Earthquake productivity law

— Supporting Information —

Peter Shebalin*[1], Clément Narteau*[2], Sergey Baranov[1,3]

[1]Institute of Earthquake Prediction Theory and Mathematical Geophysics, 84/32 Profsouznaya, Moscow 117997, Russia.

[2]Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France.

[3]Kola Branch, Federal Research Center Geophysical Survey, Russian Academy of Sciences, Apatity 184209, Russia.

*Corresponding authors: shebalin@mitp.ru and narteau@ipgp.fr

Contents

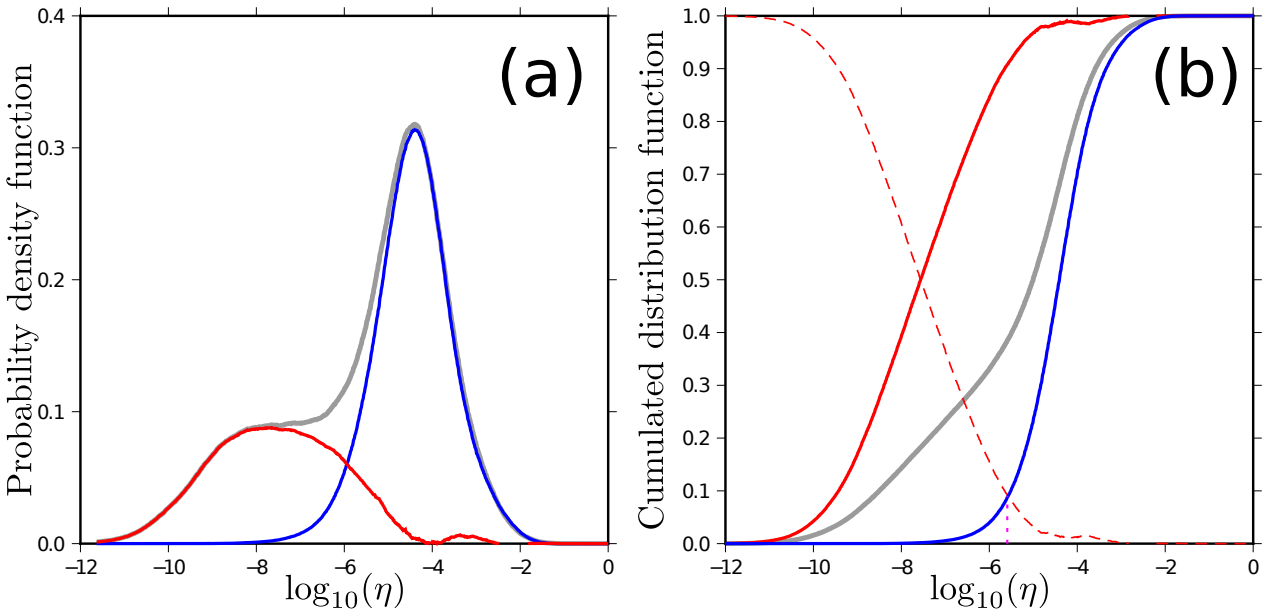

Figure S1: Determination of the threshold η_0-value using the proximity function of the main manuscript and the global catalog (see Tab. S1). **(a)** Probability distribution p_{real} of the η-value between nearest-neighbors (gray line) and its decomposition into a random and a clustered component: κp_{random} (blue line) and $(1-\kappa)\, p_{\text{clustered}} = p_{\text{real}} - \kappa p_{\text{random}}$ (red line). **(b)** Cumulative distribution F_{real} of the η-value between nearest-neighbors (gray line) and its random and clustered counterparts: F_{random} (blue line) and $F_{\text{clustered}}$ (red line). The red dashed line is the complementary cumulative distribution $1 - F_{\text{clustered}}$. The threshold η_0-value is found at the intersection of F_{random} and $1 - F_{\text{clustered}}$ (dotted line).

1 Non parametric determination of the threshold η_0-value

For a given proximity function, the standard procedure for separating related events from independent events is to assume that the distribution of η-values between nearest-neighbors follows a Gaussian mixture model, a linear combination of two Gaussian distributions[1, 2]. Then, the η-value equalizing the densities of the two weighted Gaussian distributions has been considered as the threshold η_0-value. Gamma and Weibull mixture models have also been proposed as alternatives to this standard procedure to provide a better fit to the data[3]. Obviously, the threshold η_0-value depends on the model under consideration, but there is no theoretical argument to distinguish between these different types of models.

Here, we introduce a model-independent method for selecting the threshold η_0-value whatever the choice of the proximity function. The principle of this new method is based on the assumption that the distribution of η-values for non-clustered earthquakes is similar to that of a random version of the catalog under consideration. This random catalog is obtained by mixing the time of the various events keeping unchanged their magnitudes and hypocenter coordinates. Hence, space-time clustering can also occur in the random version of the entire catalog. To overcome this problem, we use a random declustered catalog.

The complete procedure for selecting the threshold η_0-value consists of the following steps:

Step 1: Declustering of the initial catalog. Using only earthquakes above the magnitude of completeness of the initial catalog and a given proximity function, $p_{\text{real}}(\eta)$ is the distribution

of η-values between nearest-neighbors (gray curve in Fig. S1). $\eta_{\rm m}$ is the largest mode of the distribution. By definition, the probability mass of the non-clustered earthquakes preferentially lies to the right of this modal value. Then, $\eta_{1/2} > \eta_{\rm m}$ is the η-value for which $p_{\rm real}(\eta) = p_{\rm real}(\eta_{\rm m})/2$, By symmetry, we define the transient threshold $\eta_0^{\rm t}$-value

$$\eta_0^{\rm t} = \eta_{\rm m} - (\eta_{1/2} - \eta_{\rm m}).$$

Using this $\eta_0^{\rm t}$-value, we identify the hierarchical clustering trees, which may eventually consist of just one earthquake (see main manuscript).

Step 2: A random declustered catalog. While keeping together the hypocenter coordinates and the magnitude of each event, the earthquake times of the declustered catalog are redistributed at random. Using this random catalog and the same proximity function as in Step 1, $p_{\rm random}(\eta)$ and $F_{\rm random}(\eta)$ are the distribution and the cumulative distribution of η-values between nearest-neighbors, respectively.

Step 3: Decomposition of the initial catalog. Assuming that $p_{\rm random}(\eta)$ reproduces the distribution of η-values for non-clustered earthquakes, we decompose $p_{\rm real}(\eta)$ into two parts:

$$p_{\rm real}(\eta) = (1-\kappa)\, p_{\rm clustered}(\eta) + \kappa p_{\rm random}(\eta) \tag{1}$$

The same equation can be written for the cumulative distributions $F_{\rm real}(\eta)$, $F_{\rm clustered}(\eta)$ and $F_{\rm random}(\eta)$. Red and blue curves in Fig. S1 show these clustered and random components, respectively. The optimized weight κ minimizes the difference between $\kappa p_{\rm random}(\eta)$ and $p_{\rm real}(\eta)$ for η-values larger than $\eta_{4/5}$. As above for $\eta_{1/2}$, $\eta_{4/5} > \eta_{\rm m}$ is the η-value for which $p_{\rm real}(\eta) = 4p_{\rm real}(\eta_{\rm m})/5$. The mode of $\kappa p_{\rm random}(\eta)$ is smaller than $p_{\rm real}(\eta_{\rm m})$, so that $p_{\rm clustered} > 0$.

Step 4: Determination of the threshold η_0-value. The threshold η_0-value is defined as the η-value for which the proportion of clustered earthquake with nearest-neighbors $\eta > \eta_0$ (type I error) is equal the proportion of non-clustered events with nearest-neighbors $\eta \leq \eta_0$ (type II error):

$$1 - F_{\rm clustered}(\eta_0) = 1 - \frac{F_{\rm real}(\eta_0) - \kappa F_{\rm random}(\eta_0)}{1-\kappa} = F_{\rm random}(\eta_0) \tag{2}$$

Fig. S1 illustrates this approach for the worldwide catalog using the proximity function of the main manuscript. We perform a similar analysis in 7 areas using regional earthquakes catalogs (Fig. S2).

An alternative to Step 4 is to consider a number instead of the proportion of earthquakes for the two types of errors. In this case, we obtain a threshold η_1-value for which

$$F_{\rm real}(\eta_1) = 1 - \kappa. \tag{3}$$

The η_1-value may be more suitable for catalog declustering in seimic-hazard assessment because it compensates the number of independent earthquakes considered as clustered by an equivalent number of clustered earthquakes considered as independent[4]. In the vast majority of cases, the obtained η_0 and η_1-values are close to each other. Tab. S1 gives the $\{\eta_0,\, \eta_1\}$-values and the corresponding estimates of the $\{\Lambda_2,\, \Lambda_2'\}$-values. Only the results obtained from the threshold η_0-value are used in the main manuscript.

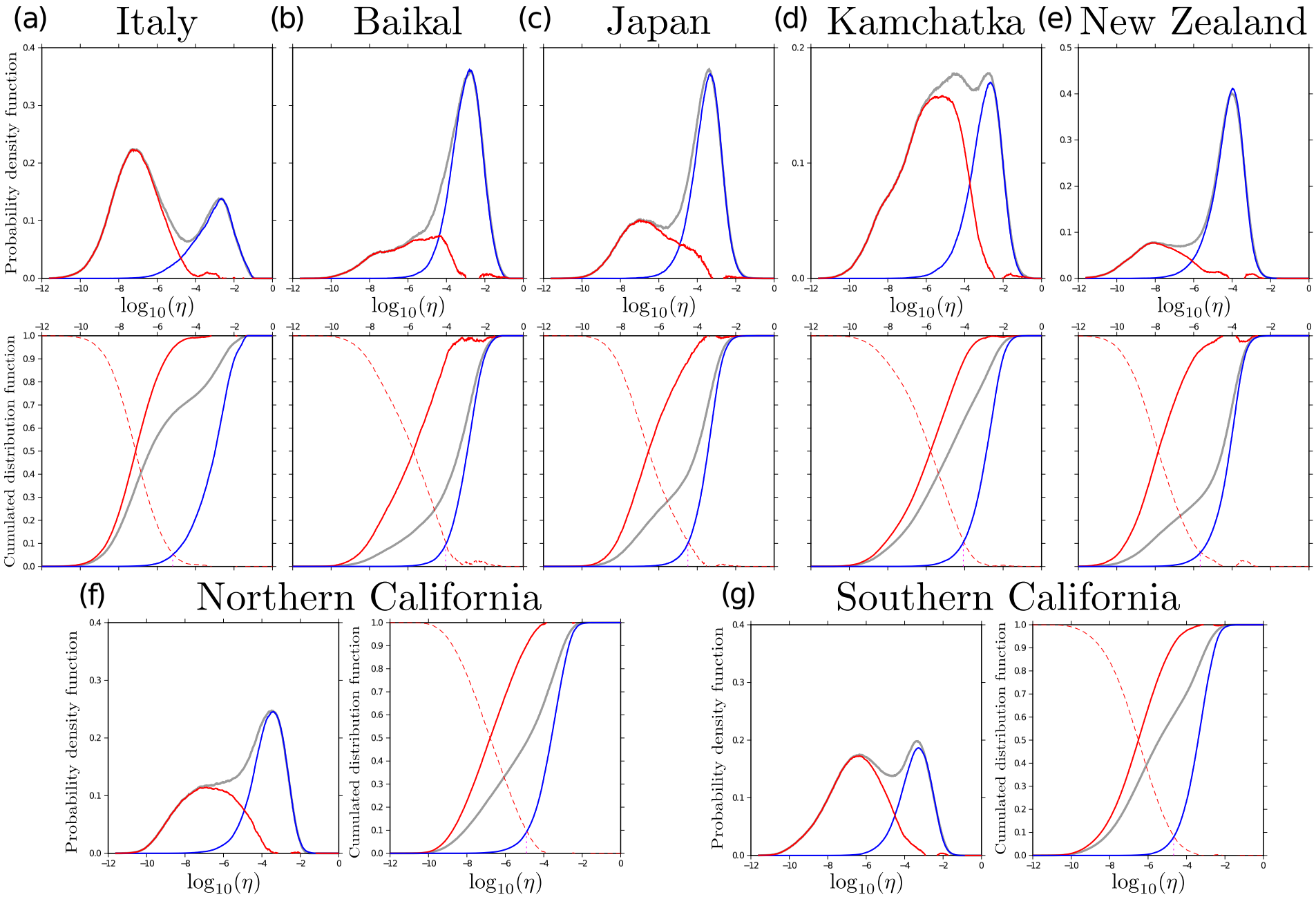

Figure S2: Determination of the threshold η_0-value using the proximity function of the main manuscript and seven different regional catalogs: **(a)** Italy, **(b)** Baikal, **(c)** Kamchatka, **(d)** Japan, **(e)** New Zealand, **(f)** Northern California and **(g)** Southern California. See Tab. S1 for catalog properties and caption of Fig. S1 for the definition of the probability distributions and the cumulative distribution functions.

Table S1: Catalog properties, threshold η-values and mean earthquake productivity $\Lambda_{\Delta M}$, $\Delta M = 2$.

Catalog	Period	$M_m^{(*)}$	$b^{(\dagger)}$	$d_f^{(\ddagger)}$	κ	$10^6\eta_0$	Λ_2	$10^6\eta_1$	Λ_2'
Global ComCat [(1)]	1981–2019	6.5	1.15	1.79	0.64	2.6	4.32	1.7	4.05
Italy [(2)]	2008–2019	3.9	1.17	1.85	0.35	6.5	5.02	11.0	5.10
Baikal [(3)]	1960–2014	4.2	0.92	1.2	0.70	90.0	1.31	54.0	1.18
Kamchatka [(4)]	1994–2019	5.5	1.05	1.98	0.62	32.0	5.28	23.0	5.04
Japan [(5)]	1997–2019	5.0	0.92	1.83	0.335	85.0	2.80	140.0	2.93
New Zealand [(6)]	1990–2019	5.0	1.14	1.78	0.73	2.2	3.72	1.0	3.38
Northern California [(7)]	1984–2019	3.8	1.0	1.4	0.52	13.0	3.06	12.0	3.04
Southern California [(8)]	1981–2019	3.8	1.0	1.6	0.37	21.0	3.46	30.0	3.54

[(*)] Minimal magnitude of triggering earthquakes. The completeness magnitude M_c is estimated using the Maximum curvature method[5]: $M_{\text{triggering}} \geq M_m = M_c + \Delta M$ with $\Delta M = 2$.

[(†)] Aki maximum likelihood estimate of the b-value[6].

[(‡)] Correlation dimension[7] measured in a range of distances from 1 to 100 km in 3D.

[(1)] U.S. Geological Survey, Earthquake Hazards Program, 2017, Advanced National Seismic System (ANSS) Comprehensive Catalog of Earthquake Events and Products, available at `https://earthquake.usgs.gov/earthquakes/search/` (doi: 10.5066/F7MS3QZH).

[(2)] The catalog of the National Center of Earthquakes, National Institute of Geophysics and Volcanology, available at `http://iside.rm.ingv.it/en`.

[(3)] Catalog of the Baikal Regional Seismological Center, Irkutsk, Geophysical Survey, Russian Academy of Sciences, available at `http://seis-bykl.ru/modules.php?name=Data&da=1`.

[(4)] Kamchatka Branch of GS RAS, Earthquakes Catalog for Kamchatka and the Commander Islands, available at `http://sdis.emsd.ru/info/earthquakes/catalogue.php`.

[(5)] The Seismological Bulletin of Japan, Japan Meteorological Agency, JMA, available at `https://www.data.jma.go.jp/svd/eqev/data/bulletin/index_e.html`.

[(6)] Earthquake catalog of the GeoNet Geological hazard information for New Zealand, available at `https://www.geonet.org.nz/data/types/eq_catalogue`.

[(7)] Double-difference relocated catalog[8, 9] of the Northern California Earthquake Data Center (NCEDC), available at `http://www.ncedc.org/ncedc/catalog-search.html`.

[(8)] Waveform relocated earthquake catalog[10], available at `http://service.scedc.caltech.edu/ftp/catalogs/hauksson/Socal`.

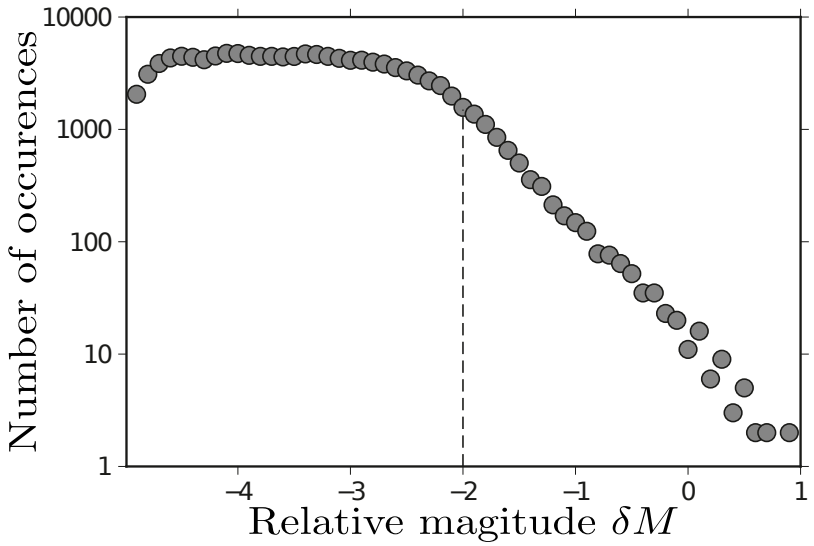

Figure S3: Earthquake-size distribution using $\delta M = M_{\text{triggered}} - M_{\text{triggering}}$, the difference in magnitude between the triggered and the triggering earthquakes. A ΔM-value of 2 ensures completeness.

2 Magnitude of completeness

To estimate completeness, we study the earthquake-size distribution using $\delta M = M_{\text{triggered}} - M_{\text{triggering}}$, the difference in magnitude between the triggered and the triggering earthquakes (Fig. S3). As throughout our analyses, we only consider $M_{\text{triggering}} \geq 6.5$ earthquakes. However, without lower limit in magnitude, we consider here all their triggered events in the catalogs according to our hierarchical declustering procedure. Fig. S3 shows that the earthquake-size distribution using the relative magnitude δM follows a Gutemberg-Richter distribution for $\delta M \geq -2$.

3 Hierarchical clustering trees of nearest-neighbour events

Using $\Delta M = 2$, Figs. S4 and S5 show the hierarchical clustering trees associated with the 2005 M8.6 Sumatran and the 1997 M7.8 Kronotsky earthquakes. Both are primary triggering events (level 0). In these clusters, we recognize the stochastic nature of the productivity. The largest event is not the one that will trigger the largest number of events in the corresponding magnitude range. For the Sumatran earthquake (Fig. S4), the most productive event belongs to level 2, it has a magnitude of 6.5 and it triggers 57 $M \geq 4.5$ events. For the Kronotsky earthquake, the most productive event belongs to level 1, it has a magnitude of 6.6 and it triggers 29 $M \geq 4.6$ events. It is the largest aftershock considering the traditional classifications.

4 Shape of the productivity distribution with respect to depth

In catalogs of seismicity, depth is a poorly constrained parameter often characterized by great uncertainty. In many cases, arbitrary depths are assigned to earthquakes based on their depth ranges. We test here how this classifications affects the dependency of the ΔM-productivity on depth. ANSS Comprehensive earthquake catalog ComCat provides depth values for all hypocenters.

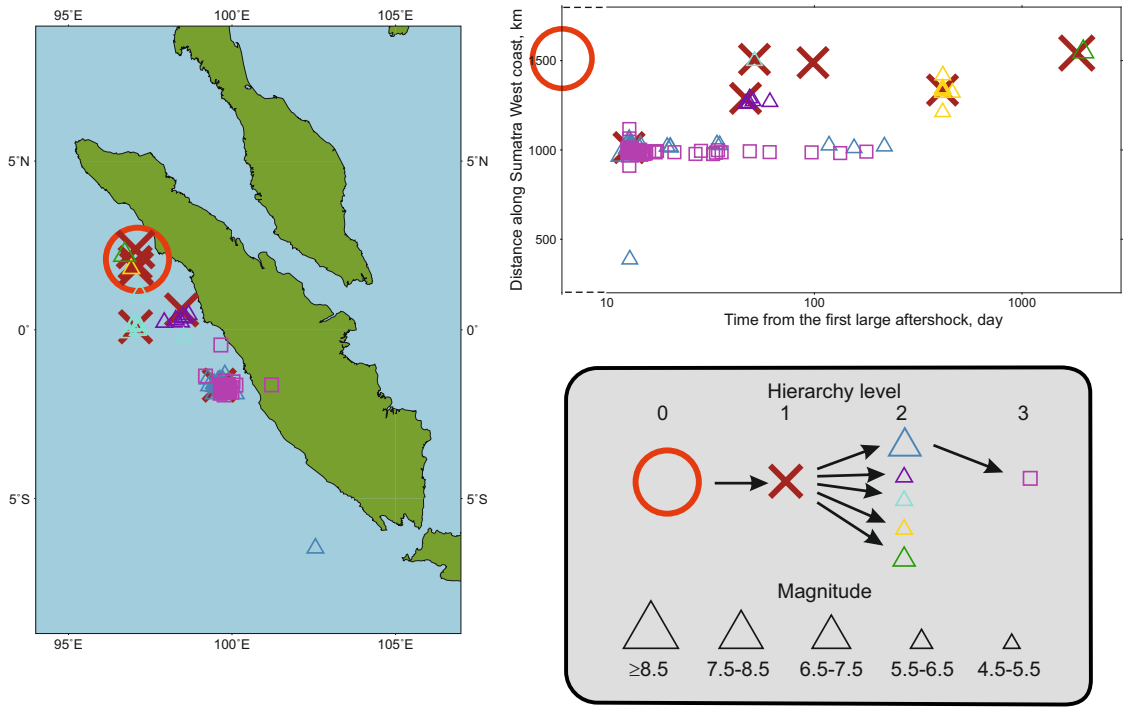

Figure S4: Hierarchical clustering tree of nearest-neighbor events for the 2005 M8.6 Sumatran earthquake using $\Delta M = 2$. This earthquake is a primary triggering event. Note the logarithmic time scale of the space-time diagram. The first triggered $M \geq 6.6$ event occurs only after 10 days.

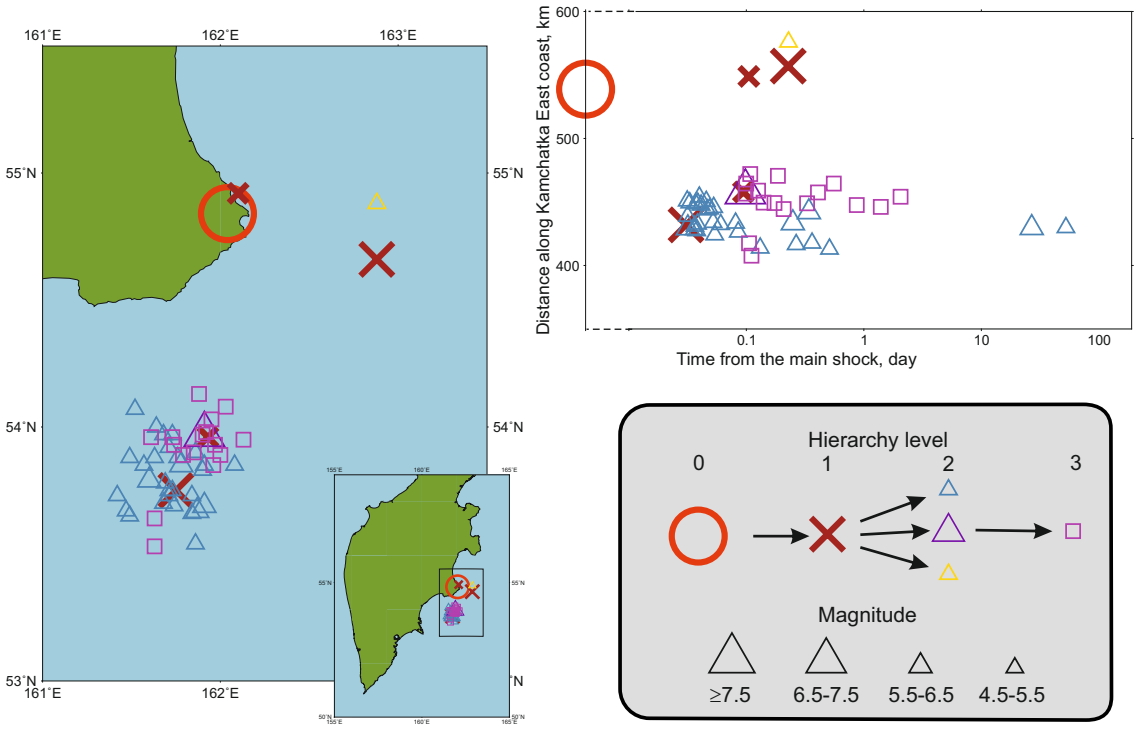

Figure S5: Hierarchical clustering tree of nearest-neighbor events for the 1997 M7.8 Kronotsky earthquake using $\Delta M = 2$. This earthquake is a primary triggering event. Note the logarithmic time scale of the space-time diagram.

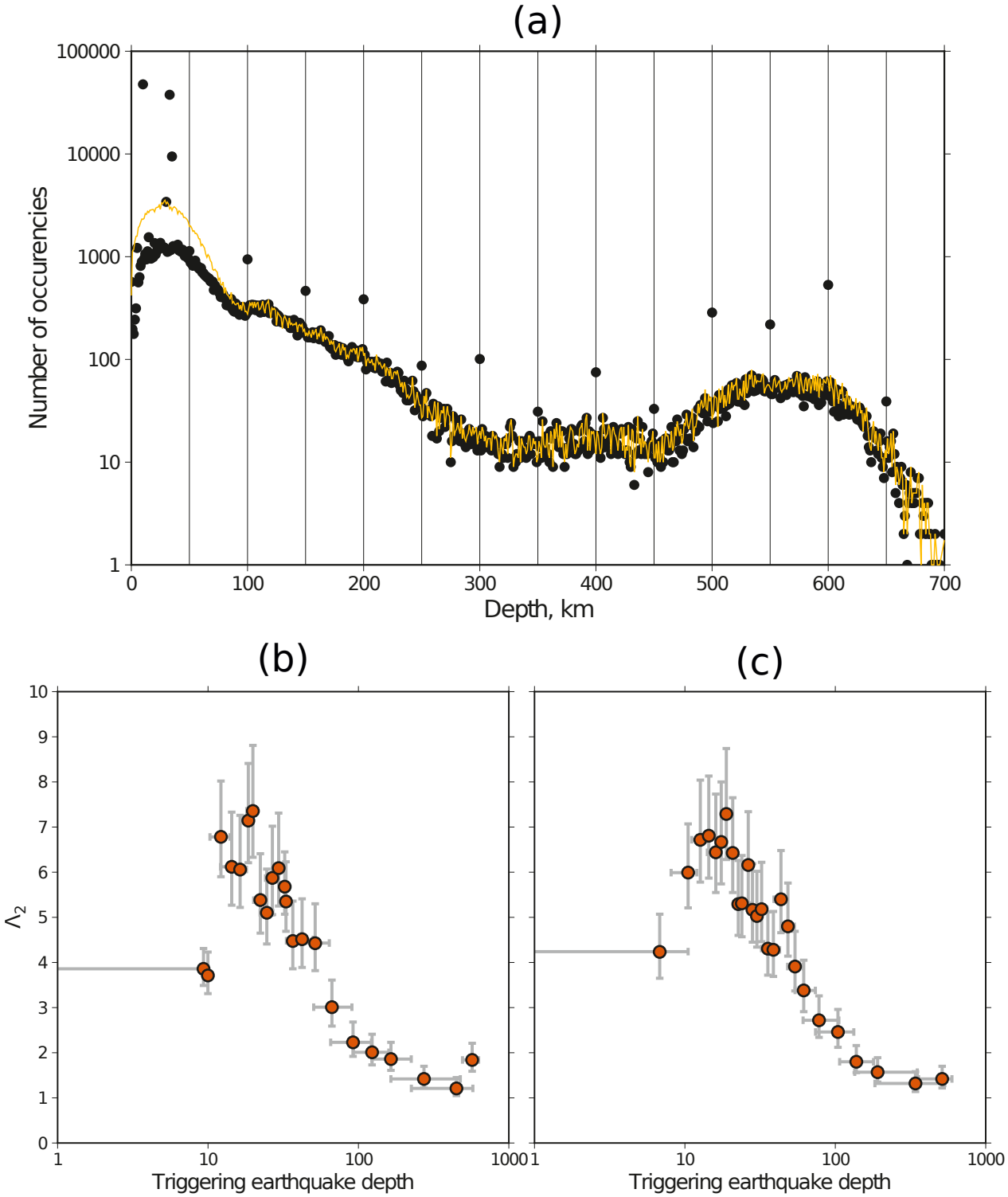

Figure S6: Impact of depth uncertainty on earthquake productivity in the worldwide catalog. (a) Distribution of depths of $M \geq 4.5$ earthquakes in the original ComCat catalog (dots) and in a modified catalog with randomized depths (orange line, see text). (b) and (c) show the average number of triggered events with respect to depth for $M \geq 4.5$ earthquakes using a relative magnitude threshold $\Delta M = 2$ in the original (b) and the randomized (c) catalogs, respectively. Triggering events are chosen with respect to depth using an overlapping sliding window of 100 events with a step of 50 events. Horizontal and vertical errorbars show the depth interval and the 90% credibility intervals of the likelihood function, respectively.

The depth distribution of $M \geq 4.5$ earthquakes shows that values of {10, 30, 33, 35} km were chosen to represent specific depth ranges (Fig. S6a). Multiples of 50 km and, to a lesser extent, multiples of 5 km are also over-represented depths. To eliminate these discontinuities in the depth distribution, we randomized shallow earthquake depth of {10, 30, 33, 35} km using Gaussian distribution with mean 30 km and standard deviation 20 km. For depths that are multiples of 50 and 5 km, we use these depths as the mean and standard deviation of 20 and 2 km, respectively. If we get a negative depth, we make a new random draw. Using this randomized depth, we apply the same declustering procedure and estimate the average ΔM-productivity as a function of the depth of the triggering earthquakes. Fig. S6 shows that the dependency on depth of the productivity is the same in the original and the randomized catalogs.

The clustering factor $\Lambda_{\Delta M}$ varies with depth. However, Fig. S7 shows that the distributions always keep an exponential shape for different depth ranges.

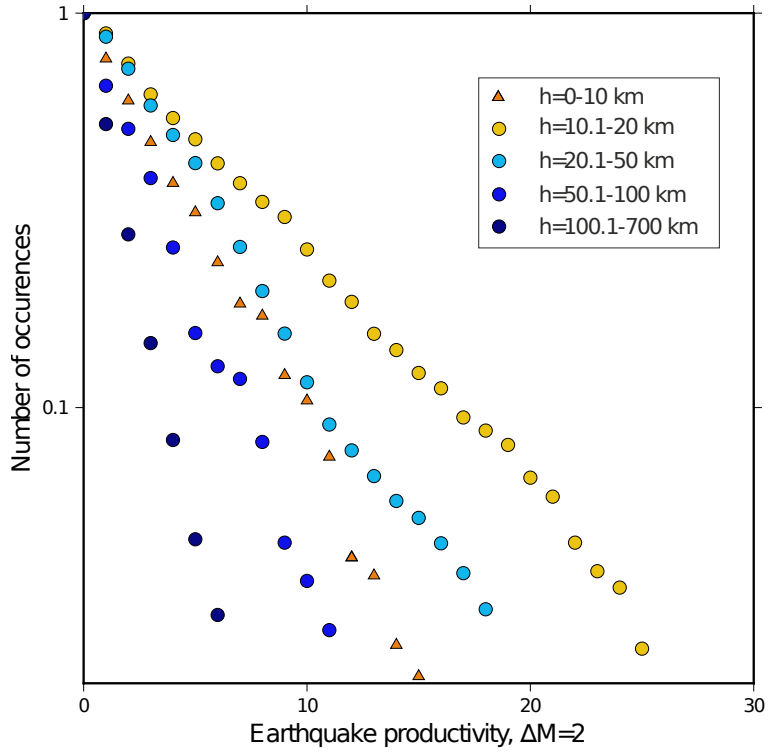

Figure S7: Normalized cumulative distribution of the number of triggered events in the global catalog for different depth ranges. As in the main manuscript, triggered events of $M \geq 6.5$ earthquakes are selected using a minimum relative magnitude $\Delta M = 2$.

5 The mean number of triggered earthquakes with a higher magnitude

The Λ_0-value gives the mean number of triggered earthquakes with the same or higher magnitude (i.e., foreshock, $M_{\text{triggered}} \geq M_{\text{triggering}}$). We calculate the Λ_0-value for all the 8 earthquake catalogs analyzed in the main manuscript using the thresholds η_0-values shown in Tab. S1. All Λ_0-values are near 0.1, the smallest 0.06 for New Zealand, the largest 0.115 for Kamchatka.

Catalog	Λ_0
Global ComCat	0.068
Italy	0.087
Baikal	0.087
Kamchatka	0.115
Japan	0.100
New Zealand	0.060
Northern California	0.097
Southern California	0.100

Table S2: Mean number of triggered events with a higher magnitude for $M_{\text{triggering}} \geq 4.5$ in all catalogs presented in Tab. S1. This is the $\Lambda_{\Delta M}$-values using a minimum relative magnitude $\Delta M = 0$.

6 Earthquake productivity in synthetic catalogs

It is critical to verify that our hierarchical declustering procedure is not the source of the observed exponential behaviour of the ΔM-productivity. With this purpose in mind, we apply our declustering procedure to synthetic catalogs.

We construct synthetic catalogs using a simple isotropic spatio-temporal Epidemic Type Aftershock sequence (ETAS) model[1, 11, 12, 13, 14]. The model consists of background events occurring according to a homogeneous stationary Poisson process of rate μ. Each earthquake in the catalog triggers first generation events, these events trigger their own sequences of events, and so on. The magnitudes of the events are assumed to be independent and to follow the Gutenberg-Richter distribution with a constant b-value. The resulting seismic rate is a compound of the background and triggered events from all generations, defined by intensity

$$\lambda(t,x,y,z) = \mu + \sum_{i:t_i<t} \frac{1}{(t-t_i+c)^p} \times \frac{K\,10^{\alpha(m_i-M_0)}}{(\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}+d)^q} \tag{4}$$

where M_0 is a magnitude threshold and t_i, x_i, y_i, z_i, m_i time, coordinates and magnitude of the i^{th} event. The temporal component corresponds to the Omori-Utsu law. Its spatial counterpart follows a similar power-law distribution. The model is specified by 8 scalar parameters $\{\mu, b, K, c, p, \alpha, d, q\}$. The parameter values are chosen with respect to the properties of seismicity in the global catalog. We take $M_0 = M_c = 4.5$, $b = 1.15$, $c = 0.0356$ day and $p = 1.08$. Under the assumptions that $\alpha = b$ and $p > 1$, Eq. 4 writes

$$\lambda(t,x,y,z) = \mu + \sum_{i:t_i<t} \frac{1}{(t-t_i+c)^p} \times \frac{(p-1)c^{p-1} \times 10^{-b\Delta M}\Lambda_{\Delta M}}{(\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}+d)^q} \tag{5}$$

Here we take $\Delta M = 2$ and $\Lambda_2 = 4.32$. We chose $\mu = 10$ events per day in an area of 5000×5000 km^2 to obtain approximately the same spatial and temporal density of earthquakes as in the global catalog. We simulate the catalog with total duration of 20,000 days testing different q-values in a range of 1.5 to 3 and different d-values in a range of 1 to 30.

For all versions of the synthetic catalog, distributions of the ΔM-productivity are in good agreement with Poisson distributions with modal values that are significantly different from zero. These modal values are close to the $\Lambda_{\Delta M}$-value. Fig. S8a shows an example of the productivity for

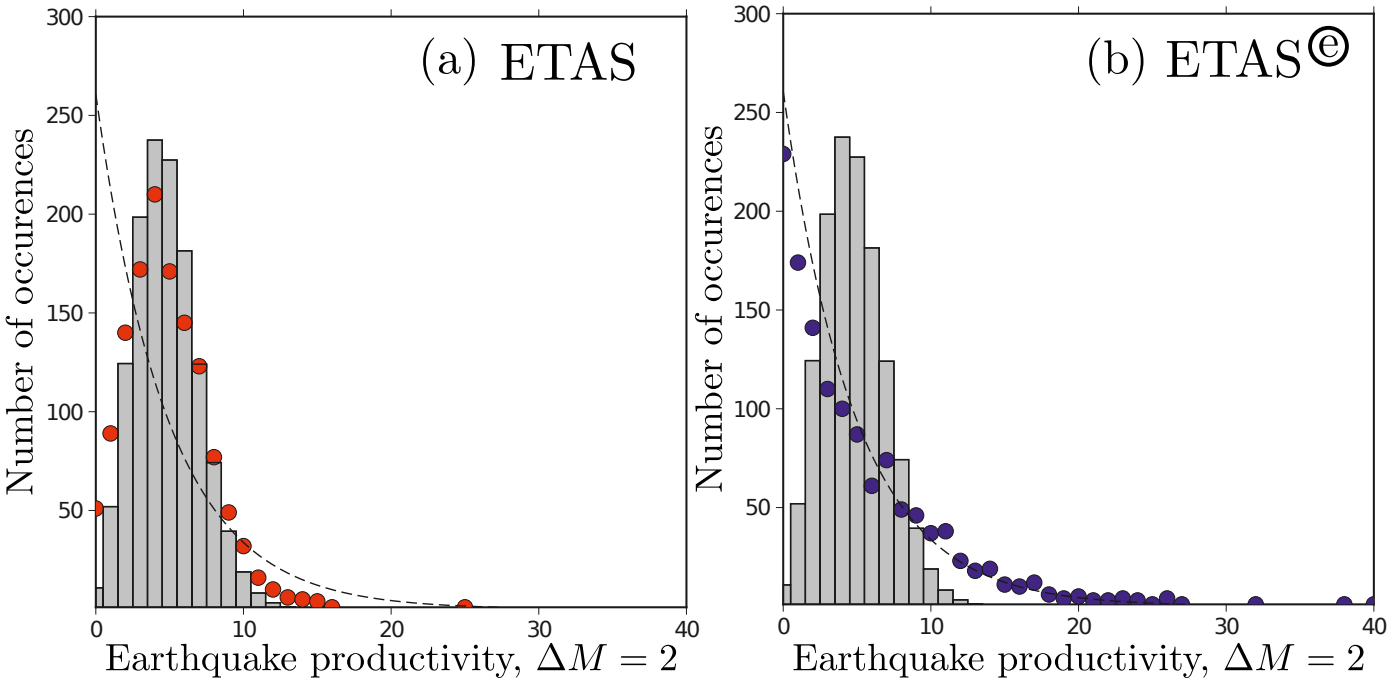

Figure S8: Distribution of the earthquake ΔM-productivity in synthetic catalogs using (a) ETAS and (b) ETAS^{e} models (see text). Dots show the distribution of the number of triggered events for $M \geq 6.5$ earthquakes using a relative magnitude threshold $\Delta M = 2$. The dashed line is the exponential law with parameter Λ_2, the mean number of triggered events derived from the data. The histogram shows the Poisson distribution with parameter Λ_2.

$\Delta M = 2$, $q = 2$ and $d = 1$ km and compared it to the Poisson and exponential distributions with parameter Λ_2.

In a new set of numerical experiments, we modify a single ingredient of the spatio-temporal ETAS model (Eq. 5). The productivity $\Lambda_{\Delta M}$ is no longer a deterministic quantity but the mean value of a random variable following an exponential distribution. When it comes to generating the synthetic catalog, this new ingredient takes the form of an additional random draw associated with each event. We call this modified ETAS model the ETAS^{e} model. It does not have an explicit form of the conditional intensity function, but we can still numerically simulate this process. When we apply our declustering procedure to the synthetic catalogs produced by the ETAS^{e} model, the ΔM-productivities are in good agreement with the exponential distribution and have maximum values at zero. Fig. S8b shows an example of the productivity keeping the same parameter as in Fig. S8a.

These tests demonstrate that our hierarchical declustering procedure is able to recover both the predefined Poisson and exponential distributions of the productivity in synthetic catalogs produced by epidemic models of seismicity. Hence, we can reject the hypothesis that the exponential behaviour observed in real catalogs is an artefact of our declustering procedure.

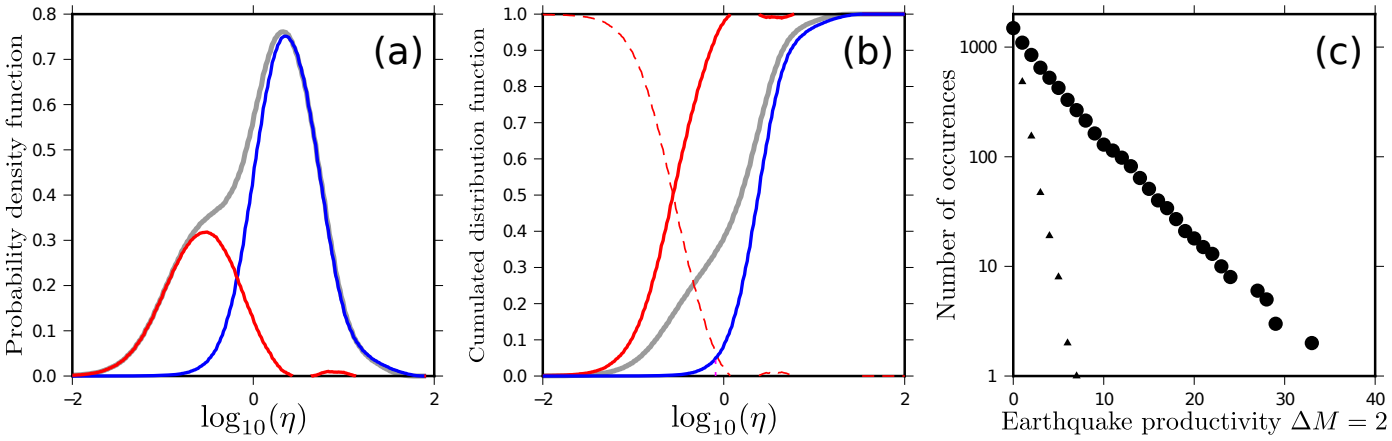

Figure S9: Determination of the earthquake productivity distribution of the global catalog (see Tab. S1) using the proximity function of Eq. 6. **(a)** Probability distribution p_{real} of the η-value between nearest-neighbors (gray line). Blue and red lines show its decomposition into a random and a clustered component, respectively. We used the same random catalog as for Fig. S1. Here, $\kappa = 0.67$. **(b)** Determination of the threshold η_0-value using the cumulative distributions F_{real} of the η-value between nearest-neighbors (gray line) and its random (F_{random}, blue line) and clustered ($F_{\text{clustered}}$, red line) counterparts. The red dashed line shows the complementary cumulative distribution $1 - F_{\text{clustered}}$. The threshold η_0-value is found at the intersection of F_{random} and $1 - F_{\text{clustered}}$ (dotted line). The threshold η_1-value is derived from the equation $F_{\text{real}}(\eta_1) = 1 - \kappa$. **(c)** Cumulative distributions of the earthquake productivity for $\Delta M = 2$ in the original (circles) and random (triangles) catalogs.

7 Impact of the proximity function on the productivity distribution

We study how the exponential form of the productivity distribution is sensitive to the proximity function. In addition to the proximity function used in the main manuscript, we test five other functions.

Following *Frohlich and Davis*[15], the first alternative is to add a scaling factor according to the magnitude of the triggering earthquake:

$$\eta_{ij} = \begin{cases} \sqrt{(c_1\ t_{ij})^2 + (r_{ij})^2}\, 10^{-c_2 m_i} & \text{for } t_{ij} > 0, \\ +\infty & \text{for } t_{ij} \leq 0, \end{cases} \tag{6}$$

where time is expressed in days, $c_1 = 1$ km/day and $c_2 = 0.3$.

Four other proximity functions are adapted from the space-time window approaches widely used for catalog declustering since *Gardner and Knopoff*[16]. The nearest neighbor is defined as the

1. closest in time,

$$\eta_{ij} = \begin{cases} t_{ij} & \text{for } t_{ij} \leq c_{\text{t}} 10^{p_{\text{t}} m_i} \text{ and } r_{ij} \leq c_{\text{r}} 10^{p_{\text{r}} m_i} \text{ and } t_{ij} > 0, \\ +\infty & \text{for } t_{ij} > c_{\text{t}} 10^{p_{\text{t}} m_i} \text{ or } r_{ij} > c_{\text{r}} 10^{p_{\text{r}} m_i} \text{ or } t_{ij} \leq 0, \end{cases} \tag{7}$$

2. closest in space,

$$\eta_{ij} = \begin{cases} r_{ij} & \text{for } t_{ij} \leq c_{\text{t}} 10^{p_{\text{t}} m_i} \text{ and } r_{ij} \leq c_{\text{r}} 10^{p_{\text{r}} m_i} \text{ and } t_{ij} > 0, \\ +\infty & \text{for } t_{ij} > c_{\text{t}} 10^{p_{\text{t}} m_i} \text{ or } r_{ij} > c_{\text{r}} 10^{p_{\text{r}} m_i} \text{ or } t_{ij} \leq 0, \end{cases} \tag{8}$$

3. closest in terms of space-time distance[15],

$$\eta_{ij} = \begin{cases} \sqrt{(c_1\ t_{ij})^2 + (r_{ij})^2} & \text{for } t_{ij} \leq c_{\rm t} 10^{p_{\rm t} m_i} \text{ and } r_{ij} \leq c_{\rm r} 10^{p_{\rm r} m_i} \text{ and } t_{ij} > 0, \\ +\infty & \text{for } t_{ij} > c_{\rm t} 10^{p_{\rm t} m_i} \text{ or } r_{ij} > c_{\rm r} 10^{p_{\rm r} m_i} \text{ or } t_{ij} \leq 0, \end{cases} \tag{9}$$

4. the largest earthquake:

$$\eta_{ij} = \begin{cases} m_i - m_j & \text{for } t_{ij} \leq c_{\rm t} 10^{p_{\rm t} m_i} \text{ and } r_{ij} \leq c_{\rm r} 10^{p_{\rm r} m_i} \text{ and } t_{ij} > 0, \\ +\infty & \text{for } t_{ij} > c_{\rm t} 10^{p_{\rm t} m_i} \text{ or } r_{ij} > c_{\rm r} 10^{p_{\rm r} m_i} \text{ or } t_{ij} \leq 0, \end{cases} \tag{10}$$

Note that, in contrast to the two first proximity functions (Eqs. 6 and 7), the three last functions (Eqs. 8 to 10) do not use a scaling in magnitude to find the nearest-neighbors.

Using the proximity function of Eq. 6, we determine the threshold η_0-value as before using a random version of the catalog. Applying this procedure to the global catalog (see Tab. S1) with $\Delta M = 2$, we find $\kappa = 0.67$, $\eta_0 = 0.81$, $\eta_1 = 0.72$, $\Lambda_2 = 3.55$ and $\Lambda_2' = 3.39$ (see Tab. S1 and Fig. S9). For the random catalog, $\Lambda_2 = 0.68$ and $\Lambda_2' = 0.52$.

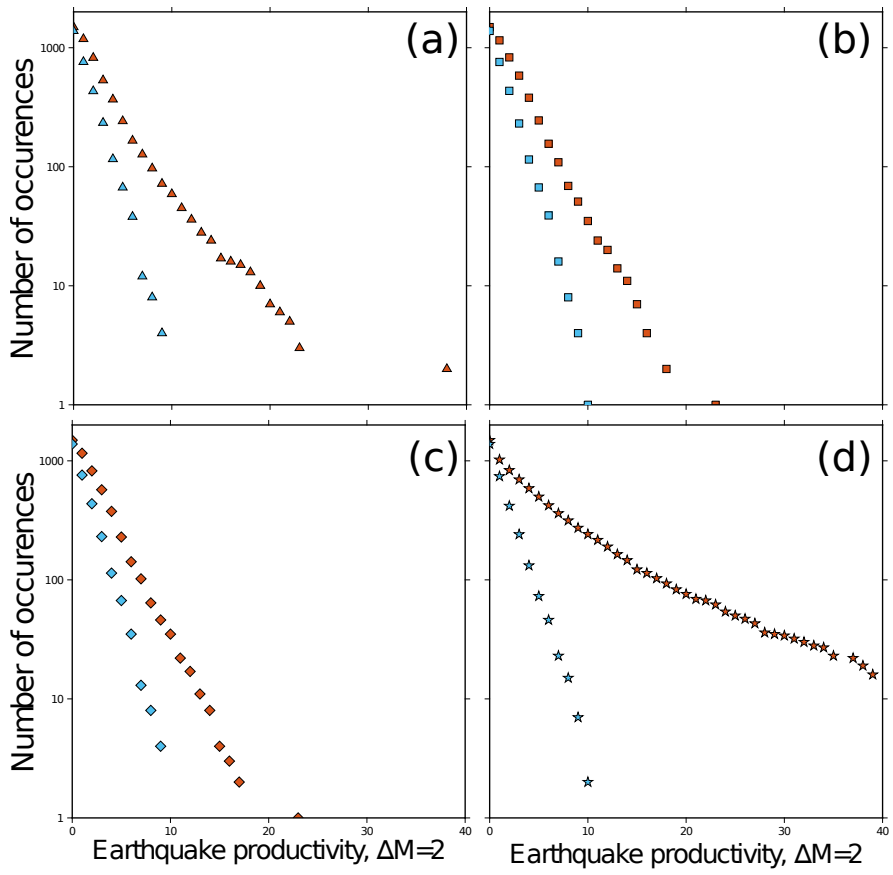

Figure S10: Earthquake productivity distributions of $M \geq 6.5$ earthquakes in the global catalog (see Tab. S1) using a minimum relative magnitude $\Delta M = 2$ and different proximity functions: Eq. 7 to Eq. 10 from **(a)** to **(d)**. The earthquake productivity distributions of the original and random catalogs are shown with brown and blue symbols, respectively.

We note that the distributions of earthquake productivity obtained in the main manuscript and those obtained using Eq. 6 are quite similar despite the difference in the scaling in magnitude: 1.15 and 0.3 respectively. It indicates that the exponential shape of the productivity distribution is an intrinsic property of seismicity and not a consequence of a specific scaling exponent.

For the four last proximity functions (Eqs. 7 to 10), the values of the parameters $c_{\rm t}$, $c_{\rm r}$, $p_{\rm t}$, and $p_{\rm r}$ are the same. The difference between these four proximity functions lies only in the choice of the

nearest-neighbor. This is critical in our analysis because we select pairs of triggering and triggered events, and not the complete hierarchical trees. For Eq. 9, $c_1 = 1$ km/day. By varying the values of the parameters c_t, c_r, p_t, and p_r, we maximize the function $(F_{real} - \kappa F_{random})$ (see Sec. 1). In the global catalog (see Tab.S1), we use $\kappa = 0.67$ and the optimal values $c_t = 0.75$ days, $p_t = 0.3$, $c_r = 0.137$ km and $p_r = 0.44$. Then, as in the main text, the threshold η_0-values are found using the procedure based on the comparison between nearest-neighbors in the original and random catalogs. Then, hierarchical trees of nearest-neighbors are built and the number of triggered events of each $M \geq 6.5$ earthquake is found using a minimum relative magnitude $\Delta M = 2$. As a result, Fig. S10 shows the cumulative non-normalized distributions of the earthquake productivity using Eqs. 7 to 10 in the global catalog. Tab. S3 gives the mean earthquake productivity Λ_2 in the original and random catalogs for this various proximity functions.

Fig. S10 and Tab. S3 show the impact of the proximity function on the productivity distribution. A scaling in magnitude as in the equation used in the main manuscript and Eq. 6 is preferable to capture a more stable behavior. However, the mode of all distributions shown in Fig. S10 is equal to 0 and, in all cases, significantly different from the mean productivity, thus demonstrating the systematic inadequacy of the Poisson distribution.

Proximity function	Original catalog	Random catalog
Eq. 7	2.64	1.20
Eq. 8	2.48	1.21
Eq. 9	2.43	1.21
Eq. 10	5.23	1.23

Table S3: $\Lambda_{\Delta M}$-values computed from the global and random catalogs using different proximity functions, $M \geq 6.5$ and a minimum relative magnitude $\Delta M = 2$.

8 Impact of the threshold η_0-value on the productivity distribution

The productivity distribution of non-clustered seismicity can be derived from random catalogs. This distribution also has an exponential shape, but with a Λ_2-value at least one order of magnitude smaller than the one derived from the original catalog.

By increasing the threshold η_0-value, the productivity distribution keeps its exponential decay rate for events triggering more than 2 events, but does not exhibit anymore such an exponential decay for a smaller number of triggered events (Figs. S11a-c). This transition from a pure exponential regime to a hybrid distribution occurs at η_0-value higher than the estimates obtained by the non-parametric procedure described in Sec. 1.

Fig. S11d shows the dependence of the Λ_2-value on the threshold η_0-value for the original global catalog and its random version. The Λ_2-value systematically increases with respect to the η_0-value, but there is a clear acceleration when this value is higher than the estimate obtained by the non-parametric procedure described in Sec. 1. Above this value, the mode of the productivity distribution is not 0 anymore (Fig. S11d).

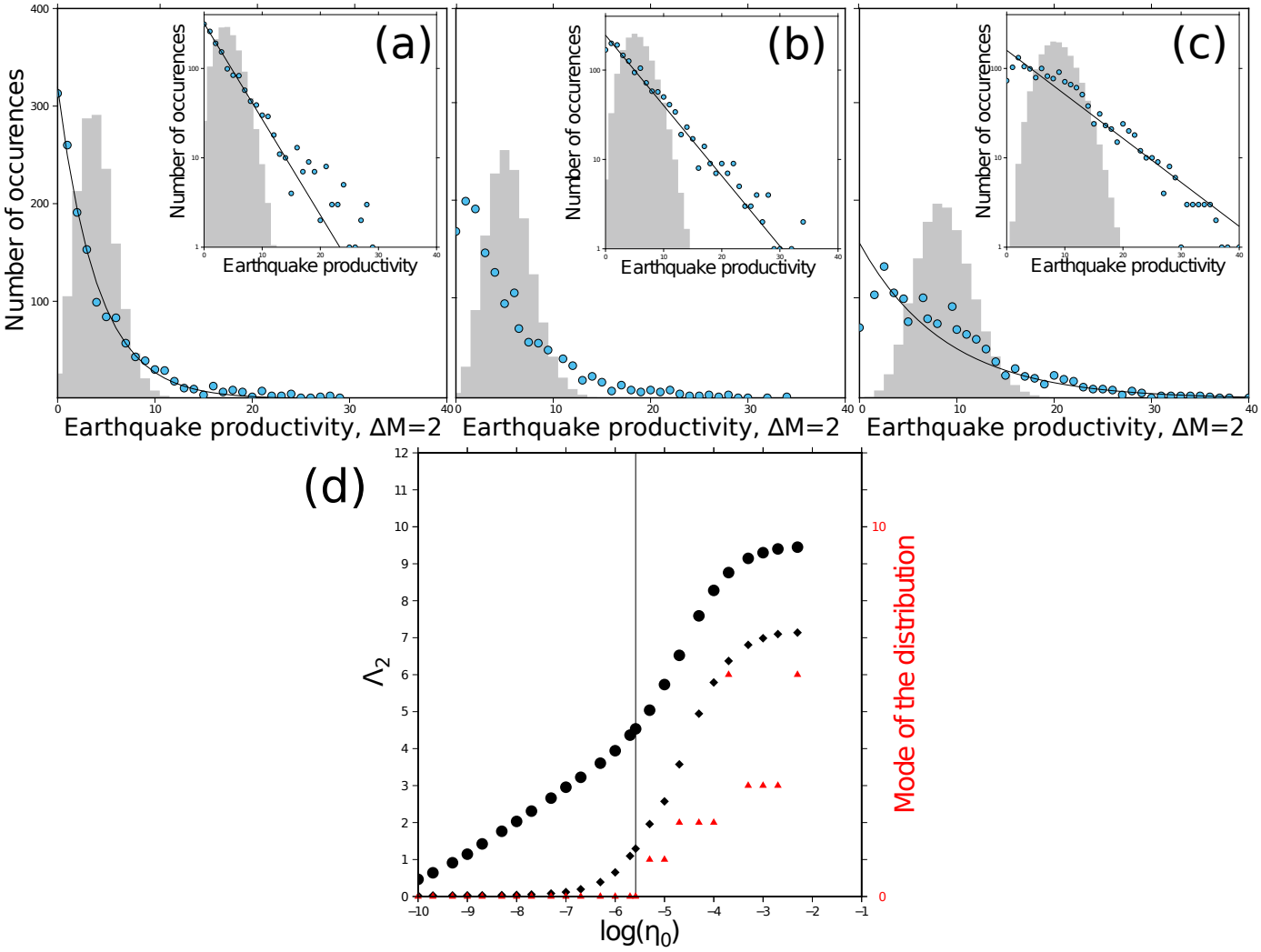

Figure S11: Dependence of the productivity distribution on the threshold η_0-value: **(a)** $\eta_0 = 2.6\ 10^{-6}$; **(b)** $\eta_0 = 10^{-5}$; **(c)** $\eta_0 = 10^{-4}$. For the global catalog using $M \geq 6.5$ earthquakes and a minimum relative magnitude $\Delta M = 2$, each figure compares the number of triggered events (blue circles) with the exponential (solid line) and the Poisson (gray histogram) distributions with the same rate parameter Λ_2. Insets are the same figures with a logarithmic scale for the number of occurrences. **d** Dependence of the Λ_2-value on the threshold η_0-value for the global catalog (circles) and the random catalogs (diamonds). Red triangles show the mode of the productivity distributions derived from the original catalog.

References

[1] Zaliapin, I. & Ben-Zion, Y. Earthquake clusters in southern california i: Identification and stability. *J. Geophys. Res.: Solid Earth* **118**, 2847–2864 (2013).

[2] Zaliapin, I. & Ben-Zion, Y. A global classification and characterization of earthquake clusters. *Geophys. J. Int.* **207**, 608–634 (2016).

[3] Bayliss, K., Naylor, M. & Main, I. G. Probabilistic identification of earthquake clusters using rescaled nearest neighbour distance networks. *Geophys. J. Int.* **217**, 487–503 (2019).

[4] Molchan, G. M. & Dmitrieva, O. E. Aftershock identification: methods and new approaches. *Geophys. J. Int.* **109**, 501–516 (1992).

[5] Wiemer, S. & Wyss, M. Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan. *Bull. Seism. Soc. Am.* **90**, 859–869 (2000).

[6] Aki, K. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence level. *Bull. Earthquake Res. Inst.* **43**, 237–239 (1965).

[7] Grassberger, P. & Procaccia, I. Measuring the strangeness of strange attractors. *Phys. D* **9**, 189 – 208 (1983).

[8] Waldhauser, F. & Schaff, D. P. Large-scale relocation of two decades of northern california seismicity using cross-correlation and double-difference methods. *J. Geophys. Res.: Solid Earth* **113** (2008).

[9] Waldhauser, F. Near-real-time double-difference event location using long-term seismic archives, with application to northern california. *Bull. Seism. Soc. Am.* **99**, 2736–2848 (2009).

[10] Hauksson, E., Yang, W. & Shearer, P. M. Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011)Short Note. *Bull. Seismol. Soc. Am.* **102**, 2239–2244 (2012).

[11] Ogata, Y. Seismicity analysis through point-process modeling; a review. *PAGEOPH* **155**, 471–508 (1999).

[12] Zhuang, J., Ogata, Y. & Vere-Jones, D. Stochastic declustering of space-time earthquake occurrences. *J. Am. Stat. Ass.* **97**, 369–380 (2002).

[13] Felzer, K. R., Abercrombie, R. E. & Ekström, G. A common origin for aftershocks, foreshocks, and multiplets. *Bull. Seismol. Soc. Am.* **94**, 88–98 (2004).

[14] Helmstetter, A. & Sornette, D. Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J. Geophys. Res.: Solid Earth* **107**, 2237 (2002).

[15] Frohlich, C. & Davis, S. How variable is the number of triggered aftershocks? *Geophys. J. Int.* **100**, 19–32 (1990).

[16] Gardner, J. & Knopoff, L. Is the sequence of earthquakes in southern California with aftershocks removed Poissonian? *Bull. Seismol. Soc. Am.* **5**, 1363–1367 (1974).