Identifying the occurrence time of an impending major earthquake by means of the fluctuations of the entropy change under time reversal

E. S. Skordas\textsuperscript{1,2}, N. V. Sarlis\textsuperscript{1,2} and P. A. Varotsos\textsuperscript{1,2}

\textsuperscript{1} Section of Solid State Physics, Physics Department, National and Kapodistrian University of Athens Panepistimiopolis, Zografos 157 84, Athens, Greece, EU
\textsuperscript{2} Solid Earth Physics Institute, Physics Department, National and Kapodistrian University of Athens Panepistimiopolis, Zografos 157 84, Athens, Greece, EU

received 17 September 2019; accepted in final form 26 November 2019; published online 20 January 2020

PACS 91.30.Ab – Theory and modeling, computational seismology
PACS 89.75.Da – Systems obeying scaling laws
PACS 95.75.Wx – Time series analysis, time variability

Abstract – In natural time analysis, the complexity measure that quantifies the fluctuations of the entropy change under time reversal plays a key role in identifying when a system approaches a critical point (dynamic phase transition). Here, a new procedure is presented through which this complexity measure can be used for the identification of the occurrence time of a major earthquake. As an example, we apply this procedure to the case of the Tohoku mega-earthquake that occurred on 11 March 2011 in Japan with magnitude 9.0, which is the largest magnitude event recorded in Japan.

Introduction. – In the 1980s, a short-term earthquake prediction method was introduced based on the observation of Seismic Electric Signals (SES), which are low-frequency transient changes of the electric field of the Earth preceding earthquakes (EQs) \cite{1,2}. Several SES recorded within a short time are termed SES activity \cite{3}. Major EQs are preceded by intense SES activities accompanied by evident Earth’s magnetic field variations \cite{4} mainly recorded on the z-component \cite{5,6}. Once a SES activity has been recorded before a major EQ (with a lead time from a few weeks to around 5 to 6 months \cite{7}), the candidate epicentral area can be estimated on the basis of the ratio of the two SES components and the so-called selectivity map of the station at which the SES was recorded, \textit{e.g.}, refs. \cite{3,8}. This method was motivated by a physical model for SES generation \cite{1,8–10} which suggests the following: In the Earth’s crust, electric dipoles always exist \cite{9} due to lattice imperfections (point and linear defects, \textit{e.g.}, ref. \cite{11}) in the ionic constituents of rocks. In the future focal region of an EQ, where the electric dipoles have initially random orientations, the stress starts to gradually increase due to an excess stress disturbance. When this gradually increasing stress reaches a critical value, the electric dipoles exhibit a cooperative orientation resulting in the emission of a transient SES. Such a cooperativity is one of the main characteristics of critical phenomena \cite{12}.

Earthquakes exhibit complex correlations in time, space and magnitude (\textit{M}) \textit{(e.g.,} \cite{13–19}). It is widely accepted \cite{7,20,21} that the observed EQ scaling laws \cite{22} indicate the existence of phenomena closely associated with the proximity of the system to a critical point. A new procedure for the analysis of complex time series, termed natural time analysis was introduced in the beginning of the 2000s (\textit{e.g.,} see ref. \cite{23} and references therein) by means of which we can determine when the system approaches the critical point. This analysis also uncovers unique dynamic features hidden behind the time series of complex systems and has found applications in diverse fields compiled in ref. \cite{7}. Natural time is currently considered as the basis for a new methodology to estimate the seismic risk by Turcotte and coworkers \cite{19,24–26} termed (seismic) nowcasting.

In a time series comprising \textit{N} EQs, the natural time $\chi_k = k/N$ serves as an index for the occurrence of the \textit{k}-th EQ. This index together with the energy $Q_k$ released during the \textit{k}-th EQ of magnitude $M_k$, \textit{i.e.}, the pair ($\chi_k, Q_k$), is studied in natural time analysis. Alternatively, one studies...
the pair \((\chi_k, p_k)\), where

\[
p_k = \frac{Q_k}{\sum_{n=1}^{N} Q_n}
\]

stands for the normalized energy released during the \(k\)-th EQ. The variance of \(\chi\) weighted for \(p_k\), labeled \(\kappa_1\), is given by \([7,23,27,28]\)

\[
\kappa_1 = \sum_{k=1}^{N} p_k (\chi_k)^2 - \left( \sum_{k=1}^{N} p_k \chi_k \right)^2
\]

where \(Q_k\), and hence \(p_k\), for EQs is estimated through the usual relation [29]: \(\log_{10}(E) = 1.5M_w + 4.8\) for the seismic energy \(E\) in joules as a function of the moment magnitude \(M_w\) leading to

\[
Q_k \propto 10^{1.5M_w}.
\]

Here, as in refs. [30,31], we used the EQ catalog of the Japan Meteorological Agency (JMA) for \(M_{JMA} \geq M_{bres}\) with magnitude threshold \(M_{bres} = 3.5\) and in order to obtain \(Q_k\) we coverted the reported magnitude \(M_{JMA}\) to \(M_e\) according to the formulae suggested in ref. [32]. Thereafter, for reasons of brevity when we use the symbol \(M\) we refer to \(M_{JMA}\).

The quantity \(\kappa_1\) can be considered [27] as an order parameter for seismicity (a mainshock is the new phase). At least 6 EQs are needed for obtaining reliable \(\kappa_1\) [27]. Upon considering a sliding natural time window comprising \(i\) consecutive events sliding through the EQ catalog, event by event, the computed \(\kappa_1\) values enable the calculation of their average value \(\mu(\kappa_1)\) and their standard deviation \(\sigma(\kappa_1)\). We then determine the quantity \(\beta_i\) [33],

\[
\beta_i = \frac{\sigma(\kappa_1)}{\mu(\kappa_1)}.
\]

termed variability of \(\kappa_1\) that corresponds to this natural time window of length \(i\). To compute the time evolution of \(\beta_i\) we apply the procedure explained in refs. [7,28,30,31].

The fluctuations of the order parameter \(\kappa_1\) of seismicity were found [34] to exhibit a minimum \(\beta_{min}\) when a SES activity [3,8] initiates. The existence of \(\beta_{min}\) has been subsequently confirmed in ref. [30] for all shallow mainshocks of magnitude 7.6 or larger that occurred in Japan during 1984–2011. The minimum \(\beta_{min}\) of the fluctuations of the order parameter of seismicity before the M9 Tohoku EQ (that occurred on 11 March 2011 in Japan) was observed [30] on 5 January 2011 being the deepest minimum during the period from 1 January 1984 until the M9 Tohoku EQ occurrence. This date almost coincides with the detection of anomalous magnetic field variations on the \(z\)-component during the period 4 to 14 January 2011 at two measuring sites (Esashi (ESA) and Mizusawa (MIZ), see fig. 1) lying at epicentral distances of around 130 km [35–37] pointing to the initiation of an SES activity. (This is in agreement with our experimental findings in Greece that SES activities, see, e.g., pp. 8–9 of ref. [8] —and the associated magnetic field variations—are clearly detectable at epicentral distances up to around 150–200 km for EQs of magnitude 6.0 or larger.) A spatiotemporal study of the minimum \(\beta_{min}\) revealed [31] an estimate of the epicentral area of the impending major EQ. In such a study, we work as follows: By dividing the entire Japanese region into small areas, we carry out the \(\beta\) calculation on them. In practice, the calculation was made at a multitude of circular small areas of radius 250 km whose center was sliding with steps of 0.1° in longitude and latitude. It was found that some small areas show \(\beta_{min}\) almost simultaneously with the large area, i.e., the entire Japanese region, and such small areas clustered within a few hundred km from the actual epicenter of the related mainshock. Such a study for the M9 Tohoku EQ led to the estimate of the candidate epicentral area depicted with the blue-green area in fig. 1, where we show the color contours of the quantity \(n_e(x_i, y_i)\) defined as follows [31]: When “local” \(\beta\) minima (i.e., those obtained in the small areas) appear simultaneously (±2 days) with the \(\beta_{min}\) of the large area, we investigate their spatial distribution by counting how many of their centers lie within 250 km from each point \((x_i, y_i)\) of a 0.05° × 0.05° grid. This number is the quantity \(n_e(x_i, y_i)\), the contours of which are shown according to the color scale of fig. 1.

By starting from 5 January 2011 and computing the \(\kappa_1\) values in the blue-green area delimited by the red line in fig. 1 we found that the critical condition \(\kappa_1 = 0.070\) [38] (which signals that the mainshock was going to occur within the next few days or so) was fulfilled, as shown...
by the gray shaded area in fig. 6(b) of ref. [39], for $M_{\text{thres}} = 4.2$ to 5.0 in the morning of 10 March 2011 upon the occurrence of the EQs from 08:36 to 13:14 LT, i.e., almost one day before the Tohoku EQ. This happened almost a day after the occurrence of the M7.3 EQ on 9 March 2011, thus identifying that this M7.3 EQ was a foreshock.

It is the scope of this paper to suggest a procedure to identify the approach of the critical point (mainshock) without making use of the condition $k_1 = 0.070$. In particular, we shall show that the approach of the critical point may be identified by studying the evolution in natural time of the complexity measure that quantifies the fluctuations of the entropy change $\Delta S$ of seismicity under time reversal to which we now turn.

**Entrophy in natural time and the complexity measure associated with the fluctuations of the entropy change under time reversal.** – The entropy $S$ in natural time is defined [7] as the derivative with respect to $q$ of the fluctuation function $\langle \chi^q \rangle - \langle \chi \rangle^q$ at $q = 1$, which results in

$$S \equiv \langle \chi \ln \chi \rangle - \langle \chi \rangle \ln \langle \chi \rangle,$$

(5)

where the brackets $\langle \ldots \rangle \equiv \sum \langle \ldots \rangle_p f$ denote averages with respect to the distribution $p_k, i.e., \langle f(\chi) \rangle \equiv \sum f(\chi_k)p_k$. It is a dynamic entropy exhibiting [40] concavity, positivity and Lesche stability [41,42]. Upon considering the time reversal $\tilde{T}$, i.e., $\tilde{T}_{pk} = p_{N-k+1}$, the value $S$ changes to a value $S_-$:

$$S_- = \sum_{k=1}^{N} p_{N-k+1} \frac{k}{N} \ln \left( \frac{k}{N} \right) - \left( \sum_{k=1}^{N} \frac{k}{N} p_{N-k+1} \right) \ln \left[ \sum_{l=1}^{N} \frac{l}{N} p_{N-l+1} \right],$$

(6)

$S_-$ is different from $S$ and hence there exists a change $\Delta S \equiv S - S_-$ in natural time under time reversal. Thus, $S$ does satisfy the condition to be time-reversal asymmetric [7,40,43]. Using a natural time window of length $i$ sliding, event by event, through the time series of $L$ consecutive events the entropy in natural time is determined for each position $j = 1, 2, \ldots, L-i$ of the sliding window. Thus, a time series of $S_i$ is constructed [43]. By employing eq. (6), we also construct the time series of $(S_-)_i$. By computing the standard deviation $\sigma(\Delta S_i)$ of the time series of $\Delta S_i \equiv S_i - (S_-)$, we define [7,44] the complexity measure $\Lambda_i$

$$\Lambda_i = \frac{\sigma(\Delta S_i)}{\sigma(\Delta S_{100})},$$

(7)

where the denominator has been selected to correspond to the standard deviation $\sigma(\Delta S_{100})$ of the time series of $\Delta S_i$ of $i = 100$ events as in ref. [45] (of course, the selection of a different scale $i$ would change the numerical values obtained but it would not change the whole behavior and physical picture of the results concerning the time evolution of $\Lambda_i$). In other words, the complexity measure $\Lambda_i$ quantifies how the statistics of $\Delta S_i$ time series varies upon changing the scale from 100 to another scale $i$.

$\Delta S$ is a measure that may be used for the identification [7] when the system approaches the critical point (dynamic phase transition). For example, $\Delta S$ has been applied to identify the time of an impending sudden cardiac death risk [43]. Furthermore, it has been used [46] for the study of the predictability of the Olami-Feder-Christensen model for EQs [47], which is [48] the most studied non-conservative self-organized criticality model. In particular, it was found that $\Delta S$ exhibits a clear minimum [7] (or maximum if we define, e.g., see ref. [46] $\Delta S \equiv S - S_-$ instead of $\Delta S \equiv S - S_-$) before a large EQ. For example, by analyzing in natural time the seismicity during 2012–2017 in the Chiapas region of Mexico in which the M8.2 EQ occurred on 7 September 2017, we found [49] that $\Delta S$ of seismicity was minimized on 14 June 2017. Furthermore, almost three months before the M9 Tohoku EQ, i.e., on 22 December 2010, the following facts have been observed: First, the complexity measure $\Lambda_i$ exhibited abrupt increase which conforms to the seminal work by Lifshitz and Slyozov [50] and independently by Wagner [51] for phase transitions showing that the characteristic size of the minority phase droplets exhibits a scaling behavior in which time growth has the form $A(t - t_0)^{1/3}$ [52]. Second, a statistically significant minimum $\Delta S_{\text{min}}$ of $\Delta S$ of seismicity in the entire Japanese region under time reversal was found in ref. [53]. In particular, the probability to observe by chance such a deep (or even deeper) minimum was estimated [53] to be close to 3%, while the fact that it can be considered as a precursor to the M9 Tohoku EQ had a much smaller probability ($<1\%$) to occur by chance as shown in ref. [54] when employing the recently introduced method of Event Coincidence Analysis (ECA), see, e.g., refs. [55,56]. Third, by investigating the fluctuations $\beta$ of $\kappa_i$ of seismicity in the entire Japanese region $N_{\text{Japan}}$, vs. the conventional time from 1 January 1984 until the Tohoku EQ occurrence on 11 March 2011, we identified [54] a large fluctuation of $\beta$ upon the occurrence of the M7.8 EQ on 22 December 2010 accompanying the minimum $\Delta S_{\text{min}}$ which is unique [53].

**Results and discussion.** – Setting a magnitude threshold $M_{\text{thres}} = 3.5$ to assure data completeness [30], there exist 47204 EQs in the entire Japanese region, i.e., during the period from 1 January 1984 until the M9 Tohoku EQ occurrence, which is about 326 months. Thus, we have on the average $\sim 145$ EQs per month.

To study the time evolution of $\Delta S_i$ of seismicity with $M \geq 3.5$ during this almost 27-year period, we select proper scales $i$ as follows: We first recall that the minimum $\beta_{\text{min}}$ of the fluctuations of the order parameter of seismicity is observed simultaneously with the initiation of an SES activity [3,8,57]. Since an SES activity exhibits critical behavior [23,58,59], it is observed during a period in which long-range correlations prevail between EQs.
magnitudes [60]. On the other hand, before the initiation of the SES activity, and hence before \( \beta_{\min} \), another stage appears in which the temporal correlations between EQ magnitudes exhibit an anticorrelated behavior [60]. Thus, the temporal correlations between EQ magnitudes exhibit a significant change between the two stages that correspond to the periods before and after the initiation of the SES activity. This change is likely to be reflected on the time evolution of \( \Delta S_i \); thus our study of \( \Delta S_i \) is focused on scales that capture the change between these stages, i.e., of the order of \( i \sim 10^3 \) events, which correspond to the number of seismic events \( M \geq 3.5 \) that occur during a period of at least around the maximum lead time of SES activities (as mentioned above we have \( \sim 145 \) EQs per month in the entire Japanese region and a period of at least 5.5 months).

In fig. 2, we plot the \( \Lambda_i \) values computed for all \( M \geq 3.5 \) EQs in the entire Japanese region \( N_{35}^{42}E_{125}^{148} \) vs. the scale \( i \) (number of events) by starting the calculation almost 5 years (a), 3 years (b), 2 years (c), and around 14 months (d) before the M9 Tohoku EQ occurrence. For each scale, the \( \Lambda_i \) values are depicted at the following dates: 30 November 2010 (just before the M7.1 EQ on this date, pluses in red), 1 December 2010 (crosses in green), 22 December 2010 (asterisks in blue), 1 January 2011 (open squares in magenta), 1 February 2011 (solid circles in cyan), 1 March 2011 (open circles in brown), 8 March 2011 (solid circles in black), 9 March 2011 (open triangles in orange), 10 March 2011 (gray filled triangles) and 11 March (inverted red triangles), almost 10 min before the M9 Tohoku EQ occurrence. The time format in the figure keys is YYYYMMDDHHMMSS in Japan Standard Time.

Leaving aside the details, an inspection of figs. 2(a)–(d), reveals the following general feature: For each of the scales that are markedly longer than 2000 events, e.g., \( i = 3000, 4000 \) and \( 5000 \) events, the dates show a tendency to be clearly clustered into two groups: The one group that comprises markedly larger \( \Lambda_i \) values corresponding to dates
later than the date 22 December 2010 at which $\Delta S_{\text{min}}$ has been observed, thus being closer to the occurrence date of the Tohoku EQ. The other group that comprises appreciably lower $\Lambda_i$ values corresponding to earlier dates. Practically the same behavior is observed in fig. 3 upon increasing the magnitude threshold to 4.0, i.e., by considering only the $M \geq 4.0$ EQs in our computations, and using scales that are smaller by a factor of 2.5 in view of the smaller number of EQs per month we have for this threshold (cf. fig. S6 of ref. [53]). Furthermore, we note that a similar behavior is observed before a smaller EQ, for example, before the M8.2 EQ in Mexico that occurred on 7 September 2017 in the Chiapas region. In this case, after considering all EQs in this region since 2012, the $\Delta S_{\text{min}}$ has been observed [49] on 14 June 2017 upon the occurrence of a M7 EQ, and afterwards the $\Lambda_i$ values started [45] to increase, see fig. 4, which depicts the results for $M_{\text{thres}} = 3.5$ and 4.0.

The situation concerning the identification of the occurrence time of the Tohoku EQ becomes more clear if we repeat the computation that led to fig. 2, but by considering the EQs ($M \geq 3.5$) that occur in the candidate epicentral area (instead of considering the entire Japanese region that was used in the derivation of the results of fig. 2) depicted with the blue-green area delimited by the red line in fig. 1, as mentioned in the Introduction. These results obtained after considering the EQs in the candidate area alone, are depicted in fig. 5 (since the EQ rate in this area is 3.7 times smaller than that of the whole Japanese

Fig. 3: The same as fig. 2, but here we plot the $\Lambda_i$ values vs. the scale $i$ by considering only the $M \geq 4.0$ EQs.

Fig. 4: The values of $\Lambda_i$ vs. the scale $i$ (number of events) for all $M \geq 3.5$ EQs as well as for all $M \geq 4.0$ EQs in the Chiapas region, Mexico, since 1 January 2012. These $\Lambda_i$ values are calculated at the following dates: 1 June 2017 (yellow solid circles), 14 June 2017 (cyan squares), 1 July 2017 (magenta plus), 1 August 2017 (blue star), 1 September 2017 (green cross) and 7 September 2017 (red circle, until the last event before the M8.2 earthquake on 7 September 2017).
Fig. 5: The same as fig. 2, but here we plot the $\Lambda_i$ values vs. the scale $i$ by considering only the $M \geq 3.5$ EQs that occur inside the estimated epicentral area shown in fig. 1 (instead of taking into account the seismicity in the entire Japanese region $N_{\text{area}}$ that we did in fig. 2).

area all scales $i$ have been divided by this factor). Their inspection reveals that each of the scales $i$ longer than about $i = 400–500$ events and smaller than around 2000 events, the dates 10 March and 11 March led to $\Lambda_i$ values that are larger than those corresponding to earlier dates. In other words, for each of the scales longer than around 400–500 events and smaller than around 2000 events, we find markedly larger $\Lambda_i$ values when approaching the date of the EQ occurrence.

**Summary and conclusions.** – Analyzing the seismic data of Japan ($M \geq 3.5$) in natural time and calculating the complexity measure $\Lambda_i$ that quantifies the fluctuations of the entropy change $\Delta S$ under time reversal, the following results have been found:

First, if the computation is made for all EQs ($M \geq 3.5$) occurring in the entire Japanese region, for each of the longer scales $i$ (e.g., $i = 3000, 4000$ and 5000 events) the resulting $\Lambda_i$ values are distinctly larger after the date at which $\Delta S$ exhibited a minimum on 22 December 2010 almost three months before the $M_9$ Tohoku EQ. Strikingly, this increase is very clear after the occurrence of the $M_7.3$ foreshock on 9 March 2011.

**REFERENCES**

Identifying the occurrence time by the entropy change under time reversal