# Event-Feature-Based Clustering Reveals Continuous Distribution of Tectonic Tremors of 0.3-100 s: Application to Western Japan

Seiya Yano<sup>1</sup> and Satoshi  $Ide^2$ 

<sup>1</sup>The University of Tokyo <sup>2</sup>University of Tokyo

March 13, 2024

#### Abstract

We develop a methodology to compile an objective tremor catalog by combining the envelope cross-correlation method with clustering, utilizing distinctive event features that can differentiate tectonic tremors from non-tremor events. This approach enables tremor extraction without subjective criteria, allowing for the detection of short-duration tremors overlooked in previous studies. The event features are depth, the mean amplitudes at high and low frequencies, the ratio of these two amplitudes, and event duration, which is defined as the minimum period within which 50 % of the seismic energy is concentrated. The application of this method for western Japan detects 1.7 times more tremors compared with previous studies, with durations of 0.3 to  $\sim 100$  s. The events with short durations are considered to be low-frequency earthquakes. The relationship between seismic moment and the duration of the detected tremors adheres to the scaling law of slow earthquakes.

#### Hosted file

Manuscript.docx available at https://authorea.com/users/749852/articles/721012-event-feature-based-clustering-reveals-continuous-distribution-of-tectonic-tremors-of-0-3-100-s-application-to-western-japan

#### Hosted file

SupportingInformation.docx available at https://authorea.com/users/749852/articles/721012event-feature-based-clustering-reveals-continuous-distribution-of-tectonic-tremors-of-0-3-100-s-application-to-western-japan

1	
2	Event-Feature-Based Clustering Reveals Continuous Distribution of
3	Tectonic Tremors of 0.3–100 s: Application to Western Japan
4	
5	Seiya Yano <sup>1</sup> , Satoshi Ide <sup>1</sup>
6	
7	<sup>1</sup> Department of Earth and Planetary Science, The University of Tokyo, Tokyo, Japan
8	
9	Corresponding author: Seiya Yano ( <u>s-yano@eps.s.u-tokyo.ac.jp</u> )
10	
11	Key Points:
12	• We compile a more complete tectonic tremor catalog for western Japan using a
13	clustering method based on event features.
14	• Event duration, newly defined using energy radiation, clearly separates tectonic
15	tremors from fast earthquakes.
16	• Tectonic tremors, ranging in duration from 0.3 to 100 s, are consistent with the
17	scaling law of slow earthquakes.
18	
19	
20	

#### 21 Abstract

22 We develop a methodology to compile an objective tremor catalog by combining the 23 envelope cross-correlation method with clustering, utilizing distinctive event features that can 24 differentiate tectonic tremors from non-tremor events. This approach enables tremor 25 extraction without subjective criteria, allowing for the detection of short-duration tremors 26 overlooked in previous studies. The event features are depth, the mean amplitudes at high and 27 low frequencies, the ratio of these two amplitudes, and event duration, which is defined as the 28 minimum period within which 50 % of the seismic energy is concentrated. The application of 29 this method for western Japan detects 1.7 times more tremors compared with previous studies, 30 with durations of 0.3 to  $\sim 100$  s. The events with short durations are considered low-frequency 31 earthquakes. The relationship between seismic moment and duration of the detected tremors 32 adheres to the scaling law of slow earthquakes.

### 33 Plain Language Summary

34 Slow earthquakes are characterized by very slow underground deformation compared with 35 regular (fast) earthquakes and are important in understanding the preparation period prior to 36 large earthquakes. Tectonic tremors are a type of slow earthquakes. These tremors radiate tiny 37 seismic waves with frequencies of several Hz, occur episodically and densely in space and 38 time, and may last for long durations of up to several hundred seconds, which is much longer 39 than the durations of fast earthquakes of equivalent magnitude. In this study, we detect 40 tectonic tremors and differentiate them from fast earthquakes and anthropogenic events using 41 a set of event features without relying on subjective criteria. The durations of tremors range 42 from 0.3 to ~100 s and appear to be consistent with a previously proposed scaling 43 relationship for slow earthquakes, which suggests that fast earthquakes and slow earthquakes 44 have different physical mechanisms.

45

#### 46 **1 Introduction**

The recent development of seismological and geodetic observation networks has led to the discovery of slow earthquakes, which include phenomena such as tectonic tremors and low-frequency earthquakes (LFEs) observed in a seismological frequency range of 1–10 Hz, very low-frequency earthquakes (VLFEs) in a frequency range of 0.02–0.05 Hz, and geodetically observed slow slip events (SSEs). The history of the discovery of these events has been summarized by Beroza and Ide (2011), and their potential relationships with large or giant earthquakes have been discussed by Obara and Kato (2016). Recent developments in slow earthquake research have been reviewed by Nishikawa et al. (2023), and geological
aspects of slow earthquakes have been reviewed by Behr and Bürgmann (2021).

56 Although the time scales vary by more than 6–7 orders of magnitude from LFEs (<1 s) to SSEs (>1 d), these slow earthquakes are strongly correlated both spatially and 57 58 temporally, and have similar deformation mechanisms; consequently, they have been 59 hypothesized to be different manifestations of a unified phenomenon across various 60 frequency ranges (Ide et al., 2007). One feature that distinguishes these unified slow 61 earthquakes from fast earthquakes is the relationship between seismic moment and duration. 62 For slow earthquakes, a linear relationship between seismic moment and duration delineates 63 the upper limit of slow earthquakes (Ide et al., 2007; Ide & Beroza, 2023), whereas the 64 seismic moment of fast earthquakes is proportional to the cube of duration. In accordance with the hypothesis of a unified slow earthquake process, we may use seismically detected 65 tremors as indicators of a longer geodetic process, as assumed in many recent studies (Aguiar 66 et al., 2009; Bletery et al., 2017; Gomberg et al., 2016; Gombert & Hawthorne, 2023; Wech 67 68 et al., 2010).

69 One of the most commonly employed methods to detect tremors is the envelope 70 cross-correlation method (Ide, 2010; Obara, 2002; Wech & Creager, 2008), in which the 71 inter-station similarity of S-wave envelopes is used to detect and locate tremors. However, 72 this method detects all seismic events and cannot distinguish between slow and fast 73 earthquakes. Therefore, additional selection criteria are required to compile a tremor catalog 74 from detected candidate events. Previously adopted criteria include spatiotemporal event 75 density (Wech & Creager, 2008), duration threshold (Mizuno & Ide, 2019), and even visual 76 inspection (Romanet & Ide, 2019). Although these criteria rely on empirical knowledge or 77 trial and error, better criteria may be established using more fundamental event features that 78 differentiate slow earthquakes from fast earthquakes.

79 It is already known that (1) tremor depths are limited in some ranges, (2) tremors 80 may last for seconds to minutes, (3) tremors lack high-frequency seismic wave energy, and 81 (4) tremors tend to cluster spatiotemporally. Based on these observations, this study develops 82 a method to compile a tremor catalog without subjective criteria by establishing a consistent 83 framework for event detection, event feature calculation, and tremor extraction through 84 machine learning. In this paper, we introduce the evaluation of event features and the 85 clustering method in the Method section. We apply the evaluation and method to continuous 86 seismic records for western Japan, one of the most intensively studied slow earthquakes

87 regions. The findings are given in the Result section. The tremors we detect have durations of 88 0.3–100 s and show a scaling law between duration and seismic energy. In the Discussion 89 section, we consider the significance of this scaling law compared with the previously 90 suggested relationship between duration and seismic moment. Conclusion section 91 summarizes our study.

92

### 93 **2 Method**

94 We compile a tremor catalog through four steps (Figure 1). In the first step, we 95 detect tremor candidates from continuous waveform data using the envelope cross-correlation 96 method (Mizuno & Ide, 2019). In the second step, we compute event features for all detected 97 candidate events. In the third step, we apply a density-based clustering algorithm to the set of 98 event features to extract tremors. Finally, in the fourth step, to improve the completeness of 99 the catalog, we reclassify noise events, which are not assigned to any meaningful clusters by 100 the clustering algorithm, through supervised learning. The training data for supervised 101 learning are prepared by removing isolated events from the established tremor clusters using 102 the modified spatiotemporal clustering method of Zaliapin and Ben-Zion (2013). 103



104

Figure 1. Flow chart of the procedure for compiling tremor catalog. The four steps explained
in the text correspond to the parts surrounded by red (step 1), orange (step 2), green (step 3),
and blue (step 4).

108

### 109 2.1 Detection of Candidate Events

110 We employ the method of Mizuno and Ide (2019) to detect seismic events. Those 111 authors improved the conventional tremor detection methods (Wech & Creager, 2007, 2008) by integrating envelope cross-correlation and the maximum likelihood method, assuming that 112 113 the observed envelope waveform at each seismic station results from the sum of a common 114 envelope waveform emitted from the source and Gaussian noise. The input data are velocity 115 seismograms resampled at 20 samples per second (sps), bandpass filtered at 2–8 Hz, squared, and resampled at 1 sps with a lowpass filter below 0.2 Hz. The square root of these data 116 117 effectively approximates the envelopes.

118 We search for events in each half-overlapping time window, spanning 300 s and 119 shifted by 150 s. A single time window may contain multiple events because the method considers local maxima rather than a global maximum of the likelihood function. The 120 121 bootstrap method is applied to assess the error in source determination, and events with 122 standard errors exceeding 2.0 km are excluded from further analysis. We also compute an 123 azimuthal gap for each event from all stations used to determine the hypocenter and remove 124 the event if the azimuthal gap exceeds 180°. Owing to the overlap of the time window, one 125 event is often identified twice in consecutive windows. We removed such multiple detections. 126 The result of the procedure described above is a catalog of candidate events.

127 The hypocentral time and size of each event are measured as follows. We 128 backproject the waveforms at each station employed in the estimation of the hypocenter by 129 the time-shift caused by the S-wave travel time and square the waveforms multiplied by the 130 hypocentral distance. When multiplied by  $4\pi\rho\beta$ , the mean of these functions approximates the seismic energy rate, which we denote as  $\dot{E}(t)$ , where  $\rho$  and  $\beta$  are the density and 131 S-wave velocity, respectively. The hypocentral time  $t_{hypo}$  is identified as the moment when 132  $\dot{E}(t)$  attains its maximum. We define the stacked envelope duration  $\tau_s$  as the period during 133 134 which  $\dot{E}(t)$  exceeds one-quarter of its maximum value, and we quantify the observed seismic energy  $E_s^{obs}$  as the integral of  $\dot{E}(t)$  over the period  $\tau_s$ . We also divide  $\tau_s$  into 135  $\tau_{pre}$  and  $\tau_{pos}$ , representing the periods before and after the peak of  $\dot{E}(t)$ , respectively. 136

137

### 138 2.2 Event Features

We extract tremors by applying a clustering method to a set of features of detected candidates. One of the features is hypocentral depth, as tremor depth is known to be limited to a certain range. We introduce other features related to event duration and signal amplitudes, 142 as described below.

143 Event duration is useful in selecting tectonic tremors. However, the stacked envelope 144 duration  $\tau_s$  is usually longer than a few seconds because we measure it using the smoothed seismic wave energy rate resampled at 1 sps after lowpass filtering below 0.2 Hz. The energy 145 146 rate is also the average over all stations, which may include elongated signals caused by later phases, known to be critical for estimating the duration of shallow tectonic tremors (Toh et al., 147 2023). Given this problem,  $\tau_s$  tends to overestimate the duration of the real source process, 148 149 if it is shorter than a few seconds. Therefore, we propose a new measure denoted as the 150 minimum energy duration  $\tau_e$ , which represents the shortest duration over which half the 151 seismic energy is concentrated. For each event, each original velocity seismogram is 152 bandpass filtered at 2–8 Hz, squared, and integrated for the period  $2\tau_s$  (Figure S1 in Supporting Information). The resulting function  $E_i(t)$  approximates the seismic energy 153 154 measured at the *i*-th channel. Given  $E_i(t)$ , we measure the shortest period  $\tau_i$  that covers half the total seismic energy at the channel. We define the minimum of  $\{\tau_i\}$  as the minimum 155 156 energy duration  $\tau_e$ . Computational details are explained in Text S1 in Supporting 157 Information.

As tremors usually lack high-frequency seismic energy, information about the relative amplitude between high- and low-frequency signals is also useful for selecting tremors (e.g. Nakajima & Hasegawa, 2021; Sit et al., 2012). In the present study, we define two features of the frequency domain: high-frequency amplitude  $A_H$  and low-frequency amplitude  $A_L$ . These are the representative values of the spectral amplitude of original velocity waveforms within high- and low-frequency bands, respectively. Details of the calculation are explained in Text S2 in Supporting Information.

We use depth,  $\tau_e$ ,  $A_H$ ,  $A_L$ , and  $A_H/A_L$  as features for clustering. Each feature may have a distribution that differs markedly from a normal distribution. To facilitate the clustering, we subject all event features to the Box–Cox transformation (Box & Cox, 1964) and subsequently standardize them to achieve a mean of 0 and a variance of 1 (Text S3 in Supporting Information).

170

171 2.3 Clustering of Candidate Events

We extract tremors from the detected candidate events using HDBSCAN (Campello et al., 2013), a density-based clustering method. HDBSCAN utilizes a mutual reachability distance instead of a Euclidean distance to efficiently form clusters with disparate densities.

The algorithm constructs a minimum spanning tree and subsequently determines whether 175 176 each node and following trees should be considered as parts of distinct clusters or the same 177 cluster, based on a stability index. We need to specify only one parameter, denoted as  $m_{pts}$ , which represents the minimum number of elements required for cluster generation. The 178 179 optimal value of  $m_{pts}$  can be objectively determined by assessing the density-based 180 clustering validation (DBCV) score (Moulavi et al., 2014). The DBCV score is an index that 181 compares the density in the sparsest region among clusters with that in the densest region among inter-cluster regions and deems a model favorable if the former surpasses the latter. In 182 our study, we define  $m_{pts}$  as the number corresponding to p % of the total number of 183 184 candidate events. The value of p is selected by varying it from 0.05 to 5.00 in increments of 185 0.01 and choosing the p that yields the best DBCV score.

186

### 187 2.4 Reclassification Using Supervised Learning

Density-based clustering algorithms tend to classify events in low-density regions as noise events that do not belong to any of the meaningful clusters. This raises the possibility of improper clustering for the end-members of tremors, such as events with very long durations. Given that the clustering algorithm yields clusters  $C_{tremor}$ ,  $C_1$ ,  $C_2$ , ..., and  $C_{noise}$ , where  $C_{tremor}$  and  $C_{noise}$  represent the tremor cluster and noise event group, respectively, we attempt to reclassify the events in  $C_{noise}$  using supervised learning to construct a more comprehensive tremor catalog.

195 As the outcome of supervised learning is strongly influenced by the quality of 196 training data, it is preferable to ensure that the tremor cluster for training data does not 197 include non-tremor events. Given that tremors tend to occur close to each other, both spatially 198 and temporally (e.g., Obara, 2002), we remove spatiotemporally isolated events from 199  $C_{tremor}$  and add them to  $C_{noise}$ . We denote the clusters of tremors and the set of noise events after this change as  $C'_{tremor}$  and  $C'_{noise}$ , respectively. We designate the tremors in 200  $C'_{tremor}$  as "Quality A" tremors. We use  $C_1, C_2, \dots$ , and  $C'_{tremor}$  (not  $C_{tremor}$ ) for training 201 data and reclassify the events in  $C'_{noise}$  into all the established clusters  $C_1, C_2, \dots$ , and 202 203  $C'_{tremor}$ .

We identify isolated events using the modified nearest-neighbor clustering method of Zaliapin and Ben-Zion (2013), in which the association between seismic events in a region can be quantified by the product of spatial and temporal distances between events. Using this spatiotemporal distance, we can distinguish between isolated events following a Poisson process and triggered events (Baiesi & Paczuski, 2004; Zaliapin et al., 2008; Zaliapin & Ben-Zion, 2013). In the present study, we modify the definition of temporal distance and define it as the time difference between events divided by event duration  $\tau_e$ . The details are given in Text S4 in Supporting Information.

We employ a neural network for the supervised learning reclassification. A neural network is a multi-layered network comprising one input layer, one or multiple hidden layers, and one output layer (Figure S2 in Supporting Information). By optimizing the parameters of the hidden layers using training data, the network can transform new input data linearly or nonlinearly in the hidden layers and output the probability of the data belonging to a specific group. In this study, we first optimize the neural network using the clusters  $C'_{tremor}$ ,  $C_1$ ,  $C_2$ , ..., and then reclassify each event in  $C'_{noise}$ .

219 As the result of the neural network has a weak dependence on the initial values of 220 weights, which represent the coupling strength between two neurons, we perform 221 reclassification five times while changing the initial values. Only those events that are 222 reclassified as tremors with a probability of more than 99 % for all five times are assigned as 223 "Quality B" tremors. Events whose probability of belonging to the tremor cluster is less than 224 99 % at least once are assigned as neither Quality A nor Quality B tremors. The event information for training and reclassification includes  $\tau_e$ ,  $A_H$ , and  $A_L$ , with depth and 225  $A_H/A_L$  being excluded. Hyperparameters of the neural network, such as the number of 226 227 neurons, the number of hidden layers, and activation function, are determined through the 228 procedure explained in Text 5 in Supporting Information.

Therefore, the final tremor catalog comprises two categories of events: those in  $C'_{tremor}$  (Quality A) and those classified as tremors through the reclassification by neural network (Quality B). The former events are more reliable, and the latter events are included to improve the completeness of the tremor catalog.

- 233
- 234 3 Data

We apply the method described above to continuous seismograms recorded at 366 Hi-net stations maintained by the National Research Institute for Earth Science and Disaster Resilience in western Japan (Figure 2a and Table S1 in Supporting Information). We use two horizontal components of velocity data originally recorded at 100 sps. The analysis period covers one year (2010). The travel time is calculated based on the travel time table of the Japan Meteorological Agency (JMA) (Ueno et al., 2002).





242

243 Figure 2. Distributions of (a) tremors detected in a previous study (Mizuno & Ide, 2019), (b) 244 all seismic events detected in this study, (c) Quality A tremors, and (d) Quality B tremors. 245 Red triangles and inversed gray triangles in (a) represent volcanoes and the Hi-net seismic 246 stations used in this study, respectively. Dashed lines are contours of the upper surface of the subducting Philippine plate (Baba et al., 2002; Hirose et al., 2008; Nakajima & Hasegawa, 247 248 2007). Seafloor topography in (a) is based on Tozer et al. (2019).

249

#### 250 **4 Results**

251 The envelope cross-correlation method detects 56,775 candidate events (Figure 2b). 252 The depth distribution of these events has two distinct peaks (Figure 3a). The shallow (~5 km) peak corresponds to shallow crustal earthquakes, tremors, and anthropogenic events,
whereas the deep peak (~35 km) represents deep crustal earthquakes and tremors around the
plate boundary. The events deeper than 60 km are intraslab earthquakes.

The minimum energy duration  $\tau_e$  is shorter than the stacked envelope duration  $\tau_s$ (Figure S3 in Supporting Information).  $\tau_e$  has a conspicuous low-density region around  $\tau_e = 0.4$  s (Figure 3b), clearly separating fast and slow earthquakes. The separation cannot be observed in  $\tau_s$  (Figure 3d), as also reported in the observation of tectonic tremors in New Zealand (Romanet & Ide, 2019).

261 The two-dimensional density profile with  $A_L$  and  $A_H$  exhibits three distinct modes 262 characterized by different values of  $A_H/A_L$  (Figure 3c), where we define  $A_H$  as the mean 263 amplitude of 10–20 Hz and  $A_L$  as that of 2–8 Hz, given that the dominant frequencies for 264 Mw 1 fast earthquakes and tremors are 10–20 Hz and 1–8 Hz, respectively (Beroza & Ide, 2011), and the influence of microseism extends up to 2 Hz (Peterson, 1993). Following the 265 Box-Cox transformation and standardization of depth,  $\tau_e$ ,  $A_H$ ,  $A_L$ , and  $A_H/A_L$ , the 266 HDBSCAN algorithm identifies four meaningful clusters: Cshallow, comprising 11,795 267 events corresponding to fast earthquakes in the crust;  $C_{deep}$ , encompassing 2,116 events 268 corresponding to fast earthquakes near the plate interface; C<sub>tremor</sub>, comprising 24,881 events 269 indicative of tremors; and Canthropogenic, including 7,218 events corresponding to 270 271 anthropogenic seismic events (Figure S4 in Supporting Information). HDBSCAN categorizes 272 the remaining 10,765 events as noise events. Here, the clustering parameter  $m_{pts}$  is 34, 273 corresponding to p = 0.06 %.

274 In the tremor cluster  $C_{tremor}$ , 1,859 events are identified as isolated events, resulting in 23,022 Quality A tremors. Subsequently, employing the neural network trained 275 276 with  $C_{shallow}$  (11,795 events),  $C_{deep}$  (2,116 events),  $C'_{tremor}$  (23,022 events), and 277  $C_{anthropogenic}$  (7,218 events), we reclassify the 7,582 events that occurred at depths from 10 km to 60 km out of 12,624 events in  $C'_{noise}$ . Events in  $C'_{noise}$  with depths shallower than 10 278 279 km or deeper than 60 km are excluded before the reclassification because almost all of the 280 events in the established four clusters occur within 10-60 km. We employ the neural network 281 with eight neurons, two hidden layers, and the activation function tanh(x) for the 282 reclassification, resulting in 830 Quality B tremors. Consequently, we obtain 23,022 Quality 283 A tremors and 830 Quality B tremors (Figure 2c, d, Data Set S1 in Supporting Information).

Of 23,852 tremors (Quality A and B), 1,756 events are also listed in the JMA catalog, and 99.6 % of these events are labeled as LFE by the JMA; for the Quality A tremors, the rate of correct classification exceeds 99.7 %. The events that are also listed in the JMA catalog
tend to have a relatively large seismic energy among all tremors detected in this study. Figure
S5 in Supporting Information shows an example of events classified as tremor in this study
but not labeled as LFE in the JMA catalog.

290



291

Figure 3. Distributions of (a) depth, (b) minimum energy duration  $\tau_e$ , (c) low-frequency amplitude  $A_L$  and high-frequency amplitude  $A_H$ , and (d) stacked envelope duration  $\tau_s$ . Dashed lines in (c) are  $A_H = A_L$ ,  $A_H = 0.4 \times A_L$ , and  $A_H = 0.2 \times A_L$ .

295

### 296 **5 Discussion**

297 5.1 Duration Statistics

We detected 1.7 times more tremors compared with Mizuno and Ide (2019). This increase is due mainly to the detection of many short-duration events that were excluded in the previous study. The frequency distribution of duration in both  $\tau_s$  and  $\tau_e$  follows an exponential law (Figure S6 in Supporting Information). This result is consistent with previous studies (Chestler & Creager, 2017; Hiramatsu et al., 2008; Watanabe et al., 2007; Yabe & Ide, 2014), suggesting a scale limit in the tremor source (Aki & Koyanagi, 1981; Benoit et al.,
2003).

305

306 5.2 Scaling Relationship

The duration of tremors in this study ranges between 0.3 to 100 s. Figure 4a compares observed seismic energy  $E_s^{obs}$  and minimum energy duration  $\tau_e$ , revealing that  $\tau_e$  of tremors increases almost linearly with  $E_s^{obs}$ , whereas the relationship for fast earthquakes is almost constant. However, caution is needed when interpretating the diagram as  $E_s^{obs}$  is affected by bandwidth limitation and attenuation.

For fast earthquakes, we estimate the seismic moment  $M_0$  and event duration  $\tau$ 312 313 from  $E_s^{obs}$  by assuming the omega-square model and circular faults (Aki, 1967; Madariaga, 1976). Details of the calculation are described in Text S6 in Supporting Information. 314 However, the effects of bandwidth limitation and attenuation on tremors are almost negligible 315 316 owing to the low predominant frequency and its coincidence with the frequency band we 317 used. Several previous studied have demonstrated that the scaled energy of LFEs, VLFEs, and SSEs is around  $10^{-10}$  to  $10^{-9}$  (Maeda & Obara, 2009; Maury et al., 2018; Yabe & Ide, 318 2014). Given these reports, we convert  $E_s^{obs}$  of tremors to  $M_0$  by multiplying by  $10^{-9.5}$ . 319 320 Figure 4b compares the estimated seismic moment and event duration. As we assume an 321 omega-square model with constant stress drop, it is unsurprising that all fast earthquakes fall 322 along a single straight line.

323 The scaling law of slow earthquakes within the time scale range of 0.5-20 s 324 (between LFEs and VLFEs) has not been comprehensively observed or discussed. Previous studies examining a time scale of 1 s or less have produced varying conclusions, with reports 325 suggesting that the scaling of slow earthquakes is close to  $M_0 \propto \tau^3$  (Supino et al., 2020), or 326 neither  $M_0 \propto \tau^3$  nor  $M_0 \propto \tau$  (Bostock et al., 2015; Farge et al., 2020). However, our results 327 328 suggest that tremors follow the scaling relationship  $M_0 \propto \tau$  even for events with a duration 329 of less than 1 s. Those events with the shortest durations are impulsive and can be recognized 330 as simple LFEs, whereas those with duration of more than several tens of seconds would 331 include many LFEs. Therefore, our catalog contains a seamless distribution of LFEs and 332 longer tremors.

333





Figure 4. (a) Comparison of observed seismic energy  $E_s^{obs}$  and minimum energy duration  $\tau_e$ . Light and dark blue dots correspond to  $C_{shallow}$  and  $C_{deep}$ , respectively. Orange dots depict Quality A and B tremors. (b) Comparison of estimated seismic moment  $M_0$  and event duration, where the duration of fast earthquakes is estimated from  $E_s^{obs}$  using the omega-square model, whereas that of tremors is  $\tau_e$  itself. The two gray lines represent  $M_0 = \tau^3 \times 10^{15-16}$  for fast earthquakes and  $M_0 = \tau \times 10^{12-13}$  for slow earthquakes. The dashed ellipse shows the region containing 95 % of the tremors.

342

#### 343 6 Conclusion

We improved the envelope cross-correlation method proposed by Mizuno and Ide (2019) to compile a more comprehensive catalog of tectonic tremors. We introduced a new definition of duration (minimum energy duration) and calculated mean amplitudes at highand low-frequency ranges. Using these event features, we developed a scheme for compiling an objective tremor catalog using a density-based clustering algorithm and supervised learning. Detected tremors are categorized into Quality A and Quality B tremors, constituting more reliable and more complete catalogs, respectively.

351 The application of our method to continuous seismograms for western Japan detects 352 1.7 times more tremors compared with the original method. The spatial distribution of 353 tremors is similar to that of previous results. Comparison with the JMA catalog reveals that 354 the accuracy of our tremor detection exceeds 99 %. Detected tremors have a wide range of 355 durations, from 0.3 to 100 s, encompassing 8,415 events with short durations that were 356 previously rejected by the conventional method. Some of these short-duration events are 357 LFEs that have been manually identified by the JMA. The duration distribution is described 358 well by an exponential law, suggesting the existence of a scale limit in the background 359 physical process.

360 This study is the first to provide a comprehensive scaling relationship for slow 361 earthquakes, ranging from LFEs with durations of  $\sim 0.3$  s to tremors with durations of more 362 than 100 s. This relationship bridges the 1–10 s gap, in which seismological observation is 363 difficult due to large microseism noise. Given that the proportionality between seismic energy 364 and seismic moment holds for this scale range, the seismic moment is proportional to the duration, with a proportionality coefficient of approximately  $10^{12-13}$  Nm/s, which is 365 consistent with the relationship suggested for the very wide range of time scales from LFEs 366 367 to SSEs.

368

### 369 Acknowledgements

This research is supported by JSPS Kakenhi (21H04505), MEXT Kakenhi
(21H05200), and Earthquake and Volcano Hazards Observation and Research.

- 372
- 373 **Open Research**
- 374

The seismic data of Hi-net are available at https://www.hinet.bosai.go.jp/?LANG=en

after obtaining an ID from https://hinetwww11.bosai.go.jp/nied/registration/?LANG=en. The code of seismic event detection by Mizuno and Ide (2019) is available at https://github.com/not522/MizunoIde2019. Our algorithm is implemented using NLopt (Johnson, 2014), GNU Parallel (Tange, 2018), and Scikit-learn (Pedregosa et al., 2018). The figures in this manuscript are prepared using Matplotlib (Hunter, 2007) and GMT (Wessel et al., 2019).

381

## 382 **References**

- Aguiar, A. C., Melbourne, T. I., & Scrivner, C. W. (2009). Moment release rate of Cascadia
  tremor constrained by GPS. *Journal of Geophysical Research: Solid Earth*, *114*(B7).
  https://doi.org/10.1029/2008JB005909
- 386 Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research* 387 (1896-1977), 72(4), 1217–1231. https://doi.org/10.1029/JZ072i004p01217
- Aki, K., & Koyanagi, R. (1981). Deep volcanic tremor and magma ascent mechanism under
  Kilauea, Hawaii. *Journal of Geophysical Research: Solid Earth*, 86(B8), 7095–7109.
  https://doi.org/10.1029/JB086iB08p07095
- 391 Baba, T., Tanioka, Y., Cummins, P. R., & Uhira, K. (2002). The slip distribution of the 1946 392 Nankai earthquake estimated from tsunami inversion using a new plate model. 393 **Physics** ofthe Earth and **Planetary** Interiors, 132(1), 59-73. 394 https://doi.org/10.1016/S0031-9201(02)00044-4
- Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks.
   *Physical Review E*, 69(6), 066106. https://doi.org/10.1103/PhysRevE.69.066106
- Behr, W. M., & Bürgmann, R. (2021). What's down there? The structures, materials and
  environment of deep-seated slow slip and tremor. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2193),
  20200218. https://doi.org/10.1098/rsta.2020.0218
- 401 Benoit, J. P., McNutt, S. R., & Barboza, V. (2003). Duration-amplitude distribution of
  402 volcanic tremor. *Journal of Geophysical Research: Solid Earth*, 108(B3).
  403 https://doi.org/10.1029/2001JB001520
- Beroza, G. C., & Ide, S. (2011). Slow Earthquakes and Nonvolcanic Tremor. *Annual Review of Earth and Planetary Sciences*, *39*(1), 271–296.
  https://doi.org/10.1146/annurev-earth-040809-152531
- 407 Bletery, Q., Thomas, A. M., Hawthorne, J. C., Skarbek, R. M., Rempel, A. W., & Krogstad, R.

- 408 D. (2017). Characteristics of secondary slip fronts associated with slow earthquakes in
  409 Cascadia. *Earth and Planetary Science Letters*, 463, 212–220.
  410 https://doi.org/10.1016/j.epsl.2017.01.046
- Bostock, M. G., Thomas, A. M., Savard, G., Chuang, L., & Rubin, A. M. (2015). Magnitudes
  and moment-duration scaling of low-frequency earthquakes beneath southern
  Vancouver Island. *Journal of Geophysical Research: Solid Earth*, *120*(9), 6329–6350.
  https://doi.org/10.1002/2015JB012195
- Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211–252.
- 417 Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-Based Clustering Based on
  418 Hierarchical Density Estimates. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu
  419 (Eds.), Advances in Knowledge Discovery and Data Mining (pp. 160–172). Berlin,
- 420 Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37456-2\_14
- 421 Chestler, S. R., & Creager, K. C. (2017). Evidence for a scale-limited low-frequency
  422 earthquake source process. *Journal of Geophysical Research: Solid Earth*, *122*(4),
  423 3099–3114. https://doi.org/10.1002/2016JB013717
- 424 Farge, G., Shapiro, N. M., & Frank, W. B. (2020). Moment-Duration Scaling of
  425 Low-Frequency Earthquakes in Guerrero, Mexico. *Journal of Geophysical Research:*426 Solid Earth, 125(8), e2019JB019099. https://doi.org/10.1029/2019JB019099
- Gomberg, J., Wech, A., Creager, K., Obara, K., & Agnew, D. (2016). Reconsidering
  earthquake scaling. *Geophysical Research Letters*, 43(12), 6243–6251.
  https://doi.org/10.1002/2016GL069967
- Gombert, B., & Hawthorne, J. C. (2023). Rapid Tremor Migration During Few Minute-Long
  Slow Earthquakes in Cascadia. *Journal of Geophysical Research: Solid Earth*, *128*(2),
  e2022JB025034. https://doi.org/10.1029/2022JB025034
- Hiramatsu, Y., Watanabe, T., & Obara, K. (2008). Deep low-frequency tremors as a proxy for
  slip monitoring at plate interface. *Geophysical Research Letters*, 35(13).
  https://doi.org/10.1029/2008GL034342
- Hirose, F., Nakajima, J., & Hasegawa, A. (2008). Three-dimensional seismic velocity
  structure and configuration of the Philippine Sea slab in southwestern Japan estimated
  by double-difference tomography. *Journal of Geophysical Research: Solid Earth*, *113*(B9). https://doi.org/10.1029/2007JB005274
- 440 Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science &

441 *Engineering*, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

- Ide, S. (2010). Striations, duration, migration and tidal response in deep tremor. *Nature*,
  443 466(7304), 356–359. https://doi.org/10.1038/nature09251
- Ide, S., & Beroza, G. C. (2023). Slow earthquake scaling reconsidered as a boundary between
  distinct modes of rupture propagation. *Proceedings of the National Academy of Sciences*, *120*(32), e2222102120. https://doi.org/10.1073/pnas.2222102120
- Ide, S., Beroza, G. C., Shelly, D. R., & Uchide, T. (2007). A scaling law for slow earthquakes.
   *Nature*, 447(7140), 76–79. https://doi.org/10.1038/nature05780
- 449 Johnson, S. G. (2014). The NLopt nonlinear-optimization package, 450 2024. http://github.com/stevengj/nlopt. Retrieved January 11. from 451 https://nlopt.readthedocs.io/en/latest/Citing\_NLopt/
- 452 Madariaga, R. (1976). Dynamics of an Expanding Circular Fault. Bulletin of the
  453 Seismological Society of America, 66, 639–666.
  454 https://doi.org/10.1785/BSSA0660030639
- Maeda, T., & Obara, K. (2009). Spatiotemporal distribution of seismic energy radiation from
  low-frequency tremor in western Shikoku, Japan. *Journal of Geophysical Research: Solid Earth*, *114*(B10). https://doi.org/10.1029/2008JB006043
- Maury, J., Ide, S., Cruz-Atienza, V. M., & Kostoglodov, V. (2018). Spatiotemporal Variations
  in Slow Earthquakes Along the Mexican Subduction Zone. *Journal of Geophysical Research: Solid Earth*, *123*(2), 1559–1575. https://doi.org/10.1002/2017JB014690
- Mizuno, N., & Ide, S. (2019). Development of a modified envelope correlation method based
  on maximum-likelihood method and application to detecting and locating deep
  tectonic tremors in western Japan. *Earth, Planets and Space*, *71*(1), 40.
  https://doi.org/10.1186/s40623-019-1022-x
- Moulavi, D., Jaskowiak, P. A., Campello, R. J. G. B., Zimek, A., & Sander, J. (2014).
  Density-Based Clustering Validation. In *Proceedings of the 2014 SIAM International Conference on Data Mining* (pp. 839–847). Society for Industrial and Applied
  Mathematics. https://doi.org/10.1137/1.9781611973440.96
- 469 Nakajima, J., & Hasegawa, A. (2007). Subduction of the Philippine Sea plate beneath
  470 southwestern Japan: Slab geometry and its relationship to arc magmatism. *Journal of*471 *Geophysical Research: Solid Earth*, *112*(B8). https://doi.org/10.1029/2006JB004770
- 472 Nakajima, J., & Hasegawa, A. (2021). Prevalence of Shallow Low-Frequency Earthquakes in
  473 the Continental Crust. *Journal of Geophysical Research: Solid Earth*, 126(4),

- 474 e2020JB021391. https://doi.org/10.1029/2020JB021391
- 475 Nishikawa, T., Ide, S., & Nishimura, T. (2023). A review on slow earthquakes in the Japan
  476 Trench. *Progress in Earth and Planetary Science*, 10(1), 1.
  477 https://doi.org/10.1186/s40645-022-00528-w
- 478 Obara, K. (2002). Nonvolcanic Deep Tremor Associated with Subduction in Southwest Japan.
  479 *Science*, 296(5573), 1679–1681. https://doi.org/10.1126/science.1070378
- 480 Obara, K., & Kato, A. (2016). Connecting slow earthquakes to huge earthquakes. *Science*,
  481 353(6296), 253–257. https://doi.org/10.1126/science.aaf1512
- 482 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2018,
  483 June 5). Scikit-learn: Machine Learning in Python (Version 4). arXiv.
  484 https://doi.org/10.48550/arXiv.1201.0490
- Peterson, J. R. (1993). Observations and modeling of seismic background noise (USGS
  Numbered Series No. 93–322). Observations and modeling of seismic background
  noise (Vol. 93–322). U.S. Geological Survey. https://doi.org/10.3133/ofr93322
- 488 Romanet, P., & Ide, S. (2019). Ambient tectonic tremors in Manawatu, Cape Turnagain,
  489 Marlborough, and Puysegur, New Zealand. *Earth, Planets and Space*, *71*(1), 59.
  490 https://doi.org/10.1186/s40623-019-1039-1
- 491 Sit, S., Brudzinski, M., & Kao, H. (2012). Detecting tectonic tremor through frequency
  492 scanning at a single station: Application to the Cascadia margin. *Earth and Planetary*493 *Science Letters*, 353–354, 134–144. https://doi.org/10.1016/j.epsl.2012.08.002
- 494 Supino, M., Poiata, N., Festa, G., Vilotte, J. P., Satriano, C., & Obara, K. (2020).
  495 Self-similarity of low-frequency earthquakes. *Scientific Reports*, 10(1), 6523.
  496 https://doi.org/10.1038/s41598-020-63584-6
- 497 Tange, O. (2018). GNU Parallel 2018. Ole Tange. https://doi.org/10.5281/zenodo.1146014
- Toh, A., Capdeville, Y., Chi, W.-C., & Ide, S. (2023). Strongly Scattering Medium Along
  Slow Earthquake Fault Zones Based on New Observations of Short-Duration Tremors. *Geophysical Research Letters*, 50(8), e2022GL101851.
  https://doi.org/10.1029/2022GL101851
- Tozer, B., Sandwell, D. T., Smith, W. H. F., Olson, C., Beale, J. R., & Wessel, P. (2019).
  Global Bathymetry and Topography at 15 Arc Sec: SRTM15+. *Earth and Space Science*, 6(10), 1847–1864. https://doi.org/10.1029/2019EA000658
- 505 Ueno, H., Hatakeyama, S., Aketagawa, T., Funasaki, J., & Hamada, N. (2002). Improvement
   506 of hypocenter determination procedures in the Japan Meteorological Agency. *Quater.*

- 507 *J. Seismol.*, 65, 123–134.
- Watanabe, T., Hiramatsu, Y., & Obara, K. (2007). Scaling relationship between the duration
  and the amplitude of non-volcanic deep low-frequency tremors. *Geophysical Research Letters*, 34(7). https://doi.org/10.1029/2007GL029391
- Wech, A. G., & Creager, K. C. (2007). Cascadia tremor polarization evidence for plate
  interface slip. *Geophysical Research Letters*, 34(22).
  https://doi.org/10.1029/2007GL031167
- Wech, A. G., & Creager, K. C. (2008). Automated detection and location of Cascadia tremor. *Geophysical Research Letters*, 35(20). https://doi.org/10.1029/2008GL035458
- Wech, A. G., Creager, K. C., Houston, H., & Vidale, J. E. (2010). An earthquake-like
  magnitude-frequency distribution of slow slip in northern Cascadia. *Geophysical Research Letters*, 37(22). https://doi.org/10.1029/2010GL044881
- Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019).
  The Generic Mapping Tools Version 6. *Geochemistry, Geophysics, Geosystems*, 20(11), 5556–5564. https://doi.org/10.1029/2019GC008515
- Yabe, S., & Ide, S. (2014). Spatial distribution of seismic energy rate of tectonic tremors in
  subduction zones. *Journal of Geophysical Research: Solid Earth*, *119*(11), 8171–8185.
  https://doi.org/10.1002/2014JB011383
- Zaliapin, I., & Ben-Zion, Y. (2013). Earthquake clusters in southern California I:
  Identification and stability. *Journal of Geophysical Research: Solid Earth*, *118*(6),
  2847–2864. https://doi.org/10.1002/jgrb.50179
- Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering Analysis of
  Seismicity and Aftershock Identification. *Physical Review Letters*, *101*(1), 018501.
  https://doi.org/10.1103/PhysRevLett.101.018501
- 531
- 532
- 533
- 534
- 551
- 535
- 536
- 537