



A machine learning diagnostic test for nonstationary ETAS aftershock models

Giuseppe Petrillo^{a,b,c}, Luca Dal Zilio^{a,b,*}

^a Earth Observatory of Singapore, Nanyang Technological University, 50 Nanyang Ave, Singapore, 639798, Singapore

^b Asian School of the Environment, Nanyang Technological University, 50 Nanyang Ave, Singapore, 639798, Singapore

^c CNRS, ENS de Lyon LPENSL, UMR5672, Lyon Cedex 07, Lyon, 69342, France

ARTICLE INFO

Editor: Dr H Thybo

Keywords:

Statistical seismology
Aftershock sequences
ETAS model
Nonstationary modelling
Anomaly detection
Unsupervised learning

ABSTRACT

Nonstationary extensions of the Epidemic Type Aftershock Sequence (ETAS) model, in which the background rate and triggering productivity vary in time, are increasingly used to infer transient post-seismic processes from earthquake catalogs. Their flexible parameterization, however, raises an identifiability problem: an improved statistical fit may stem from parameter trade-offs rather than from a genuine change in the seismic regime. We present a diagnostic that pairs penalized nonstationary ETAS inference with an independent, model-agnostic anomaly-detection analysis, and ask whether inferred parameter transients are accompanied by observable changes in catalog properties. We apply it to the aftershock sequence of the 2025 M_w 7.1 Dingri earthquake in southern Tibet, for which independent geodetic and mechanical constraints are available. An empirical-Bayes nonstationary ETAS inversion, with smoothing weights selected by ABIC, yields time-varying background rate and triggering productivity, with the strongest variations about 8–10 days after the mainshock. Using an Isolation Forest applied to temporal, magnitude, depth, and spatial features, we find no coherent anomaly during this interval. Semi-synthetic experiments that perturb only the temporal occurrence rate, consistent with the temporal ETAS formulation, together with feature-ablation tests, show that the detector responds primarily to temporal-rate changes and recovers the injected transient when present. The absence of a corresponding anomaly in the real catalog indicates that the inferred parameter variations should be interpreted with caution. More broadly, the results show the value of validating nonstationary ETAS inferences against independent, model-agnostic diagnostics.

1. Introduction

Aftershock sequences are among the most information-rich records of fault mechanics and post-seismic stress evolution available from seismological observations. The temporal and spatial organization of aftershocks reflects a complex interplay between coseismic static stress changes, dynamic stress perturbations from passing seismic waves, and a spectrum of post-seismic processes including afterslip (Avouac, 2004; Lippiello et al., 2019; Petrillo et al., 2020), viscoelastic relaxation (Jagla et al., 2014a; Jagla, 2014b), and pore-fluid pressure diffusion (Wenousky, 2006; Wang et al., 2012; Freed, 2005; Gentili et al., 2024). Understanding how these processes govern aftershock occurrence carries direct implications for post-mainshock seismic hazard assessment and short-term probabilistic forecasting (Gerstenberger et al., 2005; Falcone et al., 2020; Mizrahi et al., 2024). Statistical models of aftershock clustering play a central role in this effort, providing the quantitative framework through which observable seismicity patterns are translated

into estimates of time-dependent hazard and constraints on fault system behaviour.

The Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988, 1998) has become the standard statistical tool for characterizing earthquake clustering. In this self-exciting point process formulation, the conditional seismicity rate is decomposed into a stationary background term and a sum of triggering contributions from all past events, with productivity scaling exponentially with magnitude and temporal decay governed by the Omori-Utsu law (Utsu et al., 1995). ETAS models have been applied across a broad range of tectonic environments, from continental and subduction-zone mainshock-aftershock sequences (Ogata, 1988; Helmstetter and Sornette, 2003; Zhuang et al., 2002; Petrillo and Zhuang, 2024; Schneider et al., 2025) to earthquake swarms (Hainzl and Ogata, 2005; Petrillo and Taroni, 2025; Petrillo et al., 2026) and induced seismicity (Shcherbakov, 2024). Beyond descriptive characterization, ETAS has also been used for short-term earthquake forecasting, beginning with the pioneering work of Console et al. (2003), and

* Corresponding author.

E-mail address: luca.dalzilio@ntu.edu.sg (L. Dal Zilio).

<https://doi.org/10.1016/j.epsl.2026.120308>

Received 23 April 2026; Received in revised form 19 August 2026; Accepted 21 August 2026

Available online 3 September 2026

0012-821X/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

later forming the basis of several operational forecasting approaches (Gerstenberger et al., 2005; Mizrahi et al., 2024), and its parameters — background seismicity rate, triggering productivity, and temporal decay exponents — serve as effective macroscopic descriptors of the seismic response to stress perturbations (Helmstetter and Sornette, 2003; Marsan and Lengliné, 2008).

The classical ETAS formulation assumes that the background rate and triggering parameters remain stationary throughout the observation window. This assumption is likely violated in the dynamic environment that follows a large mainshock, where transient post-seismic processes — accelerated afterslip, pore-pressure diffusion, and progressive viscoelastic relaxation — can modulate background seismicity levels over time (Marsan and Lengliné, 2008; Helmstetter and Sornette, 2003). Nonstationary extensions of the ETAS model, in which selected parameters are allowed to vary continuously with time, have been developed to address this limitation (Kumazawa and Ogata, 2013; Hainzl and Ogata, 2005). Such formulations have proved valuable for detecting fluid-triggered swarm sequences (Hainzl and Ogata, 2005; Petrillo et al., 2024), characterizing the temporal evolution of post-seismic seismicity (Kumazawa and Ogata, 2013), and identifying transient increases in background activity that may signal evolving physical forcing. In principle, time-varying ETAS parameters can serve as indirect statistical proxies for physical processes that are difficult to observe directly.

Despite this interpretive appeal, nonstationary ETAS models face a fundamental identifiability challenge. The background rate μ and triggering productivity K are structurally coupled in the ETAS conditional intensity, and time-varying parameterizations can redistribute likelihood gain between these terms without implying any observable change in the seismicity. Information-theoretic criteria such as the Akaike Information Criterion (AIC) or its Bayesian counterpart (ABIC) penalize model complexity, but they do not distinguish between genuine physical transients and statistical artifacts of a flexible parameterization (Kumazawa and Ogata, 2013). The risk of overinterpreting flexible statistical models — attributing physical significance to parameter variations that may reflect internal trade-offs, spline flexibility, or residual catalog incompleteness effects — has received limited systematic attention in the ETAS literature. This represents a critical methodological gap. As nonstationary ETAS formulations are increasingly applied to infer post-seismic processes, robust protocols are needed to distinguish physically meaningful nonstationarity from model-induced artifacts, and to clarify under which conditions improved likelihood can be interpreted as evidence of genuine structural change in seismicity.

A natural approach to closing this gap is to complement parametric ETAS inference with diagnostic tools that operate independently of the statistical model used for fitting. Observable multivariate properties of an earthquake catalog — including seismicity rate, magnitude statistics, depth distribution, and spatial dispersion — constitute a model-agnostic basis for detecting genuine structural changes in a seismic sequence. Unsupervised anomaly detection algorithms, which identify outliers in high-dimensional feature spaces without requiring prior labeling or assumed distributions, offer an effective complement to parametric inference: if inferred parameter transients correspond to a genuine change in the physical seismic regime, they should leave a detectable imprint in directly observable catalog statistics. Conversely, the absence of any coherent multivariate anomaly coincident with inferred parameter variations provides independent evidence that the apparent nonstationarity may be a statistical artifact. To our knowledge, such cross-validation between model-based inference and model-agnostic diagnostics has not been systematically developed within the ETAS framework.

The aftershock sequence of the 2025 M_w 7.1 Dingri earthquake in southern Tibet provides an ideal case study for developing and testing this framework. The earthquake struck Dingri County on 7 January 2025 within the Himalayan-Tibetan orogen, a region of active north-south extension and east-west shortening driven by the ongoing convergence of the Indian and Eurasian plates (Armijo et al., 1986; Dal Zilio et al., 2021). Focal mechanism solutions and geodetic analy-

ses consistently identify northward unilateral rupture on a west-dipping normal fault — the Dengme Co Fault, marking the eastern boundary of the Dengme Co Graben within the southern segment of the Shenzha-Dingjie rift system (Qiao et al., 2025; Xu et al., 2025; Ma et al., 2026). ALOS-2 line-of-sight surface displacements reveal a pronounced asymmetric deformation pattern, with strong range-increase on the hanging wall to the east and range-decrease to the west, diagnostic of west-dipping normal faulting (Fig. 1(a)). Field investigations reported fault scarps up to 3 m high and widespread en échelon tension cracks, while geodetic and mechanical analyses indicate that coseismic stress perturbations activated more than 50 previously unmapped fault segments and induced an early viscoelastic response in the weak middle-to-lower crust below approximately 20 km depth (Ma et al., 2026; Wei et al., 2026). The earthquake generated an exceptionally rich and spatially coherent aftershock sequence, tightly elongated along the north-south fault trace (Fig. 1(b)), documented through rapidly deployed dense seismic arrays and advanced detection techniques that produced high-resolution catalogs resolving tens of thousands of early aftershocks (Yao et al., 2025). This combination of well-constrained rupture geometry, independent geodetic constraints on post-seismic deformation (Ma et al., 2026; Wei et al., 2026), and catalog richness makes the Dingri sequence particularly well-suited for testing whether statistically inferred nonstationarity can be robustly attributed to post-seismic physical forcing, rather than to model flexibility or parameter trade-offs.

In this study, we present and test a diagnostic framework that integrates penalized nonstationary ETAS inference with model-agnostic unsupervised anomaly detection. We fit both stationary and nonstationary ETAS formulations to the Dingri aftershock catalog and compare their performance using information-theoretic criteria. We then assess, independently of the ETAS model, whether the inferred temporal parameter variations coincide with observable structural changes in the catalog, using an Isolation Forest algorithm (Liu et al., 2008) applied to a multivariate feature representation of the seismicity. To evaluate the sensitivity of the anomaly-detection step, we perform controlled tests on semi-synthetic catalogs with injected temporal-rate perturbations of progressively increasing amplitude. Although the present application is limited to the Dingri sequence, the framework is general in principle and potentially applicable to other aftershock sequences. Systematic evaluation across multiple sequences and tectonic settings will nevertheless be required to assess its robustness and transferability. The paper is organized as follows: Section 2 describes the earthquake catalog and the methods for stationary and nonstationary ETAS modeling; Section 3 presents the model fits and model-agnostic diagnostics; Section 4 reports the semi-synthetic sensitivity tests; and Section 5 discusses the broader implications and conclusions.

2. Data and methods

2.1. Earthquake catalog and magnitude of completeness

We analyze the seismic sequence associated with the 2025 M_w 7.1 Dingri earthquake in southern Tibet, using the catalog of Yao et al. (2025), which provides origin times and magnitudes for all events recorded by the regional seismic network during the sequence (Fig. 1(b), (c)).

A reliable estimate of the magnitude of completeness M_c is a prerequisite for ETAS modeling, since catalog incompleteness below the detection threshold biases both the inferred background rate and the triggering productivity (Helmstetter and Sornette, 2003). We estimate M_c using the coefficient of variation (CV) stabilization method (Godano et al., 2024), in which the CV — defined as the ratio of the standard deviation to the mean of the inter-event magnitude distribution — converges to unity when the Gutenberg-Richter frequency-magnitude relation holds (Davidsen and Green, 2011; Petrillo and Zhuang, 2023). The completeness magnitude is identified as the lowest threshold above

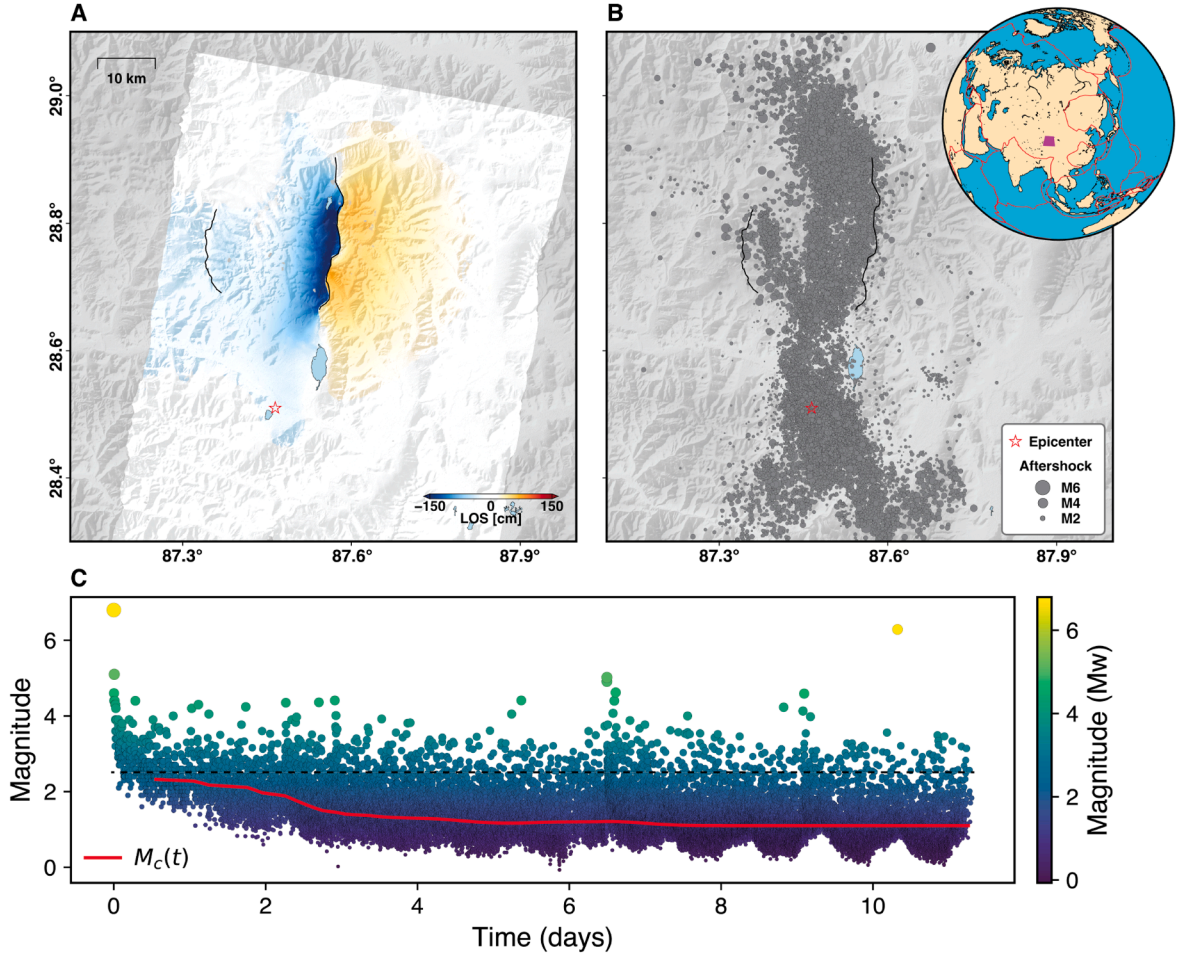


Fig. 1. Seismotectonic context and earthquake catalog for the 2025 Dingri aftershock sequence. (a) ALOS-2 coseismic line-of-sight (LOS) surface displacement. The red star marks the mainshock epicenter; the black line indicates the mapped surface fault trace. (b) Spatial distribution of the aftershock sequence. Gray circles show recorded events with marker size scaled to magnitude. The red star marks the mainshock epicenter. The aftershock cloud is tightly elongated in the north-south direction along the fault system, consistent with the northward unilateral rupture propagation inferred from geodetic and seismic analyses. The inset shows the major plate boundaries (red) and the location of the study area (purple box) within the central Himalaya and southern Tibet. (c) Temporal evolution of the catalog. Events are plotted as a function of time since the mainshock, colored by magnitude. The red line shows the smoothed time-dependent completeness magnitude $M_c(t)$, estimated using the coefficient of variation (CV) stabilization method applied within a sliding window of 250 events. The completeness magnitude threshold ($M_c = 2.5$) is indicated by the black dashed line. All ETAS analyses are restricted to events above the local $M_c(t)$ threshold.

which the estimated b -value stabilizes simultaneously with the CV approaching unity.

Because waveform overlap and coda interference in the immediate post-mainshock phase can suppress the detection of small events, we allow M_c to vary in time. The time-dependent completeness $M_c(t)$ is estimated within a sliding window of 250 events and decreases monotonically as the aftershock rate decays and catalog detectability recovers (De Arcangelis et al., 2018). The time-dependent estimate reaches a maximum value of $\max_t M_c(t) = 2.5$. To treat the earliest post-mainshock interval conservatively, and because the first $M_c(t)$ estimate is only available after the first 250-event window, all ETAS analyses are performed using the constant threshold $M_c = 2.5$. Thus, only events with $M \geq 2.5$ are retained over the entire observation window.

2.2. Stationary ETAS model

We model the temporal clustering of the aftershock sequence using the temporal ETAS model (Ogata, 1988, 1998), in which seismicity is represented as a self-exciting point process: each earthquake can trigger further events, and the rate of occurrence at any instant depends on the full history of past events. The conditional intensity function takes the

form

$$\lambda(t | \mathcal{H}_t) = \mu + \sum_{j: t_j < t} K \exp[\alpha(m_j - m_0)](t - t_j + c)^{-p}, \quad (1)$$

where $\mathcal{H}_t$ denotes the history of events prior to time t , μ is the stationary background seismicity rate, K is the triggering productivity, α controls the scaling of productivity with earthquake magnitude, and c and p are the parameters of the Omori-Utsu temporal decay kernel (Utsu et al., 1995). The reference magnitude m_0 is set equal to the completeness threshold.

The five parameters $\theta = (\mu, K, \alpha, c, p)$ are estimated by maximizing the log-likelihood

$$\log L(\theta) = \sum_{i=1}^N \log \lambda(t_i | \mathcal{H}_{t_i}) - \int_{T_0}^T \lambda(t | \mathcal{H}_t) dt, \quad (2)$$

where $[T_0, T]$ is the observation window and N is the total number of events. Approximate parameter uncertainties are estimated from the observed Fisher information matrix, computed as the inverse of the numerical Hessian of the negative log-likelihood evaluated at the maximum-likelihood solution. Model selection is based on the Akaike Information

Criterion ($AIC = -2 \log L + 2k$, where k is the number of free parameters). Parameter estimation was performed by numerical maximum-likelihood optimization. We note that ETAS likelihood surfaces may exhibit parameter trade-offs and weak identifiability, potentially leading to relatively flat regions near the optimum (Veen and Schoenberg, 2008; Petrillo and Zhuang, 2024; Schneider et al., 2025). These issues are generally more pronounced in high-dimensional space-time ETAS formulations than in purely temporal models such as the one considered here. To assess the robustness of the inversion, the optimization was repeated using different initial parameter values, yielding very similar parameter estimates and log-likelihood values. Model adequacy is assessed through residual analysis using the time-rescaling theorem (Zhuang and Ogata, 2006). Event times are mapped to transformed times

$$\tau_i = \int_{T_0}^{t_i} \lambda(t | H_t) dt, \quad (3)$$

which, for a correctly specified model, should constitute a homogeneous Poisson process with unit rate. Systematic departures of the $\{\tau_i\}$ from this expectation indicate model misspecification.

2.3. Nonstationary ETAS model

The stationary formulation assumes that the background rate μ and triggering productivity K remain constant throughout the observation window. In the post-mainshock environment, however, transient physical processes — including accelerated afterslip, pore-pressure diffusion, and viscoelastic stress relaxation — may modulate seismicity at rates that cannot be captured by a single constant background term. To allow for this possibility, we consider a nonstationary extension of the ETAS model in which μ and K are allowed to vary continuously in time, while the Omori-Utsu decay parameters (c, p) and the productivity-magnitude scaling parameter α are held constant. Fixing the decay parameters limits identifiability issues arising from the structural trade-off between K and α , and avoids over-parameterization (Kumazawa and Ogata, 2014).

The conditional intensity of the nonstationary model is

$$\lambda(t | H_t) = \mu(t) + \sum_{j: t_j < t} K(t_j) \exp[\alpha(m_j - M_0)] g(t - t_j), \quad (4)$$

where $g(\Delta t)$ is the Omori-Utsu temporal kernel. In the stationary model, the normalized kernel

$$g(\Delta t) = \frac{(p-1)c^{p-1}}{(\Delta t + c)^p}, \quad \Delta t > 0, \quad (5)$$

is used. In the nonstationary implementation, we adopt the unnormalized form $g(\Delta t) = (\Delta t + c)^{-p}$ and absorb the normalization constant into $K(t)$; this reparameterization preserves the analytic tractability of the integral of the triggering term while leaving inference on temporal variations of $K(t)$ unchanged.

The time dependence of μ and K is parameterized via multiplicative anomaly factors relative to constant reference values ($\mu_{\text{ref}}, K_{\text{ref}}$),

$$\mu(t) = \mu_{\text{ref}} q_\mu(t), \quad K(t) = K_{\text{ref}} q_K(t), \quad (6)$$

following the approach of Kumazawa and Ogata (2014). Following this empirical-Bayes formulation, the reference values ($\mu_{\text{ref}}, K_{\text{ref}}$) and the parameters (c, α, p) are held fixed at the stationary maximum-likelihood solution (the kernel normalization being absorbed into K_{ref}), so that only the two anomaly functions $q_\mu(t)$ and $q_K(t)$ are estimated in the nonstationary inversion. Fixing the stationary parameters provides common reference scales for the nonstationary inversion and prevents additional trade-offs associated with jointly re-estimating c, α , and p . The reference values μ_{ref} and K_{ref} , however, are multiplicative scaling constants rather than temporal means or centering constraints. Because non zero-mean constraint is imposed on $s_\mu(t)$ or $s_K(t)$, their constant components may shift the effective levels of $\mu(t)$ and $K(t)$ relative to the corresponding stationary estimates. To enforce positivity and enable smooth, data-driven

inference, the anomaly factors are expressed as exponentials of cubic spline functions,

$$q_\mu(t) = \exp(s_\mu(t)), \quad q_K(t) = \exp(s_K(t)), \quad (7)$$

where $s_\mu(t)$ and $s_K(t)$ are represented by natural cubic interpolating splines defined on six equally spaced knots spanning the analysis interval $[T_0, T]$, including both endpoints. The knot locations are

$$t_k = T_0 + \frac{k-1}{5}(T - T_0), \quad k = 1, \dots, 6.$$

The estimated coefficients correspond to the values of each log-anomaly function at the six knots. The two spline functions therefore contribute six free coefficients each, for a total of 12 estimated spline coefficients. Between adjacent knots, each function is represented by a piecewise cubic polynomial with continuous first and second derivatives. Natural boundary conditions are imposed at the endpoints,

$$s''_\mu(T_0) = s''_\mu(T) = 0, \quad s''_K(T_0) = s''_K(T) = 0.$$

No additional constraints are imposed on the spline coefficients. Positivity of the anomaly factors is ensured by the exponential transformation in Eq. (7). The same knots and spline parameterization are used for every pair of smoothing weights in the ABIC search.

2.3.1. Likelihood

The log-likelihood of the nonstationary model has the same form as Eq. (2), with λ now given by Eq. (4):

$$\log L(\theta) = \sum_{i=1}^N \log \lambda(t_i | H_{t_i}) - \int_{T_0}^T \lambda(t | H_t) dt. \quad (8)$$

The integral of the intensity function is decomposed into background and triggering contributions. The background integral is evaluated numerically on a fixed time grid using trapezoidal quadrature:

$$\int_{T_0}^T \mu(t) dt = \mu_{\text{ref}} \int_{T_0}^T q_\mu(t) dt. \quad (9)$$

The triggering integral admits a closed-form expression by exploiting the analytic antiderivative of $(t - t_j + c)^{-p}$:

$$\int_{T_0}^T \sum_{j: t_j < t} K(t_j) e^{\alpha(m_j - M_0)} (t - t_j + c)^{-p} dt = \sum_{j=1}^N K(t_j) e^{\alpha(m_j - M_0)} I_j, \quad (10)$$

with

$$I_j = \begin{cases} \frac{(T - t_j + c)^{1-p} - c^{1-p}}{1-p}, & p \neq 1, \\ \log\left(\frac{T - t_j + c}{c}\right), & p = 1. \end{cases} \quad (11)$$

2.3.2. Penalized inference and smoothing selection

The spline representation introduces a large number of degrees of freedom and therefore requires regularization to avoid overfitting local fluctuations in the earthquake catalog. Separate roughness penalties are imposed on the background-rate and productivity anomaly functions through the integrated squared curvature of their logarithms,

$$\mathcal{R}(w_\mu, w_K) = w_\mu \int_{T_0}^T [s''_\mu(t)]^2 dt + w_K \int_{T_0}^T [s''_K(t)]^2 dt, \quad (12)$$

where $w_\mu \geq 0$ and $w_K \geq 0$ are independent smoothing hyperparameters that control the regularization of $s_\mu(t)$ and $s_K(t)$ separately. The spline coefficients are obtained by maximizing the penalized log-posterior

$$Q(\theta) = \log L(\theta) - \frac{1}{2} \mathcal{R}(w_\mu, w_K), \quad (13)$$

with ($\mu_{\text{ref}}, K_{\text{ref}}, c, \alpha, p$) fixed at the stationary maximum-likelihood solution and the curvature integrals in Eq. (12) evaluated numerically on a dense grid. Eq. (12) is equivalent to a zero-mean Gaussian smoothness prior $\pi(\theta | w_\mu, w_K) \propto \exp[-\frac{1}{2} \theta^T \Lambda(w_\mu, w_K) \theta]$ on the spline coefficients,

where the precision matrix Λ collects the discretized curvature operators scaled by w_μ and w_K .

Following Kumazawa and Ogata (2014), the smoothing hyperparameters are selected by minimizing the Akaike Bayesian Information Criterion (ABIC), an empirical-Bayes criterion defined through the marginal likelihood

$$\text{ABIC}(w_\mu, w_K) = -2 \log \int L(\theta) \pi(\theta | w_\mu, w_K) d\theta + 2 n_{\text{hyper}}, \quad (14)$$

where $n_{\text{hyper}} = 2$ is the number of smoothing hyperparameters. For each pair (w_μ, w_K) the marginal likelihood is approximated by a Laplace expansion about the maximum-a-posteriori estimate $\hat{\theta}$, which yields the explicit, directly computed form

$$\begin{aligned} \text{ABIC}(w_\mu, w_K) = & -2 \log L(\hat{\theta}) + \mathcal{R}(\hat{\theta}) - \log \det^+ \Lambda(w_\mu, w_K) \\ & + \log \det H(\hat{\theta}) + 2 n_{\text{hyper}}, \end{aligned} \quad (15)$$

where $H(\hat{\theta})$ is the Hessian of the penalized negative log-posterior $-Q$ at $\hat{\theta}$ (w -independent additive constants are omitted). Here, $\det^+$ denotes the generalized determinant of the positive-semidefinite spline precision matrix. The constant and linear components of a natural cubic spline have zero curvature and therefore span a two-dimensional null space of each roughness matrix. These components were retained as unpenalized coefficients, whereas only the positive eigenvalues of the roughness matrices entered the prior-normalization term. Eigenvalues were retained when $\rho_j > \max(10^{-10}, 10^{-8} \max_j |\rho_j|)$. With six coefficients for each spline, R_μ and R_K each have rank 4 and nullity 2; hence the combined 12×12 matrix $\Lambda = \text{blockdiag}(w_\mu R_\mu, w_K R_K)$ has rank 8 and nullity 4 for positive smoothing weights. No ridge regularization was applied to Λ . Thus,

$$\log \det^+ \Lambda = \sum_{\rho_{\mu,j} > 0} \log(w_\mu \rho_{\mu,j}) + \sum_{\rho_{K,j} > 0} \log(w_K \rho_{K,j}),$$

and this quantity enters Eq. (15) through the prior-normalization term $-\log \det^+ \Lambda$. The unpenalized components remain included in the full posterior Hessian $H(\hat{\theta})$. The prior-normalization term $-\log \det^+ \Lambda$ and the posterior-Hessian term $+\log \det H$ jointly account for the effective flexibility of the spline representation. Here, $\det^+$ denotes the generalized determinant obtained from the positive eigenvalues of the roughness-penalty precision matrix, with its null-space components excluded. These two terms balance improvement in the penalized likelihood against the effective complexity associated with each pair of positive smoothing weights. The hyperparameters were explored on the grid $w_\mu, w_K \in \{10^{-3}, 5 \times 10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}, 5 \times 10^{-1}, 1, 5, 10\}$, spanning four orders of magnitude, and the pair yielding the minimum ABIC was retained.

2.4. Model-agnostic anomaly detection

To assess independently whether statistically inferred parameter transients correspond to genuine structural changes in the seismicity, we complement the ETAS analysis with a model-agnostic unsupervised anomaly detection procedure. This step does not use the ETAS model in any way: it operates directly on observable catalog properties and provides an external consistency check on the interpretation of time-varying parameters.

The earthquake catalog is partitioned into consecutive non-overlapping time windows of duration $\Delta t = 0.5$ days. For each window k we construct a $p = 13$ -dimensional feature vector $\mathbf{x}_k$ comprising: (1) event count N_k and seismicity rate $N_k/\Delta t$; (2) magnitude statistics: mean magnitude, maximum magnitude, and fraction of events with $M \geq 4.5$; (3) depth statistics: mean, median, and standard deviation of focal depths; (4) spatial statistics: mean latitude and longitude, their standard deviations, and a spatial dispersion proxy $\sigma_{\text{lat}} \sigma_{\text{lon}}$. Although the event count and seismicity rate are linearly related for a fixed bin width, both quantities are retained as a simple consistency check within the feature set. We verified that removing the rate feature produces negligible changes in the resulting anomaly scores. The threshold $M = 4.5$

was selected as a pragmatic indicator of the occurrence of relatively large events while retaining sufficient statistics within individual temporal bins.

This representation transforms the point process into a multivariate time series $\{\mathbf{x}_k\}$ that captures the macroscopic evolution of the seismic sequence across multiple observable dimensions simultaneously.

To reduce the influence of heavy-tailed fluctuations, each feature is standardized using robust median absolute deviation (MAD) normalization,

$$z_k^{(j)} = \frac{0.6745 (x_k^{(j)} - \text{median}(x^{(j)}))}{\text{MAD}(x^{(j)})}, \quad (16)$$

yielding a robustly standardized feature matrix $\mathbf{Z} \in \mathbb{R}^{K \times p}$.

Anomaly detection is performed using the Isolation Forest algorithm (Liu et al., 2008), an ensemble method based on random recursive partitioning of the feature space. Each tree in the ensemble isolates individual observations by randomly selecting a feature dimension and a random split value. Anomalous observations, which occupy sparse regions of the feature space, tend to require fewer splits to isolate and therefore have shorter average path lengths. For an ensemble of $T = 500$ trees, the anomaly score is

$$s(\mathbf{z}_k) = 2^{-E[h(\mathbf{z}_k)]/c(K)}, \quad (17)$$

where $E[h(\mathbf{z}_k)]$ is the mean path length over trees and $c(K) = 2H_{K-1} - 2(K-1)/K$ is a normalization constant based on the harmonic number H_{K-1} . Higher scores indicate greater anomalousness. We use the `scikit-learn` implementation of the Isolation Forest with $T = 500$ trees (`n_estimators=500`), subsampling `max_samples=256`, automatic contamination (`contamination='auto'`), all features considered at each split (`max_features=1`), and a fixed random seed (`random_state=42`) for reproducibility. In practice we report the detector output through the `scikit-learn` decision function, which is an affine, monotonically decreasing transform of the path-length score of Eq. (17) centered so that the average bin has a value near zero. To preserve the convention that larger values denote stronger anomalies, we plot the negative of the decision function. Consequently, negative plotted values do not correspond to “negative anomalies”; they simply identify bins that are more typical of the catalog than the average bin.

The resulting anomaly score time series $\{s_k\}$ provides a model-independent indicator of multivariate departures from the typical seismic regime. If the nonstationary ETAS model captures genuine transient processes, elevated anomaly scores are expected in the same temporal interval where $\mu(t)$ or $K(t)$ exhibit strong variations. Conversely, the absence of a coherent anomaly in that interval would indicate that the inferred parameter variations are not associated with a macroscopic change in observable catalog properties, and may instead reflect parameter trade-offs or statistical overfitting.

3. Results

Fig. 1(c) shows the temporal evolution of earthquake magnitudes during the Dingri sequence together with the estimated time-dependent completeness magnitude $M_c(t)$. Immediately after the main-shock, M_c reaches values up to 2.5, reflecting strong waveform overlap and reduced detectability in the early post-seismic phase. As the sequence evolves, $M_c(t)$ decreases progressively as aftershock rates decay and catalog detectability recovers, consistent with the transient detection effects described by De Arcangelis et al. (2018). All subsequent ETAS analyses are restricted to events exceeding the local $M_c(t)$ threshold.

3.1. ETAS model fits

The stationary ETAS model provides a good description of the temporal clustering observed during the Dingri sequence. Maximum likelihood estimation yields the following parameter values: background rate $\mu =$

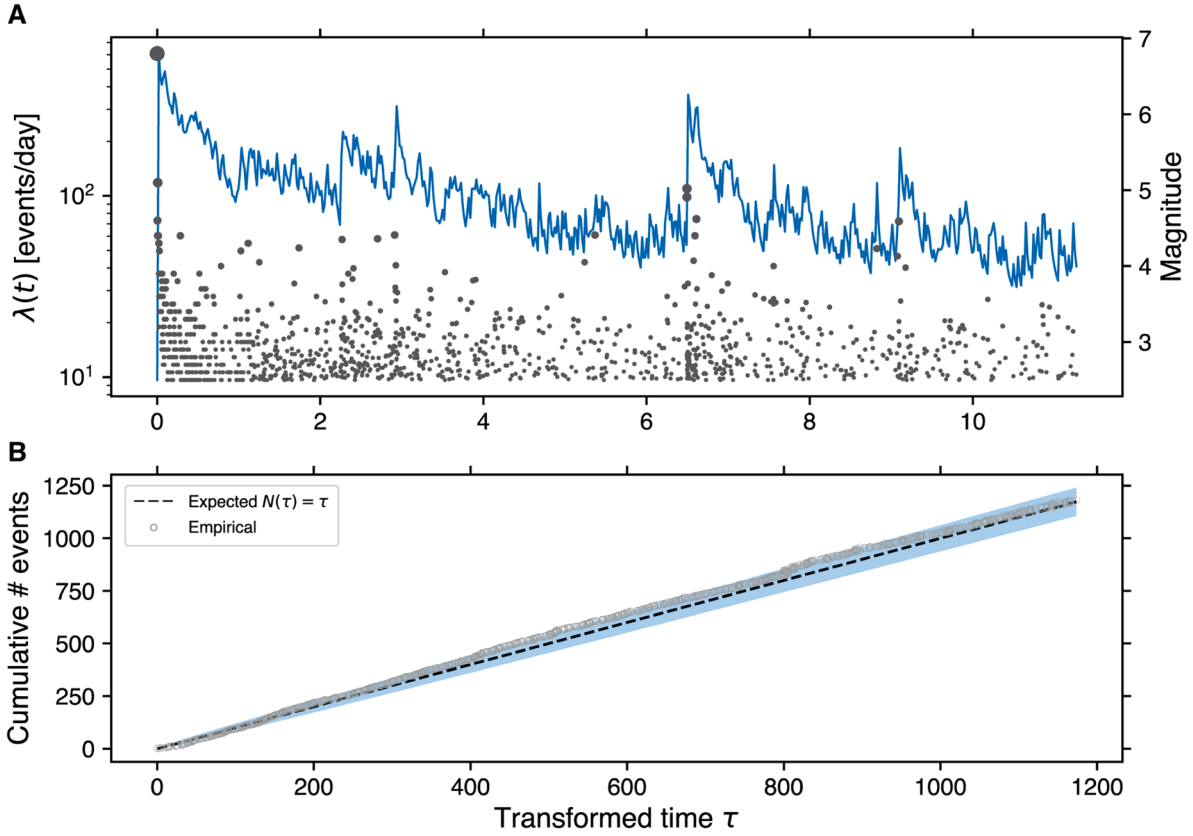


Fig. 2. Stationary ETAS model fit and residual diagnostics for the Dingri sequence. (a) Maximum-likelihood conditional intensity $\lambda(t)$ (blue line, left axis, logarithmic scale) as a function of time since the mainshock. The intensity was obtained from a single stationary ETAS model fitted over the entire observation window using all events satisfying $M \geq M_c(t)$ and was evaluated on a dense temporal grid using the full event history available prior to each evaluation time. Gray circles show earthquake occurrence times, with marker size proportional to magnitude (right axis). The model captures the sharp post-mainshock seismicity rate increase and its subsequent Omori-Utsu decay. (b) Time-rescaling residual analysis. Black dots show the cumulative event count $N(\tau)$ as a function of transformed time τ ; the dashed line gives the expectation $N(\tau) = \tau$ for a correctly specified model, and the shaded band indicates the approximate 95% confidence envelope for a unit-rate Poisson process. The close agreement indicates that the stationary ETAS model provides an adequate description of the temporal clustering.

9.639 ± 2.73 events/day, productivity parameter $K = 0.0495 \pm 0.0125$, Omori-Utsu parameter $c = 0.0209 \pm 0.0114$ days, magnitude scaling exponent $\alpha = 0.97 \pm 0.084$, and temporal decay exponent $p = 1.42 \pm 0.192$, with $\text{AIC}_{\text{stat}} = -9074.4$. The value $p > 1$ indicates a rapid decay of aftershock activity following the mainshock, while the relatively large value of c suggests a delayed onset of Omori-Utsu decay compared to typical aftershock sequences (Utsu et al., 1995). The productivity parameter K is notably elevated, pointing to efficient secondary triggering throughout the sequence.

The stationary ETAS parameters were estimated once over the entire observation window using all events satisfying the time-dependent completeness criterion $M \geq M_c(t)$. The conditional intensity shown in Fig. 2(a) was subsequently evaluated on a dense temporal grid using this single maximum-likelihood parameter set and the complete event history available prior to each evaluation time. No moving-window, segmented, or piecewise ETAS inversions were performed. Fig. 2(a) shows the resulting conditional intensity $\lambda(t)$ from the fitted stationary model. The model captures the sharp seismicity rate increase immediately following the mainshock and its subsequent gradual decay, consistent with Omori-Utsu behavior. The background rate μ contributes substantially to the total intensity throughout the observation window, reflecting a level of sustained seismic activity that exceeds what can be attributed to direct aftershock triggering alone. Residual analysis confirms model adequacy: the transformed event times $\{\tau_i\}$ are consistent with a homogeneous Poisson process with unit rate, with no systematic departures from the expected exponential distribution of inter-event intervals (Fig. 2(b)).

We next fit the nonstationary ETAS model described in Section 2.3, allowing both $\mu(t)$ and $K(t)$ to vary through log-spline anomaly factors while holding (c, α, p) constant.

ABIC minimization over the explored two-dimensional smoothing grid selects finite smoothing weights,

$$w_\mu = 0.01, \quad w_K = 0.01,$$

indicating that the data support a moderately regularized representation of both anomaly functions rather than the unsmoothed solution obtained in the original implementation. The corresponding minimum value of the criterion is $\text{ABIC}_{\text{min}} = -9115.4$. The resulting conditional intensity (Fig. 3(a)) reproduces the overall post-mainshock evolution of the sequence while remaining substantially smoother than the unconstrained estimate of the original submission. The inferred background-rate anomaly function (Fig. 3(c)) is largest immediately after the mainshock and decreases rapidly during the first few days of the sequence. Around 6–8 days after the mainshock, the model infers a modest increase in the productivity anomaly function, coinciding with a secondary burst of seismic activity visible in the catalog. Subsequently, $\mu(t)$ exhibits a moderate increase between approximately 8 and 10 days after the mainshock, although this feature is not associated with a comparably pronounced seismic activation and therefore remains difficult to interpret physically. Time-rescaling diagnostics confirm that the nonstationary representation provides an adequate fit across the full observation window (Fig. 3(b)).

Because the nonstationary solution is obtained by penalized maximum-a-posteriori inference rather than ordinary maximum likelihood

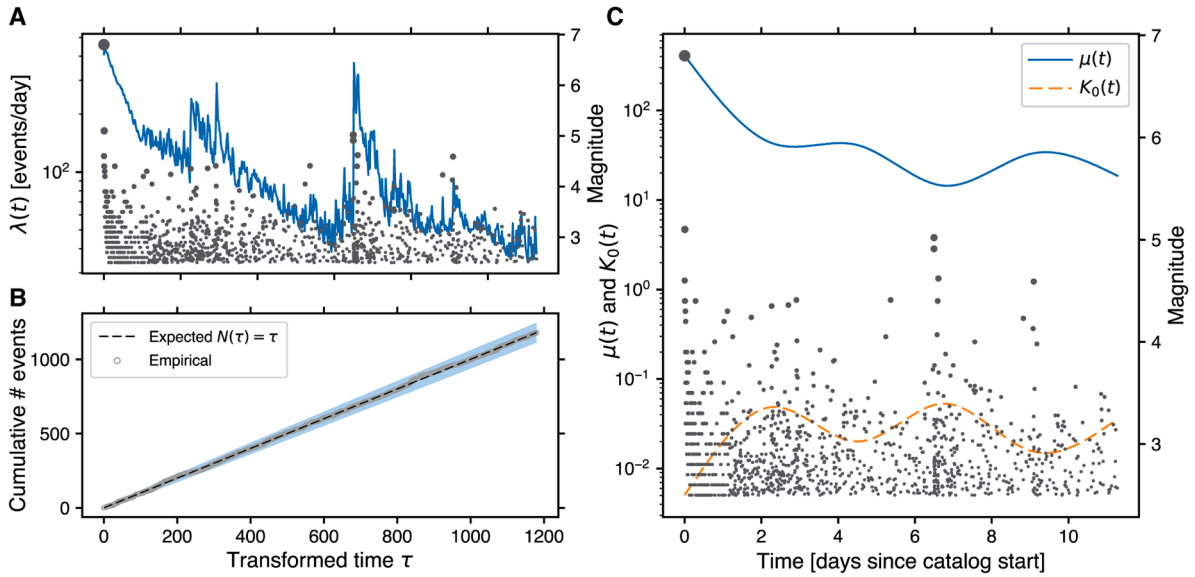


Fig. 3. Nonstationary ETAS model results for the Dingri sequence. (a) Conditional intensity $\lambda(t)$ (blue line, logarithmic scale) as a function of time, with the earthquake catalog in the background (gray circles, marker size proportional to magnitude). The pronounced short-timescale variability reflects the combined contribution of time-varying background rate and triggering productivity. (b) Time-rescaling residual analysis, as in Fig. 2(b), confirming adequate model fit across the full observation window. (c) Estimated time-dependent background rate $\mu(t)$ (blue solid line) and productivity parameter $K_0(t)$ (orange dashed line), both on a logarithmic scale, as a function of time since catalog start. $\mu(t)$ decreases rapidly over the first days following the mainshock, while $K_0(t)$ increases during the early sequence before evolving more slowly. The seismic catalog is shown in the background for reference. The stationary values μ_{ref} and K_{ref} are fixed multiplicative reference scales rather than centering constraints; consequently, the time-dependent estimates are not required to fluctuate around or intersect their stationary counterparts.

hood, a formal AIC comparison with the stationary model is not strictly justified, and we do not report one. Both models yield transformed-time residuals consistent with a unit-rate Poisson process (Figs. 2(b) and 3(b)). The nonstationary formulation has, by construction, additional temporal flexibility and therefore at least as much descriptive capacity as the stationary model; whether this added flexibility corresponds to a genuine, physically distinct change in the seismic regime is precisely the question addressed by the independent diagnostic of the following subsection.

As discussed in Section 2.3, an improved fit achieved through increased model flexibility does not in itself establish that the inferred parameter variations reflect genuine physical changes in the seismic regime. Parameter trade-offs and identifiability limitations can produce apparent nonstationarity even in the absence of external physical forcing, and the following subsection addresses this ambiguity through independent, model-agnostic diagnostics.

3.2. Model-agnostic anomaly detection

We apply the Isolation Forest procedure described in Section 2.4 to the Dingri catalog to assess whether the parameter transients inferred by the nonstationary ETAS model are accompanied by observable structural changes in the seismicity.

Fig. 4 shows the anomaly score time series as a function of time since the mainshock. The interval between 8 and 10 days – where the nonstationary ETAS model infers the most pronounced variations in $\mu(t)$ – does not exhibit any elevated or coherent cluster of anomaly scores. The temporal evolution of the score is relatively smooth across the full observation window, with fluctuations of comparable magnitude distributed throughout the sequence rather than concentrated in the 8–10 day interval.

The heatmap of robust feature z-scores (Fig. 5) corroborates this picture across individual catalog dimensions. Seismicity rate, magnitude statistics, depth distribution, and spatial dispersion all display variability over time, but none of these features undergoes a systematic shift specifically during the 8–10 day window. In particular, the absence of a

coherent, cross-feature anomaly pattern during this period indicates that no distinct seismic regime transition is detectable in first-order catalog statistics.

The discrepancy between the statistical evidence for nonstationarity provided by the ETAS model and the absence of any corresponding multivariate anomaly in directly observable properties suggests that the inferred parameter variations reflect internal parameter trade-offs or the capacity of the spline representation to absorb subtle rate fluctuations, rather than a macroscopic change in the physical seismic regime. We emphasize that this finding does not invalidate the nonstationary ETAS fit: the model remains a valid statistical description of the data. Rather, it demonstrates that a favorable information criterion is not sufficient grounds for attributing physical significance to time-varying ETAS parameters, and that independent diagnostic validation is an essential complementary step.

4. Validation on semi-synthetic catalogs

4.1. Construction of temporal-rate semi-synthetic catalogs

A potential limitation of the anomaly-detection analysis is that the absence of a pronounced anomaly in the Dingri sequence could reflect insufficient detector sensitivity rather than the absence of a genuine structural change. To evaluate this possibility, we constructed a family of semi-synthetic catalogs designed specifically to mimic the type of perturbation that a temporal ETAS model would naturally generate. Starting from the original Dingri catalog, we selected the interval between 8 and 10 days after the mainshock, corresponding to the period where the nonstationary ETAS inversion inferred the largest temporal deviations from the stationary solution. Within this interval, additional synthetic events were generated according to a Poisson process with rate

$$\lambda_{\text{add}} = a \lambda_{\text{obs}}, \quad (18)$$

where λ_{obs} denotes the observed seismicity rate in the selected window and $a = (0.5, 1.0, 1.5, 2.0, 2.5)$ is a prescribed amplification factor controlling the perturbation strength. For each synthetic event, the occurrence

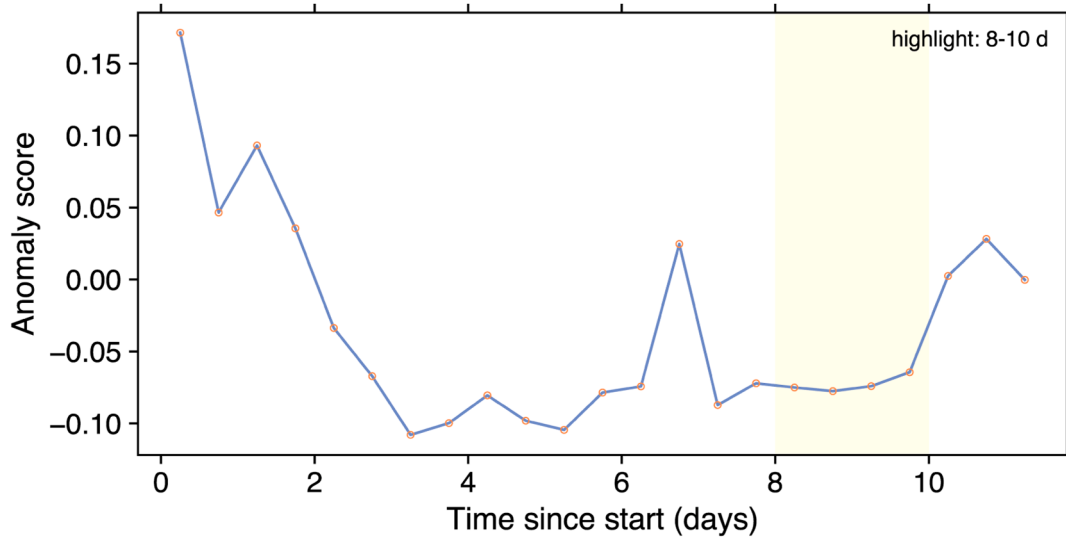


Fig. 4. Model-agnostic anomaly detection applied to the Dingri aftershock sequence using Isolation Forest. The anomaly score is shown as a function of time since the mainshock, computed in non-overlapping temporal bins of width $\Delta t = 0.5$ days; higher values indicate greater multivariate departure from the background seismic regime. The shaded band marks the 8–10 day interval where the nonstationary ETAS model infers the most pronounced variations in $\mu(t)$. Anomaly scores within this window are below the sequence average, with no distinct peak or elevated cluster, indicating that the inferred parameter transients are not accompanied by a detectable structural change in directly observable catalog properties.

time was sampled uniformly within the 8–10 day interval. To avoid introducing artificial anomalies in any other catalog property, magnitudes, focal depths, and epicentral coordinates were inherited from randomly selected real earthquakes already present in the same time window. As a result, the synthetic catalogs differ from the original catalog only through an increase in the temporal occurrence rate. No spatial migration, depth shift, magnitude perturbation, or other structural modification was imposed. Five semi-synthetic realizations (Sim1–Sim5) were generated using progressively larger amplification factors, producing a family of catalogs characterized by increasingly strong temporal-rate perturbations while preserving the original magnitude, depth, and spatial distributions.

4.2. Feature-ablation validation

To identify which catalog properties control the anomaly-detection response, we performed a feature-ablation analysis in which the Isolation Forest was applied independently to different groups of catalog features. Four feature families were considered:

1. Temporal features: event count and seismicity rate.
2. Magnitude features: mean magnitude, maximum magnitude, and fraction of events with $M \geq 4.5$.
3. Depth features: mean depth, median depth, and depth standard deviation.
4. Spatial features: mean latitude and longitude, their standard deviations, and the spatial-dispersion proxy $\sigma_{\text{lat}} \sigma_{\text{lon}}$.

For each feature family, the catalog was divided into non-overlapping windows of duration $\Delta t = 0.5$ days, robustly standardized using median absolute deviation normalization, and analyzed using the same Isolation Forest configuration adopted throughout this study. Fig. 6 compares the resulting anomaly-score evolution for the original catalog and the five semi-synthetic realizations. The strongest response is observed for the temporal-feature subset (Fig. 6(a)). As the perturbation amplitude increases from Sim1 to Sim5, the anomaly score systematically rises within the 8–10 day interval, demonstrating that the Isolation Forest successfully detects localized increases in seismicity rate. By contrast, the magnitude, depth, and spatial feature subsets

(Fig. 6(b)–(d)) exhibit only minor fluctuations and show no systematic increase with perturbation strength. This behaviour is expected because these catalog properties were intentionally preserved during the generation of the semi-synthetic catalogs and therefore do not contain any injected anomaly. The feature-ablation experiment demonstrates that the Isolation Forest is sensitive to the type of perturbation expected from a temporal ETAS model while remaining largely insensitive to feature groups that are unaffected by the synthetic modification. The detector therefore responds specifically to temporal-rate anomalies rather than producing spurious multivariate signals. These results provide an important validation of the anomaly-detection framework. The method successfully identifies temporal-rate perturbations of the same type that would be associated with transient changes in ETAS parameters. Consequently, the absence of a comparable anomaly in the real Dingri catalog cannot be attributed to insufficient detector sensitivity, but instead suggests that the temporal variations inferred by the nonstationary ETAS model are not associated with a macroscopic change in observable catalog properties.

5. Discussion

5.1. Stationary model performance and parameter interpretation

The stationary ETAS model provides a statistically adequate description of the Dingri aftershock sequence, with well-behaved time-rescaling residuals and parameter estimates that are physically interpretable. The estimated temporal decay exponent $p = 1.42$ is above the typical range of 1.0–1.3 reported for continental aftershock sequences (Utsu et al., 1995; Helmstetter and Sornette, 2003), indicating a relatively rapid decay of aftershock activity following the mainshock. The Omori onset parameter $c = 0.0209$ days reflects a slight delay in the onset of power-law decay, which is common in sequences where catalog incompleteness masks the earliest post-mainshock events (De Arcangelis et al., 2018). The elevated background rate $\mu = 9.639 \pm 2.73$ events/day suggests that a substantial fraction of the observed seismicity cannot be attributed to direct aftershock cascades, possibly reflecting the tectonic loading environment of the southern Tibetan rift system and the activity of adjacent fault segments (Armijo et al., 1986; Bischoff and Flesch, 2018). Moreover, this

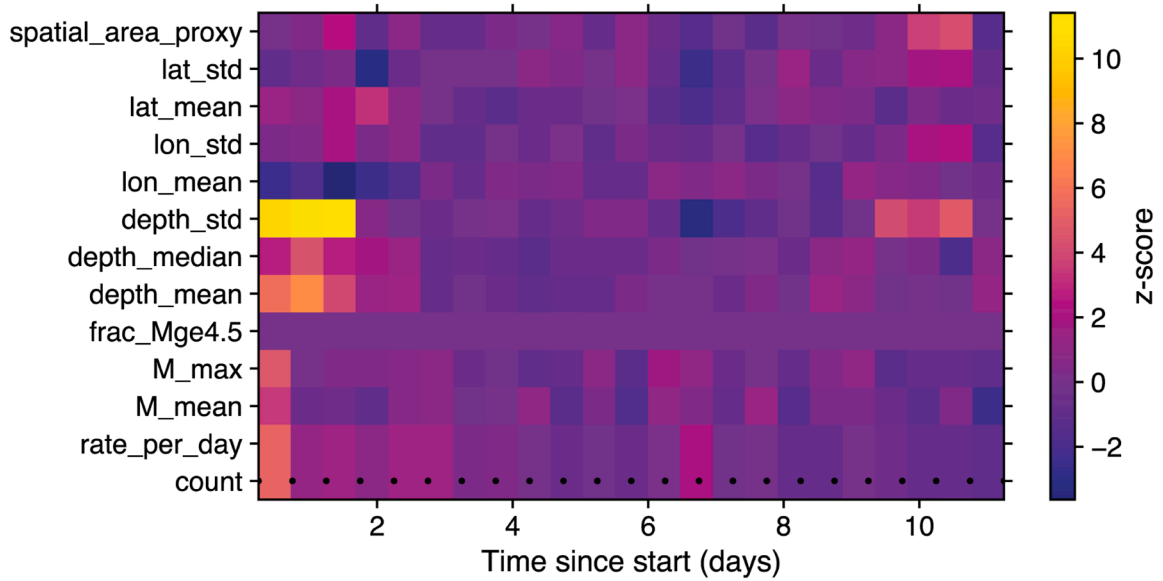


Fig. 5. Temporal evolution of multivariate seismic features used for anomaly detection, expressed as robust z-scores computed using median and MAD normalization within each 0.5-day bin. Rows correspond to the 13 catalog features: event count, seismicity rate, magnitude statistics (mean, maximum, fraction $M \geq 4.5$), depth statistics (mean, median, standard deviation), and spatial dispersion proxies (mean latitude and longitude, their standard deviations, and $\sigma_{\text{lat}} \sigma_{\text{lon}}$). Columns represent successive temporal bins. Warm and cool colors indicate positive and negative deviations from the median background. Elevated z-scores are concentrated in the first two days after the mainshock, particularly in depth variability and spatial spread, consistent with the initial diversity of the early aftershock distribution. No coherent cross-feature anomaly pattern is observed in the 8–10 day interval, supporting the absence of a distinct multivariate regime shift during that period.

value should be interpreted as the effective ETAS background parameter for the analyzed post-mainshock catalog, rather than as an estimate of the long-term tectonic seismicity rate of the region. Because the catalog begins immediately after the mainshock, part of the seismicity assigned to the background component may reflect the well-known trade-off between background and triggering terms in ETAS models.

5.2. Nonstationary model: improved fit or parameter trade-off?

The empirical-Bayes nonstationary ETAS model yields a stable, moderately regularized solution: the ABIC criterion selects finite smoothing weights for both anomaly functions, and the inferred $\mu(t)$ and $K(t)$ vary smoothly in time while leaving the transformed-time residuals consistent with a unit-rate Poisson process. A central question is whether these inferred temporal variations correspond to a physically distinct change in the seismic regime, or whether they instead reflect the intrinsic descriptive flexibility of the spline parameterization and the residual non-uniqueness of the background–triggering decomposition.

Three considerations bear on this question.

First, although the empirical-Bayes ABIC analysis selects finite smoothing weights for both anomaly functions, the inferred temporal variations remain relatively modest and do not correspond to obvious changes in independently observable catalog properties. The introduction of temporal flexibility therefore improves the statistical description of the occurrence process without necessarily identifying a physically distinct seismic phase. This distinction is important because ETAS parameters are effective descriptors of seismicity and need not map uniquely onto individual physical mechanisms.

Second, following Kumazawa and Ogata (2014), the stationary ETAS parameters (μ_{ref} , K_{ref} , c , α , p) are held fixed during the nonstationary inversion, and only the anomaly functions $q_{\mu}(t)$ and $q_K(t)$ are estimated. This choice reduces the trade-offs between the background rate, productivity, and Omori-law parameters that can otherwise arise when all components are allowed to vary simultaneously. Nevertheless, the separation between $q_{\mu}(t)$ and $q_K(t)$ remains an effective statistical decomposition rather than a unique physical attribution, and the resulting temporal variations must therefore be interpreted with caution. To further

assess the potential influence of parameter trade-offs, we performed an additional nonstationary inversion in which the triggering productivity was held fixed and only the background-rate anomaly function was allowed to vary. This choice has been advocated in situations where external forcing processes, such as fluid migration, geothermal forcing, or earthquake swarms, are expected to primarily affect the background seismicity rate rather than the triggering efficiency (Llenos and Michael, 2013, 2016). It also reduces the trade-offs between productivity and background rate that have been discussed in previous ETAS studies (Hainzl et al., 2013). The resulting model provided residual diagnostics comparable to those obtained when both anomaly functions were allowed to vary. This result suggests that the absence of a catalog-scale anomaly is robust with respect to the specific parameterization adopted for the nonstationary ETAS model and is not driven by the simultaneous estimation of both background-rate and productivity anomalies. The approximately antisymmetric evolution of $\mu(t)$ and $K(t)$ in Fig. 3(c) is also consistent with residual covariance between the background and triggering components. A full analysis of the joint uncertainty and covariance of the nonstationary spline coefficients could help quantify this dependence and assess whether the individual parameter variations are statistically distinguishable within the assumed ETAS formulation. The model-agnostic diagnostic proposed here is not intended to replace such an analysis, but instead addresses the complementary question of whether the inferred temporal variations coincide with a detectable change in directly observed catalog properties. Future applications should therefore combine propagation of the full covariance of the nonstationary coefficients with model-agnostic catalog diagnostics.

Third, and most importantly, the model-agnostic Isolation Forest analysis provides an independent external check that is free of these structural trade-offs. The absence of any coherent multivariate anomaly in seismicity rate, magnitude statistics, depth distribution, or spatial dispersion during the 8–10 day interval where $\mu(t)$ varies most strongly indicates that the inferred parameter transient is not accompanied by a macroscopically observable change in the character of the seismicity. The semi-synthetic validation experiments confirm that the detection framework has sufficient sensitivity to recover anomalies of comparable

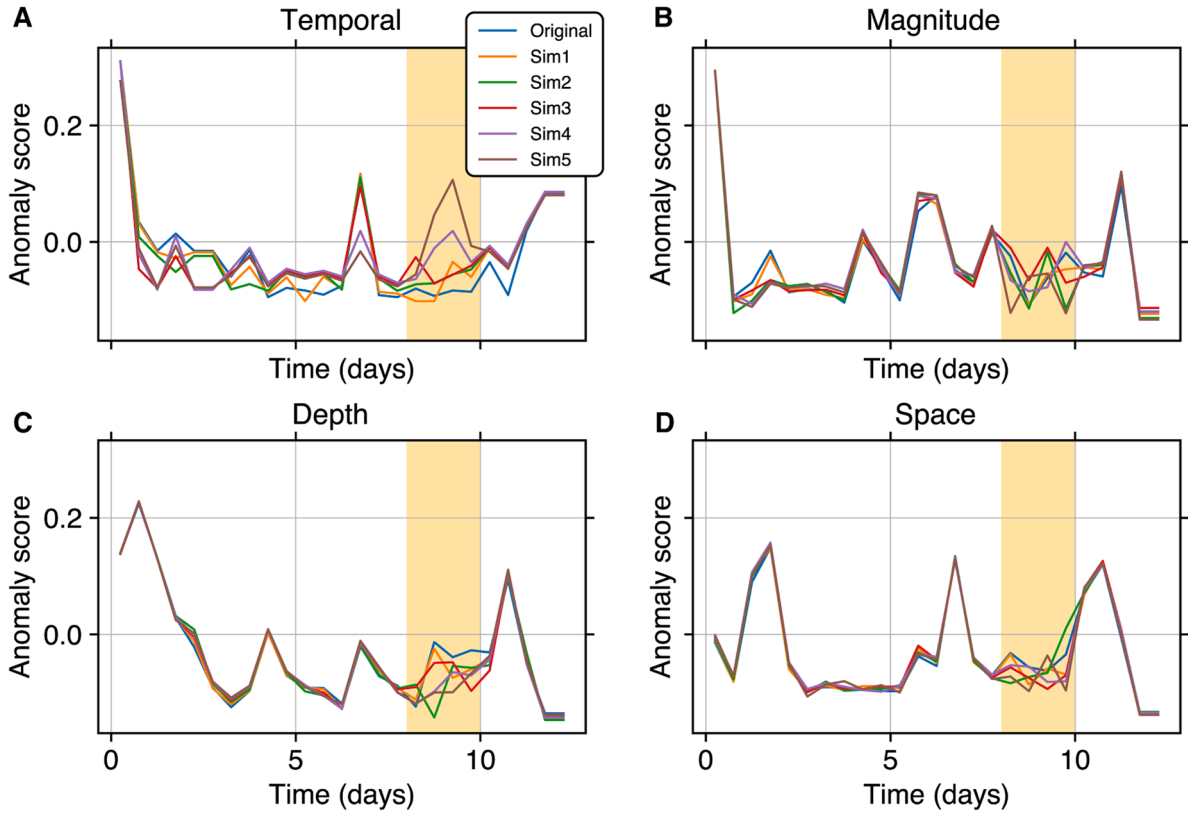


Fig. 6. Feature-ablation analysis of the Isolation Forest anomaly detector. The anomaly score is computed separately using four feature subsets: (a) temporal features, including event count and seismicity rate; (b) magnitude features, including mean magnitude, maximum magnitude, and the fraction of events with $M \geq 4.5$; (c) depth features, including mean, median, and standard deviation of focal depth; and (d) spatial features, including mean latitude and longitude, their standard deviations, and the spatial-dispersion proxy $\sigma_{\text{lat}} \sigma_{\text{lon}}$. The shaded band marks the 8–10 day interval in which the nonstationary ETAS model infers its strongest temporal parameter variations. The original Dingri catalog does not exhibit a coherent anomaly localized within this interval. By contrast, the semi-synthetic catalogs show a systematic increase in anomaly score from Sim1 to Sim5 within the shaded interval, primarily in the temporal-feature subset (a), demonstrating successful detection of the injected temporal-rate perturbation. The magnitude, depth, and spatial subsets (b–d) show no comparable systematic response, as these catalog properties were preserved in the semi-synthetic experiments. The anomaly score is defined as the negative of the `scikit-learn` Isolation Forest decision function, so that larger values indicate more anomalous time bins.

amplitude when they are genuinely present, ruling out methodological insensitivity as an alternative explanation.

Taken together, these considerations indicate that the statistical preference for the nonstationary model should not be interpreted as direct evidence for a physically distinct post-seismic regime. The inferred $\mu(t)$ and $K(t)$ variations are an effective, non-unique statistical description of the occurrence process; the decisive point is that, although the validated detector is sensitive to temporal-rate transients of the relevant type and amplitude, it identifies no corresponding change in any directly observable catalog property during the interval where the ETAS parameters vary most.

5.3. Relation to independent geodetic and mechanical constraints

The Dingri earthquake was accompanied by complex post-seismic deformation, including stress redistribution and transient crustal loading documented through geodetic and mechanical analyses (Ma et al., 2026). Such processes are precisely the kind of physical forcing that nonstationary ETAS models are designed to detect, and the geodetic evidence provides an important independent reference against which the statistical inferences can be evaluated.

The fact that the nonstationary ETAS model does not recover a statistically robust, physically interpretable signal despite the presence of documented post-seismic deformation is itself informative. It suggests that the statistical imprint of transient stress redistribution on the tem-

poral organization of the Dingri aftershock catalog is subtle enough to be indistinguishable from internal model artefacts at the level of temporal clustering statistics alone. This does not mean that post-seismic processes did not influence the seismicity; rather, it highlights a fundamental limitation of temporal ETAS models in resolving low-amplitude physical forcing against the background of statistical uncertainty and parameter trade-offs. Richer model formulations that simultaneously incorporate geodetic constraints or physics-based stress histories — rather than inferring them indirectly from seismicity statistics — may be better positioned to detect such signals (Freed, 2005).

5.4. Implications for nonstationary ETAS modeling practice

The results of this study have direct methodological implications for the growing body of work that uses nonstationary ETAS extensions to infer time-dependent physical processes from earthquake catalogs.

Information-theoretic model selection criteria such as AIC and ABIC are widely used to justify flexible parameterizations, but they evaluate goodness of fit relative to model complexity without assessing whether the inferred variations are physically interpretable or externally observable. Our results show that an ABIC-selected nonstationary ETAS representation can improve the statistical description of the occurrence process while leaving no detectable imprint in independent catalog observables. This is a sobering result: it implies that the bar for claiming physi-

cal significance of nonstationary ETAS parameters should be higher than a favorable information criterion alone.

We propose that independent, model-agnostic diagnostic validation — of the kind implemented here through unsupervised anomaly detection — should become a standard component of nonstationary ETAS analyses. The specific implementation using Isolation Forest is straightforward, requires no assumptions about the functional form of parameter variations, and can be applied to any earthquake catalog for which multi-dimensional metadata (depth, location, magnitude) are available alongside event times. The semi-synthetic validation protocol introduced here provides a principled way to calibrate the sensitivity of such diagnostics to the specific catalog at hand, ensuring that negative results can be interpreted with confidence.

More broadly, this work underscores a fundamental distinction that is sometimes conflated in the literature: statistical nonstationarity — in the sense of a significant improvement in model fit achieved by allowing parameters to vary — is a necessary but not sufficient condition for inferring physically distinct seismic phases. The sufficient condition requires that the inferred variations are consistent with independent, model-agnostic evidence and are not attributable to structural properties of the statistical model itself. Establishing this distinction is particularly important in the context of operational seismic hazard assessment, where nonstationary parameter estimates may be used to update time-dependent seismicity forecasts (Gerstenberger et al., 2005).

6. Conclusions

We have developed and applied a diagnostic framework for evaluating the physical robustness of nonstationary ETAS aftershock models, integrating penalized likelihood inference with model-agnostic unsupervised anomaly detection. Application to the aftershock sequence of the 2025 M_w 7.1 Dingri earthquake in southern Tibet leads to the following principal conclusions.

1. The stationary ETAS model provides a statistically adequate description of the Dingri aftershock sequence, with physically plausible parameter estimates and well-behaved residual diagnostics. The estimated parameters are consistent with a rapidly decaying aftershock sequence superimposed on an elevated background seismicity rate characteristic of the active southern Tibetan rift environment.
2. The nonstationary ETAS extension, fitted by empirical-Bayes penalized inference, yields a stable solution in which the ABIC criterion selects finite smoothing weights for both anomaly functions, producing moderately regularized estimates of the time-varying background rate and triggering productivity. Because this solution is obtained by penalized maximum-a-posteriori inference rather than ordinary maximum likelihood, it is not compared with the stationary model through a formal information criterion; both models nonetheless yield time-rescaling residuals consistent with a unit-rate Poisson process.
3. Independent, model-agnostic Isolation Forest analysis of multivariate catalog features reveals no coherent anomaly in seismicity rate, magnitude statistics, depth distribution, or spatial dispersion during the interval where the nonstationary model infers the strongest parameter variations. Controlled semi-synthetic validation experiments confirm that the detection framework has sufficient sensitivity to recover anomalies of comparable amplitude when they are genuinely present in the catalog, ruling out insufficient detector sensitivity as an explanation for the negative result.
4. The inferred nonstationarity in the Dingri sequence is not accompanied by a macroscopically observable change in any directly measured catalog property, and is robust to the nonstationary parameterization adopted (including an inversion in which only the background rate is allowed to vary). The time-varying ETAS parameters should therefore be regarded as an effective statistical description

rather than as direct evidence for a physically distinct seismic regime.

5. A favorable fit is not sufficient evidence for attributing physical significance to time-varying ETAS parameters. We propose that model-agnostic diagnostic validation should become standard practice in nonstationary ETAS analyses, particularly when parameter transients are used to support claims about post-seismic physical processes. The diagnostic proposed here is general in principle and can be applied to other aftershock sequences for which event times, magnitudes, depths, and locations are available; a systematic, multi-sequence evaluation against cases with and without independent evidence for post-seismic forcing is an important direction for future work.

CRedit authorship contribution statement

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

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

Giuseppe Petrillo and Luca Dal Zilio 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). Giuseppe Petrillo acknowledges support from the DyDéCo project and from the CPJ Chair at ENS de Lyon and CNRS. We are grateful to Zhangfeng Ma for providing us with a GMT script and data to plot Fig. 1. We thank Andy Michael and an anonymous reviewer for their constructive comments, which helped us improve the clarity and methodological presentation of the manuscript.

Data availability

The earthquake catalog used in this study is that of Yao et al. (2025), to which readers are referred for full acquisition and processing details. The complete numerical framework developed for this study, including earthquake catalog, scripts for stationary and nonstationary ETAS inference, model-agnostic Isolation Forest anomaly detection, semi-synthetic catalog generation, and reproduction of all figures, is openly available at: <https://github.com/Computational-Geophysics/ETAS-diagnostic-framework>. The repository contains all earthquake catalogs (original and semi-synthetic), analysis codes, and documentation to reproduce the results presented here.

References

- Armijo, R., Tapponnier, P., Mercier, J., Han, T.L., 1986. Quaternary extension in southern Tibet: field observations and tectonic implications. *J. Geophys. Res. Solid Earth* 91 (B14), 13803–13872.
- Avouac, J.-P., 2004. Postseismic relaxation driven by brittle creep: a possible mechanism to reconcile geodetic measurements and the decay rate of aftershocks, application to the chi-chi earthquake, Taiwan. *J. Geophys. Res. Solid Earth* 109 (B2).
- Bischoff, S.H., Flesch, L.M., 2018. Normal faulting and viscous buckling in the Tibetan Plateau induced by a weak lower crust. *Nat. Commun.* 9 (1), 4952.
- Console, R., Murru, M., Lombardi, A.M., 2003. Refining earthquake clustering models. *J. Geophys. Res. Solid Earth* 108 (B10).
- Dal Zilio, L., Hetényi, G., Hubbard, J., Bollinger, L., 2021. Building the Himalaya from tectonic to earthquake scales. *Nat. Rev. Earth Environ.* 2 (4), 251–268.

- Davidson, J., Green, A., 2011. Are earthquake magnitudes clustered? *Phys. Rev. Lett.* 106 (10), 108502.
- De Arcangelis, L., Godano, C., Lippiello, E., 2018. The overlap of aftershock coda waves and short-term postseismic forecasting. *J. Geophys. Res. Solid Earth* 123 (7), 5661–5674.
- Falcone, G., Mendicelli, A., Mori, F., Fabozzi, S., Moscatelli, M., Occhipinti, G., Peronace, E., 2020. A simplified analysis of the total seismic hazard in Italy. *Eng. Geol.* 267, 105511.
- Freed, A.M., 2005. Earthquake triggering by static, dynamic, and postseismic stress transfer. *Annu. Rev. Earth Planet. Sci.* 33, 335–367.
- Gentili, S., Brondi, P., Rossi, G., Sukan, M., Petrillo, G., Campanella, S., 2024. Seismic clusters and fluids diffusion: a lesson from the 2018 molise (Southern Italy) earthquake sequence. *Earth Planets Space* 76 (1), 157.
- Gerstenberger, M.C., Wiemer, S., Jones, L.M., Reasenberg, P.A., 2005. Real-time forecasts of tomorrow's earthquakes in California. *Nature* 435 (7040), 328–331.
- Godano, C., Petrillo, G., Lippiello, E., 2024. Evaluating the incompleteness magnitude using an unbiased estimate of the b value. *Geophys. J. Int.* 236 (2), 994–1001.
- Hainzl, S., Ogata, Y., 2005. Detecting fluid signals in seismicity data through statistical earthquake modeling. *J. Geophys. Res. Solid Earth* 110 (B5). <https://doi.org/10.1029/2004JB003247>
- Hainzl, S., Zakharova, O., Marsan, D., 2013. Impact of aseismic transients on the estimation of aftershock productivity parameters. *Bull. Seismol. Soc. Am.* 103 (3), 1723–1732.
- Helmstetter, A., Sornette, D., 2003. Importance of direct and indirect triggered seismicity in the ETAS model of seismicity. *Geophys. Res. Lett.* 30 (11), 1576. <https://doi.org/10.1029/2003GL017670>
- Jagla, E.A., Landes, F.P., Rosso, A., 2014a. Viscoelastic effects in avalanche dynamics: a key to earthquake statistics. *Phys. Rev. Lett.* 112 (17), 174301.
- Jagla, E.A., , 2014b. Aftershock production rate of driven viscoelastic interfaces. *Phys. Rev. E* 90 (4), 042129.
- Kumazawa, T., Ogata, Y., 2013. Quantitative description of induced seismic activity before and after the 2011 Tohoku-Oki earthquake by nonstationary ETAS models. *J. Geophys. Res. Solid Earth* 118 (11), 6165–6182. <https://doi.org/10.1002/2013JB010259>
- Kumazawa, T., Ogata, Y., 2014. Nonstationary ETAS models for nonstandard earthquakes. *Ann. Appl. Stat.* 8. <https://doi.org/10.1214/14-AOAS759>
- Lippiello, E., Petrillo, G., Landes, F., Rosso, A., 2019. Fault heterogeneity and the connection between aftershocks and afterslip. *Bull. Seismol. Soc. Am.* 109 (3), 1156–1163.
- Liu, F.T., Ting, K.M., Zhou, Z.H., 2008. Isolation forest. In: *Proceedings of the 8th IEEE International Conference on Data Mining (ICDM)*. IEEE, pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>
- Llenos, A.L., Michael, A.J., 2013. Modeling earthquake rate changes in Oklahoma and Arkansas: possible signatures of induced seismicity. *Bull. Seismol. Soc. Am.* 103 (5), 2850–2861.
- Llenos, A.L., Michael, A.J., 2016. Characterizing potentially induced earthquake rate changes in the Brawley seismic zone, Southern California. *Bull. Seismol. Soc. Am.* 106 (5), 2045–2062.
- Ma, Z., Li, C., Zeng, H., Chen, H., Lyu, M., Zhang, Y., Dal Zilio, L., Shan, X., Wei, S., 2026. Deciphering stress perturbations throughout the 2025 M_w 7.1 Dingri, southern Xizang earthquake. *Nat. Commun.*
- Marsan, D., Lengliné, O., 2008. Extending earthquakes' reach through cascading. *Science* 319 (5866), 1076–1079.
- Mizrahi, L., Dallo, I., Van Der Elst, N.J., Christophersen, A., Spassiani, I., Werner, M.J., Iturrieta, P., Bayona, J.A., Iervolino, I., Schneider, M., et al., 2024. Developing, testing, and communicating earthquake forecasts: current practices and future directions. *Rev. Geophys.* 62 (3), e2023RG000823.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* 83 (401), 9–27.
- Ogata, Y., 1998. Space-time point-process models for earthquake occurrences. *Ann. Inst. Stat. Math.* 50 (2), 379–402.
- Petrillo, G., Convertito, V., Tramelli, A., Godano, C., 2026. Seismic swarms occurrence rate and b -value mapping at Taupo volcanic zone. *Seismica* 5 (1), 1
- Petrillo, G., Kumazawa, T., Napolitano, F., Capuano, P., Zhuang, J., 2024. Fluid-triggered swarm sequence supported by a nonstationary epidemic-like description of seismicity. *Seismol. Res. Lett.* 95 (6), 3207–3220.
- 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 (1), 3010.
- 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 (2).
- Petrillo, G., Zhuang, J., 2023. Verifying the magnitude dependence in earthquake occurrence. *Phys. Rev. Lett.* 131 (15), 154101.
- Petrillo, G., Zhuang, J., 2024. Bayesian earthquake forecasting approach based on the epidemic type aftershock sequence model. *Earth Planets Space* 76 (1), 78.
- Qiao, X., Lu, Z., Yan, S., Shi, H., Zhi, M., Zhao, D., 2025. The 2025 Mw7.0 Dingri earthquake: conjugate normal faulting of a graben structure in the southern Xainza-Dinggye rift. *Geophys. Res. Lett.* 52 (12), e2025GL116154.
- Schneider, M., Barall, M., Guttorp, P., Hardebeck, J., Michael, A.J., Page, M., van der Elst, N., 2025. Bayesian ETAS modeling for the pacific northwest: uncovering effects of tectonic regimes, regional differences, and swarms on aftershock parameters. *Bull. Seismol. Soc. Am.* 115 (5), 2219–2236.
- Shcherbakov, R., 2024. A stochastic model for induced seismicity at the geothermal systems: a case of the Geysers. *Seismol. Res. Lett.* 95 (6), 3545–3556.
- Utsu, T., Ogata, Y., Matsu'ura, R.S., 1995. The centenary of the Omori formula for a decay law of aftershock activity. *J. Phys. Earth* 43 (1), 1–33.
- Veen, A., Schoenberg, F.P., 2008. Estimation of space-time branching process models in seismology using an em-type algorithm. *J. Am. Stat. Assoc.* 103 (482), 614–624.
- Wang, K., Hu, Y., He, J., 2012. Deformation cycles of subduction earthquakes in a viscoelastic Earth. *Nature* 484 (7394), 327–332.
- Wei, G., Chen, K., Li, M., Huang, K., Zhu, Q., Yuan, L., Dal Zilio, L., 2026. Pre-and co-seismic stress loading promoted low-angle splay fault during the 2025 mw7. 1 Tingri earthquake. *Commun. Earth Environ.* 7 (1), 426.
- Wesnousky, S.G., 2006. Predicting the endpoints of earthquake ruptures. *Nature* 444 (7117), 358–360.
- Xu, X., Wang, S., Cheng, J., Wu, X., 2025. Shaking the Tibetan plateau: insights from the Mw 7.1 Dingri earthquake and its implications for active fault mapping and disaster mitigation. *NPJ Nat. Hazards* 2 (1), 16.
- Yao, J., Yao, D., Chen, F., Zhi, M., Sun, L., Wang, D., 2025. A preliminary catalog of early aftershocks following the 7 January 2025 M_s 6.8 Dingri, Xizang earthquake. *J. Earth Sci.* 36 (2), 856–860.
- Zhuang, J., Ogata, Y., 2006. Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. *Phys. Rev. E* 73, 046134. <https://doi.org/10.1103/PhysRevE.73.046134>
- Zhuang, J., Ogata, Y., Vere-Jones, D., 2002. Stochastic declustering of space-time earthquake occurrences. *J. Am. Stat. Assoc.* 97 (458), 369–380.