Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Simple statistics for complex Earthquakes' time distribution

Teimuraz Matcharashvili[1], Takahiro Hatano[2], Tamaz Chelidze[1], and Natalia Zhukova[1]

[1]M. Nodia Institute of Geophysics, Tbilisi State University, Tbilisi, Georgia
[2]Earthquake Research Institute, the University of Tokyo, Tokyo, Japan

Correspondence to: T.Matcharashvili (matcharashvili@gtu.ge)

Abstract. Here we investigated a statistical feature of earthquakes time distribution in southern Californian earthquake catalogue. As a main data analysis tool, we used simple statistical approach based on the calculation of integral deviation times (IDT) from the time distribution of regular markers. The research objective is to define whether the process of earthquakes time distribution approaches to randomness. Effectiveness of the IDT calculation method was tested on the set of simulated color noise data sets with the different extent of regularity. Standard methods of complex data analysis have also been used, such as power spectrum regression, Lempel and Ziv complexity and recurrence quantification analysis as well as multi-scale entropy calculation. After testing the IDT calculation method for simulated model data sets, we have analyzed the variation of the extent of regularity in southern Californian earthquake catalogue. Analysis was carried out for different time periods and at different magnitude thresholds. It was found that the extent of the order in earthquakes time distribution is fluctuating over the catalogue. Particularly, we show that the process of earthquakes' time distribution becomes most random-like in periods of decreased local seismic activity.

1 Introduction

Time distribution of earthquakes remain one of the important questions in nowadays geophysics. At present, the results of theoretical research and the analysis of features of earthquakes' temporal distributions from different seismic regions with different tectonic regimes carried worldwide can be found in (e.g., Matcharashvili et al., 2000; Telesca, 2001, 2012; Corral, 2004; Davidsen and Goltz, 2004; Martinez et al., 2005; Lennartz et al., 2008; Chelidze and Matcharashvili, 2007). Such analyses among others often aims to the assessment of the strength of correlations or the extent of the determinism/regularity in the earthquakes time distribution. One of the main conclusions of such analysis is the understanding that earthquake generation in general does not follow the patterns of random process. Exactly, well established clustering, at least in time (and spatial domains), suggests that earthquakes are not independent completely and that seismicity is characterized by slowly decaying correlations (named long-range correlations): such behavior is commonly exhibited by non-linear dynamical systems far from equilibrium (Peng et al., 1994, 1995). Moreover, it was shown that in the temporal and spatial domains earthquakes' distribution may reveal some features of a low-dimensional, nonlinear structure, while in the energy domain (magnitude distribution) it is

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

Nonlinear Processes in Geophysics
Discussions
Open Access

close to a random-like high dimensional process (Goltz, 1998; Matcharashvili et al., 2000). Moreover, according to present views, the extent of regularity of the seismic process should vary in time and space (Goltz, 1998; Matcharashvili et al., 2000, 2002; Abe and Suzuki, 2004; Chelidze and Matcharashvili, 2007; Iliopoulos et al., 2012).

At the same time, the details of how the extent of randomness (or non-randomness) of seismic process changes over the time and space still remain unclear. In the present work, on the basis of southern California earthquake catalogue, we aimed to be focused on this question and analyzed earthquake time distribution to find where it is closer to randomness.

2 Data and used Methods

Our analysis is based on the southern Californian earthquakes catalogue from 1975 to 2017. In order to discern features of the EQ time distribution we compared it with the time distribution of regular markers according to scheme shown in Fig. 1.

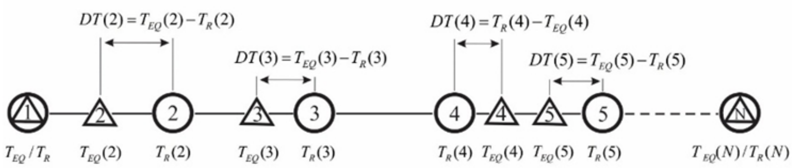

Figure 1. Explanation to used approach. Triangles - time location of original earthquakes ($T_{EQ}(i)$), circles – time location of regular markers ($T_R(i)$). $DT_{(i)}$ denotes the difference between the time of earthquake occurrence ($T_{EQ}(i)$) in catalogue and the time location of regular marker ($T_R(i)$) .

Namely, knowing the time duration of whole period of observation (22167178 minutes, from 01.01.1975 to 23.02. 2017) and the number of earthquakes (34020) with the magnitude above a representative threshold (M2.6), we calculated time step between consecutive regular markers (651.6 min). Then, for each of earthquakes in the catalogue we calculated difference between original event occurrence time and time point of regular marker. We denoted as $DT(i)$ the time interval (delay or deviation time) between occurrence of original earthquake $T_{EQ(i)}$ and corresponding $i-th$ regular marker $T_{R(i)}$. It is clear that original earthquake (EQ_i) may occur prior or after of corresponding regular marker (R_i), so by $DT_{(i)}$ with minus or plus sign we understand earthquakes occurred prior or after regular markers accordingly. Logically, for any random sequence, the sum of the deviation times should approach zero when $n \rightarrow \infty$. Thus, the main assumption is that the integral of deviation times (IDT) or the sum Sum $\sum_{i=1}^{N} DT_{(i)} \rightarrow \infty$ when the time distribution of events is random-like. From this point of view the used approach looks close to Cumulative Sums (Cusum) test, where for a random sequence, the sum of excursions of the random walk should be near zero (Rukhin et.al., 2010).

In order to test whether used approach is sensitive to occurred slight dynamical changes in the seismic process, we started from the analysis of model data sets with a different extent of randomness. Exactly, we used simulated noise data sets of different color with power spectrum function $1/f^{\beta}$, where scale exponent (β) varied from about 0 to 2. These noise data sets have been investigated by several data analysis methods, often used to assess different aspects of changes occurred in dynamical process

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

of interest. Exactly, power spectrum regression, Lempel and Ziv algorithmic complexity calculation, as well as recurrence quantification analysis and Multi-scale entropy calculation method have been used for simulated model data sets. All these popular methods of time series analysis are well described in number of research articles and we will just briefly mention their main principles.

Power spectrum regression exponent calculation enables to elucidate scaling features of data set in the frequency domain. By this method a fractal property is reflected as a power law dependence between the spectral power $(S(f))$ and the frequency (f) by spectral exponent β:

$$S(f) \sim \frac{1}{f^{\beta}} \tag{1}$$

β often is regarded as a measure of the strength of the persistence or anti-persistence in data set. As easily calculated from log-log power spectrum plot, β is related to the type of correlations present in time series (Malamud, 1999; Munoz-Diosdado et al., 2005; Stadnitski, 2012). For example, $\beta = 0$, corresponds to the uncorrelated white noise, and processes with some extent of memory or long-range correlations are characterized by nonzero values of spectral exponents.

Next, we proceeded to the Lempel and Ziv algorithmic complexity (LZC) calculation (Lempel and Ziv, 1976; Aboy et al., 2006; Hu and Gao, 2006), which is a common method for quantification of the extent of order (or randomness) in data sets of different origin. LZC is based on the transformation of analyzed sequence into new symbolic sequence. For this original data are converted into a 0, 1, sequence by comparing to certain threshold value (usually median of the original data set). Once the symbolic sequence is obtained, it is parsed to obtain distinct words, and the words are encoded. Denoting the length of the encoded sequence for those words, the LZ complexity can be defined as

$$C_{LZ} = \frac{L(n)}{n} \tag{2}$$

where $L(n)$ is the length of the encoded sequence and n is the total length of sequence (Hu and Gao, 2006). Parsing methods can be different (Cover, 1991; Hu and Gao, 2006). In this work, we used scheme described in (Hu and Gao, 2006).

Next, in order to further quantify changes in dynamical structure of simulated data sets, we have used recurrence quantification analysis (RQA) approach (Zbilut and Webber, 1992; Webber and Zbilut, 1994; Marwan et al., 2007). RQA is often used for analysis different type of data sets and represents a quantitative extension of Recurrent Plot (RP) construction method. RP, itself, is based on the fact that returns (recurrence) to the certain condition of the system (or state space location) is a fundamental property of any dynamical system with quantifiable extent of determinism in underlying laws (Eckman, 1987). In order to successfully fulfill RQA calculations, at first the phase space trajectory should be reconstructed from the given scalar data sets.. It is important the proximity of points of the phase trajectory to be tested by the condition that the distance between them is less than a specified threshold ε (Eckman, 1987). In this way, we obtain two-dimensional representation of the recurrence features of dynamics, which is embedded in a high-dimensional phase space. Then a small-scale structure of recurrence plots can be quantified (Zbilut and Webber, 1992; Webber and Zbilut, 1994, 2005; Marwan et al., 2007; Webber et al., 2009; Webber and Marwan, 2015). Namely, RQA method enables to quantify features of a distance matrix, of recurrence plot, by means of several measures of complexity. These measures of complexity are based on the quantification of diagonally and vertically oriented lines in the recurrence plot. In this research we calculated one of such measures: the percent determinism ($\%DET$) which is defined as the fraction of recurrence points that form diagonal lines of recurrence plots and shows changes in the extent of determinism in analyzed data sets.

An additional test, which we used to quantify the extent of regularity in modeled data sets, is the composite multi-scale entropy (CMSE) calculation (Wu et al., 2013a). CMSE method represents expansion of multi-scale entropy (MSE) [Costa, et al. 2005] method, which in turn originates from the concept of sample entropy (SampEn) (Richman and Moorman, 2000). SampEn is regarded as an estimator of complexity of data sets for a single time scale. In order to capture the long-term structures in the time series, further Costa (2015) proposed the above mentioned multi-scale entropy (MSE) algorithm, which uses sample entropies (SampEns) of a time series at multiple scales. At the first step of this algorithm, often used in different fields, a coarse-graining procedure is used to derive the representations of a system's dynamics at different time scales; at the next step, the SampEn algorithm is used to quantify the regularity of a coarse-grained time series at each time scale factor. Main problem of MSE is that, for a shorter time series, the variance of the entropy estimator grows very fast as the number of data points is reduced. In order to avoid this problem and reduce the variance of estimated entropy values at large scales, a method of the composite multi-scale entropy (CMSE) calculation was developed by Wu and colleagues (Wu et al., 2013a).

3 Results and discussion

3.1 Analysis of model data sets

First, we needed to ascertain whether calculation of IDT values is sensitive to dynamical changes, occurred in analyzed data sets. As was mentioned above, for this purpose, we decided to generate artificial data sets of the same type, e.g. noises, which are anyway quantifiably different, i.e. represent different type of colored noises. We have started from the analysis of 34020 data length sequences of noise data sets with different long-range correlation features. For clarity we add here that to test the robustness of results, the same analyses were performed on much longer data sets, but here we show results for simulated noise data sets, which are of the same length as original data sets from the used seismic catalogue. The noise data sets have been generated according to concepts described in Kasdin (1995), Milotti (2007) and Beran et al. (2013). As a metrics for these data sets we have used mentioned above power spectrum exponents (β), also referred to as the spectral indexes (Schaefer et al., 2014). Exactly, we have analyzed eight of such data sets having spectral exponents in the range: 0.001, 0.275, 0.545, 0.810, 1.120, 1.387, 1.655, 1.970. Values of β, often used as a metric for the fractal characteristics of data sequences, in our case indicated that simulated noise data sets are different by their dynamical features (Shlesinger, 1987; Schaefer et al., 2014). This means that though all simulated data sets are noises, they are different by extent of correlations (Schaefer et al., 2014). Indeed, the first one, with the $\beta = 0.001$ (Fig.2, a), was closest to random-like white noise and the last one, with the $\beta = 1.932$(Fig.2, b), indicated the features closer to colored noises of red or Brownian type, with detectable dynamical structure. For further analysis of changes in the extent of randomness in considered simulated data sets, we used algorithmic complexity

© Author(s) 2018. CC BY 4.0 License.

(LZC) and recurrence quantification analysis methods as well as testing based on multi-scale entropy (MSE) analysis.

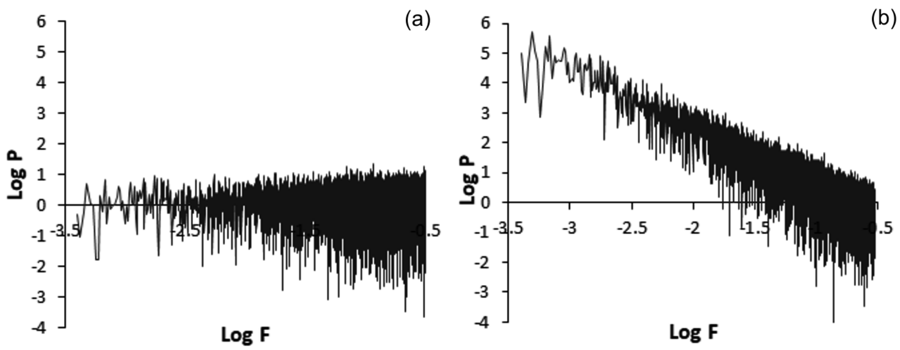

Figure 2. Typical plot of the power spectrum of simulated data sets with different spectral regression, a)$\beta = 0.001$ and b)$\beta = 1.932$

In Fig. 3, we show results of LZC and $\%DET$ calculation. Both methods, though based on different principles, help to answer question, how similar or dissimilar are considered data sets by the extent of randomness. We see that, Lempel and Ziv complexity measure decreases from 0.98 to 0.07, when β of noises increases. It means that the extent of regularity in simulated data sets increases. The same conclusion is drawn from RQA: the percentage of determinism increases from 8.5 to 96.1 when the spectral exponent increases.

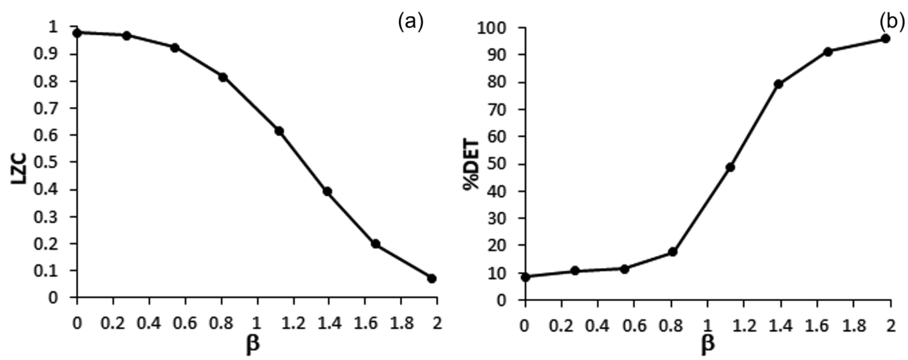

Figure 3. LZC and $\%DET$ values calculated for eight noise data sets with different spectral indexes.

Thus, according to Fig. 3, the extent of regularity in simulated noise sequences clearly increases from white ($\beta = 0$) to Brownian ($\beta = 2$) noise. Thus, the results of LZC and $\%DET$ calculations confirm the result of power spectrum exponents calculations, pointing that considered noise data sets are indeed different by the extent of regularity.
Additional multi-scale, CMSE, analysis also shows (Fig. 4) that the extent of regularity in model noise data sets increases, when they become "more" colored (from $\beta = 0.001$ to $\beta = 1.932$). We see that for small scales (exactly for scale one and partly scale two), noise data sets reveal decrease in the entropy values for simulated data sets, when spectral indexes rise from $\beta = 0.001$ to $\beta = 1.932$. This is logical for simulated data sets, where the extent of order, according to the above analysis, slightly increases.

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.
Nonlinear Processes in Geophysics
Open Access
Discussions

At the same time, at larger scales, the value of entropy for the noise data set with $\beta = 0.001$ monotonically decreases like for coarse-grained white noise time series (Costa, 2015). On the other hand the value of entropy for 1/f type processes with the β values close to pink noises (0.81, 1.12) remained almost constant for all scales. As noticed by Costa (2015), this fact was confirmed in different articles on multi-scale entropy calculation [see e.g. Chou (2012); Wu et al. (2013a, b)]. Costa and coauthors explained this result by the presence of complex structures across multiple scales for 1/ f type of noises.

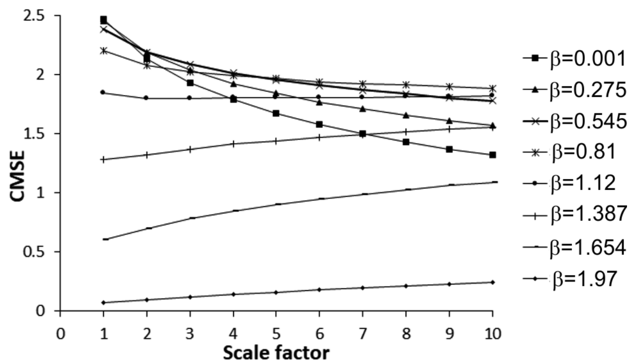

Figure 4. CMSE values versus scale factor for simulated data sequences with different spectral indexes.

From this point of view, in color noises closer to a Brownian type process, the emerging complex dynamical structures become more and more organized. Apparently, these structures are preserved over multiple scales including small ones. This is clearly indicated by the gradual decrease of calculated values of entropy from $\beta = 1.387$ to $\beta = 1.654$ and $\beta = 1.97$ at all considered scales (see Fig. 4).

Thus, CMSE analysis additionally confirms the slight dynamical differences in simulated data sets observed by other above mentioned methods.

Once we had targeted the data sets with quantifiable differences in their dynamical structures we started to test the robustness of IDT calculation for these simulated data sets. For this, we created cumulative sum sequences of these eight data sets and regarded them as models of time occurrences of consecutive events. Exactly, we treated these, 34020 data, sets as described above for time occurrence sequence of real earthquakes and results are presented in the upper curve (circles) in Fig. 5. As we see, absolute values of IDTs calculated for the model data sets indicate stronger deviation from zero, when the extent of order in simulated noise data sets increases (according to the above analysis). It needs to be pointed that comparing to results obtained by used above methods, IDT calculation is even more sensitive to slight dynamical changes occurred in simulated data sets; note more than 4 order difference between sequences with $\beta = 0.001$ and $\beta = 1.97$ for the entire length of data sets (circles in Fig.5).

For the further analysis, it is important to mention that results of above calculations do not practically depend on the length of used data sets. As we see in Fig. 5 (bottom curve), even in the case of short (100 data) sequences IDT calculations enables indicate clearly dynamical changes, occurred in simulated data sets (three orders of difference between sequence with $\beta =$

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

0.001 and $\beta = 1.97$).

Thus, we see that IDT calculation method is effective to detect small changes occurred in analyzed data sets.

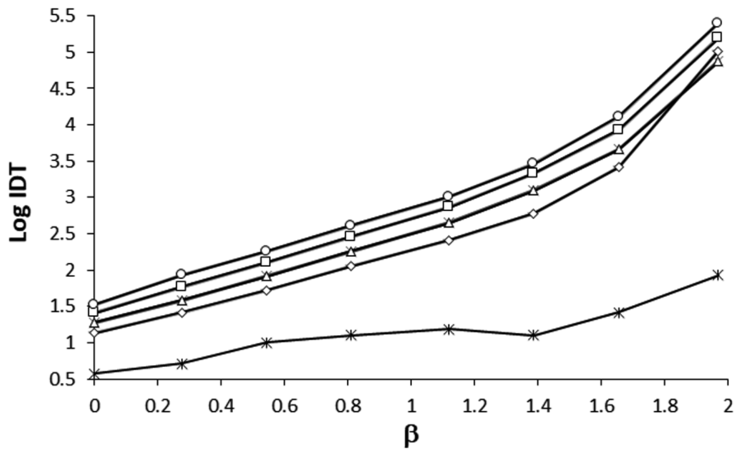

Figure 5. Logarithms of normed to the length of window absolute values of IDT for different length of simulated data sets (circles-34020, squares-20000, diamonds-10000, triangles-5000 data, asterisks-100 data) vs. spectral indexes.

3.2 Analysis of data sets from south California catalogue

In this section we proceeded to the analysis of original data sets drawn from the south California seismic catalogue using IDT calculation approach.

As it was said above, for any random sequence, the sum of the deviation times should approach zero in the infinite length limit. This was clearly shown in previous sections for simulated data sets. In the case of real earthquake catalogues with restricted number of events, we should assume that the integral of deviations times (IDTs) for the more random-like earthquake time distribution will be closer to zero compared to the less random one.

The used seismic catalogue of south California is recognized as the most trustworthy data base for analysis like targeted in this research. According to common practice [see e.g. Christensen et al. (2002); Corral (2004)] we regard the seismic processes in this catalogue as a whole, irrespective of the details of tectonic features, earthquakes location or their classification as main-shocks or aftershocks. For the whole duration of considered catalogue (at the representative threshold M2.6) we found that the number of earthquakes occurred prior and after of corresponding regular markers is different. Exactly, 55% of all earthquakes occurred prior and 45% after the regular time markers (defined in methods section). Obtained IDT value, calculated for the whole southern Californian catalog and for the entire observation period is: -14611458375 minutes (as mentioned above sign minus here denotes the direction of summary deviation along time axis).

Next, we compared this calculated for the original catalogue IDT value with calculated for the set of randomized catalogues values of integral deviation times. In these randomized catalogues original earthquakes' data have been randomly shuffled (i.e. earthquakes' time and space locations as well as magnitudes have been randomly changed). We have generated 150 of such

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

randomized catalogues and for each of them IDT values have been calculated for the whole catalogue time span. It was found that for the whole period of observation, like in the case of original catalogue, in randomized catalogues (at representative threshold M2.6) most of earthquakes (58%) occurred prior to regular markers and consequently prevailed summary deviation times (IDTs) with minus sign. Average IDT, calculated from IDTs of all (150) randomly shuffled catalogues, was also with minus sign (-139755608 min). Thus,, comparing the average of integral deviation times, calculated for the entire length of randomized catalogues with the IDT value of the original catalogue, we see that last one is two orders of magnitude larger. The difference between IDT of the whole original catalogue and the average IDT of randomized catalogues was statistically significant, with Z score =11.2, corresponding to p=0.001 (Bevington and Robinson, 2002; Sales-Pardo et.al., 2007).

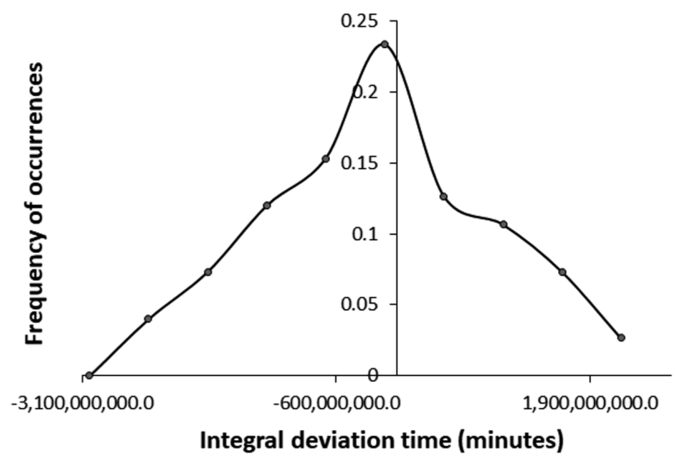

Figure 6. Frequency of occurrence of integral deviation time values calculated for each of 150 randomized catalogues for the whole catalogue length.

It deserves to be pointed that as it follows from Fig. 6, at least 50% of IDTs, calculated for randomized catalogues is even three orders smaller than what we found for the original catalogue (for which integral deviation time equals to -14611458375 minutes). This result undoubtedly shows that in general, the time distribution of earthquakes in south California for the entire considered period should be regarded as a strongly non-random process. The result of this simple statistical analysis is in complete agreement with our earlier results obtained by contemporary nonlinear data analysis methods, as well as with the results of similar analysis reported by other authors using different methodological approaches [see e.g. Goltz (1998); Matcharashvili et al. (2000, 2016); Abe and Suzuki (2004); Telesca (2012); Iliopoulos et al. (2012)]. Next, considering the differences in the number of earthquakes occurred prior and after of corresponding regular markers, we decided to carry out additional calculation of summary deviation times separately for each of these groups. To avoid confusion with IDT, we named summary deviation times for earthquakes occurred prior to markers - SDTp and SDTa – for those occurred after regular markers. Such analysis performed for original and randomized catalogues, for the whole observation period (1975-2017), represented general interest and was also important to test results of the above IDT calculation.

In Fig. 7, we show SDTp and SDTa, sums of deviation times calculated separately for earthquakes occurred prior (2) and after

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

(1) of regular markers. In complete accordance with previous results we see that the sum of deviation times, normed to the number of corresponding earthquakes (i.e. those occurred prior or after of regular markers), are essentially smaller (about 400 times) in the case of randomized catalogues than in the original catalogue.

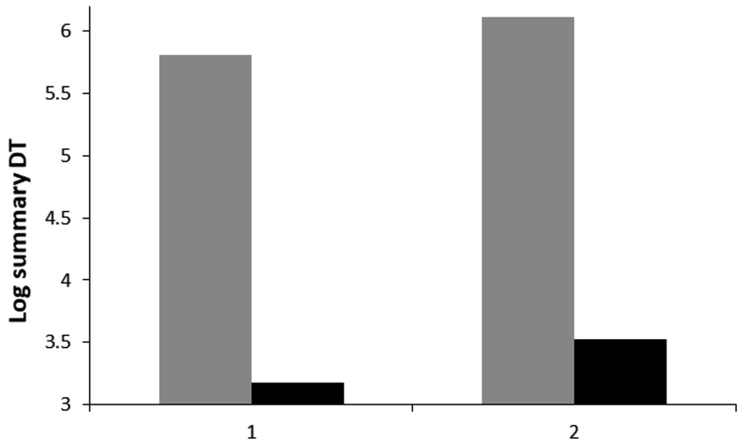

Figure 7. Logarithms of, normed to the number of earthquakes, summary deviation times, calculated for earthquakes occurred after (SDTa, columns 1) and prior (SDTp, columns 2) to regular markers (shown without a sign, i.e. without indication of direction of deviation). Calculations were carried out for the original (grey) and randomized (black) south California earthquake catalogues.

It is interesting that for both original and randomized catalogues the sum of deviation times, normed to the number of corresponding earthquakes is twice larger for events, occurred prior to regular markers than ones occurred after markers (2.022 and 2.207 accordingly). Also, it deserves to be underlined that, for the entire original catalogue, opposite to the summary deviation times, the total amount of seismic energy released by earthquakes which occurred after regular markers is more than three times (3.33) larger comparing to the summary seismic energy released by earthquakes that occurred prior to corresponding regular markers. In the case of randomized catalogues, the total amount of released energy normed to the number of events is almost the same for earthquakes occurred prior or after corresponding regular markers. Thus, according to the results obtained for all the earthquakes in the catalogue as well as ones considered separately prior or after the regular markers, integral deviation times calculated for catalogues with randomly distributed earthquakes is indeed closer to zero, compared with the original catalogue. Based on the presented results we assume that more random the considered process of earthquake time distribution is, the corresponding integral deviation time is closer to zero.

As mentioned above, calculation of the number of occurred earthquakes and the sign to the integral value of deviation time indicates that in the south Californian earthquake catalogue (above representative threshold M2.6), for the entire considered period, prevailed EQs occurred prior to corresponding regular markers (see also the last point in the upper curve of Fig. 8b). At the same time, it was quite expectable that the number of earthquakes occurred prior or after corresponding regular markers as well as the values of the integral deviations times may change depending on the time span of analyzed catalogue.

In order to investigate the character of the time variation of IDT values of South California earthquake catalogue in different

© Author(s) 2018. CC BY 4.0 License.

periods we fulfilled calculation for the expanding time windows. Exactly, we have calculated IDT values starting from the first 100 data (earthquakes), expanding initial window by the consecutive 10 data to the end of catalogue. In Fig. 8, we see that the number and the time location of earthquakes (relative to regular markers), undergoes essential change, when the length of analyzed part of catalogue (analyzed window's) gradually expands to the end of catalogue (in our case from 01.01.1975 to 23.02. 2017).

As it is shown in Fig. 8a, in most of the analyzed windows the majority of earthquakes occurred after regular markers, although there are windows with opposite behavior . So far as in most windows prevail earthquakes which occurred after regular markers, it is not surprising that calculated for consecutive windows integral deviation times, are mostly positive. This is clear from Fig. 8b, (upper curve), where we see windows with negative IDTs too. Thus, the values of IDTs, calculated for extended windows in different periods vary in a wide range, increasing or decreasing and sometimes coming close to zero.

We suppose that when IDT value approaches zero, the dynamical features of originally nonrandom seismic process undergoes qualitative changes and becomes random-like or at least is closer to randomness. In other cases, when IDT value changes over time, but is far from zero, we observe quantitative changes in the extent of regularity of nonrandom earthquakes time distribution. Interestingly, from this point of view, earthquakes' time distribution looks random-like for the relatively quiet (by the amount of released seismic energy) periods, prior or after strongest earthquakes. Indeed, in the lower curve of Fig. 8b, we present cumulative values of seismic energy, calculated for consecutive windows, expanding by 10 events to the end of catalogue. We see that in South California, from 1975 to 2017, the strongest earthquakes never occurred in periods, when IDT curve comes close to zero value or crosses abscissa line.

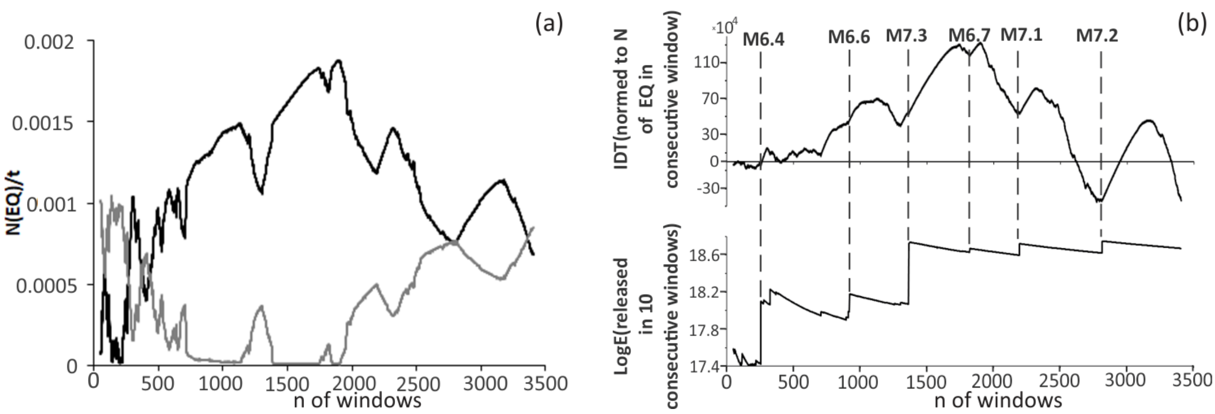

Figure 8. Calculated for extending by 10 consecutive data windows in the South California earthquake catalogue, a) normed to the time duration of window, the number of earthquakes occurred prior (grey) and after (black) regular markers, b) integral deviation times (top) and cumulative seismic energies (bottom).

To avoid misunderstanding because of restricted visibility in Fig. 8b, we point here that M6.4 earthquake has occurred in the window 256, from the start of catalogue, and IDT curve crossed abscissa later, in the window 265, i.e. 100 events later. Results in Fig. 8, also provides interesting knowledge about relation between IDT and the amount of released seismic energy. As we see, three strongest earthquakes in southern Californian earthquake catalogue (1975-2017) occurred on the rising branch of IDT curve close or immediately after local minima. For further clarity, and in order to avoid doubts related to the fixed starting point of the above analysis, we have carried out the same calculation of IDT values for catalogues, which started in 1985, 1990, 1995, 2000 and 2005.

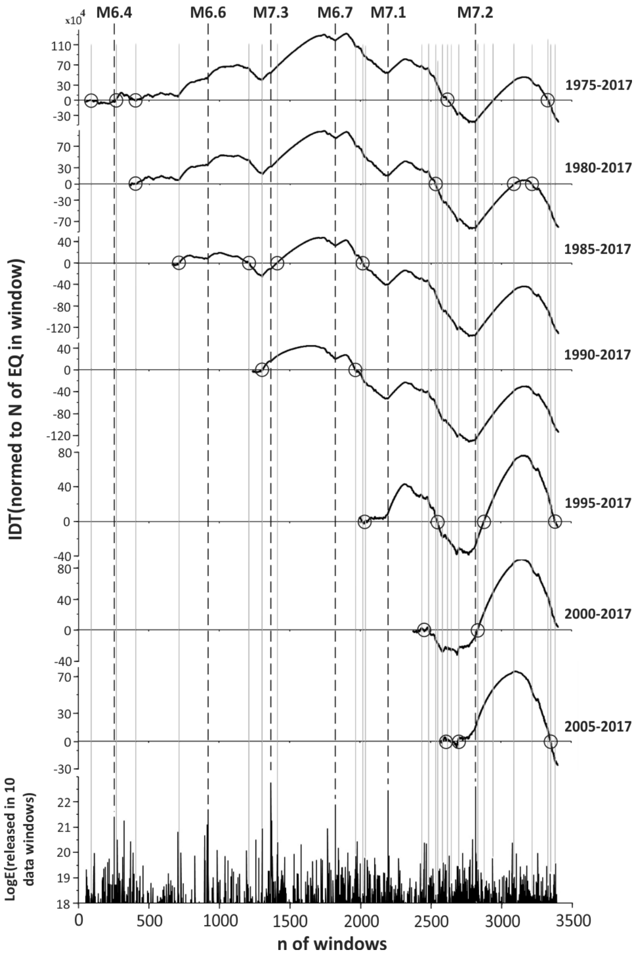

Figure 9. Calculated for the expanding (by consecutive 10 data) windows, integral deviation times (top 7 curves) and the released by consecutive 10 events seismic energies (bottom curve) obtained from the South California earthquakes catalogue (above threshold M2.6). By the grey circles and grey vertical lines we show, where IDT curves cross abscissa axis. Dashed lines show the occurrence of largest earthquakes in the catalogue.

As we see in Fig. 9, analysis carried out on shorter catalogues confirm the result obtained for the entire period of observation (1975-2017) and convinces that the curve of IDT values crosses abscissa at periods of relatively decreased seismic energy release. For better visibility of changes in the process of energy release, in Fig. 9 (bottom), opposite to Fig. 8, we present the seismic energies calculated for consecutive 10 data windows.

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.
Nonlinear Processes in Geophysics
Discussions

Thus we saw that shortening time of the analyzed part of catalogue does not influence obtained results. Next, it was necessary to check how the change of representative threshold will influence obtained results.

This was a very important aspect of our analysis, because there is a well known point of view that the time distribution of large (considered as independent events - coupling between which is exception rather than a rule) and medium-small earthquakes (for which time distribution may be governed or triggered by the interaction between events) is significantly different (Lombardi and Marzocchi, 2007).

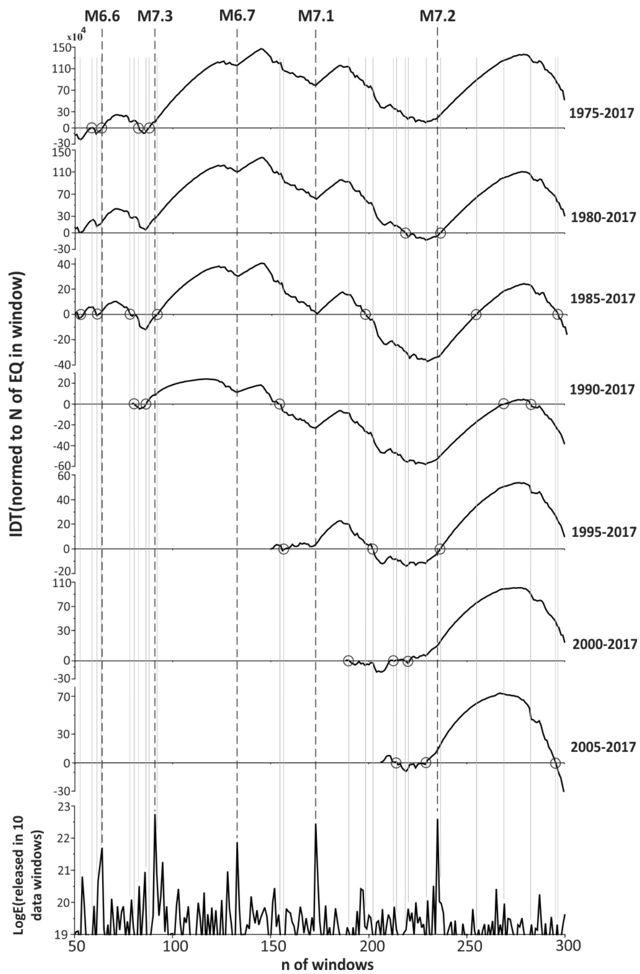

Figure 10. Calculated for the expanding (by consecutive 10 data) windows integral deviation times (top 7 curves) and the released by consecutive 10 events seismic energies (bottom curve) obtained from the South California earthquakes catalogue (above threshold M3.6). By the grey circles and grey vertical lines we show, where IDT curves cross abscissa axis. Dashed lines show the occurrence of largest earthquakes in the catalogue.

To see how the results of integral deviation times analysis may be influenced by considering smaller or stronger earthquakes we have carried out analysis of south California catalogue for earthquakes above M3.6 and M4.6 thresholds. Analysis (see results

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

Nonlinear Processes in Geophysics
Discussions
Open Access

in Figs. 10 and 11) has been accomplished in a manner, similar to the scheme for threshold M2.6, i.e. for the entire available period 1975-2017 and for shorter periods (from 1985, 1990, 1995, 2000, 2005 to 2017).

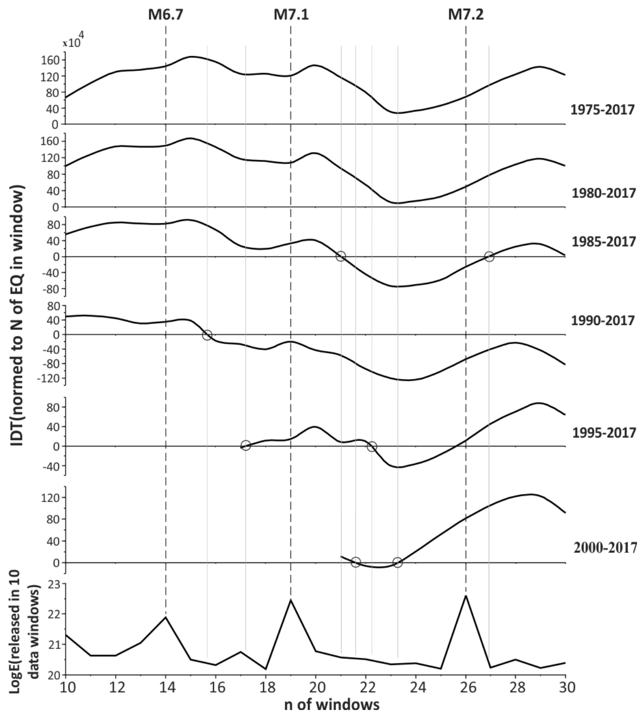

Figure 11. Calculated for the expanding (by consecutive 10 data) windows integral deviation times (here top 6 curves, because for the 7^{th} curve, corresponding to period 2005-2017, the number of events at the threshold M4.6 is small) and the released by consecutive 10 events seismic energies (bottom curve) obtained from the south California earthquakes catalogue (above threshold M4.6). By the grey circles and grey vertical lines we show, where IDT curves cross abscissa axis. Dashed lines show the occurrence of largest earthquakes in the catalogue.

Further analysis by the same scheme for higher (e.g. M5.6) threshold magnitudes was impossible because of the scarce number of large earthquakes in the considered seismic catalogue (just 29 earthquakes above M5.6). At the same time we point out that even for M5.6 representative threshold, for the entire period 1975-2017, the results obtained for two or three existed windows (29 events at windows expanding by 9 or 10 data) agree with the above results showing that a lower IDT value corresponds to period with decreased seismic activity.

Thus, we conclude that the increase of magnitude threshold (Figs. 10 and 11) practically do not change the results found for lower representative threshold. This may be caused by the facts that, increasing representative threshold we still deal with the catalogue in which relatively small and medium size events prevail, while the number of strong events is not enough to carry out correct analysis for such events only. Therefore, conclusions drawn from the analysis for original representative threshold (M2.6) remain correct for the case, when we consider a catalogue with relatively stronger events; thus it seems that there is no principal difference in the character of the contribution of smaller and stronger events to the results of IDT calculation.

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.
Nonlinear Processes in Geophysics
Discussions

After the above analysis, carried out on the expanding windows, we have decided to test obtained results for the sliding windows with fixed number of events. Exactly, in the South Californian earthquake catalogue we have calculated IDT values for non overlapping windows of 100 consecutive events shifted by 100 data. We have used short sliding windows of 100 data for two reasons, a) to have good resolution of changes occurred in the time distribution of earthquakes and because, b) even relatively short, 100 data span, windows also provide good enough discrimination in the IDT calculation as it was shown in Figure 5.

In Fig. 12 (top), we see that for the entire period of analysis there are dozens of IDT values that are not far from one tenth of the standard deviation (σ) from zero (given by black circles in the top figure). Most importantly, among them 8 of IDT values are within of 0.01σ to zero.

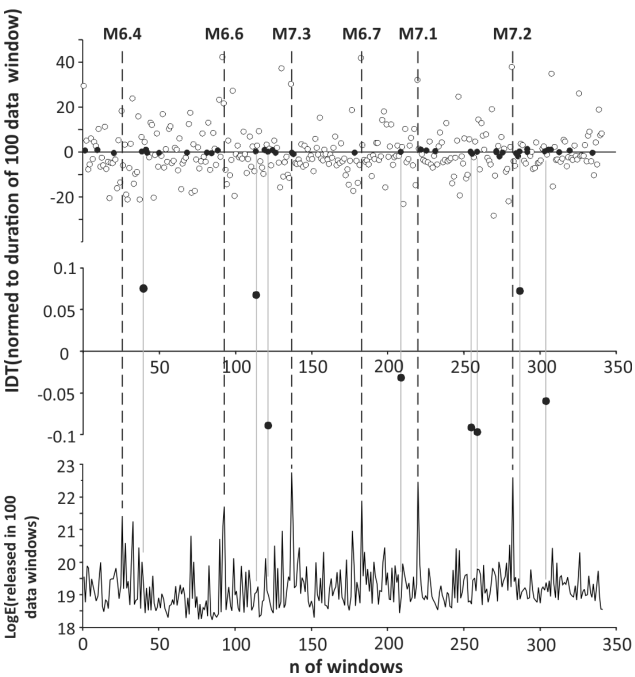

Figure 12. Calculated for the non-overlapping 100 data windows (shifted by 100 data), integral deviation times (circles in the upper and middle curves) and the released seismic energies (bottom curve). IDT values in vicinity of 0.1σ to zero are given by black circles in the top figure. IDT values in vicinity of 0.01σ to zero are given by black circles in the middle figure. By grey lines, we show location of closest to zero IDT values relative to the released seismic energy. Dashed lines show the occurrence of largest earthquakes in the south California catalogue (above threshold M2.6).

These values of IDT, (shown in black in the middle figure), can be regarded practically equal to zero. According to the above consideration, seismic process in the windows with close to zero values of IDT can be regarded as random. If we compare occurrence of these practically zero IDT values with the amount of seismic energy released in consecutive windows (bottom in Fig. 12) it becomes clear that they occur in periods of decreased seismic energy release (grey vertical lines). Similar conclusion is drawn from the analysis of catalogues for earthquakes above M3.6 threshold (Fig.13). Because of the restricted number of strong events in the catalogue, further increase of the threshold magnitude was impossible for the case of 100 data non-overlapping sliding windows.

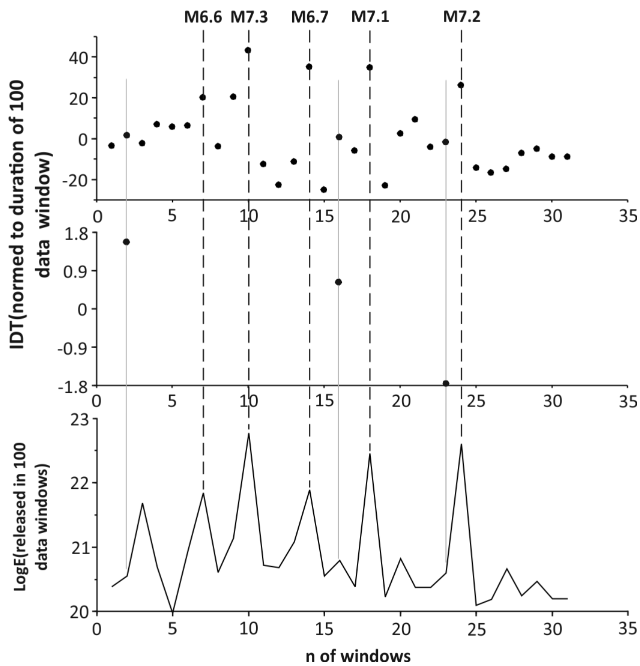

Figure 13. Calculated for the non-overlapping 100 data windows (shifted by 100 data), integral deviation times (circles in the upper and middle curves) and the released seismic energies (bottom curve). IDT values in vicinity of 0.1σ to zero, are given by black circles in the top figure. IDT values in vicinity of 0.01σ to zero are given by black circles in the middle figure. By grey lines, we show location of closest to zero IDT values relative to the released seismic energy. Dashed lines show the occurrence of largest earthquakes in the South California catalogue (above threshold M3.6).

Results obtained for non-overlapping sliding windows of fixed length also confirm the results obtained for expanding windows. Simple statistical approach used here thus shows that extent of randomness in the earthquakes time distribution is changing over time and that it is most random-like at decreased seismic activity. The results of this analysis provide additional indirect arguments in favor of our earlier suggestion that the extent of regularity in the earthquakes time distribution should decrease in seismically quiet periods and increase in the periods of strong earthquakes preparation (Matcharashvili et al., 2011, 2013).

4 Conclusions

We investigated earthquakes time distribution in the Southern Californian earthquake catalogue by the method of calculation of integral deviation times relative to regular time marks. The main goal of research was to quantify, when the time distribution of earthquakes become closer to the random process. Together with IDT calculation, standard methods of complex data analysis

Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-77
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 19 January 2018
© Author(s) 2018. CC BY 4.0 License.

Nonlinear Processes in Geophysics
Discussions
Open Access

such as power spectrum regression, Lempel and Ziv complexity and recurrence quantification analysis, as well as multi-scale entropy calculation have also been used. Analysis was accomplished for different time intervals and for different magnitude thresholds. Based on a simple statistical analysis results, we infer that the temporal distribution of earthquakes in Southern Californian catalogue is most random-like at the periods of decreased local seismic activity.

Competing interests. The authors declare that they have no conflict of interest." This statement should come prior to the acknowledgements.

Acknowledgements. This work was supported by Shota Rustaveli National Science Foundation (SRNSF), grant 217838 “Investigation of dynamics of earthquake’s temporal distribution”.

© Author(s) 2018. CC BY 4.0 License.

References

Abe, S. and Suzuki, N.: Scale-free network of earthquakes, EPL (Europhysics Letters),65(4), 2004.

Aboy, M., Hornero, R., Abásolo, D. and Álvarez, D.: Interpretation of the lempel-ziv complexity measure in the context of biomedical signal analysis, IEEE Trans. Biomed. Eng.,53(11), 2282–2288, 2006.

Beran, J., Feng, Y., Ghosh, S., and Kulik, R.: Long-Memory Processes: Probabilistic Properties and Statistical Methods, Springer, 2013.

Bevington, P., Robinson, D. K.: Data Reduction and Error Analysis for the Physical Sciences.3rd edition, McGraw-Hill, N.Y., 2002.

Chelidze, T. , Matcharashvili, T.: Complexity of seismic process; measuring and applications, A review, Tectonophysics, 431, 49-60, 2007.

Chou, C. M.: Applying Multiscale Entropy to the Complexity Analysis of Rainfall-Runoff Relationships, Entropy 2012, 14, 945-957, doi:10.3390/e14050945, 2012.

Christensen, K., Danon, L., Scanlon, T., Bak, P.: Unified scaling law for earthquakes, Proc. Natl. Acad. Sci. USA 99, 2509, 2002.

Corral, A.: Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys. Rev. Lett. 92, 108501, 2004.

Costa, M.,Goldberger, A. L., Peng, C. K.: Multiscale entropy analysis of biological signals, Phys. Rev. E Stat. Nonlin Soft Matter Phys, 71, 021906, 2005.

Cover T. M., and Thomas J. A.: Elements of Information Theory, New York, Wiley, 1991.

Davidsen, J., Goltz, C.: Are seismic waiting time distributions universal?, Geophys. Res. Lett. 31, L21612, doi:10.1029/2004GL020892, 2004.

Eckmann, J. P., Kamphorst, S., Ruelle, D.: Recurrence plots of dynamical systems, Europhysics Letters, 4,973-977, 1987.

Goltz ,C.: Fractal and Chaotic Properties of Earthquakes, in: Lecture Notes in Earth Sciences, Springer, Berlin, 1998.

Hu, J., Gao, J.: Analysis of Biomedical Signals by the Lempel-Ziv Complexity: the Effect of Finite Data Size, IEEE Transaction on Biomedical Engineering, 53(12), 2006.

Iliopoulos, A. C., Pavlos, G. P., Papadimitriou, P. P., Sfiris, D. S., Athanasiou, M. A., Tsoutsouras, V. G.: Chaos, selforganized criticality, intermittent turbulence and nonextensivity revealed from seismogenesis in north Aegean area, Int. J. Bifurcation Chaos, 22, 9, doi:1250224,10.1142/S0218127412502240, 2012.

Kasdin, N. J.: Discrete simulation of colored noise and stochastic processes and 1/f power law noise generation, Proceedings of the IEEE, 83(5),802–827, 1995.

Lempel A. and Ziv, J.: On the complexity of finite sequences, IEEE Trans. Inf. Theory, IT-22(1), 75–81, 1976.

Lennartz, S., Livina, V.N., Bunde, A. and Havlin, S.: Long-term memory in earthquakes and the distribution of interoccurrence times, Europhys. Lett. 81, 69001, doi:10.1209/0295-5075/81/69001, 2008.

Lombardi, A. M., and Marzocchi, W.: Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes, J. Geophys. Res., 112, B02303, doi:10.1029/2006JB004568, 2007.

Malamud, B. D., Turcotte, D. L.: Self-affine time series I: Generation and analyses, Adv. Geophysics. 40, 1-90, 1999.

Martınez, M. D., Lana X., Posadas, A. M. and Pujades, L.: Statistical distribution of elapsed times and distances of seismic events: the case of the Southern Spain seismic catalogue, Nonlinear Processes in Geophysics, 12, 235–244, 2005.

Marwan, N., Romano, M. C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex system, Phys. Rep. 438, 237–329, 2007.

Matcharashvili, T., Chelidze, T., Javakhishvili, Z.: Nonlinear analysis of magnitude and interevent time interval sequences for earthquakes of Caucasian region, Nonlinear Processes in Geophysics,7, 9–19, 2000.

Matcharashvili, T., Chelidze, T., Javakhishvili Z., Ghlonti, E.: Detecting differences in dynamics of small earthquakes temporal distribution before and after large events, Computers and Geosciences, 28(5), 693-700, 2002.

Matcharashvili, T., Chelidze, T., Javakhishvili, Z., Jorjiashvili, N. and FraPaleo, U.: Non-extensive statistical analysis of seismicity in the area of Javakheti, Georgia, Computers and Geosciences, 37(10), 1627-1632, 2011.

Matcharashvili, T., Telesca, L., Chelidze, T., Javakhishvili, Z., Zhukova, N.: Analysis of temporal variation of earthquake occurrences in Caucasus from 1960 to 2011, Tectonophysics, 608, 857-865, 2013.

Matcharashvili, T., Chelidze, T., Javakhishvili, Z., Zhukova, N.: Variation of the scaling characteristics of temporal and spatial distribution of earthquakes in Caucasus, Physica A, 449, 136-144, 2016.

Milotti, E.: New version of PLNoise: a package for exact numerical simulation of power-law noises, Computer Physics Communications, 177(4), 391-398, 2007.

Munoz-Diosdado, A., Guzman-Vargas, L., Rairez-Rojas, A., Del Rio-Correa, J. L., Angulo-Brown, F.: Some cases of crossover behavior in heart interbeat and electoseismic time series, Fractals ,13(4), 253-263, 2005.

Peng, C.K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L.: Mosaic organization of DNA nucleotides, Phys. Rev. E, 49, 1685–1689, doi:10.1103/PhysRevE.49, 1685, 1994.

Peng, C.K., Havlin, S., Stanley, H. E., and Goldberger, A. L: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, Chaos, 5, 82–87, doi:10.1063/1.166141, 1995.

Richman, J. S., Moorman, J. R.: Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol. Heart Circul. Physiol., 278, H2039–H2049, 2000.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., Vo, S.: The Statistical Tests Suite for Random and Pseudorandom Number Generators for Cryptographic application, NIST Special Publication 800-22revla, 2010.

Sales-Pardo, M., Guimer ,R., Moreira, A. A., NunesAmaral, L. A.: Extracting the hierarchical organization of complex systems, Proceedings of the National Academy of Sciences of the United States of America, 104(39), 14224–15229, 2007.

Schaefer, A., Brach, J.S., Perera, S., Sejdic, E.: A comparative analysis of spectral exponent estimation techniques for 1/f processes with applications to the analysis of stride interval time series, J Neurosci Methods, 222, 118–130, 2014.

Shlesinger, M. F.: Fractal time and 1/f noise in complex systems. Annals of the New York Academy of Sciences, 504(1), 214–228, 1987.

Stadnitski T.: Measuring fractality, Front. Physio, 3(127), doi:10.3389/fphys.2012.00127, 2012.

Telesca, L., Cuomo, V., Lapenna, V., Macchiato, M.: Identifying space–time clus-tering properties of the 1983–1997 Irpinia-Basilicata (Southern Italy) seismicity, Tectonophysics, 330, 93–102, 2001.

Telesca, L., Matcharashvili, T., Chelidze, T.: Investigation of the temporal fluctuations of the 1960–2010 seismicity of Caucasus, Nat. Hazard. Earth Syst. Sci. 12, 1905–1909, 2012.

Webber, C. L. and Zbilut, J.: Dynamical assessment of physiological systems and states using recurrence plot strategies, J. Appl. Physiol., 76, 965–973, 1994.

Webber, C. L. and Zbilut, J. P.: Recurrence quantification analysis of nonlinear dynamical systems, In Tutorials in contemporary nonlinear methods for the behavioral sciences (ed. Riley M. A. and Van Orden G. C.), National Science Foundation Program in Perception, Action and Cognition, 26-94, 2005.

Webber, C. L., Marwan, N., Facchini, A., Giuliani, A.: Simpler methods do it better: Success of Recurrence Quantification Analysis as a general purpose data analysis tool, Physics Letters A, 373, 3753–3756, 2009.

© Author(s) 2018. CC BY 4.0 License.

Webber, C. L., Marwan, N. (Editors): Recurrence Quantification Analysis Theory and Best Practices, Springer, 2015.

Wu, S. D., Wu, C. W., Lin, S. G., Wang, C. C., Lee, K. Y.: Time Series Analysis Using Composite Multiscale Entropy, Entropy 2013a, 15, 1069-1084, 2013a.

Wu, S. D., Wu, C. W., Lee, K. Y., Lin, S. G.: Modified multiscale entropy for short-term time series analysis, Physica A, 392, 5865–5873, 2013b.

Zbilut, J. P., and Webber, C. L.: Embeddings and delays as derived from quantification of recurrence plots, Phys. Lett. A, 171, 199–203, 1992.