Long-Term Clustering, Scaling, and Universality in the Temporal Occurrence of Earthquakes

Álvaro Corral*

*Departament de Física, Facultat de Ciències, Universitat Autònoma de Barcelona,
Edifici Cc, E-08193 Bellaterra, Barcelona, Spain*
(Received 14 October 2003; published 10 March 2004)

Analyzing diverse seismic catalogs, we have determined that the probability densities of the earthquake recurrence times for different spatial areas and magnitude ranges can be described by a unique universal distribution if the time is rescaled with the rate of seismic occurrence, which therefore fully governs seismicity. The shape of the distribution shows the existence of clustering beyond the duration of aftershock bursts, and scaling reveals the self-similarity of the clustering structure in the space-time-magnitude domain. This holds from worldwide to local scales, for quite different tectonic environments and for all the magnitude ranges considered.

Although earthquakes are a phenomenon of great complexity, certain simple general laws govern the statistics of their occurrences [1–5]; however, for the time interval between successive events a unified description has not yet been established [6–9]. On the contrary, the rich variability intrinsic to earthquakes has promoted that all possibilities have been proposed for their temporal properties, from totally random occurrence to the periodic ticking of great quakes. The most extended view is that of two separated processes, one for mainshocks, which ought to follow a Poisson distribution [10] (or not [6–9]), and one for an independent process to generate aftershocks. Consequently, the "standard practice" for this approach consists first of delimiting the spatial area to be studied, on the basis of its tectonic characteristics, and then of carefully (not so standard) identifying aftershocks, in order to separate them from the main sequence.

Here we take an alternative perspective, complementary to the previous reductionist view. We try to look at the system as a whole, irrespective of tectonic features, and place all the events on the same footing, whether these would be classified as mainshocks or aftershocks [11–13]. This follows one of the key guidelines of complexity philosophy, which is to find descriptions on a general level; the existence of general laws fulfilled by all the earthquakes unveil a degree of unity in an extremely complex phenomenon [14].

We analyze a global catalog, the PDE from the NEIC [15], as well as several local catalogs: that of the SCSN (Southern California) [16], the JUNEC (Japan) [17], the Bulletins of the IGN (the Iberian Peninsula and the North of Africa) [18], and the BGS catalog (the British Islands and the North Sea) [19]. Catalogs generally characterize each earthquake by three main quantities: time of occurrence, magnitude, and a vector of spatial coordinates for the hypocenter; these are then the variables that we focus on.

Without concerning ourselves with the tectonic properties, as Bak *et al.* [11–13], we consider spatial areas delimited by a window of L degrees in longitude and L degrees in latitude (this corresponds to a square region if these angles are translated onto a rectangular coordinate system [20]). For each one of these regions, only events with magnitude M above a certain threshold M_c are considered (the threshold should be larger than the minimum magnitude for which the catalog is considered complete for the spatial and temporal windows considered). In this way, we transform a time process in four dimensions (spatial coordinates and magnitude) into a simple process on a line for which events occur at times t_i, with $i = 0, 1, 2, \ldots$, and therefore, the time between successive events can be obtained as $\tau_i = t_i - t_{i-1}$, $i = 1, 2, \ldots$. These are the recurrence times in a given L^2 region for events above M_c, which can also be referred to as interoccurrence or waiting times. Note that, with this transformation, we have lost the structure in space and in the magnitude scale; nonetheless, the change in the process properties with the variation of L and M_c will allow us to recover some of this information.

Because of the multiple time scales involved (from seconds or minutes to many years), the probability density of the recurrence time must be calculated with care. We could work with the logarithm of τ, but an equivalent and more direct possibility is to define the bins over which the probability density is calculated exponentially growing as c^n, with $c > 1$ and n labeling consecutive bins. This ensures an appropriate bin size for each time scale (we usually take $c = 2.5$, although this particular value is totally irrelevant). We then count the number of pairs of consecutive events separated by a time whose value lies in a given bin and divide by the total number of pairs of events (number of events minus one) and by the size of the bin to attain the estimation of the probability density $D_{xy}(\tau)$ over that bin, where xy denotes the spatial coordinates of the region. (D_{xy} also depends on L and M_c, but for the sake of simplicity in the notation, we obviate this dependence.) Moreover, because of the incompleteness of the catalogs in the short-time scale, we do not display in the plots recurrence times smaller than 2 min.

The entire Earth has been analyzed by this method. Figure 1(a) shows the results for $D_{xy}(\tau)$ for worldwide earthquakes in the NEIC-PDE catalog for the 1973–2002 period, using M_c from 5 to 6.5 and L from 180° to 2.8° (about 300 km), for many regions of different x, y coordinates (see the figure's caption). Note the variation of the recurrence time across several orders of magnitude. Figure 1(b) shows the rescaling of all the distributions with the mean rate R in the region, defined as the total number of events divided by the total time interval over which these events span. The perfect data collapse implies that we can write

$$D_{xy}(\tau) = R_{xy}f(R_{xy}\tau),$$

where R_{xy} stresses that the rate refers to the (x, y) region (of size L^2 and with $M \geq M_c$; also here we eliminate these variables from the notation). The scaling function f can be well fit by a generalized gamma distribution,

$$f(\theta) = C\frac{1}{\theta^{1-\gamma}}\exp(-\theta^\delta/B),$$

with parameters $\gamma = 0.67 \pm 0.05$, $\delta = 0.98 \pm 0.05$, $B = 1.58 \pm 0.15$, and $C = 0.50 \pm 0.10$, which yields a coefficient of variation $CV \simeq 1.2$. In fact, the value of δ can be approximated to 1, which corresponds to the standard gamma distribution; therefore, we have essentially a decreasing power law with exponent about 0.3, up to the largest values of the argument, $\theta = R_{xy}\tau$ about 1, where the exponential factor comes into play.

The fit of the rescaled distributions by the gamma scaling function is surprisingly good for intermediate and large values of the recurrence time, about $\tau > 0.01/R_{xy}$ (this usually contains from 90% to 95% of probability). The deviations are considerable for small values of τ. Although the statistic is low in this case (few events in the small bins being considered), for certain regions there is a clear tendency for the distribution to exceed the value given by the scaling function; that is, there is an excess of very short recurrence times, in the form of another power law but with the exponent much closer to $(-)1$. This occurs when the rate in the region is not stationary, due to the sudden increase and slow decay of the activity provoked by aftershock sequences. Furthermore, these increments become more apparent when the size of the region decreases, in such a way that, for $L \leq 22.5°$, not all the regions in the world verify the scaling law; this can be solved in some cases by rescaling with the mean rate in the region calculated, not over the whole time span of the catalog, but only over the period for which the rate is stationary (no activity peaks). Nevertheless, the aftershocks can be so important for certain particular regions that a stationary period may not exist; these cases are then not included in Fig. 1, but are addressed below.

The same analysis is performed on local catalogs and identical results hold, as Fig. 1(c) illustrates. For Southern California for the 1984–2001 period, a number of small regions with stationary activity are shown; for larger
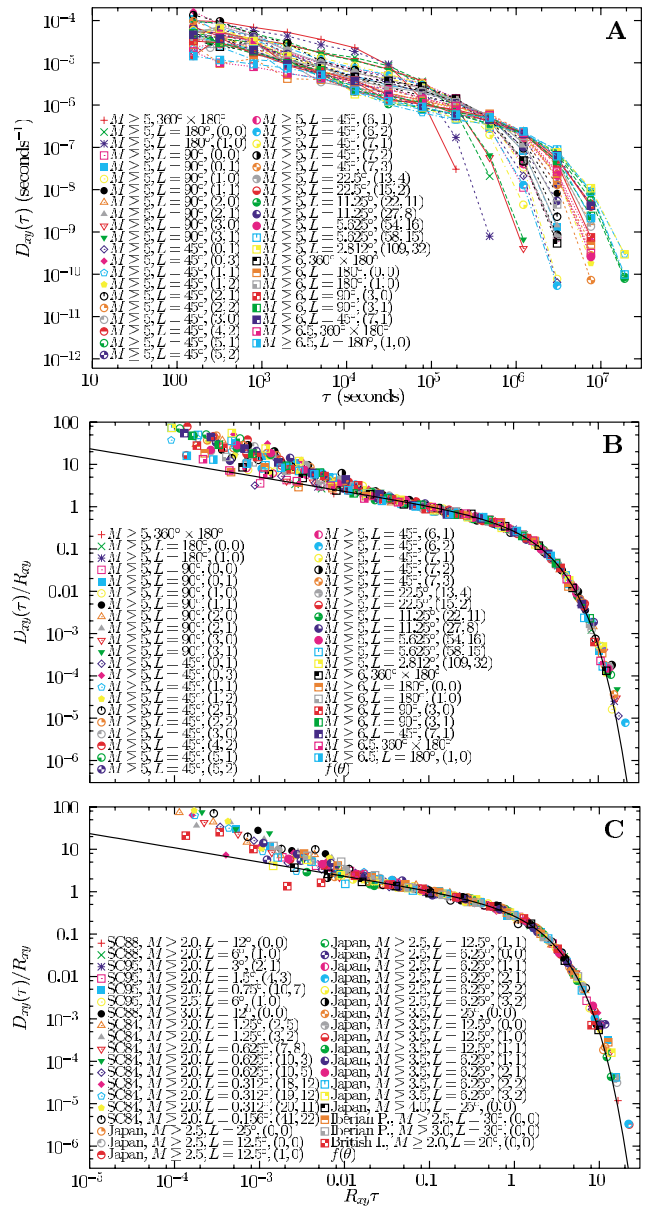

FIG. 1 (color online).   Recurrence-time distributions without and with rescaling. (a) Probability densities from the NEIC-PDE worldwide catalog for several regions, L, and M_c. For $L \geq 45°$, all the regions with more than 500 events are shown, whereas for $L \leq 22.5°$, only regions with moderate aftershock activity are displayed (for these, the total number of events versus time shows a dominant linear behavior). The vector (k_x, k_y) labels the different regions, for which the coordinates of the center can be obtained as $x = x_{min} + (k_x + 0.5)L$, $y = y_{min} + (k_y + 0.5)L$, with $x_{min} = -180°$, $y_{min} = -90°$. (The 360° × 180° region, which covers the whole planet, has been included for completeness.) (b) Previous data, after rescaling, with a fit of the scaling function f. (c) Rescaled distributions from local catalogs. SC88, SC95, and SC84 refer to Southern California for the years 1988–1991, 1995–1998, and 1984–2001. To obtain region coordinates, use the previous formula with $(x_{min}, y_{min}) = (-124°, 29°)$, $(-123°, 30°)$, $(125°, 25°)$, $(-20°, 30°)$, and $(-10°, 45°)$ for SC88-95, SC84, Japan, the Iberian Peninsula, and the British Islands, respectively. The function displayed is the fit obtained from the NEIC-PDE catalog.

regions, the time window must be reduced to 1988–1991 or to 1995–1998, for instance, in order to find stationarity. For Japan, we also analyze large regions for the 1995–1998 period; for the Iberian Peninsula the period is 1993–1997, and for the British Islands, 1991–2001. The magnitude thresholds range from 2 to 4, and L from 30° to 0.16° (approximately 3300 to 17 km).

In all cases, the shape of the distribution $D_{xy}(\tau)$ (clearly different from an exponential), indicates that the memory of the last earthquake is conserved up to the largest times, with the probability of a subsequent event being maximum immediately after the last shock, and slowly decreasing with time. This constitutes a clustering effect [21], in which earthquakes attract each other, and has as a counterintuitive consequence the fact that the longer it has been since the last earthquake, the longer the expected time will be till the next [6,22,23]. Since this effect occurs beyond the duration of aftershock sequences, i.e., when the rate of seismic activity has returned to its background value (for the region considered), and extends for the largest times, it constitutes a long-term clustering (in opposition to the short-term clustering of aftershocks, the time scale being set by $1/R_{xy}$).

On the other hand, the scaling of $D_{xy}(\tau)$ under changes in M_c, L, and region coordinates implies that the clustering structure is self-similar over different regions and magnitude ranges. The robustness of the distribution under such changes is therefore noteworthy. It is also remarkable that if the region is kept fixed and only M_c varies, the scaling with the rate R_{xy} can be substituted by the factor 10^{-bM_c}, where b refers to the b value of the frequency-magnitude relation [1,2] in that particular region [11,12]; despite the regional variability of b, the universality of the scaling function f remains valid.

At the outset of this exposition, we suggested that our results are valid for all events, including aftershocks; however, aftershocks can break the scaling of the distribution up to very large time values (in $1/R_{xy}$ scale), as we have also mentioned. One could then conclude that the universal distribution does not describe the short-time intervals over which aftershocks replicate; this may appear to be correct, but can turn out to be mistaken in the following way: for aftershock sequences, the mean rate is not stationary, but changes with time; therefore, we should rescale the recurrence times using the "instantaneous" rate $r_{xy}(t)$. For many sequences, $r_{xy}(t)$ is found to decay following the modified Omori law for long t,

$$r_{xy}(t) = \frac{A}{t^p},$$

where t is the time elapsed since the mainshock (and the parameters depend on x, y, L, and M_c). Figure 2(a) shows precisely this for several important earthquakes in Southern California [16]. The data present clear "holes" in the occurrence of small earthquakes ($M < 3$) for about a couple of days after the main shock; however, for the periods for which the power-law decay of the rate

is fulfilled the catalog appears reasonably complete for $M \geq 2$, as the validity of the Gutenberg-Richter law suggests (in fact, the deviations from the power law for short times may be due to only the incompleteness of the
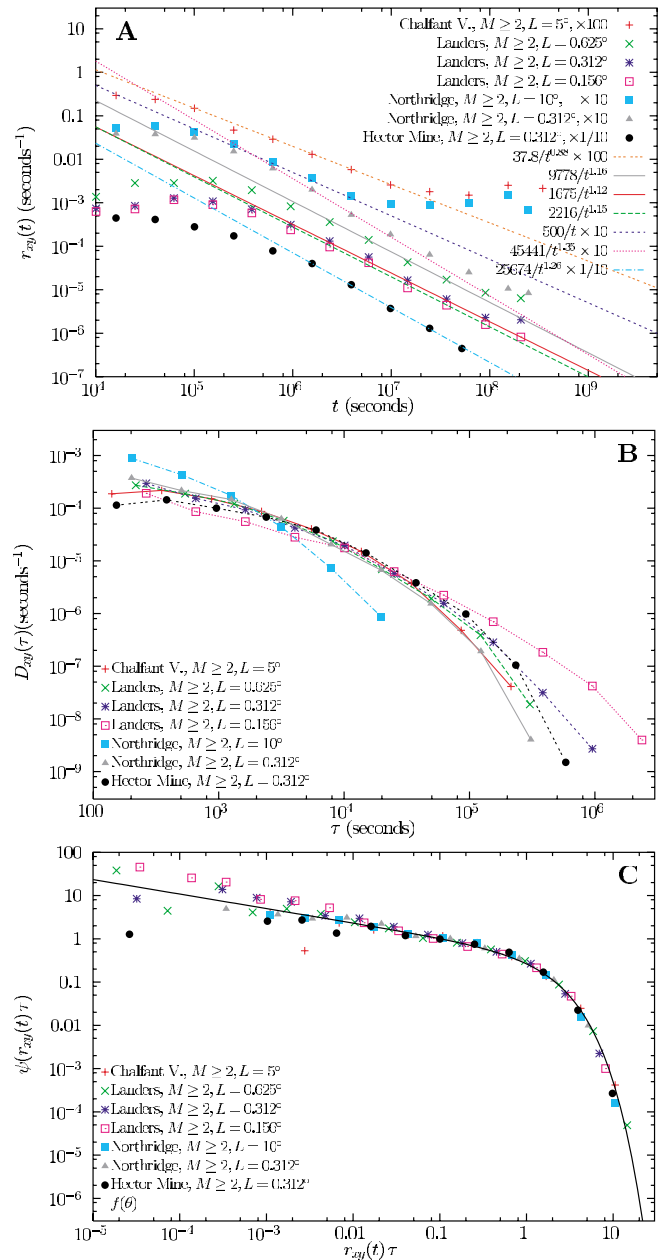

FIG. 2 (color online). Analysis of aftershock sequences. (a) Decay of the rate after a mainshock and illustration of the Omori law for the following earthquakes in Southern California: Chalfant Valley (July 21, 1986, $M = 5.9$), Landers (June 28, 1992, $M = 7.3$), Northridge (Jan. 17, 1994, $M = 6.7$), and Hector Mine (Oct. 16, 1999, $M = 7.1$). Regions of diverse size L are considered, all of these including the mainshock. Some curves are shifted for the sake of clarity. (b) Distributions of recurrence times for the previous sequences. (c) Distributions ψ of the dimensionless time $r_{xy}(t)\tau$, in total agreement with the universal scaling function f derived previously for the NEIC-PDE data.

catalogs). The corresponding recurrence-time distributions (for the power-law-decay period) are displayed in Fig. 2(b); the results after rescaling with $r_{xy}(t)$ appear in Fig. 2(c), again in agreement with the universal distribution. The scaling is outstanding, taking into account that the p values spread from 0.9 to 1.35. Note also that this scaling implies the existence of a secondary clustering inside the primary clustering of the aftershock sequence, and therefore the process is not a nonhomogeneous Poisson process [as the recurrence times are not governed by an exponential distribution with decaying rate $r_{xy}(t)$]. In summary, it is only the rate $r_{xy}(t)$ [given by the Omori law or being constant (R_{xy} plus fluctuations) in the stationary case] that controls seismic occurrence.

The results we obtain are extremely robust against the incompleteness of the catalogs in magnitude, space, and time. The variation of the minimum magnitude M_c, region spatial coordinates, size of the region L, and temporal window of observation (avoiding or explicitly including the periods corresponding to important aftershock sequences), spanning a very broad spectrum of values, allows one to test the robustness of our conclusions. This test is at the core of the procedure and makes it clear the genuine physical origin of the universal scaling law for earthquake recurrence times. We also note that this law is obtained without any underlying model of earthquake occurrence.

The differences between our approach and that of Bak *et al.* [11,12] are worth mentioning. Instead of measuring the recurrence-time distribution for a single L^2 region of coordinates *x, y*, as we do, Bak *et al.* perform a mixture of all the recurrence times for all the regions; i.e., the values of τ coming from regions with different coordinates are counted together in the same distribution. In some sense, they are measuring the return times to different parts of a large region (with heterogeneities also in time), whereas we deal with plain recurrence-time distributions (in the time-homogeneous case, or making the sequence stationary). Naturally, the results and the conclusions of both approaches are also different. For an in-depth discussion see Ref. [13].

The present characterization of the stochastic spatio-temporal occurrence of earthquakes by means of a unique law indicates the existence of universal mechanisms in the earthquake-generation process, governed only by the seismic rate [24]. The understanding of this, however, is still far beyond us; nevertheless, the context of self-organized critical phenomena [14] provides a coherent framework at this stage. These findings can also be relevant to continuous-time random-walk models of seismicity [25,26], time-dependent hazard, and forecasting in general [27–29].

*Email address: Alvaro.Corral@uab.es

[1] B. Gutenberg and C. F. Richter, *Seismicity of the Earth* (Hafner Publishing Co., New York, 1965).
[2] D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics* (Cambridge University Press, Cambridge, 1997), 2nd ed.
[3] T. Utsu, Y. Ogata, and R. S. Matsu'ura, J. Phys. Earth **43**, 1 (1995).
[4] P. A. Reasenberg and L. M. Jones, Science **243**, 1173 (1989).
[5] Y. Y. Kagan and L. Knopoff, Geophys. J. R. Astron. Soc. **62**, 303 (1980).
[6] D. Sornette and L. Knopoff, Bull. Seismol. Soc. Am. **87**, 789 (1997).
[7] J.-H. Wang and C.-H. Kuo, J. Seismol. **2**, 351 (1998).
[8] W. L. Ellsworth *et al.*, U.S. Geological Survey Open-File Report No. 99-522, 1999.
[9] R. F. Smalley, J.-L. Chatelain, D. L. Turcotte, and R. Prévot, Bull. Seismol. Soc. Am. **77**, 1368 (1987).
[10] J. K. Gardner and L. Knopoff, Bull. Seismol. Soc. Am. **64**, 1363 (1974).
[11] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).
[12] K. Christensen, L. Danon, T. Scanlon, and P. Bak, Proc. Natl. Acad. Sci. U.S.A. **99**, 2509 (2002).
[13] A. Corral, Phys. Rev. E **68**, 035102(R) (2003).
[14] P. Bak, *How Nature Works: The Science of Self-Organized Criticality* (Copernicus, New York, 1996).
[15] National Earthquake Information Center, Preliminary Determination of Epicenters catalog, http://wwwneic.cr.usgs.gov/neis/epic/epic_global.html.
[16] Southern California Seismographic Network, http://www.scecdc.scec.org/ftp/catalogs/SCSN.
[17] Japan University Network Earthquake Catalog, http://wwweic.eri.u-tokyo.ac.jp/CATALOG/junec/monthly.html.
[18] Instituto Geográfico Nacional, Boletín de Sismos Próximos, http://www.geo.ign.es/servidor/sismo/cnis/terremotos.html.
[19] British Geological Survey, catalog available upon request.
[20] Nevertheless, the shape of the region is totally irrelevant in our study.
[21] Y. Y. Kagan and D. D. Jackson, Geophys. J. Int. **104**, 117 (1991).
[22] P. M. Davis, D. D. Jackson, and Y. Y. Kagan, Bull. Seismol. Soc. Am. **79**, 1439 (1989).
[23] A. Corral, cond-mat/0310407.
[24] S. Toda, R. S. Stein, and T. Sagiya, Nature (London) **419**, 58 (2002).
[25] Y. Ogata, Pure Appl. Geophys. **155**, 471 (1999).
[26] A. Helmstetter and D. Sornette, Phys. Rev. E **66**, 061104 (2002).
[27] I. Main, Nature (London) **385**, 19 (1997).
[28] R. J. Geller, D. D. Jackson, Y. Y. Kagan, and F. Mulargia, Science **275**, 1616 (1997).
[29] T. Parsons, S. Toda, R. S. Stein, A. Barka, and J. H. Dieterich, Science **288**, 661 (2000).