Maximum Earthquake Size and Seismicity Rate from an ETAS Model with Slip Budget

David Marsan[*1] and Yen Joe Tan[2,3]

ABSTRACT

We define a seismicity model based on (1) the epidemic-type aftershock sequence model that accounts for earthquake clustering, and (2) a closed slip budget at long timescale. This is achieved by not permitting an earthquake to have a seismic moment greater than the current seismic moment deficit. This causes the Gutenberg–Richter law to be modulated by a smooth upper cutoff, the location of which can be predicted from the model parameters. We investigate the various regimes of this model that more particularly include a regime in which the activity does not die off even with a vanishingly small spontaneous (i.e., background) earthquake rate and one that bears strong statistical similarities with repeating earthquake time series. Finally, this model relates the earthquake rate and the geodetic moment rate and, therefore, allows to make sense of this relationship in terms of fundamental empirical law (the Gutenberg–Richter law, the productivity law, and the Omori law) and physical parameters (seismic coupling, tectonic loading rate).

KEY POINTS

- A model that accounts for both earthquake clustering and seismic moment budget is proposed.
- An upper smooth cut-off of the Gutenberg-Richter law is caused by the seismic moment budget.
- A remarkable, self-organized critical regime is found, in which all earthquakes are aftershocks.

INTRODUCTION

By combining together well-accepted empirical laws, the epidemic-type aftershock sequence (ETAS) model (Ogata, 1988; see Zhuang *et al.*, 2011, for a review) has proved to be a valuable approach for studying various aspects in earthquake studies. It provides a general framework for analyzing complex earthquake time series, make sense of them, and hence allows to identify in the data anomalous features that are not "built-in" the model. Using ETAS as a null hypothesis, one can objectively address the issue of measuring whether a specific earthquake pattern is at odds with what can be considered as a normal behavior. This is particularly useful for evaluating whether a cluster of earthquakes, for example related to swarm of foreshock activity, could simply emerge from normal earthquake triggering, hence clustering, laws, or if it is rather the signature of an aseismic phenomenon, for example, a fluid pressure transient or slow slip (e.g., Hainzl and Ogata, 2005; Lombardi *et al.*, 2010; Marsan *et al.*, 2013; Reverso *et al.*, 2015).

ETAS models are based on a stochastic approach and do not contain any mechanical information related to the state of the fault zone (e.g., slip, strain, or stress). This is particularly powerful, as dealing with numbers of earthquakes only makes this approach a simple and efficient one, which can deal with very large earthquake datasets that nowadays seismic networks and more and more efficient detection methods quasi-routinely produce. It is also a limitation, as no mechanical constrains originating from either frictional or bulk properties are incorporated in the data. A simple illustration of this limitation comes from the widespread use in ETAS models of an unbounded Gutenberg–Richter (G-R) law to generate synthetic earthquake time series. It is straightforward to show that the mean seismic moment of such a distribution is infinite (see the Observed Upper Cutoff in Magnitude section; Knopoff and Kagan, 1977; Sornette and Sornette, 1999), as long as the *b*-value that characterize the G-R law is less than 1.5—as is very often observed with real data. Although this has no consequence when dealing with just numbers or rates of earthquakes, this becomes a critical problem when one needs to model seismic slip on a fault: an ETAS-generated time series with no upper bound in the G-R law will result in a cumulative slip or moment (Brune, 1968; Kostrov, 1974) with no statistical upper bound, hence with the potential to reach any arbitrary (large) slip-rate value even after very long observation times. Thus, while

1. Université Savoie Mont-Blanc, CNRS, IRD, IFSTTAR, ISterre, Le Bourget-du-Lac, France; 2. Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York, U.S.A.; 3. Now at Department of Geophysics, Stanford University, Stanford, California, U.S.A.

*Corresponding author: david.marsan@univ-smb.fr

Cite this article as Marsan, D., and Y. J. Tan (2020). Maximum Earthquake Size and Seismicity Rate from an ETAS Model with Slip Budget, *Bull. Seismol. Soc. Am.* **110**, 874–885, doi: 10.1785/0120190196

mimicking important features found in earthquake catalogs (i.e., clustering), it cannot describe how the total slip evolves on a fault. This unfortunately prohibits possibly fruitful comparisons with geodetic information. The simple addition of an upper cutoff to the G-R law would solve this issue, but would be *ad hoc*, as the choice of the location and shape of this upper cutoff would not be constrained.

Besides ETAS models, statistical approaches have been developed to infer how the G-R law must be (upper) bounded by exploiting geodetic and seismological constraints (Kagan, 2002b; D'Agostino, 2014; Avouac, 2015; Stevens and Avouac, 2017). These studies estimate a maximum magnitude and can infer a mean return time for the largest earthquakes of a given tectonic region, based on seismicity and geodetic data observed at the scale of years to tens of years (Molnar, 1979; McCaffrey, 1994). Given (1) λ the earthquake rate above a minimum magnitude m_0, (2) the b-value of the G-R law, (3) the seismic moment rate $\dot{M}_0$, the requirement that over long timescales $\dot{M}_0 = \lambda\, \bar{M}_0$ with $\bar{M}_0$ the mean seismic moment for any random earthquake put a constraint on the maximum magnitude Ω, yielding

$$\Omega = \frac{1}{\gamma - \beta}\left(-\beta m_0 - 9.1\ln 10 + \ln \dot{M}_0 - \ln\lambda + \ln\frac{\gamma-\beta}{\beta}\right), \qquad (1)$$

for $\dot{M}_0$ in $\mathrm{N}\cdot\mathrm{m}$ per unit time, with $\gamma = 1.5\ln 10$ and $\beta = b\ln 10$, in the case of a sharp cutoff (see Kagan, 2002a,b, for alternative models with smoother cutoffs), and in the $\Omega \gg m_0$ limit. This is crucial for seismic hazard assessment. Somewhat similar approaches, recognizing that seismic moment release is dominated by the largest shocks (provided $b < 1.5$), have shown that one can reconcile seismic coupling inferred from earthquake rates (including paleoseismic observations) and geodetic interseismic coupling (Scholz and Campos, 2012). Both methods are, however, known to be prone to undersampling (McCaffrey, 1997): for example, the pre-2011 estimate of the earthquake-based seismic coupling for the offshore Honshu subduction was 1/2 of the geodetic estimate (Pacheco *et al.*, 1993), this discrepancy vanishing when adding the M_w 9.0 Tohoku earthquake (Scholz and Campos, 2012). With a return time of about 1000 yr or more, the chance to miss an M_w 9.0 earthquake in less than 100 yr of seismicity observation is high. More generally, unknown past occurrences of large earthquakes and nonstationarity of the earthquake rate can result in biased estimations of λ and therefore of the maximum magnitude Ω. Because earthquakes cluster in time, there exists period's intense activity intertwined with lulls. The intermittency of earthquake occurrence rates must thus be accounted for to avoid such undersampling. More fundamentally, these methods assume that λ constrains Ω—although the converse is also true—and thus they do not explain the seismicity rate value, even though the two are coupled. Modeling both Ω and λ from more fundamental (i.e., stable) variables can be performed by considering a seismicity model constrained by a closed long-term slip budget.

We here propose to reconcile both approaches: (1) ETAS models that, on their own, cannot resolve the slip budget and (2) estimations of the maximum magnitude, which rely on an inadequate Poisson description of earthquake occurrences. The connection between the two is very simply made by constraining the cumulative seismic moment release to always be less than the cumulative moment imposed by the tectonic forcing, the rate of which can be independently estimated by geodesy. Several important consequences naturally emerge from adding this constraint: the G-R law is then modulated by a smooth upper cutoff, location of which can be predicted from model parameters. In particular, this cutoff depends on the background rate of earthquakes, which is a much less intermittent, nearly constant, hence more fundamental variable than the highly complex earthquake rate—which includes aftershock sequences; estimation of the background rate can however be delicate, and the gain of describing the seismicity by a more stable background rate rather than by the more intermittent raw rate has to be balanced against potentially severe estimation biases, as will be explained later. Remarkably, our approach makes original predictions about the relationships between (1) seismic coupling times relative plate velocity, (2) the mean earthquake rate, and (3) the background earthquake rate, into nonlinear ones. This thus allows to directly relate observed earthquake rates and geodetic information in a nontrivial way, as will be described later.

This article is structured as follows: we first describe the model, including a simple modification of the model, to produce simulations that quickly reach steady state. The upper cutoff on the G-R relationship that naturally arises in the modeled time series is then analyzed. We show how the earthquake rates (both overall and background) and the imposed moment rate are nonlinearly related and point out to specific regimes of interest. Finally, we discuss the potential pitfalls when estimating model parameters from the data, especially when the p exponent of the Omori–Utsu law is close to 1.

MODEL

Definition

We model earthquakes with magnitudes $\{m_i\}$ above a minimum magnitude m_0; their occurrence times $\{t_i\}$ follow the simplest version of the ETAS model (Ogata, 1988): earthquakes result from a nonstationary Poisson process with rate

$$\lambda_t = \mu + \sum_{i/t_i<t} n_0 e^{\alpha(m_i - m_0)}\frac{p-1}{c}\left(1 + \frac{t-t_i}{c}\right)^{-p}, \qquad (2)$$

in which μ is the background rate, assumed constant here, and the sum corresponds to aftershock triggering. We do not consider earthquake spatial locations or rupture extents, that is, the model is purely temporal. We further assume that the earthquakes follow a strictly bounded G-R law with density

$$f_t(m) = \frac{\beta e^{-\beta m}}{e^{-\beta m_0} - e^{-\beta \Omega_t}}, \quad (3)$$

for $m_0 < m < \Omega_t$, in which $\beta = b \ln 10$, and Ω_t is the maximum magnitude conditioned on time t. We suppose that $\Omega_t \geq m_0$ at all time. This upper bound is taken such that the moment deficit

$$\Delta M_0(t) = \dot{M}_0 t - \sum_{i/t_i<t} M_{0,i}, \quad (4)$$

that is, the departure from the Brune (1968) moment estimate with the geodetic one, would be exactly filled by an earthquake of magnitude Ω_t. We emphasize that the model does not impose a seismic moment conservation, only an upper, time-varying maximum magnitude Ω_t that prevents ruptures to relax more slip than the cumulative slip deficit. We will, however, show in the Observed Upper Cutoff in Magnitude section that a long-term seismic moment conservation is effectively verified by the model, as the mean of Ω_t indeed agrees with the prediction based on this principle. We use the Hanks and Kanamori (1979) relationship $M_0 = 10^{1.5m+9.1}$ relating seismic moment M_0 (in $\mathrm{N \cdot m}$) with magnitude m, so that $M_{0,i} = 10^{1.5m_i+9.1}$ and $\Delta M_0(t) = 10^{1.5\Omega_t+9.1}$. We will here make the hypothesis that the forcing moment rate $\dot{M}_0$ is constant with time.

The branching ratio $\bar{n}$ is a key parameter in ETAS models and is defined as the mean number of aftershocks directly triggered by an earthquake taken at random (Helmstetter and Sornette, 2002). Because the total number of aftershocks grows as $n_0 e^{\alpha(m-m_0)}(1 + \bar{n} + \bar{n}^2 + \cdots)$ when adding aftershocks of the second, third, …, generations, the condition $\bar{n} < 1$ is required to stabilize the model; $\bar{n} > 1$ would lead to an infinite number of earthquakes in a finite time, hence λ growing to infinity. Here, it is straightforward to show that $\bar{n}$ depends on time because it is conditioned by Ω_t

$$\bar{n}_t = \int_{m_0}^{\Omega_t} dm\, n_0 e^{\alpha(m-m_0)} f_t(m) = \frac{\beta n_0}{\beta - \alpha} \times \frac{e^{(\alpha-\beta)(\Omega_t - m_0)} - 1}{e^{-\beta(\Omega_t - m_0)} - 1}, \quad (5)$$

if $\alpha = \beta$, or $\bar{n}_t = \beta n_0 (\Omega_t - m_0)/(1 - e^{-\beta(\Omega_t - m_0)})$ if $\alpha = \beta$. We note that, contrary to the ETAS model, the cutoff at Ω_t allows $\alpha > \beta$; however, because analyses suggest that α is less than β (compare to Hainzl *et al.*, 2013), we will assume this inequality to hold true and thus restricts our attention to $\alpha < \beta$ models, with the exception of the Repeating Earthquakes section. For $\Omega_t \to \infty$ (and $\alpha < \beta$), one finds that $\bar{n}_t = \frac{\beta n_0}{\beta - \alpha}$ and stability requires that $\bar{n}_t < 1$, hence $n_0 < 1 - \frac{\alpha}{\beta}$. As we will show later, this stability criterion is relaxed in our model because Ω_t is finite; however, we find different seismicity regimes depending on whether $n_0 < 1 - \frac{\alpha}{\beta}$ or not (compare to the Dependence of Seismicity Rate λ on Background Rate μ and on Seismic Moment Rate M_0 section), so that the critical value $n_c = 1 - \frac{\alpha}{\beta}$ remains of particular importance.

All model parameters (μ, α, β, p, c, n_0, $\dot{M}_0$) are assumed constant. This implies that we do not account for seismogenic processes that could depend on a time-varying mean level of stress acting on the faults. Such dependencies would render the ETAS model nonlinear and would therefore prevent the analytical developments described in this work.

Accounting for prior earthquake activity

A classical issue when fitting an ETAS model to data is the initial edge effect: knowing earthquakes only past an origin time $t = 0$ implies that, initially, most earthquakes will wrongly be considered as background (Wang *et al.*, 2010; Seif *et al.*, 2017). This is particularly problematic when searching for possible slow changes in background rate, as this will cause a spurious deceleration of this rate (Marsan *et al.*, 2017). Adding data prior to the target period is a way to mitigate this effect (Wang *et al.*, 2010; Harte, 2018). A related problem comes with simulations: generating a synthetic dataset with no activity prior to $t = 0$ leads to a preliminary period during which the rate slowly grows, as more and more aftershock sequences add up. Unfortunately, this transient acceleration can be very long, especially as p is close to 1, so that the steady state is practically never reached. It is thus desirable to generate synthetic datasets that quickly reach steady state. We here describe how this can be achieved, developing on ideas detailed in van der Elst (2017).

Instead of starting a simulated dataset ab nihilo, one can simply assume that it was preceded by seismicity activity with a rate constant with time, equal to λ_∞ defined as the mean seismicity rate for infinitely long time periods. These prior earthquakes have magnitudes distributed according to the G-R law $f(m)$ of equation (3) bounded by Ω_t, which can be approximated as a fixed mean value $\bar{\Omega}$. We will detail in the Observed Upper Cutoff in Magnitude section how the rate λ_∞ and the upper bound $\bar{\Omega}$ can be directly computed from model parameters. This prior activity generates aftershocks during the simulated $t > 0$ period, a prior earthquake at time $t' < 0$ triggering aftershocks with rate:

$$\lambda(t; t') = n_0 \frac{p-1}{c} \left(1 + \frac{t - t'}{c}\right)^{-p} \int_{m_0}^{\bar{\Omega}} dm f(m) e^{\alpha(m-m_0)}. \quad (6)$$

Integrating this rate yields $\int_{-\infty}^{0} dt' \lambda(t; t') = \bar{n}(1 + \frac{t}{c})^{1-p}$. Thus, starting at time $t = 0$, the simulated dataset is an outcome of the rate of equation (2), to which a term is added to account for this prior activity:

$$\lambda_t = \mu + \nu_t + \sum_{i/t_i<t} n_0 e^{\alpha(m_i - m_0)} \frac{p-1}{c} \left(1 + \frac{t - t_i}{c}\right)^{-p}, \quad (7)$$

with

$$\nu_t == \lambda_\infty \bar{n} \left(1 + \frac{t}{c}\right)^{1-p}. \quad (8)$$

It is important to note that v_t decays very slowly with time when p is close to 1; this will have consequences on estimation strategies, as discussed later. Also, we set the initial moment deficit to its expected value (when averaged over infinitely long times) $M_0(\bar{\Omega})$ so that the system is already loaded when starting the simulation.

Example

To illustrate the model, we run a 2000-year-long simulation that roughly mimics seismicity in the northern Japanese subduction zone (compare to Fig. 1). We follow the parameters used and obtained by Marsan *et al.* (2017) in their study of the 1990–2011, $m \geq 3$, subduction-related activity (138° < longitude < 145°, 34° < latitude < 42°, and depth within 20 km of the upper surface of the Pacific plate, as given by Hayes *et al.*, 2012): $m_0 = 3$, $\alpha = 2$, $p = 1.1$, $c = 10^{-5}$ days, $b = 0.95$, and $\mu = 0.33$ per day when averaged over the 21 yr. A more thorough explanation about how model parameters can be estimated from the data is given in the Estimating Model Parameters section. The moment rate $\dot{M}_0$ is taken as identical to the Honshu region of Scholz and Campos (2012), where it is assumed that a 8.3 cm/yr convergence is accommodated with a seismic coupling of 0.6, a rigidity of $4\,10^{10}$ Pa, and a coupling width of 110 km, yielding $\dot{M}_0 = 3.75\ 10^{17}$ N · m/day. Taking $n_0 = 0.106$ gives an expected mean rate λ_∞ of 2.17 per day, close to the observed 1990–2011 average of 2.15 per day. The mean branching ratio is then $\bar{n} = 0.85$, and the mean maximum magnitude is $\bar{\Omega} = 9.17$ (see later on how these quantities are inferred from the model parameters).

We count two earthquakes of magnitude 9 and above, and 24 with $m \geq 8$ within the simulated 2000 yr. The evolution with time of Ω_t is also shown: we find that its average is $\bar{\Omega} = 9.03$, close to the expected long time average of 9.17, in agreement with the estimate of Stevens and Avouac (2017). The branching ratio of equation (5) is found to be equal to 0.83 on average more than 2000 yr, whereas the expected long-term average is 0.85. If the G-R law were effectively unbounded, replacing $\Omega_t \to \infty$ in equation (5) would give $\bar{n} = \frac{\beta n_0}{\beta - \alpha} = n_0/n_c = 1.24$, hence a time series with a rate growing to infinity. The bounding of the G-R law constrained by the maximum mobilizable slip therefore provides a way to stabilize the model.

The mean return time for $m > \bar{\Omega}$ earthquakes (with $\bar{\Omega} = 9.17$, the long term expected average) is 926 yr, so that using equation (1) to estimate a maximum magnitude should be sufficiently robust only if observation is sufficiently longer than these 926 yr. As a matter of fact, injecting the observed $\bar{\lambda} = 2.15$ per day in equation (1) yields $\Omega = 9.18$. However, taking shorter and more realistic observation durations helps pointing out the undersampling issue that is typically encountered when exploiting the raw seismicity rate. As an example, we split the simulation into 50 intervals of 40 yr, and show in Figure 1 how the equation (1) estimate varies, in relation to the maximum magnitude effectively observed in the sample.

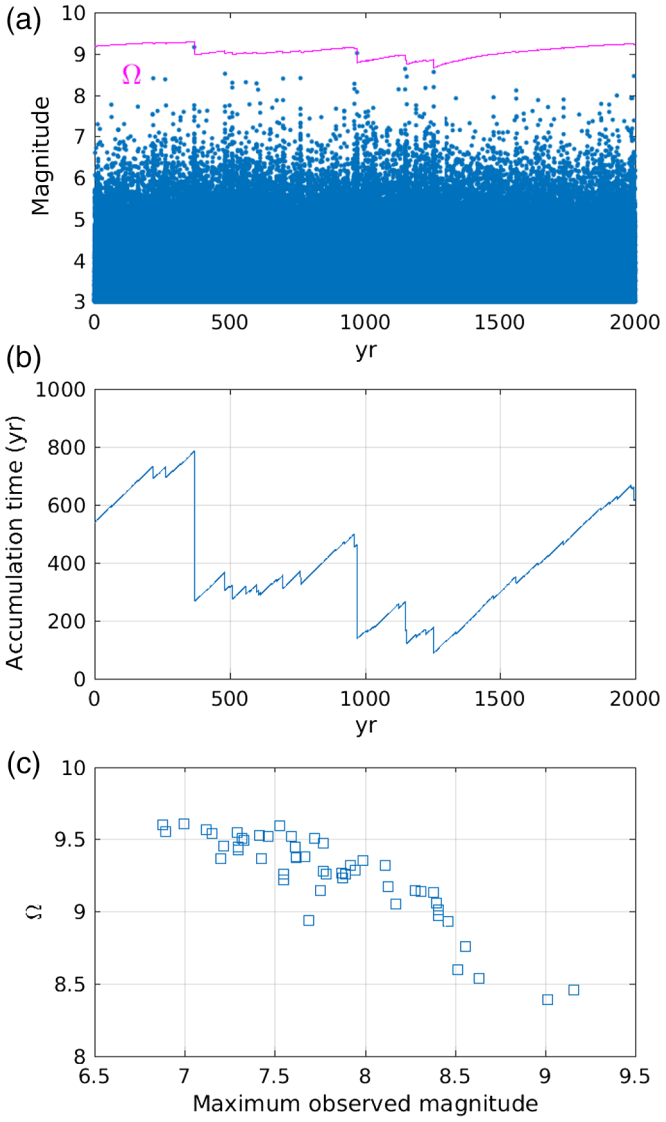

Figure 1. Synthetic earthquake time series generated over 2000 yr, with parameters intended to mimic the northern Japanese subduction zone seismicity: $m_0 = 3$, $\alpha = 2$, $b = 0.95$, $p = 1.1$, $c = 10^{-5}$ days, $n_0 = 0.106$, $\mu = 0.33$ per day, and $\dot{M}_0 = 3.75\ 10^{17}$ N · m/day. (a) The maximum magnitude Ω_t is shown in magenta, and averages 9.02 (the value expected when averaging over an infinitely long time is $\bar{\Omega} = 9.12$, compare to the Observed Upper Cutoff in Magnitude section). (b) The moment deficit $\Delta M_0(t)$ of equation (4) can be transformed into a equivalent time $t_{acc} = \Delta M_0(t)/\dot{M}_0$, called here the accumulation time, proportional to the average shear loading stress acting on the faults. (c) Mean maximum magnitude Ω estimated from equation (1) for 40-year-long windows versus the sample maximum magnitude; Ω varies from 8.38 to 9.60. The color version of this figure is available only in the electronic edition.

A negative correlation is expected given the $\Omega \sim -\frac{1}{\gamma - \beta} \ln \lambda$ dependence of equation (1).

OBSERVED UPPER CUTOFF IN MAGNITUDE

We now show that in this model the magnitude distribution over long periods departs from the G-R law at large magnitudes,

the cutoff being characterized by a location that can be predicted from the model parameters. Most of the derivation is given in Appendix A, where we prove that (1) the model follows the moment conservation principle $\bar{\lambda} \times \bar{M}_0 = \dot{M}_0$ even though it is not imposed *a priori*; (2) the long-term average maximum magnitude $\bar{\Omega}$ can be computed knowing the model parameters; and (3) the departure at large magnitudes of the observed G-R law from a simple exponential law is directly related to the distribution of Ω_t.

Provided that the observation time is long enough, the departure of the observed G-R law with an unbounded one can be clearly seen in simulations. We show this departure in Figure 2 for a synthetic catalog generated with $\mu = 1$ per day, $\alpha = 2$, $b = 1$, $n_0 = 0.2$, $m_0 = 3$, $p = 1.1$, $c = 10^{-5}$ days, and $\dot{M}_0 = 10^{17}$ N · m/day. The high n_0 value (1.66 times the critical n_c value) implies strong aftershock triggering, hence a relatively low $\bar{\Omega} = 6.97$ and a characteristic return time well shorter than T_{obs} (we here define the characteristic return time as the mean return time for $m > \bar{\Omega}$ events, here equal to 2.22 yr). The observed magnitude distribution is related to the distribution of Ω_t, which is unknown to the observer. The relationship between these two distributions is described in Appendix A, item (3). We checked in Figure 2b the validity of equation (A7), see Appendix A, on the simulation of Figure 2a.

Looking at a simulated catalog like those of Figures 1 and 2, it can be seen that Ω_t mostly stays within a limited range centered on the averaged $\bar{\Omega}$, so that $g(\Omega)$ can be at first order modeled by a Gaussian with mean $\bar{\Omega}$ and standard deviation σ. Our simple $\frac{d\Phi(M)}{dM} \simeq -g(M)$ result of Appendix A, which is valid here (we indeed have $\beta(\bar{\Omega} - m_0) \gg 1$), indicates that this *a priori* unknown distribution $g(\Omega)$ directly translates into the observed departure from the G-R law at large magnitudes, with $\Phi(M)$ the complementary error function (erfc) with mean $\bar{\Omega}$ and standard deviation σ. This two-parameter model is different from the simpler (one parameter), tapered G-R law found by Kagan (2002a) to be a good model for global seismicity (Global Centroid Moment Tensor catalog up to 1999; Dziewonski *et al.*, 1981; Ekström *et al.*, 2012). We, however, note that (1) the erfc cutoff model found here could emerge from the somewhat arbitrary choice of a strict magnitude cutoff Ω_t; (2) the estimates of Kagan (2002a) are obtained for a worldwide catalog, so for tectonic regions that have very different seismic moment rates. This mixing is likely to change the overall shape of the magnitude–frequency relationship, that is, one could possibly have different cutoff shapes at small and large scales. These two arguments suggest that a direct comparison with the results of Kagan (2002a) is not straightforward.

As explained previously, $\bar{\Omega}$ can be readily predicted from the model parameters; however, the standard deviation σ cannot, or at least we could not find a way to do so. This unfortunately implies that the value of σ can only be obtained by running simulations and quantifying the observed departure from the G-R law, with (so far) little understanding on how σ effectively depends on model parameters. We fitted the erfc cutoff to the simulated dataset of Figure 1; namely, we estimated the mean and standard deviation of the best Gaussian density g such that $\frac{d\Phi(M)}{dM}$ is given by $-g(M)$. Because b and $\bar{\lambda}$ are known, these estimates of $\bar{\Omega}$ and σ allow to compute the mean return time for large ($m > \bar{\Omega}$) earthquakes. Compared to the approach reviewed in Avouac (2015), this estimation can help constraining the probability and return times of unexpectedly very large earthquakes, that is, with magnitudes greater than the mean $\bar{\Omega}$, therefore refining seismic hazard evaluations.

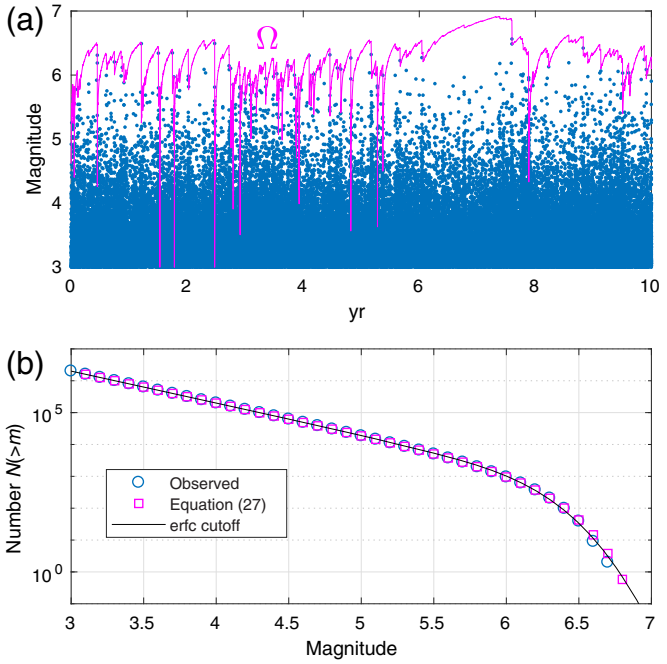

Figure 2. (a) Simulated time series, for $m_0 = 3$, $\alpha = 2$, $b = 1$, $p = 1.1$, $c = 10^{-5}$ days, $n_0 = 0.2$, $\mu = 1$ per day, and $\dot{M}_0 = 10^{17}$ N · m/day, and (b) number of occurrences versus magnitude, showing the departure of a straight unbounded Gutenberg–Richter law at large magnitude. The black curve gives the best fit using an error function. The color version of this figure is available only in the electronic edition.

DEPENDENCE OF SEISMICITY RATE $\bar{\lambda}$ ON BACKGROUND RATE μ AND ON SEISMIC MOMENT RATE $\dot{M}_0$

The mean earthquake rate in a given region is a direct observable that carries information about how the fault system is loaded, how strong the seismic coupling is, and how the faults mechanically interact. However, a relationship that link the mean earthquake rate to stress rate, and so on, is largely missing. In the framework of the present model, such a relationship can however be investigated, although a systematic study of how the two key quantities (mean seismicity rate $\bar{\lambda}$ and mean maximum magnitude $\bar{\Omega}$) depend on model parameters

$\{\mu, \alpha, n_0, \dot{M}_0, \beta\}$ is beyond the scope of this work. We note that the two extra parameters p and c only affect the aftershock triggering kernel shape, and have no impact on the time-averaged $\bar{\lambda}$ and $\bar{\Omega}$. We recall (compare to Appendix A, equations A2–A5) that a set of four equations provides constraints on the four variables $\{\bar{\lambda}, \bar{M}_0, \bar{n}, \bar{\Omega}\}$, for sufficiently long observation times. We rewrite these equations assuming that $\bar{\Omega}$ is sufficiently greater than m_0, and β is sufficiently smaller than γ (which is the case for $b = 1$):

$$\bar{\lambda} \times \bar{M}_0 = \dot{M}_0 \tag{9}$$

$$\bar{\lambda} = \frac{\mu}{1 - \bar{n}} \tag{10}$$

$$\bar{n} \simeq n_0 \frac{\beta}{\beta - \alpha}\left(1 - e^{(\alpha-\beta)(\bar{\Omega}-m_0)}\right) \tag{11}$$

$$\bar{M}_0 \simeq M_0(m_0)\frac{\beta}{\gamma - \beta} e^{(\gamma-\beta)(\bar{\Omega}-m_0)}. \tag{12}$$

A numerical solution can be computed, given the model parameters. However, the exploration of a 5D space is practically nontrivial. Instead, we here more particularly (1) study the specific dependence of $\bar{\lambda}$ (which can be partly constrained directly from observed earthquake rates) on μ and $\dot{M}_0$, allowing us to isolate distinct relationships of interest, and (2) further explore the model by pointing out remarkable regimes.

Dependence of $\bar{\lambda}$ on μ

The $\bar{\lambda} = \frac{\mu}{1-\bar{n}}$ relationship of Helmstetter and Sornette (2003) is valid for infinitely long observation times, and implies a simple proportionality between μ and $\bar{\lambda}$. However, introducing the additional constraint of a fixed seismic moment release rate adds an extra layer of complexity, as changing μ, hence $\bar{\lambda}$, also affects $\bar{n}$ through its dependence on $\bar{\Omega}$. As an example, coming back to the simulation shown in Figure 1, we found that $\bar{\lambda} =$ 2.17 per day; a 10-fold decrease in background rate μ would yield $\bar{\lambda} = 0.46$ per day, hence a decrease by a factor of 4.7 in $\bar{\lambda}$, clearly distinct from the simple proportionality.

Equations (9)–(12) can be solved to infer the dependence of $\bar{\lambda}$ on μ. From there, some algebra (compare to Appendix B) gives the nonlinear dependence

$$\mu = A\bar{\lambda} + B\bar{\lambda}^{\delta}, \tag{13}$$

with A, B, and δ depending on the model parameters (these dependencies are given in Appendix B). An important point is that A can be either negative (if $n_0 > n_c$, in which $n_c = 1 - \frac{\alpha}{\beta}$ is the stability threshold in classical ETAS models) or positive (if $n_0 < n_c$). These two cases correspond to strong or weak aftershock triggering cases, respectively. It is worth emphasizing that $n_0 > n_c$ is not allowed in classical ETAS models. This constraint is relaxed in our model: taking $n_0 > n_c$ yields a regime in which there exists periods of high $\bar{n}_t$, leading to increased activity, hence more large shocks that reduce the moment deficit, hence lowering $\bar{n}_t$. This stabilizing loop ensures that whatever the value of $n_0 > n_c$, the model is stable and does not produce running-off sequences.

The case $A < 0$ (i.e., $n_0 > n_c$) is worth further examination. In the limit $\mu \to 0$, equation (13) gives that $\bar{\lambda} \to \lambda_c$, with

$$\lambda_c = \left(-\frac{A}{B}\right)^{\frac{1}{\delta-1}} = \left(1 - \frac{n_c}{n_0}\right)^{\frac{1}{\delta-1}} \frac{\dot{M}_0}{M_0(m_0)} \times \frac{\gamma - \beta}{\beta}, \tag{14}$$

which is finite (and nonzero). This regime is obtained for $n_0 > n_c$ and $\mu \ll \mu_c$ with $\mu_c = \lambda_c(\frac{n_0}{n_c} - 1)$. It implies that the seismicity rate becomes practically independent of μ. This is a critical regime in which aftershock triggering sustains itself over infinitely long periods of time. Then, very little spontaneous (background) activity can maintain potentially very high seismicity rates. Remarkably, the model naturally reaches $\bar{n} \to 1$ without having to precisely tune model parameters, in a way similar to a self-organized critical system.

Still for $n_0 > n_c$ and $\mu \ll \mu_c$, we find that

$$\bar{\Omega} = m_0 + \frac{1}{\alpha - \beta}\ln\left(1 - \frac{n_c}{n_0}\right), \tag{15}$$

which implies an undesirable dependence of the mean maximum magnitude $\bar{\Omega}$ on m_0. This issue can be addressed by considering that among all model parameters $\{\alpha, \beta, p, c, n_0, \dot{M}_0, \mu, m_0\}$ both n_0 and μ depend on the choice of m_0. Namely, changing the lower cutoff m_0 requires changing n_0 so that

$$n_0 = \frac{n_c}{1 - e^{(\alpha-\beta)(\bar{\Omega}-m_0)}}. \tag{16}$$

The seismicity rate $\bar{\lambda}$ then follows the expected G-R scaling, that is, $\bar{\lambda} \sim e^{-\beta m_0}$. We note that, in the limit $m_0 \to -\infty$, equation (16) gives that $n_0 \to n_c$, independently of $\bar{\Omega}$. We summarize in Figure 3 the various regimes in the (n_0, μ) parameter space. Figure 4 shows how $\bar{\lambda}$ depends on μ, for the parameters of Figure 1.

Dependence of $\bar{\lambda}$ on $\dot{M}_0$

We now examine how $\bar{\lambda}$ depends on $\dot{M}_0$, with all other variables (including the background rate μ) fixed. Building on the same set of equations as earlier, that is, equations (9)–(12), we again find that two cases must be distinguished:

1. for $n_0 < n_c$, then $\bar{\lambda}$ asymptotically saturates at $\bar{\lambda}_\infty = \frac{\mu}{1-n_0/n_c}$ when $\dot{M}_0 \to \infty$. Because a large moment rate must be

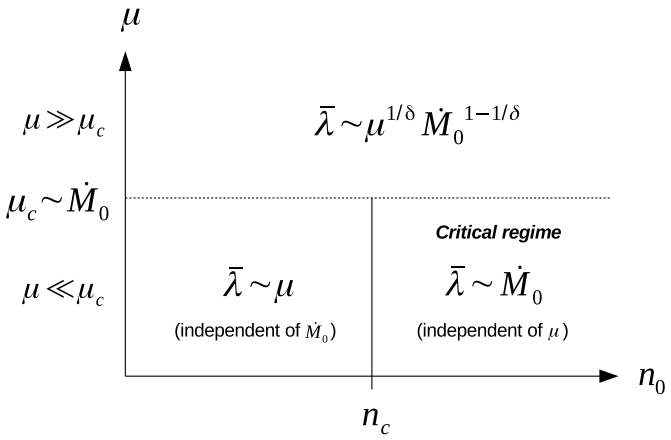

Figure 3. Summary of the $\mu \ll \mu_c$ and $\mu \gg \mu_c$ regimes, function of n_0. We recall that the critical n_c is $n_c = 1 - \frac{\alpha}{\beta}$. These regimes require that $\bar{\Omega} \gg m_0 + \frac{1}{\beta}$; this condition does not hold true any longer when $\dot{M}_0$ is too small.

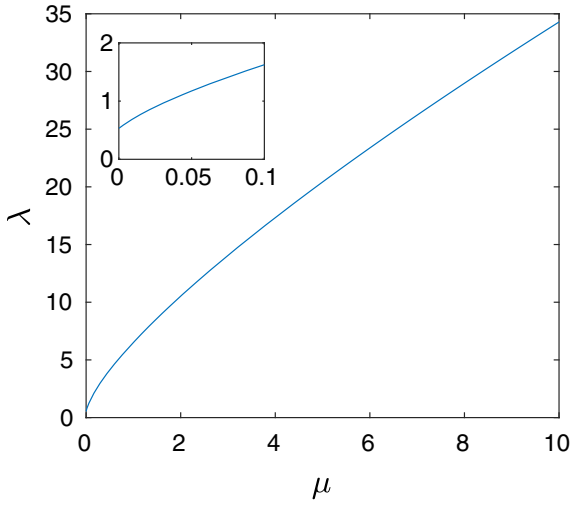

Figure 4. Top graph: dependence of the seismicity rate $\bar{\lambda}$ on the background rate μ, for model parameters $m_0 = 3$, $\alpha = 2$, $b = 0.95$, $n_0 = 0.106$ (hence greater than $n_c = 0.086$), and $\dot{M}_0 = 3.75\ 10^{17}$ N · m/day (same as in Fig. 1). The inset shows a zoom in at small values of μ, with $\lambda \to \lambda_c = 0.53$/day when $\mu \to 0$, corresponding to a self-organized critical regime. Bottom: relative increase of λ vs the relative increase of μ (i.e., $(\frac{\Delta\lambda}{\lambda})/(\frac{\Delta\mu}{\mu})$), function of μ, for the same set of parameters. The sensitivity of $\bar{\lambda}$ on μ becomes null at very low μ-values, in this regime ($n_0 > n_c$), and saturates at $\frac{1}{\delta} = 0.86$ large μ. The color version of this figure is available only in the electronic edition.

released with a finite seismicity rate, this implies that $\bar{\Omega} \to \infty$ when $\dot{M}_0 \to \infty$;

2. for $n_0 > n_c$, then $\dot{M}_0 \to \infty$ causes $\bar{\lambda} \to \infty$, which implies $\tilde{n} \to 1$ as already commented (critical case). Because $\tilde{n}$ depends on Ω_t (hence $\bar{\Omega}$) but not on $\dot{M}_0$, see equations (5) and (11), this implies that $\bar{\Omega}$ does not depend on $\dot{M}_0$, and thus neither does $\bar{M}_0$. As a consequence, the budget of equation (9) imposes that $\bar{\lambda}$ is simply proportional to $\dot{M}_0$.

In both cases, $\bar{\lambda} \simeq \mu$ when $\dot{M}_0 \to 0$. The separation between the low ($\bar{\lambda} \simeq \mu$) and high ($\bar{\lambda} \to \infty$) regimes at $n_0 > n_c$ occurs when $\dot{M}_0$ is of the order of $\mu \frac{\beta M_0(m_0)}{\gamma - \beta} (1 - \frac{n_c}{n_0})^{1/1-\delta}$. This transition value decreases to a low rate when n_0 is kept increasing: for $n_0 \to \infty$, this value is $\mu \frac{\beta M_0(m_0)}{\gamma - \beta}$, hence only a factor $\frac{\beta}{\gamma - \beta}$ greater than the rate dissipated by just the background earthquakes.

A third transitional regime is found for n_0 greater but sufficiently close to n_c that marks the transition between the two $\bar{\lambda} \simeq \mu$ (when $\dot{M}_0 \to 0$) and $\bar{\lambda} \sim \dot{M}_0$ (at greater $\dot{M}_0$ values) regimes, in which $\bar{\lambda}$ is proportional to the $1 - \frac{1}{\delta} = \frac{\beta - \alpha}{\gamma - \alpha}$ power of $\dot{M}_0$. For typical β and α values, that is, $b = 1$ and $\alpha = 0.6 \ln 10$ to $0.9 \ln 10$, compare to Hainzl *et al.* (2013), this power ranges from 0.16 to 0.44, implying a relatively weak dependence of $\bar{\lambda}$ on $\dot{M}_0$. This transitional regime only applies to an ever-narrower interval of $\dot{M}_0$, as n_0 is increased. Figure 5 shows examples of the various $\bar{\lambda}$ versus $\dot{M}_0$ regimes for three cases: (1) $n_0 < n_c$, (2) n_0 greater but close to n_c, and (3) $n_0 > n_c$.

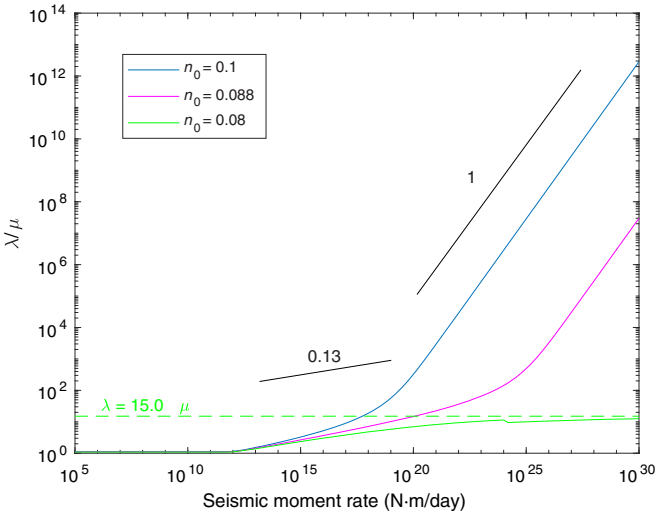

Figure 5. Dependence of $\bar{\lambda}$ (in units of μ) on $\dot{M}_0$, for three values of n_0. We used the same parameters as in Figure 1, with $\mu = 0.01$ per day. The critical n_c is 0.085. For $n_0 < n_c$, $\bar{\lambda}$ ranges between μ and λ_c (dashed lines). For $n_0 > n_c$, an intermediate nonlinear regime $\bar{\lambda} \sim \dot{M}_0^{0.13}$ is found (best seen for $n_0 = 0.088$), followed (at large $\dot{M}_0$) by a simple linear regime. The color version of this figure is available only in the electronic edition.

Dependence on both $\dot{M}_0$ and μ

So far only $\dot{M}_0$ or μ have been left to vary, the other quantity being held constant. However, it is clear that both $\dot{M}_0$ and μ are related; for example, increasing the area A under study implies increasing (1) $\dot{M}_0$ in direct proportion of A (provided that the average tectonic slip rate and seismic coupling do not change when including the new area), and (2) μ as the number of nucleation sites must necessarily grow with A. Moreover, both quantities depend on the slip rate $\dot{\delta}$, this dependence being a

simple proportionality for $\dot{M}_0$. The dependence of μ on $\dot{\delta}$ is less trivial: a linear $\mu \sim \dot{\delta}$ relationship was found for subduction zones by Ide (2013), and to characterize changes during slow-slip events in Boso (Reverso *et al.*, 2016), but previous work by Bird *et al.* (2009) suggests that more complex nonlinear relationships could emerge owing to nonelastic effects (e.g., thermal, poroelastic, or viscous). If μ is in proportion of the number of velocity-weakening patches that are big enough to nucleate slip accelerating to seismic velocities, then increasing the loading rate is known from laboratory friction experiments to shrink the size of the nucleation zones of future dynamic instabilities, so that more patches could nucleate earthquakes (McLaskey and Yamashita, 2017; Guerin-Marthe *et al.*, 2019). This effect would result in a nonlinear dependence of μ with $\dot{\delta}$, as the number of sites increases and as the rate of sites reaching the nucleation threshold and starting the nucleation phase grows proportionally with $\dot{\delta}$. This dependence is, however, likely only weakly nonlinear, Guerin-Marthe *et al.* (2019) finding that the nucleation radius decays in $-\ln \dot{\delta}$.

The $\mu \sim \dot{\delta}$ case leads to a particularly simple result. Then, all regimes of Figure 3 give that $\bar{\lambda} \sim \dot{\delta}$. We emphasize that this simple relationship is only strictly valid for a fault undergoing a change in slip rate (all other model parameters being the same), but cannot be used to relate distinct faults that could be characterized by different values of the model parameters.

A similar simple conclusion is obtained when considering the dependence of the model on the area A. Extending the area implies increasing both $\dot{M}_0$ and μ. If we neglect the variations of other model parameters (i.e., α, β, n_0) with the spatial extent, and assume that both $\dot{M}_0$ and μ are simply proportional to A, then $\bar{\lambda}$ becomes also proportional to A, so that the ratio $\bar{\lambda}/\dot{M}_0$ and $\bar{\lambda}/\mu$ are independent of the spatial extent. According to equations (9) and (12), a constant ratio $\bar{\lambda}/\dot{M}_0$ implies a constant $\bar{M}_0$, hence a constant $\bar{\Omega}$: over infinitely long observation times, the G-R cutoff remains the same whatever the size of the considered area.

Repeating earthquakes

We have assumed that $\alpha < \beta$, as is observed at regional scale when inverting ETAS parameters. Relaxing this constraint, we now discuss the specific case of the $\alpha \geq \beta$ and $\mu \to 0$ regime. Developing the same argument as above (compare to Appendix B), we obtain that, in this $\mu \to 0$ limit, $\bar{\lambda} \to \lambda_c$ with

$$\lambda_c = \left(1 - \frac{n_c}{n_0}\right)^{\frac{\gamma-\beta}{\beta-\alpha}} \frac{\dot{M}_0}{M_0(m_0)} \frac{\gamma-\beta}{\beta}, \tag{17}$$

if $\alpha > \beta$, or

$$\lambda_c = e^{\frac{\beta-\gamma}{n_0\beta}} \frac{\dot{M}_0}{M_0(m_0)} \frac{\gamma-\beta}{\beta}, \tag{18}$$

if $\alpha = \beta$. This is a self-organized critical regime, as discussed earlier, but here without any requirement on n_0. We recall that this regime was found in the $\beta > \alpha$ case when $n_0 > n_c$; here, $\alpha \geq \beta$ implies that $n_c \leq 0$, so that this $n_0 > n_c$ condition is always verified, and the self-organized regime is thus found whatever the value of n_0.

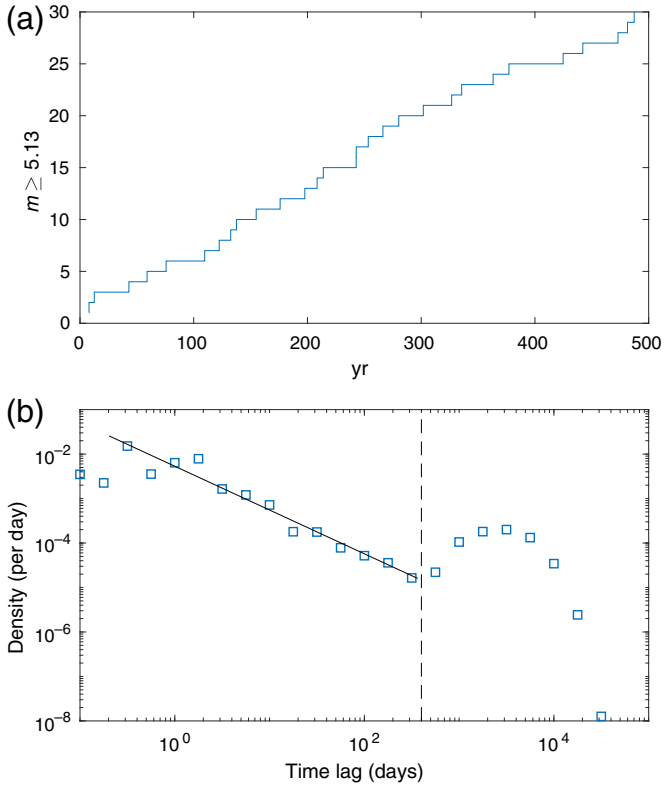

Figure 6. Results of a simulation ran over 10^5 yr, for $m_0 = 3$, $\mu = 10^{-3}$ per day, $n_0 = 0.02$, $b = 1$, $\alpha = 4$, and $\dot{M}_0 = 10^{14}$ N · m/day. The expected $\bar{\Omega}$ is 5.13. (a) Time series zoomed over 500 yr of $m > \bar{\Omega}$ earthquakes, which are considered as "asperity-sized," displaying quasi periodicity. The coefficient of variation is 0.65. (b) Distribution of the interevent times for these asperity-sized earthquakes. The regime at short timescales gives a density in 1/time gap (black line) and accounts for 3% of these earthquakes. The second regime is the quasi-periodic mode centered at ≃9.5 yr. The color version of this figure is available only in the electronic edition.

Interestingly, this regime exhibits a relatively simple dynamics, in which characteristic earthquakes occur quasi-periodically. As an example, we show in Figure 6 results from a simulation ran for $m_0 = 3$, $\mu = 10^{-3}$ per day, $n_0 = 0.02$, $b = 1$, $\alpha = 4$, and $\dot{M}_0 = 10^{14}$ N · m/day, for which the expected $\bar{\Omega}$ is 5.13. Considering all $m > \bar{\Omega}$ earthquakes, the interevent times display a clear bimodal distribution with aftershock triggering at short timescales (about 3% of the events) and a well-peaked mode centered at ≃9.5 yr. Such a bimodal distribution and proportion of aftershocks are similar to what Lengliné and Marsan (2009) found for repeating earthquake interevent times in Parkfield. Here, this quasi-periodic dynamics naturally arises because $\alpha \geq \beta$ allows for very large $\bar{n}_t$ values, which cause the earthquake interactions to quickly cascade during aftershock sequences, exhausting the moment deficit

and resetting the loading near zero. Equation (17) implies a simple $\bar{\lambda} \sim \dot{M}_0$ scaling; because this regime displays vanishing dependence on μ, this directly translates into a $\bar{\lambda} \sim \dot{\delta}$ relationship, so that the repeating failure of such an asperity allows to gauge possible changes in local slip rates, as initially proposed by Nadeau and McEvilly (1999).

This model of a repeating earthquake thus assumes that (1) the asperity is isolated (it must accommodate a constant $\dot{M}_0$ on its own); (2) it constantly undergoes partial ruptures, with sizes smaller than the asperity size; (3) this inner dynamics is controlled by "regular" earthquake laws, that is, the Omori law, the G-R law, the productivity law, albeit with a large productivity exponent α (or, alternatively, a small b-value, because the key requirement is $\alpha \geq \beta$). Then, the branching ratio can temporally be much greater than 1, causing episodes of cascading stress release, which culminate with an asperity-size failure. Remarkably, the critical regime of such an asperity implies that all its failures, either partial or complete, are effectively aftershocks of previous failures, because $\mu \to 0$. This inner dynamics model of an asperity complements the frictional model of Dublanchet *et al.* (2013) that sees the failure of an asperity as resulting from microfailures of microasperities, provided the latter have a high enough density; see also Cattania (2019) for a similar approach at the scale of an isolated fault.

ESTIMATING MODEL PARAMETERS

As already evidenced by van der Elst (2017), it can be difficult to distinguish between background earthquakes and aftershocks from long past mainshocks, the more so as T_{obs} is short and p is close to 1. This has undesirable consequences regarding to our capacity at estimating model parameters from the data.

Based on the model of equations (7) and (8), the maximum-likelihood estimation of parameters $\{\mu, \alpha, n_0, p, c\}$ from an earthquake dataset—assuming that both $\dot{M}_0$ and β can be independently estimated on their own—implies minimizing the cost or misfit function (Ogata, 1998)

$$J = \int_0^{T_{\text{obs}}} dt\ \lambda_t - \sum_{i/0<t_i<T_{\text{obs}}} \ln \lambda_t. \tag{19}$$

For $p \simeq 1$, ν_t of equations (7) and (8) decays very slowly with time, and its temporal signature becomes practically indistinguishable from the constant background rate μ. The model can then be simplified by defining $\mu_{\text{eff}} = \mu + \nu$ as an effective background rate, with ν being for example the time average of ν_t over the observed time interval, or its initial $\lambda_\infty \bar{n}$ value (taking $\nu = \lambda_\infty \bar{n}$ implies that $\mu_{\text{eff}} = \lambda_\infty$). However, this prevents a definite determination of μ itself, which only accounts for an unknown portion of μ_{eff}. This is an issue, as the μ-value is required to estimate λ_∞, $\bar{n}$, and $\bar{\Omega}$, as well as the standard deviation σ of the erfc cutoff model.

As an illustration, we refer to the simulation of Figure 1, for which $\mu = 0.33$ per day, $\lambda_\infty = 2.15$ per day, and $\bar{n} = 0.85$. The initial (maximum) value of ν_t is then $\lambda_\infty \bar{n} = 1.82$ per day, hence about five times the background rate. Based on data alone, one could therefore find that $\mu_{\text{eff}} = \lambda_\infty = 2.15$ per day, but would not be able to isolate μ. Taking extreme guesses $\mu \to 0$ and $\mu = \mu_{\text{eff}}$, we would estimate that $\bar{\Omega}$ ranges between 8.0 and 11.8, hence a very large, unuseful spread that includes seemingly unrealistic values. Likewise, the estimated λ_∞ would range between 0.077 and 8.9 earthquakes per day. In the $\mu \to 0$ scenario (giving $\lambda_\infty = 0.077$ and $\bar{\Omega} = 11.8$), all the observed earthquakes would be aftershocks of long past mainshocks (critical regime, see the Dependence of $\bar{\lambda}$ on μ section). Extra independent knowledge could help better constraining μ, although this information is generally missing or incomplete. We recall that a direct estimate of $\bar{\Omega}$ based on the raw seismicity rate (see equation 1) would yield $8.38 < \bar{\Omega} < 9.60$ if splitting the 2000 yr into 40-year-long windows. This shows that, at least, for the purpose of estimating $\bar{\Omega}$ and in the $p \to 1$ limit, the use of the background rate rather than the raw seismicity rate is questionable: the undersampling issue when taking the observed λ is replaced by an estimation issue for μ (again, in the $p \to 1$ limit). Further work is required to better understand these limits and point to solutions.

CONCLUSIONS

We have defined a seismicity model that can effectively account for earthquake clustering, and which verifies the moment conservation principle at long timescales. This can be done by simply adding the requirement that no earthquake can have a moment greater than the current moment deficit at the time of its occurrence. We have explored how this requirement affects the seismicity. The important consequences are listed as follows:

1. The G-R law observed over long times has a smooth upper cutoff, even though the instantaneous G-R law has a sharp cutoff. This smooth cutoff has a location that can be predicted from the model parameters.
2. The seismicity (ETAS) model becomes stable over the whole parameter space, thanks to the existence of a (time-varying) maximum magnitude.
3. A remarkable regime in which earthquake occurrence is dominated by aftershock triggering can be spontaneously reached by the model, provided that $\mu \to 0$ and $n_0 > n_c$ (this $\mu \to 0$ regime is always found, whatever the value of n_0, if $\alpha \geq \beta$). This regime is characterized by a mean branching ratio that naturally tends to 1, without having to tune it precisely with the model parameters. We recall that a branching ratio of 1 implies that every earthquake on average (i.e., taking the average over all the possible mainshock magnitudes) triggers exactly one aftershock, so that the earthquake activity never dies off and remains permanent over infinitely long durations. This occurs without having to reactivate the system by adding new background earthquakes (as $\mu \to 0$).

Because of this spontaneous $\bar{n} \to 1$ behavior, this regime is a self-organized critical regime.

4. Provided that $\dot{M}_0$ is proportional to the tectonic displacement rate $\dot{\delta}$, and keeping all other parameters fixed, it is expected that the seismicity rate becomes simply proportional to $\dot{\delta}$ over the whole parameter space. This is only valid if $\mu \sim \dot{\delta}$, which, given our current knowledge, appears to be a reasonable assumption. Transient changes in $\dot{\delta}$, or alternatively in seismic coupling, must therefore translate on average into proportional changes in seismicity rate.

Several questions naturally stem from these consequences. First of all, it is still unclear whether the description of seismicity with a clustering model and a background rate brings a significant gain with regard to estimating the mean maximum magnitude, as compared to already existing methods that more simply use the observed (raw) seismicity rate. This issue is potentially severe in the $p \to 1$ limit, as the estimation of the background rate μ becomes challenging.

A daring working hypothesis is that earthquake systems could be in the self-organized critical regime $\bar{n} \to 1$ described earlier. This would much simplify the modeling and estimation as the seismicity would then become insensitive to μ, as long as it remains small enough. This hypothesis will be tested on actual datasets in a future work.

Finally, although the location of the upper cutoff of the G-R law can be predicted from the model parameters, its shape and width cannot. Further work would be required to investigate this issue, which, in terms of seismic hazard estimation, is of importance.

DATA AND RESOURCES

No data were used in this article.

ACKNOWLEDGMENTS

The authors would like to thank the Associate Editor Matthew Gerstenberger and two anonymous reviewers for their suggestions and for allowing us to improve this article. Yen Joe Tan was supported by a Science, Technology, Engineering, Mathematics and Biology-Health (STEM) Chateaubriand Fellowship and the Make our Planet Great Again Initiative.

REFERENCES

Avouac, J. P. (2015). From geodetic imaging of seismic and aseismic fault slip to dynamic modeling of the seismic cycle, *Annu. Rev. Earth Planet. Sci.* **43,** 233–271.

Bird, P., Y. Y. Kagan, D. D. Jackson, F. P. Schoenberg, and M. J. Werner (2009). Linear and nonlinear relations between relative plate velocity and seismicity, *Bull. Seismol. Soc. Am.* **99,** 3097–3113.

Brune, J. N. (1968). Seismic moment seismicity and rate of slip along major fault zones, *J. Geophys. Res.* **73,** 777–784.

Cattania, C. (2019). Complex earthquake sequences on simple faults, *Geophys. Res. Lett.* **46,** 10,384–10,393.

D'Agostino, N. (2014). Complete seismic release of tectonic strain and earthquake recurrence in the Apennines (Italy), *Geophys. Res. Lett.* **41,** 1155–1162.

Dublanchet, P., P. Bernard, and P. Favreau (2013). Interactions and triggering in a 3-D rate-and-state asperity model, *J. Geophys. Res.* **118,** 2225–2245.

Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.* **86,** 2825–2852.

Ekström, G., M. Nettles, and A. M. Dziewonski (2012). The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. In.* **200/201,** 1–9.

Guerin-Marthe, S., S. Nielsen, R. Bird, S. Giani, and G. Di Toro (2019). Earthquake nucleation size: Evidence of loading rate dependence in laboratory faults, *J. Geophys. Res.* **124,** 689–708.

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

Hainzl, S., O. Zakharova, and D. Marsan (2013). Impact of aseismic transients on the estimation of aftershock productivity parameters, *Bull. Seismol. Soc. Am.* **103,** 1723–1732.

Hanks, T. C., and H. Kanamori (1979). Moment magnitude scale, *J. Geophys. Res.* **84,** 2348–2350.

Harte, D. S. (2018). Effect of sample size on parameter estimates and earthquake forecasts, *Geophys. J. Int.* **214,** 759–772.

Hayes, G. P., D. J. Wald, and R. L. Johnson (2012). Slab1.0: A three-dimensional model of global subduction zone geometries, *J. Geophys. Res.* **117,** doi: 10.1029/2011JB008524.

Helmstetter, A., and D. Sornette (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, *J. Geophys. Res.* **107,** doi: 10.1029/2001JB001580.

Helmstetter, A., and D. Sornette (2003). Importance of direct and indirect triggered seismicity in the ETAS model of seismicity, *Geophys. Res. Lett.* **30,** doi: 10.1029/2003GL017670.

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

Kagan, Y. Y. (2002a). Seismic moment distribution revisited: I. Statistical results, *Geophys. J. Int.* **148,** 520–541.

Kagan, Y. Y. (2002b). Seismic moment distribution revisited: II. Moment conservation principle, *Geophys. J. Int.* **149,** 731–754.

Knopoff, L., and Y. Kagan (1977). Analysis of theory of extremes as applied to earthquake problems, *J. Geophys. Res.* **82,** 5647–5657.

Kostrov, V. V. (1974). Seismic moment and energy of earthquakes, and seismic flow of rock, *Izv. Earth Phys.* **1,** 23–40.

Lengliné, O., and D. Marsan (2009). Inferring the coseismic and post-seismic stress changes caused by the 2004 $M_w = 6$ Parkfield earthquake from variations of recurrence times of microearthquakes, *J. Geophys. Res.* **114,** doi: 10.1029/2008JB006118.

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.

Marsan, D., M. Bouchon, B. Gardonio, H. Perfettini, A. Socquet, and B. Enescu (2017). Change in seismicity along the Japan trench, 1990-2011, and its relationship with seismic coupling, *J. Geophys. Res.* **122,** 4645–4659.

Marsan, D., E. Prono, and A. Helmstetter (2013). Monitoring aseismic forcing in fault zones using earthquake time series, *Bull. Seismol. Soc. Am.* **103,** 169–179.

McCaffrey, R. (1994). Dependence of earthquake size distributions on convergence rates at subduction zones, *Geophys. Res. Lett.* **21,** 2327–2330.

McCaffrey, R. (1997). Statistical significance of the seismic coupling coefficient, *Bull. Seismol. Soc. Am.* **87,** 1069–1073.

Mclaskey, G. C., and F. Yamashita (2017). Slow and fast ruptures on a laboratory fault controlled by loading characteristics, *J. Geophys. Res.* **122,** 3719–3738.

Molnar, P. (1979). Earthquake recurrence intervals and plate tectonics, *Bull. Seismol. Soc. Am.* **69,** 115–133.

Nadeau, R. M., and T. V. McEvilly (1999). Fault slip rates at depth from recurrence intervals of repeating microearthquakes, *Science* **285,** 718–721.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.* **83,** 9–27.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.* **50,** 379–402.

Pacheco, J. F., L. R. Sykes, and C. H. Scholz (1993). Nature of seismic coupling along simple plate boundaries of the subduction type, *J. Geophys. Res.* **98,** 14,133–14,159.

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

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

Scholz, C. H., and J. Campos (2012). The seismic coupling of subduction zones revisited, *J. Geophys. Res.* **117,** doi: 10.1029/2011JB009003.

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

Sornette, D., and A. Sornette (1999). General theory of the modified Gutenberg-Richter law for large seismic moments, *Bull. Seismol. Soc. Am.* **89,** 1121–1130.

Stevens, V. L., and J. P. Avouac (2017). Determination of M_{max} from background seismicity and moment conservation, *Bull. Seismol. Soc. Am.* **107,** 2578–2596.

van der Elst, N. J. (2017). Accounting for orphaned aftershocks in the earthquake background rate, *Geophys. J. Int.* **211,** 1108–1118.

Wang, Q., P. P. Schoenberg, and D. D. Jackson (2010). Standard errors of parameter estimates in the ETAS model, *Bull. Seismol. Soc. Am.* **100,** 1989–2001.

Zhuang, J., M. J. Werner, S. Hainzl, D. Harte, and S. Zhou (2011). Basic models of seismicity: Spatiotemporal models, *Community Online Resource for Statistical Seismicity Analysis*, doi: 10.5078/corssa-07487583.

APPENDIX A

Conservation of moment

We here show that the existence of a cutoff magnitude Ω_t, even if it is time-varying, imposes that seismic moment is conserved, that is, $\bar{\lambda} \times \bar{M}_0 = \dot{M}_0$, in which the $\bar{\cdot}$ symbol denotes long-term averaged quantities. We recall that this conservation is not *a priori* imposed in the model (we only constrain every earthquake to have a magnitude less than Ω_t, in which $\Delta M_0(t) = 10^{1.5\Omega_t + 9.1}$ is the moment deficit at time t).

The mean seismic moment $\overline{M_{0,t}}$ for an earthquake drawn at random from the strictly bounded density of equation (3) is

$$\overline{M_{0,t}} = \int_{m_0}^{\Omega_t} dm\, M_0(m) f_t(m) = M_0(m_0) \frac{\beta}{\beta - \gamma} \times \frac{e^{(\gamma-\beta)(\Omega_t - m_0)} - 1}{e^{-\beta(\Omega_t - m_0)} - 1}, \tag{A1}$$

in which $\gamma = 1.5 \log 10$. This mean depends on time through Ω_t. As already commented, when $\beta < \gamma$, that is, $b < 1.5$, then $\lim_{\Omega_t \to \infty} \overline{M_{0,t}} = \infty$. Suppose that the model is run from $t = 0$ to $t = T_{\text{obs}}$. Then, because the moment deficit stays finite, the difference $\frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt \lambda_t \overline{M_{0,t}} - \dot{M}_0$ must tend to 0 as $T_{\text{obs}} \to \infty$. Approximating the integral by $\frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt \lambda_t \overline{M_{0,t}} \simeq \frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt \lambda_t \times \frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt \overline{M_{0,t}} = \bar{\lambda} \times \bar{M}_0$, then this product $\bar{\lambda} \times \bar{M}_0 \to \dot{M}_0$ for infinitely long T_{obs}, so that moment is conserved over sufficiently long time periods.

Estimation of $\bar{\Omega}$

We further approximate $\bar{M}_0$ as $M_0(m_0) \frac{\beta}{\beta - \gamma} \times \frac{e^{(\gamma-\beta)(\bar{\Omega} - m_0)} - 1}{e^{-\beta(\bar{\Omega} - m_0)} - 1}$, that is, equation (A1) with Ω_t replaced by the time-averaged $\bar{\Omega} = \frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt \Omega_t$. This $\bar{\Omega}$ can be estimated directly from the model parameters: the mean rate being $\bar{\lambda} = \frac{\mu}{1 - \bar{n}}$ (Helmstetter and Sornette, 2003), with $\bar{n} = \frac{1}{T_{\text{obs}}} \int_0^{T_{\text{obs}}} dt\ \bar{n}_t \simeq \frac{n_0 \beta}{\beta - \alpha} \times \frac{e^{(\alpha-\beta)(\bar{\Omega} - m_0)} - 1}{e^{-\beta(\bar{\Omega} - m_0)} - 1}$, the condition $\bar{\lambda} \times \bar{M}_0 = \dot{M}_0$ provides a closure on $\bar{\Omega}$, as both $\bar{\lambda}$ and $\bar{M}_0$ depend on model parameters $\{\mu, \alpha, \beta, n_0, m_0, p, c\}$ and $\bar{\Omega}$. We effectively must verify a closed set of four equations, with $\bar{\lambda}$, $\bar{M}_0$, $\bar{n}$, and $\bar{\Omega}$ the four unknowns:

$$\bar{\lambda} \times \bar{M}_0 = \dot{M}_0 \tag{A2}$$

$$\bar{\lambda} = \frac{\mu}{1 - \bar{n}} \tag{A3}$$

$$\bar{n} = n_0 \frac{\beta}{\beta - \alpha} \times \frac{1 - e^{(\alpha-\beta)(\bar{\Omega} - m_0)}}{1 - e^{-\beta(\bar{\Omega} - m_0)}} \tag{A4}$$

$$\bar{M}_0 = M_0(m_0) \frac{\beta}{\beta - \gamma} \times \frac{1 - e^{(\gamma-\beta)(\bar{\Omega} - m_0)}}{1 - e^{-\beta(\bar{\Omega} - m_0)}}. \tag{A5}$$

All other parameters (μ, n_0, α, β, m_0, $\dot{M}_0$) are (fixed) model parameters. The numerical estimation of $\bar{\Omega}$ is thus straightforward. Note that approximating $\bar{M}_0$ as $M_0(m_0) \frac{\beta}{\gamma - \beta} e^{(\gamma-\beta)(\bar{\Omega} - m_0)}$,

which is valid if $\gamma > \beta$ and $\bar{\Omega} \gg m_0$, implies that the slip budget $\lambda \bar{M}_0 = \dot{M}_0$ directly leads to equation (1).

Departure from the Gutenberg–Richter (G-R) law at large magnitudes

Denoting by Ω_i the maximum magnitude at the time t_i of the ith earthquake, that is, $\Omega_i = \Omega_{t_i}$, the probability that any earthquake among N observed events has a magnitude greater than M is

$$P(m > M) = \frac{1}{N} \sum_{i=1}^{N} P(m_i > M | \Omega_i > M) \tag{A6}$$

$$= \frac{1}{N} \sum_{i=1/\Omega_i > M}^{N} \frac{e^{-\beta(\Omega_i - m_0)} - e^{-\beta(M-m_0)}}{e^{-\beta(\Omega_i - m_0)} - 1}. \tag{A7}$$

In real data, Ω_t is unknown, thus the question remains as to how the observed cutoff on earthquake magnitudes could carry information on the distribution of Ω_t. Calling $g(\Omega)$ the density of Ω_t, equation (A7) can be rewritten as

$$P(m > M) = \int_M^{\infty} d\Omega \; g(\Omega) \frac{e^{-\beta(\Omega - m_0)} - e^{-\beta(M-m_0)}}{e^{-\beta(\Omega - m_0)} - 1}, \tag{A8}$$

and the corresponding density $f(M)$ on magnitudes is thus

$$f(M) = -\frac{P(m > M)}{dM} = -\int_M^{\infty} d\Omega \; g(\Omega) \frac{\beta e^{-\beta(M-m_0)}}{e^{-\beta(\Omega - m_0)} - 1}, \tag{A9}$$

or equivalently

$$f(M) = \beta e^{-\beta(M-m_0)} \times \int_M^{\infty} d\Omega \; \frac{g(\Omega)}{1 - e^{-\beta(\Omega - m_0)}} \tag{A10}$$

$$= \beta e^{-\beta(M-m_0)} \times \Phi(M), \tag{A11}$$

in which $\Phi(M)$ is a cutoff function continuously decreasing from 1 (at small M) to 0 (at large M). Its derivative is

$$\frac{d\Phi(M)}{dM} = \frac{-g(M)}{1 - e^{-\beta(M-m_0)}}, \tag{A12}$$

which, when $\beta(M - m_0) \gg 1$, gives $\frac{d\Phi(M)}{dM} \simeq -g(M)$. This proves that the modulation $\Phi(M)$ of the G-R law at large magnitudes is directly related to the integral of the density of Ω_t. Constraining this modulation from the data should therefore help estimating $g(\Omega)$. It must be noted that equation (A12) relates two distributions; to infer $g(\Omega)$ from $\Phi(M)$, the latter must be well constrained by the data, which requires a sufficiently long observation span.

APPENDIX B

We here detail how the dependence of $\bar{\lambda}$ on μ can be calculated from the set of equations (A2)–(A5). We will make the simplifying assumption that $e^{-\beta(\bar{\Omega} - m_0)} \ll 1$ and $e^{(\gamma-\beta)(\bar{\Omega} - m_0)} \gg 1$, which is true if $\bar{\Omega}$ is sufficiently greater than m_0 and b is not too close to 1.5.

Starting with equation (A3), we have that $\bar{\lambda} - \bar{n}\bar{\lambda} = \mu$. Using equation (A4), this can be rewritten as

$$\bar{\lambda} - n_0 \frac{\beta}{\beta - \alpha} (1 - e^{(\alpha-\beta)(\bar{\Omega} - m_0)}) \bar{\lambda} = \mu. \tag{B1}$$

Equations (A2) and (A5) can be combined to find that

$$e^{(\bar{\Omega} - m_0)} = \left(\frac{\dot{M}_0}{M_0(m_0)} \times \frac{\gamma - \beta}{\beta} \times \frac{1}{\bar{\lambda}} \right)^{\frac{1}{\gamma - \beta}}. \tag{B2}$$

Inserting this in equation (B1), we finally get that

$$\mu = A\bar{\lambda} + B\bar{\lambda}^{\delta}, \tag{B3}$$

with

$$A = 1 - \frac{n_0}{n_c}, \tag{B4}$$

which can be either positive or negative

$$B = \frac{n_0}{n_c} \left(\frac{\dot{M}_0}{M_0(m_0)} \times \frac{\gamma - \beta}{\beta} \right)^{1-\delta}, \tag{B5}$$

so that $\beta > \alpha$ implies $B > 0$, and

$$\delta = 1 - \frac{\alpha - \beta}{\gamma - \beta}, \tag{B6}$$

so $\delta > 1$ for $\beta > \alpha$. The sign of A is conditioned on whether $n_0 < n_c$ or not. We recall that the critical value $n_c = 1 - \frac{\alpha}{\beta}$ is equal to the stability threshold in classical ETAS models, so in these models $n_0 > n_c$ is not allowed.

In the $\alpha > \beta$ case, these results remain valid, but we notice that $n_c = 1 - \frac{\alpha}{\beta}$ is then negative, so that n_0 is always greater than n_c: the only regime is the self-organized critical one. Noting that $A < 0$, we get for $\mu = 0$ that equation (B3) yields $|A|\bar{\lambda} = B\bar{\lambda}^{\delta}$; replacing A and B by their expressions, we obtain that

$$\bar{\lambda} = \left(1 - \frac{n_c}{n_0}\right)^{\frac{1}{\delta - 1}} \frac{\dot{M}_0}{M_0(m_0)} \times \frac{\gamma - \beta}{\beta}. \tag{B7}$$

Manuscript received 5 August 2019
Published online 3 March 2020