Chaos, Solitons and Fractals 170 (2023) 113419

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos

Global versus local clustering of seismicity: Implications with earthquake prediction

Davide Zaccagnino [a,*], Luciano Telesca [b], Carlo Doglioni [a,c]

[a] Earth Science Department, Sapienza University, Piazzale Aldo Moro, 5, Rome, 00185, Italy
[b] Institute of Methodologies for Environmental Analysis, National Research Council, Zona Industriale C.P. 27, Tito Scalo, 85050, Italy
[c] National Institute of Geophysics and Volcanology, Via di Vigna Murata, 605, Rome, 00143, Italy

ARTICLE INFO

Dataset link: https://quakesearch.geonet.org.n/, https://earthquake.usgs.gov/earthquakes/search/

Keywords:
Seismogenic potential
Seismic clustering
Maximum expected magnitude
Clustering coefficients
Fractal dimension
b-value

ABSTRACT

The estimation of the maximum expected magnitude is crucial for seismic hazard assessment. It is usually inferred via Bayesian analysis; alternatively, the size of the largest possible event can be roughly obtained from the extent of the seismogenic source and the depth of the brittle–ductile transition. However, the effectiveness of the first approach is strongly limited by catalog completeness and the intensity of recorded seismicity, so that it can be of practical use only for aftershocks, while the second is affected by extremely large uncertainties. In this article, we investigate whether it may be possible to assess the magnitude of the largest event using some statistical properties of seismic activity. Our analysis shows that, while local features are not appropriate for modeling the emergence of peaks of seismicity, some global properties (e.g., the global coefficient of variation of interevent times and the fractal dimension of epicenters) seem correlated with the largest magnitude. Unlike several scientific articles suggest, the *b*-value of the Gutenberg–Richter law is not observed to have a predictive power in this case, which can be explained in the light of heterogeneous tectonic settings hosting fault systems with different extension.

1. Introduction

According to the classical view of seismicity, earthquakes occur as a consequence of stress loading on faults accumulated by the long-term action of tectonic strain [1]. The elastic rebound model predicts that the stress be released as soon as its value overcomes the local shear resistance of the fault interface. Stress drop can be almost complete because of the weakening processes during slip, which produce a sudden lowering of fault friction just after strengthening during the transition from static to dynamic configurations [2]. So, at a first level of approximation, the evolution of seismic activity can be described as an iterative sequence of energy accumulation, with few or no seismic events, development of instability, featured by minor breakdowns, i.e., foreshocks, and seismic sequences, driving the system towards stability [3]. As much as this recursive pattern is just one of several possible dynamic evolutions, this simple framework is routinely invoked as the foundational mechanism for earthquake occurrence [4]. A more refined model conceives seismicity as a self-excited dynamics taking place in a complex [5], disordered [6] and self-organized setting [7]; so, characterized by a certain number of emergent properties that cannot be derived from the fundamental laws ruling the basic components of the system. From this viewpoint, the trends in stress and strain accumulation play a background role, i.e., they just provide a certain amount of energy for avalanches, but they do not take actively part in the spatial and temporal evolution of earthquakes [8]. In this context, major seismic events can be forewarned by foreshocks, seismic quiescence or happen unexpectedly, according to the different kind of transition associated with the peculiar considered geological environment, e.g., sub-, super- or critical dynamics [9]. Investigating seismicity in the light of the physics of complex systems also provides a natural explanation for its key property: earthquake clustering. Previous research clearly shows that seismic activity tends to occur clustered both in time and space at both short and long scales [10–13]. Although at its core, clustering, especially the long-term one, is usually neglected in modeling seismicity [14], except for the Omori–Utsu decay after major seismic events, e.g., [15]. The assumption of negligible long-term correlation in the occurrence of earthquakes makes models easier, but, in the other hand, also removes the possibility of long-lasting memory of past events, which is indeed the other crucial feature of seismic activity [16]. Therefore, in order to better understand how earthquakes occur in space and time we need to analyze how they are organized. Furthermore, the lack of comprehension of long-term mechanisms producing the emergence of large-scale crustal instability,

* Corresponding author.
E-mail address: davide.zaccagnino@uniroma1.it (D. Zaccagnino).

https://doi.org/10.1016/j.chaos.2023.113419
Received 2 February 2023; Received in revised form 15 March 2023; Accepted 24 March 2023

unavoidably implies a bias in our ability to forecast them. As a matter of fact, the current approaches to seismic hazard only take into account of the recently happened seismic events [17] – at least compared to the geological time scales of fault activity – without considering whether the geophysical setting under study rests on a state of higher or lower stability, i.e., it is more or less prone to close, iterated large earthquakes, or not. Hence, the underestimation of the intrinsic non-linear, self-excited nature of seismicity may constitute a serious threat for a reliable assessment of seismic hazard, above all for what concerns the maximum expected magnitude, e.g., [18]. Nowadays, two different methods exist to evaluate the probability of occurrence of large earthquakes: the first is a pure mathematical approach based on Bayesian probability, so that, given a seismic catalog, the chances of occurrence of an unobserved event can be calculated [19]; the second rests on empirical relationships between the size of faults and the largest hosted earthquake, e.g., [20]. However, the first method is grounded on the hidden hypothesis that the available information be representative of the whole fault system dynamics, usually thousand-years lasting; while the second is intrinsically riddled with several sources of uncertainty [21] also concerning paleoseismic records [22]. Moreover, the latter case also requires a rather advanced state of maturation of the seismological sources, otherwise also the possibility of fault coalescence should be considered [23], while the first case is not easily adaptable to the idea of faults with long memory and seismic super-cycles [24] already discussed above. Therefore, a deep understanding of the mechanisms driving the slow dynamics of seismic activity towards major, widespread instability is an essential tool for both the foundations of earthquake physics and seismic hazard. Here, we investigate some local and global properties of seismic catalogs (the GeoNet earthquake catalog for New Zealand in this case [25]) such as the local and global coefficients of variation, namely L_V and C_V, the scaling exponent of the Gutenberg–Richter law (the so-called *b*-value), the fractal dimension of spatial distribution of epicenters D_f [26], the seismic rate (the amount of energy nucleated by earthquakes during a certain reference time interval) and the number of events (above the completeness magnitude) and the maximum magnitudes (observed in catalog and historical). Although certainly not exhaustive, this list is a first attempt to understand what kind of information can be stored in such collective parameters and the mutual relationships among them. What is more innovative, we apply clustering and fractal analysis to infer the properties of large earthquakes, in particular the maximum expected magnitude, from the observation of small and intermediate events.

2. Methods

This section is devoted to the procedures applied for the analysis of catalogs and for the calculation of relevant collective parameters. More detailed information is reported in the Supplementary Materials.

2.1. Catalog completeness and conversions

In this research, only earthquakes above the completeness magnitude are considered. We apply the Wiemer–Wyss method [27] and add a correction of +0.2 magnitude units, as suggested in [28]. The completeness magnitude is computed for samples of one thousand earthquakes each in order to take into account of the different phases of seismic activity (i.e., preseismic, seismic, post-seismic and interseismic periods) which unavoidably imply a change of the catalog completeness. The GeoNet catalog is mainly composed of local magnitudes that we convert to moment magnitudes according to the empirical law $M_w = 1.14M_L - 0.83$, if the hypocentral depth is smaller 33 km, and $M_w = 0.92M_L - 0.05$, if the hypocenter is deeper than 33 km [29].

2.2. Coefficients of variation

The global coefficient of variation of interevent times, C_V, defined by [10]

$$C_V = \frac{\sigma_{\Delta T}}{\langle \Delta T \rangle} \tag{1}$$

where $\langle \Delta T \rangle$ is the mean value of the inter-event time and $\sigma_{\Delta T}$ is its standard deviation, is used to evaluate the time-clustering of seismicity as a whole; so, without providing information about the temporal scales at which clustering occurs. The physical meaning of the coefficient of variation is the following: if $C_V < 1$, then the dynamics is regular; on the contrary, if $C_V > 1$, the time series is clustered. The condition $C_V = 1$ stands for a completely random, Poisson process [30]. Conversely, the local coefficient of variation, L_V defined by [31]

$$L_V = \frac{3}{N-1} \sum_{i=1}^{N-1} \frac{(T_i - T_{i+1})^2}{(T_i + T_{i+1})^2} \tag{2}$$

is applied for the description of local variability of inter-event time series. The interpretation of the values of L_V is analogous to those of C_V.

2.3. b-value

We apply the Tinti–Mulargia algorithm [32] for the estimation of the *b*-value of the Gutenberg–Richter law, summarized by the following formulas:

$$\hat{b}_{TM} = \frac{\ln\left(1 + \frac{\delta_M}{\langle M_w \rangle - M_{wc}}\right)}{\ln(10)\delta_M} \tag{3}$$

where δ_M is the binning interval for magnitudes, and

$$\sigma_{\hat{b}_{TM}} = \frac{\frac{\delta_M}{\langle M_w \rangle - M_{wc}}}{\ln(10)\delta_M \sqrt{N\left(1 + \frac{\delta_M}{\langle M_w \rangle - M_{wc}}\right)}}. \tag{4}$$

This method is particularly suitable for limited catalogs; moreover, it takes into account of magnitude binning. In order to utilize it, $\langle M_w \rangle$ and the completeness magnitude M_{wc} are required.

The first is obtained by definition of arithmetic mean of N magnitudes in catalog, while the second is given by the Wiemer–Wyss method [27] with an additional correction of +0.2 magnitude units, as described in the first paragraph of methods.

2.4. Fractal dimension of epicenters

The correlation function is defined by [33]

$$C(r) = \lim_{N\to\infty} \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \theta\left(r - \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}\right) \tag{5}$$

where x and y stand for the epicentral longitude and latitude, r is an arbitrary (small) distance, θ is the Heaviside function and N is the total number of earthquakes selected. We study the correlation dimension D_f of the distribution of epicenters defined as [34]

$$D_f = \lim_{r\to 0} \frac{\log C(r)}{\log r}. \tag{6}$$

The practical determination of $C(r)$ is realized following the Grassberger&Procaccia algorithm [35], which consists in constructing spheres (circles in our case) around N epicenters in order to calculate the number of events occurred within a varying threshold distance r. If the spatial distribution of epicenters is fractal, $C(r)$ behaves as a power-law

$$C(r) \sim r^{\zeta} \tag{7}$$

where $\zeta \to D_f$ as $r \to 0$. Therefore, we obtain D_f as the output of a linear fit of the log–log scale plot of $C(r)$. Details about the

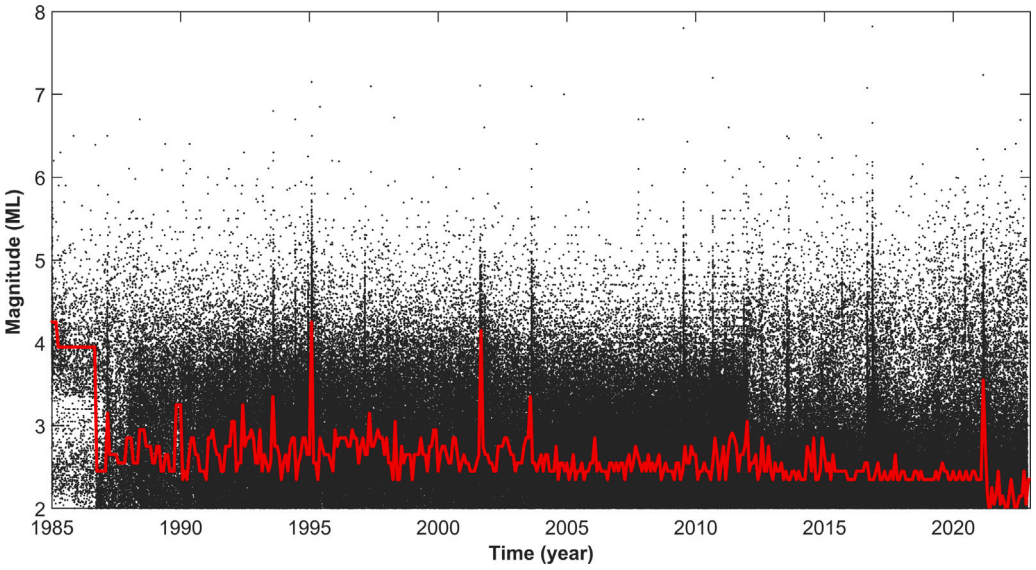

Fig. 1. Seismicity in New Zealand and completeness of the seismic catalog (GeoNet earthquake catalog accessible at https://quakesearch.geonet.org.nz/, crustal events with depth ≤ 50 km, localizations in the range 35–47° S of latitude with local magnitude $M_L \geq 2.0$). The red line represents the completeness magnitude during different stages of seismic activity computed for samples of one thousand earthquakes each. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

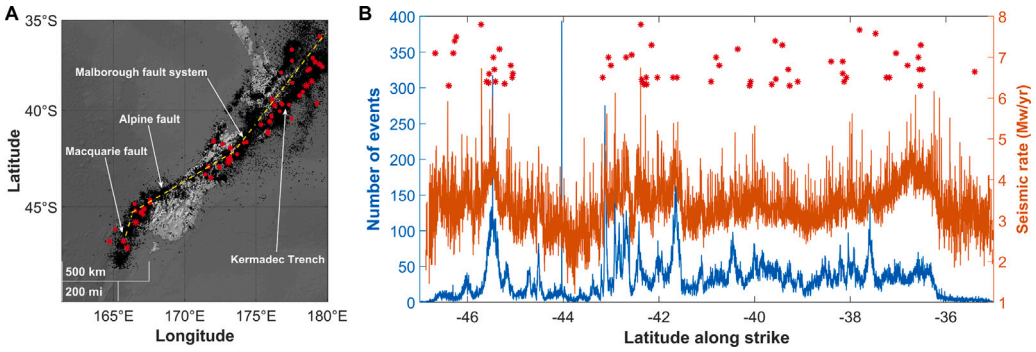

Fig. 2. Seismic activity in New Zealand. (**A**) The map shows the spatial distribution of seismic events from 1985 to 2022 as black points, while large earthquakes (M_w 6+) occurred between 1920 and 2022 are represented by red asterisks. The yellow, dashed lines stand for the principal strike direction of the most important fault systems in New Zealand. (**B**) The number of earthquakes and energy nucleated by seismic events in New Zealand along the main seismogenic structures from 1985 to 2022. Large earthquakes (1920–2022) are represented as red stars. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

calculation of the uncertainty of the fractal dimension are discussed in the Supplementary Material. We do not consider hypocentral series in this case because, while the epicentral uncertainties are usually moderate, depths are featured by extremely large errors which make fractal analysis unreliable unless large relocated seismic catalogs are available. The spatial segmentation of the seismogenic region is made so that the number of relative distances between epicenters is always larger (usually much higher, in the order of 1–10 millions) than 50'000. This number, corresponding to ~250 events, is considered enough to guarantee reliable values of D_f [33]; in fact, the correlation dimension is feasible also for small data sets, in contrast with other fractal dimensions (e.g., capacity and information dimensions) based on the method of covering, usually requiring several thousand data.

3. Analysis and results

New Zealand is one of the most seismically active regions in the world, with several M_w 7+ recorded during the last century and numerous seismic sequences occurring along all its longitudinal extension. Moreover, the country is well equipped with an efficient seismic network, so that the GeoNet seismic catalog lists almost all the M_L 3+ events in the last thirty years except for peaks of seismic activity observed after large mainshocks, when the completeness magnitude is significantly larger (up to $M_c \sim 4.3$). However, the catalog is often complete even at lower magnitudes during quiescent and more recent periods (down to $M_c \sim 2$ recently). For this reason, it is necessary to accurately assess the time-dependent completeness magnitude in order to reduce errors in the estimation of the statistical properties of seismicity such as the *b*-value of the frequency-size Gutenberg–Richter law. Compare with Fig. 1 for a detailed analysis of the completeness magnitude of the GeoNet catalog from 1985 to 2022. New Zealand is located at the boundary between Pacific and Australian plates. Their relative motion, ranging in between 30 and 60 mm/yr [36], is responsible for mainly transpressive tectonic settings in the North Island, while transcurrent zones dominate to South (Fig. 2 A). The southern portion of the Alpine fault system hosted several M_w 6+ earthquakes in the last century, with also a M_w 7.8 in 2009, the Great Fiordland earthquake. This region is featured by elevated seismic rate, while the central and northern part of the Alpine range is characterized by diffuse, moderate magnitude seismicity, without M_w 6+ events recorded during the recent few decades. Conversely, the Canterbury and Otago regions are prone to megaquakes, like the M_w 7.8 Kaikoura earthquake in 2016. The Malborough-North Island zone and the Wellington region are also frequently hit by large events accompanied by a moderate background

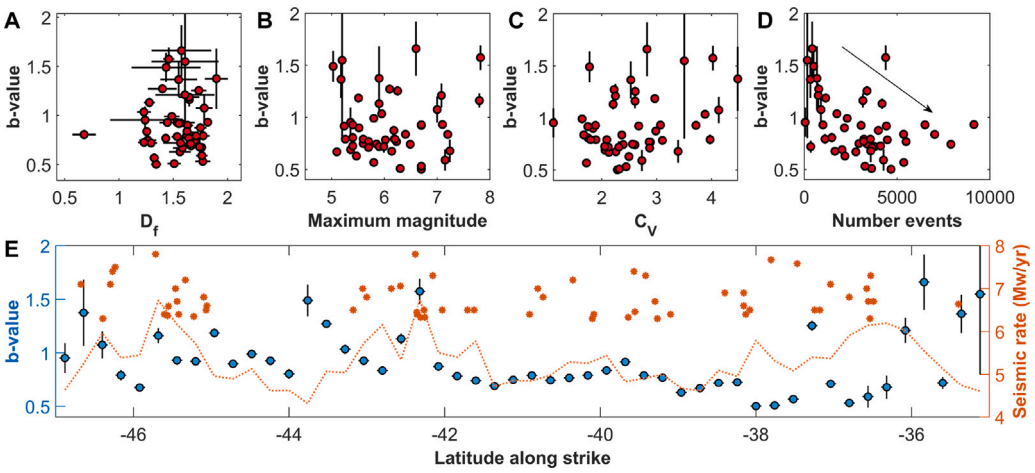

Fig. 3. (**A, B, C**) The *b*-value does not correlate with the fractal dimension of the epicenters, neither with the maximum magnitude nor the global coefficient of variation in New Zealand. (**D**) Conversely, the *b*-value is negatively related to the number of events (above the completeness magnitude) since thrust-faulting regions, usually featured by lower values, are those hosting larger earthquakes, hence more seismic events. (**E**) Distribution of the *b*-value along the strike of the most important seismogenic sources in New Zealand (compare with Fig. 2 A). The orange line represents the seismic rate calculated using seismic events above the completeness magnitude (compare with Fig. 1), while the orange asterisks stand for the largest earthquakes occurred since 1920. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

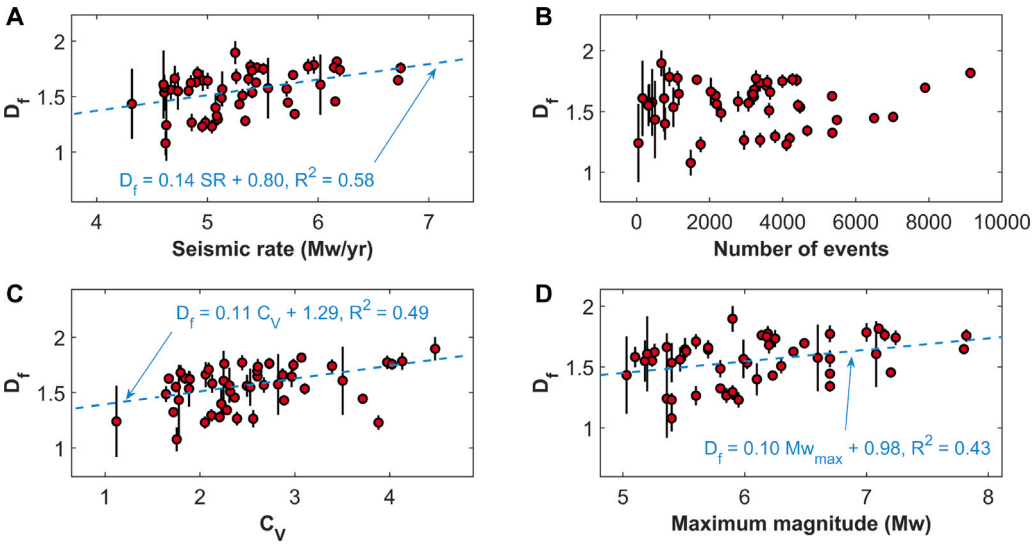

Fig. 4. The relationships of the fractal dimension of epicenters (D_f) with other properties of seismicity. (**A**) D_f is positively related to the nucleated seismic rate. (**B**) D_f vs number of events above the completeness magnitude. (**C**) The fractal dimension of the epicentral series is positively correlated with the global coefficient of variation. (**D**) D_f is also positively related to the maximum observed magnitude in catalog.

seismic rate. Compare with Fig. 2 B. We divide the main fault belt into 50 contiguous segments (compare with the first paragraph of the Supplementary Material for a detailed description), this number is chosen to allow a better assessment of the statistical properties of seismicity and their variation along the catalog. For each of them, we calculate the *b*-value and other physical quantities of interest, as shown in Fig. 3. The *b*-value of the Gutenberg–Richter law is a collective, routinely investigated parameter of seismicity, often put into relation with spatial and temporal variations of fault stress and transition from a seismic regime to a new one, e.g., from interseismic to preseismic phase. Surprisingly, we do not find any significant correlation neither with the maximum magnitude, nor with C_V or the seismic rate. No correlation is also unveiled between the *b*-value and the fractal dimension of epicenters, likely due to the large uncertainties affecting both quantities. We just draw attention to an inverse relationship between the *b*-value and the number of events above the completeness magnitude, which can be simply explained in the light of different maximum magnitudes in various tectonic settings, i.e, higher in compressive than transcurrent and rifting zones. More rewarding results are instead achieved by analyzing the correlation between the fractal correlation dimension of epicentral series and other quantities such as the seismic rate, C_V and the maximum observed magnitude in catalog (Fig. 4). They are all positively correlated, in particular, the following empirical formula holds

$$Mw_{max} = 10.0\, D_f - 9.8 \tag{8}$$

Compare with Fig. 4 D. We also investigate the role of the size of the seismic catalog in shaping our output. In order to do this, we separate our database into different fractions, respectively ∼2%, ∼7%, ∼1/3, ∼2/3 and full catalog in chronological order without temporal superposition for each of our 50 segments. Our results, summarized in Fig. 5 (compare also with Figure 6 of the Supplementary Material), show that extremely short earthquake catalogs tend to be globally Poissonian ($C_V \approx 1$); on the other hand, databases with intermediate or extended length become more and more locally Poissonian or weakly clustered as their size increases, while they keep their global clustering behavior

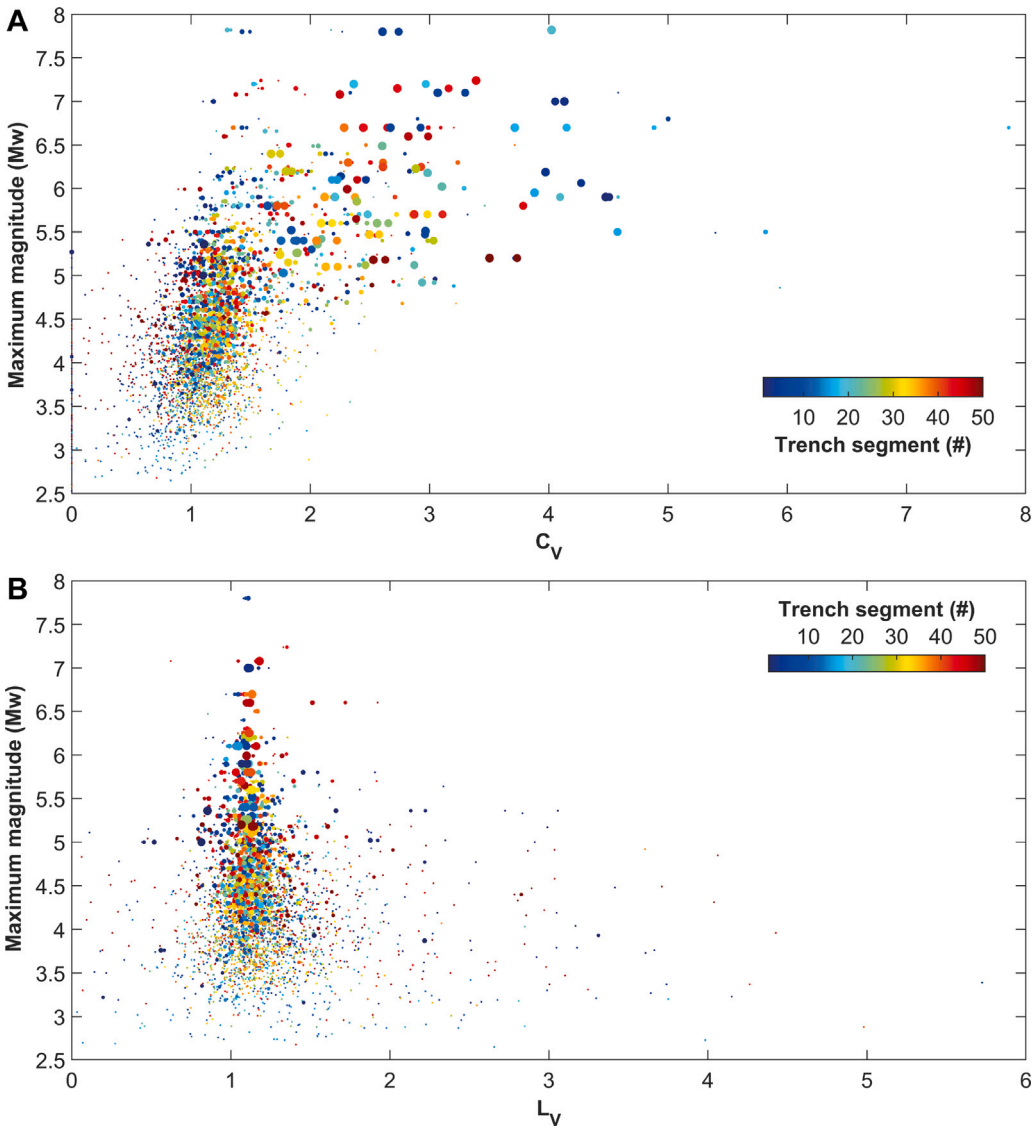

Fig. 5. (A) Maximum observed magnitude in catalog vs global coefficient of variation for fifty segments of seismogenic sources in New Zealand (represented by the color) calculated using all the portions of the seismic catalog (M_w>M_c, depth < 50 km,1985–2022) segmented into parts of ~2%, ~7%, ~1/3, ~2/3 and full catalog. The size of the colored points represents the length of the considered sub-catalog. The behavior of seismic activity in the long-term tends to be globally clustered. Therefore, regions with higher C_V, fixed the size of the catalog, may be prone to larger earthquakes. **(B)** Maximum observed magnitude in catalog vs local coefficient of variation. In the long-term, seismicity is locally Poissonian. Colors and sizes of the points as in the upper panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

unchanged after an initial transient. C_V also preserves its property of positive correlation with the maximum magnitude in catalog.

4. Discussion

4.1. Seismic clustering and its implications for seismic hazard

Our analysis of the clustering and fractal properties of seismicity in New Zealand shows that, although apparently featured by similar patterns of occurrence, earthquakes are organized in time and space in a quite heterogeneous manner. In the southern part of the Alpine fault system, as well as in the Canterbury region, seismicity is globally clustered but locally Poissonian. These areas host violent seismic events with maximum magnitudes approaching M_w 8. Conversely, the central South Island and the northernmost segment of North Island host seismic activity with low C_V and relatively elevated L_V. In these regions, just one M_w 6+ event took place in the last century. We suggest, on the base of such results, that sequences with large mainshocks produce long-term clustering of seismic time series; nevertheless, locally, seismicity is almost Poissonian, i.e., it is not synchronized from external perturbations nor from surrounding events. A possible explanation is that, once extremely large earthquakes happen, great part of the accumulated strain is released, moving crustal volumes towards more stable conditions; therefore, they are less prone to further destabilization [2]. Oppositely, where strain is only partially released by moderate earthquakes, fault systems are kept closer to the edge of instability, which makes them easily synchronized by stress modulations and responsive to additional strain sources [37]. This phenomenon might also explain the reason of global regularity. These results are in agreement with previous research about the clustering properties of regular [38] and silent earthquakes [39]. Our research also shows that clustering analysis has a predictive power about the seismogenic potential at local and regional spatial scales, provided a sufficiently long database, since the global properties of seismic clusters appear roughly self-similar; hence, strongly clustered low-magnitude seismicity may be a valuable indicator for analogously long-term clustered large events in the same region. Compare with Fig. 5 and with our results reported in the Supplementary material (Figure S3 and Figure S6). We also consider the *b*-value of the

Gutenberg–Richter law. Despite we apply a quite accurate procedure for the determination of the scaling exponent of the frequency-size relationship, our analysis is not able to highlight any significant correlation between it and other parameters of interest. We used the Tinti–Mulargia algorithm for calculating the b-value since it is proven to be robust for limited data sets, even though it also has its limitations. In particular, the b-value shows no correlation with C_V and the maximum magnitude in catalog (Fig. 3 B). This output sounds odd, since it is well known that regions hosting seismicity with higher b-value are usually hit by events with lower extreme sizes and vice-versa, which is also observed in different tectonic settings [40]. A possible solution to the puzzle can be found in the negative correlation between b-value and number of events above the completeness magnitude (Fig. 3 D). In fact, in New Zealand different tectonic regimes are mixed in the same region, i.e., thrust-, strike-slip or normal-faulting. New Zealand is featured by an evident crustal heterogeneity [41], various thickness of the seismogenic layer [42], fault systems with macroscopically different size, hence different maximum magnitudes, variable b-value and fractal dimension of the epicentral series. Therefore, the lack of correlation between b-value and other observables may be justified in the light of a mixture of tectonic settings, crustal heterogeneity and too various corner magnitudes, which should be investigated separately to distinguish between contrasting effects, in agreement with [43–47]. Furthermore, particular attention and criticism have emerged in recently published articles on the reliability of the results obtained for the calculation of the b-value as a function of the number of events used and the method adopted, e.g., [48], and on the spatial and temporal window at which it is estimated, also suggesting a multi-scale approach [49]. At last, the positive correlation between fractal dimension of the epicenters and seismic rate, C_V and maximum magnitudes is a further evidence of the crucial role of clustering in shaping seismic dynamics. Globally clustered seismicity dissipates a larger amount of energy with respect to sequences with lower C_V (Fig. 4 A) even though no clear difference is observed in the number of events above the completeness magnitude (Fig. 4 B). So, large earthquakes tend to occur where more intense seismicity has already taken place. In the present work, we limited our investigation to epicentral series in order to keep bounded the uncertainty of D_f, which otherwise would have been made larger by the hypocentral depth error. Thus, further studies are required to better understand the role of fractal dimensions and their predictive power about unobserved extreme events. In order to do so, large, relocated catalogs based on 3D local velocity models for seismic waves are needed.

4.2. Physical justification and implications

So far, we have characterized seismicity finding mutual relationships between statistical parameters and, above all, between the features of small-to-intermediate events and unobserved large ones; however, we did not focus on the physical mechanisms producing such differences. In this paragraph, we suggest a possible physical justification of our result.

The number of faults, N, in a certain region as a function of their surface, S, is well described by the empirical law [50,51]

$$N \propto S^{-\zeta} \tag{9}$$

where ζ is a suitable scaling exponent. Each fault can be connected to the magnitude of the largest event it can nucleate via the following relationship [52]

$$M_w = \frac{2}{3}\log(\mu S u) - 6.1 \tag{10}$$

where u represents the average cumulative slip over the fault surface occurred during the earthquake and μ is the shear modulus describing the resistance of the interface (usually in the order of 30 GPa). Fixed the size of the investigated region, so that also the extent of the largest seismogenic source is hypothesized to be somehow constrained, crustal volumes with higher values of ζ are featured by a more fractured tectonic setting with a large number of smaller faults; conversely, a lower ζ corresponds to a poorly developed fracturing with stress mainly accommodated by a few main surfaces of dislocation. The first situation is often observed in rifting zones, e.g., [53], transcurrent boundaries, e.g., [54], and intraplate slow deforming regions; while the second is common along subduction zones. In the first kind of setting, off-fault seismicity is quite common and occurs along small faults distributed throughout the crustal volume; while, in the second one, in-fault events dominate statistics, being mainly nucleated along large planar faults. So, in the former case, seismicity is mainly given by the occurrence of many smaller events that tend to fill the entire crustal volume; this implies that the proportion $C(r)$ of pairs of events with interdistance lower than r tends to increase, and, as a consequence, D_f is higher than in the latter case. Since the events tend to fill the crustal volume, spatial correlations would be likely to be established. Although a clear physical justification of the positive relationship between D_f and C_V cannot be assessed, it has been, however, observed that aftershocks, characterized by high time clustering, are mainly enhanced by static and post-seismic triggering [55–57], whose effectiveness is controlled by a power-law decaying kernel as a function of distance from the perturbing event, e.g., [58]. So, $D_f \propto C_V$. Therefore, broadly fractured zones are also prone to higher probability of complex cascading rupture of different fault segments whose activation is triggered by stress transfer, with respect to fault systems with concentrated fracturing. The seismological consequence of this property is that complex fault systems host longer seismic sequences featured by several events with large magnitude, e.g., in agreement with [59]; therefore, over long-lasting intervals, seismicity appears to be more clustered in time with respect to what is expected to happen along structurally simple plate boundaries, where large, Poissonian run-away ruptures have larger chances to occur [40]. Moreover, it is worthy to notice that, since at the scale of resolution ξ the surface of interface follows $S(\xi) \propto \xi^{-D_f}$ [60], where D_f in this case represents the fractal dimension of hypocenters, taking as a reference an ideal fault system with just one planar dislocation surface, if S_0 is its area, then, for a generic fault system

$$S(\xi) \propto S_0\, \xi^{-D_f + D_{f(0)}}. \tag{11}$$

For large-scale faulting $D_f = 2.20 \pm 0.05$ [33] is an accepted value. However, a wide variability has been found with a significant difference between tectonic settings, as supported by the empirical relationship $D_f \approx 2b$ [61], and depending on depth. Being the number of earthquakes as a function of the seismic moment M_0 power-law distributed (the Gutenberg–Richter law), taking into consideration Eq. (9) and the moment conservation principle [62], a higher seismic rate is expected in structurally complex settings. In addition, it is widely accepted by the scientific community that seismicity is controlled by spatial rheology gradients [63], rock physics on faults [64] and geometrical complexities of the seismogenic sources [65,66]. A mechanically weak interface (low μ) in poorly-competent settings is not able to bear an elevated spatial concentration of stress, producing diffuse small magnitude seismicity along the interface, e.g., it is the case of silent events along the shallow section of megathrusts [67]; conversely, strong rheology (elevated μ values) and highly-coupled interfaces favor strain and stress accumulation and, therefore, large seismic events. In the case corresponding to low D_f, b-value and C_V, even major seismic events tend to be featured by relatively low stress drops, e.g., [68]. All these properties are observed along subduction zones and faults with an elevated degree of maturation, in agreement with [23,69]. Opposite conclusions can be achieved for the other type of setting. In the light of this, we suggest that a connection between geophysical properties of crustal volumes, state of fracturing, statistical and clustering features of seismicity exists in agreement with previous research, e.g., [70], with implications for seismic hazard. Since the geological properties change over extremely long time periods, the seismic behavior can be

considered almost invariant for the temporal intervals we are interested in. Lastly, at the spatial scales involved by seismicity from micro-events to large ones, the mechanical properties of rocks can be roughly considered the same since rheology gradients usually act at regional lengths. This implies that also the mechanism of stress accumulation and release can be considered almost scale invariant by a seismological point of view. Therefore, where small magnitude events showcase more clustered behavior in space and time, it is also likely that the same may be true for major seismicity in the same area, even though still not reported in catalog.

5. Conclusions

Clustering is the essence of earthquake occurrence. Although not exhaustive, our analysis is a first attempt to understand what information may be hidden in partial, limited earthquake catalogs only containing mid-size and a few large seismic events (or even no one) about the largest possible ones. We consider the local and global coefficients of variation, the scaling exponent of the Gutenberg–Richter law, the fractal dimension of epicentral series D_f, the seismic rate and the number of events. We find that the largest earthquakes occur in locally Poissonian systems ($L_V \approx 1$) with globally clustered dynamics ($C_V > 1$). What is more interesting, we highlight that, while the local clustering behavior is strongly dependent on the size of catalog, so that longer databases tend to be less regular and more Poissonian than shorter ones, the global coefficient seems to be a reliable parameter even in cases of rather limited available information (few thousand events). This observation can be explained in the light of self-similar dynamics, so, analogous patterns producing both intermediate, large and extreme earthquakes. The fractal dimension of spatial series is positively correlated with the seismic rate, C_V and, therefore, with the maximum observed magnitude in catalog. Conversely, the *b*-value does not show any correlation with the principal observables except for the number of earthquakes. This phenomenon is due to the different sizes of mainshocks in various tectonic settings, namely higher in thrust-faulting than strike-slip- and normal-faulting regions, which are known to be characterized, respectively, by increasing *b*-values. Further research must be done in order to understand the real potential of these statistical parameters in the field of earthquake forecasting. We propose that their predictive power stems from the self-similar nature of slow dynamics producing the emergence of avalanches and breakdown in complex systems such as the brittle crust. In the light of this, global observables are the best candidates for modeling the long-term physical processes ultimately responsible for large earthquakes. By a physical viewpoint, they are also the most appropriate to grab and to describe mathematically the long-range interactions developing in disordered and critical settings while approaching dynamical transitions. So, our attention should be drawn to find new global collective observables showing predictive power and to apply them in order to set up new models of earthquake occurrence. Prospectively, this approach can be of great interest, once tuned, to extrapolate the features of extreme, still unobserved events given a limited database.

Funding

This research is funded by Sapienza University of Rome via the Ph.D. students Starting Research Fund 2022 (Type 1) awarded by Sapienza to Davide Zaccagnino for his research project “Statistical analysis of earthquake clusters for enhanced seismic hazard assessment”.

CRediT authorship contribution statement

Davide Zaccagnino: Conceptualization, Methodology, Software, Investigation, Writing–original draft, Visualization, Project administration. **Luciano Telesca:** Methodology, Visualization, Supervision. **Carlo Doglioni:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data available at https://quakesearch.geonet.org.n/ and https://earthquake.usgs.gov/earthquakes/search/

Acknowledgments

Authors acknowledge fruitful discussions with James Neely, Boris Rösler and Seth Stein and thank Rita Di Giovambattista and two anonymous reviewers for their comments which greatly improved the manuscript. All authors have read and agreed with the published version of the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.chaos.2023.113419.

References

[1] Kanamori H, Brodsky EE. The physics of earthquakes. Rep Progr Phys 2004;67(8):1429.
[2] Brodsky EE, Mori JJ, Anderson L, Chester FM, Conin M, Dunham EM, et al. The state of stress on the fault before, during, and after a major earthquake. Annu Rev Earth Planet Sci 2020;48(1):49–74.
[3] Reid HF. The elastic-rebound theory of earthquakes. Univ Calif Publ Bull Dept Geol 1911;6(19):413–44.
[4] Field EH. Computing elastic-rebound-motivated earthquake probabilities in unsegmented fault models: A new methodology supported by physics-based simulators. Bull Seismol Soc Am 2015;105(2A):544–59.
[5] Rundle JB, Turcotte DL, Shcherbakov R, Klein W, Sammis C. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. Rev Geophys 2003;41(4).
[6] Zaccagnino D, Telesca L, Doglioni C. Scaling properties of seismicity and faulting. Earth Planet Sci Lett 2022;584:117511.
[7] Bak P, Tang C. Earthquakes as a self-organized critical phenomenon. J Geophys Res: Solid Earth 1989;94(B11):15635–7.
[8] Zaccagnino D, Doglioni C. Earth’s gradients as the engine of plate tectonics and earthquakes. La Rivista Del Nuovo Cimento 2022;1–81.
[9] Sornette D. Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: Concepts and tools. Springer Science & Business Media; 2006.
[10] Kagan YY, Jackson DD. Long-term earthquake clustering. Geophys J Int 1991;104(1):117–33.
[11] Marco S, Stein M, Agnon A, Ron H. Long-term earthquake clustering: A 50,000-year paleoseismic record in the Dead Sea Graben. J Geophys Res: Solid Earth 1996;101(B3):6179–91.
[12] Corral Á. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys Rev Lett 2004;92(10):108501.
[13] Zaliapin I, Ben-Zion Y. Earthquake clusters in southern California I: Identification and stability. J Geophys Res: Solid Earth 2013;118(6):2847–64.
[14] Neely JS, Salditch L, Spencer BD, Stein S. A more realistic earthquake probability model using long-term fault memory. Bull Seismol Soc Am 2022.
[15] Ogata Y, Zhuang J. Space–time ETAS models and an improved extension. Tectonophysics 2006;413(1–2):13–23.
[16] Salditch L, Stein S, Neely J, Spencer BD, Brooks EM, Agnon A, et al. Earthquake supercycles and long-term fault memory. Tectonophysics 2020;774:228289.
[17] Mulargia F, Stark PB, Geller RJ. Why is probabilistic seismic hazard analysis (PSHA) still used? Phys Earth Planet Inter 2017;264:63–75.
[18] Stein S, Geller RJ, Liu M. Why earthquake hazard maps often fail and what to do about it. Tectonophysics 2012;562:1–25.
[19] Baker J, Bradley B, Stafford P. Seismic hazard and risk analysis. Cambridge University Press; 2021.
[20] Wells DL, Coppersmith KJ. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 1994;84(4):974–1002.
[21] Stein S, Geller R, Liu M. Bad assumptions or bad luck: Why earthquake hazard maps need objective testing. Seism Res Lett 2011;82(5):623–6.

[22] Kempf P, Moernaut J. Age uncertainty in recurrence analysis of paleoseismic records. J Geophys Res: Solid Earth 2021;126(8). e2021JB021996.
[23] Thakur P, Huang Y. Influence of fault zone maturity on fully dynamic earthquake cycles. Geophys Res Lett 2021;48(17). e2021GL094679.
[24] Rundle JB, Stein S, Donnellan A, Turcotte DL, Klein W, Saylor C. The complex dynamics of earthquake fault systems: New approaches to forecasting and nowcasting of earthquakes. Rep Progr Phys 2021;84(7):076801.
[25] GNS-Science. GeoNet Aotearoa New Zealand Earthquake Catalogue. 2023, http://dx.doi.org/10.21420/0S8P-TZ38.
[26] Kagan YY. Earthquake spatial distribution: the correlation dimension. Geophys J Int 2007;168(3):1175–94.
[27] Wiemer S, Wyss M. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. Bull Seismol Soc Am 2000;90(4):859–69.
[28] Woessner J, Wiemer S. Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 2005;95(2):684–98.
[29] Ristau J. Comparison of magnitude estimates for New Zealand earthquakes: moment magnitude, local magnitude, and teleseismic body-wave magnitude. Bull Seismol Soc Am 2009;99(3):1841–52.
[30] Telesca L, Lovallo M, Lopez C, Molist JM. Multiparametric statistical investigation of seismicity occurred at El Hierro (Canary Islands) from 2011 to 2014. Tectonophysics 2016;672:121–8.
[31] Chelidze T, Vallianatos F, Telesca L. Complexity of seismic time series: Measurement and application. Elsevier; 2018.
[32] Tinti S, Mulargia F. Confidence intervals of b values for grouped magnitudes. Bull Seismol Soc Am 1987;77(6):2125–34.
[33] Goltz C. Fractal and chaotic properties of earthquakes. Fractal and Chaotic Properties of Earthquakes 1997;3–164.
[34] Vulpiani A, Cecconi F, Cencini M. Chaos: From simple models to complex systems. vol. 17, World Scientific; 2009.
[35] Grassberger P, Procaccia I. Characterization of strange attractors. Phys Rev Lett 1983;50(5):346.
[36] Haines AJ, Wallace LM. New Zealand-wide geodetic strain rates using a physics-based approach. Geophys Res Lett 2020;47(1). e2019GL084606.
[37] Zaccagnino D, Telesca L, Doglioni C. Different fault response to stress during the seismic cycle. Appl Sci 2021;11(20):9596.
[38] Zaccagnino D, Telesca L, Doglioni C. Correlation between seismic activity and tidal stress perturbations highlights growing instability within the brittle crust. Sci Rep 2022;12(1):1–14.
[39] Zaccagnino D, Telesca L, Doglioni C. Variable seismic responsiveness to stress perturbations along the shallow section of subduction zones: the role of different slip modes and implications for the stability of fault segments. Front Earth Sci 2022;2192.
[40] Bird P, Kagan YY. Plate-tectonic analysis of shallow seismicity: Apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings. Bull Seismol Soc Am 2004;94(6):2380–99.
[41] Montuori C, Falcone G, Murru M, Thurber C, Reyners M, Eberhart-Phillips D. Crustal heterogeneity highlighted by spatial *b*-value map in the Wellington region of New Zealand. Geophys J Int 2010;183(1):451–60.
[42] Michailos K, Smith EG, Chamberlain CJ, Savage MK, Townend J. Variations in seismogenic thickness along the Central Alpine Fault, New Zealand, revealed by a decade's relocated microseismicity. Geochem Geophys Geosyst 2019;20(1):470–86.
[43] Mori J, Abercrombie RE. Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation. J Geophys Res: Solid Earth 1997;102(B7):15081–90.
[44] Kagan Y, Bird P, Jackson D. Earthquake patterns in diverse tectonic zones of the globe. Pure Appl Geophys 2010;167(6):721–41.
[45] Spassiani I, Marzocchi W. An energy-dependent earthquake moment–frequency distribution. Bull Seismol Soc Am 2021;111(2):762–74.
[46] Godano C, Tramelli A, Petrillo G, Bellucci Sessa E, Lippiello E. The dependence on the moho depth of the b-value of the Gutenberg–Richter law. Bull Seismol Soc Am 2022;112(4):1921–34.
[47] Zaccagnino D, Doglioni C. The impact of faulting complexity and type on earthquake rupture dynamics. Commun Earth Environ 2022;3(1):1–10.
[48] Marzocchi W, Spassiani I, Stallone A, Taroni M. How to be fooled searching for significant variations of the *b*-value. Geophys J Int 2020;220(3):1845–56.
[49] García-Hernández R, D'Auria L, Barrancos J, Padilla GD, Pérez NM. Multiscale temporal and spatial estimation of the *b*-value. Seism Res Lett 2021;92(6):3712–24.
[50] Kim Y-S, Sanderson DJ. The relationship between displacement and length of faults: a review. Earth-Sci Rev 2005;68(3–4):317–34.
[51] Torabi A, Berg SS. Scaling of fault attributes: A review. Mar Pet Geol 2011;28(8):1444–60.
[52] Hanks TC, Kanamori H. A moment magnitude scale. J. Geophys Res: Solid Earth 1979;84(B5):2348–50.
[53] Peacock DCP. Propagation, interaction and linkage in normal fault systems. Earth-Sci Rev 2002;58(1–2):121–42.
[54] Ross ZE, Idini B, Jia Z, Stephenson OL, Zhong M, Wang X, et al. Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. Science 2019;366(6463):346–51.
[55] Hardebeck JL, Nazareth JJ, Hauksson E. The static stress change triggering model: Constraints from two southern California aftershock sequences. J Geophys Res: Solid Earth 1998;103(B10):24427–37.
[56] Lippiello E, De Arcangelis L, Godano C. Role of static stress diffusion in the spatiotemporal organization of aftershocks. Phys Rev Lett 2009;103(3):038501.
[57] Toda S, Stein RS, Beroza GC, Marsan D. Aftershocks halted by static stress shadows. Nat Geosci 2012;5(6):410–3.
[58] Zhuang J, Werner MJ, Hainzl S, Harte DS, Zhou S. Basics models of seismicity: spatiotemporal models. 2011.
[59] Stein S, Liu M. Long aftershock sequences within continents and implications for earthquake hazard assessment. Nature 2009;462(7269):87–9.
[60] Theiler J. Estimating fractal dimension. J Opt Soc Amer A 1990;7(6):1055–73.
[61] Aki K. A probabilistic synthesis of precursory phenomena. Earthq Predict: An Int Rev 1981;4:566–74.
[62] Kagan YY. Seismic moment distribution revisited: II. Moment conservation principle. Geophys J Int 2002;149(3):731–54.
[63] Howarth JD, Barth NC, Fitzsimons SJ, Richards-Dinger K, Clark KJ, Biasi GP, et al. Spatiotemporal clustering of great earthquakes on a transform fault controlled by geometry. Nat Geosci 2021;14(5):314–20.
[64] Malservisi R, Furlong K, Dixon TH. Influence of the earthquake cycle and lithospheric rheology on the dynamics of the eastern California shear zone. Geophys Res Lett 2001;28(14):2731–4.
[65] Pondard N, Armijo R, King GC, Meyer B, Flerit F. Fault interactions in the Sea of Marmara pull-apart (North Anatolian Fault): earthquake clustering and propagating earthquake sequences. Geophys J Int 2007;171(3):1185–97.
[66] Romanet P, Bhat HS, Jolivet R, Madariaga R. Fast and slow slip events emerge due to fault geometrical complexity. Geophys Res Lett 2018;45(10):4809–19.
[67] Beroza GC, Ide S. Slow earthquakes and nonvolcanic tremor. Annu Rev Earth Planet Sci 2011;39:271–96.
[68] Courboulex F, Vallée M, Causse M, Chounet A. Stress-drop variability of shallow earthquakes extracted from a global database of source time functions. Seism Res Lett 2016;87(4):912–8.
[69] Sagy A, Brodsky EE, Axen GJ. Evolution of fault-surface roughness with slip. Geology 2007;35(3):283–6.
[70] Liu Y-K, Ross ZE, Cochran ES, Lapusta N. A unified perspective of seismicity and fault coupling along the San Andreas Fault. Sci Adv 2022;8(8):eabk1167.