassiani, Falcone, et al., 2023; Spassiani, Yaghmaei-Sabegh, et al., 2023).

Retrospective analyses are mostly focused on aftershock sequences following a large event, named *mainshock*, and the results obtained are used for analyses of perspective type. However, a lot of attention has been paid in the literature also to the largest event in the earthquake sequence: the forecast of extreme events is crucial to prevent disasters causing the death of several people. More precisely, in Saichev and Sornette (2005), Zhuang and Ogata (2006), Vere-Jones and Zhuang (2008) and Luo and Zhuang (2016), the distribution of the largest event in the critical ETAS model is related to the magnitude distribution of *foreshocks*, which are defined as background events that have at

least one offspring, direct or indirect, with a larger magnitude. Although the difference between mainshocks and foreshocks is still not rigorously defined and accepted in the literature (Helmstetter et al., 2002; Helmstetter & Sornette, 2003; Felzer et al., 2004; Mignan, 2014; Lippiello et al., 2020; Petrillo & Lippiello, 2021), the derivation of the probability of foreshocks under a clustering model can be used as a test to find distinguishing features between these kinds of events. The distribution of foreshocks is in fact considered essential in the earthquake prediction (Abercrombie & Mori, 1994; Savage & Rupp, 2000; Merrifield et al., 2004; Marzocchi et al., 2019), as it allows us to evaluate the probability of occurrence of the largest event in a sequence.

In this paper, we extend the results of Saichev and Sornette (2005), Zhuang and Ogata (2006), Vere-Jones and Zhuang (2008) and Luo and Zhuang (2016) by deriving the probability for the largest event, in any ETAS cluster, to occur in time t and have magnitude m. It is related to all the samples and to the extreme events in the sequence, as here we do not consider a fixed initial event's magnitude and we include the temporal component of the process. The same procedure can be adopted to include the spatial component and then, thanks to the separability between space and time in the ETAS rate, to derive the probability for the largest event in any ETAS cluster to occur in (t, x, y, m), (x, y) being the epicenter coordinates, that is, the final expression for that probability in the complete ETAS setting.

2 The ETAS model

The Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988, 1989, 1998) represents a benchmark in statistical seismology and belongs to the class of linear, marked, self-exciting Hawkes processes of branching type: each seismic event represents an element of the process, it is identified with its space-time location and it is marked by the magnitude. According to the branching framework, all the events may give birth to their own progeny independently of the others, resulting in a cluster-type structure. The clusters' elements represent the aftershocks component of the process, associated with the stress perturbations due to previous shocks and modeled as a space-time non-homogeneous Poisson process. Instead, the initial events constitute the background component of the process, not triggered by precursory shocks and modeled as a time-stationary Poisson process, non-homogeneous in space. The conditional intensity of the space-time-magnitude ETAS model, completely characterizing the process, is

$$\begin{aligned} \lambda(t,x,y,m|\mathcal{H}_t) &= \lambda(t,x,y|\mathcal{H}_t)\, s(m), \\ \text{with} \quad \lambda(t,x,y|\mathcal{H}_t) &= \mu(x,y) + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i) f(x-x_i, y-y_i), \end{aligned} \tag{1}$$

where $\mathcal{H}_t$ is the past history of the event (t, x, y, m) and:

1. $\mu(x, y)$ is the time-stationary background rate;
2. $\kappa(m) = A \exp\{\alpha(m - m_c)\}$, $m \geq m_c$, is the productivity law, representing the fertility of the event with magnitude m. The magnitude m_c is the completeness threshold, that is the value such that all the events with a higher magnitude are surely recorded in the seismic catalog;
3. $g(t) = \frac{p-1}{c}\left(1 + \frac{t}{c}\right)^{-p}$, $t > 0$, is the Omori-Utsu law for the aftershocks' decay;
4. $f(x, y)$ is the probability density function (PDF) for the location of the event, typically of power law type, that is $f(x,y) = \frac{q-1}{\pi D e^{\gamma(m-m_c)}}\left[1 + \frac{x^2+y^2}{D e^{\gamma(m-m_c)}}\right]^{-q}$;
5. $s(m) = \beta \exp\{-\beta(m-m_c)\}$, $m \geq m_c$, is the decreasing exponential Gutenberg-Richter (GR) law for the magnitudes of all the events in the process.

All the parameters in the ETAS conditional intensity are positive and typically estimated through maximum likelihood techniques, where the log-likelihood function is easily derived as

$$\ln L = \sum_i \lambda(t_i, x_i, y_i, m_i|\mathcal{H}_{t_i}) - \int\int\int_{\mathcal{T}\times\mathcal{S}} \lambda(t, x, y|\mathcal{H}_t)dtdxdy + \sum_i \ln s(m_i).$$

A seismic sequence modeled by the space-time-magnitude ETAS is non-explosive under the conditions $p > 1$, $\beta > \alpha$ and $\bar{n} = \frac{A\beta}{\beta-\alpha} > 1$, where the latter is the critical parameter of the process; in the specific case of the ETAS model, it coincides also to the branching ratio: it is the average number of first generation aftershocks per mother event, but can be defined also as the proportion of triggered events with respect to all the shocks (Helmstetter & Sornette, 2002; Zhuang et al., 2012).

3 Theoretical derivation of the probability for the extreme event (t, m)

In Saichev and Sornette (2005), Zhuang and Ogata (2006), Vere-Jones and Zhuang (2008) and Luo and Zhuang (2016), the authors consider the probability $\zeta(m, m')$ for an event with magnitude m to have no offspring larger than a given magnitude m'. It is obtained as

$$\begin{aligned} \zeta(m, m') &= \exp\{-\kappa(m)F(m')\}, \\ \text{where} \quad F(m') &= 1 - \int_{m_c}^{m'} s(m^*)\exp\{-\kappa(m^*)F(m')\}\, dm^* \end{aligned}$$

represents the probability for the largest earthquake in an arbitrary cluster, including the initial event and all its descendants, to be greater than m'. In the above formulations, $\kappa(\cdot)$ and $s(\cdot)$ are respectively the productivity law and the GR law introduced in the previous section.

Here, we want to integrate over the magnitude of the mother event, and we want to add the temporal component of the ETAS rate. We then focus on the probability for the largest event to have magnitude m and to occur in time t or, briefly, to be the (t, m) event:

$$\eta(t, m) = \mathbb{P}\{\text{The largest event in any cluster is the } (t, m) \text{ event}\}.$$

The probability $\eta(t, m)$ can be obtained as the union of two disjoint conditions:

$I_1 =$ "the largest event is the mother, occurring in $t = 0$, and none of its children has magnitude greater than m; that is, the largest event is $(0, m)$ and any other $(\bar{m}, t > 0)$ event is such that $\bar{m} < m$";

$I_2 =$ "the largest event (t, m) belongs to one subcluster, that is, $t > 0$ and neither the mother nor any other event in all the other subclusters have magnitude greater than m".

These two events represent a partition of $\Omega = \{$The largest event in any cluster is the (t, m) event$\}$, therefore $\eta(t, m)$ is obtained as the sum of their probabilities. In other words, it holds

$$\eta(t, m) = \mathbb{P}\{I_1\} + \mathbb{P}\{I_2\}. \tag{2}$$

To derive the two probabilities $\mathbb{P}\{I_1\}$ and $\mathbb{P}\{I_2\}$ of above, we have to compute the probability $R(m')$ for the largest event in a cluster to be less than m', which is obviously $R(m') = 1 - F(m')$, in fact:

$$
\begin{aligned}
R(m') &= \mathbb{P}\{\text{The largest event in a cluster is less than } m'\},\\
&= \int_0^{m'} \mathbb{P}\{\text{The largest event in the cluster generated by initial event with mag}\\
&\qquad m^* \text{ is less than } m' \,|\, \text{The initial event in the cluster has mag } m^*\}\\
&\qquad \cdot \mathbb{P}\{\text{The initial event in the cluster has mag } m^*\}\, dm^*\\
&= \int_0^{m'} \sum_{n=0}^{\infty} [R(m')]^n \, \frac{[\kappa(m^*)]^n \, e^{-\kappa(m^*)}}{n!} \, s(m^*)\, dm^*\\
&= \int_0^{m'} e^{-\kappa(m^*)[1-R(m')]} \, s(m^*)\, dm^*, \qquad (3)
\end{aligned}
$$

where we have set $m_c = 0$ to shorten notations (hereafter), and we summed over all the subclusters of the initial event's n children.

Now, recalling that the number of events is Poisson distributed, we derive the first term in the Right-Hand-Side (RHS) of (2) as follows:

$$
\mathbb{P}\{I_1\} = \delta(t)\, s(m) \sum_{n=0}^{\infty} [R(m)]^n \, \frac{[\kappa(m)]^n}{n!} \, e^{-\kappa(m)} = \delta(t)\, s(m)\, e^{-\kappa(m)[1-R(m)]}, \qquad (4)
$$

where the term $\delta(t)\, s(m)$ says that the initial event occurred in $t = 0$ and is randomly selected from a GR distribution. Instead, as regards the second term in the RHS of (2), we have to randomly select one of the subclusters of the initial event's n children, which is assumed to contain the largest event, and to impose that both the mother and all the other shocks have a lower magnitude:

$$
\begin{aligned}
\mathbb{P}\{I_2\} &= \int_0^m s(m^*) \left[\sum_{n=1}^{\infty} \frac{[\kappa(m^*)]^n}{n!} \, e^{-\kappa(m^*)} \binom{n}{1} \int_0^t \eta(t-t_i, m) \, [R(m)]^{n-1} g(t_i)\, dt_i \right] dm^*\\
&= \int_0^m s(m^*) \left[\sum_{n=1}^{\infty} \frac{\kappa(m^*)\, [\kappa(m^*)]^{n-1}}{(n-1)!} \, e^{-\kappa(m^*)} [R(m)]^{n-1} \int_0^t \eta(t-t_i, m)\, g(t_i)\, dt_i \right] dm^*\\
&= \int_0^m s(m^*)\kappa(m^*) \sum_{n=0}^{\infty} \frac{[\kappa(m^*)]^n}{n!} \, e^{-\kappa(m^*)} \, [R(m)]^n \int_0^t \eta(t-t_i, m)\, g(t_i)\, dt_i \, dm^*\\
&= \int_0^m s(m^*)\, \kappa(m^*)\, e^{-\kappa(m^*)[1-R(m)]} \, dm^* \int_0^t \eta(t-t_i, m)\, g(t_i)\, dt_i\\
&= \int_0^m s(m^*)\, \kappa(m^*)\, e^{-\kappa(m^*)[1-R(m)]} \, dm^* \, \cdot \big(\eta \circ g\big)\,(t, m), \qquad (5)
\end{aligned}
$$

where we used the symbol $\circ$ for the convolution. Then, setting

$$
\begin{aligned}
A(m) &= s(m)\, e^{-\kappa(m)[1-R(m)]},\\
B(m) &= \int_0^m s(m^*)\, \kappa(m^*)\, e^{-\kappa(m^*)[1-R(m)]} \, dm^*. \qquad (6)
\end{aligned}
$$

We can rewrite the probability $\eta(t, m)$ in (2) as a recursive equation:

$$
\begin{aligned}
\eta(t,m) &= \delta(t)\,A(m) + B(m)\left(\eta \circ g\right)(t,m) \\
&= \delta(t)\,A(m) + B(m)\left\{\left[\delta(t)\,A(m) + B(m)\left(\eta \circ g\right)\right] \circ g\right\}(t,m) \\
&= \delta(t)\,A(m) + A(m)\,B(m)\,g(t) + B^2(m)\left(\eta \circ g^{(2)}\right)(t,m) \\
&= \delta(t)\,A(m) + A(m)\,B(m)\,g(t) + A(m)\,B^2(m)\,g^{(2)}(t) + B^3(m)\left(\eta \circ g^{(3)}\right)(t,m) + \dots \\
&= A(m)\sum_{i=0}^{\infty} B^i(m)\,g^{(i)}(t), \qquad (7)
\end{aligned}
$$

where $g^{(0)}(t) = \delta(t)$, and we used that the Delta function is the identity for convolution power; furthermore, $g^{(i)}(\cdot)$ indicates the i^{th} convolution of $g(\cdot)$ with itself.

By the Leibniz integral rule, one easily obtains a relationship between $\dfrac{d\,R(m)}{d\,m} = R'(m)$, $A(m)$ and $B(m)$, that is

$$
R'(m) \;=\; A(m) + B(m)\,R'(m) \quad \Leftrightarrow \quad R'(m) \;= \frac{A(m)}{1 - B(m)},
$$

which is a first order non-homogeneous differential equation that can be solved with the standard techniques of differential calculus.

To the aim of solving the implicit equation in $\eta(t,m)$, we stop at the first equality in (7), and we apply the Laplace transform with respect to t to both the terms (we use the hat sign to indicate the transformed functions):

$$
\hat{\eta}(s,m) = A(m) + B(m)\,\hat{\eta}(s,m)\,\hat{g}(s) \quad \Leftrightarrow \quad \hat{\eta}(s,m) = \frac{A(m)}{1 - B(m)\hat{g}(s)}, \qquad (8)
$$

where

$$
\hat{g}(s) = (p-1)\,(sc)^{p-1} e^{sc}\,\Gamma(1-p, sc) \qquad (9)
$$

is the Laplace transform of the Omori-Utsu function introduced in Section 2, and $\Gamma(u,v) = \int_v^{\infty} x^{u-1}e^{-x}dx$ is the upper incomplete Gamma function (Bateman, 1953; Temme, 1996). In Fig.(1) we plot the function $\hat{g}(s)$ defined above for different values of its parameters p, c. It shows quasi-constancy for small values of the variable, turning into a decreasing trend as s becomes larger. The rate of decrease does not change with p, whose variability only induces a vertical shift: the smaller p, the larger $\hat{g}(s)$; the increase of the parameter c induces instead a faster decrease of the function as s gets larger.

Since the inverse Laplace transform of the left hand side in (8) is difficult to compute analytically, we follow the same approach as in Molchan (2005) and we make use of the Tauberian theorems (Feller, 1971) to obtain an approximation of the probability $\eta(t,m)$ for the largest event in any time-magnitude ETAS cluster to be the (t,m) event.

More precisely, given

$$
\hat{\eta}(s,m) = \int_0^{\infty} e^{-st}\eta(t,m)\,dt = \int_0^{\infty} e^{-st}\,H(dt,m),
$$

where $H(\cdot,m)$ is the cumulative distribution function (c.d.f.) corresponding to $\eta(\cdot,m)$, we observe that for every $y > 0$ and for $p > 1$ it holds

$$\begin{aligned}
\lim_{s\to 0}\frac{\hat{\eta}(sy,m)}{\hat{\eta}(s,m)} &= \lim_{s\to 0}\frac{1-B(m)\hat{g}(s)}{1-B(m)\hat{g}(sy)}\\
&= \lim_{s\to 0}\frac{1-B(m)(p-1)\,(sc)^{p-1}e^{sc}\Gamma(1-p,sc)}{1-B(m)(p-1)\,(scy)^{p-1}e^{scy}\Gamma(1-p,scy)}\\
&= \lim_{s\to 0}\frac{1-B(m)(p-1)\,e^{sc}\dfrac{1}{p-1}}{1-B(m)(p-1)\,e^{scy}\dfrac{1}{p-1}}\\
&= \frac{1-B(m)}{1-B(m)}=1,
\end{aligned}$$

where we used that $\Gamma(z,x)\sim\Gamma(z)-x^z/z$ for $x\to 0^+$ and $\Re(z)<0$ if $z\neq -1$. In fact, in the case of $\Re(z)<0$, from Bender and Orszag (1999) we can write

$$\begin{aligned}
\Gamma(z,x) &= \int_x^{+\infty} t^{z-1}\left(1-t+\cdots+(-1)^N\frac{t^N}{N!}\right)dt-\int_0^x t^{z-1}\left(\exp(-t)-\sum_{n=0}^{N}\frac{(-t)^N}{n!}\right)dt\\
&\quad+\int_0^{+\infty} t^{z-1}\left(\exp(-t)-\sum_{n=0}^{N}\frac{(-t)^N}{n!}\right)dt,
\end{aligned}$$

where N is the largest integer less than $-z$. Solving the first two integrals term by term and integrating the last one by part, we obtain

$$\Gamma(z,x)=\Gamma(z)-\sum_{n=0}^{+\infty}(-1)^N\frac{x^{z+n}}{n!(z+n)}.$$

Then, we can apply the first Tauberian theorem (Feller, 1971), which states as follows.

Tauberian Theorem 1

Let U be a measure (c.d.f.) and $\hat{u}$ its Laplace transform, i.e., $\hat{u}(s)=\int_0^\infty e^{-st}\,U(dt)$. For $y>0$, $\varrho\geq 0$ and with $s\cdot t=1$, it holds:

$$\lim_{s\to 0}\frac{\hat{u}(sy)}{\hat{u}(s)}=\left(\frac{1}{y}\right)^{\varrho}\quad\Leftrightarrow\quad\lim_{t\to\infty}\frac{U(tx)}{U(t)}=x^{\varrho}\quad\Leftrightarrow\quad\hat{u}(s)\sim U\left(\frac{1}{s}\right)\Gamma(\varrho+1).$$

In our case, we have to set $\varrho=0$, obtaining that the approximated solution $H(t,m)$ for the implicit equation (8) verifies

$$\begin{aligned}
H(t,m) &\sim \hat{\eta}\left(\frac{1}{t},m\right)\\
&= \frac{A(m)}{1-B(m)\hat{g}\left(\frac{1}{t}\right)}\\
&= \frac{s(m)\,e^{-\kappa(m)[1-R(m)]}}{1-\left[\int_0^m s(m^*)\,\kappa(m^*)\,e^{-\kappa(m^*)[1-R(m)]}\,dm^*\right](p-1)\,\left(\frac{c}{t}\right)^{p-1}e^{\frac{c}{t}}\Gamma(1-p,\frac{c}{t})},
\end{aligned}\tag{10}$$

where we used (6) and (9).

To obtain $\eta(t,m)$, we calculate the derivative of H with respect to t, i.e.,

$$\eta(t,m) = \frac{\partial H(t,m)}{\partial t} \sim \frac{d\hat{g}(1/t)}{dt} \frac{A(m)\,B(m)}{\left(1 - B(m)\,\hat{g}\left(\frac{1}{t}\right)\right)^2}. \tag{11}$$

It is straightforward to show that $\eta(t,m)$ is asymptotically separable. In fact, using that $\Gamma(z,x) \sim \Gamma(z) - x^z/z$ for $x \to 0$ and $\Re(z) < 0$ if $z \neq -1$, we obtain

$$\hat{g}\left(0^+\right) = (p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\Gamma(1-p, c/t) \sim (p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\left(\Gamma(1-p) - \frac{\left(\frac{c}{t}\right)^{1-p}}{1-p}\right) = 1.$$

Therefore, for $t \to \infty$, we have

$$\eta(t,m) \sim \frac{d\hat{g}(1/t)}{dt}\,\frac{A(m)\,B(m)}{\left(1 - B(m)\right)^2} = \eta(t)\;h(m). \tag{12}$$

As regards the temporal function, we get

$$\begin{aligned}
\frac{d\hat{g}\left(\frac{1}{t}\right)}{dt} &= \hat{g}\left(\frac{1}{t}\right)\left(-\frac{p-1}{t} - \frac{c}{t^2} + \frac{\Gamma'(1-p,c/t)}{\Gamma(1-p,c/t)}\right) \\
&= (p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\,\Gamma(1-p,c/t)\left(\frac{1-p}{t} - \frac{c}{t^2}\right) + (p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\,\Gamma'(1-p,c/t) \\
&= (p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\left(\Gamma(1-p) - \frac{(c/t)^{p-1}}{1-p}\right)\left(\frac{1-p}{t} - \frac{c}{t^2}\right) + \frac{p-1}{t} \\
&= \left((p-1)\left(\frac{c}{t}\right)^{p-1} e^{c/t}\,\Gamma(1-p) + e^{c/t}\right)\left(\frac{1-p}{t} - \frac{c}{t^2}\right) + \frac{p-1}{t} \\
&= -\frac{(p-1)^2 c^{p-1}\Gamma(1-p)}{t^p} \sim t^{-p}.
\end{aligned} \tag{13}$$

It is very interesting to observe that the asymptote of $\eta(t)$ depends only on $g(t)$ and that $\kappa(m)$ and $s(m)$ have no effect on it.

To obtain instead the final expression for $h(m)$, the productivity and GR functions introduced in Section 2 can also be substituted, as well as the explicit expressions of $F(m) = 1 - R(m)$, that can be derived from Luo and Zhuang (2016) and Vere-Jones and Zhuang (2008), depending on the specific case considered. Precisely, we obtain the following.

***Subcritical case*, $\mu < 1$**

From Vere-Jones and Zhuang (2008), $F(m) = \dfrac{1}{1-\mu} e^{-\beta m}[1 + o(1)]$ for $m \to \infty$, then:

$$h(m) \sim \frac{\beta\, e^{-\beta m}\, e^{-Ae^{\alpha m}\frac{1}{1-\mu}e^{-\beta m}} \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\frac{1}{1-\mu}e^{-\beta m}\right\}}\, dm^*}{\left(1 - \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\frac{1}{1-\mu}e^{-\beta m}\right\}}\, dm^*\right)^2}. \tag{14}$$

***Critical case*, $\mu = 1$, $\alpha < \dfrac{\beta}{2}$**

From Vere-Jones and Zhuang (2008), $F(m) = \left(1-\frac{\alpha}{\beta}\right)^{-1}\sqrt{2\left(1-2\frac{\alpha}{\beta}\right)}e^{-\frac{\beta m}{2}}[1+o(1)]$ for $m \to \infty$, then:

$$h(m) \sim \frac{\beta\, e^{-\beta m}\, e^{-Ae^{\alpha m}\left\{\left(1-\frac{\alpha}{\beta}\right)^{-1}\sqrt{2\left(1-2\frac{\alpha}{\beta}\right)}e^{-\frac{\beta m}{2}}\right\}} \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\left(1-\frac{\alpha}{\beta}\right)^{-1}\sqrt{2\left(1-2\frac{\alpha}{\beta}\right)}e^{-\frac{\beta m}{2}}\right\}} dm^*}{\left[1-\int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\left(1-\frac{\alpha}{\beta}\right)^{-1}\sqrt{2\left(1-2\frac{\alpha}{\beta}\right)}e^{-\frac{\beta m}{2}}\right\}} dm^*\right]^2}. \tag{15}$$

***Critical case*, $\mu = 1$, $\alpha = \frac{\beta}{2}$**

From Luo and Zhuang (2016), $F(m) = \frac{2\sqrt{2}}{\sqrt{\beta m}}e^{-\alpha m}[1+o(1)]$ for $m \to \infty$, then:

$$h(m) \sim \frac{\beta\, e^{-\beta m}\, e^{-Ae^{\alpha m}\left\{\frac{2\sqrt{2}}{\sqrt{\beta m}}e^{-\alpha m}\right\}} \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\frac{2\sqrt{2}}{\sqrt{\beta m}}e^{-\alpha m}\right\}} dm^*}{\left[1-\int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\frac{2\sqrt{2}}{\sqrt{\beta m}}e^{-\alpha m}\right\}} dm^*\right]^2}. \tag{16}$$

***Critical case*, $\mu = 1$, $\beta > \alpha > \frac{\beta}{2}$**

From Luo and Zhuang (2016), $F(m) = \xi e^{-\alpha m}[1+o(1)]$ for $m \to \infty$, where ξ is the constant to which the function $F(m)$ converges; then:

$$h(m) \sim \frac{\beta\, e^{-\beta m}\, e^{-Ae^{\alpha m}\left\{\xi e^{-\alpha m}\right\}} \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\left\{\xi e^{-\alpha m}\right\}} dm^*}{\left[1-\int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}\{\xi e^{-\alpha m}\}} dm^*\right]^2}. \tag{17}$$

***Supercritical case*, $\mu > 1$**

From Zhuang and Ogata (2006), $F(m) = Const.\, C > 0$, for $m \to \infty$, then:

$$h(m) \sim \frac{\beta\, e^{-\beta m}\, e^{-Ae^{\alpha m}C} \int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}C}\, dm^*}{\left[1-\int_0^m A\beta\, e^{-(\beta-\alpha)m^*}\, e^{-Ae^{\alpha m^*}C}\, dm^*\right]^2}. \tag{18}$$

4 Numerical results

It follows from Eq.(12) that $\eta(t,m)$ is asymptotically separable, since it can be written as the product $\eta(t)\, h(m)$.

The function $h(m)$ represents the probability to observe, as largest event, the magnitude m event at any time in any cluster, and in Fig.(2) we show the trend of the function $h(m)$ as the criticality varies: from the subcritical to supercritical regime. We show both the analytical and the numerical results. The implementation of the ETAS model is carried out as follows.

1. The first step is to choose the background seismicity. This represents the zero-th order generation in a self-exiting point process and a certain number n_0 of events is created.

2. Each of the background element generates a certain number of offspring. The number n_1 and the occurrence space-time position of the offspring depends of the functional form implemented in $\lambda(t, x, y)$. This is the first order generation of events.
3. The previous step is repeated considering $n_j = n_{j-1}$ and is iterated until we have $n_{j^*} = 0$, for a certain j^*.

We observe that the simulated results fit very well the theoretical prediction.

The temporal decay from Eq.(13) is plotted in Fig.(3) together with the ETAS numerical simulations. Also in this case we confirm that the asymptotic theoretical prediction fits very well the results from the ETAS synthetic catalogue.

5 Inclusion of the spatial rate in the extreme event distribution

In this section we will include the spatial component in the ETAS rate to derive the probability for the largest event in any cluster to have magnitude m, occur in t and nucleate in the spatial location (x, y), that is,

$$\eta(t, x, y, m) = \mathbb{P}\{\text{The largest event in any cluster is the } (t, x, y, m) \text{ event}\}.$$

Since the ETAS rate is separable in the space-time-magnitude domain, we can repeat exactly the same reasoning as in Section 3 to obtain an implicit, recursive expression for $\eta(t, x, y, m)$:

$$\begin{aligned} \eta(t, x, ym) &= \delta(t, x, y)\, A(m) + B(m) \left(\eta \circ (g \cdot f)\right)(t, x, y, m) \\ &= \dots \\ &= A(m) \sum_{i=0}^{\infty} B^i(m)\, (g \cdot f)^{(i)}(t, x, y), \end{aligned} \tag{19}$$

where $(g \cdot f)^{(0)}(t, x, y) = \delta(t, x, y)$ and $\cdot$ is the usual multiplicative sign. To solve the above equation, we will stop at the first equality in (19), apply the Laplace transform and make use of the Tauberian Theorem 1. Recalling again that the temporal and spatial components in the ETAS rate are separated, the solution of (19) is simply

$$\hat{\eta}(s, u, v, m) = A(m) + B(m)\, \hat{\eta}(s, u, v, m)\, \hat{g}(s) \hat{f}(u, v),$$

that is,

$$\hat{\eta}(s, u, v, m) = \frac{A(m)}{1 - B(m) \hat{g}(s) \hat{f}(u, v)}, \tag{20}$$

where $\hat{g}(s)$ is the Omori-Utsu Laplace transform of equation (9) and we used (8). We need then to explicitly derive the two-dimensional Laplace transform of the spatial component $f(x, y)$ in the ETAS rate, that is (Voelker, 2013)

$$\hat{f}(u, v) = \int_0^{\infty} \int_0^{\infty} e^{-ux - vy} f(x, y)\, dxdy.$$

We implement here the formulation typically adopted in the literature for the spatial component (Zhuang, 2011; Zhuang et al., 2019):

$$f(x, y) = \frac{q - 1}{\pi D e^{\gamma(m - m_c)}} \left[1 + \frac{x^2 + y^2}{D e^{\gamma(m - m_c)}}\right]^{-q}. \tag{21}$$

In order to simplify the calculus and to interpreter better the results, we set $x^2+y^2=z^2$, $De^{\gamma(m-m_c)}=\delta_1$ and $\frac{q-1}{\pi De^{\gamma(m-m_c)}}=\delta_2$. Eq.(21) then reads

$$f(z)=\delta_2\left[1+\frac{z^2}{\delta_1}\right]^{-q}. \tag{22}$$

Recalling that the Laplace transform of a function $f(z)$ is defined as $\hat{f}(w)=\int_0^\infty e^{-wz}f(z)dz$, by performing the substitution $x=z/\sqrt{\delta_1}$ we can compute the Laplace transform of $f(z)$ in Eq.(22) as

$$\begin{aligned}\hat{f}(w) &= \delta_2\sqrt{\delta_1}\int_0^\infty e^{-w\sqrt{\delta_1}x}\left[1+x^2\right]^{-q}dx \\ &= \frac{1}{\Gamma(q)}\pi^{3/2}\delta_2 2^{-q-\frac{1}{2}}\left(\frac{1}{\delta_1}\right)^{-\frac{q}{2}-\frac{1}{4}}w^{q-\frac{1}{2}} \\ &\quad \left(\csc(\pi q)\boldsymbol{H}_{\frac{1}{2}-q}\left(\sqrt{\delta_1}w\right)+2\csc(2\pi q)J_{q-\frac{1}{2}}\left(\sqrt{\delta_1}w\right)-\sec(\pi q)J_{\frac{1}{2}-q}\left(\sqrt{\delta_1}w\right)\right), \end{aligned}\tag{23}$$

where $J_n(\zeta)$ is the Bessel function of the first type and $\boldsymbol{H}_n(\zeta)$ is the Struve function, which is the solution of a given non-homogeneous Bessel's differential equation (Abramowitz & Stegun, 1972). In Fig.(4) we plot the trend of $\hat{f}(w)$ for different values of its parameters D,γ,q: the function is shown to decrease with its variable w, and to assume smaller values for smaller q and larger γ and D, with the largest (smallest) differences observed when varying D (q).

Consequently, we can rewrite the Eq.(20) as

$$\hat{\eta}(s,w,m)=\frac{A(m)}{1-B(m)\hat{g}(s)\hat{f}(w)}, \tag{24}$$

where $\hat{g}(s)$ and $\hat{f}(w)$ are respectively defined in Eqs. (9) and (23). Now, in order to find the solution $\eta(t,z,m)$ from Eq.(24), we state and prove the following Tauberian theorem specific for functions in two-variables (see also Diamond (1987) and Bender and Orszag (2013)).

Tauberian Theorem 2

Let $U(x,y)$ be a monotone non-decreasing function and $\hat{u}$ its Laplace transform, i.e., $\hat{u}=\int_0^\infty\int_0^\infty e^{-xp-yq}dU(x,y)$. If, for $\alpha,\beta,a,b>0$, it holds

$$\lim_{(p,q)\to(0,0)}\frac{\hat{u}(\alpha p,\beta q)}{\hat{u}(p,q)}=\alpha^{-a}\beta^{-b},$$

then:

$$U\left(\frac{\xi}{p},\frac{\eta}{q}\right)=\xi^a\eta^b\frac{\hat{u}(p,q)}{\Gamma(a+1)\Gamma(b+1)}.$$

Proof. Given the function $G(x,y)=\dfrac{x^a y^b}{\Gamma(a+1)\Gamma(b+1)}$, it holds

$$\alpha^{-a}\beta^{-b}=\int_0^\infty\int_0^\infty e^{-x\alpha-y\beta}dG(x,y).$$

It then follows that

$$\begin{aligned}\frac{\hat{u}(\alpha p,\beta q)}{\hat{u}(p,q)} &= \frac{\int_0^\infty\int_0^\infty e^{-\alpha p x-\beta q y}dU(x,y)}{\hat{u}(p,q)}\\ &= \frac{\int_0^\infty\int_0^\infty e^{-\alpha\xi-\beta\eta}dU(\xi/p,\eta/q)}{\hat{u}(p,q)}\\ &= \int_0^\infty\int_0^\infty e^{-\alpha\xi-\beta\eta}d\bar{U}(\xi,\eta),\end{aligned}$$

where in the second equality we imposed $x=\xi/p$ and $y=\eta/q$, while in the third $\bar{U}(\xi,\eta)=\frac{U(\xi/p,\eta/q)}{\hat{u}(p,q)}$. The last equality is the Laplace transform of the function $\bar{U}(\xi,\eta)$. Since for $(p,q)\to(0,0)$ it holds $\dfrac{\hat{u}(\alpha p,\beta q)}{\hat{u}(p,q)}\to\alpha^{-a}\beta^{-b}$ by hypothesis, we can deduce that

$$\bar{U}(\xi,\eta)\to G(\xi,\eta)\quad\forall(\xi,\eta),$$

that is

$$U\left(\frac{\xi}{p},\frac{\eta}{q}\right)\sim\xi^a\eta^b\frac{\hat{u}(p,q)}{\Gamma(a+1)\Gamma(b+1)}.$$

□

In our case one can show that the Tauberian Theorem 2 is satisfied, in fact, from Eq.(24), it holds

$$\lim_{(s,w)\to(0,0)}\frac{\hat{\eta}(\alpha s,\beta w,m)}{\hat{\eta}(s,w,m)}=1. \tag{25}$$

Therefore, setting $a,b=0$ and $\xi,\eta=1$, we can obtain the approximated solution

$$U(t,z,m)\sim\hat{\eta}\left(\frac{1}{t},\frac{1}{z},m\right)=\frac{A(m)}{1-B(m)\hat{g}\left(\dfrac{1}{t}\right)\hat{f}\left(\dfrac{1}{z}\right)}, \tag{26}$$

where $\hat{g}$ and $\hat{f}$ are defined in Eqs. (9) and (23), respectively. The graphical representation of the product $\hat{g}\left(\dfrac{1}{t}\right)\hat{f}\left(\dfrac{1}{z}\right)$ is given in Fig.(5) with respect to the temporal variable $\dfrac{1}{t}$ and fixing five different values of the spatial variable $\dfrac{1}{z}$ in the left panel, vice-versa in the right panel. Note that the x-axes contain the "Laplace variables" $\left(\dfrac{1}{t},\dfrac{1}{z}\right)=(s,w)$. By looking at the left panel, we can see that the product of the two Laplace transforms $\hat{g}\left(\dfrac{1}{t}\right)\hat{f}\left(\dfrac{1}{z}\right)$ increases with time t, getting quasi-constant upward vertical shifts as we use lower values for the space variable z. Instead, the right panel shows that the product decreases with this latter variable, but the upward vertical shifts observed when considering higher t become progressively smaller, until blurring the lines for $t>50$ (equivalently, $\dfrac{1}{t}<0.02$). Finally, we stress that the final expression of $U(t,z,m)$ in Eq. (26) will rely on the specific choice of the function $F(m)=1-R(m)$, which is needed to obtain $A(m)$ and $B(m)$ as given in Eq. (6), depending on the stability cases discussed in Section 3. In general, for typical values of m,m_c,α,β we expect $A(m),B(m)>0$, but with a negative denominator in Eq. (26). Therefore, the trend of the function $U(t,z,m)$

is that shown in Fig.(5), but with inverted monotonicity, plus a vertical shift corresponding to the constant obtained from the explicit values of $A(m)$ and $B(m)$.

Since $U(t, z, m)$ is the joint CDF of t and z, the joint PDF η can be obtained through

$$\eta(t, z, m) = \frac{\partial^2}{\partial t \partial z} U(t, z, m). \tag{27}$$

However, due to the high complexity of the functions involved, the analytical explicit expression is hard to be obtained, thus making the numerical procedure the fast track. We expect it will certainly depend again on the specific stability case considered.

6 Conclusions

Earthquake forecasting is a challenging task that plays a crucial role in both social safety and research science. The self-organized criticality of the earthquake phenomenon piles on the complexity of the problem. Seismic hazard assessment, based on the Epidemic Type Aftershock Sequence (ETAS) model, has been instrumental in providing reliable forecasts for operational purposes. While retrospective analyses have mainly focused on aftershock sequences, attention has also been given to the largest event in the sequence, as predicting extreme events is vital for preventing catastrophic disasters. The distribution of foreshocks, which are background events with offspring of larger magnitudes, is considered essential in earthquake prediction and assessing the probability of occurrence of the largest event in a sequence.

This paper extends previous research by deriving the probability for the largest event within any ETAS cluster to occur at a specific space, time and magnitude, considering both the temporal and spatial components of the process. By examining the distinguishing features between mainshocks and foreshocks, these findings contribute to a better understanding of earthquake prediction and provide valuable insights into assessing the probability of extreme events within seismic clusters.

Open Research

All data and software used in this paper will be available by the time of publication in a specific repository.

Acknowledgments

G.P and J.Z. would like to thank the MEXT Project for Seismology TowArd Research innovation with Data of Earthquake (STAR-E Project), Grant Number: JPJ010217. I.S. would like to thank the Seismic Hazard Center - WP3 Short-Term Probabilistic Seismic Hazard (CPS, INGV) and the NEar real-tiME results of Physical and StatIstical Seismology for earthquakes observations, modeling and forecasting (NEMESIS) Project (INGV).

References

Abercrombie, R., & Mori, J. (1994). Local observations of the onset of a large earthquake: 28 June 1992 Landers, California. *Bull. Seismol. Soc. Am.*, *84*(3), 725-734. doi: https://doi.org/10.1785/BSSA0840030725

Abramowitz, M., & Stegun, I. A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover Publ., Inc.

Bak, P., & Tang, C. (1989). Earthquakes as a self-organized critical phenomenon. *J. Geophys. Res. Solid Earth*, *94*(B11), 15635-15637. doi: https://doi.org/10.1029/JB094iB11p15635

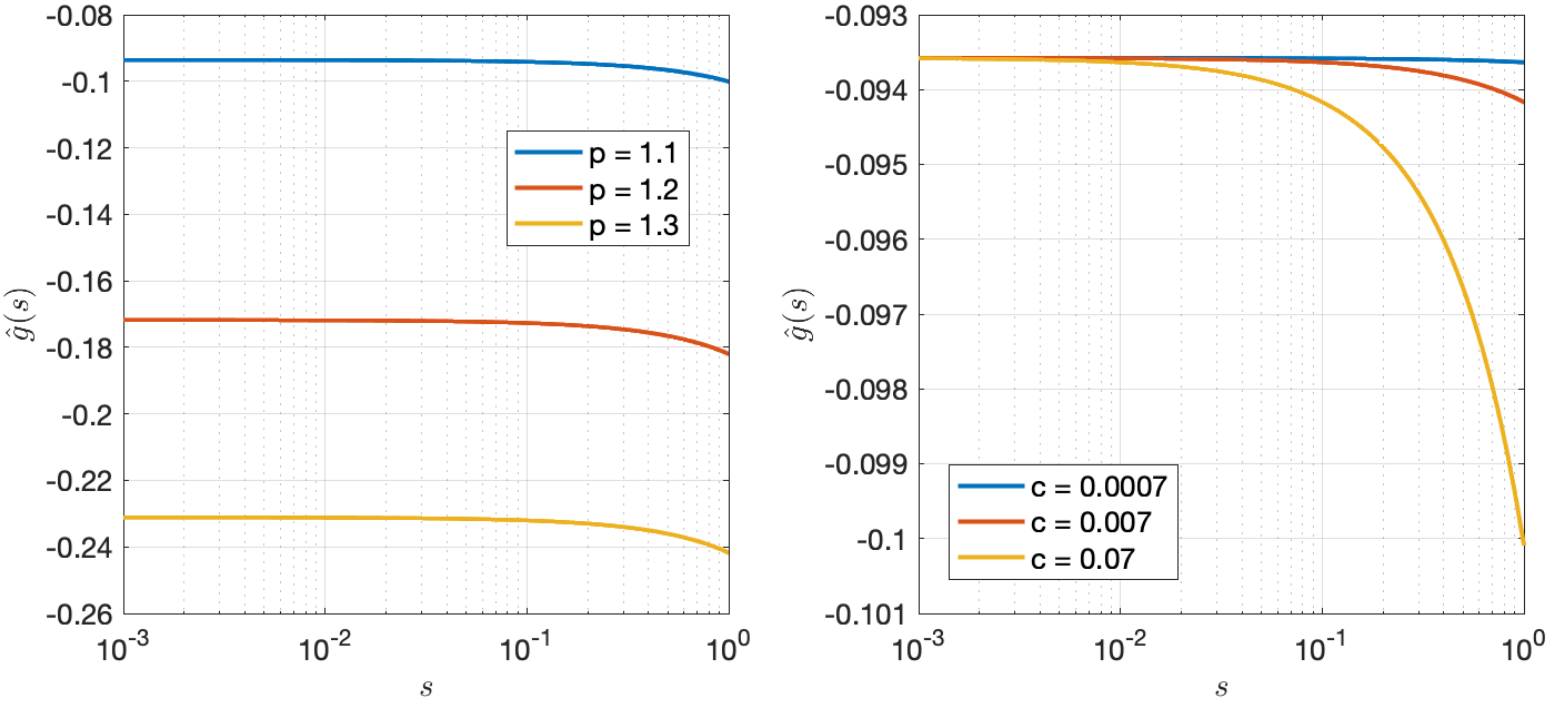

Figure 1. The time function $\hat{g}(s)$, given in Eq. (9), varying the parameter p keeping fixed $c = 0.07$ (left), and varying the parameter c keeping fixed $p = 1.1$ (right).

Bateman, H. (1953). *Higher Transcendental Functions [Volumes I-III].* McGraw-Hill Book Company.

Bender, C., & Orszag, S. (1999). *Advanced Mathematical Methods for Scientists and Engineers I - Asymptotic Methods and Perturbation Theory.* Springer.

Bender, C., & Orszag, S. (2013). Advanced Mathematical Methods for Scientists and Engineers I. *Springer*, *1*. doi: https://doi.org/10.1007/978-1-4757-3069-2

Diamond, P. (1987). Slowly varying functions of two variables and a tauberian theorem for the double Laplace transform. *Appl. Anal.*, *23*(4), 301-318. doi: https://doi.org/10.1080/00036818708839649

Feller, W. (1971). *An introduction to probability theory and its applications. Vol. II.* New York: John Wiley & Sons Inc.

Felzer, K., Abercrombie, R., & Ekstrom, G. (2004). A Common Origin for Aftershocks, Foreshocks, and Multiplets. *Bull. Seismol. Soc. Am.*, *94*, 88-98. doi: https://doi.org/10.1785/0120030069

Helmstetter, A., & Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J. Geophys. Res. Solid Earth*, *107*(B10), ESE–10. doi: https://doi.org/10.1029/2001JB001580

Helmstetter, A., & Sornette, D. (2003). Foreshocks explained by cascades of triggered seismicity. *J. Geophys. Res. Solid Earth*, *108*(B10). doi: https://doi.org/10.1029/2003JB002409

Helmstetter, A., Sornette, D., & Grasso, J. (2002). Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. *J. Geophys. Res. Solid Earth*, *108*. doi: https://doi.org/10.1029/2002JB001991

Jagla, E. A., Landes, F. P., & Rosso, A. (2014). Viscoelastic effects in avalanche dynamics: A key to earthquake statistics. *Phys. Rev. Lett.*, *112*(17), 174301. doi:

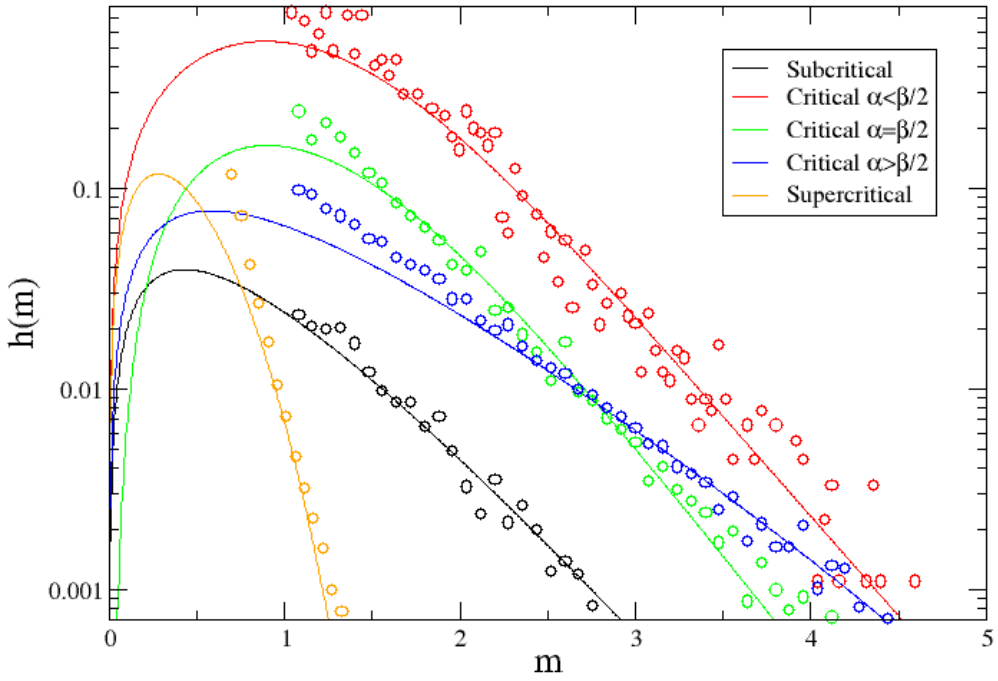

Figure 2. The magnitude dependent function $h(m)$ of Eq. (12) for different cases: subcritical, $\mu < 1$ (black); critical, $\mu = 1$, $\alpha < \beta/2$ (red); critical, $\mu = 1$, $\alpha = \beta/2$ (green); critical, $\mu = 1$, $\alpha > \beta/2$ (blue); supercritical, $\mu > 1$ (orange). The lines represent the analytical results, whereas the circles are the ETAS numerical simulations.

https://doi.org/10.1103/PhysRevLett.112.174301

Jordan, T., Chen, Y., Gasparini, P., Madariaga, R., Main, I., Marzocchi, W., ... Zschau, J. (2011). Operational Earthquake Forecasting: State of Knowledge and Guidelines for Utilization. *Ann. Geophys.*, *54*. doi: https://doi.org/10.4401/ag-5350

Jordan, T., Marzocchi, W., Michael, A. J., & Gerstenberger, M. C. (2014). Operational Earthquake Forecasting Can Enhance Earthquake Preparedness. *Seismol. Res. Lett.*, *85*, 955-959. doi: https://doi.org/10.1785/0220140143

Lippiello, E., Petrillo, G., & Godano, C. (2020). Recognizing the waveform of a foreshock. *ResearchSquare*. (Preprint) doi: https://doi.org/10.21203/rs.3.rs-57209/v1

Lippiello, E., Petrillo, G., Landes, F., & Rosso, A. (2019). Fault heterogeneity and the connection between aftershocks and afterslip. *Bull. Seismol. Soc. Am.*, *109*(3), 1156–1163. doi: https://doi.org/10.1785/0120180244

Lippiello, E., Petrillo, G., Landes, F., & Rosso, A. (2021). The genesis of aftershocks in spring slider models. *Stat. Methods Model. Seismogenesis*, *1*, 131–151.

Llenos, A. L., & Michael, A. J. (2017). Forecasting the (un) productivity of the 2014 M 6.0 South Napa aftershock sequence. *Seismol. Res. Lett.*, *88*(5), 1241–1251. doi: https://doi.org/10.1785/0220170050

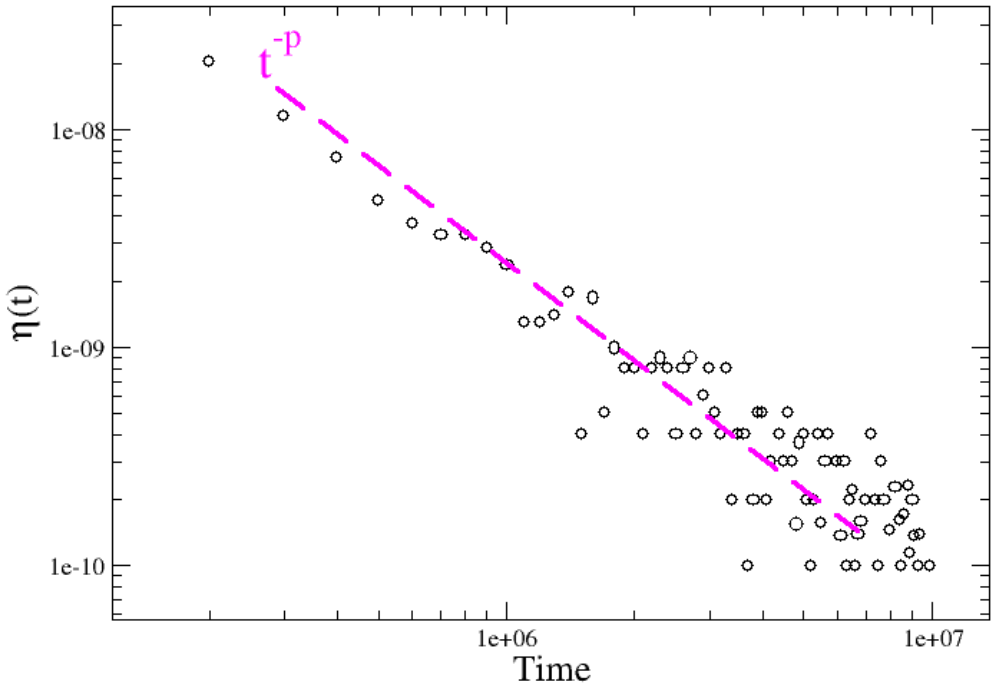

Figure 3. The function $\eta(t)$ of Eq. (12) computed from an ETAS synthetic catalogue (black circles). The magenta dashed line represents the asymptotic prediction for the temporal decay $\eta(t)$. In our simulation we set the parameter p of the Omori-Utsu law equal to 1.2.

Luo, J., & Zhuang, J. (2016). Three Regimes of the Distribution of the Largest Event in the Critical ETAS ModelShort Note. *Bull. Seismol. Soc. Am.*, *106*(3), 1364-1369. doi: https://doi.org/10.1785/0120150324

Marzocchi, W. (2008). Earthquake forecasting in Italy, before and after Umbria-Marche seismic sequence 1997. A review of the earthquake occurrence modeling at different spatio-temporal-magnitude scales. *Ann. Geophys.*, *51*(2-3), 405-416. doi: https://doi.org/10.4401/ag-4451

Marzocchi, W., & Lombardi, A. M. (2009). Real-time forecasting following a damaging earthquake. *Geophys. Res. Lett.*, *36*(21). (L21302) doi: https://doi.org/10.1029/2009GL040233

Marzocchi, W., Lombardi, A. M., & Casarotti, E. (2014). The Establishment of an Operational Earthquake Forecasting System in Italy. *Seismol. Res. Lett.*, *85*(5), 961-969. doi: https://doi.org/10.1785/0220130219

Marzocchi, W., Murru, M., Lombardi, A. M., Falcone, G., & Console, R. (2012). Daily earthquake forecasts during the May-June 2012 Emilia earthquake sequence (northern Italy). *Ann. Geophys.*, *55*(4). doi: https://doi.org/10.4401/ag-6161

Marzocchi, W., Spassiani, I., Stallone, A., & Taroni, M. (2019). How to be fooled searching for significant variations of the b-value. *Geophys. J. Int.*, *220*(3), 1845-1856. doi: https://doi.org/10.1093/gji/ggz541

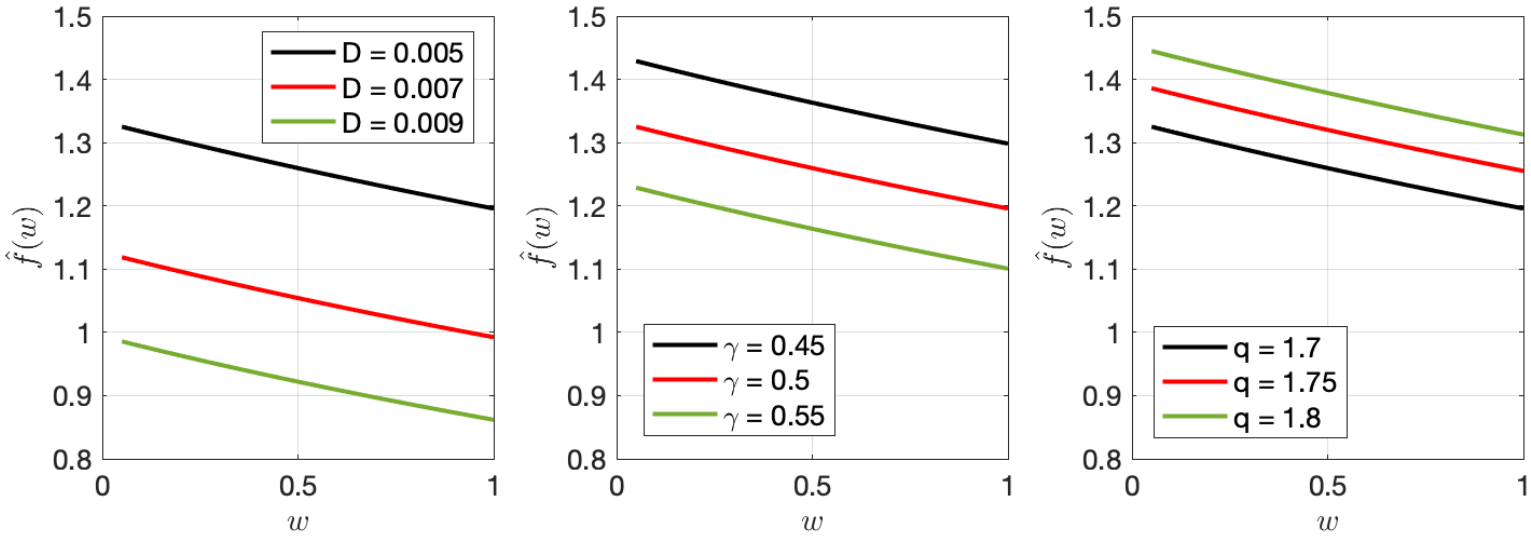

Figure 4. The space function $\hat{f}(z)$, given in Eq. (23), varying the parameter D keeping fixed $\gamma = 0.50$ and $q = 1.7$ (left), varying the parameter γ keeping fixed $D = 0.005$ and $q = 1.7$ (central) and varying the parameter q keeping fixed $D = 0.005$ and $\gamma = 0.50$ (right). The values of m_c and m used to obtain δ_1 and δ_2 are 2 and 5, respectively.

Marzocchi, W., Taroni, M., & Falcone, G. (2017). Earthquake forecasting during the complex Amatrice-Norcia seismic sequence. *Sci. Adv.*, *3*(9). doi: https://doi.org/10.1126/sciadv.1701239

Merrifield, A., Savage, M. K., & Vere-Jones, D. (2004). Geographical distributions of prospective foreshock probabilities in New Zealand. *N. Z. J. Geol. Geophys.*, *47*(2), 327-339. doi: https://doi.org/10.1080/00288306.2004.9515059

Mignan, A. (2014). The debate on the prognostic value of earthquake foreshocks: A meta-analysis. *Sci. Rep.*, *4*(1), 4099. doi: https://doi.org/10.1038/srep04099

Molchan, G. (2005). Interevent time distribution in seismicity: a theoretical approach. *Pure Appl. Geophys.*, *162*(6-7), 1135–1150. doi: https://doi.org/10.1007/s00024-004-2664-5

Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *J. Am. Stat. Assoc.*, *83*(401), 9-27. doi: https://doi.org/10.1080/01621459.1988.10478560

Ogata, Y. (1989). Statistical model for standard seismicity and detection of anomalies by residual analysis. *Tectonophysics*, *169*(1-3), 159-174. doi: https://doi.org/10.1016/0040-1951(89)90191-1

Ogata, Y. (1998). Space-Time Point Process Models for Earthquake Occurrences. *Ann. Inst. Stat. Math.*, *50*(2), 379-402. doi: https://doi.org/10.1023/A:1003403601725

Omi, T., Ogata, Y., Shiomi, K., Enescu, B., Sawazaki, K., & Aihara, K. (2018). Implementation of a real-time system for automatic aftershock forecasting in Japan. *Seismol. Res. Lett.*, *90*(1), 242–250. doi: https://doi.org/10.1785/0220180213

Page, M., van der Elst, N., Hardebeck, J., Felzer, K., & Michael, A. J. (2016). Three Ingredients for Improved Global Aftershock Forecasts: Tectonic Region, Time-Dependent Catalog Incompleteness, and Intersequence Variability. *Bull. Seismol. Soc. Am.*, *106*, 2290-2301. doi: https://doi.org/10.1785/0120160073

Petrillo, G., & Lippiello, E. (2021). Testing of the foreshock hypothesis within an

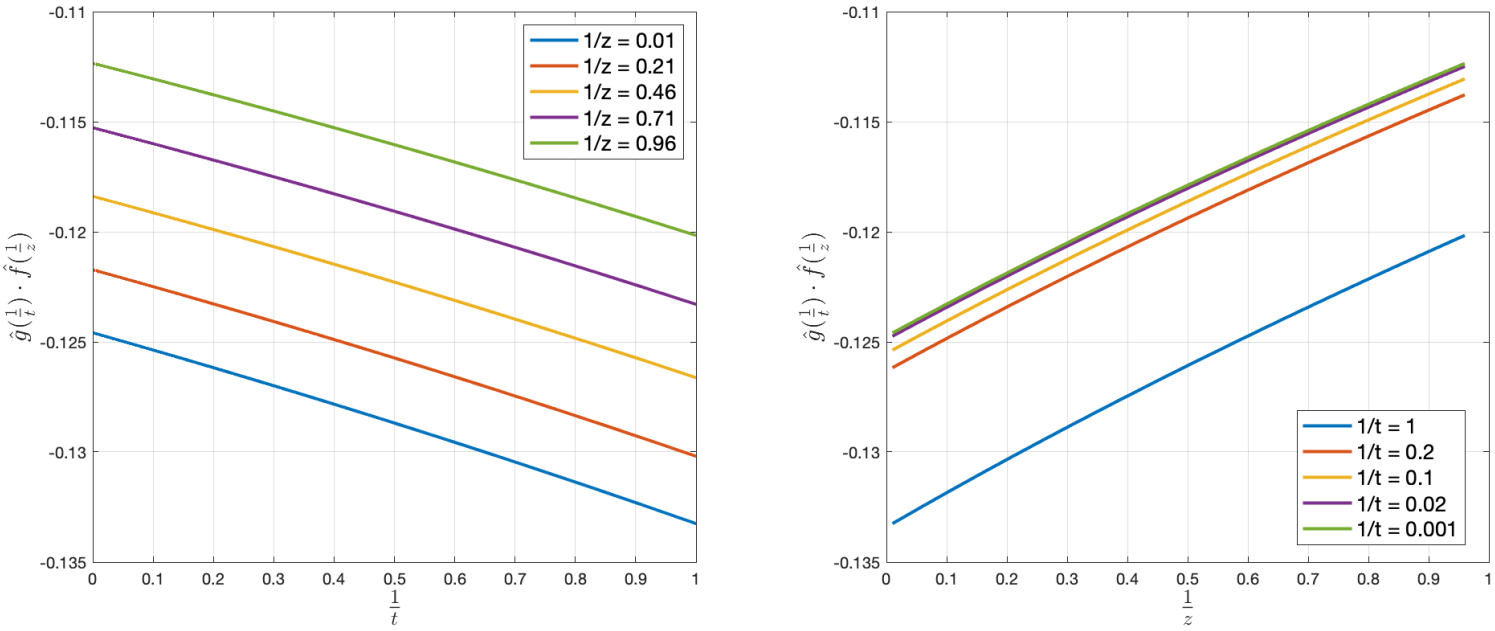

Figure 5. The product function $\hat{g}\left(\frac{1}{t}\right)\hat{f}\left(\frac{1}{z}\right)$ of Eq. (26) - $\hat{g}, \hat{f}$ are respectively defined in Eqs. (9) and (23) - with respect to: the temporal variable $\frac{1}{t}$ and fixing five different values of the spatial variable $\frac{1}{z}$ (left panel); the spatial variable $\frac{1}{z}$ in the and fixing five different values of the temporal variable $\frac{1}{t}$ (right panel). The plot has been obtained by setting $(p, c, D, q, \gamma, m_c, m) = (1.1, 0.07, 0.005, 1.7, 0.5, 2, 5)$.

epidemic like description of seismicity. *Geophys. J. Int.*, *225*(2), 1236–1257. doi: https://doi.org/10.1093/gji/ggaa611

Petrillo, G., Lippiello, E., Landes, F. P., & Rosso, A. (2020). The influence of the brittle-ductile transition zone on aftershock and foreshock occurrence. *Nat. Commun.*, *11*(1), 3010. doi: https://doi.org/10.1038/s41467-020-16811-7

Saichev, A., & Sornette, D. (2005). Distribution of the largest aftershocks in branching models of triggered seismicity: Theory of the universal Båth law. *Phys. Rev. E*, *71*, 056127. doi: https://doi.org/10.1103/PhysRevE.71.056127

Savage, M. K., & Rupp, S. H. (2000). Foreshock probabilities in New Zealand. *N. Z. J. Geol. Geophys.*, *43*(3), 461-469. doi: https://doi.org/10.1080/00288306.2000.9514902

Sornette, A., & Sornette, D. (1989). Self-organized criticality and earthquakes. *Europhys. Lett.*, *9*(3), 197–202. doi: https://doi.org/10.1209/0295-5075/9/3/002

Spassiani, I., Falcone, G., Murru, M., & Marzocchi, W. (2023). Operational Earthquake Forecasting in Italy: validation after 10 years of operativity. *Geophys. J. Int.*, ggad256. doi: https://doi.org/10.1093/gji/ggad256

Spassiani, I., Yaghmaei-Sabegh, S., Console, R., Falcone, G., & Murru, M. (2023). Comparison analysis of the ETAS model with Gutenberg–Richter (GR), Tapered-GR and characteristic magnitude distributions. *Geophys. J. Int.*, *232*(1), 413–428. doi: https://doi.org/10.1093/gji/ggac347

Temme, N. M. (1996). *Special functions. An introduction to the classical functions of mathematical physics*. New York: John Wiley and Sons Inc.

Vere-Jones, D., & Zhuang, J. (2008). Distribution of the largest event in the criti-

cal epidemic-type aftershock-sequence model. *Phys. Rev. E*, *78*, 047102. doi: https://doi.org/10.1103/PhysRevE.78.047102

Voelker, D. (2013). *Die Zweidimensionale Laplace-Transformation: Eine Einführung in Ihre Anwendung zur Lösung von Randwertproblemen Nebst Tabellen von Korrespondenzen*. Birkhäuser Basel.

Zhuang, J. (2011). Next-day earthquake forecasts for the Japan region generated by the ETAS model. *Earth Planets Space*, *63*(3), 207–216. doi: https://doi.org/10.5047/eps.2010.12.010

Zhuang, J., Harte, D. S., Werner, M. J., Hainzl, S., & Zhou, S. (2012). Theme V - Models and Techniques for Analyzing Seismicity. Basic models of seismicity: Temporal models. *CORSSA*. doi: https://doi.org/10.5078/corssa-79905851

Zhuang, J., Murru, M., Falcone, G., & Guo, Y. (2019). An extensive study of clustering features of seismicity in Italy from 2005 to 2016. *Geophys. J. Int.*, *216*(1), 302–318. doi: https://doi.org/10.1093/gji/ggy428

Zhuang, J., & Ogata, Y. (2006). Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. *Phys. Rev. E*, *73*, 046134. doi: https://doi.org/10.1103/PhysRevE.73.046134