

Dependence of earthquake recurrence times and independence of magnitudes on seismicity history

Álvaro Corral

Departament de Física, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain

Received 22 July 2005; received in revised form 6 October 2005; accepted 25 March 2006

Available online 9 June 2006

Abstract

The fulfillment of a scaling law for earthquake recurrence–time distributions is a clear indication of the importance of correlations in the structure of seismicity. In order to characterize these correlations we measure conditional recurrence–time and magnitude distributions for worldwide seismicity as well as for Southern California during stationary periods. Disregarding the spatial structure, we conclude that the relevant correlations in seismicity are those of the recurrence time with previous recurrence times and magnitudes; in the latter case, the conditional distribution verifies a scaling relation depending on the difference between the magnitudes of the two events defining the recurrence time. In contrast, with our present resolution, magnitude seems to be independent on the history contained in the seismic catalogs (except perhaps for Southern California for very short time scales, less than about 30 min for the magnitude ranges analyzed).

© 2006 Elsevier B.V. All rights reserved.

Keywords: Statistical seismology; Recurrence times; Intervent times; Correlations; Scaling and universality; Renormalization group

1. Introduction

Although seismology has intensively analyzed the characteristics of single earthquakes with great success, a clear description of the properties of seismicity (i.e., the structures formed by all earthquakes in space, time, magnitude, etc.) is still lacking. It is true that a number of empirical laws have been established since the end of the XIX century, but these laws are not enough to provide a coherent phenomenology for the overall picture of earthquake occurrence (Mulargia and Geller, 2003). In general, the approach to this phenomenon has to be statistical, as the Gutenberg–Richter law for the number of occurrences as a function of magnitude

exemplifies. Moreover, for some studies it is necessary to assume that an earthquake can be reduced to a point event in space and time; despite this drastic simplification, the resulting point process for seismicity retains a high degree of complexity. In this context, the problem turns out to be very similar to those usually studied in statistical physics. Remarkably, in the last years the so-called ETAS model (which combines the Gutenberg–Richter law, the Omori law, and the law of productivity of aftershocks) has been widely studied (Ogata, 1988, 1999; Helmstetter and Sornette, 2002); however, this model is still a first order approximation to real seismicity, useful mainly as a null-hypothesis model.

In contrast to the issues of the size of earthquakes or the temporal variations of the rate of occurrence, the timing of earthquakes (despite its potential interest for forecasting and risk assessment) has been the subject of

E-mail address: Alvaro.Corral@uab.es.

URL: <http://einstein.uab.es/acorral/>.

relatively few empirical statistical studies, with the additional drawback that no unifying law for the time occurrence has emerged, at variance with the general acceptance of the Gutenberg–Richter law and the Omori law (Gardner and Knopoff, 1974; Udías and Rice, 1975; Koyama et al., 1995; Wang and Kuo, 1998; Ellsworth et al., 1999; Helmstetter and Sornette, 2004), see also the citations of Smalley et al. (1987) and Sornette and Knopoff (1997). Rather, results turn out to be in contradiction with each other, from the paradigm of regular cycles and characteristic earthquakes to the view of totally random occurrence (Bakun and Lindh, 1985; Sieh et al., 1989; Stein, 1995; Sieh, 1996; Murray and Segall, 2002; Stein, 2002; Kerr, 2004; Weldon et al., 2004, 2005). In the opposite side of cycle regularity is the view of earthquake occurrence in the form of clusters, which has received more support from catalog analysis (Kagan and Jackson, 1991, 1995; Kagan, 1997).

A first step towards the solution of the time-occurrence problem and its insertion in a unifying description of seismicity was taken by Bak et al., who related, by means of a unified scaling law, the distribution of earthquake recurrence times with the Gutenberg–Richter law and the fractal distribution of epicenters (Bak et al., 2002; Christensen et al., 2002; Corral, 2003, 2004b; Corral and Christensen, 2006; Davidsen and Goltz, 2004). One of the goals of this line of research (apart from the applications in forecasting) should be to build a theory of seismicity that clarifies which kind of dynamical process we are dealing with. Simple behaviors such as periodicity, random occurrence, or even chaos seem too simplistic to account for the complexity of seismicity; in contrast, self-organized criticality fits better with this purpose (Bak, 1996; Jensen, 1998; Turcotte, 1999; Sornette, 2000; Hergarten, 2002), although different alternatives have been proposed more recently, see Rundle et al. (2003) and Sornette and Ouillon (2005).

Nevertheless, the recurrence–time distribution introduced by Bak et al. (2002) was defined in a somewhat complicated way through the mixing of recurrence times coming from (more or less) squared equally-sized regions with disparate seismic rates. A simpler, alternative perspective, in which each (arbitrary) region has its own recurrence–time distribution, has been developed (Corral, 2004a, 2005b). In this case, a universal scaling law involving the Gutenberg–Richter relation describes these distributions for a wide range of magnitudes and sizes of the regions, as long as seismicity shows a stationary behavior.

In this paper we extend this procedure to study conditional probability distributions. These distributions

will provide important information about correlations in seismicity, which are fundamental for the existence of the scaling law for recurrence–time distributions. It turns out that recurrence times depend on previous recurrence times and magnitudes, in such a way that small values of the recurrence time and large magnitudes lead to small recurrence times, and vice versa. In contrast, the magnitude is fairly independent on history (although obviously we cannot reject the hypothesis of a dependence weaker than our uncertainty).

2. Scaling of recurrence-time distributions and renormalization-group transformations

The main subject of our study is the earthquake recurrence time (also called waiting time, interevent time, interoccurrence time, etc.), defined as the time between consecutive events, once a window in space, time and magnitude has been selected. In this way, the i -th recurrence time is defined as

$$\tau_i \equiv t_i - t_{i-1}, \quad i = 1, 2, \dots$$

where t_i and t_{i-1} denote, respectively, the time of occurrence of the i -th and $i-1$ -th earthquakes in the selected space–time–magnitude window. Note that, following Bak et al., once the window has been selected, no further elimination of events is performed and then, all events are equally considered, independently of their hypothetical consideration as mainshocks, aftershocks or foreshocks. This is an important difference with respect to the “standard practice,” usually more focused in declustered catalogs.

The recurrence time is a variable quantity and it is convenient to work with its probability density. Let us recall that a probability density (in this case of recurrence times) is defined as the number of recurrence times within a small interval of values, normalized to the total number of recurrences and divided by the length of the small interval (to remove the dependence on it); to be precise,

$$D(\tau) \equiv \frac{\text{Prob}[\tau \leq \text{rec. time} < \tau + d\tau]}{d\tau}, \quad (1)$$

where Prob denotes probability, τ refers to a precise value of the recurrence time, and $d\tau$ is the length of the interval. Strictly, $d\tau$ should tend to zero, but in practice it is necessary a compromise to reach some statistical significance for each interval, as the number of data is not infinite. Moreover, as there are multiple scales involved in the process (from seconds to years) it is more convenient to consider a variable $d\tau$, with the appropriate size for each scale.

An easy recipe is to consider the different intervals (in seconds with a minimum recurrence time of 1 s, for instance) growing as $[\tau, \tau+d\tau]=[1, c), [c, c^2), [c^2, c^3)$, etc., with $c>1$. This procedure is sometimes referred to as logarithmic binning, as in log scale the intervals $d\tau$ have the same length. Note that the widely used cumulative probability distribution is just the integral of $D(\tau)$; nevertheless, we find that the characteristics of the distributions are more clearly reflected in the density.

The key point of the procedure introduced by Bak et al., which we adopt, is the systematic analysis of the dependence of the probability density on the parameters that define the magnitude and spatial windows. For a spatial region of arbitrary coordinates, shape and size, independent of tectonic divisions or provinces, only events with magnitude M larger than a threshold value M_c are selected (naturally, the value M_c has to be high enough to guarantee the completeness of the data). This means that the recurrence-time density, $D(\tau)$, depends on M_c and on the size, shape, and coordinates of the region; in order to stress this dependence we add a subscript w (from window) to the density, which turns out $D_w(\tau)$.

An additional point to take into account is the heterogeneity of seismicity in time, with large variations in the number of events depending on the occurrence of large earthquakes (what is usually referred to as aftershock sequences, and also earthquake swarms). In a first step we only consider time windows with

stationary seismicity or, more precisely, seismicity homogeneous in time. This means that, for a particular time window, the statistical properties of the process do not change when the window is divided into shorter windows, except for the fluctuations associated to the finite size of the windows; in particular, the mean value defined for any property of the process will be practically the same for any of the shorter time windows, provided that these windows are not too short. If, for a given window, we define its mean seismic rate R_w as the number of earthquakes divided by the time period covered by the window, i.e.

$$R_w = \frac{\text{number of events in } w}{\text{time extent of } w},$$

then, stationarity imposes that the mean rate is roughly constant for different windows and a unique seismic rate R can be defined for a long period (however, note that there can be different stationary periods with different rates, although experience tells us that these rates are also similar). When we plot the accumulated number of earthquakes versus time, as in Fig. 1, stationarity is easily identified by a linear behavior in this plot, as the rate is just the slope of the curve. In order to quantify stationarity, we may fit straight lines to some part of the behavior of the accumulated number of earthquakes as a function of time. For the examples in the figure, which will be the data analyzed in this paper as we will explain next, the correlation coefficient for a minimum-square

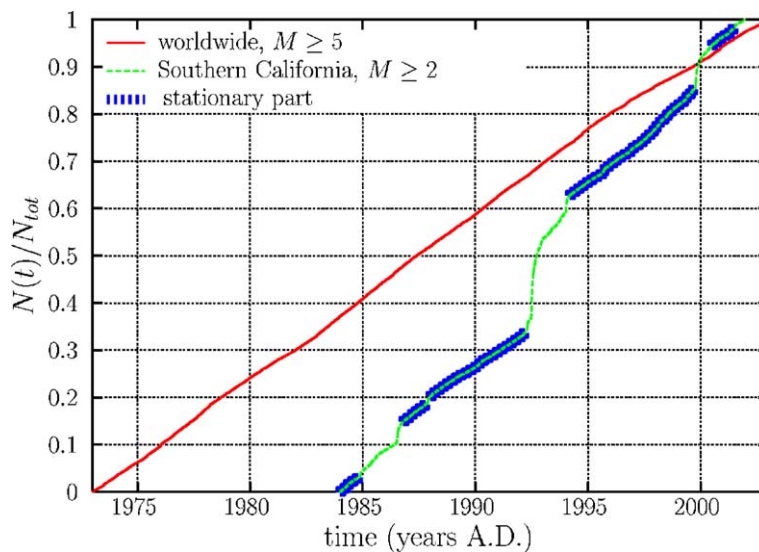


Fig. 1. Cumulative number of earthquakes $N(t)$ (normalized by the total number N_{tot} in the whole period considered) as a function of time, for worldwide seismicity with $M \geq 5$ and for Southern California seismicity with $M \geq 2$ ($N_{\text{tot}}=46,054$ and $84,771$ respectively). The linear behavior is an indication of stationary occurrence, which clearly applies to worldwide seismicity. In contrast, for Southern California we observe a nonstationary behavior, except for some periods of time. Some of these stationary periods have been studied in the text and are marked by a thick line.

procedure is always in between 0.995 and 0.9997, although the quality of the fit can still be improved by a better selection of the time windows. Nevertheless, this will be sufficient for our purposes.

It is illustrative to start our analysis with worldwide earthquakes. In this case, the spatial region of study is fixed to the whole Earth, and only the magnitude threshold M_c is changed. The main advantage of worldwide seismicity is that earthquakes occur at a uniform rate (at least for the last 30 years). In particular, we use the NEIC-PDE catalog (National Earthquake Information Center, Preliminary Determination of Epicenters, available at http://www.neic.cr.usgs.gov/neis/epic/epic_global.html) for the years 1973 to 2002.

In order to confirm the stationarity of worldwide earthquake occurrence, we plot in Fig. 1 the cumulative number of earthquakes with $M \geq 5$ since 1/1/1973 versus time. A linear trend, characteristic of stationary behavior, is clearly seen, with a rate of occurrence given roughly by $R_w \approx 4$ events per day.

Fig. 2 (top) shows the probability densities of the recurrence time in this case for several magnitude thresholds M_c . As M_c increases, the densities become broader (due to the decrease in the number of events and the corresponding increase of the time between them), but all of them share the same qualitative behavior: a slow decrease for most of its temporal range ending with a much sharper decay.

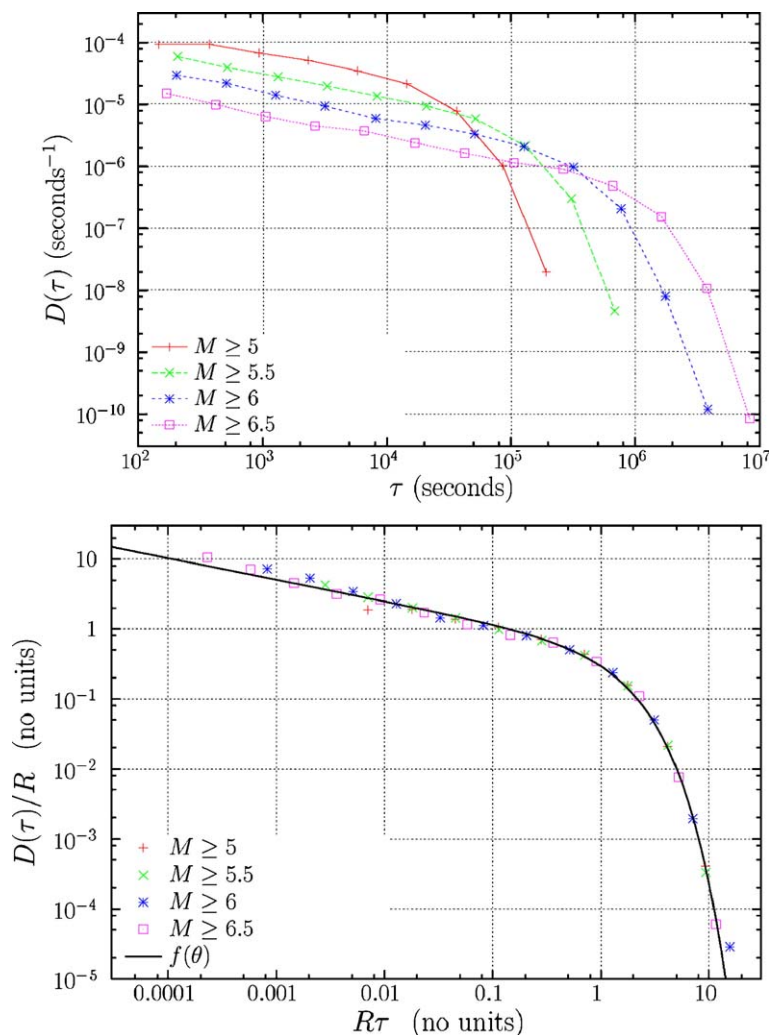


Fig. 2. (top) Probability densities of recurrence times of earthquakes worldwide during 30 years (from 1973 to 2002), with $M \geq 5$, $M \geq 5.5$, $M \geq 6$, and $M \geq 6.5$. (bottom) Same as before but with rescaled axis, $D(\tau)/R$ versus $R\tau$. The collapse of all the densities onto a single curve shows the shape of the scaling function f . Note that the deficit of occurrences for recurrence times smaller than 500 s in the $M \geq 5$ case could be an indication of the incompleteness of the catalog at the shortest time scale.

Remarkably, when the x -axis is rescaled as $R_w\tau$ and the y -axis as $D_w(\tau)/R_w$, all the different densities collapse onto a single curve, see Fig. 2 (bottom). This data collapse allows to write the recurrence-time probability density in the form of a scaling law,

$$D_w(\tau) = R_w f(R_w\tau), \quad (2)$$

where the scaling function f is the same for all the different M_c 's. Then, the above mentioned qualitative similarity between the recurrence-time distributions becomes clearly quantitative.

Analyzing a great number of windows with stationary seismicity, for worldwide as well as for local seismicity, for spatial scales as small as 20 km, and with magnitude threshold ranging from 1.5 to 7.5 (which is a factor 10^9 in the variation of the minimum energy), it was found that the scaling law (2) is of general validity for stationary seismicity, with the same scaling function f in any case (Corral, 2004a, 2005b). This means that one single scaling function can describe reasonably well the recurrence-time distribution in any stationary region, with small differences in each case. In this way, f can be considered a universal scaling function, reflecting the self-similarity in space and magnitude of the stationary temporal occurrence of earthquakes. Despite the large variability associated with the earthquake phenomenon ("there are no two identical earthquakes"), it is surprising that the timing of earthquakes is governed by a unique simple law.

The functional form of f turns out to be close to a generalized gamma distribution,

$$f(\theta) = \frac{C|\delta|}{a^\gamma \Gamma(\gamma/\delta)} \frac{1}{\theta^{1-\gamma}} e^{-(\theta/a)^\delta}, \quad \text{for } \theta > \theta_{\min} \geq 0, \quad (3)$$

where the variable $\theta \equiv R_w\tau$ constitutes a dimensionless recurrence time; the parameter γ turns out to be around 0.7, whereas δ is very close to 1 (strictly, $\delta \equiv 1$ for the usual gamma distribution). This means that f decays as a power law $\propto 1/\theta^{0.3}$ for $\theta < 1$ and the decay is accelerated by the exponential factor for $\theta > 1$. The constant a is just a dimensionless scale parameter and its value is around 1.4, Γ is the usual gamma function, which normalizes the distribution, and C is a correction to normalization due to the fact that this description may not be valid for very short times, $\theta < \theta_{\min}$. Let us remark that the modeling of the scaling function by means of a gamma distribution does not make the scaling law (2) depend on the gamma modeling; the validity of the scaling law is much more general and independent on any model of seismic occurrence, in particular, no renewal model is implicit in our approach.

The slow decay provided by the decreasing power law is a signature of *clustering*, which means that the events tend to be closer to each other in the short time scale, in comparison with a Poisson process (given by f exponential, i.e., $\gamma = \alpha = \delta = 1$). In fact, clustering is more clearly identified by looking at the hazard rate function (containing the same information than the probability density), which shows a high value for short recurrence times and decreases when the recurrence time increases; so, in some sense earthquakes tend to attract each other (Corral, 2005d). This leads to a somewhat paradoxical consequence for the time one has to wait for the next earthquake: it turns out that the longer one has been waiting for an earthquake, the longer one has still to wait. Note that the clustering we are dealing with takes place for stationary seismicity, at variance with the clustering associated to aftershock sequences due mainly to the larger seismic rate close to the mainshock. Large rates tend to concentrate events, whereas a subsequent decrease of rate dilutes the events. This mechanism does not operate in stationary seismicity, where the rate of the process stays constant (despite the presence of local aftershock sequences originated by small events, which are intertwined with the seismicity of the rest of the area to give rise to the observed stationary behavior).

Nevertheless, it is interesting to point out that the scaling law (2) is valid beyond the case of stationary seismicity. For this purpose we have considered Southern-California seismicity for the years 1984–2001, using the SCSN catalog (Southern California Seismographic Network) available at the Southern California Earthquake Data Center, <http://www.scecdc.sceec.org/ftp/catalogs/SCSN>. It is well known that several important earthquakes occurred in Southern California during the selected period, generating increased seismic activity and breaking the stationary behavior of seismicity (see Fig. 1); this leads to recurrence-time distributions dominated by aftershock sequences. Fig. 3 displays these recurrence-time densities for different values of M_c , rescaled by the mean rate R_w (which now is not stationary). The nearly perfect data collapse implies the validity of the scaling law (2), with a scaling function f still given by Eq. (3), although with a decreasing power-law exponent $1-\gamma$ closer to 1 ($1-\gamma \simeq 0.8$), clearly different than the one for stationary seismicity, $1-\gamma \simeq 0.3$. (Surprisingly, it was shown in Corral (2004a) that when time is nonlinearly transformed to obtain a stationary process, the universal scaling function corresponding to stationary seismicity is recovered; this can be done for a single aftershock sequence rescaling not by R_w but by the instantaneous Omori rate, $r_w(t) \propto 1/t^p$).

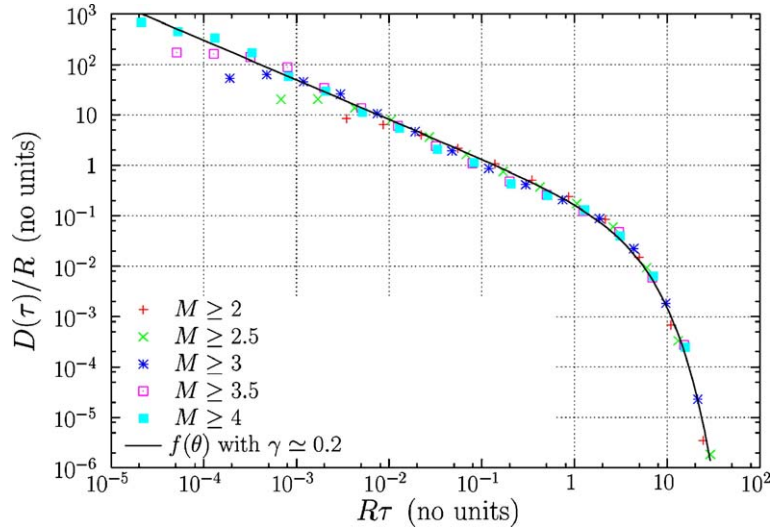


Fig. 3. Rescaled probability densities of recurrence times of earthquakes in Southern California, $(123^\circ\text{W}, 113^\circ\text{W}) \times (30^\circ\text{N}, 40^\circ\text{N})$, for the years 1984–2001, using different magnitude thresholds, from 2 to 4. The good data collapse indicates that a scaling law is valid, although the scaling function is different than in the stationary case. In concrete, the decay of the power law is faster, with an exponent $1 - \gamma \approx 0.8$. (The distributions only show recurrence times greater than 10 s.)

In the case in which the spatial window is kept fixed and only M_c is changed, the seismic rate R_w can be related with the magnitude threshold by means of the Gutenberg–Richter law,

$$R_w = R_0 10^{-bM_c}, \quad (4)$$

where b is the well-known b -value of the Gutenberg–Richter law (usually close to 1) and R_0 is an extrapolation of the seismic rate to $M_c=0$, depending only on the spatial part of the window w . Therefore, the scaling law (2) can be written as

$$D_w(\tau) = R_0 10^{-bM_c} f(R_0 10^{-bM_c} \tau), \quad (5)$$

and including here the gamma fit (with $\delta=1$) we get

$$D_w(\tau) = \frac{C}{a^\gamma \Gamma(\gamma)} \frac{R_0^\gamma 10^{-\gamma b M_c}}{\tau^{1-\gamma}} \exp(-R_0 10^{-bM_c} \tau/a), \quad (6)$$

which is a complicated formula for a simple idea.

In addition, it is remarkable the similarity between our scaling law (2) with Eq. (4) and the scaling of angular correlation functions in the distribution of galaxies (Jones et al., 2004). Also, it has been noted that the scaling relation (2) corresponds to the Cox–Oakes accelerated-life model, studied in the context of medical statistics (Cox and Oakes, 1984; German, 2005).

An important consequence of the scaling law (2) is that it implies that the Gutenberg–Richter law is not only valid over long periods of time (in the form

$R_w = R_0 10^{-bM_c}$) but it is also fulfilled continuously in time, provided the times are compared in the appropriate way. Indeed, selecting $\tau_w = A/R_w$, where A is an arbitrary fixed constant and R_w changes as a function of M_c , we get $D_w(\tau_w) = R_w f(A) = f(A) R_0 10^{-bM_c}$. In other words, selecting two different thresholds, M_u and M_v , the Gutenberg–Richter law is obtained in the form

$$\frac{D_v(\tau_v)}{D_u(\tau_u)} = 10^{-b(M_v - M_u)},$$

when times τ_u and τ_v are compared for each threshold, these given by

$$\frac{\tau_v}{\tau_u} = 10^{b(M_v - M_u)}.$$

To make it more concrete, if, for $M \geq 5$ we count the number of events coming after a recurrence time $\tau = 10^4$ s ($A \approx 0.5$ in the case of Fig. 2), and compare with the number of events with $M \geq 6$ coming after $\tau = 10^5$ s, these numbers are given by the Gutenberg–Richter relation. This constitutes somehow an analogous of the well-known law of corresponding states in condensed-matter physics (see for instance Guggenheim, 1966 or Lee et al., 1973). Two pairs of consecutive earthquakes in different magnitude windows would be in “corresponding states” if their rescaled recurrence times are the same ($\theta = R_u \tau_u = R_v \tau_v$).

Interestingly, the scaling relation (2) can be understood as arising from a renormalization-group transformation analogous to those studied in the context of

critical phenomena (Corral, 2005c). Let us investigate deeper the meaning of the scaling analysis we have performed and its relation with a renormalization-group transformation. Fig. 4 displays the magnitude M versus the occurrence time t of all worldwide earthquakes with $M \geq M_c$ for different periods of time and M_c values. Fig. 4 (top) is for earthquakes larger than (or equal to) $M_c=5$ for the year 1990. If we rise the threshold up to $M_c=6$ we get the results shown in Fig. 4 (medium). Obviously, as there are less earthquakes in this case, the distribution of recurrence times (time interval between consecutive “spikes” in the plot) becomes broader with respect to the

previous case. The rising of the threshold can be viewed as a mathematical transformation of the seismicity point process, which is referred to as *thinning* in the context of stochastic processes (Daley and Vere-Jones, 1988) and is also equivalent to the common decimation performed for spin systems in renormalization-group transformations (Kadanoff, 2000; McComb, 2004; Christensen and Moloney, 2005). Fig. 4 (bottom) shows the same as Fig. 4 (medium) but for ten years, 1990–1999, and represents a scale transformation of seismicity (also as in the renormalization group), contracting the time axis by a factor 10 (this factor comes from the Gutenberg–

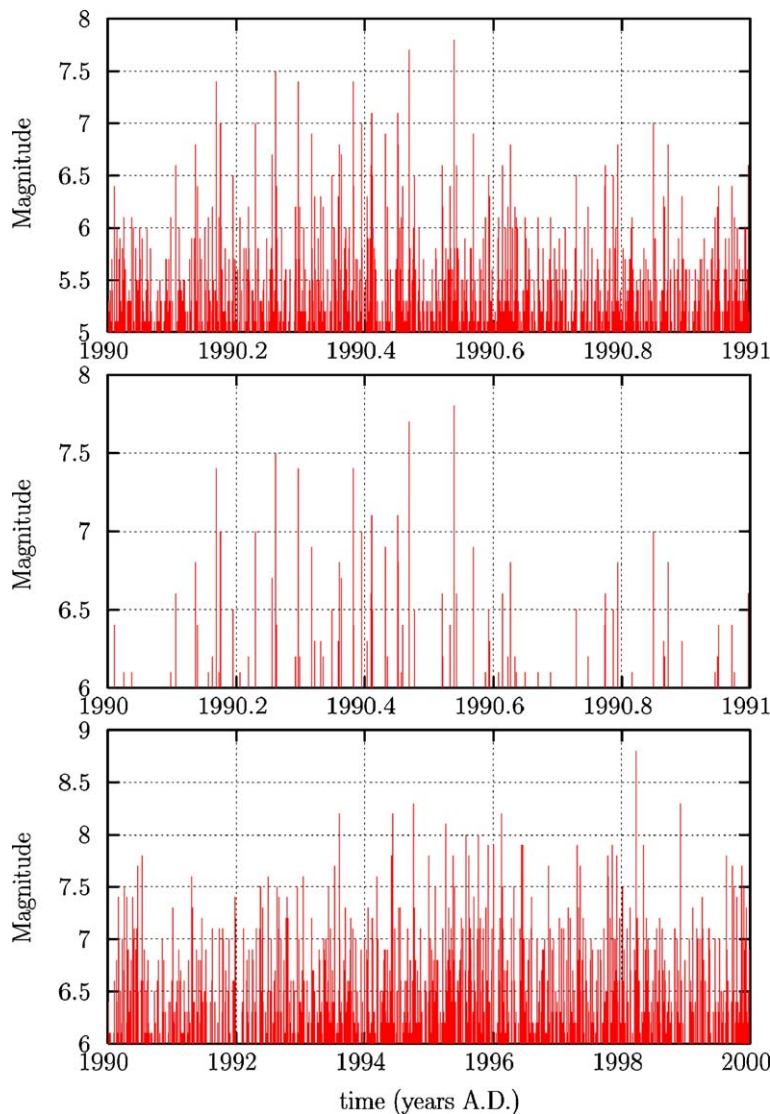


Fig. 4. Magnitude of worldwide earthquakes as a function of time of occurrence, for different time and magnitude windows to illustrate the analogies with renormalization-group transformations. (top) Earthquakes with $M \geq 5$ during year 1990. (medium) Same as before but restricted to $M \geq 6$ (this is in fact a decimation of events). (bottom) Earthquakes with $M \geq 6$ expanded to 10 years, 1990–1999 (this is equivalent to a contraction of time axis by a factor 10). Visual comparison between the top and bottom figures shows that both figures “look the same.”

Richter law with $b=1$, as it is the ratio of events with $M \geq 5$ to $M \geq 6$, i.e., $R_{w|M_c=5}/R_{w|M_c=6} \approx 10$). The similarity between Fig. 4 (top) and 4 (bottom) is apparent, and is confirmed when the probability densities of the corresponding recurrence times are calculated and rescaled following Eq. (2) (see again Fig. 2 (bottom)).

In this way, the existence of the scaling law (2) implies the invariance of the recurrence-time distribution under renormalization-group transformations, i.e., the rescaled distribution is a fixed point of the transformation, constituting then a very special function (as in general, a thinning procedure implies a change not only on the scale of the distribution but also on its shape). One can realize that there exists a trivial fixed-point solution to this problem, where the recurrence times are given by a Poisson process (i.e., independent identically distributed exponential times) and the magnitudes are generated with total independence as well; this constitutes what is called a marked Poisson process.

Moreover, it is not difficult to show that the marked Poisson process is the only process without correlations which is invariant under our renormalization-group transformation (Corral, 2005c). But even further, the Poisson process is not only a fixed point of the transformation, but a stable one, i.e., an attractor: when events are randomly removed from a process (random thinning) the resulting process converges to a Poisson process, under certain conditions (Daley and Vere-Jones, 1988).

The fact that for real seismicity the scaling function f is not an exponential tells us that our renormalization-group transformation is not performing a random thinning; this means that the magnitudes are not assigned independently on the rest of the process and therefore there exists correlations in seismicity. This of course has been pointed out before, see for instance Kagan and Jackson (1991), but let us stress that correlations are not only fundamental for the existence of the scaling law (2), but are intimately related to it; the knowledge of correlations determine the precise form of the recurrence-time distribution. In any case, it is striking that seismicity is precisely at the fixed point of a renormalization-group transformation.

At this point it is worth clarifying that the term “correlations” is generally used with two different meanings: First, when the recurrence-time distribution is not an exponential (and the time process is not Poisson), for example in a renewal process. This in fact implies the existence of a memory in the process, as events do not occur independently at any time, but remember the occurrence of the last event. And second,

when the recurrence time (whatever its distribution) depends on the history of the process, in particular the occurrence time and magnitude of previous events (of course, this is another form of memory in the process); these are the correlations we are talking about.

There are other arguments that make the existence of the scaling law (2) striking. For example, one can argue that the seismicity of large areas consists on the superposition of many more or less independent processes, and it is known that the pooled output of independent processes converges to a Poisson process (Cox and Lewis, 1966). But Fig. 2 (bottom) demonstrates for the largest possible scale that this is not the case: worldwide seismicity shows the limited applicability of the theorems that guarantee convergence to a Poisson process. (On the other hand, one could argue that the sum does not have enough samples of unit independent processes to ensure the convergence.)

In the same line, Molchan (2005) has shown that if the scaling law (2) is fulfilled in two regions separately, and also in the “superregion” formed by the two regions, then the scaling function f can only be exponential, provided that seismicity in both regions is independent on each other. As in practice one finds that the scaling function clearly departs from an exponential, one may conclude that seismicity is not independent in the regions considered in our previous analysis. Therefore, the interactions between the regions are important enough to make the boundary effects not negligible; this could be related with long-range earthquake triggering (Huc and Main, 2003).

Finally, it is interesting to mention the results of Hainzl et al. (2006). Following theoretical calculations of Molchan (2005), who assumed that seismicity is formed by Poisson mainshocks generating sequences of aftershocks, Hainzl et al. use the parameters of the scaling function f to evaluate the ratio between the number of mainshocks and the total number of events, given by $1/a \approx \gamma$, for $\delta=1$. Analyzing a number of regions these authors find different proportions, but because most of the regions analyzed were not stationary their results are not in contradiction with the universality of the proposed scaling law (2). Rather, this suggests that stationary seismicity may keep a constant proportion of mainshocks and aftershocks (in case that both entities could be unambiguously defined).

3. Conditional probability densities and correlations

In the previous section we have seen how the fulfillment of a scaling law for the recurrence-time distributions with a non-exponential scaling function is

an indication of the existence of important correlations in seismicity. Once a spatial area and a magnitude range have been selected for study, we can consider the process as taking place in the magnitude–time domain (intentionally disregarding the spatial degrees of freedom, which will be the subject of future research, see also [Davidsen and Paczuski \(2005\)](#)). Given an event i , its magnitude M_i may depend on the previous magnitudes M_j with $j < i$ as well as on the previous occurrence times t_j , or equivalently, on the previous recurrence times, $\tau_j = t_j - t_{j-1}$, with $j \leq i$ now. In its turn, the recurrence time to the next event, τ_{i+1} may depend on M_i , τ_i , M_{i-1} , τ_{i-1} , etc. (see [Fig. 5](#) for a visual explanation). This dependence is of course statistical and the most direct way to measure it is by means of the conditional probability density. In the case of recurrence times, the conditional density is defined as

$$D(\tau|X) \equiv \frac{\text{Prob}[\tau \leq \text{rec. time} < \tau + d\tau|X]}{d\tau}, \quad (7)$$

where the symbol $|$ states that the probability is only considered for the cases in which the event or condition X holds, X being a range of values of previous recurrence times, or magnitudes, or whatever. In an analogous way one can define the probability density of the magnitude conditioned to X .

It is a fundamental result of probability theory that if a random variable, for instance the recurrence time, is statistically independent of an event X , then, the distribution of the variable conditioned to X is identical to the unconditional distribution, i.e., $D(\tau|X) = D(\tau)$ (in fact, this is the definition of independence). On the other side, if both distributions are different, $D(\tau|X) \neq D(\tau)$, there is a dependence between τ and X and some correlation coefficient different than zero may be

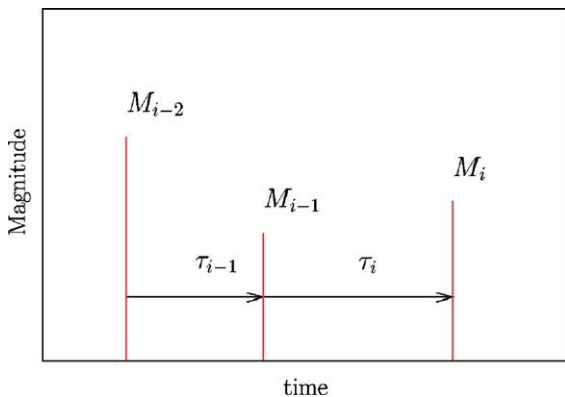


Fig. 5. Notation used to describe the seismicity point process in time and magnitude, where only events $i-2$, $i-1$, and i are shown.

defined (in general the correlation coefficient must be nonlinear; as it is well known dependent variables may yield a zero value for the linear correlation coefficient).

In this section we will measure the dependence of seismicity on its history by means of conditional distributions of recurrence times and magnitudes, concentrating mainly on the first step backwards, i.e., we will study the dependence of τ_i on M_{i-1} and τ_{i-1} , and the dependence of M_i on τ_i and M_{i-1} . These dependences will be given by the distributions $D(\tau_i|M_{i-1} \geq M'_c)$, $D(\tau_i|\tau_a \leq \tau_{i-1} < \tau_b)$, $D(M_i|\tau_a \leq \tau_i < \tau_b)$, and $D(M_i|M_{i-1} \geq M'_c)$, respectively (note that now M_i , τ_i , M_{i-1} , etc. do not denote the values of τ and M for a particular event, they refer to all events i and the corresponding preceding events $i-1$). Note also that when the condition is imposed on the magnitude, we denote its threshold as M'_c , in order to distinguish it from the generic condition, $M \geq M_c$ (with $M_c \geq M'_c$), which always holds. The application of our method to more steps backwards ($i-2$, $i-3$, etc.) is straightforward.

Nevertheless, instead of analyzing $D(M_i|\tau_a \leq \tau_i < \tau_b)$ we have found clearer to study $D(\tau_i|M_i \geq M'_c)$; although causality is not preserved in this case (as the event of magnitude M_i happens after the time τ_i), from the point of view of probability theory the equivalence is complete, i.e., if we find $D(\tau_i|M_i \geq M'_c) \neq D(\tau_i)$, then τ_i and M_i are correlated.

The imposition of a condition X reduces the statistics and increases the fluctuations of the measured distributions; for this reason we will associate an error or uncertainty to each bin of the densities. Let us consider the case of the recurrence time τ , with no extra condition. The number of data n in a particular interval $[\tau, \tau + d\tau)$ can be considered as given by the binomial distribution if the total number of data N is much larger than the range of the correlations in the data; therefore, its standard deviation will be, as it is well known, $\sigma_n = \sqrt{Np(1-p)}$. As $D(\tau) = n/(Nd\tau)$, the error in $D(\tau)$ will be

$$\sigma_D = \frac{1}{d\tau} \sqrt{\frac{p(1-p)}{N}},$$

where the value of p can be estimated simply as $p = n/N = D(\tau)d\tau$. The error in conditional distributions is obtained in the same way, obviously.

The data analyzed in this section comes from the two catalogs referred to previously. For the worldwide NEIC-PDE catalog we consider the period 1973–2002, which can be considered stationary with a rate of about $4.2 \cdot 10^{-(M_c-5)}$ earthquakes per day. As the fulfillment of the Gutenberg–Richter law points out, this catalog is reasonably complete for $M \geq 5$ (see below,

nevertheless); this yields 46,055 events in the considered magnitude–time window. In contrast, Southern-California seismicity is not stationary in general, but some stationary periods exist in between large earthquakes, see Fig. 1. The longest stationary period for the last years spans from 1994 to 1999, in particular we have considered the period from May 1, 1994 to July 31, 1999, for which the number of earthquakes with $M \geq 1.5$ in the spatial area $(123^\circ\text{W}, 113^\circ\text{W}) \times (30^\circ\text{N}, 40^\circ\text{N})$ is 46,576; this gives 24.3 earthquakes per day. Other stationary periods have also been analyzed and will be detailed below; it is remarkable that all of them share more or less the same rate of occurrence.

3.1. Correlations between recurrence times

Let us start our study of correlations in seismicity with the temporal sequence of occurrences. Using the tools described above we obtain the conditional distributions $D(\tau_i | \tau_a \leq \tau_{i-1} < \tau_b)$; in particular we distinguish two cases; first, the recurrence-time density conditioned to short preceding recurrence times, $D(\tau_i | \tau_{i-1} < \tau_b)$, where τ_b is smaller than the mean $\langle \tau \rangle$ of the unconditional distribution, $D(\tau)$; and second, the recurrence-time density conditioned to long preceding recurrences, $D(\tau_i | \tau_{i-1} \geq \tau_a)$, with τ_a larger than $\langle \tau \rangle$ (or at least larger than the τ_b in the first case).

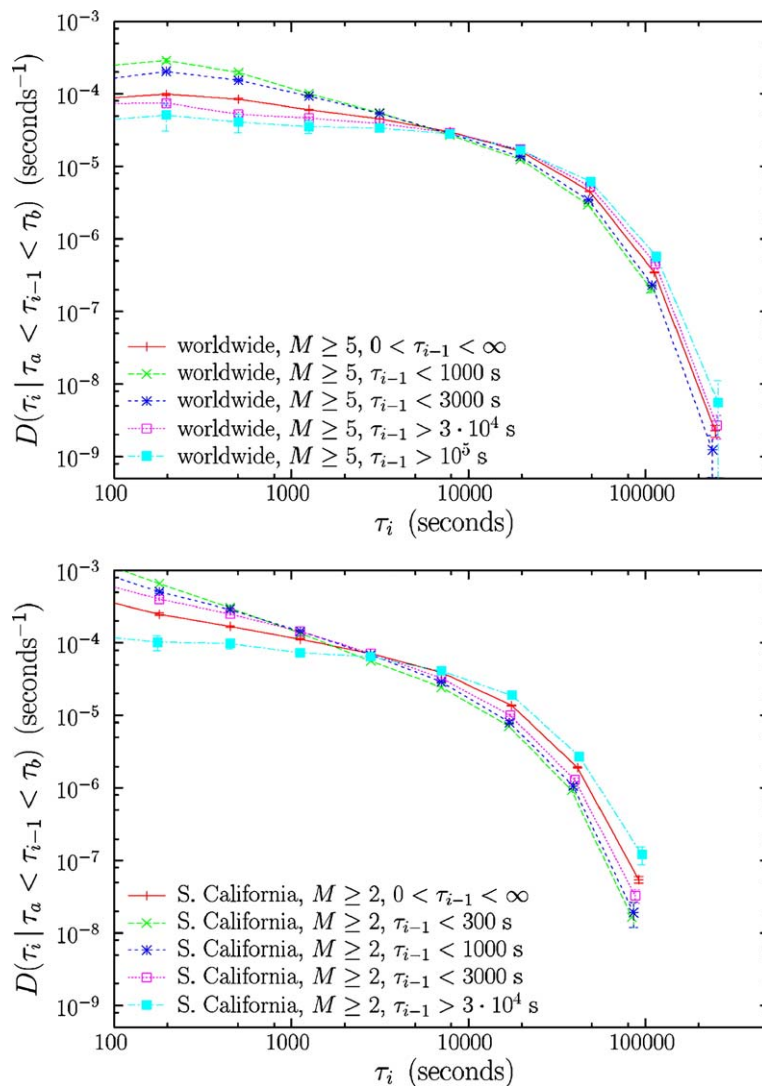


Fig. 6. Probability densities of earthquake recurrence times τ_i conditioned to the value of the preceding recurrence time, τ_{i-1} . Different ranges for τ_{i-1} lead to a systematic variation in the distributions, clearly beyond the uncertainty limits given by the error bars, leading to the conclusion that τ_{i-1} and τ_i are positively correlated. (top) Results for NEIC worldwide catalog using $M \geq 5$. (bottom) Southern-California catalog for the stationary period reported in the text, with $M \geq 2$.

The results, both for worldwide seismicity and for Southern-California stationary seismicity, turn out to be equivalent (see Fig. 6). In any case, a gamma distribution is able to fit the data, with different parameters though, depending on the condition. For short τ_{i-1} (given by short τ_b), a relative increase in the number of short τ_i and a decrease of long τ_i is obtained, which leads to a more steep power-law decay. In the opposite case, long τ_{i-1} imply a decrease in the number of short τ_i and an increase in the longer ones, in such a way that the power-law part of the gamma distribution becomes flatter, tending to an

exponential distribution for the longest τ_{i-1} . This means that short τ_{i-1} imply an average reduction on τ_i and the opposite for long τ_{i-1} and then both variables are positively correlated.

Note that this behavior corresponds to a clustering of events, in which short recurrence times tend to be close to each other, forming clusters of events, while longer times tend also to be next each other. This clustering effect is different from the clustering reported previously, associated to the non-exponential nature of the recurrence-time distribution (Corral, 2004a, 2005d), and also from the usual clustering of aftershock sequences.

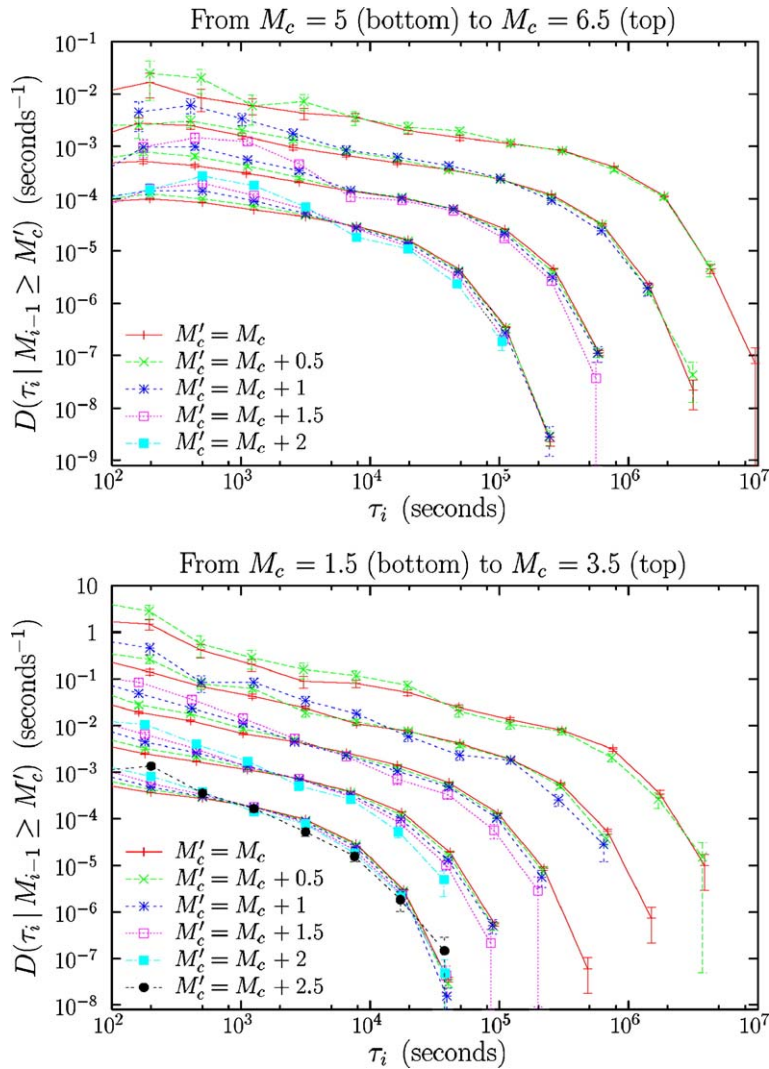


Fig. 7. Probability densities of earthquake recurrence times τ_i conditioned to the value of the preceding magnitude, M_{i-1} . The differences in the distributions are significant (except for the top curves for which the statistics was poorer), showing that the larger M_{i-1} , the shorter τ_i , implying a negative correlation between both variables. (top) Results for the NEIC worldwide catalog, using different thresholds M_c for M_i and multiplicative factors in the density for clarity sake: $M_i \geq 5$, factor 1; $M_i \geq 5.5$, factor 10; $M_i \geq 6$, factor 100; $M_i \geq 6.5$, factor 1000; the threshold M'_c for M_{i-1} ranges from $M'_c = M_c$ (unconditioned distribution) to $M'_c = 7$. (bottom) Results for Southern California stationary seismicity, using $M_i \geq 1.5$, multiplicative factor 1; $M_i \geq 2$, factor 10; $M_i \geq 2.5$, factor 100; $M_i \geq 3$, factor 1000; $M_i \geq 3.5$, factor 1000; the threshold M'_c ranges from M_c to 4.

The findings reported here are in agreement with Livina et al. (2005), who studied the same problem for Southern California, but without separating stationary periods from the rest. These authors explain this behavior in terms of the persistence of the recurrence time, which is a concept equivalent to clustering. The results could be similar to the long-term persistence observed in climate records, but this has not been studied up to now. The linear correlation coefficient in this case was calculated by Helmstetter and Sornette (2004), turning out to be 0.43, clearly supporting the positive correlation.

We have extended the study to more steps backwards in time, to be concrete, we have obtained from the catalogs the distributions $D(\tau_i | \tau_a \leq \tau_j < \tau_b)$, using different values of τ_j , with $j=i-2, i-3$, etc., up to $j=i-10$. In all cases there are significative differences between the conditional distributions and the unconditional distribution, implying that the temporal correlations extend to up to 10 events, or, in other words, that the recurrence time is dependent on its own history.

3.2. Correlations between recurrence time and magnitude

When, in addition to the recurrence times, the magnitudes are taken into account, we can study the dependences between these variables. We consider first how the magnitude of one event influences the

recurrence time to the next; in this way we measure $D(\tau_i | M_{i-1} \geq M'_c)$. The results for the case of worldwide seismicity and for Southern California (in the stationary case) are again similar and show a clear correlation between M_{i-1} and τ_i (Corral, 2005a).

Indeed, Fig. 7 shows how larger values of the preceding magnitudes, given by $M_{i-1} \geq M'_c$, lead to a relative increase in the number of short τ_i and a decrease in long τ_i ; this means that M_{i-1} and τ_i are anticorrelated. For the cases for which the statistics is better, the densities show the behavior typical of the gamma distribution, with a power law that becomes steeper for larger M'_c ; the different values of the exponent are given in the caption of Fig. 8. In a previous work (Corral, 2005c) we showed that the conditional distributions we are dealing with could be related with the unconditional one by means of the scaling relation

$$D(\tau_i | M_{i-1} \geq M'_c) \simeq \Lambda D(\tau_i) = R_w \Lambda f(R_w \Lambda \tau_i),$$

where R_w is the rate of the process for $M \geq M_c$ and Λ is a factor depending on the difference $M'_c - M_c$ with $\Lambda=1$ for $M'_c - M_c=0$. This is in disagreement with the results reported here; indeed, the above relation is only approximately valid for small $M'_c - M_c$, but for large values we have shown clearly that the shape of the distribution changes, and not only its mean, as we previously supposed. Remarkably, this change can be described by a scaling law for which the scaling function depends now on the difference between the threshold

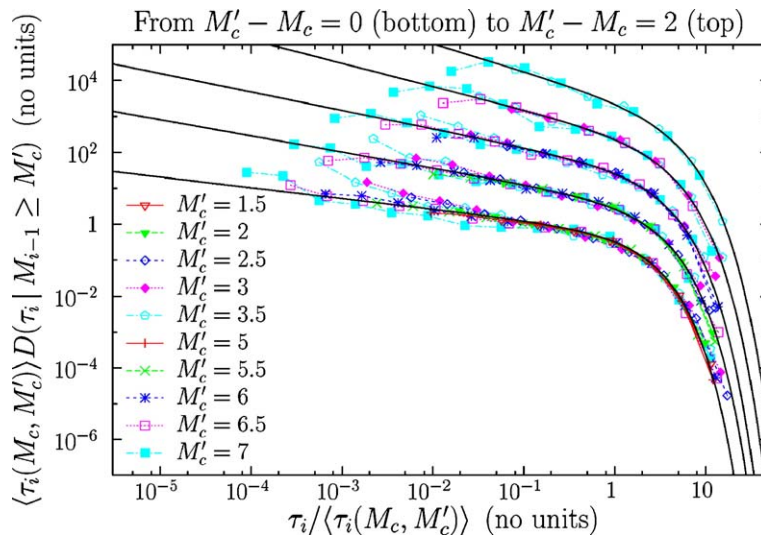


Fig. 8. Same as previous figure but rescaling the axis and separating the densities by the values of $M'_c - M_c$ (rather than by M_c). From bottom to top, $M'_c - M_c$ takes the values 0, 0.5, 1, 1.5, and 2, and the densities in each case have been multiplied by 1, 10, 100, 1000, and 10^4 , respectively. M'_c ranges from 1.5 to 3.5 for Southern California, and from 5 to 7 for the worldwide data. The figure shows how a scaling law is valid for $D(\tau_i | M_{i-1} \geq M'_c)$ but with a scaling function that depends on $\Delta M_c \equiv M'_c - M_c$. The straight lines correspond to the fits of $f_{\Delta M_c}$ for different ΔM_c , yielding gamma distributions with parameter $\gamma \simeq 0.70, 0.55, 0.48, 0.35$, and 0.23 , for $\Delta M_c = 0, 0.5, 1, 1.5$, and 2 , respectively.

magnitude for the $i-1$ event, M'_c , and the threshold for i , M_c , i.e.,

$$D(\tau_i | M_{i-1} \geq M'_c) = R_w A f_{\Delta M_c} (R_w A \tau_i),$$

with $\Delta M_c = M'_c - M_c$, $A = R_w^{-1} \langle \tau_i (M_c, M'_c) \rangle^{-1}$, and $\langle \tau_i (M_c, M'_c) \rangle$ the mean of the distribution $D(\tau_i | M_{i-1} \geq M'_c)$. Fig. 8 illustrates this new scaling law, setting clearly that there is not just a similarity between the distributions for worldwide and Southern-California seismicity, but an identical behavior described by the same scaling functions with the same exponents if ΔM_c is the same.

Further, as in the previous subsection, we have gone several more steps backwards in time, measuring conditional distributions up to $D(\tau_i | M_{i-10} \geq M'_c)$ and again we find significant dependence between τ_i and the previous magnitudes, up to M_{i-10} .

Next, we can consider how the recurrence time to one event τ_i influences its magnitude M_i , or, equivalently, how the recurrence time to one event depends on the magnitude of this event. For this purpose, we measure $D(\tau_i | M_i \geq M'_c)$ and the results are shown in Fig. 9. From there it is clear that in most of their range the distributions are nearly identical, and when some

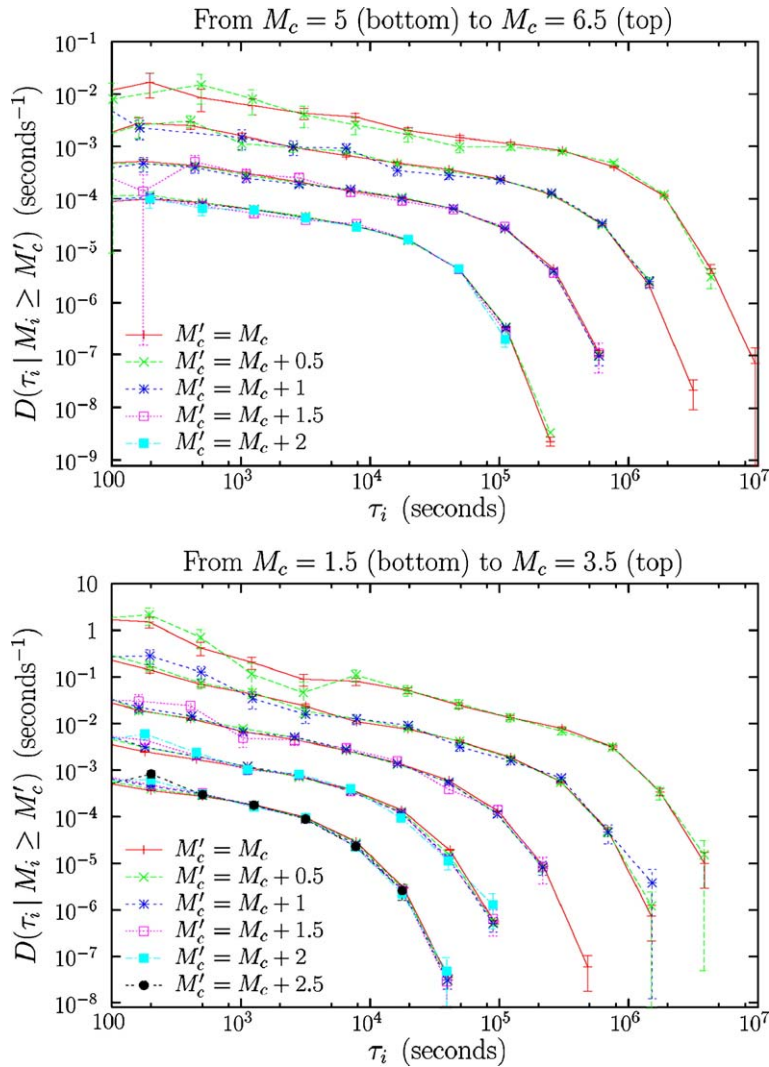


Fig. 9. Probability densities of earthquake recurrence times τ_i conditioned to the value of the magnitude after the recurrence, M_i . Contrary to the previous figure, now the differences in the distributions are negligible (except for very short times in Southern California, see text). So we can consider τ_i and M_i statistically independent. (top) Results for the NEIC worldwide catalog, using different thresholds M_c for M_{i-1} and multiplicative factors in the density (as in the previous plot): $M_{i-1} \geq 5$, factor 1; $M_{i-1} \geq 5.5$, factor 10; $M_{i-1} \geq 6$, factor 100; $M_{i-1} \geq 6.5$, factor 1000; the threshold M'_c for M_i ranges from $M'_c = M_c$ (unconditioned distribution) to $M'_c = 7$. (bottom) Results for the Southern California stationary period, using $M_{i-1} \geq 1.5$, factor 1; $M_{i-1} \geq 2$, factor 10; $M_{i-1} \geq 2.5$, factor 100; $M_{i-1} \geq 3$, factor 1000; $M_{i-1} \geq 3.5$, factor 104; the threshold M'_c ranges from M_c to 4.

difference is present this is inside the uncertainty given by the error bars. However, there is one exception: in California, very short times, $\tau_i \approx 200$ s, seem to be favored by larger magnitudes, or, in other words, short times lead to larger events. What we are seeing in this case may be the foreshocks of small events, which are restricted to these very short times, but we cannot exclude that this is an artifact due to catalog incompleteness at short time scales. With the exception of this, which has a very limited influence, we can consider the recurrence time and the magnitude after it as independent from a statistical point of view, with our present resolution (it might be that a very weak dependence is hidden in the error bars of the distributions).

In addition, the extension of the conditional distributions to the future, measuring $D(\tau_i | M_j \geq M'_c)$ with $j > i$, leads to similar results, and from here we can conclude the independence of the magnitude with the sequence of previous recurrence times, at least with our present statistics and resolution.

3.3. Correlations between magnitudes

Finally, we study the correlations between consecutive magnitudes, M_{i-1} and M_i , by means of the distribution $D(M_i | M_{i-1} \geq M'_c)$. As the analysis of the 1994–1999 period for Southern California did not provide enough statistics for the largest events we

considered a set of stationary periods, these being: Jan 1, 1984–Oct 15, 1984; Oct 15, 1986–Oct 15, 1987; Jan 1, 1988–Mar 15, 1992; Mar 15, 1994–Sep 15, 1999; and Jul 1, 2000–Jul 1, 2001; all of them displayed in Fig. 1.

Again we find that both worldwide seismicity and stationary Southern-California seismicity share the same properties, but with a divergence for short times. Fig. 10 shows the distributions corresponding to the two regions, and whereas for the worldwide case the differences in the distributions for different M'_c are compatible with their error bars (see bottom-right set of curves), for the California case there is a systematic deviation, implying a possible correlation (bottom-left set).

In order to find the origin of this discrepancy we include an extra condition, which is to restrict the events to the case of large enough recurrence times, so we impose $\tau_i \geq 30$ min. In this case, the differences in Californian distributions become insignificant (top-left curve), which means that the significant correlations between consecutive magnitudes are restricted to short recurrence times (Corral, 2005a). Therefore, we conclude that the Gutenberg–Richter law is valid independently of the value of the preceding magnitude, provided that short times are not considered. This is in agreement with the usual assumption in the ETAS model, in which magnitudes are generated from the Gutenberg–Richter distribution with total independence

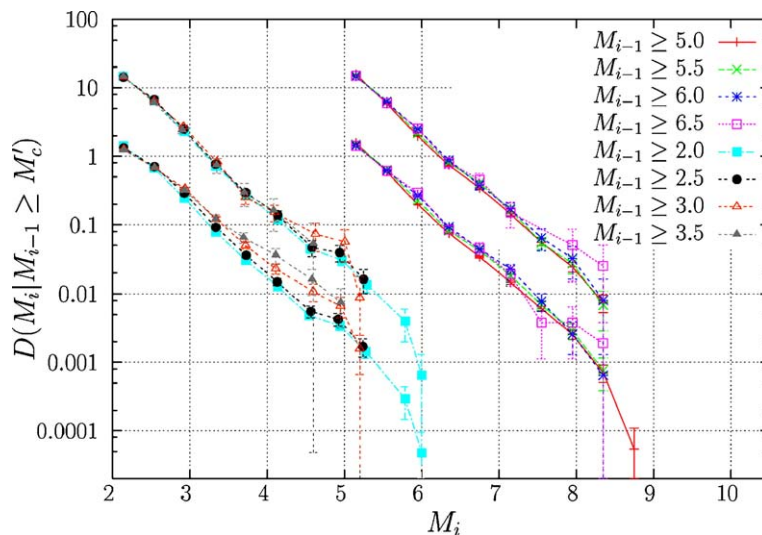


Fig. 10. Probability densities of earthquake magnitudes M_i conditioned to the magnitude of the immediately previous event, M_{i-1} , both for worldwide seismicity and Southern-California stationary periods (see text). Different thresholds M'_c for M_{i-1} have been used: for the worldwide case (right) M'_c ranges from 5 to 6.5, whereas $M_i \geq M'_c$ with $M'_c = 5$ always. For the California case (left), M'_c ranges from 2 to 3.5, with $M'_c = 2$ always. Note that the case $M'_c = M_c$ gives the unconditioned distribution $D(M)$. The top set of curves corresponds to the extra condition $\tau_i \geq 30$ min, for which the distributions have been multiplied by 10 for clarity sake. In this case the differences turn out to be not significant and therefore we cannot rule out the statistical independence of consecutive magnitudes.

of the rest of the process (Ogata, 1999; Helmstetter and Sornette, 2002). Of course, this independence is established within the errors associated to our finite sample. It could be that the dependence between the magnitudes is weak enough for that the changes in distribution are not larger than the uncertainty. With our analysis, only a much larger data set could unmask this hypothetical dependence.

The deviations for short times may be an artifact due to the incompleteness of earthquake catalogs at short time scales, for which small events are not recorded. Helmstetter et al. (2006) propose a formula for the magnitude of completeness in Southern California as a function of the time since a mainshock and its magnitude; applying it to our stationary periods, for which the larger earthquakes have magnitudes ranging from 5 to 6, we obtain that a time of about 5 h is necessary in order that the magnitude of completeness reaches a value below 2 after a mainshock of magnitude 6. After mainshocks of magnitude 5.5 and 5 this time reduces to about 1 h and 15 min, extrapolating the results of Helmstetter et al. In any case, it is perfectly possible that large mainshocks (not necessarily the preceding event) induce the loss of small events in the record and are the responsible of the deviations from the Gutenberg–Richter law at small magnitudes for short times. If an additional physical mechanism is behind this behavior, this is a question that cannot be answered with this kind of analysis.

As in the previous subsection, we have performed measurements of the conditional distributions for worldwide seismicity involving different M_i and M_j , separated up to 10 events, with no significant variations in the distributions, as expected. These results confirm the independence of the magnitude M_i with its own history.

4. Conclusions

We have argued how a scaling law for earthquake recurrence-time distributions can be understood as an invariance under a transformation analogous to those of the renormalization group (as in the critical phenomena found in condensed matter physics). Such invariance can only exist due to an interplay between the probability distribution and the correlations of the seismicity process. We have seen how the correlations are reflected on conditional probability densities, whose study for stationary seismicity (worldwide or local) leads to the results summarized here:

- The bigger the size of an earthquake, the shortest the time till next, because of the anticorrelation between

magnitudes and forward recurrence times. Note that this result is not trivial, as we are dealing with stationary seismicity.

- The shortest the time to get an earthquake, the shortest the time till the next, as the correlation between consecutive recurrence times is positive.
- An earthquake does not “know” how big is going to be, due to the independence of the magnitude on the history of seismicity (at least from the information recorded at the catalogs, disregarding spatial structure, and with our present resolution).

These sentences, together with the previously established: the longer since the last earthquake, the longer the expected time till the next (which does not deal with correlations but with a somewhat less straightforward concept, see Davis et al., 1989; Sornette and Knopoff, 1997; Corral, 2005d) form a collection of aphorisms for seismicity properties. Note however that the belief that the longer the time to get an earthquake, the bigger its size, is false, as it would imply an a priori knowledge of how big the earthquake is going to be, in contradiction with a previous statement, i.e., with the uncorrelated behavior of the magnitude.

This law of independence of the magnitude may seem very striking; in particular, our results, established for the seismicity recorded at the modern instrumental catalogs for extended spatial regions (Southern California and worldwide), contrast with common beliefs derived from the earthquake-cycle paradigm. However, historical or paleoearthquake data for single faults, for which no correlation is found between the magnitude of a paleoearthquake and the preceding recurrence time, are surprisingly in concordance with our findings (Weldon et al., 2004; Cisternas et al., 2005).

Other works have found a local variation in the Gutenberg–Richter b value for different time periods, depending on the occurrence or not of large earthquakes (Wiemer and Katsumata, 1999). In this case, the main difference with our approach is that we deal with stationarity seismicity, for which no large earthquakes occur in Southern California. Nevertheless, one may wonder if the consideration of larger values of M'_c in the magnitude distributions could lead to a variation in these distributions above the statistical uncertainties. In any case, at this point our method does not allow to distinguish between physical phenomena and catalog-incompleteness artifacts.

Finally, let us mention that very recently the magnitude of an event has been found to be correlated not with the history of seismicity but with the focal mechanism of that event, in such a way that normal

faulting events yield large b values, whereas thrust events yield small b values (Schorlemmer et al., 2005). In a related work it has been shown that a large earthquake is likely to occur in areas of local low b value (in comparison with surrounding areas of high b) (see Schorlemmer and Wiemer, 2005). Although one could think that this implies that the distribution of magnitudes is correlated with the magnitude of a future large event, this is not the case, as distributions for different areas are involved. So, there is no disagreement between these recent works and our results.

Acknowledgements

The author is grateful to Y. Y. Kagan for his referral of a previous work and other interesting comments. M. Kukito, S. Santitoh, and M. Delduero are also acknowledged for quick inspiration, as well as the Ramón y Cajal program, the National Earthquake Information Center, the Southern California Earthquake Data Center, and the participation at the research projects BFM2003-06033 (MCyT) and 2001-SGR-00186 (DGRGC).

References

- Bak, P., 1996. *How Nature Works: The Science of Self-Organized Criticality*. Copernicus, New York.
- Bak, P., Christensen, K., Danon, L., Scanlon, T., 2002. Unified scaling law for earthquakes. *Phys. Rev. Lett.* 88, 178501.
- Bakun, W.H., Lindh, A.G., 1985. The Parkfield, California, earthquake prediction experiment. *Science* 229, 619–624.
- Christensen, K., Moloney, N.R., 2005. *Complexity and Criticality*. Imperial College Press, London.
- Christensen, K., Danon, L., Scanlon, T., Bak, P., 2002. Unified scaling law for earthquakes. *Proc. Natl. Acad. Sci. U. S. A.* 99, 2509–2513.
- Cisternas, M., et al., 2005. Predecessors of the giant 1960 Chile earthquake. *Nature* 437, 404–407.
- Corral, A., 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Phys. Rev. E* 68, 035102.
- Corral, A., 2004a. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys. Rev. Lett.* 92, 108501.
- Corral, A., 2004b. Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A* 340, 590–597.
- Corral, A., 2005a. Comment on “Do earthquakes exhibit self-organized criticality?”. *Phys. Rev. Lett.* 95, 159801.
- Corral, A., 2005b. Mixing of rescaled data and Bayesian inference for earthquake recurrence times. *Nonlinear Process. Geophys.* 12, 89–100.
- Corral, A., 2005c. Renormalization-group transformations and correlations of seismicity. *Phys. Rev. Lett.* 95, 028501.
- Corral, A., 2005d. Time-decreasing hazard and increasing time until the next earthquake. *Phys. Rev. E* 71, 017101.
- Corral, A., Christensen, K., 2006. Comment on “Earthquakes Descaled: On Waiting Time Distributions and Scaling Laws”. *Phys. Rev. Lett.* 96, 109801.
- Cox, D.R., Lewis, P.A.W., 1966. *The Statistical Analysis of Series of Events*. Methuen, London.
- Cox, D.R., Oakes, D., 1984. *Analysis of Survival Data*. Chapman and Hall, London.
- Daley, D.J., Vere-Jones, D., 1988. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York.
- Davidson, J., Goltz, C., 2004. Are seismic waiting time distributions universal? *Geophys. Res. Lett.* 31, L21612.
- Davidson, J., Paczuski, M., 2005. Analysis of the spatial distribution between successive earthquakes. *Phys. Rev. Lett.* 94, 048501.
- Davis, P.M., Jackson, D.D., Kagan, Y.Y., 1989. The longer it has been since the last earthquake, the longer the expected time till the next? *Bull. Seismol. Soc. Am.* 79, 1439–1456.
- Ellsworth, W.L., Matthews, M.V., Nadeau, R.M., Nishenko, S.P., Reasenber, P.A., Simpson, R.W., 1999. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities. *U.S. Geol. Surv. Open-File Rep.* 99–522.
- Gardner, J.K., Knopoff, L., 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed Poissonian? *Bull. Seismol. Soc. Am.* 64, 1363–1367.
- German, V., 2005. Analysis of Temporal Structures of Seismic Events on Different Scale Levels. *arxiv.org cond-mat/0501249*.
- Guggenheim, E.A., 1966. *Applications of Statistical Mechanics*. Clarendon Press, Oxford.
- Hainzl, S., Scherbaum, F., Beauval, C., 2006. Estimating background activity based on interevent-time distribution. *Bull. Seismol. Soc. Am.* 96, 313–320.
- Helmstetter, A., Sornette, D., 2002. Diffusion of epicenters of earthquake aftershocks, Omori’s law, and generalized continuous-time random walk models. *Phys. Rev. E* 66, 061104.
- Helmstetter, A., Sornette, D., 2004. Comment on “Power-law time distribution of large earthquakes”. *Phys. Rev. Lett.* 92, 129801.
- Helmstetter, A., Kagan, Y.Y., Jackson, D.D., 2006. Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bull. Seismol. Soc. Am.* 96, 90–106.
- Hergarten, S., 2002. *Self-Organized Criticality in Earth Systems*. Springer, Berlin.
- Huc, M., Main, I.G., 2003. Anomalous stress diffusion in earthquake triggering: correlation length, time dependence, and directionality. *J. Geophys. Res.* 108B, 2324.
- Jensen, H.J., 1998. *Self-Organized Criticality*. Cambridge University Press, Cambridge.
- Jones, B.J.T., Martínez, V.J., Saar, E., Trimble, V., 2004. Scaling laws in the distribution of galaxies. *Rev. Mod. Phys.* 76, 1211–1266.
- Kadanoff, L.P., 2000. *Statistical Physics: Statics, Dynamics and Renormalization*. World Scientific, Singapore.
- Kagan, Y.Y., 1997. Statistical aspects of Parkfield earthquake sequence and Parkfield prediction experiment. *Tectonophysics* 270, 207–219.
- Kagan, Y.Y., Jackson, D.D., 1991. Long-term earthquake clustering. *Geophys. J. Int.* 104, 117–133.
- Kagan, Y.Y., Jackson, D.D., 1995. New seismic gap hypothesis: five years after. *J. Geophys. Res.* 100B, 3943–3959.
- Kerr, R.A., 2004. Parkfield keeps secrets after a long-awaited quake. *Science* 306, 206–207.
- Koyama, J., Shaoxian, Z., Ouchi, T., 1995. In: Koyama, J., Deyi, F. (Eds.), *Advance in Mathematical Seismology*. Seismological Press, Beijing.
- Lee, J.F., Sears, F.W., Turcotte, D.L., 1973. *Statistical Thermodynamics*, 2nd ed. Reading, Addison-Wesley.
- Livina, V., Tuzov, S., Havlin, S., Bunde, A., 2005. Recurrence intervals between earthquakes strongly depend on history. *Physica A* 348, 591–595.

- McComb, W.D., 2004. *Renormalization Methods*. Clarendon Press, Oxford.
- Molchan, G., 2005. Interevent time distribution in seismicity: a theoretical approach. *Pure Appl. Geophys.* 162, 1135–1150.
- Mulargia, F., Geller, R.J., 2003. *Earthquake Science and Seismic Risk Reduction*. Kluwer, Dordrecht.
- Murray, J., Segall, P., 2002. Testing time-predictable earthquake recurrence by direct measurement of strain accumulation and release. *Nature* 419, 287–291.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* 83, 9–27.
- Ogata, Y., 1999. Seismicity analysis through point-process modeling: a review. *Pure Appl. Geophys.* 155, 471–507.
- Rundle, J.B., Turcotte, D.L., Shcherbakov, R., Klein, W., Sammis, C., 2003. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Rev. Geophys.* 41, 1019.
- Schorlemmer, D., Wiemer, S., 2005. Microseismicity data forecast rupture area. *Nature* 434, 1086–1086.
- Schorlemmer, D., Wiemer, S., Wyss, M., 2005. Variations in earthquake-size distribution across different stress regimes. *Nature* 437, 539–542.
- Sieh, K., 1996. The repetition of large-earthquake ruptures. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3764–3771.
- Sieh, K., Stuiver, M., Brillinger, D., 1989. A more precise chronology of earthquakes produced by the San Andreas fault in Southern California. *J. Geophys. Res.* 94, 603–623.
- Smalley, R.F., Chatelain, J.-L., Turcotte, D.L., Prévot, R., 1987. A fractal approach to the clustering of earthquakes: applications to the seismicity of the New Hebrides. *Bull. Seismol. Soc. Am.* 77, 1368–1381.
- Sornette, D., 2000. *Critical Phenomena in Natural Sciences*. Springer, Berlin.
- Sornette, D., Knopoff, L., 1997. The paradox of the expected time until the next earthquake. *Bull. Seismol. Soc. Am.* 87, 789–798.
- Sornette, D., Ouillon, G., 2005. Multifractal scaling of thermally activated rupture processes. *Phys. Rev. Lett.* 94, 038501.
- Stein, R.S., 1995. Characteristic or haphazard? *Nature* 370, 443–444.
- Stein, R.S., 2002. Parkfield's unfulfilled promise. *Nature* 419, 257–258.
- Turcotte, D.L., 1999. Self-organized criticality. *Rep. Prog. Phys.* 62, 1377–1429.
- Udías, A., Rice, J., 1975. Statistical analysis of microearthquake activity near San Andreas Geophysical Observatory, Hollister, California. *Bull. Seismol. Soc. Am.* 35, 809–827.
- Wang, J.-H., Kuo, C.-H., 1998. On the frequency distribution of interoccurrence times of earthquakes. *J. Seismol.* 2, 351–358.
- Weldon, R., Scharer, K., Fumal, T., Biasi, G., 2004. Wrightwood and the earthquake cycle: what a long recurrence record tells us about how faults work. *GSA Today* 14 (9), 4–10.
- Weldon, R.J., Fumal, T.E., Biasi, G.P., Scharer, K.M., 2005. Past and future earthquakes on the San Andreas fault. *Science* 308, 966–967.
- Wiemer, S., Katsumata, K., 1999. Spatial variability of seismicity parameters in aftershock zones. *J. Geophys. Res.* 104, 13135–13151.