Order parameter fluctuations of seismicity in natural time before and after mainshocks

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2010 EPL 91 59001
(http://iopscience.iop.org/0295-5075/91/5/59001)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 195.134.94.87
The article was downloaded on 02/12/2010 at 07:49

Please note that terms and conditions apply.
Order parameter fluctuations of seismicity in natural time before and after mainshocks

N. V. Sarlis, E. S. Skordas and P. A. Varotsos\(^{(a)}\)

Solid State Section and Solid Earth Physics Institute, Physics Department, University of Athens Panepistimiopolis, Zografos 157 84, Athens, Greece, EU

received 27 April 2010; accepted in final form 16 August 2010
published online 21 September 2010

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 – It is widely accepted that the observed earthquake scaling laws indicate the proximity of the system to a critical point. Using the order parameter (OP) for seismicity suggested on the basis of natural time as well as the detrended fluctuation analysis of the magnitude time-series, we investigate the behavior of seismicity before and after significant earthquakes. The analysis reveals that the fluctuations of the OP before major earthquakes exhibit a characteristic feature which, if quantified properly, may be used as decision variable to predict the occurrence of a large earthquake in the next time step based solely on the magnitudes of previous earthquakes.

Copyright © EPLA, 2010

Earthquakes (EQs) do exhibit complex correlations in time, space and magnitude (m), the study of which has been the object of several recent works, e.g. [1–18]. Additionally, it has been repeatedly proposed that the observed EQ scaling laws indicate the existence of phenomena associated with proximity of the system to a critical point (e.g., [10,19], see also ref. [20] and references therein). In this frame, it has been suggested [21] that an order parameter (OP) for seismicity can be defined (see below) on the basis of natural time [22–24] (see fig. 1). It was through the probability density function (pdf) of this parameter, that it was possible to demonstrate [25] that the seismicities in various regions (as well as the worldwide seismicity) give rise to pdf curves that collapse on the same master (universal) curve. Moreover, the study of the OP fluctuations, relative to the standard deviation of its distribution, revealed [21] that the scaled distributions collapse on the same curve, which interestingly exhibits, over four orders of magnitude, features similar to those [26–31] in several equilibrium critical phenomena as well as in nonequilibrium systems. Such a behavior is strikingly reminiscent of the one found earlier in the analysis of non-stationary biological signals including heart rate [32], locomotor activity [33] etc, where pdf curves obtained for different scales of observation fall onto a single master curve.

\(^{(a)}\)E-mail: pvaro@otenet.gr

Fig. 1: Schematic diagram showing how a series of earthquakes is read in conventional time (upper panel) and in natural time (lower panel).

In a recent study [14], it has been undoubtedly shown that, in the regimes of stationary seismic activity, long-range correlations exist between successive EQ magnitudes and interoccurrence times. Moreover, a separate study [16] showed that the fluctuations of seismic activity, defined as the detrended cumulated sum of the magnitude time-series, exhibit Family-Vicsek dynamic scaling. In both studies, the sequence index \(k\), i.e., the sequential order in which an EQ had occurred, has been used for the detection of the long-range correlations (e.g., see fig. 3 that will be discussed later). Notice that it is the combination of this index with the released seismic energy \(E_k\) during the \(k\)-th EQ (\(E_k\) is interrelated [34] with the magnitude through \(E_k \sim 10^c m_k\) —where \(c\) is around 1.5) that constitutes the two quantities \(k, E_k\) which are in fact used in the analysis in natural time [21,22,35].
In a time-series comprising of \( N \) earthquakes (see fig. 1) the natural time \( \chi_k = k/N \) serves as an index \([22,35]\) for the occurrence of the \( k \)-th EQ. The evolution of the pair \((\chi_k, E_k)\) is studied \([21,22,35–39]\) by means of the normalized power spectrum given by

\[
\Pi(\omega) = \left| \sum_{k=1}^{N} p_k \exp \left( i \omega \frac{k}{N} \right) \right|^2 , \tag{1}
\]

where \( p_k \) stands for \( p_k = E_k / \sum_{n=1}^{N} E_n \), \( \omega = 2\pi \phi \) and \( \phi \) denotes the natural frequency. In natural time analysis the properties of \( \Pi(\omega) \) or \( \Pi(\phi) \) are studied \([21,22,35,36]\) for natural frequencies \( \phi \) less than 0.5. This is so, because in this range of \( \phi \), \( \Pi(\omega) \) or \( \Pi(\phi) \) reduces to a characteristic function for the probability distribution \( p_k \) in the context of probability theory. According to the probability theory, the moments of a distribution and hence the distribution itself can be approximately determined once the behavior of the characteristic function of the distribution is known around zero. For \( \omega \to 0 \), eq. (1) leads to \([21–23]\)

\[
\Pi(\omega) \approx 1 - \kappa_1 \omega^2 , \tag{2}
\]

where \( \kappa_1 \) is the variance of \( \chi \) given by

\[
\kappa_1 = \frac{N}{N-1} \left( \sum_{k=1}^{N} \chi_k^2 - \left( \sum_{k=1}^{N} \chi_k \right)^2 \right) . \tag{3}
\]

The quantity \( \Pi(\omega) \) for \( \omega \to 0 \) (or \( \kappa_1 \)) can be considered \([21]\) as an order parameter for seismicity since its value changes abruptly when a main shock occurs. In a seismic catalogue comprising \( W \) events, the following procedure was followed: starting from the first EQ, we calculate the \( \kappa_1 \)-values using \( N = 6 \) to 40 consecutive events (including the first one). We next turn to the second EQ, and repeat the calculation of \( \kappa_1 \). After sliding, event by event, through the whole earthquake catalogue, the calculated \( \kappa_1 \) values enable the construction of the pdf \( P(\kappa_1) \). For example, upon using the Southern California Earthquake catalogue, the pdf \( P(\kappa_1) \) depicted with black plus symbols in fig. 2 is obtained from the natural time analysis of all the \( W = 85862 \) EQs with \( m \geq 2 \) that occurred during the period 1981–2003 within the area \( N_{SCEC} = 114 \) (hereafter called SCEC). In ref. \([21]\), the statistical properties of \( \kappa_1 \) (i.e., the OP of seismicity in natural time) have been studied by means of the scaled \([26]\) distribution: \( \sigma(\kappa_1) P[(\kappa_1 - \mu(\kappa_1))/\sigma(\kappa_1)] \), where \( \mu(\kappa_1) \) and \( \sigma(\kappa_1) \) stand for the average value and the standard deviation of the \( \kappa_1 \) values calculated. This is the scaled distribution for seismicity that exhibits, as mentioned above, features similar to those obtained by studying \([26–31]\) the OPs of several equilibrium critical phenomena and in nonequilibrium systems.

We now turn to the problem of the detection of temporal correlations in the EQ magnitude time-series \( m_k \). The existence of temporal correlations in seismicity has been already treated on the basis of natural time analysis in refs. \([38,39]\), but here we focus on the recent study \([14]\) of the detrended fluctuation analysis (DFA) \([40,41]\) for the detection of long-range temporal correlations in \( m_k \).
DFA has been established as a robust method suitable for detecting long-range power law correlations embedded in non-stationary signals [for recent applications see refs. [42,43]] and can be summarized as follows: we first find the average value $m$ and determine the profile $y(i) = \sum_{k=1}^W (m_k - m)$, $i = 1, \ldots, W$. We then divide this profile of length $W$ into $W/l(= W/l)$ non-overlapping fragments of $l$-observations. Next, we define the detrended process $y_{l,\nu}(m)$, in the $\nu$-th fragment, as the difference between the original value of the profile and the local linear trend. We then calculate the mean variance of the detrended process: $F_{DFA}(l)=\frac{1}{W} \sum_{i=1}^W f^2(l,\nu)$, where

$$f^2(l,\nu) = \frac{1}{l} \sum_{i=1}^l y_{l,\nu}(m).$$

If $F_{DFA}(l) \sim l^\alpha$, the slope of the log $F_{DFA}(l)$ vs. log $l$ plot, leads to the value of the exponent $\alpha_{DFA} = \alpha$. If $\alpha_{DFA} = 0.5$, there is no correlation and the signal is uncorrelated (white noise); if $\alpha_{DFA} < 0.5$, the signal is anti-correlated; if $\alpha_{DFA} > 0.5$, the signal is correlated and specifically the case $\alpha_{DFA} = 1.5$ corresponds to the Brownian motion (integrated white noise).

Figure 3 depicts with black plus symbols the resulting log$_{10}[F_{DFA}(l)]$ vs. log$_{10}(l)$ for the whole SCEC data. A crossover is observed at $k = l \sim 200$, below which the $\alpha$-value is close to 0.61 ($\equiv \alpha_{low}$). This value agrees fairly well with the one $\alpha = 0.59(5)$ obtained in ref. [14] when solely analyzing the periods of stationary seismic activity. Thus, the substantially higher value ($\alpha_{high} = 0.93$) obtained for scales larger than $l \approx 200$ emerged upon the inclusion of the “non-stationary” periods of seismic activity in the present study. To further shed light on the origin of this crossover, we examined the behavior of the magnitude time-series before and after the two most significant earthquakes reported in SCEC, i.e., the Landers EQ (fig. 3(a)) with magnitude 7.3 (that occurred at 11:57 UT on June 28, 1992 with an epicenter at 34.2°N 116.4°W) and the Hector Mine EQ (fig. 3(b)) with magnitude 7.1 (that occurred at 09:46 UT on October 16, 1999 with an epicenter at 34.6°N 116.3°W). The aftershock magnitude time-series for both these mainshocks, as identified [44] by examining the corresponding Omori law regimes, have been analyzed in natural time [38]. The corresponding pdfs $P(k_i)$, depicted with the red plus symbols in figs. 2(a) and (b), respectively, were found [38] to almost coincide with the $P(k_i)$ for the whole SCEC. This could be interpreted as a “return” of the seismic activity to its mean behavior after the completion of the aftershock sequence. The application of DFA to both aftershock time-series is shown by the red plus symbols in figs. 3(a) and (b), respectively. Interestingly, an inspection of this figure also reveals that the scaling behavior of DFA in both aftershock series is close to that of the whole SCEC, thus being in accordance with the previous result of the natural time analysis. Notice that the crossover still persists.

In order to further shed light on the analysis, we examined the magnitude time-series with lengths $W = 5000, 3000$ and 1000 EQs just before and just after these two EQs in SCEC. The corresponding results of DFA are given in figs. 3(a) and (b) for the Landers and the Hector Mine EQ, respectively. They show that the high value of the DFA exponent $\alpha$ at longer scales should be attributed to the highly correlated “immediate” aftershocks (e.g., for $W = 1000$ see the blue circles in fig. 3). We now proceed to the study of the magnitude time-series solely before these two mainshocks. These DFA results suggest that the $\alpha$ value for scales longer than the crossover is now (see the squares, triangles and inverted triangles in fig. 3) significantly smaller than in the case of aftershocks, and much closer to that for scales shorter than the crossover. Thus, the crossover effect is definitely smoothed in the magnitude time-series that end just before a mainshock. A closer inspection of the inverted triangles, i.e., the results obtained from $W = 1000$ EQs just before the mainshock, indicates that the DFA scaling exponent becomes even smaller than $\alpha_{low} (= 0.61)$ and the values obtained are $\alpha = 0.53(2)$ and $\alpha = 0.50(2)$ for the Landers and the Hector Mine EQ, respectively. Thus, the “foreshocks” appear to

![Fig. 3: (Color online) Detrended fluctuation analysis of SCEC (black plus) along with those (1) of the aftershock sequence as reported in ref. [44] for the Landers EQ (a) and the Hector Mine EQ (b). The points (2) to (7) depict the DFA of W = 5000, 3000, 1000 EQs immediately after and W = 5000, 3000, 1000 EQs immediately before the Landers EQ (a) and the Hector Mine EQ (b), respectively. The straight lines with $\alpha = 0.61$ and 0.93 have been drawn as a guide to the eye.](image-url)
exhibit correlations that are somewhat weaker than those found [14] in the stationary seismicity (cf. recall ref. [14] reported \( \alpha = 0.59(5) \) for stationary periods).

The results of the natural time analysis of the time-series with \( W = 5000, 3000 \) and 1000 EQs just before and just after Landers and Hector Mine EQ are shown in fig. 2(a) and (b), respectively. In particular, we observe that the pdf \( P(\kappa_1) \) vs. \( \kappa_1 \) curves differ in general from the relevant curve obtained from whole SCEC or from the aftershock time-series identified in ref. [44]. This reveals that either before or after a significant EQ, the seismicity deviates from its mean behavior in natural time. The most important difference, however, is noticed when we compare the pdfs \( P(\kappa_1) \) before and after a significant EQ. For example, in fig. 2(c) when plotting \( P(\kappa_1) \) vs. \( \kappa_1 \), for \( W = 1000 \) EQs, before and after the Landers EQ with the thick red and the thin red line, respectively, they are markedly different. Thus, natural time analysis made it possible to reveal the following picture of profound importance: Before the Landers EQ a significant bimodal feature appears in \( P(\kappa_1) \) vs. \( \kappa_1 \), which is strikingly reminiscent of the bimodal feature observed in the pdf of the order parameter when approaching (from below) \( T_c \) in equilibrium critical phenomena (e.g., see fig. 13(b) of ref. [38]). Since it has been suggested [21] that \( \kappa_1 \) is an order parameter for seismicity, a similar behavior should be expected before every mainshock in SCEC. For example, in fig. 2(c), we also depict \( P(\kappa_1) \) vs. \( \kappa_1 \) for \( W = 5000 \) events before and after the Hector Mine EQ with the thick blue and the thin blue line, respectively. We again observe that the bimodal feature is much more significant in the curve before the mainshock. In what remains we assume that the variability \( \beta \equiv \sigma(\kappa_1)/\mu(\kappa_1) \) constitutes, at a first approximation, a measure to quantify the presence of this bimodal behavior.

If the presence of a bimodal feature in \( P(\kappa_1) \) vs. \( \kappa_1 \) actually signifies the occurrence of an impending mainshock, then the quantity \( \beta \) can, in principle, be used as a decision variable to predict the occurrence of a large earthquake solely based on the past magnitudes, in a fashion analogous to that followed by the authors of ref. [45] where in self-organized critical sandpiles the past avalanches were used for the prediction of a future avalanche exceeding a certain threshold to occur in the next step. Such a “prediction” scheme should not be confused with the one achieved when the seismic data are enriched (supplemented) with geoelectrical data which enable the detection of precursory electric field changes termed Seismic Electric Signals (SES) [46–51] on the basis of which the epicentral area and the magnitude of the impending mainshock can be determined [52,53]. Once the latter are available, the natural time analysis of the seismicity (in the future epicentral area) that occurs after the SES detection, leads to the identification of the time window of the impending mainshock within a narrow range of a few days to one week [37,38,54–57].

Hereafter, we proceed as follows: For each event \( m_k \) in SCEC, we estimate \( \pi_1 = \sum_{N=6}^{40} \kappa_1(N)/35 \), i.e., the average value of the \( \kappa_1 \)'s that we calculated upon considering \( N = 6 \) to 40 consecutive EQs (including the \( k \)-th event). Next, we assign this value to \( k' \)-th, \( k' = k + 40 \), element of the time-series the \( \kappa(k') \). Following this way, \( \kappa(k') \) has no information of the event with magnitude \( m_k \) which is the 40-th EQ that occurred after \( m_k \). We can now estimate, for various windows of \( W \) EQs, the time-series of the average values

\[
\mu_k(W) = \frac{1}{W} \sum_{n=k′-W+1}^{k′} \kappa(n),
\]

which is equivalent to \( \mu(\kappa_1) \) obtained when considering a catalog consisting of the \( W \) EQs that occurred just before \( m_k \). In addition, the time-series of standard deviations can be obtained from

\[
\sigma_k(W) = \sqrt{\frac{1}{W} \sum_{n=k′-W+1}^{k′} [\kappa(n) - \mu_k(W)]^2},
\]

and the variability time-series is given by

\[
\beta_k(W) = \frac{\sigma_k(W)}{\mu_k(W)}.
\]

We will now examine whether \( \beta_k \) can be used as a decision variable for binary “predictions”. Following the terminology of Keilis-Borok and coworkers [58,59], the time increased probability (TIP) is turned on when \( \beta_k(W) \geq \beta_c \), where \( \beta_c \) is a given threshold in the prediction. If the magnitude \( m_k \) is greater than or equal to a target threshold \( M_{thres} \), we have a successful prediction. For the present case of binary predictions, the prediction of events becomes a classification task, with two type of errors: missing an event and giving a false alarm. We therefore choose, following ref. [45], the receiver operating characteristics (ROC) [60] as the method to analyze here the prediction quality. This is a plot of the hit rate vs. the false alarm rate, which here is tuned by the threshold \( \beta_c \). Only if in between the hit rate exceeds the false alarm rate, the predictor is useful. Random predictions generate equal hit and false alarm rate, and hence they lead to the diagonal in ROC plot, see the black straight lines in figs. 4 and 5. (If \( \beta_c \) is maximum, both hit rate and false alarm rate are zero, while for very small \( \beta_c \) values both rates tend to unity.) Thus, only when the points lie above this diagonal the predictor is useful. Figure 4 depicts the ROC curves, for various values of \( M_{thres} = 3.0 \) to 4.5, together with the results obtained when using, for example, two randomly shuffled copies (green and red circles) of SCEC. The results for various \( W \) values are shown, i.e., \( W = 70, 300, 1000 \) and 3000, in figs. 4(a) to (d), respectively. In all cases, the results are better (i.e., points lying above the diagonal) for the original SCEC catalogue than the randomly shuffled ones. This indicates that the predictive power of \( \beta_k(W) \) stems from temporal correlations present in the actual seismicity.
In order to further examine the statistical significance of this “prediction” scheme, we depict in fig. 5 the results for \( W = 1000 \) together with the results of \( 10^2 \) runs of the same catalogue when using as decision variable a uniformly distributed random number in the same range as \( \beta_k(1000) \). We observe that none of these runs outperforms \( \beta_k(1000) \) for false alarm rates from 20\% to 60\%. Thus, \( \beta_k(1000) \) has predictive power which is statistically significant. The inset of fig. 5 depicts the ratio of the true positive rate (hit rate) vs. \( M_{\text{thres}} \), which shows that the prediction results become better upon increasing \( M_{\text{thres}} \). For example, when \( M_{\text{thres}} = 4 \) (light blue line with squares) the hit rate is approximately 60\% when the false alarm rate is 50\%. The TIP can be visualized in fig. 6, where the red shaded areas correspond to the periods when the TIP is on (i.e., \( \beta_k(1000) \) \( \geq 0.35 \)). The results convincingly outperform chance, but are not spectacular. This (which remains so when using, instead of \( \beta_k \), the kurtosis, see the black dots in fig. 6) is not unreasonable in view of the fact that when using a constant \( W \) it may not correspond to the time at which the focal area of the impending mainshock enters into the critical regime which is captured however by the SES detection when available.

In conclusion, we made use of the order parameter \( \kappa_1 \) of seismicity defined \cite{21} in natural time together with the DFA of the magnitude time-series to investigate the period before and after a significant mainshock. The study was focused on two significant EQs that occurred.
in Southern California in 1992 and 1999, i.e., the Landers and the Hector Mine earthquakes. Magnitude time-series for various lengths of W EQs that occurred before or after the main shock have been considered. Quite interestingly, the natural time analysis of these time-series, reveals that “foreshocks” exhibit a behavior characteristic of systems close to their critical point: i.e., upon considering the order parameter of seismicity $k_1$ the probability distribution function $P(k_1)$ versus $k_1$ exhibits a bimodal feature. In an attempt to quantify this feature, we considered the variability of $k_1$, which was then used as decision variable for the “prediction” of the occurrence of a large earthquake in the next time step based solely on the magnitudes of previous earthquakes. The results outperform chance.

REFERENCES