J Seismol
https://doi.org/10.1007/s10950-019-09820-7

ORIGINAL ARTICLE

Monitoring earthquake network measures between main shocks in Greece

D. Chorozoglou · E. Papadimitriou

Received: 23 November 2018 / Accepted: 31 January 2019

Abstract The monitoring of complex earthquake networks that are formed from Greek seismicity based on the evolution of their measures, such as degree centrality, characteristic path length, and clustering coefficient is performed, aiming to identify whether and when these networks exhibit distinct evolution between main shocks. As network nodes, the 17 seismic zones in which the study area was appropriately divided are considered and their connections are given by the significant correlation computed on the time series of each node seismic activity. The data are taken from a seismic catalog comprising crustal earthquakes (focal depth less than 50 km) of magnitude $M \geq 3.0$ that occurred in the territory of Greece in between 1999 and 2017. During this period twenty one (21) main shocks of $M \geq 6.0$ occurred, but for only six (6) of them, the interevent time since the last one was adequate for the network measures calculation. The earthquake networks are formed on sliding windows of different durations for monitoring the network measures variation. To assess whether the values of network measures are statistically significant, the construction of randomized networks is required, and the same network measures are calculated for comparison purposes. The monitoring of network measures revealed that their values were found statistically significantly different from the corresponding values of the randomized networks shortly before the main shocks.

Keywords Network measures · Randomized networks · Time series · Main shocks · Seismicity of Greece

D. Chorozoglou (✉) · E. Papadimitriou
Department of Geophysics, School of Geology, Aristotle University of Thessaloniki, GR54124 Thessaloniki, Greece
e-mail: chorozod@geo.auth.gr

E. Papadimitriou
e-mail: ritsa@geo.auth.gr

1 Introduction

The investigation of the complex seismicity behavior constitutes a major scientific challenge and an indispensable component in improving our knowledge concerning seismogenesis and earthquake forecasting. The earthquake distribution in space and time is considerably complex, forming temporal clusters on faults that remain active for some time and then becoming quiescent for long. Episodic occurrence has been often documented worldwide and in Greece (Papadimitriou and Karakostas 2005), clustered in time at certain seismic zones (Papadimitriou 2002; Scholz 2010) and migrating (Stein et al. 1997; Papadimitriou et al. 2004). Research efforts include both deterministic and stochastic modeling for approaching and deciphering the earthquake process, including investigation of a complex system of interacting faults through their stress field, and statistical tools to examine the whole system that exhibits a more complex behavior than expected from each individual fault.

From the early scientific steps in Seismology about 130 years ago, the statistical properties of seismicity

Published online: 20 February 2019

formed the basis for studying the complex physical mechanisms leading to strong earthquake occurrence (Utsu 2002). The empirical statistical scaling laws such as the Gutenberg-Richter (Gutenberg and Richter 1944) and the Omori law (Omori 1894; Utsu 1961; Utsu et al. 1995) allowed us to understand some fields of the complexity of the seismological issues. The Gutenberg-Richter (GR) law establishes that the frequency of occurrence of earthquakes follows a power law in respect to the energy released ($log n = a - bM$), whereas the Omori law presents the decaying rate of aftershocks after in strong earthquakes. The slope b of the GR law relates to the spatial fractal dimension, D, of earthquake epicenters or faults, through the simple formula $D = 3b/c = 2b$ (for a typical value of $c = 1.5$, Aki 1981).

Seismogenesis alike other geophysical processes are neither random nor deterministic, and therefore complexity analysis is engaged for revealing the structures and processes that lie between randomness and determinism (Chelidze 2017). Routine statistical analysis is inadequate to unfold certain nonlinear spatio-temporal structures but can be seen by application of modern tools. One such approach for investigating the spatial and temporal complexity of seismicity is through the construction of earthquake networks. Graph theory provides a framework to investigate the structure and dynamics of a complex system. The network nodes are usually assumed to represent distinct subsystems and the connections represent the interactions among them. In recent years, network theory was successfully applied in different disciplines, such as Economics (Billio et al. 2012; Heiberger 2014; Fiedor 2014; Papana et al. 2017), Biology (Jeong et al. 2001; Girvan and Newman 2002; Wang and Chen 2003), Climatology (Donges et al. 2009; Bialonski et al. 2010; Palus et al. 2011), Neuroscience (Rubinov and Sporns 2010; Bullmore et al. 2016; Fornito et al. 2016; Kugiumtzis et al. 2017), Physiology (Porta and Faes 2016), Transportation (Wang et al. 2017, Aydin et al. 2017), and virus diffusion (Zhang and Gan 2018). The complex network analysis was introduced in Seismology by Abe and Suzuki (2004a) to study seismicity as a spatiotemporal complex system. In the earthquake network approach, a seismic region can be correlated with another far away, which is coherent with observations of "remotely triggered seismicity" (Hill et al. 1993). Such interpretation is also confirmed with the hypothesis that seismicity is self-organized phenomenon (Bak and Tang 1989). The complex network theory provided more insight and perspective in the seismicity patterns Abe and Suzuki (2007) in California to reveal that the values of the clustering coefficient remain stationary before main shocks and suddenly jumped up at the main shocks, Baiesi and Paczuski (2004, 2005) for Southern California and recently by Daskalaki et al. (2016) for Italian area.

The study of complex networks has revealed interesting non-trivial properties that signify their specific structures, such as the small-world (Watts and Strogatz 1998) and the scale-free (Albert and Barabasi 2002) property, which were found to indicate their underlying organization principles. Abe and Suzuki (2004a, 2006) and Abe et al. (2011) have proved the universality of the small-world property for earthquake networks constructed for four different study areas (California, Japan, Iran and Chile). This property was also studied by Jimenez et al. (2008) for California, Baek et al. (2011) for Korean peninsula, Daskalaki et al. (2014) and Chorozoglou et al. (2018) for the Greek area, and León et al. (2018) for the Colombian area where the small-world structure was confirmed. The scale-free property was firstly investigated by Abe and Suzuki (2004b) and by Baiesi and Paczuski (2004) for the southern California and lastly by Pastén et al. (2016) for Chile and Janer et al. (2017) for the Philippines and southern California.

The various statistical approaches and the non-trivial network properties were adequately investigated as mentioned above. Hence, examination of some basic network measures, such as degree centrality, global efficiency, and clustering coefficient, will unveil properties of spatio-temporal seismicity structure and give a new perspective for the seismic hazard assessment. The scope of this work is the monitoring of evolution of nine (9) network measures between six (6) main shocks with $M \geq 6.0$ that occurred during 1999–2017 in the Greek area and for which an adequate time for the network measures to be statistically robustly estimated has elapsed since the last main shock. The target is the identification of potential patterns in the distinct evolution of the earthquake network structure. For the distinct evolution, we are not interested in the value of network measure decreases or increases abruptly for a certain time interval between main shocks (Abe and Suzuki 2009; Daskalaki et al. 2016; Chorozoglou et al. 2017) but whether it is statistically significant even if it does not get too low or too high values. In our analysis, the nodes are represented by the seismic zones into which the study area was divided and the connections between them are given by the long-term significant correlation of earthquake activity, based on creation of time series, of the corresponding nodes. Thus, the introduction of the connections, are not based on successive earthquakes as in Abe and Suzuki (2004a), but in the times between any possible pair, since it has been shown (Omori 1894, Corral 2004, Livina et al. 2005, Lippiello et al. 2008, Lennartz et al. 2008) that the distribution of these times is not Poissonian. As a result, the successive earthquakes may not be the result of uncorrelated independent probability, but dependent on the long-term history for each seismic area (node). The investigation whether the values of network measures are statistically significant, i.e., whether a distinct evolution appears, requires the formation of randomized networks for comparison with the original (observed) ones. The standard methods of Maslov and Sneppen (2002) and Newman (2010) which are based on the randomization of the original network connections, the method of Erdős and Rényi (1959) model that the probability of connections is the same as in the original networks, and the method that applies the randomization to the time series (Chorozoglou and Kugiumtzis 2014, 2018) are used for the construction of randomized networks. The evolution of network measures revealed that the values for the original earthquake networks are differ, i.e., there is statistically significance and distinct evolution, from the corresponding values for randomized networks in the last time interval before the main shocks.

We first present in "Methodology" the construction of the earthquake networks, both the original and the randomized ones, as well as the network measures which are used and the significance test that is required for comparing the network performance. In "Application," the results for the evolution of network measures which are computed on each sliding window for the six (6) pairs of main shocks are presented and discussed. The concluding remarks are given in "Discussion and concluding remarks."

2 Methodology

In the following, we present firstly the construction of the earthquake networks and then the network measures followed by the methods of network randomization and lastly the significance test to check whether the values of network measures are statistically significant.

2.1 Building the earthquake networks

For the earthquake networks construction, we firstly assigned as nodes in the study area the K seismic zones that were defined on the basis of the homogeneity of their seismotectonic properties (faulting type, seismic moment release), and maximum observed magnitude (Papazachos et al. 1998; Vamvakaris et al. 2016).

The earthquake networks connections are given by the significant correlation of the seismic activity in two nodes. The observed variables of the time series are either the cumulative of seismic moment M_0 or the number of earthquakes (Jimenez et al. 2008) within each seismic zone. The scalar seismic moment M_0 of an earthquake of magnitude M is given by $logM_0 = 1.5M + 16.01$ (Kanamori and Anderson 1975). The connections are given by the linear zero-lag cross-correlation, which is actually the Pearson correlation coefficient. Suppose we have a set of K random variables observed at n time points $\{X_{1,\ t}, \dots, X_{K,\ t}\}$, for $t = 1, \dots, n$, where each variable is the cumulative seismic moment M_0 or the number of earthquakes at each node. For two variables $X = X_i$ and $Y = Y_j$, $i, j \in \{1, .., K\}$, the Pearson correlation coefficient is defined as.

$$r_{X,Y} = \frac{S_{XY}}{\sqrt{S_X^2 S_Y^2}}, \tag{1}$$

where $S_{XY} = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \overline{x})(y_t - \overline{y})$ is the sample covariance of (X, Y), $S_X^2 = \frac{1}{n-1}\sum_{t=1}^{n}(x_t - \overline{x})^2$ and $S_Y^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \overline{y})^2$ are the sample variances of X and Y, and $\overline{x}$ and $\overline{y}$ are the sample means of X and Y, respectively.

The symmetric matrix of weighted connections $W = \{w_{ij}\}$, $i, j \in \{1, \dots, K\}$, is simply formed by the absolute value of $r_{i,j} = r_{X_i,X_j} = r_{X,Y}$. For simple connections, the adjacency matrix $A = \{a_{ij}\}$, $i, j \in \{1, \dots, K\}$, acquires components assigned to one (1) when the zero-cross correlation is found significant, and zero (0), otherwise. The decision for the statistical significance is given by the significance test for the correlation coefficient (Horvath 2011). The null hypothesis is $H_0 : \rho_{X,\ Y} = 0$, where $\rho_{X,\ Y}$ is the true correlation Pearson coefficient. The sample cross-correlation coefficient $r_{X,Y}$ is transformed to the test statistic

$$t = r_{X,Y}\sqrt{\frac{n-2}{1-r_{X,Y}^2}} \quad (2)$$

that follows Student distribution with $n-2$ degrees of freedom. A connection is significant and equals to one if the p value of the test is less than a given significance level α (here $\alpha = 0.05$).

2.2 Network measures

The graph $G = (K, E)$ is defined by the set of K nodes and the set of E connections among them. The connections here are undirected and weighted or binary. For any two nodes, i and j, the distance between them, denoted as d_{ij} for simple connections and d_{ij}^w for weighted connections, is defined as the length of the shortest path from i to j if the nodes are connected, otherwise $d_{ij} = \infty$ or $d_{ij}^w = \infty$. For the monitoring of earthquake network structure between the examined main shocks, nine (9) different network measures are considered and computed, either on the adjacency matrix A or on the weighted matrix W, the mathematical expressions of which are given in Table 1 and are briefly described below.

1. Degree centrality is the average over all nodes either of the number or the strength of connections, respectively.
2. Clustering coefficient is the average over all nodes by the fraction of connections between the nodes within their neighborhood, divided by the number of connections that could possibly exist among them.
3. Characteristic path length is the average of the shortest path lengths in the network-computed over all pairs of nodes.
4. Global efficiency is the average inverse shortest path length in the network-computed over all pairs of nodes.
5. Betweenness centrality is the average over all nodes of the node betweenness centrality, which is the fraction of all shortest paths in the network that contain the node, divided by the number of all paths that could possibly exist among them.
6. Eigenvector centrality is the average over all nodes of the node eigenvector centrality, which, for a node indexed i, is the corresponding component i of the eigenvector of the adjacency A or weight matrix W.
7. Assortativity is the correlation coefficient of the degrees of all pairs of connected nodes.
8. Eccentricity is the average over all nodes of the maximal shortest path length between the node and any other node.
9. Diameter is the maximum over all nodes of the node eccentricity.

For the calculation of the nine (9) network measures, the Matlab functions of the Brain Connectivity Toolbox were used (https://sites.google.com/site/bctnet/measures).

2.3 Methods for network randomization

The investigation of evolution of the 9 network measures values between the 6 main shocks, requires check if these values are statistically significant in each time interval (sliding window) with the comparison of them with the corresponding values of randomized networks. A randomized network is created through a randomization procedure on the original one. A simple setting for randomization is the preservation of the total number of connections or the total strength if the connections are weighted (Newman 2010). A more elaborative randomization setting requires either the preservation of the degree of each node in the original network (Molloy and Reed 1995, Maslov and Sneppen 2002, Del Genio et al. 2010), or respectively the strength for weighted connections (Opsahl et al. 2008). In a different approach, the random network is not formed by randomizing the connections of the original network but it is built according to the Erdős and Rényi (1959) model with preset probability of connections as in the original network, which essentially corresponds to the preservation of the average degree. The degree distribution of this randomized network is Poissonian.

For networks derived from time series, the randomized networks are given by the correlation measure (Chorozoglou and Kugiumtzis 2018). Each of the K time series $\{X_{1,t}, \ldots, X_{K,t}\}$ for $t = 1, \ldots, n$, is randomized separately under the condition of preserving the marginal distribution and the autocorrelation function (Kugiumtzis 2002), or equivalently the power spectrum (Schreiber and Schmitz 1996). The procedure forming B randomized correlation networks with weighted

Table 1 The mathematical expressions of nine network measures with simple and weighted undirected connections, where ρ is the number of shorter paths between corresponding nodes for the betweenness centrality and λ is the greatest eigenvalue of the solution of the equation $AX = \lambda X$ (for simple connections) or $WX = \lambda X$ (for weighted connections) for the eigenvector centrality

Simple connections	Weighted connections
Degree centrality or strength	
$\bar{k} = \sum_{i=1}^{K} k_i, \quad k_i = \sum_{j=1}^{K} a_{ij}$	$\bar{k}^w = \sum_{i=1}^{K} k_i^w, \quad k_i^w = \sum_{j=1}^{K} w_{ij}$
Clustering coefficient	
$C = \frac{1}{K}\sum_{i=1}^{K} c_i = \frac{1}{K}\sum_{i=1}^{K} \frac{\sum_{j,\square \in K} a_{ij} a_{i\square} a_{j\square}}{k_i(k_i - 1)}$	$C^W = \frac{1}{K}\sum_{i=1}^{K} c_i^w = \frac{1}{K}\sum_{i=1}^{K} \frac{\sum_{j,\square \in K} w_{ij} w_{i\square} w_{j\square}}{k_i^w(k_i^w - 1)}$
Characteristic path length	
$L = \frac{1}{K}\sum_{i=1}^{K} L_i = \frac{1}{K}\sum_{i=1}^{K} \frac{\sum_{j \in K, j \neq i} d_{ij}}{K-1}$	$L^W = \frac{1}{K}\sum_{i=1}^{K} L_i^w = \frac{1}{K}\sum_{i=1}^{K} \frac{\sum_{j \in K, j \neq i} d_{ij}^w}{K-1}$
Global efficiency	
$E = \frac{1}{K}\sum_{i \in K} E_i = \frac{1}{K}\sum_{i \in K} \frac{\sum_{j \in K, j \neq i} d_{ij}^{-1}}{K-1}$	$E^W = \frac{1}{K}\sum_{i \in K} E_i^W = \frac{1}{K}\sum_{i \in K} \frac{\sum_{j \in K, j \neq i} (d_{ij}^w)^{-1}}{K-1}$
Betweenness centrality	
$b_i = \frac{1}{(K-1)(K-2)} \sum_{\substack{h,j \in \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}}$	$b_i^w = \frac{1}{(K-1)(K-2)} \sum_{\substack{h,j \in \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}^w(i)}{\rho_{hj}^w}$
Eigenvector centrality	
$X_i = \lambda^{-1} \sum_{j \in K} a_{ij} x_j$	$X_i^w = \lambda^{-1} \sum_{j \in K} w_{ij} x_j$
Assortativity	
$r = \frac{E^{-1}\sum_{(i,j) \in E} k_i k_j - [E^{-1}\sum_{(i,j) \in E} \frac{1}{2}(k_i^2 + k_j^2)]^2}{E^{-1}\sum_{(i,j) \in E} \frac{1}{2}(k_i^2 + k_j^2) - [E^{-1}\sum_{(i,j) \in E} \frac{1}{2}(k_i^2 + k_j^2)]^2}$	$r^W = \frac{E^{-1}\sum_{(i,j) \in E} w_{ij} k_i^w k_j^w - [E^{-1}\sum_{(i,j) \in E} \frac{1}{2} w_{ij}(k_i^w + k_j^w)]^2}{E^{-1}\sum_{(i,j) \in E} \frac{1}{2} w_{ij}((k_i^w)^2 + (k_j^w)^2) - [E^{-1}\sum_{(i,j) \in E} \frac{1}{2} w_{ij}(k_i^w + k_j^w)]^2}$
Eccentricity	
$e_i = max_{x \in K} \{d(i,x)\}$	$e_i^W = max_{x \in K} \{d^W(i,x)\}$
Diameter	
$d_G = max_{i \in K}\, e_i$	$d_G^W = max_{i \in K}\, e_i^W$

connections is given in the following steps: In the first step, for each time series $\{X_{k,\,t}\}$, $t = 1, \ldots, n$, a randomized (surrogate) time series, $\{X_{k,t}^*\}$, is generated by the algorithm of Iterative Amplitude Adjusted Fourier Transform (IAAFT) (Schreiber and Schmitz 1996). By repeating this for each of the K time series, the surrogate multivariate time series $\{X_{1,t}^*, \ldots, X_{K,t}^*\}$ is obtained. In the second step, the correlation matrix is computed on $\{X_{1,t}^*, \ldots, X_{K,t}^*\}$ and a proper weight matrix $W^* = \left\{w_{ij}^*\right\}$, $i, \; j \in \{1, \ldots, K\}$, is formed, where e.g., $w_{ij}^* = |r_{ij}^*|$ and

r^*_{ij} is the zero-lag cross-correlation computed on the surrogate time series $\left\{X^*_{i,t}\right\}$ and $\left\{X^*_{j,t}\right\}$ of X_i and X_j, respectively. This randomization procedure is repeated B times to generate B randomized correlation networks. For simple connections, an adjacency matrix A^* is derived from the weight matrix W^*, and two approaches for forming A^* are considered. Firstly, the same threshold criterion (significance test at the significance level $\alpha = 0.05$) is applied as for the original network, which, however, does not preserve the total degree $\overline{k}$. Then, for preserving the total degree $\overline{k}$, the threshold that gives the total number of the original network connections is used.

2.4 Randomization significance test

Since the asymptotic null distribution of each statistic is not known, meaning each one of the 9 network measures, a randomization significance test is applied and the empirical null distributions for each statistic is formed from the values of the statistic computed on the B randomized networks (for any of the 6 approaches for network randomization which are used). The null hypothesis H_0 that the network measure values of both the original and randomized networks are similar (i.e., $p-value > 0.05$, see Eq. (3)) is considered. For the null hypothesis H_0 to be accepted, the values of the original network measures are required to lie within the range of the corresponding values for B randomized networks. To establish the statistical significance of the network measures values, the test should reject the null hypothesis H_0 (i.e., $p-value < 0.05$, see Eq. (3)). The p value is calculated, and the test decision is reached at the significance level $\alpha = 0.05$. Denoting q_0, the test statistic computed for the original network, and $q_1, \dots, q_B$ for the B randomized ones (in the computations we set $B = 100$), the p value is

$$p = \begin{cases} \frac{2r0}{B+1}, \text{if } r0 \leq \frac{B+1}{2} \\ \frac{2(1-r0)}{B+1}, \text{if } r0 > \frac{B+1}{2} \end{cases} \quad (3)$$

where r_0 is the rank of q_0 in the ordered list of q_0, $q_1, \dots, q_B$.

3 Application

The data and the application of the methodology are firstly described and then the results are given. The final goal is to identify potential temporal patterns that the earthquake networks follow a distinct evolution, namely whether there exist time intervals before the occurrence of each one of the investigated main shock, when the values of network measures are statistically significant.

3.1 Data and computational setting

The earthquake catalog used in this study has been complied in the Geophysics Department of the Aristotle University of Thessaloniki (http://geophysics.geo.auth.gr/ss/). Crustal earthquakes (focal depth less than 50 km) of magnitude $M \geq 3.0$, which occurred during 1999–2017 in Greece and surrounding neighboring areas, are only considered and the condition for the catalog completeness has been checked and fulfilled (Fig. 1). During this period, twenty one (21) main shocks with $M \geq 6.0$ occurred, however, for only six (6) of them, the interevent period was sufficiently long for the network measures to be robustly computed.

The network nodes are represented by the $K = 17$ seismic zones, and the observed variable of the time series is either the cumulative seismic moment M_0 of earthquakes or the number of earthquakes as has been already mentioned above. The connections are given by the weight matrix, W, which is formed from the cross-correlation value $|r_{X,\ Y}|$ for each pair of observed variables X and Y, and by the adjacency matrix A from the significance test given by Eq. (2) for the cross-correlation. The randomized networks are generated by the three approaches in which the randomization is performed on time series (Chorozoglou and Kugiumtzis 2018), by the two approaches in which the randomization is performed directly on the original network connections (Maslov and Sneppen 2002; Opsahl et al. 2008) as well as by the method with a preset probability (Erdős and Rényi 1959) as are presented in "Methods for network randomization."

Three different settings of time series length were considered, in particular $n = 60$, $n = 90$, and $n = 120$, to investigate firstly whether the results are affected by the n and then whether there appear specific temporal patterns that the values of network measures between main shocks differ from the randomness, i.e., they are statistically significant. We do not choose a shorter or longer

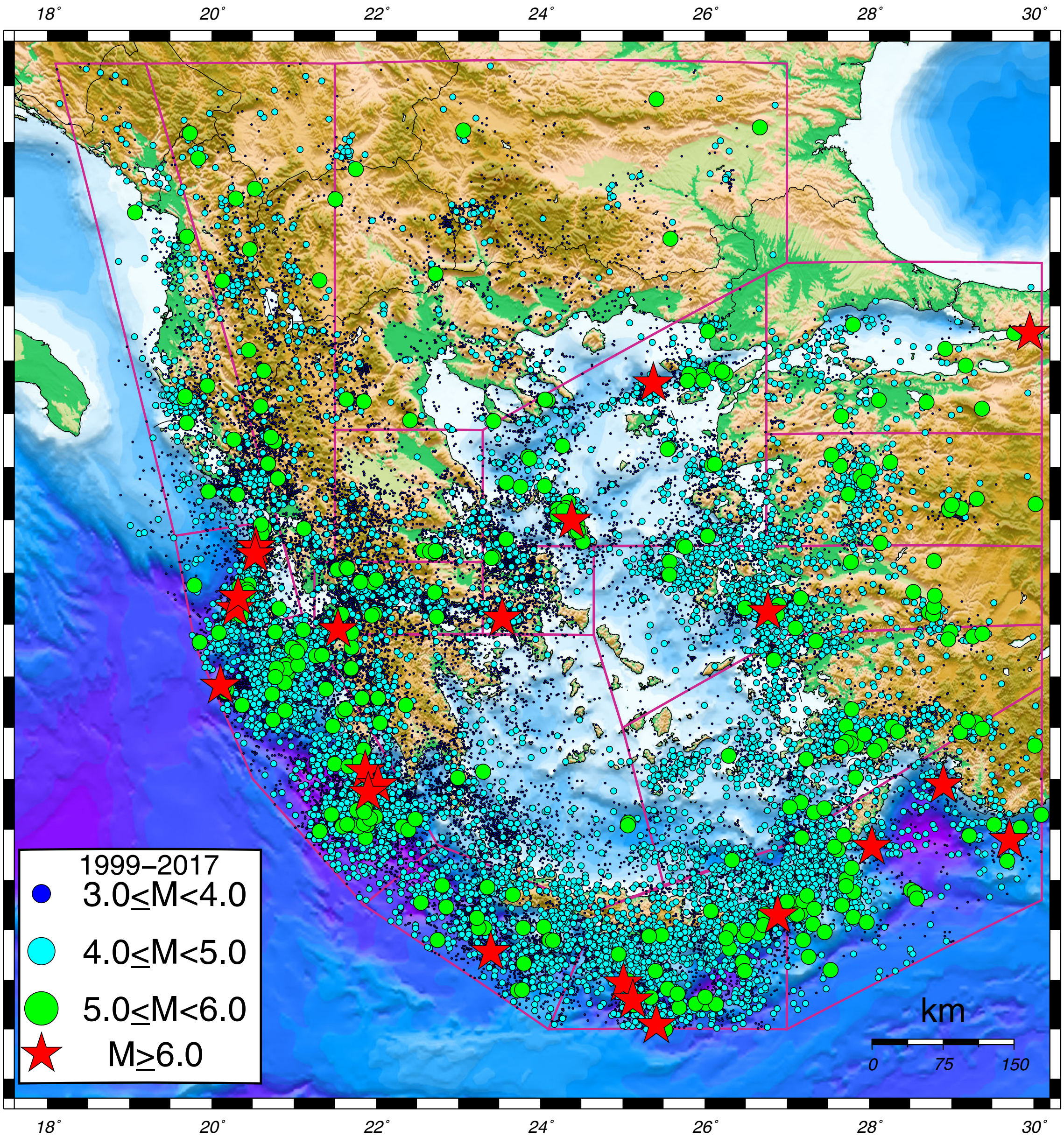

Fig. 1 Epicentral distribution of the earthquakes of $M \geq 3.0$ that occurred in 1999–2015 in the broader area of Greece divided in 17 seismic zones

length n of the time series, because either the assessment of the correlation among the seismic zones for the connections introduction could not be robust or the number of time windows would be particularly low. The high levels of the seismic activity in the study area allow to consider the sampling time equal to one (1) day. Although the sampling time is short, the zeros in the earthquake time series (no occurrence of earthquake within the sampling time) are minimized because the threshold of magnitude completeness is particularly low ($M \geq 3.0$) during 1999–2017. Therefore, the monitoring of the 9 network measures values is performed per 60, 90, and 120 days on each non-overlapping sliding window and compared with the values of the corresponding $B = 100$ randomized networks, constructed by the 6 randomization approaches, based on the relation (3), which checks the statistical significance of the values. We consider the cases where the interevent temporal distance is over 360 days, for achieving the formation of at least three time windows, for the purpose that the investigation of the specific time patterns will have great importance. The days between successive main shocks are not the same for all examined cases (see Table 2), hence, the total number of the time windows differs, although the time series length for each n is kept the same.

4 Results

The percentages of rejection of the null hypothesis, H_0, that the values of the original and randomized network measures are similar, for different length n and observed variables of the time series for each one of the 6 examined cases (bold font in Table 2) are shown in Figs. 2, 3, 4, 5, 6, and 7, without making a distinction about network measures and the randomization methods. Thus, it is scrutinized with an aggregated manner whether for each of the six cases the null hypothesis H_0 tends to be rejected, i.e., the values of the nine network measures are different from the correspondence of the B randomized networks, and in which time window this distinct evolution of values is observed. The percentages of rejection of H_0 are computed from the average of the corresponding rejections of H_0 that are derived from all the six randomization approaches, for all the nine network measures in each non-overlapping time window and for each one of the six examined case, which are discussed in more detail below.

The first main shock occurred on 26/01/2001 with $M = 6.3$. The percentage of rejection of H_0 is up to 25% in the case when the time series length is $n = 60$, the time window is the third (i.e., between 120 and 180 days before the upcoming main shock and as observed variable of time series the number of earthquakes is considered (Fig. 2). The percentages of rejection of H_0 when the seismic moment is considered as observed variable are lower than in the case when the earthquakes frequency. Nevertheless, there is not a specific time window in which high level of rejection percentages of H_0

Table 2 The details of the 21 main shocks of magnitude $M \geq 6.0$ that occurred in 1999–2015 in the broader area of Greece. The main shocks which are used for the study are in bold font

Index of $M \geq 6.0$	Date	Longitude	Latitude	Magnitude	Days of succession
1	7/9/1999	38.06	23.53	6.0	–
2	**26/1/2001**	**38.99**	**24.38**	**6.3**	**688**
3	**14/8/2003**	**38.74**	**20.53**	**6.2**	**749**
4	17/3/2004	34.77	23.39	6.3	216
5	23/1/2005	35.89	29.70	6.2	312
6	31/1/2005	37.41	29.10	6.2	8
7	20/10/2005	38.12	26.76	6.1	262
8	**14/2/2008**	**36.57**	**21.86**	**6.8**	**847**
9	14/2/2008	36.43	22.02	6.5	0
10	20/2/2008	36.36	21.90	6.2	6
11	8/6/2008	37.95	21.53	6.4	109
12	15/7/2008	35.83	28.03	6.4	37
13	**1/7/2009**	**34.04**	**25.41**	**6.4**	**361**
14	**10/6/2012**	**36.44**	**28.90**	**6.1**	**1075**
15	**15/6/2013**	**34.46**	**25.01**	**6.3**	**370**
16	16/6/2013	34.29	25.12	6.1	1
17	26/1/2014	38.15	20.28	6.1	224
18	3/2/2014	38.26	20.32	6.0	8
19	24/5/2014	40.28	25.37	6.2	110
20	16/4/2015	35.14	26.88	6.1	327
21	17/11/2015	38.67	20.53	6.5	215

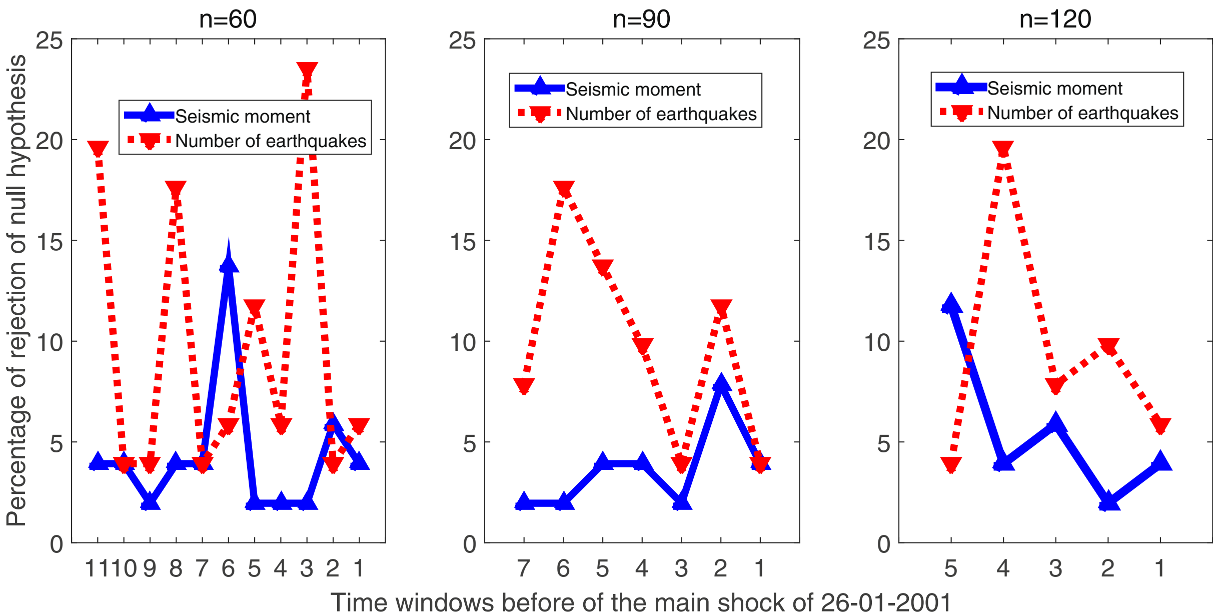

Fig. 2 The percentage of rejection of H_0 that the values for the network measures of original and randomized networks are similar, for the main shock of 26-01-2001, in the left panel for $n = 60$, in the middle panel for $n = 90$ and in the right panel for $n = 120$ based on the average of rejections of H_0 for all of the nine network measures and the six randomization approaches in each time window which are used. The different number of time windows (1, in $x - axis$, means the last time interval before the upcoming main shock) for each panel is due in the different time series length n which is considered

systematically appear. The last time window that could warn of the upcoming main shock presents very low percentages of rejection of H_0, near to 5% for both observed variables and any time series length n.

In the second case, the main shock occurred on 14/08/2003 with $M = 6.2$. The profile of the rejection of H_0 also follows the same pattern as in the previous case in terms of the percentages which go up to 20% (Fig. 3). In one case, when the time series length is $n = 120$, the time window is the fifth (i.e., between 600 and 720 days before the upcoming main shock) and when as observational variable the seismic moment is considered, the percentage of rejection of H_0 reached 35% and is associated with the aftershock sequence of the main shock of 26/1/2001. Here, there is a distinct evolution, i.e., the values of the nine network measures differ from the randomness, in the last time interval before the upcoming main shock of 14/08/2003 in relation to the previous case. Thus, when the time series length is $n = 60$ for both observed variables, the percentage of rejection of H_0 reached 20% which far exceeds the values in the remaining windows.

The third case of the study, ending up to a main shock occurred on 14/02/2008 with $M = 6.8$, has also the same pattern as the previous cases (Fig. 4). Hence, the percentages of rejection of H_0 are lower when the seismic moment is considered as the observed variable than the number of earthquakes. There is also one case when the time series length is $n = 90$, the time window is the seventh (i.e., between 630 and 720 days just before the main shock) and as the observed variable, the number of earthquakes is considered, with high percentage of rejection of H_0 because of the aftershock sequence of the main shock of 20/10/2005. In addition, there is a distinct evolution in the last time interval. Therefore, when the time series length is $n = 120$ for both of observed variables which are considered, the percentage of rejection of H_0 reached the 20%, which far exceeds the remaining windows especially in the case of seismic moment. The same situation as regards the temporal patterns, for the last time interval, is shown when the time series length is $n = 90$ but only when the seismic moment is considered as the observed variable of time series.

In the fourth case, the main shock that occurred on 01/07/2009 with $M = 6.4$ is studied and, here, the lowest number of the time windows in relation with the other cases (Fig. 5) is set because of the shorter temporal distance between this and the latter main shock (the previous main shock occurred on 15/7/2008). Similarly, with the previous cases, the percentages of rejection of

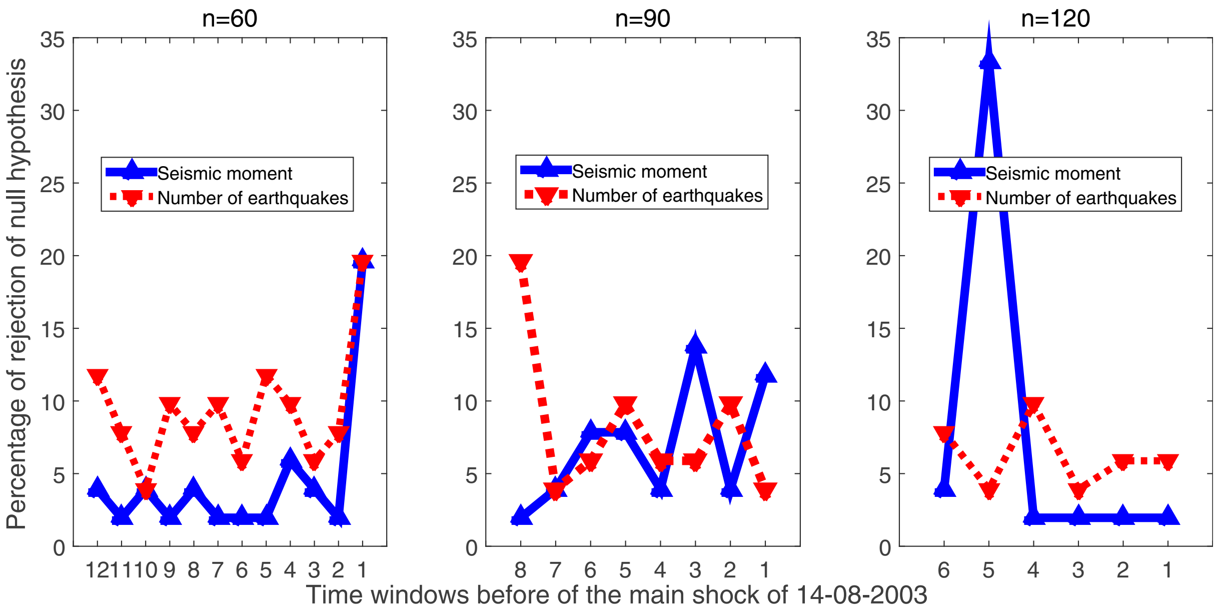

Fig. 3 The same as Fig. 2 but for the main shock of 14-08-2003

H_0 are lower when the seismic moment is considered as the observed variable, in comparison with the number of earthquakes and reached up to 25%. Furthermore, the distinct evolution also appears in the last time window just before the main shock. Hence, when the time series length is $n = 120$ for both observed variables, the percentage of rejection of H_0 is much higher than the remaining windows, up to 25% for the seismic moment and 10% for the number of earthquakes. The same behavior is observed for $n = 90$ when the seismic moment is considered as observed variable but with lower percentage than in the case of $n = 120$.

In the fifth case of the study, the main shock that occurred on 10/06/2012 with $M = 6.1$ is investigated. The percentages of rejection of H_0 when the seismic moment is considered as the observed variable are

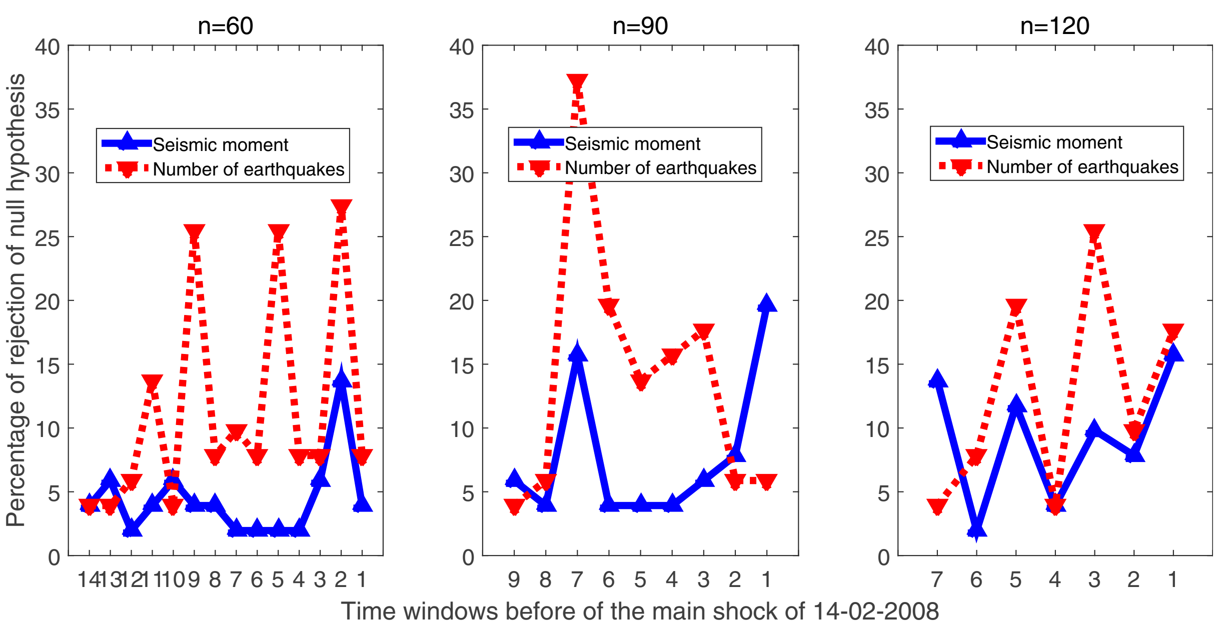

Fig. 4 The same as Fig. 2 but for the main shock of 14-02-2008

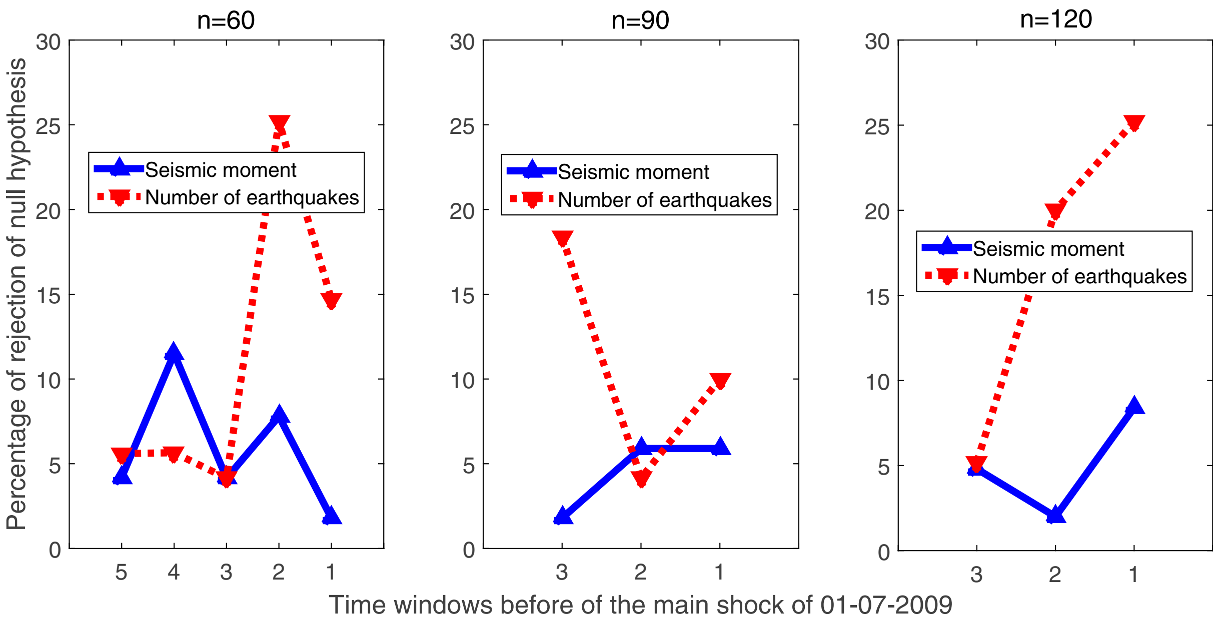

Fig. 5 The same as Fig. 2 but for the main shock of 01-07-2009

smaller than the number of earthquakes especially for $n = 60$ (Fig. 6). The encouraging indication for this case is the distinct evolution of the network measures values in the last time interval for $n = 120$ for both observed variables of time series. Especially, for seismic moment, the values of the network measures diverge 15% by the randomness in relation with the other time windows of which the percentages of rejection of H_0 arrived at 5%.

In the last case, the main shock that occurred on 15/06/2013 with $M = 6.3$ (Fig. 7) is considered. The important point in relation to the previous five cases is that when the number of earthquakes is considered as the observed variable the last time interval, just before the main shock exhibits the highest percentages of rejection of H_0 in relation with the other time windows regardless of time series length n.

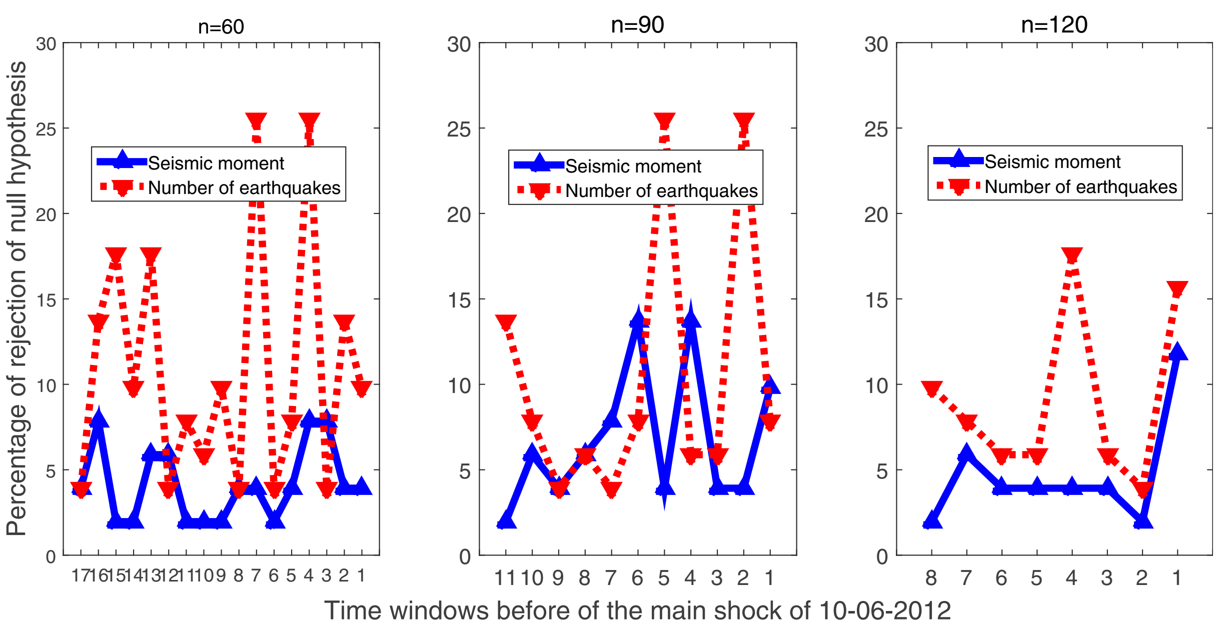

Fig. 6 The same as Fig. 2 but for the main shock of 10-06-2012

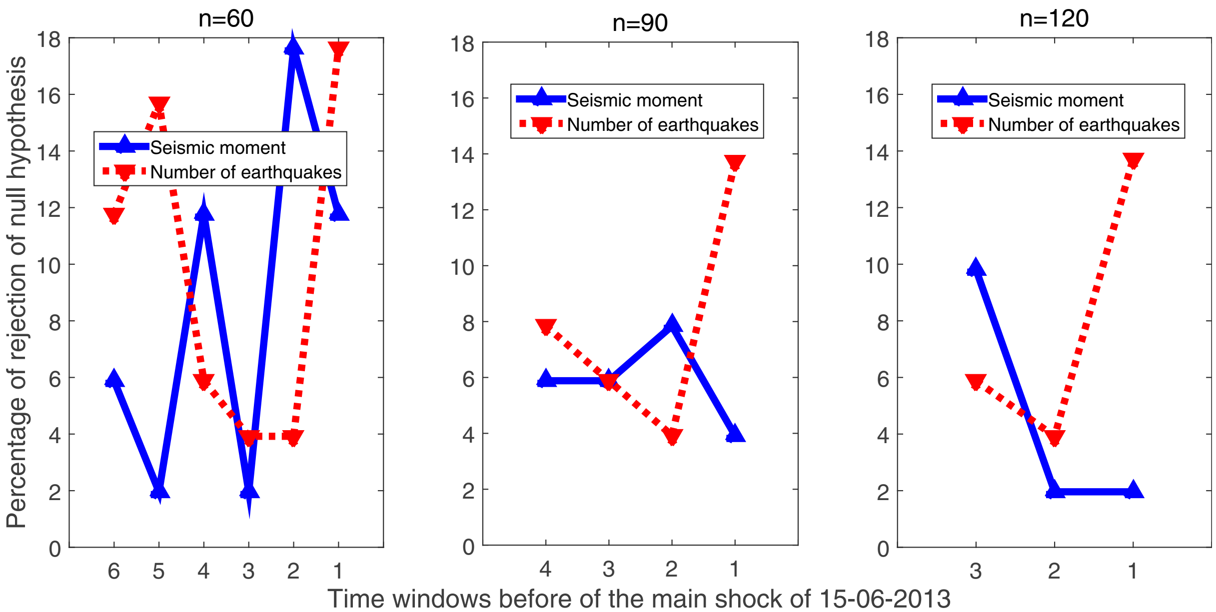

Fig. 7 The same as Fig. 2 but for the main shock of 15-06-2013

The study is not focused on the comparison of randomization methods, as this has been done by Chorozoglou and Kugiumtzis (2018), but on the network measures as to whether their values are statistically significant in certain time interval just before the main shocks occurrence. Hence, Fig. 8 shows the percentage of rejection of H_0 for each of the nine network measures and for all main shocks, and the different time series lengths n. The percentage of rejection of H_0 for each time window is computed by averaging the corresponding rejections for the specific time window that defined for each one of the six main shocks, with different time series length n. The percentages of rejection of H_0 for the network measures, namely of the degree centrality or strength, clustering coefficient, characteristic path length, global efficiency, eccentricity, and diameter in the last time interval before upcoming main shocks are very high, reaching over 10%, for both observed variables of time series, than the previous windows. This happens if we exclude some of cases that in the first windows the percentages of rejections of H_0 are higher than the last windows because of aftershock sequence of main shocks. This finding, that the values of the six network measures for original earthquake networks are different from corresponding of B randomized networks in the last time interval before upcoming main shocks, might be considered promising for the seismic hazard assessment. The remaining three network measures, namely betweenness, eigenvector centrality, and assortativity do not exhibit any distinct evolution.

5 Discussion and concluding remarks

The area of Greece accommodates a large number of strong earthquakes that often produced severe damage and even loss of life. To address this problem, we use the network theory which is a relatively recent field, using some basic network measures, to identify potential temporal patterns aiming to assess the seismic hazard in the study area. An exhaustive analysis was performed aiming to identify whether the earthquake networks which are constructed, for the time interval between six main shocks, exhibit a distinct evolution as far as the values of the nine basic network measures concerns. The importance point which is attempted here in relation to the previous studies is that we do not interested in the abrupt variation in the value of a certain network measure but whether this value is statistically significant in a certain time window, and in particular, just before the upcoming main shocks.

The application of the network theory is found to be a powerful tool for the investigation of complex phenomena, such as seismic activity as the changes in the network structure can reveal certain seismicity behavior a few days before a main shock occurrence. The

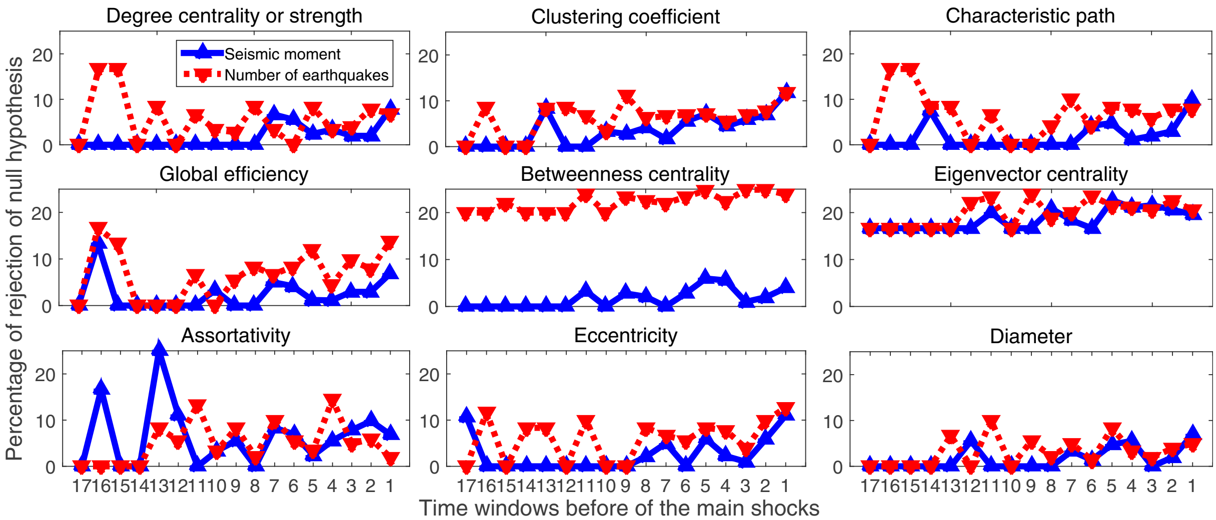

Fig. 8 The percentage of rejection of H_0 that the values for each of the nine network measures of original and randomized networks are similar, based on all the six6 main shocks and the different time series lengths n which are considered

investigation of values of two basic network measures, namely the clustering coefficient and the betweenness centrality, before an upcoming main shock revealed that they tend to be stable with small variations before the main shock and then they have an abrupt jump just before the main shock occurrence (Abe and Suzuki 2009; Daskalaki et al. 2016; Chorozoglou et al. 2017). Hence, the values of the nine (9) basic network measures are computed, for each of the observed variable and length n of the time series on each time window for the six (6) main shocks and compared with the corresponding values of the randomized networks that constructed with the six (6) different randomization approaches. It was found that the values of the six network measures, namely the degree centrality or strength, clustering coefficient, characteristic path length, global efficiency, eccentricity and diameter, are different than the corresponding values of the randomized networks in the last time window, with higher percentages of rejection of H_0 than in the other time intervals, before the upcoming main shocks. Therefore, the complex network theory reveals that these network measures could serve as potential indices for short-term seismic hazard assessment. These results suggest the effectiveness of seismicity network analysis in an earthquake prone area, such as Greece and its surroundings. It is anticipated that the development of these tools to characterize and provide useful information on the seismic regions could eventually provide insight about the development of robust seismic forecasting tools.

An open issue arising from this study is the application of the same analysis but this time the investigation of the values of the network measures will be extended to the studying of the values of each node instead of average of all nodes of network, with purpose of the spatial estimation of the upcoming main shocks. In addition, earthquake networks with directed connections need to be considered, i.e., causality networks, and to use more network measures for a more complete and thorough analysis. Hence, we acknowledge that the analysis in this work can be extended in the future to the investigation of the values of each node using directed network and possibly increase also the number of network measures to be considered.

Funding information The financial support by the European Union and Greece (Partnership Agreement for the Development Framework 2014-2020) for the project "Development and application of time-dependent stochastic models in selected regions of Greece for assessing the seismic hazard" is gratefully acknowledged, MIS5004504. The authors also acknowledge the Geophysics Department Contribution, 915.

References

Abe S, Suzuki N (2004a) Small-world structure of earthquake network. Phys A 337:357–362

Abe S, Suzuki N (2004b) Scale-free network of earthquakes. Europhys Lett 65:581–586

Abe S, Suzuki N (2006) Complex-network description of seismicity. Nonlinear Proc Geoph 13:145–150

Abe S, Suzuki N (2007) Dynamical evolution of clustering in complex network of earthquakes. Eur Phys J B 59:93–97

Abe S, Suzuki N (2009) Main shocks and evolution of complex earthquake networks. Braz J Phys 39(2A):428–430

Abe S, Pasten D, Munoz V, Suzuki N (2011) Universalities of earthquake-network characteristics. Ch Sc Bul 56(34):3697–3701

Aki K (1981) A probabilistic synthesis of precursory phenomena. In: Simpson DW, Richards PG (eds) Earthquake prediction: an international review, Maurice Ewing series 4. American Geophysical Union, Washington, D.C., pp 566–574

Albert R, Barabasi AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97

Aydin NY, Duzgun HS, Wenzel F, Heinimann HR (2017) Integration of stress testing with graph theory to assess the resilience of urban road networks under seismic hazards. Nat Hazards. 1–32

Baek WH, Lim G, Kim K, Chang KH, Jung JW, Seo SK, Yi M, Lee DI, Ha DH (2011) Robustness of the topological properties of a seismic network. J Korean Phys Soc 58(6):1712–1714

Baiesi M, Paczuski M (2004) Scale-free networks of earthquakes and aftershocks. Phys Rev E 69(6):066106

Baiesi M, Paczuski M (2005) Complex networks of earthquakes and aftershocks. Nonlinear Process Geophys 12:1–11

Bak P, Tang C (1989) Earthquakes as a self-organized critical phenomenon. J Geophys Res 94(B11):635–615 637

Bialonski S, Horstmann MT, Lehnertz K (2010) From brain to earth and climate systems: small-world interaction networks or not? Chaos 20:1–9

Billio M, Getmansky M, Lo AW, Pelizzon L (2012) Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J Financ Econ 104(3):535–539

Bullmore ET, Fornito A, Zalesky A (2016) Fundamentals of brain network analysis, Academic Press, Elsevier

Chelidze T (2017) Complexity of seismic process: a mini-review. Phys Astron Intern J 1(6):1–8

Chorozoglou D, Kugiumtzis D (2014) Testing the randomness of causality networks from multivariate time series, International Symposium on Nonlinear Theory and its Applications, NOLTA, Luzern, Switzerland, September 14-18; 229–232

Chorozoglou D, Kugiumtzis D (2018) Testing the randomness of correlation networks from multivariate time series. J Complex Net. https://doi.org/10.1093/comnet/cny020

Chorozoglou D, Kugiumtzis D, Papadimitriou E (2017) Application of complex network theory to the recent foreshock sequences of Methoni (2008) and Kefalonia (2014) in Greece. Acta Geoephys 65(3):543–553

Chorozoglou D, Kugiumtzis D, Papadimitriou E (2018) Testing the structure of earthquake networks from multivariate time series of successive main shocks in Greece. Phys A 499C: 28–39

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

Daskalaki E, Papadopoulos GA, Spiliotis K, Siettos C (2014) Analysing the topology of seismicity in the Hellenic arc using complex networks. J Seismol 18:37–46

Daskalaki E, Spiliotis K, Siettos C, Minadakis G, Papadopoulos GA (2016) Foreshocks and short-term hazard assessment of large earthquakes using complex networks: the case of the 2009 L'Aquila earthquake. Nonlinear Process Geophys 23: 241–256

Del Genio C, Kim H, Toroczkai Z, Bassler K (2010) Efficient and exact sampling of simple graphs with given arbitrary degree sequence. PLoS One 5(4):e10012

Donges JF, Zou Y, Marwan N, Kurths J (2009) The backbone of the climate network export. Europhys Lett 87:48007

Erdős P, Rényi A (1959) On random graphs, pub. Math (Debrecen); 6: 290–297

Fiedor P (2014) Networks in financial markets based on the mutual information rate. Phys Rev E 89:052801

Fornito A, Bullmore ET, Zalesky A (2016) Fundamentals of brain network analysis; Cambridge Massachusetts; United States; Academic Press; Elsevier

Girvan M, Newman ME (2002) Community structure in social and biological networks. Proc Natl Acad Sci 99:7821–7826

Gutenberg B, Richter CF (1944) Frequency of earthquakes in California, Bulletin of the Seismological Society of America. Bull Seismol Soc Am 34:185–188

Heiberger RH (2014) Stock network stability in times of crisis. Phys A 393:376

Hill DP, Reasenberg PA, Michael A, Arabaz WJ, Beroza G, Brumbaugh D, Brune JN, Castro R, Davis R, DePolo D (1993) Seismicity remotely triggered by the magnitude7.3 Landers, California, Earthq. Science 260(5114):1617–1623

Horvath S (2011) Weighted network analysis, applications in genomics and systems biology. Springer, New York

Janer C, Biton D, Batac R (2017) Incorporating space, time, and magnitude measures in a network characterization of earthquake events. Acta Geophys 65:1153–1166

Jeong H, Mason SP, Barabasi AL, Oltvai ZN (2001) Lethality and centrality in protein networks. Nature 411:41

Jimenez A, Tiampo KF, Posadas AM (2008) Small world in a seismic network: the California case. Nonlinear Process Geophys 15:389–395

Kanamori H, Anderson L (1975) Theoretical basis of some empirical relations in seismology. Bull Seismol Soc Am 65(5): 1073–1095

Kugiumtzis D (2002) Statistically transformed autoregressive process and surrogate data test for nonlinearity. Phys Rev E 66: 025201

Kugiumtzis D, Koutlis C, Tsimpiris A, Kimiskidis VK (2017) Dynamics of epileptiform discharges induced by transcranial magnetic stimulation in genetic generalized epilepsy. Int J Neural Syst 27(7):1750037

Lennartz S, Livina VN, Bunde A, Havlin S (2008) Long-term memory in earthquakes and the distribution of interoccurrence times. Europhys Lett 81:69001

León DA, Valdivia JA, Bucheli VA (2018) Modeling of Colombian seismicity as small-world networks. Seismol Res Lett 89(5):1807–1816

Lippiello E, Arcangelis L, Godano C (2008) Influence of time and space correlations on earthquake magnitude. Phys Rev Lett 100:038501

Livina VN, Havlin S, Bunde A (2005) Memory in the occurrence of earthquakes. Phys Rev Lett 95:208501

Maslov S, Sneppen K (2002) Specificity and stability in topology of protein networks. Sc 296:910–913

Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. Random Struct Algoritm 6(2–3): 161–180

Newman M (2010) Networks, an introduction, Oxford University Press

Omori F (1894) On the aftershocks of earthquakes. Sc Imp Univ Tokyo 7: 111–120

Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: the weighted rich-club effect. Phys Rev Lett 101: 168702

Palus M, Hartman D, Hlinka J, Vejmelka M (2011) Discerning connectivity from dynamics in climate networks. Nonlinear Process Geophys 18:751–763

Papadimitriou E (2002) Mode of strong earthquake recurrence in the Central Ionian Islands (Greece): possible triggering due to coulomb stress changes generated by the occurrence of previous strong shocks. Bull Seismol Soc Am 92(8):3293–3308

Papadimitriou E, Karakostas V (2005) Occurrence patterns of strong earthquakes in Thessalia area (Greece) determined by the stress evolutionary model B: a case study. Earth Plan Sc Lett 235(3):395–409

Papadimitriou E, Wen X, Karakostas V, Jin X (2004) Earthquake triggering along the Xianshuihe Fault Zone of Western Sichuan, China. Pure Appl Geophys 161(8):1683–1707

Papana A, Kyrtsou C, Kugiumtzis D, Diks C (2017) Financial networks based on Granger causality: a case study. Phys A 482:65–73

Papazachos BC, Papadimitriou EE, Kiratzi AA, Papazachos CB, Louvari EK (1998) Fault plane solutions in the Aegean Sea and the surrounding area and their tectonic implication. Boll Geof Teor App 39:199–218

Pastén D, Torres F, Toledo B, Muñoz V, Rogan J, Valdivia JA (2016) Time-based network analysis before and after the Mw 8.3 Illapel earthquake 2015 Chile. Pure Appl Geophys 173(7):2267–2275

Porta A, Faes L (2016) Wiener-Granger causality in network physiology with applications to cardiovascular control and neuroscience. Proc IEEE 104:282–309

Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52:1059–1069

Scholz CH (2010) Large earthquake triggering, clustering, and the synchronization of faults. Bull Seismol Soc Am 100:901–909

Schreiber T, Schmitz A (1996) Improved surrogate data for nonlinearity tests. Phys Rev Lett 77(4):635–638

Stein R, Barka A, Dieterich J (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. Geophys J Int 128:593–604

Utsu T (1961) A statistical study on the occurrence of aftershocks. Geophys Mag 30:521–605

Utsu T (2002) Statistical features of seismicity. International Handbook of Earthquake & Engineering Seismology; Part A San Diego. Academic; 719–732

Utsu T, Ogata Y, Matsura RS (1995) The centenary of the Omori formula for a decay law of aftershock activity. J Phys Earth 43:1

Vamvakaris DA, Papazachos CB, Papaioannou CA, Scordilis EM, Karakaisis GF (2016) A detailed seismic zonation model for shallow earthquakes in the broader Aegean area. Nat Hazards Earth Syst Sci 16:55–84

Wang X, Chen G (2003) Complex networks: small-world, scale-free and beyond. Feature. 6–20

Wang X, Koç Y, Derrible S, Ahmad SN, Pino WJA, Kooij RE (2017) Multi-criteria robustness analysis of metro networks. Phys A 474:19–31

Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. Nature 393:440–442

Zhang X, Gan C (2018) Global attractivity and optimal dynamic countermeasure of a virus propagation model in complex networks. Phys A 490:1004–1018