FULL PAPER

Comparison of statistical low-frequency earthquake activity models

Tomoaki Nishikawa[1*]

Abstract

Slow earthquakes are slow fault slip events. Quantifying and monitoring slow earthquake activity characteristics are important, because they may change before large earthquakes occur. Statistical seismicity models are useful for quantifying seismicity characteristics. However, no standard statistical model exists for slow earthquake activity. This study used a high-quality catalog of low-frequency earthquakes (LFEs), a type of slow earthquake, in the Nankai subduction zone from April 2004 to August 2015 and conducted the first comparison of existing statistical LFE activity models to determine which model better describes LFE activity. Based on this comparison, this study proposes a new hybrid model that incorporates existing model features. The new model considers the LFE activity history in a manner similar to the epidemic-type aftershock sequence (ETAS) model and represents the LFE aftershock rate (subsequent LFE occurrence rate) with a small number of model parameters, as in the Omori–Utsu aftershock law for regular earthquakes. The results show that the proposed model outperforms other existing models. However, the new model cannot reproduce a feature of LFE activity: the sudden cessation of intense LFE bursts. This is because the new model superimposes multiple aftershock activities and predicts extremely high seismicity rates during and after the LFE bursts. I suggest that reproducing and successfully predicting the sudden cessation of intense LFE bursts is critical for the further improvement of statistical LFE activity models. In addition, the empirical equations formulated in this study for the LFE aftershock rates may be useful for future statistical and physical modeling of LFE activity.

Keywords Low-frequency earthquake, Statistical model, Seismicity, Nankai Trough, Subduction zone, Slow earthquake
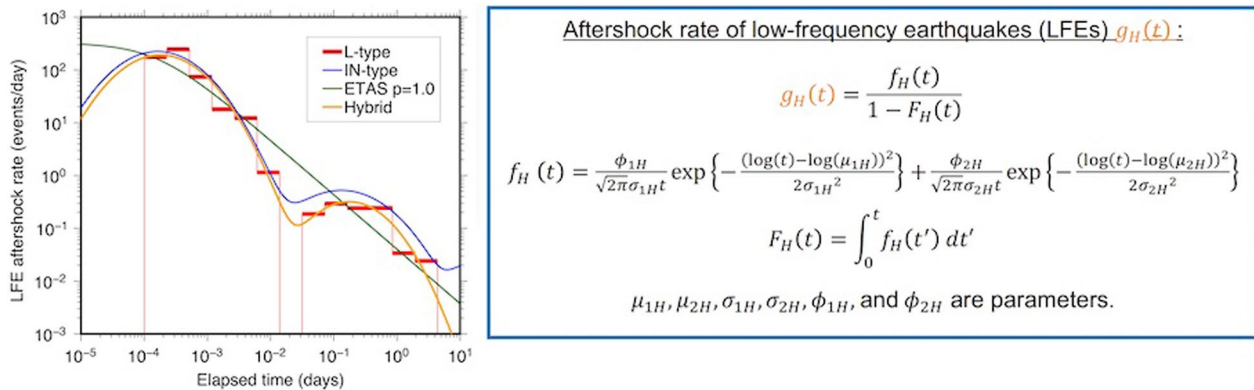
---

*Correspondence:
Tomoaki Nishikawa
nishikawa.tomoaki.2z@kyoto-u.ac.jp
Full list of author information is available at the end of the article

Graphical Abstract

Introduction

Slow earthquakes are a general term used to describe various spontaneous slow fault-slip events, such as low-frequency earthquakes, tectonic tremors, very-low-frequency earthquakes, and slow slip (Ide et al. 2007; Ide and Beroza 2023). They are characterized by longer duration than regular (fast) earthquakes of comparable seismic moment. Slow earthquakes have been observed or inferred to occur at several subduction and transform plate boundaries (e.g., Dragert et al. 2001; Obara 2002; McGuire et al. 2005; Nadeau and Dolenc 2005; Kuna et al. 2019). In particular, very detailed slow earthquake activity has been revealed in subduction zones such as the Nankai Trough, Cascadia subduction zone, and the Hikurangi and Japan Trenches (Beroza and Ide 2011; Obara and Kato 2016; Wallace 2020; Nishikawa et al. 2019).

In subduction zones, slow earthquakes occur mainly on the immediate deep and/or shallow sides of megathrust earthquake sources. The relationship between slow and megathrust earthquakes has been actively studied (e.g., Obara and Kato 2016; Nishikawa et al. 2023). Simulation studies have indicated that slow earthquake activity characteristics may change prior to megathrust earthquakes (e.g., Matsuzawa et al. 2010; Luo and Liu 2019). Therefore, quantifying and monitoring slow earthquake activity characteristics are important for gaining insight into the imminence of future megathrust earthquakes.

Statistical seismicity models are useful for quantifying seismicity characteristics. For example, the epidemic-type aftershock sequence (ETAS) model is widely used to quantify the characteristics of regular earthquake activity (Ogata 1988; Zhuang et al. 2002; Kumazawa and Ogata 2013; Nishikawa and Nishimura 2023). This model was also used as a standard model in an international earthquake forecasting experiment named the Collaboratory for the Study of Earthquake Predictability (Schorlemmer et al. 2018). In contrast to the regular earthquake activity, there is no standard statistical model for slow earthquake activity. Statistical modeling of low-frequency earthquakes (LFEs), a type of slow earthquake, has only recently begun. LFEs are characterized by low dominant frequencies (1–10 Hz) compared to regular microearthquakes of comparable seismic moment (Shelly et al. 2007; Nishikawa et al. 2023). Lengliné et al. (2017) and Tan and Marsan (2020) proposed statistical LFE activity models similar to the ETAS model. In their models, the LFE occurrence rate was assumed to be the sum of the stationary background rate and the effect of past LFEs or external processes behind past LFEs inducing future LFEs.

Ide and Nomura (2022) proposed a statistical model for tectonic tremors (i.e., LFE swarms) (Shelly et al. 2007). Their model is based on an approach different from that of Lengliné et al. (2017) and Tan and Marsan (2020). Ide and Nomura (2022) described the probability distribution of tremor interevent times using a mixture of lognormal and Brownian passage-time (BPT) distributions and forecasted tremor interevent times. Furthermore, their model depends only on the time elapsed since the last event and does not depend on detailed tremor activity history.

Constructing a statistical model that successfully describes the LFE activity is important for better characterization and forecasting. However, existing statistical LFE activity models (Lengliné et al. 2017; Tan and Marsan 2020; Ide and Nomura 2022) have never been compared, and it is unclear which model best describes LFE activity. This study applies the statistical LFE and ETAS models (Ogata 1988;

Lengliné et al. 2017; Ide and Nomura 2022) to LFE activity along the Nankai Trough and compare their performances using Akaike's information criterion (AIC; Akaike 1974) (Sect. "Results"). Based on the model comparison, I propose a new model that incorporates existing model features and discuss LFE activity occurrence patterns (Sect. "Discussion").

Data and methods

LFE catalog and study regions

A high-quality catalog of LFEs along the Nankai Trough in southwest Japan created by Kato and Nakagawa (2020) (Fig. 1a) was used. In the Nankai Trough, the Philippine Sea Plate subducts under the Amurian Plate. This plate subduction causes diverse slow and fast earthquakes at the plate boundary (Obara and Kato 2016; Takemura et al. 2023). Kato and Nakagawa (2020) detected LFEs along the Nankai Trough from April 2004 to August 2015 using a matched-filter technique, using LFEs in the Japan Meteorological Agency (JMA) catalog as templates. This catalog contained approximately 510,000 LFEs from M −1.1 to M 1.4 (Fig. 1b), which is approximately 23 times the number of LFEs in the JMA catalog over the same period.

When analyzing the LFE activity, it is necessary to determine the minimum magnitude of the LFEs to be analyzed. In regular earthquake activity analyses, the earthquake catalog completeness is typically evaluated, and the completeness magnitude is used as the minimum magnitude (e.g., Nishikawa and Nishimura 2023). However, no method currently exists for evaluating the completeness of an LFE catalog. Therefore, as a compromise, the goodness-of-fit test (GFT) method (Wiemer and Wyss 2000; Woessner and Wiemer 2005), which has been used to evaluate the completeness magnitude of regular earthquake catalogs, was applied to the LFE catalog, and the obtained magnitude was used as a reference to select the minimum magnitude. All LFEs in the Kato and Nakagawa (2020) catalog were used regardless of their source locations. See Woessner and Wiemer (2005) for details of the GFT method.

In the GFT method, the R value is a measure of the goodness-of-fit between the observed number of earthquakes and the number of earthquakes expected based on the Gutenberg–Richter relationship (Gutenberg and Richter 1944). Specifically, the R value is the absolute difference between the observed and expected number of events in magnitude bins and is given by

$$R(M) = 100 \cdot \left\{ 1 - \left(\frac{\sum_{M_i=M}^{M_{\max}} |B_i - S_i|}{\sum_i B_i} \right) \right\}, \qquad (1)$$

where B_i and S_i are the observed and expected number of events in each magnitude bin ($M_i - \Delta M/2$ to $M_i + \Delta M/2$), respectively, M is the magnitude for which the R value is calculated, and $M_{\max}$ is the catalog maximum magnitude. ΔM is the width of the magnitude bins and was set to $\Delta M = 0.1$ in this analysis. A perfect fit between the observed and expected values yields R of 100%. See Wiemer and Wyss (2000) for more details on R value calculation.

I searched for the smallest magnitude with an $R > 95\%$ (Fig. 1c). The resulting magnitude was M 0.4. Considering this estimate of the GFT method, LFEs with M 0.6 or greater were used, which is M 0.2 greater than the estimate. This is because I prefer the minimum magnitude to be as conservative (or as large) as possible. However, if the minimum magnitude is too large, the number of LFEs will be too small for statistical analysis. In the following analyses, I used 3545 M 0.6 or larger LFEs from Kato and Nakagawa's (2020) catalog. In the above analysis, I followed Bostock et al. (2015) and assumed that the LFE magnitude–frequency distribution (especially for large events in the catalog) obeys the Gutenberg–Richter law. I further discuss this assumption and the influence of the minimum magnitude selection on my results in Sect. "Selection of the minimum magnitude".

In the Nankai Trough, LFEs occur in band-like regions downdip of the megathrust seismogenic zones (Fig. 1a). These band-like regions were divided into subregions in a manner similar to that of Ide and Nomura (2022). In and around the band-like regions, rectangles of 0.2° in latitude and longitude were placed at 0.1° intervals to create subregions (Fig. 1a). Each subregion was required to contain at least 100 $M \geq 0.6$ LFEs, and LFE activity in each subregion was analyzed. Consequently, 43 subregions were included in the analysis.

Statistical LFE activity models

Lengliné et al. (2017) model

Lengliné et al. (2017) proposed a statistical LFE activity model similar to the ETAS model, see Sect. "ETAS model" for details on the ETAS model. Their model assumed the rate of LFE occurrence λ at time t to be the sum of a stationary background rate μ_L and the effect of past LFEs or external processes behind the LFEs inducing future LFEs:

$$\lambda(t) = \mu_L + \sum_{t_i < t} a_i \cdot g(t - t_i), \qquad (2)$$

where a_i (≥ 0) are parameters controlling the magnitude of the effect of the ith event inducing following events, t_i is the occurrence time of the ith event, and g is the normalized aftershock rate kernel. In this study, an event refers to the combined phenomenon of an LFE and possible external processes behind it. Lengliné et al. (2017) interpreted the external process as a transient slip rate
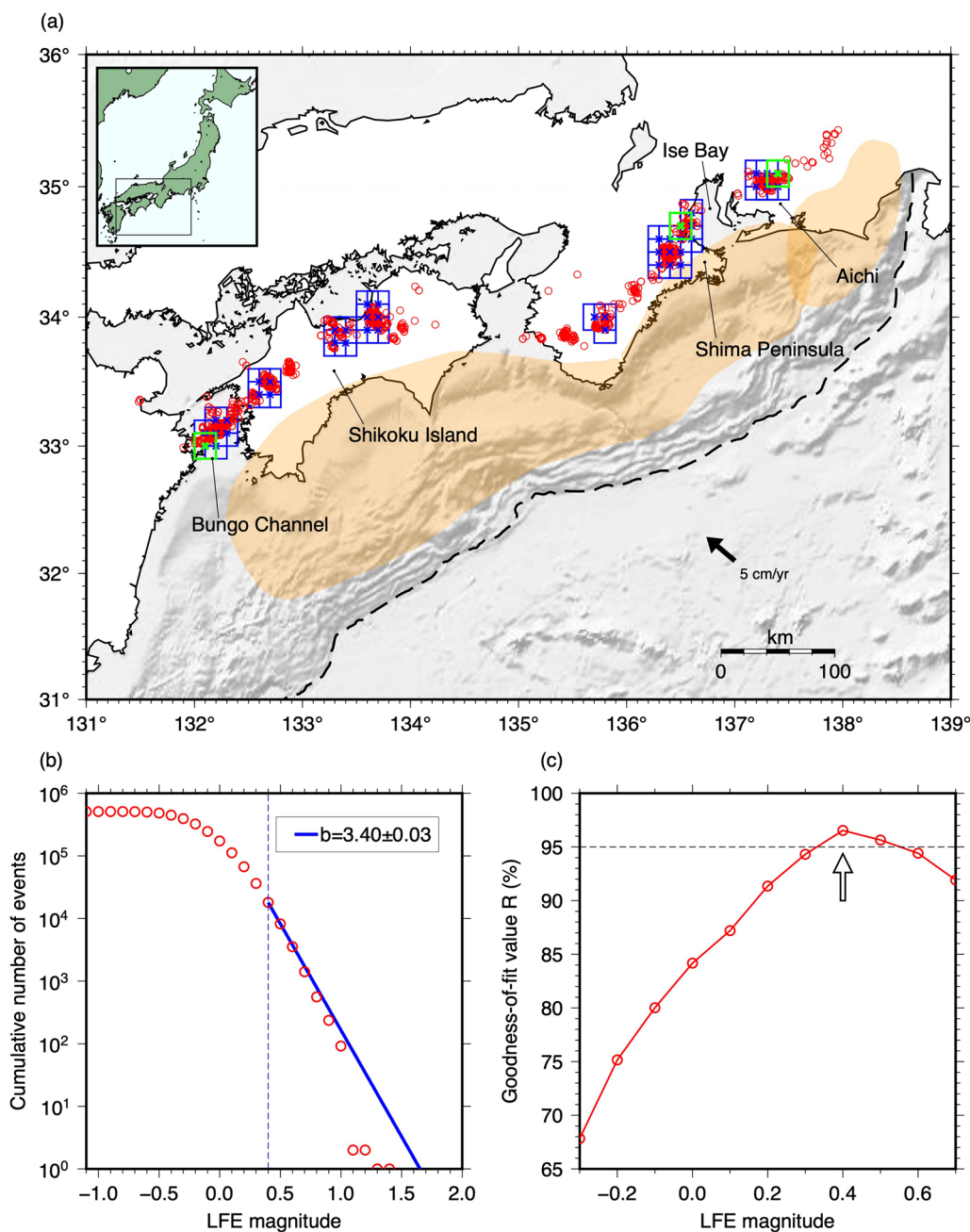
Fig. 1 LFEs along the Nankai Trough. **a** LFE epicenters and 43 rectangular subregions used in this study. Red circles denote LFE epicenters from April 2004 to August 2015 (Kato and Nakagawa 2020). Blue rectangles and crosses indicate the subregions and their centers, respectively. The light green rectangles and crosses denote the subregion locations selected in Sect. "Subregion examples". The dashed black line denotes the trench axis of the Nankai Trough. The orange areas indicate the Nankai Trough megathrust seismogenic zones (Obara and Kato 2016). The small inset map shows the location of the Nankai Trough. **b** LFE size-frequency distribution. The blue solid line indicates the number of events expected from Gutenberg–Richter's relationship (Gutenberg and Richter 1944) ($b = 3.40$). **c** Goodness-of-fit value R for the goodness-of-fit test method (Wiemer and Wyss 2000; Woessner and Wiemer 2005). The arrow indicates the smallest magnitude with an $R > 95\%$ (M 0.4)

increase in the area surrounding LFE-generating asperities. In the above formulation, a_i is equivalent to the number of events induced by the ith event.

The most significant difference between the above model and the ETAS model is that in the former, the shape of the normalized kernel g is not assumed a priori but is determined based on the observed data. In

contrast, the ETAS model assumes a simple power-law decay ($\sim t^{-p}$) for the aftershock rate kernel. Lengliné et al. (2017) used piecewise constant discretization for g:

$$g(T_k < t < T_{k+1}) = g_k, \qquad (3)$$

where T_k is the time interval used for discretization and $k \in [1:N]$, with $T_1 = 0$ (days) and $T_N = 10$ (days). Furthermore, when $T_N < t$, g is assumed to be zero:

$$g(T_N < t) = 0. \qquad (4)$$

The aftershock rate kernel g is normalized to satisfy the following equation:

$$\int_0^\infty g(t)dt = \sum_{k=1}^{N-1} g_k \cdot (T_{k+1} - T_k) = 1. \qquad (5)$$

Lengliné et al. (2017) used approximately 30 piecewise constants to discretize the kernel g. They applied their model to LFE activity in the Mexican and Cascadia subduction zones and on the San Andreas Fault.

The Lengliné et al. (2017) model has the drawback of numerous parameters (usually hundreds or more), because the number of parameters a_i increases with the number of events. This is a major disadvantage when compared to other models using AIC, because AIC substantially increases with the number of model parameters, see Sect. "Model performance evaluation" for details on AIC. In my preliminary analyses, the original Lengliné et al. (2017) model had a much larger AIC than the Ide and Nomura (2022) and ETAS models, because it had too many model parameters. This suggests that the original Lengliné et al. (2017) model is generally inferior to other models. Therefore, in this study, the following model with significantly reduced model parameters was used:

$$\lambda(t) = \mu_L + \sum_{t_i < t} a \cdot g(t - t_i). \qquad (6)$$

The triggering effect magnitude (or aftershock rate amplitude) was assumed to be constant for all events (i.e., $a_i = a$). Tan and Marsan (2020) reduced the parameters of Lengliné et al. (2017) model in a similar manner. However, the model proposed by Tan and Marsan (2020) is a spatiotemporal LFE model, whereas the above model is purely temporal. I used the temporal model, because Ide and Nomura (2022), which is used for comparison, is a temporal model. Hereafter, the above model is referred to as the L-type model.

New parameters g'_k and kernel function g' were defined as follows:

$$g'_k \equiv a \cdot g_k, \qquad (7)$$

$$g'(T_k < t < T_{k+1}) = g'_k. \qquad (8)$$

Using the new kernel function g', Eq. 6 can be rewritten as

$$\lambda(t) = \mu_L + \sum_{t_i < t} g'(t - t_i). \qquad (9)$$

The likelihood $L(\theta|X)$ of this model is given by

$$\log L(\theta|X) = \sum_{0 \le t_i \le T} \log \lambda(t_i) - \int_0^T \lambda(t)dt, \qquad (10)$$

where θ represents a set of the model parameters and $X = \{t_i | 0 \le t_i \le T\}$ are the data. I obtained the model parameters (μ_L and g'_k) that maximized the above likelihood via the Powell method (Powell 1964). In this study, I used 15 piecewise constants (g'_1 to g'_{15}) to discretize the aftershock rate kernels. I assigned g'_1 in the period from 0 to 10^{-4} days. I then divided the period from 10^{-4} to 10 days into 14 divisions evenly spaced in logarithmic space and assigned g'_2 to g'_{15}, respectively. In addition, Eq. 7 was used to obtain the parameter a, which indicates the number of events induced by a single event.

Lengliné et al. (2017) apply their model to LFEs belonging to the same family. An LFE family is defined as a group of LFEs with similar waveforms. However, as Lengliné et al. (2017) acknowledge, there is no theoretical restriction that this model should only be applied to events belonging to the same LFE family. Therefore, in this study, the model was applied to all LFEs that occurred within the same subregion (see Sect. "LFE catalog and study regions").

Ide and Nomura (*2022*) model

Ide and Nomura (2022) represented the probability distribution of inter-event times using a mixture of log-normal and BPT distributions, which correspond to short-term clustering and long-term recurrence of events, respectively. Ide and Nomura (2022) applied their model to deep tectonic tremors along the Nankai Trough. Upon applying the Ide and Nomura (2022) model to LFE activity along the Nankai Trough, an additional log-normal distribution was added to the model. This is because, as discussed in Sect. "Subregion examples", short-term LFE clustering often displays two characteristic timescales (tens of seconds and a few hours). Kato and Nakgawa (2020) refer to this LFE activity feature as intermittency. Tan and Marsan (2020) also reported two characteristic timescales for short-term LFE clustering along the San Andreas Fault.

The model consists of two lognormal distributions (the first and second terms on the right-hand side) and

one BPT distribution (the third term on the right-hand side):

$$
\begin{aligned}
f(\Delta t) = & \frac{\phi_1}{\sqrt{2\pi}\sigma_1 \Delta t} \exp\left\{ -\frac{\left(\log(\Delta t) - \log(\mu_1)\right)^2}{2\sigma_1^2} \right\} \\
& + \frac{\phi_2}{\sqrt{2\pi}\sigma_2 \Delta t} \exp\left\{ -\frac{\left(\log(\Delta t) - \log(\mu_2)\right)^2}{2\sigma_2^2} \right\} \\
& + (1 - \phi_1 - \phi_2)\sqrt{\frac{\mu_3}{2\pi\sigma_3^2 \Delta t^3}} \exp\left\{ -\frac{(\Delta t - \mu_3)^2}{2\mu_3\sigma_3^2 \Delta t} \right\},
\end{aligned}
\tag{11}
$$

where Δt is the interevent time, and $f(\Delta t)$ is the probability density function of Δt, μ_1, μ_2, and μ_3 ($\mu_1 < \mu_2 < \mu_3$) are parameters representing characteristic time scales of the log-normal and BPT distributions, σ_1 and σ_2 represent $\log(\Delta t)$ variances of the first and second log-normal distributions, respectively, σ_3 controls the variation in LFE burst long-term recurrence periods, and ϕ_1 and ϕ_2 are parameters controlling the proportions of the first and second log-normal distribution. We call the above model the IN-type model.

An important feature of the IN-type model is that the probability density distribution $f(\Delta t)$ does not depend on detailed LFE activity history. This differs markedly from the L-type and ETAS models, which consider the superimposition of multiple LFE aftershock rates (see Eq. 6).

The likelihood $L(\theta|X)$ of the IN-type model is given by

$$
\log L(\theta|X) = \sum_{i=1}^{n} \log\left\{ f(\Delta t_i) \right\},
\tag{12}
$$

$$
\Delta t_i \equiv t_i - t_{i-1},
\tag{13}
$$

where t_i is the ith event occurrence time. Eight model parameters (μ_1, μ_2, μ_3, σ_1, σ_2, σ_3, ϕ_1, and ϕ_2) were obtained that maximize the likelihood via Powell's method (Powell 1964).

The event occurrence rate expected from the IN-type model (Ide and Nomura 2022) is given by

$$
h(\Delta t) = \frac{f(\Delta t)}{1 - F(\Delta t)},
\tag{14}
$$

$$
F(\Delta t) = \int_0^{\Delta t} f(t')dt'.
\tag{15}
$$

These equations were used to compare the expected LFE aftershock rates from the different models in Sect. "Results".

ETAS model

The ETAS model (Ogata 1988), which is similar to the L-type model (Lengliné et al. 2017), assumes that the event occurrence rate is the sum of the stationary background rate and the effect of past events inducing future events. However, the aftershock rate kernel of the ETAS model differs from that of the L-type model, and the ETAS model assumes power-law decay a priori (Omori–Utsu's aftershock law; Utsu 1957; Utsu et al. 1995). In the ETAS model, the event occurrence rate λ at time t is written as follows:

$$
\lambda(t) = \mu_E + \sum_{t_i < t} \frac{Ke^{\alpha(M_i - M_c)}}{(t - t_i + c)^p},
\tag{16}
$$

where μ_E is the stationary background rate, the second term of the right-hand side is the aftershock rate summation controlled by four parameters (α, c, K, and p), t_i and M_i are the occurrence time and ith event magnitude, respectively, and M_c is the catalog minimum magnitude. As with the L-type model, the likelihood of the ETAS model $L(\theta|X)$ is given by Eq. 10. The ETAS parameters that maximized the likelihood were obtained using Powell's method (Powell 1964).

Model performance evaluation

The L-type, IN-type, and ETAS models were compared based on AIC (Akaike 1974), which is defined as follows:

$$
\text{AIC} = -2\log L(\theta|X) + 2k,
\tag{17}
$$

where k denotes the number of adjusted parameters. The L-type, IN-type, and ETAS models had 16, 8, and 5 adjusted parameters, respectively. A model with a smaller AIC was considered to be significantly better than a model with a larger AIC when the AIC difference (ΔAIC) between the models was 2 or greater.

Results

Subregion examples

This section presents examples of the analyses of the three selected subregions. The selected subregion locations are shown in Fig. 1a (light-green rectangles and crosses).

The first subregion is in the Bungo Channel (33.0°N, 132.1°E; subregion #1). Figure 2a shows the magnitude–time diagram of this subregion. The clear increase in the frequency of LFEs in 2010 corresponds to a period of a long-term slow-slip event (SSE) in the Bungo Channel (Takagi et al. 2019; Kato and Nakagawa 2020). Model parameter estimates and AICs for this subregion
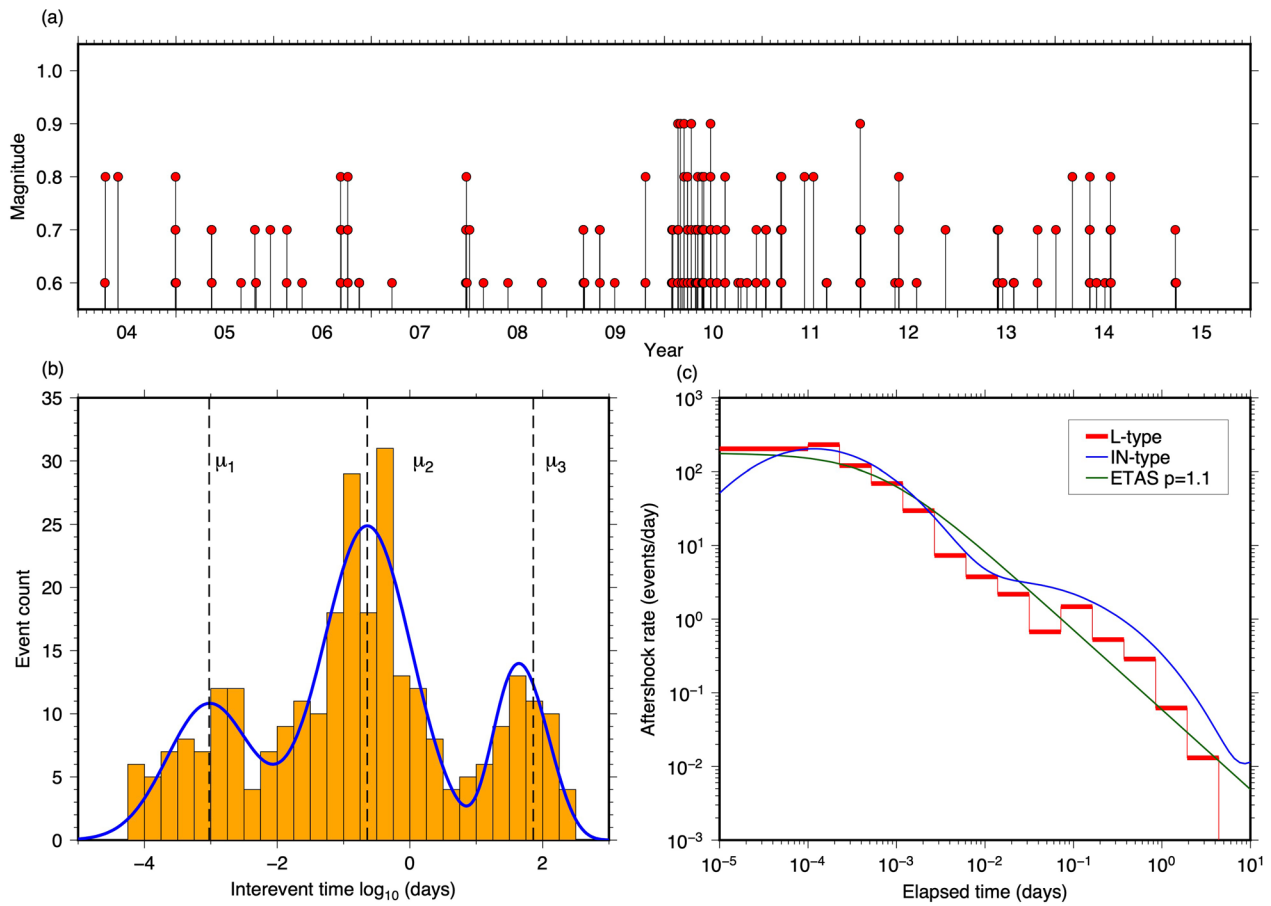
Fig. 2 LFE activity in a subregion near the Bungo Channel (33.0°N, 132.1°E; subregion #1). **a** LFE magnitude–time diagram. **b** Interevent time histogram (orange bars). The blue curve indicates the interevent time distribution of the IN-type model. Vertical dashed lines denote the three characteristic time constants (μ_1, μ_2, and μ_3) of the IN-type model. **c** Aftershock rates (LFE occurrence rates after a single LFE) predicted by the L-type, IN-type, and ETAS models (red, blue, and green lines, respectively). The maximum likelihood estimate of the ETAS model p value is 1.1

Table 1 Model parameter estimates and AICs for the selected subregions

Model	Subregion	Parameter								AIC
		μ_L	a							
L-type	#1 (Bungo Channel)	1.6×10^{-2}	0.78							524
	#34 (Shima Peninsula)	6.8×10^{-3}	0.83							317
	#43 (Aichi Prefecture)	1.5×10^{-2}	0.52							618
		μ_1	σ_1	μ_2	σ_2	μ_3	σ_3	ϕ_1	ϕ_2	
IN-type	#1 (Bungo Channel)	9.5×10^{-4}	1.5	0.23	1.5	72	1.0	0.24	0.58	499
	#34 (Shima Peninsula)	1.6×10^{-3}	1.3	0.18	1.6	174	0.5	0.17	0.69	282
	#43 (Aichi Prefecture)	7.2×10^{-4}	1.2	0.62	1.3	73	1.3	0.23	0.34	626
		μ_E	α	c	K	p				
ETAS	#1 (Bungo Channel)	5.6×10^{-3}	0.00	6.2×10^{-4}	5.9×10^{-2}	1.1				585
	#34 (Shima Peninsula)	6.6×10^{-3}	0.00	4.5×10^{-1}	3.4×10^{-1}	3.0				358
	#43 (Aichi Prefecture)	1.1×10^{-2}	0.00	1.6×10^{-4}	4.0×10^{-2}	1.0				638

g'_1 to g'_{15} of the L-type model are shown in Additional file 2: Table S1

are shown in Table 1 (see subregion #1). Model parameter estimates and AICs for all 43 subregions are shown in Additional files 2, 3, 4, 5: Tables S1–S4 in Additional files. In this subregion, the AICs of the L-type, IN-type, and ETAS models were 524, 499, and 585, respectively (Table 1). The AIC of the IN-type model was significantly lower than that of the other models, indicating that the IN-type model was superior to the other two models in this subregion.

Figure 2b shows a histogram of the interevent times for this subregion. There are maxima at approximately 10^{-3}, 10^{-1}, and $10^{1.5}$ days (approximately 30 days). This feature is difficult to reproduce using a model with only two characteristic timescales such as the original Ide and Nomura (2022) model. In contrast, the IN-type model (blue curve in Fig. 2b) had three characteristic time constants (μ_1, μ_2, and μ_3), and they were estimated to be 9.5×10^{-4}, 2.3×10^{-1}, and 72 days, respectively. Note that the AIC (= 540) for the original Ide and Nomura (2022) model was significantly greater than the AIC (= 499) of the IN-type model in this study, suggesting that the model with the three time constants was significantly better.

Figure 2c shows the temporal variation in the aftershock rate (LFE occurrence rate after a single LFE) for up to 10 days in this subregion, as predicted by the L-, IN-type, and ETAS models. For the ETAS model, the LFE aftershock rate after a single M 0.6 LFE was presented. However, because parameter α of the ETAS model was zero in almost all subregions, including this subregion (Table 1 and Additional file 4: Table S3), it is safe to assume that the aftershock rate does not depend on the mainshock magnitude. In Fig. 2c, the aftershock rate of the ETAS model decays with a power law of $p = 1.1$. For regular earthquake activity, p values typically range from 0.9 to 1.5 (Utsu et al. 1995). It is intriguing that the p value estimated from LFE activity in this subregion is comparable to that of regular earthquake activity.

In contrast to the simple power-law decay of the ETAS model, the aftershock rates expected from the L- and IN-type models exhibit more complex temporal variations. In the IN-type model, the aftershock rate decay stagnates between 1.0×10^{-2} and 2.0×10^{-1} days. This was because of the second log-normal distribution of the IN-type model (Fig. 2b). This feature was also observed in the aftershock rate of the L-type model (red line in Fig. 2b). As shown in Fig. 2c, the aftershock occurrence rate of the L-type model tended to be systematically lower than that of the IN-type model. This is because the L-type model includes the effect of the summation of multiple aftershock rates (Eq. 6). This effect resulted in lower individual aftershock rates compared to the IN-type model.

In this subregion, the ETAS model was the most inferior in terms of AIC. Considering that the only difference between the ETAS and L-type models is the functional form of the aftershock rate kernel (Eqs. 6 and 16), the difference in AIC between the L-type (AIC = 524) and ETAS models (AIC = 585) can be attributed to the difference in the functional form. This result suggests that the simple power-law decay assumed by the ETAS model is insufficient to describe the LFE aftershock rate.

The second example is a subregion near the Shima Peninsula (34.7N°, 136.5E°; subregion #34) (Figs. 1a and 3). In this subregion, the AICs of the L-type, IN-type, and ETAS models were 317, 282, and 358, respectively; the IN-type model had the lowest AIC and outperformed the other two models (Table 1). The AIC value of the ETAS model was the highest. Figure 3b shows a histogram of the interevent times in this subregion. The three characteristic time constants of the IN-type model (μ_1, μ_2, and μ_3) (blue curve in Fig. 2b) were estimated to be 1.6×10^{-3}, 1.8×10^{-1}, and 174 days, respectively.

Figure 3c shows the LFE aftershock rates for the L-type, IN-type, and ETAS model. The p value for the ETAS model is estimated as 3.0. This is substantially greater than the typical p value range for regular earthquake activity (0.9–1.5) (Utsu et al. 1995). Owing to this extremely high p value, the aftershock rate decayed rapidly from 1×10^{-1} to 10 days (from 2 to 3×10^{-4} events/ days). However, up to 1×10^{-1} days, the aftershock rate was almost constant owing to the large c value (= 0.45 d). A rapid decay from 1×10^{-1} to 10 days was also observed for the L- and IN-type models (red and blue lines, respectively, in Fig. 3c).

The IN- and L-type model aftershock rates showed weaker decay between approximately 1.0×10^{-2} and 2.0×10^{-1} days. This is similar to the results for the Bungo Channel subregion (Fig. 2c). The aftershock rates predicted by the IN- and L-type models showed a maximum value (approximately 80 events/days) of approximately 3×10^{-4} days, in contrast to the ETAS model, which predicted an almost constant aftershock rate until approximately 1×10^{-1} days.

The third example is a subregion in the Aichi Prefecture (35.1°N, 137.4°E; subregion #43) (Figs. 1a and 4). In this subregion, the AICs values of the L-type, IN-type, and ETAS models were 618, 626, and 638, respectively, with the L-type model having the lowest AIC (Table 1). This result contrasts with the results of the first and second examples in the Bungo Channel and near Shima Peninsula, where the IN-type model performed best. However, the fact that the ETAS model had the highest AIC was common to all the three examples.

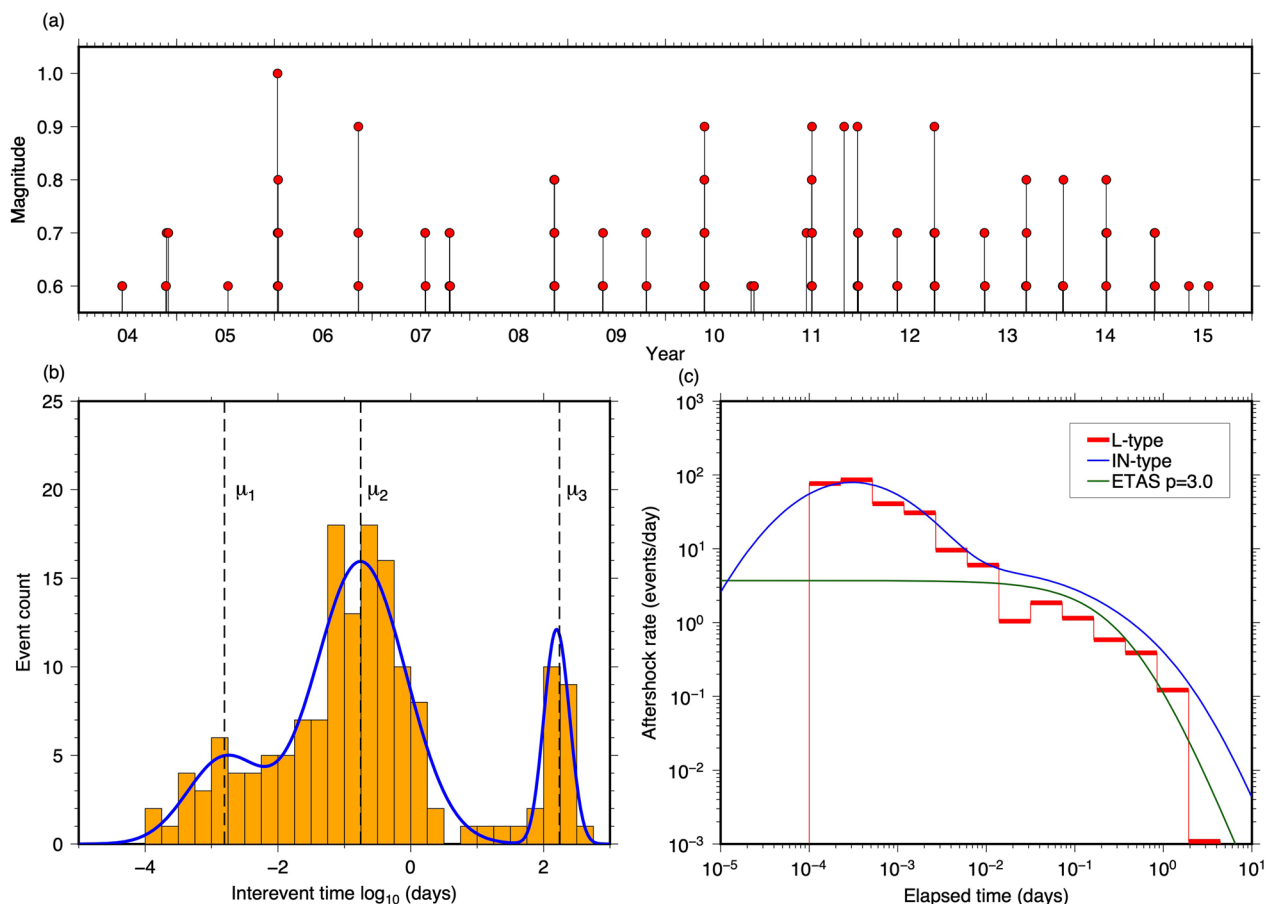Figure 4b shows a histogram of the interevent times in this subregion. The three characteristic time constants of

Fig. 3 LFE activity in a subregion near the Shima Peninsula (34.7°N, 136.5°E; subregion #34). **a** LFE magnitude–time diagram. **b** Interevent time histogram (orange bars). The blue curve indicates the interevent time distribution of the IN-type model. Vertical dashed lines denote the three characteristic time constants of IN-type model (μ_1, μ_2, and μ_3). **c** Aftershock rates predicted by the L-type, IN-type, and ETAS models (red, blue, and green lines, respectively). The maximum likelihood estimate of the ETAS model p value is 3.0

the IN-type model (μ_1, μ_2, and μ_3) were estimated to be 7.2×10^{-4}, 6.2×10^{-1}, and 73 days, respectively. Figure 4c shows the aftershock occurrence rates predicted using the L-type, IN-type, and ETAS models. The ETAS model p value was estimated at 1.0. This value is comparable to the typical p values for regular earthquake activity (Utsu et al. 1995). The aftershock rate decay of the L- and IN-type models stagnated between approximately 3×10^{-2} and 8×10^{-1} days. This feature is similar to that of the subregions in the Bungo Channel and near the Shima Peninsula.

LFE aftershock rates

The results for the three subregions in Sect. "Subregion examples" suggest that the LFE aftershock rates (Figs. 2c, 3c, and 4c) cannot be described by a simple power-law decay and that there is decay stagnation (1.0×10^{-2} to 2.0×10^{-1} days). This is a general feature of the LFE activity in the Nankai Trough. Figure 5 shows the L-type

model median aftershock rates for all 43 subregions. This figure also shows that the decay of the aftershock rate stagnates between 1.0×10^{-2} and 2.0×10^{-1} days. The complex shape of the LFE aftershock rate probably explains why the ETAS model, which assumes a simple power-law decay a priori, is significantly inferior to the L- and IN-type models.

A similar shape for the LFE aftershock rate was suggested in LFE analyses of the San Andreas Fault by Lengliné et al. (2017) and Tan and Marsan (2020). However, for LFEs in the Mexican and Cascadia subduction zones, Lengliné et al. (2017) reported a simple power law decay of $p = 1.3$. The prevalence of stagnation in the LFE aftershock rate decay (Fig. 5) for LFE activity worldwide remains unclear and should be studied in the future.

The physical mechanism of the observed aftershock rate decay stagnation (Fig. 5) is not clear. However, I suggest two hypotheses. One is that triggering between LFEs might involve a characteristic time scale of
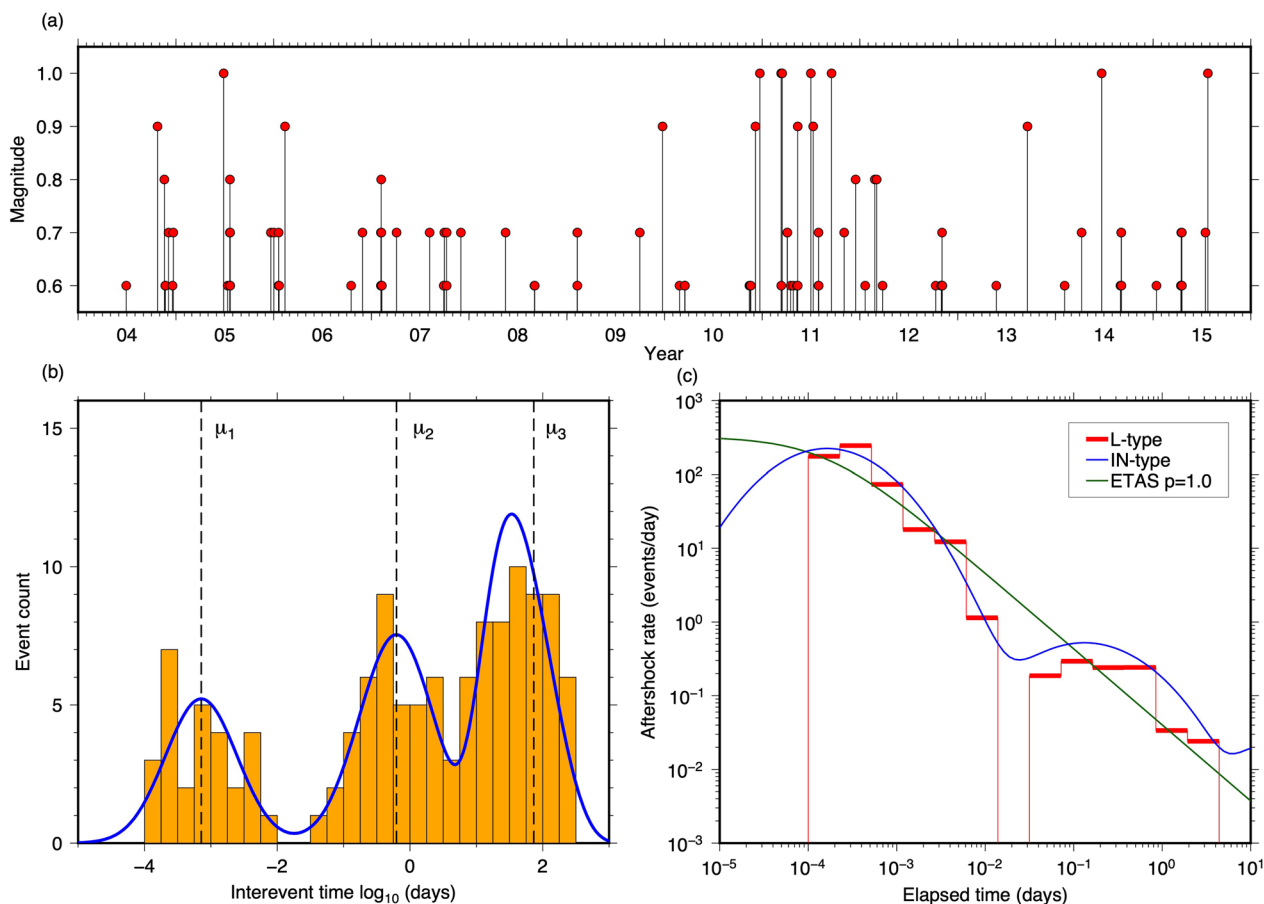
Fig. 4 LFE activity in a subregion in Aichi Prefecture (subregion #43) (35.1°N, 137.4°E). **a** LFE magnitude–time diagram. **b** Interevent time histogram (orange bars). The blue curve indicates the interevent time distribution of the IN-type model. Vertical dashed lines denote the three characteristic time constants of the IN-type model (μ_1, μ_2, and μ_3). **c** Aftershock rates predicted by the L-type, IN-type, and ETAS models (red, blue, and green lines, respectively). The maximum likelihood estimate of the ETAS model *p* value is 1.0
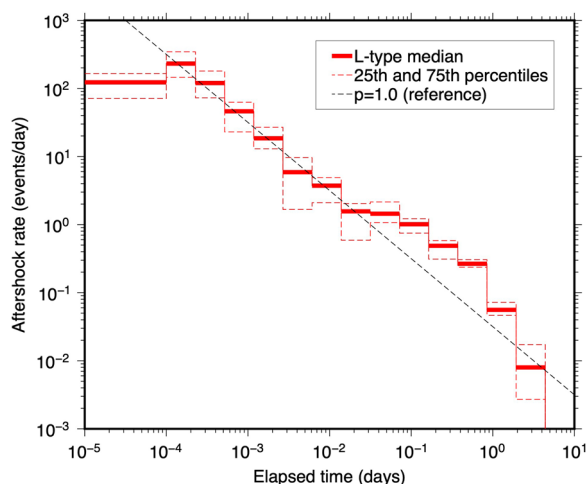
Fig. 5 L-type model median aftershock rates. The solid red line shows the median aftershock rates of all 43 subregions. The dashed red line indicates the 25th and 75th percentiles. The dashed black line indicates a line with an Omori–Utsu *p* value of 1.0

10^{-2}–2×10^{-1} days in its physical mechanism. However, such a characteristic time scale has not been reported for triggering between regular earthquakes, whose primal physical mechanism is considered to be stress transfer (e.g., Ellsworth and Bulut 2018). Physical models of triggering between regular earthquakes (e.g., Dieterich 1994) also do not suggest the existence of such a time scale. Therefore, the physical mechanism of event-to-event triggering may be different for LFEs and regular earthquakes. Another possibility is that the processes behind the LFEs (e.g., slow slip in the region surrounding the LFE asperities) may be intermittent, with characteristic time scales of 10^{-2}–2×10^{-1} days (Kato and Nakagawa 2020). Although slow slips are usually modeled as smooth slow fault-slip events (e.g., Im et al. 2020), actual slow slips may be a series of rapid acceleration/deceleration events with a characteristic recurrence period of 10^{-2}–2×10^{-1} days.
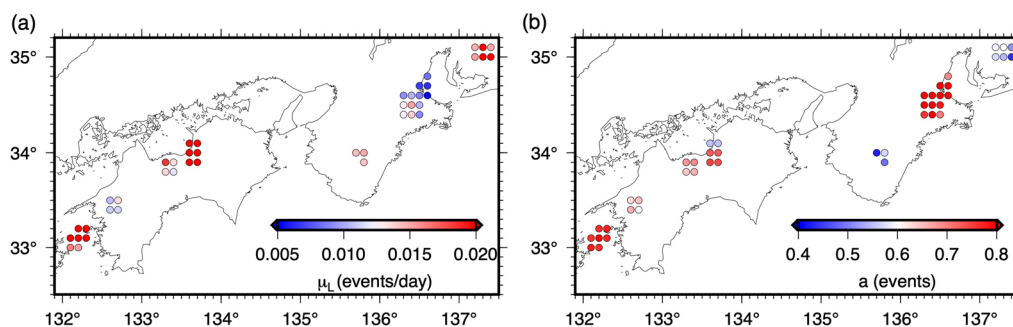
Fig. 6 L-type model parameter spatial distributions. **a**, **b** Distributions of μ_L and a, respectively, see Sect. "Lengliné et al. (2017) model" for parameter details

Spatial model parameter distribution

Figures 6, 7, 8 show the spatial parameter distributions for the L-type, IN-type, and ETAS models. It is difficult to compare parameter values of different models directly, because they depend on different model assumptions. However, common features were observed in the spatial parameter distributions of different models.

One common feature is the large change in LFE activity parameters between the east and west of Ise Bay (approximately 136.8°E). In the L-type model (Fig. 6), both the background seismicity rate μ_L and the number of LFE aftershocks a differed between the east and west of Ise Bay. East of Ise Bay (Aichi Prefecture), the background seismicity rate was higher and the number of aftershocks was lower. In contrast, to the west of Ise Bay (near the Shima Peninsula), the background seismicity rate was low and the number of aftershocks was high. In the IN-type model (Fig. 7), there was a marked contrast in μ_3, σ_3, and ϕ_2 between the east and west of Ise Bay. East of Ise Bay, the long-term clustering time constant (μ_3) was short, and the second short-term clustering fraction (ϕ_2) was small. In contrast, west of Ise Bay, the long-term clustering time constant was longer, and the secondary short-term clustering proportion was greater. Similarly, the ETAS model (Fig. 8) also shows contrasts in μ_E, c, and p between the east and west of Ise Bay. These changes in the LFE activity characteristics may reflect changes in the interplate slip behavior at a spatial scale of approximately 100 km around Ise Bay.

In addition, several parameters show contrasts between the Bungo Channel and western Shikoku (approximately 132.4°E): μ_L in the L-type model; μ_1, μ_3, and ϕ_1 in the IN-type model; and μ_E, c, and p in the ETAS model. μ_L, μ_3, and μ_E are parameters associated with long-term recurrences of LFE bursts. These differences may be related to the occurrence of long-term SSEs in the Bungo Channel (Takagi et al. 2019; Kato and Nakagawa 2020). μ_1, ϕ_1, c, and p are parameters related to LFE short-term clustering, suggesting that there are also differences in short-term LFE clustering features between the Bungo Channel and western Shikoku.

Model performance

Figure 9 shows the AIC difference (ΔAIC) between the IN- and L-type models. The IN-type, L-type, and ETAS models were superior in 25, 13, and 1 of the 43 subregions, respectively. The IN- and L-type models were superior in three of the remaining four subregions, with no significant differences between them. In the remaining subregion, the ETAS and L-type models were superior, with no significant difference between them. These results indicate that although there were more subregions in which the IN-type model performed better, the IN- and L-type model performances were competitive. In particular, the L-type model was superior to the IN-type model in several subregions of Shikoku and east of Ise Bay.

Discussion
Hybrid model

In Sect. "Results", although the IN-type model is superior in more subregions, the IN- and L-type model performances are competitive. Based on these results, this section examines whether the incorporation of both IN- and L-type model features results in a superior model. The IN-type model has the characteristic of using a small number of parameters to represent the LFE occurrence rate (Eqs. 11, 14, and 15). The L-type model considers the LFE activity history (Eq. 9). I propose a new model incorporating both features. Specifically, a model was developed in which the L-type aftershock rate kernel was represented by a small number of parameters, similar to the IN-type model.

In Figs. 2c, 3c, and 4c, the IN- and L-type model aftershock rates show a common feature: stagnation in the decay of the LFE aftershock rate at approximately
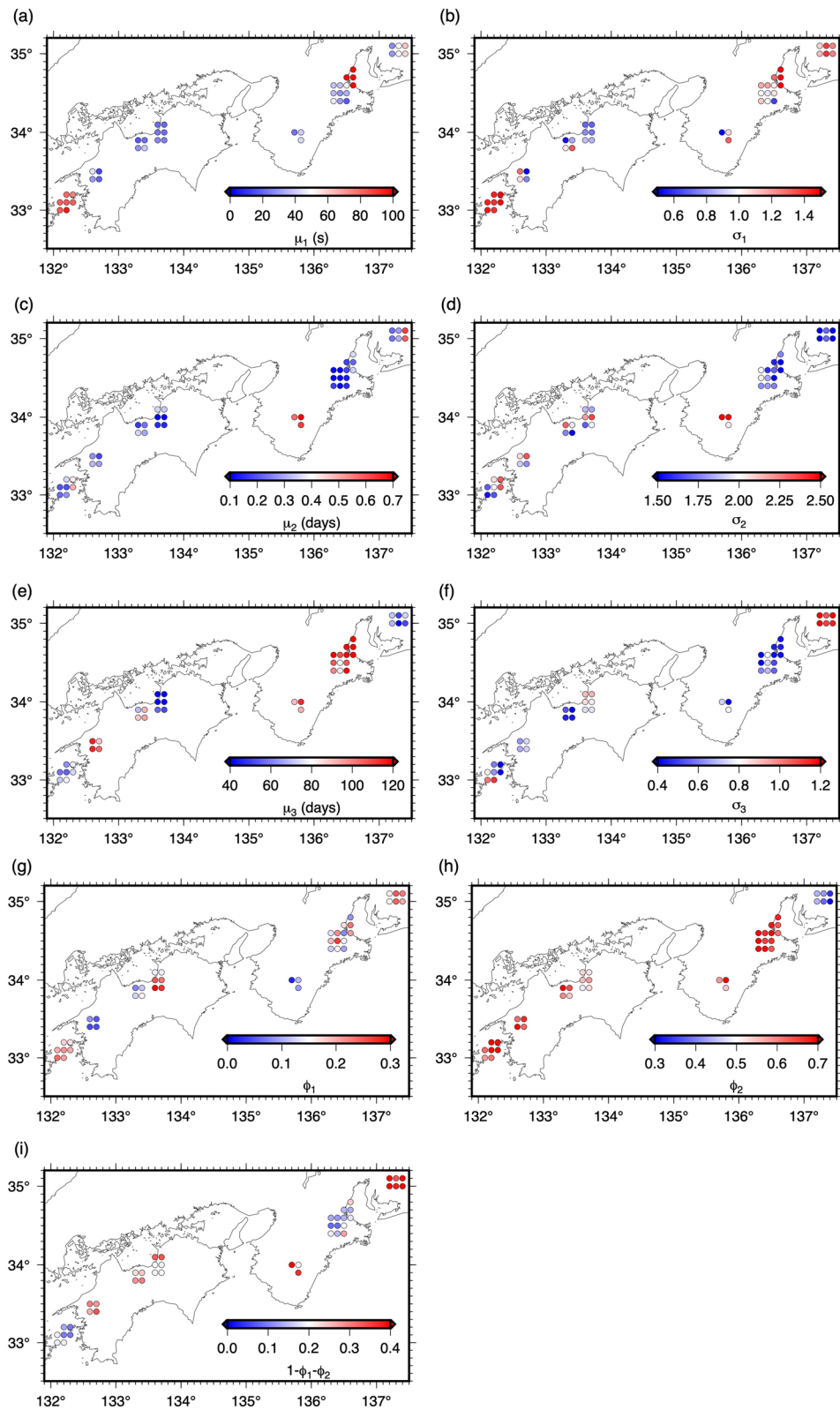
Fig. 7 IN-type model parameter spatial distributions. **a–i** Distributions of μ_1, σ_1, μ_2, σ_2, μ_3, σ_3, ϕ_1, ϕ_2, and $1-\phi_1-\phi_2$, respectively, see Sect. "Ide and Nomura (2022) model" for parameter details
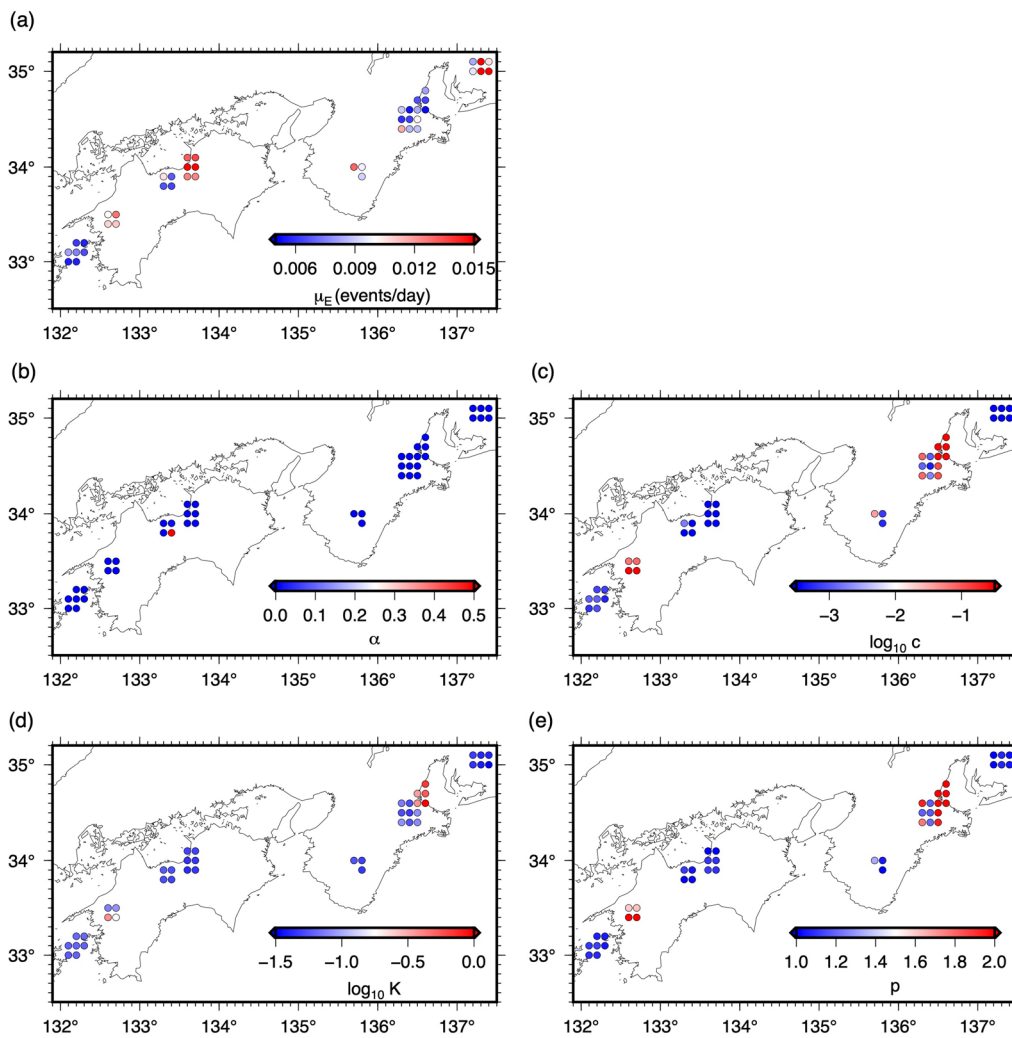
Fig. 8 ETAS model parameter spatial distributions. **a**–**i** Distributions of μ_E, α, c, K, and p, respectively, see Sect. "ETAS model" for parameter details
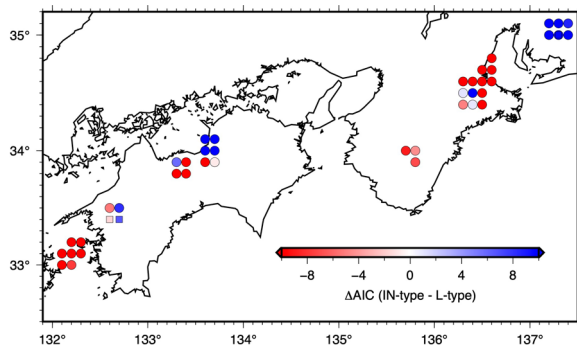

Fig. 9 AIC differences (Δ AIC) between the IN- and L-type models. Colored circles indicate regions where the IN- or L-type model AICs are less than the AICs of the other models. Colored squares indicate regions where the ETAS model AIC is less than the IN- and L-type model AICs

0.1 days. Therefore, a functional form similar to that of the IN-type model may represent the shape of the L-type model aftershock rate kernel well. Note that in the L-type model, the shape of the aftershock rate kernel is not assumed a priori, and it is not necessarily a trivial task to find a functional form with a small number of parameters that represents the L-type model aftershock rates well.

Based on these ideas, the following formulation was developed. First, the following equations for the aftershock rate kernel $g_H(t)$ were assumed:

$$g_H(t) = \frac{f_H(t)}{1 - F_H(t)}, \tag{18}$$

$$f_H(t) = \frac{\phi_{1H}}{\sqrt{2\pi}\sigma_{1H}t}\exp\left\{-\frac{(\log(t)-\log(\mu_{1H}))^2}{2\sigma_{1H}^2}\right\}$$
$$+ \frac{\phi_{2H}}{\sqrt{2\pi}\sigma_{2H}t}\exp\left\{-\frac{(\log(t)-\log(\mu_{2H}))^2}{2\sigma_{2H}^2}\right\},$$
(19)

$$F_H(t) = \int_0^t f_H(t')dt'.$$
(20)

These equations represent the LFE occurrence rate predicted by the IN-type model, excluding the long-term LFE recurrence effect (Eqs. 11, 14, and 15). Specifically, only short-term clustering with two time constants was considered. Using the above equations, the LFE occurrence rate can be written as

$$\lambda(t) = \mu_H + \sum_{t_i < t} g_H(t - t_i).$$
(21)

The above formulation is characterized by a small number of parameters (seven) to represent the LFE occurrence rate and consider the LFE activity history. I call this model a hybrid model. Seven model parameters were estimated (μ_H, μ_{1H}, μ_{2H}, σ_{1H}, σ_{2H}, ϕ_{1H}, and ϕ_{2H}) using the maximum likelihood method (Additional files 1, 6: Fig. S1 and: Table S5).

Figure 10 shows the aftershock rates expected from the hybrid model for the three subregions selected in Sect. "Subregion examples" (orange lines). The hybrid model aftershock rates also showed a stagnation of aftershock rate decay at approximately 0.1 days. The aftershock rates of the L-type and hybrid models are similar, as indicated by the proximity between the red and orange lines in Fig. 10. This suggests that the functional form of the hybrid model aftershock rate kernel (Eqs. 18, 19, 20) well represents the shape of the aftershock rate kernel of the L-type model with fewer parameters.

The AIC was used to compare the model performance. Figure 11a shows the AIC differences (ΔAICs) between the hybrid and L-type models. In most subregions (41 of the 43 subregions), the ΔAIC was < -2, indicating that the hybrid model was significantly superior to the L-type model. In the remaining two subregions, the L-type model performed significantly better than the hybrid model. This may be because in these subregions (subregions #18 and #21), the LFE inter-event time histograms have a more complex shape than the other subregions and cannot be represented by a model with only two time constants for short-term clustering, such as the IN-type and hybrid models (see Additional file 1: Figs. S2 and S3).
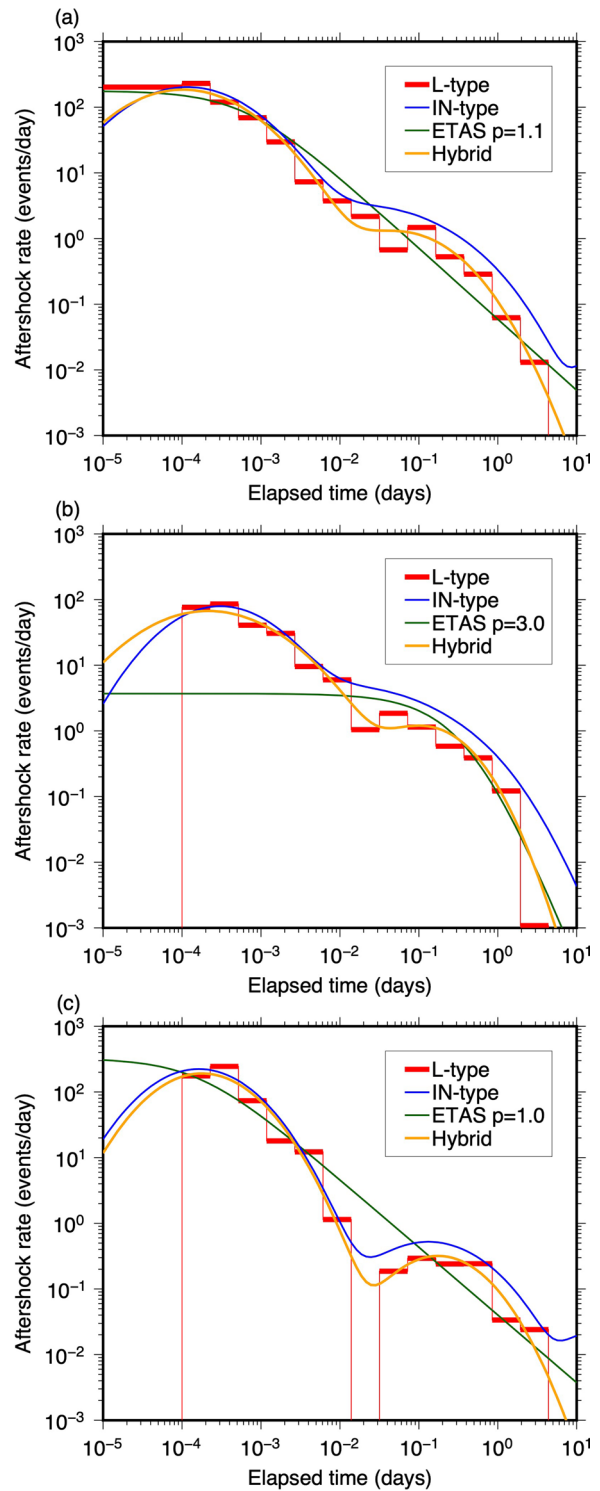

Fig. 10 Aftershock rates predicted by the hybrid model. **a–c** Aftershock rates for subregions in the Bungo Channel (subregion #1; Fig. 2), near the Shima Peninsula (subregion #34; Fig. 3), and in Aichi Prefecture (subregion #43; Fig. 4)
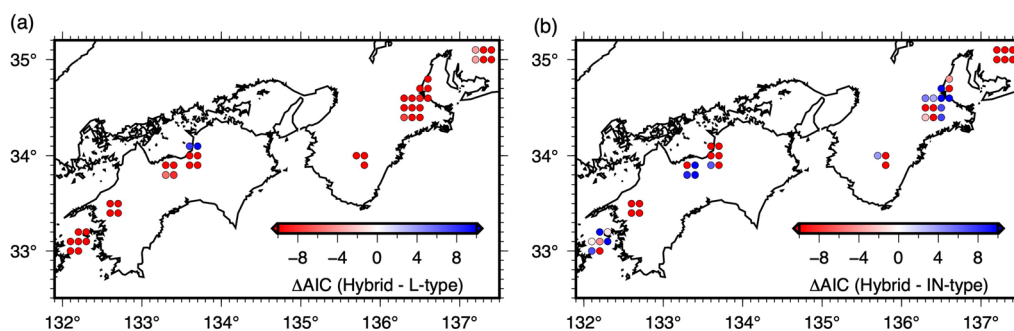
Fig. 11 AIC differences (Δ AIC) between the hybrid, IN-type, and L-type models. **a** Δ AIC between the hybrid and L-type models. **b** Δ AIC between the hybrid and IN-type models

Figure 11b shows the AIC differences between the hybrid and IN-type models. The hybrid model was superior in 26 of the 43 subregions. The IN-type model was superior in 15 of the remaining subregions. There were no significant differences in performance between the remaining two subregions. Compared with the results in Sect. "Results", the number of subregions in which the IN-type model performed the best decreased substantially (from 25 to 15).

This analysis demonstrates that the hybrid model is superior to the L-type model. Furthermore, the hybrid model outperformed the IN-type model in many subregions and could be considered the best model in this study. However, there were still subregions where the IN-type model performed better. This indicates that considering the LFE activity history in a manner similar to the L-type and ETAS models (Eqs. 9, 16, and 21) does not necessarily improve LFE forecasts. This point is discussed in detail in Sect. "Toward further model improvement".

Toward further model improvement
Among the four models used in this study (L-type, IN-type, ETAS, and hybrid), the hybrid model was superior in many regions (Fig. 11). However, there are regions in which the IN-type model outperforms the hybrid model. The hybrid model includes the summation of multiple LFE aftershock rates (Eq. 21). This contrasts with the IN-type model, in which aftershock rates are not affected by detailed LFE activity history and are determined solely by the time elapsed since the last event (Eqs. 11, 14, and 15). Given this difference between the hybrid and IN-type models, in some subregions where the IN-type model outperformed the hybrid model, considering the effect of summing multiple LFE aftershock rates may worsen LFE forecasting.

To further understand the possible absence of the effect of summing multiple aftershock rates, the LFE activity was investigated in detail in the subregions where the

IN-type model outperformed the hybrid model. As a result, it was discovered that, in those subregions, intense LFE bursts were often followed by sudden cessations in LFE activity. In models that apply the summation of multiple aftershock rates, such sudden cessations are difficult to reproduce, because the subsequent predicted seismicity rate increases significantly after an intense burst.

Figure 12 shows an LFE burst in a subregion near the Shima Peninsula (subregion #34; Fig. 3). As shown in Fig. 12a, at approximately 3 days, eight LFEs occurred in 5 h, and the hybrid model predicted a much higher seismicity rate than the IN-type model because of the summation of the LFE aftershock rates (Fig. 12b). Contrary to the hybrid model prediction, this burst stopped at approximately 3.3 days, resulting in a significant decrease in the likelihood of the hybrid model (Fig. 12c). This reduction in likelihood is the main reason the hybrid model is considered inferior to the IN-type model in this subregion.

Considering the above, I suggest that it is important to incorporate the sudden LFE burst cessations for better LFE activity statistical model development. Models such as the L-type, ETAS, and hybrid models, which simply apply the summation of multiple LFE aftershock rates, cannot reproduce these sudden cessations. New models are required that consider LFE activity history differently from these models and successfully predict LFE burst cessations.

The sudden cessations could be reproduced if I consider a different variable (a_i) for each LFE and change the amplitude of the LFE aftershock rates, as in the original Lengliné et al. (2017) model (Sect. "Lengliné et al. (2017) model"). However, such a model would be inferior in terms of AIC, because the number of model parameters would increase substantially. Therefore, when varying the amplitude of the LFE aftershock rates, it is necessary to model the variation with a small number of parameters.
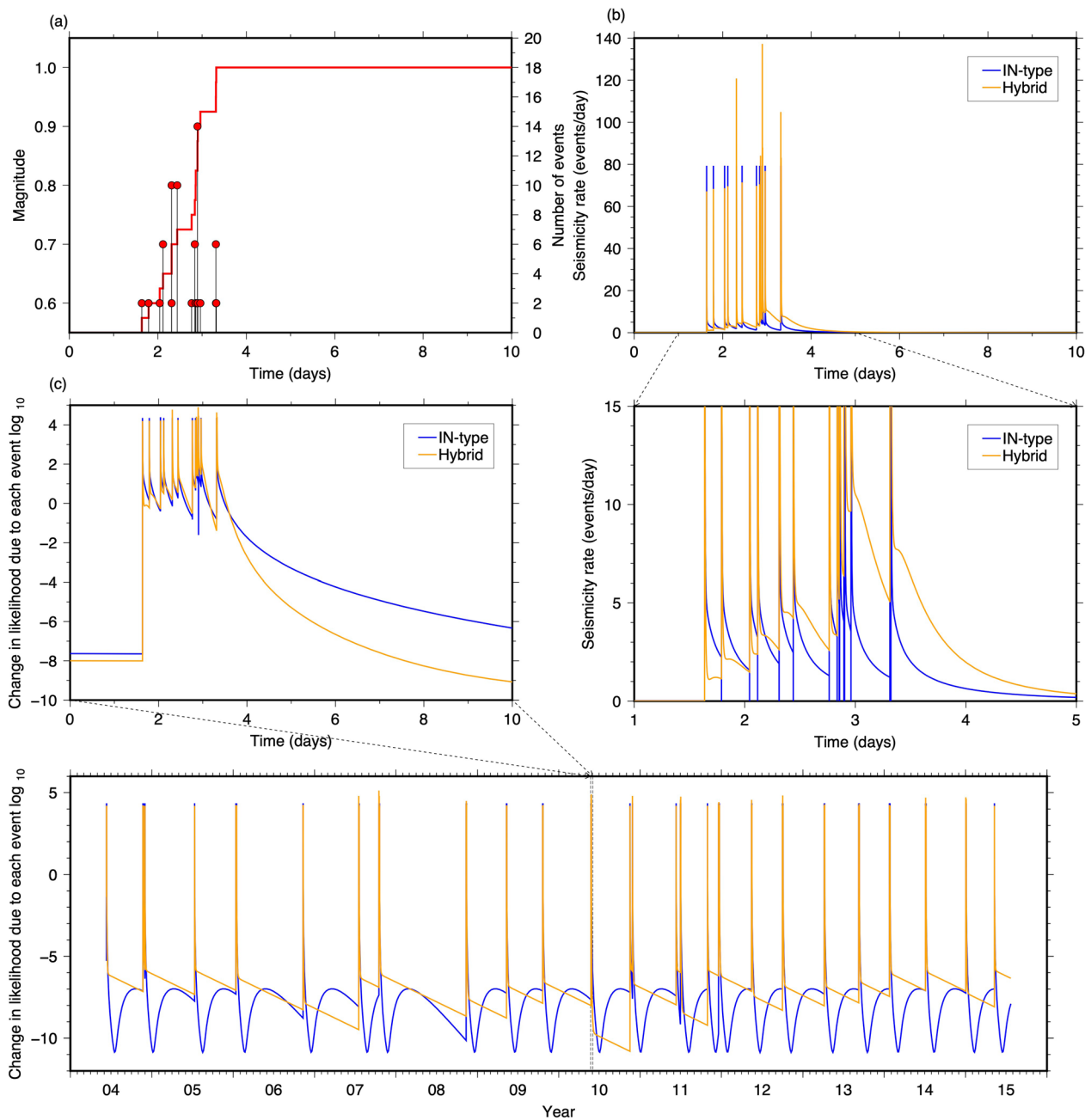
Fig. 12 Example of LFE activity in a subregion where the IN-type model outperforms the hybrid model. **a** LFE magnitude–time diagram and cumulative LFE counts in subregion #34. The date of zero is May 24, 2010. **b** LFE seismicity rates predicted by the IN-type and hybrid models. The lower panel is an enlarged view of the upper panel. **c** Change in the likelihood of the IN-type and hybrid models due to the occurrence of each event. Spike-shaped sections correspond to LFE occurrence times. Two successive spikes correspond to two successive LFEs, and the curves between them show how the model likelihood changes with the later LFE occurrence time. The actual change in likelihood due to the occurrence of the later LFE can be read from the value that the curves indicate immediately prior to the later spike. The upper left panel is an enlarged view of the lower panel

In addition to the additive nature of the LFE aftershock rates, there was another difference between the hybrid and IN-type models. The hybrid model assumes a stationary background seismicity rate (Eq. 21), whereas the IN-type model considers long-term LFE burst recurrences using a BPT distribution (Eq. 11). This difference may also cause a difference in the performance of the two models.

Transformed time sequence

The number of events $\Lambda(t)$ predicted from a model from t_0 to t is obtained by integrating the event occurrence rate $\lambda(t)$:

$$\Lambda(t) = \int_{t_0}^{t} \lambda(t')dt'. \tag{22}$$

I use Eq. 22 to transform the occurrence time of the ith event t_i into a transformed time $\tau_i \equiv \Lambda(t_i)$ (Ogata 1988, 1992).

If the model can predict events well, the transformed time sequence follows a standard Poisson process, and the plot of the cumulative event count against the transformed time is expected to be linear with a slope of unity (Ogata 1988, 1992). Proximity to the straight line with a slope of unity is usually considered as an indicator of good fitting. However, in this study, the transformed time sequence did not seem to be suitable for visualizing the superiority of model performance.

The proximity to the line with a slope of unity does not necessarily indicate large likelihood or small AIC. For example, if a model repeatedly underestimates and overestimates the event occurrence rate over a short time period, then if the event occurrence rate is integrated over a long period, the predicted and observed number of events will be close and follow the straight line with a slope of unity. However, the model likelihood becomes small, and the AIC becomes large, because the event occurrence rate is underestimated or overestimated for each event. This phenomenon was often observed in this study: in Additional file 1: Fig. S4a, b, the L-type, ETAS, and hybrid models were closer to the straight line with a slope of unity than the IN-type model, but the IN-type model had the smallest AIC. This is because the L-type, ETAS, and hybrid models repeatedly underestimated the event occurrence rate when LFE clusters occurred and overestimated it when the clusters stopped. In other words, the L-type, ETAS, and hybrid models failed to predict the occurrence of individual events. These results are consistent with the point made in Sect. "Toward further model improvement" that the L-type, ETAS, and hybrid models do not predict LFE cluster activity well.

Selection of the minimum magnitude

Whether the LFE magnitude–frequency distribution (especially for large events in a catalog) follows the Gutenberg–Richter law (a power law distribution) or an exponential distribution is still under debate and inconclusive (Bostock et al. 2015; Chestler and Creager 2017). Hence, no method currently exists for evaluating the completeness of an LFE catalog.

In seismicity analysis, the use of incomplete catalogs makes it unclear whether the results obtained are true seismicity features or artifacts due to the incompleteness of the catalogs. In light of this, in this study, I followed Bostock et al. (2015), assumed the Gutenberg–Richter law, and used as large a minimum magnitude as possible (M 0.6). This is as conservative an analysis as possible. My intention is to eliminate artifacts as much as possible and to discuss the true features of LFE activity, even if my analysis is limited to the activity of large LFEs.

However, the conservative approach of this study has the problem of ignoring the activity of many LFEs smaller than the minimum magnitude. Therefore, I conducted a reanalysis by changing the LFE minimum magnitude to M 0.4 and M 0.2. The number of events used in the analysis increased to 18,103 (5 times greater than the number of M 0.6 or larger LFEs) and 67,193 (19 times greater than the number of M 0.6 or larger LFEs), respectively. Because the number of LFEs increased substantially as the minimum magnitude was lowered, the computational cost of this study also increased substantially (on the order of the square of the number of LFEs). Therefore, the analysis could not be performed for the minimum magnitude below M 0.2.

As a result, I found two common features among the results for the minimum magnitudes of M 0.6, M 0.4, and M 0.2. The first is that the performance of the IN-type model and the L-type model are competitive, although the IN-type model performs better in more subregions (Figs. 9, 13a, b). The second is that the Hybrid model is generally superior to the L-type model (Figs. 11a, 13c, d).

In contrast, the comparisons between the hybrid and IN-type models (Figs. 11b, 13e, f) showed an increase in subregions where the IN-type model outperformed the hybrid model for the minimum magnitude of M 0.2. For the minimum magnitude of M 0.2, the performance of the hybrid and IN-type models is comparable: the hybrid model was superior in 23 subregions, and the IN-type model was superior in 20 subregions. This may be because the effect of the summation of multiple aftershock rates assumed by the hybrid model does not describe the clustering of LFEs well, as suggested in Sect. "Toward further model improvement". When the minimum magnitude is small, the likelihood and AIC strongly reflect the ability to predict short-term LFE clustering,

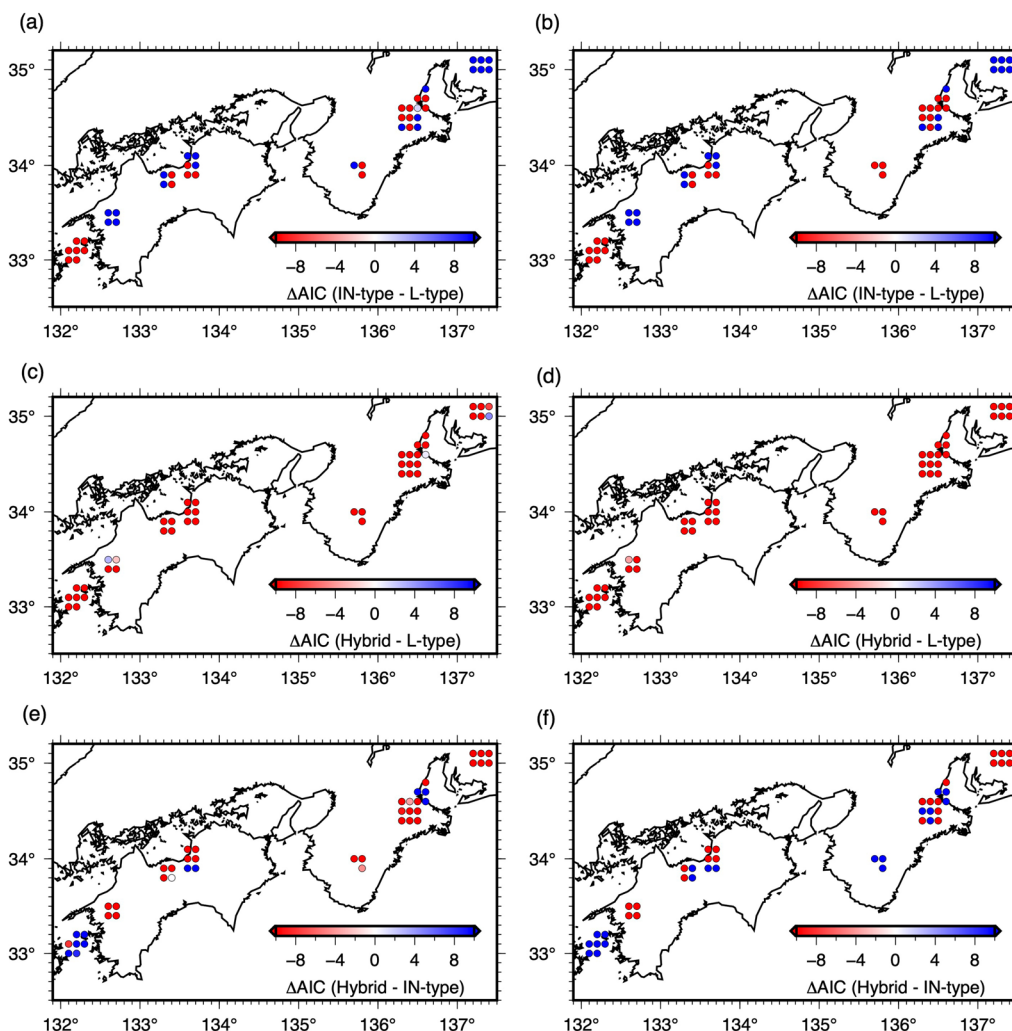**Fig. 13** AIC differences (Δ AIC) between the hybrid, IN-type, and L-type models for the minimum magnitudes of *M* 0.4 and *M* 0.2. **a** and **b** Δ AIC between the IN-type and L-type models for the minimum magnitudes of *M* 0.4 and *M* 0.2. **c** and **d** Δ AIC between the hybrid and L-type models for the minimum magnitudes of *M* 0.4 and *M* 0.2. **e** and **f** Δ AIC between the hybrid and IN-type models for the minimum magnitudes of *M* 0.4 and *M* 0.2

because numerous small LFEs are concentrated over short time periods.

The selection of the optimal minimum magnitude for statistical analysis of LFE activity is an important future issue. It is desirable to use as many LFEs as possible in the analysis while ensuring the completeness of the LFE catalog. In future studies, it is important to develop a method to evaluate the completeness of LFE catalogs and to select an appropriate minimum magnitude.

Short-term aftershock incompleteness
Overlooking small aftershocks due to intense aftershock activity after large mainshocks is called short-term aftershock incompleteness (STAI) (e.g., Petrillo and Zhuang 2023). For regular earthquake activity, we can easily visually identify STAI in the magnitude–time diagram. In contrast, for LFE activity, STAI is not visually clear in the magnitude–time diagram, because the distinction between mainshocks and aftershocks is unclear for LFE activity.

However, smaller LFEs may be overlooked during intense LFE bursts. To examine this possibility, for *M* 0.4 and *M* 0.8 LFEs, I checked the elapsed time from the immediately preceding event to the occurrence of each LFE. In Additional file 1: Fig. S5, I showed the histograms of the elapsed time. Counts were summed and

normalized for all 43 subregions. Note that the minimum magnitude of M 0.4 was used in this analysis.

The results showed that M 0.4 LFEs had a smaller percentage of events occurring in less than 10^{-3} days (approximately 90 s or less) compared to M 0.8 LFEs. This is consistent with the hypothesis that during intense LFE bursts (i.e., periods characterized by very short interevent times), small LFEs may be overlooked.

Much is still not known about the temporal variations of the LFE catalog completeness and LFE magnitude–frequency distribution. Further research on STAI for LFE activity is needed.

Factors not considered by the models used in this study

Some important LFE activity features are not considered in the models used in this study (L-type, IN-type, ETAS, and hybrid models). For example, the LFE activity exhibits a high tidal response (e.g., Royer et al. 2015). It is also known that LFE epicenters migrate diffusively on a timescale of approximately 10 days and a spatial scale of approximately 100 km; they also migrate diffusively or at a constant speed on a timescale of approximately 30 min and a spatial scale of 15 km (Kato and Nakagawa 2020). Statistical LFE activity models could be improved by considering these factors. However, as stated in Sect. "Introduction", this study aims to compare existing statistical LFE activity models, identify which model better describes LFE activities, and create a new model based on this comparison. It is beyond the scope of this study to incorporate new factors that are not considered in existing models into each model. This is an important issue that should be addressed in future studies.

Conclusions

This study applied the existing statistical LFE activity models (L- and IN-type models) (Lengliné et al. 2017; Ide and Nomura 2022) to LFEs along the Nankai Trough (Kato and Nakagawa 2020) and conducted the first comparison based on AIC (Sect. "Results"). The aftershock rate decay predicted by the L- and IN-type models stagnated near an elapsed time of 0.1 days. This is the primary reason why the ETAS model (Ogata 1988), which a priori assumes a simple power-law decay, does not perform well for LFE activity.

Although there were more subregions in which the IN-type model performed better, the IN- and L-type model performances were competitive. Based on these results, a new hybrid model was created that incorporates the features of both models (Sect. "Discussion"). The hybrid model considers the LFE activity history in a manner similar to the L-type and ETAS models, and represents the LFE aftershock rate with a small number of parameters, similar to the IN-type model. The results showed that the hybrid model outperformed the L- and IN-type models in many subregions, and the hybrid model was considered to be the best currently available model.

However, I found that the hybrid model could not reproduce the sudden LFE burst cessations, because the model summed multiple LFE aftershock rates. In subregions where such a cessation often occurs, the IN-type model, which does not consider the LFE activity history, outperformed the hybrid model. For better LFE activity forecasting, new models that consider LFE activity history differently from the hybrid model and successfully predict the abrupt LFE burst cessations are required.

Moreover, the empirical equations formulated in this study for LFE aftershock rates (Eqs. 18, 19, 20) may be useful for future investigations concerning statistical and physical modeling of LFE activity. Specifically, they could be an important observational fact that future physical LFE activity models should reproduce.

Abbreviations
AIC   Akaike's information criterion
BPT   Brownian passage time
ETAS   Epidemic-type aftershock-sequence
GFT   Goodness-of-fit test
JMA   Japan Meteorological Agency
LFE   Low-frequency earthquake
SSE   Slow-slip event
ΔAIC   AIC difference

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-024-02007-6.

Additional file 1. Supplementary figures.

Additional file 2: Table S1. L-type model parameter estimates.

Additional file 3: Table S2. IN-type model parameter estimates.

Additional file 4: Table S3. ETAS model parameter estimates.

Additional file 5: Table S4. AICs for all models.

Additional file 6: Table S5. Hybrid model parameter estimates.

Availability of data and materials
The LFE catalog is from Kato and Nakagawa (2020) and is included in its supplementary material. The seafloor topography data (Smith and Sandwell 1997) are available at https://topex.ucsd.edu/marine_topo/. The figures were prepared used the Generic Mapping Tools software package (Wessel et al. 2019), which are available at https://www.generic-mapping-tools.org.

Declarations

Ethics approval and consent to participate
Not applicable.

Author details
[1]Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723. https://doi.org/10.1109/TAC.1974.1100705

Beroza GC, Ide S (2011) Slow earthquakes and nonvolcanic tremor. Annu Rev Earth Planet Sci 39:271–296. https://doi.org/10.1146/annurev-earth-040809-152531

Bostock MG, Thomas AM, Savard G, Chuang L, Rubin AM (2015) Magnitudes and moment-duration scaling of low-frequency earthquakes beneath southern Vancouver Island. J Geophys Res 120(9):6329–6350. https://doi.org/10.1002/2015JB012195

Chestler SR, Creager KC (2017) Evidence for a scale-limited low-frequency earthquake source process. J Geophys Res 122(4):3099–3114. https://doi.org/10.1002/2016JB013717

Dieterich J (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. J Geophys Res 99(B2):2601–2618. https://doi.org/10.1029/93JB02581

Dragert H, Wang K, James TS (2001) A silent slip event on the deeper Cascadia subduction interface. Science 292(5521):1525–1528. https://doi.org/10.1126/science.106015

Ellsworth WL, Bulut F (2018) Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. Nat Geosci 11:531–535. https://doi.org/10.1038/s41561-018-0145-1

Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34(4):185–188. https://doi.org/10.1785/BSSA0340040185

Ide S, Beroza GC (2023) Slow earthquake scaling reconsidered as a boundary between distinct modes of rupture propagation. Proc Natl Acad Sci 120(32):e2222102120. https://doi.org/10.1073/pnas.2222102120

Ide S, Nomura S (2022) Forecasting tectonic tremor activity using a renewal process model. Prog Earth Planet Sci 9:67. https://doi.org/10.1186/s40645-022-00523-1

Ide S, Beroza GC, Shelly DR, Uchide T (2007) A scaling law for slow earthquakes. Nature 447(7140):76–79. https://doi.org/10.1038/nature05780

Im K, Saffer D, Marone C, Avouac JP (2020) Slip-rate-dependent friction as a universal mechanism for slow slip events. Nat Geosci 13(10):705–710. https://doi.org/10.1038/s41561-020-0627-9

Kato A, Nakagawa S (2020) Detection of deep low-frequency earthquakes in the Nankai subduction zone over 11 years using a matched filter technique. Earth Planet Space 72:128. https://doi.org/10.1186/s40623-020-01257-4

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 118(12):6165–6182. https://doi.org/10.1002/2013JB010259

Kuna VM, Nábělek JL, Braunmiller J (2019) Mode of slip and crust–mantle interaction at oceanic transform faults. Nat Geosci 12(2):138–142. https://doi.org/10.1038/s41561-018-0287-1

Lengliné O, Frank WB, Marsan D, Ampuero JP (2017) Imbricated slip rate processes during slow slip transients imaged by low-frequency earthquakes. Earth Planet Sci Lett 476:122–131. https://doi.org/10.1016/j.epsl.2017.07.032

Luo Y, Liu Z (2019) Slow-slip recurrent pattern changes: perturbation responding and possible scenarios of precursor toward a megathrust earthquake. Geochem Geophys Geosyst 20(2):852–871. https://doi.org/10.1029/2018GC008021

Matsuzawa T, Hirose H, Shibazaki B, Obara K (2010) Modeling short- and long-term slow slip events in the seismic cycles of large subduction

earthquakes. J Geophys Res 115:B12301. https://doi.org/10.1029/2010JB007566

McGuire JJ, Boettcher MS, Jordan TH (2005) Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults. Nature 434(7032):457–461. https://doi.org/10.1038/nature03377

Nadeau RM, Dolenc D (2005) Nonvolcanic tremors deep beneath the San Andreas Fault. Science 307(5708):389–389. https://doi.org/10.1126/science.110714

Nishikawa T, Nishimura T (2023) Development of an epidemic-type aftershock-sequence model explicitly incorporating the seismicity-triggering effects of slow slip events. J Geophys Res 128(5):e2023JB026457. https://doi.org/10.1029/2023JB026457

Nishikawa T, Matsuzawa T, Ohta K, Uchida N, Nishimura T, Ide S (2019) The slow earthquake spectrum in the Japan Trench illuminated by the S-net seafloor observatories. Science 365(6455):808–813. https://doi.org/10.1126/science.aax5618

Nishikawa T, Ide S, Nishimura T (2023) A review on slow earthquakes in the Japan Trench. Prog Earth Planet Sci 10:1. https://doi.org/10.1186/s40645-022-00528-w

Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. Science 296(5573):1679–1681. https://doi.org/10.1126/science.107037

Obara K, Kato A (2016) Connecting slow earthquakes to huge earthquakes. Science 353(6296):253–257. https://doi.org/10.1126/science.aaf1512

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. J Am Stat Assoc 83(401):9–27. https://doi.org/10.1080/01621459.1988.10478560

Ogata Y (1992) Detection of precursory relative quiescence before great earthquakes through a statistical model. J Geophys Res 97(B13):19845–19871. https://doi.org/10.1029/92JB00708

Petrillo G, Zhuang J (2023) Verifying the magnitude dependence in earthquake occurrence. Phys Rev Lett 131(15):154101. https://doi.org/10.1103/PhysRevLett.131.154101

Powell MJ (1964) An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput J 7(2):155–162. https://doi.org/10.1093/comjnl/7.2.155

Royer AA, Thomas AM, Bostock MG (2015) Tidal modulation and triggering of low-frequency earthquakes in northern Cascadia. J Geophys Res 120(1):384–405. https://doi.org/10.1002/2014JB011430

Schorlemmer D, Werner MJ, Marzocchi W, Jordan TH, Ogata Y, Jackson DD et al (2018) The collaboratory for the study of earthquake predictability: achievements and priorities. Seismol Res Lett 89(4):1305–1313. https://doi.org/10.1785/0220180053

Shelly DR, Beroza GC, Ide S (2007) Non-volcanic tremor and low-frequency earthquake swarms. Nature 446(7133):305–307. https://doi.org/10.1038/nature05780

Smith WH, Sandwell DT (1997) Global sea floor topography from satellite altimetry and ship depth soundings. Science 277(5334):1956–1962. https://doi.org/10.1126/science.277.5334.1956

Takagi R, Uchida N, Obara K (2019) Along-strike variation and migration of long-term slow slip events in the western Nankai subduction zone, Japan. J Geophys Res 124(4):3853–3880. https://doi.org/10.1029/2018JB016738

Takemura S, Hamada Y, Okuda H, Okada Y, Okubo K, Akuhara T et al (2023) A review of shallow slow earthquakes along the Nankai Trough. Earth Planet Space 75:164. https://doi.org/10.1186/s40623-023-01920-6

Tan YJ, Marsan D (2020) Connecting a broad spectrum of transient slip on the San Andreas fault. Sci Adv 6(33):abb2489. https://doi.org/10.1126/sciadv.abb2489

Utsu T (1957) Magnitudes of earthquakes and occurrence of their aftershocks. Zisin 10(2):35–45

Utsu T, Ogata Y, Matsu'ura R (1995) The centenary of the Omori formula for a decay law of aftershock activity. J Phys Earth 43(1):1–33. https://doi.org/10.4294/jpe1952.43.1

Wallace LM (2020) Slow slip events in New Zealand. Annu Rev Earth Planet Sci 48:175–203. https://doi.org/10.1146/annurev-earth-071719-055104

Wessel P, Luis JF, Uieda L, Scharroo R, Wobbe F, Smith WHF, Tian D (2019) The generic mapping tools version 6. Geochem Geophys Geosyst 20:5556–5564. https://doi.org/10.1029/2019GC008515

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(4):859–869. https://doi.org/10.1785/0119990114

Woessner J, Wiemer S (2005) Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty. Bull Seismol Soc Am 95(2):684–698. https://doi.org/10.1785/0120040007

Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. J Am Stat Assoc 97(458):369–380. https://doi.org/10.1198/016214502760046925

Publisher's Note