Tectonophysics 926 (2026) 231137

Contents lists available at ScienceDirect

Tectonophysics

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

Graph-based probabilistic earthquake clustering and forecasting in Sumatra

Giuseppe Petrillo [a,*], Stefania Gentili [b], Luca Dal Zilio [a,c]

[a] Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Ave, Singapore, 639798, Singapore
[b] National Institute of Oceanography and Applied Geophysics (OGS), Via Treviso, Udine (UD), 33100, Italy
[c] Asian School of the Environment, Nanyang Technological University, 50 Nanyang Ave, Singapore, 639798, Singapore

ARTICLE INFO

Dataset link: https://repogempa.bmkg.go.id/, https://doi.org/10.5281/zenodo.17153233

Keywords:
Earthquake clustering
ETAS model
Graph-based stochastic declustering
Sumatra subduction zone
b-value estimation
Earthquake forecasting

ABSTRACT

Earthquake clustering is a fundamental feature of seismicity and underpins many short-term forecasting models. However, conventional clustering techniques based on fixed space–time windows often fail in tectonically complex regions such as the Sumatra subduction zone, where seismicity is heterogeneous, offshore station coverage is sparse, and location uncertainties are substantial. Here, we develop a region-specific, physics-informed framework for clustering and forecasting earthquakes in Sumatra, combining stochastic modelling with empirical spatial and temporal constraints. We apply GRETAS (GRaph-based approach to ETAS), a graph-theoretical method based on the Epidemic-Type Aftershock Sequence (ETAS) model, enhanced with physically motivated filtering. Compared to traditional window-based methods, GRETAS produces more compact and physically consistent clusters, with improved performance measured by the Silhouette coefficient and Davies–Bouldin index. We also evaluate the forecasting potential of the Gutenberg–Richter *b*-value using a probabilistic, weighted version of the *b*-more-positive estimator, which accounts for classification uncertainty and magnitude incompleteness. Our results show no statistically significant difference in *b*-values between background and triggered events, and no consistent precursory trend prior to large earthquakes. These findings underscore the importance of probabilistic, regionally tailored approaches for robust seismic analysis in complex tectonic settings.

1. Introduction

Earthquakes rarely occur in isolation. Instead, they tend to cluster in space and time, forming complex sequences such as foreshocks (Mignan, 2014; Petrillo and Lippiello, 2021), aftershocks (Omori, 1894), and swarms (Hainzl, 2004; Godano et al., 2023). These patterns reflect a wide range of interacting physical processes, including stress transfer (Petrillo et al., 2020), pore pressure perturbations (Petrillo et al., 2024a), aseismic slip (Radiguet et al., 2016), and dynamic triggering (Brodsky and van der Elst, 2014). Deciphering the origin and structure of these clusters remains a central goal in seismology, with significant implications for real-time hazard assessment and earthquake forecasting. Indeed, many operational and research-driven forecasting models (Lippiello et al., 2019; Gentili and Di Giovambattista, 2017, 2020, 2022; Anyfadi et al., 2023; Brondi et al., 2024; Gentili et al., 2025b; Lippiello et al., 2025b) rely on the accurate identification of clustered seismicity, a process that is itself highly sensitive to the choice of clustering methodology (Gentili et al., 2024).

Over the past two decades, statistical models – most notably the Epidemic-Type Aftershock Sequence (ETAS) model (Ogata, 1988, 1989; Console et al., 2007, 2010; Omi et al., 2014) and its many extensions (Petrillo and Lippiello, 2021, 2023; Petrillo and Zhuang, 2024; Petrillo et al., 2024b; Molkenthin et al., 2024; Ross, 2021; Petrillo and Taroni, 2025) – have transformed how earthquake interactions are modelled. These frameworks represent seismicity as a branching process, allowing a more realistic and probabilistic characterization of earthquake cascades beyond deterministic mainshock–aftershock classifications. However, the practical identification of earthquake clusters from real catalogues remains methodologically challenging. Many widely used techniques rely on fixed space–time windows or empirical rules, with limited grounding in physics and little consideration of catalogue uncertainties (Reasenberg, 1985; Uhrhammer, 1986; Gentili et al., 2025a). This issue is particularly acute in tectonically complex or poorly instrumented regions, where observational limitations (e.g., station density, offshore coverage) introduce spatially variable uncertainty in event detection and location.

The Sumatra subduction zone illustrates these challenges. As part of the Sunda arc, Sumatra accommodates the convergence of the Indo-Australian and Eurasian plates, hosting some of the largest and most destructive earthquakes of the past century, including the 2004 Mw

* Corresponding author.
E-mail address: giuseppe.petrillo@ntu.edu.sg (G. Petrillo).

https://doi.org/10.1016/j.tecto.2026.231137
Received 3 November 2025; Received in revised form 18 February 2026; Accepted 21 February 2026
Available online 21 February 2026

9.1 Aceh-Andaman and the 2005 Mw 8.6 Nias events (Lay et al., 2005; Borrero et al., 2011). Despite its tectonic significance, the region has received comparatively limited attention in the statistical seismology literature. Previous studies have tended to focus on individual megathrust events or have applied globally calibrated models, often overlooking regional features such as elongated rupture geometries, spatiotemporal variations in catalogue completeness, and the elevated location uncertainties associated with offshore seismicity (Widiyantoro et al., 2024; Diantari et al., 2018; Haridhi et al., 2018). These limitations underscore the need for tailored approaches that integrate both physical constraints and probabilistic modelling adapted to the Sumatran context.

In this study, we present a comprehensive framework for clustering and forecasting earthquakes in Sumatra. Central to our approach is GRETAS (GRaph-based approach to ETAS), a novel method that combines stochastic declustering (Zhuang et al., 2002) with graph theory to flexibly identify clusters of interacting events (Gentili et al., 2025b). Unlike conventional declustering techniques or window-based methods, GRETAS constructs a network of triggering probabilities and identifies clusters based on the persistence of these links across multiple stochastic realizations. The resulting structure captures the probabilistic nature of seismic triggering and is further refined using empirically calibrated spatial and temporal constraints, ensuring that the identified clusters are both statistically robust and physically plausible. This step is particularly important in Sumatra, where uncertainties in earthquake locations can otherwise lead to spurious long-range associations. Alongside physics-based and statistical approaches, a growing number of studies have explored data-driven and machine learning-based methods for earthquake clustering and forecasting. These include pattern-recognition techniques based on supervised (Shcherbakov and Kothari, 2025) or unsupervised classifiers (Scarfì et al., 2025; Piegari et al., 2022; Birant and Kut, 2007) designed to identify clustered seismicity. Such methods have shown encouraging performance in specific tectonic settings, particularly where long, homogeneous catalogs are available and the seismic regime is relatively stable. Nevertheless, the application of purely data-driven approaches to regions such as Sumatra remains challenging. Strong spatial heterogeneity, variable detection capability, offshore seismicity, and evolving catalog completeness can limit the transferability and interpretability of machine learning models. Moreover, training labels in supervised approaches often rely on prior declustering assumptions, potentially propagating methodological biases. For these reasons, we adopt here a probabilistic, physics-informed framework based on stochastic seismicity modelling, while recognizing machine learning approaches as a complementary avenue for future developments.

Beyond clustering, we also revisit a long-standing and controversial question in earthquake forecasting: can the Gutenberg–Richter b-value (Gutenberg and Richter, 1944) serve as a reliable precursor to large earthquakes? The b-value is often interpreted as a proxy for crustal stress state (Amitrano, 2003), and numerous studies have reported systematic decreases in b preceding major events (Papadopoulos et al., 2018, 2010; Gulia and Wiemer, 2019; Nanjo et al., 2012). However, other investigations have failed to find consistent precursory signals, or have shown that apparent changes in b may reflect biases arising from catalogue incompleteness, regional heterogeneity, or assumptions in declustering (Lombardi, 2023; Godano et al., 2024b; Herrmann and Marzocchi, 2021). The reliability of b-value variations as a forecasting tool thus remains uncertain and may depend strongly on both region-specific factors and methodological choices.

Traditional b-value estimators also neglect key sources of uncertainty, such as whether a given event is part of the background or triggered population, or whether the catalogue is complete above a uniform magnitude threshold. To address these issues, we adopt the recently proposed b more positive method (Lippiello and Petrillo, 2024; Lippiello et al., 2025a), which is based on inter-event magnitude differences and provides greater robustness against such biases. We further extend this method by introducing a probabilistic, weighted estimator that explicitly incorporates the posterior classification probabilities of each event derived from ETAS-based declustering (Zhuang et al., 2002). This formulation allows us to directly account for classification uncertainty and to construct principled confidence intervals on b-value estimates.

We apply this methodology to the regional catalogue of the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG), covering the period 2008–2024. We find that GRETAS produces more compact and geophysically meaningful clusters than conventional methods, as quantified by standard clustering metrics such as the Silhouette coefficient (Rousseeuw, 1987) and the Davies–Bouldin index (Davies and Bouldin, 2009). Our probabilistic b-value analysis reveals no statistically significant difference between background and triggered events, and a retrospective analysis of large earthquakes shows no systematic decrease in b prior to major ruptures. These findings suggest that the predictive power of the b-value, at least in the Sumatran context, may be limited. More broadly, our results highlight the need for region-specific, probabilistic models that explicitly incorporate observational and methodological uncertainties. By integrating stochastic simulations, physical constraints, and robust statistical estimators, the framework developed here contributes to a more transparent and reliable understanding of clustered seismicity in one of the world's most hazardous tectonic environments.

2. Seismic catalogue characterization

We begin by characterizing the key statistical features of the earthquake catalogue compiled by the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG), and highlighting the limitations of conventional clustering methods when applied to this tectonically complex region. We focus on the Sumatra segment of the Sunda Arc, where the Indo-Australian plate subducts beneath the Eurasian Sunda plate (Lay et al., 2005; Haridhi et al., 2018; Widiyantoro et al., 2024). The region of interest spans from 95°E to 141°E longitude and from 11°S to 6°N latitude, encompassing the main seismogenic segments of the megathrust and associated tectonic domains, including the accretionary prism and the back-arc region.

The catalogue spans the period from 1 November 2008 to 24 November 2024 and aggregates seismic records from an extensive network of broadband and short-period stations operated by BMKG, supplemented in some regions by international networks (Haridhi et al., 2018). Earthquake magnitudes are reported using multiple scales, including local magnitude (M_L), short-period body-wave magnitude (m_b), and moment magnitude (M_w). To ensure consistency across the dataset, all magnitudes were converted to moment magnitude (M_w) using empirical regression models calibrated by Diantari et al. (2018), who fitted least-squares relationships between BMKG-reported values and the GFZ reference catalogue. These conversions yielded coefficients of determination ranging from 0.54 to 0.92, depending on the original magnitude type (see Table 2 in Diantari et al. (2018)).

The magnitude of completeness (M_c) of the catalogue was estimated using the Coefficient of Variation (CV) stabilization method recently proposed by Godano et al. (2024a) (see Methods). To account for potential temporal fluctuations in detectability – due to changes in network configuration, background noise levels, or seismic rate – we computed $M_c(t)$ over a moving window of 250 events. The completeness threshold for the entire dataset was conservatively defined as $M_c = \max(M_c(t)) = 4.7$, consistent with previous assessments for this region (Diantari et al., 2018). Only events with $M_w \geq 4.7$ were retained for subsequent analyses. This threshold ensures that spatial and temporal variations in seismicity are not biased by incomplete detection at lower magnitudes (see, e.g., Petrillo and Zhuang (2022, 2023)). The spatial distribution of the selected events is shown in Fig. 1. Each epicentre is colour-coded by magnitude, and larger events ($M_w \geq 7$) are marked with yellow stars. A

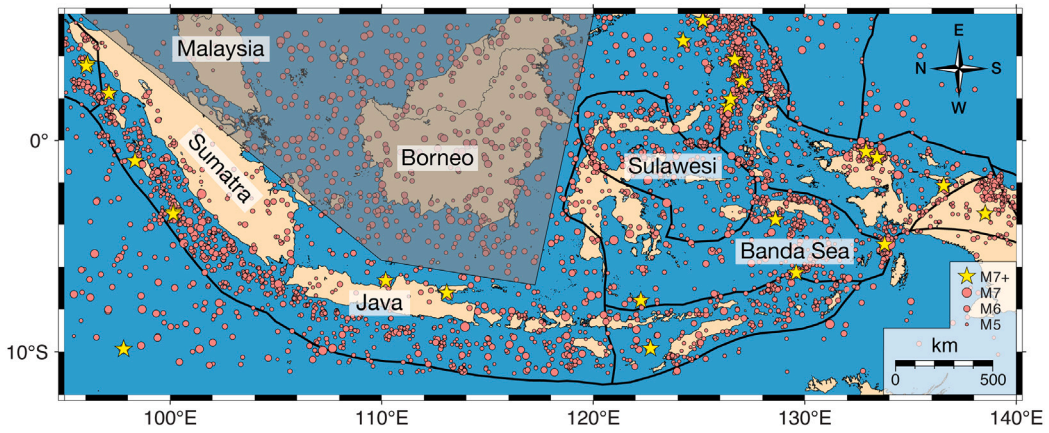

Fig. 1. Epicentres from the BMKG catalogue (1 November 2008 to 24 November 2024), for events with $M_w \geq 4.7$, corresponding to the estimated magnitude of completeness. Dot size scales with magnitude; yellow stars denote $M_w \geq 7$ events. The shaded region was excluded due to a locally elevated magnitude of completeness.

subset of the region, shaded in grey, was excluded from further analysis due to locally elevated completeness thresholds.

To ensure catalogue reliability, we conducted a systematic inspection for potential duplicate events, which can arise when an earthquake is relocated but the original entry is not removed, leading to artificial repetitions. Each event was checked for the occurrence of a nearby event within a spatial window of 0.5° in both latitude and longitude and within 5 s in origin time. Suspected duplicates were manually cross-checked against the ISC catalogue (Storchak et al., 2013); if no distinct events were confirmed, the earlier record-presumed to correspond to the initial, uncorrected location-was removed. This filtering step eliminated 32 events with magnitudes between 4.7 and 6.9.

3. Methods

In the following, we describe in detail the methods we used for declustering and cluster identification (stochastic declustering and GRETAS), and to estimate of cluster quality (Silhouette mean and Davies–Bouldin index), the completeness magnitude (Godano et al. unbiased b value estimate), and b value (b-more positive). For further details please refer to the original publications in literature listed in the following sub sections.

3.1. ETAS model and stochastic declustering

The Epidemic Type Aftershock Sequence (ETAS) model is a branching process model designed to represent how seismic events are triggered by other earthquakes in both space and time (Ogata, 1988, 1998). It views each event as a potential "ancestor" that can produce "offspring", depending on its magnitude, time, and location. The model's conditional intensity function for an earthquake occurring at time t, location (x, y) and magnitude M, given the earthquake history up to a time t, is formulated as:

$$\lambda(t,x,y,M|\mathcal{H}_t) = \mu(x,y,M) + \sum_{i:t_i<t} \kappa(M_i) g(t-t_i) f(x-x_i, y-y_i, M_i) h(M, M_i), \quad (1)$$

where $\lambda(t,x,y,M|\mathcal{H}_t)$ is the conditional intensity at given the earthquake history $\mathcal{H}_t$. $\mu(x,y,M)$ is the time-independent space-dependent background seismic rate. The sum represents the triggered part, i.e., seismicity caused by past events, $\kappa(M_i|\alpha) = A\exp(\alpha(M_i - M_0))$ is the productivity of the ith event, which depends on its magnitude M_i and M_0 is the lower magnitude; $g(t-t_i|p,c)$ is the Omori–Utsu law (Utsu et al., 1995); $f(x-x_i, y-y_i, M_i|q,d,\gamma) = [(x-x_i)^2 + (y-y_i)^2 + d_0 10^{\gamma M_i}]^{-q}$ is the spatial distribution function with the aftershock rupture area included (Ogata and Zhuang, 2006; Kagan, 2002), representing how far aftershocks are distributed from their parent event; $h(M, M_i|b)$ is the magnitude distribution of the triggered events. We assume that the magnitude distribution of the background events is independent on its location (Ogata and Zhuang, 2006), $\mu(x,y,M) = \mu(x,y)h_\mu(M|b)$ and is identical for background events and their offspring (Petrillo and Zhuang, 2022, 2023; Taroni, 2024; Corrado et al., 2024), that is,

$$h(M, M_i|b) = h_\mu(M|b) = b\ln(10)10^{-b(M-M_0)}. \quad (2)$$

Although this formulation implies a spatially uniform b-value across the study region, we explicitly tested the validity of this assumption for the Sumatra catalog. To this end, we subdivided the study area into four longitudinal sectors (105E, 118E, 130E) and independently estimated the b-value in each subregion using the same magnitude threshold and estimation procedure. The resulting b-values overlap within their respective uncertainties, and no statistically significant spatial variations were observed. This supports the use of a spatially uniform magnitude distribution in the ETAS formulation and indicates that the declustering results and subsequent analyses are robust with respect to plausible regional variations in b. We set the form of the background intensity as in Zhuang et al. (2002), namely $\mu(x,y) = \nu u(x,y)$.

The set of parameters $\vec{\theta} = \{\alpha, p, c, q, d, \gamma, d_0, \mu\}$ is optimized by means of Maximum Likelihood Estimation (MLE). Which means that we need to maximize the log-likelihood function

$$\log L = \sum_{i=1}^{N} \log \lambda_\theta(t_i, x_i, y_i|\mathcal{H}_{t_i}) - \iiint_{T\times\Omega} \lambda_\theta(t,x,y|\mathcal{H}_t)dtdxdy \quad (3)$$

However, it is possible to use alternative methods for parameter estimation, such as Bayesian inversion (Ross, 2021; Ross and Kolev, 2022; Molkenthin et al., 2022; Petrillo and Zhuang, 2024) or other techniques use for models where there is no analytical form of the likelihood function (Petrillo and Lippiello, 2021, 2023).

Once the parameters have been optimized, it is necessary to reconstruct the family tree of events, that is, to determine which events are "immigrants" (background) and which are offsprings (triggered). Probabilistic approaches, such as the stochastic declustering method (Zhuang et al., 2002), allow us to estimate the probability that an event is spontaneous or triggered by others. Specifically, the probability that event j is triggered by event i is given by

$$\rho_{ij} = \Pr\{j \leftarrow i \mid \mathcal{H}_t\} = \frac{\kappa(m_i)\, g(t_j - t_i)\, f(x_j - x_i, y_j - y_i)}{\mu(x_j, y_j) + \sum_{k:t_k<t_j} \kappa(m_k)\, g(t_j - t_k)\, f(x_j - x_k, y_j - y_k)}, \quad (4)$$

where $j \leftarrow i$ denotes that event j is a direct offspring of event i, and $\mathcal{H}_t$ is the history of the process up to time t. In other words, ρ_{ij} measures the conditional probability that i is the parent of j, given all previous events. Accordingly, we denote by $j \in \mathcal{I}$ the case where event j is an

immigrant (background) event. Its probability is

$$\psi_j = \Pr\{j \in \mathcal{I} \mid \mathcal{H}_i\} = 1 - \sum_{i=1}^{j-1} \rho_{ij}. \quad (5)$$

Thus, each event is assigned a probability of being spontaneous or triggered, and the collection of these probabilities provides a probabilistic reconstruction of the underlying branching structure of the catalog.

3.2. GRETAS: Graph-based approach to ETAS clustering

To identify seismic clusters in a flexible yet physically motivated way, we developed a hybrid deterministic–probabilistic method referred to as GRETAS. This method combines the stochastic declustering capabilities of the ETAS model (Zhuang et al., 2002) with a graph-theoretical framework that captures the persistent connectivity between events.

Traditional clustering techniques, such as window-based methods, rely on fixed spatial and temporal thresholds (e.g., Uhrhammer (1986)) and are prone to over-clustering or misclassification in complex tectonic regions. In contrast, GRETAS leverages the probability of causal interaction between earthquake pairs to identify clusters in a more adaptive and data-driven manner.

Step 1: Stochastic declustering via ETAS. We first apply the stochastic declustering method (Zhuang et al., 2002), which estimates the probability ρ_{ij} that a given event j was triggered by a previous event i, and the complementary probability ψ_j that event j is part of the background (i.e., not triggered) (see Eqs. (4), (5))

The declustering algorithm is then executed $k = 10$ times, each time assigning events to background or triggered categories based on a random number $U_j \in [0, 1]$ compared to ψ_j. The resulting cluster configurations vary across runs due to this stochastic assignment.

Step 2: Graph construction and event connectivity. We interpret the declustering output as an oriented graph, where nodes represent events and directed edges (with weight ρ_{ij}) represent possible triggering links. Strongly connected events, i.e., those consistently linked to the same mainshock across multiple stochastic realizations, are considered robust members of a seismic cluster.

Step 3: Clustre stabilization and mainshock assignment. For each of the k declustering realizations, we identify clusters and assign to each cluster a representative mainshock, the largest magnitude event (or the earliest if magnitudes are equal). We then retain only those events that consistently appear in the same cluster across all k realizations. This intersection yields a stable and unique set of cluster members for each mainshock. Importantly, the graph-based nature of GRETAS is not intended to recover a unique parent–child genealogy, as in classical stochastic ETAS declustering. Instead, it uses the triggering graph to define *stable* cluster membership across an ensemble of stochastic realizations. While individual parent links can change from one realization to another, an event is interpreted as a robust member of a given cluster if it remains connected to the same mainshock in a consistent way across realizations, possibly through different intermediate events. In this ensemble view, cluster membership is treated separately from specific ancestry, and events supported only by weak or unstable connectivity are naturally excluded from the final clusters.

Step 4: Post-processing via empirical spatial constraints. This step consists in applying empirically calibrated spatial and temporal filters to exclude outliers that fall beyond plausible bounds of cluster extent. Formally, for each event associated to a mainshock of magnitude M_m, we define admissible spatial and temporal windows:

$$\text{Cluster} = \left\{ j : r_{ij} \le D(M_m), \quad \Delta t_{ij} \le T(M_m) \right\},$$

where r_{ij} is the spatial (e.g., epicentral) distance between event j and the mainshock i, and Δt_{ij} is the time difference. The functions $D(M_m)$ and $T(M_m)$ represent empirically derived distance and duration thresholds, which can be calibrated for each region based on observational data or physical heuristics (e.g., rupture scaling laws, typical aftershock decay times). Note that in this application, no temporal filtering was necessary.

This filtering acts as a deterministic envelope that constrains the output of the probabilistic clustering. It serves two complementary purposes: (i) it removes events that are spuriously linked to a cluster via long-range weak associations, and (ii) it ensures that the final clusters respect known physical characteristics of aftershock sequences, such as spatial compactness and temporal decay.

Note that this step is modular: the functional form of $D(M_m)$ and $T(M_m)$ is not fixed by the GRETAS method itself, but is specified depending on the region and the seismicity under study. This modularity allows GRETAS to adapt to different tectonic settings without compromising methodological consistency.

It is important to clarify that Steps 1–3 constitute the core of the GRETAS algorithm. These steps define cluster membership through an ensemble-based probabilistic framework grounded in the ETAS model, without relying on fixed space–time windows or empirical thresholds. Step 4 is introduced as an optional post-processing procedure adopted in this application to facilitate comparison with classical window-based declustering methods and to enforce additional spatial constraints motivated by location uncertainties. As such, Step 4 is not a fundamental component of the GRETAS framework and does not affect the probabilistic definition of clusters provided by Steps 1–3.

To facilitate comparison with classical window-based declustering approaches, we explicitly note how each step of GRETAS addresses their main limitations. Step 1 replaces fixed space–time windows with probabilistic, model-based triggering relationships, avoiding arbitrary threshold choices. Step 2 propagates these relationships through multiple generations, overcoming the inability of window methods to account for cascade triggering. Steps 3 and 4 jointly account for uncertainty in cluster membership and clustering variability, whereas window-based methods typically produce a single, deterministic partition of the catalog.

3.3. Silhouette mean and Davies–Bouldin index

To quantitatively compare the clustering quality produced by different methods (GRETAS without correction, GRETAS with spatial correction, and window-based clustering), we evaluated two well established internal validation metrics: the Silhouette mean and the Davies–Bouldin index.

Silhouette mean. For each event i assigned to cluster C_k, the silhouette coefficient s_i is defined as:

$$s_i = \frac{\beta_i - a_i}{\max(a_i, b_i)} \quad (6)$$

where:

- a_i is the average distance between event i and all other events within the same cluster C_k (intra-cluster cohesion),
- β_i is the minimum of the average distances between event i and all events in other clusters C_l, with $l \neq k$ (inter-cluster separation).

The silhouette value s_i ranges from −1 to 1, where values close to 1 indicate well-clustered events, values around 0 suggest overlapping clusters, and negative values indicate possible misclassification. The Silhouette mean is the average of s_i over all events in the dataset, providing an overall measure of clustering quality.

Davies–Bouldin index. This index evaluates clustering compactness and separation by comparing intra-cluster scatter with inter-cluster distances. It is computed as follows:

1. For each cluster C_i, compute the average intra-cluster distance (scatter):

$$S_i = \frac{1}{|C_i|} \sum_{x \in C_i} \|x - \xi_i\|,$$

where ξ_i is the centroid of cluster C_i.

2. For each pair of clusters (i, j), compute the inter-centroid distance:

$$M_{ij} = \|\xi_i - \xi_j\|.$$

3. Compute the similarity between clusters:

$$R_{ij} = \frac{S_i + S_j}{M_{ij}}.$$

4. For each cluster i, take the worst-case (maximum) similarity:

$$R_i = \max_{j \neq i} R_{ij}.$$

5. The Davies–Bouldin index is then defined as:

$$\mathrm{DB} = \frac{1}{N_c} \sum_{i=1}^{N_c} R_i,$$

where N_c is the total number of clusters.

Lower values of the Davies–Bouldin index indicate better clustering performance, as they correspond to compact and well separated clusters.

Distance definition for spatio-temporal clustering. Since our clustering is performed in both space and time, we defined a composite metric between events i and j as:

$$\tau_{ij}^{(ST)} = \sqrt{\Delta t_{ij}^2 + \left(\eta\, \Delta x_{ij}\right)^2},$$

where Δt_{ij} is the time difference in days, Δx_{ij} is the spatial (epicentral) distance in km. Here, η has dimensions of time per unit length (days/km), corresponding to the inverse of a characteristic propagation velocity. Consequently, the term $\eta \times \Delta x_{ij}$ has dimensions of time, ensuring that the spatio-temporal distance metric is dimensionally consistent. We tested values of η ranging from 0.1 to 1000 days/km to assess the robustness of the clustering methods across different space–time scales.

3.4. Completeness magnitude

Seismic catalogs often suffer from incompleteness, meaning that many smaller earthquakes go unreported. The extent of this incompleteness depends on factors such as the ability to distinguish seismic events from background noise and the proximity of seismic stations to the earthquake epicentre, which determines whether an event is recorded in the catalog (Schorlemmer and Woessner, 2008; Mignan, 2012). This issue is particularly pronounced in the early stages of aftershock sequences, where coda waves from large earthquakes and their aftershocks obscure subsequent smaller events (Kagan, 2004; Helmstetter et al., 2006; Peng et al., 2007; de Arcangelis et al., 2016; Hainzl, 2016). Accurately estimating the completeness magnitude, M_c, is crucial for a reliable evaluation of the b value in seismic analysis but also for simulating stochastic models like ETAS. Several methods exist to estimate M_c, including the maximum curvature technique, which identifies based on the peak of the non-cumulative GR distribution (Wiemer and Wyss, 2000) and the goodness-of-fit method, which performs an exponential fit of the GR distribution as a function of threshold magnitude and selects as the smallest yielding an accurate fit (Wiemer and Wyss, 2000). The b-value stability approach identifies as the point where the estimated b-value becomes stable (Cao and Gao, 2002; Godano and Petrillo, 2023). Other techniques include the entire magnitude range method, utilizing the full magnitude range modified by a complementary error function (Ogata and Katsura, 1993), and the harmonic mean-based method (Godano, 2017).

In this study, we employ the method developed by Godano et al. (2024a), where the completeness magnitude is estimated by analysing the relationship between the estimated b-value and the variability coefficient CV. More precisely, we define the quantity $m^* = m - M_{th}$ and evaluate its variability coefficient

$$CV = \frac{\sigma_{th}}{\langle m^* \rangle} \tag{7}$$

where σ_{th} is the standard deviation of m^* and $\langle m^* \rangle$ is its mean. For an exponential distribution $CV = 1$. Therefore, CV is evaluated as a function of M_{th}. When $CV \simeq 1 \rightarrow M_c = M_{th}$.

3.5. Weighted b-more-positive estimation

From the stochastic declustering we can disentangle two subset from the seismicity: the set of background events S_{bg} and the set of triggered events S_{trig}. For each pair (i, j) that belongs to the *background* subset S_{bg}, we define a weight

$$w_{ij}^{(\mathrm{bg})} = \psi_i \cdot \psi_j.$$

where ψ is defined in Eq. (5)

Analogously, for a pair (i, j) in the *triggered* subset S_{trig},

$$w_{ij}^{(\mathrm{trig})} = (1 - \psi_i)(1 - \psi_j).$$

Below, we illustrate the procedure for the *background* case; the triggered case is handled the same way, simply using its respective weights $w_{ij}^{(\mathrm{trig})}$.

We accumulate the following sums over all valid pairs (i, j) in S_{bg}:

$$(1)\quad \Sigma_w = \sum_{(i,j) \in S_{\mathrm{bg}}} w_{ij}, \qquad (2)\quad \Sigma_{w\Delta M} = \sum_{(i,j) \in S_{\mathrm{bg}}} w_{ij}\, \Delta M_{ij},$$

$$(3)\quad \Sigma_{w(\Delta M)^2} = \sum_{(i,j) \in S_{\mathrm{bg}}} w_{ij}\, (\Delta M_{ij})^2, \qquad (4)\quad \Sigma_{w^2} = \sum_{(i,j) \in S_{\mathrm{bg}}} (w_{ij})^2.$$

Therefore the weighted mean of ΔM is given by

$$\overline{\Delta M} = \frac{\Sigma_{w\Delta M}}{\Sigma_w}.$$

We denote it as $\mu_{\Delta M}$ whenever needed:

$$\mu_{\Delta M} = \frac{\Sigma_{w\Delta M}}{\Sigma_w}.$$

To capture the variance of ΔM, we define the weighted variance in a manner consistent with typical weighted least squares or physics-based conventions:

$$\mathrm{Var}_w(\Delta M) = \frac{1}{\Sigma_w} \sum_{(i,j) \in S_{\mathrm{bg}}} w_{ij} \left(\Delta M_{ij} - \mu_{\Delta M}\right)^2.$$

We use an algebraic trick to avoid summing in two passes:

$$\sum w\,(\Delta M)^2 - 2\,\mu_{\Delta M} \sum w\,(\Delta M) + \mu_{\Delta M}^2 \sum w = \sum_{(i,j)} w_{ij} \left(\Delta M_{ij} - \mu_{\Delta M}\right)^2.$$

Thus:

$$\mathrm{Var}_w(\Delta M) = \frac{\Sigma_{w(\Delta M)^2} - 2\,\mu_{\Delta M}\, \Sigma_{w\Delta M} + \mu_{\Delta M}^2\, \Sigma_w}{\Sigma_w}.$$

Since we have a set of weights w_{ij}, we can define an "effective number of independent observations" via

$$N_{\mathrm{eff}} = \frac{\left(\Sigma_w\right)^2}{\Sigma_{w^2}}.$$

This quantity accounts for the fact that some pairs have larger or smaller weights than others.

The standard deviation of $\overline{\Delta M}$ can be approximated as

$$\sigma_{\overline{\Delta M}} \approx \sqrt{\frac{\mathrm{Var}_w(\Delta M)}{N_{\mathrm{eff}}}} = \sqrt{\frac{\mathrm{Var}_w(\Delta M)}{\left(\Sigma_w\right)^2}\,\Sigma_{w^2}}\quad .$$

Once we have the weighted mean difference of magnitudes,

$$\mu_{\Delta M} = \overline{\Delta M},$$

we define

$$b = \frac{1}{\ln(10)\,\mu_{\Delta M}}.$$

Denote $\sigma_{\overline{\Delta M}}$ as the standard error of $\overline{\Delta M}$. We seek σ_b such that

$$\sigma_b^2 \approx \left(\frac{db}{d\mu_{\Delta M}}\right)^2 \sigma^2_{\overline{\Delta M}}.$$

Since

$$b(\mu_{\Delta M}) = \frac{1}{\ln(10)\,\mu_{\Delta M}},$$

its derivative is

$$\frac{db}{d\mu_{\Delta M}} = -\frac{1}{\ln(10)\,\mu^2_{\Delta M}}.$$

Hence,

$$\sigma_b = \frac{1}{\ln(10)\,\mu^2_{\Delta M}}\,\sigma_{\overline{\Delta M}} = b^2\,\ln(10)\,\sigma_{\overline{\Delta M}}.$$

4. Results

4.1. Physically informed probabilistic clustering of seismicity

Traditional window-based declustering methods remain widely used in seismology, yet their performance is highly sensitive to the region and catalogue properties. These methods require tuning of empirical laws and often struggle to separate closely spaced events in time and space, leading to ambiguities in cluster assignment. To address these limitations, we apply a hybrid deterministic–probabilistic clustering framework tailored to the Sumatran context. Building on our recent work in Japan (Gentili et al., 2025b), we employ GRETAS (GRaph-based approach to ETAS; see Methods), a flexible graph-theoretical method based on ETAS stochastic declustering (Zhuang et al., 2002). Unlike window-based techniques, GRETAS avoids arbitrary thresholds and instead constructs clusters using probabilistic triggering links inferred from ETAS simulations. However, in the absence of additional physical constraints, this flexibility may lead to unrealistic associations between spatially or temporally distant events, particularly in regions with significant location uncertainty.

We illustrate this issue in Fig. 2a, where GRETAS is applied without spatial or temporal filtering to identify clusters associated with $M_m \geq 7$ mainshocks. Although many aftershock-rich sequences are captured successfully, several clusters (e.g., labels 1–3) include events located more than 100–200 km from the mainshock. These long-range links, while permitted by ETAS triggering probabilities, are likely artefacts caused by offshore location uncertainty or misclassified background events. To improve spatial realism, we introduce a modified version of the classic Uhrhammer law, adding a tolerance of 55 km to account for the elevated location uncertainties characteristic of the Indonesian region. The resulting spatial threshold is:

$$D = \exp(-1.024 + 0.804 \times M_m) + 55 \quad \text{(km)} \tag{8}$$

The additional 55 km term in Eq. (8) is introduced tentatively as a tolerance margin accounting for epicentral location uncertainties, which are particularly pronounced for offshore earthquakes in the Sumatran region. This value corresponds approximately to half a degree in latitude or longitude and was empirically motivated by the distribution of cluster-member distances obtained from the unconstrained

Algorithm 1 Procedure for computing the weighted b-more-positive

1: **Read the event catalog:** For each event, read $(M_i, p_i, \mathrm{lat}_i, \mathrm{lon}_i)$.
2: **Thresholding:** Filter events by magnitude: keep only $M_i \geq M_{\mathrm{thr}}$.
3: **For each event** i**:**

- Look up to l subsequent events $(i+1), \ldots, (i+l)$.
- For each candidate j, compute:
  - Epicentral distance r_{ij} (km) between events i and j.
  - Magnitude difference $\Delta M_{ij} = M_j - M_i$.
- **Selection criteria:** Select the first j such that:

$$r_{ij} \leq d_{\mathrm{th}}, \qquad \Delta M_{ij} > d\,M_{\mathrm{c}}.$$

- If such j is found, form the pair (i, j) and compute the *shifted* magnitude difference:

$$\Delta M'_{ij} = \Delta M_{ij} - d\,M_{\mathrm{c}}.$$

4: **Background vs. triggered:**

- Accumulate sums in the *background* subset, i.e. $\Sigma_w, \Sigma_{w\Delta M'}, \Sigma_{w(\Delta M')^2}, \Sigma_{w^2}$, with weight $w_{ij} = p_i \cdot p_j$.
- Similarly, accumulate in the *triggered* subset, with $w_{ij} = (1 - p_i)(1 - p_j)$.

5: **Compute the weighted mean** of $\Delta M'$ for each subset:

$$\mu_{\Delta M'} = \frac{\Sigma_{w\Delta M'}}{\Sigma_w}.$$

6: **Compute the weighted variance** $\mathrm{Var}_w(\Delta M')$ and the effective sample size N_{eff}.
7: **Derive the standard error** of the mean:

$$\sigma_{\overline{\Delta M'}} = \sqrt{\frac{\mathrm{Var}_w(\Delta M')}{N_{\mathrm{eff}}}}.$$

8: **Compute** b **and its error**:

$$b = \frac{1}{\ln(10)\,\mu_{\Delta M'}}, \qquad \sigma_b = \frac{1}{\ln(10)\,\mu^2_{\Delta M'}}\,\sigma_{\overline{\Delta M'}}.$$

9: **Repeat** for the *background* and *triggered* subsets separately.

GRETAS clustering (Fig. 3a), where many events systematically fall beyond the classical Uhrhammer limit but remain within an additional ~50–60 km envelope. The impact of this spatial constraint is shown in Figs. 2(b–d). Compared to the unconstrained version, clusters become more compact, and distant outliers are removed. To validate the filtering radius, we analysed the distribution of distances between cluster events and their mainshock as a function of mainshock magnitude (Fig. 3a). Many events lie just beyond the classical Uhrhammer limit but within the extended +55 km envelope. These results confirm that applying a tolerance is crucial for reducing false associations in a catalogue with variable location quality. Next, we turn to the temporal extent of seismic clusters. Standard declustering algorithms often rely on fixed or magnitude-dependent time windows, which may inadequately represent the evolving nature of earthquake sequences, particularly in the presence of swarm-like activity. Rigid time cutoffs can lead to inclusion of unrelated background events or the omission of delayed aftershocks that remain physically connected to the mainshock. To address this, we propose a new empirically derived time-windowing rule, based on observations from the Sumatran catalogue (Fig. 3b):

$$T = \begin{cases} 10^{M_m - 4.5} & \text{if } M_m \leq 6.5 \\ 100 & \text{otherwise} \end{cases} \tag{9}$$

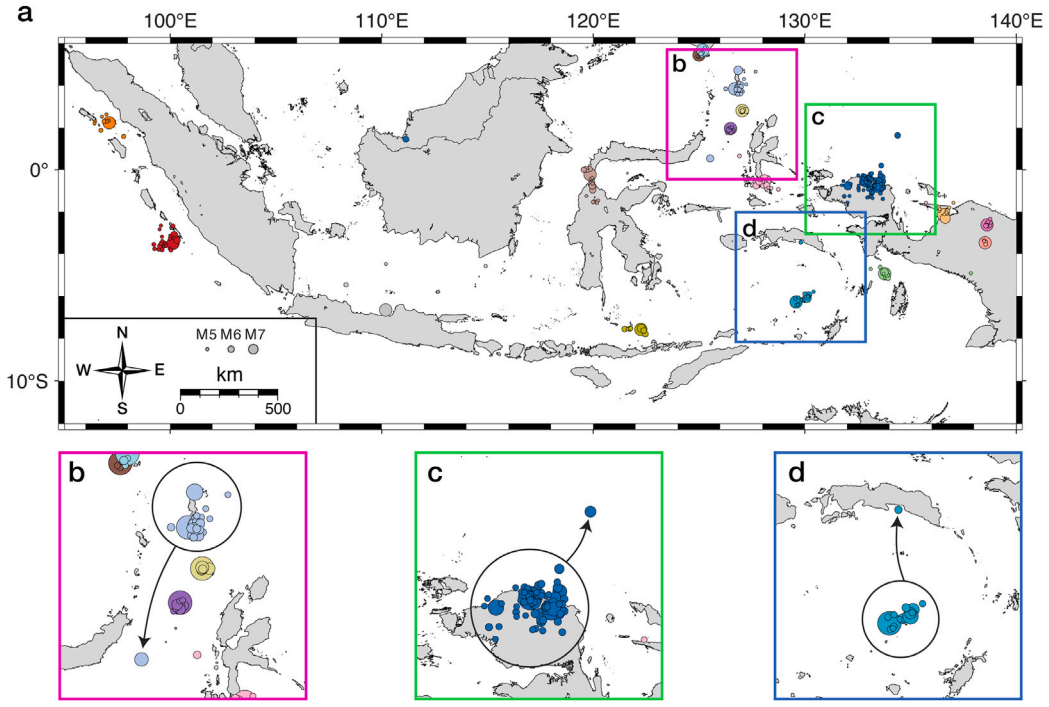

Fig. 2. Comparison of seismic clusters identified with different levels of spatiotemporal filtering, using the GRETAS method for mainshocks with $M_m \geq 7$. (a) Map showing clusters from ETAS-based stochastic declustering without additional filters. Several clusters (e.g., b–d) include spatial outliers. (b–c) Applying the modified Uhrhammer law (Eq. (8)) removes distant outliers and yields more compact clusters.

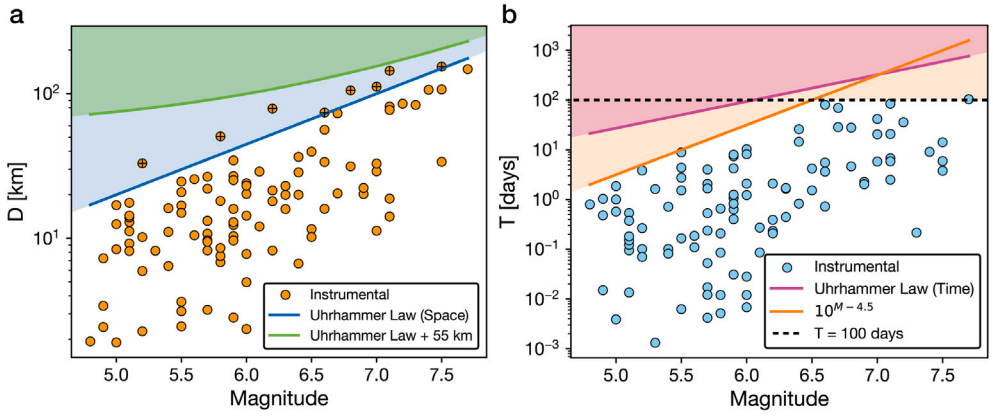

Fig. 3. Empirical justification for the spatial and temporal filters. (a) Distance from mainshock vs. magnitude for all cluster events. Dots show observations; the blue and green lines correspond to the original and modified spatial laws, respectively. Black crosses mark events excluded by spatial filtering. (b) Cluster duration vs. mainshock magnitude. The proposed law (orange) captures the observed temporal evolution and saturates at 100 days.

This scaling captures the typical increase in aftershock duration with mainshock magnitude, consistent with earlier studies (Gentili and Bressan, 2008), but saturates at 100 days—likely due to the relatively high completeness threshold of the catalogue ($M_w \geq 4.7$), which excludes later, lower-magnitude aftershocks. To test this rule, we implemented a window-based clustering scheme that includes all events occurring within the spatial radius (Eq. (8)) and within time T from the mainshock. This test evaluates whether a physically informed yet simple windowing strategy can approximate more sophisticated probabilistic clustering.

Different cluster methods outline similarly general characteristics of the region. First of all, the level of seismic hazard in the area: the region on the west of Sumatra around longitude 100 (Sumatra Subduction zone) and the one around longitude 125–130 (Philippine Trench) is the more prone to earthquakes with magnitude greater than 7 (see Fig. 1). In particular, from the cluster analysis, the smaller difference in magnitude between the mainshock and the strongest aftershock is recorded in the Philippine Trench; a difference in magnitude smaller or equal to 1 correspond to the most dangerous clusters, because the strong aftershock occurs when structures have been already damaged by the mainshock (Gentili et al., 2025b). These results are in good agreement with analyses available in literature (Petersen et al., 2004; Hutchings and Mooney, 2021).

To compare the different clustering approaches quantitatively, we computed the Silhouette coefficient (Rousseeuw, 1987) and the Davies–Bouldin index (Davies and Bouldin, 2009) for clusters with $M_m \geq 7$ (see Methods). The Silhouette coefficient measures the degree to which events are more similar to their own cluster than to others (ranging from −1 to 1, with higher values indicating better separation), while the Davies–Bouldin index evaluates intra-cluster compactness and inter-cluster separation (lower values indicate better performance). Because our clusters span space and time, we define a composite distance

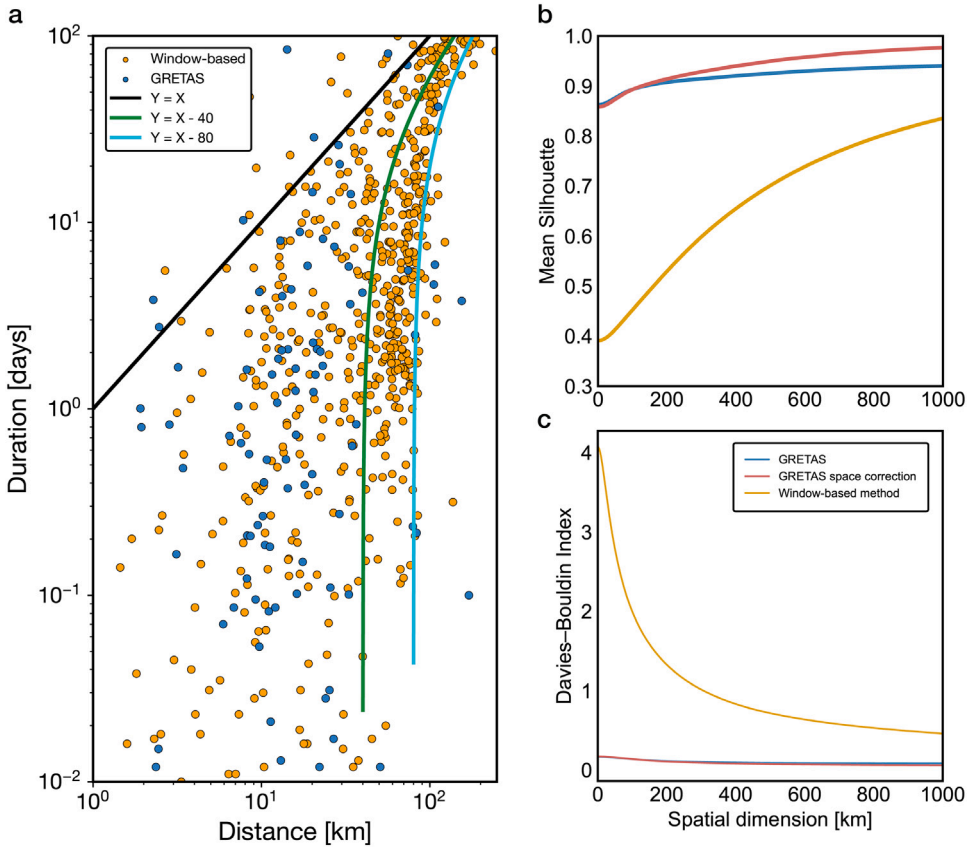

Fig. 4. Comparison of clustering performance and scaling relations. (a) Cluster duration versus maximum distance from the mainshock for clusters associated with $M \geq 7$ earthquakes. Orange circles indicate window-based clustering and blue circles GRETAS. Solid lines show reference trends ($Y = X$, $Y = X - 40$, $Y = X - 80$). The broader spread of the window-based method reflects the inclusion of unrelated background events, inflating cluster durations. (b) Mean silhouette coefficient as a function of spatial dimension. Higher values denote more coherent clusters; GRETAS and its space-corrected variant consistently outperform the window-based method. (c) Davies–Bouldin index versus spatial dimension. Lower values indicate more compact and well-separated clusters. Both GRETAS variants achieve superior performance compared to the window-based approach.

metric that combines spatial and temporal coordinates, expressed in kilometres and days, respectively. We examine the effect of rescaling the spatial dimension using a multiplicative factor from 0.1 to 1000, thereby tuning the relative importance of space and time (e.g., 1 day corresponds to 0.1–1000 km).

Fig. 4a shows the relationship between cluster duration and spatial extent for the window-based method and the unconstrained GRETAS method. The broader spread observed in the window-based clusters reflects the inclusion of temporally unrelated events. This suggests that rigid time–space windows tend to overestimate cluster duration due to background contamination. Figs. 4b and 4c show how performance metrics vary with the assumed space–time scaling. The GRETAS-based methods consistently outperform the window-based approach across a broad range of scaling factors. At small space-to-time ratios, the unconstrained GRETAS method performs best, whereas at larger ratios the spatially constrained variant is superior. This behaviour can be explained by the presence of spatial outliers that are temporally close to mainshock events: when time dominates the metric, these events remain in the cluster; when space dominates, they are excluded. This phenomenon is consistent with known limitations of other space–time clustering methods, such as the Nearest-Neighbour approach (Zaliapin et al., 2008; Zaliapin and Ben-Zion, 2013; Gentili et al., 2025a).

In summary, while empirical windowing rules (e.g., Eq. (9)) may offer marginal improvements over classical methods, they remain limited by their rigidity and inability to account for the diversity of seismic sequences. Our results demonstrate that probabilistic, data-driven methods such as GRETAS offer significantly greater robustness and flexibility. By incorporating physical constraints into a stochastic framework, GRETAS provides a more realistic representation of clustered seismicity in complex tectonic environments.

4.2. Evaluating the forecasting potential of the Gutenberg–Richter *b*-value

Whether the Gutenberg–Richter *b*-value differs systematically between background and triggered earthquakes remains a subject of ongoing debate. Previous global studies, including analyses of the GCMT catalogue (Taroni et al., 2025), found no statistically significant difference between the *b*-values of background and triggered events. However, global catalogues aggregate data from a wide range of tectonic settings, which may obscure region-specific behaviours. Moreover, earlier analyses often relied on estimators that are susceptible to biases introduced by catalogue incompleteness and magnitude uncertainties. In this study, we revisit this question using a refined methodology that combines the *b*-more-positive estimator (Lippiello and Petrillo, 2024) with recent advances in probabilistic *b*-value inference (Pastoressa et al., 2023; Taroni et al., 2024, 2025). This approach explicitly incorporates magnitude uncertainties and the posterior probabilities of event classification derived from stochastic declustering. By applying this method to the Sumatra subduction zone, we assess whether regional variations exist and whether *b*-values can serve as reliable indicators of impending large earthquakes.

We analyse seismicity from the BMKG instrumental catalogue, selecting all events with moment magnitude $M_w \geq M_c = 4.7$ and focal depths $z < 70$ km. This selection yields a dataset of $N_{\text{sub}} = 6378$ events. Events are classified as background or triggered based

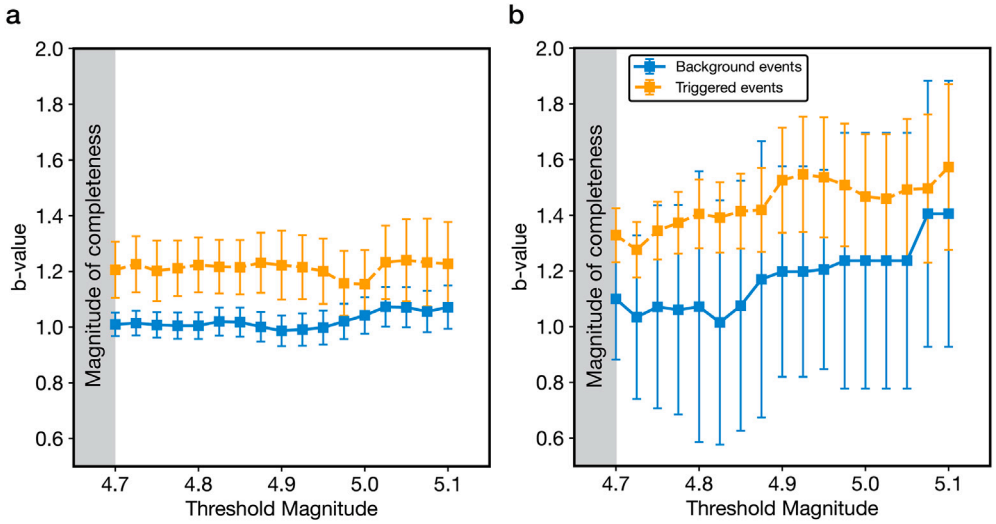

Fig. 5. Estimated b-values for background and triggered events as a function of threshold magnitude. (a) Synthetic ETAS catalogue results. The estimator correctly recovers the input b-values for background (blue) and triggered (red) events. (b) BMKG catalogue results. No statistically significant difference is observed between the b-values of background and triggered events across magnitude thresholds. Error bars denote 1σ standard errors derived from uncertainty propagation of weighted magnitude differences.

on stochastic declustering using the original method of Zhuang et al. (2002). Unlike full clustering, declustering only requires assigning a probability that an event was triggered, without identifying its specific parent. For this reason, we adopt the classical ETAS model for this analysis, calibrated for Sumatra with optimized parameters $\bar{\theta} = (\mu, K, c, \alpha, p, d, q, \gamma) = (0.98, 0.54, 0.014, 1.67, 1.07, 0.0008, 1.07, 0.66)$. Using a posterior probability threshold of 0.5, we classify 62% of events ($N_b = 3,965$) as background and 38% ($N_t = 2,413$) as triggered. To estimate the b-value, we implement a new weighted version of the b-more-positive method (Lippiello and Petrillo, 2024) (see Methods), which assigns probabilistic weights to each event based on its classification likelihood. This formulation provides a more robust estimate of the b-value and its uncertainty, accounting for both classification ambiguity and magnitude measurement errors, while reducing biases associated with magnitude truncation and catalogue incompleteness.

The accuracy of this estimator was validated using synthetic ETAS catalogues with a known input b-value and a magnitude threshold $M_{\text{th}} = 4.7$. Stochastic declustering yields classification probabilities for each synthetic event, and the weighted estimator is then applied to compute separate b-values for background and triggered populations. As shown in Fig. 5a, the method successfully recovers the input values: $b \simeq 1.0$ for background and $b \simeq 1.2$ for triggered events, consistent with the simulation parameters. We then apply the same weighted estimator to the real BMKG data. Following stochastic declustering, we reconstruct the probabilistic ancestry of each event and compute b-values for background and triggered events over a range of magnitude thresholds. The results, presented in Fig. 5b, reveal no statistically significant difference between the two populations. This finding confirms, in a region-specific context, the results previously obtained from global catalogues (Taroni et al., 2025). Next, we investigate whether temporal variations in the b-value can serve as a precursor to large earthquakes. Since our earlier analysis showed no significant b-value differences between background and triggered events, we treat the entire dataset as a single population. We use the classical (unweighted) b-more-positive estimator (Lippiello and Petrillo, 2024) to track changes in b preceding each $M \geq 7$ mainshock.

For each mainshock, we compute the b-value within spatial windows defined by Eq. (8) and temporal windows of 4, 3, 2, and 1 years preceding the event. This approach mirrors that of earlier forecasting studies (Gulia and Wiemer, 2019; Godano et al., 2024b), but is systematically applied to a regionally homogenized, completeness-corrected catalogue. While some mainshocks are preceded by low b-values that could be interpreted as possible precursors, the pattern is inconsistent across the dataset. Furthermore, as the temporal window shortens, the number of events decreases, leading to larger uncertainties and more variable estimates. These results suggest that, in this region, b-value fluctuations do not provide a reliable or consistent forecasting signal for large earthquakes.

5. Discussion and conclusions

We introduced a new framework for clustering and forecasting seismicity in the Sumatra subduction zone—an area of complex tectonics, spatially heterogeneous seismicity, and limited offshore observational coverage. Our approach integrates stochastic modelling of earthquake interactions with graph-based classification techniques, culminating in the development and application of GRETAS. By combining probabilistic triggering with empirically calibrated spatial and temporal constraints, GRETAS addresses key limitations of traditional window-based clustering and provides a more flexible, physically informed means of identifying seismic sequences. Quantitative benchmarking using the Silhouette coefficient and Davies–Bouldin index demonstrates that window-based clustering methods tend to overassign events to clusters. In particular, background events falling within arbitrary space–time windows are often misclassified as aftershocks. This effect, previously noted in studies on declustering artefacts and spatial biases (Peresan and Gentili, 2020; Gentili et al., 2025a), becomes especially pronounced in regions with elongated rupture geometries, non-uniform station coverage, or complex spatiotemporal seismicity patterns (Reasenberg, 1985; Uhrhammer, 1986). In contrast, GRETAS, especially when supplemented with spatial constraints based on location uncertainty, produces more coherent and compact clusters while preserving the inherent flexibility of stochastic modelling.

Our analysis also shows that the durations of seismic sequences identified by GRETAS are typically shorter than those predicted by classical rules (e.g., the Uhrhammer law), and better captured by the empirical time window defined in Eq. (9). Even when the same empirical time window is used with a traditional window-based method, the resulting clusters tend to be longer in duration and more spatially dispersed than those obtained by GRETAS (Fig. 4a). These results confirm the added value of probabilistic, data-driven clustering frameworks in regions where traditional rules are poorly suited. Beyond clustering,

we assessed the forecasting potential of the Gutenberg–Richter *b*-value. While previous studies have reported that a systematically decreasing *b*-value may signal increased hazard (Gulia and Wiemer, 2019), our analysis of the BMKG catalogue reveals that temporal fluctuations in *b* preceding large earthquakes are inconsistent and subject to substantial statistical uncertainty. These results reinforce earlier findings from global studies (e.g., Taroni et al. (2025)) and suggest that the predictive utility of the *b*-value is both region-dependent and highly sensitive to methodological assumptions, including the declustering approach and magnitude of completeness.

A key methodological advance in our forecasting analysis is the implementation of a probabilistic *b*-value estimator that incorporates classification uncertainty and magnitude measurement error (Lippiello and Petrillo, 2024; Pastoressa et al., 2023). By using posterior probabilities from stochastic declustering (Zhuang et al., 2002), we compute weighted *b*-values for background and triggered populations and validate the method on synthetic data. Both in synthetic and real datasets, we find no statistically significant difference between the *b*-values of background and triggered events. Overall, our results underscore the need for regionally calibrated, physically informed, and probabilistically rigorous tools for analysing clustered seismicity and assessing seismic hazard. See the Data Availability section for the link to download the clusters obtained using the three different approaches. The flexibility of GRETAS makes it suitable for application to other regions and catalogues, particularly in offshore or data-sparse environments where conventional clustering approaches fail to perform reliably. Future work should focus on incorporating additional physical constraints into the clustering and forecasting pipeline. This includes integrating geodetic data, Coulomb stress changes, and rate-and-state frictional properties, as well as exploring hybrid frameworks that combine statistical indicators such as the *b*-value with physics-based models of fault loading and rupture nucleation. Moreover, extending the probabilistic framework to explicitly include epistemic uncertainty in model structure—such as through ensemble approaches or Bayesian model averaging (Petrillo and Zhuang, 2024; Laporte et al., 2025), may offer a path toward more robust and generalizable earthquake forecasting strategies.

CRediT authorship contribution statement

Giuseppe Petrillo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Stefania Gentili:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Luca Dal Zilio:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Funding acquisition.

Additional information

Code Availability

The stochastic declustering software is free to download at https://bemlar.ism.ac.jp/zhuang/software.html while the code used for the postprocessing is archived on Zenodo with the DOI https://doi.org/10.5281/zenodo.16749793.

Declaration of competing interest

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

Acknowledgements

L.D.Z. and G.P. were supported by the Earth Observatory of Singapore (EOS), the Singapore Ministry of Education Tier 3b project "Investigating Volcano and Earthquake Science and Technology (InVEST)" (Award No. MOE-MOET32021-0002) and the Nanyang Assistant Professorship (NAP) (Award No. 025244-00001). S.G. would like to acknowledge the funding by a grant from the Italian Ministry of Foreign Affairs and International Cooperation, the co-funding within the RETURN Extended Partnership which received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan - NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005) and the co-funding by the NEar realtiME results of Physical and StatIstical Seismology for earthquakes observations, modelling and forecasting (NEMESIS) Project (INGV). We thank Giuseppe Davide Chiappetta for his help in developing the GRETAS algorithm.

Data availability

The regional catalogue of the Indonesian Agency for Meteorology, Climatology and Geophysics (BMKG) can be downloaded at https://repogempa.bmkg.go.id/. The clusters obtained in this paper are archived on Zenodo with the DOI https://doi.org/10.5281/zenodo.17153233.

References

Amitrano, D., 2003. Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b value. J. Geophys. Res.: Solid Earth 108.

Anyfadi, E.A., Gentili, S., Brondi, P., Vallianatos, F., 2023. Forecasting strong subsequent earthquakes in Greece with the machine learning algorithm nestore. Entropy 25, 797.

de Arcangelis, L., Godano, C., Grasso, J.R., Lippiello, E., 2016. Statistical physics approach to earthquake occurrence and forecasting. Phys. Rep. 628, 1–91.

Birant, D., Kut, A., 2007. ST-DBSCAN: An algorithm for clustering spatial–temporal data. Data Knowl. Eng. 60, 208–221.

Borrero, J.C., McAdoo, B., Jaffe, B., Dengler, L., Gelfenbaum, G., Higman, B., Hidayat, R., Moore, A., Kongko, W., Lukijanto, Peters, R., Prasetya, G., Titov, V., Yulianto, E., 2011. Field survey of the March 28, 2005 nias-simeulue earthquake and tsunami. Pure Appl. Geophys. 168, 1075–1088. http://dx.doi.org/10.1007/s00024-010-0218-6.

Brodsky, E.E., van der Elst, N.J., 2014. The uses of dynamic earthquake triggering. Annu. Rev. Earth Planet. Sci. 42, 317–339.

Brondi, P., Gentili, S., Giovambattista, R.D., 2024. Forecasting strong subsequent events in the italian territory: a national and regional application for nestorev1.0. Nat. Hazards 121, http://dx.doi.org/10.1186/s40623-024-02096-3.

Cao, A., Gao, S.S., 2002. Temporal variation of seismic b-values beneath Northeastern Japan island arc. Geophys. Res. Lett. 29, 48–1–48–3.

Console, R., Jackson, D.D., Kagan, Y.Y., 2010. Using the ETAS model for catalog declustering and seismic background assessment. Pure Appl. Geophys. 167, 819–830. http://dx.doi.org/10.1007/s00024-010-0065-5.

Console, R., Murru, M., Catalli, F., Falcone, G., 2007. Real time forecasts through an earthquake clustering model constrained by the rate-and-state constitutive law: Comparison with a purely stochastic ETAS model. Seismol. Res. Lett. 78, 49–56. http://dx.doi.org/10.1785/gssrl.78.1.49.

Corrado, P., Herrmann, M., Marzocchi, W., 2024. Earthquake magnitude correlations expose short-term catalog incompleteness. Seismol. Res. Lett. http://dx.doi.org/10.1785/0220240277.

Davies, D.L., Bouldin, D.W., 2009. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 224–227.

Diantari, H., Suryanto, W., Anggraini, A., Irnaka, T., Susilanto, P., Ngadmanto, D., 2018. Preliminary magnitude of completeness quantification of improved BMKG catalog (2008–2016) in indonesian region. In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012026.

Gentili, S., Bressan, G., 2008. The partitioning of radiated energy and the largest aftershock of seismic sequences occurred in the Northeastern Italy and Western Slovenia. J. Seism. 12, 343–354. http://dx.doi.org/10.1007/s10950-007-9075-x.

Gentili, S., Brondi, P., Rossi, G., Sugan, M., Petrillo, G., Zhuang, J., 2025a. Seismic clusters and fluids diffusion: a lesson from the 2018 molise (Southern Italy) earthquake sequence. Earth, Planets Space 76, 3499–3531. http://dx.doi.org/10.1007/s11069-024-06913-6.

Gentili, S., Brondi, P., Rossi, G., Sugan, M., Petrillo, G., Zhuang, J., Campanella, S., 2024. Seismic clusters and fluids diffusion: a lesson from the 2018 molise (Southern Italy) earthquake sequence. Earth, Planets Space 76, 157.

Gentili, S., Chiappetta, G., Petrillo, G., Brondi, P., Zhuang, J., 2025b. Forecasting strong subsequent earthquakes in Japan using an improved version of nestore machine learning algorithm. Geosci. Front. 16, 102016. http://dx.doi.org/10.1016/j.gsf.2025.102016, URL https://www.sciencedirect.com/science/article/pii/S1674987125000167.

Gentili, S., Di Giovambattista, R., 2017. Pattern recognition approach to the subsequent event of damaging earthquakes in Italy. Phys. Earth Planet. Inter. 266, 1–17.

Gentili, S., Di Giovambattista, R., 2020. Forecasting strong aftershocks in earthquake clusters from Northeastern Italy and Western Slovenia. Phys. Earth Planet. Inter. 303, 106483.

Gentili, S., Di Giovambattista, R., 2022. Forecasting strong subsequent earthquakes in California clusters by machine learning. Phys. Earth Planet. Inter. 327, 106879.

Godano, C., 2017. A new method for the estimation of the completeness magnitude. Phys. Earth Planet. Inter. 263, 7–11.

Godano, C., Petrillo, G., 2023. Estimating the completeness magnitude mc and the b-values in a snap. Earth Space Sci. 10.

Godano, C., Petrillo, G., Lippiello, E., 2024a. Evaluating the incompleteness magnitude using an unbiased estimate of the b value. Geophys. J. Int. 236, 994–1001.

Godano, C., Tramelli, A., Mora, M., Taylor, W., Petrillo, G., 2023. An analytic expression for the volcanic seismic swarms occurrence rate. A case study of some volcanoes in the world. Earth Space Sci. 10, e2022EA002534.

Godano, C., Tramelli, A., Petrillo, G., Convertito, V., 2024b. Testing the predictive power of b value for italian seismicity. Seismica 3.

Gulia, L., Wiemer, S., 2019. Real-time discrimination of earthquake foreshocks and aftershocks. Nature 574, 193–199.

Gutenberg, B., Richter, C.F., 1944. Frequency of earthquakes in California. Bull. Seismol. Soc. Am. 34, 185–188.

Hainzl, S., 2004. Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering. Geophys. J. Int. 159, 1090–1096.

Hainzl, S., 2016. Apparent triggering function of aftershocks resulting from rate-dependent incompleteness of earthquake catalogs. J. Geophys. Res.: Solid Earth 121, 6499–6509.

Haridhi, H.A., Huang, B.S., Kuo-Liang, W., Denzema, D., Prasetyo, R.A., Chao-Shing, L., 2018. A study of large earthquake sequences in the Sumatra subduction zone and its possible implications. TAO: Terr. Atmos. Ocean. Sci. 29, 6.

Helmstetter, A., Kagan, Y.Y., Jackson, D.D., 2006. Comparison of short-term and time-independent earthquake forecast models for Southern California. Bull. Seismol. Soc. Am. 96, 90–106.

Herrmann, M., Marzocchi, W., 2021. Inconsistencies and lurking pitfalls in the magnitude–frequency distribution of high-resolution earthquake catalogs. Seism. Soc. Am. 92, 909–922.

Hutchings, S.J., Mooney, W.D., 2021. The seismicity of Indonesia and tectonic implications. Geochem. Geophys. Geosystems 22.

Kagan, Y.Y., 2002. Aftershock zone scaling. Bull. Seismol. Soc. Am. 92, 641–655.

Kagan, Y.Y., 2004. Short-term properties of earthquake catalogs and models of earthquake source. Bull. Seismol. Soc. Am. 94, 1207–1228.

Laporte, M., Durand, S., Gardonio, B., Marsan, D., 2025. B-bayesian: The full probabilistic estimate of b-value temporal variations for non-truncated catalogs. J. Geophys. Res.: Solid Earth 130, e2024JB029973.

Lay, T., Kanamori, H., Ammon, C.J., Nettles, M., Ward, S.N., Aster, R.C., Beck, S.L., Bilek, S.L., Brudzinski, M.R., Butler, R., et al., 2005. The Great Sumatra-Andaman earthquake of 26 December 2004. Science 308, 1127–1133.

Lippiello, E., Godano, C., Petrillo, G., 2025a. Improve the estimate of the b-value in regional catalogs by means of the the b-more positive method. arXiv preprint arXiv:2504.04663.

Lippiello, E., Petrillo, G., 2024. B-more-incomplete and b-more-positive: Insights on a robust estimator of magnitude distribution. J. Geophys. Res.: Solid Earth 129, e2023JB027849.

Lippiello, E., Petrillo, G., Godano, C., Dal Zilio, L., 2025b. Toward recognizing the waveform of foreshocks. Geophys. Res. Lett. 52, e2025GL115466.

Lippiello, E., Petrillo, G., Godano, C., Tramelli, A., Papadimitriou, E., Karakostas, V., 2019. Forecasting of the first hour aftershocks by means of the perceived magnitude. Nat. Commun. 10, 2953.

Lombardi, A., 2023. Anomalies and transient variations of b-value in Italy during the major earthquake sequences: what truth is there to this? Geophys. J. Int. 232, 1545–1555.

Mignan, A., 2012. Functional shape of the earthquake frequency-magnitude distribution and completeness magnitude. J. Geophys. Res.: Solid Earth 117.

Mignan, A., 2014. The debate on the prognostic value of earthquake foreshocks: A meta-analysis. Sci. Rep. 4, 4099.

Molkenthin, C., Donner, C., Reich, S., Zöller, G., Hainzl, S., Holschneider, M., Opper, M., 2022. Gp-etas: semiparametric bayesian inference for the spatio-temporal epidemic type aftershock sequence model. Stat. Comput. 32, 29.

Molkenthin, C., Zöller, G., Hainzl, S., Holschneider, M., 2024. Bayesian earthquake forecasting using gaussian process modeling: Gp-etas applications. Seismol. Res. Lett. 95, 3532–3544.

Nanjo, K., Hirata, N., Obara, K., Kasahara, K., 2012. Decade-scale decrease inb value prior to the m9-class 2011 Tohoku and 2004 Sumatra quakes. Geophys. Res. Lett. 39.

Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. J. Amer. Statist. Assoc. 83, 9–27.

Ogata, Y., 1989. Statistical model for standard seismicity and detection of anomalies by residual analysis. Tectonophysics 169, 159–174.

Ogata, Y., 1998. Space-time point process models for earthquake occurrences. Ann. Inst. Statist. Math. 50, 379–402.

Ogata, Y., Katsura, K., 1993. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues. Geophys. J. Int. 113, 727–738.

Ogata, Y., Zhuang, J., 2006. Space–time etas models and an improved extension. Tectonophysics 413, 13–23.

Omi, T., Ogata, Y., Hirata, Y., Aihara, K., 2014. Estimating the ETAS model from an early aftershock sequence. Geophys. Res. Lett. 41, 850–857. http://dx.doi.org/10.1002/2013GL058958.

Omori, F., 1894. On aftershocks of earthquakes. J. Coll. Sci. Imp. Univ. TOkyo 7, 111–200.

Papadopoulos, G., Charalampakis, M., Fokaefs, A., Minadakis, G., 2010. Strong foreshock signal preceding the l'aquila (Italy) earthquake (m w 6.3) of 6 April 2009. Nat. Hazards Earth Syst. Sci. 10, 19–24.

Papadopoulos, G., Minadakis, G., Orfanogiannaki, K., 2018. Short-term foreshocks and earthquake prediction. In: Pre-earthquake processes: a multi-disciplinary approach to earthquake prediction studies. pp. 125–147.

Pastoressa, A.E., Murru, M., Taroni, M., Console, R., Montuori, C., Falcone, G., Di Stefano, R., 2023. Temporal variations of seismicity rates and Gutenberg–Richter b-values for a stochastic declustered catalog: An example in central Italy. Seism. Soc. Am. 94, 1566–1578.

Peng, Z., Vidale, J.E., Ishii, M., Helmstetter, A., 2007. Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan. J. Geophys. Res.: Solid Earth 112.

Peresan, A., Gentili, S., 2020. Identification and characterisation of earthquake clusters: a comparative analysis for selected sequences in Italy and adjacent regions. Boll. di Geofis. Teor. Ed Appl. 61, 57–80. http://dx.doi.org/10.4430/bgta0249.

Petersen, M.D., Dewey, J., Hartzell, S., Mueller, C., Harmsen, S., Frankel, A.D., Rukstales, K., 2004. Probabilistic seismic hazard analysis for Sumatra, Indonesia and across the Southern Malaysian Peninsula. Tectonophysics 390, 141–158.

Petrillo, G., Kumazawa, T., Napolitano, F., Capuano, P., Zhuang, J., 2024a. Fluids-triggered swarm sequence supported by a nonstationary epidemic-like description of seismicity. Seismol. Res. Lett. 95, 3207–3220.

Petrillo, G., Lippiello, E., 2021. Testing of the foreshock hypothesis within an epidemic like description of seismicity. Geophys. J. Int. 225, 1236–1257.

Petrillo, G., Lippiello, E., 2023. Incorporating foreshocks in an epidemic-like description of seismic occurrence in Italy. Appl. Sci. 13.

Petrillo, G., Lippiello, E., Landes, F.P., Rosso, A., 2020. The influence of the brittle-ductile transition zone on aftershock and foreshock occurrence. Nat. Commun. 11, 3010.

Petrillo, G., Lippiello, E., Zhuang, J., 2024b. Including stress relaxation in point-process model for seismic occurrence. Geophys. J. Int. 236, 1332–1341.

Petrillo, G., Taroni, M., 2025. Adding strain rate information into a short-term seismicity model improves forecasting performances: the case of campi flegrei, Italy. Seismica 4.

Petrillo, G., Zhuang, J., 2022. The debate on the earthquake magnitude correlations: A meta-analysis. Sci. Rep. 12, 20683.

Petrillo, G., Zhuang, J., 2023. Verifying the magnitude dependence in earthquake occurrence. Phys. Rev. Lett. 131, 154101.

Petrillo, G., Zhuang, J., 2024. Bayesian earthquake forecasting approach based on the epidemic type aftershock sequence model. Earth Planets Space 76, 78.

Piegari, E., Herrmann, M., Marzocchi, W., 2022. 3-D spatial cluster analysis of seismic sequences through density-based algorithms. Geophys. J. Int. 230, 2073–2088.

Radiguet, M., Cotte, N., Gualandi, A., Kostoglodov, V., Lhomme, T., Walpersdorf, A., Cabral Cano, E., Campillo, M., 2016. Triggering of the 2014 m w 7.3 papanoa earthquake by a slow slip event in guerrero, mexico. Nat. Geosci. 9, 829–833.

Reasenberg, P., 1985. Second-order moment of central California seismicity, 1969–1982. J. Geophys. Res.: Solid Earth 90, 5479–5495.

Ross, G.J., 2021. Bayesian estimation of the ETAS model for earthquake occurrences. Bull. Seismol. Soc. Am. 111, 1473–1480.

Ross, G.J., Kolev, A.A., 2022. Semiparametric Bayesian forecasting of SpatioTemporal earthquake occurrences. Ann. Appl. Stat. 16, 2083–2100. http://dx.doi.org/10.1214/21-AOAS1554.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.

Scarfì, L., Barreca, G., Cassisi, C., Langer, H., 2025. Earthquake clustering and structural modelling unravel volcano-tectonic complexity beneath Mount Etna. Sci. Rep. 15, 18607.

Schorlemmer, D., Woessner, J., 2008. Probability of detecting an earthquake. Bull. Seismol. Soc. Am. 98, 2103–2117.

Shcherbakov, R., Kothari, S., 2025. Earthquake declustering using supervised machine learning. Bull. Seismol. Soc. Am.

Storchak, D.A., Di Giacomo, D., Bondár, I., Engdahl, E.R., Harris, J., Lee, W.H., Villaseñor, A., Bormann, P., 2013. Public release of the isc–gem global instrumental earthquake catalogue (1900–2009). Seismol. Res. Lett. 84, 810–815.

Taroni, M., 2024. Are the magnitudes of earthquakes in Southern California, with incompleteness removed, correlated? Geophys. J. Int. 236, 1596–1600.
Taroni, M., Console, R., Montuori, C., Murru, M., Falcone, G., Chiaraluce, L., Pastoressa, A.E., 2024. Statistically significant difference between earthquake size distributions of independent and triggered seismicity. Commun. Earth & Environ. 5, 193.
Taroni, M., Petrillo, G., Lippiello, E., 2025. Earthquake size distributions of strong worldwide seismicity are similar for background and triggered events. Seismol. Res. Lett. http://dx.doi.org/10.1785/0220240481.
Uhrhammer, R., 1986. Characteristics of northern and central California seismicity. Earthq. Notes 57, 21–37.
Utsu, T., Ogata, Y., et al., 1995. The centenary of the omori formula for a decay law of aftershock activity. J. Phys. Earth 43, 1–33.
Widiyantoro, S., Supendi, P., Rawlinson, N., Daryono, M., Rosalia, S., 2024. A note on the seismicity of Sumatra, Western Sunda arc, Indonesia, in relation to the potential for back-arc thrusting. Sci. Rep. 14, 13115.
Wiemer, S., Wyss, M., 2000. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the Western United States, and Japan. Bull. Seismol. Soc. Am. 90, 859–869.
Zaliapin, I., Ben-Zion, Y., 2013. Earthquake clusters in Southern California i: Identification and stability. J. Geophys. Res.: Solid Earth 118, 2847–2864.
Zaliapin, I., Gabrielov, A., Keilis-Borok, V., Wong, H., 2008. Clustering analysis of seismicity and aftershock identification. Phys. Rev. Lett. 101, 018501.
Zhuang, J., Ogata, Y., Vere-Jones, D., 2002. Stochastic declustering of space-time earthquake occurrences. J. Amer. Statist. Assoc. 97, 369–380.