Physica A 481 (2017) 224–234

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

# Network similarity and statistical analysis of earthquake seismic data

Krishanu Deyasi [a], Abhijit Chakraborty [b,*], Anirban Banerjee [a,c]

[a] Department of Mathematics and Statistics, Indian Institute of Science Education and Research Kolkata, Mohanpur-741246, India
[b] Graduate School of Simulation Studies, University of Hyogo, Kobe 650-0047, Japan
[c] Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Mohanpur-741246, India

## HIGHLIGHTS

- Complex network and statistical analysis of distinct earthquake catalogs.
- Peakedness of graph spectra of different earthquake regions varies significantly.
- Hierarchical clustering of earthquake networks is explained by tectonic plates.
- We have calculated the conditional probabilities for forthcoming earthquake events.

## ARTICLE INFO

*Article history:*
Received 25 November 2016
Received in revised form 15 March 2017
Available online 19 April 2017

*Keywords:*
Earthquake network
Seismic catalogs
Hierarchical clustering
Jensen–Shannon divergence
Graph spectra
Conditional probability

## ABSTRACT

We study the structural similarity of earthquake networks constructed from seismic catalogs of different geographical regions. A hierarchical clustering of underlying undirected earthquake networks is shown using Jensen–Shannon divergence in graph spectra. The directed nature of links indicates that each earthquake network is strongly connected, which motivates us to study the directed version statistically. Our statistical analysis of each earthquake region identifies the hub regions. We calculate the conditional probability of the forthcoming occurrences of earthquakes in each region. The conditional probability of each event has been compared with their stationary distribution.

## 1. Introduction

Earthquakes are one of the most devastating natural calamities that can shatter human civilization in large extent. Naturally, scientists are investigating this phenomena in great detail. The most robust empirically established facts in the phenomenology of earthquakes are the Gutenberg–Richter law [1] and the Omori law [2]. While the Gutenberg–Richter law expresses the relationship between frequency and magnitude of tremors in a region, the Omori law describes the temporal rate of decay of aftershocks. Different models have been proposed to study the phenomena of earthquakes. One well-known example is the Burridge–Knopoff model that demonstrates the statistical properties of earthquakes using friction on a fault surface as a stickslip process [3]. It is considered that large faults in the Earth's crust are formed due to the action of plate tectonic forces.

Plate tectonic theory, which is based on Alfred Wegener's continental drift theory [4], is considered to be the fundamental theoretical framework in the field of Earth Science and plays the pivotal role to explain tremors. This theory states that Earth's

---

[*] Corresponding author.
*E-mail address:* abhiphyiitg@gmail.com (A. Chakraborty).

http://dx.doi.org/10.1016/j.physa.2017.04.050

outer shell, the lithosphere is divided into several rigid pieces, called plates. There are mainly eight major plates: African, Antarctic, Eurasian, North American, South American, Pacific, and Indo-Australian. When a pair of plates move with respect to each other they do not deform internally rather they deform along their edges and create earthquakes and volcanoes along the edges of the plates.

Recently a complex network approach has been applied to observe the universal features of the earthquake phenomena from seismic catalogs. Based on the work of Bak et al. [5], Baiesi and Paczuski, using a correlation metric between a pair of events, have constructed a network in which tremors are the nodes and a pair of nodes are linked if the correlation between them is higher than a certain threshold [6,7]. It is shown that the earthquake network exhibits a scale-free nature with highly heterogeneous degree distribution characterized by a power law. In another study, Abe and Suzuki have constructed the network using a grid, covering the entire earthquake events over a region [8–15]. A cell of the grid is considered to be a node if at least one epicenter occurs within the cell, and the cell size is a tunable parameter for the model. A pair of nodes is connected by a link if two successive events occur on those two nodes. Using this network model, they found various robust features of earthquake networks. They have shown that earthquake networks exhibit a power-law degree distribution [8] small-world phenomena [9], assortativity [10] and scaling in local clustering [15]. Subsequently, different structural properties of weighted earthquake network have also been studied extensively in [16].

Although universal features of earthquake networks of different regions have been studied extensively [6,8–15], but no study has been devoted to capture the similarity and dissimilarity between earthquake networks of different regions. A study of hierarchical clustering [17] of earthquake networks will be very useful in this regard. Hierarchical clustering is usually represented by a dendrogram that pictorially shows the similar entities are clustered together in groups. Quantitatively, the similarity between networks could be measured based on different properties, viz., Euclidean distance, structural properties and dynamical behavior.

Primarily, earthquake networks are directed sequence of consecutive events, so analyzing the directed structure of earthquake networks can give more information than the underlying structure of earthquake networks [18]. Not only the directed structure of earthquake network, studying the directed sequence of earthquake events can also give us important insight, such as, prediction of consecutive earthquakes.

In this paper, we divide our study of earthquake data into two parts. In the first part, we follow the method of Abe and Suzuki to construct earthquake networks as it is connected to a universal law [19]. The similarity between each pair of spectral probability functions of eleven earthquake networks is measured using a probabilistic measure viz., Jensen–Shannon divergence. Using the similarity distances, the hierarchical clustering between earthquake networks is shown as a dendrogram. The hierarchical clustering of earthquake networks reveals the similar or dissimilar nature of different earthquake regions based on their positions on different tectonic plates. We also measure the frequency of earthquake events on a node and identify the earthquake prone locations for different regions. In the second part, we find the pair of regions where consecutive earthquake events have occurred with relatively higher frequency. For this, we calculate the conditional probability between two successive events. We also compare the conditional probability of two consecutive earthquakes with their stationary conditional probabilities to predict the occurrence of an earthquake event at a node consecutively after one occurs at a certain node.

## 2. Data

We have analyzed eleven distinct earthquake catalogs for different parts of the world, namely the Southern California Earthquake Data Center catalog (SC), Northern California Earthquake Catalog (NC), Japan University Network Earthquake catalog (JAP), Canada's National Earthquake Database catalog (CAN), International Institute of Earthquake Engineering and Seismology—Iran catalog (IRAN), Institute of Geodynamics—Greece catalog (GR), Center for Earthquake Research and Information—New Madrid catalog (NM) and British Geological Survey Earthquake Database around the British Isles catalog (BI), Geoscience Australia catalog(AUS), Swiss Seismological Service catalog (SZ) and GeoNet—the official source of geological hazard information for New Zealand catalog (NZ). Each catalog contains the geographical positions of the epicenters, specified by their latitudes ($\theta$) and longitudes ($\phi$) and the exact occurrence times of tremors. The positions of epicenters of the different earthquake regions are shown in Fig. 1. All the related parameters are mentioned in Table 1. The minimum and maximum values of the latitude–longitude coordinates, i.e., $(\theta_{\min}, \theta_{\max})$ and $(\phi_{\min}, \phi_{\max})$ characterize the extent of a earthquake region. The entire earthquake region is divided into a large number of square cells, following the approach in [8]. In this method, the cell size $L$ is the parameter of the model. We use the definition [14] of the dimensionless cell size parameter $\ell = L/(L_{lat}L_{lon})^{1/2}$. Here, the total extent along the north–south and the east–west directions of the entire earthquake region are $L_{lat}$ and $L_{lon}$, respectively. The North–South distance between $(\theta_i, \phi_i)$ and $(\theta_{\min}, \phi_{\min})$ is $d_{NS} = R(\theta_i - \theta_{\min})$ and the East–West distance is $d_{EW} = R(\phi_i - \phi_{\min}) \cos \theta_{av}$, where the radius of the earth is $R = 6370$ km and $\theta_{av}$ is $(\theta_{\min} + \theta_{\max})/2$.

## 3. Construction of the earthquake network

We use the method of Abe and Suzuki [8] to construct an earthquake network. Although the Abe–Suzuki method is not limited to two dimensional cells, but also to three dimensional cells. We consider a two dimension version of the model for our study. An entire earthquake region has been discretized into a two dimensional rectangular grid with the dimensionless