

Earthquake correlations and networks: A comparative study

T. R. Krishna Mohan* and P. G. Revathi†

CSIR Centre for Mathematical Modelling and Computer Simulation (C-MMACS) Bangalore 560017, India

(Received 23 March 2010; published 15 April 2011)

We quantify the correlation between earthquakes and use the same to extract causally connected earthquake pairs. Our correlation metric is a variation on the one introduced by Baiesi and Paczuski [M. Baiesi and M. Paczuski, *Phys. Rev. E* **69**, 066106 (2004)]. A network of earthquakes is then constructed from the time-ordered catalog and with links between the more correlated ones. A list of recurrences to each of the earthquakes is identified employing correlation thresholds to demarcate the most meaningful ones in each cluster. Data pertaining to three different seismic regions (viz., California, Japan, and the Himalayas) are comparatively analyzed using such a network model. The distribution of recurrence lengths and recurrence times are two of the key features analyzed to draw conclusions about the universal aspects of such a network model. We find that the unimodal feature of recurrence length distribution, which helps to associate typical rupture lengths with different magnitude earthquakes, is robust across the different seismic regions. The out-degree of the networks shows a hub structure rooted on the large magnitude earthquakes. In-degree distribution is seen to be dependent on the density of events in the neighborhood. Power laws, with two regimes having different exponents, are obtained with recurrence time distribution. The first regime confirms the Omori law for aftershocks while the second regime, with a faster falloff for the larger recurrence times, establishes that pure spatial recurrences also follow a power-law distribution. The crossover to the second power-law regime can be taken to be signaling the end of the aftershock regime in an objective fashion.

DOI: [10.1103/PhysRevE.83.046109](https://doi.org/10.1103/PhysRevE.83.046109)

PACS number(s): 89.75.Da, 91.30.Px, 64.60.aq, 05.65.+b

I. INTRODUCTION

Many complex processes in nature have the intrinsic tendency of self-organizing themselves into a critical state. For example, earthquakes, forest fires, avalanches, biological evolution, and so on have this feature in common. These processes are characterized by long-range correlations in space and time and, typically, also exhibit power-law distributions in many of their variables [1]. However, the dynamics of the processes that self-organize the system to the critical state is still not understood clearly [2]. In recent times, many studies have sought to understand seismicity from the viewpoint of complex networks [3–8]. The spatiotemporal properties of the seismicity of a region are analyzed from the patterns exhibited by a network constructed from the earthquake catalog of the region. Such an approach which focusses on the spatial and temporal links between nodes represented by the events in the catalog, without considering the causes of such linkages, is particularly useful in seismicity where the underlying dynamics is still obscure. We adopt this method in this study and compare different networks constructed from earthquake catalogs of three different seismic regions of the globe, California, Japan and the Himalayas.

The clustering of earthquakes in space and time suggests that events that follow in time are, to a certain extent, causally related to the earlier ones. However, restricting the causality connection to a single predecessor or to the somewhat arbitrary mainshock-aftershock scenario may not be enough. Rather, the causality connection can be extended to a cluster of events that are strongly correlated based on data analysis. With this in mind, Baiesi and Paczuski [5] introduced a metric to

quantify correlations between earthquakes. We use a similar metric here, but with a change which will be clarified below. Both the metrics are obtained by combining two of the most robust statistical laws that characterize earthquake data, the Gutenberg-Richter (GR) law [9] and the fractal distribution of earthquake epicenters (see, for example, Ref. [10]). The former law states that the number N of earthquakes of magnitude $\geq m$ vary as

$$N(m) \sim 10^{-bm}, \quad (1)$$

where b is a constant ≈ 1 , but does vary a little with the region and the catalog (see, for example, Ref. [11]). The latter is based on fractal analysis and is of recent origin, but has been shown to be quite robust through the analysis of different data sets from various regions of the globe (see, for example, Refs. [10,12])

$$N'(l) \propto l^{d_f}, \quad (2)$$

where N' is the number of pairs of points separated by a distance l and d_f is the fractal dimension.

If we combine the above two laws, we may state that the average number of events that can occur in the region within a distance l_{ij} separating two events i and j is

$$n_{ij} = K l_{ij}^{d_f} 10^{-bm_i} \Delta m_i, \quad (3)$$

where K is a constant of proportionality that may be related to the seismic activity of the region [13] and Δm_i is the accuracy in measurement of the magnitude m_i . Note that the events of the catalog are assumed to be time ordered with $j > i$ and $t_j > t_i$. We may then define a correlation relation between earthquakes as

$$c_{ij} = \frac{1}{n_{ij}}, \quad (4)$$

*kmohan@cmmacs.ernet.in; <http://www.cmmacs.ernet.in/~kmohan>

†revathi.pg@gmail.com

and note that the correlation between two events is a maximum when n_{ij} is at a minimum [i.e., “unexpected” (few) events are likely to be more correlated than many events in the radius of reckoning which lead to them being “expected”]. Importantly, these unexpected events will be, as seen from the equation, characterized by larger magnitudes and smaller distances resulting in very small values for n_{ij} . This is what we would also expect intuitively. The metric used here ensures that spatially close by events are more correlated and, also, a higher magnitude event contributes a larger value to correlation than a lower magnitude one at that same location.

We construct the network now by making the linkages between correlated events in the time-ordered list of earthquakes. For this purpose, correlation thresholds c_{th} are defined and only those event pairs with $c_{ij} \geq c_{th}$ are linked together in the network. The basic premise is that all subsequent events in the catalog are recurrences to the selected event, but a pruning is being done based on c_{th} to ensure that only the most correlated are selected.

Our metric for measuring correlation is similar to the one used by Baiesi and Paczuski [5] (hereafter referred to as BAPA). However, we will be focusing on spatial recurrences rather than aftershocks. It is the latter which was studied in BAPA. Thus, they had chosen to include additionally a time factor t_{ij} ($= t_j - t_i$) in their expression corresponding to Eq. (3). This reduces the temporal radius from within which correlated events will be sought. Our metric does not, unlike in BAPA, leave out events that are separated by large time intervals, as long as they satisfy the correlation relation, Eq. (4) (see Krishna Mohan and Revathi [14] for a more detailed comparison).

It is to be also noted that, in BAPA, the question that was asked was as to which previous event the present event is an aftershock of. To this end, Eq. (3) (with the added factor of t_{ij}) was evaluated in BAPA by holding j fixed and varying i from the set of all previous events. In such a procedure, m_i as well as l_{ij}^{df} keep changing in each evaluation. In our case, we go *forward* in time and ask which event is returning to the same location. Therefore, i and m_i are fixed, j varies between evaluations, and only l_{ij} keeps varying. Thus, in our case, each event selects a spatial window proportional to its magnitude and the subsequent events closer to it get more weight in the computed correlation value. In the case of BAPA, the size of the spatial window depends not on the magnitude of the event whose associations are being looked for, but the earlier events which are being sampled to identify associations. Our procedure identifies all spatial recurrences whereas only aftershocks are sought out in BAPA.

We studied in a recent work [14] (hereafter referred to as CALNTWRK) the spatial recurrences of earthquakes in the California region by following the above procedure. In this paper, we adopt the same procedure to construct networks of correlated events from an earthquake catalog of Japan and from a catalog for the Himalayan belt. We analyze the seismicity of these regions in terms of the topology of these networks and compare the same with the findings from CALNTWRK to extract the robust features of such correlated networks of seismic events.

II. THE REGION AND CATALOGS

The regions chosen for the study here are both regions of high seismic potential. Japan is located in a region of considerable seismic risk. Japan lies on the cusp of the Pacific-Philippine-Eurasian triple plate junction, where the complex interactions of three tectonic plates are unpredictable and loaded with potential activity. Seismicity is dominated by the subduction of the Pacific Plate under the Okhotsk Plate to the north and, in southern Japan, the subduction of the Philippine Sea Plate under the Amurian Plate and the Okinawa Plate [15]. The complex interaction of these plates has produced a long history of damaging earthquakes. The recurrence interval of earthquakes along most crustal faults in Japan is typically quite long, while the recurrence intervals of events along subduction zones is usually much shorter [16]. The region covered in this analysis lies between 126.433°E–148.0°E longitudes and 25.730°N–47.831°N latitudes.

The Himalayas are among the most seismically active regions of the world. The seismicity of the Himalayas is contributed mainly by the north-south convergence of the Indian and Eurasian plates, the east-west convergence of the Indo-Burmese mountain, and the underthrusting of the Indian Plate below the Eurasian Plate [17]. Several studies of this region indicate that shallow focus earthquakes dominate this region and the fault plane solutions indicate the dominance of thrust faulting and strike slip in this region [17]. Though the whole Himalayan belt may be considered as one seismic belt, we have carried out the analysis based on a division into three zones: Western Himalayas (WH) (between 70°E–78°E longitudes and 30°N–38°N latitudes), Central Himalayas (CH) (between 78°E–98°E longitudes and 28°N–38°N latitudes), and North Eastern Himalayas (NEH) (between 88°E–98°E longitudes and 20°N–28°N latitudes). Due to the paucity of data, CH is not analyzed here.

The data sources are (1) Japan University Network Earthquake Catalog [18] for the Japanese data (January 1, 1993 to December 31, 1998) and (2) Advanced National Seismic System (ANSS) catalog [19] (January 1, 1973 to December 31, 2007) for the Indian Himalayas. The data were analyzed against Eq. (1) to confirm adherence to the GR law, by varying the minimum magnitudes (m_{min}) considered in the data sets. The minimum of m_{min} , which gives a reasonable power-law behavior, was subsequently chosen (Fig. 1) and only data with m greater than that considered for the analysis presented here. The m_{min} chosen was 3.0 for the Japanese data and 4.5 for the Indian Himalayas. The Japanese data use the Japan Meteorological Agency (JMA) scale to express magnitude while the ANSS data are expressed in moment magnitude. The ANSS catalog is for a much longer period. However, since the minimum magnitude for the catalog is higher at 4.5, the number of events is much less and thus the data set much poorer.

III. METHODOLOGY

We have chosen to retain the values for K and Δm_i as was used in CALNTWRK for all the regions studied here. Note that the absolute value of these parameters are not critical for this study since only the relative correlation values matter in

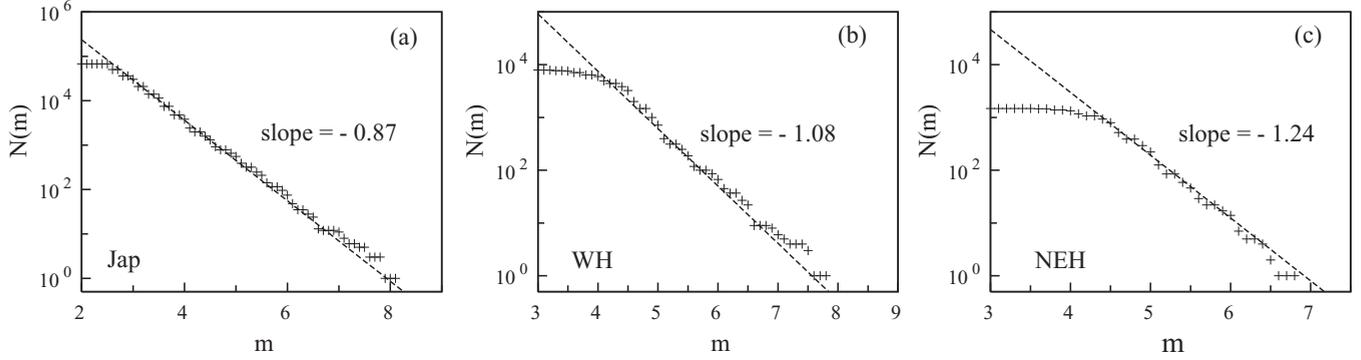


FIG. 1. The log-log plot of Eq. (1) for the different regions studied here. The panels have been labeled after the corresponding seismic regions: (a) “Jap” for Japanese data, (b) “WH” for Western Himalayas, and (c) “NEH” for North Eastern Himalayas.

separating the most correlated from the less correlated. The values of b were obtained from a linear fit to the the log-log plots of Eq. (1) (Fig. 1) and were 0.87 for Japan, 1.08 for WH, and 1.24 for NEH. d_f was set equal to 1.6 for Japan [12] while, for WH and NEH, we estimated d_f ourselves as 1.85 and 1.6, respectively, using the Grassberger-Procaccia algorithm [20].

The distribution of correlation values between all pairs of events in each catalog is similar. Excepting at the upper and lower limits, the distribution is a power law over many decades of the log scale. The range of the power law will be more for a region if the m_{th} value of the catalog is smaller since a lower m_{th} value leads to bigger data sets. However, the range will be different for different regions even if the m_{th} values are the same because of differing sizes of data sets which result because of differing rates of seismicity. The distribution for Japan ($m_{th} = 3.0$) shows a range of about seven decades for the power-law regime, while the two Himalayan regions (both with $m_{th} = 4.5$) have ranges less than five decades. CALNTWRK ($m_{th} = 2.5$) had a range of about eight decades.

The exponents of the power law for all the regions are in the range -1.6 to -1.9 and are given in the plots (Fig. 2); CALNTWRK had an exponent of -1.7 . Again, whether the exponent depends on the region can only be established if we compare the values obtained from catalogs with similar degrees of completeness, homogeneity, and accuracy. For example, we noticed that, in general, for the same region and same catalog, if we increase m_{th} , the magnitude of the exponent

of the power law decreases. Nevertheless, for different regions, the exponent does not agree even if they have the same m_{th} values.

In the analysis below, we have tested the sensitivity of our results as c_{th} and m_{th} values are varied. For this purpose, based on Fig. 2, it was decided to explore a range of c_{th} values from 10^6 to 10^9 for the Japanese data. The different m_{th} values used in the analysis of the Japanese data are 3.0, 3.5, 4.0, and 4.5. Likewise, we employed a range of c_{th} values from 10^6 to 10^9 for the WH data and 10^7 to 10^{10} for the NEH data. For both of these data sets, the m_{th} values tried were 4.5, 5.0, and 5.5.

IV. RESULTS

A. Degree distributions

The maximum out-degree (k_{max}) and maximum in-degree (j_{max}) distributions for the Japanese seismic network [Figs. 3(a) and 3(d), respectively] shows almost similar features as CALNTWRK. In particular, a hub structure is present as far as out-degree is concerned while it is absent with the in-degree distribution; in-degree distribution is more density dependent. The former can be deduced from the fact that k_{max} values are consistently much higher than the j_{max} values. Large earthquakes will tend to have more events associated with them through higher correlation values since Eq. (4) is positively correlated with magnitude. k_{max} is, for the Japanese data, at least an order of magnitude higher than

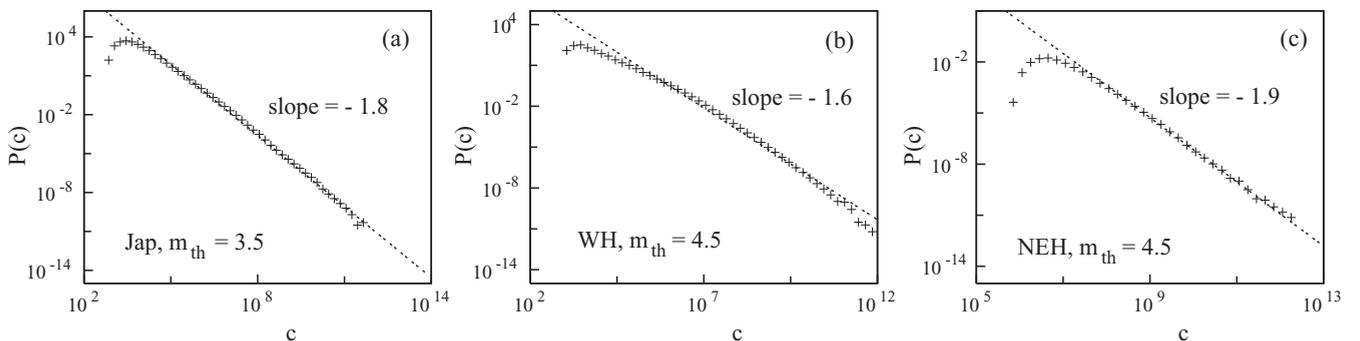


FIG. 2. The distribution of correlation values, evaluated using Eq. (4), for all pairs of events of the corresponding earthquake catalogs. The panels have been labeled after the catalogs: (a) is for Japan, (b) is for WH, and (c) is for NEH.

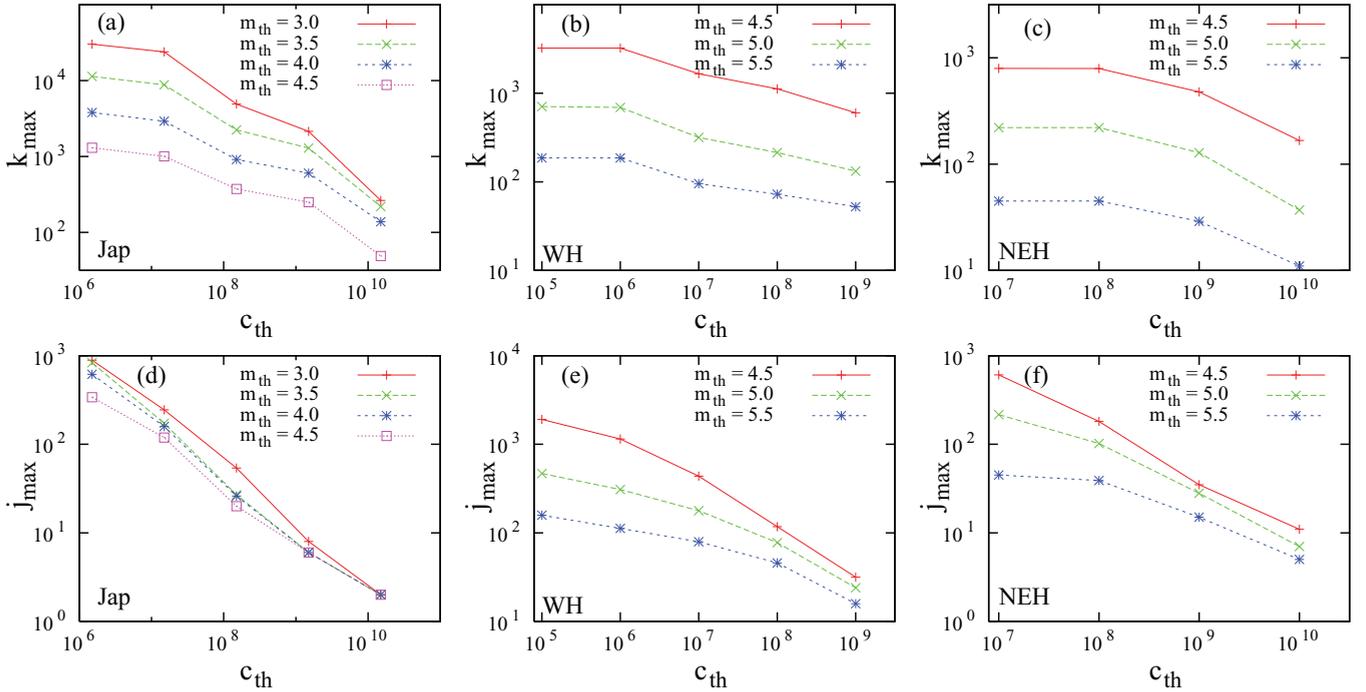


FIG. 3. (Color online) The maximum out-degree (k_{\max}) (top panels) and maximum in-degree (j_{\max}) (bottom panels) values are plotted against increasing c_{th} values for the different seismic regions considered. The seismic regions pertaining to the panels have been labeled therein with (a,d) for Japan, (b,e) for WH and (c,f) for NEH.

the j_{\max} for the corresponding m_{th} values. On the other hand, lower magnitude events require small l_{ij} 's to obtain larger correlation values (i.e., the density of points become important for such events). In the case of both the Himalayan regions [Figs. 3(b) and 3(e) for WH and, Figs. 3(c) and 3(f) for NEH], we observe that the k_{\max} values are not so much larger than the corresponding j_{\max} values as in the Japanese (or California) case. This needs to be investigated with better data sets to try to understand whether the distribution of different magnitudes across the region is in any way significantly different from the other regions.

As c_{th} is increased, k_{\max} values falloff in approximately a power-law fashion. This was observed with CALNTWRK and we see the same behavior with these data sets as well. However, the falloff in the Japanese case is slightly less uniform. Whereas with the CALNTWRK we observed a slower initial falloff followed by a faster falloff for the higher c_{th} values, we find this pattern repeated twice for the Japanese data. The parallelism between the different graphs for different m_{th} values during the falloff, which confirms that a scale invariance with respect to magnitude is maintained as c_{th} is increased, is mostly present with the Japanese data, excepting only in the case of the graph for $m_{\text{th}} = 3.0$. The latter converges to the same value as that of the graph with $m_{\text{th}} = 3.5$ for $c_{\text{th}} = 10^{10}$.

We had observed, in the case of CALNTWRK, a convergence of j_{\max} values as c_{th} is increased resulting in a common value for all m_{th} values at higher c_{th} values. This is in accordance with the observation that the in-degree is controlled by the density of points. As c_{th} is increased, only the higher magnitude events and close by events are left in the clusters and, given the sparse network, all clusters have more or less similar densities. We observe the same behavior here with all

data sets. The only difference is that, in the case of Japanese data, the difference in j_{\max} values, for different m_{th} values, are not as much for low c_{th} values as for the other regions [cf., for example, Figs. 3(d) with 3(e) and 3(f)]. This tendency to uniformity in the density of j_{\max} values across different m_{th} values needs to be explored further because it suggests a uniform distribution of magnitudes with the same j_{\max} values.

As far as the general out-degree distribution is concerned [Fig. 4(a) for the Japanese data, Fig. 4(b) for WH, and Fig. 4(c) for NEH], all the regions considered have similar behaviors with an approximate power-law falloff with increasing k values. For the intermediate range of k values, the exponent is close to -2.0 for the CALNTWRK and Japanese data. This is also in agreement with the exponent value quoted by BAPA. On an average (for $c_{\text{th}} > 10^7$), the same exponent value is obtained for the Himalayan region as well. For $c_{\text{th}} = 10^7$, a departure is observed from this value in the case of WH and NEH and, in particular, NEH shows a significant plateau region (see below). The graphs depart from the general trend for very small k values as well as very large k values. This is understandable since very small k values indicate isolated events and very large k values indicate very large magnitude events which are also sparse as attested to by the GR law [Eq. (1)]. There is, however, some new features present in the distribution of k values for the Himalayas. This is prominent in the case of NEH where we see that, for the smallest c_{th} value, a significant plateau is observed in the distribution. We can identify a similar tendency in the case of WH, even if it is not so prominent. Since smaller c_{th} values are dominated by smaller magnitudes and larger distances, we see in these cases a tendency toward a uniform density with respect to the spatial distribution of such events. Perhaps this may be

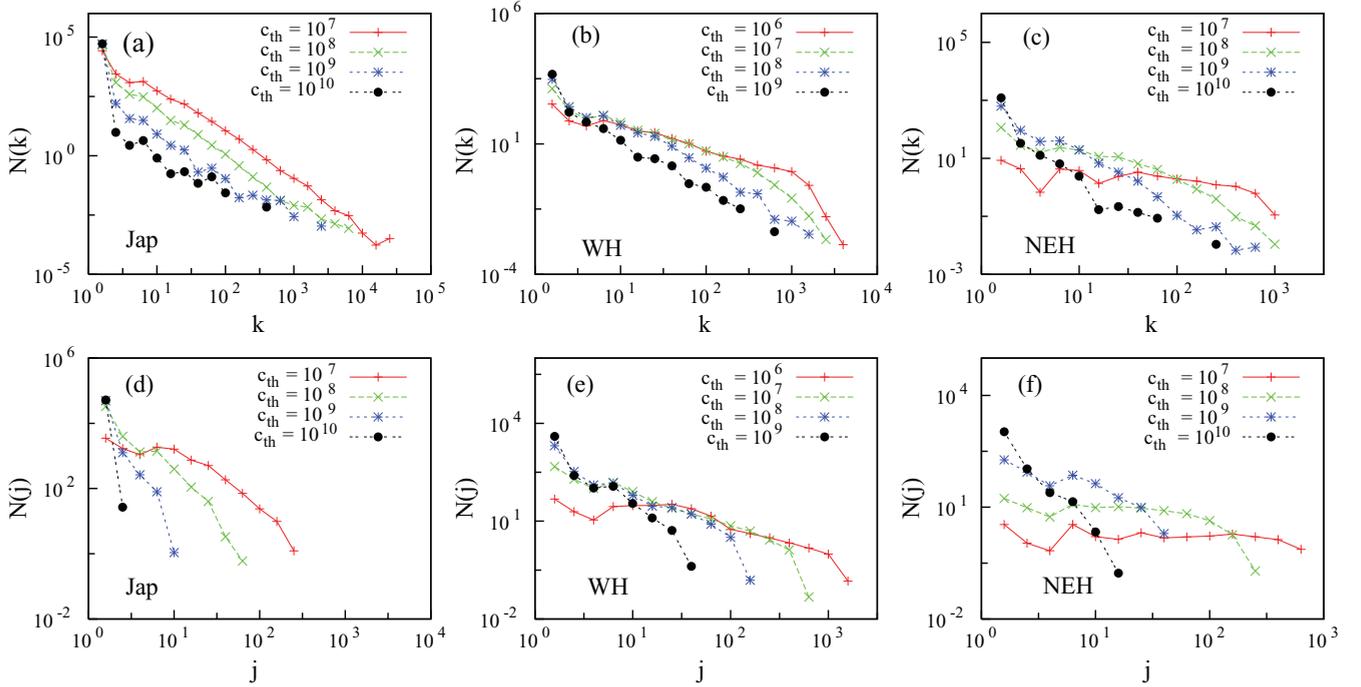


FIG. 4. (Color online) The general out-degree (k) (top panels) and in-degree (j) (bottom panels) distributions are plotted for the different seismic regions with the panel labels indicating the region analyzed. (a,d) are for Japan, (b,e) are for WH and (c,f) are for NEH.

indicating a uniform distribution of new fractures generated by plate tectonics. A more complete and homogeneous data set needs to be studied with respect to these areas to clarify this behavior further.

The general in-degree distribution [Fig. 4(d) for the Japanese data, Fig. 4(e) for WH, and Fig. 4(f) for NEH] is similar across all regions of California, Japan, and the Himalayas. They all show a significant plateau region for the lowest c_{th} value suggesting a uniformity with respect to the distribution of intermediate j values across all the clusters, a prominent density-dependent feature. When the network gets sparser with higher c_{th} values, the number of available j values decrease, and the distribution has a steeper falloff in an approximate power-law fashion.

B. Recurrent time distributions

An important feature of earthquake data is that aftershocks decay according to the Omori law [21,22]

$$n(t) \sim \frac{K}{(c+t)^p}, \quad (5)$$

where c and K are constants in time but depend on magnitude m and $p \approx 1$; n is the number of aftershocks per unit time. Earthquakes of all magnitudes have aftershocks which decay according to the Omori law. It is widely recorded in the literature that even intermediate magnitude events can have aftershocks that persist for years [5,22,23]. We investigate this by studying the distribution of *recurrence times* (also referred to as “waiting times” in the literature [24,25]) in the data.

We define here the recurrence time (τ) as the time between successive events in a cluster surrounding an event. τ ranges from seconds to tens of years in all the clusters. We have left

out $\tau < 180$ s because of the likely error margins. It can be seen from Fig. 5 that the graphs for the different m_{th} values, for a fixed c_{th} value [Figs. 5(d), 5(e), and 5(f)], overlap indicating that *all recurrence times* follow the same pattern regardless of the m_{th} values (i.e., there is no dependency on m_{th} at all). This is pretty much the same for the graphs corresponding to different c_{th} values for a fixed m_{th} value [Figs. 5(a), 5(b), and 5(c)], except for the Himalayan graphs which are considerably noisy due to data inadequacy which has already been noted. But even for these graphs the insets which show rescaled graphs indicate a collapse onto a single graph (see below for details).

In the case of CALNTWRK, we had observed that, in a log-log plot of the distribution of τ , a power law holds for most of the range of recurrence times and, in particular, for about five to six decades. In the case of the Japanese data [Figs. 5(a) and 5(d)], we see that such a situation holds very well (for about four decades in this case). However, in addition, we also see a second power-law regime with a different exponent before noise sets in. The Himalayan data, due to its sparsity, show a lot of scatter and do not give visually appealing graphs [Figs. 5(b) and 5(e) for WH and Figs. 5(c) and 5(f) for NEH]. Nevertheless, we can again identify a similar trend of two linear regimes in the log-log plots. We revisited the data for CALNTWRK and concluded that a two linear-regimes analysis does hold in that case also even though the first linear regime is quite prominent in that case and lasts for almost five decades (out of seven decades).

In the insets provided with the plots we have shown rescaled graphs which bring out the linear aspect better. From the inset of, for example, Fig. 5(d), we see that the appearance of about four decades of linearity in the main graph can be made more precise. The plot has two linear regimes, with the first one till $\log \tau \approx 4.5$ and the second for $\log \tau > 4.5$. The Japanese data

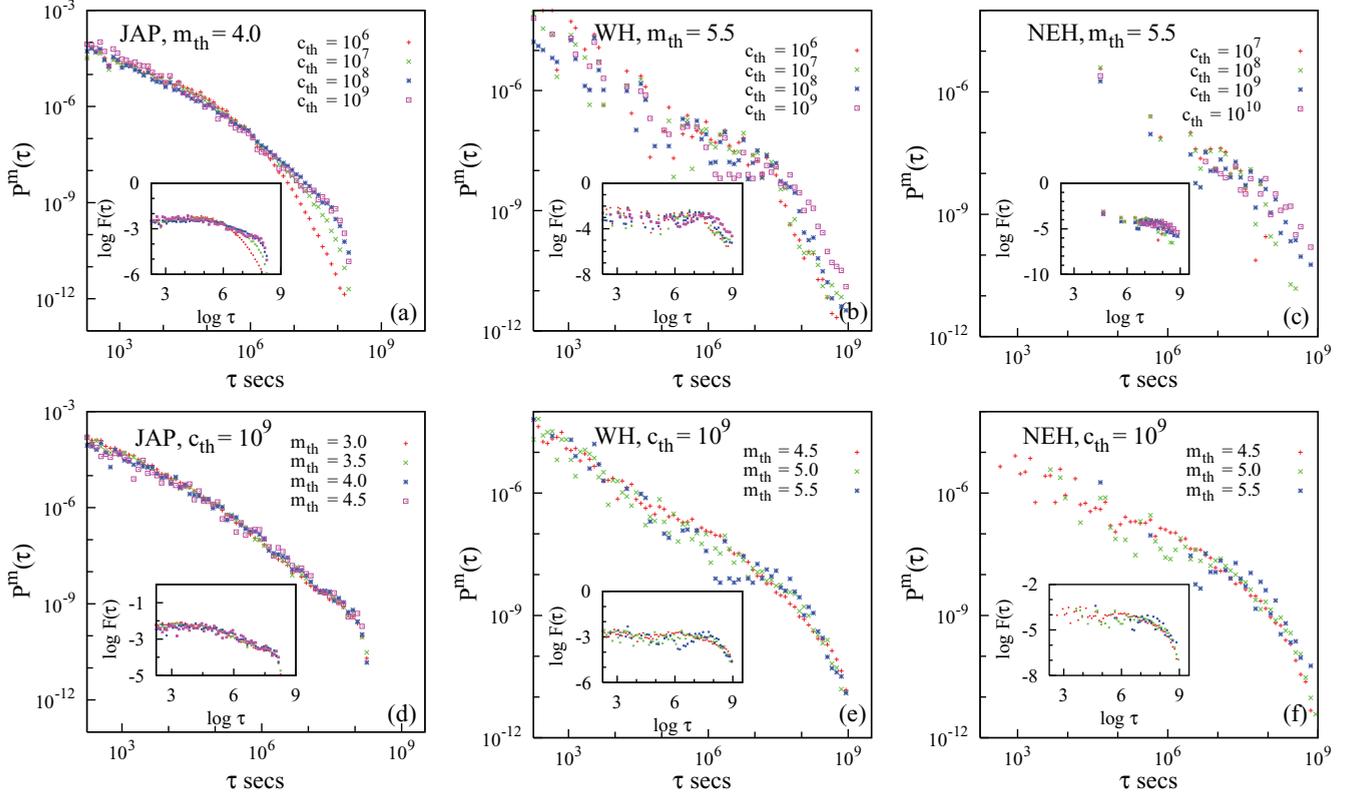


FIG. 5. (Color online) Recurrence time (τ) distribution is plotted for fixed m_{th} and varying c_{th} in the upper row and vice versa in the bottom row. The regions have been labeled in the panels with (a,d) for Japan, (b,e) for WH, and (c,f) for NEH. The insets to the plots show a rescaled version which brings out the linear regimes in the main log-log plots better. Two different linear regimes are clearly seen in the insets. Rescaling function $F(\tau) = \tau^\beta P^m(\tau)$ where β is the slope of the first linear regime in the main graph.

do show that the low correlation data, for c_{th} values 10^6 and 10^7 , do not overlap all the way with the graphs for other c_{th} values [Fig. 5(a)]. Since this occurs for low correlation values, we do not consider it a significant departure from the general trend, but the data will nevertheless be examined later in more detail.

Since we are dealing with only correlated earthquakes, the above trend of two linear regimes suggests that the recurrences belong to two different classes. It is known that the recurrence times display the Omori law in the initial part, for shorter time periods, up to a cutoff time and subsequently the power law changes [26]. It is then clear that the two classes of recurrences we find here are (i) aftershocks in the first part and it is followed by (ii) the general spatial recurrences. It follows that these plots can be used to identify aftershocks in an objective fashion. The time at which the first power law crosses over to the second may be used as the limit of the aftershock time zone. We recall here that no objective criterion has, so far, been suggested in the literature to identify the aftershock time zone, and this has been a serious flaw underpinning the concept of foreshocks, mainshocks, and aftershocks in earthquakes. Quite often there is debate in the literature on whether earthquakes can be identified to be aftershocks when they are well separated in time, of the order of months and years. Our method makes it possible to state the aftershock time zone in an unambiguous fashion. In addition, we have also succeeded in extending the Omori law (with a different exponent, i.e., in so far as there is a power law) to general spatial recurrences not limited by time.

Recently, analysis of the distribution of τ has gained momentum and several works have appeared in the literature [24–33] dealing with it. The trend was initiated by a work of Bak *et al.* [24] where the waiting times in the California region were analyzed. The novelty in their approach was in dividing the analysis region into square (in latitude-longitude units) grids and noting how the distribution varied, if at all, with varying unit cell sizes. No distinction was made between the earthquakes in terms of foreshocks, mainshocks, or aftershocks and all were treated on par with the only restriction being that they employed *complete* catalogs, with no missing events greater in magnitude than a minimum. They came to the conclusion that a scaling law of the form

$$\tau^\alpha P_{S,L}(\tau) = f(\tau^{-b} L^{d_f}), \quad (6)$$

holds where P is the waiting time distribution, $m = \log S$, and L the size of unit cells of the grid. The function f was seen to consist of a constant part and an exponentially decaying part, separated by a sharp kink. They concluded that the constant part consisted of the aftershocks (correlated earthquakes) and the decaying part with uncorrelated earthquakes (i.e., independent earthquakes seeded by the nonzero driving rate of plate tectonics). The scaling function was claimed to be universal and some of the parameters as well; for example, α which was identified with the Omori exponent p of Eq. (5).

If we assume that seismicity is a mixture of random events, each with their own set of triggered aftershocks, as is often assumed in the literature, then the mixture of these

two categories of events can be in arbitrary ways in a given catalog depending on the region and the period of the data analyzed. For example, a larger region is likely to have the Poissonian events interspersed at short waiting times itself in the catalog. On the other hand, a smaller region may have the full set of aftershocks running out before the next Poissonian event strikes (see, for example, Touati *et al.* [27]). Such manners of mixing of the two categories of events may vary with the period of the catalog and rate of seismicity. The epidemic-type aftershock sequences (ETAS) [28] model for seismicity explicitly builds in the above structure for the distribution of events and generates appropriate time series. Based on the above idea, it has been argued that the interpretation of the kink in the distribution by Bak *et al.* [24] may not be appropriate since no distinct range of waiting times can be attributed to either of the two categories of events [29,30].

It has also been shown that another power-law regime exists for small waiting times [26]. This regime was not seen by Bak *et al.* because they had left out τ values < 38 s. This introduces a kink at smaller waiting times in the full distribution which scales differently from the second kink which was the one seen by Bak *et al.* [24]. This was attributed by Davidsen and Goltz [26] to multifractality in epicenter distribution contrary to the single fractal distribution assumed by Bak *et al.* [24]. The former concluded that the waiting time distribution is not universal and depends in a nontrivial way on the region under consideration.

Following up on Bak *et al.* [24], Corral [25] studied the τ distributions for different regions, from very small to very large (global level), and concluded that all of them can be made to collapse onto a single graph by rescaling them using the mean seismic rate. He came to the conclusion that the τ distribution can be modeled as a gamma distribution

$$p(\Delta t) = C \cdot (\Delta t)^{\gamma-1} \cdot e^{-\Delta t/\beta}, \quad (7)$$

with $C = 0.5 \pm 0.1$, $\gamma = 0.67 \pm 0.05$, and $\beta = 1.58 \pm 0.15$. In the above, Δt is the normalized waiting time (i.e., $\Delta t = \lambda \tau$ where λ is the seismic event rate). He also argued for clustering beyond the aftershock bursts (i.e., clustering of the independent events seeded by plate tectonics, as well as self-similarity of such clustering in the space-time-magnitude domain). A key criticism of this work has been that it requires the data to be stationary failing which a mean seismic rate may not define (see Corral [31] for an approach based on an instantaneous rate of seismicity). This necessitates that the initial set of aftershocks (short waiting times) be ignored because of fast change in the rate of seismicity.

In general, it has been argued that the τ distribution is bimodal. The correlated events (aftershocks) are expected to follow the gamma distribution and the Poissonian events an exponential distribution with a crossover region between the two which may resemble a power law [27]. The exact form of $p(\Delta t)$ depends on the percentage of Poissonian events and thus is not universal. This has been shown analytically [32] and using ETAS model simulations [33].

Note that we are selecting only correlated earthquakes for our analysis; we are not dealing with the Poissonian events at all. We also note that our correlation metric derived from Eq. (3) combines Eqs. (1) and (2) in a similar fashion to

Eq. (6). While we use such a metric to cull the correlated events, Bak *et al.* [24] tried to rescale the raw data using it and thereby distinguish between correlated and uncorrelated events. However, such an approach is questionable because, as stated before, no distinct range of waiting times can be attributed to either of the two categories of events [29,30]. Our focused approach to correlated events show that a kink appears in the τ distribution for correlated events itself which separates the aftershocks from the purely spatial recurrences and both are distributed in a power-law fashion. While new fracture zones will be opened up by repeated earthquakes, a majority of earthquakes will tend to visit old fault zones which are repeatedly subjected to a healing process before once again fracturing (see, for example, Nielsen [34,35]). We believe this process induces long-term correlations which show up in the data in the form of correlated earthquakes (i.e., purely spatial recurrences are a ubiquitous feature). The latter populate the second power-law regime in plots like Fig. 5. We may also note here that the variation in grid size (L) effected by Bak *et al.* [24] is akin to variation in c_{th} in our case; higher c_{th} values imply smaller distances between earthquake pairs and lower c_{th} values imply larger distances between earthquake pairs. We have, as already noted, confirmed that the same pattern exists for all c_{th} sizes [for fixed m_{th} values; Figs. 5(a), 5(b), and 5(c)].

C. Recurrence length distributions

By recurrence length, we mean the distance from an event to each of the other events in the same cluster. In other words, we are measuring the scatter in the way the subsequent events in a cluster are distributed around the event in question. While there is no established law like the Omori law for recurrence lengths, studies have indicated that a power law seems to hold for recurrence lengths as well [8]. To investigate this, we use the great circle distance r_{ij} for computing the distance between two event locations. This is calculated using the Haversine formula [36] with the radius of the earth taken as 6367 km.

The general appearance and behavior of the graphs for the Japanese as well as Himalayan data (Fig. 6) are similar to the CALNTWRK. The probability distribution of recurrence lengths is unimodal in nature and peaks at a recurrence length (r_{ch}^m) which is larger for increasing values of m_{th} , for a fixed c_{th} . The sparsity of Himalayan data gives rise to much scatter and does not give clean graphs [Figs. 6(b) and 6(e) for WH and Figs. 6(c) and 6(f) for NEH]. Nevertheless, the above behavior can be clearly seen there as well. Keep in mind that the distribution with lower m_{th} value contains the data points appearing in the distribution with higher m_{th} values. Also, not all the data points appearing with the lower m_{th} value is present with the higher m_{th} value. Thus, as m_{th} increases, the data points corresponding to the lower m values get left out of the distribution. The shifting of r_{ch}^m to the right, as m_{th} is increased, therefore points to the presence of a characteristic r_{ch}^m associated with particular m values and, this clearly is larger for larger m values. The rupture is larger with larger earthquakes and we may think of r_{ch}^m as representative of a typical rupture length associated with events of that magnitude (see also Ref. [8]).

While the graphs for different m_{th} values are separated before the rise to the peak, they are collapsed onto a single

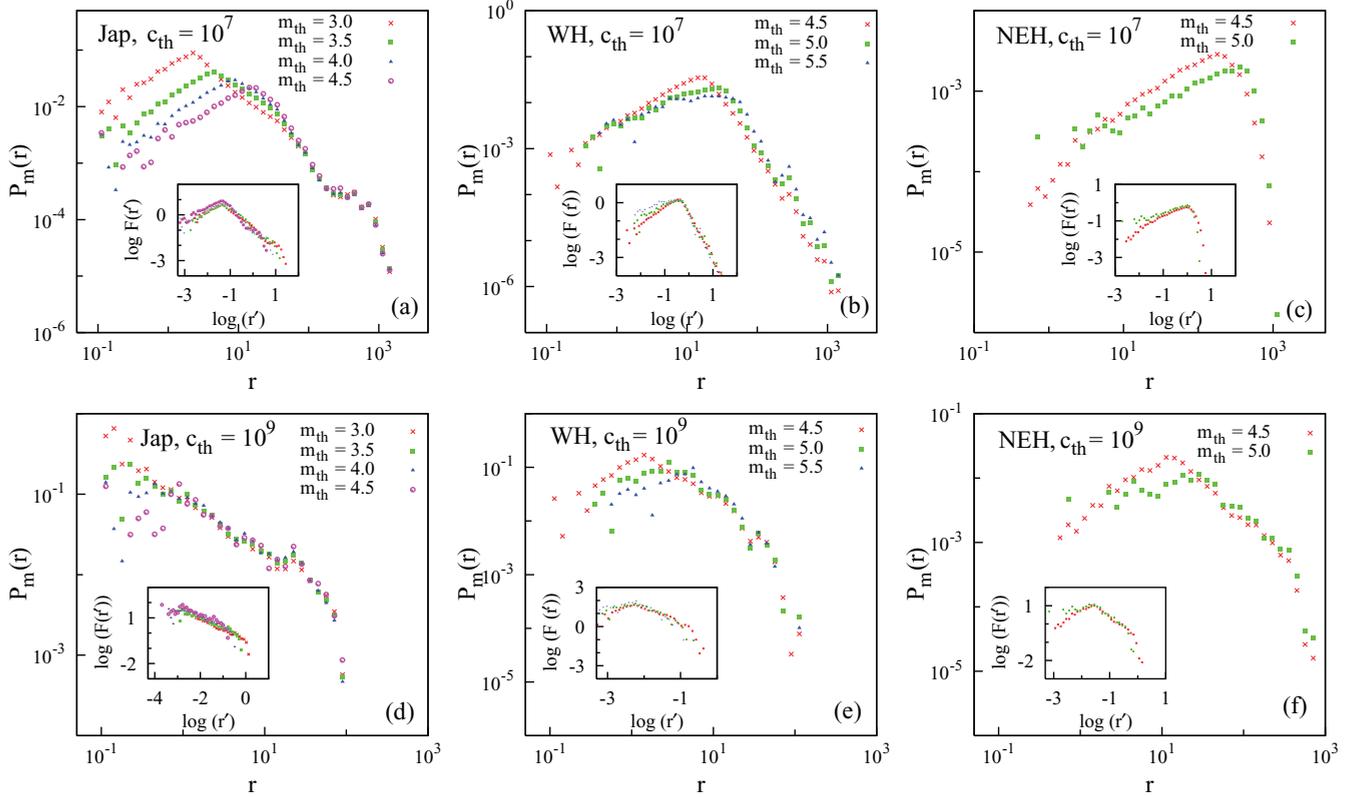


FIG. 6. (Color online) Recurrence length (r) distribution is plotted for two different c_{th} values for the three different seismic regions (indicated in the respective panels) considered here. Top panels are for $c_{th} = 10^7$ and the bottom panels are for $c_{th} = 10^9$. (a,d) are for Japan, (b,e) are for WH, and (c,f) are for NEH. The insets show rescaled graphs, where $r' = r/10^{\alpha m_{th}}$ and $F(r') = 10^{\alpha m_{th}} P_m(r)$; α is a function of the data set and is computed from a linear fit to the plot of r_{ch}^m values (the peaks in the plots) against m_{th} .

curve after their peaks. We had also observed the same in our study of CALNTWRK. This trend is easily enough observed in the case of the Japanese data, but more difficult to confirm for the Himalayan data because of the scatter induced by data paucity. The presence of the power law, before the peak, with a common exponent for the different m_{th} values, suggest that the location errors for these small recurrence length values are not random in nature and, if any, are systematic in nature. This is contrary to what has been claimed in the literature [8], as already suggested by us in an earlier work [14].

As c_{th} is increased, the part of the graphs before the peak is shortened (compare the top panels with the corresponding bottom panels in Fig. 6), eventually disappearing for sufficiently high c_{th} values (not shown here). Note that we are plotting only pairs of events with $r_{ij} > 100$ m. As such, the peaks disappear for larger values of c_{th} since the peak values are now < 100 m. The behavior seen with r_{ch}^m is completely in accordance with Eq. (3). As c_{th} increases, we are selecting event pairs of larger magnitudes separated by smaller distances.

The part of the graphs, for different m_{th} values, after the peaks, are collapsed onto a single curve till data paucity forces scatter for the higher recurrence length values. The former may be expected from the fact, already stated, that the data points for the lower m_{th} value contain the data points for the higher m_{th} values. In the case of the Japanese data [Figs. 6(a) and 6(d)], we see a small plateau appearing after the falloff just before the very high recurrence length values. In the Himalayan case

[Figs. 6(b) and 6(e) for WH, Figs. 6(c) and 6(f) for NEH], the falloff is seen to be much steeper compared to the other cases because of the smaller range of recurrence lengths.

The plots for the different values of m_{th} collapse on to a single curve when the x axis is rescaled as $r' = r/10^{\alpha m_{th}}$ and the y axis by $10^{\alpha m_{th}} P_m(r)$. α is a function of the data set and is the rate of variation of r_{ch}^m with m_{th} . It is computed from a linear fit to the plot of r_{ch}^m values against m_{th} . α was seen to be the same regardless of the c_{th} value in the case of CALNTWRK. We observe the same feature in the case of the Japanese data as well. However, in the case of the Himalayan data, different α values were obtained for the different c_{th} values. This was not surprising because, given the sparse data set and with just three m_{th} values to play with, it was difficult to spot the r_{ch}^m peaks correctly. We have, however, still used a single α value for any particular Himalayan region, and with reasonable success as seen from Fig. 6, with this value being an averaged value over the different α values obtained for the different c_{th} values.

V. CONCLUSION

In a continuation of our earlier study, we have compared the network features of seismic data from three different areas, California, Japan, and the Himalayas. The networks were constructed according to our algorithm, based on a modification of the algorithm given by BAPA. Both these algorithms are based on a metric for correlations between

earthquakes and our algorithm differs from BAPA in omitting the time factor from their definition to enable the spatial recurrences to be identified, as opposed to the identification of aftershocks by BAPA. Also, we evaluate the correlations going forward in time as against the backward associations in BAPA. The approach developed in BAPA locates the most correlated predecessor which is identified as the mainshock of the present event which, in turn, is therefore one of the aftershocks of that predecessor. Our algorithm goes forward in time and identifies all spatial recurrences to the present event from the subsequent events in time.

Our comparisons show that there are some robust features present in networks constructed from seismic data from around the globe. These include a hub structure in the out-degree distribution accompanied by a density-dependent in-degree distribution, unimodal recurrence length distribution with a peak value that increases with m_{th} indicating that characteristic rupture lengths for different magnitudes may be present and a recurrence time distribution, which has two power-law regimes with different exponents. Note that we are employing only correlated earthquakes for the analysis. Hence, the two different power-law regimes in the recurrent time distribution

suggests an initial aftershock regime (following Omori's law) followed by a pure spatial recurrences regime. We have therefore proposed that the recurrence time at which the first power-law regime crosses over to the second one be used as an objective criterion for the upper limit of the time window used for selecting aftershocks. This is a significant new result since no objective criterion was available till now for identifying aftershocks. Furthermore, this also establishes that recurrence times for spatial recurrences, after the initial aftershocks, also fall off in a power-law fashion akin to the Omori law for aftershocks, albeit with a different exponent (higher in magnitude).

It is clear from this comparative study that earthquake networks provide us with a useful analysis tool to study the seismicity patterns of the globe. Further work in this area should lead us to a better understanding of the event distribution and, eventually, the dynamics behind it.

ACKNOWLEDGMENTS

P.G.R. would like to acknowledge University Grants Commission, India for support under the Faculty Improvement Programme.

-
- [1] M. Newman, *Contemp. Phys.* **46**, 323 (2007).
 - [2] J. Davidsen and M. Paczuski, *Phys. Rev. Lett.* **94**, 048501 (2005).
 - [3] S. Abe and N. Suzuki, *Europhys. Lett.* **65**, 581 (2004).
 - [4] S. Abe and N. Suzuki, *Nonlinear Processes in Geophysics* **13**, 145 (2006).
 - [5] M. Baiesi and M. Paczuski, *Phys. Rev. E* **69**, 066106 (2004).
 - [6] M. Paczuski, in *Complexity, Metastability and Nonextensivity* (World Scientific, London, 2005).
 - [7] J. Davidsen, P. Grassberger, and M. Paczuski, *Geophys. Res. Lett.* **33**, L11304 (2006).
 - [8] J. Davidsen, P. Grassberger, and M. Paczuski, *Phys. Rev. E* **77**, 066104 (2008).
 - [9] B. Gutenberg and C. Richter, *Seismicity of the Earth* (New York: the Society, Princeton, NJ, 1941).
 - [10] D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics* (Cambridge University Press, Cambridge, England, 1997).
 - [11] Y. Bayrak, A. Yilmaztürk, and S. Öztürk, *Journal of Geodynamics* **34**, 653 (2002).
 - [12] T. Hirata, *Pure Appl. Geophys.* **131**, 157 (1989).
 - [13] M. Baiesi, *Physica A* **360**, 534 (2006).
 - [14] T. R. Krishna Mohan and P. G. Revathi, *Journal of Seismology* **15**, 71 (2011).
 - [15] M. Mahdyiar, Air Currents, Feb 19, 2009 (No. 2, 2009), Air Worldwide Corporation [<http://www.air-worldwide.com/PublicationsItem.aspx?id=16678>].
 - [16] Earthquake Research Committee, *Seismic Activity in Japan—Regional Perspectives on the Characteristics of Destructive Earthquakes*, The Headquarters for Earthquake Research Promotion (HERP), Japan (2005).
 - [17] S. Srilakshmi and R. K. Tiwari, *Geofísica Internacional* **46**, 65 (2007).
 - [18] [<http://kea.eri.u-tokyo.ac.jp/catalog/junec/monthly.html>].
 - [19] [<http://www.ncedc.org/>].
 - [20] P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983).
 - [21] F. Omori, *J. Coll. Sci., Imp. Univ. Tokyo* **7**, 111 (1894).
 - [22] T. Utsu, Y. Ogata, and R. Matsuura, *J. Phys. Earth* **43**, 1 (1995).
 - [23] Y. Ogata and K. Shimazaki, *Bull. Seismol. Soc. Am.* **74**, 1757 (1984).
 - [24] P. Bak, K. Christensen, L. Danon, and T. Scanlon, *Phys. Rev. Lett.* **88**, 178501 (2002).
 - [25] Á. Corral, *Phys. Rev. E* **68**, 035102 (2003).
 - [26] J. Davidsen and C. Goltz, *Geophys. Res. Lett.* **31**, L21612 (2004).
 - [27] S. Touati, M. Naylor, and I. G. Main, *Phys. Rev. Lett.* **102**, 168501 (2009).
 - [28] Y. Ogata, *J. Am. Stat. Assoc.* **83**, 9 (1988).
 - [29] M. Lindman, K. Jonsdottir, R. Roberts, B. Lund, and R. Bødvarsson, *Phys. Rev. Lett.* **94**, 108501 (2005).
 - [30] K. Jonsdottir, M. Lindman, R. Roberts, B. Lund, and R. Bdvarsson, *Tectonophysics* **424**, 195 (2006).
 - [31] Á. Corral, *Phys. Rev. Lett.* **92**, 108501 (2004).
 - [32] G. Molchan, *Pure Appl. Geophys.* **162**, 1135 (2005).
 - [33] S. Hainzl, F. Scherbaum, and C. Beauval, *Bull. Seismol. Soc. Am.* **96**, 313 (2006).
 - [34] S. Nielsen and R. Madariaga, *Bull. Seismol. Soc. Am.* **93**, 2375 (2003).
 - [35] S. Nielsen, J. Taddeucci, and S. Vinciguerra, *Geophy. J. Int.* **180**, 697 (2010).
 - [36] R. W. Sinnott, *Sky and Telescope* **68**, 159 (1984).