HAL
open science

Small world in a seismic network: the California case

A. Jiménez, K. F. Tiampo, A. M. Posadas

To cite this version:

A. Jiménez, K. F. Tiampo, A. M. Posadas. Small world in a seismic network: the California case. Nonlinear Processes in Geophysics, 2008, 15 (3), pp.389-395. hal-00302996

HAL Id: hal-00302996

https://hal.science/hal-00302996

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nonlin. Processes Geophys., 15, 389–395, 2008
www.nonlin-processes-geophys.net/15/389/2008/
© Author(s) 2008. This work is distributed under
the Creative Commons Attribution 3.0 License.

Nonlinear Processes in Geophysics

Small world in a seismic network: the California case

A. Jiménez[1], **K. F. Tiampo**[1], **and A. M. Posadas**[2]

[1]Department of Earth Sciences Biological and Geological Sciences, University of Western Ontario, London, Canada
[2]Department of Applied Physics, University of Almería, Almería, Spain

Received: 31 January 2008 – Revised: 4 April 2008 – Accepted: 5 April 2008 – Published: 9 May 2008

Abstract. Recent work has shown that disparate systems can be described as complex networks i.e. assemblies of nodes and links with nontrivial topological properties. Examples include technological, biological and social systems. Among them, earthquakes have been studied from this perspective. In the present work, we divide the Southern California region into cells of 0.1°, and calculate the correlation of activity between them to create functional networks for that seismic area, in the same way that the brain activity is studied from the complex network perspective. We found that the network shows small world features.

1 Introduction

Physics, a major beneficiary of reductionism, has developed an arsenal of successful tools to predict the behavior of a system as a whole from the properties of its constituents. The success of these modeling efforts is based on the simplicity of the interactions between the elements: there is no ambiguity as to what interacts with what, and the interaction strength is uniquely determined by the physical distance. We are at a loss, however, in describing systems for which physical distance is irrelevant, or there is ambiguity whether two components interact (Albert and Barabási, 2002).

Historically, the study of networks has been mainly the domain of a branch of discrete mathematics known as graph theory. Since its birth in 1736, when the Swiss mathematician Leonhard Euler published the solution to the Königsberg bridge problem (consisting in finding a round trip that traversed each of the bridges of the Prussian city of Königsberg exactly once), graph theory has witnessed many exciting developments and has provided answers to a series of practical questions such as: what is the maximum flow per unit time from source to sink in a network of pipes, how to color the regions of a map using the minimum number of colors so that neighboring regions receive different colors, or how to fill n jobs by n people with maximum total utility. In addition to the developments in mathematical graph theory, the study of networks has seen important achievements in some specialized contexts, as for instance in the social sciences. Social networks analysis started to develop in the early 1920s and focuses on relationships among social entities such as communication between members of a group, trades among nations, or economic transactions between corporations (Boccaletti et al., 2006).

Recent work has shown that disparate systems can be described as complex networks, that is, assemblies of nodes and links with nontrivial topological properties, examples of which include technological, biological and social systems (Eguíluz et al., 2005). Among them, earthquakes also have been studied from this perspective (Abe and Suzuki, 2004, 2006a; Baiesi and Paczuski, 2005). In the past few years, the discovery of small-world and scale-free properties of many natural and artificial complex networks has stimulated a great deal of interest in studying the underlying organizing principles of various complex networks, and has led to dramatic advances in this emerging and active field of research (Wang and Chen, 2003). Here we present for earthquake fault systems a similar approach to that of Eguíluz et al. (2005) for functional brain networks, and find that the analyzed catalog has small-world behavior.

The small world concept, in simple terms, describes the fact that despite their often large size, in many networks there is a relatively short path between any two nodes. The distance between two nodes is defined as the number of edges along the shortest path connecting them. The most popular manifestation of small worlds is the six degrees of separation concept, uncovered by the social psychologist Milgram (Milgram, 1967), who concluded that there was a path of acquaintances with typical length about six between most pairs of people in the United States (Kochen, 1998). Watts and

Correspondence to: A. Jiménez
(ajlloret@ual.es)

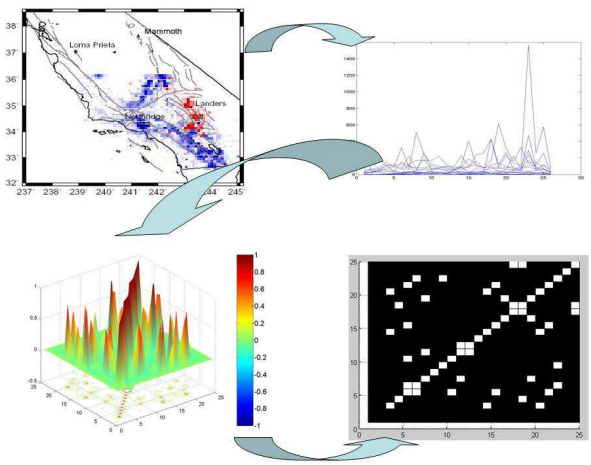

Fig. 1. Scheme of the method.

Strogatz (1998), in their seminal paper, have proposed to define small-world networks as those networks having both a small value of Lp (characteristic path length), like random graphs, and a high clustering coefficient C, like regular lattices. They consider a one-dimensional graph with N nodes, each vertex being connected to its k nearest neighbors (where $Nkln[N]$). The number k of edges per vertex is also called the degree of the graph. Next, with a probability P, a random edge is chosen and rewired to connect to a randomly chosen vertex. By varying P between 0 and 1 graphs can be created which span the whole range from regular (P=0) to random (P=1).

Two measures were introduced to characterize such graphs: the characteristic path length Lp is the mean of the shortest path (expressed in number of edges) connecting any two vertices on the graph. The cluster coefficient Cp is the likelihood (between 0 and 1) that the kv neighbors of vertex v are also connected to each other, averaged over all vertices. Regular networks or graphs have a high Cp ($Cp \approx 3/4$) but a long characteristic path length ($Lp \approx N/2k$); random graphs have a low $Cp(k/N)$ but the shortest possible path length ($Lp \approx ln(N)/ln(k)$). The discovery of Watts and Strogatz was that some networks with $0 < P \ll 1$, thus regular networks with only a very small number of random edges, have a path length that is much smaller than that of a regular network, while the Cp is still close to that of a regular network. This dramatic drop in Lp for P only slightly higher than 0 implies that any vertex on the graph can be reached from any other vertex in only a small number of steps. This is equivalent to the small-world phenomenon and this type of graph (Cp close to regular network; Lp close to random network) was called a small-world graph by Watts and Strogatz. They showed that many real world networks such as networks of actors playing in the same movies, the power grid of North America, and the neuronal network of *Caenorhabditis elegans* have small-world features. Furthermore, they suggested that such networks may be optimal for information processing in complex systems. Since then it has been shown that many real networks display small world features and that these may reflect an optimal architecture for information processing (Stam, 2004).

2 Method

For a network (or graph) representation, first we have to define the *nodes* and the *edges*. Figure 1 shows a scheme of the method. The seismic region is divided into squared cells (for latitude and longitude only in this particular case), which will be the nodes; the time is divided into intervals. At each time step (we will try some, from days to several years), the activity, $a(x,t)$ of the cell is calculated as the number of earthquakes at that cell and time. Now we have a time series for each cell. For each pair of cells, $x1$ and $x2$, we calculate their correlation coefficient in this way:

$$r(x1,x2) = \frac{\langle a(x1,t)a(x2,t)\rangle - \langle a(x1,t)\rangle \langle a(x2,t)\rangle}{\sigma(a(x1))\sigma(a(x2))} \quad (1)$$

where $\sigma^2(a(x)) = \langle a(x,t)^2\rangle - \langle a(x,t)\rangle^2$, and $\langle\cdot\rangle$ represents temporal averages.

Then, a threshold matrix is calculated for different values of the correlation coefficient, r_c, so that when the correlation between two cells (nodes) is higher than the threshold value (positive values of r_c only), we say that they are positively correlated, and the nodes (cells) are connected by an edge. Once our network is defined, we proceed to analyze its properties.

2.1 Node degree and degree distributions

The degree (or connectivity) d_i of a node i is the number of edges incident with the node, and is defined in terms of the adjacency matrix A as:

$$d_i = \sum_{j=1}^{N} a_{ij} \quad (2)$$

The most basic topological characterization of a graph can be obtained in terms of the degree distribution $p(d)$, defined as the probability that a node chosen uniformly at random has degree d or, equivalently, as the fraction of nodes in the graph having degree d.

2.2 Shortest path lengths and diameter

The shortest path is the geodesic distance between vertex pairs in a network. The mean geodesic l is then:

$$l = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} g_{ij} \quad (3)$$

where g_{ij} is the geodesic distance from vertex i to vertex j, and n is the number of nodes. The maximum value of

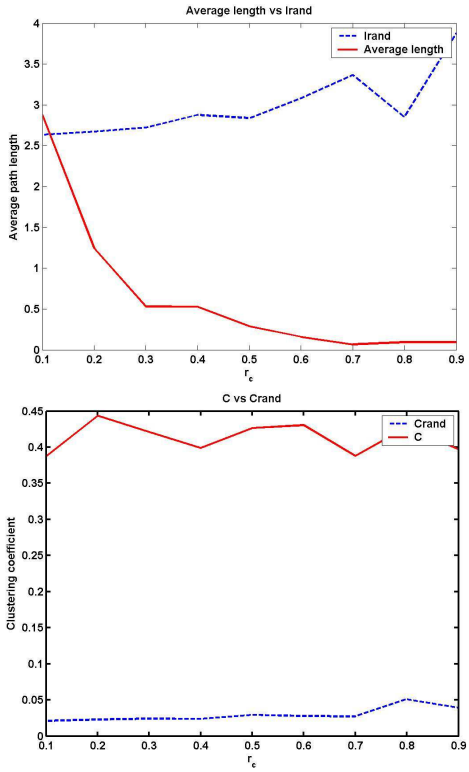

Fig. 2. Average length and clustering coefficient of the networks with different thresholds (r_c) compared to those corresponding to a random graph with the same number of nodes and the same average node degree, for time lags of 1 day.

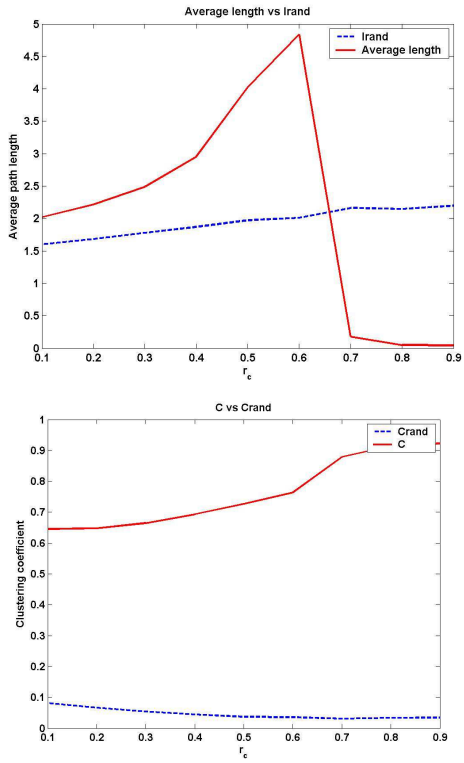

Fig. 3. Average length and clustering coefficient of the networks with different thresholds (r_c) compared to those corresponding to a random graph with the same number of nodes and the same average node degree, for time lags of 100 days.

g_{ij} is called the diameter of the graph. We used Dijkstra's algorithm to implement this calculation (Dijkstra, 1959).

2.3 Clustering

A clear deviation from the behavior of the random graph can be seen in the property of network transitivity, sometimes also called clustering (Newman, 2003). Here we use the definition by (Watts and Strogatz, 1998), that has found wide use in numerical studies and data analysis (Newman, 2003):

$$C = \frac{1}{n}\sum_i C_i \tag{4}$$

$$C_i = \frac{\text{number of triangles connected to vertex i}}{\text{number of triples centered on vertex i}} \tag{5}$$

where *triple* means a single vertex with edges running to an unordered pair of others. The clustering coefficient measures the average density of triangles in a network. For random networks, C tends to zero as n^{-1} in the limit of large system size.

3 Data

The catalog belongs to the Southern California Earthquake Center (SCEC) and contains the seismic data for the period 1 January 1984 to 3 July 2001. The analyzed area ranges from 32–37 N, and 115–121 W. The magnitude spans from 3.0 to 8.0. The catalog is complete above magnitude 3.

As explained in Sect. 2, the catalog has to be discretized, in order to translate it into nodes and edges. We used a 2D approximation, with box size of 0.1°, which is reasonable taking into account the typical size of a small fault and the accuracy in the hypocentral locations. We also need a discretized time, so that the activity at each time step is the number of events at that interval, and then we obtain a time series of activity for each cell. This is necessary in order to correlate the different activities by means of Eq. (1). We tested different time intervals: 1 day, 100 days, and 1000 days. 1 day is a natural selection to obtain almost continuous time series seismic activity; between 100 and 1000 days we have the commonly accepted time scale for aftershock sequences and can be related to nucleation processes (Dieterich, 1994).

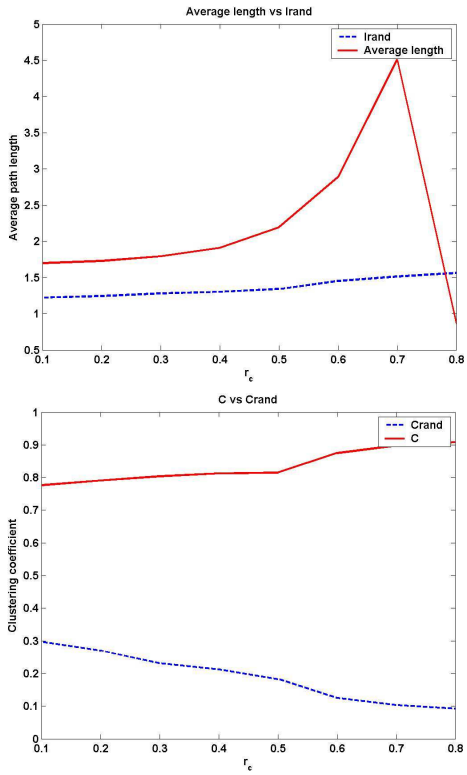

Fig. 4. Average length and clustering coefficient of the networks with different thresholds (r_c) compared to those corresponding to a random graph with the same number of nodes and the same average node degree, for time lags of 1000 days.

4 Results

As can be seen in Figs. 2–4, the clustering coefficient for the connected components is always much higher than that of a random network. It also shows that for correlations between the cells higher than 0.8, the average path is always lower than that of a random network. So, when we apply a threshold for r_c higher than that 0.8, we obtain a complex network which behaves as a small world, as defined in (Watts and Strogatz, 1998). Note that we are interested in studying highly correlated cells, and $r_c > 0.8$ is therefore a good lower threshold for our seismicity network. Note also that the threshold r_c affects the connectivity of generated networks. Larger r_c will result in the disconnected network whose average short path length will become very large in the sense of graph theory, which would be different from the results in the figures. In Fig. 5 we show the degree distribution for r_c=0.8 and 1, 100, and 1000 days, respectively. They are not scale free. This result is opposite to that found previously in (Abe and Suzuki, 2004, 2006a; Baiesi and Paczuski, 2005). Thus, the scale invariance is violated by thresholding. This implies that thresholding eliminates an important element of complexity of a seismic network. We also analyze the scaling relationship between the clustering coefficient and the degree

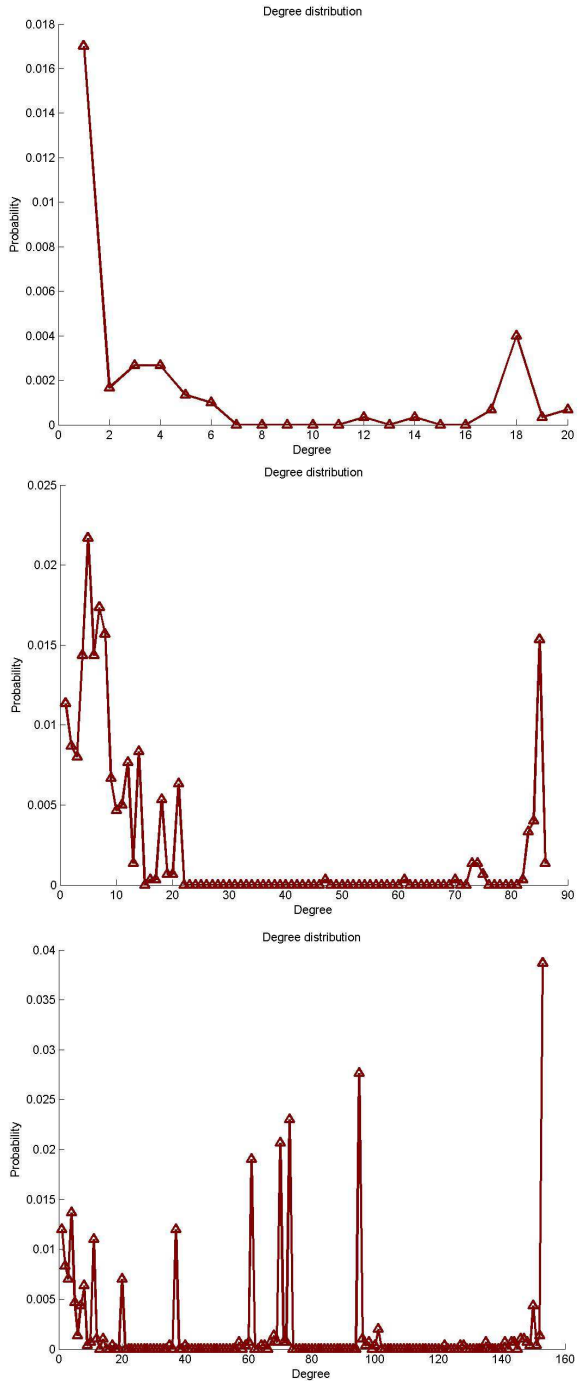

Fig. 5. Degre distribution for r_c=0.8 and 1 day, 100 days, 1000 days lag, respectively. As can be seen, they are not scale free.

(Abe and Suzuki, 2006b). We see that it is also violated by thresholding (Fig. 6).

In Fig. 7 we present the networks obtained with $r_c > 0.8$ and 1 day time interval. The main component is relating the Landers earthquake of 1992. All the components are related to big earthquakes with their corresponding aftershocks. It

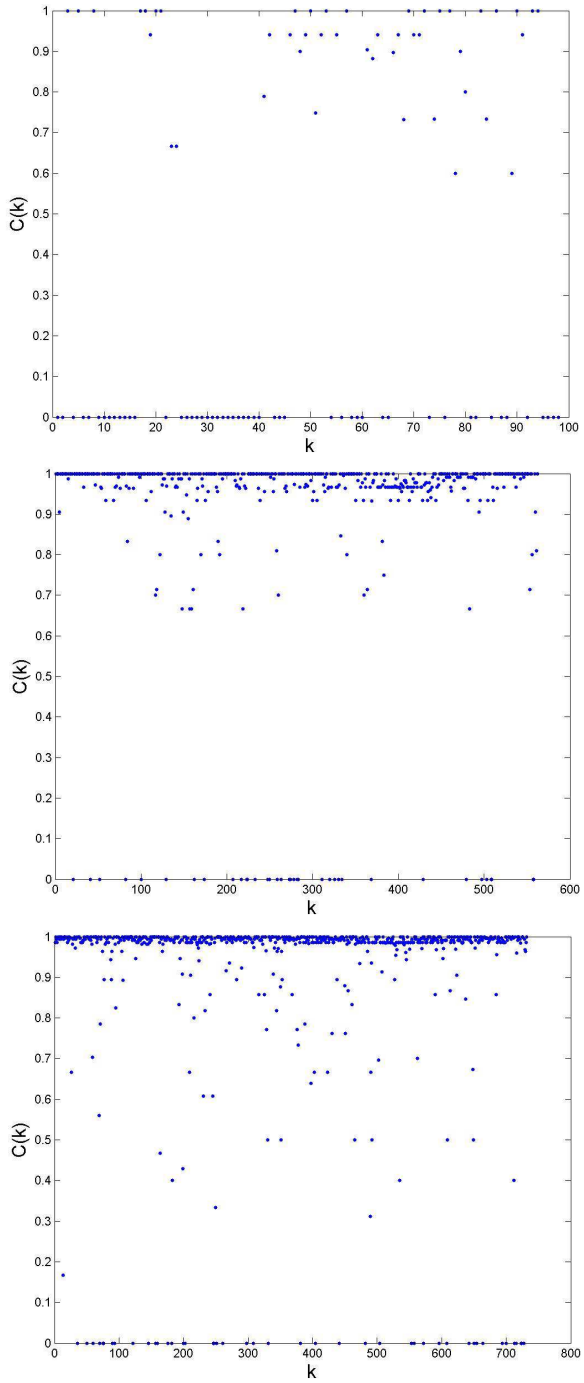

Fig. 6. Clustering in function of the degre for r_c=0.8 and 1 day, 100 days, 1000 days lag, respectively. As can be seen, they are not scale free.

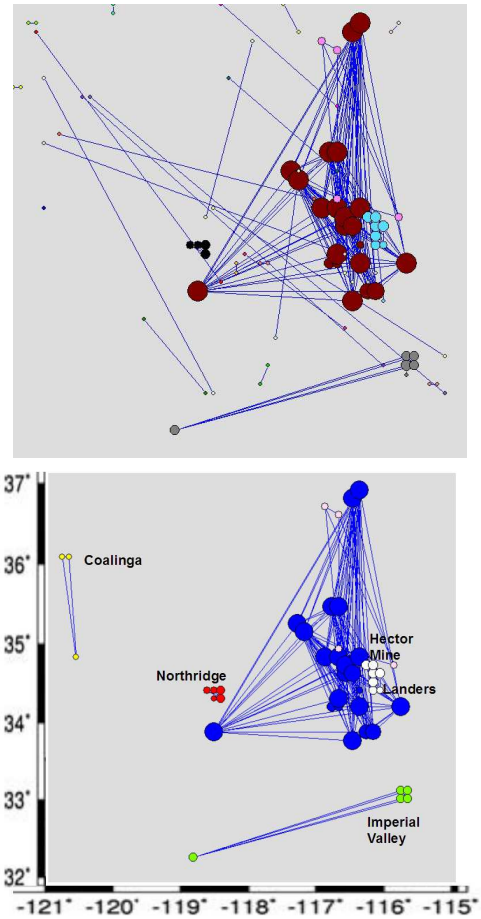

Fig. 7. Networks obtained for r_c>0.8, by using a betweenness and a hierarchical algorithm (with some of the main earthquakes in the region) to find the components, with Pajek (Batagelj and Mrvar, 1998), for time lags of 1 day.

is interesting to note that those Coalinga and Imperial Valley earthquakes have cells relatively far from them, but with a high correlation in the seismicity rate series.

When we visualize the network for 100 days lag, we can see that there are much more main components (56) than before (6) by using the hierarchical algorithm. The main component (first in Fig. 8) is the same as the main component for 1 day lag. Other particular components are related to different earthquakes. Most of the clusters' links outline clearly the San Andreas fault direction.

The number of clusters decreases with respect to the network found for the 100 days lag. In Fig. 9, the main component for the 1000 days lag relates the whole area.

The way of constructing the seismic network is different than that proposed by Abe and Suzuki (2006a), but one conclusion obtained is very similar: the mainshock plays a role of a "hub" with large degrees of connectivity. For these highly correlated events, the network obtained represents a small world. It is also interesting to note that those components can be viewed as the different activities a brain processes. Each stimulus is answered by different cells in the area. However, the networks obtained are not scale-free (the degree distributions were not found to follow a power law). So, the model for the network seems to be as follows: links are much more likely to connect "neighbor nodes" than

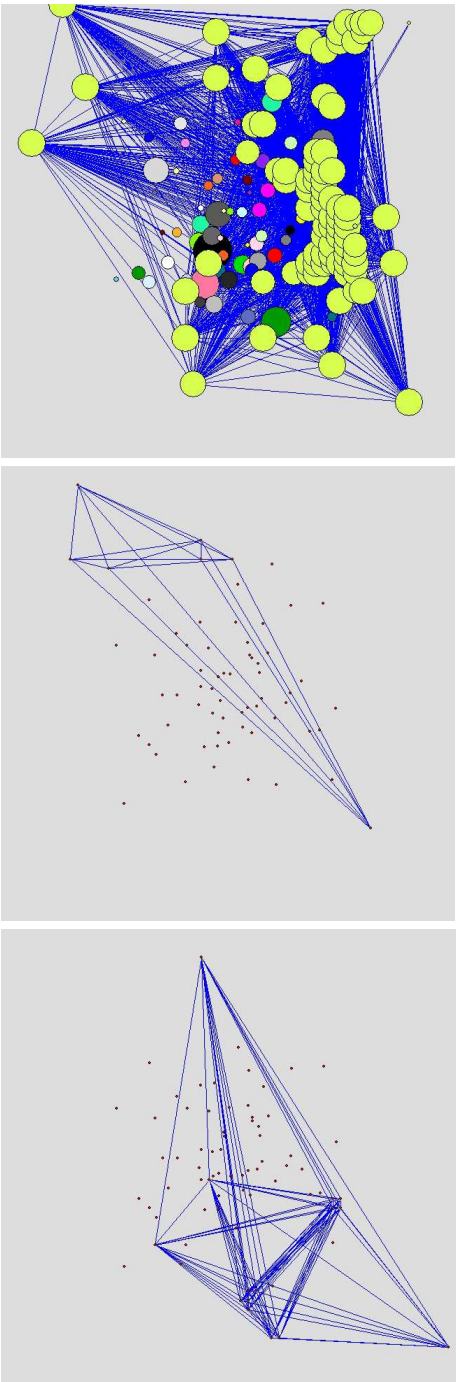

Fig. 8. Some of the 56 components for the 100 days time interval and $r_c > 0.8$, by using a hierarchical algorithm to find the components, with Pajek (Batagelj and Mrvar, 1998).

distant nodes. However, as can be deduced from the networks for 1 day lag (Fig. 7), there are some long range links, in particular following the San Andreas fault. Since the rate of earthquakes is related to the stress transfer (Helmstetter et al., 2005), the networks reflect the way stresses are diffused in the area.

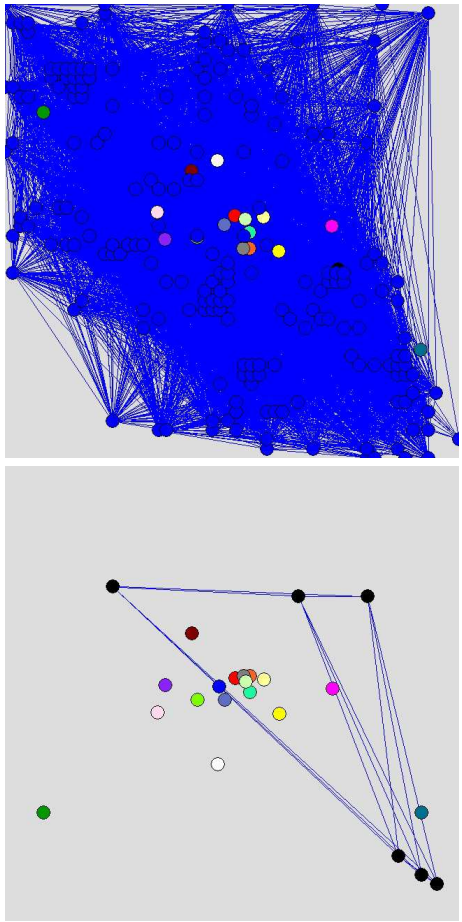

Fig. 9. Some of the 18 components for the 1000 days time interval and $r_c > 0.8$, by using a hierarchical algorithm to find the components, with Pajek (Batagelj and Mrvar, 1998).

5 Conclusions

We propose a different analysis of seismicity in terms of complex networks. They are obtained in a way similar to the way brain functional networks are studied (Eguíluz et al., 2005). In our preliminary results, we see that the different components of the obtained networks act as different responses to the stimulus given by the general plate motions in the region. The results of this study show that the functional connectivity matrix of seismic activity recordings can be converted into a sparsely connected graph by applying a suitable threshold of r_c. So, it can be said that the highest correlated cells in the region form a small world network. This small world property means that there are long-range connections in the seismic network. These connections might be related to the San Andreas or other large faults in the region, that transfer the stresses. This method could be useful to find triggered earthquakes, as well as for declustering the catalogs. This is related to the correlation coefficients, that may offer some important information about the

earthquake activities. Another choice would be to let these "small-world" networks to be weighted. In the present study we only made our analysis in two dimensions, due to limitations in the computations.

Acknowledgements. This work was partially supported by the MCYT project CGL2005-05500-C02-02/BTE, the MCYT project CGL2005-04541-C03-03/BTE, and the Research Group "Geofísica Aplicada" RNM194 (Universidad de Almería, España) belonging to the Junta de Andalucía. The work of K. F. Tiampo was funded under the NSERC and Benfield/ICLR Industrial Research Chair in Earthquake Hazard Assessment. The work of A. Jiménez was funded by a "Fundación Ramón Areces" Grant.

Edited by: J. Kurths
Reviewed by: two anonymous referees

References

Abe, S. and Suzuki, N.: Scale-free network of earthquakes, Europhys. Lett., 65, 4, 581–586, 2004.

Abe, S. and Suzuki, N.: Complex-network description of seismicity, Nonlin. Processes Geophys., 13, 145–150, 2006, http://www.nonlin-processes-geophys.net/13/145/2006/.

Abe, S. and Suzuki, N.: Complex earthquake networks: Hierarchical organization and assortative mixing, Phys. Rev. E., 74, 026113, 2006.

Albert, R. and Barabási, A. L.: Statistical mechanics of complex networks, Rev. Mod. Phys., 74, 47–97, 2002.

Baiesi, M., Paczuski, M.: Complex networks of earthquakes and aftershocks, Nonlin. Processes Geophys., 12, 1–11, 2005, http://www.nonlin-processes-geophys.net/12/1/2005/.

Batagelj, V. and Mrvar, A.: Pajek – program for large network analysis, Connections, 21, 47–57, 1998.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U.: Complex networks: Structure and dynamics, Physics Reports, 424, 175–308, 2006.

Dieterich, J.: A constitutive law for rate of earthquake production and its application to earthquake clustering, J. Geophys. Res., 99, 2601–2618, 1994.

Dijkstra, E. W.: A note on two problems in connection with graphs, Numerische Math., 1, 269–271, 1959.

Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., and Apkarian, A. V.: Scale-free brain functional networks, Phys. Rev. Lett., 94, 018 102–018 104, 2005.

Helmstetter, A., Kagan, Y. Y., and Jackson, D. D.: Importance of small earthquakes for stress transfers and earthquake triggering, J. Geophys. Res., 110, doi:10.1029/2004JB003286, B05S08, 2005.

Kochen, M.: The Small World, Ablex, Norwood 1989.

Milgram, S.: The small world problem, Physch. Today, 2, 60–67, 1967.

Newman, M. E. J.: The structure and function of complex networks, SIAM Review, 45, 2, 167–256, 2003.

Southern California Earthquake Center: available at http://www.data.scec.org/.

Stam, C. J.: Functional connectivity patterns of human magnetoencephalographic recordings: a "small-world" network?, Neuroscience Lett., 355, 25–28, 2004.

Wang, X. F. and Chen, G.: Complex networks: small-world, scale-free and beyond, IEEE Circuits and Systems Magazine, 3, 6–20, 2003.

Watts, D. J. and Strogatz, S. H.: Collective dynamics of "small-world" networks, Nature, 393, 440–442, 1998.