Is the Earth crust operating at a critical point?

Shyam Nandan¹, Sumit Kumar Ram², Guy Ouillon³ and Didier Sornette²,⁴

¹Swiss Seismological Service, ETH Zürich, Sonneggstrasse 5, 8092, Zürich, Switzerland
e-mail address: shyam4iiser@gmail.com
²Department of Management, Technology and Economics, ETH Zürich, Scheuchzerstrasse 7, 8092, Zürich, Switzerland
³Lithophyse, 4 rue de l’Ancien Sénat, 06300 Nice, France
⁴Institute of Risk Analysis, Prediction and Management (Risks-X), Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen, China

Seismicity and faulting within the Earth crust are characterized by many scaling laws that are usually interpreted as qualifying the existence of underlying physical mechanisms associated with some kind of criticality in the sense of phase transitions. Using an augmented Epidemic-Type Aftershock Sequence (ETAS) model that accounts for the spatial variability of the background rates \( \mu(x, y) \), we present a direct quantitative test of criticality. We calibrate the model to the ANSS catalog of the entire globe, the region around California, and the Geonet catalog for the region around New Zealand using an extended Expectation-Maximization (EM) algorithm including the determination of \( \mu(x, y) \). We demonstrate that the criticality reported in previous studies is spurious and can be attributed to a systematic upward bias in the calibration of the branching ratio of the ETAS model, when not accounting correctly for spatial variability. We validate the version of the ETAS model which possesses a space varying background rate \( \mu(x, y) \) by performing pseudo prospective forecasting tests. The non-criticality of seismicity has major implications for the prediction of large events.

Earthquakes have fascinated and continue to capture the imagination and interest of physicists, as they express how the huge, unbridled forces of Nature can be organized according to remarkable regular statistical laws obeying power-law statistics. The late Per Bak, one of the fathers of the concept of “self-organized criticality,” was fond of exclaiming in his talks: “I love this law,” when referring to (i) the Gutenberg-Richter (GR) distribution of earthquake seismic moments because it is valid over several decades larger than most known power laws in physical and social sciences. Several other scaling laws further characterize seismicity: (ii) the Omori law (the rate of aftershocks decays as \( \frac{1}{(t-t_m)^\gamma} \) after a main shock that occurred at a time \( t_m \)), (iii) a spatial Green function quantifying the power-law decay of the influence of the main shock as a function of the distance to its aftershocks, (iv) a power selection law of the average number of aftershocks triggered as a function of the magnitude of the main shock, (v) power-law distributions of the lengths of the faults on which earthquakes occur, (vi) fractal, multifractal or hierarchical scaling of the set of earthquake epicenters as well as fault networks, and so on. For physicists, these laws suggest the existence of underlying physical mechanisms associated with some kind of criticality in the sense of phase transitions and field theory with zero mass. Indeed, many proposals in this spirit have been put forward to rationalize these power laws: self-organized criticality [1, 2], critical point behavior before large earthquakes [3-8], based on accelerating seismic release [9-12], combinations of the two [13-16].

The evidence for the criticality of the Earth’s crust is thus generally inferred from the presence of scale invariance and power-law scaling. But it is well-known that many other mechanisms can be at the origin of power-law scaling without the need for invoking criticality [17, 18]. In this context, the class of Epidemic-Type Aftershock Sequence (ETAS) models offer a direct path to calibrate and quantify the distance to the criticality of earthquake catalogs. ETAS models combine the scaling laws (i)-(iv) to formulate a statistical framework that reproduces many of the statistical features observed in seismicity catalogs and is often used as the benchmark to assess the skills of competing earthquake forecasting models [19]. The ETAS models have an exact branching process representation [20], in which background earthquakes supposed to be driven by the forces of plate tectonics (immigrants in the language of epidemic branching processes) can trigger cohorts of earthquakes (first-generation “daughters”), themselves triggering second-generation events and so on [21]. The ETAS models exhibit a transition determined by the control parameter \( n \), called the branching ratio, which is both the average number of triggered events of first-generation per immigrant and the fraction of all triggered events of any generation in the catalog [22]. The transition occurring at the critical value \( n = 1 \) separates the sub-critical regime \( (n < 1) \), where the sequence of events is stationary, from the super-critical regime \( (n > 1) \) for which the number of earthquakes explodes exponentially with time with a
finite probability [21]. The transition at \( n = 1 \) is characterized by the standard signatures of criticality, such as diverging rate \( \mu/(1-n) \) of events where \( \mu \) is the background rate, nonlinear “susceptibility” at \( n = 1 \) [23] and various power-law statistics [24-27]. A lot of attention has been devoted to determining the empirical value of \( n \), with most studies suggesting that \( n \) is very close to 1 [28, 29]. In contrast, some other studies averaging over broad tectonic areas find lower numbers in the range 0.35 to 0.65 [30]. If \( n \approx 1 \), a significant portion of the Earth’s crust would then function close to or even exactly at criticality, so that previous earthquakes trigger most observed earthquakes. If confirmed, this has far-reaching consequences for modeling, predicting and managing seismic risks.

Within the representation offered by ETAS models, determining whether the Earth crust is at criticality or not is crucially dependent on the ability to partition the observed seismicity clustering in space and time between the background rates and the triggered rates resulting from previous earthquakes. For instance, a misclassification of earthquakes as background events due to incomplete catalogs biases the calibrated branching ratio \( n \) downward [31, 32]. On the other hand, failing to account for spatial and temporal variations of the background rates leads to upward biases for \( n \), as the variability may be incorrectly attributed to triggering. This effect has been recently demonstrated in the time domain for financial time series [33, 34]. Here, we revisit this question of the criticality of the Earth’s crust by offering an augmented ETAS model that improves on state of the art by accounting self-consistently for the possible spatial dependence of the key parameters.

Before presenting our method and results, we briefly review previous related studies. [35] and several related studies [36-38] have modeled the spatial dependence of the background rate \( \mu(x,y) \) using a kernel density estimation. Generally, these studies use gaussian kernels with adaptive bandwidth. These kernel density estimates are obtained iteratively using the algorithm proposed by [35]. However, for all practical purposes, these studies fix the parameter on which the adaptive bandwidth of kernels depends, stating that the choice of this crucial parameter, which dictates the smoothing of background intensity in space, is relatively unimportant. Other studies such as those of [30] and [39] pre-delineate regions in which all the parameters are assumed to be uniform and the parameters are then individually or jointly inferred in each of these regions. In some other studies, such as [40], the spatial probability density function (PDF) of background earthquakes is pre-estimated and then fed into the ETAS inversion machinery. Furthermore, [41, 29, 42] infer the optimal space variation of all ETAS parameters jointly, while more recently, [43] have proposed an alternative approach to non-parametrically model the space variation of background rate using a Gaussian process prior.

Here, we extend the iterative algorithm of [35] to jointly estimate the ETAS parameters and the kernel parameters used for obtaining the spatially variable background rate using the Expectation-Maximization (EM) algorithm of [39]. This extended algorithm leads to an augmented ETAS model that accounts specifically for the space variation of \( \mu(x,y) \). Using this model, we show that the value of the branching ratio and other triggering parameters depend crucially on how the background rate is modeled. We demonstrate, using synthetic tests, that ignoring the spatial variation of the background rate leads to estimated parameters that are highly biased. However, when the spatial variation of the background rate is accounted for, the calibrated parameters are close to their true values and unbiased. We also document this bias effect in the case of three real catalogs. We then perform rigorous pseudo-prospective experiments and show that the ETAS model with spatially varying background rates significantly overperforms the ETAS model with uniform background rates. To the best of our knowledge, such direct comparisons have never been made between these two models. This gives us confidence that the extended ETAS model is superior to its spatially invariant counterpart, and thus reveals genuine characteristics of seismicity.

### Self-consistent estimation of spatially variable background rate using an extended EM algorithm

Our implementation of the ETAS model expresses the seismicity rate \( \lambda(t,x,y|H_t) \), at any time \( t \) and location \( (x,y) \) and conditional upon the history \( H_t \) of seismicity up to \( t \), as

\[
\lambda(t,x,y|H_t) = \mu(x,y) + \sum_{i \leq t} g(t-t_i, x-x_i, y-y_i, m_i)
\]

\( g(t-t_i, x-x_i, y-y_i, m_i) = \frac{K \exp[a(m_i - M_c)](t-t_i + c)^{-1-\omega} e^{-\frac{t-t_i}{\tau}} \times T_{norm} \times S_{norm}}{((x-x_i)^2 + (y-y_i)^2 + d \exp[y(m_i - M_c)])^{1+p}} \)

This expression combines the fertility law \( P(m) = K \exp[a(m_i - M_c)] \) that quantifies the expected number of first-generation aftershocks (\( \geq M_c \)) triggered by an earthquake with magnitude \( m \), the Omori-Utsu law \( (t-t_i + c)^{-1-\omega} e^{-\frac{t-t_i}{\tau}} \) and the spatial Green function, leading to the set \( \phi = \{ \mu, K, a, c, \omega, \tau, d, y, \rho \} \) of parameters that characterize the ETAS model. \( T_{norm} \) and \( S_{norm} \) are normalization constants for the time and space kernels, ensuring that they are proper probability density functions (PDF).

With this formulation, the branching ratio is

\[
n = \int_{M_{c}}^{M_{max}} P(m) \times f(m) \]

defined as the expected number of aftershocks of the first generation triggered by an earthquake, averaged over all magnitudes. The averaging over magnitude is thus performed using the GR distribution \( f(m) = b \ln(10) 10^{b - bm} \), \( \forall M_c \leq m \leq M_{max} \). Denoting \( \alpha = a/\ln 10 \), this yields

\[
n = \begin{cases} K b \left( 1 - 10^{-(b-a)(M_{max}-M_c)} \right) & \text{if} \; \alpha \neq b \\ (b-a)(1-10^{-b(M_{max}-M_c)}) & \text{if} \; \alpha = b \\ K b \ln(10) (M_{min} - M_c) & \end{cases}
\]

We consider two variants of the ETAS model: \( ETAS_\mu \), which features a spatially uniform background rate \( \mu \), and \( ETAS_{\mu(x,y)} \).
which possesses a space varying background rate $\mu(x, y)$. In ETAS$_{\mu(x,y)}$, $\mu(x, y)$ is informed by the spatial positions of previous earthquakes. The proposed parameterization given below in Eq. (5) should not be confused with the triggering part of the ETAS model given by the second term in the r.h.s. of Eq. (1), which also involves a summation over previous earthquakes. Here, the guiding idea is that observed earthquakes occur more frequently where the background intensity is larger because the background events are, by definition, the sources of all observed seismicity. Even if many earthquakes are triggered by previous earthquakes, their locations are related to that of their background sources [44, 22].

To estimate $\mu(x, y)$, we extend the EM algorithm proposed by [39]. In general, the EM algorithm allows us to obtain the so-called independence probability $P_i$ for each earthquake $i$, defined as the probability that it is a background event (and thus has not been triggered by any previous earthquake). We smooth out its contribution to $\mu(x, y)$ according to a regularized inverse power law function of the distance between that earthquake and the point $(x, y)$ of interest, and sum over all earthquakes in the training catalog:

$$\mu(x, y) = T^{-1} \sum_{i=1}^{N} P_i \pi^{-1} Q D^Q (x - x_i)^2 + (y - y_i)^2 + D^2)^{1-q} (5)$$

The normalization by the duration $T$ of the training catalog ensures that $\mu(x, y)$ represents the seismicity rate per unit time. The factor $\pi^{-1}Q D^Q$ ensures the normalization of $\mu(x, y)$ per unit area at the location $(x, y)$. The choice of the power-law kernel is motivated by the findings of [45] that power-law kernels are superior to the more commonly used Gaussian kernels. This functional form (5) of $\mu(x, y)$ adds two free parameters $(D, Q)$ to the set $\phi$, which are estimated along with $\phi$ in an extended EM algorithm. This improves on [35, 36] by embedding the determination of $\mu(x, y)$ into the EM method, thus making it entirely self-consistent and optimal. This extended algorithm is described in details in Text S1 of the supplementary material (SM).

The power-law kernel used in Eq. (5) features a non-adaptive bandwidth. However, the algorithm proposed in Text S1 can easily be extended for the adaptive bandwidth case with the parameter to optimize being the number of nearest neighbors instead of the non-adaptive bandwidth $D$.

**Dataset.** We apply the extended EM algorithm (Text S1 of SM) to catalogs obtained from two sources: the Advanced National Seismic System (ANSS) and GeoNet. The ANSS catalog is used for two study regions: the entire globe and the region around California. For the area around New Zealand, we use the GeoNet catalog. The location of ∼600,000 earthquakes ($M \geq 3$) for the entire globe between 1975-2020, as reported in the ANSS catalog, and of ∼1.2 million and ∼600,000 earthquakes ($M \geq 1$) between 1975-2020 in the study regions surrounding California and New Zealand, as reported by the ANSS and Geonet catalogs respectively, are shown in Figures S1-3 of the SM. Text S2 in SM also presents the method to select the magnitude of completeness $M_c$ for each catalog.

**Parameter calibration.** The results of the calibration of the two ETAS models, ETAS$_{\mu}$ and ETAS$_{\mu(x,y)}$, on the global, Californian, and the New Zealand catalogs are presented in Table 1. When going from the ETAS$_{\mu}$ model to the ETAS$_{\mu(x,y)}$ model, the most remarkable changes are that (i) the overall background rate increases by nearly 29, 3, and 16 times for the three catalogs, respectively, while consequently (ii) the branching ratio $n$ is substantially smaller for the ETAS$_{\mu(x,y)}$ model compared to the ETAS$_{\mu}$ model. The other parameters also show substantial changes. More specifically, the branching ratio is remarkably close to 1 for the three catalogs when calibrated with the ETAS$_{\mu}$ model, which would lead to the erroneous conclusion that the Earth crust is critical, as discussed in the introduction. In contrast, the ETAS$_{\mu(x,y)}$ model gives $n$ = 0.4, 0.8, and 0.6 for the global, Californian, and the New Zealand catalogs, respectively, clearly excluding criticality and qualifying the Earth crust in the sub-critical regime. The difference between the two spatial intensity of background earthquakes inferred from the calibrations of the two models is vividly illustrated in Figure 1(a and b) in the case of the global catalog (and Figures S4 and S5 in the SM for the Californian and New Zealand catalogs). We find that not only is the overall background rate is different between the two models, but also that the spatial patterns of the density of background earthquakes differ substantially between the two modeling choices, as one can observe the much more refined localization of the background rates along plate boundaries in ETAS$_{\mu(x,y)}$.

**Synthetic tests of the bias in branching ratio $n$ due to uniform background rate.** We now demonstrate that a realistic synthetic catalog with a relatively smaller branching ratio as found with the full ETAS$_{\mu(x,y)}$ model calibrated on the real catalogs yields a spuriously large branching ratio when calibrated with the ETAS$_{\mu}$ model, and recovers the correct value when calibrated with the full ETAS$_{\mu(x,y)}$ model. For this, using the full ETAS$_{\mu(x,y)}$ model, we simulate a 50-year long synthetic catalog [46] with earthquakes of magnitudes $M \geq 5$ for the entire globe using the parameters that are listed in Table 1 corresponding to ETAS$_{\mu(x,y)}$. Using the first ten years of the catalog as the auxiliary period and the remaining as the primary period [47], we calibrate the ETAS$_{\mu}$ model on this synthetic catalog. The obtained parameters are: $N_{bg} = 22.27$ year$^{-1}$, $\log_{10} K = -0.18$, $a = 1.07$, $\log_{10} d = 2.17$, $\rho = 0.73$, $\gamma = 0.78$, $\log_{10} c = -3.09$, $\omega = 0.80$, $\log_{10} T = 3.9$, $n = (1.17, 0.98, 1.11)$ (see Table 1). This should be compared with the true input parameters $N_{bg} = 1.013.2$ year$^{-1}$ for the background rate and $n = (0.45, 0.39, 0.41)$ for the branching ratio obtained with the theoretical, empirical, and semi-empirical methods (see Table 1 for definition). The background rate inverted with the ETAS$_{\mu}$ model is too low, and the inferred branching ratio is too large, being remarkably close to the value inferred by inverting this model with uniform background rate $\mu$ on the real catalog, on nearly all the parameters. For instance, the Omori exponents inferred for the synthetic catalog and the real catalog are 0.80 and 0.79, respectively. This provides an excellent self-consistent test and further supports the validity of our hypothesis that the background rate $\mu(x, y)$ is strongly varying in space. This suggests that it is absolutely essential to account for the non-
uniform background rate to obtain unbiased parameter estimates. These synthetic tests also show that the biases are predictable.

Pseudo-prospective forecasting experiments as a validation step for $\text{ETAS}_\mu(x,y)$. We now proceed to show that $\text{ETAS}_\mu(x,y)$ leads to operationally better forecasts of future seismic activity by setting up pseudo-prospective forecasting experiments. In these experiments, we use the early part of the data to calibrate the models and leave future data unseen. Subsequently, we use the calibrated models to forecast the future and use the left out future data to evaluate the performance of the models. Starting on January 1, 1990, we perform 368, 368, and 371 pseudo prospective experiments for the three catalogs. For details on these pseudo-prospective experiments, we refer the readers to Text S3 in the SM. For more general discussions on the importance of these experiments and their design, we refer the readers to [46, 47].

At a given spatial resolution and magnitude threshold ($M_i$), the log-likelihood score of a model during a given testing period is defined as $LL = \sum_{i=1}^{N} \ln P(n_i)$, where $P(n_i)$ is the probability of observing $n_i$ earthquakes in the $i^{th}$ pixel during the testing period, and $N$ is the total number of equal area pixels that tile the study region. In any pixel, the probability is constructed using the number of earthquakes observed in different simulated catalogs as described in Text S3 in the SM.

Once the likelihoods for two models are calculated, the information gain of one over the other is simply the difference of their log-likelihoods. Figure 2(a-b) shows the time-series of cumulative information gain (CIG) that the $\text{ETAS}_\mu(x,y)$ model obtains over the $\text{ETAS}_\mu$ model in the 368 experiments that we perform with the global catalog (see similar figures S6-8 of the SM for the time series of CIG at different spatial resolutions and $M_t$ for Global, Californian and New Zealand catalogs). At all spatial resolutions and magnitude thresholds, the $\text{ETAS}_\mu(x,y)$ substantially outperforms the $\text{ETAS}_\mu$ model for all the three study regions.

To quantify if this over-performance of the $\text{ETAS}_\mu(x,y)$ is statistically significant, we define the mean information gain as the average information that $\text{ETAS}_\mu(x,y)$ obtains over the 368 testing periods. We then test the null hypothesis that this mean information gain is significantly larger than 0 against the alternative that it is not, using the student’s t-test. Figure 2(c-d) (and similar figures S9-11 for all spatial resolutions and $M_t$ for the Global, Californian, and New Zealand catalogs) confirm that the mean information gain of the $\text{ETAS}_\mu(x,y)$ over $\text{ETAS}_\mu$ is significantly larger than 0, thus rejecting the null hypothesis at significance levels much smaller than 0.01 in all cases.

Discussion and Conclusions. Our results demonstrate that, when the spatial variation of background rate is appropriately accounted for, we get a superior forecasting model and a branching ratio that is much smaller than 1. This lower value of the branching ratio $\eta$ reported here can be rationalized by the long-term localization of seismicity at different scales along faults and plate boundaries. This localization process leads to a highly intermittent spatial distribution of the possible locations of earthquakes. Thus, even pairs of time-independent events occur on highly clustered spatial structures, which have slowly grown and organized over tens of millions of years. In return, this spatial localization increases the probability for each event to be clustered around each other, i.e., to be (spatially and apparently) dependent on each other, leading to a substantial underestimation of the actual rate of independent (background) events in models not accounting for spatial variability.

At a more fundamental level, the small value of the branching ratio $\eta$ invites a reexamination of the physical picture we have of the brittle rupture process in the Earth’s crust. Until now, large values of $\eta$ close to unity have been reported, suggesting the loaded fault network to be in a permanent critical state, compatible with the popular concept of self-organized criticality. The much lower value of $\eta$ that we estimate using more appropriate assumptions and a superior algorithm suggests that fault networks mainly evolve far from a critical point. This has major implications for the prediction of large events. Indeed, in the self-organized critical scenario, each event is no different from all others from a generating process viewpoint, making prediction impossible [48]. In contrast, if the fault network remains most of the time far from criticality, more sporadic singularities may appear via various possible mechanisms and announce upcoming catastrophic events. This suggests, for instance, the need for a reevaluation of the Accelerated Moment Release hypothesis [11], benefitting from the prior use of the $\text{ETAS}_\mu(x,y)$ fitting model to better eliminate the contribution of the uncorrelated part of seismicity to the total seismicity rate.
two magnitude thresholds (different panels) of the testing catalog. The error bars indicate the 99 and 95 % confidence interval of the mean information gain (MIG). The two numbers indicate the p-value and t-statistic resulting from the student’s t-test, in which we test the null hypothesis that the MIG is equal to 0 against the alternative that it is larger than 0. When the p-value is smaller than 0.05, the null hypothesis is rejected.

Table 1: ETAS parameters inverted for three catalogs (first column) using the two models ETAS_µ and ETAS_µ(x,y). The branching ratio n is inferred in three ways: theoretically, empirically, and semi-empirically. The theoretical estimate of the branching ratio is obtained using Eq.(4) with M_{max} = 10, 8.5 and 9 for the Global, Californian, and New Zealand catalog, respectively. The empirical estimate is obtained as the fraction of earthquakes identified as triggered events by summing independence probabilities. The semi-empirical estimate is obtained by computing the average number of earthquakes triggered from the triggering probability matrix, corrected for the finite duration of the catalog using the inverted parameters of the time kernel.

| Catalog          | Model type | N_{bkg} (year^{-1}) | \log_{10} \gamma | \log_{10} d (km²) | \rho | \gamma | \log_{10} \delta (days) | 1 + \omega | \log_{10} \delta | Branching ratio n | \text{D} (km) |
|------------------|------------|---------------------|------------------|------------------|-----|--------|------------------------|-------------|------------------|------------------|-------------|
| Globe            | ETAS_µ     | 35.35               | -0.12            | 0.8              | 2.20| 0.7    | 0.5                    | -3.16       | 0.7              | 3.81             | 1.15        |
|                  | ETAS_µ(x,y)| 1,013.2             | -0.72            | 1.3              | 1.93| 1.0    | 0.9                    | -2.08       | 1.0              | 3.41             | 0.45        |
| California       | ETAS_µ     | 60.73               | -0.22            | 0.9              | 0.59| 0.5    | 1.1                    | -2.94       | 0.9              | 3.87             | 1.00        |
|                  | ETAS_µ(x,y)| 182.42              | -0.37            | 1.1              | 0.73| 0.5    | 1.2                    | -2.64       | 1.0              | 3.67             | 0.79        |
| New Zealand      | ETAS_µ     | 9.08                | -0.17            | 1.0              | 1.60| 0.7    | 0.0                    | -3.01       | 0.8              | 3.76             | 1.14        |
|                  | ETAS_µ(x,y)| 149.00              | -0.58            | 1.5              | 1.17| 0.7    | 0.4                    | -2.28       | 1.0              | 3.37             | 0.61        |

[1] P. Bak and C. Tang., J. Geophys. Res. 94 (B11), 15635, 1989.
[2] A. Sornette and D. Sornette, Europhys. Lett. 9 (3), 197, 1989.
[3] D. Sornette and C. G. Sammis, J. Phys. I 5 (5),607, 1995.
[4] V. I. Keilis-Borok. PEPI, 111 (3-4), 179, 1999.
[5] D. S. Fisher, K. Dahmen, S. Ramanathan, and Y. Ben-Zion. Phys. Rev. Lett., 78 (25), 4885, 1997.
[6] G. Zöller, S. Hainzl, and J. Kurths. J. Geophys. Res. 106 (B2), 2167, 2001.
[7] K. F. Tianmo and M. Anghel. Tectonophys., 1(413):1–3, 2006.
[8] M. F. Xia, Y. J. Wei, F. J. Ke, and Y. L. Bai. In Earth. Proc.: Phys. Model., Num. Sim. and Data Anal. Part II, 2491. Springer, 2002.
[9] C. G. Bufe and D. J. Varnes. J. Geophys. Res 98 (B6), 9871, 1993.
[10] D.D. Bowman et al., J. Geophys. Res, 103 (B10), 24359, 1998.
[11] A. Mignan, Tectonophys. 505 (1-4), 1, 2011.
[12] M. Acosta et al., Geophys. Res. Lett. 46(22), 12927, 2019.
[13] Y Huang et al., Europhys lett. 41(1), 43, 1998.
[14] A. Helmstetter, S. Hergarten, and D. Sornette. Phys. Rev. E, 70(4):046120, 2004.
[15] Y. Ben-Zion. Rev. of Geophys., 46(4), RG4006, 2008.
[16] J Kazemiet al., Phys. Rev. Lett., 114(8):088501, 2015.
[32] A Saichev and D. Sornette. Eur. Phys. J. B, 51(3):443, 2006.
[33] A. Wehrli, S. Wheatley, and D. Sornette. Quant. Fin. doi: 10.1080/14697688.2020.1838602, 2020.
[34] S. Wheatley, A. Wehrli, and D. Sornette. Quant. Fin., 19(7):1165–1178, 2019.
[35] J. Zhuang, Y. Ogata, and D. Vere-Jones. J. Am. Stat. Assoc., 97(458):369–380, 2002.
[36] A. M. Lombardi. Sci. Rep., 5:8417, 2015.
[37] J. Zhuang, Y. Ogata, and D. Vere-Jones. J. of Geophys. Res., 109(B5), 2004.
[38] J. Zhuang. Earth planets and space, 63(3):207–216, 2011.
[39] A. Veen and F. P. Schoenberg. J. Am. Stat. Assoc., 103(482):614–624, 2008.
[40] A. Helmstetter, Y. Y. Kagan, and D. D. Jackson. Bull. Seismol. Soc. Am., 96(1):90–106, 2006.
[41] Y. Ogata. Earth, planets and space, 63(3):6, 2011.
[42] S. Nandan, et al., J. of Geophys. Res., 124(8):8404–8425, 2019.
[43] C. Molkenthin, et al., arXiv preprint arXiv:2005.12857, 2020.
[44] A Helmstetter and D Sornette. Phys. Rev. E, 66(6):061104, 2002.
[45] A. Helmstetter, Y. Y. Kagan, and D. D. Jackson. Seism. Res. Lett., 78(1):78–86, 2007.
[46] S. Nandan, et al., Seismological Research Letters, 90(4), 1650-1659, 2019.
[47] S. Nandan, et al., Eur. Phys. J. Special Topics, DOI:10.1140/epjst/e2020-000259-3, 2020.
[48] R J Geller et al., Science, 275(5306):1616–1616, 1997.
Supplementary material for “Is the Earth crust operating at a critical point?”

Shyam Nandan¹, Sumit Kumar Ram², Guy Ouillon³, and Didier Sornette²,⁴

¹ETH Zürich, Swiss Seismological Service, Sonneggstrasse 5, 8092 Zürich, Switzerland
²ETH Zürich, Department of Management, Technology and Economics, Scheuchzerstrasse 7, 8092 Zürich, Switzerland
³Lithophyse, 4 rue de l’Ancien Sénat, 06300 Nice, France
⁴Institute of Risk Analysis, Prediction and Management (Risks-X), Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen, China

Corresponding author: Shyam Nandan (shyam4iiser@gmail.com)

Contents
Text S1-3
Table S1
Figures S1-11
References

Text S1: Extended Expectation-Maximization Algorithm for joint estimation of ETAS parameters and space varying background rate

We start with an initial guess of the independence probability, \( I_P \), for all the earthquakes in the primary catalog. This initial guess can be created by simply drawing a uniform random number between 0 and 1 corresponding to each earthquake. We also make initial guesses of the triggering parameters \{K, a, c, \( \omega \), \( \tau \), d, \( \gamma \), \( \rho \)\} as well as of the smoothing parameters \( (D, Q) \) for the background kernel.

**E step**: Using the current value of the parameters \( K, a, c, \omega, \tau, d, \gamma, \rho \) and the IP’s, we define the probability that the \( i \)th earthquake triggered the \( j \)th earthquake as:

\[
P_{ij} = \frac{g(t_j - t_i, x_j - x_i, y_j - y_i, m_i)}{\mu(x_j, y_j) + \sum_{l:\theta_l<j}g(t_l - t_i, x_l - x_i, y_l - y_i, m_i)}
\]

where \( \mu(x_j, y_j) \) is given by a modification of expression (5) obtained by removing the term \( i=j \) in the sum in order to avoid over-influencing the determination of the background rate by the specific location of the earthquakes in the catalog (in a way similar to the leave-one-out validation scheme of [Error! Reference source not found.] p. 127). The new estimates of the independence probabilities can be obtained as:

\[
I_{P, new} = 1 - \frac{\sum_{i \neq j} P_{ij}}{n-1}
\]

**M step 1**: Using the current estimate of the \( I_P \), we obtain the value of the PDF of the locations of background earthquakes at the location of the \( j \)th background earthquake as:

\[
\mu_{PDF}(x_j, y_j) = \frac{\sum_{i \neq j} I_P \ n^{-1} Q D^{2Q} \left((x_j - x_i)^2 + (y_j - y_i)^2 + D^2\right)^{-1-Q}}{\sum_{i \neq j} I_P}
\]

Using \( \mu_{PDF}(x_j, y_j) \) and \( I_P \), we can define the complete data loglikelihood for the spatial distribution of the background earthquakes as:
\[ LL_{bkg} = \sum_{j=1}^{N} IP_j \times \ln \mu_{PDF}(x_j,y_j) \]  

\( LL_{bkg} \) can be maximized with respect to the parameters \( D \) and \( Q \) to obtain their new estimates. To ensure consistency with the data quality, the minimum value of \( D \) is set to the average location error. We also remark that the expression of \( \mu_{PDF}(x_j,y_j) \) omits the term \( i = j \) in the sum in the r.h.s of Eq. (5). Otherwise, the optimization of \( LL_{bkg} \) leads to \( D \) being very close to 0 and to very large \( Q \) values, creating a Dirac function at the location of the \( j^{th} \) earthquake, thus leading to the maximal possible value of \( LL_{bkg} \). This singularity is avoided by using all earthquakes except the \( j^{th} \) earthquake to explain the background rate at its location [Silverman, 1986].

**M step 2:** Maximize the loglikelihood function \( LL_{trig} \) for the triggering part of the seismicity rate, given by the sum in the r.h.s. of Eq.(1), with respect to the parameters \{\( K, a, c, \omega, \tau, d, \gamma, \rho \}\).

Update the current estimates of \( (K, a, c, \omega, \tau, d, \gamma, \rho, D, Q) \) to the new estimates obtained in **M step 1** and **M step 2**. Repeat steps **E step, M step 1**, and **M step 2** until convergence. The latter requires that the loglikelihood increase is smaller than 1 (a sufficient criterion as we shall only consider datasets with thousands of samples in the primary catalog) and that the relative variation of the sum of inverted kernel parameters is less than 0.01.

**Text S2: The three data sets and their magnitude of completeness**

Locations of ~600,000 earthquakes \((M \geq 3)\) for the entire globe between 1975-2020, as reported in the ANSS catalog, are shown in Figure S1a. Figures S2a and S3a show the location of ~1.2 million and ~600,000 earthquakes \((M \geq 1)\) between 1975-2020 in the study region surrounding California and New Zealand, as reported by the ANSS and Geonet catalogs respectively. Some important metadata related to these catalogs is reported in Table S1.

Figures S1-3(b) show the time series of the cumulative number of earthquakes (black curve) corresponding to earthquakes shown in panel (a). All these curves show significant convexity, which expresses the increase in reporting rate of earthquakes at the chosen magnitude threshold. In essence, this convexity indicates that the global, Californian, and New Zealand catalogs are not complete above \( M \geq 3, M \geq 1, \) and \( M \geq 1, \) respectively. To decide an appropriate completeness level for each of these catalogs, we estimate how the \( b \)-value changes with different choices of \( M_c \). Before this, we bin the magnitude reported in the catalogs at 0.1 units. With these binned magnitudes, we then compute the \( b \)-values for different choices of \( M_c \), using the formula proposed by Tinti and Mulargia [1987]. We find that the estimated \( b \)-value first increases and then stabilizes. The stability of the \( b \)-value with change in chosen \( M_c \) has been proposed as a proxy for the magnitude above which the catalog is complete. Based on the stability of the \( b \)-value for each catalog, we conservatively choose a magnitude of completeness of 5, 3, and 4, respectively, for the global, Californian, and New Zealand catalogs. In panel d of each of the three figures (S1-3), we show the fit of the theoretical Gutenberg-Richter (GR) law, with parameters estimated using magnitudes above corresponding \( M_c \) to the empirical magnitude distribution. Due to an appropriate choice of \( M_c \), the theoretical GR law seems to describe the observed magnitude distribution above \( M_c \) quite well. As a last check of the appropriateness of the estimated \( M_c \), we again plot the time series of the cumulative number of earthquakes above \( M_c \) in panel (b) in figures S1-3 using the orange curve. We see that the convexity present in the black curve now disappears in the orange curve, further indicating that our choice of \( M_c \) for the three catalogs is appropriate.

**Text S3: Setup of pseudo-prospective experiments conducted on the three regions**
Starting on January 1, 1990, we perform 368, 368, and 371 pseudo prospective experiments for the three catalogs. In these experiments, the duration of the testing period is always 30 days long, and all the testing periods are non-overlapping. Only the data before the beginning of the testing periods are used to calibrate the models. Each model simulates 400,000 catalogs for each of the testing periods. The forecasts are made as stochastic event sets as this allows us to express the full PDF of the forecasts and not just the mean (see Nandan et al. [2019] for a discussion on the importance of the PDF). Since the parametric form of this dressed PDF is not known, it is expressed by sampling many times from the underlying ETAS model, justifying the need for many stochastic catalogs. The stochastic catalogs are then used to construct the forecast of the models at any spatial resolution and magnitude threshold during the testing periods. Given the size of the region and the computational complexity that follows, the global catalog is evaluated at nine spatial resolutions: $389 \text{ km}^2$, $4 \times 389 \text{ km}^2$, $4^2 \times 389 \text{ km}^2$, ..., $4^8 \times 389 \text{ km}^2$ and six magnitude thresholds: 5, 5.2, 5.4, ..., 6. For the forecast experiments, the Californian catalog is evaluated at five spatial resolutions: $4^{-1} \times 389 \text{ km}^2$, $389 \text{ km}^2$, $4 \times 389 \text{ km}^2$, $4^2 \times 389 \text{ km}^2$, $4^3 \times 389 \text{ km}^2$ and six magnitude thresholds: 3, 3.2, 3.4, ..., 4. The New Zealand catalog is evaluated at five spatial resolutions: $4^{-1} \times 389 \text{ km}^2$, $389 \text{ km}^2$, $4 \times 389 \text{ km}^2$, $4^2 \times 389 \text{ km}^2$, $4^3 \times 389 \text{ km}^2$ and six magnitude thresholds: 4, 4.2, 4.4, ..., 5. For more details on the testing setup, we refer the readers to [Nandan et al., 2020].

Table S1: Main characteristics of the catalog used in this study.

| Catalog Source | Region Details | Auxiliary Period | Primary Period | $N_{\text{evt}}$ (Primary) $\geq M_c$ | $N_{\text{evt}}$ (Auxiliary) $\geq M_c$ | $M_c$ (Primary) | b-value |
|----------------|----------------|------------------|----------------|--------------------------------------|---------------------------------------|-----------------|---------|
| ANSS Globe     | Entire Globe   | 1975-1980        | 1981-24 March  2020 | 65,200     | 9,609                                | 5                | 1.06    |
| ANSS California| Collection Polygon [Schorlemmer & Gerstenberger, 2007] | 1975-1980 | 1981-24 March 2020 | 27,413     | 5,610                                | 3                | 1.04    |
| Geonet New Zealand | Centre: [174.3, -41] Radius: 278 km | 1975-1980 | 1981-23 May  2020 | 12,875      | 1,041                                | 4                | 1.14    |
Figure S1: (a) Spatial distribution of ~600,000, $M \geq 3$ earthquakes, distributed over the entire globe and reported in the ANSS catalog; size and colors of the earthquakes scale with their magnitude according to the color code on the left; (b) time series of the cumulative number of $M \geq 3$ (black) and $M \geq 5$ (orange) earthquakes since 1975; (c) b-value as a function of the assumed magnitude of completeness $M_c$; solid black line indicates the chosen magnitude of completeness ($M_c = 5$); (d) Empirical magnitude distribution (black circles) and best fit to a Gutenberg-Richter law (red line) obtained using $M \geq M_c$.

Figure S2: (a) Spatial distribution of ~1.1 million, $M \geq 1$ earthquakes, distributed around California and reported in the ANSS catalog; size and colors of the earthquakes scale with their magnitude according to the color code on the left; (b) time series of the cumulative number of $M \geq 1$ (black) and $M \geq 3$ (orange) earthquakes since 1975; (c) b-value as a function of the assumed magnitude of completeness $M_c$; solid black line indicates the chosen magnitude of completeness ($M_c = 3$); (d) Empirical magnitude distribution (black circles) and best fit to a Gutenberg-Richter law (red line) obtained using $M \geq M_c$. 
Figure S3: (a) Spatial distribution of ~600,000, $M \geq 1$ earthquakes, distributed around New Zealand and reported in the GeoNet catalog; size and colors of the earthquakes scales with their magnitude according to the color code on the left; (b) time series of the cumulative number of $M \geq 1$ (black) and $M \geq 4$ (orange) earthquakes since 1975; (c) $b$-value as a function of the assumed magnitude of completeness $M_c$; solid black line indicates the chosen magnitude of completeness ($M_c = 4$); (d) Empirical magnitude distribution (black circles) and best fit to a Gutenberg-Richter law (red line) obtained using $M \geq M_c$.

Figure S4: Spatial density of the earthquakes identified as background by the ETAS$_\mu$ and ETAS$_{\mu(x,y)}$ models for the Californian catalog. The title indicates the date up to which earthquakes in the catalog were used to calibrate the two models and obtain these maps.
Figure S5: Spatial density of the earthquakes identified as background by the ETAS$_\mu$ and ETAS$_\mu(x,y)$ models for the New Zealand catalog. The title indicates the date up to which earthquakes in the catalog were used to calibrate the two models and obtain these maps.

Figure S6: Time series of cumulative information gain of ETAS$_\mu(x,y)$ over the ETAS$_\mu$ model in 368 pseudo prospective experiments for the Global ANSS catalog, at different spatial resolutions (different colors whose meaning is given in the inset) and magnitude threshold (different panels) of the testing catalog.
Figure S7: Time series of cumulative information gain of $\text{ETAS}_{\mu(x,y)}$ over the $\text{ETAS}_\mu$ model in 368 pseudo prospective experiments for the Californian ANSS catalog, at different spatial resolutions (different colors whose meaning is given in the inset) and magnitude threshold (different panels) of the testing catalog.

Figure S8: Time series of cumulative information gain of $\text{ETAS}_{\mu(x,y)}$ over the $\text{ETAS}_\mu$ model in 371 pseudo prospective experiments for the New Zealand catalog, at different spatial resolutions (different colors whose meaning is given in the inset) and magnitude threshold (different panels) of the testing catalog.
Figure S9: Mean information gain (MIG) of \( \eta_{\mu(x,y)} \) over the \( \eta_{\mu} \) model in 368 pseudo prospective experiments for the global ANSS catalog, at different spatial resolutions and magnitude threshold (different panels) of the testing catalog. The error bars indicate the 99 and 95% confidence interval of the mean information gain (MIG). The two numbers indicate the p-value and t-statistic resulting from the student’s t-test, in which we test the null hypothesis that the MIG is equal to 0 against the alternative that it is larger than 0. When the p-value is smaller than 0.05, the null hypothesis is rejected, and the circles are green. The notations e-xx means \( 10^{-xx} \).
Figure S11: Mean information gain (MIG) of $ETAS\mu(x,y)$ over the $ETAS\mu$ model in 371 pseudo prospective experiments with New Zealand catalog, at different spatial resolutions and magnitude threshold (different panels) of the testing catalog. The error bars indicate the 99 and 95 % confidence interval of the mean information gain (MIG). The two numbers indicate the p-value and t-statistic resulting from the student’s t-test, in which we test the null hypothesis that the MIG is equal to 0 against the alternative that it is larger than 0. When the p-value is smaller than 0.05, the null hypothesis is rejected, and the circles are green. The notations e-$xx$ means $10^{-xx}$.

References

1. Nandan, S., Ouillon, G., Sornette, D., & Wiemer, S. (2019). Forecasting the Full Distribution of Earthquake Numbers Is Fair, Robust, and Better. Seismological Research Letters, 90(4), 1650-1659.
2. Nandan, S., Kamer, Y., Ouillon, G., Hiemer, S., & Sornette, D. (2020). Global models for short-term earthquake forecasting and predictive skill assessment. Eur. Phys. J. Special Topics, https://doi.org/10.1140/epjst/e2020-000259-3.
3. Schorlemmer, D., & Gerstenberger, M. C. (2007). RELM testing center. Seismological Research Letters, 78(1), 30-36.
4. Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). CRC press.
5. Tinti, S., & Mulargia, F. (1987). Confidence intervals of b values for grouped magnitudes. Bulletin of the Seismological Society of America, 77(6), 2125-2134.