Optimal observation time window for forecasting the next earthquake

Takahiro Omi,[1,*] Ido Kanter,[2,†] and Shigeru Shinomoto[1,‡]

[1]*Department of Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan*
[2]*Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan, 52900 Israel*

We report that the accuracy of predicting the occurrence time of the next earthquake is significantly enhanced by observing the latest rate of earthquake occurrences. The observation period that minimizes the temporal uncertainty of the next occurrence is on the order of 10 hours. This result is independent of the threshold magnitude and is consistent across different geographic areas. This time scale is much shorter than the months or years that have previously been considered characteristic of seismic activities.

I. INTRODUCTION

Earthquakes occur unexpectedly in an instant, causing catastrophic damage. A great deal of effort has gone into understanding the timing of earthquake occurrences in a causal manner, for example, to understand how the main shock induces aftershocks [1–8]. By classifying earthquakes into groups of cause and result, theories such as the Omori law [1,2], which indicates that the aftershock frequency decreases roughly in proportion to the inverse time after the main shock, have succeeded in reproducing some statistical aspects of earthquake occurrence. However, the main shocks are determined *ex post facto*, and their timing is simply attributed to an unpredictable random process [8]. In contrast, a simple statistical analysis based on equal treatment of all earthquake occurrences has revealed that the distributions of interevent intervals, for different geographic areas and over a range of earthquake magnitudes, may collapse onto a unique distribution after rescaling time by the mean occurrence rate [9–14].

If interevent intervals are independent and identically distributed, the occurrence time of the next earthquake depends only on the time of the most recent occurrence. However, several researchers have recently reported that consecutive interevent intervals are correlated, such that a short (long) interval tends to be followed by another short (long) interval, which indicates that a memory effect extends over a single interval [15–22]. In other words, knowledge of the interevent interval between the ultimate and penultimate occurrences makes it possible to better assess the probability of the next occurrence time. This assessment consists of a conditional distribution of interevent intervals given the preceding interval.

In this paper we present a factor on which to base the prediction of the next earthquake occurrence that can be more relevant than the preceding interevent interval. For this discussion, we adopt the latest occurrence rate, evaluate its relevance to the next interevent interval, and optimize the duration of the observation time window over which this occurrence rate is defined. This optimization is performed by maximizing the mutual information function that relates the latest occurrence rate with the interevent interval between the most recent (i.e., past) occurrence and the next (i.e., future) occurrence.

II. SEISMIC CATALOG

We analyzed seismic data from four catalogs, Southern California Seismic Network (SCSN) [23], GeoNet [24], Japan Meteorological Agency (JMA) [25], and National Earthquake Information Center (NEIC) [26], which contain the time, spatial coordinates, and magnitude of earthquakes that occurred in southern California, New Zealand, Japan, and over the entire earth, respectively. From each catalog, we collected the occurrence times of earthquakes whose magnitude M was greater than or equal to a given threshold M_c [Fig. 1(a)]. In collecting these data, we ignored geographic or spatial information. The interevent interval preceding the most recent event is denoted by τ_{-1} [see Fig. 1(a)]. The new interevent interval from the most recent occurrence to the next occurrence is denoted simply as τ without a subscript. The latest occurrence rate λ, which may be defined over an arbitrary observation window w, is simply the number of earthquakes that occurred (within the given observation window, which terminates at the most recent event) divided by the duration w of the observation window.

III. SELECTING THE SIZE OF AN OBSERVATION TIME WINDOW

The relevance of the latest rate λ to the next interevent interval τ may be defined by the degree to which knowledge of the latest rate reduces uncertainty in the next interevent interval. Figure 1(b) shows how the raw distribution of interevent intervals $P(\tau)$ is condensed into the conditional distribution $P(\tau|\lambda)$ due to information regarding the latest rate λ. The reduced uncertainty can be gauged in terms of the mutual information that quantifies the statistical interdependence between two quantities [27], which is given by

$$I(\tau,\lambda) = \int d\tau \int d\lambda \, P(\tau,\lambda) \log_2 \left[\frac{P(\tau,\lambda)}{P(\tau)P(\lambda)} \right], \quad (1)$$

where $P(\tau,\lambda)$ is the joint-probability density function of the interevent interval τ and the rate λ, and $P(\tau)$ and $P(\lambda)$ are their respective marginal density functions. Given the catalog data, the rate λ is determined for a given observation window

----

*omitakahiro@gmail.com
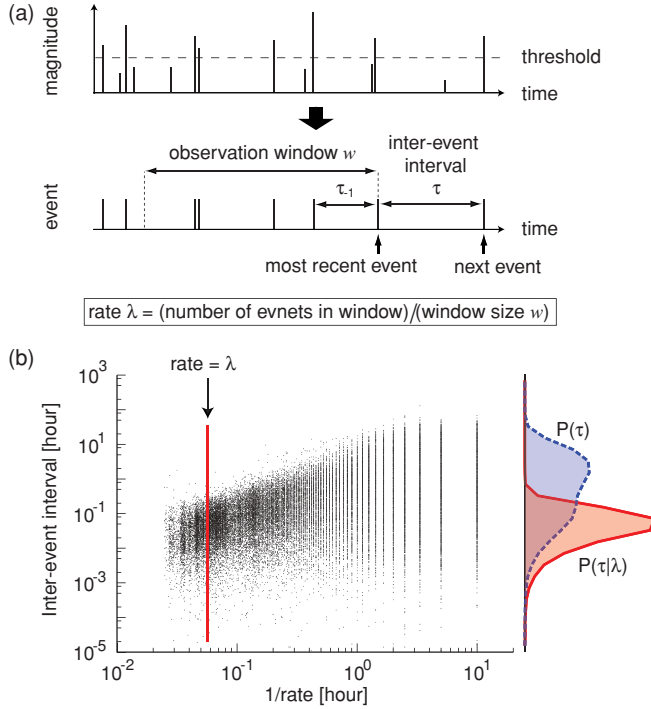†kanter@mail.biu.ac.il
‡shinomoto@scphys.kyoto-u.ac.jp

FIG. 1. (Color online) Manner in which information on the latest occurrence rate reduces the uncertainty in the interevent interval before the next earthquake. (a) Earthquakes with magnitudes greater than or equal to a given threshold ($M \geqslant M_c$) are treated here as unit events. Their occurrence rate λ is defined as the number of events in the observation window that terminates at the most recent event divided by the duration w of the window, and τ is the interevent interval between the most recent event and the next event. (b) A scatter plot of the interevent interval τ versus the latest rate λ. (Data were taken from the SCSN catalog for a threshold magnitude $M_c = 2.0$ and for an observation window $w = 10$ hours) The raw distribution of interevent intervals $P(\tau)$ is condensed into the conditional distribution $P(\tau|\lambda)$ due to information regarding the latest rate λ, as shown by the red line in (b).

w, so the mutual information $I(\tau, \lambda)$ depends on w. We select the duration of the observation window w so that the mutual information $I(\tau, \lambda)$ is maximized.

By applying the best conceivable remedy for minimizing potential bias in the numerical estimation [28–30], we estimate the mutual information for data from the four catalogs using a variety of threshold magnitudes M_c (see Appendix for details). We find that the mutual information for all the examined data peaks at an observation window w on the order of 10 hours (Fig. 2). Note that for each data catalog, approximately 10 earthquakes typically occur over a 10 hours period [Fig. 3(a)]. The correlation coefficient between $\log(\tau)$ and $\log(\lambda)$ also exhibits a similar tendency (Fig. S2 in [31]). Because the correlation coefficient depends on the coordinates, we adopt here the mutual information function that is invariant under arbitrary coordinate transformation from (τ, λ) to $(f(\tau), g(\lambda))$ for monotonous functions f and g. Although the mutual information is maximized (for all the data) for observation windows on the order of 10 hours, its value at the optimal window does depend on the geographic size of the area and on the threshold magnitude M_c; there is a tendency
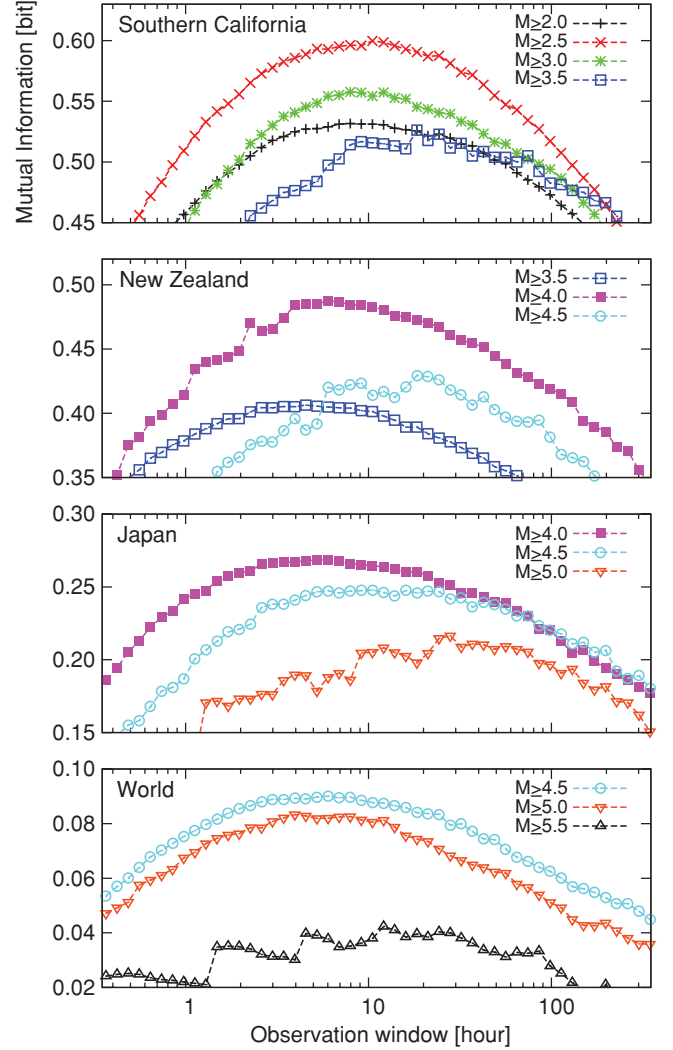

FIG. 2. (Color online) Dependence of the mutual information $I(\tau, \lambda)$ between the interevent interval τ before the next earthquake and the latest occurrence rate λ on the observation window w (over which the occurrence rate is defined). Various threshold magnitudes M_c are examined for four seismic catalogs: southern California, New Zealand, Japan, and the entire earth.

for the mutual information to decrease with the geographic size of the area (the graphs in Fig. 2 are arranged from top to bottom in ascending order of geographic size) and with the threshold magnitude.

This optimum observation window to predict the occurrence time of earthquakes would indicate a characteristic time scale of the latent process governing seismicity. In this respect, 10 hours may appear strangely short in comparison to the time scales of months or years that have heretofore been considered characteristic of seismic events [1,2]. To check if the time scale on the order of 10 hours is due simply to aftershocks that occur shortly after mega-earthquakes, we apply the analysis to active and inactive periods segmented according to whether they belong to the aftershock periods of mega-earthquakes or not [Fig. 4(a)] [32]. It is observed that the mutual information peaks at approximately 10 hours for both active and inactive periods and for various main shock threshold magnitudes [Figs. 4(b) and 4(c)].
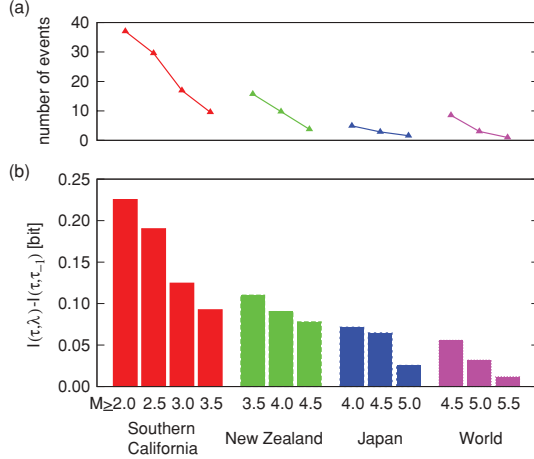
FIG. 3. (Color online) Comparison of the latest rate λ with the preceding interevent interval τ_{-1}, with respect to the relevance to the next interevent interval τ. (a) The average number of events contained in the 10 hours observation window. (b) The difference between $I(\tau,\lambda)$ and $I(\tau,\tau_{-1})$ for various seismic data examined in Fig. 2. That the values are positive indicates that information based on the latest rate λ is superior to information based on the preceding interevent interval τ_{-1} for reducing the uncertainty in the occurrence of the next earthquake.
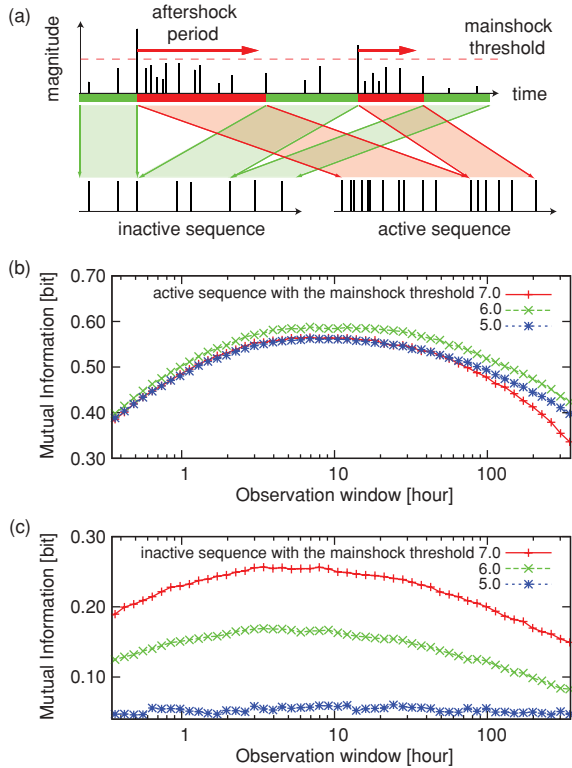

FIG. 4. (Color online) (a) A method for segmenting an entire sequence to construct active and inactive sequences according to whether the event intervals belong to the aftershock periods of mega-earthquakes or not. (b) and (c) The mutual information computed for the active and inactive sequences, respectively. Various aftershock periods defined for a range of mega-earthquakes are tested [32]. Data were taken from the SCSN catalog with $M_c = 2.0$.

Next, we compare information on the interevent interval gained from the latest occurrence rate with the preceding interevent interval. Figure 3(b) shows that gaps exist between the mutual information $I(\tau,\lambda)$ obtained from the rate determined from the optimal observation window and the mutual information $I(\tau,\tau_{-1})$ obtained for the consecutive interevent intervals. This demonstrates that in assessing the occurrence time of the next earthquake knowledge of the latest occurrence rate λ is more relevant than knowledge of the preceding interevent interval τ_{-1}. The difference is more pronounced for cases with a small threshold magnitude in which more
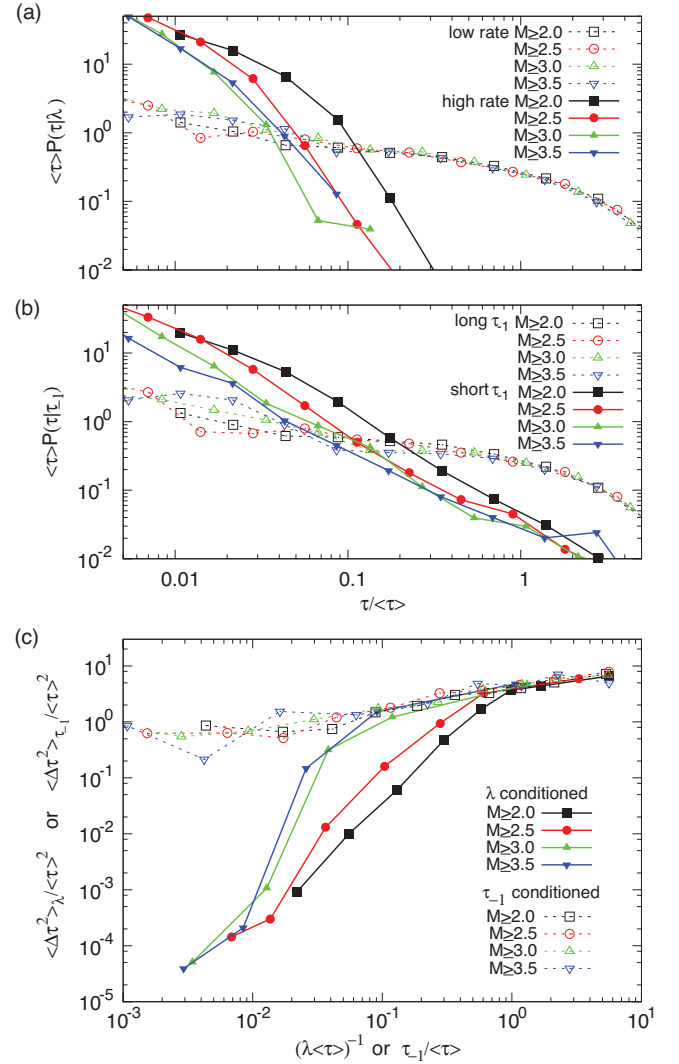

FIG. 5. (Color online) (a) Conditional distributions of interevent intervals given low and high latest rates, $P(\tau\,|\,\text{the lowest } 1/8\ \lambda)$ and $P(\tau\,|\,\text{the highest } 1/8\ \lambda)$. (The optimal observation window sizes are 8.0, 10.5, 8.0, and 18.4 h for $M_c = 2.0$, 2.5, 3.0, 3.5, respectively.) (b) Those given short and long preceding interevent intervals, $P(\tau\,|\,\text{the lowest } 1/8\ \tau_{-1})$ and $P(\tau\,|\,\text{the highest } 1/8\ \tau_{-1})$. In (a) and (b), an open (solid) symbol represents the distribution for a low (high) value for the latest occurrence rate λ (a) or the long (short) preceding interevent interval τ_{-1} (b). (c) The variances of the interevent intervals given the latest rate, $\langle\Delta\tau^2\rangle_\lambda$ (solid line), and those given the previous interevent interval, $\langle\Delta\tau^2\rangle_{\tau_{-1}}$ (dashed line). Data were taken from the SCSN catalog.

earthquakes are contained in the time window on the order of 10 hours.

IV. CONDITIONAL INTEREVENT INTERVAL DISTRIBUTION

In addition to comparing the amounts of mutual information, we show how the probability distribution of the interevent interval to the next event is narrowed by knowledge of an area's seismic history. To this end, we measure the conditional distribution of the interevent interval τ given the latest occurrence rate λ, which is given by $P(\tau|\lambda) = P(\tau,\lambda)/P(\lambda)$, or the conditional distribution of τ given the preceding interevent interval τ_{-1}, which is given by $P(\tau,\tau_{-1}) = P(\tau,\tau_{-1})/P(\tau_{-1})$. We impose on the conditional distribution $P(\tau|\lambda)$ two extreme conditions, namely, the distributions of interevent intervals τ given the latest rate λ belonging to the lowest and the highest $1/8$ among all samples. Figure 5(a) shows the conditional distributions $P(\tau|$ the lowest $1/8$ $\lambda)$ and $P(\tau|$ the highest $1/8$ $\lambda)$ rescaled by the mean interevent interval $\langle\tau\rangle$ averaged over a single long sequence of events. It is apparent that the interevent interval strongly depends on the latest rate, with a high (low) rate being followed by a short (long) interevent interval. This result is in contrast to the other conditional distribution $P(\tau|\tau_{-1})$ with analogous conditions applied (i.e., the preceding interevent interval τ_{-1} belongs to the longest or the shortest $1/8$ from all samples). Figure 5(b) shows the conditional distributions $P(\tau|$ the longest $1/8$ $\tau_{-1})$ and $P(\tau|$ the highest $1/8$ $\tau_{-1})$ rescaled by the mean interevent interval $\langle\tau\rangle$. Consecutive interevent intervals are correlated, but the difference between the two conditional distributions is not as distinct as that for the rate-conditioned distributions. Figure 5(c) compares the variances of the interevent intervals estimated given the latest rate, $\langle\Delta\tau^2\rangle_\lambda$, and those estimated given the previous interevent interval, $\langle\Delta\tau^2\rangle_{\tau_{-1}}$. The variance of the interevent intervals given $(\lambda\langle\tau\rangle)^{-1}$ is smaller than the one given the corresponding preceding interevent interval $\tau_{-1}/\langle\tau\rangle$, indicating that the latest rate λ can better predict the future interevent interval τ than the previous interevent interval τ_{-1}.

V. CONCLUSION

To conclude, we find that the occurrence rate of earthquakes based on a time window on the order of 10 hours preceding an earthquake occurrence is a more efficient factor for reducing the uncertainty in the interevent interval before the next earthquake irrespective of the spatial and magnitude range. The size of the time window optimized for forecasting the next event does not largely depend on the range of earthquake magnitudes, implying that common mechanisms are underlying the generation of different scales of earthquakes. The optimal time scale does not vary between seismic catalogs of different areas, including the worldwide catalog. It would be interesting to observe whether distant events contributed to the peak of the mutual information, that is, to estimate the spatial scale for correlated events. It should be noted that the characteristic time scale on the order of 10 hours is much shorter than the months or years that have previously been considered to determine seismic activities. We believe that this finding provides cause to revise the framework for analyzing seismic data and may lead to more accurate predictions of earthquakes. Furthermore,

the method presented here may be applicable to the analysis of a wide range of bursty phenomena, including natural hazards, the economy, or human communication [33–37].

APPENDIX: NUMERICAL ESTIMATION OF MUTUAL INFORMATION

The mutual information between two variables x and y is defined as

$$I(x,y) = \int dx \int dy\, P(x,y) \log_2\left[\frac{P(x,y)}{P(x),P(y)}\right], \quad \text{(A1)}$$

where $P(x,y)$ is the joint-probability density function of x and y and $P(x)$ and $P(y)$ are the marginal density functions of x and y, respectively [27]. The mutual information is a general measure that quantifies the statistical dependence and takes a value of 0 iff x and y are statistically independent; that is, $P(x,y) = P(x)P(y)$. Otherwise, $I(x,y) > 0$.

A method to estimate the mutual information from m samples $\{x_i,y_i\}$ $(i = 1,\ldots,m)$ is summarized as follows. First, we calculate Eq. (A1) by estimating the probability density functions via a histogram method. Although it is common to divide x and y into equal-sized bins to construct a histogram, such uniform binning results in a poor estimate of the mutual information. Therefore, we employ herein adaptive binning [28,29]. We divide $x[y]$ into $N_x[N_y]$ bins so that the number of the points in each bin is identical. If $D_x(i)[D_y(j)]$ denotes the ratio of the points belonging to the ith$[j$th$]$ bin of $x[y]$, then $D_x(i) = 1/N_x[D_y(j) = 1/N_y]$. Throughout this paper, we use $N_x = N_y = [\sqrt{m/10}]$. However, if the variable takes on discrete values, it is often not possible to have the same number of points in each bin. For such cases, we divide x or y so that the number of points in each bin is greater than or equal to $[\sqrt{10m}]$. Let $D_{xy}(i,j)$ denote the ratio of the points in the intersection of the ith bin of x with the jth bin of y. For $x[y]$ belonging to ith$[j$th$]$ bin of $x[y]$, by setting $P(x) = D_x(i)/\delta_x(i)[P(y) = D_y(j)/\delta_y(j)]$ and $P(x,y) = D_{xy}(i,j)\delta_x(i)^{-1}\delta_y(j)^{-1}$, where $\delta_x(i)[\delta_y(j)]$ is the size of the ith$[j$th$]$ bin of $x[y]$, Eq. (A1) becomes

$$I_{\text{hist}}(x,y) = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} D_{xy}(i,j)\log_2\left[\frac{D_{xy}(i,j)}{D_x(i)D_y(j)}\right]. \quad \text{(A2)}$$

Using Eq. (A2), we estimate the mutual information $I(\tau,\lambda)$ for SCSN, the catalog with $N_x = N_y = \sqrt{m/\alpha}$ ($\alpha = 5,8,10,15,20$). Clearly, the estimated information strongly depends on the choice of α (see left panels in Fig. S1 in [31]), which is because the probability density functions are approximated by a histogram and because the probability is

estimated from a finite-sized data set. To reduce the bias in Eq. (A2), we use

$$I_{\text{correct}}(x,y) = I_{\text{hist}}(x,y) + \frac{B_x + B_y - B_{xy} - 1}{2m \ln(2)}, \quad \text{(A3)}$$

where $B_x[B_y]$ is the number of bins where $D_x(i) \neq 0[D_y(j) \neq 0]$ and B_{xy} is the number of bins where $D_{xy}(i,j) \neq 0$ [30]. This correction can largely eliminate the dependence on α (see right panels in Fig. S1 in [31]).Throughout this study, we estimate the mutual information by using Eq. (A3) with $\alpha = 10$.

[1] F. Omori, J. Coll. Sci. Imp. Univ. Tokyo **7**, 111 (1894).
[2] T. Utsu, Y. Ogata, and R. S. Matsu'ura, J. Phys. Earth **43**, 1 (1995).
[3] Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988).
[4] P. A. Reasenberg and L. M. Jones, Science **243**, 1173 (1989).
[5] M. Baiesi and M. Paczuski, Phys. Rev. E **69**, 066106 (2004).
[6] D. Marsan and O. Lengline, Science **319**, 1076 (2008).
[7] M. Bottiglieri *et al.*, J. Geophys. Res. **114**, B03303 (2009).
[8] J. K. Gardner and L. Knopoff, Bull. Seismol. Soc. Am. **64**, 1363 (1974).
[9] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).
[10] A. Corral, Phys. Rev. Lett. **92**, 108501 (2004).
[11] J. Davidsen, S. Stanchits, and G. Dresen, Phys. Rev. Lett. **98**, 125502 (2007).
[12] A. Saichev and D. Sornette, Phys. Rev. Lett. **97**, 078501 (2006).
[13] S. Touati, M. Naylor, and I. G. Main, Phys. Rev. Lett. **102**, 168501 (2009).
[14] M. Bottiglieri, L. de Arcangelis, C. Godano, and E. Lippiello, Phys. Rev. Lett. **104**, 158501 (2010).
[15] A. Helmstetter and D. Sornette, Phys. Rev. Lett. **92**, 129801 (2004).
[16] V. Livina, S. Tuzov, S. Havlin, and A. Bunde, Phys. A **348**, 591 (2005).
[17] V. N. Livina, S. Havlin, and A. Bunde, Phys. Rev. Lett. **95**, 208501 (2005).
[18] A. Corral, Tectonophysics **424**, 177 (2006).
[19] S. Lennartz, V. N. Livina, A. Bunde, and S. Havlin, Europhys. Lett. **81**, 69001 (2008).
[20] T. Akimoto, T. Hasumi, and Y. Aizawa, Phys. Rev. E **81**, 031133 (2010).
[21] X. Zhao, T. Omi, N. Matsuno, and S. Shinomoto, New J. Phys. **12**, 063010 (2010).

[22] J. F. Eichner, J. W. Kantelhardt, A. Bunde, and S. Havlin, Phys. Rev. E **75**, 011128 (2007).
[23] Southern California Seismic Network [http://www.data.scec.org/ftp/catalogs/SCSN/], 1990–2008, 32°–37°N, 114°–122°W.
[24] GeoNet Project [http://magma.geonet.org.nz/resources/quakesearch/], 1990–2008, 34°–49°S, 164°–182°E.
[25] Japan Meterogical Agency [ftp://ftp.eri.u-tokyo.ac.jp/pub/data/jma/mirror/JMA_HYP/], 1990–2008, 25°–50°N, 125°–150°E.
[26] National Earthquake Information Center [http://neic.usgs.gov/neis/epic/epic_global.html], 1990–2008.
[27] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[28] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[29] C. J. Cellucci, A. M. Albano, and P. E. Rapp, Phys. Rev. E **71**, 066208 (2005).
[30] M. S. Roulston, Phys. D **125**, 285 (1999).
[31] See supplemental material at [http://link.aps.org/supplemental/10.1103/PhysRevE.83.026101] for two supplementary figures.
[32] Aftershock period is defined as the 2-yr interval after earthquakes with $M \geqslant 7.0$, the 1-yr interval after earthquakes with $7.0 > M \geqslant 6.0$, and the half-year interval after earthquakes with $6.0 > M \geqslant 5.0$.
[33] A.-L. Barabási, *Bursts: The Hidden Pattern Behind Everything We Do* (Dutton, New York, 2010).
[34] R. D. Malmgren, D. B. Stouffer, A. S. L. O. Campanharo, and L. A. N. Amaral, Science **325**, 1696 (2009).
[35] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).
[36] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **102**, 9424 (2005).
[37] K. I. Goh and A.-L. Barabási, Europhys. Lett. **81**, 48002 (2008).