Correlation properties in worldwide and synthetic earthquake networks

Jennifer Ribeiro*[1,2], Paulo S. L. Oliveira Júnior[1,3], Leandro O. Pereira[1], Renato P. Freitas[1], André R. Pimenta[1], Andrés Papa[4,5] and Douglas S. R. Ferreira[1], [1]LISComp – IFRJ, [2]UFRRJ, [3]UFF, [4]UERJ, [5]ON

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 8-11 November 2021.

Abstract

In this work, we have studied correlation properties of earthquake networks, by the assortativity analysis of global and synthetic earthquakes networks. We have used data from the worldwide earthquake catalog for the period between 2002 and 2016 of earthquakes with magnitude greater than or equal to 4.5, with shallow events (depth up to 70 km) and deep events (depth greater than 70 km) being analyzed separately, and an earthquakes catalog simulated with a modified version of the OFC model. The study was conducted for two methodologies of connections: the successive model and the time window model. We calculated the nearest-neighbors average connectivity of nodes and the degree correlation coefficient for all networks. It was observed that only the networks built from shallow earthquakes and synthetic data have assortative mixing. These results are in concordance with previous works; however, the present work is more comprehensive. We observed that the networks of deep earthquakes, are neutral. Our results contribute to a better understanding of the seismological dynamics and the correlations existing between different seismological regions.

Introduction

In the last years, several works have implemented the concept of complexity in the study of earthquakes, aiming to better understand and characterize the seismological dynamics and properties. Complex networks are a powerful tool to investigate the topological structure and statistical behavior of complex systems and have been successfully applied in many real-world networked systems, such as the Internet, economic market, spread of diseases, solar flares and social relationships (Watts and Strogatz, 1998; Barabási and Albert, 1999; Albert and Barabási, 2000, 2002; Dorogovtsev and Mendes, 2003; Gheibi et al., 2017; Newman, 2003).

Some of the most common and fundamental features of networks are the clustering coefficient, the average shortest path and the degree distribution. However, another crucial characteristic to be studied is the correlation property given by the networks' assortativity, which indicates the likelihood of a node with degree k to be connected to other nodes of the same degree k. For example, in social networks it is observed that people tend to relate to other people belonging to the same group as themselves (Newman, 2003) and for this reason this network is called *assortative*. However, the protein-interaction network of yeast has the opposite property: it is *disassortative* (Jeong et al., 2001). The proteins with larger values of degree interact much more with small-degree proteins. This way, the study of degree correlation is important, since the assortative (or disassortative) patterns of a network provide a much more complete information on the structure of the network, as, for example, it can describe the robustness of a network against selective node failure (Newman, 2002; D'Agostino et al., 2012).

In Abe and Suzuki (2004a, 2004b), the authors constructed earthquake networks for seismic data from California and Japan using a successive model to create the links between the nodes. They observed that these networks are small-world and scale-free. Moreover, the same authors analyzed these networks with respect to their assortative mixing (Abe and Suzuki, 2006) and found that they are assortative.

Once prediction is one of the goals of earthquake research, it makes sense to look for interactions due to long-range correlations. Therefore, intending to study global earthquakes, and not just those of a certain region of the world, the authors in Ferreira et al. (2014) constructed a successive network of worldwide epicenters, where they found that this network is also scale-free and small-world, showing evidence of long-range correlations across the planet. Furthermore, these authors proposed in Ferreira et al. (2018, 2020) a refined model, called the "time window" model, to construct networks of epicenters from all around the globe, which improves the previous successive methodology.

This way, in order to corroborate the understanding of the earthquakes phenomenon and to observe the agreement between the results found for specific regions of the world and those obtained for global earthquakes data, in this paper we analyze the assortativity of the networks created for worldwide seismic events using both the successive model and the time window model of connections for two datasets: one for shallow earthquakes and other for deep earthquakes. This division aims to make comparisons between earthquakes with similar seismic origins. Furthermore, a network of successive connections was also created using earthquake data simulated with a modified version of the Olami-Feder-Christensen (OFC) model (Ferreira et al., 2015) and the result was compared with those of real events.

This paper is organized as follows. First, we present information about correlation properties by the study of assortativity analysis, considering the nearest-neighbors average connectivity of nodes and the degree correlation coefficient. In the sequence, we give information about our worldwide and synthetic earthquake data catalogs. Then, the methods employed to construct our networks are described. Finally, we show and discuss the results obtained, ending with the conclusions.

Correlation properties in networks

The study of correlation properties in networks, using assortativity, has been implemented in many different real-world networks in the last years. Assortativity (or assortative mixing) refers to the tendency of nodes in a network to connect to other nodes with similar properties. Here, we focus on assortativity in terms of a node's degree. The analysis of this property allows us to investigate the relation between the connectivity degrees of the nodes that link to each other. A statistical measure that is commonly used to analyze this preference is the nearest-neighbors average connectivity of nodes, or also denominated degree correlation function (Pastor-Satorras et al., 2001; Vázquez, 2002; Barabási and Pósfai, 2016), expressed as

$$k_{nn}(k) = \sum_{j} jP(j|k) \qquad (1)$$

where $P(j|k)$ is the conditional probability that an arbitrary selected edge links a j-degree node with a k-degree node. This function considers the average degree of the neighbors of a node as a function of its degree k. If it is independent of k, the network has no obvious correlation of degree and is called neutral. When, however, $k_{nn}(k)$ increases with k, the network is assortative. This means that the hubs (nodes with high degrees) of the network tend to connect to other hubs and nodes with low degrees tend to be linked to other low degrees nodes. On the other hand, if $k_{nn}(k)$ decreases with k, the network is disassortative, i.e., the hubs prefer to link to nodes with low degrees (Barabási and Pósfai, 2016).

Thus, the degree correlation function can help to detect the presence or absence of correlations in real networks. However, a useful way to capture the magnitude of the correlations present in the networks is the use of a unique number. This number can be computed from the fitting of the $k_{nn}(k)$ plot or by calculating the degree correlation coefficient, defined as follows.

The degree correlation coefficient, which is the Pearson correlation coefficient between the degrees found at the two ends of the same link, is a complementation of the analysis of the degree correlation function and gives to us a quantitative characterization. We calculate this coefficient by

$$r = \sum_{jk} \frac{jk(e_{jk} - q_j q_k)}{\sigma^2}, \qquad (2)$$

where e_{jk} is the probability of finding a node with degrees j and k at the two ends of a randomly selected link, q_k is the probability of existing a node with degree k at the end of a randomly selected link and

$$\sigma^2 = \sum_k k^2 q_k - \left[\sum_k k q_k\right]^2, \qquad (3)$$

is the variance of q_k. The value of r varies from -1 (perfect disassortativity) to 1 (perfect assortativity). If r = 0, then the network has no assortative (or disassortative) mixing and, therefore, is neutral.

Data

In this study, we used real and synthetic earthquakes data catalogs.

Worldwide earthquakes

The dataset was obtained from the World Catalog of Earthquakes of Advanced National Seismic System (ANSS)[1] and it covers earthquakes from the entire world between 2002 and 2016. For the record, we only considered earthquakes with magnitude (m) larger or equal to 4.5, because in that catalog the events with magnitudes less than 4.5 are not completely registered for the whole world. The total of events is 101746, in which 80520 are shallow earthquakes (earthquakes with depth up to 70 km) and 21226 are deep earthquakes (events occurred at depths greater than 70 km).

These data had a good agreement with the Gutenberg-Richter law (GR) (Gutenberg and Richter, 1942), with a *b-value* exponent equal to 1.080 ± 0.003, for the shallow seismic events, and 1.080 ± 0.010, for the deep ones. The values of b were close to 1.00, as expected, since we have excluded earthquakes with small magnitudes.

Synthetic earthquakes

To generate the synthetic seismic catalog, we have used a modified version of the original OFC model (Olami et al.,1992). This model is widely used because, despite its simplicity, it is capable of reproducing several characteristics found in real data, such as the Gutenberg-Richter law. The original model can be represented by a bi-dimensional square ℓ x ℓ lattice of $N = \ell^2$ blocks (sites) interconnected by springs, where each block is also connected through a spring to a single rigid driven plate and by friction to other rigid fixed plate on which they stay. This is the regular topology of the lattice. Due to the relative motion between the plates (imposed by the model), all the blocks will be subjected to an elastic force which tends to put them in motion and other frictional force opposite to the first. When the resulting force in one of the blocks is greater than the maximum static friction force, the block slides and relaxes to a position of zero force, so that there is a rearrangement of forces in its first neighbors, which can cause other slippages and the emergence of a chain reaction. The first block to move is the epicenter of the earthquake and we measure the magnitude s of this earthquake by the number of blocks that skidded.

In this work, we used a lattice with small-world topology instead of the regular one. As done by Ferreira et al. (2015), this new small-world topology is built from the regular topology, where each edge of the network is

[1] https://earthquake.usgs.gov/data/comcat/

reconnected randomly with probability p, keeping fixed the original degree of each site.

Our analysis was conducted for a lattice of size $\ell = 400$, a dissipation coefficient $\alpha = 0.20$ and $p = 0.001$. The number of events generated was 2×10^7 after the transient regime. It is relevant to highlight that we have excluded epicenters of earthquakes with magnitude s = 1 to construct the network, due to the fact that these events seem to obey their own statistics (Grassberger, 1994).

Method

To construct the network of global epicenters, following the definition used in Ferreira et al. (2014), we divide the planet into equal square cells of size $L \times L$, with L= 20 km, and a cell becomes a node of the network every time the epicenter of an earthquake is located therein. To create the links between the nodes, we used two methods, which are described below.

Successive model

This methodology is the same created by Abe and Suzuki (2004a, 2004b) and employed in Ferreira et al. (2014), which basically consists of connecting a node to its subsequent one, in the temporal order, by a directed edge

Time window model

This improved model was proposed in Ferreira et al. (2018) and showed evidences of being a better approach to construct networks of earthquakes from all over the world. It consists of defining a *time window, T*, which is placed on the chronologically ordered data, to create the links between the nodes. This way, the first node inside the window is connected to all other nodes within that window,

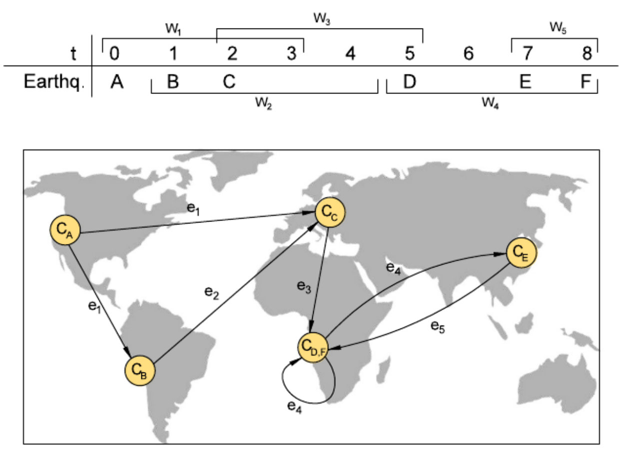


Figure 1 *- Example of the network's construction for the time window model (taken from Ferreira et al. (2020)). The time windows are represented by W_i, where i is the window number and all the time windows must have the same value (in this example, T = 3, in arbitrary units). Events in the same window are connected as explained in the text. We can see that there are 6 earthquakes (A, B, C, D, E, F), but the network of epicenters has only 5 nodes (C_A, C_B, C_C, C_D, C_E), because $C_D = C_F$. It can also be observed that the link between C_D and C_F is a self-link.*

by directed edges. Thereafter, the window is moved forward to the next event and the connections procedure is repeated. Figure 1 illustrates an example of this process.

We built networks using both the successive and the time window model for the shallow and deep earthquakes collected. The time window, T, values were T = 3800s, for shallow seismic events, and T = 16500s, for deep earthquakes. These values were calculated in Ferreira et al. (2018, 2020), respectively.

Regarding the data generated with the OFC model, each epicenter was defined as a node and we constructed a network using the successive model of connections, as it was done in Ferreira et al. (2015).

Results and discussion

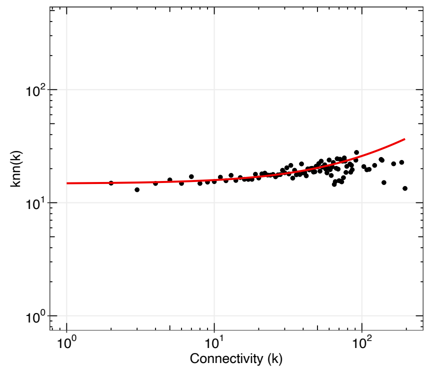
From eq. (1), we calculated the degree correlation function of our networks, using the degree k of the nodes. Figure 2 shows a comparation between the network of shallow earthquakes built using the successive model and the time window model. It is observed that, in both distributions, the nearest-neighbors average connectivity of nodes, $k_{nn}(k)$, increases linearly with k, which means these networks are assortative. Therefore, the nodes with high degree connect, on average, to nodes with high degree. This result was the same found in networks of earthquakes from California and Japan (Abe and Suzuki, 2006), which makes sense, since most earthquakes that occur in these areas have depth up to 70 km (shallow earthquakes). However, the network constructed with the time window model is more assortative than the one constructed with the successive model.

It is also interesting to note that, in Figure 2, the linear growth does not hold for high degree-nodes. This is expected in scale-free networks, like ours, since the system is unable to sustain assortativity for high-degree, because there are only a few nodes with large values of degree (hubs). The same happens for other real networks, as citation networks.

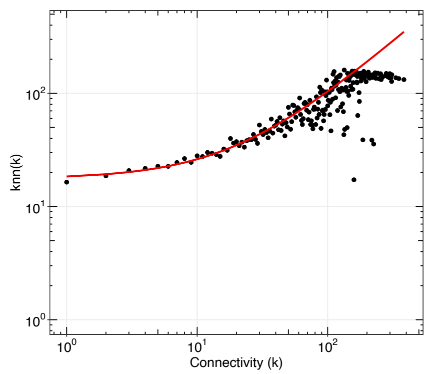
The results for the networks of deep earthquakes are shown in Figure 3. In both cases, successive and time window model, the networks are neutral, i.e., $k_{nn}(k)$ is independent of k.

Finally, as shown in Figure 4, we found that the nearest-neighbors average connectivity of nodes, $k_{nn}(k)$, for the network of earthquakes simulated with the modified OFC model has an increasing behavior with k, in agreement with the results found for shallow earthquakes, in Figure 2. Similar results were also found in a previous work conducted for networks built using the standard OFC model (Peixoto and Prado, 2006).

In addition, the degree correlation coefficient was calculated from eq. (2) for each of our networks and the values are shown in Table 1. As it can be seen by the positive values obtained, the networks of shallow events (for both successive and time window models) and the network of earthquakes of the modified OFC model are assortative, in agreement with the results for the nearest-neighbors average connectivity of nodes. Once again, we



(a)



(b)

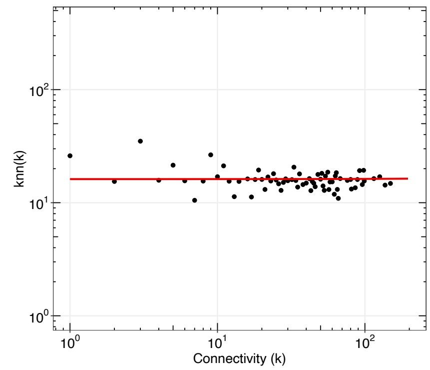
Figure 2 *- Nearest-neighbors average connectivity of nodes, $k_{nn}(k)$, for the network of shallow earthquakes using (a) the successive model and (b) the time window model. It can be observed that both distributions follow a crescent linear fit (red line). This shows that both networks have assortative mixing, being the network constructed with the time window model much more assortative.*

have noticed that the network of shallow earthquakes constructed with the time window model is much more assortative.

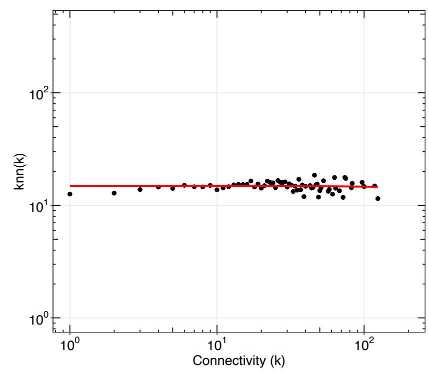
It is noteworthy that the r value found in our shallow earthquake network, for the time window model, is consistent with an analogous finding for Japan data, using a different technique from ours (Tenembaum et al., 2012). As the Japan region has a predominance of shallow earthquakes, our results make sense.

The networks of deep earthquakes for both models, successive and time window, present $r \approx 0$, indicating they are neutral, as it was found in the nearest-neighbors average connectivity of nodes analyzes.

The assortative behavior exhibited in the networks of shallow and synthetic earthquakes is an interesting result. Firstly, because it makes the giant component of the networks not able to reach large values. Once the giant component is the largest subset of nodes in the network (where each of its nodes must be connected to at least one other node), the assortative result means that in earthquake networks the regions with higher degrees are forced to connect much more to each other, than to small



(a)



(b)

Figure 3 *- Nearest-neighbors average connectivity of nodes, $k_{nn}(k)$, for the network of deep earthquakes constructed with (a) the successive model and (b) the time window model. No correlation between $k_{nn}(k)$ and k is presented in both distributions, which means that both networks are neutral.*

degree nodes, making the hubs have more probability to belong to the giant component. Another point is that in assortative networks hub removal makes less damage because the hubs form a core group, which means that in our earthquake network the features described by our results will not experience significant changes if, in some period of time, do not occur earthquakes in one of the regions of high degree of the network.

We also emphasize that the correlation properties found in the present work concern to both temporal and spatial correlations, since we use time sequences to establish links between different spatial regions of the globe. It reinforces the idea present in previous works about long-range correlations in seismic activities.

Conclusions

In this work, the assortativity of networks of worldwide and synthetic earthquakes were studied, in order to better characterize correlation properties and understand the spread of information in the system of earthquakes. With respect to the global events, we analyzed shallow and deep earthquakes separately, since they have different seismic origins. From these datasets, we built networks of

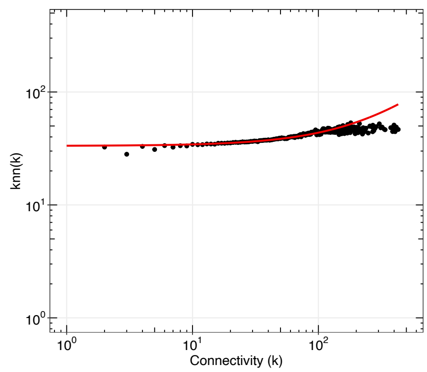


Figure 4 *- Nearest-neighbors average connectivity of nodes, $k_{nn}(k)$, for the network of earthquakes generated with the modified OFC model using the successive model. This network has $k_{nn}(k)$ increasing linearly with k (red line), therefore it is assortative.*

epicenters by using two methodologies: the successive model and the time window model of connections. For the shallow earthquakes, the networks of both models of linking nodes presented assortative mixing, with the difference that the network constructed with the time window model was found to be much more assortative. These results were similar to the ones obtained for networks of seismic events from California and Japan (regions with predominance of shallow earthquakes).

The assortative mixing was also found for the network created using a catalog produced by an improved version of the computational model proposed by Olami, Feder and Christensen (OFC model), which shows agreement between real data for shallow events and synthetic data catalogs. On the other hand, the deep earthquakes networks, for both the successive model and the time window model, presented no correlation between the degree of the nodes.

Our results suggest that shallow and deep earthquakes have different temporal and spatial correlation properties. While for shallow earthquakes we have positive degree correlations, for deep earthquakes these correlations seem to not exist. Because of that, the shallow earthquakes networks have a tendency to link high-degree regions with other high-degree regions, making more difficult to change the seismological behavior of the earthquake networks if a certain region stops to have earthquakes for a period of time.

The present study shows that, when analyzing an earthquake network, more than just degree distributions must be investigated, but it is also necessary to observe how the nodes of the network are related. This way, we have shown that the presence of mixing patterns in earthquake networks have a profound effect on the topological properties of the network as it affects the detailed wiring of links among nodes.

The results obtained in this study contribute to a better understanding of the dynamics of the seismological phenomenon and a possible spatial and temporal correlation in global and synthetic data.

Table 1 *- The number of nodes (N) and the values of the degree correlation coefficient (r) of the networks of worldwide and synthetic earthquakes used in this study.*

Network	***N***	***r***
Shallow earthquakes (successive)	28471	0.0711
Shallow earthquakes (time window)	23380	0.508
Deep earthquakes (successive)	8958	0.000763
Deep earthquakes (time window)	7675	0.0152
OFC model earthquakes	115510	0.0750

For future works, we intend to study the earthquake clusters and communities performed from deep and shallow earthquakes.

Acknowledgments

J.R. thanks for the scholarship of FAPERJ, Brazil in this research, P.S.L.O thanks for the scholarship of CNPq, Brazil, R.P.F thanks FAPERJ, Brazil for productivity grant, A.P. thanks CNPq, Brazil for productivity grant and D.S.R.F, R.P.F and A.R.P thanks FAPERJ, Brazil for grant number 241029.

References

ABE, S.; SUZUKI, N. 2004a. Scale-free network of earthquakes. *EPL (Europhysics Letters)*, 65(4), 581.

ABE, S.; SUZUKI, N. 2004b. Small-world structure of earthquake network. *Physica A: Statistical Mechanics and its Applications*, 337(1-2), 357-362.

ABE, S.; SUZUKI, N. 2006. Complex earthquake networks: Hierarchical organization and assortative mixing. *Physical Review E*, 74(2), 026113.

ALBERT, R.; BARABÁSI, A. L. 2000. Topology of evolving networks: local events and universality. *Physical Review Letters*, 85, 5234.

ALBERT, R.; BARABÁSI, A. L. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74, 47.

BARABÁSI, A. L.; ALBERT, R. 1999. Emergence of scaling in random networks. *Science*, 286, 509-512.

BARABÁSI, A.-L.; PÓSFAI, M. 2016. *Network science*. Cambridge: Cambridge University Press. ISBN: 9781107076266 1107076269

D'AGOSTINO, G.; SCALA, A.; ZLATIĆ, V.; CALDARELLI, G. 2012. Robustness and assortativity for diffusion-like processes in scale-free networks. *EPL (Europhysics Letters)*, 97(6), 68006.

DOROGOVTSEV, S. N.; MENDES, J. F. 2003. *Evolution of networks: From biological nets to the Internet and WWW*. OUP Oxford.

FERREIRA, D. S.; PAPA, A. R.; MENEZES, R. 2014. Small world picture of worldwide seismic events. *Physica A: Statistical Mechanics and its Applications*, 408, 170-180.

FERREIRA, D. S.; PAPA, A. R.; MENEZES, R. 2015. On the agreement between small-world-like OFC model and real earthquakes. *Physics Letters A*, 379(7), 669-675.

FERREIRA, D.; RIBEIRO, J.; PAPA, A.; MENEZES, R. 2018. Towards evidence of long-range correlations in shallow seismic activities. *EPL (Europhysics Letters)*, 121(5), 58003.

FERREIRA, D. S.; RIBEIRO, J.; OLIVEIRA, P. S.; PIMENTA, A. R.; FREITAS, R. P.; PAPA, A. R. 2020. Long-range correlation studies in deep earthquakes global series. *Physica A: Statistical Mechanics and its Applications*, 560, 125146.

GRASSBERGER, P. 1994. Efficient large-scale simulations of a uniformly driven system. *Physical Review E*, 49(3), 2436.

GHEIBI, A.; SAFARI, H.; JAVAHERIAN, M. 2017. The solar flare complex network. *The Astrophysical Journal*, *847*(2), 115.

GUTENBERG, B.; RICHTER, C. F. 1942. Earthquake magnitude, intensity, energy, and acceleration. *Bulletin of the Seismological society of America*, 32(3), 163-191.

JEONG, H.; MASON, S. P.; BARABÁSI, A. L.; OLTVAI, Z. N. 2001. Lethality and centrality in protein networks. *Nature*, 411(6833), 41-42.

NEWMAN, M. E. 2002. Assortative mixing in networks. *Physical Review Letters*, 89(20), 208701.

NEWMAN, M. E. 2003. Mixing patterns in networks. *Physical Review E*, 67(2), 026126.

OLAMI, Z.; FEDER, H. J. S.; CHRISTENSEN, K. 1992. Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical review letters*, 68(8), 1244.

PASTOR-SATORRAS, R.; VÁZQUEZ, A.; VESPIGNANI, A. 2001. Dynamical and correlation properties of the internet. *Physical Review Letters*, 87(25), 258701.

PEIXOTO, T. P.; PRADO, C. P. 2006. Network of epicenters of the Olami-Feder-Christensen model of earthquakes. *Physical Review E*, 74(1), 016126.

TENENBAUM, J. N.; HAVLIN, S.; STANLEY, H. E. (2012). Earthquake networks based on similar activity patterns. *Physical Review E*, 86(4), 046107.

VÁZQUEZ, A.; PASTOR-SATORRAS, R.; VESPIGNANI, A. 2002. Large-scale topological and dynamical properties of the Internet. *Physical Review E*, 65(6), 066130.

WATTS, D. J.; STROGATZ, S. H. 1998. Collective dynamics of ‘small-world’networks. *Nature*, 393, 440-442.