Nonlinear Processes
in Geophysics

Open Access

Scale free properties in a network-based integrated approach to earthquake pattern analysis

M. Suteanu

Department of Mathematics & Computing Science, Saint Mary's University, Halifax, Canada

Correspondence to: M. Suteanu (mirela.suteanu@gmail.com)

Abstract. This paper proposes a network-based method for the assessment of earthquake relationships in space–time–magnitude patterns. It is shown that networks with high values for the minimum edge weight W_{min} enjoy strong scaling properties, as opposed to networks with low values for W_{min}, which exhibit no such properties. The scaling behavior along the spectrum of W_{min} values, in conjunction with the robustness regarding parameter variations, endorse the idea of a relationship between fundamental properties of seismicity and the scaling properties of the earthquake networks. Results of this method are further applied for the study of temporal changes in volcanic seismicity patterns.

1 Introduction

Extensive research is dedicated to earthquake pattern analysis in an ongoing effort to understand the laws that govern seismicity. Correlations in earthquake patterns have been found in magnitude (Gutenberg and Richter, 1954; Lippiello et al., 2012b), time (Omori, 1894; Shcherbakov et al., 2004; Shcherbakov et al., 2006), and space (Turcotte, 1977; Felzer and Brodsky, 2006; Lippiello et al., 2009). Integrated approaches have been developed to find space–time–magnitude patterns (Bak et al., 2002). Network-based approaches have shown not only that networks of correlated earthquakes can be created, but also that these networks enjoy scaling properties (Baiesi and Paczuski, 2004, 2005; Davidsen et al., 2008; Suteanu and Suteanu, 2011).

A space–time–magnitude metric defined for directed networks of earthquakes was proposed by Baiesi and Paczuski with their seismicity declustering method (Baiesi and Paczuski, 2004, 2005). The study in our paper also creates directed weighted networks of earthquakes with the purpose

of assessing relationships between them. Given the ubiquity of power laws governing earthquake distributions (Nanjo and Nagahama, 2000; Lapenna et al., 2000; Shcherbakov et al., 2004, 2006; Carbone et al., 2005; Felzer and Brodsky, 2006; Bunde and Lennartz, 2012; Lippiello et al., 2012a, b), power law forms are used to estimate quantitatively the relationships between events in a space–time–size perspective. However, there are major differences between Baiesi and Paczuski's model and the work presented in this paper. Not only do the two methods use different metrics, but, most importantly, they use different criteria for the discrimination of interrelated earthquakes from the rest of the set: in the method of Baiesi and Paczuski (2004, 2005) the criterion is the maximization of a correlation function, while in this study series of networks are created, assessed and searched for scale free properties. Since the main component in Baiesi and Paczuski's metric is a function that is exponential in magnitude, their method effectively addresses the identification of event clusters around the largest shocks, while our method addresses earthquakes of all sizes that are considered close enough in space–time–magnitude to be interrelated.

Although a new quantitative metric is defined in our study and a new type of network is built, the results show power law properties that are consistent with previous work of Baiesi and Paczuski (2005) and with their interpretation that the underlying correlations of the seismicity structure are unambiguous, sufficiently strong to survive the approximation of the metric, and can be reliably detected.

Our method is applied to seismicity associated with hotspot volcanism in Hawaii. The earthquakes are seen as sets of space–time–magnitude events that can be related to each other, while the quality of the interactions among earthquakes can vary over time. In order to assess these interactions and their change in time, an integrative approach that

maps seismic information to directed weighted networks is developed. Different classes of networks of earthquakes are studied, and results show scale free properties that are robust with respect to certain variations in the definition of the networks. Networks with values of the minimum edge weight W_{min} in the middle to upper range of the spectrum of edge weight values enjoy strong scaling properties, as opposed to networks with W_{min} in the lower range, which exhibit poor or no such properties. It is shown that network parameters studied for successive event windows are able to reflect the way the relationships between earthquakes are changing over time, and that patterns of change can be related to important events in the life of the volcanic system.

2 Construction of the earthquake networks

The epicenters of earthquakes that could be related to other earthquakes are seen as network nodes that are connected through directed edges. The edge direction is given by the temporal succession of the events. Ideally, only interrelated earthquakes can be nodes of this network, and the edges that link them to other nodes carry a space–time–magnitude weight. Therefore, a combination of three factors is evaluated before deciding whether or not any two earthquakes belong to the network: the size (magnitude) of the first occurring event, and its proximity to future events in space and in time. There are many possible combinations of these three factors. Even small earthquakes may be related to subsequent events if the latter were close enough in space and time.

In order to assess quantitatively the relationship between earthquakes, three weight variables are defined: the weight in distance w_d, the weight in time w_t, and the weight in magnitude w_m. A total weight W characterizes every edge as a combination of the previous three variables. Considering the Gutenberg–Richter law, the Omori law, and other scaling relationships regarding the distributions of earthquakes in space, time, and magnitude (e.g. Lei and Kusunose, 1999; Richards-Dinger et al, 2010; Felzer and Brodsky, 2006; Shcherbakov et al., 2006; Lennartz et al., 2008; Lippiello et al., 2009, 2012a; Sanchez and Shcherbakov, 2012), the following forms for the node weights of any one edge have been chosen:

(a) Distance weight:

$$w_d = cd^r, \ r < 0, \tag{1}$$

where d is the spatial distance between the two nodes of the edge measured in km, and c is a positive constant.

(b) Time weight:

$$w_t = st^p, \ p < 0, \tag{2}$$

where t is the time interval between the two nodes of the edge measured in hours, and s is a positive constant.

(c) Magnitude weight:

$$w_m = \frac{m}{m_{max}} \quad , \tag{3}$$

where m is the magnitude of the first occurring node of the edge, and m_{max} is the maximum magnitude value in the data set.

The total weight of an edge is calculated as the product of the weights in space, time, and magnitude. Only the nodes that carry enough total weight belong to the network, which means that only edges that have a value of the total weight W higher than a minimum threshold W_{min} are selected for the network:

$$W = \begin{cases} w_d \cdot w_t \cdot w_m, \ W \geq W_{min} \\ 0, \quad\quad\quad\quad W < W_{min} \end{cases} . \tag{4}$$

The generality of this definition allows various combinations of space–time–magnitude correlations between any two events and includes the possibility of multiple interactions for any given event: any node can have any number of edges that enter the node and any number of edges that leave the node, as long as these edges carry enough total weight.

For practical reasons and with the purpose of avoiding singularities, a small cutoff value is used for the weights in space and time (it is also reasonable to assume that all earthquakes that are very close in space or in time could be related to each other). Therefore, modified forms of Eqs. (1) and (2) are used in the actual construction of the networks:

$$w_d = \begin{cases} 1, \quad d \leq d_{min} \\ cd^r, \ d \geq d_{min}, \ r < 0 \end{cases} \tag{5}$$

$$w_t = \begin{cases} 1, \quad t \leq t_{min} \\ st^p, \ t \geq t_{min}, \ p < 0 \end{cases} . \tag{6}$$

Various values for the exponents r and p, and for the cutoff values d_{min} and t_{min}, are explored. The constants c and s are calculated using the boundary conditions

$$w_d = cd_{min}^r = 1 \tag{7}$$

and

$$w_t = st_{min}^p = 1. \tag{8}$$

An essential difference between the total edge weight defined in this paper and Baiesi and Paczuski's (2004, 2005) metric consists in the contribution of each of the three factors (time interval, space interval, and magnitude). In contrast with their approach, in this paper the three components (time interval, spatial distance, and magnitude) are seen independently, as separate components that can have comparable contributions to the total edge weight W. This is accomplished by limiting the upper value of each of the three components to 1. The definition of a magnitude weight proportional to m, and not exponential in m, is therefore meant to support a balance of

factors in the total edge weight formula. The resulting networks and network distributions are governed by statistical contributions of each of the three components. This choice is especially important in the study of volcano-tectonic seismicity, where the seismic sources are associated not only with tectonic stress, but also with thermodynamic processes and the dynamics of gas, fluid and solid.

To simplify the computation, a maximum interval of influence in time T_{max}, and a maximum interval of influence in space D_{max}, are assigned. Although fixed values of T_{max} and D_{max} are chosen to create the initial network, making this choice is different from making the parameter choice in window declustering methods or in Reasenberg's cluster method (Reasenberg, 1985). For example, in order to identify aftershocks, Knopoff and Gardner define space–time windows that are functions of the mainshock magnitude (Knopoff and Gardner, 1972; Gardner and Knopoff, 1974); various choices of parameter values lead to significant variations in the aftershock identification. Reasenberg's algorithm (Reasenberg, 1985) identifies foreshocks and aftershocks within a cluster based on Omori's law for the cluster's time extension and on a window-type function for the cluster's spatial extension; also in this case, different choices of fixed parameter values may lead to substantially different estimates of the correlations between earthquakes. In this paper, T_{max} and D_{max} receive fixed values only with the purpose of simplifying the computation. In principle, T_{max} and D_{max} could cover the whole extent of the catalogue in time and space. The study shows that the final outcome is not affected by the initial choice of T_{max} and D_{max}, since large distances and long time intervals between events result in very small values of the edge weights w_d and w_t, and therefore lead to small values of the total edge weight W. The links carrying small weights are eliminated from the network in the next step anyway, when the network definition (4) is applied.

Different values for the maximum interval of influence in time T_{max}, and for the maximum interval of influence in space D_{max}, are explored, as well as various values for r, p, d_{min} and t_{min}. This way, a series of network classes are generated with the purpose of creating a structured framework for the analysis: all networks that belong to a certain class B, C, ..., N share the same initial choice of D_{max}, r, d_{min}, T_{max}, p, t_{min}. Letters are used to name the classes of earthquake networks, and the choice of every letter has only a classification purpose. A description of the classes that have been studied is shown in Table 1.

In each class, an initial network is created when assigning the specific values to parameters. For clarity, an index $_0$ is used to describe these initial networks: B_0, C_0, etc. The highest value of the total edge weight in each class, H, is the highest value of the total edge weight in the initial network (NETWORK CLASS)$_0$, while the lowest value of the total edge weight in the class, L, is the lowest value of the total edge weight in the initial networks (NETWORK CLASS)$_0$. For example, the highest value in network B_0 is $H = 1$, and

Table 1. All classes. T_{max} is the maximum time interval between events, D_{max} is the maximum distance between events, r is the exponent of the distance weight w_d (Eq. 1), p is the exponent of the time weight w_t (Eq. 2), d_{min} and t_{min} represent cutoff values, H is the highest value of the total edge weight in the class, and L is the lowest value of the total edge weight in the class.

Class	T_{max} (days)	D_{max} (km)	r	p	t_{min} (h)	d_{min} (km)	H	L
B	10	30	−1.35	−1	1	1	1.00	1.01×10^{-5}
C	10	30	−1.35	−1	0.5	0.2	0.65	5.76×10^{-7}
D	30	30	−1.35	−1	1	1	1.00	3.39×10^{-6}
E	40	50	−1.35	−1	0.05	0.2	6.07	7.24×10^{-9}
F	7	10	−1.35	−1	0.05	0.1	2.72	1.44×10^{-7}
G	7	10	−1.35	−1	0.05	0.025	2.72	2.22×10^{-8}
H	8	10	−1.35	−1	0.5	0.2	0.65	3.19×10^{-6}
I	8	11	−1.35	−1	0.05	0.1	2.72	1.11×10^{-7}
J	8	10	−1.35	−1	1	1	1.00	5.60×10^{-5}
L	7	10	−1	−0.5	14	2	1.00	1.39×10^{-2}
M	7	10	−1	−0.5	1	1	1.00	1.86×10^{-3}
N	40	50	−1	−0.5	0.5	0.2	0.66	2.19×10^{-5}

the lowest value in network B_0 is $L = 1.01 \times 10^{-5}$. In general, the first networks B_0, C_0, ..., N_0 are simply collections of earthquakes, and not networks of interrelated events, and they serve for the operational initiation of the method.

Inside each class, specific values for the threshold W_{min} define distinct networks.

3 Data

The data source for this study is the Advanced National Seismic System (ANSS) catalog for the Big Island of Hawaii, with events ranging from 1 January 1989 to 31 December 2012. Figure 1 shows an example of a network of earthquakes. By zooming in, the actual nodes and edges can be seen (Fig. 2).

The number of events in the catalog is 64 392. For catalog completeness, only events with magnitude $m \geq 1.6$ are used in the analysis (37 451 earthquakes); the b value in the Gutenberg–Richter magnitude-frequency distribution for this data set is $b \approx 0.99$ (Fig. 3).

Different sets of networks in various network classes have been analyzed. The following values of the parameters have been studied:

– D_{max}: 10 km; 30 km; 50 km.

– T_{max}: 7 days; 8 days; 10 days; 30 days; 40 days.

– d_{min}: 0.025 km; 0.1 km; 0.2 km; 1 km; 2 km.

– t_{min}: 3 min; 30 min; 1 h; 14 h.
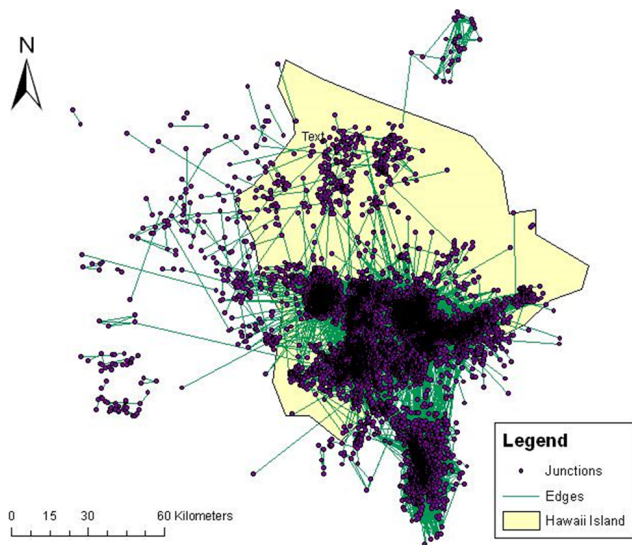
– r: −1; −1.35.

– p: −0.5; −1.

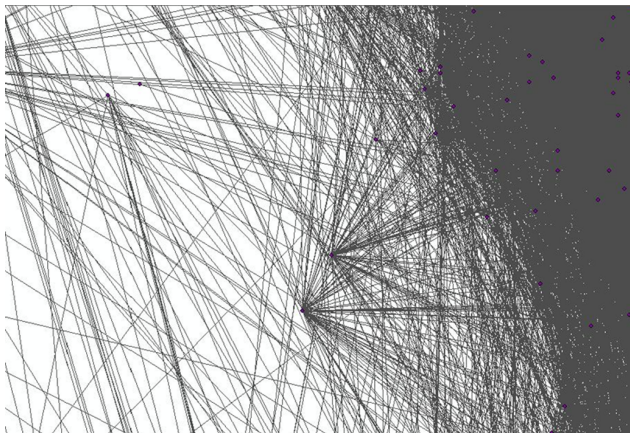**Fig. 1.** The earthquake network: an example for the Big Island of Hawaii.

Fig. 2. Zoomed-in example of nodes and edges of the earthquake network presented in Fig. 1.

Depending on the lowest edge weight value L and the highest edge weight value H in each class, specific values for the threshold $W_{\min}$ are chosen in order to create and analyze distinct networks inside each class.

4 Results

4.1 Network parameters and analysis

Network parameters such as node connectivity and node weight distribution (Boccaletti et al., 2006) are assessed for sets of earthquake networks in different classes. The node connectivity (degree) represents the total number of edges in the node. The number of edges that enter the node (in-connectivity), and the number of edges that go out of the
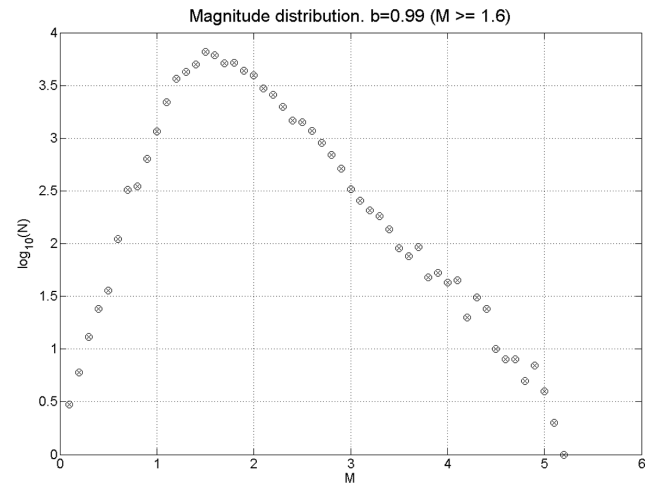

Fig. 3. Gutenberg–Richter magnitude-frequency distribution for the Hawaii data set ranging from 1 January 1989 to 31 December 2012.

node (out-connectivity) are also studied. Similarly, the node weight, in-weight, and out-weight are calculated and analyzed. The node weight is given by the sum of the weights of all edges in the node, the in-weight is the weight of all edges that enter the node, and the out-weight is the weight of all edges that go out of the node.

The results of the studies that were performed on all classes in Table 1 show that, for each class of networks, the connectivity distribution enjoys power law properties for all networks that have $W_{\min}$ in the upper range of the interval between the lowest edge weight value L and the highest edge weight value H in the class, while for networks that have values of $W_{\min}$ in the lower range of the interval between L and H, the connectivity distribution is irregular and scattered. In many of the networks with irregular shapes, constants with a power law tail are present.

For example, a study on the large class E of networks is shown in Fig. 4. The class definition, the initial network E_0 and the network characteristics are presented in Table 2. The large size of the class (8 488 767 edges) originates in the assumption that, for any earthquake of the network, the interval of influence may go up to 40 days in time, and up to 50 km in space. This is a broad supposition for the active volcanic system of Hawaii; for the majority of these volcanic earthquakes, which do not have large magnitudes (Fig. 3), correlations with earthquakes so far away in space and time are quite unlikely. This situation is suggestively illustrated in Fig. 4. The set of six images in Fig. 4 shows the change in shape of the connectivity distribution when the minimum value of the total weight changes from low values, such as in networks E1, E2, E3, towards higher values, as in networks E10, E11, from highly irregular and scattered shapes to well-organized shapes that exhibit significant power law properties. This behavior is characteristic of networks in all the other classes. When weak links are included (low values of $W_{\min}$), most
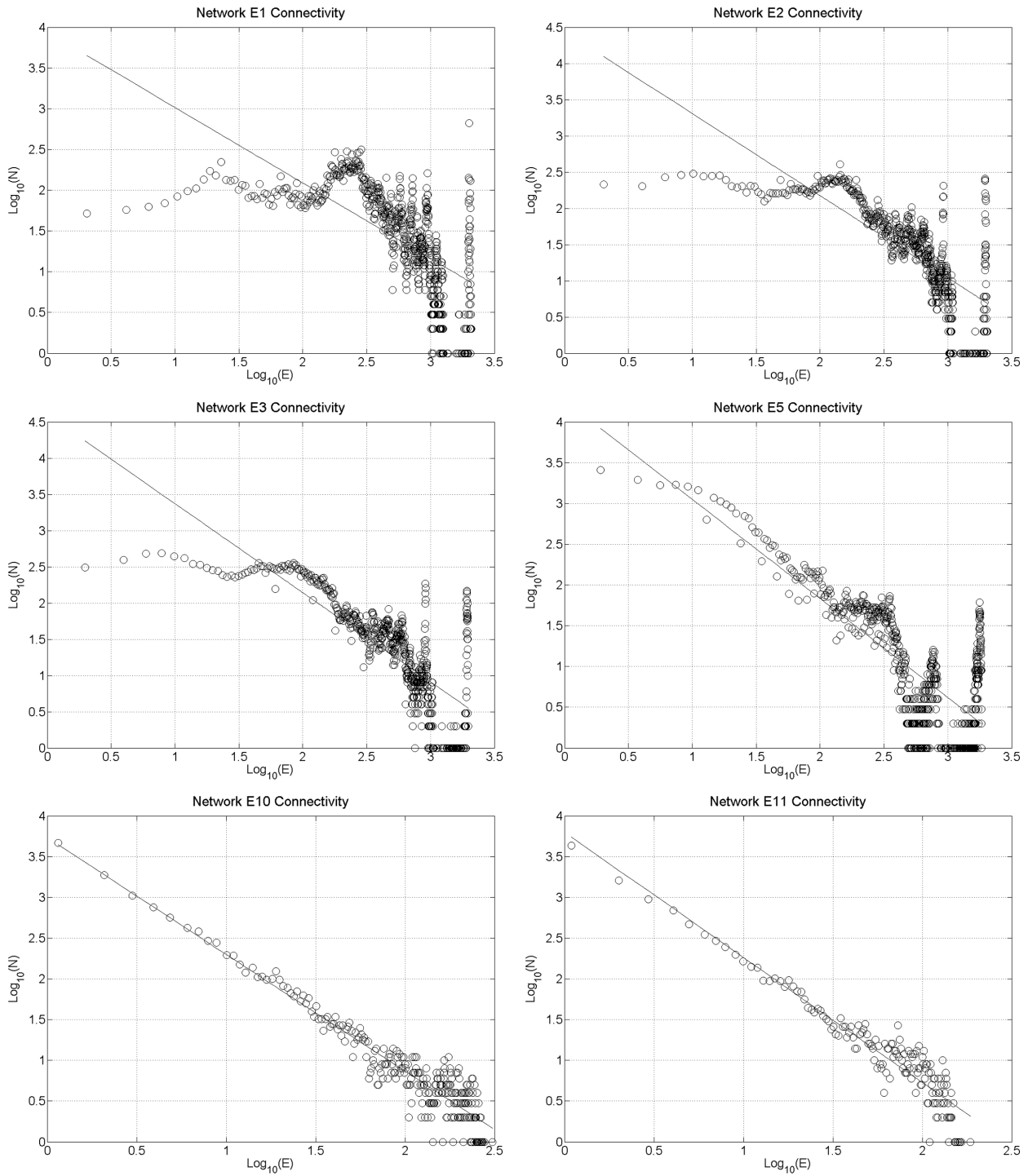
Fig. 4. Connectivity distribution for networks E1, E2, E3, E5, E10, and E11. N (on the Y axis) is the number of nodes that have a number E of edges (on the X axis).

of the nodes in the emerging networks have little or no relationship with each other, and results show that this choice for network nodes translates into irregular and scattered shapes of the node connectivity. When only strong links are retained (high values of W_{min}), the events selected to participate in

the network are primarily earthquakes that are related to each other, and the results show that the underlying properties of seismicity manifest themselves in the well-organized, scale free appearance of the node connectivity.

Table 2. Class E networks. E_0 is the initial network that was generated using the parameter values shown in the first column. See Table 1 for the meaning of T_{max}, D_{max}, r, p, d_{min}, t_{min}, H, and L.

Class E definition	Characteristics of the initial network E_0	Network		Number of nodes	Number of edges
		Name	W_{min}		
– $T_{max} = 40$ days	– 37 441 nodes	E1	$W \geq 10^{-8}$	37 441	8 461 301
– $D_{max} = 50$ km	– 8 488 767 edges	E2	$W \geq 5 \times 10^{-8}$	37 419	6 628 575
– $r = -1.35$	– $H = 6.07$	E3	$W \geq 10^{-7}$	37 357	5 660 485
– $p = -1$	– $L = 7.24 \times 10^{-9}$	E4	$W \geq 5 \times 10^{-7}$	36 808	3 782 245
– $d_{min} = 0.2$ km		E5	$W \geq 10^{-6}$	36 172	3 102 765
– $t_{min} = 3$ min		E6	$W \geq 5 \times 10^{-6}$	32 886	1 784 667
		E7	$W \geq 10^{-5}$	30 377	1 330 446
		E8	$W \geq 5 \times 10^{-5}$	23 032	584 548
		E9	$W \geq 10^{-4}$	20 063	376 852
		E10	$W \geq 5 \times 10^{-4}$	14 091	117 835
		E11	$W \geq 10^{-3}$	11 955	68 408

The threshold values W_{min} are evaluated in the context of the whole set of earthquakes in the class, and not individually. W_{min} values that identify networks of interrelated events emerge from the global assessment of network properties: they are found as those values for which scale free properties appear and become stronger when subsequent networks are created using increasing values of W_{min}. In this sense, the threshold W_{min} is a global parameter; it belongs to a range of values in the middle to upper zone of the interval between the lowest edge weight value L and the highest edge weight value H in any given class of networks. This represents another major difference between this method and the method of Baiesi and Paczuski (2004, 2005).

Moreover, the generality of this method allows a variety of correlations between earthquakes: any event of the network can have any number of predecessors and any number of successors if the corresponding edges carry enough space–time–magnitude weight, with no arbitrary limitation on magnitude, time, or distance.

Power law properties can also be found in the distributions of time intervals and distances between nodes. Figure 5 shows the distribution of times intervals in the initial network E_0, i.e., the distribution of all the time intervals between any two earthquakes within a space–time window that is quite large in the Hawaii volcano-tectonic context. As shown in Fig. 5, there is a distinct scale free zone that goes up to 7 days with a power law exponent of -0.5. The peak between 7 and 15 days with a maximum at 11 days is consistent with studies that show that the precursory sequences in Hawaii follow a power law acceleration with 10–15 days before eruption (Chastin and Main, 2003).

The distribution of the distances between events also has scaling properties. Figure 6 shows the distribution of space intervals in the initial network E_0, which exhibits an exponent of -1 for the power law interval between 1 and 10 km.
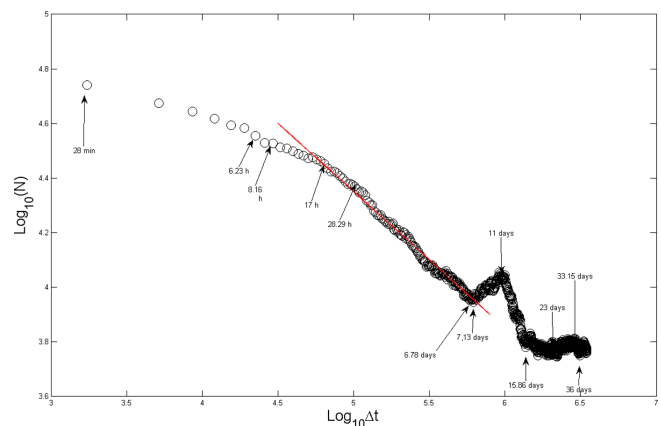

Fig. 5. Distribution of time intervals (in s) between any two nodes in the initial network E_0. N is the number of time intervals of Δt seconds between any two earthquakes in a space–time window of 50 km and 40 days. The red line represents a reference line with the slope of -0.5.

Although the context is volcano-tectonic and not all earthquakes in this space–time window are interrelated, the overall shape of the distribution shows remarkable similarities with the distribution of distances of aftershocks from the mainshock in Lippiello et al. (2009): an increase up to a maximum value (1 km in this case), followed by a power law decrease. We believe that the peaks with maximums at 13, 18, etc. km refer to events that are spatially clustered around the neighboring volcanoes, distinct vents and fracture zones, and reflect the spatial characteristics of the Hawaii volcanic system: as shown in Fig. 1, the network exhibits intense clusters spatially centered on the volcanoes, their vents and fracture zones, which are situated at distances compatible with the peaks in Fig. 6.

Table 3. Class M networks. M_0 is the initial network that was generated using the parameter values shown in the first column. See Table 1 for the meaning of T_{max}, D_{max}, r, p, d_{min}, t_{min}, H, and L.

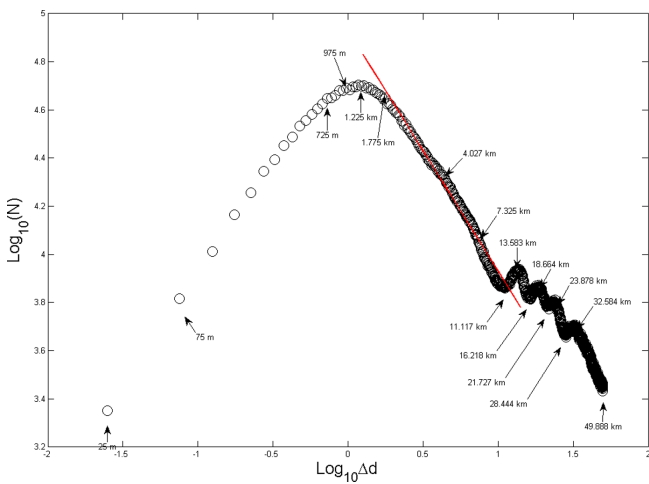| Class M definition	Characteristics of the initial network M_0	Network		Number of nodes	Number of edges
		Name	W_{min}		
$- T_{max} = 7$ days	$- 33\,065$ nodes	M1	$W \geq 1.5 \times 10^{-2}$	26 919	1 270 458
$- D_{max} = 10$ km	$- 1\,913\,280$ edges	M2	$W \geq 2 \times 10^{-2}$	25 396	1 088 015
$- r = -1$	$- H = 1$	M3	$W \geq 3 \times 10^{-2}$	22 682	782 659
$- p = -0.5$	$- L = 1.86 \times 10^{-3}$	M4	$W \geq 4 \times 10^{-2}$	20 440	584 686
$- d_{min} = 1$ km		M5	$W \geq 5 \times 10^{-2}$	18 825	443 533
$- t_{min} = 1$ h		M6	$W \geq 10^{-1}$	14 006	149 235

Fig. 6. Distribution of distances (in km) between any two nodes in the initial network E_0. N is the number of distances of Δd km between any two earthquakes in a space–time window of 50 km and 40 days. The red line represents a reference line with the slope of -1.

Fig. 7. Class M networks. **(a)** The dependency of the number N of edges (circles) and the number N of nodes (squares) on the minimum weight W_{min}. **(b)** The dependency of the number of edges on the number of nodes.

As another example, one of the network classes, class M, was created with characteristics drawn from the two distributions shown in Figs. 5 and 6: $T_{max} = 7$ days, $p = -0.5$, $D_{max} = 10$ km, $r = -1$. The full description of class M is summarized in Table 3.

The choice of the parameter values in the definition of class M was largely based on specific statistical characteristics of the data set; however, the analysis reveals the same behavior and the same qualitative patterns as those found in all the other classes. For example, Fig. 7a shows the exponential character of the dependency of the number of edges and number of nodes on the minimum weight in the network, and Fig. 7b shows the power law dependency of the number of edges on the number of nodes for class M of networks.

In the example of class M, the connectivity, in-connectivity and out-connectivity distributions are assessed for networks M1–M6. Since the maximum intervals in time and space T_{max} and D_{max} that define class M have rather low
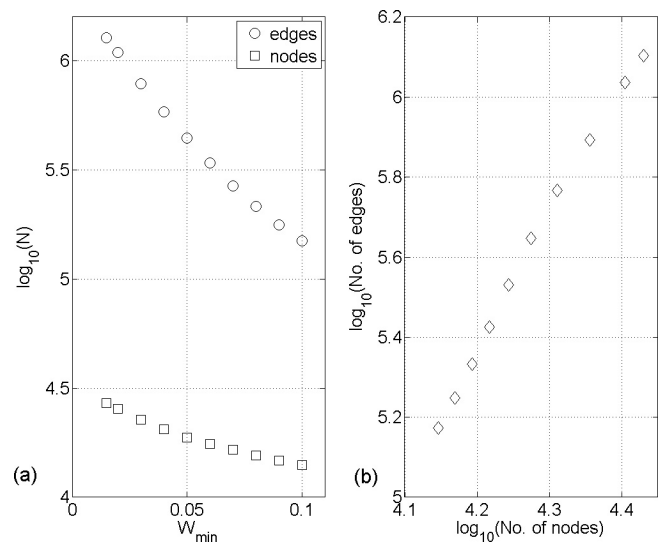
values, inside class M even networks with relatively low values of W_{min} display power law properties, as illustrated in Fig. 8a. The connectivity distributions for the other five networks M2–M6 are quite similar; another example is shown for network M5 in Fig. 9a. The exponent β for each type of connectivity distribution (all, in, out) is calculated and a graph with all β values is shown in Fig. 10. Similarly, power law properties are detected for the weight, in-weight and out-weight distributions in networks M1–M6, as shown in Figs. 8b and 9b, and the exponent γ (Albert and Barabasi, 2002) for each weight distribution is illustrated in Fig. 11.

In general, the analysis shows that the scale free behavior observed in all networks with superior values of W_{min} inside their class is remarkably robust with respect to variations of parameter values r, p, d_{min}, and t_{min}. Moreover, networks of the same class that have lower values of W_{min} exhibit poor scaling characteristics or even no such characteristics at all.
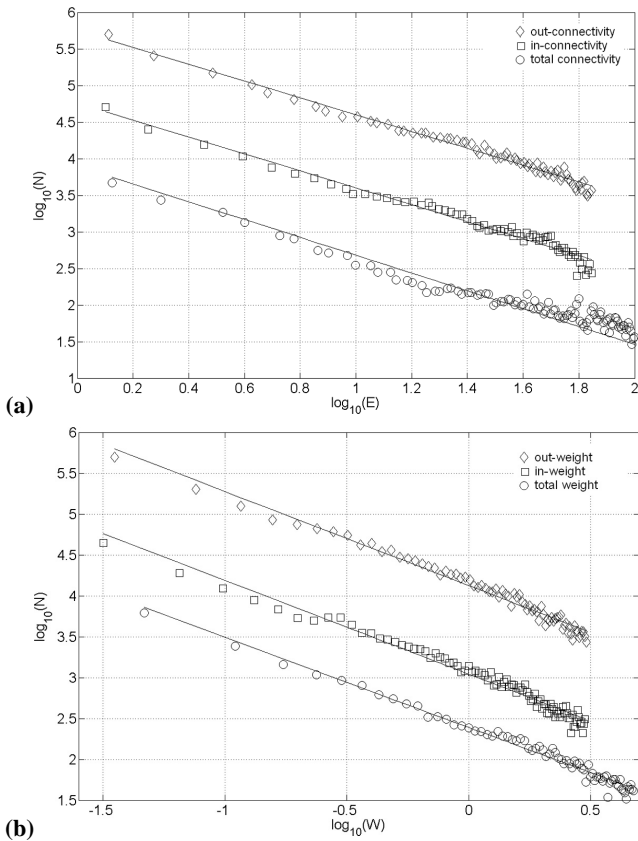
Fig. 8. Connectivity and weight distributions for network M1. The curves are shifted by 10 units each along the Y axis. **(a)** Total connectivity (circles), in-connectivity (squares), out-connectivity (diamonds). N is the number of nodes that have a number of edges E (on the X axis). **(b)** Total weight (circles), in-weight (squares), out-weight (diamonds). N is the number of nodes that have a weight of W (on the X axis).

Fig. 9. Connectivity and weight distributions for network M5. The curves are shifted by 10 units each along the Y axis. **(a)** Total connectivity (circles), in-connectivity (squares), out-connectivity (diamonds). N is the number of nodes that have a number of edges E (on the X axis). **(b)** Total weight (circles), in-weight (squares), out-weight (diamonds). N is the number of nodes that have a weight of W (on the X axis).

The scaling behavior along the spectrum of W_{min} values, in conjunction with the robustness regarding parameter variations, endorse the idea of a relationship between fundamental properties of seismicity and the scaling characteristics found in earthquake networks.

This observation suggests that a way of testing the reliability of the method is to question the identity of the earthquakes selected in networks that possess scaling properties. If the method is reliable, networks with strong scaling properties should retain only the nodes that correspond to earthquakes that are truly related to each other, regardless of the choice in the parameter values. Since a series of parameter values were explored in this study, a discussion of a few samples of results would be relevant. As shown in Table 1, each specific set of parameter values corresponds to the definition of a class. Table 4 compares four networks with strong scale free properties (D9, J6, M6, and E10) from four different combinations of initial parameters, i.e., from four different classes, with the purpose of showing that all four of them identify the same correlated earthquakes. D9 is the most
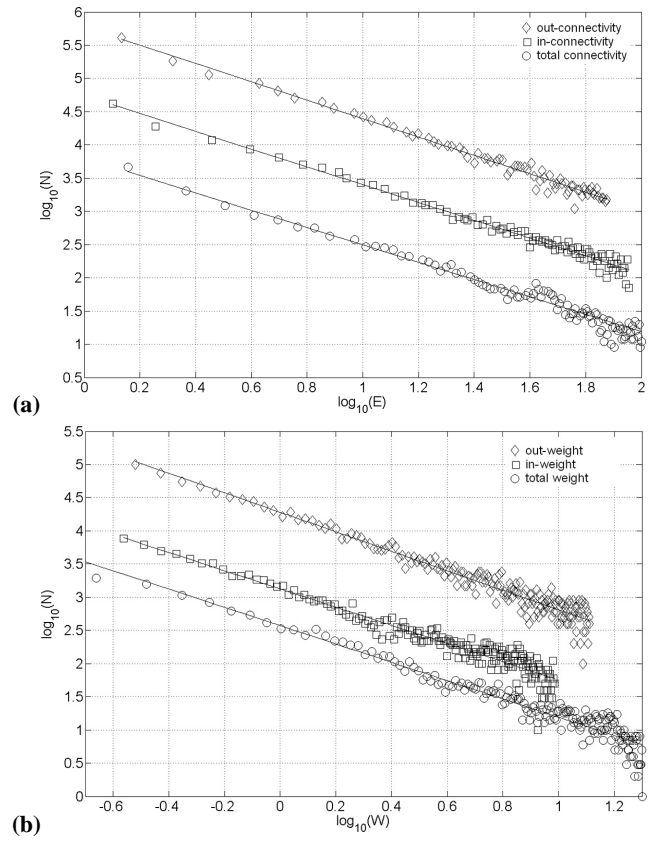
selective network, having 11 966 nodes (the lowest number of nodes), the next selective is J6, with 13 523 nodes, while M6 and E10 have almost the same number of nodes, 14 006, and 14 091, respectively. Although they come from different classes, with different characteristics, all these networks possess a high value of W_{min} inside their class, and their connectivity distributions enjoy significant power law properties. The question is how many earthquakes selected in the smallest network, D6, have also been selected in the slightly larger networks J6, M6, and E10, then how many earthquakes chosen for network J6 have also been accepted in networks M6 and E10, and how many earthquakes included in network M6 have also been included in network E10. The results in Table 4 show that all the earthquakes selected in network D9 have also been selected in networks J6 and M6, and 98.72 % have also been included in E10, regardless of the variations in the choice of the parameter values. In network J6, 97.27 % of the earthquakes have also been selected by the slightly larger network M6, and 93.93 % of the events have also been chosen for network E10. Finally, 91.34 % of the events included
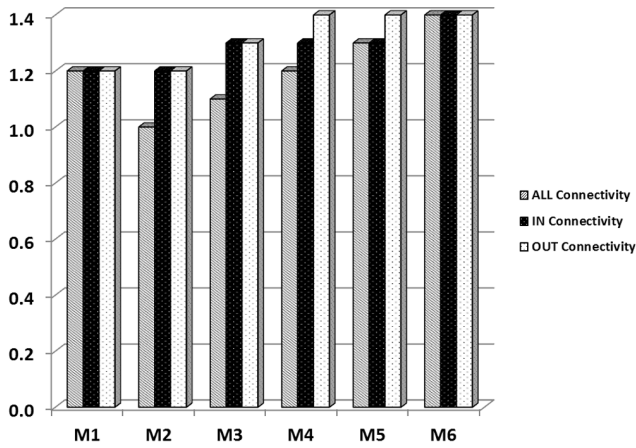
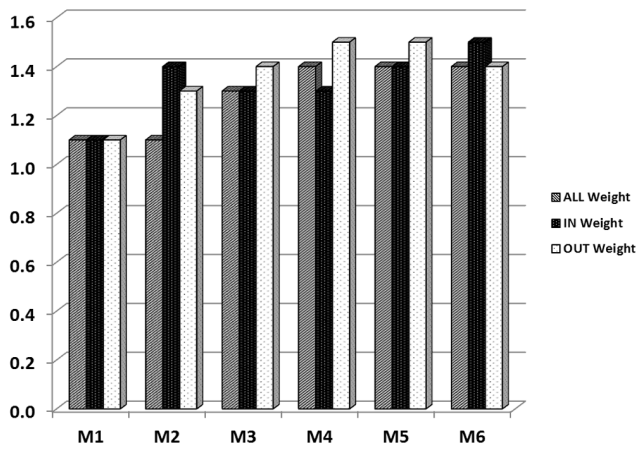**Fig. 10.** Exponent of connectivity distributions for networks M1–M6.

Fig. 11. Exponent of the weight distribution for networks M1–M6.

Fig. 12. Temporal variation of the connectivity distribution exponent (absolute value) in successive temporal windows of network M2.

Fig. 13. Temporal variation of the weight distribution exponent (absolute value) in successive temporal windows of network M2.

in network M6 have also been included in network E10. We are actually looking, in each of these cases, at the same statistical population of earthquakes. These are the earthquakes that, being close enough in space–time–magnitude, are most likely to be related to each other.

This is an interesting result, indicating that the method is reliable, robust with respect to variations in parameter values, and reflects fundamental properties of seismicity. Consequently, the process of identification of correlations between earthquakes can start with a certain choice of parameter values (class definition), and end when networks with scale free properties are found.

4.2 Evolution of network properties over time

Results of this method are further applied for the study of the way the relationships between earthquakes change over time. The network is split up into successive event windows, each window having the same number of successive events. The first objective of the analysis is to determine whether scaling
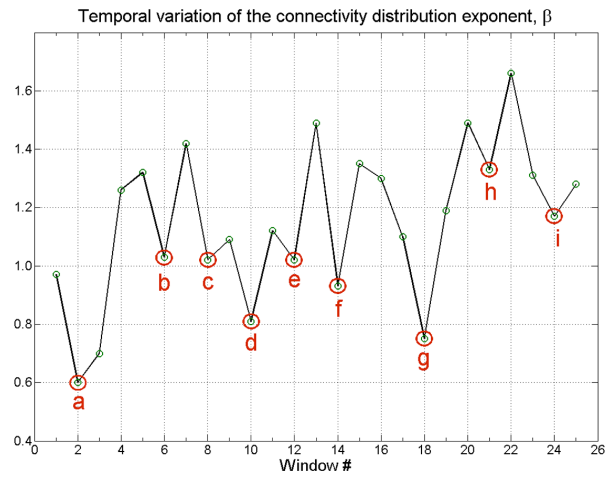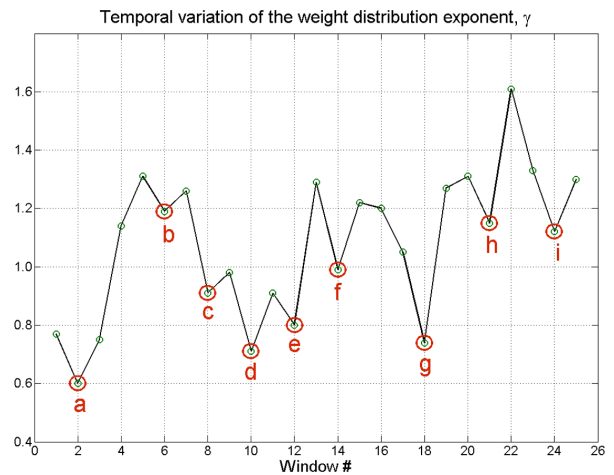
properties can be identified in the temporal windows. If that is the case, the next objective is to study whether changes in scaling properties in successive temporal windows can be related to real-life changes in the volcanic system.

Numerous networks from different classes have been studied. The networks chosen for the analysis were those with strong scaling properties. They were split into event windows of 1000 successive events and sub-networks of 1000 nodes were generated accordingly. In each case, the node connectivity distribution and the node weight distribution were assessed.

The results (not shown) confirm that also these distributions manifest power law characteristics; for each scaling regime, the corresponding exponents β and γ were calculated. As an example, the study of successive event windows in network M2 is presented in Fig. 12 (the variation of the

Table 4. Robustness of the method: regardless of various choices of parameter values shown under column "Class definition", the method identifies the same statistical population of events as being the earthquakes that are interconnected. In this example, the overlap between the nodes selected by networks D9, J6, M6, and E10 ranges between 91.34 and 100 %.

Network	Network characteristics	Comparative node selection	Class	Class definition	Class characteristics
D9	11 966 nodes $W \geq 7.5 \times 10^{-3}$	11 966 nodes in J6 – 100 % 11 966 nodes in M6 – 100 % 11 813 nodes in E10 – 98.72 %	D	$T_{\max} = 30$ days $D_{\max} = 30$ km $r = -1.35$ $p = -1$ $d_{\min} = 1$ km $t_{\min} = 1$ h	37 325 nodes 5 801 083 edges $H = 1.00$ $L = 3.39 \times 10^{-6}$
J6	13 523 nodes $W \geq 5 \times 10^{-2}$	13 154 nodes in M6 – 97.27 % 12 702 nodes in E10 – 93.93 %	J	$T_{\max} = 8$ days $D_{\max} = 10$ km $r = -1.35$ $p = -1$ $d_{\min} = 1$ km $t_{\min} = 1$ h	33 447 nodes 2 035 257 edges $H = 1.00$ $L = 5.60 \times 10^{-5}$
M6	14 006 nodes $W \geq 10^{-1}$	12 793 nodes in E10 – 91.34 %	M	$T_{\max} = 7$ days $D_{\max} = 10$ km $r = -1$ $p = -0.5$ $d_{\min} = 1$ km $t_{\min} = 1$ h	33 065 nodes 1 913 280 edges $H = 1.00$ $L = 1.86 \times 10^{-3}$
E10	14 091 nodes $W \geq 5 \times 10^{-4}$		E	$T_{\max} = 40$ days $D_{\max} = 50$ km $r = -1.35$ $p = -1$ $d_{\min} = 0.2$ km $t_{\min} = 3$ min	37 441 nodes 8 488 767 edges $H = 6.07$ $L = 7.24 \times 10^{-9}$

connectivity distribution exponent β) and Fig. 13 (the variation of the weight distribution γ). Although the graphs of the two distributions are not identical, they show the same trend in their evolution. In this example, the minimum values in the variation of β correspond to the windows where γ also has minimum values; therefore, the same lower-case letters from "a" to "i" were used on both graphs to tag the corresponding minimum values of the two exponents.

Figure 14 presents the graph of the cumulative number of earthquakes from January 1989 to December 2012. On this graph, the areas corresponding to the temporal windows tagged with letters from "a" to "i" in Figs. 13 and 14 were tagged with the same letters. Each of the labeled areas in Fig. 14 is therefore associated with minima in the absolute values of β and γ in successive temporal windows. A steep increase in the number of earthquakes can be noticed in Fig. 14 for the tagged areas. The history of the volcano shows that sudden events, with important discharges of energy, such as rapid openings of new fissures, violent massive eruptions or explosions, occurred in the volcano in the corresponding time intervals.

For example, minimum "a", corresponding to window number 2, can be related to the braking of the Kupaianaha tube system in 1989 that caused massive surface lava flows;
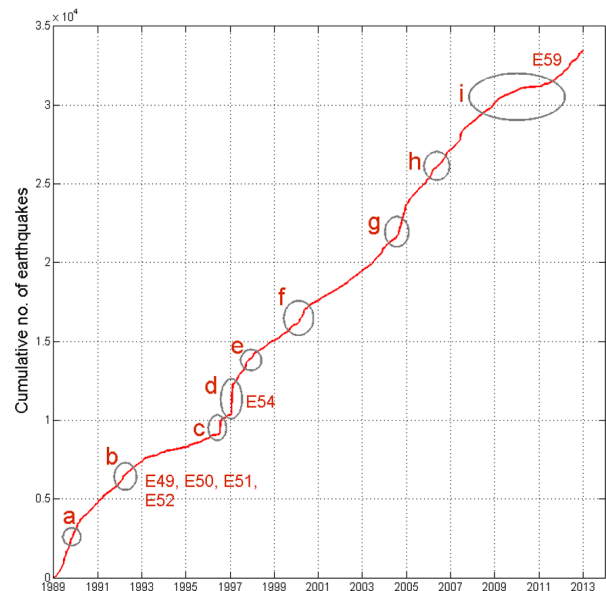

Fig. 14. Cumulative number of earthquakes between January 1989 and December 2012. The small letters tag the areas corresponding to the minimum absolute values of the exponents β and γ in successive temporal windows of network M2.

these surface lava flows spectacularly invaded new territory, overran the Wahaʻula Visitor Center and residences in Hawaii Volcanoes National Park. Minimum "b", corresponding to window number 6, can be related to the four episodes that occurred between the end of 1991 and the end of 1992: episodes 49, 50, 51, and 52. New fissures developed with these four episodes, and were accompanied by important seismic phenomena, as for example the 4.5 magnitude earthquake that preceded episode 52 from October 1992. As another example, minimum "d", corresponding to window number 10, can be related to episode 54 from 1997, when another new fissure developed, lava fountains reached tens of meters in height, and a period with 2000–4000 earthquakes per day followed.

Each of these minima can be related to such sudden events, with surges in the activity of the volcano, important discharges of energy, and changes in seismicity. However, the steep slope situated between "h" and "i" in Fig. 14, which is probably due to episode 58 from July 2007, cannot be related to a minimum in the values of β and γ. A possible explanation for this exception is the process of artificially breaking down network M2 into sub-networks with an equal number of nodes. A study on the optimization of the temporal windows selection should address this issue and is subject to further research.

The meaning of the minima in the exponents β and γ is an increased connectivity in the corresponding networks; the proportion of nodes that have high connectivity is larger. In the studied context, energy dissipates through various processes such as magma flows, lava effusion, explosions, heat emission, tectonic phenomena, degassing, etc. (Wright and Pilger, 2008). Although the increased connectivity in minima of the exponents β and γ is not consistently related to higher dissipation in tectonic energy, it could be related to peaks in the overall energy emitted by the volcanic system during eruptions or large outpourings of lava. An analysis of the clustering coefficient and of associations with energy dissipation in the system is the subject of future research.

Overall, the study shows that variations in the values of the exponents β and γ are able to reflect the way the relationships between earthquakes are changing over time. Minimum absolute values of β and γ in successive temporal windows can be related to important events in the life of the volcanic system and the associated seismicity.

5 Conclusions

A new type of directed network has been proposed for the assessment of relationships between earthquakes. The method was applied to volcanic seismicity in Hawaii. The nodes of the networks are epicenters of earthquakes; the edges that link the nodes carry space–time–magnitude weights, and have a direction given by the temporal succession of the events. The generality of the definition of the edge weight, W, as a combination of a factor in time, a factor in space,

and a factor in magnitude, is comprehensive and permits various combinations of space–time–magnitude correlations between earthquakes. Since any node can have any number of edges that enter the node and any number of edges that leave the node, any given event may have multiple predecessors, and any given event can contribute to multiple future events, as long as its edges carry enough weight.

Parameters and formulas used in the calculation of the weights take into consideration well-established properties of seismicity. High values of W are associated with strong relationships between earthquakes, while low values of W are associated with either weak relationships or no relationships at all. Various classes of networks can be generated based on distinct values of the parameters. Inside each class, different networks can be created by setting different thresholds for the minimum edge weight $W_{\min}$.

It is shown that networks that have $W_{\min}$ in the middle to upper range of the interval between the lowest edge weight value L and the highest edge weight value H in their class manifest significant scaling properties of node connectivity distributions, as opposed to networks with low values of $W_{\min}$, which exhibit poor or no scaling characteristics. Since high values of weight describe the strong links, the events selected in the networks with high values of $W_{\min}$ are primarily the earthquakes that are most likely to be related to each other. Therefore, it is reasonable to see a relationship between the fundamental characteristics of seismicity and the well-organized, scale free distributions of node connectivity. In networks with low values of $W_{\min}$, most of the nodes have little or no relationship with each other. In this context, the irregular and scattered shapes of their connectivity distributions are not a surprise.

It is also shown that the scale free behavior observed in networks with superior values of $W_{\min}$ is robust with respect to variations in parameter values. Tests performed on networks that manifest strong power law properties, but originating in different choices of parameter values, confirm the reliability of the method. They show that the same statistical population of earthquakes is chosen to participate in these networks, i.e., the earthquakes most likely to be interrelated. The results indicate that the method is reliable, robust with respect to variations of parameter values, and reflects fundamental properties of seismicity.

The threshold values $W_{\min}$ that identify networks of interrelated events are assessed in the context of all the earthquakes in the class: they are found as those values for which scale free properties appear and become stronger when subsequent networks are created using increasing values of $W_{\min}$. It can be said that the threshold $W_{\min}$ is a global parameter that characterizes the set of earthquakes and its values are meaningful only inside that set.

There are also other significant scaling properties that are detected in the analysis of the classes of networks. Node weight distributions also enjoy scaling properties. For each class, the dependency of the number of edges on the number

of nodes is a power law. The distribution of the distances between events exhibits distinct regimes with scale free properties. Similarly, the distribution of the time intervals between events is characterized by different domains with scaling properties.

It is also shown that the evolution of the relationships among earthquakes over time can be studied by splitting up the network into successive event windows with an equal number of nodes. The distributions of node connectivity and node weight in the emerging sub-networks manifest scaling properties that can be used to follow the change of seismicity over time. The exponents β and γ of these distributions have a similar evolution over the temporal windows. The increased connectivity in minima of β and γ can be associated with sudden, important discharges of energy in the life of the volcanic system and accompanying earthquakes. It is shown that the exponents of connectivity and weight distributions for successive event windows are able to reflect the way the relationships between earthquakes are changing over time. Aspects regarding the clustering coefficient, energy dissipation and optimization of the selection of the temporal windows are subject to further research.

Edited by: J. Kurths
Reviewed by: S. Lennartz-Sassinek and one anonymous referee

References

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

Baiesi, M. and Paczuski, M.: Scale-free networks for earthquakes and aftershocks, Phys. Rev. E, 69, 066106-1–8, 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.

Bak, P., Christensen, K., Danon, L., and Scanlon, T.: Unified scaling law for earthquakes, Phys. Rev. Lett., 88, 178501, doi:10.1103/PhysRevLett.88.178501, 2002.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U.: Complex networks: structure and dynamics, Phys. Rep., 424, 175–308, 2006.

Bunde, A. and Lennartz, S.: Long-term correlations in earth sciences, Acta Geophys., 60, 562–588, 2012.

Carbone, V., Sorriso-Valvo, L., Harabaglia, P., and Guerra, I.: Unified scaling law for waiting times between seismic events, EPL-Europhys. Lett., 71, 1036–1042, 2005.

Chastin, S. F. M. and Main, I. G.: Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions, Geophys. Res. Lett., 30, 1671, doi:10.1029/2003GL016900, 2003.

Davidsen, 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.

Felzer, K. R. and Brodsky, E. E.: Decay of aftershock density with distance indicates triggering by dynamic stress, Nature, 441, 735–738, 2006.

Gardner, J. K. and Knopoff, L.: Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, B. Seismol. Soc. Am., 64, 1363–1367, 1974.

Gutenberg, B. and Richter, C. F.: Seismicity of the Earth, Princeton University Press, Princeton, 1954.

Knopoff, L. and Gardner, J. K.: Higher seismic activity during local night on the raw worldwide earthquake catalogue, Geophys. J. Roy. Astr. S., 28, 311–313, 1972.

Lapenna, V., Macchiato, M., Piscitelli, S., and Telesca, L.: Scale invariance properties in seismicity of Southern Apennine Chain (Italy), Pure Appl. Geophys., 157, 589–602, 2000.

Lei, X. and Kusunose, K.: Fractal structure and characteristic scale in the distributions of earthquake epicentres, active faults and rivers in Japan, Geophys. J. Int., 139, 754–762, 1999.

Lennartz, S., Livina, V. N., Bunde, A., and Havlin S.: Long-term memory in earthquakes and the distribution of interoccurrence times, EPL-Europhys. Lett., 89, 69001, doi:10.1209/0295-5075/81/69001, 2008.

Lippiello, E., de Arcangelis, L., and Godano, C.: The role of static stress diffusion in the spatio-temporal organization of aftershocks, Phys. Rev. Lett., 103, 038501, doi:10.1103/PhysRevLett.103.038501, 2009.

Lippiello, E., Corral, A., Bottiglieri, M., Godano, C., and de Arcangelis, L.: Scaling behavior of the earthquake intertime distribution: Influence of large shocks and time scales in the Omori law, Phys. Rev. E, 86, 066119, doi:10.1103/PhysRevE.86.066119, 2012a.

Lippiello, E., Godano, C., and de Arcangelis, L.: The earthquake magnitude is influenced by previous seismicity, Geophys. Res. Lett., 39, L05309, doi:10.1029/2012GL051083, 2012b.

Nanjo, K. and Nagahama, H.: Spatial distribution of aftershocks and the fractal structure of active fault systems, Pure Appl. Geophys., 157, 575–588, 2000.

Omori, F.: On the aftershocks of earthquakes, J. College of Science, Imperial University of Tokyo, Japan, 7, 111–200, 1894.

Reasenberg, P.: Second-order moment of central California seismicity, 1969–1982, J. Geophys. Res., 90, 5479–5495, 1985.

Richards-Dinger, K., Stein, R., and Toda, S.: Decay of aftershock density with distance does not indicate triggering by dynamic stress, Nature, 467, 583–586, 2010.

Sanchez, L. and Shcherbakov, R.: Temporal scaling of volcanic eruptions, J. Volcanol. Geoth. Res., 247–248, 115–121, 2012.

Shcherbakov, R., Turcotte, D. L., and Rundle, J. B.: A generalized Omori's law for earthquake aftershock decay, Geophys. Res. Lett., 31, L11613, doi:10.1029/2004GL019808, 2004.

Shcherbakov, R., Turcotte, D. L., and Rundle, J. B.: Scaling properties of the Parkfield aftershock sequence, B. Seismol. Soc. Am., 96, S376–S384, doi:10.1785/0120050815, 2006.

Suteanu, C. and Suteanu, M.: Identification of change in spatio-temporal patterns: two multiscale approaches to volcanic seismicity, Romanian Geophysical Journal, 55, 7–16, 2011.

Turcotte, D.: Fractals and Chaos in Geology and Geophysics, Cambridge University Press, Cambridge, 1977.

Wright, R. and Pilger, E.: Radiant flux from Earth's subaerially erupting volcanoes, Int. J. Remote Sens., 29, 6443–6466, 2008.