

Effect of significant data loss on identifying electric signals that precede rupture by detrended fluctuation analysis in natural time

E. S. Skordas,¹ N. V. Sarlis,¹ and P. A. Varotsos^{1,*}

¹*Solid State Section and Solid Earth Physics Institute, Physics Department, University of Athens, Panepistimiopolis, Zografos 157 84, Athens, Greece*

Electric field variations that appear before rupture have been recently studied by employing the detrended fluctuation analysis (DFA) as a scaling method to quantify long-range temporal correlations. These studies revealed that seismic electric signals (SES) activities exhibit a scale invariant feature with an exponent $\alpha_{DFA} \approx 1$ over all scales investigated (around five orders of magnitude). Here, we study what happens upon significant data loss, which is a question of primary practical importance, and show that the DFA applied to the natural time representation of the remaining data still reveals for SES activities an exponent close to 1.0, which markedly exceeds the exponent found in artificial (man-made) noises. This, in combination with natural time analysis, enables the identification of a SES activity with probability 75% even after a significant (70%) data loss. The probability increases to 90% or larger for 50% data loss.

Keywords: detrended fluctuation analysis, complex systems, scale invariance

PACS numbers: 91.30.-f, 05.40.-a

Complex systems usually exhibit scale-invariant features characterized by long-range power-law correlations, which are often difficult to quantify due to various types of non-stationarities observed in the signals emitted. This also happens when monitoring geoelectric field changes aiming at detecting Seismic Electric Signals (SES) activities that appear before major earthquakes. To overcome this difficulty the novel method of detrended fluctuation analysis (DFA) has been employed, which when combined with a newly introduced time domain termed natural time, allows a reliable distinction of true SES activities from artificial (man-made) noises. This is so, because the SES activities are characterized by infinitely ranged temporal correlations (thus resulting in DFA exponents close to unity) while the artificial noises are not. The analysis of SES observations often meet the difficulty of significant data loss caused either by failure of the data collection system or by removal of seriously noise-contaminated data segments. Thus, here we focus on the study of the effect of significant data loss on the long-range correlated SES activities quantified by DFA. We find that the remaining data, even after a considerable percentage of data loss (which may reach $\sim 80\%$), may be correctly interpreted, thus revealing the scaling properties of SES activities. This is achieved, by applying DFA *not* to the original time series of the remaining data but to those resulted when employing natural time.

I. INTRODUCTION

The output signals of complex systems exhibit fluctuations over multiple scales^{1,2} which are characterized by

absence of dynamic scale, i.e., scale-invariant behavior³. These signals, due to the nonlinear mechanisms controlling the underlying interactions, are also typically non-stationary and their reliable analysis cannot be achieved by traditional methods, e.g., power-spectrum and auto-correlation analysis⁴⁻⁶. On the other hand, the detrended fluctuation analysis (DFA)^{7,8} has been established as a robust method suitable for detecting long-range power-law correlations embedded in non-stationary signals. This is so, because a power spectrum calculation assumes that the signal is stationary and hence when applied to non-stationary time series it can lead to misleading results. Thus, a power spectrum analysis should be necessarily preceded by a test for the stationarity of the (portions of the) data analyzed. As for the DFA, it can determine the (mono) fractal scaling properties (see below) even in non-stationary time series, and can avoid, in principle, spurious detection of correlations that are artifacts of non-stationarities. DFA has been applied with successful results to diverse fields where scale-invariant behavior emerges, such as DNA⁹⁻¹⁶, heart dynamics¹⁷⁻²⁷, circadian rhythms²⁸⁻³¹, meteorology³² and climate temperature fluctuations³³⁻³⁷, economics³⁸⁻⁴⁴ as well as in the low-frequency (≤ 1 Hz) variations of the electric field of the earth that precede earthquakes⁴⁵⁻⁴⁷ termed Seismic Electric Signals⁴⁸⁻⁵⁴ and in the relevant⁵⁵⁻⁵⁸ magnetic field variations⁴⁷.

Monofractal signals are homogeneous in the sense that they have the same scaling properties, characterized locally by a single singularity exponent h_0 , throughout the signal. Thus, monofractal signals can be indexed by a single global exponent, e.g., the Hurst exponent $H \equiv h_0$, which suggests that they are stationary from viewpoint of their local scaling properties (see Ref.²² and references therein). Since DFA can measure only one exponent, this method is more suitable for the investigation of monofractal signals. In several cases, however, the records cannot be accounted for by a single scaling expo-

nent (i.e., do not exhibit a simple monofractal behavior). In some examples, there exist crossover (time-) scales separating regimes with different scaling exponents. In general, if a multitude of scaling exponents is required for a full description of the scaling behavior, a multifractal analysis must be applied. Multifractal signals are intrinsically more complex, and inhomogeneous, than monofractals (see Ref.²² and references therein). A reliable multifractal analysis can be performed by the Multifractal Detrended Fluctuation Analysis, MF-DFA^{59,60} or by the wavelet transform (e.g., see Ref.⁶¹).

DFA has been applied, as mentioned, to the SES activities. It was found⁴⁵ that when DFA is applied to the original time series of the SES activities and artificial (man-made) noises, both types of signals lead to a slope at short times (i.e., $\Delta t \leq 30s$) lying in the range $\alpha=1.1-1.4$, while for longer times the range $\alpha=0.8-1.0$ was determined without, however, any safe classification between SES activities and artificial noises. On the other hand, when employing natural time (see Section II), DFA enables the distinction between SES activities and artificial noises in view of the following difference: for the SES activities the α -values lie approximately in the range 0.9 - 1.0 (or between 0.85 to 1.1, if a reasonable experimental error is envisaged), while for the artificial noises the α -values are markedly smaller, i.e., $\alpha=0.65-0.8$. In addition, MF-DFA has been used^{45,46} and it was found that this multifractal analysis, when carried out in the conventional time frame, did not lead to any distinction between these two types of signals, but does so, if the analysis is made in the natural time domain.

The aforementioned findings of DFA for SES activities, are consistent with their generation mechanism^{62,63} which could be summarized as follows. Beyond the usual intrinsic lattice defects⁶⁴⁻⁶⁹ that exist in solids, in ionic solids in particular, when doped with aliovalent impurities, extrinsic defects are formed for the sake of charge compensation. A portion of these defects are attracted by the nearby impurities, thus forming electric dipoles the orientation of which can change by means of a defect migration. In the focal area of an impending earthquake the stress gradually increases and hence affects the thermodynamic parameters of this migration, thus it may result in a gradual decrease of their relaxation time. When the stress (pressure) reaches a *critical* value, a *cooperative* orientation of these dipoles occurs, which leads to the emission of a transient signal. This signal constitutes the SES activity and, since it is characterized by *critical* dynamics, should exhibit infinitely range temporal correlations. This is consistent with the above findings of DFA that for SES activities $\alpha_{DFA} \approx 1$.

Hereafter, we will solely use DFA in view of its simplicity and its ability to reliably classify SES activities. It is the basic aim of this study to investigate how significant data loss affects the scaling behavior of long-range correlated SES activities inspired from the new segmentation approach introduced recently by Ma et al⁷⁰ to generate surrogate signals by randomly removing data segments

from stationary signals with different types of long-range correlations. The practical importance of this study becomes very clear upon considering that such a data loss is inevitable mainly due to the following two reasons: First, failure of the measuring system in the field station, including the electric measuring dipoles, electronics and the data collection system, may occur especially due to lightning. Second, noise-contaminated data segments are often unavoidable due to natural changes such as rainfall, lightning, induction of geomagnetic field variations and ocean-earth tides besides the noise from artificial (man-made) sources including the leakage currents from DC driven trains. The latter are common in Japan where at some sites they may last for almost 70% of the time every day. We clarify, however, that even at such noisy-stations in Japan, several clear SES activities have been unambiguously identified⁷¹ during the night (when the noise level is low). In addition, prominent SES activities were recently reported⁷² at noise-free stations (far from industrialized regions) having long duration, i.e., of the order of several weeks. As we shall see, our results described in Section IV, are in essential agreement with those obtained in the innovative and exhaustive study of Ma et al⁷⁰. Before proceeding to our results, we will briefly summarize DFA and natural time analysis in Section II, and then present in Section III the most recent SES data along with their analysis in natural time. In Section V, we summarize our conclusions.

II. CONVENTIONAL DETRENDED FLUCTUATION ANALYSIS. NATURAL TIME.

We first sum up the original time series and determine the profile $y(i), i = 1, \dots, N$; We then divide this profile of length N into $N/l (\equiv N_l)$ non overlapping fragments of l -observations. Next, we define the detrended process $y_{l,\nu}(m)$, in the ν -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:

$$F^2(l) = \frac{1}{N_l} \sum_{\nu=1}^{N_l} f^2(l, \nu) \quad (1)$$

where

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

If $F(l) \sim l^\alpha$, the slope of the $\log F(l)$ versus $\log l$ plot, leads to the value of the exponent $\alpha_{DFA} \equiv \alpha$. (This scaling exponent is a self-similarity parameter that represents the long-range power-law correlations of the signal). 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).

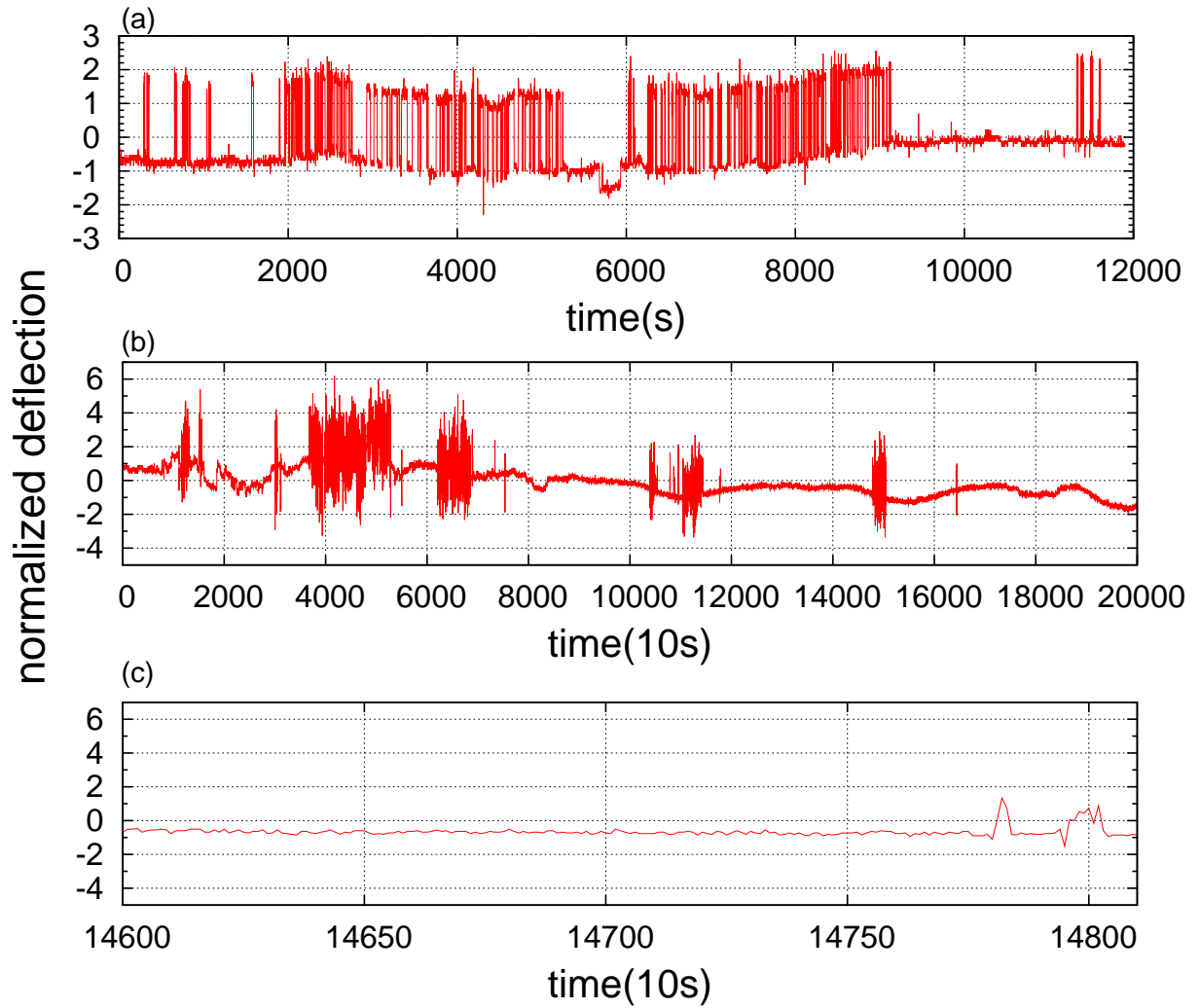


FIG. 1: (color online) Examples of the electric field recordings in normalized units, i.e., by subtracting the mean value μ and dividing by the standard deviation σ . The following SES activities are depicted: (a) the one recorded on April 18, 1995 at Ioannina station, (b) the long duration SES activity recorded from December 27, 2010 to December 30, 2009 at Lamia station. (c) is an excerpt of (b) showing that, after long periods of quiescence, the electric field exhibits measurable excursions (transient pulses).

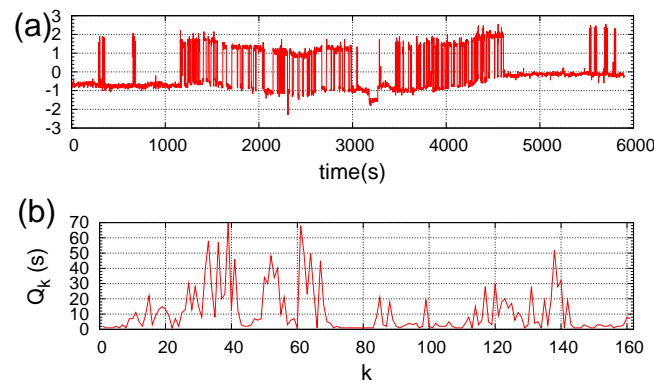


FIG. 2: (color online) (a): Example of a surrogate time-series (in normalized units as in Fig.1) obtained by removing segments of length $L = 200$ from the signal of Fig.1(a) with 50% data loss (i.e., $p = 0.50$). (b): The natural time representation of (a). The values obtained from the analysis of (b) in natural time are $\kappa_1 = 0.067(4)$, $S = 0.076(4)$, $S_- = 0.071(4)$ and $a_{DFA} = 0.90(5)$.

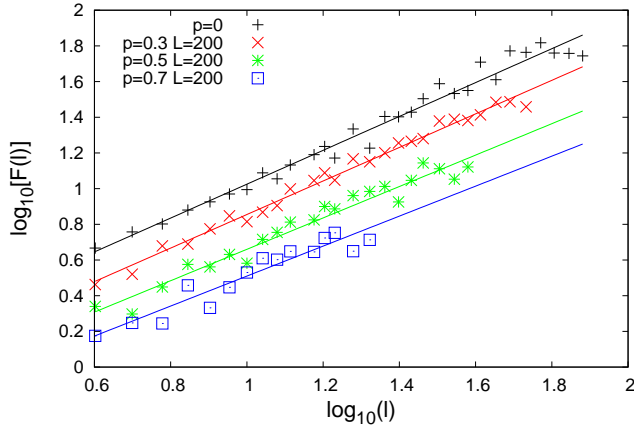


FIG. 3: (color online) The dependence of the DFA measure $F(l)$ versus the scale l in natural time: we increase the percentage of data loss p by removing segments of length $L = 200$ samples from the signal of Fig.1(a). The black (plus) symbols correspond to no data loss ($p=0$), the red (crosses) to 30% data loss ($p=0.3$), the green (asterisks) to 50% data loss ($p=0.5$) and the blue (squares) to 70% data loss ($p=0.7$). Except for the case $p=0$, the data have been shifted vertically for the sake of clarity. The slopes of the corresponding straight lines that fit the data lead to $\alpha_{DFA} = 0.95, 0.94, 0.88$ and 0.84 from the top to bottom, respectively. They correspond to the average values of α_{DFA} obtained from 5000 surrogate time-series that were generated with the method of surrogate by Ma et al.⁷⁰ (see the text).

We now summarize the background of natural time χ . In a time series comprising N events, the natural time $\chi_k = k/N$ serves as an index⁷³ for the occurrence of the k -th event. The evolution of the pair (χ_k, Q_k) is studied^{45,46,73-78}, where Q_k denotes a quantity proportional to the *energy* released in the k -th event. For dichotomous signals, which is frequently the case of SES activities, the quantity Q_k stands for the duration of the k -th pulse. By defining $p_k = Q_k / \sum_{n=1}^N Q_n$, we have found⁷³ that the variance

$$\kappa_1 = \langle \chi^2 \rangle - \langle \chi \rangle^2, \quad (3)$$

where $\langle f(\chi) \rangle = \sum_{k=1}^N p_k f(\chi_k)$, of the natural time χ with respect to the distribution p_k may be used^{45,46} for the identification of SES activities. In particular, the following relation should hold

$$\kappa_1 \approx 0.070. \quad (4)$$

The entropy S in the natural time-domain is defined as⁴⁶

$$S \equiv \langle \chi \ln \chi \rangle - \langle \chi \rangle \ln \langle \chi \rangle. \quad (5)$$

It exhibits⁷⁵ Lesche^{79,80} (experimental) stability, and for SES activities (critical dynamics) is smaller⁴⁶ than the value $S_u (= \ln 2/2 - 1/4 \approx 0.0966)$ of a “uniform” (u) distribution (as defined in Refs.^{45,46,74}, e.g. when all p_k

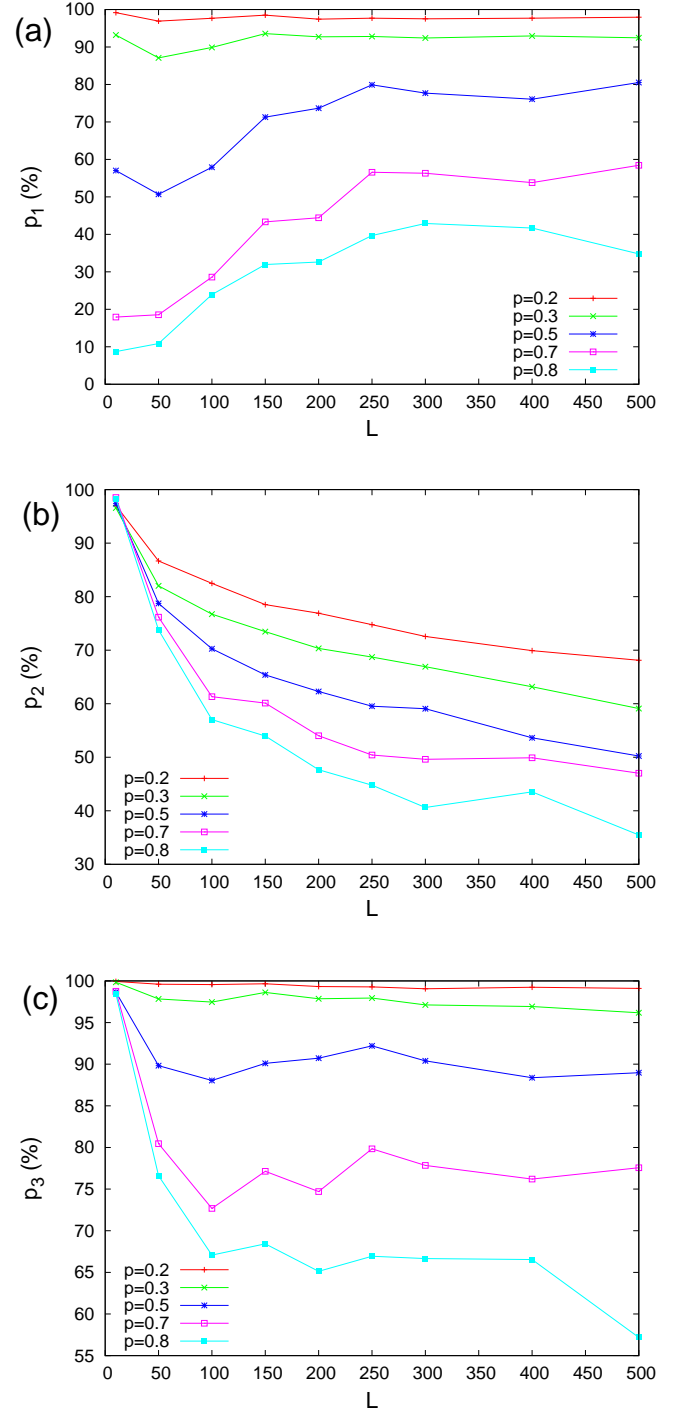


FIG. 4: (color online) The probabilities p_1 (a), p_2 (b) and p_3 (c) to recognize the signal of Fig.1(a) as true SES activity when considering various percentages of data loss $p=0.2, 0.3, 0.5, 0.7$ and 0.8 as a function of the length L of the contiguous samples removed. The removal of large segments leads to better results when using DFA in natural time (a), whereas the opposite holds when using the conditions of Eqs.(6) and (7) for κ_1 , S and S_- (b). The optimum selection (c) for the identification of a signal as SES activity consists of a proper combination of the aforementioned procedures in (a) and (b), see the text. The values presented have been obtained from 5000 surrogate time-series (for a given value of p and L), and hence they have a plausible error 1.4% ($\approx 1/\sqrt{5000}$).

are equal or Q_k are positive independent and identically distributed random variables of finite variance. In this case, κ_1 and S are designated $\kappa_u (= 1/12)$ and S_u , respectively. Thus, $S < S_u$. The same holds for the value of the entropy obtained^{75,77} upon considering the time reversal \hat{T} , (the operator \hat{T} is defined by $\hat{T}p_k = p_{N-k+1}$), which is labelled by S_- .

In summary, the SES activities, in contrast to the signals produced by man-made electrical sources, when analyzed in natural time (see Section I) exhibit *infinitely* ranged temporal correlations^{46,73} and obey the conditions⁷⁷:

$$\kappa_1 \approx 0.07 \quad (6)$$

and

$$S, S_- < S_u. \quad (7)$$

It should be recalled that SES activities are publicized *only* when the magnitude of the impending earthquake is $M_s(\text{ATH}) \geq 6.0$

III. THE EXPERIMENTAL DATA

In Fig.1(a), we depict the SES activity recorded at Ioannina station (Northwestern Greece) on 18 April, 1995. It preceded the 6.6 earthquake on 13 May, 1995. Since this earthquake was the strongest one in Greece during the 25 year period 1983-2007, we focus on this example in the next Section to present the DFA results.

In addition, the most recent SES activity in Greece is depicted in Figs.1(b),(c). This has been recorded at Lamia station located in Central Greece during the period 27 December - 30 December, 2009. Almost three weeks later, two strong earthquakes of magnitude $M_s(\text{ATH})=5.7$ and 5.6 occurred in central Greece with an epicenter at 38.4°N 22.0°E (but see also Refs.^{81,82}).

The two signals in Figs. 1(a),(b) have been classified as SES activities after analyzing them in natural time. In particular, for the signal in Fig. 1(a), straightforward application of natural time analysis leads to the conclusion that the conditions (6) and (7) are satisfied (see Table I of Ref.⁷⁵). Further, the DFA analysis of the natural representation of this signal gives an exponent $\alpha_{DFA} = 0.95(4) \approx 1$. (We also note that the classification of this signal as SES activity had been previously achieved by independent procedures discussed⁸³ in the Royal Society Meeting that was held during May 11-12, 1995 before⁸⁴ the occurrence of the 6.6 earthquake on May 13, 1995.) For the long duration signal of Fig. 1(b), the procedure explained in detail in Ref.⁴⁷ was followed.

IV. DATA ANALYSIS AND RESULTS

Following Ma et al⁷⁰, we now describe the segmentation approach used here to generate surrogate signals $\tilde{u}(i)$

by randomly removing data segments of length L from the original signal $u(i)$. The percentage p of the data loss, i.e., the percentage of the data removed, also characterizes the signal $\tilde{u}(i)$. The procedure followed is based on the construction of a binary time-series $g(i)$ of the same length as $u(i)$. The values of $u(i)$ that correspond to $g(i)$ equal to unity are kept, whereas the data of $u(i)$ when $g(i)$ equals zero are removed. The values of $u(i)$ kept, are then concatenated to construct $\tilde{u}(i)$.

The binary time-series $g(i)$ is obtained as follows⁷⁰: (i) We first generate the lengths $l_j = L$ with $j = 1, 2, \dots, M$ of the removed segments, by selecting M to be the smallest integer so that the total number of removed data satisfies the condition $\sum_{j=1}^M l_j \geq pN$. (ii) We then construct an auxiliary time-series $a(k)$ with $a(k) = L$ when $k = 1, 2, \dots, M$ and $a(k) = 1$ when $k = M + 1, \dots, N - M(L + 1)$ of size $N - M(L + 1)$. (iii) We shuffle the time-series $a(k)$ randomly to obtain $\tilde{a}(k)$. (iv) We then append $\tilde{a}(k)$ to obtain $g(i)$: if $\tilde{a}(k) = 1$ we keep it, but we replace all $\tilde{a}(k) = L$ with L elements of value '0' and one element with value '1'. In this way, a binary series $g(i)$ is obtained, which has a size equal to the one of the original signal $u(i)$. We then construct the surrogate signal $\tilde{u}(i)$ by simultaneously scanning the original signal $u(i)$ and the binary series $g(i)$, removing the i -th element of $u(i)$ if $g(i) = 0$ and concatenating the segments of the remaining data to $\tilde{u}(i)$.

The resulting signal $\tilde{u}(i)$ is later analyzed in natural time, thus leading to the quantities κ_1 , S and S_- as well as to the DFA exponent α_{DFA} in natural time. Such an example is given in Fig. 2. In what remains we present the results focusing hereafter, as mentioned, on the example of the SES activity depicted in Fig. 1(a).

Typical DFA plots, obtained for $L=200$ and $p = 30, 50$ and 70% are given in Fig. 3. For the sake of comparison, this figure also includes the case of no data loss (i.e., $p = 0$). We notice a gradual decrease of α_{DFA} upon increasing the data loss, which affects our ability to classify a signal as SES activity.

In order to quantify, in general, our ability to identify SES activities from the natural time analysis of surrogate signals with various levels of data loss, three procedures have been attempted:

Let us call procedure 1, the investigation whether α_{DFA} , resulted from the DFA analysis of the natural time representation of a signal, belongs to the range $0.85 \leq \alpha_{DFA} \leq 1.0$. If it does, the signal is then classified as SES activity. Figure 4(a) shows that for a given amount of data loss ($p=\text{const}$), upon increasing the length L of the randomly removed segment, the probability p_1 of achieving, after making 5000 attempts (for a given value of p and L), the identification of the signal as SES activity is found to gradually increase versus L at small scales and stabilizes at large scales. For example, when considering the case of 70% data loss (magenta color in Fig. 4(a)) the probability p_1 is close to 20% for $L=50$; it increases to $p_1 \approx 30\%$ for $L=100$ and finally stabilizes around 50% for lengths $L = 300$ to 500.

Let us now label as procedure 2, the investigation whether the quantities κ_1 , S and S_- (resulted from the analysis of a signal in natural time) obey the conditions (6) and (7), i.e., $|\kappa_1 - 0.070| \leq 0.01$ and $S, S_- < 0.0966$. If they do so, the signal is classified as SES activity. Figure 4(b) shows that for a given amount of data loss, the probability p_2 of achieving the signal identification as SES activity -that results after making 5000 attempts for each L value- gradually decreases when moving from the small to large scales. Note that for the smallest length scale investigated, i.e., $L=10$ (which is more or less comparable -if we consider the sampling frequency of 1 sample/sec- with the average duration ≈ 11 sec of the transient pulses that constitute the signal), the probability p_2 reaches values close to 100% even for the extreme data loss of 80%. This is understood in the context that the quantities κ_1 , S and S_- remain almost unaffected when randomly removing segments with lengths comparable to the average pulse's duration. This is consistent with our earlier finding⁷⁵ that the quantities κ_1 , S and S_- are experimentally stable (Lesche's stability) meaning that they exhibit only slight variations when deleting (due to experimental errors) a small number of pulses. On the other hand, at large scales of L , p_2 markedly decreases. This may be understood if we consider that, at such scales, each segment of contiguous L samples removed, comprises on the average a considerable number of pulses the removal of which may seriously affect the quantities κ_1 , S and S_- . As an example, for 80% data loss (cyan curve in Fig. 4(b)), and for lengths $L=400-500$, the probability p_2 of identifying a true SES activity is around 40%.

Interestingly, a closer inspection of the two figures 4(a) and 4(b) reveals that p_1 and p_2 play complementary roles. In particular, at small scales of L , p_1 increases but p_2 decreases versus L . At large scales, where p_1 reaches (for considerable values of data loss) its largest value, the p_2 value becomes small. Inspired from this complementary behavior of p_1 and p_2 , we proceeded to the investigation of a combined procedure, let us call it procedure 3. In this procedure, a signal is identified as SES activity when *either* the condition $0.85 \leq \alpha_{DFA} \leq 1.0$ *or* the relations (6) and (7) are satisfied. The probability p_3 of achieving such an identification, after making 5000 attempts (for a given value of p and L), is plotted in Fig. 4(c). The results are remarkable since, even at significant values of data loss, e.g., $p = 70\%$ or 80% , the probability p_3 of

identifying a SES activity at scales $L=100$ to 400 remains relatively high, i.e., $p_3 \approx 75\%$ and 65% , respectively (cf. note also that the value of p_3 reaches values close to 100% at small scales $L=10$). This is important from practical point of view, because it states for example the following: Even if the records of a station are contaminated by considerable noise, say 70% of the time of its operation, the remaining 30% of the non-contaminated segments have a chance of $\sim 75\%$ to correctly identify a SES activity. The chances increase considerably, i.e., to $p_3 \approx 90\%$, if only half of the recordings are noisy.

The aforementioned results have been deduced from the analysis of a SES activity lasting around three hours. In cases of SES activities with appreciably longer duration, e.g., a few to several days^{47,85} detected in Greece or a few months in Japan⁷², the results should become appreciably better.

V. CONCLUSIONS

We start our conclusions by recalling that the distinction between SES activities (critical dynamics, infinitely ranged temporal correlations) and artificial (man-made) noise remains an extremely difficult task, even without any data loss, when solely focusing on the original time series of electrical records which are, of course, in conventional time. On the other hand, when combining natural time with DFA analysis, such a distinction becomes possible even after significant data loss. In particular we showed for example that even when randomly removing 50% of the data, we have a probability (p_3) around 90%, or larger, to identify correctly a SES activity. This probability becomes somewhat smaller, i.e., 75%, when the data loss increases to 70%. To achieve this goal, the proper procedure is the following: the signal is first represented in natural time and then analyzed in order to deduce the quantities κ_1 , S and S_- as well as the exponent α_{DFA} from the slope of the log-log plot of the DFA analysis in natural time. We then examine whether the latter slope has a value close to unity *or* the conditions $\kappa_1 \approx 0.070$ and $S, S_- < S_u$ are obeyed. In other words, the consequences caused by an undesirable severe data loss can be markedly reduced upon taking advantage of the DFA and natural time analysis.

* Electronic address: pvaro@otenet.gr

¹ J. B. Bassingthwaighe, L. Liebovitch, and B. J. West, *Fractal Physiology* (Oxford University Press, Oxford UK, 1994).

² M. M. Malik and A. Camm, *Heart Rate Variability* (Futura, Armonk NY, 1995).

³ H. E. Stanley, *Nature* **378**, 554 (1995).

⁴ H. E. Hurst, *Trans. Am. Soc. Civ. Eng.* **116**, 770 (1951).

⁵ B. B. Mandelbrot and J. R. Wallis, *Water Resources Re-*

search **5**, 321 (1969).

⁶ R. L. Stratonovich, *Topics in the Theory of Random Noise, vol.I* (Gordon and Breach, New York, 1981).

⁷ C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys. Rev. E* **49**, 1685 (1994).

⁸ M. S. Taqqu, V. Teverovsky, and W. Willinger, *Fractals* **3**, 785 (1995).

⁹ C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin,

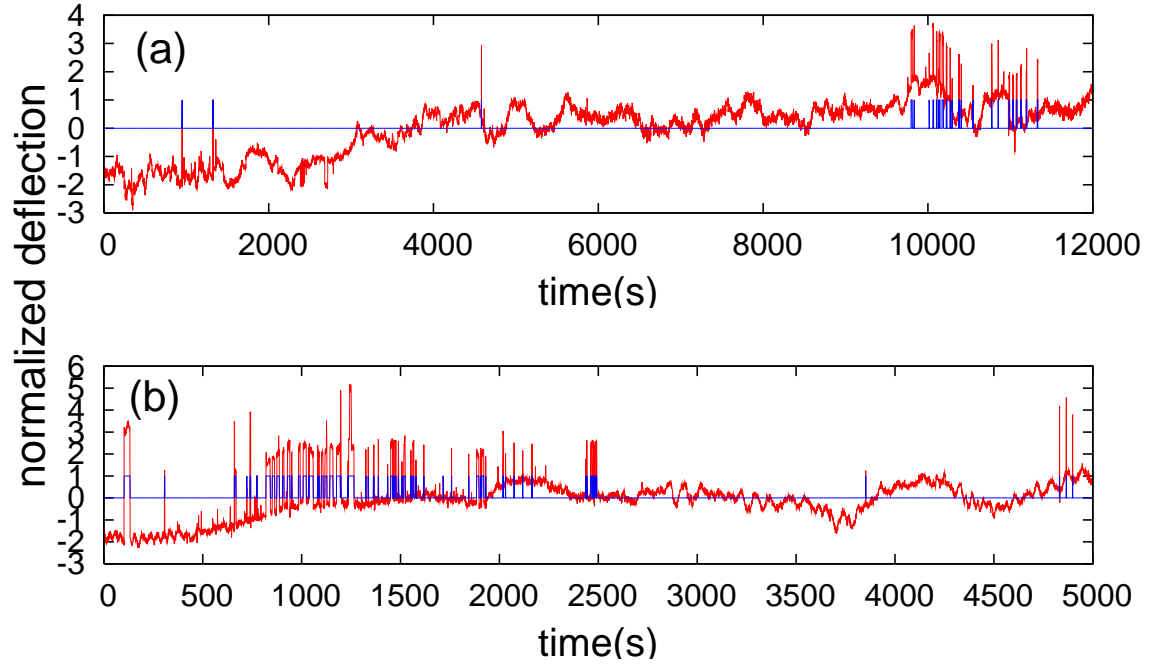


FIG. 5: (color online) Electric field recordings at PAT in normalized units, i.e., by subtracting the mean value μ and dividing by the standard deviation σ . The following SES activities are depicted in red: (a) on August 9, 2010, (b) on August 10, 2010. The dichotomous representations (blue) lead to the following values: $\kappa_1 = 0.072(4)$, $S = 0.068(6)$, $S_- = 0.088(5)$ and $\kappa_1 = 0.064(5)$, $S = 0.088(8)$, $S_- = 0.062(3)$, for (a) and (b) respectively.

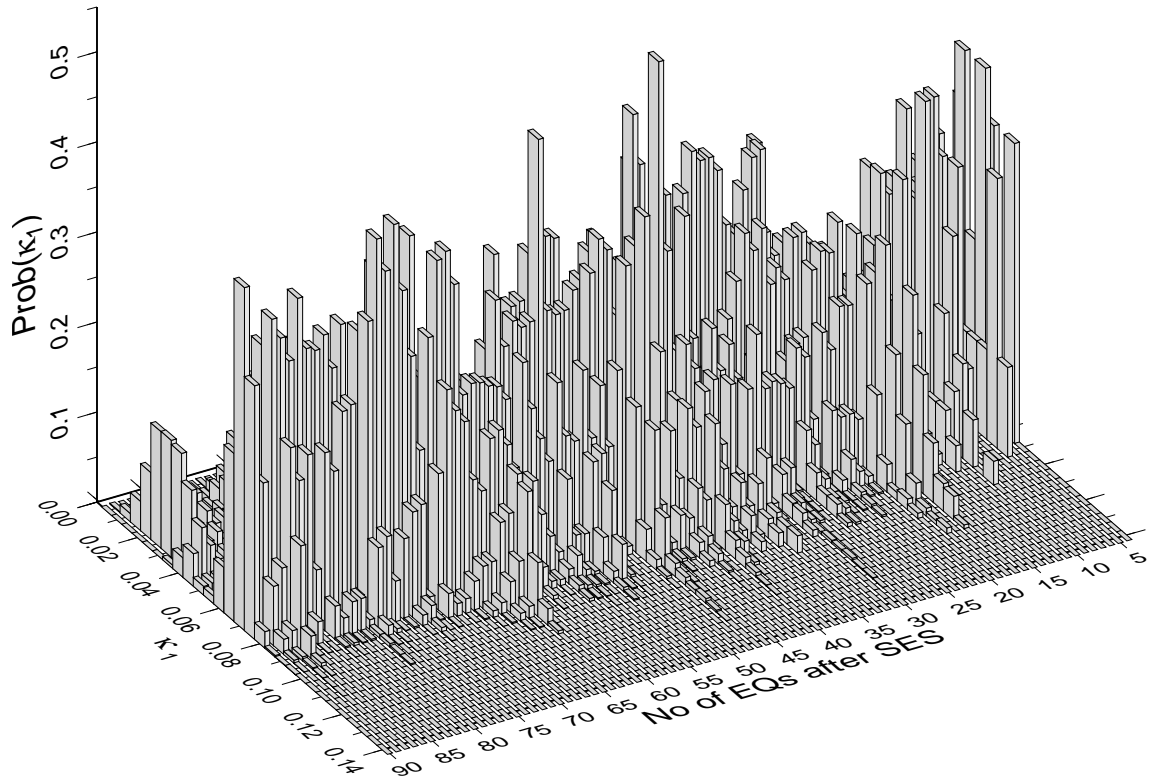


FIG. 6: The probability $\text{Prob}(\kappa_1)$ vs κ_1 upon considering the seismicity ($M_{\text{thres}}=3.1$) after August 10, 2010 until 15:42 UT on September 29, 2010 within the area $N_{37.5}^{38.8} E_{19.8}^{23.3}$.

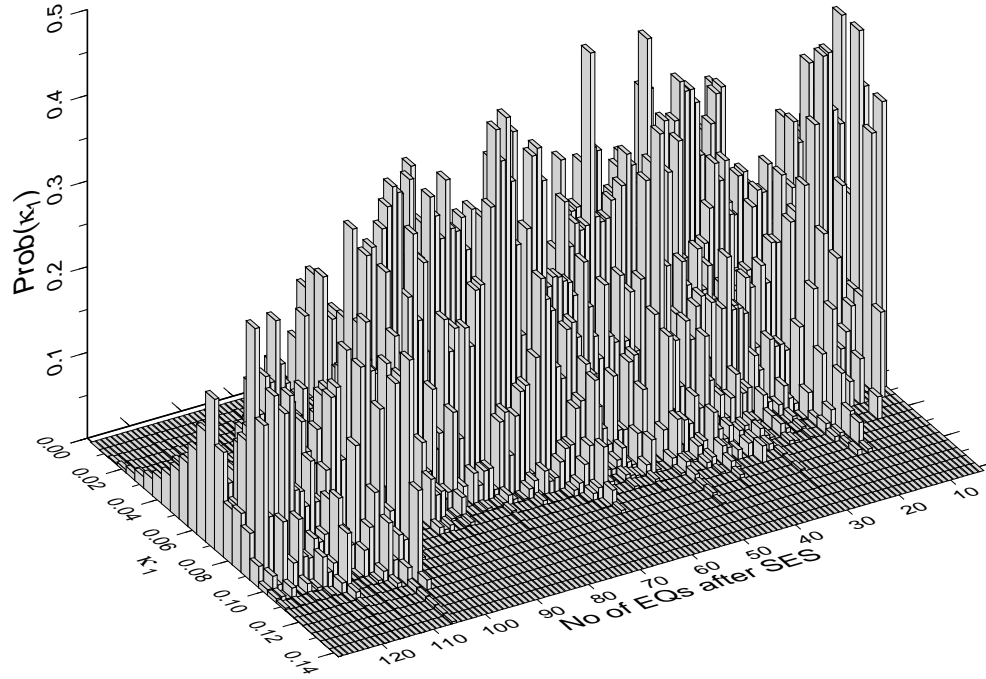


FIG. 7: The probability $\text{Prob}(\kappa_1)$ vs κ_1 upon considering the seismicity ($M_{\text{thres}}=3.1$) after August 10, 2010 until 04:04 UT on October 28, 2010 within the area $N_{37.5}^{38.8} E_{19.8}^{23.3}$.

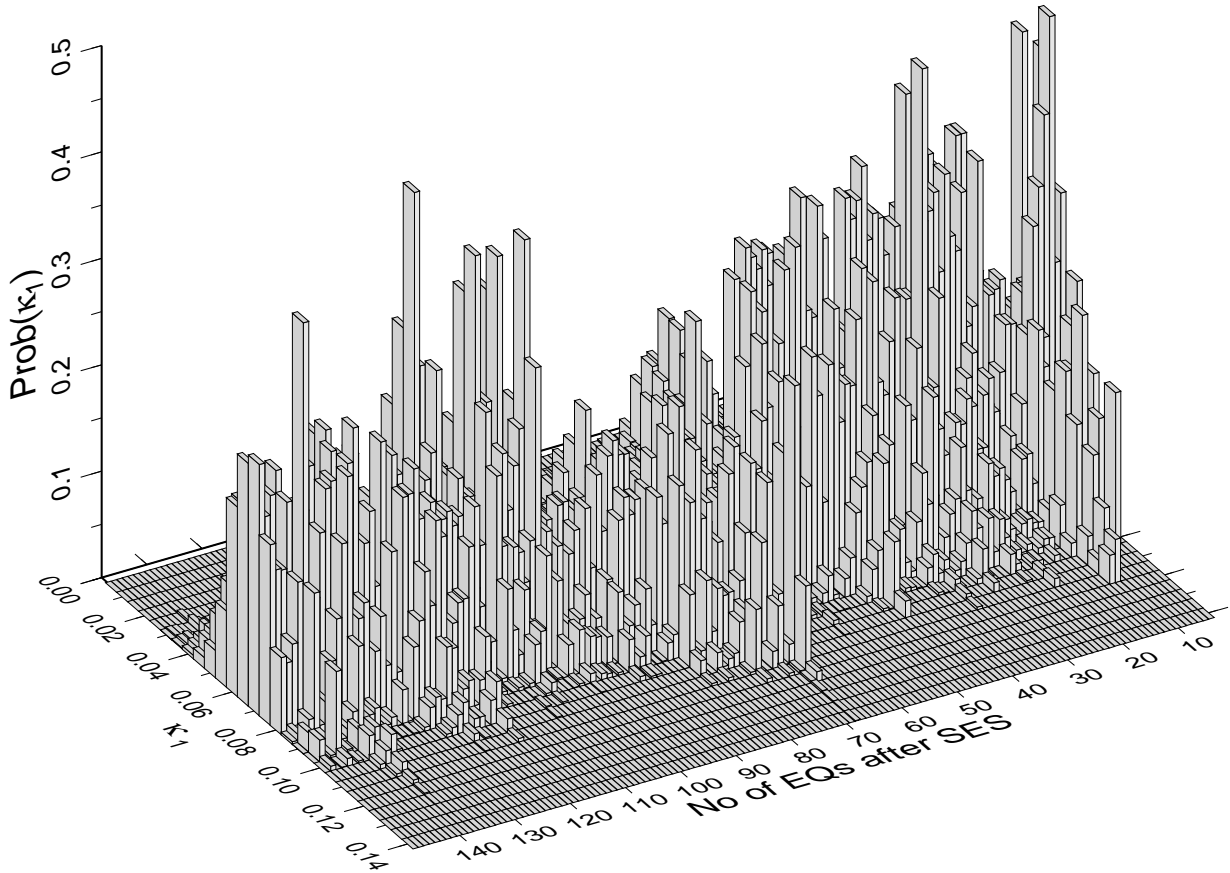


FIG. 8: Continuation of the study of the analysis of seismicity in natural time as described in the Note added on January 27, 2011.

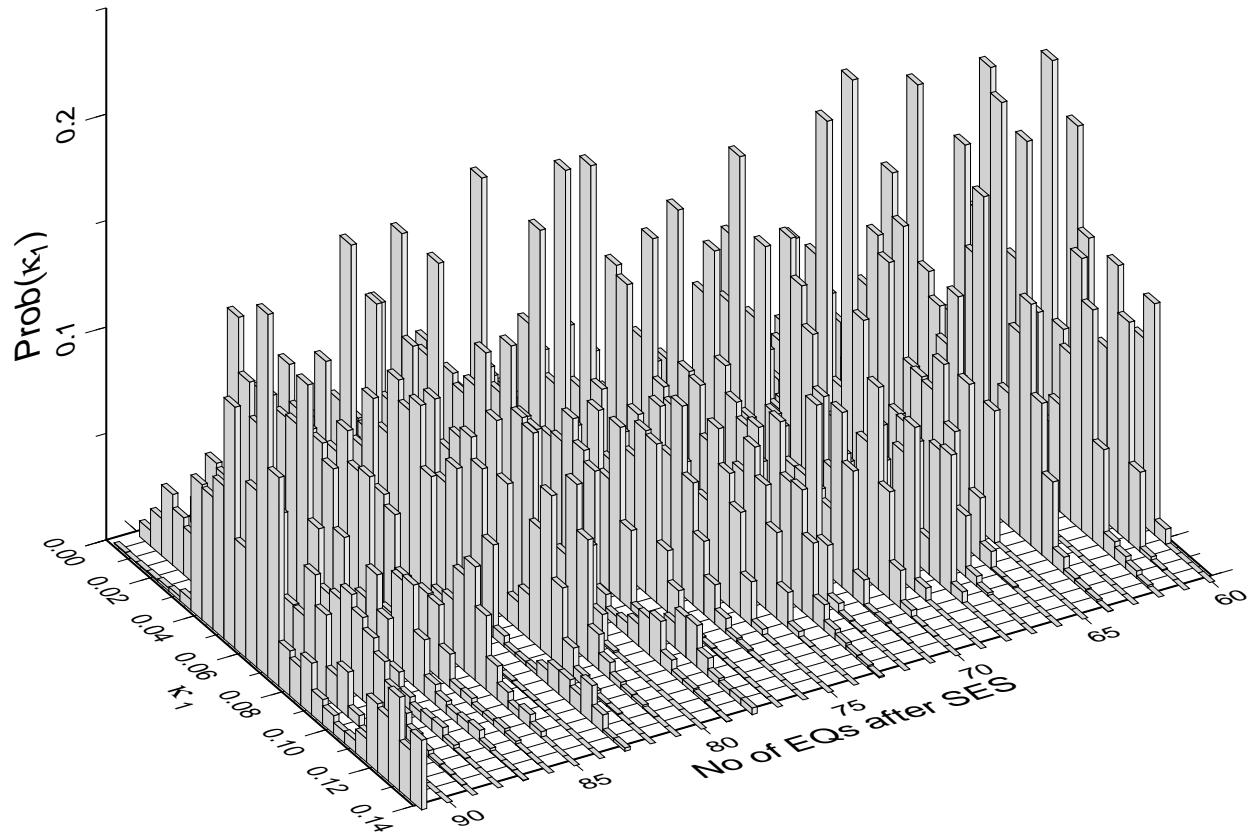


FIG. 9: The probability $\text{Prob}(\kappa_1)$ vs κ_1 upon considering the seismicity ($M_{\text{thres}}=1.9$) after January 22, 2011 until 17:10 UT on March 7, 2011 within the area $N_{37.8}^{38.8} E_{22.5}^{24.1}$.

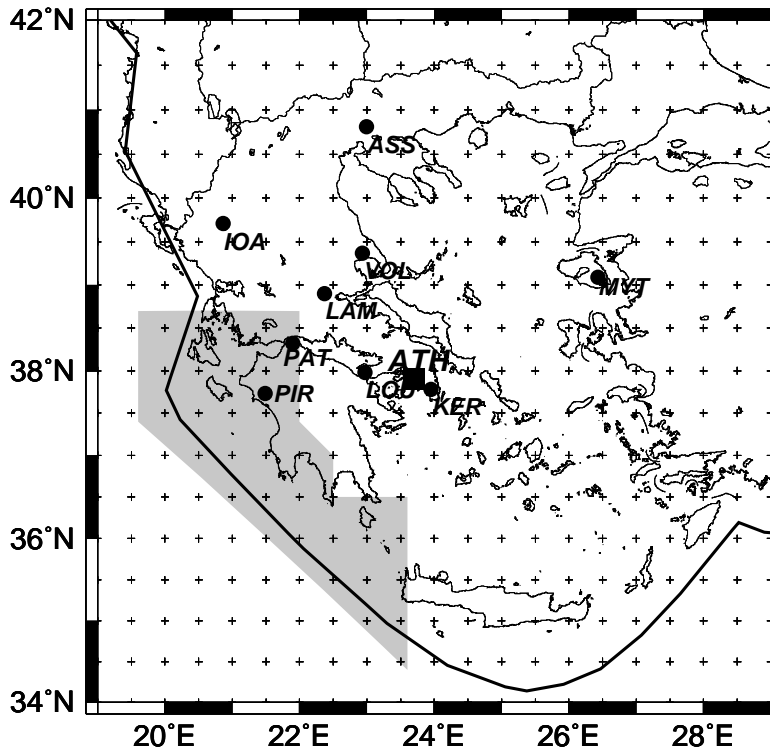


FIG. 10: The area within which the seismicity was considered in natural time analysis after the initiation of the SES activity at PIR lasting from 23:00 UT on May 25 to around 02:00 on May 26, 2011.

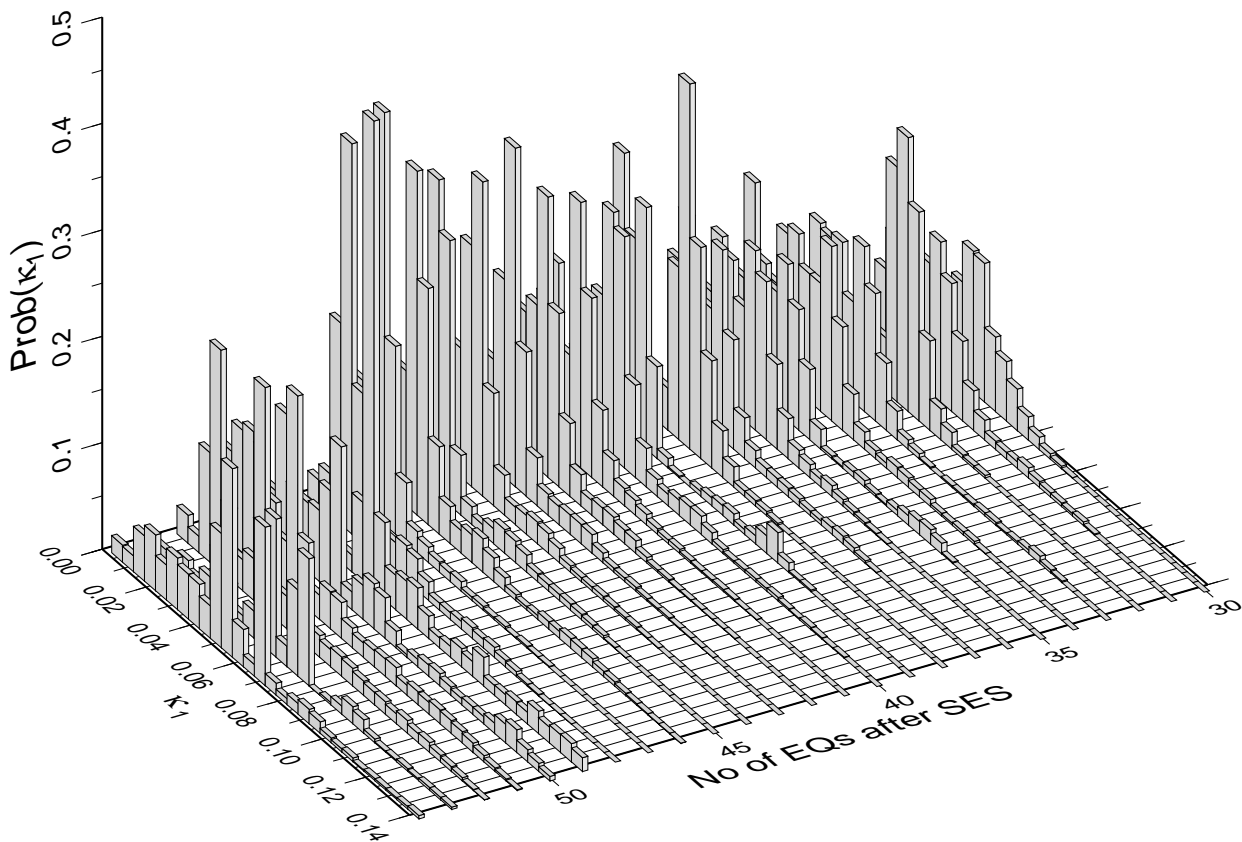


FIG. 11: The probability $\text{Prob}(\kappa_1)$ vs κ_1 upon considering the seismicity ($M_{\text{thres}}=2.7$) after the initiation of the SES activity at PIR (lasting from 23:00 UT on May 25 to around 02:00 on May 26, 2011) until 07:11 UT on June 24, 2011 within the area depicted in Fig. 10. A maximum of $\text{Prob}(\kappa_1)$ at $\kappa_1 \approx 0.070$ is also observed for $M_{\text{thres}}=2.8$ upon the occurrence of the $M_L=4.2$ event at 06:53 UT on June 24, 2011.

- M. Simons, and H. E. Stanley, *Phys. Rev. E* **47**, 3730 (1993).
- ¹⁰ R. N. Mantegna, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C. K. Peng, M. Simons, and H. E. Stanley, *Phys. Rev. Lett.* **73**, 3169 (1994).
- ¹¹ S. Havlin, S. V. Buldyrev, A. L. Goldberger, R. N. Mantegna, S. M. Ossadnik, C. K. Peng, M. Simon, and H. E. Stanley, *Chaos Soliton Fract.* **6**, 171 (1995).
- ¹² C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, R. N. Mantegna, M. Simon, and H. E. Stanley, *Physica A* **221**, 180 (1995).
- ¹³ S. Havlin, S. V. Buldyrev, A. L. Goldberger, R. N. Mantegna, C. K. Peng, M. Simon, and H. E. Stanley, *Fractals* **3**, 269 (1995).
- ¹⁴ R. N. Mantegna, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C.-K. Peng, M. Simons, and H. E. Stanley, *Phys. Rev. Lett.* **76**, 1979 (1996).
- ¹⁵ S. V. Buldyrev, N. V. Dokholyan, A. L. Goldberger, S. Havlin, C. K. Peng, H. E. Stanley, and G. M. Viswanathan, *Physica A* **249**, 430 (1998).
- ¹⁶ H. E. Stanley, S. V. Buldyrev, A. L. Goldberger, S. Havlin, C. K. Peng, and M. Simon, *Physica A* **273**, 1 (1999).
- ¹⁷ C. K. Peng, S. Havlin, and H. E. Stanley, *CHAOS* **5**, 82 (1995).
- ¹⁸ K. K. L. Ho, G. B. Moody, C.-K. Peng, J. E. Mietus, M. G. Larson, D. Levy, and A. L. Goldberger, *Circulation* **96**, 842 (1997).
- ¹⁹ P. Ch. Ivanov, A. Bunde, L. A. N. Amaral, S. Havlin, J. Fritsch-Yelle, R. M. Baevsky, H. E. Stanley, and A. L. Goldberger, *EPL (Europhysics Letters)* **48**, 594 (1999).
- ²⁰ Y. Ashkenazy, P. Ch. Ivanov, S. Havlin, C. K. Peng, Y. Yamamoto, A. L. Goldberger, and H. E. Stanley, *Comput. Cardiol.* **27**, 139 (2000).
- ²¹ Y. Ashkenazy, P. Ch. Ivanov, S. Havlin, C.-K. Peng, A. L. Goldberger, and H. E. Stanley, *Phys. Rev. Lett.* **86**, 1900 (2001).
- ²² P. Ch. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, H. E. Stanley, and Z. R. Struzik, *CHAOS* **11**, 641 (2001).
- ²³ J. W. Kantelhardt, Y. Ashkenazy, P. Ch. Ivanov, A. Bunde, S. Havlin, T. Penzel, J.-H. Peter, and H. E. Stanley, *Phys. Rev. E* **65**, 051908 (2002).
- ²⁴ R. Karasik, N. Sapir, Y. Ashkenazy, P. Ch. Ivanov, I. Dvir, P. Lavie, and S. Havlin, *Phys. Rev. E* **66**, 062902 (2002).
- ²⁵ P. Ch. Ivanov, Z. Chen, K. Hu, and H. E. Stanley, *Physica A* **344**, 685 (2004).
- ²⁶ D. T. Schmitt and P. Ch. Ivanov, *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **293**, R1923 (2007).
- ²⁷ D. T. Schmitt, P. K. Stein, and P. Ch. Ivanov, *IEEE Trans. Biomed. Eng.* **56**, 1564 (2009).
- ²⁸ K. Hu, P. Ch. Ivanov, Z. Chen, M. F. Hilton, H. E. Stanley, and S. A. Shea, *Physica A* **337**, 307 (2004).
- ²⁹ P. Ch. Ivanov, K. Hu, M. F. Hilton, S. A. Shea, and H. E. Stanley, *Proc. Natl. Acad. Sci. USA* **104**, 20702 (2007).

- ³⁰ P. Ch. Ivanov, IEEE Eng. Med. Biol. **26**, 33 (2007).
- ³¹ K. Hu, F. A. J. L. Scheer, P. Ch. Ivanov, R. M. Buijs, and S. A. Shea, Neuroscience **149**, 508 (2007).
- ³² K. Ivanova and M. Ausloos, Physica A **274**, 349 (1999).
- ³³ E. Koscielny-Bunde, A. Bunde, S. Havlin, H. E. Roman, Y. Goldreich, and H.-J. Schellnhuber, Phys. Rev. Lett. **81**, 729 (1998).
- ³⁴ P. Talkner and R. O. Weber, Phys. Rev. E **62**, 150 (2000).
- ³⁵ A. Bunde, S. Havlin, E. Koscielny-Bunde, and H. J. Schellnhuber, Physica A **302**, 255 (2001).
- ³⁶ R. A. Monetti, S. Havlin, and A. Bunde, Physica A **320**, 581 (2003).
- ³⁷ A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).
- ³⁸ Y. H. Liu, P. Cizeau, M. Meyer, C. K. Peng, and H. E. Stanley, Physica A **245**, 437 (1997).
- ³⁹ N. Vandewalle and M. Ausloos, Physica A **246**, 454 (1997).
- ⁴⁰ N. Vandewalle and M. Ausloos, Phys. Rev. E **58**, 6832 (1998).
- ⁴¹ M. Ausloos, N. Vandewalle, P. Boveroux, A. Minguet, and K. Ivanova, Physica A **274**, 229 (1999).
- ⁴² N. Vandewalle, M. Ausloos, and P. Boveroux, Physica A **269**, 170 (1999).
- ⁴³ M. Ausloos, Physica A **285**, 48 (2000).
- ⁴⁴ M. Ausloos and K. Ivanova, Phys. Rev. E **63**, 047201 (2001).
- ⁴⁵ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **67**, 021109 (2003).
- ⁴⁶ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **68**, 031106 (2003).
- ⁴⁷ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, CHAOS **19**, 023114 (2009).
- ⁴⁸ P. Varotsos and K. Alexopoulos, Tectonophysics **110**, 73 (1984).
- ⁴⁹ P. Varotsos and K. Alexopoulos, Tectonophysics **110**, 99 (1984).
- ⁵⁰ P. Varotsos, K. Alexopoulos, K. Nomicos, and M. Lazaridou, Nature (London) **322**, 120 (1986).
- ⁵¹ P. Varotsos, K. Alexopoulos, K. Nomicos, and M. Lazaridou, Tectonophysics **152**, 193 (1988).
- ⁵² P. Varotsos and M. Lazaridou, Tectonophysics **188**, 321 (1991).
- ⁵³ P. Varotsos, K. Alexopoulos, and M. Lazaridou, Tectonophysics **224**, 1 (1993).
- ⁵⁴ P. Varotsos, K. Eftaxias, M. Lazaridou, G. Antonopoulos, J. Makris, and J. Poliyiannakis, Geophys. Res. Lett. **23**, 1449 (1996).
- ⁵⁵ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. Lett. **91**, 148501 (2003).
- ⁵⁶ N. Sarlis and P. Varotsos, J. Geodyn. **33**, 463 (2002).
- ⁵⁷ P. Varotsos, N. Sarlis, and E. Skordas, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci. **77**, 87 (2001).
- ⁵⁸ P. Varotsos, N. Sarlis, and E. Skordas, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci. **77**, 93 (2001).
- ⁵⁹ R. O. Weber and P. Talkner, J. Geophys. Res.-Atmos **106**, 20131 (2001).
- ⁶⁰ J. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and H. E. Stanley, Physica A **316**, 87 (2002).
- ⁶¹ J. F. Muzy, E. Bacry, and A. Arneodo, Int. J. Bifurcation Chaos **4**, 245 (1994).
- ⁶² P. Varotsos and K. Alexopoulos, *Thermodynamics of Point Defects and their Relation with Bulk Properties* (North Holland, Amsterdam, 1986).
- ⁶³ P. Varotsos, *The Physics of Seismic Electric Signals* (TER-RAPUB, Tokyo, 2005).
- ⁶⁴ P. Varotsos, Phys. Rev. B **13**, 938 (1976).
- ⁶⁵ P. Varotsos and K. Alexopoulos, J. Phys. Chem. Solids **39**, 759 (1978).
- ⁶⁶ P. Varotsos and K. Alexopoulos, J. Physics C: Solid State **12**, L761 (1979).
- ⁶⁷ P. Varotsos and K. Alexopoulos, Phys. Rev. B **30**, 7305 (1984).
- ⁶⁸ P. Varotsos, J. Appl. Phys. **101**, 123503 (2007).
- ⁶⁹ P. Varotsos, W. Ludwig, and K. Alexopoulos, Phys. Rev. B **18**, 2683 (1978).
- ⁷⁰ Q. D. Y. Ma, R. P. Bartsch, P. Bernaola-Galván, M. Yoneyama, and P. Ch. Ivanov (2010), arXiv.org:physics.data-an/1001.3641v2.
- ⁷¹ Y. Orihara, M. Kamogawa, T. Nagao, and S. Uyeda, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci. **85**, 435 (2009).
- ⁷² S. Uyeda, M. Kamogawa, and H. Tanaka, J. Geophys. Res. **114**, B02310, doi:10.1029/2007JB005332 (2009).
- ⁷³ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas, Phys. Rev. E **66**, 011902 (2002).
- ⁷⁴ P. A. Varotsos, N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, Phys. Rev. E **70**, 011106 (2004).
- ⁷⁵ P. A. Varotsos, N. V. Sarlis, H. K. Tanaka, and E. S. Skordas, Phys. Rev. E **71**, 032102 (2005).
- ⁷⁶ P. A. Varotsos, N. V. Sarlis, E. S. Skordas, and M. S. Lazaridou, Phys. Rev. E **71**, 011110 (2005).
- ⁷⁷ P. A. Varotsos, N. V. Sarlis, E. S. Skordas, H. K. Tanaka, and M. S. Lazaridou, Phys. Rev. E **73**, 031114 (2006).
- ⁷⁸ P. A. Varotsos, N. V. Sarlis, E. S. Skordas, H. K. Tanaka, and M. S. Lazaridou, Phys. Rev. E **74**, 021123 (2006).
- ⁷⁹ B. Lesche, J. Stat. Phys. **27**, 419 (1982).
- ⁸⁰ B. Lesche, Phys. Rev. E **70**, 017102 (2004).
- ⁸¹ P. A. Varotsos, N. V. Sarlis, and E. S. Skordas (2010), arXiv.org:cond-mat.stat-mech/0904.2465v10.
- ⁸² The SES activity of Fig. 1(b) originated the natural time analysis of the seismicity in the area N(38.0-39.0) E(21.5-23.7) after December 27, 2009, described in the previous reference. This type of analysis suggests that the critical point has been approached, which reveals that an impending mainshock, if any, is imminent (See also Ref.[86]).
- ⁸³ P. Varotsos, M. Lazaridou, K. Eftaxias, G. Antonopoulos, J. Makris, and J. Kopanas, in *The Critical Review of VAN: Earthquake Prediction from Seismic Electric Signals*, edited by Sir J. Lighthill (World Scientific, Singapore, 1996), pp. 29–76.
- ⁸⁴ J. Lighthill, in *The Critical Review of VAN: Earthquake Prediction from Seismic Electric Signals*, edited by Sir J. Lighthill (World Scientific, Singapore, 1996), pp. 373–376.
- ⁸⁵ N. V. Sarlis, E. S. Skordas, M. S. Lazaridou, and P. A. Varotsos, Proc. Jpn. Acad., Ser. B: Phys. Biol. Sci. **84**, 331 (2008).
- ⁸⁶ *Note added on April 16, 2010*: Actually, as proposed in Ref.[82], three days later, i.e., on March 9, 2010, a Ms(ATH)=5.6 earthquake occurred at 38.87N 23.65E inside the expected area. In addition, upon repeating the natural time analysis of seismicity⁸⁵ in the area N(37.65-39.0) E(22.2-24.1) (KER) after February 25, 2010, we again find an approach to *criticality* in the same fashion as described in Ref.[82]. The condition for criticality was found to hold only for magnitude thresholds $M_{thres} = 2.8$ and $M_{thres} = 2.9$ and hence in order to assure magnitude threshold invariance further study is in progress.
- ⁸⁷ *Note added on May 29, 2010*: The aforementioned study in

the area N(37.65-39.0) E(22.2-24.1) finally revealed magnitude threshold invariance of the *criticality* condition upon the occurrence of the ML=3.0 event at 14:32 UT on May 27, 2010, with an epicenter at 38.5N 23.4E. This can be seen when plotting the probability $\text{Prob}(\kappa_1)$ versus κ_1 for $M_{thres} = 2.8, 2.9$ and 3.0 in accordance with the procedure described in Ref.[85] identifying the occurrence time of an impending mainshock.

⁸⁸ *Note added on August 23, 2010:* The previous Note was followed on July 16, 2010 by Ms(ATH)=5.6 earthquake at 39.3N 24.0E, i.e., very close to the northeastern edge of the predicted area. Two other SES activities were recorded at Patras station (PAT) at 09:07 UT on August 9, 2010 and at 06:24 UT on August 10, 2010 depicted in Figs. 5(a) and 5(b), respectively. They resulted in an analysis, in natural time, of the seismicity in the area N(37.6-39.0) E(20.4-23.2) similar to that in Ref.[82].

⁸⁹ *Note added on October 1, 2010:* This is just a continuation of the analysis started, as mentioned in the previous Note, in the area N(37.6-39.0) E(20.4-23.2) which is now extended to N(37.5-38.8) E(19.8-23.3). At 03:53 UT on September 2, 2010 a Ms(ATH)=4.8 earthquake actually occurred with an epicenter at 38.22N 23.17E, i.e., inside the area studied. In addition, upon the occurrence of a $M_L=3.1$ event at 15:42 UT on September 29, 2010 with an epicenter at 37.56N 20.00E, when plotting the probability $\text{Prob}(\kappa_1)$ versus κ_1 , for $M_{thres}=3.1$, we find a sharp maximum at $\kappa_1 \approx 0.070$ depicted in Fig.6, which probably indicates that the critical point has been approached. To ensure that this is a true *criticality* condition, the extent to which a magnitude threshold invariance holds for this result is currently investigated.

⁹⁰ *Note added on October 28, 2010:* Actually, almost one week after the previous Note, i.e., at 19:04 UT on October 9, 2010, a Ms(ATH)=4.8 earthquake occurred with an epicenter at 38.15N 22.72E lying inside the area studied. The continuation of the study related to the SES activities recorded at PAT on August 9 & 10, 2010 (Fig.5) in the same area N(37.5-38.8) E(19.8-23.3), reveals the following: Upon the occurrence of a ML=4.1 earthquake at 10:32 UT on October 23, 2010, with an epicenter at 38.73N 21.99E, a maximum in the plot of $\text{Prob}(\kappa_1)$ versus κ_1 for $M_{thres}=3.3$ is observed at $\kappa_1 = 0.070$. The same was observed for $M_{thres}=3.2$ upon the occurrence of a ML=3.2 earthquake at 00:38 UT on October 24, 2010, with an epicenter at 38.71N 21.99E as well as for $M_{thres}=3.1$ upon the occurrence of a ML=3.6 earthquake at 04:04 UT on October 28, 2010, with an epicenter at 38.36N 22.25E. These findings are strikingly reminiscent of the following behavior: For higher magnitude threshold, the description of the real situation *approaching criticality* becomes less ac-

curate due to “coarse graining”. The plot for the aforementioned maximum observed on October 28, 2010 is depicted in Fig.7.

⁹¹ *Note added on January 27, 2011:* At $\approx 01:00$ UT on January 22, 2011 an electric disturbance started at KER which lasted for around 9 hours. This does not obey the $\Delta V/L$ criterion so that to be classified as SES activity. On the other hand, when analyzing it in natural time we find the following values $\kappa_1 = 0.072(6)$, $S = 0.075(6)$ and $S_- = 0.085(9)$ which conform with the properties of a true precursory signal. A violation however of the $\Delta V/L$ criterion may be due to the inhomogeneities in that area, and in order to resolve the possibility of an approach to criticality, we continued the investigation described in the previous Note as well as in the area $N_{37.8}^{38.8} E_{22.8}^{24.1}$. We find that (for $M_{thres}=3.2$) upon the occurrence of a ML=3.2 event at 09:31 UT on January 24, 2011, a maximum of $\text{Prob}(\kappa_1)$ versus κ_1 is observed at $\kappa_1 = 0.070$, see Fig.8, which possibly indicates the approach of the system to the critical point.

⁹² *Note added on March 8, 2011:* Upon the occurrence of the ML=3.5 event at 17:10 UT on March 7, 2011, the $\text{Prob}(\kappa_1)$ versus κ_1 exhibited a maximum at the value $\kappa_1 = 0.070$ for $M_{thres}=1.1$ to 1.9 (see Fig. 9 for $M_{thres}=1.9$). If the seismic data for such low magnitude thresholds are complete (which is questionable), the present behavior suggests that the critical point is approached (i.e., the event is expected within a few days to around 1 week or so) in the area $N_{37.8}^{38.8} E_{22.5}^{24.1}$. Since $\Delta V/L$ is of the order of 5mV/km, Ms(ATH) depending on the epicentral distance may reach 6 if $r \sim 100$ km.

⁹³ *Note added on June 24, 2011:* Since the seismic data in the previous Note for such low magnitude thresholds were *not* complete, as questioned, the time-window calculation could not be accurate. An earthquake of magnitude $M_s(ATH)=4.4$ actually occurred later on April 22, 2011 with an epicenter at 38.37°N 23.62°E, i.e., around 60-70km from KER. In addition, a SES activity was recorded at PIR lasting for around 3 hours, i.e., from 23:00 UT on May 25 to around 02:00 on May 26, 2011. Thus, a natural time analysis of the subsequent seismicity started in the shaded area depicted in Fig.10 in a similar fashion as in arXiv:0802.3329v4 (publicized on May 29, 2008 before the 6.5 on June 8, 2008) following the same procedure described in the latter publication. This revealed that, upon the occurrence of a ML=2.9 earthquake at 07:11 UT on June 24, 2011 with an epicenter at 37.6°N 21.0°E, the $\text{Prob}(\kappa_1)$ versus κ_1 exhibited a maximum at $\kappa_1 \approx 0.070$ (as shown in Fig. 11), thus pointing to the approach of the system to the critical point.