## Characteristics of Foreshocks Revealed by an Earthquake Forecasting Method Based on Precursory Swarm-like Activity

F. Hirose[1], K. Tamaribuchi[1], and K. Maeda[2]

[1]Seismology and Tsunami Research Department, Meteorological Research Institute, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan.

[2]Seismology and Volcanology Department, Japan Meteorological Agency, 3-6-9 Toranomon, Minato-ku, Tokyo 105-8431, Japan.

**Contents of this file**

## Introduction

This document contains supporting information in the form of text, figures, and tables. Text S1, Figures S1–S5, and Tables S1–S2 provide additional details about seismicity in each study region. Text S2, Figures S6–S7, and Tables S3–S4 give additional details about constructing ETAS forecast model. Text S3, Figures S8–13, and Table S5 give more information about the comparison of real data with the ETAS catalogs along the Japan trench. Text S4 give additional details about constructing ETAS catalogs. Finally, Text S5 and Figure S14 give additional explanation for Molchan's error diagram.

**Text S1. Seismicity**

**Text S1.1. Along the Japan Trench (Off Iwate and Miyagi Prefectures, and Off Ibaraki Prefecture)**
   The Pacific plate moves west toward the Japanese Islands and subducts at the Japan trench. Earthquakes occur in and around the plate boundary associated with the plate motion.

Maeda (1996) used the JMA catalog before 1993 to detect areas of foreshocks along the Kuril and Japan trenches. Subsequently, Maeda and Hirose (2016) conducted another analysis with added data and showed that foreshocks occurred in two areas in particular (specific areas hereafter), which are enclosed by dashed lines in Figure 1.

In 1960, JMA deployed type-59 seismographs throughout Japan, thereby improving the detectability of earthquakes in ocean areas. An $M_w$ 9.0 earthquake occurred off Miyagi prefecture on 11 March 2011, and intensive aftershock activity occurred along the Japan trench (Hirose et al., 2011). Because this study focused on foreshocks, we used only data up to the end of 2010 that is the same period (1961–2010) as Maeda and Hirose (2016). Although the largest foreshock, $M$ 7.3, of the $M_w$ 9.0 earthquake occurred off Miyagi prefecture two days before the mainshock, we did not include it in our evaluation. Unlike Maeda and Hirose (2016), we took into consideration the flags S, s, and a, which indicate the precision of hypocenter location (see Section 2 in the main text). We evaluated earthquakes of $M \geq 5.0$ in the specific areas. In constructing ETAS catalogs, it is better to set wide ranges for time, space, and magnitude because aftershocks of earthquakes occurring outside the specific areas may themselves occur inside those areas (Ogata & Katsura, 2014). Therefore, in constructing the ETAS catalogs, we used a magnitude threshold lower by 0.2 than $M \geq 5.0$ (i.e., $M \geq 4.8$) for earthquakes occurring from about 0.2 years before 1961 (i.e., from 1 October 1960) (Table S1). In the same way, the specific area off Iwate and Miyagi prefectures was expanded by 1.0° in each cardinal direction (Figure S1), and the specific area off Ibaraki prefecture was similarly expanded by 0.5° (Figure S2). See Section 3.2 for further information about the construction of ETAS catalogs.

Because hypocenter depths of offshore earthquakes in the JMA catalog tend to be deeper than the true hypocenter depth (Gamage et al., 2009), we used all earthquakes shallower than 100 km so as not to lose information.


**Text S1.2. Central Honshu**

The Itoigawa-Shizuoka Tectonic Line (ISTL) occupies a part of the boundary between the Amur and Okhotsk plates, and the Niigata-Kobe Tectonic Zone (NKTZ) is a concentrated deformation zone where the strain rate is 0.1 ppm/y (Sagiya et al., 2000). Earthquake activity related to these tectonics occurs in central Honshu (Figure S3). According to the Headquarters for Earthquake Research Promotion (2015a), which has divided the ISTL into four sections, the recurrence interval along the three northern sections is about 1,000 years, and it is about 5,000 years along the southernmost section. The magnitude of an anticipated earthquake affecting an individual section would be $M$ 7.4–7.7, whereas if all sections ruptured in the same event, the magnitude would be up to $M$ 7.8–8.1. The relationship between seismicity extending northeast from the ISTL and the "western marginal fault zone of the Nagano Basin" is not clear (Headquarters for Earthquake Research Promotion, 2015b). Seismicity parallel to the ISTL on its west side appears to correspond to active volcanos rather than to known active faults. Many large earthquakes of $M$ 7–8 have occurred in the NKTZ over the past several hundred years (Sagiya et al., 2000).

Since 1 October 1997, JMA has compiled earthquake records from all institutions and universities into a single catalog and determined earthquake hypocenters, thus improving the detectability of earthquakes. In this study, we evaluated earthquakes of $M \geq 2.0$ shallower than 30 km for 1998–2019 in the square area shown in Figure S3a. See Table S2 for the dependency of segment parameter $D$.

Figure S4 shows the spatiotemporal relationship between alarm earthquakes and event C (March 12, 2011, 3:59, *M*6.7 in Figure 7c) which occurred about 13 hours after the Tohoku-oki Earthquake (*M*$_w$ 9.0). Maeda's method automatically judged event C to be an alarmed target earthquake. However, only one foreshock candidate had occurred in the aftershock area of event C (crosses in Figure S4); in addition $N_f$ = 5, indicating that event C had little relation to the alarm earthquake and swarm activity around 36.8°N.

**Text S1.3. Izu Islands**

The Izu-Bonin island arc is part of the Philippine Sea plate, and it reached approximately its current location at 15–17 Ma in association with the evolution of the Shikoku Basin (Seno & Maruyama, 1984; Okino et al., 1994). Since then, the island arc has continued to move north toward the Japanese Islands, and it is colliding with the base of Izu Peninsula (Takahashi, 1994). The middle crust is thickening beneath the active basaltic volcanoes along the Shichito-Ioto ridge (Kodaira et al., 2007), and earthquakes occur along the Shichito-Ioto and Zenisu ridges (Figure 1). Two months of active seismicity, including six *M* 6-class earthquakes, in and around Miyake-jima began in late June 2000, and was accompanied by the eruption of Miyake-jima volcano (Japan Meteorological Agency, 2006). Although seismicity and eruptive activity ended after late September 2000, inhabitants of Miyake-jima were forced to remain away from the island until 2005 due to continued toxic volcanic gas emissions (Japan Meteorological Agency, 2006).

JMA deployed type-76 seismographs throughout Japan in 1976, thereby further improving the detectability of earthquakes. In this study, we evaluated earthquakes of *M* ≥ 3.0 with hypocenters shallower than 50 km occurring during 1977–2019 in the rectangular area shown in Figure S5a. Earthquakes occurring during only a few months in 2000 account for more than half of the total number during the 43 years. The seismicity rate after 2000, was only one-tenth of the rate before 2000. See Table S2 for the dependency of segment parameter *D*.

**Text S2. Earthquake forecast based on the stationary space-time ETAS model**

As a standard earthquake forecast model to be compared with Maeda's method, we adopted an earthquake forecast model based on the stationary space-time ETAS model (Ogata and Zhuang, 2006) (hereinafter ETAS forecast model). In this section, we describe how to construct and evaluate the ETAS forecast model.

**Text S2.1. Stationary space-time ETAS model**

The R "ETAS" package coded by Jalilian (2019) was used to estimate parameters $\theta = (\mu, K, \alpha, c, p, d, q, \gamma)$ that are temporally constant and the spatial smoothing function of the background seismicity rate $u(x, y)$ in the stationary space-time ETAS model (equations (S1)–(S5)).

The seismicity rate $\lambda$ at time $t$ and the location $(x, y)$ is calculated as:

$$\lambda(t, x, y|H_t) = \mu u(x, y) + \sum_{t_i < t} \kappa(m_i|K, \alpha) h(t - t_i|c, p) f(x - x_i, y - y_i, m_i|d, q, \gamma), \quad (S1)$$

where $H_t$ is the history of occurrence times $t_i$ up to time $t$ with corresponding epicenters $(x_i, y_i)$ and magnitude $m_i$. The first term on the right hand side of equation (S1) is the background seismicity rate, which is assumed to be independent of time; the units of $u(x, y)$ are event/day/degree². The second term on the right hand side, which defines the spatiotemporal decay of aftershocks, is composed of equations (S2)–(S5):

The expected number of triggered events (aftershocks) is

$$\kappa(m|K, \alpha) = Ke^{\alpha(m - M_{\text{th}})}, \qquad (S2)$$

where $M_{\text{th}}$ is the magnitude threshold and $m$ is the magnitude of an earthquake. A large value of $\alpha$ indicates a mainshock-and-aftershock-type earthquake, and a small value of $\alpha$ indicates a swarm-type earthquake.

Equation (S3) describes the probability density function of the occurrence times of triggered events.

$$h(t|c, p) = \frac{(p - 1)c^{p-1}}{(t + c)^p}, \qquad (S3)$$

where the units of $c$ are days.

Equation (S4) describes the probability density function of the locations of the triggered events.

$$f(x, y, m|d, q, \gamma) = \frac{q - 1}{\pi\sigma(m|d, \gamma)}\left(1 + \frac{x^2 + y^2}{\sigma(m|d, \gamma)}\right)^{-q}, \qquad (S4)$$

where

$$\sigma(m|d, \gamma) = de^{\gamma(m - M_{\text{th}})}, \qquad (S5)$$

and the units of $d$ are degrees².

Model parameters $\theta$ and $u(x, y)$ are estimated simultaneously using an iterative approach under the stochastic declustering framework proposed in Zhuang et al. (2002). First, for an initial $u_0(x, y)$, $\theta$ is estimated by maximum likelihood method. Then the probability of being a background event for each event in the catalog is calculated by

$$\phi_j = \frac{\mu u(x_j, y_j)}{\lambda(t_j, x_j, y_j|H_{t_j})}. \qquad (S6)$$

And the variable bandwidth kernel estimator

$$\hat{u}(x, y) = \frac{1}{T}\sum_{j=1}^{N} \phi_j \varphi(x - x_j, y - y_j; h_j) \qquad (S7)$$

provides a better estimate for $u(x, y)$, where $T$ is an interval of length, $N$ is the number of events, and

$$\varphi(x, y; h) = \frac{1}{2\pi h^2}\exp\left(-\frac{x^2 + y^2}{2h^2}\right) \qquad (S8)$$

is the bivariate isotropic Gaussian kernel function with bandwidth $h$. The bandwidth

$$h_j = \max\{h_{\text{min}}, r(j, n_p)\}, \qquad (S9)$$

4

where $h_{min}$ is a minimum threshold bandwidth value and $r(j, n_p)$ denotes the distance between the location of event $j$ and its $n_p$-th nearest neighbor. We used default values for $h_{min} = 0.05°$ and $n_p = 5$. Using equation (S7), $u_0(x, y)$ is updated. These steps are repeated until the estimates converge.

**Text S2.2. Construction of an ETAS forecast model**

1.  The entire evaluation period and area are divided by 1 day bin and a segment with size $D°$ $\times D°$, respectively, where parameter $D$ is set to the same value in Table 2. In this way, the total number of space-time cells amounts is the total number of days multiplied by the number of grids shown in Table S3. In central Honshu, when $D = 0.2°$, the evaluation area cannot be filled without gaps, and a gap of 0.1° is created in either the north or south. At the southern end, no target earthquake (Figure 7a) had occurred and seismicity was relatively low (Figure S3). Therefore, the latitude range of area evaluated by the ETAS forecast model was set to 35.7–37.1°.
    By the way, Maeda's method is not based on cells fixed temporally, but on events that satisfy the conditions for precursory swarm activity. In addition, because the segments ($D°$ x $D°$) are arranged in the evaluation area at intervals of $D/2°$, the segments are allowed to overlap each other. Although the arrangement of space-time cells differs between Maeda's method and the ETAS forecast model, the forecast performance can be evaluated in the same way by equation (4) when we introduce the threshold probability $P_{th}$ described later in step 5.

2.  Based on the model parameters $\theta$ and $u(x, y)$ estimated in Section S2.1, the seismicity rate $\lambda(t, x, y)$ [/day/degree$^2$] at midnight is calculated daily at intervals of $\Delta g$ in latitude/longitude. In this study, $\Delta g = 0.01°$. Note that in the "ETAS" package (Jalilian, 2019), spherical coordinates (long, lat) in degrees on the earth's surface are transformed to flat map (Cartesian) coordinates $(x, y)$ in degrees in order to approximate the great-circle distance on the sphere by the corresponding Euclidean distance on the flat map by equirectangular projection as:

$$x = \cos\left(\frac{\text{cnt. lat}}{180}\pi\right)(\text{long} - \text{cnt. long}), \qquad \text{(S10)}$$

$$y = \text{lat} - \text{cnt. lat}, \qquad \text{(S11)}$$

    where cnt. lat and cnt. long are the latitude and longitude of the centroid of the geographical region.

3.  Daily seismicity rate $\lambda(t)$ [/day] in each segment ($D°$ x $D°$) on spherical coordinates is calculated following equation:

$$\lambda(t) = \sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} \lambda(t, x, y) dx dy, \qquad \text{(S12)}$$

    where, $(x_1, y_1)$ and $(x_2, y_2)$ indicate coordinates of southwestern and northeastern ends of each segment, respectively. $dx$ and $dy$ depend on the grid intervals to estimate $u(x, y)$; in this study, $dx = \cos\left(\frac{\text{cnt.lat}}{180}\pi\right) \times \Delta g$ and $dy = \Delta g$.

4.  The probability $P(N \geq 1)$ that one or more $M \geq M_{m0}$ occurs within $T_a$ days from a certain day $t$ is calculated by the following equation:

5

$$P(N \geq 1) = 1 - \exp\left(-T_a \lambda(t) e^{-\beta(M_{m0} - M_{\text{th}})}\right), \tag{S13}$$

where, $\beta = b \ln 10$, and $b$ represents slope of the Gutenberg-Richter (G-R) law (Gutenberg & Richter, 1944) estimated using earthquakes with $M \geq M_{\text{th}}$. Although $\lambda(t)$ decreases essentially with time, in this study we adopted the simple assumption that $\lambda(t)$ kept constant value for $T_a = 1$ day.

5. A daily alert is issued at space-time cells where $P(N \geq 1)$ is above a certain threshold $P_{\text{th}}$. By making it a binary problem in this way, it can be evaluated by the method described in Section 3.1.2 of the main text. $P_{\text{th}}$ was changed from 0.0001 to 1.0 in 0.0001 increments, and $P_{\text{th}}$ with the maximum ΔAIC was adopted (Table S3). A similar forecast method (Console et al., 2010) was applied for seismicity in and around Italy.

**Text S2.3. Input data and estimated ETAS parameters**

As mentioned in Section 3.1.1 in the main text, Maeda's method is unaffected by missing small aftershock data immediately after a large earthquake. In contrast, because the ETAS model depends on it strongly, data replenishment method in the time direction was suggested (Zhuang et al., 2017). In this study, not only time component but also space component is an important parameter of earthquake forecast, so only temporal replenishment method by Zhuang et al. (2017) is not enough. It is also suggested that ETASI model which extends the ETAS model to incorporate short-term aftershock incompleteness (Petrillo and Lippiello, 2021). In this study, we avoided the problem of short-term aftershock incompleteness using magnitude threshold $M_{\text{th}}$ (Table S1) equal to or larger than that used by Maeda's method (Table 1). Figure S6 shows a plot of normalized sequential magnitudes versus normalized sequential numbers of the earthquake events. If the distribution is biased (upper row in Figure S6), it suggests incomplete detection capability and short-term lack of aftershocks. Incidentally, not only the blank area in low empirical magnitude but also the dense parts in high empirical magnitude implies are caused by the existence of missing data [Zhuang et al., 2017]. The heterogeneity of the distribution reduces with $M_{\text{th}}$. In the middle row of Figure S6, there is a slight bias around 0.1 on the horizontal axis off Iwate and Miyagi prefectures, around 0.2 and 0.5 on the horizontal axis in central Honshu, and around < 0.05 on the horizontal axis in the Izu Islands. In the lower row of Figure S6 where $M_{\text{th}}$ is further increased, the distribution is no longer biased, and it can be considered that incomplete detection capability and short-term lack of aftershocks has been resolved.

Taking into consideration the uncertainty of ETAS parameters (e.g., Schoenberg et al., 2010) derived from the choice of $M_{\text{th}}$, we used four cases of $M_{\text{th}}$ to check variance of forecast performance for each region: $M_{\text{th}} = 4.8, 5.0, 5.2, 5.4$ in two regions along the Japan trench, $M_{\text{th}} = 3.3, 3.5, 3.7, 3.9$ in central Honshu and the Izu Islands. Table S1 shows conditions used to estimate parameters of the stationary space-time ETAS model. We set wide ranges for time, space, and magnitude taking into consideration the influence due to earthquakes occurring outside the specific areas (Table 1) (Ogata & Katsura, 2014). Table S4 lists an example of stationary space-time ETAS parameters estimated using data of minimum $M_{\text{th}}$ shown in Table S1. Parameters $M_{m0}, D,$ and $T_a$ used by the ETAS forecast model are the same as those used by Maeda's method (Table 2). The spatial smoothing function of the background seismicity rate $u(x, y)$ in each region is shown in Figure S7. In Figure S7a, b, epicenters in a space-time ETAS catalog produced in Section S3 are also plotted. As a result, there was almost no difference in

forecast performance depending on $M_{\text{th}}$ except for off Ibaraki area where the number of target earthquake is small and thus may leads to relatively large variation (Table S3).

**Text S3. Application of Maeda's method to ETAS catalogs**

**Text S3.1. ETAS catalogs**

In and around Japan, Ogata and Katsura (2014) investigated whether synthetic catalogs based on an ETAS model (ETAS catalogs, hereafter) could simulate swarm-like activity preceding a mainshock in the real data. They constructed ETAS catalogs based on a space-time ETAS model (Ogata, 1998) using all shallow (≤100 km depth) earthquakes of magnitude (M) ≥ 4.0 or 2.0 that occurred in or around Japan from 1 January 1994 to 10 March 2011. They constructed two types of ETAS catalogs; in the first, magnitudes randomly picked from real data were used, and in the second, the actual magnitude sequences were used. The former type did not successfully simulate the temporal acceleration characteristics of the cumulative number of foreshocks before mainshocks, whereas the latter did, although it accelerated more slowly than in the real data.

Because the stress field around the Japanese Islands is complicated by the motions of four plates (Figure 1), earthquakes with different hypocentral depths and focal mechanisms occur. ETAS catalogs obtained by applying a space-time ETAS model to nationwide seismic activity despite spatially differing characteristics are not accurate enough to be used for verifying the reproducibility of real data catalogs. In addition, the physical interpretation of the space-time ETAS model that only the magnitude sequence obeys real data, whereas the spatial and temporal characteristics of earthquakes are fluctuated by the model, is difficult. In particular, implications of the magnitude sequence are not clear when the target area is large.

Therefore, in this study we focused our investigation on two regions: off Iwate and Miyagi prefectures, and off Ibaraki prefecture. For each region, we constructed ETAS catalogs in which magnitude picked by random sampling from the magnitude frequency distribution of the real data. Table S1 shows the conditions used to construct ETAS catalogs. We produced 1,000 ETAS catalogs for $M_{\text{th}}$ = 4.8 by following step A–E in Section S4 of the supporting information because it was considered that earthquakes with $M$ ≥ 4.8 could be sufficiently detected (Figures S6 and S8). ETAS parameters estimated from real data are not always true parameters (Ross, 2021). To account for the uncertainty of ETAS parameters, in constructing the ETAS catalogs (see Section S4), we gave fluctuations in time, space, and magnitude although the same parameters as in Table S4 were given. Figures S9 and S10 show the parameter distributions of each of the 1,000 ETAS catalogs, indicating that catalogs with some variations in ETAS parameters could be produced. Then, under the same conditions shown in Table 1, we investigated forecast efficiency (see Section 3.1.2) by applying Maeda's method to both real data and the ETAS catalogs. Because it is not fair to apply the optimum parameter values used for the real data (shown in Table 2) to the ETAS catalogs, for each ETAS catalog, we searched for optimum values, i.e., those resulting in a high forecast performance, by the grid search method. For the optimum values of $M_{f0}$, $M_{m0}$, and $D$, we used the values determined by Maeda and Hirose (2016). We estimated other parameters by the grid search method from 900 combinations: $T_f$ = 1, 2, …, 10 days; $N_f$ = 2, 3, …, 10 events; and $T_a$ = 1, 2, …, 10 days. See Sections S3.2 and S3.3 in the supporting information for detail results.

Figure S11 shows the cumulative number per a target earthquake with log-scale on the horizontal axis. We also superimpose 1,000 ETAS catalogs for two regions (off Iwate and Miyagi prefectures and off Ibaraki prefecture) along the Japan Trench. There are a few cases

7

that temporal acceleration of the cumulative number of ETAS catalogs is larger than that of the real data. Seismicity in the Izu Islands is characterized as swarm-type seismicity because the ETAS parameter α is small (Table S4). Accordingly, the number of foreshock candidates (see Section 3.1.1) is higher there than in the other regions (see vertical axis in Figure S11d).

Note that for central Honshu and the Izu Islands, $M_{\text{th}}$ is about 3.3 due to the aftershock incompleteness immediately after great earthquakes, which cannot be evaluated in the same evaluation conditions with $M_{\text{th}} = 2.0$ and 3.0 shown in Table 1, so we did not deal with them here.

**Text S3.2. Off Iwate and Miyagi prefectures**

Figure S12 shows the characteristics of the real data (open arrows) and those of ETAS catalogs (histograms) off Iwate and Miyagi prefectures. Solid arrows indicate the median value of the ETAS catalogs. With respect to the total number of earthquakes of $M \geq 5.0$, the number of target earthquakes of $M \geq 6.0$, and the $b$ value of the Gutenberg-Richter law, the ETAS catalog data overlap real data (Figure S12a–c). Note that the histograms in (a–c) are derived from 1,000 catalogs, but those in (d–k) are derived from 753 ETAS catalogs. In the other 247 ETAS catalogs within the range of the grid search of foreshock parameters, there was no true alarm earthquake and ΔAIC was not calculated. In other words, because there was no foreshock-mainshock sequence in 247 of the 1,000 ETAS catalogs, it was hard for these 247 ETAS catalogs to reproduce foreshock characteristics. The most frequent values for the optimum foreshock parameters of the ETAS catalogs were $T_f = 1$ day, $N_f = 2$ events, and $T_a = 1$ day (median values: $T_f = 1$ day, $N_f = 2$ events, and $T_a = 2$ days) (Figure S12d–f, Table S5). In contrast, the optimum foreshock parameters for real data were $T_f = 9$ days, $N_f = 3$ events, and $T_a = 4$ days. Thus, the real data had more than 1.6 times the number of alarm earthquakes than the median number in the ETAS catalogs (Figure S12g). In most cases, AR, TR, and the F-measure of the real data were larger than those of the ETAS catalogs (Figure S12h–j). However, PG in 329 of 753 ETAS catalogs (~44%), and in ~33% of 1,000 ETAS catalogs, was larger than that in the real data (Figure S12k). This difference is due to the shorter alarm periods $T_a$ (Figure S12f) and fewer alarm earthquakes (Figure S12g) in the ETAS catalogs. To summarize, a forecast model based on Maeda's method is more efficient than an ETAS model according to AR, TR, and the F-measure for earthquakes off Iwate and Miyagi prefectures, whereas both models are equivalent in terms of PG. Note that Maeda's method could forecast the $M$ 9 earthquake off Miyagi prefecture retrospectively, although that event occurred after the end of the study period.

**Text S3.3. Off Ibaraki prefecture**

Figure S13 shows the characteristics of the real data and the ETAS catalogs off Ibaraki prefecture. With respect to the total number of earthquakes of $M \geq 5.0$, the number of target earthquakes of $M \geq 6.0$, and the $b$ value of the Gutenberg-Richter law, the ETAS catalog data overlap the real data (Figure S13a–c). The $b$ value distribution in this area (Figure S13c) is broader than that in the other areas because the estimation errors are larger owing to the relatively small total number of earthquakes (Figure S13a). The histograms in (a–c) are derived from 1,000 ETAS catalogs, whereas those in (d–k) are derived from 387 ETAS catalogs. Because there were no true alarm earthquakes in the other 613 ETAS catalogs within the range of the grid search of the foreshock parameters, ΔAIC was not calculated. In other words, because there was no foreshock-mainshock sequence in 613 of the 1,000 ETAS catalogs, it was hard for

these 613 ETAS catalogs to reproduce foreshock characteristics. The most frequent values of the optimum foreshock parameters were $T_f$ = 1 day, $N_f$ = 2 events, and $T_a$ = 1 day in the ETAS catalogs (median values: $T_f$ = 1 day, $N_f$ = 2 events, and $T_a$ = 1 day) (Figure S13d–f, Table S5). In contrast, the optimum foreshock parameters in the real data were $T_f$ = 3 days, $N_f$ = 2 events, and $T_a$ = 1 day. The number of alarm earthquakes was about the same in both the real data and ETAS catalogs (Figure S13g). In most cases, AR, TR, the F-measure, and PG were larger in the real data than in the ETAS catalogs (Figure S13h–k). PG was larger in 50 of 387 ETAS catalogs (~13%) than in the real data, and was larger in only ~5% of the 1,000 ETAS catalogs (Figure S13k). To summarize, a forecast model based on Maeda's method was more efficient than an ETAS model off Ibaraki prefecture.


**Text S4. Synthetic ETAS catalogs**

**Text S4.1. Construction of an ETAS catalog**

In this section, we describe how to construct an ETAS catalog. See Section S2 in the supporting Information for the definition of ETAS parameters.

Step A. Origin time of background earthquakes

Parameter $\mu$ is assumed to be independent of time. Therefore, we can treat background earthquakes as a homogeneous Poisson process.

A-1. According to the background seismicity rate $\mu \sum_{i=1}^{N_{\mathrm{g}}} u(x_i, y_i) dx dy$, earthquakes, so-called first-generation earthquakes, are generated randomly in time. Here, $N_{\mathrm{g}}$ is the total number of grids.

A-2. According to equations (S2)–(S5), aftershocks of first-generation earthquakes, so-called second-generation earthquakes, are generated. See Steps B and D below for the setting of the origin time and spatial distribution of an aftershock, respectively.

A-3. Aftershocks of second-generation earthquakes, so-called third-generation earthquakes, are generated in the same way as in Step A-2.

A-4. Similarly, fourth-, fifth-, …, and $X$th-generation earthquakes are generated in the same way until no aftershocks are occurring within the target period.


Step B. Origin time of aftershocks

Equation (S3) is dependent on time. Therefore, the simulation must be treated as a non-homogeneous Poisson process. For this, we used a thinning algorithm (Ogata, 1981) as follows.

B-1. We set $v(t, m) = \kappa(m|K, \alpha)h(t|c, p)$ and $v^* = \max_{0 \leq t \leq T} v(t, m)$, based on equations (S2) and (S3), where $t = 0$ represents the origin time of a parent earthquake and $T$ indicates the end time of the target period.

B-2. According to a homogeneous Poisson process, with seismicity rate $v^*$, a sequence of time points $\{t_1, t_2, \dots\}$ is generated until $t_i$ exceeds $T$.

B-3. We generate a uniform random number $a_i$ at each $t_i$.

B-4. $t_i$ is retained if $v(t_i, m)/v^* > a_i$; otherwise, $t_i$ is excluded. A new sequence of the retained time points obeys a non-homogeneous Poisson process with intensity function $v(t, m)$.


Step C. Spatial distribution of background seismicity

The spatial distribution of background seismicity, which consists of the earthquakes generated in Step A-1, is determined by $u(x,y)$. Because $u(x,y)$ cannot be solved analytically, we solve it numerically by the following procedure.

C-1. When the lower left grid is fixed as the origin $(x,y) = (0,0)$, the integrated value up to certain coordinates $(x,y)$ is obtained as:

$$U(x,y) = \sum_{i=0}^{x} \sum_{j=0}^{y} u(i,j). \qquad (S14)$$

C-2. $U(x,y)/U_{\max}$ ranges between 0 and 1, where $U_{\max}$ is the total integrated value up to the upper right grid.

C-3. When $U(x,y)/U_{\max}$ exceeds a uniform random number, coordinates $(x,y)$ become the location of a background earthquake. In this regard, the location is shifted randomly in a range of each grid ($\Delta g = 0.01°$).

Step D. Spatial distribution of aftershocks

The spatial distribution of aftershocks, which are earthquakes generated in Steps A-2 to A-4, is determined by a probability density function $f(x,y,m|D,q,\gamma)$ (equation (8)). We solve it analytically as follows.

D-1. The coordinates are transformed from Cartesian to polar coordinates. Then, the cumulative distribution function of $f(r,m|d,q,\gamma)$, where $r$ indicates the distance from a previous earthquake, is:

$$F(r,m|d,q,\gamma) = 1 - \left(1 + \frac{r^2}{\sigma(m|d,\gamma)}\right)^{-q+1}. \qquad (S15)$$

Accordingly, the inverse function becomes:

$$r = \left[\sigma(m|d,\gamma)\left\{(1 - F(r,m|d,q,\gamma))^{\frac{1}{-q+1}} - 1\right\}\right]^{\frac{1}{2}}. \qquad (S16)$$

D-2. Because a uniform random number $a = F(X)$ is equivalent to $a = 1 - F(X)$, equation (S5) is rewritten to obtain the distance $r$ from a previous earthquake as:

$$r = \left[\sigma(m|d,\gamma)\left\{a^{\frac{1}{-q+1}} - 1\right\}\right]^{\frac{1}{2}}. \qquad (S17)$$

D-3. Simultaneously, the azimuth $\phi$ from the previous earthquake is set randomly between 0 and $2\pi$ radians.

D-4. The location $(x,y)$ of an aftershock on the equirectangular projection map is determined from the coordinates $(r,\phi)$ obtained in Steps D-2 and D-3.

Step E. Magnitude-frequency distribution

The magnitude-frequency distribution does not always obey a specific law such as the G-R law even if earthquake detectability is enough (Utsu, 1999). Therefore, in this study, we treated ETAS catalogs with magnitudes picked by random sampling from the magnitude frequency distribution of real data without obeying the G-R law.

**Text S4.2. Examples of ETAS catalogs**

Figure S7 shows spatial distribution of $u(x,y)$ with examples of ETAS catalogs for two regions along the Japan trench. Synthetic earthquakes tend to be distributed in areas with a high background seismicity rate.

**Text S5. Molchan's error diagram**

Molchan's error diagram (Molchan 1997) consists of two parameters: alarm fraction $\tau$ (= total alarmed periods / whole periods) and the rate of failures-to-predict $\nu$ (= number of missed targets / number of targets in whole periods), which generally allows to show visually how far is the forecast from a given reference model (random guess). The ideal forecast is the case where a coordinate locates near the origin, the smaller the forecast error $e(\tau, \nu) = \tau + \nu$, the better the model. The diagram can also show the relation between two phenomena, and the smaller the error function $e(\tau, \nu) = \tau + \nu$, the higher the relationship.

In the case where phenomena occur $n$ times during the whole period $T$, i.e., average occurrence rate $\lambda_0 = \frac{n}{T}$, the probability of obtaining $k$ or more hits by chance during total alarm periods ($\tau T$) is described by the binomial distribution as follows:

$$\alpha_0 = \sum_{i=k}^{n} \left[ {}_nC_i \tau^i (1-\tau)^{n-i} \right] \qquad (S18)$$

where ${}_nC_i$ indicates the combination that is the number of possible ways to select $i$ objects among $n$ things without repetition and order. By searching for $\tau$ such that $\alpha_0$ becomes 0.01 and 0.05 for a given $\nu = 1 - \frac{k}{n}$, a confidence interval of $\tau$ with a confidence coefficient of 99% and 95%, respectively, can be determined by random guess.

Figure S14 shows Molchan's error diagrams of the relationship between the alarm fraction $\tau$ based on amplitude of the sinusoidal curve with 3.09-year period estimated by Uchida et al. (2016) and the rate of failures-to-predict $\nu$ for target earthquakes and alarm earthquakes in region N off Iwate prefecture. We can confirm that Molchan's error curves are generally out of the 95% confidence interval of random guess, indicating that target and alarm earthquakes synchronized with periodic SSEs.
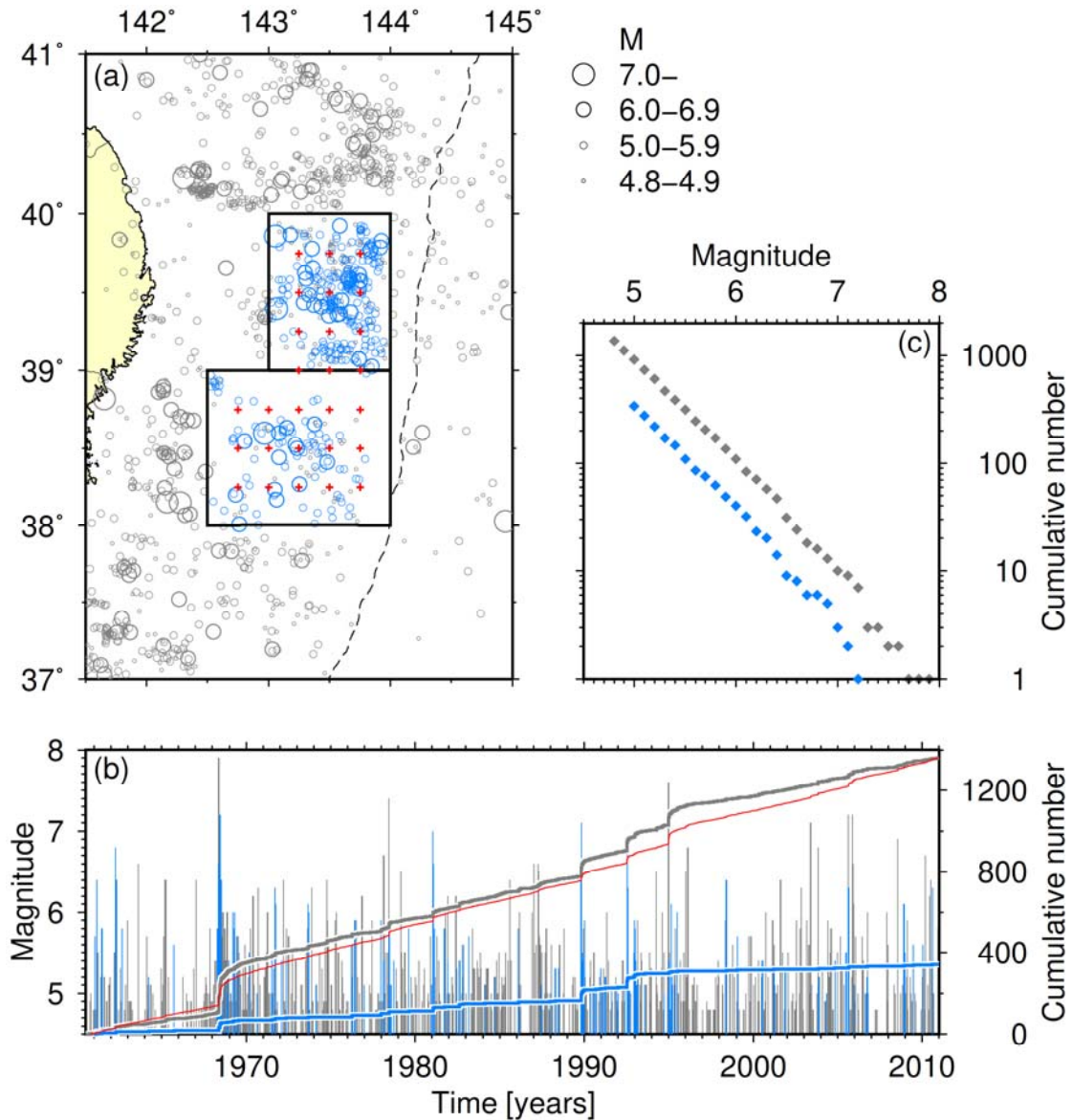
**Figure S1.** Seismicity off Iwate and Miyagi prefectures. (a) Epicenter distribution. Blue circles indicate epicenters of $M \geq 5.0$ from 1 January 1961 to 31 December 2010 within the evaluation regions (rectangles) indicated in Table 1. Gray circles indicate epicenters of $M \geq 4.8$ from 1 October 1960 to 31 December 2010 within the expanded region indicated in Table S1. Most of the gray circles within the evaluation regions are superimposed by blue circles. Red crosses indicate the calculation grid with a 0.25° spacing. (b) Plot of earthquake magnitudes (vertical bars) and the cumulative number of earthquakes (curves) versus time. The blue and gray colors correspond to those in (a). The red line indicates the space-time ETAS model fitted to the gray events. (c) Cumulative frequency-magnitude distribution. The colors correspond to those in (a).
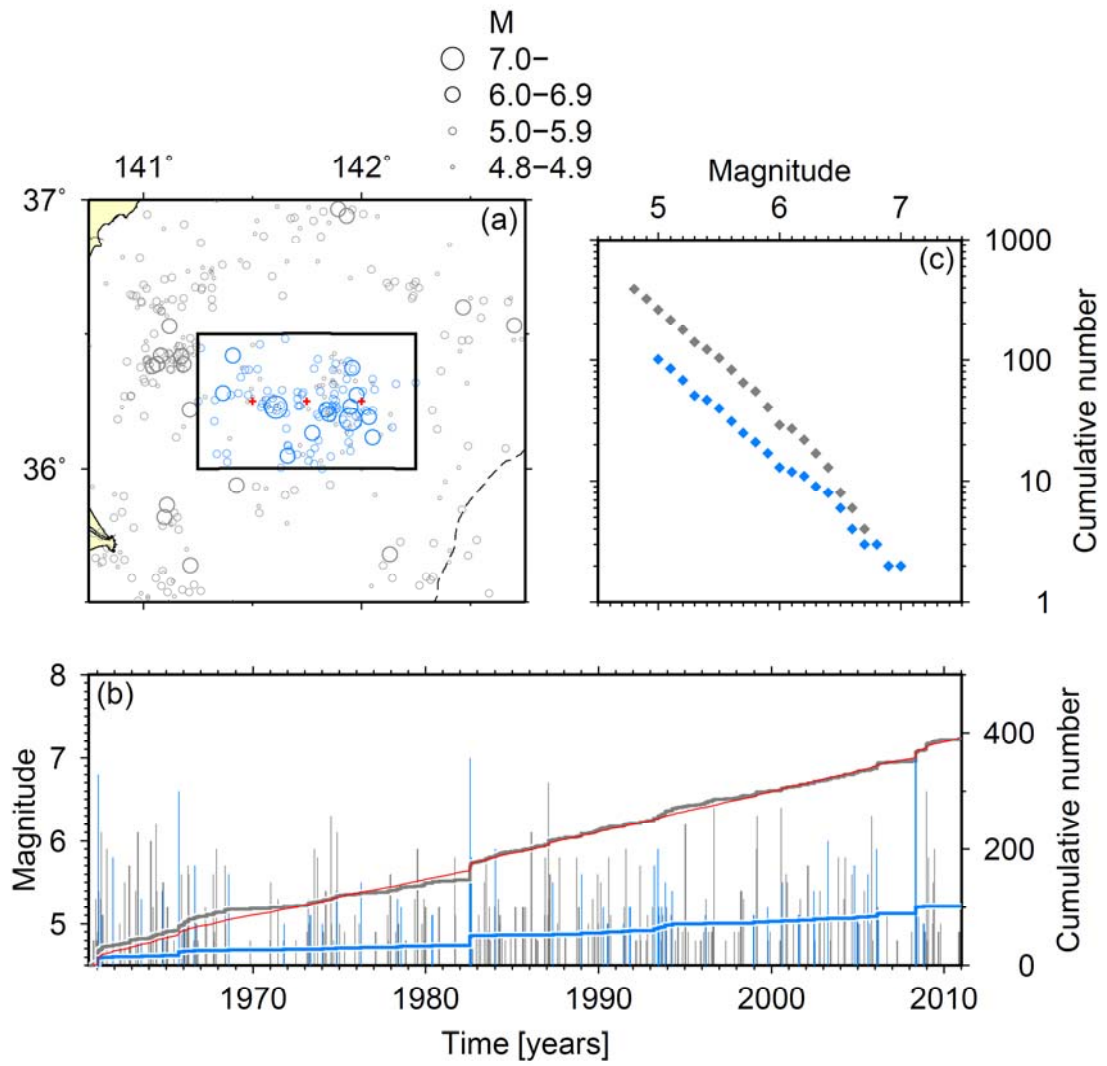
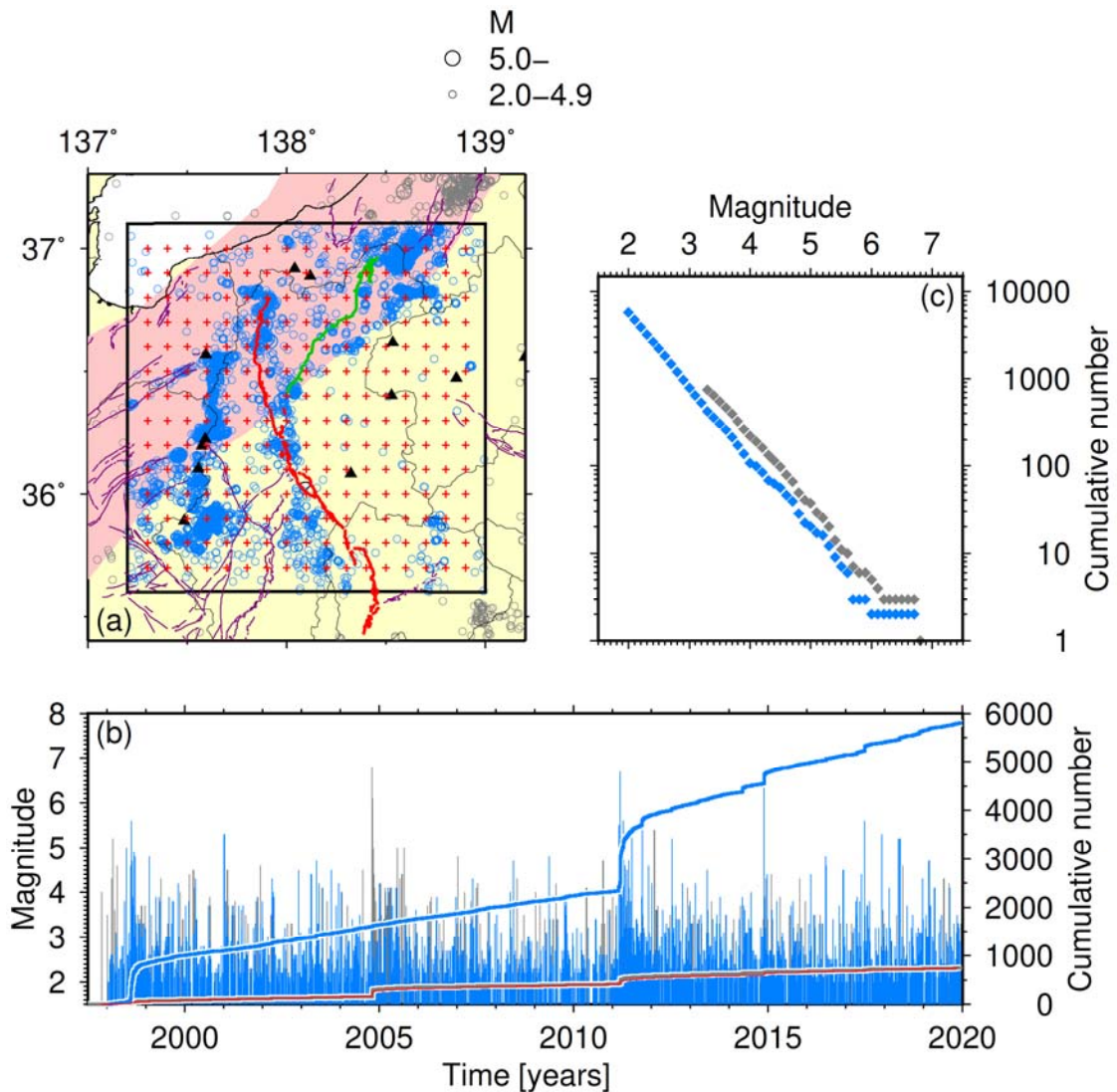**Figure S2.** Same as Figure S1, but for seismicity off Ibaraki prefecture.

**Figure S3.** Same as Figure S1 (except as noted), but for seismicity in central Honshu. Blue circles indicate epicenters of *M* ≥ 2.0 from 1 January 1998 to 31 December 2019 within the evaluation region (rectangle) indicated in Table 1. Gray circles indicate epicenters of *M* ≥ 3.3 from 1 October 1997 to 31 December 2019 within the region indicated in Table S1. Most of the gray circles within the evaluation region are superimposed by blue circles. Red crosses indicate the calculation grid with a spacing of 0.1°. Red lines indicate the Itoigawa-Shizuoka Tectonic Line (ISTL) (Headquarters for Earthquake Research Promotion, 2015a), and green lines indicate the "western marginal fault zone of the Nagano Basin" (Headquarters for Earthquake Research Promotion, 2015b). Purple lines indicate active faults. Triangles indicate active volcanoes. The pink shading indicates the Niigata-Kobe Tectonic Zone (Sagiya et al., 2000).
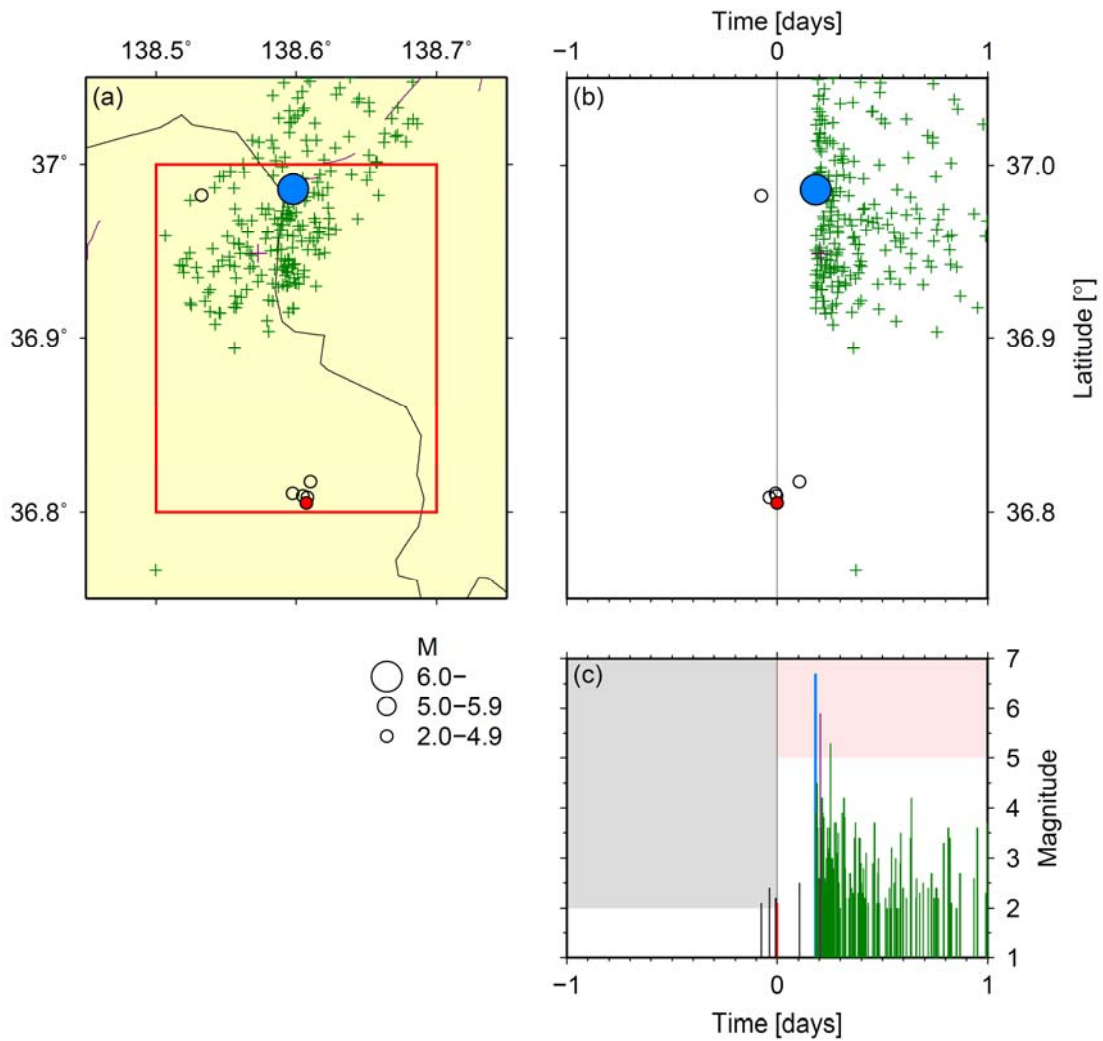
**Figure S4.** Seismicity from 1 day before to 1 day after event C (see Figure 7). (a) Epicenter distribution. Crosses indicate aftershock epicenters, and the purple cross indicates an aftershock having a magnitude difference of less than 1.0 from the previous largest earthquake. Circles indicate other epicenters: the solid blue circle indicates an alarmed target earthquake preceded by foreshocks, and the solid red circle indicates an alarm earthquake followed by a target earthquake. The red rectangle indicates the alarm area. (b) Time-space plot of the earthquakes shown in (a). (c) Plot of earthquake magnitudes versus time of events in the red rectangle in (a). The colors correspond to the symbol colors in (a).
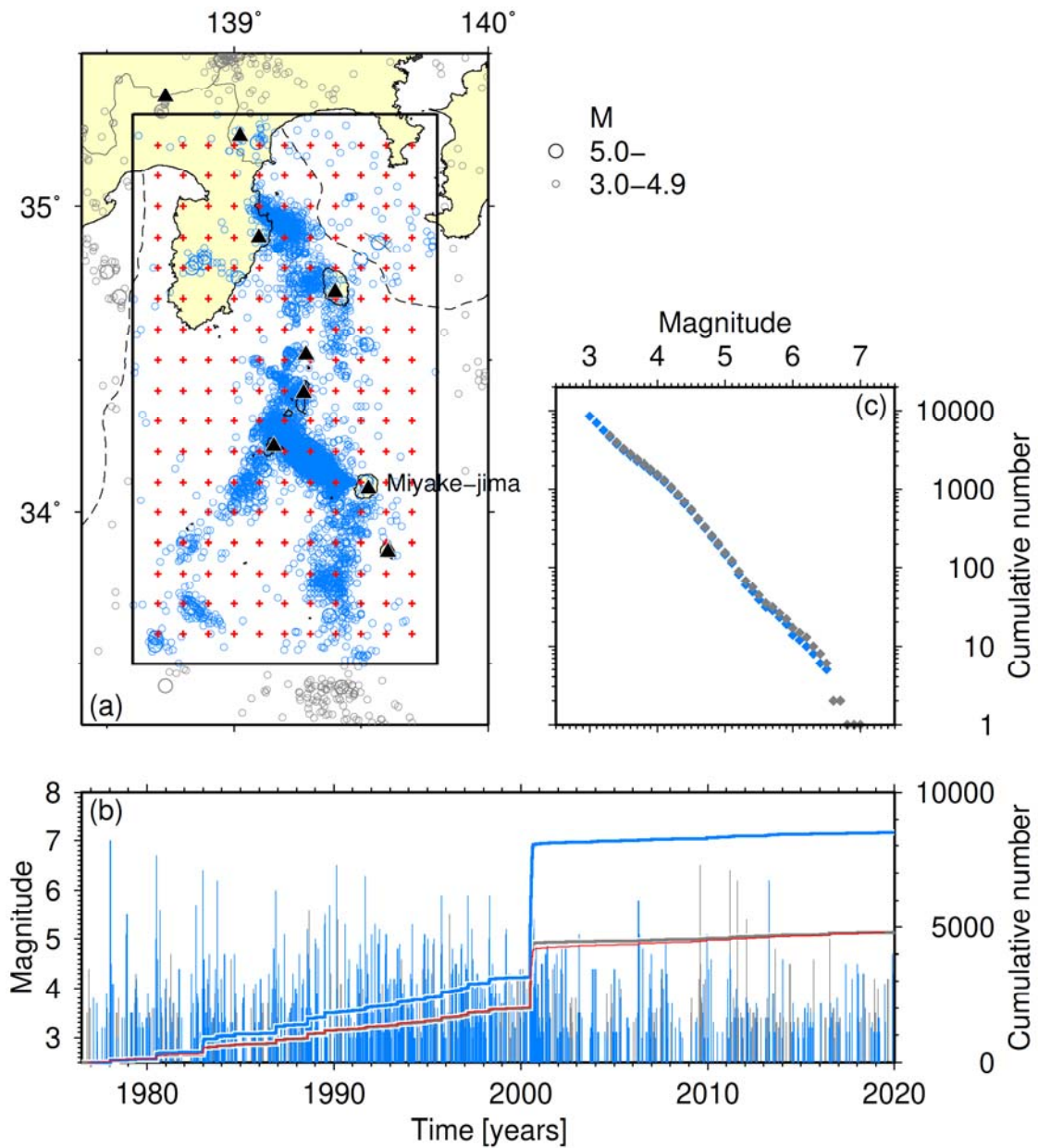
**Figure S5.** Same as Figure S1 (except as noted), but for seismicity in and around the Izu Islands. Blue circles indicate epicenters of *M* ≥ 3.0 from 1 January 1977 to 31 December 2019 within the evaluation region (rectangle) indicated in Table 1. Gray circles indicate epicenters of *M* ≥ 3.3 from 1 October 1976 to 31 December 2019 within the region indicated in Table S1. Most of the gray circles within the evaluation region are superimposed by blue circles. Red crosses indicate the calculation grid with a spacing of 0.1°. Triangles indicate active volcanoes.
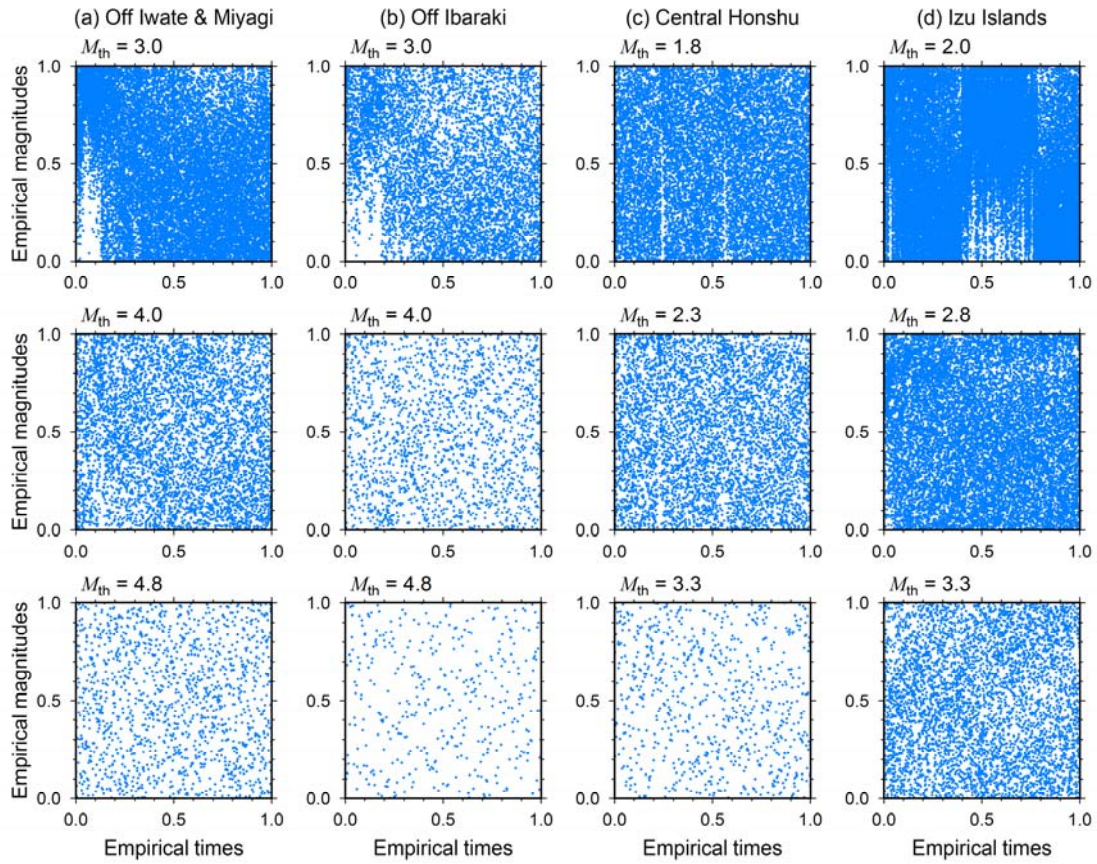
16

**Figure S6.** A plot of normalized sequential magnitudes versus normalized sequential numbers of the earthquake events. After the magnitudes were dithered with random errors uniformly in [−0.05, 0.05], the recorded events were transformed by using the bi-scale empirical transformation (Zhuang et al., 2017). (a) Off Iwate and Miyagi prefectures, (b) off Ibaraki prefecture, (c) central Honshu, and (d) Izu islands.
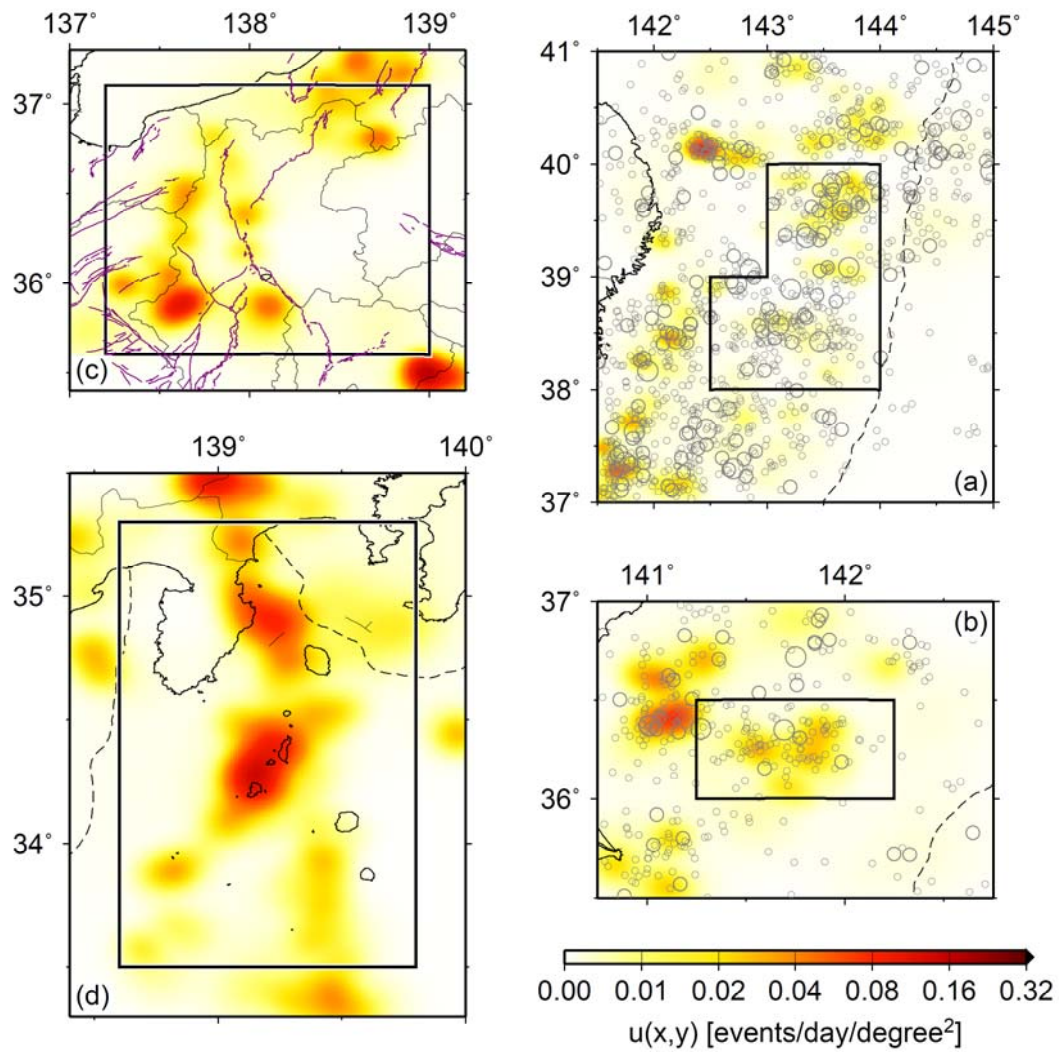
**Figure S7.** Examples of the spatial smoothing function of the background seismicity rate $u(x,y)$ on an equirectangular projection map. The background color shows $u(x,y)$ with a spacing of 0.01° on the geographic coordinate system. (a) Off Iwate and Miyagi prefectures, (b) off Ibaraki prefecture, (c) central Honshu, and (d) the Izu Islands. In (a, b), epicenters in a space-time ETAS catalog are also plotted.
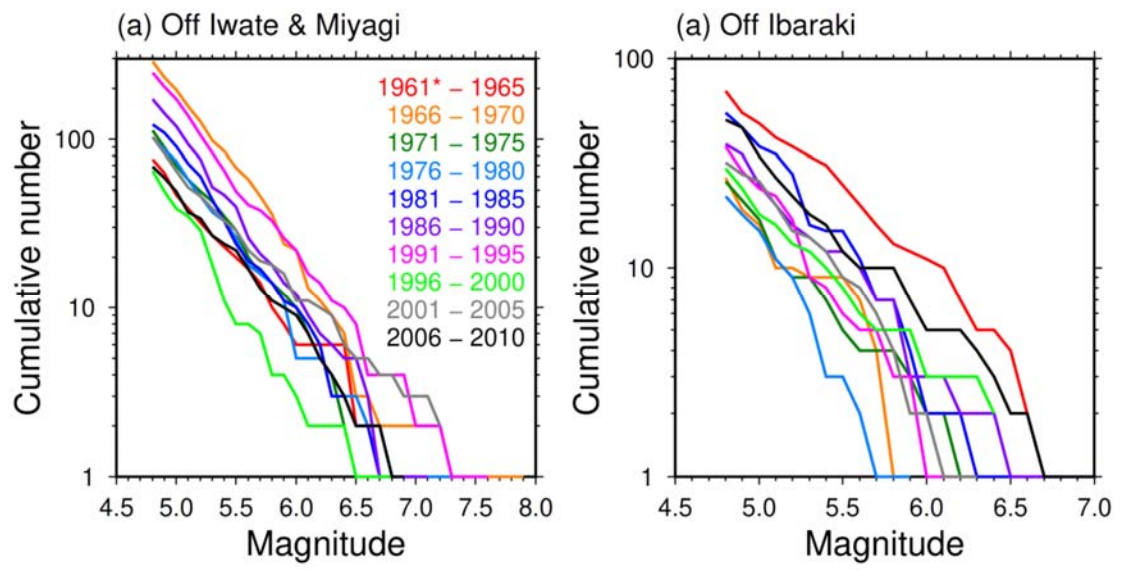
**Figure S8.** Frequency-magnitude distribution for (a) off Iwate and Miyagi prefectures, and (b) off Ibaraki prefecture. Colored lines correspond to each 5-year period. Note that 1961* indicates from 1 October 1960.
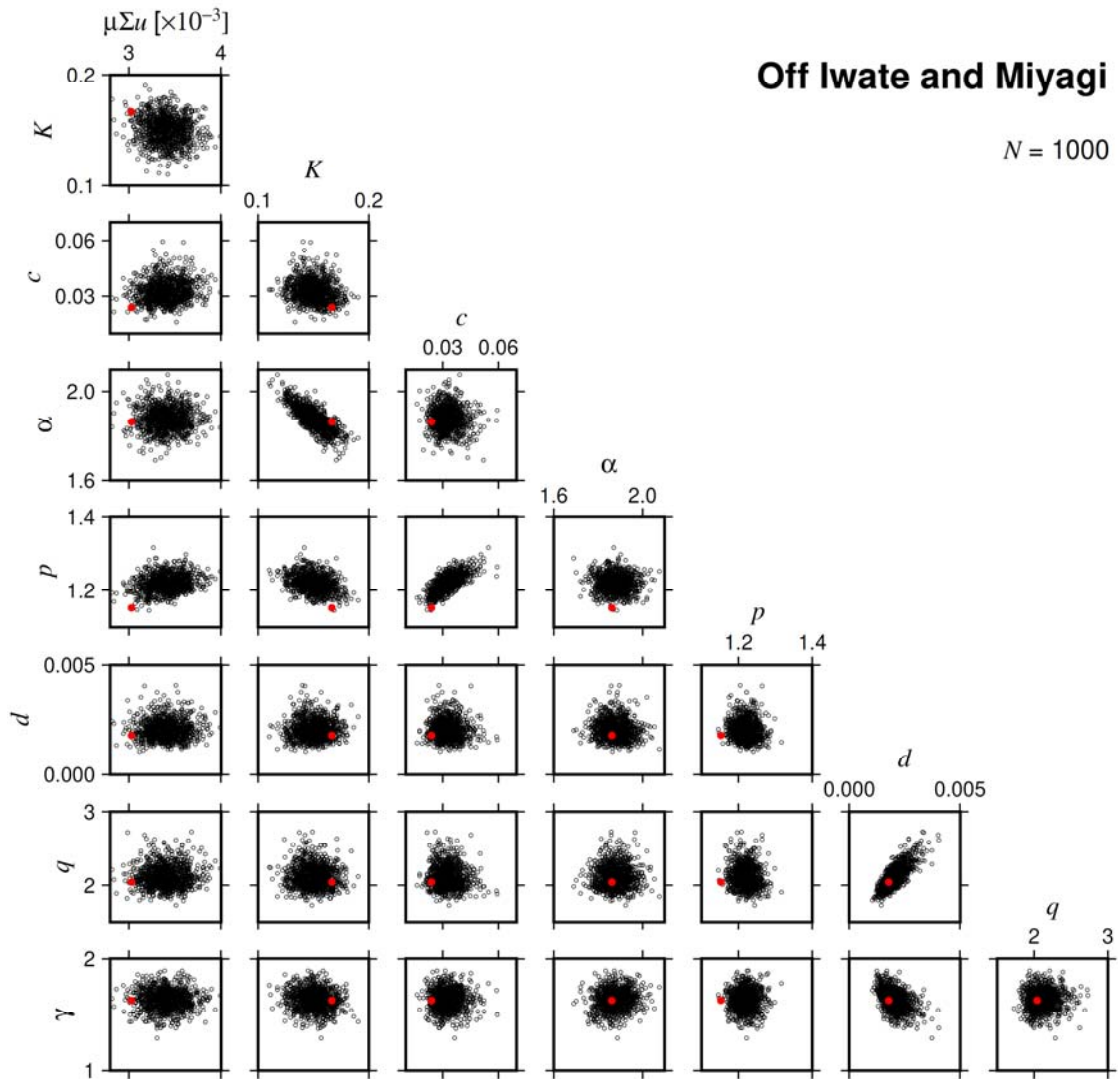
**Figure S9.** Correlations of ETAS parameters estimated for 1,000 ETAS catalogs off Iwate and Miyagi prefectures. Red dot indicates ETAS parameters estimated for the real data (Table S4). $\mu\Sigma u$ represents $\mu \times \Sigma u(x, y) \times \Delta g^2 / (\Delta \text{lon} \times \Delta \text{lat})$, where $\Delta g = 0.01°$, and $\Delta$lon and $\Delta$lat are the longitudinal and latitudinal distances shown in Table S1, respectively.

**Figure S10.** Same as Figure S9, but for off Ibaraki prefecture. The estimates converged in 941 out of 1,000 ETAS catalogs.

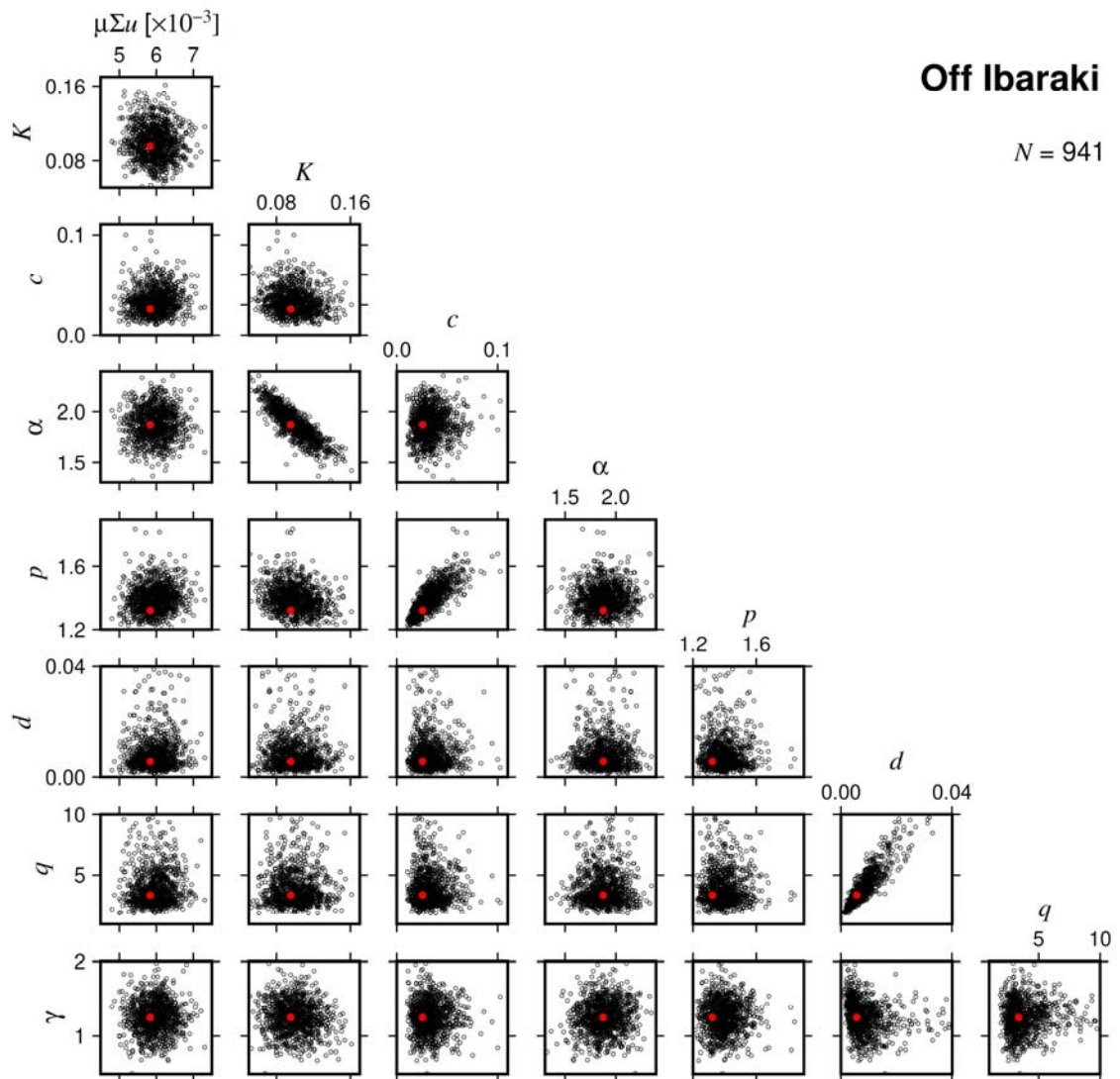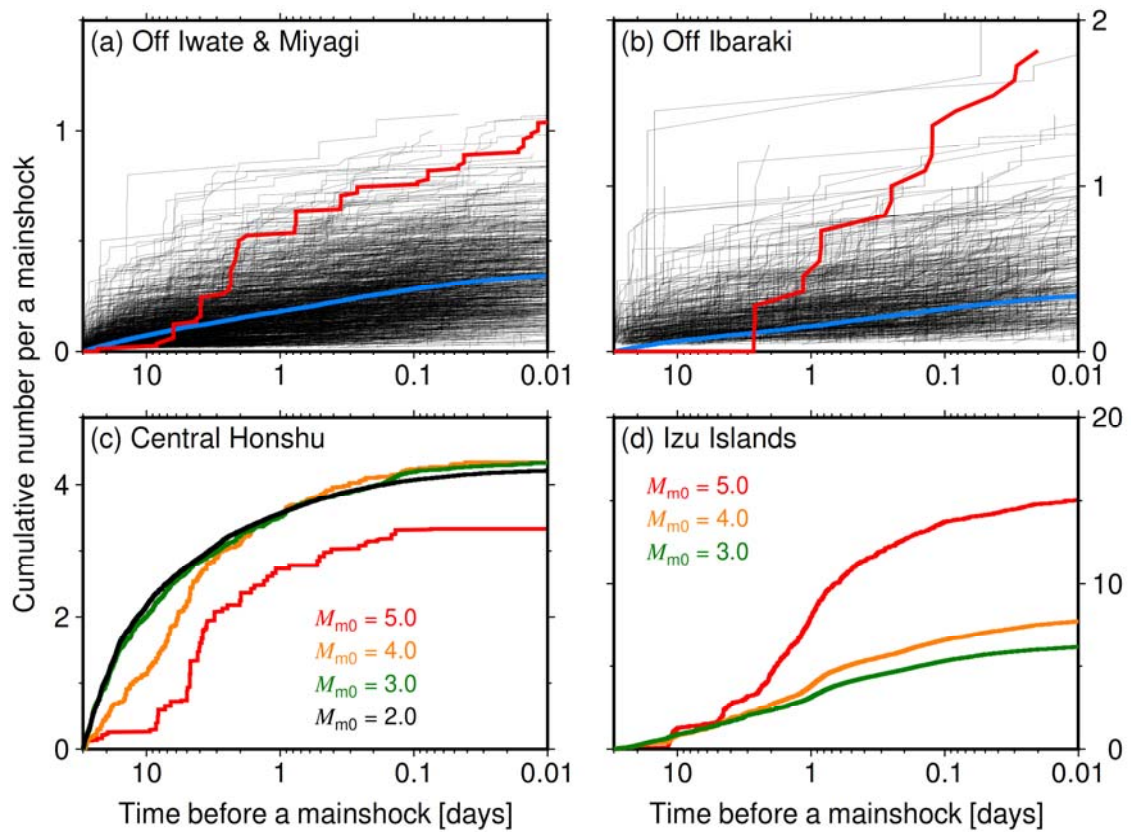**Figure S11.** The cumulative number of stacked earthquakes, normalized by the number of mainshocks, in the 30 days before a mainshock (target earthquake). Earthquakes satisfying the aftershock conditions (equations (1)–(3)) have been excluded. (a) Off Iwate and Miyagi prefectures, (b) off Ibaraki prefecture, (c) central Honshu, and (d) the Izu Islands. In (a, b), black curves are those of the 1,000 synthetic catalogs, and the blue curve represents the mean of the black curves. Other curves indicate a track depending on a magnitude threshold of target earthquakes of the JMA catalog. See Table 2 for the values of parameter $D$ in each region.
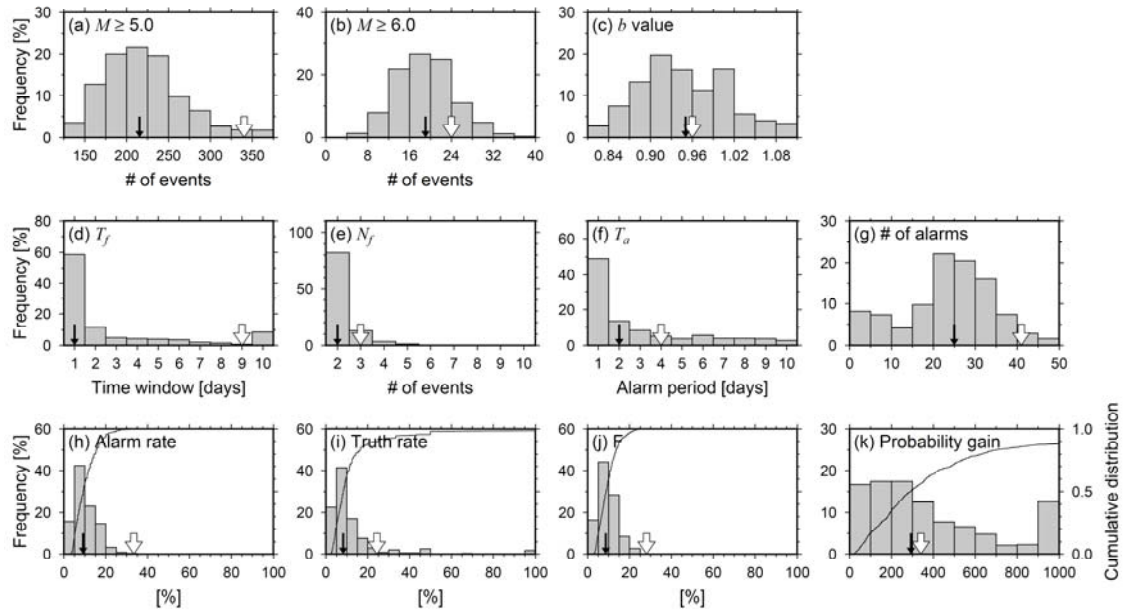
**Figure S12.** The distribution of ETAS catalog properties for 1961–2010 in the evaluation region off Iwate and Miyagi prefectures. White and black arrows indicate the JMA catalog properties and the median of ETAS catalog properties, respectively. The bins at the left and right ends include values under and over the range shown in each panel, respectively. (a) Total number of earthquakes of $M \geq 5.0$. (b) Number of target earthquakes of $M \geq 6.0$. (c) $b$ value of the Gutenberg-Richter law (Gutenberg & Richter, 1944). (d) $T_f$, the time window in which the number of earthquakes should be counted. (e) $N_f$, the cumulative number of earthquakes during the past $T_f$ days. (f) $T_a$, the alarm period in which a mainshock is expected to occur. (g) The number of alarm earthquakes. (h) Alarm rate. (i) Truth rate. (j) F-measure. (k) Probability gain. Curves in (h–k) indicate cumulative distributions. Note that the histograms in (a–c) and (d–k) are derived from 1,000 and 753 ETAS catalogs, respectively.
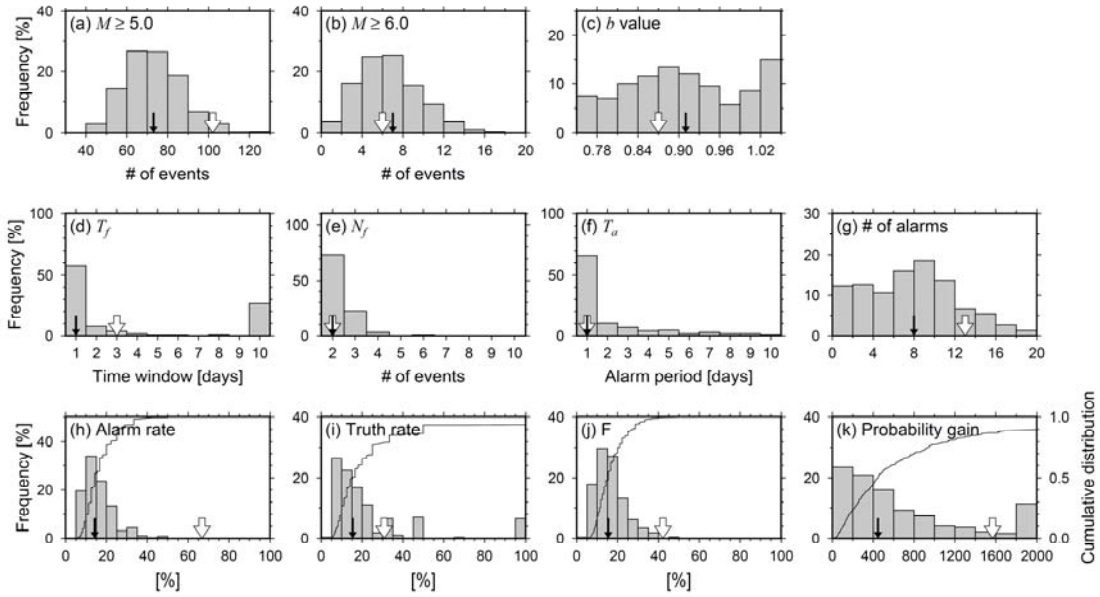
**Figure S13.** Same as Figure S12, but for off Ibaraki prefecture. Note that histograms in (a–c) and (d–k) are derived from 1,000 and 387 ETAS catalogs, respectively.
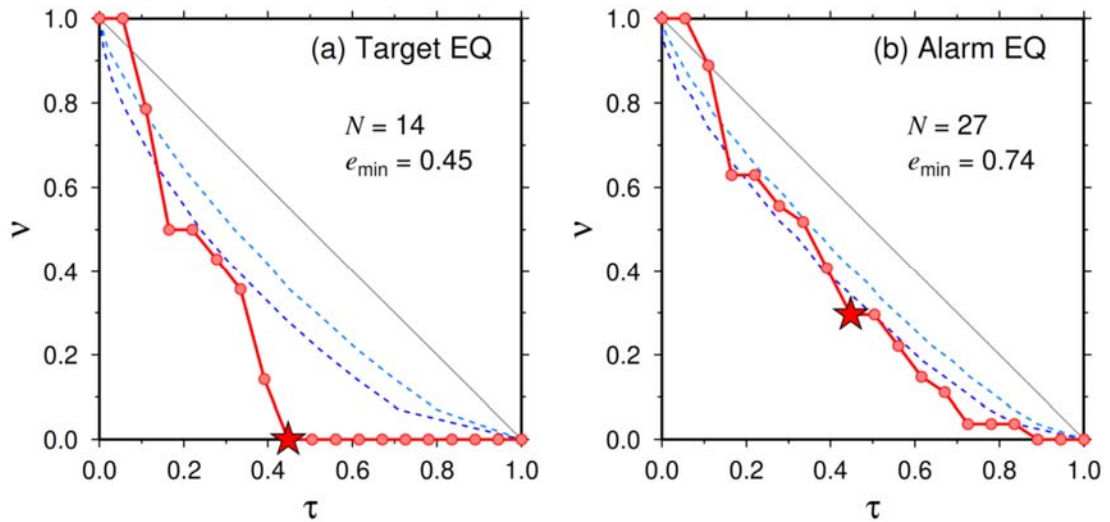


**Figure S14.** Molchan's error diagrams. The relationship between the alarm fraction $\tau$ based on the best-fitted sinusoidal function with periods of 3.09-years estimated by Uchida et al. (2016) and the rate of failures-to-predict $\nu$ for (a) target earthquakes with $M_{m0} = 6$ and (b) alarm earthquakes with $M_{f0} = 5$, $T_f = 9$ days, and $N_f = 3$ events in region N off Iwate prefecture is shown. Broken lines indicate 99% and 95% confidence interval of random guess (solid line). A red star indicates the minimum $e(\tau, \nu)$.

**Table S1.** Conditions used to estimate parameters of the stationary space-time ETAS model and to construct space-time ETAS catalogs.

| Area | Period | | Region | | | $M_{th}$ | # of earthquakes[+] | | |
|---|---|---|---|---|---|---|---|---|---|
| | Years | Days | Latitude (°N) | Longitude (°E) | Depth (km) | | Total | K, k, A | S, s, a |
| Off Iwate and Miyagi | 1960/10/1 – 2010 | 18354 | 37.0–41.0 | 141.5–145.0 | 100.0 | 4.8, 5.0, 5.2, 5.4 | 1360 | 1347 | 13 |
| Off Ibaraki | " | " | 35.5–37.0 | 140.75–142.75 | " | " | 390 | 390 | 0 |
| Central Honshu | 1997/10/1 – 2019 | 8127 | 35.4–37.3 | 137.0–139.2 | 30.0 | 3.3, 3.5, 3.7, 3.9 | 746 | 746 | 0 |
| Izu Islands | 1976/10/1 – 2019 | 15797 | 33.3–35.5 | 138.4–140.0 | 50.0 | " | 4634 | 4582 | 52 |

*Note:* [+]The number of earthquakes of minimum $M_{th}$.

**Table S2.** Optimum parameters and results for different $D$ values obtained by applying Maeda's method to the real catalog.

| Area | Period | Parameter | | | | | | Alarm rate (%) | | Truth rate (%) | | $F$ (%) | PG | ΔAIC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $D$ (°) | $M_{f0}$ | $T_f$ (days) | $N_f$ | $T_a$ (days) | $M_{m0}$ | (%) | | (%) | | | | |
| Central Honshu | 1998–2019 | 0.1 | 2.0 | 1 | 5 | 5 | 5.0 | 33.3 | (= 5/15) | 9.2 | (= 7/76) | 14.4 | 1293.5 | 61.5 |
| | | 0.2 | " | 1 | 5 | 5 | " | 40.0 | (= 6/15) | 9.6 | (= 7/73) | 15.5 | 439.9 | 61.9 |
| | | 0.3 | " | 1 | 5 | 5 | " | 40.0 | (= 6/15) | 9.5 | (= 8/84) | 15.4 | 199.7 | 52.4 |
| Izu Islands | 1977–2019 | 0.2 | 3.0 | 1 | 2 | 4 | 5.0 | 72.3 | (= 47/65) | 20.1 | (= 63/314) | 31.4 | 338.0 | 499.2 |
| | | 0.3 | " | 1 | 3 | 3 | " | 66.2 | (= 43/65) | 26.2 | (= 49/187) | 37.5 | 305.1 | 442.4 |
| | | 0.4 | " | 3 | 3 | 5 | " | 72.3 | (= 47/65) | 23.0 | (= 44/191) | 34.9 | 110.2 | 394.0 |

**Table S3.** Results of an earthquake forecast model based on the stationary space-time ETAS model.

| Area | Period | # of grids | Parameter | | | | $P_{th}$ | Alarm rate (%) | | Truth rate (%) | | $F$ (%) | PG | ΔAIC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $D$ (°) | $T_a$ (days) | $M_{m0}$ | $M_{th}$ | | (%) | | (%) | | | | |
| Off Iwate and Miyagi | 1961–2010 | 10 | 0.5 | 1 | 6.0 | 4.8 | 0.0022 | 37.5 | (= 9/24) | 0.9 | (= 9/1008) | 1.7 | 67.9 | 60.0 |
| | (18262 d) | | | | | 5.0 | 0.0020 | 37.5 | (= 9/24) | 0.9 | (= 9/1002) | 1.8 | 68.3 | 60.1 |
| | | | | | | 5.2 | 0.0017 | 37.5 | (= 9/24) | 0.9 | (= 9/1042) | 1.7 | 65.7 | 59.4 |
| | | | | | | 5.4 | 0.0019 | 37.5 | (= 9/24) | 1.0 | (= 9/932) | 1.9 | 73.5 | 61.4 |
| Off Ibaraki | 1961–2010 | 2 | 0.5 | 1 | 6.0 | 4.8 | 0.0090 | 50.0 | (= 3/6) | 5.1 | (= 3/59) | 9.2 | 309.5 | 28.3 |
| | (18262 d) | | | | | 5.0 | 0.0078 | 50.0 | (= 3/6) | 5.1 | (= 3/59) | 9.2 | 309.5 | 28.3 |
| | | | | | | 5.2 | 0.0207 | 33.3 | (= 2/6) | 8.7 | (= 2/23) | 13.8 | 529.3 | 19.8 |
| | | | | | | 5.4 | 0.0761 | 16.7 | (= 1/6) | 16.7 | (= 1/6) | 16.7 | 1014.6 | 10.0 |
| Central Honshu | 1998–2019 | 63 | 0.2 | 1 | 5.0 | 3.3 | 0.0058 | 13.3 | (= 2/15) | 3.0 | (= 2/67) | 4.9 | 1007.4 | 21.9 |
| | (8035 d) | | | | | 3.5 | 0.0052 | 13.3 | (= 2/15) | 2.9 | (= 2/69) | 4.8 | 978.2 | 21.8 |
| | | | | | | 3.7 | 0.0047 | 13.3 | (= 2/15) | 2.9 | (= 2/68) | 4.8 | 992.6 | 21.9 |
| | | | | | | 3.9 | 0.0056 | 13.3 | (= 2/15) | 3.5 | (= 2/57) | 5.6 | 1184.1 | 22.6 |
| Izu Islands | 1977–2019 | 54 | 0.2 | 1 | 5.0 | 3.3 | 0.0021 | 52.3 | (= 34/65) | 1.3 | (= 34/2521) | 2.6 | 176.0 | 303.9 |
| | (15705 d) | | | | | 3.5 | 0.0011 | 53.8 | (= 35/65) | 0.9 | (= 35/4055) | 1.7 | 112.6 | 282.6 |
| | | | | | | 3.7 | 0.0031 | 47.7 | (= 31/65) | 1.7 | (= 31/1866) | 3.2 | 216.8 | 287.6 |
| | | | | | | 3.9 | 0.0014 | 52.3 | (= 34/65) | 1.2 | (= 34/2928) | 2.3 | 151.5 | 293.7 |

**Table S4.** An example of stationary space-time ETAS parameters estimated using data of the minimum $M_{th}$ shown in Table S1.

| Parameter | Off Iwate & Miyagi | Off Ibaraki | Central Honshu | Izu Islands |
|---|---|---|---|---|
| $\mu$ | 1.045 | 1.078 | 1.094 | 1.084 |
| $K$ (events) | 0.166 | 0.095 | 0.223 | 0.655 |
| $c$ (days) | 0.024 | 0.025 | 0.003 | 0.010 |
| $\alpha$ (magnitude$^{-1}$) | 1.863 | 1.876 | 1.484 | 0.548 |
| $p$ | 1.152 | 1.322 | 1.124 | 1.405 |
| $d$ (degrees$^2$) | 0.002 | 0.006 | 1.7E-05 | 4.2E-05 |
| $q$ | 2.038 | 3.376 | 2.133 | 1.784 |
| $\gamma$ (magnitude$^{-1}$) | 1.625 | 1.243 | 1.469 | 1.022 |

**Table S5.** Medians of optimum parameters and results obtained by applying Maeda's method to ETAS catalogs.

| Area | Period | Parameter | | | | | | Alarm rate | | Truth rate | | F* (%) | PG | # of valid** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $D$ (°) | $M_{f0}$ | $T_f$ (days) | $N_f$ | $T_a$ (days) | $M_{m0}$ | (%) | # of target[+] | (%) | # of alarm | | | ETAS catalogs |
| Off Iwate and Miyagi | 1961–2010 | 0.5 | 5.0 | 1 | 2 | 2 | 6.0 | 9.1 | 20 | 8.3 | 25 | 8.7 | 292.7 | 753 |
| Off Ibaraki | " | " | " | 1 | 2 | 1 | " | 14.3 | 8 | 15.4 | 8 | 15.4 | 446.5 | 387 |

*Note:* [+]The number of target earthquakes is the median value in valid ETAS catalogs; it is not the value obtained from all ETAS catalogs. *The F-measure is the median value in valid ETAS catalogs; it is not calculated from the means of AR and TR. **The term "valid" means the number of cases that ΔAIC can be calculated, having one or more true alarm earthquakes, among the 1,000 ETAS catalogs.