Long-range correlation studies in deep earthquakes global series

Douglas S. R. Ferreira, Jennifer Ribeiro, Paulo S. L. Oliveira, André R. Pimenta, Renato P. Freitas, Andrés R. R. Papa

Abstract

In the present paper we have conducted studies on seismological properties using worldwide data of deep earthquakes (depth larger than 70 km), considering events with magnitude $m \geq 4.5$. We have addressed the problem under the perspective of complex networks, using a time window model to build the networks for deep earthquakes, which present scale-free and small-world features. This work is an extension of a previous study using a similar approach, for shallow events. Our results for deep events corroborate with those found for the shallow ones, since the connectivity distribution for deep earthquakes also follows a $q$-exponential distribution and the scaling behavior is present. Our results were analysed using both, complex networks and Nonextensive Statistical Mechanics, contributing to strengthen the use of the time window model to construct epicenters networks. They reinforce the idea of long-range correlations between earthquakes and the criticality of the seismological system.

Keywords:
Earthquakes; Complex networks; Nonextensive statistical mechanics; Long-range correlations

1. Introduction

Several phenomena in nature exhibit characteristics of complex systems such as nonlinear dynamics, fractal dimensions, power law distributions and long-range spatiotemporal memory, where the earthquakes being one example. The natural seismic phenomenon has the power to produce huge impacts in the human society, once that the results coming from that phenomenon can be devastating, thus the search for a better understanding of seismic dynamics becomes highly important. One way to improve the understanding about seismological properties is the analysis of the statistical behavior of earthquakes data. The process of seismological data mining can lead to find the arising of many patterns, which have been studied over the time many researchers, e.g. the celebrated laws known as the Gutenberg-Richter law [1], for the relationship between the magnitude and the frequency of earthquakes, and the Omori law, for the rate of aftershocks following a main shock [2].

Theories of complex systems have been applied to many different areas of knowledge for a long time, such as economics [3, 4], computer science [5], mechanical engineering [6], biology [7, 8] and chemistry [9]. In the last decades, works have used complex systems theories to perform studies on spatiotemporal properties of seismicity. In that sense, we can highlight the important approach of using complex networks concepts, which provides powerful procedures to analyse the interactions and correlations between elements of complex systems, giving efficient descriptions about the dynamics of such systems. The earthquake network formulation was developed by Abe and Suzuki [10] to define the evolving network associated with successive earthquakes. After that, a large number of
works have used earthquake networks to study different properties of the complex seismic phenomenon, for both real [11, 12, 13, 14, 15, 16, 17, 18] and synthetic data produced by computational models [19, 20, 21, 22, 23].

However, despite the various studies made using concepts of complex systems, we have a very limited set of works considering earthquakes occurred in large geographic regions, specially for the entire world. As reported by different authors, many studies suggest that earthquakes may induce other earthquakes several hundred miles away from the rupture zone, since there is no consensus about a physical length scale for aftershocks zones [24, 25, 26, 27, 28, 29]. Considering that perspective, Ferreira et al. [30] have analysed natural seismic events for the globe, investigating the epicenter network (where two regions will be connected by an edge if two successive events occur in those respective regions) and the time interval between successive earthquakes (also referred as “calm time” or “inter-event time”). The network resulting from this process has been found to being scale-free and small-world-like, with the probability distribution for the time interval between successive earthquakes presenting nonextensive characteristics. It is noteworthy that Nonextensive Statistical Mechanics (NESM) is a theory very useful to explain complex systems where the standard Boltzmann–Gibbs statistical mechanics does not seem to apply, i.e., systems with strong interactions between their elements, long-range spatiotemporal correlations and fractal geometries. Recently, in [31], the authors have improved the work done in [30] by modifying the way to build the epicenter network. To do that they have defined a “time window” to connect different regions over the world, where the region corresponding to the first event is connected to all regions within this window by directed edges but respecting the time order of earthquakes. Thereby, using data for shallow earthquakes for the entire world, the authors find nonextensive characteristics also in the epicenter network, keeping scale-free and small-world properties.

Nevertheless, there is no such work made for deep earthquakes, which restricts the complex network analysis of the worldwide seismological phenomenon. Furthermore, it is important to highlight that, in South America, for example, intermediate-deep earthquakes produced the highest numbers of fatalities of all earthquakes in the twentieth century [32], pointing once again the importance of studying deep earthquakes.

Following the methodology used in [31] for shallow earthquakes, in the present work we will consider only deep earthquakes to build the epicenter network in order to calculate topological properties that characterize the network, producing a complementary study about epicenter network studies on earthquakes all over the world. We have analysed our results under the perspective of complex network and NESM theories, also taking into consideration arguments of finite data-size scaling.

In the next sections we present the seismological data used and the methodology employed to construct the epicenter network as well as our results, analyses and considerations.

2. Data Selection

The seismological data used to perform our analyses was taken from the World Catalog of Earthquakes of Advanced National Seismic System (ANSS) covering the entire world for the period between 2002 and 2016. The magnitude types considered were $M_b$ (body-wave magnitude), $M_d$ (duration magnitude), $M_l$ (local magnitude) and $M_w$ (moment magnitude). It is relevant to highlight that we have considered only earthquakes with magnitudes larger than or equal to 4.5, once in that catalog events with magnitudes less than 4.5 are not completely registered for the whole world. In addition, we have also disconsidered all non-earthquake events, e.g. quarry blasts, sonic booms, rock bursts and nuclear explosions. With that the total number of events in our dataset is 101,746.

At this point it is noteworthy that we have separated shallow from deep earthquakes data. In the present study we will analyse only the deep ones, i.e., the events occurred at depths greater than 70 km, once the shallow earthquakes have already been studied in [31]. That separation has the objective to make comparisons between earthquakes with similar seismic origins. After this differentiation we get a number of 21,226 deep events. In order to perform a preliminary check of our data, we have verified that earthquakes’ magnitude data follows the Gutember-Richter law with a $b$-value exponent equal to 1.08, as expected that value is close to 1.00, since we have excluded events with small magnitudes.

1https://earthquake.usgs.gov/data/comcat/
3. Method

As previously mentioned, the technique used in the present paper, to study our dataset, is the construction of complex networks from earthquake data. To construct and analyse the epicenter network of our data we have followed the same method proposed in [31], so this section is dedicated to present a brief review and explanation of that method.

The most commonly method used for the construction of earthquake complex networks is based on the method proposed by Abe and Suzuki [10], where the geographical region under consideration is divided into cells of the same size, which will become vertices of the network if earthquake epicenters occur therein. Two successive events define an edge between two vertices.

However, a new procedure to build the epicenter network was proposed by Ferreira et al. [31], because the authors have observed that the standard method of construction, which considers edges only between successive events, presents problems in the cumulative probability distribution of connections for the global network, indicating a loss in the connectivity \(k\) of vertices with large number of connections [30]. A summary of that procedure is as follows. First, the Earth’s surface is divided into equal square cells of side \(L \times L\) where a cell will become a vertex of the network every time the epicenter of an earthquake is located therein. The spatial determination of the cell of each epicenter is made by using the latitude \(\theta_E\) and longitude \(\phi_E\) of each epicenter relative to a given reference, \(\theta_0\) and \(\phi_0\). Adopting those coordinates and using the spherical approximation for the Earth (with radius \(R = 6.371 \times 10^3\) km), we can calculate the distances North-South \(S_{ns}\) and East-West \(S_{ew}\) by

\[
S_{ns} = R \cdot \theta_E \\
S_{ew} = R \cdot \phi_E \cdot \cos \theta_E,
\]

where, for the sake of simplicity, we have considered \((\theta_0, \phi_0) = (0, 0)\). We can also note that, regardless the region of the planet, the cells have always the same size, \(L \times L\).

Secondly, it is defined a time window to create the connections (edges) between the vertices. Keeping the chronological order, the time window, \(T\), is placed on the data so that the first vertex inside the window is connected to all other vertices within that window. After that, the window is moved forward to the next event and the connections procedure is replicated. It is repeated until it is no longer possible to move the window forward. It is worth mentioning that the viewpoint adopted with this construction mechanism allows us to better study the long-range relationships between the cells, since it takes into consideration more intensely the relationships between each cell inside the window, which does not happen for the network construction model of successive events.

Thirdly and finally, it must be find the best time window size to be used in epicenter network construction. To accomplish that it is necessary to define an earthquake network, which is different from the epicenter network. The earthquake network is a network of connections between the events and the epicenter network is a spatial network of cells where the earthquakes occurred. Fig.1 brings to us an exemplification about the difference between epicenter network and earthquake network. By applying different time window values in data to construct the respective earthquake networks and by using the Louvain method to calculate the number of communities [33], it is possible to observe that the number of communities in the earthquake networks has a maximum value regarding the time window size. It means that at this maximum the network has a configuration where the connections between the vertices are distributed in the better possible way in order to detect long-range correlations between the events and cells. We highlight here that the “best time window” is a window not too large, which would create many links between earthquakes that probably are not correlated, but neither too small, once it would brings to us a very fragmented network in which many earthquakes that have correlations between them are not connected.

Once the best time window value is found in the earthquake network, this value is also applied to the spatial epicenter network, since this will be the network used for study the spatiotemporal correlations.

4. Results and discussion

As mentioned before, a way to improve our knowledge about the earthquake dynamics is through the spatiotemporal analysis of earthquakes data.

The viewpoint we have implemented to study the seismic phenomenon is the analysis of a complex network created from the earthquakes data. In this work we conducted a similar study to the one done in [31] but using data
of deep earthquakes, i.e., for events with hypocenters deeper than 70 km. We have used the procedure described in the previous section to build the networks and to find the best time window value for our deep earthquakes data. In Fig [2], it is observed that small windows will cause an excessive segmentation of the earthquakes network, represented by the low number of communities, and large windows will produce big clusters, which will also cause a decrease in the number of communities. In this way, the best time window value is given when the number of communities is maximum, i.e., when \( T = 16\, 500 \) s.

In [31], Ferreira et al. showed, for the first time, that the distribution of connectivities for the worldwide network of shallow earthquakes constructed using the “time window model” has a scale-free behavior and it is invariant with respect to the window size. The scale-free behavior was also found in studies for specific regions of the world using the successive model of connections [10, 11].

In Fig [3(a)] we have the density probability distribution of connections for global networks of epicenters with cells of size 20 km \( \times 20 \) km, constructed using both the successive and the time window models, where the last was made utilizing the best time window value found. In both cases it is possible to observe that the density probability distribution has a power law behavior, leading to the idea that the growth rule of this network follows a preferential attachment.

This result is not exactly the same as the one found in [31], since for shallow earthquakes the density probability distributions of connections for successive and time window models are different, while here, for deep events, they are very similar. However, as can be seen in Figs [3(b) and 3(c)] when we take the cumulative probability distribution we
recover the same result found for shallow earthquakes, i.e., the successive and time window models have not the same behavior. A possible explanation to this fact could be that, for deep events the earthquake network (for time window model) have the majority of its cells with connectivity (in or out) equal to one, so this fact makes the connectivity probability distributions for time window model to be not too different from the one found for the successive model, causing that the difference between the networks built from the two different models is only noticeable when we look at the cumulative distribution.

Fig 5(b) shows that for the successive model the connectivity cumulative distribution has a pure power law behavior only for small values of \( k \), because that distribution is actually better fitted by a power law with exponential cutoff, \( P(k) \sim k^{-\alpha}e^{-k/k_c} \). It is noteworthy that a very similar result was also found for California data, where the exponential cutoff does not seem to exist in probability density plots, but it appears in the cumulative ones, because in a density plot the fluctuations are higher than in a cumulative probability plot, as reported in [12, 30].

When using the time window model to construct the network of epicenters, the cumulative probability plot follows the non-traditional function \( q \)-exponential, \( P_{\geq}(k) = A[1-(1-q)\beta x]^{1/(1-q)} \), as shown in Fig 5(c). This function belongs to the scope of the Nonextensive Statistical Mechanics theory [34], which adopts the Tsallis’ entropy, that can explain the statistical properties of a variety of complex systems at their quasi-equilibrium states with characteristics such as long-range interaction between its elements and long-range temporal memory. The \( q \)-exponential function appears naturally from the maximization of the Tsallis’ entropy under appropriate constraints and it is a generalization of the exponential curve, where taking the limit \( q \to 1 \) we recover the standard exponential and for \( q > 1 \) it exhibits power law asymptotic behavior, for larger values of \( k \). That result is in close agreement with the results found in previous works for shallow earthquakes [31] and for networks built from synthetic epicenters data produced using a small-world topology version of the OFC model [23].

We have also noticed that it is possible to make the connectivity distribution invariant with respect to the value of the time window by using the scale function

\[
P_{\geq}(k, T) = T^{-\delta}f(k/T^\lambda),
\]

where \( \delta = 1.80, \lambda = 1.00, f(x) \) decays as \( x^{-\gamma} \) with \( \gamma = 1.80 \) and the data collapse is in agreement with the scaling hypothesis, as can be seen in Fig 4. It should be noted that for shallow earthquakes the scaling property was also present [31].

In order to complement our complex network study on seismological deep data we have calculated two important metrics, which are, the clustering coefficient (C) and the average shortest path (\( \ell \)). These metrics are important because they are used to characterize networks with dense connectivity areas and long jumps between these areas.
These are called small-world networks, which have small average shortest path when compared to the number of vertices \((N)\) and high clustering coefficient when compared to a similar random network.

The measurements of \(C\) and \(\ell\), were performed using the algorithms described in [13] and [16], respectively. The results for our epicenters network, using the time window model, were \(C = 4.91 \times 10^{-3}\) and \(\ell = 5.01\), with \(N = 7675\). For a random network with the same number of vertices the clustering coefficient is \(C_{\text{rand}} = 5.88 \times 10^{-4}\). These results reveal that, similarly to what occurs in shallow earthquakes, deep ones also exhibit small-world characteristics in the global network of epicenters created from the time window model, given that \(C \gg C_{\text{rand}}\) and that \(\ell \ll N\) \((\ell \approx \ln N)\) [37]. We point out here that the small-world property does not appear when considering the successive model for the network construction, once its clustering coefficient is \(C_{\text{succ}} = 9.00 \times 10^{-3}\), while for this network \(C_{\text{rand}} = 5.26 \times 10^{-4}\), which means that it does not have \(C_{\text{succ}} \gg C_{\text{rand}}\), despite of having a small average path length \((\ell = 4.73, \text{with } N = 8958)\). The complex networks properties found by the time window modelling is in strong agreement with many previous works, where the small-world and scale-free features arise in networks of earthquakes [11, 12, 13, 14, 15, 30, 16, 31, 17, 18].

Moreover, we obtained a geospatial image of our network constructed with the time window model, in order to
visualize it. Using informations such as the exact geographical location of each cell, the number of connections of each cell and to which other cells they are connected to, it was possible to create with the software Gephi the image shown in Fig 5. It was observed that the cells are located in regions where usually occur deep earthquakes (mostly in subduction zones at convergent plate boundaries) and that the cells with the highest connectivity are located in the Pacific Ocean near the Fiji Islands, in Molagavita-Colombia, in the Coral Sea near Vanuatu, in Salta-Argentina and in Jurm-Afghanistan, respectively. It seems to correspond the geographic regions with very intense deep seismic activity, but not necessarily with the greatest earthquakes in the period we have studied. The biggest hub obtained in deep earthquakes, for instance, does not hold the seism with greatest magnitude in our dataset, but it is located near the Fiji Islands, belonging to the Tonga-Kermadec region, which is known for being the region where deep earthquakes occur more than in any other place around the world, being the probable cause of this remarkable activity the subduction of the 70–100 Ma-old Pacific plate beneath the Fiji Plateau. That result is different from the one found in shallow earthquakes, where the hubs coincide with regions that have occurred the seisms with the largest magnitudes, which makes a sense, given that deep earthquake aftershock sequences generally show aftershock productivity rates of one to three orders of magnitude less than shallow earthquakes of equivalent size, making the smaller number of aftershocks observed for deep earthquakes one of the most clear difference between shallow and deep earthquakes [38].

5. Conclusions

In this work we have considered the time window model presented in [31] to construct a complex network of epicenters for deep events. Similarly to what was done for shallow earthquakes, our results show that the epicenter network built using the time window model proves to be better and more effective than the successive model in finding long-range relationships between the network vertices.

For deep earthquakes the time window model is also able to identify the regions with intense activity of deep earthquakes, even if these regions are not those with the largest earthquakes. This is a relevant result, given that for deep events, the occurrence of large events is not directly related to high aftershocks rates, where it can be demonstrated by the Tonga subduction zone example. Although the Tonga subduction zone has the largest number of deep

\[ P(k) \geq 1 \]

\[ T = 3000s, T = 5000s, T = 10000s, T = 14000s, T = 16500s, T = 18000s, T = 20000s \]

\[ T \delta \cdot P(k)/T \lambda = 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2} \]

\[ T = 3000s, T = 5000s, T = 10000s, T = 14000s, T = 16500s, T = 18000s, T = 20000s \]

\[ T = 3000s, T = 5000s, T = 10000s, T = 14000s, T = 16500s, T = 18000s, T = 20000s \]

\[ q - \text{exponential} (q = 1.42) \]

\[ q = 1.42 \pm 0.10 \]
earthquakes, South America hosts the largest-magnitude deep ruptures. However, while Tonga has about ten times more events with $m = 5.5$, the two regions have similar numbers of $m = 6.5$ events, and the biggest deep earthquakes occur in South America [39].

The present study also strengthens the idea of long-range correlations between earthquakes, once we have the presence of nonextensive characteristics in the epicenter network of deep earthquakes, as well as the scale-free and small-world properties. Besides that the connectivity distribution exhibit scaling properties with respect to the time window size chosen, since that distribution can be made independent from the size of the time window, using a scale function.

These results are in close agreement with previous works that have been indicating spatial and temporal long-range correlations in both local and worldwide views. The outcomes of this paper also become interesting because they reinforce the criticality idea in the seismological system.

For future works, we plan to do a deeper analysis about the communities structure present in the epicenter network, as well as the implications of perform studies on spatial and temporal probabilities distributions under the NESM view. The results will be published elsewhere.

Acknowledgements

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.R.R.P thanks CNPq, Brazil for productivity grant and D.S.R.F, R.PF and A.R.P thank FAPERJ, Brazil for grant number 241029.

References

References

[1] B. Gutenberg, C. Richter, Earthquake magnitude, intensity, energy, and acceleration, Bull. Seism. Soc. Am. 32 (3) (1942) 163–191.
