JGR Solid Earth

RESEARCH ARTICLE
10.1029/2023JB027337

Are Foreshocks Fore-Shocks?

Davide Zaccagnino[1], Filippos Vallianatos[2,3], Georgios Michas[2], Luciano Telesca[4], and Carlo Doglioni[1,5]

[1]Department of Earth Sciences, Sapienza University, Rome, Italy, [2]Section of Geophysics-Geothermics, Department of Geology and Geoenvironment, National and Kapodistrian University of Athens, Athens, Greece, [3]Institute of Physics of Earth's Interior and Geohazards, UNESCO Chair on Solid Earth Physics and Geohazards Risk Reduction, Hellenic Mediterranean University Research & Innovation Center, Chania, Greece, [4]Institute of Methodologies for Environmental Analysis, National Research Council, Tito, Italy, [5]Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

Key Points:

- The occurrence of mainshocks can be preceded by a wide range of seismic patterns
- Foreshocks slightly differ from swarms in Southern California, but they can be hardly distinguished before mainshocks
- Foreshocks are not reliable precursors of large seismic events

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Zaccagnino,
davide.zaccagnino@uniroma1.it

Citation:

Zaccagnino, D., Vallianatos, F., Michas, G., Telesca, L., & Doglioni, C. (2024). Are foreshocks fore-shocks? *Journal of Geophysical Research: Solid Earth*, *129*, e2023JB027337. https://doi.org/10.1029/2023JB027337

Received 23 JUN 2023
Accepted 27 JAN 2024

Author Contributions:

Conceptualization: Davide Zaccagnino, Filippos Vallianatos, Georgios Michas, Luciano Telesca, Carlo Doglioni
Formal analysis: Davide Zaccagnino, Filippos Vallianatos
Funding acquisition: Carlo Doglioni
Investigation: Davide Zaccagnino
Methodology: Davide Zaccagnino, Filippos Vallianatos, Georgios Michas, Luciano Telesca
Project administration: Carlo Doglioni
Software: Davide Zaccagnino
Supervision: Filippos Vallianatos, Georgios Michas, Luciano Telesca, Carlo Doglioni
Validation: Davide Zaccagnino, Filippos Vallianatos, Georgios Michas
Visualization: Davide Zaccagnino
Writing – original draft: Davide Zaccagnino

Abstract Foreshocks are spatially clustered seismic events preceding large earthquakes. Since the dawn of seismology, their occurrence has been identified as a possible mechanism leading to further crustal destabilization, hence, to major failures. However, several cases occurred without any previous anomalous seismic activity, so that the hypothesis of foreshocks as reliable seismic precursors fails to pass statistical tests. Here, we perform an all-round statistical comparative analysis of seismicity in Southern California to assess whether any differences can be identified between swarms and foreshocks. Our results suggest that extremely variable seismic patterns can forerun mainshocks, even though they tend to be preceded by clusters with more numerous events spread over larger areas than swarms and with a wider range of magnitudes. We provide a physical explanation of such dissimilarity and conclude, despite it, that foreshocks can hardly be reliable short-term precursors of large earthquakes in California.

Plain Language Summary Large earthquakes can be preceded by a wide range of different seismic anomalies. Among these, seismicity has been reported to increase both in magnitude and frequency, but, on the other hand, it can also undergo a short period of reduced intensity before major events. The first pattern corresponds to the occurrence of foreshocks, that is, small to moderate quakes forewarning an upcoming larger one, while the second behavior is called seismic quiescence. In our research, we focus on foreshock activity. We perform an analysis of seismicity in Southern California, for which a well-provided relocated earthquake catalog is available. While several studies have been conducted so far about what happens before large earthquakes after their occurrence and also there are some works about foreshocks discrimination, a systematic analysis comparing properties of clusters of "swarms", seismicity without a major event, and "foreshocks" before their mainshock is missing. Are foreshocks different from swarms before the occurrence of the main event? Are foreshocks fore-shocks? Our results suggest that foreshocks can hardly be distinguished from swarms until the largest event takes place. On the base of this analysis and theoretical modeling, we think that foreshocks have limited reliability, if considered alone, for short-term forecasts.

1. Introduction

Earthquakes are dynamical instabilities in the brittle crust which tend to propagate in space and time via stress transfer (Belardinelli et al., 2003). While the mechanisms producing earthquake clustering, at least in their key features, are well established both in terms of static (Hainzl et al., 2010; King et al., 1994) and dynamic triggering (Kilb et al., 2000; Velasco et al., 2008), it is still debated whether the occurrence of earthquake clusters may be informative about the impending nucleation of larger events in their surroundings or not (Ellsworth & Bulut, 2018; Mignan, 2014; Ogata et al., 1996; Reasenberg, 1999; Trugman & Ross, 2019; van den Ende & Ampuero, 2020). Since the dawn of earthquake science increased seismic activity with respect to previously observed rate has been occasionally reported before large events (e.g., Dodge et al., 1996; Picozzi et al., 2023; Yagi et al., 2014; Zhu et al., 2022); nonetheless, an almost similar or even larger number of cases is known with ambiguous or clearly no foreshock activity (e.g., Gentili et al., 2017; Wiemer & Wyss, 1994; Wu & Chiao, 2006).

Moreover, the hypothesis of foreshocks as reliable precursors has more than once failed statistical tests (Hardebeck et al., 2008; Kagan & Knopoff, 1987; Marzocchi & Zhuang, 2011; Rhoades & Evison, 1993). Here, the term "seismic precursor" means whatever physical observable (a quantity that can be an output of measures or their mathematical elaboration) whose value can be tested to be causatively and positively correlated with the

Writing – review & editing: Davide Zaccagnino, Filippos Vallianatos, Georgios Michas, Luciano Telesca, Carlo Doglioni

magnitude of the mainshock and negatively related to the time to failure. On this concern, seismic quiescence has also been observed and tested successfully (Huang, 2006).

Therefore, the lack of systematic precursory patterns and the limited number of observations of pre-event anomalously high rate of occurrences - which is far to be ubiquitous - seriously question the routinely proposed prognostic value of foreshock activity by an observational point of view. Unquestionably, the occurrence of seismic events fosters the destabilization of crustal volumes at spatial ranges depending on their magnitudes, which is, after all, the physical content behind the epidemic type aftershock sequence (ETAS) models which predict higher seismic rates after a shock; however, this is quite different to claim that such instability will entail a higher probability of larger earthquakes. Updated ETAS models have also been proven to be effective in modeling foreshocks (Petrillo & Lippiello, 2021) and to explain other features of seismic sequences such as the magnitude of the second largest event, that is, the Båth's law (Zhuang, 2021). However, no evidence has been provided that foreshocks are generated by different physical mechanisms or they are essentially different from other seismic events (Felzer et al., 2004). For instance, the intensity and duration of "foreshock activity" has so far been found independent of the magnitude of the mainshock (Helmstetter & Sornette, 2003). Even by a geological point of view, the nucleation of earthquakes is conditional to a suitable combination of available energy budget and interface weakness to allow the dislocation to occur. Without the becoming internal condition, an earthquake cannot be nucleated, even though nearby stress perturbation produces destabilization: no seismic event can be informative about the state of stability of faults far from those on which it happens.

This viewpoint is in agreement with bifurcation theory and statistical mechanics of disordered and critical systems which predict two distinct patterns preceding large seismic events and corresponding to a sub-critical and super-critical dynamics (Sornette, 2006). The first one is featured by a progressive cascading destabilization due to a strong stress interaction between unstable fault patches, which causes clustered seismic activity (foreshocks); while the second scenario is characterized by a completely locked system with power-law increase of the amplitudes of fluctuations of order parameters before major breakdowns. A mix between these two extreme behaviors is possible. Anyway, while the super-critical dynamics is characterized by precursors (increasing correlations within the region approaching the transition from stability to rupture), sub-critical patterns do not show clear signals marking the evolution toward widespread destabilization. As a consequence, there is no clear theoretical evidence that some seismic events may be used as a proxy for the occurrence of larger ones. Friction and fracture mechanics support the idea of a preseismic nucleation phase with preslip or foreshocks (Cattania & Segall, 2021; Main & Meredith, 1989); however, classical physical frameworks are neither the standard nor the most reliable choice to modeling complex systems featured by long-range interactions and their predictive potential should be painstakingly tested if their use would not be limited to the coseismic phase.

Nevertheless, the interest of scientific community is high on this topic, further even more recently, because of the hope that the unprecedented amount of data provided by automated localization techniques can finally allow to detect even the tiniest seismic anomaly before a major event (Beroza et al., 2021). This trust is rekindled by the results of laboratory earthquake experiments, providing compelling and widespread evidence of precursory slip events before failures (McLaskey, 2019). Nonetheless, also in this concern, it is worth to remind that the behavior of complex systems is strongly affected by the spatial and temporal scale of investigation and definitely labquakes are just distant relatives of tectonic ruptures, being the first ones mainly characterized by short-range interactions and exponential decays (Marone, 1998; Rice et al., 2001), while the second ones by power-laws (Corral, 2004; Vallianatos et al., 2018; Zaccagnino et al., 2022) and by a far higher number of degrees of freedom; moreover, evidently, temporal and spatial scales are different (e.g., Nielsen et al. (2016)). Hence, a straightforward transfer of observations from the laboratory to real fault systems and the way round is not physically justified except for peculiar conditions (Saleur et al., 1996). This matter of fact is as well-known as consciously ignored. Even differences and similarities in the behavior of small and large earthquakes are still not definitely understood (Pacheco et al., 1992; Sornette, 2009).

At last, looking for common tiny effects when the macroscopic pattern of events preceding the largest one can vary in a wide range of possibilities, that is, from seismic quiescence to intense seismic activity, is not a reasonable and scientific way to tackle with the issue. Of course, the lower the completeness magnitude, the higher the chance that before a large seismic event a few smaller ones can be detected; however, this does not guarantee any causative relationship between them: the overwhelming majority of clusters does not trigger

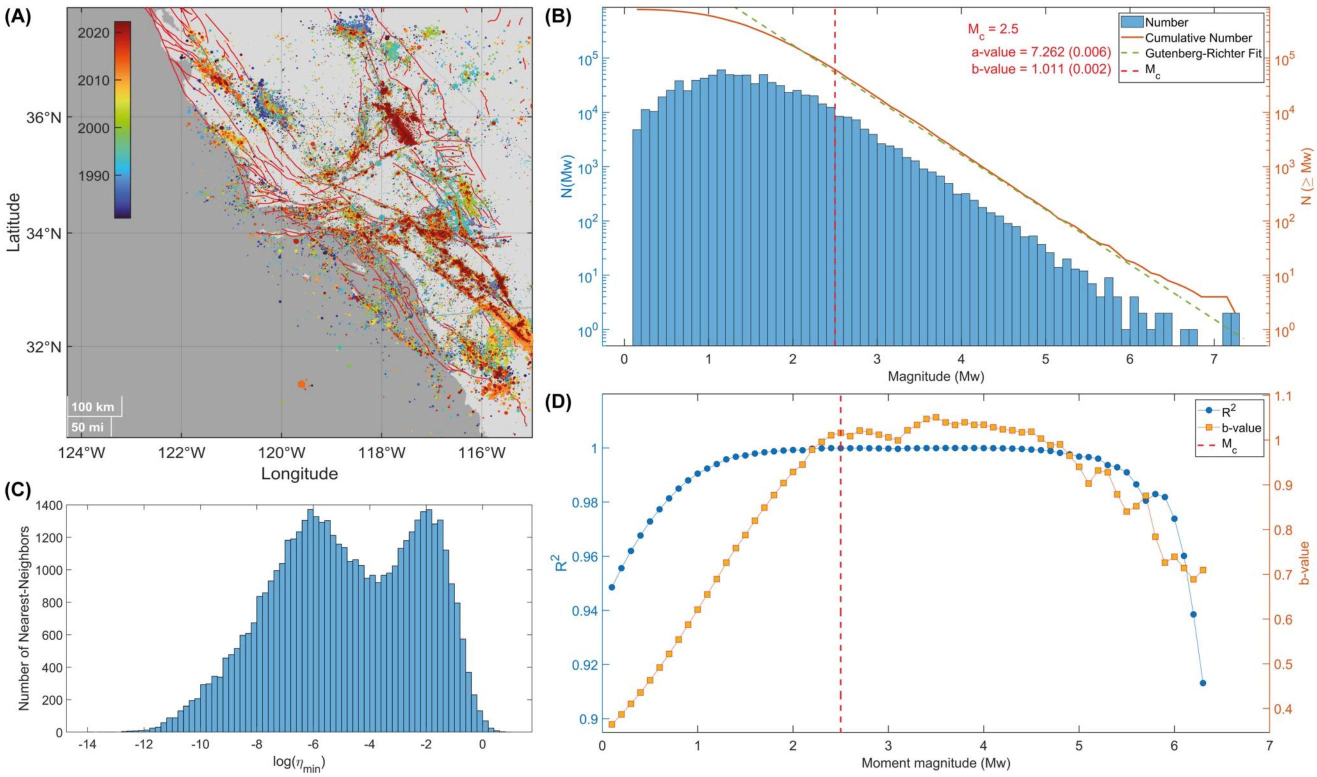

Figure 1. (a) Map of shallow (depth smaller than 30 km) seismicity in Southern California listed in the Waveform Relocated Earthquake Catalog for Southern California (1981–2022). (b) Frequency-size distribution of seismicity in Southern California; data are fitted using the Gutenberg-Richter law (green dashed line) above the completeness magnitude (red dashed vertical line). (c) Plot of the rescaled distance of nearest-neighbor seismic events in catalog. (d) Goodness of fit and stability analysis in the assessment of the b-value of the Gutenberg-Richter law in (b) as a function of the minimum considered magnitude. The red dashed vertical line represents the selected completeness magnitude. The R^2 is a measure of how successful the fit is in explaining the trends of data. $R^2 = 1$ means that the fit perfectly overlaps data.

relevant seismic events (e.g., Agnew & Jones, 1991; Ogata & Katsura, 2012; Savage & Rupp, 2000). They are just "swarms."

So, while there are plenty of studies about what happens before large earthquakes a posteriori (e.g., Brodsky and Lay (2014)), there are sparse analyses about foreshocks discrimination (Ogata & Katsura, 2012), and few rigorous investigations about the significance of foreshocks occurrence have been realized (e.g., Seif et al. (2019)), a research comparing statistical and structural properties of clusters of "swarms" and "foreshocks" before their mainshocks is lacking. Are there any differences? Do they share similarities? Can they be someway distinguished before the occurrence of the main event? Our work aims at providing some evidence to answer such questions and to find physical explanations of our results.

2. Data and Methods

In our research we analyze tectonic seismic events reported in the Waveform Relocated Earthquake Catalog for Southern California from 1981 to 2022 (Hauksson et al., 2012) above the completeness magnitude and with hypocentral depth shallower than 30 km (Figure 1a). The catalog is featured by high quality localizations with absolute errors in the order of a few kilometers and mostly affecting events with magnitude lower than 3.5 (Zaliapin & Ben-Zion, 2015). We select, following the methods hereafter, 5,486 clusters, but we focus our analysis only on those containing at least five events before the mainshock (the latter excluded). We classify clusters of events occurring before the largest shock into "swarms" and "foreshocks" according to the magnitude of the impending mainshock. We choose $M_w = 4.5$ as a suitable threshold to distinguish minor seismic sequences from larger ones. Our choice is based on the following reasons.

- If we pick the threshold larger than M_w 4.5, the number of clusters of foreshocks becomes too small to get statistically significant results in the comparison with swarms.

- If the threshold is selected smaller than M_w 4.5, we are too close to the completeness magnitude (see the next paragraph) and the range of magnitudes available for swarms becomes too tight.

So, we need a compromise. However, our analysis does not require a rigid definition and we also check other possible solutions in the range 4.0–5.5. We verified that our results did not change significantly, but their statistical quality is poorer.

2.1. Catalog Completeness

Even though the selection of the minimum magnitude does not affect significantly the scaling properties of clusters (Zaliapin & Ben-Zion, 2013), in order to get reliable comparison of their structural properties, a trustworthy estimate of the catalog completeness must be achieved. In our case, we do not require a strict respect of the Gutenberg-Richter law; however, a minimal stability of the scaling parameters has to be guaranteed. Classical goodness of fit methods return quite low completeness magnitudes, for example, $M_c = 1.5$ and $b \approx 0.8$ following the maximum curvature method (Wiemer & Wyss, 2000) with an additional +0.2 correction, but a clear underestimation of the true value is suggested by unreliably low b-value. On the other hand, more unyielding techniques, although producing extremely reliable estimation of the b-value (e.g., $M_c = 3.3$, $b \approx 1.04$ using the magnitude–frequency distribution method (Marzocchi et al., 2020)), strongly reduce the number of earthquakes that can be considered in our analysis (less than eight thousand). For this reason, we apply a simple goodness-of-fit test for different minimum magnitudes and, contextually, we also require the estimation of the b-value to be stable (percentage variation $\Delta_{\%} b \leq 1\%$ in our case). Therefore, the completeness magnitude is simply defined as the lowest magnitude allowing a stable calculation of the b-value and high goodness-of-fit quality defined as in Wiemer and Wyss (2000). See (Figures 1b and 1d). Our estimate returns $M_c = 2.5$ and, via maximum likelihood method, we get $b = 1.011 \pm 0.002$, which is compatible with results obtained by selecting higher M_c with the advantage that this choice allows to gain more than forty thousand events. At last, it has been noticed that the catalog also suffers from short-term incompleteness after large events even in the case in which only magnitudes larger than 2.5 are considered (Zaliapin & Ben-Zion, 2015). Moreover, some sparse events are not reported. However, we are only interested in events occurring before the mainshocks, so that $M_c = 2.5$ can be considered reliable and cluster statistics should not be affected by incompleteness.

2.2. Cluster Selection

Using routinely applied clusters detection methods, for example, the Zaliapin and Ben-Zion algorithm (Zaliapin & Ben-Zion, 2013), we would obtain about twenty thousand clusters with a mean of less than three events each, which prevents the possibility to compare their internal structures. In addition, even such rigorous techniques suffer from a certain degree of arbitrariness and limitations (Bayliss et al., 2019) (compare with the supplementary material for in-depth discussion). Therefore, we slightly modify it to allow to group seismicity into larger clusters.

The distance η_{ij} between the event i and the subsequent j is a generalization from the definition given in Zaliapin and Ben-Zion (2013):

$$\eta_{ij} = R_{ij}T_{ij} = \left[\left((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2\right)^{D_f/2} \times (t_j - t_i)\right] 10^{-bM_i}, \tag{1}$$

where each earthquake is featured by a hypocentral localization (x_i, y_i, z_i), being z its depth, and a catalog magnitude M_i, $D_f = 2.2$ is a good compromise for the fractal dimension of hypocentral series and rock fractures (Kagan, 1991; Sahimi et al., 1993) and $b = 1.01$. R_{ij} and T_{ij} represent the spatial and temporal distances re-scaled by the magnitudes of the triggered event (Zaliapin & Ben-Zion, 2013). In our analysis, we use just one spatially homogeneous b-value corresponding to the scaling exponent of the Gutenberg-Richter law of the whole catalog (which is not far from $b = 1$ used by Zaliapin and Ben-Zion). Indeed, the b-value of each seismic cluster cannot be reliably estimated because of the limited number of events. Moreover, its temporal variations, mostly due to artifacts (Zaliapin & Ben-Zion, 2015), make our effort to better assess its spatial value almost impractical. Moreover, we are interested in the b-value only in order to apply Equation 1, hence, for selecting the parametric distance between two seismic events, whose value is dominated by the magnitude of the triggering event. We do not consider nearest neighbors (represented Figure 1c). We aggregate clusters assuming that the distance between the first event and the triggered ones occur within $\eta \leq \eta_0 = 10^{\log(R_{ij})+\log(T_{ij})} = 10^{-1.46} = 0.035$ and $R_{ij} \leq R_0 = 35$.

R_0 corresponds to a lower saddle point in the distribution of spatial distances between seismic events in the catalog (Figure 2a), while the relationship for η_0 is chosen to reconcile the scaling trend of re-scaled times with the condition for R_{ij}. The detailed motivation, advantage, stability and physical implications of our method are discussed in the supplementary material.

2.3. Cluster Area

For each cluster i containing at least five events we calculate the surface A_i of the region hit by seismicity. We define it as the area of the convex envelope of the epicentral coordinates.

2.4. Cluster Entropy

Given the empirical distribution of probability of inter-event times and magnitudes in each cluster, p_i, approximated by their frequency, we calculate the Shannon entropy (Shannon, 1948; Telesca & Lovallo, 2017) according to its definition

$$S(p_i) = H(p_i) = -\sum_i p_i \log(p_i) \tag{2}$$

and the Tsallis entropy (Tsallis, 1988)

$$S_q(p_i) = T(p_i) = \frac{1}{1-q}\left(1 - \sum_i p_i^q\right) \tag{3}$$

with $q = 1.491$, in agreement with the output of the Sotolongo-Costa and Posadas fit (Sotolongo-Costa & Posadas, 2004) (see Figures S3a and S3b in Supporting Information S1) and compatible with previous estimations (e.g., Telesca (2011)).

2.5. Cluster Seismic Rate

Each cluster is featured by a seismic rate of events per unit area and time calculated using

$$N_R = \frac{N}{A_i \Delta t}, \tag{4}$$

with N standing for the number of earthquakes, and a seismic moment rate defined as

$$S_R = \frac{1}{A_i \Delta t} \sum_i 10^{1.5(M_i+6.1)} \tag{5}$$

where Δt is the duration of cluster activity from its onset until the mainshock.

2.6. Cluster Global Coefficient of Variation

The global coefficient of variation of interevent times, C_V (Kagan & Jackson, 1991), is given by

$$C_V = \frac{\sigma_\tau}{\langle \tau \rangle}, \tag{6}$$

where $\langle \tau \rangle$ is the mean value of the inter-event times and σ_τ is its standard deviation. It 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 (Telesca et al., 2016).

10.1029/2023JB027337

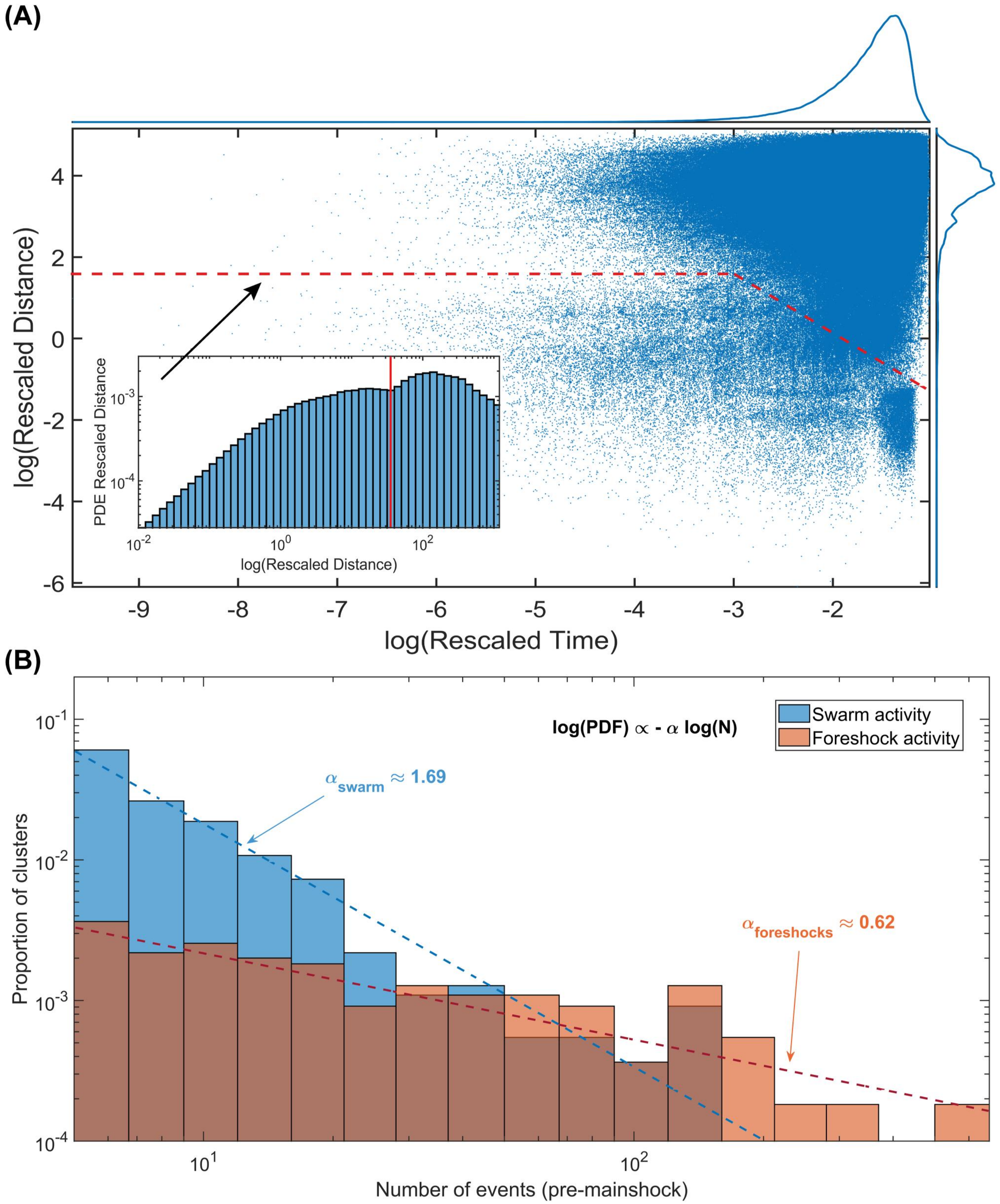

Figure 2.

2.7. Trends Within Clusters

In our analysis, we also investigate whether seismic activity shows peculiar trends within clusters or not. In order to do this, we calculate the mean value of consecutive differences of interevent times τ, $\tau_{i+1} - \tau_i$ and magnitudes, $M_{i+1} - M_i$, and their squared values as

$$\langle x \rangle = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i) \tag{7}$$

and

$$\langle x^2 \rangle = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 \tag{8}$$

respectively, where x stands for the variation of interevent times or consecutive magnitudes.

2.8. Testing Differences in Distributions

In order to assess whether there are any statistically significant differences in the distributions of some parameters of interest of foreshocks and swarms until the mainshock, we compare their empirical cumulative probability functions given by

$$F(X) = \int_{-\infty}^{X} f(x)\, dx, \tag{9}$$

where x represents the investigated variable and $f(x)$ is its empirical density function, using the two-sided non-parametric two-sample Smirnov test (Massey Jr, 1951). This procedure evaluates the quantity

$$\mathcal{D}\left[F_{swarm}, F_{foreshocks}\right] = \max_x\left(\left|F_{swarm}(x) - F_{foreshocks}(x)\right|\right) \tag{10}$$

which works as a functional distance between the two distributions. The Smirnov test detects a wide range of differences between two distributions, while other usually applied tests, such as t-tests, are sensitive to variations of means or medians but cannot highlight other differences in the shape of distributions. The test returns a p-value which is used to assess the probability that the two samples follow the same distribution. We take a significance level $\alpha = 0.001$ to reject the null hypothesis that foreshocks and swarms follow the same distribution. We use the smallest conventionally adopted α because of the large source of uncertainty in the estimation of the parameters of interest.

3. Results and Discussion

3.1. Analysis and Results

In our work we consider seismic events in clusters before the earthquake with the largest magnitude; seismicity preceding a mainshock with magnitude larger than or equal to 4.5 is classified as foreshocks, otherwise it is a swarm. A detailed motivation of our choice is provided in the section devoted to methods. Our analysis shows that larger earthquakes are forewarned, on average, by clusters with a higher number of events N, being the scaling exponent of the frequency-size relationship, $\propto 10^{-\alpha N}$, lower for foreshocks ($\alpha_{foreshocks} \approx 0.62$) than for swarms ($\alpha_{swarm} \approx 1.69$) in Southern California (compare with Figure 2b). However, the distribution of clusters as a

Figure 2. (a) Distribution of the relative distance of events in Southern California listed in the Waveform Relocated Earthquake Catalog (Hauksson et al., 2012). (b) Distribution of cluster sizes (pre-largest event, at least five seismic events) of swarms (blue bars) and foreshocks (orange bars). Only clusters with at least five events (mainshock excluded) are plotted in the histogram. If the magnitude of the largest seismic event in each cluster is higher than M_w 4.5, then it is classified as potential foreshock activity, otherwise it is a swarm. α represents the scaling exponent fitting the probability density function of the number of clusters as a function of the number of seismic events until the mainshock.

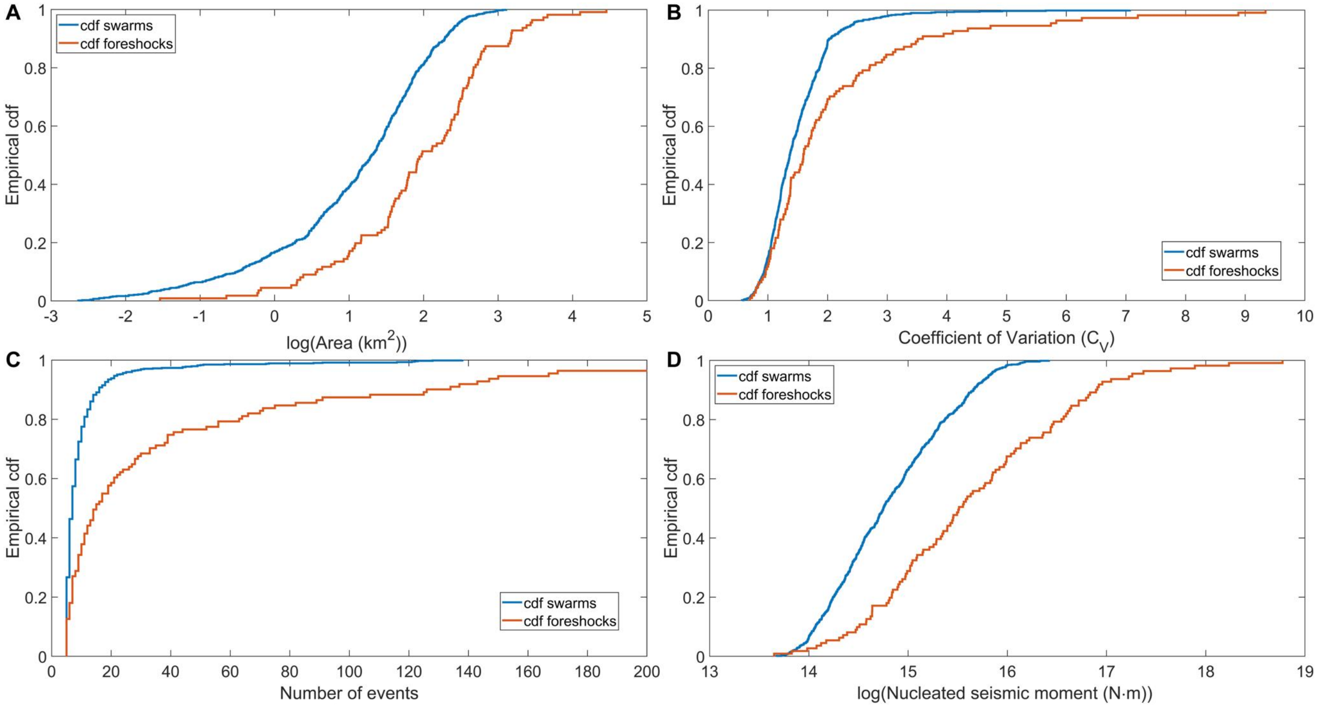

Figure 3. Cumulative distribution of the number of clusters of seismic events ($M_w \geq 2.5$) until the mainshock in Southern California (1981–2022) as a function of their various features: area (a), global coefficient of variation of interevent times (b), number of events (c), cumulative seismic moment (d).

function of their area (Figure 3a), global coefficient of variation of interevent times (Figure 3b), number of events (Figure 3c), nucleated seismic moment (Figure 3d) shows significant differences between foreshocks and swarms before the occurrence of the mainshock (see Table 1). Foreshocks tend to take place over larger areas, they are more numerous and with slightly higher magnitudes. Foreshocks also seem to be internally more clustered in time

Table 1
Results of the Non-Parametric Two Sample Smirnov Test for the Comparison of Distributions of Different Features of Seismic Clusters Preceding Mainshocks With Magnitude Smaller Versus Larger Than M_w 4.5 in Southern California

Parameter	p-value	Figure
Cluster area	$\mathbf{5.9 \times 10^{-11}}$	3a, 4c, and 8a
C_V interevent times	1.3×10^{-3}	3b
Number of events	$\mathbf{3.4 \times 10^{-14}}$	2b and 3c
Cumulative seismic moment	$\mathbf{9.7 \times 10^{-14}}$	3d
Tsallis entropy magnitudes	$\mathbf{5.9 \times 10^{-14}}$	9d and 10d
Tsallis entropy interevents	3.0×10^{-3}	9c and 10b
Shannon entropy magnitudes	$\mathbf{2.3 \times 10^{-9}}$	9b and 10c
Shannon entropy interevents	1.9×10^{-3}	9a and 10a
Seismic rate	0.028	4b, 5a, 7a and 7b
Seismic moment rate	0.68	4a, 5b, 7c and 7d
Duration	0.39	4d, 8a and 8b
Magnitude trend	0.081	6a
Interevents trend	$\mathbf{2.6 \times 10^{-10}}$	6b
Magnitude variance	$\mathbf{1.1 \times 10^{-26}}$	6c
Interevents variance	$\mathbf{2.8 \times 10^{-6}}$	6d

Note. The first column reports the analyzed parameter, the second one the p-value of the test and the third one the reference to figures. Significant p-values are in bold ($p \leq 0.001$).

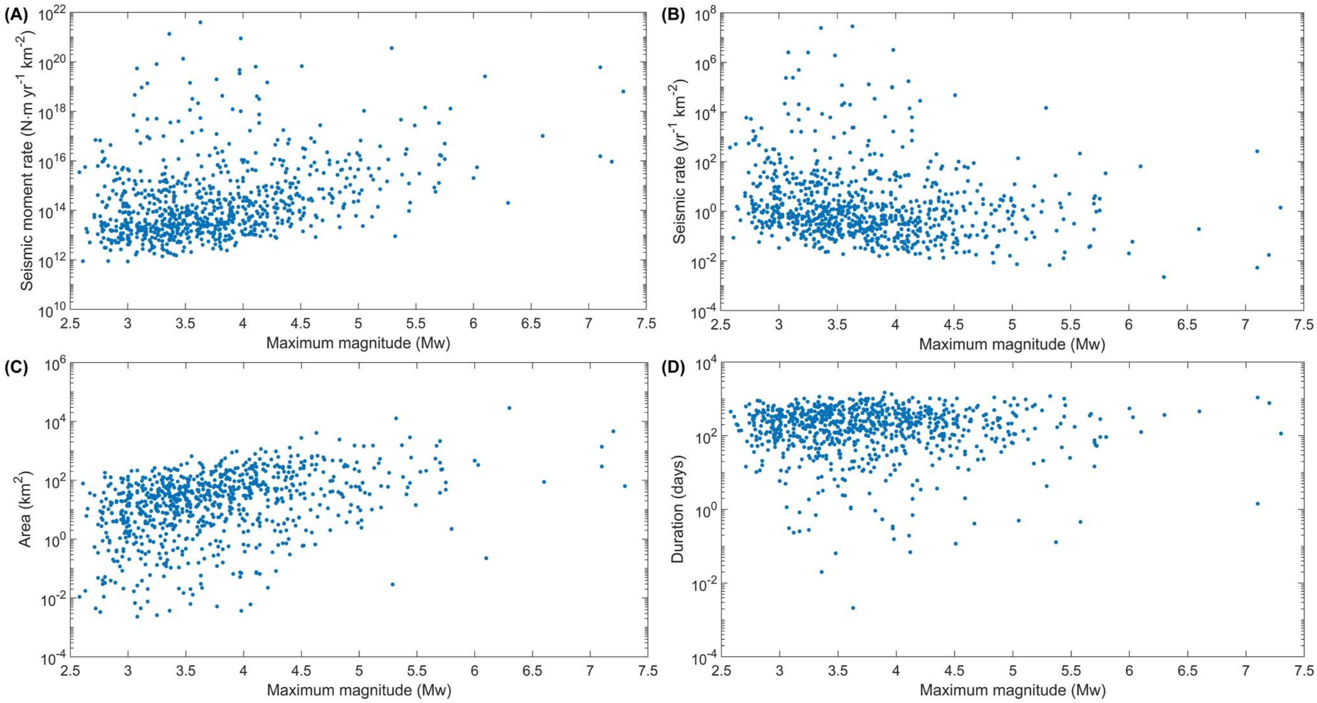

Figure 4. The seismic moment rate is positively correlated with the maximum magnitude in the cluster (a), while the number of earthquakes nucleated in unit measures of time and space weakly negatively related to the size of the largest seismic event (b). Moreover, while it is positively correlated with the surface where the events take place (c), the maximum magnitude is independent of the duration of previous seismic activity (d).

than swarms, but the Smirnov test reports a p-value higher than the chosen α significance value and does not allow us to reject the null hypothesis; therefore, the observed difference between the distributions is not statistically significant in this case. The seismic moment rate is very weakly correlated with the maximum magnitude of each cluster (Figure 4a), while the number of earthquakes nucleated per unit measure of time and space is weakly negatively related to the size of the largest seismic event (Figure 4b) without any difference between swarms and foreshocks. Moreover, the maximum magnitude is independent of the duration of previous clustered seismic activity (Figure 4d), which suggests that no relation exists between the temporal extent and intensity of "foreshocks" and the incoming mainshock. We also observe that the maximum magnitude is positively correlated with the size of the region where the events take place (Figure 4c). The cumulative distributions of the seismic rate and the seismic moment rate of foreshocks and swarms (represented in Figures 5a and 5b respectively) are almost overlapping. A quantitative analysis provided by the Smirnov test does not allow to reject the null hypothesis, that

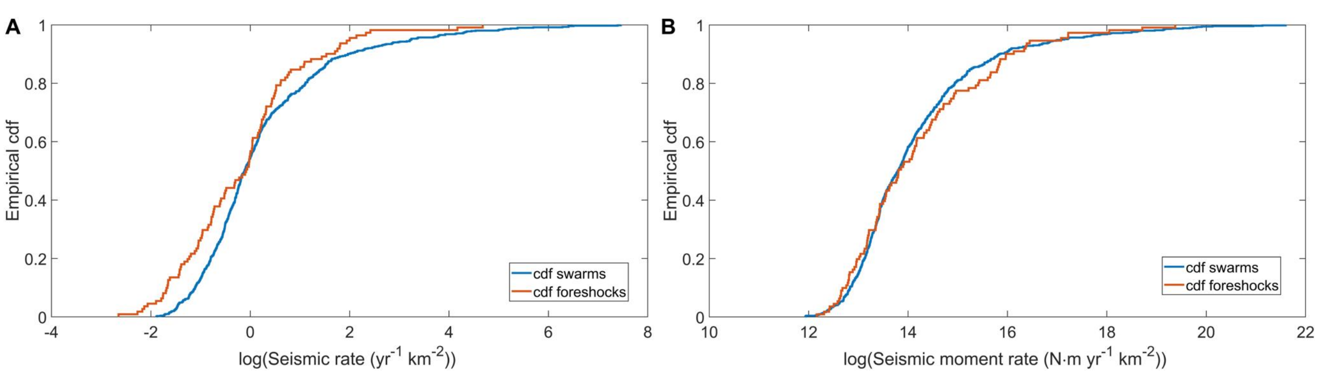

Figure 5. Cumulative probability distribution of the seismic rate (a) and of the seismic moment rate (b) of clusters of events ($M_w \geq 2.5$) occurring before the mainshocks in Southern California (1981–2022). The cumulative density function of foreshocks is plotted in orange, while the blue line represents swarms.

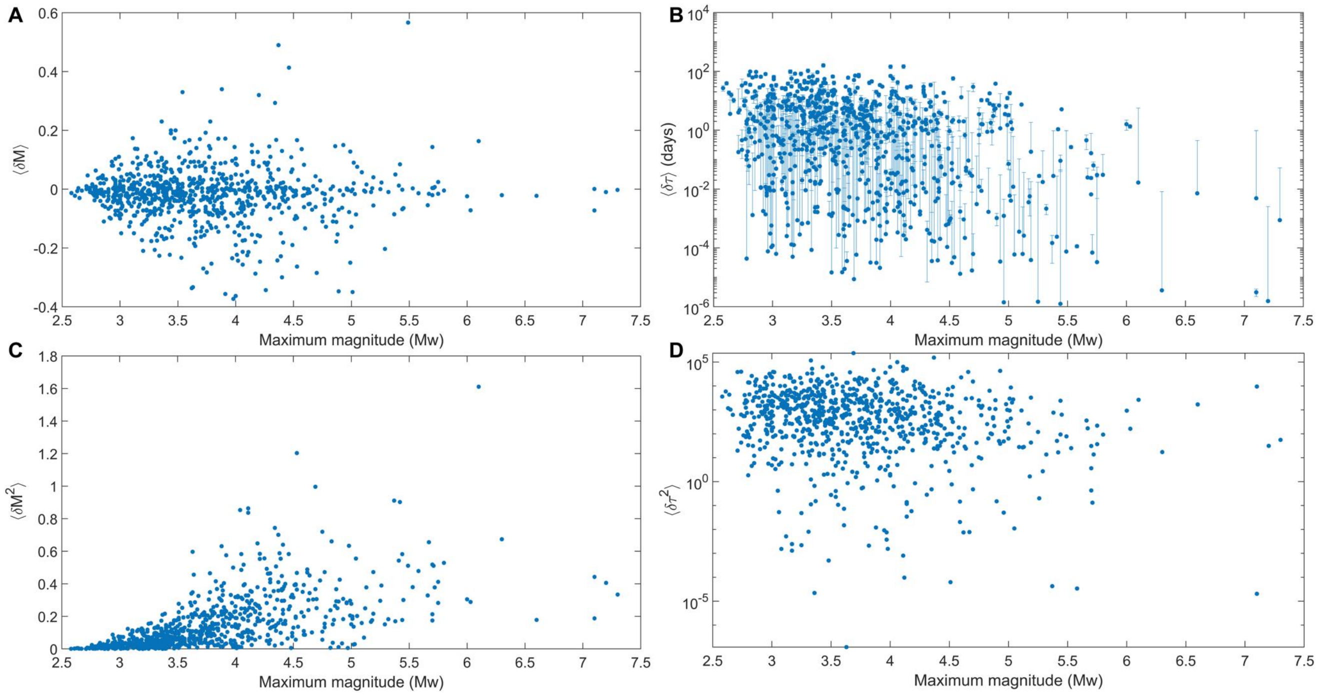

Figure 6. Mean variation of consecutive magnitudes (a), interevent times (b), squared magnitudes (c), and squared interevent times (d) as a function of the mainshock magnitude, according to the definitions given in Equations 7 and 8.

is, the two samples follow the same distribution. Results are listed in Table 1. During the temporal evolution of clusters, no clear trend in magnitudes is observed as a function of the size of the mainshock neither in swarm-like seismic activity nor in foreshocks (Figure 6a). Therefore, magnitudes do not tend to increase as the mainshock is approaching. This result is also supported by the output of the Smirnov test, which returns a high p-value (see Table 1). Conversely, the variance of magnitudes clearly increases with the size of the forthcoming mainshock (Figure 6c), while interevent times and their fluctuations decrease as a function of the largest magnitude (Figures 6b and 6d).

While different patterns are observed in foreshocks and swarms in terms of their spatial extent, amount of nucleated energy and number of seismic events, no difference is found in the scaling relationships connecting the seismic rate of clusters with the area where they occur and their total temporal duration (see Figure 7). Figure 8a shows the scaling relationship between the duration of pre-mainshock seismic activity as a function of the involved area represented by blue points in the case of swarms, while clusters of foreshocks are highlighted by orange dots. The scaling exponent, γ, is compatible with zero meaning that seismicity does not tend to diffuse significantly during the preseismic phase in California. Besides, the cumulative distribution function of the durations of clusters until the mainshock is the same for foreshocks and swarms (Figure 8b, for the output of the two-samples Smirnov test see Table 1.)

At last, no significant difference is detected between the Tsallis and Shannon entropy of interevent times of foreshocks and swarms, while the entropy of magnitudes is more smoothly distributed in the case of clusters leading to larger events (Figures 9b, 9d and 10); however, the Shannon entropy is also more sensible to fluctuations due to the (sometimes) extremely limited number of events contained in each cluster. If only data sets with more than 50 earthquakes are considered, the aforementioned difference in the information entropy distributions tends to vanish (compare with Figure S4 in Supporting Information S1). Moreover, a clear increasing trend exists in the relationship between the Shannon entropy and the size of clusters (Figure S4 in Supporting Information S1), which is only slightly present in the case of Tsallis entropy (the scaling exponent is compatible with zero within 2σ), coherently with the non-extensive nature of such physical quantity.

Our results, summarized in the outputs of the Smirnov tests listed in Table 1, suggest that foreshocks and swarms share the same scaling behaviors and are likely generated by the same physical mechanisms. Foreshocks have

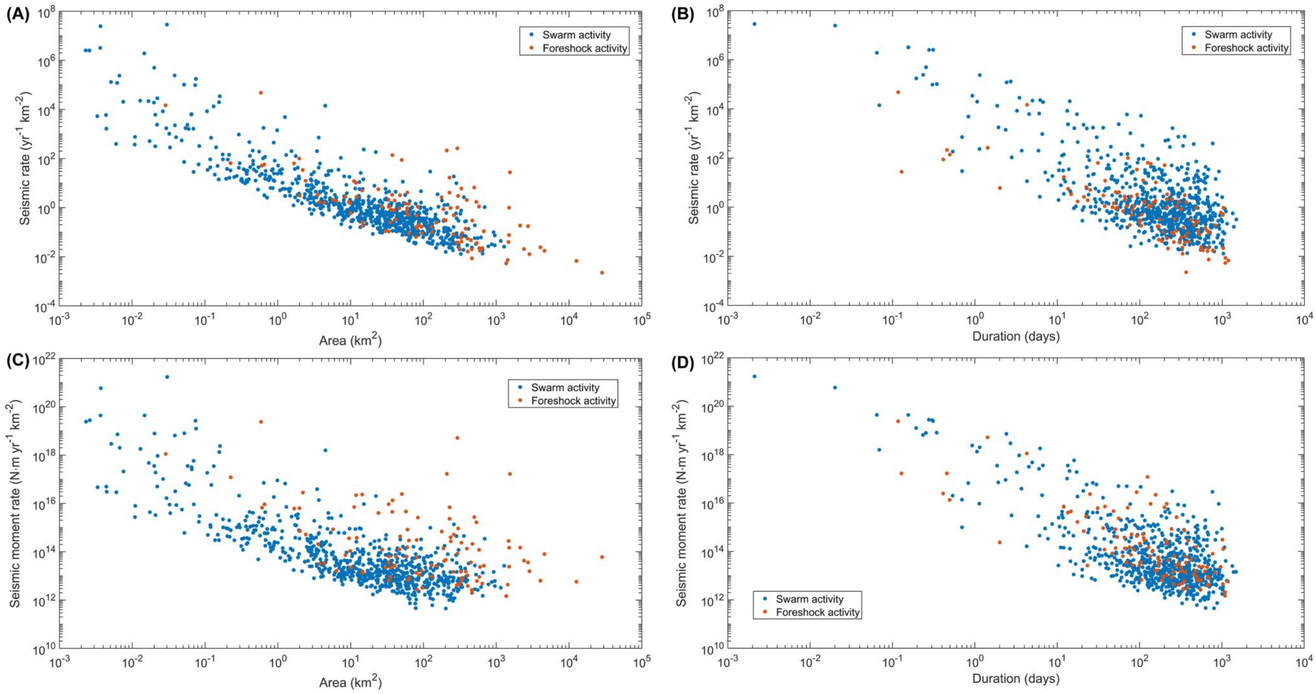

Figure 7. Seismic rate and seismic moment rate versus the involved area (a and c) and duration of clusters until the largest event occurs (b and d).

duration, seismic rate and seismic moment rate, as well as magnitude trends and clustering properties indistinguishable from swarms; however, statistical analyses highlight that foreshocks spread over larger areas, are featured by larger and more energetic clusters with also higher variance of magnitudes and relative Tsallis and Shannon entropies. We propose a possible explanation of such differences in the next paragraph.

3.2. Physical Modeling and Interpretation

Our results, as well as a long list of observational evidences about the occurrence of increased seismic activity before large earthquakes (McGuire et al., 2005; Mignan, 2014; Reasenberg, 1999; Zhu et al., 2022) prove that mainshocks can occur with or without foreshocks with extremely variable magnitudes. The classical explanation of this phenomenon advocates the development of two possible physical scenarios in the brittle crust corresponding to sub-critical or super-critical conditions. The first one is accompanied by cascading ruptures; while the second situation undergoes seismic quiescence. Here, we would like to discuss briefly a possible interpretation to understand why large earthquakes can be preceded by so different seismic patterns. Moreover, we would like to apply the same framework to explain why "foreshocks" seem to share some features with swarms, while others differ.

Clusters of seismic events are characterized by significant spatial and temporal correlations which mark the evolution of stress in the brittle crust. The occurrence of seismic events releases part of the differential stress accumulated during the interseismic phase on faults and within the adjoined crustal volumes. Each event not only modifies the state of stress locally (in the position R at time T), but also contributes to rearrange the mechanical stability state all around itself. So, theoretically, we can measure the two-point correlation function between the level of stress in (R, T) and $(R + r, T + t)$

$$C(r,t) = \frac{1}{N}\sum_{R,T} \sigma(R,T)\sigma(R + r,T + t) \tag{11}$$

Now, we are interested to understand how the past and present occurrence of seismic events might impact on the future spatial and temporal evolution of seismicity in such a correlated system. For the sake of simplicity, we start considering the crustal volume as a whole, so that we omit the spatial dependence of stress inside the crustal

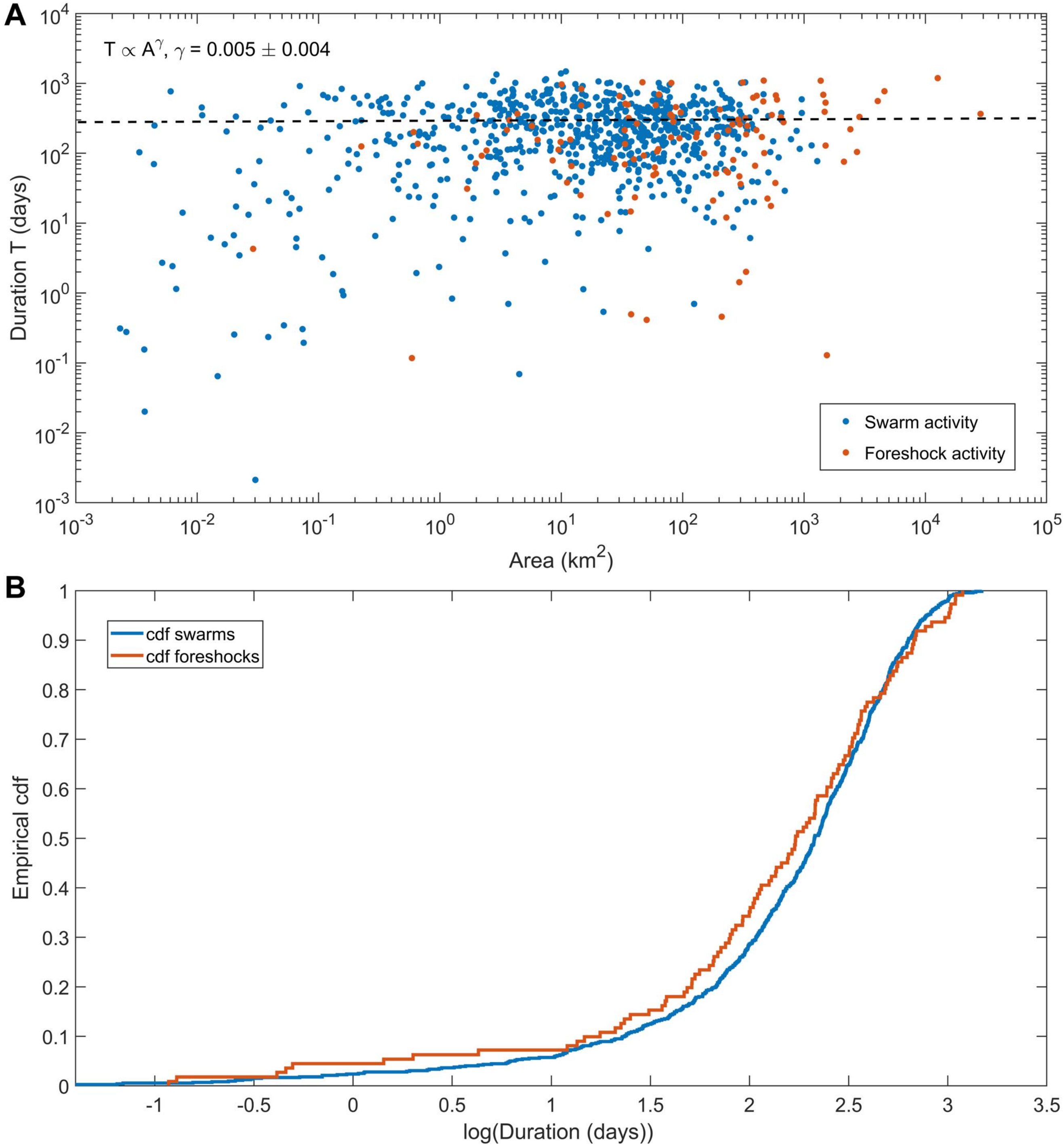

Figure 8. (a) Duration of pre-mainshock seismic activity of clusters of earthquakes as a function of the involved area (swarms are represented by blue points, foreshocks with orange dots) in Southern California from 1981 to 2022. (b) Cumulative distribution function of the durations of clusters of swarms until the mainshock (blue line) and foreshocks (orange line).

volume which is becoming unstable. However, the same derivation can be done for including the effect of spatial memory, not only the temporal one, on seismic activity. So, we write

$$\sigma(t)_{nonlocal} \sim \int_{\varepsilon}^{\omega} \sigma(t-s)\ K(\beta,\zeta;s)\ ds \tag{12}$$

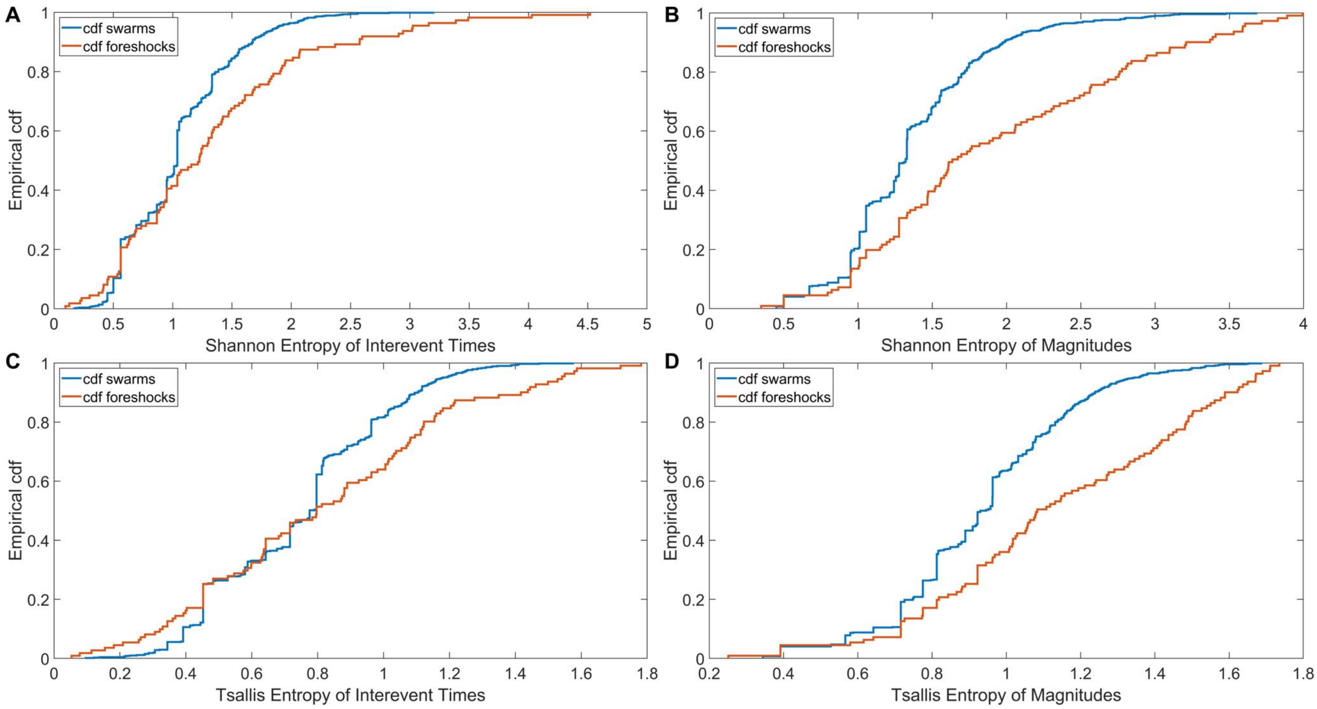

Figure 9. Cumulative distribution of the Tsallis (a and c) and Shannon entropy (b and d) of interevent times and magnitudes of clusters of seismic events before the mainshock in Southern California (1981–2022).

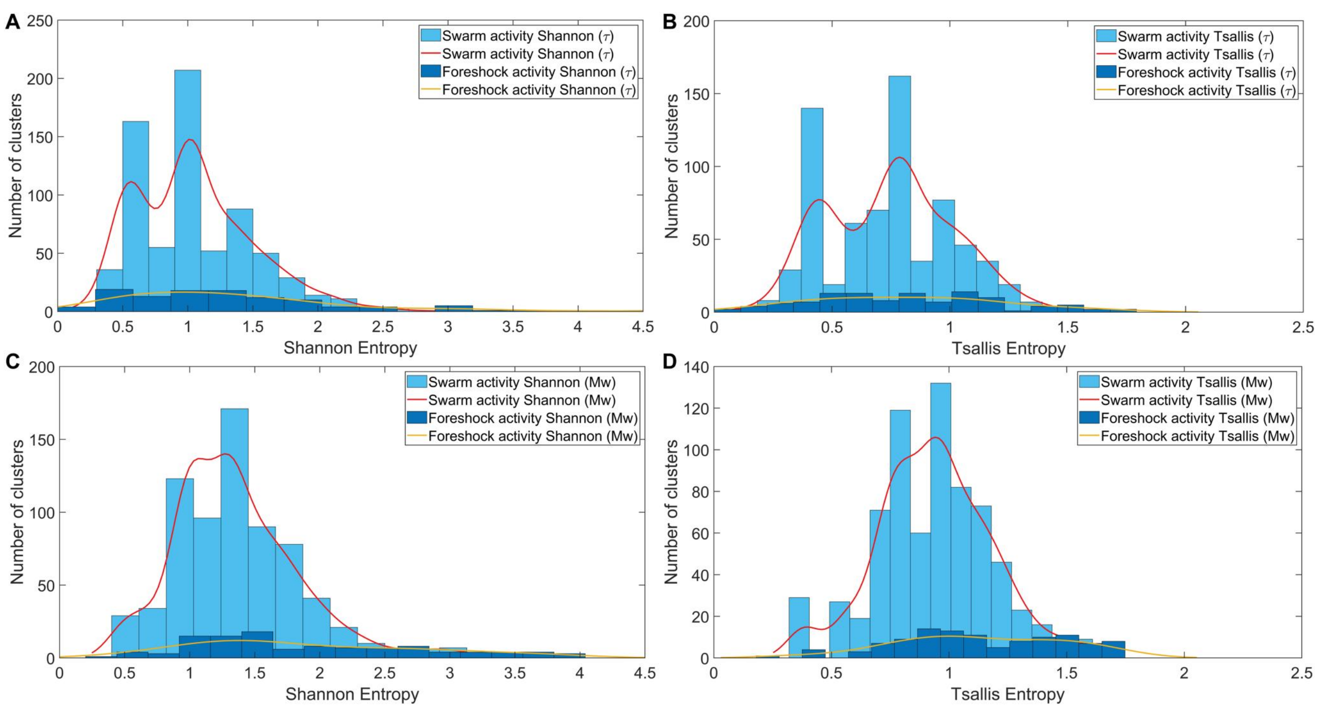

Figure 10. Shannon and Tsallis entropy distributions of interevent times (a and b) and magnitudes (c and d) of clusters of seismic events before the mainshock in Southern California (1981–2022).

where

$$K(\beta,\zeta;s) \sim e^{-\beta|s|^{\zeta}} \tag{13}$$

represents a memory kernel, which we assume to be exponential, even though this choice is not the most general one. ω represents the maximum temporal duration of the correlated system, while ε is a nominally vanishing time interval which guarantees causality. This formulation states that the value of stress featuring the whole system at a certain time is also affected by the history of recently happened variations of the stress itself, depending on the memory function. This non-local integral formula, that is, it takes into account of the contribution of the past in the evolution of the system, can be made simpler for studying the temporal variation of stress, assuming that the temporal shift can be extended to the entire past of the system; hence

$$\left.\frac{d}{dt}\right|_{nonlocal}\sigma(t) = \frac{1}{\mathbf{N}}\frac{d}{dt}\int_0^{\infty}\sigma(t-s)\ e^{-\beta|s|^{\zeta}}\ ds, \tag{14}$$

where $\mathbf{N}$ stands for an appropriate normalization. An analogous procedure can be applied to spatial memory. It is clear that if $K(\beta, \zeta; s)$ was a delta function, no memory effect would appear; therefore, fixed the initial condition, $\sigma(0)$, only one future evolution is possible. Conversely, in the case of Equation 14, the value of stress at local time t is not enough to determine the value of the overall variation of $\sigma(t)$, that is, future seismic activity. A weighted integral over (almost) all the past is necessary. This means that even the same history may result in different future evolution because of the effect of memory acting on a strongly correlated system. In short, even an identical past can result in drastically different future outputs because of a tiny change in the parameters of the system. The application to seismicity is the following: two identical seismic clusters can flow into a large mainshock, moderate events or a swarm depending on the action of tiny details in the evolution of stress gradients. On the other hand, two completely different seismic patterns can give rise to seismic events with similar features, for example, with the same magnitude. The 2009 M_w 6.3 L'Aquila earthquake and the 2016 M_w 6.2 Amatrice earthquake are emblematic cases of this. Both took place on normal faults in the Central Apennines in Italy, the first event was preceded by about 5 months of accelerating seismicity (Papadopoulos et al., 2010), while the second one hit with neither tiny forewarning shocks (Chiaraluce et al., 2022). Therefore, in our view, it is not a matter of improving earthquake detectability in order to record small precursory events; instead, it may be fundamental to identify and to model memory processes ultimately leading seismic activity in the brittle crust. Anyway, this result strongly challenges the possibility of accurate earthquake prediction, both in terms of time to failure and magnitude, at least just considering past seismic activity. Paradoxically, it is not due to limited and low-quality data sets, but to the nature of seismicity itself: fault systems do not know exactly what is going to occur until it happens. In summary, no seismic clusters should be considered to be precursors of large events, since part of the history of failures still stored in the memory of fault systems contains key information about their future.

This argument is a proposal to explain why some large earthquakes are preceded by smaller ones while others not. But what about the features of clusters we described in the previous paragraphs? Clusters covering larger areas are displays of long-range correlations within larger crustal volumes. As tectonic strain increases the level of stress, faults become more and more unstable, until a spontaneous rupture develops on the weakest interface. Static and dynamic stress variations trigger further events afterward within the crustal volume showing significant correlations with the hypocenter, that is, sensitivity to stress perturbations. The larger the region close to instability, the more seismic events can be triggered (Figure 3c) and with statistically higher magnitudes (Figure 3d). The latter relationship is due to the scaling laws connecting fault extent with earthquake size (Wells & Coppersmith, 1994). For the same reason, also larger mainshocks are possible, being the maximum possible magnitude positively correlated with the size of the unstable fault patch and crustal volume (see Figure 4c). Therefore, even a wider range of magnitudes is possible, which increases the amplitudes of fluctuations during the development of pre-mainshock clusters (Figure 6c). This effect also has an impact on magnitude series during clusters, that become less predictable (in agreement with Figures 9b and 9d). Because of the self-exciting nature of earthquakes, regions prone to larger seismic events are also likely to host clustered seismicity in time (Kagan & Jackson, 1991; Zaccagnino et al., 2023). Foreshocks are not "fore-shocks": they are not informative about the magnitude or time-to-failure of the eventually impending major events. Mainshocks tend to happen after clusters spread over larger areas (Figure 3a), with higher number of events and magnitudes not because such seismic activity ultimately triggers them. Instead, the reason of their occurrence is the size of the unstable crustal volume where ruptures propagate. Dynamic instabilities can be triggered by any kind of stress perturbations; otherwise, they can occur spontaneously (no "foreshocks"). They ultimately grow to become giant events by chance, that is, because of fine details of differential stress patterns and fault strength, regardless of previous seismic activity, if the extension of the prone-to-failure volumes is large enough. In this sense, fore-shocks do not exist, at least seismic clusters can hardly be really useful for improving short-term earthquake forecasting. Conversely, seismic hazard may benefit from the identification of regions featured by strongly correlated seismic activity, which is likely to be a mark of crustal developing instability.

4. Conclusions

Our work aims at understanding whether seismic activity preceding large earthquakes can be distinguished by swarms at least in regions equipped with highly developed seismic networks. To achieve our goal, we analyze seismicity reported in the Waveform Relocated Earthquake Catalog for Southern California from 1981 to 2022. We selected seismicity above the completeness magnitude $M_c = 2.5$ obtained combining a goodness of fit test for the power-law scaling of the frequency-size relationship and a stability condition for the b-value of the Gutenberg-Richter law. We grouped seismic events into clusters if their scaled spatial distance, R, and their distance in the time-space diagram from the triggering one, η, are respectively $R \leq 35$ and $\eta \leq 0.035$. Our procedure allows us to identify 5,486 clusters of events; among them, we only consider groups with at least five earthquakes before the mainshock. We investigate whether the relationships between maximum magnitude, involved area, duration, seismic rate, number of events, cumulative nucleated seismic moment, Shannon and Tsallis entropy, inter-event times distribution, global coefficient of variation of interevent times, trend and fluctuations of magnitudes and interevent times are different for foreshocks and swarms according to our definition and procedure. We found differences in the distributions of some features, although large variability is observed, while others cannot be distinguished. On the base of our results, precursory patterns of accelerated seismic activity can be hardly distinguished by more frequent swarms using the structural and statistical properties of clusters we considered in our study, which are likely the most significant ones. In light of this and of theoretical considerations, even though common in the language of seismologists, the term "foreshock" has little reason of being used. Each seismic event has chances to become a large earthquake if favorable local and boundary physical conditions occur, regardless the happening of "anomalous" preceding seismic activity. The observations that large seismic events are preceded by clustered seismic activity spread over larger areas and with higher magnitude variances and entropies than swarms, jointly with theory, suggest that without a reliable assessment of the large-scale stability conditions of brittle crustal volumes, the occurrence of seismicity, being it clustered or swarm-like, will be poorly informative about the probability of impending major quakes. Therefore, our conclusion is that foreshocks can have limited application to forewarn large seismic events in Southern California and that the probability of occurrence of large earthquakes is ultimately controlled by the finest details of the seismogenic source, memory effects and by long-range correlations within and between fault systems and crustal volumes.

Data Availability Statement

The Waveform Relocated Earthquake Catalog for Southern California (1981–2022) used in this research (Hauksson et al., 2012) is available at https://scedc.caltech.edu/data/alt-2011-dd-hauksson-yang-shearer.html.

Acknowledgments
Authors thank the editors Satoshi Ide and Qinghua Huang for handling the manuscript and two anonymous reviewers for valuable suggestions which greatly improved the article. This research was funded by the National Institute of Geophysics and Volcanology (INGV).

References

Agnew, D. C., & Jones, L. M. (1991). Prediction probabilities from foreshocks. *Journal of Geophysical Research*, *96*(B7), 11959–11971. https://doi.org/10.1029/91jb00191

Bayliss, K., Naylor, M., & Main, I. G. (2019). Probabilistic identification of earthquake clusters using rescaled nearest neighbour distance networks. *Geophysical Journal International*, *217*(1), 487–503. https://doi.org/10.1093/gji/ggz034

Belardinelli, M., Bizzarri, A., & Cocco, M. (2003). Earthquake triggering by static and dynamic stress changes. *Journal of Geophysical Research*, *108*(B3). https://doi.org/10.1029/2002jb001779

Beroza, G. C., Segou, M., & Mostafa Mousavi, S. (2021). Machine learning and earthquake forecasting—Next steps. *Nature Communications*, *12*(1), 4761. https://doi.org/10.1038/s41467-021-24952-6

Brodsky, E. E., & Lay, T. (2014). Recognizing foreshocks from the 1 April 2014 Chile earthquake. *Science*, *344*(6185), 700–702. https://doi.org/10.1126/science.1255202

Cattania, C., & Segall, P. (2021). Precursory slow slip and foreshocks on rough faults. *Journal of Geophysical Research: Solid Earth*, *126*(4), e2020JB020430. https://doi.org/10.1029/2020jb020430

Chiaraluce, L., Michele, M., Waldhauser, F., Tan, Y. J., Herrmann, M., Spallarossa, D., et al., (2022). A comprehensive suite of earthquake catalogues for the 2016–2017 central Italy seismic sequence. *Scientific Data*, *9*(1), 710. https://doi.org/10.1038/s41597-022-01827-z

Corral, Á. (2004). Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A: Statistical Mechanics and its Applications*, *340*(4), 590–597. https://doi.org/10.1016/j.physa.2004.05.010

Dodge, D. A., Beroza, G. C., & Ellsworth, W. (1996). Detailed observations of California foreshock sequences: Implications for the earthquake initiation process. *Journal of Geophysical Research*, *101*(B10), 22371–22392. https://doi.org/10.1029/96jb02269

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, *11*(7), 531–535. https://doi.org/10.1038/s41561-018-0145-1

Felzer, K. R., Abercrombie, R. E., & Ekstrom, G. (2004). A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America*, *94*(1), 88–98. https://doi.org/10.1785/0120030069

Gentili, S., Di Giovambattista, R., & Peresan, A. (2017). Seismic quiescence preceding the 2016 central Italy earthquakes. *Physics of the Earth and Planetary Interiors*, *272*, 27–33. https://doi.org/10.1016/j.pepi.2017.09.004

Hainzl, S., Brietzke, G. B., & Zöller, G. (2010). Quantitative earthquake forecasts resulting from static stress triggering. *Journal of Geophysical Research*, *115*(B11). https://doi.org/10.1029/2010jb007473

Hardebeck, J. L., Felzer, K. R., & Michael, A. J. (2008). Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. *Journal of Geophysical Research*, *113*(B8). https://doi.org/10.1029/2007jb005410

Hauksson, E., Yang, W., & Shearer, P. M. (2012). Waveform relocated earthquake catalog for southern California (1981 to June 2011). [Dataset]. Bulletin of the Seismological Society of America, *102*, 2239–2244. https://doi.org/10.1785/0120120010

Helmstetter, A., & Sornette, D. (2003). Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research*, *108*(B10). https://doi.org/10.1029/2003jb002409

Huang, Q. (2006). Search for reliable precursors: A case study of the seismic quiescence of the 2000 western Tottori prefecture earthquake. *Journal of Geophysical Research*, *111*(B4). https://doi.org/10.1029/2005jb003982

Kagan, Y. Y. (1991). Fractal dimension of brittle fracture. *Journal of Nonlinear Science*, *1*, 1–16. https://doi.org/10.1007/bf01209146

Kagan, Y. Y., & Jackson, D. D. (1991). Long-term earthquake clustering. *Geophysical Journal International*, *104*(1), 117–133. https://doi.org/10.1111/j.1365-246x.1991.tb02498.x

Kagan, Y. Y., & Knopoff, L. (1987). Statistical short-term earthquake prediction. *Science*, *236*(4808), 1563–1567. https://doi.org/10.1126/science.236.4808.1563

Kilb, D., Gomberg, J., & Bodin, P. (2000). Triggering of earthquake aftershocks by dynamic stresses. *Nature*, *408*(6812), 570–574. https://doi.org/10.1038/35046046

King, G. C., Stein, R. S., & Lin, J. (1994). Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America*, *84*(3), 935–953.

Main, I., & Meredith, P. (1989). Classification of earthquake precursors from a fracture mechanics model. *Tectonophysics*, *167*(2–4), 273–283. https://doi.org/10.1016/0040-1951(89)90078-4

Marone, C. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*(1), 643–696. https://doi.org/10.1146/annurev.earth.26.1.643

Marzocchi, W., Spassiani, I., Stallone, A., & Taroni, M. (2020). How to be fooled searching for significant variations of the b-value. *Geophysical Journal International*, *220*(3), 1845–1856. https://doi.org/10.1093/gji/ggz541

Marzocchi, W., & Zhuang, J. (2011). Statistics between mainshocks and foreshocks in Italy and southern California. *Geophysical Research Letters*, *38*(9). https://doi.org/10.1029/2011gl047165

Massey, F. J., Jr. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, *46*(253), 68–78. https://doi.org/10.1080/01621459.1951.10500769

McGuire, J. J., Boettcher, M. S., & Jordan, T. H. (2005). Foreshock sequences and short-term earthquake predictability on East Pacific rise transform faults. *Nature*, *434*(7032), 457–461. https://doi.org/10.1038/nature03377

McLaskey, G. C. (2019). Earthquake initiation from laboratory observations and implications for foreshocks. *Journal of Geophysical Research: Solid Earth*, *124*(12), 12882–12904. https://doi.org/10.1029/2019jb018363

Mignan, A. (2014). The debate on the prognostic value of earthquake foreshocks: A meta-analysis. *Scientific Reports*, *4*(1), 4099. https://doi.org/10.1038/srep04099

Nielsen, S., Spagnuolo, E., Smith, S., Violay, M., Di Toro, G., & Bistacchi, A. (2016). Scaling in natural and laboratory earthquakes. *Geophysical Research Letters*, *43*(4), 1504–1510. https://doi.org/10.1002/2015gl067490

Ogata, Y., & Katsura, K. (2012). Prospective foreshock forecast experiment during the last 17 years. *Geophysical Journal International*, *191*(3), 1237–1244. https://doi.org/10.1111/j.1365-246x.2012.05645.x

Ogata, Y., Utsu, T., & Katsura, K. (1996). Statistical discrimination of foreshocks from other earthquake clusters. *Geophysical Journal International*, *127*(1), 17–30. https://doi.org/10.1111/j.1365-246x.1996.tb01531.x

Pacheco, J. F., Scholz, C. H., & Sykes, L. R. (1992). Changes in frequency–size relationship from small to large earthquakes. *Nature*, *355*(6355), 71–73. https://doi.org/10.1038/355071a0

Papadopoulos, G., Charalampakis, M., Fokaefs, A., & Minadakis, G. (2010). Strong foreshock signal preceding the L'Aquila (Italy) earthquake (Mw 6.3) of 6 April 2009. *Natural Hazards and Earth System Sciences*, *10*(1), 19–24. https://doi.org/10.5194/nhess-10-19-2010

Petrillo, G., & Lippiello, E. (2021). Testing of the foreshock hypothesis within an epidemic like description of seismicity. *Geophysical Journal International*, *225*(2), 1236–1257. https://doi.org/10.1093/gji/ggaa611

Picozzi, M., Iaccarino, A. G., Spallarossa, D., & Bindi, D. (2023). On catching the preparatory phase of damaging earthquakes: An example from central Italy. *Scientific Reports*, *13*(14403), 14403. https://doi.org/10.1038/s41598-023-41625-0

Reasenberg, P. A. (1999). Foreshock occurrence before large earthquakes. *Journal of Geophysical Research*, *104*(B3), 4755–4768. https://doi.org/10.1029/1998jb900089

Rhoades, D., & Evison, F. (1993). Long-range earthquake forecasting based on a single predictor with clustering. *Geophysical Journal International*, *113*(2), 371–381. https://doi.org/10.1111/j.1365-246x.1993.tb00893.x

Rice, J. R., Lapusta, N., & Ranjith, K. (2001). Rate and state dependent friction and the stability of sliding between elastically deformable solids. *Journal of the Mechanics and Physics of Solids*, *49*(9), 1865–1898. https://doi.org/10.1016/s0022-5096(01)00042-4

Sahimi, M., Robertson, M. C., & Sammis, C. G. (1993). Fractal distribution of earthquake hypocenters and its relation to fault patterns and percolation. *Physical Review Letters*, *70*(14), 2186–2189. https://doi.org/10.1103/physrevlett.70.2186

Saleur, H., Sammis, C. G., & Sornette, D. (1996). Renormalization group theory of earthquakes. *Nonlinear Processes in Geophysics*, *3*(2), 102–109. https://doi.org/10.5194/npg-3-102-1996

Savage, M. K., & Rupp, S. H. (2000). Foreshock probabilities in New Zealand. *New Zealand Journal of Geology and Geophysics*, *43*(3), 461–469. https://doi.org/10.1080/00288306.2000.9514902

Seif, S., Zechar, J. D., Mignan, A., Nandan, S., & Wiemer, S. (2019). Foreshocks and their potential deviation from general seismicity. *Bulletin of the Seismological Society of America*, *109*(1), 1–18. https://doi.org/10.1785/0120170188

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sornette, D. (2006). *Critical phenomena in natural sciences: Chaos, fractals, selforganization and disorder: Concepts and tools*. Springer Science and Business Media.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. ArXiv preprint ArXiv:0907.4290.

Sotolongo-Costa, O., & Posadas, A. (2004). Fragment-asperity interaction model for earthquakes. *Physical Review Letters*, *92*(4), 048501. https://doi.org/10.1103/physrevlett.92.048501

Telesca, L. (2011). Tsallis-based nonextensive analysis of the southern California seismicity. *Entropy*, *13*(7), 1267–1280. https://doi.org/10.3390/e13071267

Telesca, L., & Lovallo, M. (2017). On the performance of Fisher Information Measure and Shannon entropy estimators. *Physica A: Statistical Mechanics and its Applications*, *484*, 569–576. https://doi.org/10.1016/j.physa.2017.04.184

Telesca, L., Lovallo, M., Lopez, C., & Molist, J. M. (2016). Multiparametric statistical investigation of seismicity occurred at El Hierro (Canary Islands) from 2011 to 2014. *Tectonophysics*, *672*, 121–128. https://doi.org/10.1016/j.tecto.2016.01.045

Trugman, D. T., & Ross, Z. E. (2019). Pervasive foreshock activity across southern California. *Geophysical Research Letters*, *46*(15), 8772–8781. https://doi.org/10.1029/2019gl083725

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, *52*(1–2), 479–487. https://doi.org/10.1007/bf01016429

Vallianatos, F., Michas, G., & Papadakis, G. (2018). Nonextensive statistical seismology: An overview. *Complexity of seismic time series*, 25–59.

van den Ende, M. P., & Ampuero, J.-P. (2020). On the statistical significance of foreshock sequences in southern California. *Geophysical Research Letters*, *47*(3), e2019GL086224. https://doi.org/10.1029/2019gl086224

Velasco, A. A., Hernandez, S., Parsons, T., & Pankow, K. (2008). Global ubiquity of dynamic earthquake triggering. *Nature Geoscience*, *1*(6), 375–379. https://doi.org/10.1038/ngeo204

Wells, D. L., & Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, *84*(4), 974–1002. https://doi.org/10.1785/bssa0840040974

Wiemer, S., & Wyss, M. (1994). Seismic quiescence before the landers (M=7.5) and big bear (M=6.5) 1992 earthquakes. *Bulletin of the Seismological Society of America*, *84*(3), 900–916.

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859–869. https://doi.org/10.1785/0119990114

Wu, Y.-M., & Chiao, L.-Y. (2006). Seismic quiescence before the 1999 Chi-Chi, Taiwan, Mw 7.6 earthquake. *Bulletin of the Seismological Society of America*, *96*(1), 321–327. https://doi.org/10.1785/0120050069

Yagi, Y., Okuwaki, R., Enescu, B., Hirano, S., Yamagami, Y., Endo, S., & Komoro, T. (2014). Rupture process of the 2014 Iquique Chile earthquake in relation with the foreshock activity. *Geophysical Research Letters*, *41*(12), 4201–4206. https://doi.org/10.1002/2014gl060274

Zaccagnino, D., Telesca, L., & Doglioni, C. (2022). Scaling properties of seismicity and faulting. *Earth and Planetary Science Letters*, *584*, 117511. https://doi.org/10.1016/j.epsl.2022.117511

Zaccagnino, D., Telesca, L., & Doglioni, C. (2023). Global versus local clustering of seismicity: Implications with earthquake prediction. *Chaos, Solitons and Fractals*, *170*, 113419. https://doi.org/10.1016/j.chaos.2023.113419

Zaliapin, I., & Ben-Zion, Y. (2013). Earthquake clusters in southern California I: Identification and stability. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2847–2864. https://doi.org/10.1002/jgrb.50179

Zaliapin, I., & Ben-Zion, Y. (2015). Artefacts of earthquake location errors and short-term incompleteness on seismicity clusters in southern California. *Geophysical Journal International*, *202*(3), 1949–1968. https://doi.org/10.1093/gji/ggv259

Zhu, G., Yang, H., Tan, Y. J., Jin, M., Li, X., & Yang, W. (2022). The cascading foreshock sequence of the Ms 6.4 Yangbi earthquake in Yunnan, China. *Earth and Planetary Science Letters*, *591*, 117594. https://doi.org/10.1016/j.epsl.2022.117594

Zhuang, J. (2021). Explaining foreshock and the Båth law using a generic earthquake clustering model. *Statistical Methods and Modeling of Seismogenesis*, *105*, 105–130. https://doi.org/10.1002/9781119825050.ch4