

Weighted complex networks applied to seismicity: the Itoiz dam (Northern Spain)

A. Jiménez and F. Luzón

Department of Applied Physics, University of Almería, Carretera de Sacramento s/n, Cañada de San Urbano, 04120, Almería, Spain

Received: 10 December 2010 – Revised: 11 July 2011 – Accepted: 11 July 2011 – Published: 18 July 2011

Abstract. On 18 September 2004, an earthquake of magnitude $m_{Lg} = 4.6$ was recorded near the Itoiz dam (Northern Spain). It occurred after the first impoundment of the reservoir and has been catalogued by some authors as induced seismicity. We analyzed the seismicity in the region as weighted complex networks and tried to differentiate this event from others that occurred nearby. We calculated the main topological features of the networks formed by the seismic clusters and compared them. We compared the results with a series of simulations, and showed that the clusters were better modelled with the Epidemic-Type Aftershock Sequence (ETAS) model than with random models. We found that the properties of the different clusters are grouped according to the magnitude of the main shocks and the number of events in each cluster, and that no distinct feature could be obtained for the 18 September 2004 series. We found that the nodes with the highest strength are the most important in the networks' traffic, and are associated with the events with the highest magnitude within the clusters.

1 Introduction

Earthquakes are one of the most interesting natural phenomena that can be described as complex systems. A key ingredient of a complex system is the non-linear interaction between its constituents, which under special circumstances can give rise to coherent, emergent, complex behaviour patterns with a very rich structure. A way to study these structures is to describe the system as a complex network, and analyze the topological characteristics it forms.

In the past few years, much attention has been focused on the study of complex networks, such as the World Wide Web (WWW), the Internet, scientific collaboration networks, worldwide airport networks, etc. Researchers have mainly focused on unweighted networks, that is, networks in which every pair of nodes in the network are either connected or not connected, with weights of 1 or 0, respectively. However, many technological, biological and social systems are best described by weighted networks, whose properties and dynamics depend not only on their structures but also on the weight of the connections between the nodes (Chen and Chen, 2007). Even for purely unweighted graphs, edge weights naturally emerge as dynamical properties, when transport, random walks or other processes take place on the network (Tadic et al., 2007; Barrat et al., 2008).

In seismology, several models based on complex networks have been proposed (Abe and Suzuki, 2004a, b, 2006; Baiesi and Paczuski, 2004, 2005; Davidsen, 2008; Jiménez et al., 2008, 2009). However, these models are unweighted complex networks. In our research we have used a weighted complex network, which describes the seismicity in a more realistic way than an unweighted network, and provides more characteristics from which to gain a more complete picture of the seismicity.

The network model we propose is as follows: each earthquake represents a node, and the link weight between nodes is a distance based on the ETAS model (Ogata, 1988, 1998, 1999). In this research, we analyzed a small region centred on the Itoiz dam (Northern Spain), which has been extensively studied since the 18 September 2004 earthquake that occurred nearby. For example, Ruiz et al. (2006) analyzed the seismic series corresponding to the 18 September 2004 earthquake; Luzón et al. (2009, 2010) and Durá-Gómez and Talwani (2010) studied the effect of the pore pressure due to the impoundment of the reservoir; Santoyo et al. (2010) calculated the stress produced by the reservoir near Itoiz. They



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

showed that the aftershock sequence was indeed produced by the stress transfer caused by the 18 September 2004 earthquake. Rivas-Medina et al. (2011) calculated the seismic hazard near the Itoiz dam. In a previous work (Jiménez et al., 2009), we studied the seven main seismic clusters near the Itoiz dam by modelling them as unweighted complex networks. We obtained seven clusters with small-world behaviour of the complex networks they formed. We did not find any difference between the 18 September 2004 cluster and the others, except for a higher fractal dimension of the epicentre distribution.

2 Methods

In this work, we improve on the study in Jiménez et al. (2009), because we obtain the values for the parameters of the ETAS model in the region. These are real values which allow us to use an appropriate distance between earthquakes to decluster the catalogue.

We then analyzed each cluster as a weighted complex network, with the inverse of the ETAS distance as the weight between two events. This distance depends on the magnitude, spatial separation and time between them. We characterized the networks by their average path length and the clustering coefficient. Afterwards, we classified the clusters according to these two features by means of a k -means algorithm (Teknomo, 2007). We also calculated the distribution of weights, strength and distances, and the betweenness (Gleich, 2007). We then compared the results with the average values produced by several simulations, both with a random model (Garlaschelli, 2009) and the ETAS model (Ogata, 1998, 1999; Helmstetter and Sornette, 2002).

2.1 ETAS model

The ETAS model is the most popular contemporary model for studying aftershock sequences. It is a stochastic point process model, which has a number of parameters that are usually fitted from a training period before the model can be applied for forecasting purposes. In the framework of this model, there is no difference between aftershocks, main shocks or foreshocks (Helmstetter and Sornette, 2002). The ETAS model is a point process representing the activity of earthquakes of magnitude m_0 and larger in a region over a given period of time. The model includes background activity of constant occurrence rate μ in time (i.e. stationary Poisson process) and also includes aftershocks as described below. Each earthquake, including aftershocks of another earthquake, is followed by its aftershock activity, although only aftershocks of magnitude m_0 and larger are included in the data. The aftershock activity is represented by a non-stationary Poisson process according to the modified Omori formula in such a way that the occurrence rate ($v_i(t)$) of aftershocks at time t following the i -th earthquake (t_i, m_0) is given by (Ogata, 1998):

$$v_i(t) = \frac{K}{(t - t_i + c)^p} e^{\alpha(m_i - m_0)} \quad (1)$$

for $t > t_i$, where the parameters K, α, c, p are constants common to all i , and m_0 is the minimum magnitude that produces an aftershock sequence. The rate of occurrence of the whole earthquake series at time t , called the conditional intensity function based on the history of occurrence $H_t = \{(t_i, m_i); t_i < t\}$, then becomes:

$$\lambda_0(t | H_t) = \mu + \sum_{\{i: t_i < t\}} v_i(t) \quad (2)$$

where μ is the background seismicity. As extensions of the ETAS model we confined ourselves to space-time response functions $g_\Phi(t, x, y; m)$ such that the superposed conditional intensity reads (Ogata, 1998):

$$\lambda(t) = \int \int_A \lambda(t, x, y) dx dy = \int \int_A g_\Phi(t, x, y; m) dx dy \quad (3)$$

$$\approx \int \int_{R^2} g_\Phi(t, x, y; m) dx dy = \frac{K e^{\alpha(m - m_0)}}{(t + c)^p}$$

where A is the area of the study region. Note that m in Eqs. (2–3) enters only as a parameter, not a variable. We used the response function in such a way that spatial dependency was separable from dependency on magnitude and time. We did this in order to simulate the seismicity as in Helmstetter and Sornette (2002), where the terms relating to time, space and magnitude are independent. For this purpose, we made the following hypothesis (Ogata, 1998):

$$g_\Phi(t, x, y; m) = \frac{K}{(t + c)^p} \frac{e^{\alpha(m - m_0)}}{(x^2 + y^2 + d)^q} \quad (4)$$

where x and y are the spatial coordinates of the event in 2-D. In order to choose the best model we followed the maximum likelihood procedure described in detail by Ogata (1998). We used a Genetic algorithm (GA) to maximize the likelihood of the model. We improved the method proposed by Ogata (1998, 1999) to obtain the ETAS parameters, because the GA is a more efficient way of calculating the optimum. GA are methods of global optimization, which have proved to be effective when the models are described by a few parameters, the problem is nonlocal (the global optimum is needed, but there are many local optima) and non-linear, and there is no a priori knowledge of the behaviour of the function. In geophysics, and particularly in seismology, many problems often have such features. The GA used in this paper was implemented in Jiménez et al. (2005). The definitive search strategy chosen was selection by rank wheel, crossover based on fitness, and replacement by rank wheel. This strategy is moderately elitist, because the best individuals are easily selected, but there is low selection pressure. The simple GA was improved with the reinitialization of the

population when the convergence stays blocked (Jiménez et al., 2005).

Thanks to the large number of models considered, in addition to the fittest model, a mean model and its accuracy were evaluated by means of a statistical approach based on the estimation of the Marginal Posterior Probability Density (MPPD) (Dal Moro et al., 2007). The averaged model was calculated as:

$$\theta_j = \frac{1}{n} \sum_{i=1}^n \rho(i) \theta_{ij} \quad (5)$$

where θ_j is the model parameter j , θ_{ij} is the value of the parameter in the evaluated model i , and $\rho(i)$ is the function:

$$\rho(i) = \frac{e^{\text{fitness}(i)}}{\sum_j e^{\text{fitness}(j)}} \quad (6)$$

Where the fitness is the log-likelihood calculated for each model, following Ogata (1998). The standard deviation was calculated as:

$$\sigma_j = \frac{1}{n-1} \sum_{i=1}^n \rho(i) (\theta_{ij} - \theta_j)^2. \quad (7)$$

2.2 Declustering algorithm

The purpose of declustering the catalogue is to find the main clusters in it. Window-based or link-based methods are traditionally used for declustering a seismic catalogue or identifying earthquake clusters (Baiesi and Paczuski, 2005). The main problem is to identify which earthquakes are correlated, because there is no single operational way of distinguishing between aftershocks and main shocks (Zhuang et al., 2004). Several methods based on the ETAS model have been proposed for declustering the catalogues (Zhuang et al., 2004; Jiménez et al., 2009; Console et al., 2010). We followed the method proposed in Jiménez et al. (2009), where the distance between earthquakes is given by the inverse of the probability of two earthquakes being correlated. In the present work, it is given by the function g_Φ , Eq. (4), with the parameters obtained previously for the region. We began by calculating these distances between all pairs of earthquakes and then used them to create a hierarchical cluster tree where two events are linked if their correlation is higher than a certain threshold. In order to group individual earthquakes to clusters, we used the single linkage method. The single linkage method is a fundamental agglomerative hierarchical clustering algorithm. By truncating the resulting cluster tree at a suitable threshold value for the distance being studied, a complex network was obtained in which each earthquake represents a single node. The threshold was chosen in such a way that the number of clusters was in agreement with the value obtained for the background seismicity μ . The estimation of the background seismicity following the ETAS model

gives us a natural way of choosing the threshold necessary to decluster the catalogue. In a previous work (Jiménez et al., 2009) this step was a little arbitrary, since we chose the threshold that gave us more than one cluster.

When the number of points is large, this way of finding the main clusters is less time-consuming than, for example, the k -means algorithm (MacQueen, 1967). It is also unique, unlike the k -means algorithm, which depends on the initial clusters' centres.

2.3 Characterization of weighted complex networks

Here we describe the most important parameters that describe a complex network. In an unweighted network, an important feature is the degree distribution, k_i :

$$k_i = \sum_{j=1}^N A_{i,j} \quad (8)$$

where A is the adjacency matrix and N is the number of nodes. For a weighted network, it is important to know the distribution of the weights, $w_{i,j}$, which replace the Boolean numbers $A_{i,j}$, and the strengths, s_i , that replace the degree of the node. The higher the weight and strength, the more closely related are the two nodes. We use g_Φ (Eq. 4) as the weight between the nodes.

Two of the most important quantities required to characterize a complex network are the average path length and the clustering coefficient.

The average (or characteristic) path length L is the mean length of the shortest paths (expressed in terms of the number of edges) connecting any two nodes on the graph. The shortest path between a pair (i, j) of nodes in a network is considered as their geodesic distance G_{ij} , with a mean geodesic distance L of:

$$L = \frac{2}{N(N-1)} \sum_{i < j} G_{ij} \quad (9)$$

where N is the number of nodes.

In order to calculate the average path length, we used the distance given by $1/g_\Phi$, and put it into the Dijkstra algorithm (Dijkstra, 1959). We used the implementation given in Kay (2001).

The other important quantity, the clustering coefficient, was calculated as in Onnela et al. (2005):

$$C = \frac{1}{k_i(k_i-1)} \sum_{j,k} (\hat{w}_{ij} \hat{w}_{ik} \hat{w}_{jk})^{1/3} \quad (10)$$

where $\hat{w}_{ji} = w_{ij} / \max_{i,j}(w_{ij})$, and k_i equals N if all the nodes are connected, as is the case here. With this definition, C has the same value if the weights become binary, it is defined between 0 and 1, uses a global normalization, takes into account weights of all edges in the triangle and is invariant to weight permutation for one triangle (Saramäki et al., 2007).

To further characterize the complex network, we also calculated the betweenness distribution. This was first proposed by Freeman (1977), and represents the total number of optimal paths between any pairs of nodes passing through one node (Park et al., 2004). More succinctly (Freeman, 1977):

$$B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (11)$$

where σ_{st} is the number of shortest paths from s to t , and $\sigma_{st}(v)$ is the number of shortest paths from s to t that pass through a vertex v . We used the code implemented by Gleich (2007) in order to calculate it.

2.4 Network model: random weighted networks and ETAS simulation

In order to classify the seismic networks derived in this work according to their main topological features, they can be compared with the corresponding values of characteristic network theoretic measures obtained for comparable model networks, i.e. model networks with the same number of nodes as the original seismic networks. For this purpose, we will use particularly weighted random networks and simulations of the ETAS model. Since the particular network pattern of both types of models depends on random realizations of links, the properties of the resulting networks may differ between individual realizations. In order to consider this in our analysis, 100 realizations were studied in each case to obtain estimates for the expectation values and 95 % confidence levels (the latter were approximated by ± 1.96 times the standard deviation of the respective values of the network-theoretic measures obtained for the individual realizations).

The Erds-Rényi (ER) random graph (Erdős and Rényi, 1959, 1960, 1961) is the prototype of all unweighted network models: in a graph with N vertices, an unweighted edge is drawn independently between any pair of vertices with equal probability P . The ER model provides a fundamental reference for the properties of real networks. The increasing interest in complex networks originates precisely because of the striking difference between the observed properties of real networks and the behaviour of the ER model (Garlaschelli, 2009). So, we needed a random reference model to enable us to compare its properties with those of the networks we obtained. Garlaschelli (2009) proposed a procedure to construct a weighted random model as follows: the number of nodes is fixed. For each pair of vertices, we start a series of Bernoulli trials with success probability P , which depends on the number of links and the sum of edge weights in the real network. Each success implies that the weight is increased between the same two vertices in one unit. As soon as a failure occurs for the first time, the sequence of trials stops and a new pair of vertices is selected. The process is repeated until all pairs have been considered.

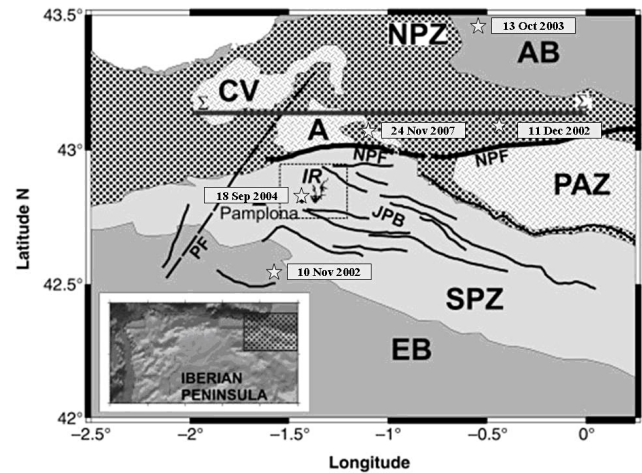


Fig. 1. Location of the main events corresponding to the main clusters. NPZ = North Pyrenean Zone; PAZ = Palaeozoic Axial Zone; SPZ = South Pyrenean Zone; NPF = North Pyrenean Fault; PF = Pamplona Fault; Palaeozoic Basque Massifs: CV = Cinco Villas Massif; A = Aldudes-Quinto Real Massif. AB = Aquitaine Basin; EB = Ebro Basin; JPB = Jaca-Pamplona Basin; IR = Itoiz Reservoir (Modified from Santoyo et al., 2010).

We needed a more realistic simulation of the complex networks. In order to achieve this, we simulated a network with the same number of nodes by using the ETAS model. The algorithm for the simulation is described in Ogata (1998, 1999) and Helmstetter and Sornette (2002). We assumed a decoupling between magnitude, space and time, because of its simplicity. Starting with the main event in each cluster, events were simulated sequentially. First, we calculated the conditional seismic rate $\lambda(t)$ defined by:

$$\lambda(t) = \sum_{t_i \leq t} \frac{K}{(t - t_i + c)^p} e^{\alpha(m_i - m_0)}. \quad (12)$$

The time of the following event was then determined according to the nonstationary Poisson process of conditional intensity $\lambda(t)$, and its magnitude was chosen according to the Gutenberg-Richter (1956) distribution with parameter b . To determine the position in space of this new event, we first chose its mother randomly among all preceding events with a probability proportional to their rate of aftershocks evaluated at the time of the new event. Once we had chosen the mother, we generated the distance r between the new earthquake and its mother according to the spatial part of the distribution given in Eq. (4). The location of the new event was determined assuming an isotropic distribution of aftershocks.

3 Tectonical setting and data

The study area (Fig. 1) is located in a region with a complex thin-skinned structure belonging to the South Pyrenean Zone

(SPZ), and takes the form of a square of $2^\circ \times 2^\circ$ with its centre in the 111 m high Itoiz dam (42.8° N , 1.35° W). The SPZ forms the external part of the Pyrenean belt and is a Tertiary unit overriding to the south of the Ebro foreland basin. This unit is north-bounded by the Palaeozoic Axial Zone (PAZ), which outcrops here as a series of individual massifs, known as the Palaeozoic Basque Massifs (PBM). The North Pyrenean Fault (NPF) is located between the PAZ and the North Pyrenean Zone – a Mesozoic unit overthrusting the Aquitaine foreland basin to the north. This fault is a major tectonic suture running east-west along the Pyrenean range that is interpreted as the superficial expression of the Iberian and European plate boundaries. Palaeomagnetic data, seismic profiles and seismicity studies suggest that the structural boundary associated with the NPF is prolonged westward through the Basque-Cantabrian basin along the Leiza Fault. Another relevant structure in the area is the Pamplona Fault (PF) that runs NNE-SSW from the Ebro basin to the PBM. It has been interpreted as a deep transverse structure separating two different structural zones, the SPZ to the east, where the most important structures mainly trend south, and the Basque-Cantabrian basin to the west, with a thicker lower Cretaceous sequence, where most structures trend northward. This structure acted as an extensional transfer fault during the Mesozoic extensional period that led to the opening of the Gulf of Biscay and the separation of the Iberian Peninsula as a subplate. The PF was subsequently involved in the Tertiary compression responsible for the Pyrenean belt uplift. The epicentral region is a Mesozoic and Tertiary cover area, located in the NE of the Pamplona basin, and composed of anticlines and synclines with the axes trending to the east, but truncated in some places by E–W to ESE–WNW fault systems (Ruiz et al., 2006).

The present day seismicity of the western Pyrenees, as reported from permanent networks, is rather moderate and mainly concentrated on the French side. In the past few decades it has been characterized by events of magnitudes up to 5.5. This activity follows an E–W oriented strip 150 km long and 5 to 15 km wide, starting at the western edge of the NPF, at a longitude of 0.1° W , and continuing through the PBM along the Leiza Fault. Historical seismicity also appears concentrated in the border region, along the PAZ and the PBM. Six destructive events with intensity greater than VII have been reported in this area over the last 200 yr approximately. Instrumental catalogues of the westernmost Pyrenean edge show a sparse and moderate to low magnitude activity, mainly concentrated westward of Pamplona, where events are related to the central segment of the Pamplona Fault and to the Aralar thrust. An E–W seismicity belt is also found related to the Roncesvalles thrust that separates the Palaeozoic Aldudes massif from the Pamplona basin. The catalogues also report a few events related to the E–W thrust systems delineating the contact between the Pamplona and Ebro basins (Ruiz et al., 2006).

The Itoiz dam (42.80° N , 1.36° W) is located in Navarre, Northern Spain, 2 km north of the village of Aoiz and 25 km east of Pamplona. Construction was completed in 2003. In January 2004 its impoundment began and 8 months later, on 18 September a clustered seismic series occurred. The epicentre was located between the city of Pamplona and the Itoiz reservoir. The main shock ($m_b L_g = 4.6$, 42.85° N , 1.45° W) and the largest aftershock (the 30 September 2004 event) were widely felt in this region causing widespread unease amongst the general public. The Itoiz reservoir has a maximum capacity of 418 hm^3 and, once filled, covers a maximum surface area of 510 km^2 . It was constructed for irrigation and electricity generation purposes. In September 2004, the reservoir was only 15 % full, after beginning impounding water in January 2004 (Ruiz et al., 2006).

The data used were recorded by the Instituto Geográfico Nacional (IGN, 2010), and cover earthquakes from 1999 to 2008. In total, 2350 earthquakes with a magnitude greater than 1 were observed. The location of the data does not include the depth of the earthquakes. The Gutenberg-Richter distribution (1956) has $b = 1.14$, with error bound equal to 0.03. This value differs slightly from that used in Jiménez et al. (2009). The difference is due to the fact that in the present work we calculated it by fitting of the cumulative distribution, and not by using the probability density function as we did in Jiménez et al. (2009). We did this because we wanted to simulate the networks as far as possible according to the method proposed by Ogata (1998).

4 Results

Following the procedure detailed in Sect. 2, we obtained the main clusters in the region. Afterwards, we translated the clusters into the language of complex networks: each node represents an event, and the weight between each pair of nodes is given by the inverse of the distance between the two earthquakes. This distance is based upon the ETAS model, so that it depends on the magnitude of the first earthquake, the time interval and the spatial separation between events. A larger distance represents a lower probability of two earthquakes being connected, so the weight of the link is lower.

Ruiz et al. (2006) also calculated some of the parameters of the ETAS model related to time and magnitude for the 18 September 2004 series, but for that cluster only, not for the region as a whole. Here, we calculated the ETAS parameters for the region of interest. The intervals of the search were the following: $p = [0.8, 1.2]$, $\alpha = [0.5, 1.5]$, $c = [10^{-3}, 10^{-2}]$, $d = [0.1, 20]$, $q = [0, 2]$, $K = [10^{-9}, 10^{-4}]$, and $\mu = [10^{-6}, 10^{-4}]$. The resolution for each parameter was (in bits, so 2 to the power of the given resolution), 6, 6, 6, 7, 4, 15, 6, respectively. The intervals were chosen on the basis of the values obtained by Ogata (1998) for a variety of regions. The result of the maximization of the log-likelihood of the model in Eq. (4) Ogata (1998) is shown in Table 1.

Table 1. Average values and error of the ETAS fit for magnitude higher than 1.

Parameter	p	α	c (days)	d (km)	q	K (shock/day/km ²)	μ (shock/day/km ²)
Value	0.975	0.8438	0.0049	0.1	1.625	$7.1323 \cdot 10^{-5}$	$5.6406 \cdot 10^{-6}$
Error MPPD	$9 \cdot 10^{-30}$	$4 \cdot 10^{-30}$	$2 \cdot 10^{-32}$	$9 \cdot 10^{-32}$	$6 \cdot 10^{-30}$	$9 \cdot 10^{-39}$	$7 \cdot 10^{-42}$
Error simulation (51)	0.31	0.83	0.0062	6.1	0.790	$5.6512 \cdot 10^{-5}$	$3.9899 \cdot 10^{-5}$

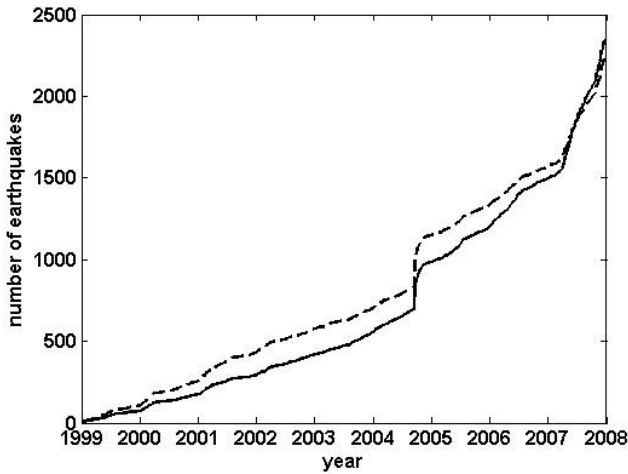


Fig. 2. In solid line, the real seismicity, and in dashed line, the ETAS model.

The Akaike Information Criterion (AIC) (Akaike, 1974) of the best fit is 1374, which is a low value if we compare it with the ones obtained in Ogata (1998). The average values coincide almost completely with the best fit, and are the same for the resolution given in Table 1. The error is given by the calculation of the Marginal Posterior Probability Density (MPPD) error bound (Dal Moro et al., 2007). The best fit was found in the second generation, and was stable until the 100th generation. In Fig. 2 we can see the difference between the real data and the model. We can see that the ETAS model underestimates the number of events. We are very confident that the GA reaches the best fit, so perhaps even the ETAS model does not fit our data set sufficiently well.

We now roughly estimated the errors in the simulations, following Wang et al. (2010). First we generated the spontaneous (background) events that are uniformly distributed in the space-time window. The total number of spontaneous events is a Poisson random variable with a mean of μTA , and the magnitudes are generated from the truncated Gutenberg-Richter magnitude distribution. This is the first generation. We then proceeded to generate the offspring events, by calculating the aftershock sequence for each event in the first generation, with a number of events equal to a Poisson random variable with mean $Ke^{\alpha(m-m_0)}TA$. The events were generated according to the algorithm proposed

Table 2. Main clusters in the catalogue. D_2 is the correlation dimension.

Main shock	Magnitude (mbLg)	Number of shocks
10-Nov-2002	3.3	55
11-Dec-2002	3.6	447
13-Oct-2003	3.3	93
18-Sep-2004	4.6	925
24-Nov-2007	3.3	137
21-Feb-2002	3.8	22
27-Oct-2007	2.4	34
16-Jul-2002	2.8	24

by Ogata (1998). This is the second generation. We then continued with the next generation until we reached a generation with no offspring or we reached the maximum number of events in the real catalogue. We then calculated the errors of the simulations as in Wang et al.(2010):

$$\text{error} = \sqrt{\frac{\sum (\text{simulated parameter} - \text{real value})^2}{\text{number of simulations}}}. \quad (13)$$

Note that these errors are calculated for the simulations, and they are not related to the actual error of the fits. They are shown in Table 1. We could only perform 51 instances of the catalogue, with a maximum number of generations equal to 10 for the GA, because of computation limits. We observed that the errors are high, and this is reflected in the errors found in Table 4 when we use the ETAS simulation to compare it with the real networks.

With the ETAS model parameters obtained previously, we proceeded to decluster the catalogue. We wanted to find the threshold in the distances that gives us a number of groups in agreement with the quantity $t\mu TA$, with T being the time interval covered by the catalogue, A the area of the region and μ the background seismicity in the ETAS model. With our data set, the quantity μTA is 447. We use a simple link algorithm with the threshold $1/g_\phi = 2 \cdot 10^6$, that gives us 440 clusters in our catalogue.

In order to ensure that we had good statistics for the analysis of the weighted complex networks formed, we only chose the clusters with more than 20 events, which are described

Table 3. Average values and error of the ETAS fit for magnitude higher than 2.

Parameter	p	α	c (days)	d (km)	q	K (shock/day/km ²)	μ (shock/day/km ²)
Value	1.1563	1.25	0.0069	0.1	1.125	$3.8618 \cdot 10^{-5}$	$2.5469 \cdot 10^{-6}$
Error MPPD	$4 \cdot 10^{-31}$	$5 \cdot 10^{-32}$	$7 \cdot 10^{-36}$	$2 \cdot 10^{-33}$	$2 \cdot 10^{-31}$	$2 \cdot 10^{-40}$	$2 \cdot 10^{-42}$

in Table 2. The clusters are named after the earthquake with the highest magnitude. In Fig. 1 we show the main shocks of the clusters analyzed. The clusters corresponding to the 18 September 2004 and the 24 November 2007 earthquakes coincide with two of the three clusters with more than 50 events described in Jiménez et al. (2009). The other one, corresponding to the 2 April 2007, $m = 3.1$ earthquake does not appear in the present work, probably because of its lower magnitude. We can also see that the cluster corresponding to the 21 February 2002, $m = 3.8$ earthquake has fewer events. In our previous research (Jiménez et al., 2009) this cluster contained 34 events, instead of the 22 it has here. As can be seen from the results, only the three major earthquakes are listed in both works. The other clusters found in Jiménez et al. (2009) seem to be less important in the new approach.

Ruiz et al. (2006) found different families within the 18 September 2004 series, based upon the different characteristics of the waves in the recorded seismogram. We did not find these families, mainly because we did not use the waveforms and instead used the spatio-temporal distribution of the earthquakes in the catalogue. Ruiz et al. (2006) used a spatio-temporal window in order to decide if an event belongs to the aftershock sequence or not. In contrast, we used a distance based on the ETAS model obtained for the whole region, which depends on the magnitude, the time interval, and the spatial separation between events.

In order to test the influence of the threshold magnitude for the calculation of the clusters, we applied the method to the same catalogue but taking into account only the events with a magnitude of more than 2. With that condition, we have a catalogue with 455 earthquakes. So, we began by calculating the ETAS parameters that best fit the data and obtained the values given in Table 3. The AIC is 907. We had to perform the fitting part again because the values of the ETAS parameters vary when different threshold magnitudes are applied (Ogata, 1998). So, once we had the new values, we applied the link algorithm with the new distance. Now the quantity $\mu T A$ was 167. With a threshold distance equal to $1.5 \cdot 10^7$, we obtained 166 clusters. We found two main clusters with more than 20 events: the 11 December 2002 and the 18 September 2004. These correspond to the two larger clusters found by using a magnitude threshold of 1. The new clusters have less events than the number of earthquakes with magnitude higher than 2 in the original clusters. We observed that 24 % (18 September 2004) and 33 % (11 December 2002) of the

events in the original clusters were not found in the new clusters. This is due to the fact that there are some earthquakes linked to others by events with magnitudes of between 1 and 2. If we compare the events in the cluster with this new magnitude threshold and the events with a magnitude higher than 2 in the clusters with a magnitude threshold higher than 1, we can see that 99 % of them are the same for the 18 September 2004 earthquake, and 76 % for the 11 December 2002 earthquake. This shows that the main influence in the catalogue is due to the 18 September 2004 event, and the method fits this main event better.

We now studied the properties of the weighted complex networks we obtained, in particular, the distribution of the distances, weights and strengths of the nodes. The strength of a node is the sum of all the weights of its connections with others. It represents the importance of the node in the whole network.

The distribution of ETAS distances in the clusters seems to follow a power law distribution, with exponents that range from -1.6 to 2.6 . We used the programs implemented by Clauset et al. (2009) in order to calculate these exponents and other values that characterize the goodness of the power law hypothesis. The results show that the hypothesis of a power law distance is very plausible. However, the number of points that follow the power law is very small, and the total number of points is not enough to allow us to obtain good statistics. Finite-size bias may be present. The same situation is found for the weight distribution (the exponents range from -1.4 to 1.9 , but not enough points are available) and for the strength distribution. We found that the highest strength corresponds to the main shocks in all cases. Other works that use unweighted complex networks find power laws for the degree distributions (Baiesi and Paczuski, 2004, 2005; Abe and Suzuki, 2004b), a quantity equivalent to strength in these networks. In Jiménez et al. (2009) we did not find power laws for the same data set. However, in that research we did not use the method proposed by Clauset et al. (2009) for testing the power law hypothesis.

Another important characteristic of the nodes is betweenness, i.e. the total number of optimal paths between all pairs of nodes passing through one node (Park et al., 2004), in which an optimal path is the shortest geodesic distance between two nodes. The highest betweenness does not always correspond to the highest magnitude. If we take into account the 10 November 2002 earthquake, the highest betweenness

Table 4. Values of the network characteristics corresponding to the main clusters in the catalogue. L is the average path length, and C is the clustering coefficient. For the random model and the ETAS model we also provide the error bounds, below the mean value.

Main shock	L	C	L_{Rand} (error)	C_{Rand} (error)	L_{ETAS} (error)	C_{ETAS} (error)
10-Nov-2002	$1.2216 \cdot 10^6$	$8.1 \cdot 10^{-5}$	$3.0938 \cdot 10^{11}$ $2.5386 \cdot 10^{10}$	$2.1304 \cdot 10^{-5}$ $4.6791 \cdot 10^{-5}$	$1.0026 \cdot 10^8$ $1.3540 \cdot 10^9$	$1.3555 \cdot 10^{-4}$ $4.6503 \cdot 10^{-5}$
11-Dec-2002	$4.0414 \cdot 10^6$	$1.1 \cdot 10^{-6}$	$1.6246 \cdot 10^{12}$ $1.0708 \cdot 10^{11}$	$4.9302 \cdot 10^{-8}$ $1.1255 \cdot 10^{-7}$	$7.6461 \cdot 10^7$ $4.3256 \cdot 10^8$	$3.3079 \cdot 10^{-6}$ $1.9742 \cdot 10^{-6}$
13-Oct-2003	$2.1470 \cdot 10^6$	$3.7 \cdot 10^{-5}$	$4.7137 \cdot 10^{10}$ $8.3205 \cdot 10^9$	$5.7867 \cdot 10^{-6}$ $1.3090 \cdot 10^{-5}$	$4.1232 \cdot 10^7$ $3.5233 \cdot 10^8$	$5.1744 \cdot 10^{-5}$ $2.1860 \cdot 10^{-5}$
18-Sep-2004	$9.3911 \cdot 10^5$	$2.0 \cdot 10^{-7}$	$3.9283 \cdot 10^{12}$ $1.7696 \cdot 10^{10}$	$4.7100 \cdot 10^{-9}$ $1.2092 \cdot 10^{-8}$	$2.0029 \cdot 10^8$ $1.4636 \cdot 10^9$	$6.7001 \cdot 10^{-7}$ $1.9297 \cdot 10^{-7}$
24-Nov-2007	$2.5787 \cdot 10^6$	$2.2 \cdot 10^{-5}$	$2.8067 \cdot 10^{11}$ $2.8649 \cdot 10^{10}$	$1.5394 \cdot 10^{-6}$ $3.1449 \cdot 10^{-6}$	$4.9252 \cdot 10^7$ $6.0062 \cdot 10^8$	$1.2899 \cdot 10^{-4}$ $4.3295 \cdot 10^{-5}$
21-Feb-2002	$7.4554 \cdot 10^5$	$5.6322 \cdot 10^{-4}$	$2.2848 \cdot 10^{10}$ $1.4022 \cdot 10^{10}$	$5.6555 \cdot 10^{-4}$ $1.100 \cdot 10^{-3}$	$1.0046 \cdot 10^8$ $1.7182 \cdot 10^9$	$5.6320 \cdot 10^{-4}$ $1.2004 \cdot 10^{-4}$
27-Oct-2007	$4.4715 \cdot 10^3$	$1.0000 \cdot 10^{-3}$	$2.4751 \cdot 10^6$ $4.5742 \cdot 10^5$	$1.2961 \cdot 10^{-4}$ $2.8709 \cdot 10^{-4}$	$2.8843 \cdot 10^8$ $5.2832 \cdot 10^9$	$4.8115 \cdot 10^{-4}$ $2.5041 \cdot 10^{-4}$
16-Jul-2002	$2.9116 \cdot 10^5$	$8.1496 \cdot 10^{-4}$	$1.9220 \cdot 10^9$ $5.1871 \cdot 10^8$	$3.3098 \cdot 10^{-4}$ $7.6452 \cdot 10^{-4}$	$5.2226 \cdot 10^7$ $4.5342 \cdot 10^8$	$8.1659 \cdot 10^{-4}$ $3.5790 \cdot 10^{-4}$

corresponds to the 17 April 2007, $m = 2$ earthquake. For the 11 December 2002, it corresponds to the 6 April 2006, $m = 2.5$ event. For the 24 November 2007, we have both the highest magnitude and the highest betweenness at the same event. For the 18 September 2004, event, the highest betweenness is for the 30 September 2004, $m = 3.9$ earthquake. For the 13 October 2003, it corresponds to the 8 July 2005, $m = 2.4$ event. The highest betweenness found for the 21 February 2002 is the 27 February 2002 event. For the 27 October 2007, this same event had the highest value, and for the 16 July 2002, the highest betweenness is for the 28 December 2003 earthquake. So, with regard to betweenness, we have two groups of clusters: one formed by the 18 September 2004, the 24 November 2007, the 21 February 2002, and the 27 October 2007 earthquakes closely concentrated around the main event, and the other made up of the remaining clusters, which seem to cover longer time-spans in the respective regions. We also tested the power law. The exponents range from -2.1 to -2.6 for the betweenness distribution. Only a few clusters could be analyzed, due to the low number of points involved. In any case, the small amount of data available means that we cannot conclude that the clusters follow a power law.

We further analyzed which links are most important, the ones with a lower weight or the ones with a higher weight. For this purpose we used the concept of metaweight (Furuya and Yakubo, 2008). We calculated the clustering coefficient following Barrat et al. (2004), as suggested in Furuya and

Yakubo (2008), for the networks with weights equal to the q -th power of the weight. The behaviour with respect to q gives us the importance of weak or strong edges. This analysis can only be provided for weighted complex networks, and not for unweighted ones. We obtained a clustering coefficient (Barrat et al., 2004) very close to 0 for $q > 0$ and near 1 for $q < 0$, which means that the networks are mainly connected by weak edges (Furuya and Yakubo, 2008), and sparsely connected by strong edges. Note that this clustering coefficient is different from that used in Eq. (10). This is because we wanted to follow the same procedure as Furuya and Yakubo (2008). The q power enhances the weak links when $q < 0$, and the strong links when $q > 0$. So, when we obtained a clustering coefficient that was higher for $q < 0$ than for $q > 0$, it meant that the nodes had low weights between them, and only a few strong connections. This may be related to the high values for the betweenness scaling exponents found. Wang et al. (2008) showed that high values of this exponent (around 2.3) can be interpreted as a high concentration of traffic on the most important links.

We also calculated the values of the average path lengths and clustering coefficients, shown in Table 4. We provide the values for the random graph (Garlaschelli, 2009) and the ETAS model (Ogata, 1998, 1999; Helmstetter and Sornette, 2002), to interpret them. In unweighted networks, it is very useful to compare the real networks to random models, so that, by comparing their main characteristics, namely, the clustering coefficient and the average path length, important

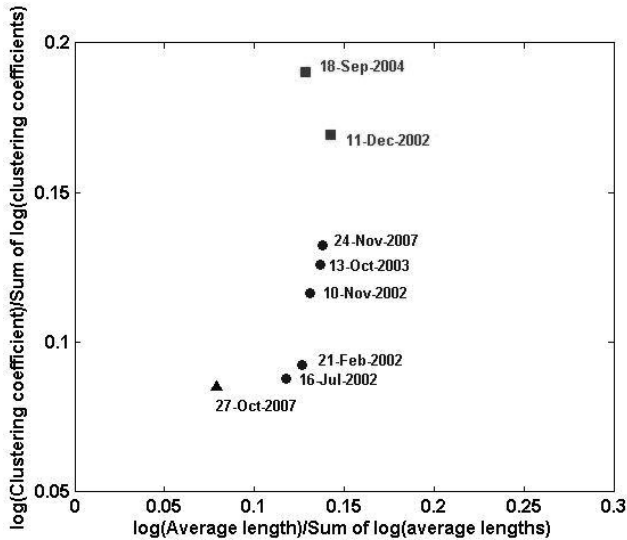


Fig. 3. Classification of the clusters. We see three clusters, one formed by the 27-Oct-2007, other formed by the 18-Sep-2004 and 11-Dec-2002 events, and the remaining. We used the logarithm of the clustering coefficient divided by the sum of the logarithms of the clustering coefficient, and of the logarithm of the average path length divided by the sum of the logarithms of the averaged path lengths.

conclusions can be reached. In our case, we can see that the values of the average path lengths are lower for the real networks. For the random graphs, the clustering coefficients are lower. We can therefore conclude that there is a small world effect (Watts and Strogatz, 1998) in the clusters because of the values for the path lengths and clustering coefficients in the real networks. We can also see that the simulations with the ETAS model give values for the clustering coefficient close to that of the real networks, but the average path lengths are larger. This means that any vertex on the graph can be reached from any other vertex in only a small number of steps, fewer steps than with the models. In Table 5 we show the result of simulating random networks merely by shuffling the weights of the real networks. If we take these values into account, we can see that, in general, the average path lengths are lower for the random model, and the clustering coefficients are lower too. Since the average path lengths of the real networks are higher, no small world behaviour is found for them. In any case, the ETAS model is better than the random model (whichever we use) for simulating the seismicity, as might be expected.

To classify the data in Table 4, we grouped them into different classes. The groups were obtained using a *k*-means algorithm (Teknomo, 2007), with the logarithm of the clustering coefficient and the logarithm of the average path length as variables, and taking into account that the mean standard deviation of the distances within each group was minimized. We normalized the variables by the sum of the values. We

Table 5. Values of the average path length and clustering coefficient for the random networks calculated by shuffling the weights.

Main shock	$L_{Rand.2}$ (error)	$C_{Rand.2}$ (error)
10-Nov-2002	$1.3667 \cdot 10^4$ $1.7561 \cdot 10^4$	$3.4335 \cdot 10^{-5}$ $3.3024 \cdot 10^{-5}$
11-Dec-2002	$6.2977 \cdot 10^4$ $1.8435 \cdot 10^4$	$7.3945 \cdot 10^{-8}$ $6.3092 \cdot 10^{-8}$
13-Oct-2003	$1.4501 \cdot 10^4$ $2.2156 \cdot 10^4$	$1.0486 \cdot 10^{-5}$ $8.6077 \cdot 10^{-6}$
18-Sep-2004	$1.4854 \cdot 10^2$ $3.8252 \cdot 10^1$	$1.5236 \cdot 10^{-8}$ $2.4113 \cdot 10^{-9}$
24-Nov-2007	$3.9554 \cdot 10^2$ $5.9992 \cdot 10^2$	$4.5301 \cdot 10^{-6}$ $2.6715 \cdot 10^{-6}$
21-Feb-2002	$1.2015 \cdot 10^4$ $2.6572 \cdot 10^4$	$4.3663 \cdot 10^{-4}$ $4.4739 \cdot 10^{-4}$
27-Oct-2007	$4.9800 \cdot 10^2$ $4.9857 \cdot 10^2$	$1.0000 \cdot 10^{-3}$ $2.6722 \cdot 10^{-4}$
16-Jul-2002	$3.0688 \cdot 10^3$ $7.5177 \cdot 10^3$	$6.8709 \cdot 10^{-4}$ $5.8363 \cdot 10^{-4}$

used the logarithm because we observed that small numbers can be very clustered, since there are different magnitude scales in both average path length and clustering coefficient. The best choice for three groups is depicted in Fig. 3, where the clusters are ordered by their magnitude and number of events. The standard deviation is 0.0106. With two groups, we have the lowest standard deviation (0.0192) for the 18 September 2004 and the 11 December 2002 in the same group, and another group with the remaining clusters. These two clusters are the ones with the highest main-shock magnitude and the most aftershocks. This analysis indicates that the topological properties of the clusters depend mainly on the magnitude of the main event, and not on the mechanism or mechanisms that originate them.

5 Conclusions

We proposed an analysis of the seismic clusters in a region based on the characteristics of weighted complex networks. In order to obtain the main clusters in a region, we used a distance based on the ETAS model (Ogata, 1998). We fitted the data to this model and performed a simple linkage clustering algorithm to obtain the main clusters. We analyzed the errors of both the fit and the simulated catalogues. In our approach, we used the ETAS model, although this is not the only seismicity-based model. Another option is the BASS model (Turcotte et al., 2007, Holliday et al., 2008), for example. The main difference between the ETAS and

BASS models is that the latter includes Båth's law (Båth, 1965) and is completely self-similar. However, it is normally used as a forward model, and for simulating one cluster only. Ogata (1998) provides several fitting functions that are based on real data, and this is why we used the ETAS model. In future research, other distances based on the BASS model could be used by implementing a similar procedure for fitting the data to a proposed BASS model.

We then analyzed and compared the clusters found to some models (one based on the ETAS model and two random models). We concluded that the ETAS model better approaches the seismicity near the Itoiz dam than the two random models we tried. The different values for the topological features analyzed seem to depend on the magnitude of the clusters and the number of events in each series, which are related. An event with higher magnitude can produce more aftershocks. We analyzed clusters both before and after the impoundment, and they were classified by their magnitude rather than the time interval within which they occurred. So, we cannot conclude that the 18 September 2004 earthquake behaves abnormally compared to others. However, only two global topological features have been analyzed in this classification. In order to discern the possible difference of this particular event we must have more clusters to analyze. A larger area must therefore be studied, with the difficulty inherent in analyzing a more complex region. Another option would be to study a larger time interval. In that case, the minimum magnitude would have to be increased to have a complete catalogue, and as a result smaller clusters would be found.

We found that the nodes with the higher strength are the most important in the traffic of the networks. The main shocks are associated with the events with the highest strength. However, the betweenness does not always coincide with the main events. We analyzed the weight, strength, and betweenness distributions, but we could not conclude that they followed a power law, because of the few data points involved.

We compared the real networks obtained with two random models in order to study the small-world phenomenon, in the same sense that it is studied for unweighted complex networks. When we used a theoretical random model (Garlaschelli, 2009), we found that there was small world behaviour. However, when all we did was shuffle the weights, results showed that the average path length was higher in the real networks, so no small world was found. No single answer was found for this question.

For future research, more properties based on the graph theory should be calculated in order to characterize the seismic networks better. We can also use this method in other places where no induced seismicity exists, and try to understand if we can differentiate the seismicity affected by artificial processes from that produced naturally.

Acknowledgements. We wish to thank the Confederación Hidrográfica del Ebro (CHE) for allowing us to use their facilities and their seismic data. We also thank the Instituto Geográfico Nacional (IGN), Spain. We thank Reik Donner and an anonymous referee for their useful comments and corrections. This work was partially supported by the Secretaría General para el Territorio y la Biodiversidad part of the Ministerio de Medio Ambiente, Rural y Marino, Spain, under grant 115/SGTB/2007/8.1, by EU with FEDER and by the research team RNM-194 of the Junta de Andalucía, Spain. The work of AJ is supported by a Juan de la Cierva grant, from the Spanish Government.

Edited by: J. Kurths

Reviewed by: S. Dana and another anonymous referee

References

- Abe, S. and Suzuki, N.: Small-world structure of earthquake network, *Physica A*, 337, 357–362, 2004a.
- Abe, S. and Suzuki, N.: Scale-free network of earthquake, *Europhys. Lett.* 65, 581–586, 2004b.
- Abe, S. and Suzuki, N.: Complex-network description of seismicity, *Nonlin. Processes Geophys.*, 13, 145–150, doi:10.5194/npg-13-145-2006, 2006.
- Akaike, H.: A new look at the statistical model identification, *IEEE T. Automat. Contr.*, 19(3), 716–72, 1974.
- Baiesi, M. and Paczuski, M.: Scale-free networks of earthquakes and aftershocks, *Phys. Rev. E*, 69, 066106, doi:10.1103/PhysRevE.69.066106, 2004.
- Baiesi, M. and Paczuski, M.: Complex networks of earthquakes and aftershocks, *Nonlin. Processes Geophys.*, 12, 1–11, doi:10.5194/npg-12-1-2005, 2005.
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A.: The architecture of complex weighted networks, *Proc. Natl. Acad. Sci. USA*, 101, 3747, doi:10.1073/pnas.0400087101, 2004.
- Barrat, A., Barthélemy, M., and Vespignani, A.: *Dynamical Processes on Complex Networks*, Cambridge University Press, Cambridge, 2008.
- Båth, M.: Lateral inhomogeneities in the upper mantle, *Tectonophysics*, 2, 483–514, 1965.
- Chen, S. and Chen, Q.: A weighted network model based on the preferential selection of edges, *Appl. Math. Sci.*, 1(23), 1145–1156, 2007.
- Clauset, A., Shalizi, C. R., and Newman, M. E. J.: Power-law distributions in empirical data, *SIAM Review*, 51, 661–703, 2009.
- Console, R., Jackson, D. D., and Kagan, Y. Y.: Using the ETAS Model for Catalog Declustering and Seismic Background Assessment, *Pure Appl. Geophys.*, 167, 819–830, 2010.
- Dal Moro, G., Pipan, M., and Gabrielli, P.: Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation, *J. Appl. Geophys.*, 61, 39–55, 2007.
- Davidson, J., Grassberger, P., and Paczuski, M.: Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity, *Phys. Rev. E*, 77, 066104, doi:10.1103/PhysRevE.77.066104, 2008.
- Dijkstra, E. W.: A Note on Two Problems in Connection with Graphs, *Numerische Math.*, 1, 269–271, 1959.

- Durá-Gómez, I. and Talwani, P.: Reservoir-induced seismicity associated with the Itoiz Reservoir, Spain: a case study, *Geophys. J. Int.*, 181, 343–356, 2010.
- Erdős, P. and Rényi, A.: On Random Graphs. I, *Pub. Math. Debrecen*, 6, 290–297, 1959.
- Erdős, P. and Rényi, A.: The Evolution of Random Graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5, 17–61, 1960.
- Erdős, P. and Rényi, A.: On the evolution of random graphs, *Bull. Int. Stat. Inst.*, 38, 343–347 1961.
- Freeman, L. C.: A set of measures of centrality based on betweenness, *Sociometry*, 40, 35–41, 1977.
- Furuya, S. and Yakubo, K.: Generalized strength of weighted scale-free networks, *Phys. Rev. E*, E78, 066104, doi:10.1103/PhysRevE.78.066104, 2008.
- Garlaschelli, D.: The weighted random graph model, *New J. Phys.*, 11, 073005, doi:10.1088/1367-2630/11/7/073005, 2009.
- Gleich, D.: MatlabBGL, available at: <http://www.mathworks.co.uk/matlabcentral/fileexchange/10922> (last access: July 2011), 2007.
- Gutenberg, B. and Richter, C. F.: Earthquake magnitude, intensity, energy and acceleration, *Bull. Seism. Soc. Am.*, 46, 105–145, 1956.
- Helmstetter, A. and Sornette, D.: Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models, *Phys. Rev. E*, 66, 061104, doi:10.1103/PhysRevE.66.061104, 2002.
- Holliday, J. R., Turcotte, D. L., and Rundle, J. B.: A Review of Earthquake Statistics: Fault and Seismicity-Based Models, ETAS and BASS, *Pure Appl. Geophys.*, 165, 1003–1024, 2008.
- Instituto Geográfico Nacional (IGN): available at: <http://www.ign.es/ign/es/IGN/home.jsp>, 2010.
- Jiménez, A., García, J. M., and Romacho, M. D.: Simultaneous Inversion of Source Parameters and Attenuation Factor Using Genetic Algorithms, *Bull. Seism. Soc. Am.*, 94, 1401–1411, 2005.
- Jiménez, A., Tiampo, K. F., and Posadas, A. M.: Small world in a seismic network: the California case, *Nonlin. Processes Geophys.*, 15, 389–395, doi:10.5194/npg-15-389-2008, 2008.
- Jiménez, A., Tiampo, K. F., Posadas, A. M., Luzón, F., and Donner, R.: Analysis of complex networks associated to seismic clusters near the Itoiz reservoir dam, *EPJ Special Topics*, 174, 181–195, 2009.
- Kay, M.: available at: <http://web.mit.edu/cocosci/isomap/code/dijkm>, 2001.
- Luzón, F., García-Jerez, A., Santoyo, M. A., and Sánchez-Sesma, F. J.: A hybrid technique to compute the pore pressure changes due to time varying loads: application to the impounding of the Itoiz reservoir, northern Spain, in: *Poromechanics-iv*, edited by: Ling, H., Smyth, A., and Betti, R., Destech Publications, Inc., Lancaster, Pennsylvania, ISBN: 978-1-60595-006-8, 1109–1114, 2009.
- Luzón, F., García-Jerez, A., Santoyo, M. A., and Sánchez-Sesma, F. J.: Numerical modelling of pore pressure variations due to time varying loads using a hybrid technique: the case of the Itoiz reservoir (Northern Spain), *Geophys. J. Int.*, 180, 327–338, 2010.
- MacQueen, J. B.: Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1, 281–297, 1967.
- Ogata, Y.: Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, 83, 401, Applications, 9–27, 1988.
- Ogata, Y.: Space-time point-process models for earthquake occurrences, *Ann. I. Stat. Math.*, 50(2), 379–402, 1998.
- Ogata, Y.: Seismicity analysis through point-process modelling: A review, *Pure Appl. Geophys.*, 155, 471–507, 1999.
- Onnela, J. P., Saramäki, J., Kertész, J., and Kaski, K.: Intensity and coherence of motifs in weighted complex networks, *Phys. Rev. E*, 71, 065103, doi:10.1103/PhysRevE.71.065103, 2005.
- Park, K., Lai, Y.-C., and Ye, N.: Characterization of weighted complex networks, *Phys. Rev. E*, 70, 026109, doi:10.1103/PhysRevE.70.026109, 2004.
- Rivas-Medina, A., Santoyo, M. A., Luzón, F., Benito, B., Gaspar-Escribano, J. M., and García-Jerez, A.: Seismic Hazard and ground motion characterization at the Itoiz dam (Northern Spain), *PAGEOPH*, in press, 2010.
- Ruiz, M., Gaspa, O., Gallart, J., Díaz, J., Pulgar, J. A., García-Sansegundo, J., López-Fernández, C., and González-Cortina, J. M.: Aftershocks series monitoring of the September 18, 2004 $M = 4.6$ earthquake at the western Pyrenees: A case of reservoir-triggered seismicity?, *Tectonophysics*, 424, 223–243, 2006.
- Santoyo, M. A., García-Jerez, A., and Luzón, F.: A subsurface stress analysis and its possible relation with seismicity near the Itoiz Reservoir, Navarra, Northern Spain, *Tectonophysics*, 482, 205–215, 2010.
- Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K., and Kertész, J.: Generalizations of the clustering coefficient to weighted complex networks, *Phys. Rev. E*, 75, 027105, doi:10.1103/PhysRevE.75.027105, 2007.
- Tadic, B., Rodgers, G. J., and Thurner, S.: Transport on Complex Networks: Flow, Jamming and Optimization, *Int. J. Bifurcation Chaos*, 17, 2363–2385, 2007.
- Teknomo, K.: K-Means Clustering Tutorials, available at: <http://people.revoledu.com/kardi/tutorial/kMean/>, 2007.
- Turcotte, D. L., Holliday, J. R., and Rundle, J. B.: BASS, an alternative to ETAS, *Geophys. Res. Lett.*, 34, L12303, doi:10.1029/2007GL029696, 2007.
- Wang, Q., Schoenberg, F. P., and Jackson, D. D.: Standard errors of parameter estimates in the ETAS model, *Bull. Seism. Soc. Am.*, 100, 1989–2001, 2010.
- Watts, D. J. and Strogatz, S. H.: Collective dynamics of “small-world” networks, *Nature*, 393, 440–442, 1998.
- Zhuang, J., Ogata, Y., and Vere-Jones, D.: Analyzing earthquake clustering features by using stochastic reconstruction, *J. Geophys. Res.*, 109(B5), B05301, doi:10.1029/2003JB002879, 2004.