Scale Free Networks of Earthquakes and Aftershocks

Marco Baiesi∗,† and Maya Paczuski†

March 22, 2022

Abstract

We propose a new metric to quantify the correlation between any two earthquakes. The metric consists of a product involving the time interval and spatial distance between two events, as well as the magnitude of the first one. According to this metric, events typically are strongly correlated to only one or a few preceding ones. Thus a classification of events as foreshocks, main shocks or aftershocks emerges automatically without imposing predefined space-time windows. To construct a network, each earthquake receives an incoming link from its most correlated predecessor. The number of aftershocks for any event, identified by its outgoing links, is found to be scale free with exponent $\gamma = 2.0(1)$. The original Omori law with $p = 1$ emerges as a robust feature of seismicity, holding up to years even for aftershock sequences initiated by intermediate magnitude events. The measured fat-tailed distribution of distances between earthquakes and their aftershocks suggests that aftershock collection with fixed space windows is not appropriate.

1 Introduction

Earthquakes exhibit complex correlations in space, time, as well as magnitude [11, 12, 13, 14, 15, 16]. Sequences of earthquakes often appear related to main shocks of large magnitude, which are followed in time by nearby smaller events. Sometimes, the main shock is also preceded by a few intermediate or smaller precursor events. Earthquakes can also cluster as swarms, where the seismic activity is not distinctly associated with a main event. Human observation tends toward labeling these events depending on their relative magnitude and their position in the space-time sequence: foreshocks, main shocks and aftershocks, respectively. However, in identifying aftershocks, it is necessary to distinguish them from what is called background seismicity, and to identify their main shock. Although an observation by eye of the evolving seismic situation can support a classification, a precise label for each event may be intrinsically impossible.

In the most popular approach, aftershocks are collected by counting all events within a fixed space-time window [12, 13, 14, 15] following a pre-assigned main event (see Fig. 1). However, this method does not define the probability that an event thereby collected is actually correlated to the main event under consideration. Maybe more importantly, one does not know whether the predefined space-time windows are too large or too small for minimizing errors in the procedure. A more subtle issue is to define aftershocks of aftershocks. If an aftershock can have more than one preceding large event, which of these should be regarded as the most important or correlated one?

A quantitative metric of the correlation between any two earthquakes, or the extent to which one can be considered an aftershock of another, may be crucial for solving these problems, and for developing a better understanding of seismicity. A reliable metric should include known statistical properties. One is the Gutenberg-Richter (G-R) distribution [4] for the number of earthquakes of magnitude $m$ in a seismic region,

$$P(m) \sim 10^{-b m},$$  \hfill (1)

with $b$ usually $\approx 1$. Another is the fractal appearance of earthquake epicenters [11, 3, 14], with fractal dimension $d_f$. Thus, the average number of earthquakes of magnitude within an interval $\Delta m$ of $m$, occurring in an area of radius $r$ over a time interval $\tau$, is

$$\bar{n} = C \tau r^{d_f} \Delta m 10^{-b m},$$  \hfill (2)
where $C$ is a constant depending on the overall seismicity in the region under consideration.

For any earthquake $j$ in the seismic region, looking backward in time, how many earthquakes of magnitude within an interval $\Delta m$ of $m$ would be expected to have occurred within a time interval $t$, and within a distance $l$, of that specific event? In fact, an $n$ value can be defined between any two events $i$ and $j$ occurring in the sequence at times $T_i$ and $T_j$, with $T_i < T_j$. If we take the magnitude $m_i$ of the $i$-th event, the spatial distance $l = l_{ij}$ between the two earthquake epicenters, and the time interval $t = t_{ij} = T_j - T_i$, the expected number of events of magnitude within $\Delta m$ of $m_i$ occurring in the particular space-time domain bounded by events $j$ and $i$ is

$$n_{ij} = C t^{d_j} \Delta m 10^{-bm_i}.$$  

(3)

Note that the domain appearing in Eq. 3 is selected by the particular history of the seismic activity in the region and not predefined by any observer.

Of all the earthquakes preceding $j$, the most unlikely to occur according to Eq. 3 is earthquake $i^*$ such that $n_{ij}$ is minimized when $i = i^*$. However, earthquake $i^*$ actually occurred relative to $j$, even though it was the least likely to have done so. Therefore, $i^*$ must be the event to which earthquake $j$ is most correlated. In general, if $n_{ij}$ is extremely small, then the correlation between $j$ and $i$ is very strong, and vice versa. By this argument, the degree of correlation between any two earthquakes $i$ and $j$ is inversely proportional to $n_{ij}$. Since the space-time-magnitude scales are selected by the actual sequence of events, the variables $n_{ij}$ can be considered to be self-organizing tags of the underlying physical process governing seismicity. Note that singularities are eliminated by taking a small scale cutoff in time (here $t_{\text{min}} = 180$ sec) and a minimum spatial resolution (here $t_{\text{min}} = 100$ meters).

The metric defined by Eq. 3 allows various classifications of aftershocks. Therefore, the question of which is the better candidate to be the foreshock of an event can be quantitatively decided. Hierarchical clusters of earthquakes emerge, in which the biggest event in the cluster is called the main event, but where possibly later aftershocks create their own sequences of aftershocks, whenever they are able to “steal” aftershocks from the main event, and so on for further generations of aftershocks. Nevertheless, earthquakes are automatically collected into hierarchically self-organized clusters, without any special pre-analysis of single event properties.

In the language of modern complex network theory [12, 13], what we achieve is a time-oriented growing network where nodes (earthquakes) have internal variables (magnitude, occurrence time, and location), and links between the nodes carry a weight (the metric $n_{ij}$) and are directed according to the time orientation, from the older to the newer nodes. Empirically, we find that both the distribution of outgoing links and the cluster size distribution are scale free. Due to the continuous nature of the link variable, $n_{ij}$, no event is a priori purely an aftershock or a main shock. However, due to the broad distribution of $n_{ij}$ observed, main shocks and aftershocks emerge as extreme limits of a continuous spectrum of the extent to which any given event can be considered to be a precursor or aftershock of other events in the sequence.

Our approach was inspired by a recent analysis of earthquake waiting times by Bak et al. [6, 13]. They introduced a space-time-magnitude scaling variable that allows a data collapse of the distribution of waiting times between subsequent earthquakes larger than a specified magnitude, occurring within grid cells of a specified size, covering non-overlapping areas of the Earth. Also, Abe and Suzuki found scale free networks for earthquakes in a completely different context, where nodes representing these grid cells were linked when subsequent earthquakes occurred in them [13]. However, neither of these works quantified the correlation between an arbitrary pair of earthquakes, or dealt with the subject of aftershock identification.

2 Data and parameters

The catalog we have analyzed is maintained by the Southern California Earthquake Data Center (it can be downloaded from the SCEDC web site [http://www.scedc.eseec.org/ftp/catalogs/scen], for which $\Delta m = 0.1$. It is considered to be complete for events with $m > 2$. We use data ranging from January 1, 1984 to December 31, 2000. In order to work with a well defined ensemble, a lower threshold on the magnitude is introduced: events with magnitude smaller than $m_c$ are discarded. For each event, its position $i$ in the sequence is used as a label, and we record the magnitude $m_i$, the occurrence time $T_i$ (measured in seconds from midnight of the first day), and the latitude and longitude of the epicenter (converted to angles measured in radians, $\theta_i$ and $\phi_i$ respectively). The distance between two events $i$ and $j$ is then measured as the arc length on the Earth’s surface, $l_{ij} = R_0 \arccos[\sin(\theta_i)\sin(\theta_j) + \cos(\theta_i)\cos(\theta_j)\cos(\phi_i - \phi_j)]$, where the Earth radius is $R_0 = 6.3763 \times 10^6$ meters.

The $b$-value of the G-R law is $b \approx 0.95$ for this data set, while $d_{ij} \approx 1.6$ was found by Corral [14] using a box counting procedure. It is consistent with the correlation dimension we measure for most of our clusters. However, many of the statistical results we find are not sensitive to the precise value of $d_{ij}$. With these units and values, the constant $C$ can be estimated using Eq. 3. However, a precise evaluation of $C$ is not possible, because $n$ is the mean of a variable with huge variations in space and time. We have measured $n$ for several circular windows well inside the zone covered by the catalog, finding $C \leq 10^{-9}$. For simplicity, our choice in this paper is $C = 10^{-9}$. Most of our results are insensitive to the precise value of $C$ because we focus on
relative, rather than absolute correlations between a pair of events. Throughout this paper we use, unless otherwise stated, the above mentioned values, and a lower threshold \( m_c = 2.5 \).

To simplify notation, we denote the probability distribution of a generic quantity \( q \) as \( P(q) \). On finding distributions decaying as power laws, a clearer result appears by binning the values of \( P(q) \) in properly normalized bins of a width that grows geometrically with \( q \).

### 3 Method

In the simplest implementation of the network, each new earthquake \( j \) is attached with a single link to the previous earthquake in the sequence that minimizes \( n_{ij} \), with a weight denoted as \( n^*_j \). Hence, each link carries the extremal \( n^*_j \) for the added node \( j \) relative to all previous nodes, and globally one obtains a growing directed tree. Links with large \( n^*_j \) indicate a stronger correlation between the emitting node and the receiving one, and are expected to identify events normally classified as aftershocks. Weak links with large \( n^*_j \) arise when none of the previous events are sufficiently strong, and close in space and time to event \( j \). Thus, the strength of the link to an event \( j \) is inversely proportional to \( n^*_j \).

A natural decomposition of the network into clusters is achieved by then removing all weak links where \( n^*_j > n_c \), and \( n_c \) is a link threshold value. The correlated events are reliably detected when \( n_c \) is less than one but not extremely small. In the latter case, correlated events detach, and a very fragmented network appears. For large \( n_c \) some uncorrelated events make links, and a giant cluster appears. The resulting space-time windows are concave (see Figure 11 and Conclusions), at variance with the convex windows usually used.

### 4 Results

A part of the network constructed using this method is shown in Fig. 2. Hierarchically organized clusters of earthquakes emerge, where the links join aftershocks with their most correlated predecessor.

In order to quantitatively assess the properties of this network, we start by analyzing the distribution of link weights \( P(n^*) \). This distribution exhibits power law behavior with a slope \( \alpha \sim -1 \) up to a cutoff, as shown in Fig. 3. Thus, the distribution of link strengths, \( S = 1/n^* \), is also a power law, \( P(S) \sim 1/S \). Such a broad, continuous distribution, without particular characteristic peaks, indicates that a division of earthquakes into rigid classes is intrinsically impossible. Instead, a continuum of possibilities ranges from clear aftershocks, which have an incoming link with small \( n^* \), to events that are independent, with an incoming link of large \( n^* \), but may emit many outgoing links with small \( n^* \), and would be called main shocks.

The resulting network of earthquakes is scale free. The number of aftershocks of an earthquake is equal to the number, \( k \), of outgoing links from the node representing that event. In the language of network theory, this is called the out-degree of the node. Fig. 4 shows that earthquakes in Southern California form a scale free network, with an out-degree distribution scaling over more than three decades, with an index \( \gamma = 2.0(1) \).

Recently, many scale free networks with \( P(k) \sim k^{-\gamma} \), have been discovered in a broad variety of contexts. These include the Internet, the world-wide web, protein interaction and genetic regulatory networks, and the solar coronal magnetic field. Both the coronal field and the earthquake network share the property that the strength of a link between a pair of nodes is also scale free.

Lowering the link threshold \( n_c \) from infinity, the fully connected network breaks into clusters, in a percolation-like transition from a giant cluster to a finite cluster regime. We estimate the transition to take place between link thresholds \( n_c = 10^{-1} \) and \( n_c = 10^{-2} \). This estimate is obtained by examining the distribution of cluster sizes, \( N \), which is the total number of earthquakes in a cluster, as a function of \( n_c \) (see Fig. 5). Near the transition, the cluster size distribution also appears to be scale

![Figure 2: Scale free earthquake network around Landers epicenter (red cluster) and Hector Mine epicenter (blue cluster). Colors fade with the aftershock generation, from darker to lighter within each cluster. Note that the big event following the Landers earthquake, giving rise to its own (orange) cluster of aftershocks at \((-116.50', 34.10')\), is not a first generation aftershock, since it has no link from Landers. Here \( m_\leq = 4 \), and \( n_c = 10^{-2} \).](image-url)

...
free, \( P(N) \sim N^{-1.7(1)} \). Furthermore, a scaling regime exists for a wide range of link thresholds, indicating a relative insensitivity to a sharp separation between what are considered to be correlated and uncorrelated events. For clarity, we use the value \( n_c^* = 10^{-2} \) to locate the transition point where the giant cluster emerges. This value is consistent with our ansatz, Eq. 3, which requires that correlated events have \( n \) values significantly less than one. Network clusters constructed with \( n_c^* \) therefore only link strongly correlated events.

In Fig. 5 we study the effect of changing the temporal span of the catalog on the distribution of link weights. The power law behavior for strong links is stable, certainly up to \( n_c^* \). However, the cutoff in \( P(n^*) \) for weak links decreases to smaller \( n^* \) values, when earthquakes can link to events at further distance in the past. For an ideal “infinite” catalog, we conjecture that the cutoff value cannot be less than \( n_c^* \). Indeed, below the transition point a finite fraction of events stop having any correlated predecessors.

We define the link length, \( l \), as the distance between the epicenter of an aftershock and its linked predecessor. The distribution of link lengths depends on the magnitude \( m \) of the predecessor, being on average greater for larger \( m \). Dividing the link length distribution into classes depending on the magnitude of the predecessor, \( P_m(l) \), a maximum in the distribution occurs, which shifts to larger \( l \) on increasing \( m \), as shown in Fig. 5. This behavior is consistent with using larger space-time windows to collect aftershocks from larger events.

However, the distribution of link lengths exhibits no cutoff at large distances, but rather decays slowly as a power law with \( l \), up to the linear extent of the seismic region covered by the catalog. The different distributions are consistent with a scaling ansatz:

\[
P_m(l) \simeq 10^{-\sigma m} F(l/10^m)
\]

where \( l \) is measured in kilometers, \( \sigma \approx 0.4 \), and \( F(x) \) is a scaling function. The tail of the scaling function is a power law; i.e. \( F(x) \sim x^{-\lambda} \) with \( \lambda \approx 2 \) for \( x \gg 1 \). A data collapse using this ansatz is shown in the inset to Fig. 6. Such a slow decay at large distances calls into question...
the use of sharply defined space windows for collecting aftershocks, as already pointed out by Ogata [21].

Figure 7 shows the rate of aftershocks for the Landers, Hector Mine, and Northridge events. Aftershocks occurring at time $t$ after one of these events are binned into geometrically increasing time intervals. The number of aftershocks in each bin is then divided by the temporal width of the bin to obtain a rate of earthquakes per second. The same procedure is applied to each remaining event, not aftershocks of these three. An average is made for the rate of aftershocks linked to events having a magnitude within an interval $\Delta m$ of $m$. Figure 7 also shows the averaged results for $m = 3$ (1710 events), $m = 4$ (161 events), $m = 5$ (28 events) and $m = 5.9$ (4 events).

The collection of aftershocks linked to earthquakes of all magnitudes is one of the main results of our method. Even intermediate magnitude events can have aftershocks that persist up to years. Earthquakes of all magnitudes have aftershocks which decay according to the Omori law [5, 22],

$$\nu(t) \sim \frac{K}{c + t}, \quad \text{for } t < t_{\text{cutoff}}$$

(5)

where $c$ and $K$ are constant in time, but depend on the magnitude $m$ [22] of the earthquake. We find that the Omori law persists up to a time $t_{\text{cutoff}}$ that also depends on $m$ as well as the link threshold, $n_c$. Estimates of the cutoff times for $n_c^*$ are $t_{\text{cutoff}} \approx 3$ months for $m = 3$, and $t_{\text{cutoff}} \approx 1$ year for $m = 4$. For larger magnitudes, it is difficult to distinguish $t_{\text{cutoff}}$ from the temporal duration of the data set.

The Omori law for aftershocks emerges as a result of our analysis, although it is not part of the original ansatz, Eq. 8 used to define aftershocks. It has been extensively investigated over decades, together with its modified version [22] involving a scaling $\sim t^{-p}$. The data shown in Figure 7 is consistent with the original Omori result, $p = 1$, for aftershocks of earthquakes of all magnitudes, once second and further generations of aftershocks are excluded. Our result is also consistent with theoretical studies on stick-slip motion [21, 28], which suggest $p \approx 1$.

5 Discussion

Convex space time windows have been used since the 1970’s [7, 8, 9, 10], often with the size of the window determined by the main shock magnitude. The performance of this procedure is satisfactory for large earthquakes, although fixed window sizes may omit relevant aftershocks. Nevertheless, as a shortcoming, it can lead to distortions if many large aftershocks occur. In this case, nothing can be said on the “ownership” of further aftershocks.

Different approaches to the problem of aftershocks collection were proposed by several authors, sometimes with the aim to cure the former shortcomings. For a review see [20]. Our method has some similarities with these approaches. For instance, Frohlich and Davis collected earthquakes in clusters [24] by means of a different linking procedure. However, their analysis was done using a metric of the form $\sim \sqrt{l^2 + \text{const} t^2}$, which does not take into
account the magnitude of events, and has a space-time form at variance with measured earthquake correlations.

Maximum likelihood methods \[23, 24\], in the context of seismicity, usually start with an ansatz on the law governing aftershocks, typically the modified Omori law. It is further assumed that seismicity is a non-stationary Poisson process. Using a likelihood analysis with space, time and magnitude, Ogata compared several forms of aftershocks distance distributions \[24\], and showed that an aftershock rate of the form

\[
\nu_{m,t}(t) \sim \frac{10^{am}}{(c_l(m) + l)^\alpha(c_t + t)^\beta},
\]

was the most appropriate among his choices \((c_l, \alpha, \beta \text{ and } \mu)\) are constant, while \(c_l(m)\) is scaling with the magnitude \(m\) of the main shock). Hence, he also concluded that fixed space windows were not the best choice. Indeed, our metric variable \(n\) in Eq. \ref{eq:metric} somewhat resembles his form of \(\nu\).

However, our method is simpler to implement than likelihood methods. Furthermore, it does not require an ansatz on the validation of the modified Omori law, or on the type of statistical process that describes seismicity. Instead, the original Omori law is found as a result of our analysis. In addition, the physical argument leading to the variable \(n_{ij}\) also fixes the parameters in its definition, without the need to evaluate them by maximizing a likelihood.

One could object that the values of \(b\) and/or \(d_f\) can depend on the region of the Earth being considered, or may fluctuate depending on the specific fault zone being studied. However, the statistical results we find, as shown in Figs. 3-7, are remarkably robust to variations in either of these parameters. Varying \(d_f\) over a wide range, from 1 to 3 (using \(d_f > 2\) requires the introduction of event depths, see below) does not alter considerably the distribution of outgoing links, which retains its power law behavior with index \(\gamma \approx 2\). The distribution of link weights, \(n^*\), is even more insensitive to variations of \(b\) and \(d_f\). Also the Omori law with \(p \approx 1\), shown in Fig. 7, does not depend sensibly on the parameters, and holds for aftershocks linked to earthquakes of all magnitudes.

The crust of the Earth has a finite width \((\approx 20 \text{ km in California})\) in which events take place according to a “three-dimensional” fractal distribution, involving their depth. It is believed that there is a qualitative difference between small earthquakes and large ones, the former producing ruptures smaller that the crust width \[2\]. Hence, our arguments may need to be corrected at distances of the order of tens of kilometers. We have computed spatial distances through the three-dimensional Euclidean metric distance, using an appropriately revised \(d_f\) in Eq. \[3\]. No significant departures from the results leading to our present conclusions were found.

The introduction of more than one correlated predecessor for an event will be the subject of a future investigation. This is accomplished by attaching links between all earthquake pairs where \(n_{ij} < n_c\). In this case, a general network, which is not tree-like emerges. The clustering of earthquakes could then be quantified in terms of the clustering coefficient of the nodes in the network \[30, 19\].

In our view, an earthquake network with nodes having multiple incoming links represents a second order modeling of seismicity, the first being the simple tree structure we have presented here.

## 6 Conclusions

We have introduced a metric to determine correlations between earthquakes that takes into account known statistical properties of seismicity. By means of an appealingly simple yet quantifiable procedure, networks of earthquakes and aftershocks emerge, where the number of aftershocks linked to any event is scale free with an index \(\gamma \approx 2\). The metric is constructed by looking backward in time from any particular event and calculating an expected number of events that would occur, compared to events that actually occurred. If this ratio is significantly less than one, then the preceding event is correlated with the particular one. This is reminiscent of Kierkegaard’s adage that life must be lived forward, but can only be understood backward.

Due to the form of the metric \(n\) measuring correlations, larger earthquakes collect aftershocks from larger space-time windows. From Eq. \[3\] these windows have a spatial radius varying with time as \(r_i(T) = [n_c(T - T_i)]^{1/d_f}\). They span an hyperbolic space-time region (see Fig. \ref{fig:space-time}), which is at variance with the usual “rectangular” or convex windows, of constant radius up to a finite time. In our method, at early times after an earthquake, its aftershock collection window is wider in space than it is at later times.

To our knowledge, the idea that an earthquake can be correlated to an event very far away, if it occurs shortly after it, is new and certainly unconventional. But it is consistent with the hypothesis that seismicity is a self-organized critical phenomenon \[31, 32, 33\]. In that case, some locations may be “on the edge of giving an earthquake” (or toppling, according to the sandpile paradigm), and even a small perturbation from an event far away could trigger them. However, we do not necessarily ascribe the correlations measured here to represent a usual cause and effect relationship. In the sandpile paradigm a completely insignificant event, like adding one grain of sand to an enormous pile, can trigger an arbitrarily large avalanche involving the whole system. Indeed, seismicity as one hierarchically correlated self-organized critical process, generates the scale free network of earthquakes and aftershocks.

Our results also suggest that modern network theory may be a useful and illuminating way to approach the complexities of seismicity, including perhaps problems related to prediction. Our metric and network construction
may also have applications to other phenomena with intermittent bursts such as, for instance, solar flares or even turbulence.

References

[1] Turcotte, D. L. (1997) Fractals and Chaos in Geology and Geophysics. (Cambridge University Press, Cambridge), 2nd edition.
[2] Scholz, C. H. (2002) The Mechanics of Earthquakes and Faulting. (Cambridge University Press, Cambridge), 2nd edition.
[3] Kagan, Y. Y. (1994) Physica D 77, 160–192.
[4] Gutenberg, B. & Richter, C. F. (1941) Geol. Soc. Am. Bull. 34, 1–131. Special papers.
[5] Omori, F. (1894) J. Coll. Sci. Imp. Univ. Tokyo 7, 111–200.
[6] Bak, P., Christensen, K., Danon., L., & Scanlon, T. (2002) Phys. Rev. Lett. 88, 178501.
[7] Gardner, J. & Knopoff, L. (1974) Bull. Seism. Soc. Am. 64, 1363–1367.
[8] Keilis-Borok, V., Knopoff, L., & Rotwain, I. (1980) Nature 283, 259–263.
[9] Knopoff, L., Kagan, Y., & Knopoff, R. (1982) Bull. Seism. Soc. Am. 72, 1663–1675.
[10] Knopoff, L. (2000) Proc. Natl. Acad. Sci. USA 97, 11880–11884.
[11] Hirata, T. (1989) J. Geophys. Res. 64, 7507–7514.
[12] Albert, R. & Barabási, A.L. (2002) Rev. Mod. Phys. 74, 47–97.
[13] Bornholdt, S. & Schuster, H. G., eds. (2002) Handbook of Graphs and Networks. (Wiley-VCH).
[14] Corral, A. (2003) Phys. Rev. E 68, 035102(R).
[15] Abe, S. & Suzuki, N. (2002) Scalefree network of earthquakes. preprint [cond-mat/0210289]
[16] Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999) Proc. ACM SIGCOMM, Comput. Commun. Rev. 29, 251-262.
[17] Barabási, A.L. & Albert, R. (1999) Science 286, 509–512.
[18] Wagner, A. (2001) Mol. Biol. Evol. 18, 1283.
[19] Maslov, S. & Sneppen, K. (2002) Science 296, 910.
[20] Hughes, D., Paczuski, M., Dendy, R. O., Helander, P., & McClements, K. G. (2003) Phys. Rev. Lett. 90, 131101.