Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL109865

Magnitude Clustering During Stick-Slip Dynamics on Laboratory Faults

Omid Khajehdehi[1], Thomas H. W. Goebel[2], Georg Dresen[3], and Jörn Davidsen[1,4]

[1]Complexity Science Group, Department of Physics and Astronomy, University of Calgary, Calgary, AB, Canada, [2]Center for Earthquake Research and Information, University of Memphis, Memphis, TN, USA, [3]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Section 4.2: Geomechanics and Scientific Drilling, Potsdam, Germany, [4]Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada

Key Points:

- Magnitude clustering of acoustic emissions (AEs) is strongest during major slip events in fluid-saturated samples of Westerly granite
- Variations in the frequency-magnitude distribution of AEs during slip events explain the strong magnitude clustering
- These variations indicate a prevalence of larger AEs right after the onset of slip events

Supporting Information:

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

Correspondence to:

J. Davidsen,
davidsen@phas.ucalgary.ca

Citation:

Khajehdehi, O., Goebel, T. H. W., Dresen, G., & Davidsen, J. (2024). Magnitude clustering during stick-slip dynamics on laboratory faults. *Geophysical Research Letters*, *51*, e2024GL109865. https://doi.org/10.1029/2024GL109865

Received 17 APR 2024
Accepted 7 OCT 2024

Author Contributions:

Conceptualization: Omid Khajehdehi, Thomas H. W. Goebel, Jörn Davidsen
Data curation: Thomas H. W. Goebel
Formal analysis: Omid Khajehdehi
Funding acquisition: Jörn Davidsen
Investigation: Omid Khajehdehi, Thomas H. W. Goebel
Methodology: Omid Khajehdehi, Jörn Davidsen
Project administration: Jörn Davidsen
Resources: Georg Dresen, Jörn Davidsen
Software: Omid Khajehdehi
Supervision: Jörn Davidsen
Validation: Omid Khajehdehi, Thomas H. W. Goebel, Georg Dresen, Jörn Davidsen
Visualization: Omid Khajehdehi

Abstract We present an analysis of magnitude clustering of microfractures inferred from acoustic emissions (AEs) during stick-slip (SS) dynamics of faulted Westerly granite samples in frictional sliding experiments, with and without fluids, under triaxial loading with constant displacement rate. We investigate magnitude clustering in time across periods during, preceding and after macroscopic slip events on laboratory faults. Our findings reveal that magnitude clustering exists such that subsequent AEs tend to have more similar magnitudes than expected. Yet, this clustering only exists during macroscopic slip events and is strongest during major slip events in fluid-saturated and dry samples. We demonstrate that robust magnitude clustering arises from variations in frequency-magnitude distributions of AE events during macroscopic slip events. These temporal variations indicate a prevalence of larger AE events right after (0.3–3 s) the SS onset. Hence, magnitude clustering is a consequence of non-stationarities.

Plain Language Summary Can we determine the size of a future earthquake based on the size of past earthquakes? This fundamental question has been controversially debated over the years, without an agreed-upon answer. Here, we tackle this question under controlled conditions in a lab setting by studying the frictional stick-slip dynamic of rough granite faults, which gives rise to mm-scale seismic events. We find that the sizes of these seismic events are not independent during periods containing a macroscopic slip but instead are clustered such that larger seismic events tend to directly follow other large seismic events. We show that this can be explained by temporal changes in the frequency of occurrence of seismic events associated with the sliding motion of the fault. Our findings link the properties of mm-scale seismic events with macroscopic slip on lab faults.

1. Introduction

Spatiotemporal clustering is a widely acknowledged characteristic of tectonic seismicity (Davidsen & Baiesi, 2016; Davidsen & Paczuski, 2005; Moradpour et al., 2014; Sornette & Werner, 2005; Turcotte, 1997; Zaliapin et al., 2008), fluid-induced seismicity in the field (Karimi & Davidsen, 2023; Kothari et al., 2020; Maghsoudi et al., 2018; Schoenball & Ellsworth, 2017; Shapiro, 2015), and seismicity in laboratory experiments (Baró et al., 2013; Davidsen et al., 2017; Goebel et al., 2023; Kwiatek et al., 2024; Mäkinen et al., 2015; Ribeiro et al., 2015; Scholz, 2002). Well-known instances across all three cases include the temporal decay of aftershocks captured by the Omori-Utsu relationship (Utsu et al., 1995; Wetzler et al., 2016) and spatial clustering of aftershocks (Kagan, 2013; Karimi & Davidsen, 2021; van der Elst & Shaw, 2015).

Earthquake magnitudes are considered clustered if the magnitude of an earthquake is related to the magnitudes of prior events. However, the question of whether earthquake magnitudes exhibit clustering or correlations in time remains a topic of ongoing debate (Petrillo & Zhuang, 2022). The presence of magnitude clustering, if confirmed, would suggest the possibility of forecasting the magnitude of forthcoming earthquakes by considering the magnitudes of past earthquakes (Davidsen et al., 2012; Helmstetter et al., 2006; Nandan et al., 2022). Beyond the Gutenberg-Richter relation describing the frequency-magnitude distribution (FMD) (Gutenberg & Richter, 1944), magnitudes were traditionally considered to be statistically independent and identically distributed. Some studies have reported clustering between the magnitudes of consecutive tectonic earthquakes (Lippiello et al., 2007, 2008; Spassiani & Sebastiani, 2016; Xiong et al., 2023), while others have reached opposite conclusions (Davidsen & Green, 2011; Kwiatek et al., 2022; Petrillo & Zhuang, 2023). Nandan et al. (2019, 2022) argued that distinct FMDs exist for mainshocks and aftershocks resulting in magnitude clustering. However, the

Writing – original draft: Omid Khajehdehi
Writing – review & editing: Omid Khajehdehi, Thomas H. W. Goebel, Georg Dresen, Jörn Davidsen

argument has been put forth that the observed magnitude clustering and correlations could arise from catalog incompleteness, specifically short-term aftershock incompleteness (STAI), raising doubts about the significance of the observed clusterings (Davidsen & Green, 2011; Petrillo & Zhuang, 2023). Notably, sequences of microseismicity, nanoseismicity, and picoseismicity (corresponding to earthquakes with moment magnitudes Mw in the range [−4, 0]) recorded at the Mponeng mine in South Africa, which were not affected by STAI, have shown no evidence of magnitude clustering (Davidsen et al., 2012). Conversely, microseismicity associated with hydraulic fracturing in Canada demonstrated strong magnitude clustering, potentially due to the specific geological setting (Maghsoudi et al., 2016). In controlled laboratory settings, a recent study investigated magnitude clustering of mm-scale microfractures as measured by acoustic emissions (AE) during rock fracture under uniaxial loading (Xiong et al., 2023). This study observed evidence of magnitude clustering in shear rock fracture tests on two marble samples, with one sample featuring a single angled flaw and the other containing two angled flaws. In cases where rock fractures were induced primarily by tensile stress, such as tensile bending and hydraulic fracturing, no evidence of magnitude clustering was found. This led Xiong et al. (2023) to propose that the presence of shear stress induced by at least one angled flaw relative to the loading axis is necessary to observe magnitude clustering in rock fracture.

Here, we test the hypothesis that stick-slip (SS) dynamics along rough faults under triaxial loading causes magnitude clustering of AEs in the lab under conditions that more closely mimic those encountered in natural earthquakes (Dresen et al., 2020; Goebel et al., 2023). Specifically, we focus on laboratory faults in Westerly granite under both dry and fluid-saturated conditions. In triaxial compression tests, we can identify *macroscopic slip events* as periods of combined slip and stress relaxation that include abrupt, dynamic failure and gradual stress relaxation over several seconds, where the entire fault at the macro-scale was affected by a measurable stress drop rather than a smaller confined patch. The dynamics at the macro-scale are accompanied by mm-scale microfractures and we present the first comprehensive analysis of magnitude clustering of these AE events. Our findings reveal significant magnitude clustering of AE events during macroscopic slip events, while its presence is notably absent during preparatory stages before slip initiation as well as during post slip stages. We provide compelling evidence that robust magnitude clustering is a consequence of temporal variations in the FMD of AE events during macroscopic slip events along the fault.

2. Materials and Method

2.1. Samples and Experiments

Triaxial compression experiments were carried out on nine cylindrical Westerly granite samples, with diameters of 40 or 50 mm and a height of 105 mm. These experiments were conducted under confining pressures (P_c = 150 MPa). Each sample was enclosed within an elastic rubber jacket, ensuring separation from the confining oil, and subjected to a constant loading at a displacement rate of 0.3 μ m/sec, reaching a maximum vertical displacement of 6 mm. More information about the experiment is in Table S1 in Supporting Information S1. Axial force and strain were monitored using an external load cell and two vertical and horizontal strain gauges, all operating at a sampling rate of 10 Hz. To create natural fault roughness, we initially fractured the originally intact rocks under a confining pressure of P_c = 75 MPa. Subsequently, the samples were reloaded at P_c = 150 MPa with and without pore fluid, which led to SS behavior. Subjecting rough and heterogeneous laboratory faults to triaxial loading results in various slip behaviors, as illustrated in Figures 1a–1c.

2.2. Detection and Characterization of Acoustic Emission Events

We monitor active and passive seismic sources at the mm-scale using a 16-channel, high-speed data acquisition system (Stanchits et al., 2006). Accurate AE locations were possible due to high-quality automated picks, high sampling rates (10 MHz) and time-dependent, anisotropic, layered velocity models generated from ultrasonic pulses emitted every 30 s. We searched for AE events within 100 μs time windows (resulting in peak detection rates of 10 ev/ms), and kept events with high signal-to-noise ratio, at least 8 station picks and travel time residuals of less than 0.5 μs, thereby minimizing the likelihood of erroneous detection and locations. AE location uncertainty was between 0.5 and 3 mm. In analogy with local magnitudes of earthquakes (Bakun & Joyner, 1984), we computed local magnitudes based on peak AE amplitudes averaged across the entire array, corrected for source-receiver distances (Kwiatek et al., 2014; Zang et al., 1998), as follows:

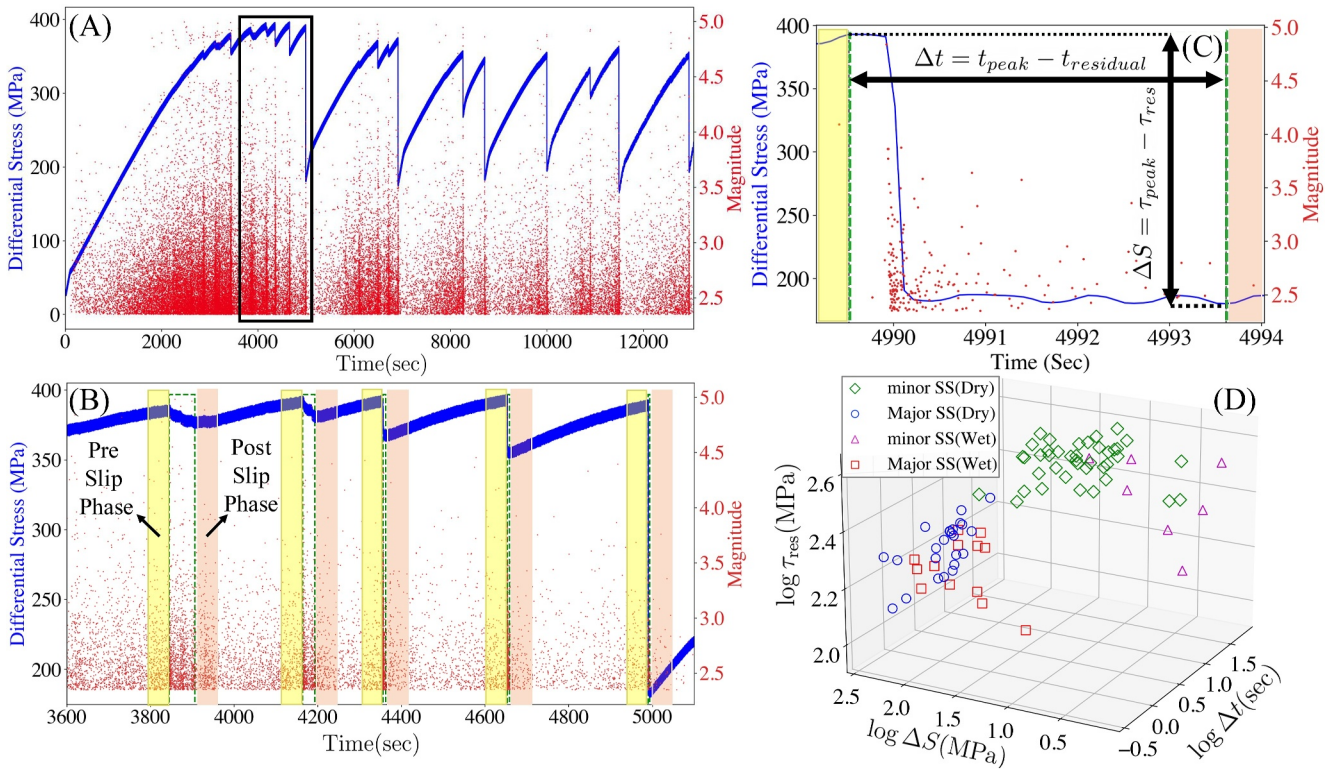

Figure 1. (a–c): Example of experimental loading of a rough fault and acoustic emission (AE) response for a dry sample (see Figure S1 in Supporting Information S1 for a wet sample). Blue curves represent the differential stress, and the red dots correspond to relative local magnitudes of AE events (see text) with $m > m_c = 2.35$. Six major stick-slip (SS) events and multiple minor SS events are visible in panel (a). (b) Is a zoom-in of the experiment from 3600 − 5100 sec in panel (a). It contains multiple SS events highlighted by green dashed rectangles. For each SS event, we show the associated preparatory phase (yellow shaded area) and post-slip phase (pink shaded area), defined here as the $t_p = 50$ sec interval prior to or after the SS event, respectively. (c) Is a zoom-in of the major SS event in panel (b), which starts at the local maximum differential stress and ends when the differential stress consistently increases again, containing the abrupt, dynamic failure, and gradual stress relaxation over several seconds. Both the duration of the SS event (Δt) and the stress drop (ΔS) are highlighted. (d) Shows the relationship between Δt, ΔS, and τ_{res} of all minor and major SS events in our dry and wet samples.

$$m = \log_{10}\sqrt{\frac{1}{n}\sum_{j=1}^{n}(A_j D_j)^2}, \tag{1}$$

where A_j represents the maximum amplitude of the first P-wave pulse detected at sensor j, corrected for the distance D_j between the hypocenter of the AE event and the sensor. AE events occur at the mm-scale, with double-couple (shear type) or isotropic (tensile, pore space collapse) source types (Goebel et al., 2023). The local magnitude scale provides good estimates of relative event sizes for the vast majority of the recorded AE events but likely underestimates the largest event magnitudes relative to the magnitudes of the rest of the catalog, that is, AE events directly associated with dynamic slip (Goebel et al., 2013; Kwiatek et al., 2024). We define the magnitude of completeness, denoted as m_c, as the lowest magnitude threshold at which the estimate of the b-value in the FMD stabilizes (Davidsen et al., 2015), and for all Westerly granite samples, $m_c = 2.35$. Note that directly comparing magnitude scales and, hence, b-values obtained in laboratory settings to those measured on tectonic scales is not possible without the appropriate calibration of AE sensors (Goodfellow et al., 2015; McLaskey & Glaser, 2012).

2.3. Classification of Macroscopic Slip Events

In our experiments, we define macroscopic slip events as a combination of SS and stress relaxation processes. These processes include abrupt, dynamic failure, and gradual stress relaxation over several seconds, leading to a measurable stress drop. The onset of a macroscopic slip event corresponds to the time of the local maximum

differential stress before its sharp decrease and the event ends once the differential stress consistently increases again as illustrated in Figure 1. Each macroscopic slip event can, thus, be characterized by its duration, Δt, its residual stress, τ_{res}, and its stress drop, ΔS. The relationship between ΔS, Δt, and τ_{res} across all macroscopic slip events allows us to divide macroscopic slip events into major and minor SS events for dry and wet samples separately as follows from Figure 1d (see also Figure S2 in Supporting Information S1). Among wet samples, major SS events exhibit relatively more significant stress drops and shorter durations compared to minor SS events. Among the dry samples, minor SS events exhibit larger residual stress and lower stress drops compared to major SS events. On average, the stress drop during SS events (major or minor) in the dry samples is larger than in the wet samples. Nevertheless stress drop alone allows an almost identical separation between major and minor SS events across all wet and dry samples (see Figure S2 in Supporting Information S1).

We also analyze the preparatory phase preceding each macroscopic slip event as illustrated in Figure 1. In principle, this phase could extend back to the previous slip event and it is variable between slip events. To ensure consistency and considering that the minimum time between the initiation of two slip events among all samples is about 70 s, we only consider fixed time intervals of length $t_p \leq 70$ s directly preceding each (minor or major) slip event as the duration of the preparatory phase. Analogously, we can define a post-slip phase (see Figure 1). While we show our findings for both phases for $t_p = 50$ s in the following, these are robust for $30 \leq t_p \leq 70$ s.

After classification of the macroscopic slip events, we end up with 12 AE catalogs for our analysis: Four during macroscopic slip events (major SS in dry samples, major SS in wet samples, minor SS in dry samples, and minor SS in wet samples), and four each for the associated preparatory and post-slip phases (see Table S2 in Supporting Information S1 for a summary).

2.4. Magnitude Clustering Analysis

To determine the presence of magnitude clustering, we focus on events with magnitudes larger than the magnitude of completeness ($m > m_c$), and we compute the differences between the magnitude of successive AEs: $\Delta m_i = m_{i+1} - m_i$ (Davidsen & Green, 2011; Davidsen et al., 2012; Lippiello et al., 2008; Xiong et al., 2023). Then, we shuffle the magnitudes of the initial catalog and compute the differences between the magnitudes of the shuffled and initial catalog denoted by $\Delta m_i^* = m_{i^*} - m_i$, where i^* represents the index of a randomly chosen AE in the catalog. By doing this, the magnitude of the following earthquake is independent of its predecessors in Δm_i^*. Since our AE catalogs aggregate AE events from all macroscopic slip events of a given type, our shuffling process exclusively involves the AE events within each individual slip event to avoid mixing of AEs between slip events. That is when shuffling an AE catalog, we only shuffle the events within each AE subcatalog associated with an individual slip event and refrain from shuffling across the entire catalog.

If magnitudes are clustered, there should be substantial statistical distinctions between the distribution of Δm and Δm^*. To measure magnitude clustering, we calculate the differences in their (cumulative) distributions $\delta P(m_0) = P(\Delta m < m_0) - P(\Delta m^* < m_0)$ (see Figure S3 in Supporting Information S1 for an illustration), where $P(\Delta m < m_0)$ is the probability that successive magnitude difference (Δm) will have a value less than m_0. We use $10^3 - 10^4$ independent realizations of the shuffled catalogs and calculate $P(\Delta m^* < m_0)$ across an ensemble of independent random realizations of shuffled catalogs, resulting in the distribution $\rho[P(\Delta m^* < m_0)]$ with a mean denoted as $\mu^*(m_0)$. In the absence of clustering, $\delta P(m_0)$ should be statistically indistinguishable from zero for all m_0. To assess the significance of the deviation from zero, we calculate the percentile of the observed difference $\delta P(m_0) = P(\Delta m < m_0) - \mu^*(m_0)$ for each m_0. If the calculated percentile is less than 5% or exceeds 95%, magnitude clustering for that specific m_0 is deemed significant.

2.5. Empirical Cumulative Density Functions

To determine whether any observed magnitude clustering is homogeneous across magnitude ranges, we utilize the Empirical Cumulative Density Functions (ECDF) method (Xiong et al., 2023). This method involves arranging the magnitudes in ascending order and assigning an ECDF value between 0 and 1 that equals the count divided by the total number. For the original catalog, the ECDF value of each event (i) is compared to the ECDF value of the following event ($i + 1$). The results are grouped into 0.2×0.2 bins of the two ECDF values representing the count of events falling into each bin. To identify variations beyond what is expected based on the magnitude distribution, ECDF values are also computed for a time-randomized catalog, and the mean count

across the bins in the randomized version is determined. The percentage difference of values in each bin relative to the mean derived from the randomized catalog is then calculated and is depicted in a color map. The larger the percentage difference of values in a bin, the stronger the clustering of magnitude in that magnitude range, and if the percentage difference of values in a bin is very close to zero, it signifies a lack of clustering of magnitudes within that range.

3. Results

3.1. Magnitude Clustering

Drawing upon the methodologies outlined in Sections 2.4 and 2.5, our analysis of the eight AE catalogs described in Section 2.3 reveal substantial and robust magnitude clustering during major slip events for both dry and wet samples and minor slip events for dry samples, indicating that larger AE events tend to be followed by AE events of similar magnitude. Specifically, Figure 2a illustrates pronounced clustering of magnitudes of AE events during slip events across a broad range of m_0, indicating that subsequent AE events during slip events tend to resemble the magnitudes of preceding events. For instance, during major slip events in wet samples, which show the strongest deviation from zero in $\delta P(m_0)$, there is an approximately 8% higher likelihood that the magnitude of a subsequent event falls within an interval of ± 0.5 of the preceding event. Figure 2b shows that the deviations of $\delta P(m_0)$ from zero are significant, especially during major slip events in both wet and dry samples, since the percentile of the observed difference $\delta P(m_0)$ either falls below 5% or surpasses 95% for almost all m_0. Notably, among the AE catalogs during slip event, only the one for minor slip events in wet samples lacks significant magnitude clustering across a wider range of m_0. This could be related to the more limited statistics in this case (see Table S2 in Supporting Information S1). All these findings are robust. Neither variations in the magnitude of completeness (Figure S5 in Supporting Information S1) nor the aforementioned detector limitations for high AE event rates during slip events (Figure S4 in Supporting Information S1) (both of which could lead to STAI) do significantly impact the observed magnitude clustering and the associated significance levels.

The observed magnitude clustering can be further characterized by the ECDF analysis as illustrated in Figure 2c for major slip events in dry samples (see Figure S6 in Supporting Information S1 for other catalogs). The color-coded map of ECDF$[m_{i+1}]$ versus ECDF$[m_i]$ shows, for example, that the top right corner bin in Figure 2c has a much higher rate than shuffled catalogs, surpassing 50%. This increase in rate is significant at a very high level as follows from the percentiles given in Figure 2c. This indicates that magnitude clustering is particularly strong between the largest AE events such that large magnitudes are the most highly clustered. This observation likewise holds for major slip events in wet samples and minor slip events in dry samples, while it is less pronounced for minor slip events in wet samples (Figure S6 in Supporting Information S1). Moreover, six out of nine bins in Figure 2c involving the largest AE events (top row and right column) indicate very significant magnitude clustering. Overall, these findings explain in particular why variations in m_c do not notably affect magnitude clustering for major slip events (Figure S5 in Supporting Information S1).

In contrast to the slip events, we do not observe significant magnitude clustering during any of the preparatory or post-slip phases. Figures 2d and 2g show minimal deviations from zero in $\delta P(m_0)$ across almost all m_0, indicating at best weak magnitude clustering in these cases. More importantly, Figures 2e and 2h show that the vast majority of deviations in $\delta P(m_0)$ are not statistically significant. Similarly, the analysis of ECDF maps for the preparatory and post-slip phases reveal no clear evidence of magnitude clustering. For example, Figure 2f shows that the deviations from randomized catalogs for the preparatory phase of major slip events in dry samples never exceed 5%, highlighting the absence of pronounced magnitude clustering. The same conclusion holds for the post-slip phase (Figure 2i) and all other catalogs (Figure S6 in Supporting Information S1).

3.2. Frequency-Magnitude Distribution

A potential explanation for the observed robust magnitude clustering during slip events are temporal variations, or non-stationarities, in the corresponding FMDs. To test this hypothesis, we examine the non-stationarities by separating the AE events associated with a given SS event into five groups of equal numbers according to their time of occurrence such that the first quintile corresponds to the first 20% of the AE events after slip onset at peak stress, etc. Figure 3 shows the corresponding FMDs when stacked over all SS events in a given class. Indeed, in all cases, we observe a much higher frequency of large AE events during the beginning of slip events compared to later times during slip events. As we progress through subsequent quintiles, the number of larger magnitude

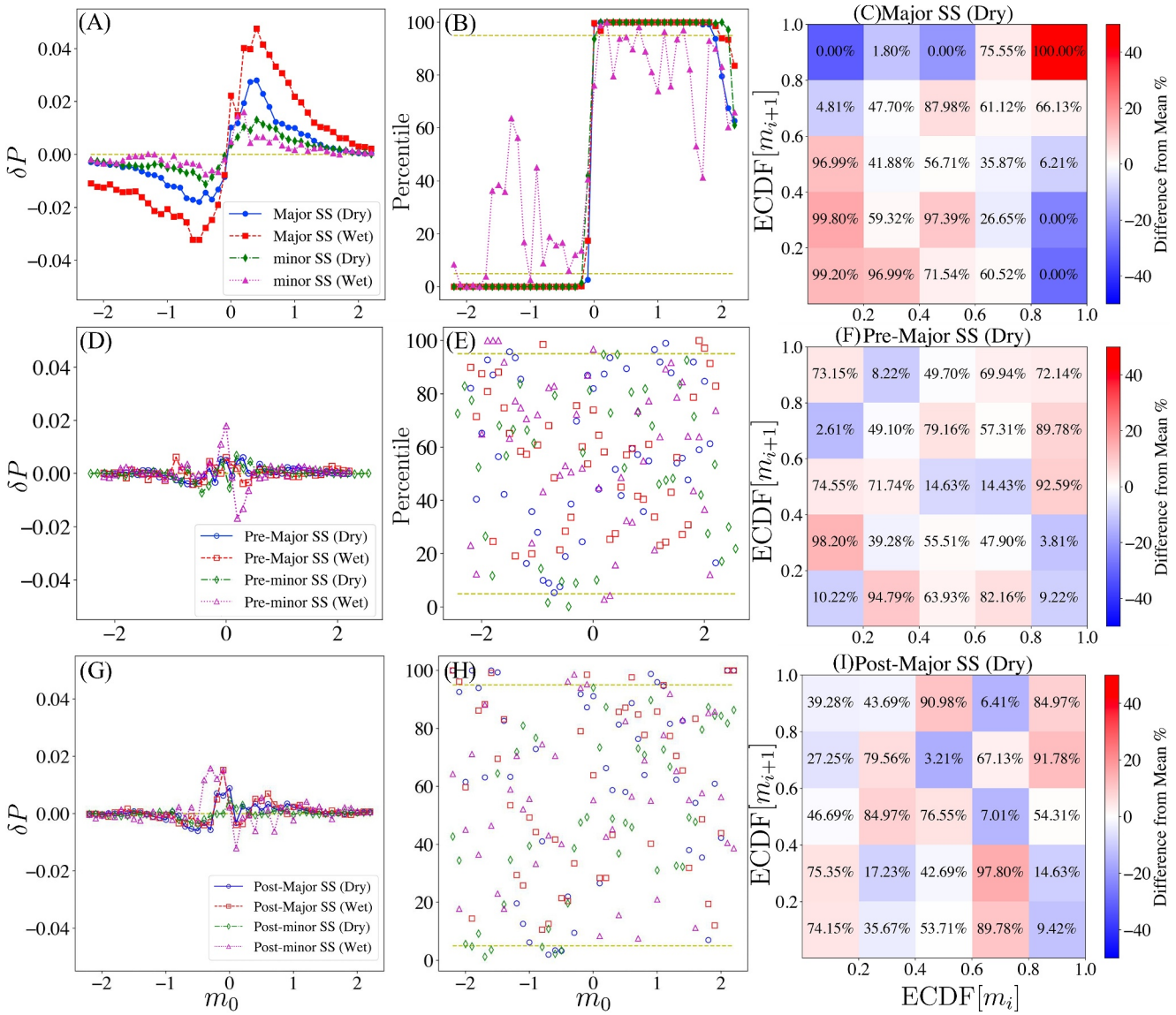

Figure 2. Investigation of magnitude clustering during stick-slip (SS) events (a–c), prior to SS events (d–f), and post SS events (g–i). (a, d, and g): Differences in the probability of observing an acoustic emission (AE) magnitude difference $\Delta m < m_0$, between the original catalogs and randomized versions; see text for details. Magnitude clustering corresponds to significant deviations from zero. To determine if a given deviation from zero is significant, (b, e, and h) show the percentile of $\delta P(m_0)$ for each m_0, where the yellow dashed lines represent the 5% and 95% percentiles. Major SS events (wet and dry) and minor SS events in dry samples exhibit significant magnitude clustering of AE events for almost all m_0. (c, f, and i): Magnitude clustering across different magnitude ranges in dry samples during major SS events (c), the associated preparatory phases (f), and post-slip phases (i) using the Empirical Cumulative Density Functions method; see text for details. The color scale indicates the percentage difference of values within each bin relative to the mean derived from randomized catalogs. The number in each color bin indicates the significance of the deviation from the mean by giving the associated percentile obtained from the randomized catalogs—percentiles below 5% or above 95% can be considered significant. The strongest magnitude clustering in terms of relative strength and significance for major SS events occurs between the largest AE events ((c), upper right corner).

events decreases, and the number of smaller magnitudes increases, representing a clear non-stationarity. To quantify this temporal variation, we utilize the two-sided Kolmogorov-Smirnov (KS) test (Hodges, 1958) to compare the FMDs associated with the first and last quintiles of a given AE catalog. In all cases, the KS test rejects the null hypothesis that the two distributions are identical at a high confidence level ($p < 10^{-6}$, see Table S3 in Supporting Information S1). This contrasts the preparatory phases (Figure S7 in Supporting Information S1) for which we cannot reject the null hypothesis ($p \geq 0.1$, see Table S3 in Supporting Information S1) as well as the post-slip phases (Figure S8 in Supporting Information S1).

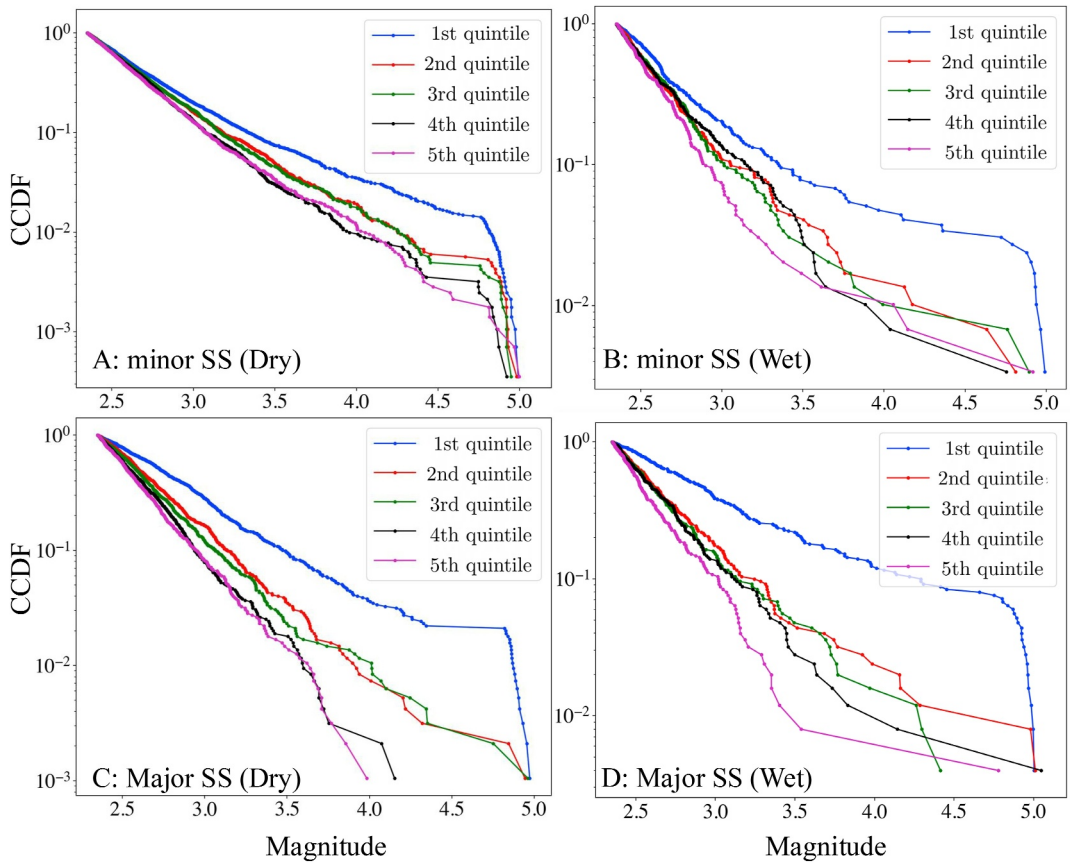

Figure 3. Complementary cumulative distribution function of magnitudes over 5 different time intervals during stick-slip (SS) events: Minor SS events in (a) dry samples and (b) wet samples. Major SS events in (c) dry samples and (d) wet samples. In all cases, the number of large magnitudes is significantly increased at early times (1st quintile) compared to later times. Note that the cut-off for the largest magnitudes is a reflection of the instrumental limitation.

4. Discussion

Our findings indicate that magnitude clustering is observed during slip events, driven by highly non-stationary behavior in the FMDs of AEs. Conversely, during the preparatory and post-slip phase, there is no significant magnitude clustering, and the FMDs remain stationary over time scales of the order of 50 s before the slip event. This is consistent with previous observations that variations in FMD can occur over longer time scales when considering the entire period before a slip event, extending back to the previous slip (Goebel et al., 2013; Kwiatek et al., 2024), which also occurs for our experiments (Figures S9 and S10 in Supporting Information S1). The observed strong non-stationarities in the FMD trivially lead to magnitude clustering. When we shuffle the magnitudes of AEs associated with each slip event, we blend AEs originating from different FMDs. This means in our case that the rate of occurrence of large magnitudes is homogenized under shuffling. The excess of large magnitudes at early times during slip events in our original catalogs is hence interpreted as magnitude clustering of predominantly large magnitudes. In particular, the larger the variation in the FMD the stronger the magnitude clustering. Indeed, Figure 3d shows that the most significant variation in FMD occurs for major slip events of wet samples (as confirmed by the KS distance, see Table S3 in Supporting Information S1), explaining why the strongest magnitude clustering is observed in this case, see Figures 2a and 2b. In all cases, magnitude clustering is strongest at early times (0.3–3 s) during a slip event, but it often remains significant even at later times (see Figures S11 and S12 in Supporting Information S1 for details). When taking the spatial distances between subsequent AE events into account, AE event pairs separated by larger distances during slip events (but still primarily confined to the fault zone) typically show more pronounced magnitude clustering, especially for minor slip events (Figure S13 in Supporting Information S1). This indicates that magnitude clustering is not predominantly driven by short-ranged interactions between AE events, but rather that the macroscopic slip directly

induces large AE events across the whole fault zone as the rupture front passes through the fault. Similarly, during preparatory and post-slip phases, conditioning on distances larger or shorter than the median distance between subsequent AE events does result in weak but significant magnitude clustering (Figure S14 in Supporting Information S1): Whereas the magnitudes tend to be more similar for AE event pairs separated by larger distances, magnitudes tend to be more dissimilar for AE event pairs separated by shorter distances. The latter might be related to the presence of event-event triggering discussed below.

In another controlled laboratory setting, a recent study investigated magnitude clustering in rock fracture under uniaxial loading, identifying magnitude clustering under shear stress (Xiong et al., 2023). Analyzing their data sets, we confirmed strong magnitude clustering (see Figures S15a and S15b in Supporting Information S1), especially among the largest events based on the ECDF maps (see Figures S15c and S15f in Supporting Information S1) and at the same time, we found significant non-stationarities in the FMD of the AE events throughout the experiment (see Figures S15d and S15e in Supporting Information S1). While the nature of the experiments differs from ours (i.e., fracture vs. stick-slip), the findings suggest that the same general conclusion holds that magnitude clustering is a consequence of non-stationarities in the FMD.

Xiong et al. (2023) also highlighted the influence of event-event triggering or aftershocks on magnitude clustering (Davidsen & Green, 2011). Aftershock sequences are manifestations of relaxation phenomena and are particularly pronounced after large seismic events and have also been observed in laboratory-scale seismicity (Baró et al., 2013; Davidsen et al., 2017, 2021; Goebel et al., 2023). This implies that large seismic events are often directly followed by a triggered seismic event such that the magnitude of the triggering event is typically larger than the magnitude of the triggered event *if no other event occurs between the two*. If we identify pairs of triggering and triggered AE events in our catalogs using a well-established method (Baiesi & Paczuski, 2004; Gu et al., 2013; Khajehdehi & Davidsen, 2023) (more details in Text S1 in Supporting Information S1), we can restrict our analysis to pairs of subsequent events that form a triggering-triggered pair (and those that do not, summarized in Table S4 in Supporting Information S1) and indeed confirm this expectation (see right column in Figures S16 and S17 in Supporting Information S1). The variations in the FMDs of different subgroups (as confirmed by the KS test) manifest themselves in significant magnitude clustering in sub-populations of the AE catalog, even stronger than (and different from) magnitude clustering in the entire catalog (see left and middle column in Figures S16 and S17 in Supporting Information S1), and consistent with the observations by Xiong et al. (2023).

Aftershock triggering can also contribute to magnitude clustering through variations in the FMD associated with the triggering itself. Based on empirical observations for natural earthquakes in Southern California, Nandan et al. (2019, 2022) argue that using the Gutenberg-Richter relation seismic events can be divided into three different classes with distinct FMDs: (a) events that are not triggered by other events (so-called background events) with b_{bkg}, (b) triggered events that are bigger than their trigger with $b_{aft>}$, and (c) triggered events that are smaller than their trigger with $b_{aft<}$. Nandan et al. (2022) found that $b_{aft>} \leq b_{bkg} \leq b_{aft<}$, which directly implies magnitude clustering and allowed them to develop an improved aftershock forecast. For our AE catalogs associated with slip events, we find a similar distinction between the three classes of events with a significant excess of large events for group (b) (see Figure S18 in Supporting Information S1). In contrast to the proposition by Nandan et al. (2022), however, the FMDs of the three classes are not stationary but vary largely to the same degree as the FMD of all events (see Figure S19 in Supporting Information S1). In particular, major slip events in wet samples exhibit the most significant non-stationarity in the FMD in both background and triggered events, manifesting as the most substantial magnitude clustering among our catalogs. This indicates again that the non-stationarities are the main driver behind the observed magnitude clustering.

Over the last decades, controlled laboratory SS experiments have significantly advanced our understanding of earthquakes and related crustal processes (Brace & Byerlee, 1966; Dresen et al., 2020; Goebel et al., 2023). In particular, the scale-free distribution of seismic event sizes as captured by the GR relation allows insights from small-scale seismicity to inform our understanding of large-scale natural seismic behavior. However, limitations such as finite sample size and fault length should be considered, along with the confinement of AE events to near-field stress transfer. Direct comparisons between AE events and earthquakes also require caution due to differences in magnitude definitions. Nonetheless, our findings suggest in general that magnitude clustering occurs in specific subsets of seismic catalogs, driven by non-stationarities in the FMD, which may extend to natural earthquakes where similar non-stationarities have been reported (Gulia et al., 2018; Matsumoto et al., 2024).

Understanding these non-stationarities is crucial for developing improved time-dependent forecast models for seismic hazard assessment including forecasting larger earthquakes (Gulia & Wiemer, 2019).

5. Conclusions

This study has provided a comprehensive analysis of magnitude clustering within the context of laboratory fault behavior, focusing on Westerly Granite samples in frictional sliding experiments, under fluid-saturated and dry conditions. We investigated magnitude clustering of AE events encompassing the periods during, following and preceding slip events on laboratory faults considering different magnitude thresholds and detection overload of the AE sensors. We demonstrated the existence of magnitude clustering, with the strongest clustering observed during major slip events in both fluid-saturated and dry Westerly Granite samples. Furthermore, we demonstrated that temporal variations in the FMDs of AE events—occurring during slip events—play a pivotal role in explaining the robust nature of this magnitude clustering. Our research results provide insight into the complex mechanisms underlying fault behavior.

Data Availability Statement

All data is available at: https://data.mendeley.com/datasets/t2cht66xpr/1. (Goebel, 2022). Data analyzed in Figure S15 in Supporting Information S1 is available at: https://doi.org/10.5281/zenodo.7328586 (Xiong et al., 2022).

Acknowledgments
This research was made possible by a Humboldt fellowship to Goebel, an NSF Career Award (Award 2142489) and an USGS NEHRP Grant (G21AP10050). OK acknowledges Alberta Innovates for financial support. JD was supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN/05221-2020).

References

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review*, *69*(6), 066106. https://doi.org/10.1103/physreve.69.066106

Bakun, W. H., & Joyner, W. B. (1984). The ml scale in central California. *Bulletin of the Seismological Society of America*, *74*(5), 1827–1843. https://doi.org/10.1785/bssa0740051827

Baró, J., Corral, Á., Illa, X., Planes, A., Salje, E. K., Schranz, W., et al. (2013). Statistical similarity between the compression of a porous material and earthquakes. *Physical Review Letters*, *110*(8), 088702. https://doi.org/10.1103/physrevlett.110.088702

Brace, W., & Byerlee, J. (1966). Stick-slip as a mechanism for earthquakes. *Science*, *153*(3739), 990–992. https://doi.org/10.1126/science.153.3739.990

Davidsen, J., & Baiesi, M. (2016). Self-similar aftershock rates. *Physical Review E*, *94*(2), 022314. https://doi.org/10.1103/physreve.94.022314

Davidsen, J., Goebel, T., Kwiatek, G., Stanchits, S., Baró, J., & Dresen, G. (2021). What controls the presence and characteristics of aftershocks in rock fracture in the lab? *Journal of Geophysical Research: Solid Earth*, *126*(10), e2021JB022539. https://doi.org/10.1029/2021jb022539

Davidsen, J., & Green, A. (2011). Are earthquake magnitudes clustered? *Physical Review Letters*, *106*(10), 108502. https://doi.org/10.1103/PhysRevLett.106.108502

Davidsen, J., Gu, C., & Baiesi, M. (2015). Generalized Omori–Utsu law for aftershock sequences in Southern California. *Geophysical Journal International*, *201*(2), 965–978. https://doi.org/10.1093/gji/ggv061

Davidsen, J., Kwiatek, G., Charalampidou, E.-M., Goebel, T., Stanchits, S., Rück, M., & Dresen, G. (2017). Triggering processes in rock fracture. *Physical Review Letters*, *119*(6), 068501. https://doi.org/10.1103/PhysRevLett.119.068501

Davidsen, J., Kwiatek, G., & Dresen, G. (2012). No evidence of magnitude clustering in an aftershock sequence of nano-and picoseismicity. *Physical Review Letters*, *108*(3), 1–4. https://doi.org/10.1103/PhysRevLett.108.038501

Davidsen, J., & Paczuski, M. (2005). Analysis of the spatial distribution between successive earthquakes. *Physical Review Letters*, *94*(4), 48501. https://doi.org/10.1103/PhysRevLett.94.048501

Dresen, G., Kwiatek, G., Goebel, T., & Ben-Zion, Y. (2020). Seismic and aseismic preparatory processes before large stick–slip failure. *Pure and Applied Geophysics*, *177*(12), 5741–5760. https://doi.org/10.1007/s00024-020-02605-x

Goebel, T. H. (2022). Fault roughness and laboratory aftershocks [Dataset]. *Mendeley Data*. https://doi.org/10.17632/t2cht66xpr.1

Goebel, T. H., Brodsky, E. E., & Dresen, G. (2023). Fault roughness promotes earthquake-like aftershock clustering in the lab. *Geophysical Research Letters*, *50*(8), e2022GL101241. https://doi.org/10.1029/2022gl101241

Goebel, T. H., Schorlemmer, D., Becker, T., Dresen, G., & Sammis, C. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophysical Research Letters*, *40*(10), 2049–2054. https://doi.org/10.1002/grl.50507

Goodfellow, S., Nasseri, M., Maxwell, S., & Young, R. (2015). Hydraulic fracture energy budget: Insights from the laboratory. *Geophysical Research Letters*, *42*(9), 3179–3187. https://doi.org/10.1002/2015gl063093

Gu, C., Schumann, A. Y., Baiesi, M., & Davidsen, J. (2013). Triggering cascades and statistical properties of aftershocks. *Journal of Geophysical Research: Solid Earth*, *118*(8), 4278–4295. https://doi.org/10.1002/jgrb.50306

Gulia, L., Rinaldi, A. P., Tormann, T., Vannucci, G., Enescu, B., & Wiemer, S. (2018). The effect of a mainshock on the size distribution of the aftershocks. *Geophysical Research Letters*, *45*(24), 13277–13287. https://doi.org/10.1029/2018gl080619

Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, *574*(7777), 193–199. https://doi.org/10.1038/s41586-019-1606-4

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, *34*(4), 185–188. https://doi.org/10.1785/bssa0340040185

Helmstetter, A., Kagan, Y. Y., & Jackson, D. D. (2006). Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bulletin of the Seismological Society of America*, *96*(1), 90–106. https://doi.org/10.1785/0120050067

Hodges, J., Jr. (1958). The significance probability of the Smirnov two-sample test. *Arkiv för matematik*, *3*(5), 469–486. https://doi.org/10.1007/bf02589501

Kagan, Y. Y. (2013). *Earthquakes: Models, statistics, testable forecasts*. John Wiley & Sons.

Karimi, K., & Davidsen, J. (2021). Aftershock triggering and spatial aftershock zones in fluid-driven settings: Discriminating induced seismicity from natural swarms. *Geophysical Research Letters*, *48*(15), e2020GL092267. https://doi.org/10.1029/2020gl092267
Karimi, K., & Davidsen, J. (2023). Separating primary and secondary mechanisms for seismicity induced by hydraulic fracturing. *Bulletin of the Seismological Society of America*, *113*(5), 1982–1991. https://doi.org/10.1785/0120220200
Khajehdehi, O., & Davidsen, J. (2023). On the potential role of viscoelasticity in fluid-induced seismicity. *Journal of Geophysical Research: Solid Earth*, *128*(12), e2023JB027349. https://doi.org/10.1029/2023jb027349
Kothari, S., Shcherbakov, R., & Atkinson, G. (2020). Statistical modeling and characterization of induced seismicity within the Western Canada Sedimentary Basin. *Journal of Geophysical Research: Solid Earth*, *125*(12), e2020JB020606. https://doi.org/10.1029/2020jb020606
Kwiatek, G., Goebel, T. H. W., & Dresen, G. (2014). Seismic moment tensor and b value variations over successive seismic cycles in laboratory stick-slip experiments. *Geophysical Research Letters*, *41*(16), 5838–5846. https://doi.org/10.1002/2014GL060159
Kwiatek, G., Martínez-Garzón, P., Davidsen, J., Malin, P., Karjalainen, A., Bohnhoff, M., & Dresen, G. (2022). Limited earthquake interaction during a geothermal hydraulic stimulation in Helsinki, Finland. *Journal of Geophysical Research: Solid Earth*, *127*(9), e2022JB024354. https://doi.org/10.1029/2022jb024354
Kwiatek, G., Martínez-Garzón, P., Goebel, T., Bohnhoff, M., Ben-Zion, Y., & Dresen, G. (2024). Intermittent criticality multi-scale processes leading to large slip events on rough laboratory faults. *Journal of Geophysical Research: Solid Earth*, *129*(3), e2023JB028411. https://doi.org/10.1029/2023jb028411
Lippiello, E., de Arcangelis, L., & Godano, C. (2008). Influence of time and space correlations on earthquake magnitude. *Physical Review Letters*, *100*(3), 38501. https://doi.org/10.1103/PhysRevLett.100.038501
Lippiello, E., Godano, C., & de Arcangelis, L. (2007). Dynamical scaling in branching models for seismicity. *Physical Review Letters*, *98*(9), 98501. https://doi.org/10.1103/PhysRevLett.98.098501
Maghsoudi, S., Baró, J., Kent, A., Eaton, D. W., & Davidsen, J. (2018). Interevent triggering in microseismicity induced by hydraulic fracturing. *Bulletin of the Seismological Society of America*, *108*(3A), 1133–1146. https://doi.org/10.1785/0120170368
Maghsoudi, S., Eaton, D. W., & Davidsen, J. (2016). Nontrivial clustering of microseismicity induced by hydraulic fracturing. *Geophysical Research Letters*, *43*(20), 10–672. https://doi.org/10.1002/2016gl070983
Mäkinen, T., Miksic, A., Ovaska, M., & Alava, M. J. (2015). Avalanches in wood compression. *Physical Review Letters*, *115*(5), 055501. https://doi.org/10.1103/physrevlett.115.055501
Matsumoto, S., Iio, Y., Sakai, S., & Kato, A. (2024). Strength dependency of frequency–magnitude distribution in earthquakes and implications for stress state criticality. *Nature Communications*, *15*(1), 4957. https://doi.org/10.1038/s41467-024-49422-7
McLaskey, G. C., & Glaser, S. D. (2012). Acoustic emission sensor calibration for absolute source measurements. *Journal of Nondestructive Evaluation*, *31*(2), 157–168. https://doi.org/10.1007/s10921-012-0131-2
Moradpour, J., Hainzl, S., & Davidsen, J. (2014). Nontrivial decay of aftershock density with distance in Southern California. *Journal of Geophysical Research: Solid Earth*, *119*(7), 5518–5535. https://doi.org/10.1002/2014JB010940
Nandan, S., Ouillon, G., & Sornette, D. (2019). Magnitude of earthquakes controls the size distribution of their triggered events. *Journal of Geophysical Research: Solid Earth*, *124*(3), 2762–2780. https://doi.org/10.1029/2018jb017118
Nandan, S., Ouillon, G., & Sornette, D. (2022). Are large earthquakes preferentially triggered by other large events? *Journal of Geophysical Research: Solid Earth*, *127*(8), e2022JB024380. https://doi.org/10.1029/2022jb024380
Petrillo, G., & Zhuang, J. (2022). The debate on the earthquake magnitude correlations: A meta-analysis. *Scientific Reports*, *12*(1), 20683. https://doi.org/10.1038/s41598-022-25276-1
Petrillo, G., & Zhuang, J. (2023). Verifying the magnitude dependence in earthquake occurrence. *Physical Review Letters*, *131*(15), 154101. https://doi.org/10.1103/physrevlett.131.154101
Ribeiro, H. V., Costa, L. S., Alves, L. G. A., Santoro, P. A., Picoli, S., Lenzi, E. K., & Mendes, R. S. (2015). Analogies between the cracking noise of ethanol-dampened charcoal and earthquakes. *Physical Review Letters*, *115*(2), 025503. https://doi.org/10.1103/PhysRevLett.115.025503
Schoenball, M., & Ellsworth, W. L. (2017). A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and Southern Kansas. *Journal of Geophysical Research: Solid Earth*, *122*(12), 10189–10206. https://doi.org/10.1002/2017JB014850
Scholz, C. H. (2002). *The mechanics of earthquakes and faulting* (2nd ed.). Cambridge University Press.
Shapiro, S. A. (2015). *Fluid-induced seismicity*. Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9781139051132
Sornette, D., & Werner, M. J. (2005). Apparent clustering and apparent background earthquakes biased by undetected seismicity. *Journal of Geophysical Research*, *110*(B9), B09303. https://doi.org/10.1029/2005jb003621
Spassiani, I., & Sebastiani, G. (2016). Magnitude-dependent epidemic-type aftershock sequences model for earthquakes. *Physical Review E*, *93*(4), 042134. https://doi.org/10.1103/physreve.93.042134
Stanchits, S., Vinciguerra, S., & Dresen, G. (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics*, *163*(5–6), 975–994. https://doi.org/10.1007/3-7643-7712-7_4
Turcotte, D. L. (1997). *Fractals and chaos in geology and geophysics*. Cambridge University Press.
Utsu, T., Ogata, Y., & Matsu'ura, R. S. (1995). The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth*, *43*, 1–33. https://doi.org/10.4294/jpe1952.43.1
van der Elst, N. J., & Shaw, B. E. (2015). Larger aftershocks happen farther away: Nonseparability of magnitude and spatial distributions of aftershocks. *Geophysical Research Letters*, *42*(14), 5771–5778. https://doi.org/10.1002/2015GL064734
Wetzler, N., Brodsky, E. E., & Lay, T. (2016). Regional and stress drop effects on aftershock productivity of large megathrust earthquakes. *Geophysical Research Letters*, *43*(23), 12012–12020. https://doi.org/10.1002/2016GL071104
Xiong, Q., Brudzinski, M., Gossett, D., Lin, Q., & Hampton, J. (2022). Seismic magnitude clustering is prevalent in field and laboratory catalogs [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.7328586
Xiong, Q., Brudzinski, M., Gossett, D., Lin, Q., & Hampton, J. (2023). Seismic magnitude clustering is prevalent in field and laboratory catalogs. *Nature Communications*, *14*(1), 2056. https://doi.org/10.1038/s41467-023-37782-5
Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, *101*(1), 018501. https://doi.org/10.1103/physrevlett.101.018501
Zang, A., Christian Wagner, F., Stanchits, S., Dresen, G., Andresen, R., & Haidekker, M. A. (1998). Source analysis of acoustic emissions in aue granite cores under symmetric and asymmetric compressive loads. *Geophysical Journal International*, *135*(3), 1113–1130. https://doi.org/10.1046/j.1365-246x.1998.00706.x