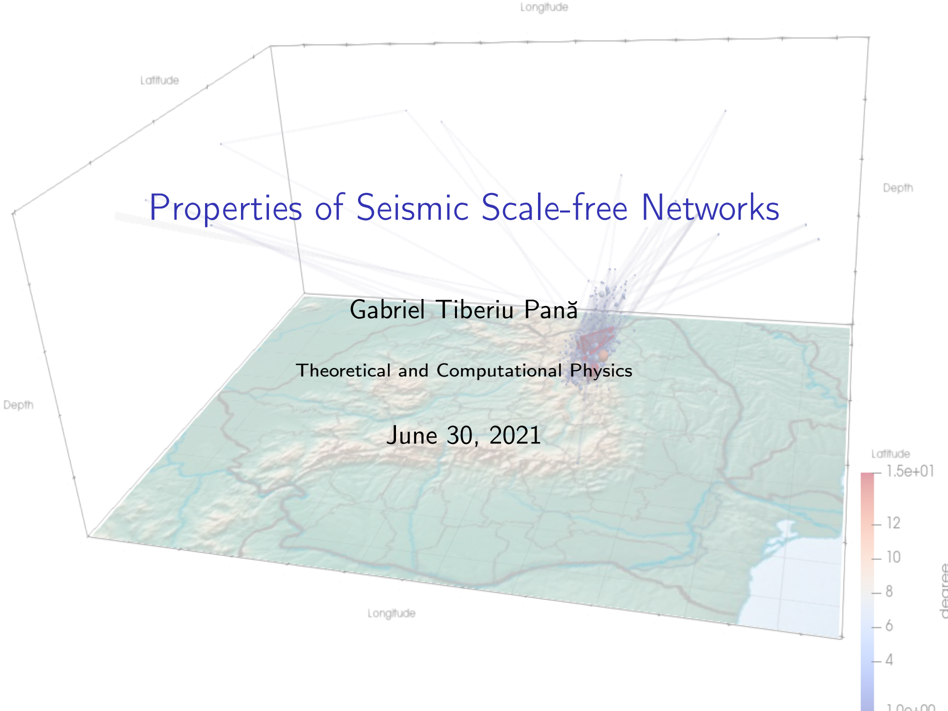
Properties of Seismic Scale-free Networks

Gabriel Tiberiu Pană

Theoretical and Computational Physics

June 30, 2021



Outline

Introduction

In this work we will familiarize ourselves with some complexity theory concepts such as complex systems, network theory, critical phenomena and self-organization to criticality.

Then we will analyze various seismic regions around the globe from the perspective of complex networks with the aid of the mathematical tools of graph theory. We expect that this would reinforce the idea that seismic zones behave as complex, self-organized critical systems.

Complex Systems

Complexity can be thought of as particular structures and patterns that cannot be easily described or predicted. A system becomes complex when *diverse* rule-following entities behave in an *interdependent* way. These entities interact over a *contact structure* or *network*. Characteristics:

- *adaptation* (eg. in a social system, individuals can learn, or in an ecosystem natural selection can take place)
- *robustness* (eg. they exhibit a certain behaviour at any spatio-temporal scale)
- *large events occurence* (eg. large earthquakes)
- *equillibrium states, fixed points, patterns or chaotic behaviour*

Stanley Milgram's Small Worlds

Figure: Stanley Milgram's Small Worlds Experiment: A letter randomly sent to a citizen in Nebraska starts on a 6 person's journey to it's target in Boston. Each person mailed the letter to an acquaintance that they thought would be closer to the target. The second to last person, mails the letter to the target because they knew him personally. Image from Elisa Baek et. al., Social Network Analysis for Social Neuroscientists"

Euler's bridges

Even earlier than Milgram's experiment, the problem that is thought to be the birth of graph theory is the Seven Bridges of Königsberg.

Problem

Is there a possible walk through the city that would cross each of the seven bridges once and only once?
"No" - Leonhard Euler, 1736

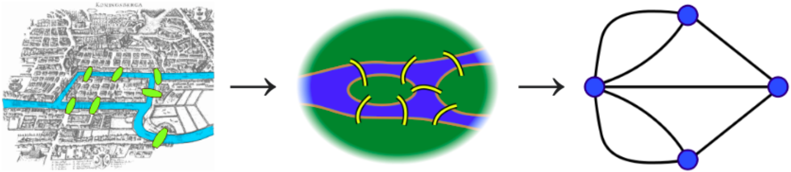
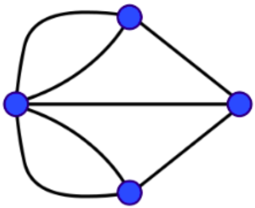


Figure: A depiction of what the problem looks like then and now. On the left, a map of Seventeenth-century Königsberg with the bridges in question highlighted. In the middle, a visualization of how Euler graphically represented the problem and on the right, how we represent the issue today, in modern graph theory. Image credit: Bogdan Giușcă, Wikipedia



Euler's Solution by walking in the graph, except for the start and finish nodes, one must enter a node as many times as he exists, so the nodes must be touched by even numbers of bridges. This made the connection between walks and node degrees in a graph, meaning that the necessary and sufficient condition for the desired walk is that the graph may have exactly zero or two nodes of odd degree.

Since in the problem, all land masses are connected by odd number of bridges, the proposed walk is impossible.

Graph Theory

Graph theory is the framework for the exact mathematical treatment of complex networks. An undirected (or directed) graph $G = (\mathcal{N}, \mathcal{L})$ consists of two sets $\mathcal{N}$ and $\mathcal{L}$:

- $\mathcal{N} \equiv \{n_1, n_2, ..., n_N\}$ = the nodes (or vertices, points) of G
- $\mathcal{L} \equiv \{l_1, l_2, ..., l_K\}$ = the links (or edges, lines) of G

$G(N, K) = (\mathcal{N}, \mathcal{L})$ represents a graph with N nodes and K edges.

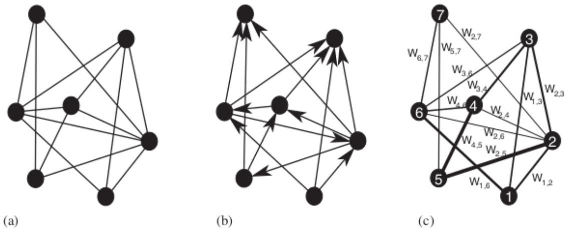


Figure: Graphical representation of a few types of graphs, each with $N = 7$ nodes and $K = 14$ edges: (a) undirected, (b) directed, the arrows showing the direction of each link and (c) weighted undirected, each link's weight $W_{i,j}$ reported on the respective line. Image credit: Vito Latore et. al., "Complex networks: Structure and dynamics"

Criticality and Self-organization

Critical phenomena refers to peculiar behaviour of a system when it is near or at the point of a continuous-phase transition, also called a *critical* point, which is a point at which the system changes from one state to another without a jump or discontinuity in it's properties such as internal energy, density or magnetization.

Self-organization is the process by which individual components of a system organize their communal behaviour to create global order by interactions amongst themselves rather than through external influence or instruction. Complex dynamic systems which have many and diverse elements interacting with each other, may display features of self-organization.
Fundamental characteristics of this organization are:

1. **decentralization** : the organization is distributed over all the components of the system
2. **robustness** : the system is able to survive or self-repair perturbations

"In chaos theory, self-organizations represent "islands" of predictability in a sea of chaotic unpredictability".

Self-Organized Criticality

In his 1996 paper "Simplest Possible Self-Organized Critical System", Flyjberg explains that a SOC system is a driven dissipative system consisting of:

1. a *medium* which has:
2. *disturbances* propagating through it, causing
3. a *modification* of the medium, such that eventually
4. the medium is in a *critical state* and
5. the medium is *modified no more*

Per Bak, Chao Tang and Kurt Wiesenfeld formulated the followinng key points when they defined Self-Organized Criticality (SOC):

- Spatial and temporal scaling must usually be unavoidably connected.
- There must be a robust, widespread spatio-temporal critical behaviour which arises from self-organization.
- Slow driven interaction and existence of a threshold.
- Dissipation has a role in maintaining a SOC state

Bak, Tang and Wiesenfeld's Sandpile Model

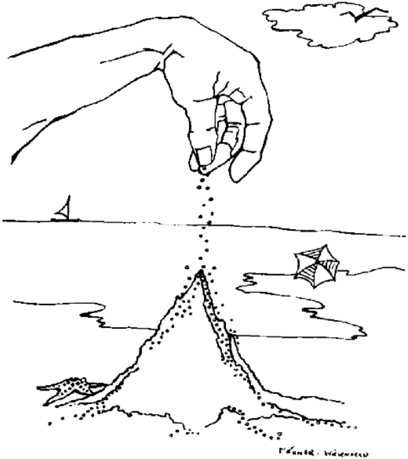


Figure: An illustration from Per Bak's "How Nature Works", a drawing by Elaine Wiesenfeld in which the dropping of grains of sand on a little pile on the beach is pictured.

Per Bak's Bureaucrats Model

In his famous book, How Nature Works, Per Bak describes a re-imagination of the sandpile model, called the bureaucrats model. The principle is similar and it is illustrated in the following picture:

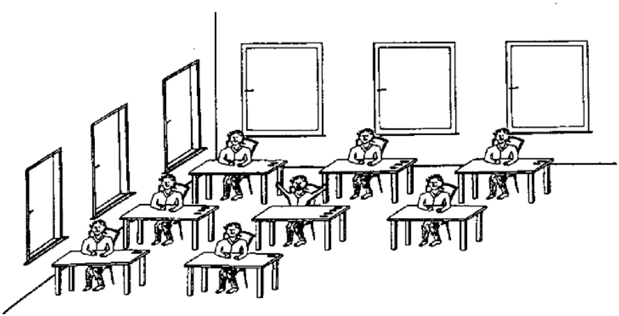


Figure: The "Office" version of the sandpile model. At each timestep a document is placed on the desk of a bureaucrat. When he finds four or more documents on his desk, he redistributes them, one to each of his neighbours, or he throws it out the window if he is at the edge of the room (the boundary conditions). Image credit: Per Bak, "How Nature Works"

Olami-Feder-Christensen Slider-Blocks Model

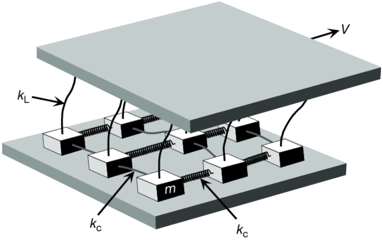


Figure: **Image credit: Robert Shcherbakov et. al., "Complexity and Earthquakes"**

The model is constructed as follows :

1. consider a 2D lattice of blocks;
2. each block is connected with springs to its four nearest nighbours and to a rigid moving plate;
3. each block interacts frictionally to a fixed plate that they move on;
4. blocks are driven by the continuous displacement of the moving plate;
5. when the force of one of the blocks reaches a threshold value F_{th}, the block slips, redefining the forces of it's neighbours;

The slip of one block may result in an avalanche of slips by the nearest blocks resulting in a chain reaction.

The model evolves as follows:

1. initialize all sites' forces to a random value between 0 and 1;
2. check if any $F_{ij} \geq F_{th}$;
3. if yes, redistribute the force according to:

$$F_{n,n} \rightarrow F_{n,n} + \alpha F_{i,j}, \tag{1}$$

$$F_{i,j} \rightarrow 0. \tag{2}$$

4. repeat from step 2 until the earthquake fully evolves;
5. locate the block with F_{max} and add $F_{th} - F_{max}$ to all blocks and return to the second step.

The probability distribution of the total number of relaxations of the slips (earthquakes) is measured. This quantitiy is proportional to the energy release during an earthquake. This cellular automaton model is found to show SOC behaviour for a large number of α values.

Slider-blocks model conclusions

OFC reached the following results with their model:

- displays robust SOC behaviour over a number of conservation levels;
- the amount of conservation impacts the obtained power-laws;
- as conservation increases, the behaviour transitions from localized to nonlocalized;
- this conservation dependence explains the variance of the parameter in the Gutenberg-Richter law.

Seismic Databases

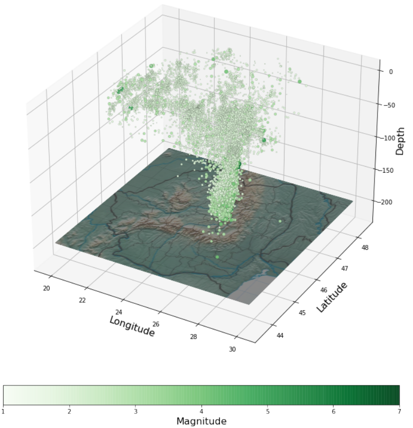
For each seismic region we wish to study, the first step is to collect the earthquakes databases available online, published by the respective region institute:

- **Vrancea(Romania)** - National Institute for Earth Phyiscs.
- **California(USA)** - Southern California Earthquake Center.
- **Italy** - National Institute for Geophysics and Vulcanology.
- **Japan** - Japan Meteorological Agency.

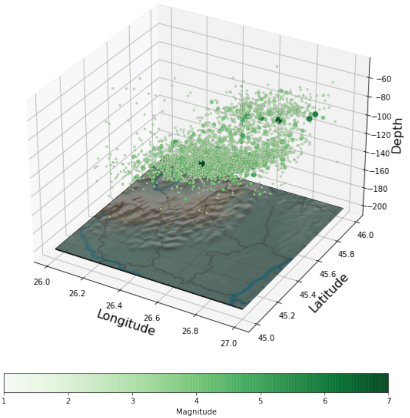
All data available in databases

Seismic Databases				
Seismic Zone	Timeframe	Latitude	Longitude	Depth
Romania	0984-01-01 2021-02-28	43.594°N 48.23°N	20.1°E 26.14°E	0 218.4
California(USA)	1932-01-02 2020-12-31	32°N 37°N	-114°W -122°W	0 51.1
Italy	1986-01-01 2020-12-31	30.61°N 47.998°N	-6.08°W 36.02°E	0 644.4
Japan	1919-01-11 2019-08-31	17.41°N 54.97°N	114.78°E 160.17°E	0 698.4

Romania and Vrancea



(a) Romania: 29186 earthquakes,
from 1976-02-03 13:29:16 to 2021-02-28 16:57:29.



(b) Vrancea: 7512 earthquakes,
from 1976-08-19 19:03:01 to 2021-02-28 00:11:55.

Figure: Earthquakes distribution for Romania (left) and Vrancea (right) seismic zones with magnitude > 1.

California(USA)

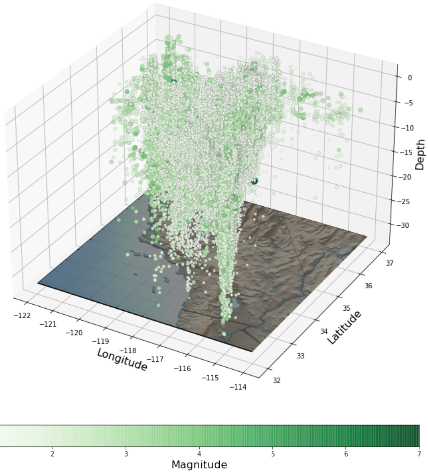


Figure: Earthquakes distribution for California(USA), seismic zone: 221113 events with magnitude > 1, from 1984-01-01 18:27:55 to 2020-12-31 23:04:53.

Italy

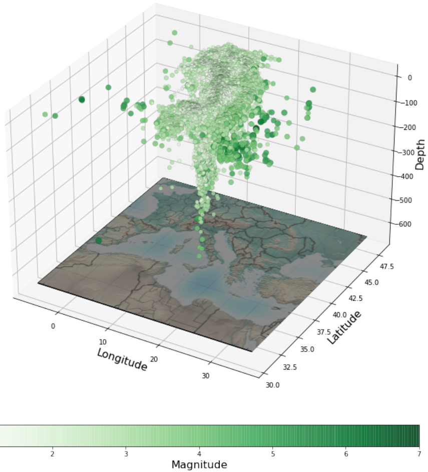


Figure: Earthquakes distribution for Italy seismic zone: 319567 events with magnitude > 1, from 1986-01-01 17:22:53 to 2020-12-31 23:41:18.

Japan

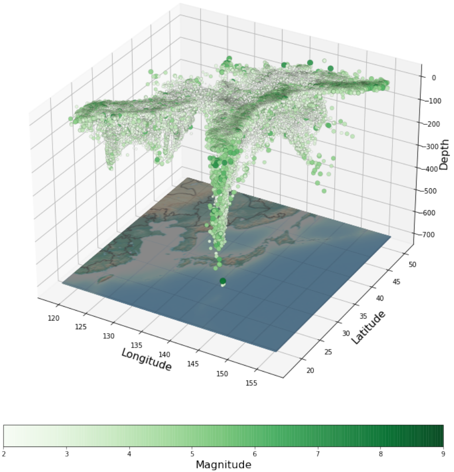


Figure: Earthquakes distribution for Japan seismic zone: 595713 events with magnitude > 2, from 1992-01-01 00:54:03 to 2019-08-31 23:54:24.

Earthquake Table

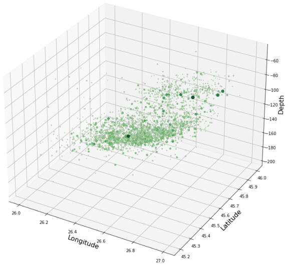
The fundamental tool we use in our analysis is the Earthquakes Table:

	date	latitude	longitude	depth	magnitude	energyRelease	x	y	z	cubeIndex	cubeLatitude	cubeLongitude	cubeDepth
0	1976-08-19 19:03:01	45.5400	26.3700	162.0	2.3	3.564511e+08	8	6	23	6434	45.5378	26.3481	162.5
1	1976-09-07 17:38:08	45.6200	26.5000	155.3	3.6	2.654606e+10	10	9	21	5914	45.6279	26.5380	152.5
2	1976-10-01 17:50:43	45.6800	26.4900	146.0	6.0	7.585776e+13	11	8	20	5609	45.6730	26.4747	147.5
3	1977-03-04 19:21:54	45.7700	26.7600	94.0	7.4	7.870458e+15	13	13	9	2533	45.7631	26.7911	92.5
4	1977-03-04 21:21:01	45.2200	26.6500	141.0	3.0	3.630781e+09	1	11	19	5365	45.2225	26.6646	142.5
...	...	...	...	...	...	...	...	...	...	...	...	...	...

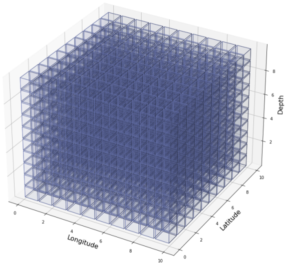
Figure: The Earthquakes Table - our fundamental tool for the following analysis, containing the basic information from the databases available online (*date, latitude, longitude, depth and magnitude*) and our computations: the *energyRelease* and the "cube parametrization" with indexing both in the "cubes space": x, y, z and the real space: *cubeLatidue*, *cubeLongitude*, *cubeDepth*.

The data in this example represents the first 5 events in the Vrancea seismic zone, starting with the year 1976, as a pandas DataFrame in Python.

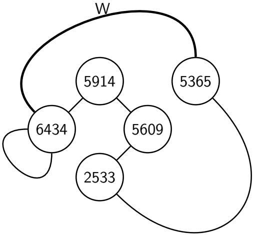
1. **energyRelease:** we can roughly estimate the energy release of each event by converting the moment magnitude M_W to energy using the equation $logE = 5.24 + 1.44M$ where M is the magnitude.
2. **Cube splitting:** in order to build our earthquakes network, we need to divide the spatial region selected into cubes and place each event in it's respective cube.



Cube split



Graph example of a Seismic Network



	date	latitude	longitude	depth	magnitude	cubeIndex
0	1976-08-19 19:03:01	45.5400	26.3700	162.0	2.3	6434
1	1976-09-07 17:38:08	45.6200	26.5000	155.3	3.6	5914
2	1976-10-01 17:50:43	45.6800	26.4900	146.0	6.0	5609
3	1977-03-04 19:21:54	45.7700	26.7600	94.0	7.4	2533
4	1977-03-04 21:21:01	45.2200	26.6500	141.0	3.0	5365
...	...	...	...	...	...	...

Figure: Seismic Network (Graph representation) example: Nodes are identified as *cubeIndex*, used to acces any information about events in that respective cube from our Earthquakes Table. Edge weight can be represented on the edge as W, where each additional link between two nodes increases this value by 1.

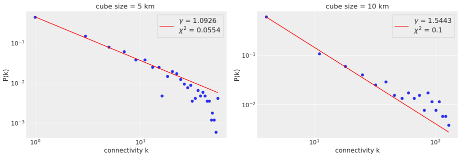
Connectivity of a Network

The fundamental measure to describe our seismic network is the connectivity distribution $P(k)$. This distribution comes naturally by realising the histogram of the nodes degree and then the regression of the distribution and show that it follows a power law:

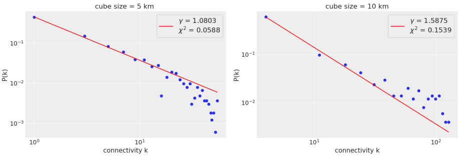
$$P(k) \sim k^{-\gamma} \tag{3}$$

The connectivity computations are made for various seismic networks (Vrancea(Romania), California(USA), Italy and Japan), with different magnitude restrictions, for 2 cube sizes: $5 \times 5 \times 5$ km and $10 \times 10 \times 10$ km, with and without edge weights.

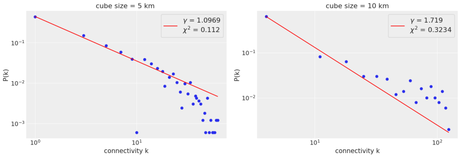
Vrancea Connectivity



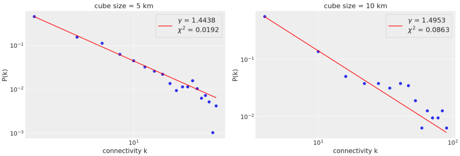
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



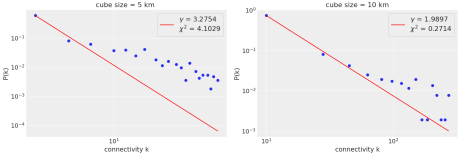
(c) $2 \leq magnitude \leq 4$



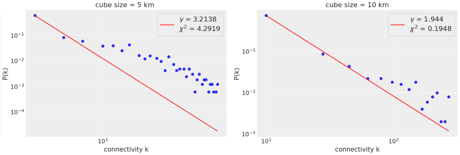
(d) $3 \leq magnitude \leq 7$

Figure: Connectivity distribution $P(k)$ for Vrancea in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 1.08 to 1.756.

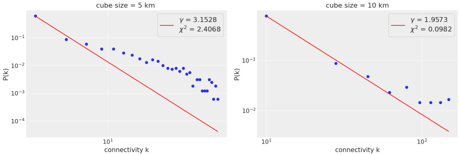
Vrancea Weighted Connectivity



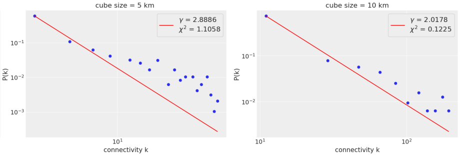
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



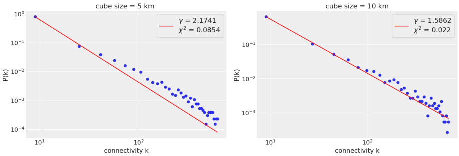
(c) $2 \leq magnitude \leq 4$



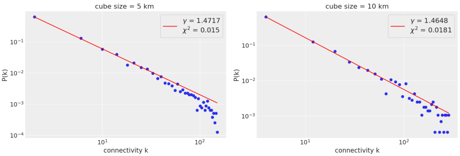
(d) $3 \leq magnitude \leq 7$

Figure: Weighted connectivity distribution $P(k)$ for Vrancea in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 1.95 to 3.26.

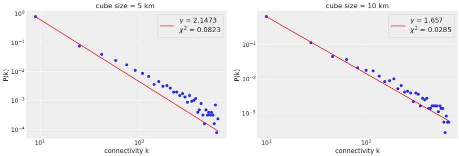
California Connectivity



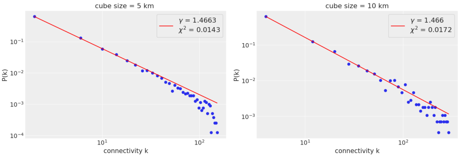
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



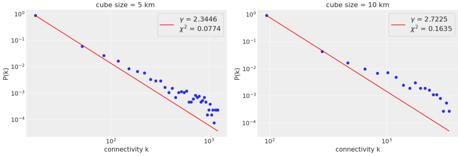
(c) $1 \leq magnitude \leq 3$



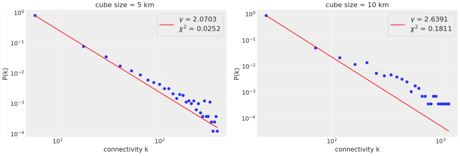
(d) $2 \leq magnitude \leq 4$

Figure: Connectivity distribution $P(k)$ for California in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 1.45 to 2.16.

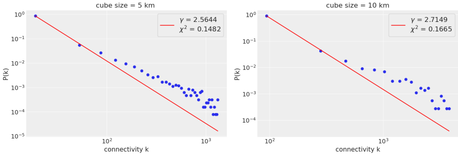
California - Connectivity Weighted



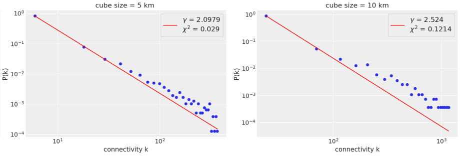
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



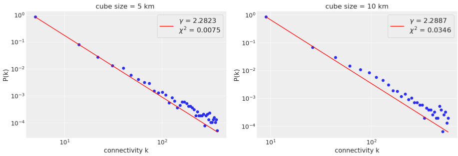
(c) $1 \leq magnitude \leq 3$



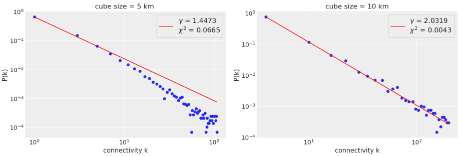
(d) $2 \leq magnitude \leq 4$

Figure: Weighted connectivity distribution $P(k)$ for California in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 2.06 to 2.71.

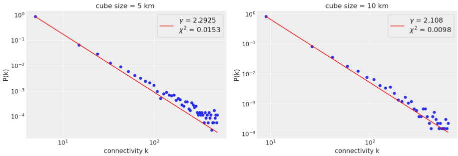
Italy - Connectivity



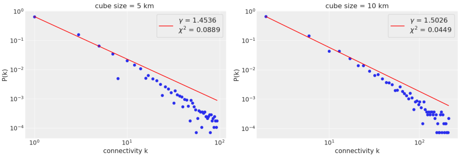
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



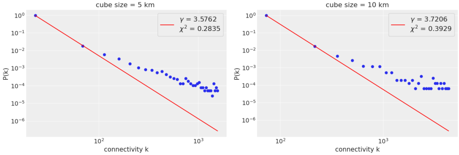
(c) $1 \leq magnitude \leq 3$



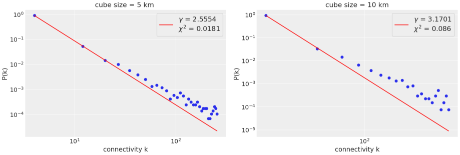
(d) $2 \leq magnitude \leq 4$

Figure: Connectivity distribution $P(k)$ for Italy in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 1.44 to 2.3.

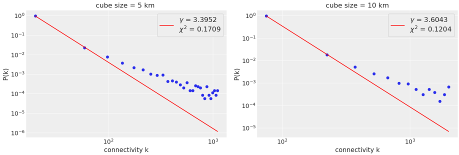
Italy - Connectivity Weighted



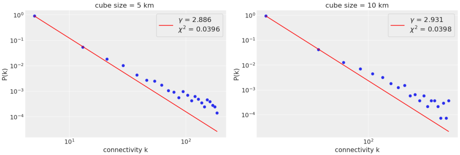
(a) $1 \leq magnitude$



(b) $1 \leq magnitude$



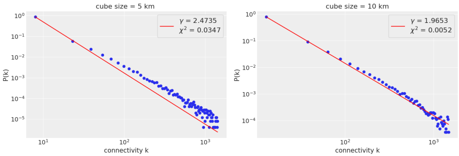
(c) $1 \leq magnitude \leq 3$



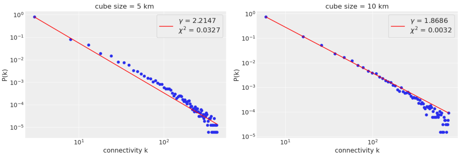
(d) $2 \leq magnitude \leq 4$

Figure: Weighted connectivity distribution $P(k)$ for Italy in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 2.8 to 3.72.

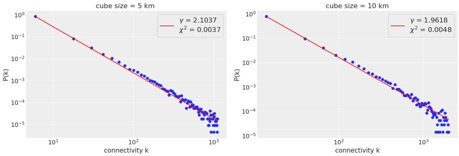
Japan - Connectivity



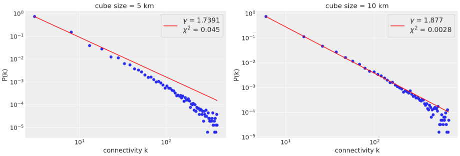
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



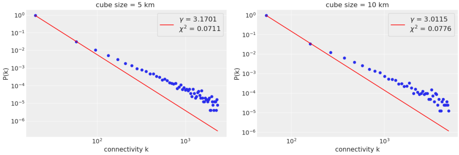
(c) $1 \leq magnitude \leq 3$



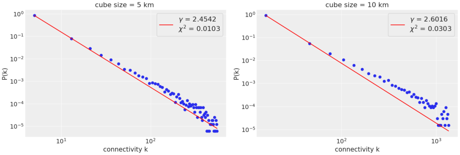
(d) $2 \leq magnitude \leq 4$

Figure: Connectivity distribution $P(k)$ for Japan in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 1.87 to 2.5.

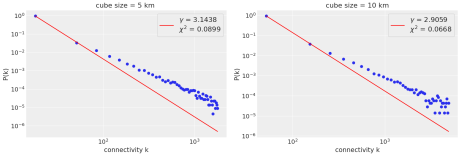
Japan - Connectivity Weighted



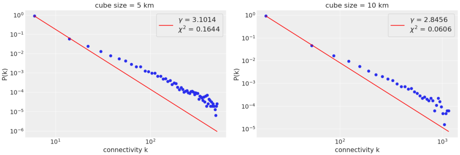
(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



(c) $1 \leq magnitude \leq 3$



(d) $2 \leq magnitude \leq 4$

Figure: Weighted connectivity distribution $P(k)$ for Japan in log-log and interpolation of the results, for a number of magnitude ranges. The exponent of the power law, γ ranges from ~ 2.45 to 3.17.

Motifs

Network motifs are sub-graphs that repeat themselves in a specific network or even among various networks. Each of these sub-graphs, defined by a particular pattern of interactions between vertices, may reflect a framework in which particular functions are achieved efficiently.

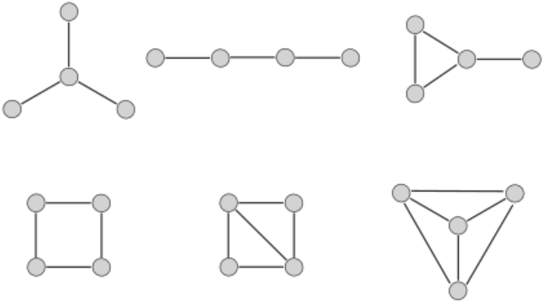


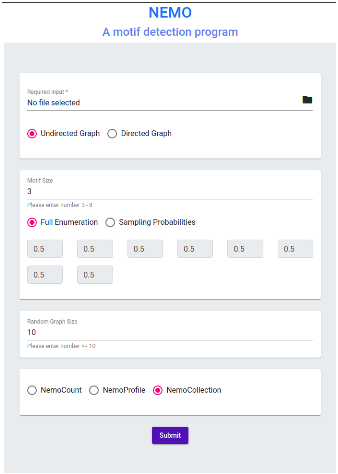
Figure: All 6 possible connected planar 4-node motifs. The most basic of them, the square, would outline a tetrahedron in 3D space

Given an undirected graph $G = (\mathcal{N}, \mathcal{L})$ and any possible small connected graph $F_{n,l}$ with n nodes and l links, we wish to find if F is a significant subgraph of G. The simplest approach to quantify the relevance of $F_{n,l}$ as a subgraph of G is based on the evaluation of the *Z-score*, defined as follows:

$$Z_F = \frac{n_f - \bar{n}_F^{rand}}{\sigma_{n_F}^{rand}} \tag{4}$$

where n_F is the number of times the subgraph $F_{n,l}$ appears in G and $\bar{n}_F^{rand}$ and $\sigma_{n_F}^{rand}$ are the mean and the standard deviation, respectively, of the number of occurences in an ensemble of graphs obtained by randomising G.

NemoSuite



NemoSuite (Network Motif Analysis in a Suite) is a web program developed and hosted online by researchers at University of Washington Bothell CSSE to detect and analyze network motifs.

A network motif is a frequent and unique subgraph pattern in an input network, and it is determined by Z-score being larger than 2.

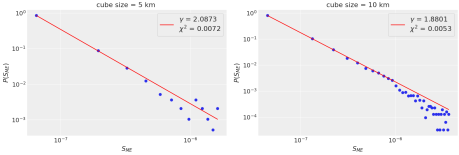
Triangle Surfaces

Our goal is to identify motifs in our networks and compute the distribution of their areas weighted by the total and mean energy that is released by earthquakes contained in them.

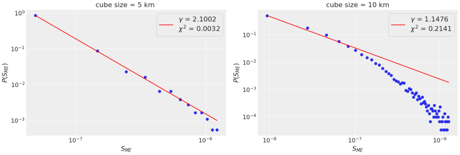
The 3 nodes motifs in 3D space (as in 2D) they outline *triangles*. Calculations proceed as follows:

- Use NemoSuite to extract all the triangles;
- Calculate mean energy and total energy in each motif;
- Calculate the total surface of each motif, using the coordinates of the nodes;
- Compute the distribution of surfaces weighted by mean/total energy;
- Compute the regression using a power-law and find *gamma* exponent.

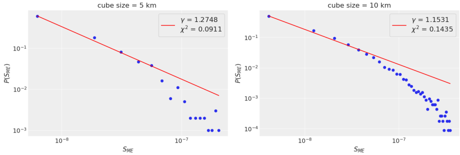
Vrancea Mean Energy Weighted Surfaces



(a) $1 \leq magnitude$



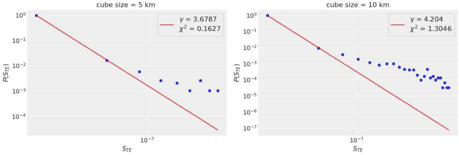
(b) $2 \leq magnitude$



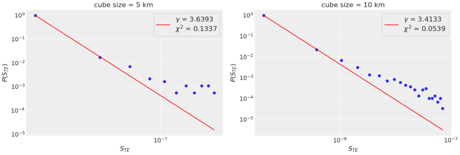
(c) $3 \leq magnitude$

Figure: $S_{ME} = Surface/Mean\ Energy$ distribution in log-log plots for triangle motifs in Vrancea for 3 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 1.14 to 2.1

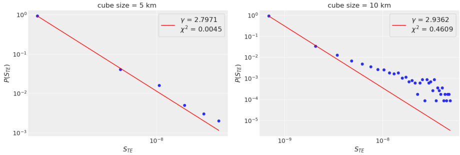
Vrancea - Total Energy Weighted Surfaces



(a) $1 \leq magnitude$



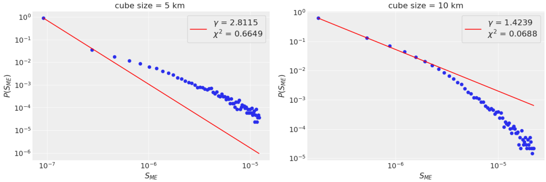
(b) $2 \leq magnitude$



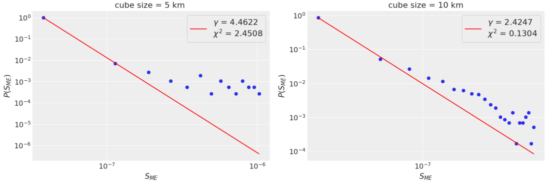
(c) $3 \leq magnitude$

Figure: $S_{TE} = Surface/Total\ Energy$ distribution in log-log plots for triangle motifs in Vrancea for 3 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 2.79 to 3.67

California - Mean Energy Weighted Surfaces



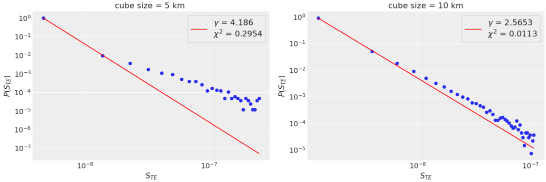
(a) $2 \leq magnitude$



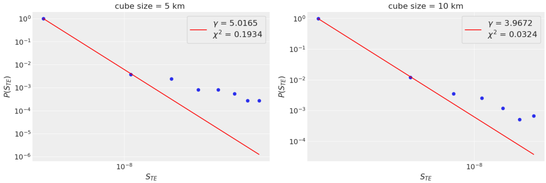
(b) $3 \leq magnitude$

Figure: $S_{ME} = Surface/Mean\ Energy$ distribution in log-log plots for triangle motifs in California for 2 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 1.42 to 4.46

California - Total Energy Weighted Surfaces



(a) $2 \leq magnitude$



(b) $3 \leq magnitude$

Figure: $S_{TE} = Surface/Total\ Energy$ distribution in log-log plots for triangle motifs in California for 2 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 2.56 to 5.01

Italy - Mean Energy Weighted Surfaces

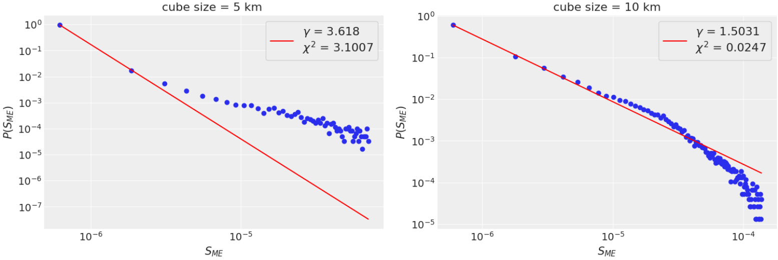


Figure: $S_{ME} = Surface/Mean\ Energy$ distribution in log-log plots for triangle motifs in Italy for $2 \leq magnitude$. The resulting interpolation shows that the distribution appears scale-free better at 10 km cube side granularization, with $\gamma = 1.503$

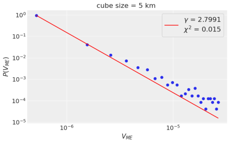
Tetrahedrons Volumes

Our goal is to identify motifs in our networks and compute the distribution of their volumes weighted by the total and mean energy that is released by earthquakes contained in them.

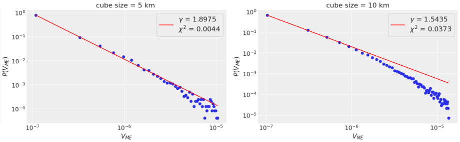
The 4 nodes motifs in 3D space they outline *tetrahedrons*. Calculations proceed as follows:

- Use NemoSuite to extract all the tetrahedrons;
- Calculate mean energy and total energy in each motif;
- Calculate the total volume of each motif, using the coordinates of the nodes;
- Compute the distribution of volumes weighted by mean/total energy;
- Compute the regression using a power-law and find *gamma* exponent.

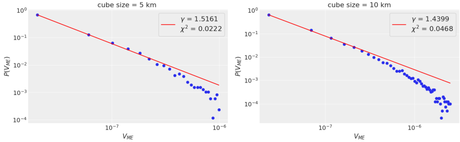
Vrancea Mean Energy Weighted Volumes



(a) $1 \leq magnitude$



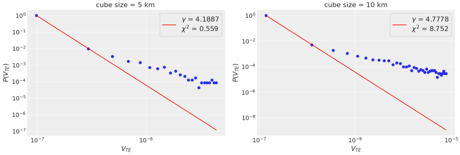
(b) $2 \leq magnitude$



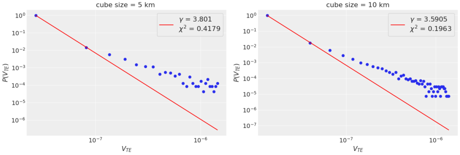
(c) $3 \leq magnitude$

Figure: $V_{ME} = Volume/Mean\ Energy$ distribution in log-log plots for tetrahedron motifs in Vrancea for 3 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 1.44 to 2.8

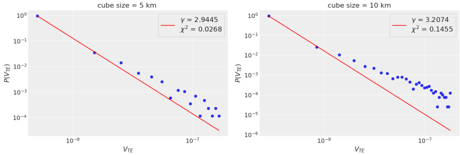
Vrancea Total Energy Weighted Volumes



(a) $1 \leq magnitude$



(b) $2 \leq magnitude$



(c) $3 \leq magnitude$

Figure: $V_{TE} = Volume/Total\ Energy$ distribution in log-log plots for tetrahedron motifs in Vrancea for 3 magnitude restrictions. The resulting interpolation shows that the distribution appears scale-free with γ ranging from ~ 2.94 to 4.77

California - Mean Energy Weighted Volumes

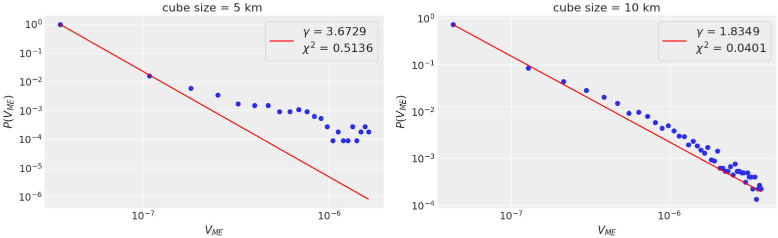


Figure: $V_{ME} = Volume/Mean\ Energy$ distribution in log-log plots for tetrahedron motifs in California for $3 \leq magnitude$. The resulting interpolation shows that the distribution appears scale-free better at 10 km cube side granularization, with $\gamma = 1.835$

Italy - Mean Energy Weighted Volumes

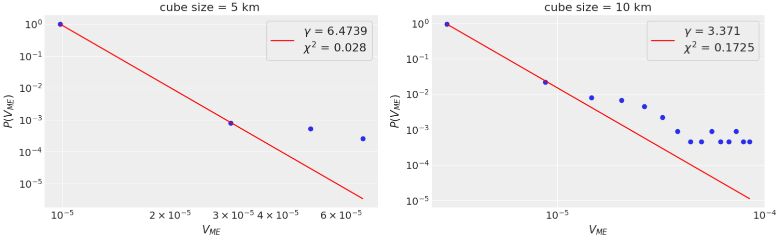


Figure: $V_{ME} = Volume/Mean\ Energy$ distribution in log-log plots for tetrahedron motifs in Italy for $3 \leq magnitude$. The resulting interpolation shows that the distribution appears scale-free better at 10 km cube side granularization, with $\gamma = 3.371$

Motifs Visualization in Paraview

As a visualization tool for the earthquakes networks and their motifs we used the application ParaView, which works together with the python library "VTK".

The Visualization ToolKit (VTK) is an open source software system for 3D graphics, visualization and image processing. VTK supports a variety of visualization algorithms : scalar, vector, tensor, texture and volumetric methods, making it easy to translate graphs into geometric form.

ParaView is an open-source application designed to visualize data of varying sizes, from small to very large. Under the hood, ParaView uses VTK as the data processing and rendering engine.

Romania - Real Network Visualization

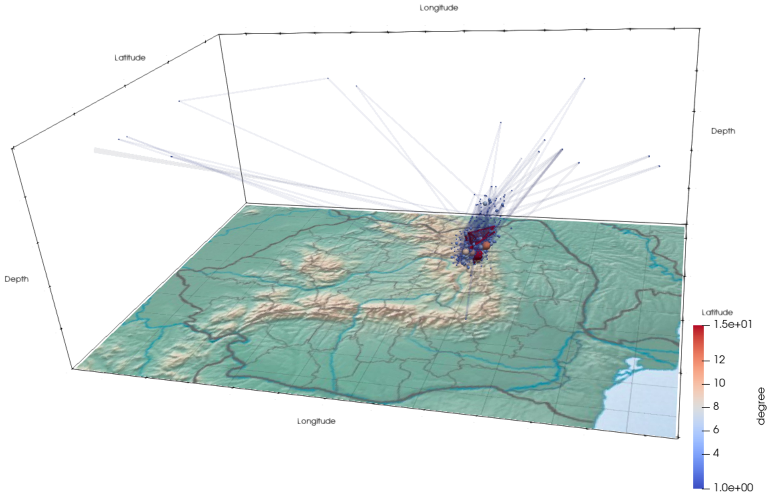
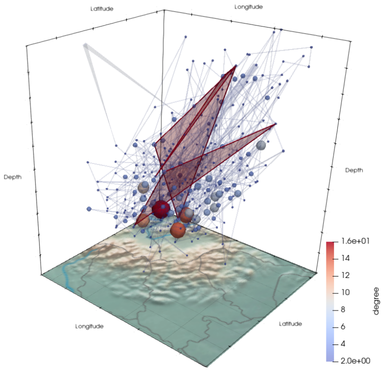
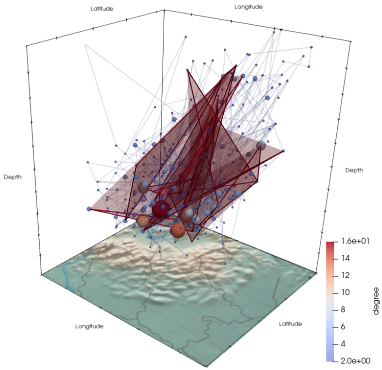


Figure: Motifs in Romania Seismic Network, earthquakes with magnitude > 4. Most of the network is concentrated in the Vrancea seismic zone.

Vrancea - Real Network Visualization



(a) Triangles in Vrancea Seismic Network



(b) Tetrahedrons in Vrancea Seismic Network

Figure: Motifs in Vrancea Seismic Network - 380 events stored in 255 nodes and connected through 375 edges. The degree of connectivity of the nodes ranges from 1 to 16. The motifs are represented as surfaces, for the triangles, or by volumes for the tetrahedrons (drawn in red)

California - Real Network Visualization

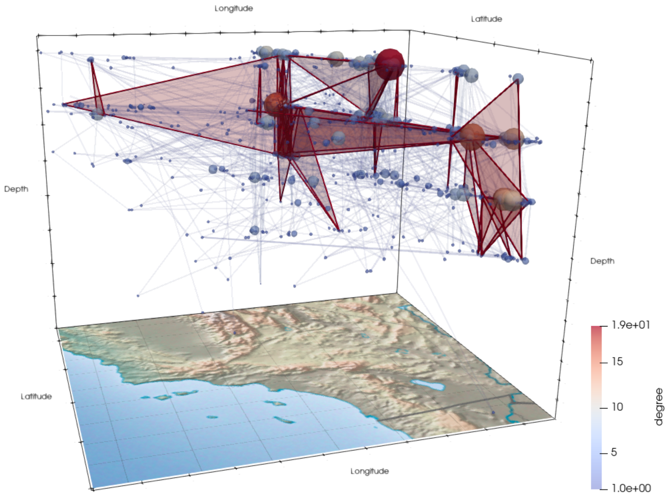


Figure: Triangle motifs in California Seismic Network, earthquakes with magnitude > 4 : 1119 events are stored in 718 nodes, which are connected by 1036 edges.

Italy - Real Network Visualization

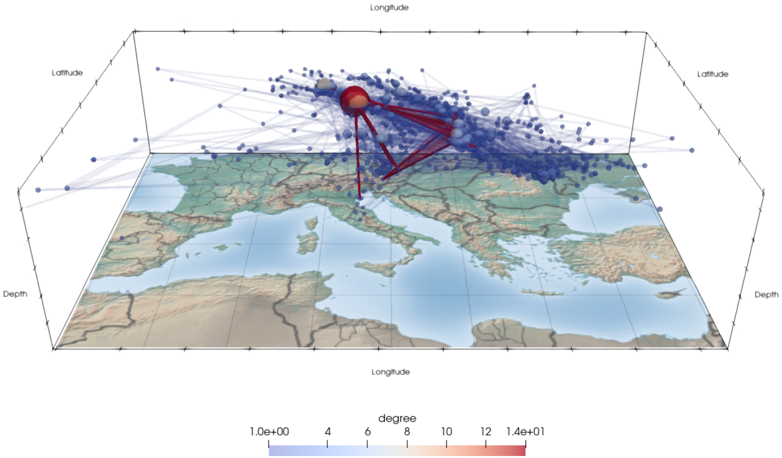


Figure: Tetrahedron motifs in Italy Seismic Network, earthquakes with magnitude > 4 : 1490 events are stored in 1325 nodes, which are connected by 1452 edges.

Japan - Real Network Visualization

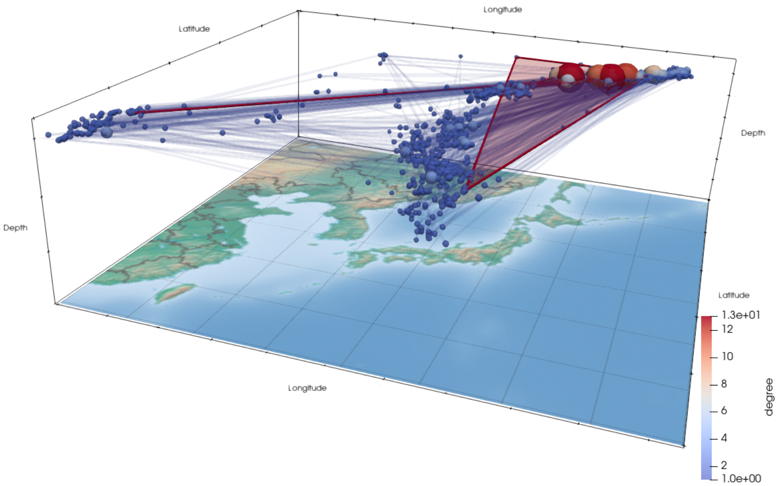


Figure: Triangle motifs in Japan Seismic Network, earthquakes with magnitude > 5 : 16948 events are stored in 14396 nodes, which are connected by 20332 edges.

Correlations and Autocorrelations

Correlations represent a measure of how one value or system responds to another. There are many different types of correlation functions which can be used to determine the correlation of two random variables or systems. For example, time correlation functions are used in the theory of noise and stochastic processes in statistical physics and spectroscopy.
Two types of correlation:

1. **cross-correlation** - when two **different** sequences are correlated (eg. comparing two different time series, allowing to see how two signals match)
2. **autocorrelation** - the correlation occurs between two of the **same** sequences (i.e you correlate the signal with itself).

Temporal Autocorrelations

In time series analysis, autocorrelation is used to correlate observations at a time step with observations at previous time steps, called *lags*.
In general, considering a time series $y_1, ..., y_n$, it's mean is:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \tag{5}$$

The autocovariance function at lag k, for $k \geq 0$, of the time series is defined by:

$$s_k = \frac{1}{n}\sum_{i=1}^{n-k}(y_i - \bar{y})(y_{i+k} - \bar{y}). \tag{6}$$

Finally, the autocorrelation function (ACF) at lag k of the time series is:

$$r_k = \frac{s_k}{s_0}. \tag{7}$$

For our networks, we need to establish a way of defining the lags, and the quantity that is to be autocorrelated. These calculations are made as such:

- Establish the time period of the network, from the first earthquake of the database to the last. For example the timeframe we have for our whole network in Vrancea(Romania) is from August 19 1976 to February 28 2021, so we would have a *timeframe* of 16263 days.
- Next we iterate through the table and find the biggest interval dt between two consecutive earthquakes.
- Then, by dividing the total *timeframe* by that interval dt we create a certain amount of time windows W, so we have split the total timeframe in equal parts and surely each part contains at least one earthquake.
- Calculate the total energy release for the earthquakes in each time window. Now we have E_i total energies (the values of our time series y_i) with $i = 1, 2, ...W$ (the lags).

Having defined our earthquake energies time series, we can now proceed to calculate the temporal autocorrelation function by adapting the formula described earlier:
Firstly, compute the mean energy, by summing over all time windows W:

$$\bar{E} = \frac{1}{W} \sum_{i=1}^{W} E_i. \tag{8}$$

Then you can compute the autocovariance function at lag k, for $k >= 0$, of the time series as:

$$s_k = \frac{1}{W - k} \sum_{i=1}^{W-k} (E_i - \bar{E})(E_{i+k} - \bar{E}). \tag{9}$$

Finally, the autocorrelation function is defined as:

$$C(k) = \frac{s_k}{s_0}. \tag{10}$$

TACF Vrancea(Romania) Seismic Network

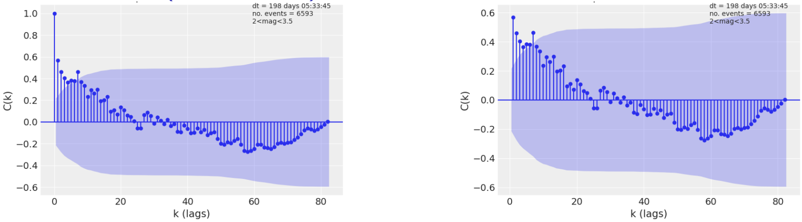


Figure: TACF for Vrancea(Romania) for mag restrictions: $2 < mag < 3.5$. A *timeframe* of 16263 days, containing a total of 6593 events, is split into $W = 83$ equal windows of $dt = 198$ days.

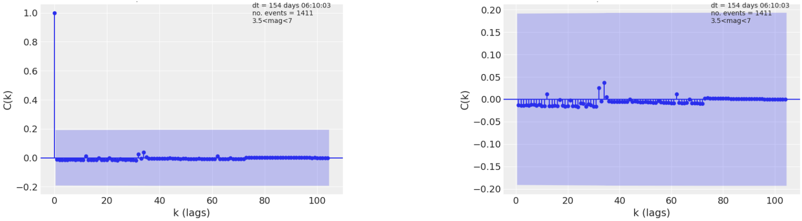


Figure: TACF for Vrancea(Romania) for mag restrictions: $3.5 < mag < 7$. A *timeframe* of 16263 days, containing a total of 1411 events, is split into $W = 105$ equal windows of $dt = 154$ days.

TACF California(USA) Seismic Network

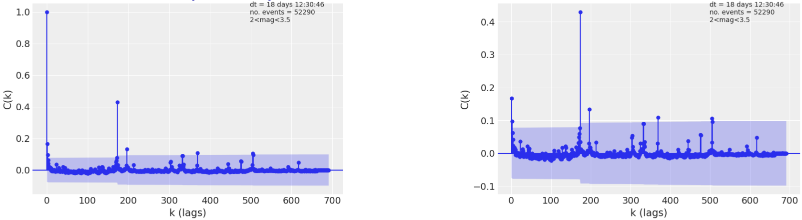


Figure: TACF for California(USA) for mag restrictions: $2 < mag < 3.5$. A *timeframe* of 13514 days, containing a total of 52290 events, is split into $W = 692$ equal windows of $dt = 18$ days.

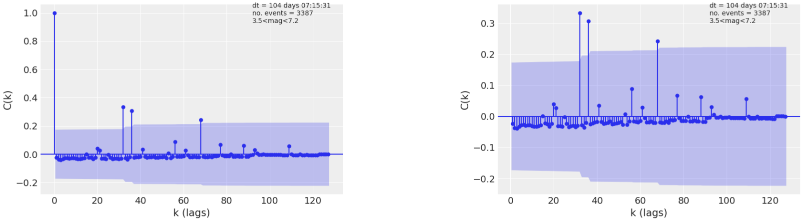


Figure: TACF for California(USA) for mag restrictions: $3.5 < mag < 7.2$. A *timeframe* of 13514 days, containing a total of 3387 events, is split into $W = 128$ equal windows of $dt = 104$ days.

TACF Italy Seismic Network

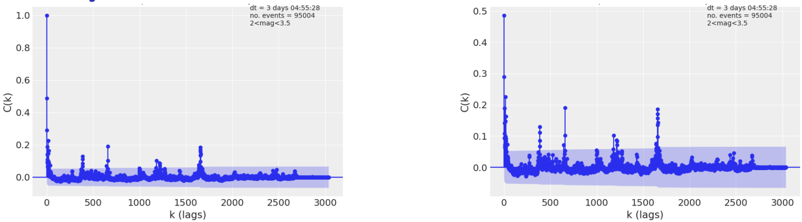


Figure: TACF for Italy for magn restrictions: $2 < mag < 3.5$. A *timeframe* of 12783 days, containing a total of 95004 events, is split into $W = 3040$ equal windows of $dt = 3$ days.

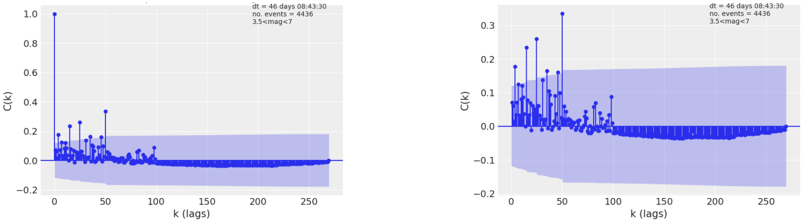


Figure: TACF for Italy for mag restrictions: $3.5 < mag < 7$. A *timeframe* of 12783 days, containing a total of 4436 events, is split into $W = 270$ equal windows of $dt = 46$ days.

TACF Japan Seismic Network

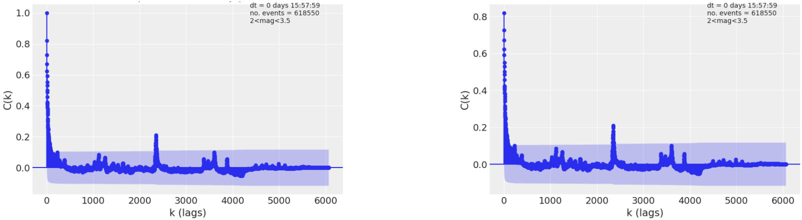


Figure: TACF for Japan for mag restrictions: $2 < mag < 3.5$. A *timeframe* of 10104 days, containing a total of 618550 events, is split into $W = 6068$ equal windows of $dt = 04:55:28$ hours.

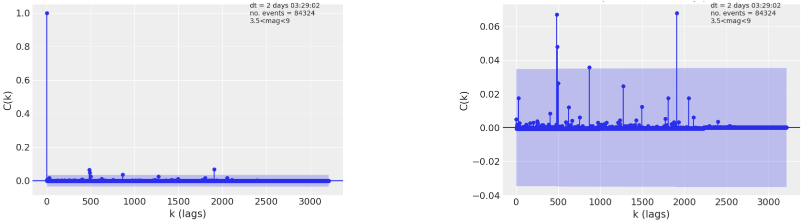


Figure: TACF for Japan for mag restrictions: $3.5 < mag < 9$. A *timeframe* of 10104 days, containing a total of 84324 events, is split into $W = 3213$ equal windows of $dt = 2$ days.

Conclusion

In conclusion, we have shown:

- How to access and extract information about earthquakes from the databases available.
- Used complex networks to construct seismic networks for regions around the globe.
- By measures of connectivity and motif discovery, we found scale-free behaviour for the node degrees and for surfaces or volumes of motifs weighted by their total and mean energy release of earthquakes in these motifs.
- By realizing these calculations for a number of magnitude ranges and cube granularization of the zone, we have shown that the scale-free networks are **very robust**.
- We presented the real networks with the visualization tool Paraview.
- We calculated temporal autocorrelations for different seismic zones.
- Every computation indicates, for each zone that the seismic phenomena may display a form of **self-organization to criticality**

Future work

In the future, with more **computational capability**, motifs can be identified for bigger and more complex networks, for example all the earthquakes in Japan.

Further study can be done in this space, with the possibility of analyzing **other seismic regions** by employing the same computational tools used in this report.

Also, from the perspective of network theory, other **centrality or community measures** may give insight about **clusters of earthquakes**, opening the possibility of splitting a region into more relevant sections.