

# Perspectives on clustering and declustering of earthquakes

Ilya Zaliapin<sup>1</sup> and Yehuda Ben-Zion<sup>2</sup>

<sup>1</sup> Department of Mathematics and Statistics, University of Nevada, Reno, 1664 N. Virginia st.,  
Reno, NV 89557

<sup>2</sup> Department of Earth Sciences and Southern California Earthquake Center, Zumberge Hall of  
Science, University of Southern California, Los Angeles, CA 90089

Corresponding author: Ilya Zaliapin ([zal@unr.edu](mailto:zal@unr.edu))

*Seismological Research Letters*

In review, 2021

## Abstract

Clustering is a fundamental feature of earthquakes that impacts basic and applied analyses of seismicity. Events included in the existing short-duration instrumental catalogs are concentrated strongly within a very small fraction of the space-time volume, which is highly amplified by activity associated with the largest recorded events. The earthquakes that are included in instrumental catalogs are unlikely to be fully representative of the long-term behavior of regional seismicity. We illustrate various aspects of space-time earthquake clustering and propose a quantitative cluster measure based on the Receiver Operating Characteristic diagram. The proposed approach allows eliminating effects of marginal space and time inhomogeneities related to the geometry of the fault network and region-wide changes in earthquake rates, and quantifying

coupled space-time variations that include aftershocks, swarms and other forms of clusters. The proposed measure is used to quantify and compare earthquake clustering in Southern California, Western United States, Central and Eastern United States, Alaska, Japan, and ETAS model results. All examined cases show a high degree of coupled space-time clustering, with the marginal space clustering dominating the marginal time clustering. Declustering earthquake catalogs can help to clarify long-term aspects of regional seismicity and increase the signal-to-noise ratio of effects that are subtler than the strong clustering signatures. We illustrate how the high coupled space-time clustering can be decreased or removed using a data-adaptive parsimonious nearest-neighbor declustering approach. We discuss several alternative declustering methods that have been used in analyses of seismicity for over 50 years, emphasizing basic questions that remain on the proper outcome and quality metrics of declustering. Declustering remains at present an exploratory tool, rather than a rigorous optimization problem, and selecting an appropriate declustering method should depend on the data and problem at hand.

## 1. Introduction

Earthquake clustering is a fundamental aspect of seismicity. It is manifested most clearly by concentration of seismicity around tectonic plate boundaries and large faults (spatial clustering; see Fig. 1a, 2a) and after large earthquakes (aftershocks; see Fig. 1b, 2b), and is also associated with foreshocks, swarms and other types of clustering (e.g. *Jones and Molnar, 1979; Kagan and Jackson, 1991; Romanowicz, 1993; Press and Allen, 1995; Utsu, 2002; Ben-Zion, 2008; Mignan, 2014*). Earthquake clustering reflects various forms of triggering mechanisms including static and dynamic stress transfers from previous earthquakes, fluid migration and aseismic slip below and along faults (e.g. *Dieterich, 1994; Console et al., 2006; Felzer and Brodsky, 2006; Lengliné et al.,*

2012; *Ross et al.*, 2017; *Kato and Ben-Zion*, 2021). Seismicity clustering is important for many purposes such as identifying the character and interaction of active fault structures (e.g. *Ruhl et al.*, 2016), estimating time-dependent probability of seismicity (e.g. *Ogata*, 1999; *Field et al.*, 2017) and more. On the other hand, efforts to clarify less obvious aspects of earthquake dynamics are best done using declustered catalogs that do not include aftershocks, foreshocks and other strong forms of clustering. Examples include detection of tidal, seasonal, hydrologic, climatic and anthropogenic triggering of seismicity (e.g. *Cochran et al.*, 2004; *Amos et al.*, 2014; *Johnson et al.*, 2019; *Goebel and Brodsky*, 2018; *Hammond et al.*, 2019; *Hsu et al.*, 2021). For some applications it is also important to consider the “background” earthquake activity, generated within a given volume by the long-term regional tectonic processes, separated from the chains of triggered seismicity that typically follow the background events. Examples include development of long-term seismic hazard maps (e.g. *Petersen et al.*, 2015, 2018), inversions of focal mechanisms for the background stress field operating on the volume under consideration (e.g. *Martínez-Garzón et al.*, 2016; *Abolfathian et al.*, 2019) and estimations of evolving localization of seismicity before large earthquakes (e.g. *Ben-Zion and Zaliapin*, 2020).

The duration of instrumental earthquake catalogs is very short (typically only a few tens of years) relative to the duration of large earthquake cycles. Accordingly, they are often dominated by aftershocks of the largest earthquakes that happened to occur during the observational period. The instrumental catalogs may also be influenced by aftershocks of non-recorded historical events and aftershocks of events outside of the examined spatial region (*Wang et al.*, 2010; *van der Elst*, 2017). As a result, the recorded activity is not necessarily representative of what might occur in the following 20 (or even 5) years. The strong clustering of seismicity masks other properties and

reduces the ability to understand the long-term large-scale dynamics of earthquakes for which local triggering effects are averaged out.

A high degree of earthquake clustering is commonly obscured (quite deceptively) by a strong space-time concentration that is not easily discernible in a visual catalog inspection. Several key earthquake clustering phenomena are described by power laws that apply to size (e.g. frequency-moment statistics related to the Gutenberg-Richter law), time (e.g. Omori law), and space (e.g. decay of seismic rate with distance from the fault core); see *Ben-Zion* (2008, Table 2) for details and further examples. The intermediate portions of these power laws are well established observationally, but the tails of the power laws - associated with small-scale (left tails) and large-scale (right tails) features - remain underexplored. Most relevant for the present paper is the fact that the majority of events in the available catalogs occur within a tiny fraction of the total examined space-time volume in the immediate vicinity of other events, in accordance with the Omori-Utsu power-law temporal decay of aftershock sequences (*Utsu and Ogata*, 1995) and the power-law decay of seismic intensity off the fault core (*Felzer and Brodsky*, 2006; *Dieterich and Smith*, 2009; *Powers and Jordan*, 2010; *Goebel et al.*, 2014). Although these laws are well known, the severity of their consequences for clustering is rarely quantified and its effects remain underappreciated. In the following sections we illustrate and quantify earthquake clustering with examples of seismicity from Southern California and other regions. We propose a simple and robust measure of space-time clustering that allows one to disentangle effects related to concentration of earthquake along a heterogeneous fault network from additional coupled space-time fluctuations. We also discuss several existing approaches for declustering seismicity.

## **2. Clustering of earthquakes**

## 2.1 Basic observations

We begin by illustrating earthquake clustering in Southern California, using the high-quality relocated catalog of *Hauksson et al.* (2012, extended to later years) for earthquakes with magnitudes  $M \geq 2$  during 1981-2020 (Fig. 1). The catalog lists 89,341 such events within  $119.5^\circ\text{W}$ - $115.5^\circ\text{W}$  and  $32.8^\circ\text{N}$ - $36.5^\circ\text{N}$ , and it features an unprecedented combination of spatial coverage, completeness magnitude, and event location quality. The average horizontal location error for the examined events is 275m; the completeness magnitude is between 2 and 3 and may slightly vary in time and space (*Hauksson et al.*, 2012). Counting events within 7,605 spatio-temporal voxels that have a square spatial projection with  $0.25^\circ$  latitude side and duration of 1 year, shows that 38% (2,857) of the voxels are empty. The space and time projections of these voxels, with the corresponding marginal event counts, are shown in Figs. 1a and 1b, respectively. The voxel event count has a heavy-tailed distribution (e.g. *Kagan and Jackson*, 2000) with a highly uneven voxel population. While the average number of events in the non-empty voxels is 18, the maximal count is 2,733, which is 145 times the mean. This maximal voxel count corresponds to aftershocks of 1992 M7.3 Landers earthquake, so about 3% (2,733) of the entire catalog is contained within this single voxel, which is only about 0.02% of the examined non-empty space-time volume.

Similar clustering is seen in other regions and with other magnitude thresholds, so the clustering effect in the Southern California catalog is not due to completeness issues and event location errors. As a complementary example, we consider 7,066 global earthquakes with  $M \geq 5.8$  and depth  $z < 70\text{km}$  during 1976 – 2017 from the ISC-GEM v8 Global Instrumental Earthquake Catalogue (*Storchak et al.*, 2013, 2015; *Di Giacomo et al.*, 2018). This selection has been reported to be complete in several independent studies (e.g., *Michael*, 2014; *Di Giacomo et al.*, 2015). The global seismicity occupies a small fraction of the available space, concentrating around the main

plate boundaries (Fig. 2a). If we tile the Earth surface by 2,592 square spatial cells with latitude size of  $5^\circ$  (Fig. 2a), only 508 (20%) cells contain at least one earthquake and 2,084 (80%) cells remain empty. The earthquake distribution is also strongly non-uniform in time (Fig. 2b) mainly because of aftershocks sequences of large earthquakes (e.g. the 2011 M9.1 Tohoku). The space and time inhomogeneities combine and amplify when one considers the joint space-time distribution of earthquakes. For instance, if we partition the space-time volume of the examined global catalog into 20,736 voxels with a square space projection of  $5^\circ$  latitude and 5-year duration, 18,866 (90%) remain empty. The number of events in the 1,870 (10%) non-empty voxels varies between 1 and 130, with a mean of 3.78. The largest count of 130 corresponds to a voxel around the epicenter of the great 2011 M9.1 Tohoku earthquake. This single voxel contains 2% of the global seismicity, while representing only 0.005% of the examined space-time.

Figure 3 offers another visual illustration of the strong clustering of earthquakes in Southern California. Panel (a) shows the time-latitude projection of all 8,140 events with  $M \geq 3$  in the catalog. Panel (b) only includes the 50% (4,083) least clustered events based on the nearest-neighbor earthquake proximity discussed in *Zaliapin and Ben-Zion (2020)* and illustrated below. Despite this substantial culling, the panels are barely distinguishable visually. They show the same relative intensity of earthquakes at different latitudes, the same temporal fluctuations, and the same major groups of events. The figure illustrates that at least half of the catalog is in the form of extremely dense clusters not easily discernible by eye. The majority of earthquakes occur within the immediate spatio-temporal vicinity of other events, and do not evenly sample the active seismogenic volume.

Figure 3b suggests that a proper declustering of this data set should remove  $> 50\%$  of events, since it still shows visible clusters associated with the 1992 M7.3 Landers, 1999 M7.1 Hector

Mine, and 2019 M7.1 Ridgecrest. Figure 3c presents a declustered version of the catalog using the method of *Zaliapin and Ben-Zion (2020)*. The declustered catalog includes 19% (1,535) least clustered events. The declustering eliminates the prominent clusters illustrated in panels (a, b) while preserving the relative rates of events in different regions. This declustering approach is based on the separation of scales in the nearest-neighbor earthquake proximity (*Baiesi and Paczuski, 2004; Zaliapin et al., 2008; Zaliapin and Ben-Zion, 2013, 2016*). The proximity from any earthquake to an earlier event is defined as a product of the interevent time  $\Delta t$  and surface distance  $r$  raised to the fractal dimension  $d$  of the events,  $\eta = \Delta t r^d$ . The distribution of the nearest-neighbor proximity is prominently bimodal in various observed and model catalogs; see *Zaliapin and Ben-Zion (2020)* for references to specific studies.

Figure 4 shows the bimodal distribution of the nearest-neighbor proximity in the examined catalog of Southern California with magnitudes  $M \geq 3$  during 1981 – 2020. Panels a and b show, respectively, the scalar proximity and the joint distribution of its time component  $T = \Delta t$  and space component  $R = r^d$  defined so that  $\eta = RT$ . Here we use  $d = 1.6$ ; but the bimodal nature of the proximity is independent of the parameter selection (*Zaliapin and Ben-Zion, 2013*). The two modes are separated by several orders on the proximity scale (Fig. 4a), emphasizing the extreme nature of earthquake clustering. The two modes are also well separated in space-time (Fig. 4b). The cluster mode (lower proximity values) is mainly comprised of aftershock sequences and swarms, and is primarily responsible for the space-time clustering discussed and quantified in this work. The declustering involves stochastic thinning based on the proximity values; this helps avoiding a sharp cut along the model separation line, which might result in artificial inhomogeneities of the declustered catalog (*Zaliapin and Ben-Zion, 2020*).

An informative way to look at earthquake clustering is to examine how many events are offspring of background events, and hence may be viewed statistically as being triggered. In the absence of clustering, each event belongs to the background activity and does not produce offspring. In contrast, a high clustering implies that a large proportion of earthquakes are offspring to a small number of parent events. We estimate the offspring relations using the nearest-neighbor approach; informally, an event is considered an offspring if its nearest neighbor (parent), according to the space-time-magnitude proximity  $\eta$  belongs to the cluster mode of the bimodal diagram of Fig. 4.

Figure 5a shows the proportion of offspring events (y-axis) vs. the proportion of the parent events ( $x$ -axis) for earthquakes with  $M \geq 2$  in the catalog of *Hauksson et al.* (2012, extended) for Southern California during 1981 - 2020. The degree of estimated offspring productivity (triggering) is striking: the single M7.3 1992 Landers earthquake has 15,389 (17%) offspring, and over 50% of the events in the catalog are estimated offspring to mere 77 earthquakes with  $M \geq 5$ . Figure 5b illustrates the offspring analysis for several other regions examined in this study – they all show a high offspring productivity (measured by the deviation from the diagonal).

The results in Figs. 1-5 suggest that the majority of events in the observed earthquake catalogs belong to clusters that occupy a small highly active part of the examined space-time volume. This active part reflects primarily the particular set of large events that occurred during the data acquisition period and generated numerous aftershocks. This part might be not representative of the long-term processes in the region. For instance, large sections of the San Andreas and other major faults in Southern California have been relatively inactive during the data acquisition period 1981-2020 in the *Hauksson et al.* (2012, extended) catalog. Similarly, the two M9+ earthquakes in the examined ISC-GEM catalog (2004 M9.3 Sumatra-Andaman and 2011 M9.1 Tohoku) both



occurred in Asia and do not fully represent regions prone to M9+ earthquakes (e.g. the Pacific Rim). We show below that earthquake clustering goes beyond the existence of relatively small active space-time domains (non-empty voxels in the analysis of Figs. 1,2) and extreme aftershock sequences of the largest events (Figs. 3,5). The non-uniform heavy-tailed event distribution is seen generically within seismically active space-time volumes, biasing statistical analyses of seismicity at a variety of scales and with different event- and region-selection criteria.

In order to reduce the bias caused by space-time clustering and study more representative long-term processes, it is important to sample the entire seismogenic volume more evenly. This is the basic goal of catalog declustering. Before discussing declustering, we introduce a formal framework for assessing the degree of earthquake clustering and show that strong clustering is a common robust feature of multiple seismically active regions. We illustrate the analysis using earthquake catalogs of Southern California, Western United States, Central and Eastern United States, Alaska, Japan, global seismicity, and catalog of the Epidemic Type Aftershock Sequence (ETAS) model.

## ***2.2. Quantifying clustering: ROC diagram of earthquake space-time distribution***

To systematically quantify the inhomogeneity of the space-time distribution of seismicity, we use the Receiver Operating Characteristic (ROC) diagrams, following *Ben-Zion and Zaliapin* (2020). Specifically, we partition the examined space-time volume into voxels, count events in the voxels, sort non-zero event counts in descending order and construct a diagram (Fig. 6a) that shows the proportion of events in the most populated voxels (y-axis) vs. the proportion of such voxels (x-axis). Considering only non-empty voxels ensures that the analysis is not affected by selection of the boundaries of the examined region, which might affect the number of empty voxels. We

observe also that using only non-empty voxels decreases the estimated degree of clustering. The sorting of counts implies that the diagram always stays above the diagonal and connects the origin to the upper-right corner (1,1). The diagonal line corresponds to a constant count within non-empty voxels. It may be useful to think of a constant count as the expected value of a model that assigns the identically distributed event counts to all examined voxels. Convexity of the diagram (deviation from the diagonal) reflects non-uniformity in the event distribution – more events are concentrated within some voxels. Observe that the ROC analysis only focuses on relative rates of events in different voxels, and not on the absolute values of the rates. We quantify the non-uniformity of a ROC diagram by the Gini coefficient ( $G$ ) defined as twice the area between the diagram and the diagonal line. All realistic values of  $G$  are within the interval (0,1), where 0 corresponds to constant counts and 1 to an extreme concentration of all events within a single voxel. Tracking the temporal evolution of  $G$  provides a way of analyzing evolving localization of seismicity, as was done in *Ben-Zion and Zaliapin (2020)*.

To develop intuition about the  $G$  scale, Fig. 7 shows three synthetic examples of non-zero event counts with different degrees of non-uniformity (panels a-c) and their ROC diagrams with respective  $G$  values (panel d). All examples correspond to 50 cells with comparable average event counts. Figure 7a shows the most clustered sequence – the event counts substantially deviate from their mean value, varying between 1 and about 800; this corresponds to  $G = 0.76$ . Figure 7b shows an intermediate clustering with event counts between 1 and roughly 200; here  $G = 0.49$ . Figure 7c shows the least clustered sequence where all counts are close to the mean of 50 and  $G$  is about 0.1.

The ROC diagram for the space-time partition of the Southern California seismicity shown in Fig. 6a (red) has  $G = 0.79$ . In particular, we observe that 50% (44,638) of the events are contained within only 2% (86) of non-empty voxels, and 80% (71,464) of the events are contained within

only 16% (478) of non-empty voxels. The corresponding diagrams for the marginal spatial (green) and temporal (blue) partitions of events illustrated in Figs. 1a and 1b have  $G$  values of 0.74 and 0.47, respectively. The spatial clustering is evidently stronger than the temporal clustering, and the joint space-time partition is more clustered than each of its marginals.

### ***2.3 Focusing on coupled space-time clustering***

The high degree of clustering in our analysis of Southern California seismicity (Fig. 6) is strongly affected by the concentration of seismicity around the main faults (Fig. 1a) and decay of earthquake rates with distance from the fault cores. Indeed, the space (green) and space-time (red) clustering diagrams in Fig. 6a are similar to each other, and the respective  $G$  values of 0.74 and 0.79 are close as well. This suggests that the space-time (red) curve in Fig. 6a would be convex, and the respective clustering would be high, even for a stationary rate of seismicity simply because some regions are more active (during the data acquisition period) than the others. Similarly, the clustering can be affected by the marginal time inhomogeneities. Such marginal inhomogeneities may represent the actual long-term dynamics of seismicity. Accordingly, our goal is to eliminate the effects of marginal space and time inhomogeneities on clustering and only quantify the coupled space-time variations.

The ROC framework is well suited for achieving this goal. This is illustrated in Fig. 6b where the axes are scaled in such a way that the product  $J(x,t) = S(x)T(t)$  of the marginal space  $S(x)$  and marginal time  $T(t)$  intensities of the estimated background seismicity corresponds to the diagonal. Formally, first we sort the voxel event counts in the examined process in descending order. The  $x$ -axis in Fig. 6b shows the proportion of the factorized rate  $J(x,t)$  within the most active cells of the examined process, and the  $y$ -axis shows the respective proportion of events in the examined

process. In this analysis, the marginal space and time inhomogeneities are reflected by  $J(x,t)$  and hence are mapped onto the diagonal. Only coupled space-time irregularities – short-term temporal fluctuations in local areas – cause deviation from the diagonal. The blue curve in Fig. 6b corresponds to the full catalog, and the red curve to a declustered version based on the nearest neighbor methodology of *Zaliapin and Ben-Zion (2020)*. We also use this declustering method for estimating the background activity used in calculating the factorized rate  $J(x,t)$ . Figure 6b demonstrates that the overwhelming majority of events in the full catalog occur within a very small fraction of the seismically active volume even after controlling for the marginal space and time irregularities. For example, 50% of the events occurs within 2% of weighted voxels (red circle). This reflects coupled spatio-temporal clustering mainly caused by aftershock sequences and swarms, and not the concentration of events along the fault network. The declustered catalog (red line) lacks this space-time coupling and its ROC diagram closely follows the factorized intensity  $J(x,t)$ .

The Supplementary Materials (Figs. S1-S3, Table S1) illustrate high stability of the ROC based cluster analysis with respect to the following choices: (i) lowest examined magnitude, (ii) space and time sizes of partition voxels, and (iii) catalog declustering. In particular, our analysis of Southern California seismicity with magnitude  $M \geq 2$  indicates that the cluster Gini coefficient  $G$  varies roughly within  $\pm 0.1$  unit when changing the lowest magnitude within 2 units (between 2 and 4), the number of examined events within 2 orders (between 89,000 and 800), the duration of time discretization within 2 orders (between 1 month and 10 years), the lengths of space cells within 1 order (between 3 km and 100 km), and the number of estimated background events between 10% and 25% of the full catalog. In the examined cases, the  $G$  coefficient was found to

be representative if the average number of events in non-empty voxels is above 5 and total number of voxels is above 300.

## **2.4 Clustering in other regions and ETAS model**

To further illustrate clustering properties, we analyze several other seismically active regions and synthetic seismicity of the ETAS model. The catalogs are described and the results are summarized in Table 1. Specifically, we use the USGS ComCat catalog to examine seismicity of the Western United States (WUS) using events with  $M \geq 3$  (Table 1 row 2); Central and Eastern United States (CEUS) east of Kansas and Nebraska (to exclude induced seismicity that dominates some areas) using events with  $M \geq 2.5$  (row 3); Japan using events with  $M \geq 4$  and depth less than 70 km (row 4); and Alaska using events with  $M \geq 4$  and depth less than 70 km (row 5). We use the ISC-GEM v8 catalog to examine the global seismicity (row 6). Finally, we examine a version of the ETAS model (row 7) with spatial parameters fitted for Southern California by *Gu et al.* (2013). We keep the same cluster analyses parameters as those used for the Southern California catalog, other than increasing the linear dimension of space cells to 1000 km for all catalogs except ETAS, to reflect their larger spatial extent. The Southern California results that we discussed and illustrated in earlier sections are listed in row 1 of Table 1.

All examined catalogs show similar general clustering patterns. The cluster coefficient  $G$  of a full catalog with respect to constant rate (column 9) has the highest value, which is above 0.6 in all examined cases. The variability of this cluster value with respect to the space-time resolution is mild, roughly within  $\pm 0.1$  unit. This mainly reflects complex geometry of the fault networks, concentration of earthquakes around the faults, and decay of earthquake rates off the faults. The clustering of a full catalog with respect to the factorized rate  $J(x,t)$  (column 10) has smaller, and

still well confined, values of  $G$ . The WUS and Japan catalogs have the highest coupled space-time clustering, with  $G = 0.38$  and  $0.36$ , respectively. The CEUS and Alaska show lower coupled space-time clustering, with  $G = 0.17$  and  $0.18$ , respectively. The global ISC-GEM catalog has the lowest clustering (among the examined natural seismicity cases), with  $G = 0.11$ . In the observed seismicity catalogs,  $G$  has a negative relation with the estimated proportion of background events, which is expected according to its definition. The clustering of background events with respect to the factorized rate (column 11) is always small, below  $0.05$  when using declustering with  $\alpha_0 = 0$ .

The ETAS model with parameters fitted for Southern California by *Gu et al.* (2013) has a substantially lower degree of clustering  $G$  than that of the observed Southern California seismicity (Table 1, row 7), despite comparable space, time, and magnitude ranges of the two catalogs. The cluster measure  $G$  of the full ETAS catalog with respect to the constant rate (column 9) is  $G = 0.66$ . The respective 95% CI ( $0.58, 0.71$ ) does not include the value  $G = 0.77$  that is observed in the real catalog of this region (Table 1, row 1). Even bigger difference with respect to the observed seismicity is seen in clustering of the full ETAS catalog with respect to the factorized rate  $J(x, t)$  (column 10): The average Gini coefficient in the model is merely  $G = 0.05$  with the CI of  $(-0.09, 0.32)$ , which is comparable to the  $G$  values in the declustered observed catalogs. This indicates that the examined ETAS catalog has no substantial space-time coupling beyond that dictated by inhomogeneous background distribution. This is related to the following three features of the ETAS model. (i) The model aftershock sequences are very short compared to actual aftershock sequences. This makes the degree of clustering to increase (up to  $G = 0.3$ ) for short time discretization and stay much lower for longer time discretization, decreasing the average value. (ii) The largest earthquakes occur within the most active space regions, which is not the case in actual observations. (iii) The rate of clustered events is proportional to the overall region activity,

since each event of a given magnitude has the same offspring productivity independently of its location or occurrence time. This is not the case in natural data, where different regions may have different productivity rates and clustering may change with time (especially for swarms). The last two features imply that the clustered activity in the examined version of the ETAS model provides a much better sampling of the model background space distribution, compared to how the observed clustered seismicity samples the observed background, leading to a lower model  $G$  value. The results indicate that properties of observed seismicity are significantly more inhomogeneous than those in an ETAS model with parameters fitted regionally. The ETAS parameters can be fitted adaptively to different regions or even to individual aftershock sequences (e.g. *Enescu et al.* 2009), which would likely improve the correspondence between the clustering properties of the ETAS model and observations.

### **3. Declustering earthquake catalogs**

The high spatio-temporal clustering of earthquakes illustrated in Figs. 1-6 and Table 1 implies that the statistical properties of a full catalog can be dominated by a miniscule space-time volume highlighted by aftershocks of the largest observed events. Our results suggest that clustering (triggering) might be responsible for the overwhelming majority of events in a catalog (Fig. 5). The actual clustering might be even higher than those estimated here, since some events may be triggered by earthquakes outside the examined space-time region (*van der Elst*, 2017). A full catalog strongly amplifies short-term triggering relative to the long-term interseismic background processes. Catalog declustering is aimed at removing this bias and facilitating analyses of additional, subtler, features of seismicity.

In addition to addressing *why* it might be essential to decluster earthquake catalogs, it is important to consider carefully *how* to decluster. At present there is no physical criterion that can separate independent (background) earthquakes from dependent (clustered, triggered) ones. More elementally, since the dynamics of seismicity is an unsolved problem (e.g. *Ben-Zion*, 2008), we do not know to what extent natural processes operate in these terms (background, clustered, mainshock, foreshock, etc.). Therefore, the problem of earthquake declustering has neither a definite physical nor mathematical formulation, and the desired form of a declustered catalog may depend on the application at hand. In some problems, like stress inversions of focal mechanisms (*Michael*, 1987; *Vavryčuk*, 2014; *Martínez-Garzón et al.*, 2016), it is useful to retain the largest number of background events likely to be produced by the remote tectonic loading, even if the resulting catalog remains non-stationary. In other problems, like evaluating long-term seismic hazard (*Peterson et al.*, 2018; *Llenos and Michael*, 2020), it may be desired to obtain a quasi-stationary background activity, even if this means removing additional events. There are numerous sources of non-stationarities in modern high-quality catalogs (e.g. *Hauksson et al.*, 2012; *Ross et al.*, 2019) including non-tectonic triggering (e.g., *Bettinelli et al.*, 2008; *Johnson et al.*, 2019; *Hsu et al.*, 2021), non-stationary postseismic relaxation below the seismogenic crust (e.g., *Ben-Zion et al.*, 1993; *Pollitz et al.*, 2014), and various catalog uncertainties (e.g. *Kagan*, 2003; *Hauksson et al.*, 2012; *Zaliapin and Ben-Zion*, 2015). Accordingly, a stationary declustered catalog – a renowned Poisson process of background events discussed in the classic works of *Gardner and Knopoff* (1964) and *Reasenber* (1985) – is not necessarily an appropriate target. More generally, it seems that any hard and fast statistical criterion of a declustered catalog can be easily falsified with modern tools, yet often in a non-informative way (*Luen and Stark*, 2013). Since there is no unique physical theory of earthquake clustering, there is no unique “right solution” to the



declustering problem. The statistical and physical criteria of a desirable declustering for a given problem should meet in the middle. Declustering should be viewed as an interactive process, allowing an expert user to resolve a trade-off between the number of background events and their emerging properties (e.g. time homogeneity, coupled space-time clustering, etc.). Examining a range of possible backgrounds, and the implied results for a problem at hand, may be more informative than a judicious selection of a single “correct” declustering. Quoting *Utsu* (1969), “... *a definition of aftershocks uniformly applicable to all kinds of aftershock problems is difficult to establish, and a suitable working definition must be formulated according to the character of the problem and the data employed.*”. Increasingly comprehensive data sets, statistical techniques, and computing power provide new opportunities and challenges for making suitable working definitions for different purposes.

Declustering significantly modifies the catalog and may alter its key statistical attributes, such as *b*-value (relative proportion of large and small events), *a*-value (earthquake annual rate), fractal dimension of hypocenters, relative rates of events in different regions, etc. (e.g. *Knopoff*, 2000; *van Stiphout et al.*, 2011; *Llenos and Michael*, 2020; *Taroni and Akinici*, 2021, *Mizrahi et al.*, 2021). Different types of declustering affect various attributes to different extents. For example, declustering may start with identifying clusters and then select a single event from each cluster as background (and remove all other events as clustered). The two most popular choices are the *first* event or the *largest* event. Selecting the *first* event from each cluster might not significantly affect the catalog *b*-value (e.g. *Llenos and Michael*, 2020), but in this case the largest event in the catalog may not belong to the background. Selecting the *largest* event from each cluster does recognize the largest events as background (mainshocks), but this will necessarily deflate the *b*-value increasing the relative proportion of large events (*Mizrahi et al.*, 2021). As mentioned, the choice

of declustering method and its implementation (e.g. parameter selection) should be made by the user. An informed decision would consider various aspects of the examined catalog – the maximal magnitude, the range of magnitudes, level of completeness, inhomogeneity of spatial distribution, degree of clustering, regional cluster style – as well as how the declustered catalog will be used. The ROC-based cluster framework presented here might be a useful tool in assessing the degree of clustering and guiding a user in quantifying an appropriate degree of declustering suitable for the problem at hand.

There are several approaches (and respective algorithms) for declustering seismicity, each with its pros and cons. We refer to *van Stiphout et al.* (2012) for a technical review of the methods up to that time. The classical window technique initiated by *Knopoff* (1964) and *Gardner and Knopoff* (1974) and its ramifications (e.g. *Uhrhammer*, 1986; *Rosson et al.*, 2019) are intuitive and simple to implement; this approach is traditionally used in the USGS National Seismic Hazard Map (*Petersen et al.*, 2018). There is an increasing awareness, however, that window techniques can be incommensurable with the quality and detail of modern catalogs (e.g., *Teng and Baker*, 2019; *Peresan and Gentili*, 2020). The most obvious drawback is the appearance of holes – regions with no activity – around large earthquakes, which may be inconvenient for further analyses. A more elaborate technique of *Reasenber*g (1985) connects events in clusters according to adaptive space-time interaction zones and provided good performance in California (for which it has been originally developed) and other regions. This technique has several parameters that might be region- and catalog-dependent.

Approaches based on the ETAS model (e.g. *Zhuang et al.*, 2002; *Console et al.*, 2010; *Llenos and Michael*, 2020) have mathematical appeal and effective for treating the major aftershock sequences (e.g. *Woessner et al.*, 2011; *Field et al.*, 2017; *Shcherbakov*, 2020), but can have many

parameters and be computationally demanding (e.g. *Veen and Schoenberg*, 2008). The ETAS model assumes a rigidly parameterized conditional intensity of earthquake clusters, so its appropriateness depends on how well this assumption is met by the examined data and how well one can estimate the multiple model parameters (*Ogata*, 1999; *Wang et al.*, 2010; *Schoenberg*, 2013; *Seif et al.*, 2017). A more flexible approach of *Marsan and Lengline* (2008) offers a possibility to work with an arbitrary (but fixed for the entire catalog) cluster model. Declustering is reduced here to estimating the triggering kernel, which is done by recursive approximation. The declustering method of *Zaliapin and Ben-Zion* (2020) used in the present work is more exploratory and parsimonious in spirit. It uses a non-parametric definition of a cluster, which is quantified using the nearest-neighbor earthquake proximity (*Baiesi and Paczuski*, 2004; *Zaliapin and Ben-Zion* 2013, 2016). The method is highly robust with respect to its parameters and various catalog imperfections, and allows users to select different degree of declustering (different number of background events) while preserving the relative spatial rates of events in different regions. The non-parametric nature of this technique makes it possible for different forms of clustering to co-exist in different parts of the catalog; at the same time, the technique can closely reconstruct a time-stationary space-inhomogeneous Poisson background of the parametric ETAS model (*Zaliapin and Ben-Zion*, 2020). This flexibility comes at a price of not having an optimization procedure, as is common in non-parametric methods (e.g. *Wasserman*, 2006).

Alternative declustering techniques can produce comparable results in selected problems and data sets. For example, the nearest-neighbor declustering of *Zaliapin and Ben-Zion* (2013, 2020) produces results comparable with that of (i) an ETAS-based ROBERE method when used for probabilistic seismic hazard assessment in the induced seismicity of Oklahoma and Kansas and tectonic seismicity of the San Francisco Bay Area (*Llenos and Michael*, 2020), (ii) *Reasenbergs*

(1985) declustering and hazard analysis of Oklahoma and Kansas (*Teng and Baker, 2019*), (iii) ETAS stochastic declustering on *Zhuang et al. (2002)* in seismicity of Northeastern Italy and Western Slovenia (*Varini et al., 2002*), and some other methods and regions (e.g. *Peresan and Gentili, 2020; Mizrahi et al., 2021*). At the same time, other problems and/or data sets may be more sensitive to specific aspects of the used declustering method (e.g. *van Stiphout et al., 2011; Mizrahi et al., 2021*).

A useful perspective on declustering is provided by the existence of two characteristic space-time scales in earthquake dynamics, that was illustrated in the bimodal distributions in Fig. 4. Multiple declustering algorithms successfully and consistently eliminate the bulk of the cluster mode (even if this mode is not explicitly defined/used in the declustering procedure) and keep the bulk of the background mode, which is facilitated by the above-mentioned scale separation. Some declustering methods still leave a collection of events with quite complex space-time patterns and remaining clustering. The main differences among alternative declustering approaches concern treating the neighborhood of the red line in Fig. 4 that separates the modes. Notwithstanding the overall scale separation, this neighborhood includes a substantial proportion of the catalog so the results of different procedures may differ significantly.

#### 4. Discussion

This paper illustrates the general existence of strong coupled space-time clustering of natural earthquakes that is retained after controlling for the marginal space and time inhomogeneities of observed catalogs. The ROC-based Gini coefficient  $G$  provides an efficient and stable tool for quantifying the degree of coupled space-time clustering. The degree of earthquake clustering is substantially stronger than what might be suggested by visual inspections (Fig. 3) and modeling

based on the ETAS framework with parameters based on a regional fitting (Table 1). Problems aiming at uncovering additional feature of seismicity beyond the strong clustering dominated by aftershocks of the largest events or expected long-term properties can benefit from catalog declustering prior to the analyses.

The ROC diagram with respect to the factorized intensity  $J(x,t)$  provides a convenient initial assessment of the space-time clustering in a catalog (Fig. 6b). The large values (close to 1) of  $G$  indicate that a large proportion of events is concentrated in a small fraction of the examined space-time volume (the entire ROC diagram gives more complete information about this concentration). This implies that some statistical analyses might be biased, having an amplified focus on the concentration regions and not reflecting general properties of examined data sets. The ROC approach may also facilitate assessing the quality of declustering – a properly declustered catalog would correspond to an almost-diagonal ROC diagram and a small degree of clustering value  $G$ . Notably, this approach is not aimed at a stationary declustered catalog, which might be an improper target, but only assesses factorization of the earthquake rates. We propose the measure  $G$  of clustering (and other possible metrics of the ROC diagram) as a complement to the currently used measures of declustering quality (*Luen and Stark, 2012; Zaliapin and Ben-Zion, 2020*). More generally, we propose that the quality of declustering should be accessed not via the properties of the final product (e.g. stationary sequence of events), but via the degree to which the catalog inhomogeneities, and related biases, have been removed.

The ROC-based cluster measure  $G$ , quantifying the degree of non-uniformity in a space-time distribution of earthquakes, is related to the Shannon entropy (*Shannon, 1948*), which is defined for a distribution  $p(i)$  as  $H = -\sum p(i) \log p(i)$  and has the same informal interpretation. The main difference between the entropy  $H$  and clustering  $G$  is that the latter is much less affected by the

number of bins in the distribution (the total number of voxels in our case). Figure 8 compares the entropy  $H$  and clustering  $G$  in a synthetic experiment that produces 400 measures with different degree of non-uniformity for each of the bin counts  $N = 3, 10, 30, 100, 300$ , and 1000. As seen, for each given  $N$  the entropy and clustering are closely related, but the entropy is affected by the number of bins, while clustering is not sensitive to this parameter. The two measures have a negative (non-linear) relation, since  $G$  increases and  $H$  decreases for a higher degree of non-uniformity. We also observe that  $H$  has weaker discriminating power for low  $G$  values, which is reflected in almost zero slopes of the curves in Fig. 8 at  $G < 0.2$ .

The increasing quality of earthquake records and the complexity of problems associated with seismicity require refined and more sophisticated methods for earthquake cluster analyses, since space-time clustering dominate natural earthquake catalogs. Declustered versions of seismicity are needed for various analyses. Until there are accepted solutions for the physics and mathematics behind different forms of earthquake clustering, the used declustering approaches remain intrinsically heuristic and exploratory. Simplicity of implementation and flexibility of output have some advantages at the current stage, but ultimately the appropriate form of the resulting product (declustered catalog) depend on the quality of the full catalog and problem at hand.

## **Data and Resources**

The relocated catalog of *Hauksson et al.* (2012) is available from the Southern California Earthquake Data Center: <https://scedc.caltech.edu/data/alt-2011-dd-hauksson-yang-shearer.html>. The USGS ComCat catalog is available at <https://earthquake.usgs.gov/data/comcat/>. The International Seismological Centre ISC-GEM Earthquake Catalogue is available at <https://doi.org/10.31905/d808b825>. The ETAS catalog is available in the Electronic Supplement

of *Gu et al.* (2013): <https://doi.org/10.1002/jgrb.50306>. The Supplemental Material describes stability analysis of the cluster measure  $G$  using data from the Southern California.

## Acknowledgements

The study was supported by the National Science Foundation, (Grants EAR-1723033 and EAR-1722561) and the Southern California Earthquake Center (based on NSF Cooperative Agreement EAR-1600087 and USGS Cooperative Agreement G17 AC00047). The paper benefitted from comments by an anonymous referee and an anonymous associate editor.

## Declaration of Competing Interests

The authors declare no competing interests.

## References

- Abolfathian, N., Martínez-Garzón, P., and Ben-Zion, Y. (2019) Spatiotemporal variations of stress and strain parameters in the San Jacinto fault zone. *Pure and Applied Geophysics*, 176(3), 1145–1168. <https://doi.org/10.1007/s00024-018-2055-y>
- Amos, C. B., Audet, P., Hammond, W. C., Bürgmann, R., Johanson, I. A., & Blewitt, G. (2014) Uplift and seismicity driven by groundwater depletion in central California. *Nature*, 509(7501), 483-486.
- Baiesi, M. and Paczuski, M. (2004) Scale-free networks of earthquakes and aftershocks. *Physical review E*, 69(6), 066106. <https://doi.org/10.1103/PhysRevE.69.066106>

524 Ben-Zion, Y. (2008) Collective Behavior of Earthquakes and Faults: Continuum-Discrete  
 525 Transitions, Evolutionary Changes and Corresponding Dynamic Regimes, *Rev.*  
 526 *Geophysics*, 46, RG4006, doi:10.1029/2008RG000260.

527 Ben-Zion, Y., Rice, J. R., and Dmowska, R. (1993) Interaction of the San Andreas fault creeping  
 528 segment with adjacent great rupture zones, and earthquake recurrence at Parkfield.  
 529 *Journal of Geophysical Research*, 98(B2), 2135–2144.  
 530 <https://doi.org/10.1029/92JB02154>

531 Ben-Zion, Y. and Zaliapin, I. (2020) Localization and coalescence of seismicity before large  
 532 earthquakes. *Geophysical Journal International*, 223(1), 561-583.

533 Bettinelli, P., J.P. Avouac, M. Flouzat, L. Bollinger, G. Ramillien, S. Rajaure, and S. Sapkota,  
 534 (2008) Seasonal variations of seismicity and geodetic strain in the Himalaya induced by  
 535 surface hydrology, *Earth and Planetary Science Letters*, 266, 332-344.

536 Cochran, E. S., Vidale, J. E., and Tanaka, S. (2004) Earth tides can trigger shallow thrust fault  
 537 earthquakes. *Science*, 306(5699), 1164-1166.

538 Console, R., Murru, M. and Catalli, F. (2006) Physical and stochastic models of earthquake  
 539 clustering. *Tectonophysics*, 417(1-2), 141-153.

540 Console, R., Jackson, D.D. and Kagan, Y.Y. (2010) Using the ETAS model for catalog  
 541 declustering and seismic background assessment. *Pure and applied geophysics*, 167(6),  
 542 819-830.

543 Di Giacomo, D., Bondár, I., Storchak, D.A., Engdahl, E.R., Bormann, P. and Harris, J. (2015)  
 544 ISC-GEM: Global Instrumental Earthquake Catalogue (1900–2009), III. Re-computed  
 545 MS and mb, proxy MW, final magnitude composition and completeness assessment.  
 546 *Physics of the Earth and Planetary Interiors*, 239, 33-47.



547 Di Giacomo, D., E.R. Engdahl and D.A. Storchak (2018) The ISC-GEM Earthquake Catalogue  
 548 (1904–2014): status after the Extension Project, *Earth Syst. Sci. Data*, 10, 1877-1899,  
 549 doi: 10.5194/essd-10-1877-2018.

550 Dieterich, J. (1994) A constitutive law for rate of earthquake production and its application to  
 551 earthquake clustering. *Journal of Geophysical Research: Solid Earth*, 99(B2), 2601-  
 552 2618.

553 Dieterich, J.H. and Smith, D.E. (2009) Nonplanar faults: Mechanics of slip and off-fault damage.  
 554 In *Mechanics, structure and evolution of fault zones* (pp. 1799-1815). Birkhäuser Basel.

555 Enescu, B., Hainzl, S. and Ben-Zion, Y. (2009) Correlations of seismicity patterns in Southern  
 556 California with surface heat flow data. *Bulletin of the Seismological Society of America*,  
 557 99(6), 3114-3123.

558 Felzer, K.R. and Brodsky, E.E. (2006) Decay of aftershock density with distance indicates  
 559 triggering by dynamic stress. *Nature*, 441(7094), 735-738.

560 Field, E. H., K. R. Milner, J. L. Hardebeck, M. T. Page, N. van der Elst, T. H. Jordan, T. H., A. J.  
 561 Michael, B. E. Shaw, and M. J. Werner (2017) A spatiotemporal clustering model for the  
 562 third Uniform California Earthquake Rupture Forecast (UCERF3-ETAS): Toward an  
 563 operational earthquake forecast. *Bulletin of the Seismological Society of America*, 107(3),  
 564 1049-1081.

565 Gardner, J. K., and Knopoff, L. (1974) Is the sequence of earthquakes in Southern California,  
 566 with aftershocks removed, Poissonian? *Bulletin of the Seismological Society of America*,  
 567 64(5), 1363–1367.

568 Goebel, T.H.W., Candela, T., Sammis, C.G., Becker, T.W., Dresen, G. and Schorlemmer, D.,  
 569 (2014) Seismic event distributions and off-fault damage during frictional sliding of saw-

570 cut surfaces with pre-defined roughness. *Geophysical Journal International*, 196(1), 612–  
571 625.

572 Goebel, T. H. W. and Brodsky, E. E. (2018) The spatial footprint of injection wells in a global  
573 compilation of induced earthquake sequences. *Science*, 361(6405), 899–904.  
574 <https://doi.org/10.1126/science.aat5449>

575 Gu, C., Schumann, A.Y., Baiesi, M. and Davidsen, J. (2013) Triggering cascades and statistical  
576 properties of aftershocks. *Journal of Geophysical Research: Solid Earth*, 118(8), 4278–  
577 4295.

578 Hammond, W. C., Kreemer, C., Zaliapin, I., & Blewitt, G. (2019) Drought-triggered magmatic  
579 inflation, crustal strain and seismicity near the Long Valley Caldera, Central Walker  
580 Lane. *Journal of Geophysical Research: Solid Earth*, 124, 6072–6091.  
581 <https://doi.org/10.1029/2019JB017354>

582 Hauksson, E., Yang, W., and Shearer, P. M. (2012) Waveform relocated earthquake catalog for  
583 Southern California (1981 to June 2011). *Bulletin of the Seismological Society of*  
584 *America*, 102(5), 2239–2244. <https://doi.org/10.1785/0120120010>

585 Hsu, Y.J., Kao, H., Bürgmann, R., Lee, Y.T., Huang, H.H., Hsu, Y.F., Wu, Y.M. and Zhuang, J.,  
586 (2021) Synchronized and asynchronous modulation of seismicity by hydrological  
587 loading: A case study in Taiwan. *Science Advances*, 7(16), p.eabf7282.

588 Johnson, C. W., Fu, Y. and Bürgmann, R. (2019) Hydrospheric modulation of stress and  
589 seismicity on shallow faults in southern Alaska. *Earth and Planetary Science Letters*,  
590 530, 115904. <https://doi.org/10.1016/j.epsl.2019.115904>

591 Jones, L.M. and Molnar, P. (1979) Some characteristics of foreshocks and their possible  
592 relationship to earthquake prediction and premonitory slip on faults. *Journal of*  
593 *Geophysical Research: Solid Earth*, 84(B7), 3596-3608.

594 Kagan, Y.Y. (2003) Accuracy of modern global earthquake catalogs. *Physics of the Earth and*  
595 *Planetary Interiors*, 135(2-3),173-209.

596 Kagan, Y.Y. and Jackson, D.D. (1991) Long-term earthquake clustering. *Geophysical Journal*  
597 *International*, 104(1), 117-133.

598 Kagan, Y.Y. and Jackson, D.D. (2000) Probabilistic forecasting of earthquakes. *Geophysical*  
599 *Journal International*, 143(2), 438-453.

600 Kato, A. and Y. Ben-Zion (2021) The generation of large earthquakes. *Nature Reviews Earth &*  
601 *Environment*, 2, 26–39, <https://doi.org/10.1038/s43017-020-00108-w>

602 Knopoff, L. (1964). The statistics of earthquakes in Southern California. *Bulletin of the*  
603 *Seismological Society of America*, 54(6A), 1871-1873.

604 Knopoff, L. (2000) The magnitude distribution of declustered earthquakes in Southern  
605 California. *Proceedings of the National Academy of Sciences*, 97(22), 11880-11884.

606 Lengliné, O., B. Enescu, Z. Peng, and K. Shiomi (2012) Decay and expansion of the early  
607 aftershock activity following the 2011, Mw9.0 Tohoku earthquake, *Geophys. Res. Lett.*,  
608 39, L18309, doi:10.1029/2012GL052797.

609 Llenos, A. L., and Michael, A. J. (2020) Regionally Optimized Background Earthquake Rates  
610 from ETAS (ROBERE) for Probabilistic Seismic Hazard Assessment, *Bull. Seismol. Soc.*  
611 *Am.* 110, 1172–1190, doi: 10.1785/0120190279

612 Luen, B. and Stark, P. B. (2012) Poisson tests of declustered catalogues. *Geophysical Journal*  
613 *International*, 189(1), 691–700. <https://doi.org/10.1111/j.1365-246X.2012.05400.x>

614 Marsan, D. and Lengline, O. (2008) Extending earthquakes' reach through cascading. *Science*,  
615 319(5866), 1076–1079. <https://doi.org/10.1126/science.1148783>

616 Martínez-Garzón, P., Ben-Zion, Y., Abolfathian, N., Kwiatak, G., & Bohnhoff, M. (2016). A  
617 refined methodology for stress inversions of earthquake focal mechanisms. *Journal of*  
618 *Geophysical Research: Solid Earth*, 121, 8666–8687.  
619 <https://doi.org/10.1002/2016JB013493>

620 Martínez-Garzón, P., Vavryčuk, V., Kwiatak, G. and Bohnhoff, M. (2016) Sensitivity of stress  
621 inversion of focal mechanisms to pore pressure changes. *Geophysical Research Letters*,  
622 43(16), 8441-8450.

623 Michael, A.J. (1987) Use of focal mechanisms to determine stress: a control study. *Journal of*  
624 *Geophysical Research: Solid Earth*, 92(B1), pp.357-368.

625 Michael, A.J. (2014) How complete is the ISC-GEM global earthquake catalog? *Bulletin of the*  
626 *Seismological Society of America*, 104(4), 1829-1837.

627 Mignan, A. (2014) The debate on the prognostic value of earthquake foreshocks: A meta-  
628 analysis. *Scientific reports*, 4(1), 1-5.

629 Mizrahi, L., Nandan, S. and Wiemer, S. (2021) The Effect of Declustering on the Size  
630 Distribution of Mainshocks. *Seismological Research Letters*.  
631 <https://doi.org/10.1785/0220200231>

632 Ogata, Y. (1999) Seismicity analysis through point-process modeling: A review. *Seismicity*  
633 *patterns, their statistical significance and physical meaning*, 471-507.

634 Peresan, A. and Gentili, S. (2020) Identification and characterisation of earthquake clusters: a  
635 comparative analysis for selected sequences in Italy and adjacent regions. *Bollettino di*  
636 *Geofisica Teorica ed Applicata*, 16(2).

637 Petersen, M. D., Moschetti, M. P., Powers, P. M., Mueller, C. S., Haller, K. M., Frankel, A. D.,  
 638 et al. (2015). The 2014 United States national seismic hazard model. *Earthquake Spectra*,  
 639 31(1\_suppl), S1–S30. <https://doi.org/10.1193/120814EQS210M>  
 640 Petersen, M. D., Mueller, C. S., Moschetti, M. P., Hoover, S. M., Rukstales, K. S., McNamara,  
 641 D. E., et al. (2018). 2018 One-year seismic hazard forecast for the Central and Eastern  
 642 United States from induced and natural earthquakes. *Seismological Research Letters*,  
 643 89(3), 1049–1061. <https://doi.org/10.1785/0220180005>  
 644 Pollitz, F.F., Bürgmann, R., Stein, R.S., and Sevilgen, V., The Profound Reach of the 11 April  
 645 2012 M 8.6 Indian Ocean Earthquake: Short-Term Global Triggering Followed by a  
 646 Longer-Term Global Shadow (2014) *Bull. Seismol. Soc. Am.*, doi: 10.1785/0120130078  
 647 Powers, P.M. and Jordan, T.H. (2010) Distribution of seismicity across strike-slip faults in  
 648 California. *Journal of Geophysical Research: Solid Earth*, 115(B5).  
 649 Press, F. and Allen, C. (1995) Patterns of seismic release in the southern California region.  
 650 *Journal of Geophysical Research: Solid Earth*, 100(B4), 6421-6430.  
 651 Reasenber, P. (1985) Second-order moment of central California seismicity, 1969–1982, *J.*  
 652 *Geophys. Res.*, 90, 5479–5496.  
 653 Romanowicz, B. (1993) Spatiotemporal patterns in the energy release of great earthquakes.  
 654 *Science*, 260(5116), 1923-1926.  
 655 Ross, Z. E., C. Rollins, E. S. Cochran, E. Hauksson, J-P. Avouac and Y. Ben-Zion (2017)  
 656 Aftershocks driven by afterslip and fluid pressure sweeping through a fault-fracture  
 657 mesh, *Geophys. Res. Lett.*, 44, 8260-8267, doi: 10.1002/2017GL074634.  
 658 Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019) Searching for hidden  
 659 earthquakes in Southern California. *Science*, 364(6442), 767-771.

660 Rosson, Z., J. I. Walter, T. Goebel, and X. Chen (2019) Narrow spatial aftershock zones for  
 661 induced earthquake sequences in Oklahoma, *Geophys. Res. Lett.* 46, 10,358–10,366, doi:  
 662 10.1029/2019gl083562

663 Ruhl, C. J., R. E. Abercrombie, K. D. Smith, and I. Zaliapin (2016) Complex spatiotemporal  
 664 evolution of the 2008 Mw 4.9 Mogul earthquake swarm (Reno, Nevada): Interplay of  
 665 fluid and faulting. *Journal of Geophysical Research: Solid Earth*, 121(11), 8196-8216.

666 Schoenberg, F.P. (2013) Facilitated estimation of ETAS. *Bulletin of the Seismological Society of*  
 667 *America*, 103(1), 601-605.

668 Seif, S., Mignan, A., Zechar, J.D., Werner, M.J. and Wiemer, S. (2017) Estimating ETAS: The  
 669 effects of truncation, missing data, and model assumptions. *Journal of Geophysical*  
 670 *Research: Solid Earth*, 122(1), 449-469.

671 Shannon, C.E. (1948) A mathematical theory of communication. *The Bell system technical*  
 672 *journal*, 27(3), 379-423.

673 Shcherbakov, R. (2020) Statistics and Forecasting of Aftershocks during the 2019 Ridgecrest,  
 674 California, Earthquake Sequence. *Journal of Geophysical Research: Solid Earth*,  
 675 e2020JB020887.

676 Taroni, M. and Akinici, A. (2021) Good practices in PSHA: declustering, b-value estimation,  
 677 foreshocks and aftershocks inclusion; a case study in Italy. *Geophysical Journal*  
 678 *International*, 224(2), 1174-1187.

679 Teng, G., and J. W. Baker (2019). Seismicity declustering and hazard analysis of the Oklahoma-  
 680 Kansas region, *Bull. Seismol. Soc. Am.* 109, 2356–2366, doi: 10.1785/0120190111.

681 Utsu, T. (1969). Aftershocks and earthquake statistics I, *J. Fac. Sci.*, 3, 129-195, Hokkaido  
 682 University of Geophysics.

683 Utsu, T. and Ogata, Y. (1995) The centenary of the Omori formula for a decay law of aftershock  
 684 activity. *Journal of Physics of the Earth*, 43(1), 1–33.

685 Utsu, T. (2002) Statistical features of seismology, in *International Handbook of Earthquake and*  
 686 *Engineering Seismology*, eds. W. HK Lee, H. Kanamori, P. C. Jennings, and C.  
 687 Kisslinger, *Part A*, 719-732, Academic Press.

688 van der Elst, N.J., (2017) Accounting for orphaned aftershocks in the earthquake background  
 689 rate, *Geophysical Journal International*, 211(2), 1108–1118.

690 van Stiphout, T., Schorlemmer, D. and Wiemer, S. (2011) The effect of uncertainties on  
 691 estimates of background seismicity rate. *Bulletin of the Seismological Society of America*,  
 692 101(2), 482-494.

693 van Stiphout, T., J. Zhuang, and D. Marsan (2012), Seismicity declustering, *Community Online*  
 694 *Resource for Statistical Seismicity Analysis*, doi:10.5078/corssa-52382934.

695 Storchak, D.A., D. Di Giacomo, I. Bondár, E.R. Engdahl, J. Harris, W.H.K. Lee, A. Villaseñor  
 696 and P. Bormann (2013) Public Release of the ISC-GEM Global Instrumental Earthquake  
 697 Catalogue (1900-2009). *Seism. Res. Lett.*, 84, 5, 810-815, doi: 10.1785/0220130034.

698 Storchak, D.A., D. Di Giacomo, E.R. Engdahl, J. Harris, I. Bondár, W.H.K. Lee, P. Bormann and  
 699 A. Villaseñor (2015) The ISC-GEM Global Instrumental Earthquake Catalogue (1900-  
 700 2009): Introduction, *Phys. Earth Planet. Int.*, 239, 48-63, doi: 10.1016/j.pepi.2014.06.009.

701 Varini, E., Peresan, A. and Zhuang, J. (2020) Topological Comparison Between the Stochastic  
 702 and the Nearest-Neighbor Earthquake Declustering Methods Through Network Analysis.  
 703 *Journal of Geophysical Research: Solid Earth*, 125(8), e2020JB019718.

704 Vavryčuk, V., 2014. Iterative joint inversion for stress and fault orientations from focal  
 705 mechanisms. *Geophysical Journal International*, 199(1), 69-77.

706 Veen, A. and Schoenberg, F. P. (2008) Estimation of space–time branching process models in  
707 seismology using an EM–type algorithm. *Journal of the American Statistical Association*,  
708 *103*(482), 614-624.

709 Uhrhammer, R. (1986) Characteristics of Northern and Central California Seismicity,  
710 *Earthquake Notes*, *57*(1), 21.9

711 Wang, Q., Jackson, D.D. and Zhuang, J. (2010) Missing links in earthquake clustering models.  
712 *Geophysical Research Letters*, *37*(21).

713 Wang, Q., Schoenberg, F.P. and Jackson, D.D. (2010) Standard errors of parameter estimates in  
714 the ETAS model. *Bulletin of the Seismological Society of America*, *100*(5A), 1989-2001.

715 Wasserman, L. (2006) *All of nonparametric statistics*. Springer Science & Business Media.

716 Woessner, J., Hainzl, S., Marzocchi, W., Werner, M. J., Lombardi, A. M., Catalli, F., Enescu, B.  
717 Cocco, M., Gerstenberger, M. C., and Wiemer, S. (2011) A retrospective comparative  
718 forecast test on the 1992 Landers sequence. *Journal of Geophysical Research: Solid*  
719 *Earth*, *116*(B5).

720 Zaliapin, I., Gabrielov, A., Keilis-Borok, V. and Wong, H. (2008) Clustering analysis of  
721 seismicity and aftershock identification. *Physical review letters*, *101*(1), 018501.

722 Zaliapin, I. and Ben-Zion, Y. (2013) Earthquake clusters in southern California I: Identification  
723 and stability. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2847-2864.

724 Zaliapin, I. and Ben-Zion, Y. (2015) Artefacts of earthquake location errors and short-term  
725 incompleteness on seismicity clusters in southern California. *Geophysical Journal*  
726 *International*, *202*(3), 1949-1968.

727 Zaliapin, I. and Ben-Zion, Y. (2016) A global classification and characterization of earthquake  
728 clusters. *Geophysical Journal International*, *207*(1), 608-634.



729 Zaliapin, I. and Ben-Zion, Y. (2020) Earthquake declustering using the nearest-neighbor  
730 approach in space-time-magnitude domain. *Journal of Geophysical Research: Solid*  
731 *Earth*, 125(4), e2018JB017120.

732 Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002) Stochastic declustering of space-time  
733 earthquake occurrences. *Journal of the American Statistical Association*, 97(458), 369–  
734 380. <https://doi.org/10.1198/016214502760046925>

**Table 1: Summary of cluster analyses** in Southern California, Western United States (WUS), Central and Eastern United States (CEUS), Alaska, Japan, global ISC-GEM catalog, and Epidemic Type Aftershock (ETAS) model.

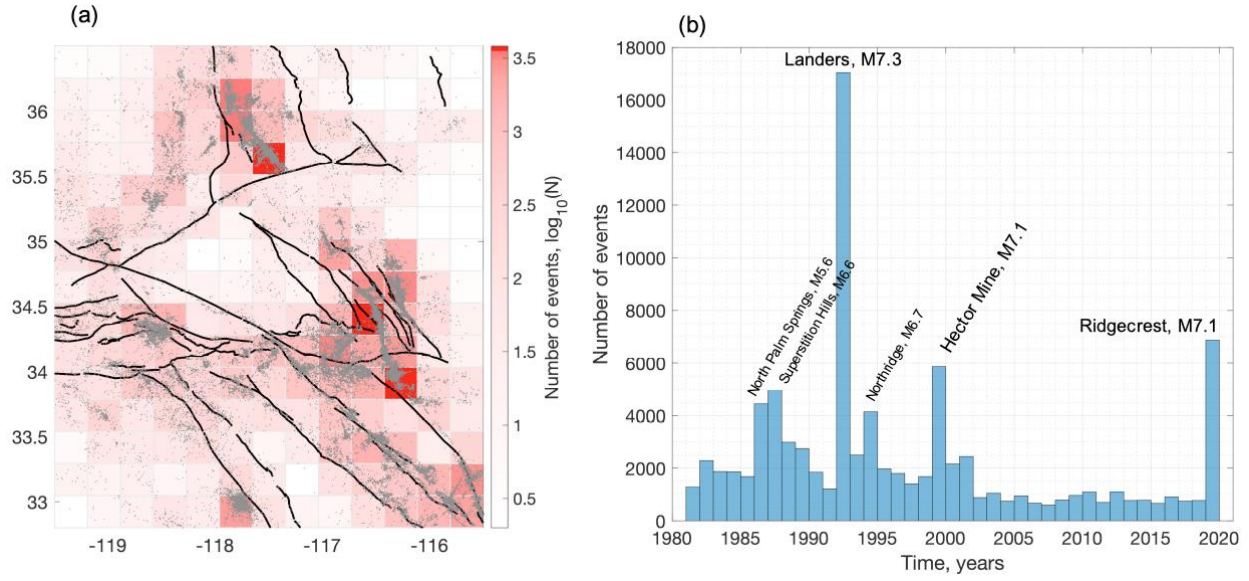
1	2	3	4	5	6	7	8	9	10	11
#	Catalog	Space range	Time range	Magnitude range	No. events	$\alpha_0^*$	No. background events <sup>†</sup>	Clustering w.r.t. constant rate <sup>‡</sup> , mean $G$ (95% CI)	Clustering w.r.t. factorized rate $J(x,t)^{\ddagger}$ , mean $G$ (95% CI)	Background clustering w.r.t. factorized rate $J(x,t)^{\ddagger}$ , mean $G$ (95% CI)
1	HYS SoCal	119.5 – 115.5 W 32.8 – 36.5 N	1981 – 2020	2.0 – 7.3	89,341 (100%)	0	$15,914 \pm 110$ (17.8 $\pm$ 0.1 %)	<b>0.77</b> (0.7, 0.82)	<b>0.55</b> (0.35, 0.63)	<b>0.031</b> (0, 0.06)
2	ComCat WUS	126 – 110 W 30 – 50 N	1981 – 2021	3.0 – 7.3	30,483 (100%)	0	$8,494 \pm 72$ (27.9 $\pm$ 0.24 %)	<b>0.74</b> (0.59, 0.82)	<b>0.38</b> (0.2, 0.48)	<b>0.029</b> (0, 0.07)
3	ComCat CEUS	95 – 60 W 30 – 50 N	1981 – 2021	2.5 – 5.9	2,591 (100%)	0	$1,172 \pm 27$ (45.2 $\pm$ 1.03 %)	<b>0.63</b> (0.51, 0.71)	<b>0.17</b> (0.07, 0.28)	<b>0.029</b> (0, 0.06)
4	ComCat Japan	138 – 146 W 32 – 44 N	1974 – 2021	4.0 – 9.1	17,823 (100%)	0	$5,885 \pm 59$ (33.0 $\pm$ 0.33 %)	<b>0.67</b> (0.57, 0.72)	<b>0.39</b> (0.07, 0.51)	<b>0.044</b> (0, 0.09)
5	ComCat Alaska	180 – 120 E 45 – 75 N	1973 – 2020	4.0 – 8.0	15,008 (100%)	0	$6,682 \pm 56$ (44.5 $\pm$ 0.37 %)	<b>0.71</b> (0.6, 0.81)	<b>0.18</b> (0.06, 0.29)	<b>0.002</b> (–0.02, 0.02)
6	ISC-GEM	World	1976 – 2017	5.8 – 9.3	7,066 (100%)	0	$3,680 \pm 41$ (52.1 $\pm$ 0.58 %)	<b>0.64</b> (0.57, 0.72)	<b>0.11</b> (0.07, 0.16)	<b>0.006</b> (0, 0.03)
7	ETAS	600 $\times$ 600 km	20.9 years	2.5 – 7.3	26,986 (100%)	0	$7,130 \pm 73$ (26.4 $\pm$ 0.27 %)	<b>0.66</b> (0.58, 0.71)	<b>0.05</b> (–0.09, 0.32)	<b>0.002</b> (–0.02, 0.03)

\* Cluster threshold that controls the number of background events (*Zaliapin and Ben-Zion, 2020*).

<sup>†</sup> Mean and 95% confidence interval (CI) according to  $10^4$  realizations of stochastic declustering. The CI is estimated as  $\pm 1.96$  sample standard deviation.

<sup>‡</sup> Estimation is done for the space-time partitions with the average of at least 5 events per non-empty voxel and at least 300 voxels total. The CI limits are the 0.025 and 0.975 sample quantiles.

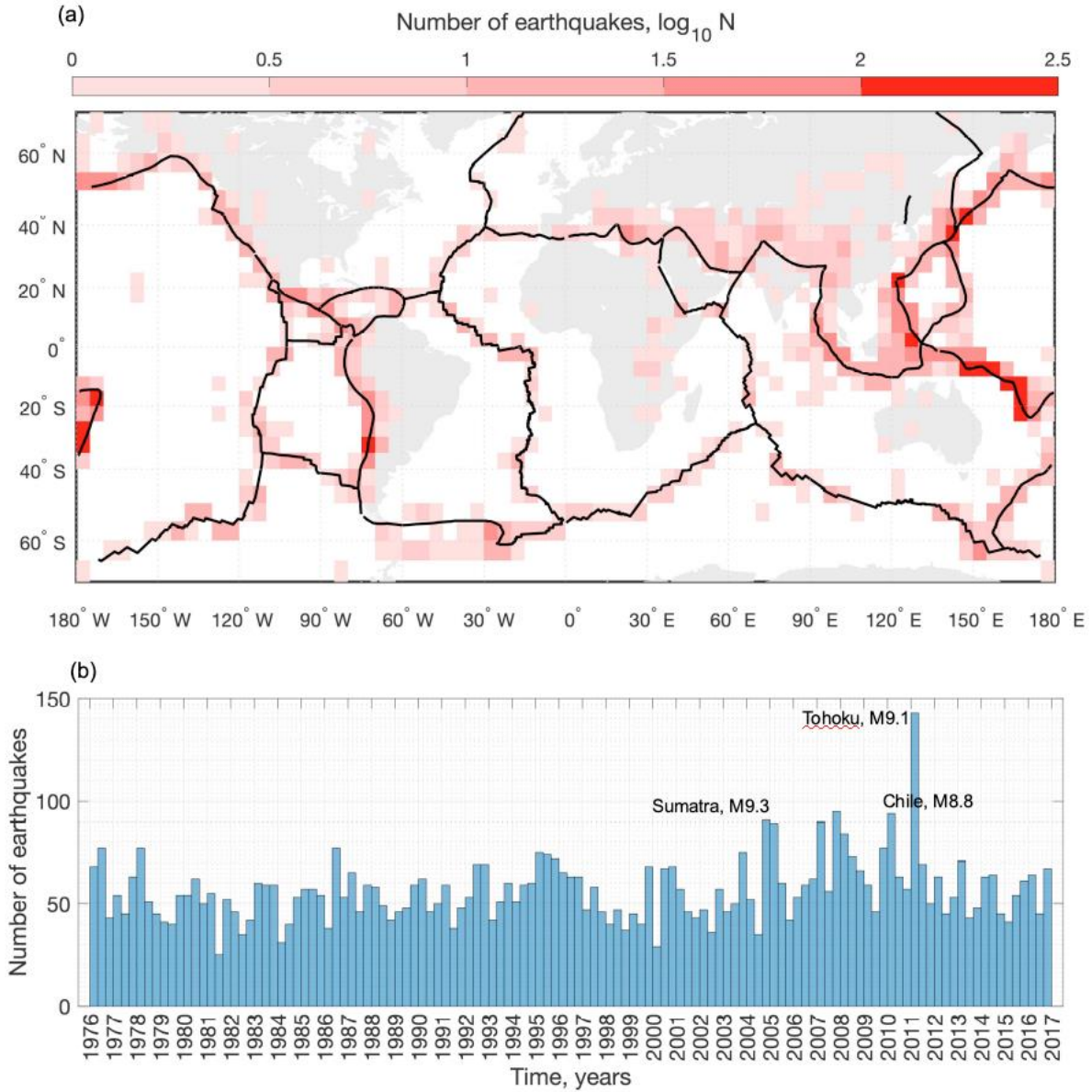
743



744

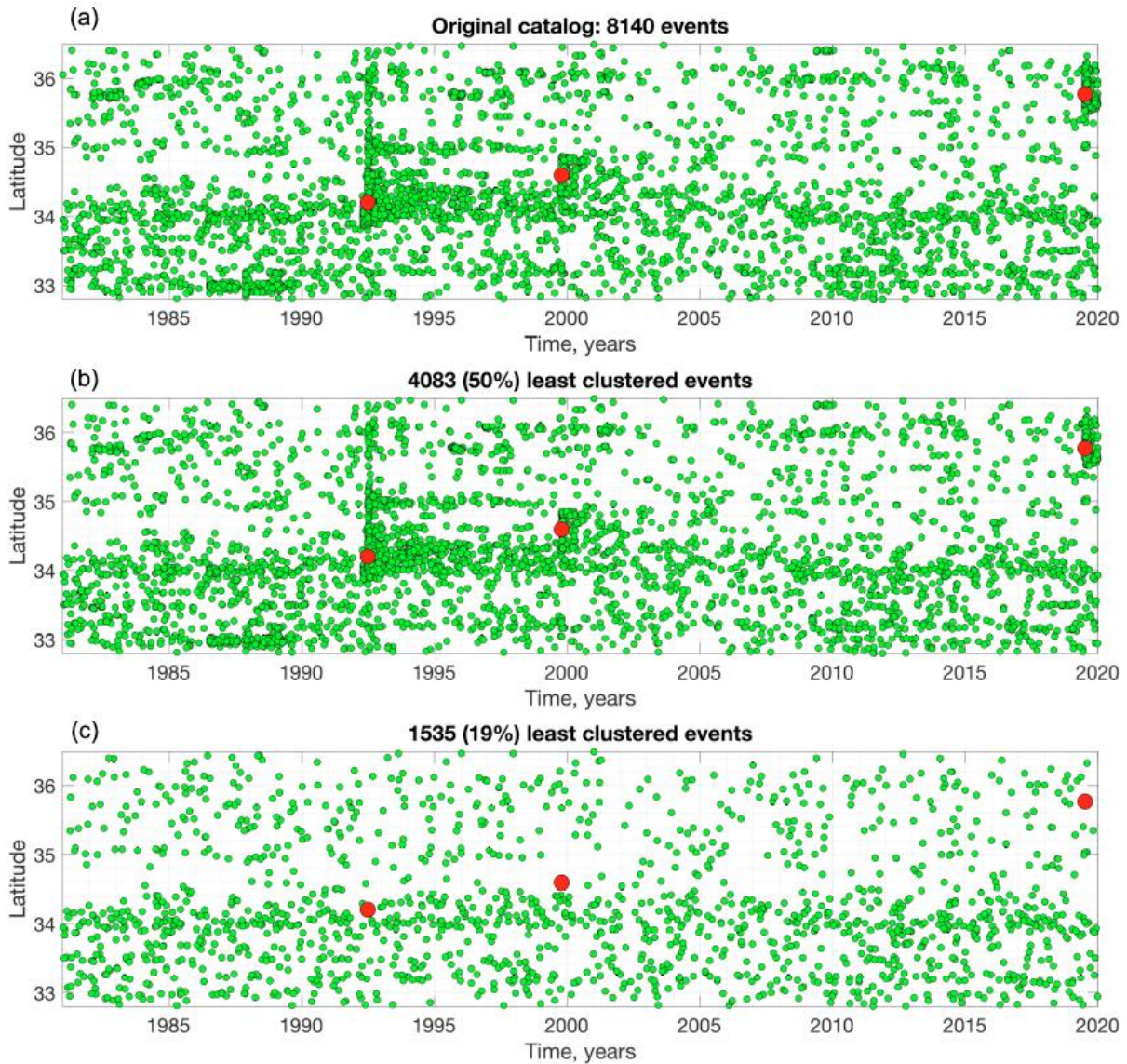
745 **Figure 1:** Spatio-temporal clustering of earthquakes in Southern California based on the catalog  
 746 of *Hauksson et al.* (2012, extended) for the period 1981 – 2020. (a) Map view of earthquakes with  
 747  $M \geq 2$  (gray). Shades of red color reflect the number of events (logarithmic scale; see colorbar) in  
 748 each square spatial cell with latitude side of  $0.25^\circ$ . Black lines show major faults. The most active  
 749 cell at  $116.4^\circ\text{W}$ ,  $33.7^\circ\text{N}$  with 3,822 events covers the aftershock sequence of 1992 M7.3 Landers  
 750 earthquake. (b) Event counts in 1-year time intervals. The largest counts correspond to the most  
 751 numerous aftershock sequences (marked in the figure).

752

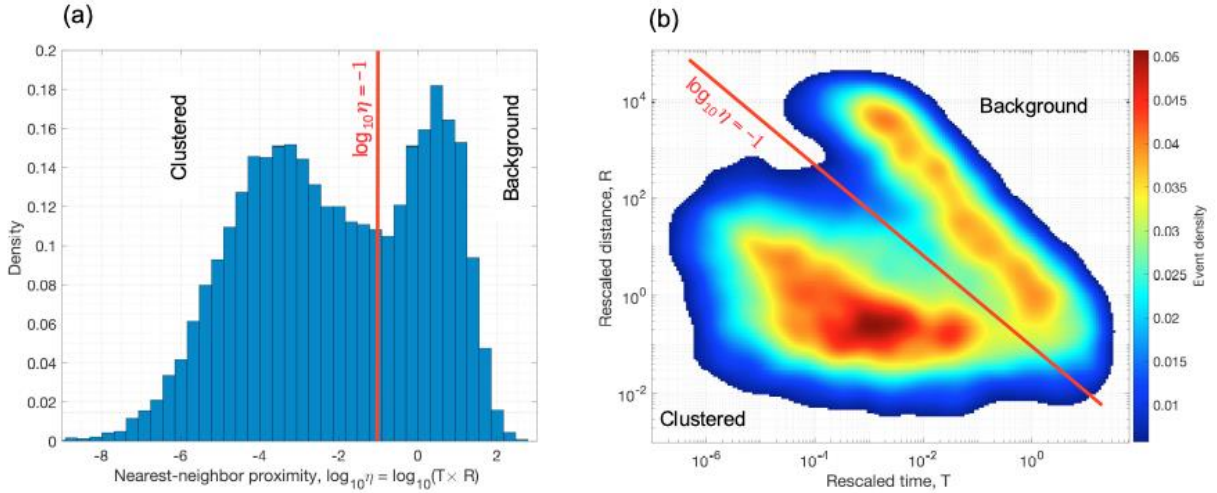


**Figure 2:** Spatio-temporal clustering of global seismicity based on the ISC-GEM v8 catalog during the period 1976 – 2017. (a) The number of earthquakes with  $M \geq 5.8$  in square spatial cell with latitude side of  $5^\circ$  (color). Black lines show plate boundaries. (b) Event counts in 4-month time intervals. Some of the spikes corresponds to aftershocks of large earthquakes indicated in the figure; the higher number of events following the M9.1 Tohoku earthquake relative to the larger M9.3 Sumatra earthquake illustrates the lack of true universality in earthquake behavior.

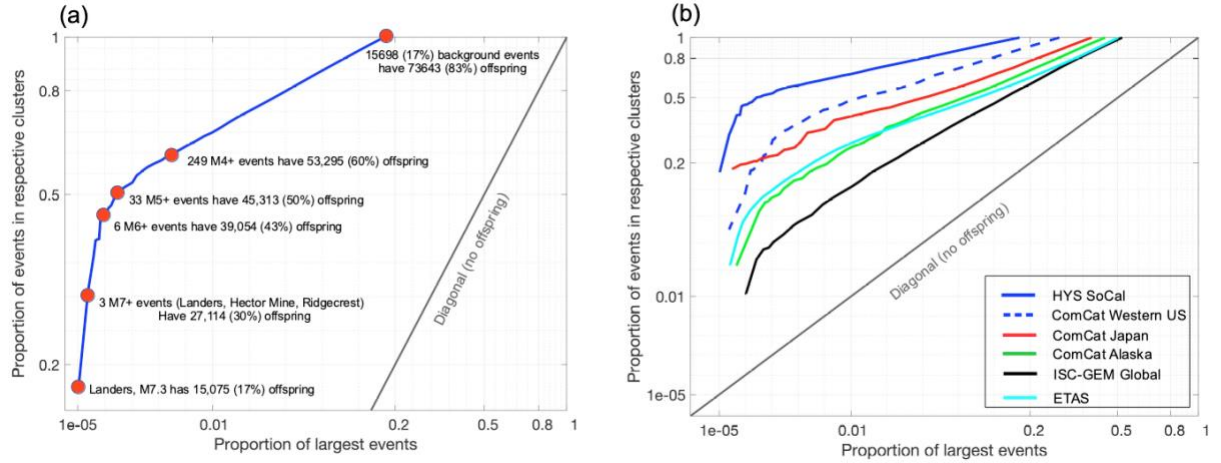




**Figure 3:** Clustering of seismicity in Southern California. Time-latitude projection of earthquakes with magnitude  $M \geq 3$  in the catalog of *Hauksson et al.* (2012, extended) during 1981-2020. The largest earthquakes (1992 M7.3 Landers, 1999 M7.1 Hector Mine, and 2019 M7.1 Ridgecrest) are marked by red circles. (a) The full catalog with 8,140 events. (b) The 50% (4,083) least clustered events. (c) A declustered catalog based on the method of *Zaliapin and Ben-Zion* (2020) with 19% (1535) least clustered events. The visual similarity of panels (a) and (b) shows that 50% of the catalog occurs in dense clusters.

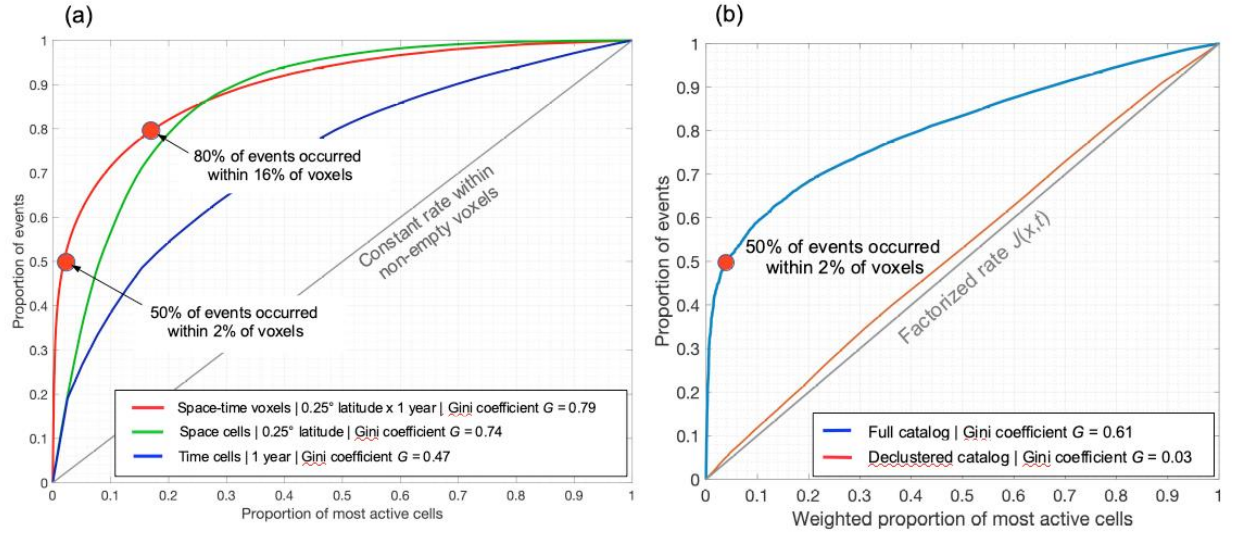


**Figure 4:** Scale separation in the earthquake nearest-neighbor proximity in Southern California, using the catalog of *Hauksson et al.* (2012, extended) with magnitudes  $M \geq 3$  during 1981 – 2020. (a) Bimodal distribution of the nearest-neighbor earthquake proximity  $\eta$ . (b) Bimodal distribution of the time ( $x$ -axis) and space ( $y$ -axis) components of the earthquake proximity. Red line ( $\log_{10} \eta = -1$ ) depicts the separation between the cluster (lower proximity) and background (higher proximity) modes in both panels. The two modes are separated by several orders on the proximity scale. This separation of scale facilitates the catalog declustering, but the large number of events around the separation line produces differences between alternative declustering approaches.



**Figure 5:** Triggering diagram for earthquakes in Southern California. Proportion of offspring (y-axis) produced by a given proportion of the largest earthquakes (x-axis). The offspring relations are identified by the method of *Zaliapin and Ben-Zion* (2020). The diagonal line corresponds to the absence of triggering (all events belong to the background). The deviation of the observed line from the diagonal signifies a large fraction of offspring events. The axes are scaled with a power of 0.25 to zoom-in the left part of the diagram, highlighting that a very small number of largest events triggers a significant fraction of the catalog. (a) Earthquakes with magnitude  $M \geq 2$  in the *Hauksson et al.* (2012, extended) catalog during 1981-2020. (b) Other catalogs examined in the study; see Table 1 for catalog parameters.

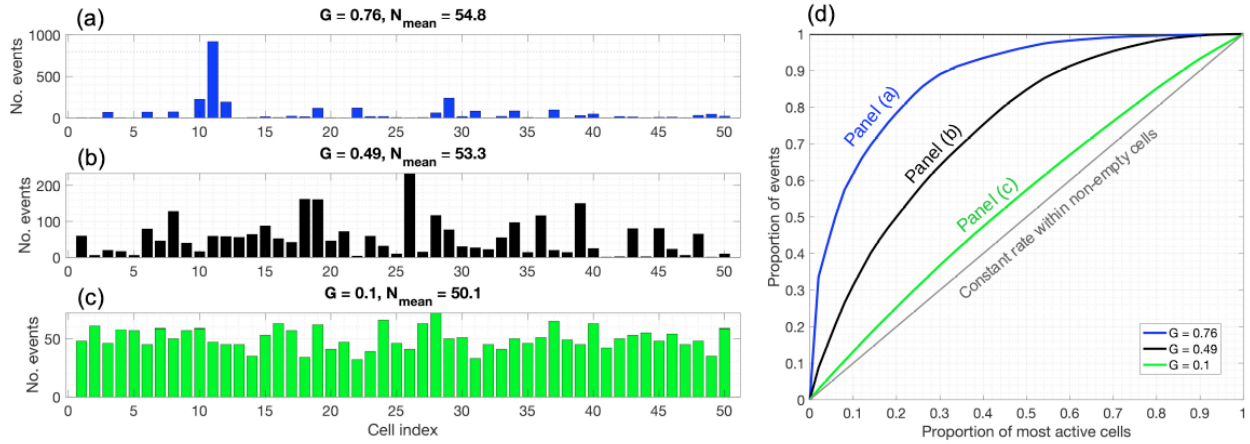




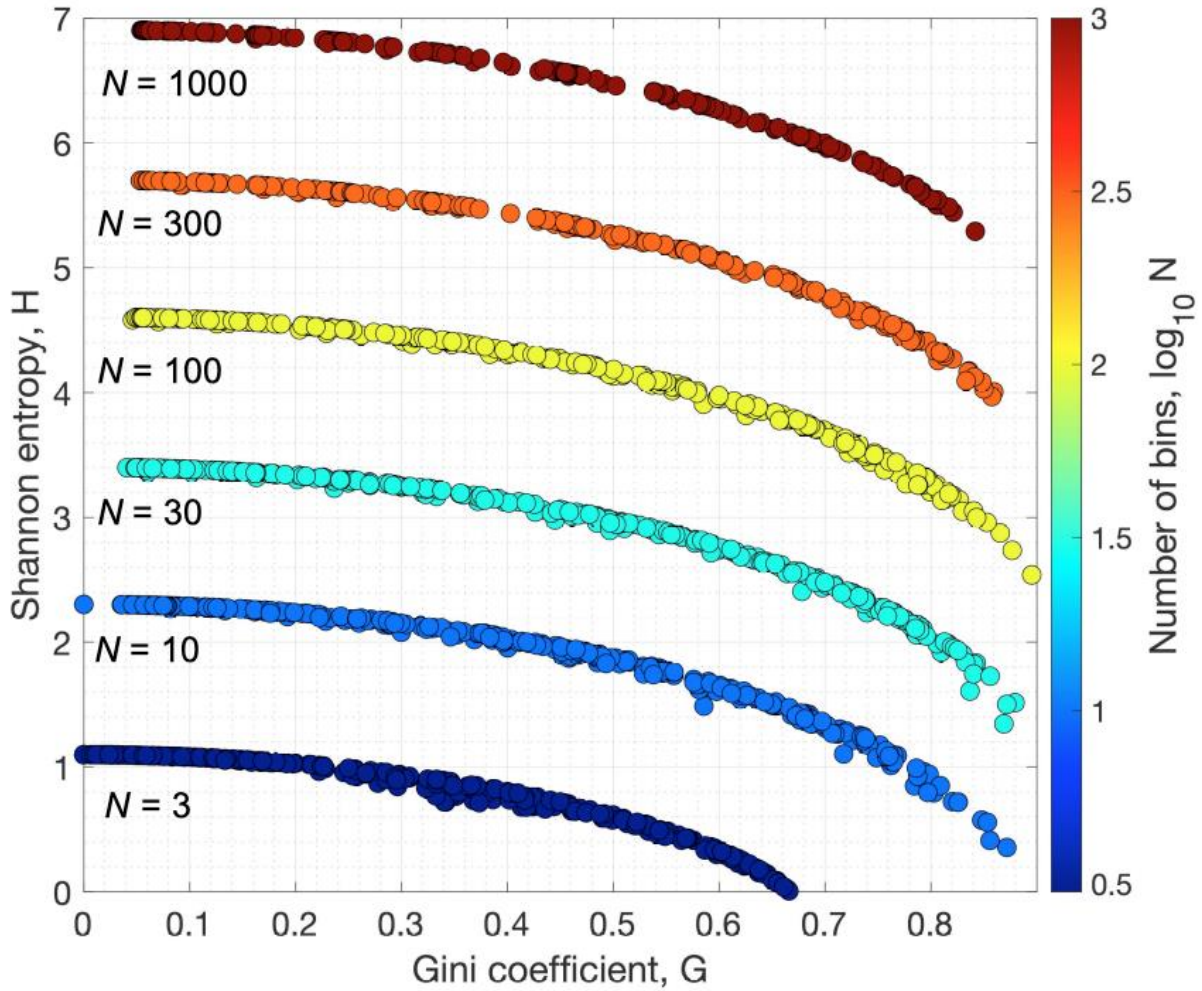
**Figure 6:** Quantifying clustering of earthquakes in Southern California with the ROC diagram.

The analysis uses the catalog of *Hauksson et al.* (2012, extended) during 1981-2020 with magnitude  $M \geq 2$ , and space-time voxels with square space projection of latitude size  $0.25^\circ$  and 1-year duration. (a) Proportion of earthquakes (y-axis) that occurred within a given fraction (x-axis) of most active voxels. The diagonal line corresponds to a constant rate of events in the non-empty voxels. The results for the joint space-time partition are shown by the red line; results for the marginal space and time partitions are shown by green and blue lines, respectively. (b) The proportion (y-axis) of earthquakes that occurred within a given weighted fraction (x-axis) of voxels. The weight of a voxel is proportional to the number of background events that occurred within the space projection of the voxel during the entire examined time interval. The diagonal line corresponds to the factorized rate  $J(x,t)$  of events with marginal space and time distribution taken from the estimated background. The deviation of the observed line (blue) from the diagonal signifies temporal deviations of event rate within local areas (coupled space-time inhomogeneity). The red line corresponds to a declustered catalog (background events). The background events are identified using the method of *Zaliapin and Ben-Zion* (2020).





**Figure 7:** Synthetic examples of ROC cluster analysis. Three synthetic sequences of non-zero cell counts with 50 distinct values and average cell count of about 50 have Gini coefficients of  $G = 0.76$  (panel a),  $G = 0.49$  (panel b), and  $G = 0.1$  (panel c). Their ROC diagrams are shown in panel d.



**Figure 8:** Comparing the Gini coefficient  $G$  and Shannon entropy  $H$  measures of non-uniformity. The curves show the Gini coefficient  $G$  ( $x$ -axis) in analysis with respect to a constant rate vs. Shannon entropy  $H$  ( $y$ -axis) for 2,400 synthetic discrete distributions with different number  $N$  of bins (color). The entropy  $H$  is significantly affected by the number  $N$  of bins. The entropy  $H$  and clustering  $G$  are closely (non-linearly) related for distributions with the same number of bins.

## Supplementary Material

### A. Stability of the cluster measure $G$

The ROC-based cluster framework of Sects. 2.2, 2.3, involves the following choices: (i) lowest examined magnitude, (ii) space and time sizes of partition voxels, and (iii) catalog declustering. Here we demonstrate that the ROC-based Gini coefficient  $G$  is stable with respect to each of these choices, and hence can be used as an objective measure of earthquake clustering.

First, we demonstrate stability with respect to the resolution of the space-time partition. Consider the ROC diagram with respect to a constant rate that has been illustrated in Fig. 5a. To test its stability, we calculate  $G$  in the same catalog of Southern California with  $M \geq 2$  events using combinations of time discretization durations between 1 month and 40 years, and space cell sides between 3 km and 125km. Specifically, we use the time size grid and space size grid that have 50 individual values each placed uniformly on a logarithmic scale between the above limits. This results in 2,500 versions of analysis with total number of voxels varying between 9 and 6,643,230, and average number of events per non-empty voxel varying between 2.46 and 9,927.8. Figure S1 shows the corresponding  $G$  values (panel a) and average number of events per non-empty voxel (panel b). The clustering measure has a stable value above 0.7 for a wide range of space-time resolutions that correspond to a range of 5-300 events per cell. In the corners of the diagram, the  $G$  value decreases to about 0.6 due to finite-size effects. Namely, the finite size of the catalog causes a decreased clustering for very fine resolutions that correspond to small event counts per voxel (e.g. the bottom left corner of the diagram corresponds to 95% of non-empty cells having no more than 3 events). The finite size of the examined region causes a decreased clustering at very coarse resolutions that correspond to a small number of examined voxels (e.g. the top right corner of the diagram corresponds to 9 voxels total).

Figure S2 illustrates the finite-size effects by showing  $G$  values (color) as a function of the average number of events per voxel ( $x$ -axis) and total number of examined voxels ( $y$ -axis). The intermediate resolutions, which roughly correspond to at least 5 events per voxel and at least 300 examined voxels, result in a close range of  $G$  values, within the interval 0.75 – 0.8. The decrease of clustering to below 0.7 is due to either insufficient number of events per voxel (finite catalog, top left corner), or insufficient total number of voxels (finite region, bottom right corner).

Figure S3 demonstrates stability of the cluster measure  $G$  to the space and time resolution in the ROC diagrams calculated with respect to the factorized rate  $J(x,t)$  introduced in Sect. 2.3, Fig. 5b. The analysis is done in the same fashion as that in Fig. S1 and is applied to the full catalog (panels a,b) and declustered catalog (panels c,d). Again, the majority of space-time resolutions show a consistent value of the Gini coefficient  $G$  – around 0.6 for the full catalog and around 0.03 for the declustered catalog – with deviations due to finite-size effects. We also note that in the ROC analysis with respect to a constant rate (Fig. S1a) the clustering is maximized at smaller space sizes and larger time sizes; while in the analysis with respect to the factorized rate (Fig. S3a,c) the clustering is maximized for larger space sizes and shorter time sizes. This is consistent with the focus of each analysis. The former mainly reflects the geometry of the fault network, which is better outlined with smaller space cells; while the latter mainly focuses on aftershock sequences, which are better depicted by shorter time discretizations.

Next, we demonstrate stability with respect to the lower magnitude. Figure S1c,d shows the cluster measure  $G$  computed at different space and time resolutions for earthquakes with magnitude  $M \geq 3$  in the ROC analysis with respect to a constant rate. The results are very similar to those in Fig. S1a,b for a lower magnitude threshold ( $M \geq 2$ ). The stable values of  $G$  above 0.7 are observed for resolutions that correspond to an average of 5-300 events per cell; the clustering

decreases to below 0.6 for very fine and very coarse resolutions. The maximal Gini coefficient of 0.75 in the analysis of  $M \geq 3$  events (Figs. S3c,d) is slightly lower than that of 0.8 in the analysis of  $M \geq 2$  events (Figs. S3a,b). This is due to the overall intensity decrease – there are 89,341 events with  $M \geq 2$  and only 8,140 events with  $M \geq 3$ . A smaller number of events is less effective in sampling the (unobservable) inhomogeneous space-time rate, resulting in a slight decrease of the cluster measure. In the analysis of 822 events with  $M \geq 4$ , the coefficient  $G$  takes a maximal value of 0.67.

Table S1 summarizes these stability results in rows 2,5,8. The table reports the average values of the Gini coefficient  $G$  and the respective 95% confidence intervals (CI) in the ROC cluster analysis of (i) the full catalog with respect to a constant rate (column 6), (ii) the full catalog with respect to the factorized rate (column 7), and (iii) a random realization of declustered catalog with respect to the factorized rate (column 8). To ensure the results are not affected by finite-size effects (Fig. S2), we only consider combinations of space-time resolutions which result in at least 5 events per non-empty voxel and at least 300 voxels in total. In summary, the average value of the cluster measure  $G$  varies within 0.13 units (between 0.77 and 0.64) when changing the lower magnitude of examined events by two units from 2 to 4 (column 2) and hence changing the total number of examined events within two orders (column 3) from roughly  $9 \times 10^4$  to  $8 \times 10^2$ .

Finally, we demonstrate stability with respect to the declustering that is used to estimate the factorized rate  $J(x,t)$  in assessing the space-time coupling of earthquakes. The used stochastic declustering (*Zaliapin and Ben-Zion, 2020*) allows a user to control the average total number of background events in a declustered catalog, preserving the relative rates of background earthquakes in different regions. This is done by varying a cluster threshold  $\alpha_0$  indicated in column 4 of Table S1. The default value of the threshold is 0, which is used in the above experiments (rows

2,5,8). Smaller (larger) values of the threshold lead to smaller (larger) number of background events. The stochastic variability of the number of background events for a fixed  $\alpha_0$  is small (Table S1, column 5) and typically within a fraction of a percent, relative to the total number of events in the full catalog; see also *Zaliapin and Ben-Zion (2020)*.

Rows 1-3 of Table S1 summarize the results for magnitudes  $M \geq 2$  for three versions of declustering that correspond to  $\alpha_0 = 0.5, 0, -0.5$ . The average number of background events changes from roughly 26% to 18% to 9% of the full catalog, respectively (Table S1, column 5). Despite this substantial change of the estimated background rate, the clustering of full catalog with respect to the factorized rate (column 7) and clustering of background events with respect to the factorized rate (column 8) remain very close for different choices of  $\alpha_0$ ; the former around  $G = 0.55$  and the latter around  $G = 0.03$ . A very slight systematic increase of the average cluster value  $G$  for the full catalog from 0.52 to 0.56 (column 7, rows 1-3) is explained by the observation that when more events are included in the background, their space-time distribution is more similar to that in the full catalog, which includes the background. A slight and systematic decrease of the average  $G$  from 0.054 to 0.025 (column 8, rows 1-3) is due to the fact that when more events are included in the background, the higher are the chances that some of them are slightly clustered. The confidence intervals reported in Table S1 further emphasize the stability of clustering with respect to the number of background events. Rows 4-6 of Table S1 refer to the same analysis of events with magnitudes  $M \geq 3$ , and rows 7-9 to analysis of events with  $M \geq 4$ . In each magnitude range, the results are very stable with respect to the number of background events; and overall results are only weakly sensitive to the magnitude selection.

In summary, the experiments illustrated in Figs. S1-S3 and Table S1 indicate high stability of the ROC-based earthquake clustering quantification with respect to the resolution of space-time

911 partition, lower magnitude of analysis, degree of declustering (number of background events), and  
912 alternative stochastic realizations of declustering. This suggests that the ROC-based coefficient  $G$   
913 can be used as an objective measure of the degree of earthquake clustering.

**Table S1: Stability of clustering results in Southern California** using the *Hauksson et al.* (2012, extended) catalog during 1981 – 2020 within 119.5 – 115.5 W and 32.8 – 36.5 N. Stability is tested with respect to the lower magnitude of analysis (columns 2,3), the number of estimated background events and different versions of stochastic declustering (columns 4,5), and resolution of the space-time partition (CI in columns 6-8).

1	2	3	4	5	6	7	8
#	Magnitude range	No. events	$\alpha_0^*$	No. background events <sup>†</sup>	Clustering w.r.t. constant rate <sup>‡</sup> , mean $G$ (95% CI)	Clustering w.r.t. factorized rate $J(x,t)^{\ddagger}$ , mean $G$ (95% CI)	Background clustering w.r.t. factorized rate $J(x,t)^{\ddagger}$ , mean $G$ (95% CI)
1	2.0 – 7.3	89,341 (100%)	0.5	23,045 $\pm$ 101 (25.8 $\pm$ 0.11 %)	<b>0.77</b> (0.7, 0.82)	<b>0.52</b> (0.31, 0.61)	<b>0.054</b> (0, 0.09)
2			0	15,914 $\pm$ 110 (17.8 $\pm$ 0.1 %)		<b>0.55</b> (0.35, 0.63)	<b>0.031</b> (0, 0.06)
3			-0.5	8,322 $\pm$ 111 (9.3 $\pm$ 0.13 %)		<b>0.56</b> (0.37, 0.65)	<b>0.025</b> (0, 0.08)
4	3.0 – 7.3	8,140 (100%)	0.5	2,081 $\pm$ 27 (25.6 $\pm$ 0.33 %)	<b>0.73</b> (0.66, 0.77)	<b>0.57</b> (0.40, 0.65)	<b>0.054</b> (0, 0.12)
5			0	1,528 $\pm$ 31 (18.8 $\pm$ 0.38 %)		<b>0.58</b> (0.44, 0.66)	<b>0.031</b> (0, 0.07)
6			-0.5	852 $\pm$ 34 (10.5 $\pm$ 0.41 %)		<b>0.59</b> (0.46, 0.66)	<b>0.031<sup>§</sup></b> (0, 0.11)
7	4.0 – 7.3	822 (100%)	0.5	222 $\pm$ 8 (27 $\pm$ 0.99 %)	<b>0.64</b> (0.59, 0.67)	<b>0.53</b> (0.43, 0.61)	<b>0.068<sup>¶</sup></b> (0, 0.11)
8			0	169 $\pm$ 10 (20.5 $\pm$ 1.2 %)		<b>0.55</b> (0.44, 0.63)	<b>0.055<sup>¶</sup></b> (0, 0.11)
9			-0.5	95 $\pm$ 11 (11.6 $\pm$ 1.4 %)		<b>0.57</b> (0.51, 0.62)	<b>0.018<sup>¶</sup></b> (0, 0.08)

\* Cluster threshold that controls the number of background events (*Zaliapin and Ben-Zion, 2020*).

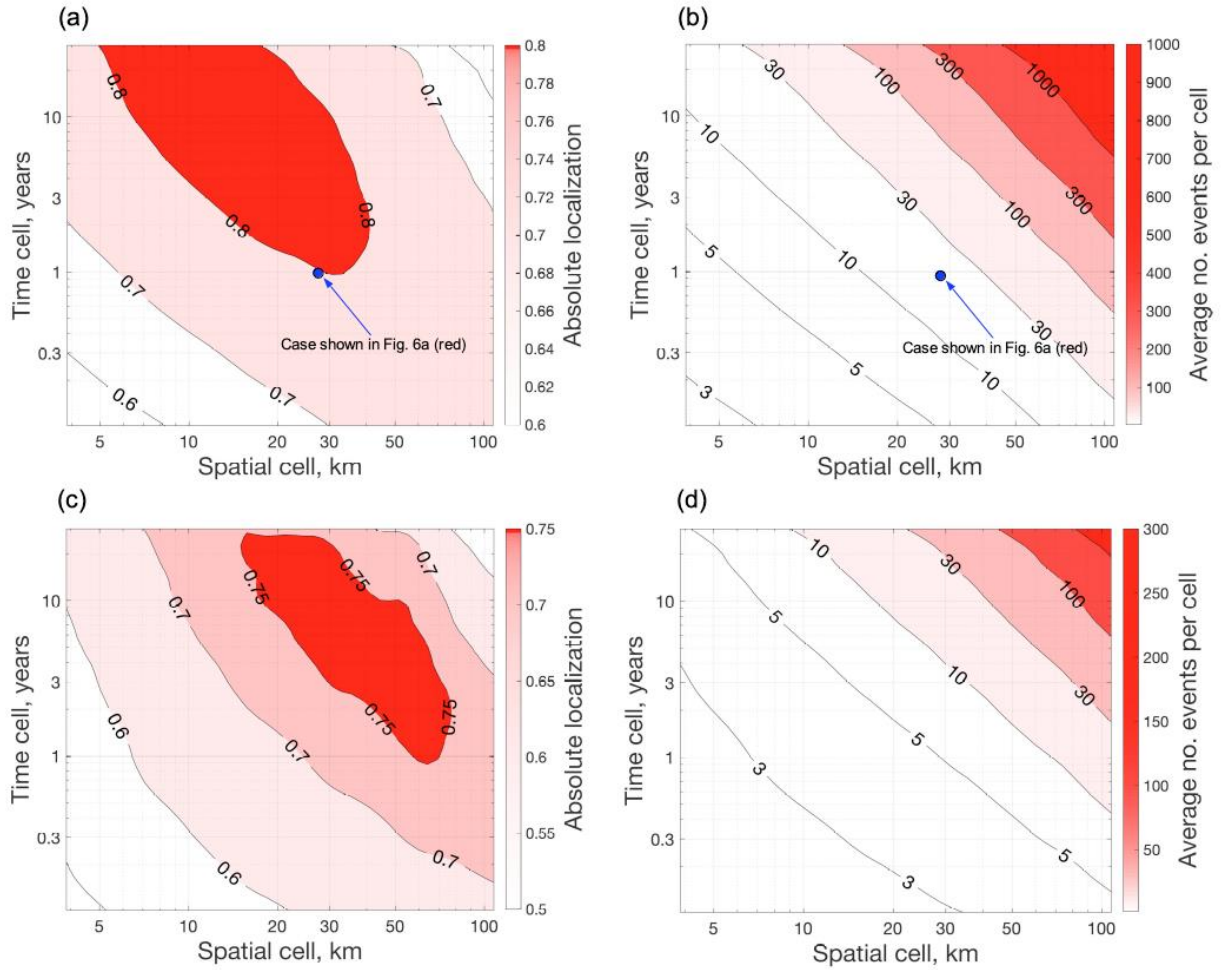
<sup>†</sup> Mean and 95% confidence interval (CI) according to  $10^4$  realizations of stochastic declustering. The CI is estimated as  $\pm 1.96$  sample standard deviation.

<sup>‡</sup> Estimation is done for the space-time partitions with the average of at least 5 events per non-empty voxel and at least 300 voxels total. The CI limits are the 0.025 and 0.975 sample quantiles.

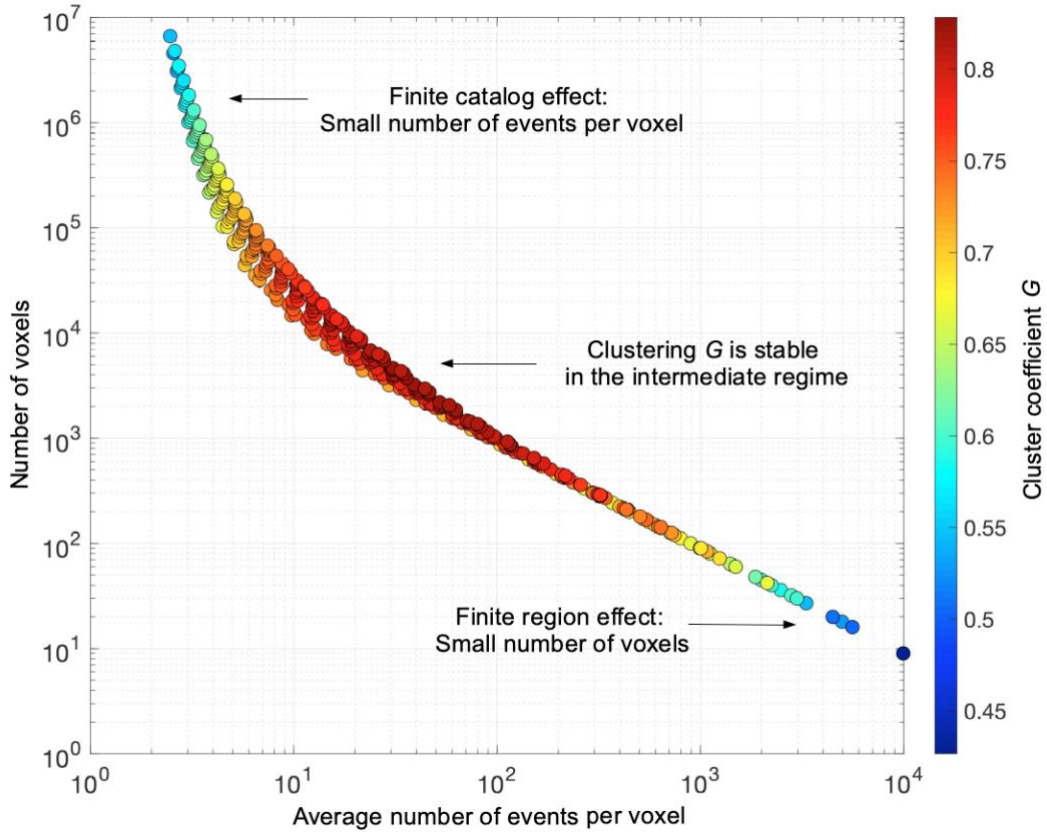
<sup>§</sup> There are no voxels satisfying the common condition <sup>‡</sup>; the condition is relaxed to the average of  $\geq 5$  events per non-empty voxel and  $\geq 100$  voxels total.

<sup>¶</sup> There are no voxels satisfying the common condition <sup>‡</sup>; the condition is relaxed to the average of  $\geq 1.5$  events per non-empty voxel and  $\geq 100$  voxels total.

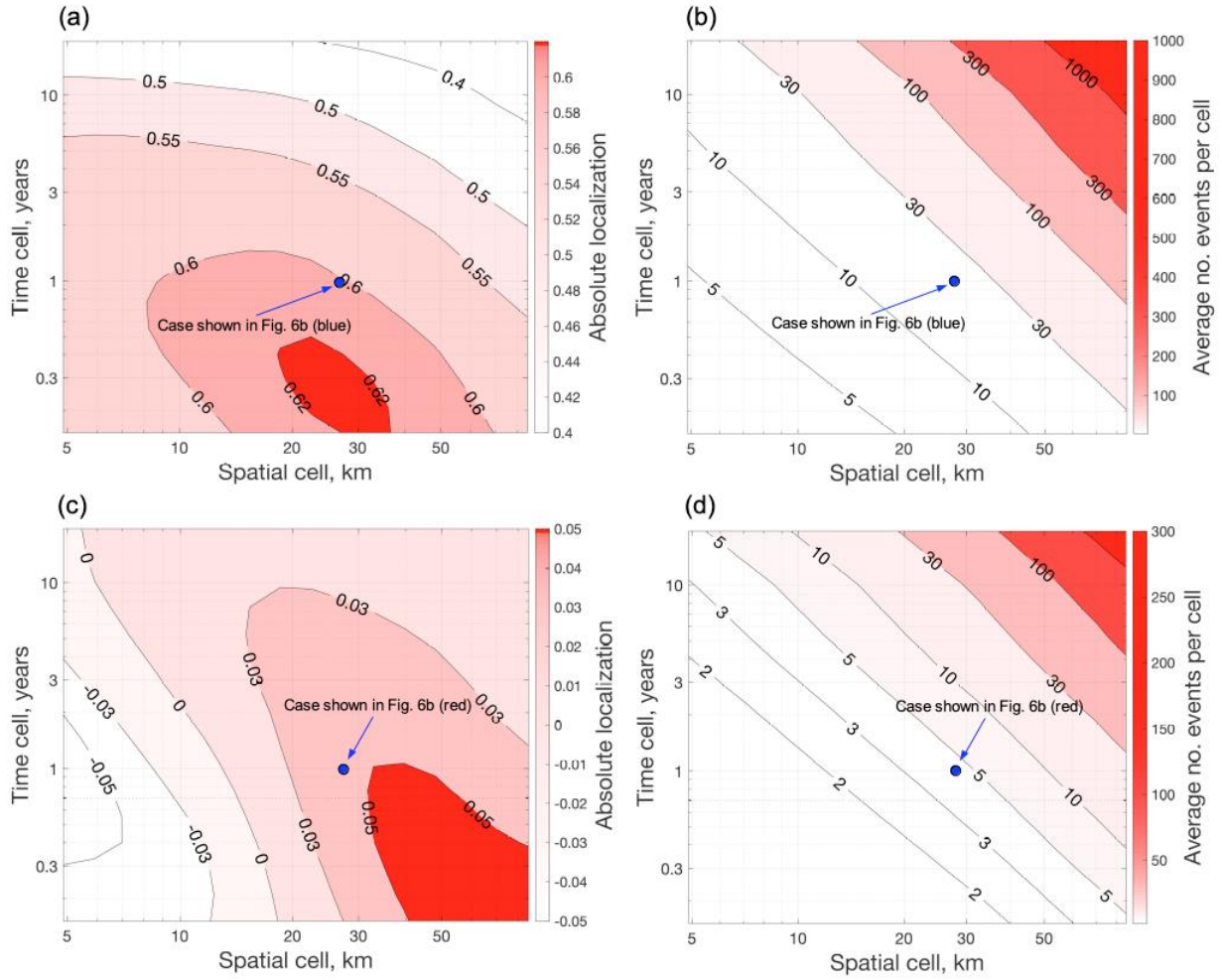




**Figure S1: Stability of cluster measure  $G$  in Southern California.** The catalog of *Hauksson et al.* (2012, extended) with magnitude  $M \geq 2$  during 1981 – 2020. Gini coefficient  $G$  in cluster analysis of earthquakes with magnitude  $M \geq 2$  (panel a) and  $M \geq 3$  (panel c) in the ROC analysis with respect to a constant rate of events for different space ( $x$ -axis) and time ( $y$ -axis) resolutions. Average number of events per non-empty voxel in the experiment of panels (a,c) is shown in panels (b,d), respectively. Blue circles in panels (a,b) corresponds to the case illustrated in Fig. 6a.



**Figure S2: Finite-size effects in cluster analysis** of *Hauksson et al.* (2012, extended) catalog during 1981 – 2020 with magnitude  $M \geq 2$ . Gini coefficient  $G$  (color) of the full catalog with respect to a constant rate within non-empty voxels as a function of the average number of events per non-empty voxel ( $x$ -axis) and the total number of voxels ( $y$ -axis). The  $G$  value is stable within the intermediate range of space-time voxel sizes: for at least 5 events per voxel and at least 300 voxels total (red colors, localization is above 0.7). The cluster value  $G$  may decrease because of either insufficient number of events per voxel (top left corner) or insufficient total number of voxels (bottom right corner).



**Figure S3: Stability of cluster measure  $G$  in Southern California.** The catalog of *Hauksson et al.* (2012, extended) with magnitude  $M \geq 2$  during 1981 – 2020. Gini coefficient  $G$  of the full catalog (panel a) and background earthquakes (panel c) in the ROC analysis with respect to the factorized rate  $J(x,t)$  for different space ( $x$ -axis) and time ( $y$ -axis) resolutions. Average number of events per non-empty voxel in the experiment of panels (a,c) is shown in panels (b,d), respectively. Blue circles correspond to the cases illustrated in Fig. 6b.