# Sensitivity analysis of the parameters of earthquake recurrence time power law scaling

**Abdelhak Talbi (✉)· Fumio Yamazaki**

Department of Urban Environment System

Graduate School of Engineering

Chiba University

1-33, Yayoi-cyo, Inage-ku,

Chiba-shi, Chiba 263-8522, Japan

Tel: +81-43-290-3528

Fax: +81-43-290-3558

e-mail: abdelhak_t@graduate.chiba-u.jp

**Abstract** The stability of the power law scaling of earthquake recurrence time distribution in a given space-time window is investigated, taking into account the magnitude of completeness and the effective starting time of aftershock sequences in earthquake catalogs from Southern California and Japan. A new method is introduced for sampling at different distances from a network of target events. This method allows the recurrence times to be sampled many times on the same area. Two power laws with unknown exponents are assumed to govern short and long recurrence time ranges. This assumption is developed analytically and shown to imply simple correlation between these power laws. In practice, the results show that this correlation structure is not satisfied for short magnitude cutoffs ($m_c$=2.5, 3.5, 4.5), and hence the recurrence time distribution departs from the power law scaling. The scaling parameters obtained from the stack of the distributions corresponding to different magnitude thresholds are quite different for different regions of study. It is also found that significantly different scaling parameters adjust the distribution for different magnitude thresholds. In particular, the power law exponents decrease when the magnitude cutoff increases, resulting in a slower decrease of the recurrence time distribution, especially for short time ranges. For example in the case of Japan, the exponent $p_2$ of the power law scaling at large recurrence times follows roughly the relation: $p_2(m_c) = -0.07m_c + 2.7;\ m_c \geq 3.5$, where $m_c$ is the magnitude cutoff. In case of Southern California, it is shown that Weibull distribution provides a better alternative fit to the data for moderate and large time scales.

**Keywords** Recurrence times · Scaling· Power laws · Universality · Magnitude of completeness.

## 1 Introduction

Quite recently, a new scaling law for earthquake recurrence time distribution, *D*, has been introduced and claimed to be universal for broad areas and different magnitude thresholds

(Bak et al. 2002; Christensen et al. 2002; Corral 2003, 2004a,b, 2007). The law we refer to in short as "Bak's scaling law" consists of two power laws (PLs), one for short time ranges and the other for long time ranges. According to its authors, this law reveals a complex spatiotemporal organization of seismicity, which may be viewed as an intermittent flow of energy released within a self-organized (possibly critical) system. However, several authors argued on the robustness of such scaling on the whole time scale (Davidsen and Goltz 2004; Carbone et al. 2005; Lindman et al. 2005, 2006; Molchan 2005; Corral and Christensen 2006; Hainzl et al. 2006; Saichev and Sornette 2006, 2007; Molchan and Kronrod 2007). In particular, Saichev and Sornette (2006, 2007) discussed quite well the distribution $D$ in the framework of the epidemiological time aftershock sequence (ETAS) model (Ogata 1988) and showed that the so-called universal scaling law of interevent times does not strictly hold and can be well reproduced by the usual laws of seismicity (Gutenberg-Richter and Omori, essentially). In addition, Corral (2004a, 2007) found the evidence that the PL at short time ranges is not unique for all areas.

It is noteworthy that the most of the previously cited studies dealing with this topic do not assess the magnitude of completeness explicitly, despite its importance in all quantitative studies based on earthquake catalogs. Besides, if we consider the data far below the effective starting times of aftershock sequences as in Bak et al. (2002), the interevent times issued from the whole catalog data (that may contain several aftershock sequences) are not complete and may lead to biased results (Kagan 2004; Helmstetter et al. 2006). Moreover, the precision (confidence intervals) of Bak's scaling law exponents has never been estimated. To account for such deficiency, we may raise two main issues. The first one ($T_1$) is the preliminary assessment of different magnitudes of completeness and minimum interevent time cutoffs for the data in use. The second ($T_2$) is the estimation of the scaling parameters precision and the study of their sensitivity, relative to the performed cutoffs. These tasks are extremely difficult and time consuming because Bak's scaling law involves the use of big data sets including catalogs from different regions and different magnitude scales.

In this study, we aim to initiate the tasks $T_1$ and $T_2$ to discuss Bak's scaling law. Namely, our objective is to study the sensitivity of Bak's scaling law parameters when $D$ is issued from different regions and calculated using different magnitude thresholds, with a particular care to its derivation from a reasonable cutoff at short interevent times and from the areas with sufficient information. We perform the process in three steps. First, different magnitudes of completeness for all available data sets are assessed and sampling strategy for interevent times is developed. Secondly, the tendency of the distribution of interevent times to PLs is empirically assessed and the PL exponents for different regions are estimated. Finally, the parameterization of the distribution $D$ is analyzed under the hypothesis of a doubly PL behavior, with a particular focus on the case where the PL exponents vary with different magnitude cutoffs.

As a result, a simple relation is established between short and long range PLs with parameters linked to the probability at the distribution tails. The PL parameter estimates obtained from the stack of the distributions corresponding to different magnitude thresholds are not stable for different regions of study. Moreover, significantly different scaling parameters fit the distribution for different magnitude thresholds. In particular,

the PL exponents tend to decrease with increasing magnitudes, resulting in a slower decrease of $D$, especially at short time ranges.

Finally, it is worth mentioning that in our analysis no assumption of a specific seismicity model (such that the ETAS used by Saichev and Sornette (2006, 2007)) is made. Thus, the results described here are supposed to describe seismicity in a broad sense.

## 2 Data and Methods

Earthquake catalog data from Southern California and Japan were used in this analysis. For Southern California, we combined the catalog files for the period 1932-2005 from Southern California National Network (http://www.data.scec.org/ftp/catalogs/SCSN/), truncated at low magnitude thresholds ($m_c$ = 2.5, 3.5). For higher thresholds, the catalog (http://moho.ess.ucla.edu/~kagan/s_cal.dat) compiled by Kagan was used (Kagan et al. 2006). The space window was considered as 32-37° N latitude and 114-122° W longitude (Fig. 1a). For Japan, the JMA catalog (The Seismological and Volcanological Bulletin for November 2005 (CD-ROM), Japan Meteorological Agency, JMA) for the period 1923-2005, and a compiled version of Utsu catalog (www5b.biglobe.ne.jp/~t-kamada/CBuilder/eqlist.htm) for the period 679-1922 were used with the space window 24-50° N latitude and 122-152° E longitude (Fig. 1b).

**(a)**

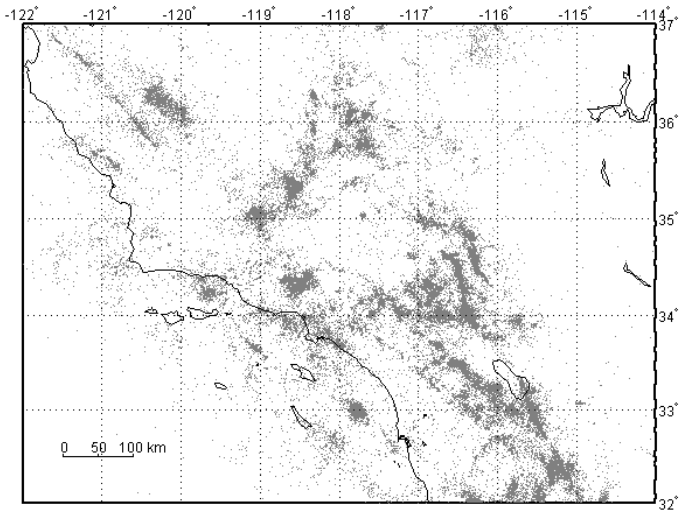
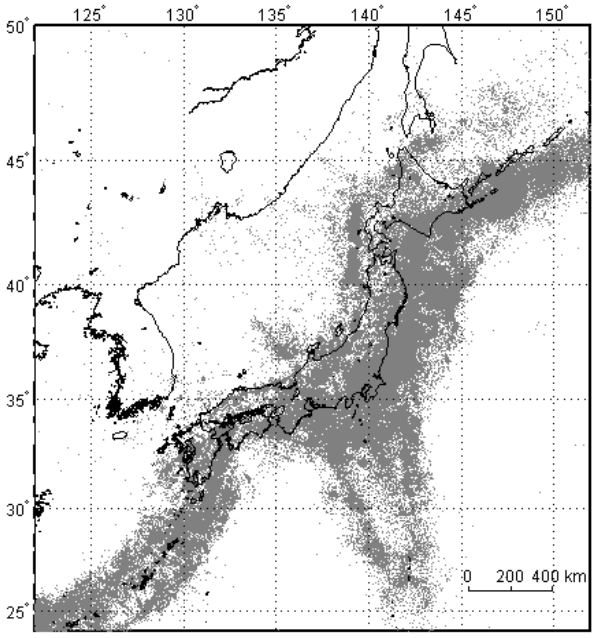



**Fig. 1** Epicenter distribution map of earthquakes with magnitude greater or equal to 2.5 occurred in **(a)** Southern California within the period 1932-2005. **(b)** Japan within the period 679-2005 as compiled from JMA and Utsu catalogs

**(b)**




Completeness periods for different magnitude thresholds were calculated using the Stepp's approach (Stepp 1971, 1972) as shown in Table 1. This method is briefly described in Appendix A,

Table 1

Seismicity models derived for Southern California and Japan and their estimated parameters.

$m_c^l$ is the magnitude threshold corresponding to the sampling scheme *l*, $N\left(M \geq m_c^l\right)$ denotes the number of events of magnitude greater or equal to $m_c^l$. For each magnitude threshold $m_c^l$, a sampling radius *R* for interevent times is considered, so that the parameters: *P*, $m_c^l$, *N* and *R* define the sampling scheme *l* (see the sampling procedure ERS described later in Figure 4).

| Southern California | | | | |
|---|---|---|---|---|
| Scheme $l$ | Time period $P$ | $m_c^l$ | $N = N\left(M \geq m_c^l\right)$ | $R$ [km] |
| 1 | 1990-2005 | 2.5 | 21257 | 50 |
| 2 | 1947-2005 | 3.5 | 6006 | 50 |
| 3 | 1932-2005 | 4.7 | 556 | 50 |
| Japan | | | | |
| Scheme $l$ | Time period $P$ | $m_c^l$ | $N = N\left(M \geq m_c^l\right)$ | $R$ [km] |
| 1 | 1990-2005 | 3.5 | 37352 | 50 |
| 2 | 1975-2005 | 4.5 | 11406 | 50 |
| 3 | 1923-2005 | 5.5 | 3664 | 100 |
| 4 | 1890-2005 | 6.5 | 590 | 200 |

For each data set, the logarithm of the cumulative event number was calculated for all the magnitude thresholds. To account for the error in the report of the magnitude $M$, the data were grouped in $n$ magnitude classes $[m_i, m_{i+1}]$ with a fixed width ($\Delta m = m_{i+1}-m_i = 0.1$; $i = 1, 2, \ldots, n$). Then, the $b$-value was estimated using the least square method weighted with the number of events in each magnitude class. The following is the sum of Gutenberg-Richter residuals to be minimized,

$$\varepsilon^2 = \sum_{l=1}^{L}\sum_{i=1}^{n} w_i^l \left(\log_{10}\left(N_l\left(M \geq m_i\right)\right) - \left(a - bm_i\right)\right)^2 \qquad (1)$$

$$w_i^l = \frac{N_l\left(m_i \leq M < m_{i+1}\right)}{N_l\left(m \geq m_c^l\right)}, \quad i = 1, 2, \ldots, n;\ l = 1, 2, \ldots, L \qquad (2)$$

where $L$ is the number of magnitude cutoffs or schemes shown in Table 1, $m_c^l$ is the cutoff magnitude corresponding to the scheme $l$, whereas $N_l\left(M \geq m_i\right)$ and $N_l\left(m_i \leq M < m_{i+1}\right)$ are respectively the number of events with magnitude greater or equal to $m_i$ and that with magnitude falling in the $i^{\text{th}}$ bin $[m_i, m_{i+1}]$, counted in the scheme $l$. The weights $w_i^l$ ($i = 1, 2, \ldots, n$; $l = 1, 2, \ldots, L$) were introduced to account for a contribution proportional to the number of events in each magnitude class. The corresponding Gutenberg-Richter distributions together with the estimated $a$ and $b$ values are shown in Figures 2.

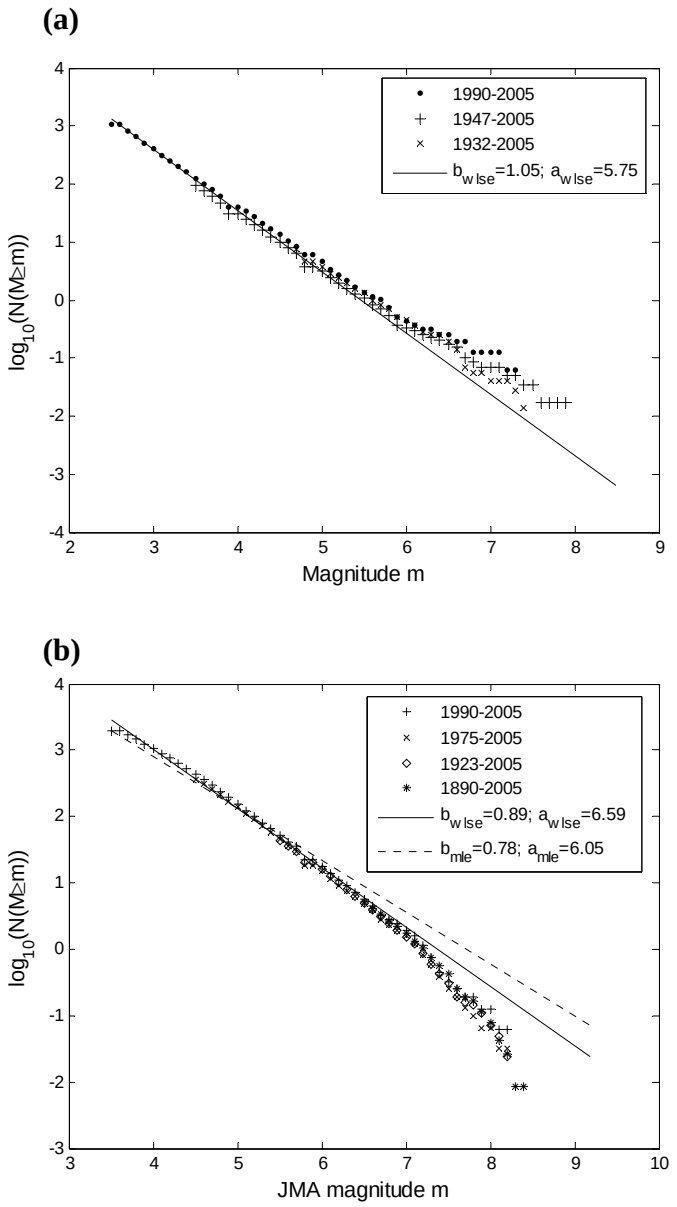



**Fig. 2** Fitted G-R distributions for different magnitude cutoffs in case of **(a)** Southern California and **(b)** Japan. $a_{\mathrm{wlse}}$ and $b_{\mathrm{wlse}}$ are the Gutenberg Richter parameter estimates obtained using the weighted least square method and $a_{\mathrm{mle}}$ and $b_{\mathrm{mle}}$ are those obtained using the maximum likelihood method

In all the cases, the Gutenberg-Richter distribution is characterized by linear trend along several magnitude scales, showing that the corresponding catalogs are complete for the periods and the threshold magnitudes shown in Table 1. Departure above magnitude 5.5 in Figure 2a denotes the existence of a second branch in the magnitude distribution of Southern California (Knopoff 2000). For comparison, the maximum likelihood estimates for unequal observation periods (Weichert 1980; Bender 1983) are also shown in Figure 2b. The derived *a* and *b* values estimates in this case and those derived using the weighted least square method are quite equivalent. Finally, the periods of completeness shown in Table 1 were checked using the maximum curvature method (MAXC) (Wiemer and Wyss 2000). We used the standard parameter predefined in the Zmap free Matlab code that is a sample window size of 500, a binning of 0.1 and 200 bootstrap samples to

calculate uncertainties. The results summarized in Figure 3a, b show that our data can be considered complete above the magnitude of completeness and within the time periods selected in Table 1.

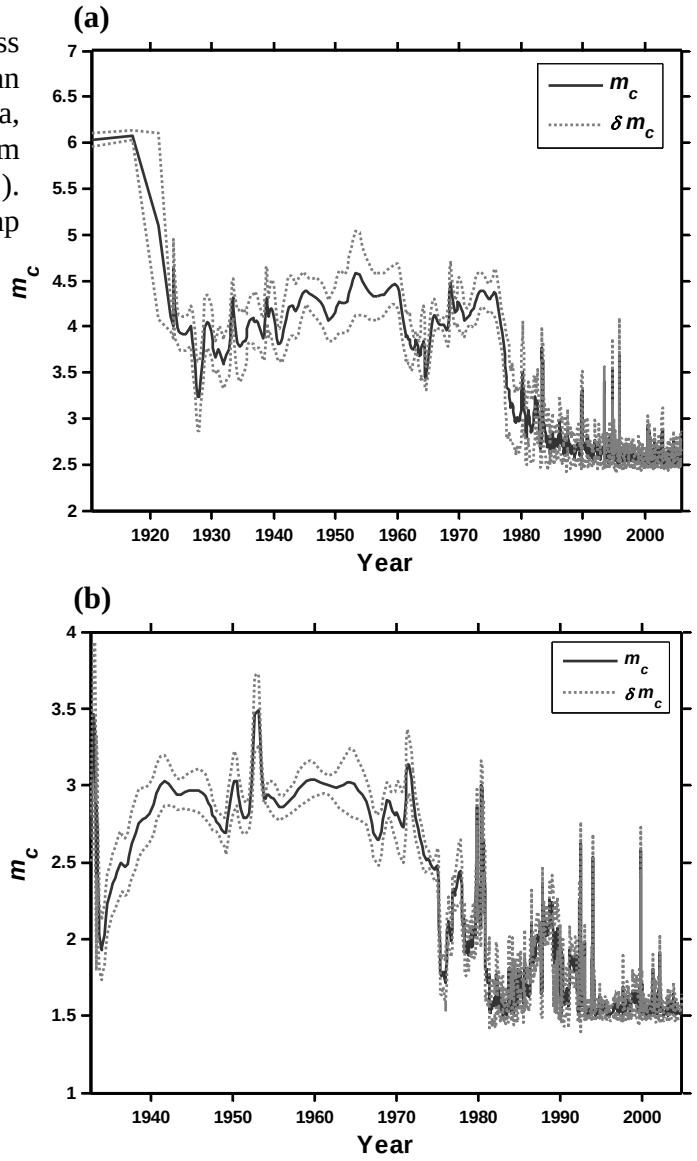



**Fig. 3** Magnitude of completeness as function of time for **(a)** Japan and **(b)** Southern California, calculated using the maximum curvature method (MAXC). Dashed lines are the bootstrap errors limits

After assessing the magnitude of completeness, waiting times were sampled using the discretization method described in Appendix B and Figure 4 (Talbi and Yamazaki 2007). In the following, this procedure is referred to as "ERS(*R*)", the abbreviation of "Earthquake Random Sampling with sampling radius *R*". ERS(*R*) uses disks with fixed radius *R* (called sampling radius) to control sampling at a given distance around each event from a set of target events. An illustration of sampling disks obtained from two ERS runs for the case of Japan scheme 3 of Table 1 is shows in Figure 5.

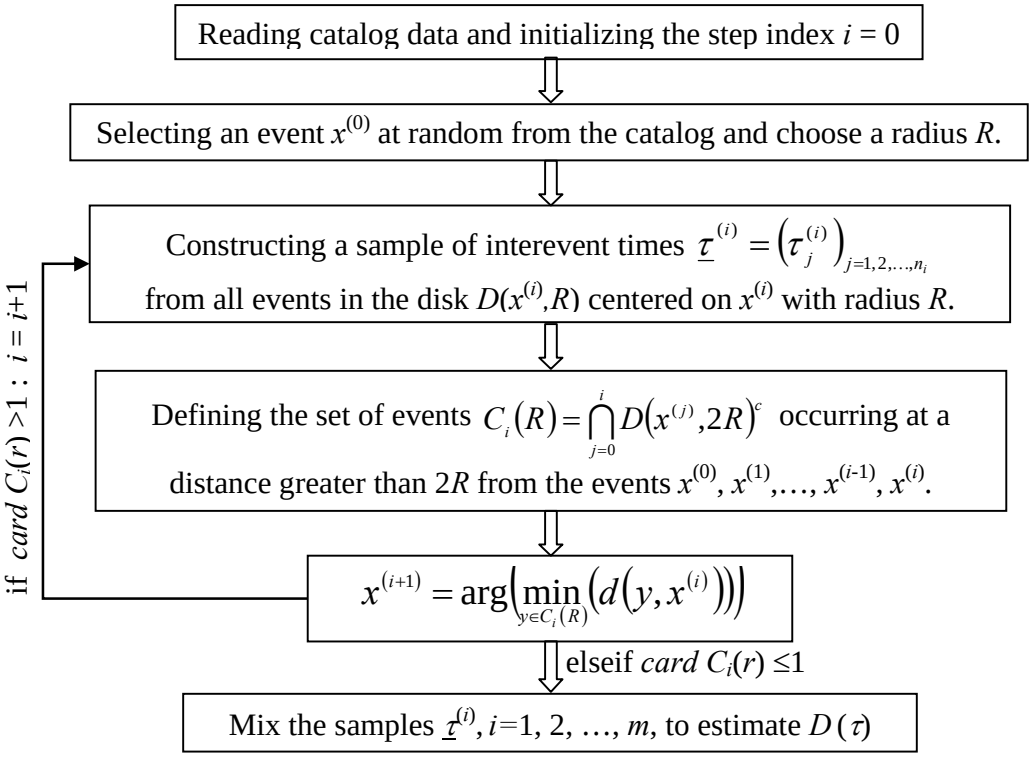
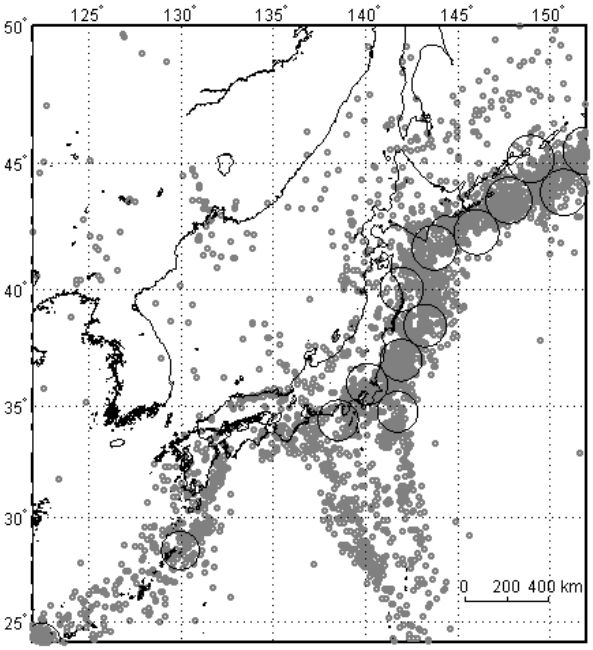



**Fig. 4** Flowchart showing basic steps in the sampling algorithm, $d$ is the Euclidian distance in $IR^2$, the exponent $c$ denotes the complementary set whereas *card* denotes the number of element in the set (cardinal number). $\left(x^{(i)}\right)_{i=0,1,\ldots,m}$, $R$ and $\left\{D\left(x^{(i)},R\right)\right\}_{i=0,1,\ldots,m}$ are the target events, the sampling radius and the sampling disks, respectively

**(a)**




**(b)**

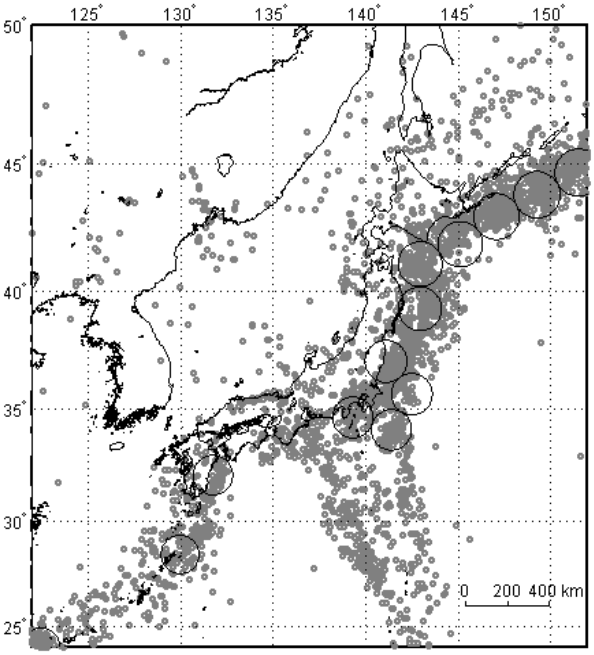



**Fig. 5** ERS(100) strategy showing sampling disks with 100 km radius for scheme 3 of JMA data in Table 1. **(a)** and **(b)** correspond to two different runs. Only sampling disks containing more than 50 events are shown

To describe Bak's scaling law analytically, the recurrence times $\tau$ are scaled with the inverse of their mean $1/\bar{\tau}$. Then, the distribution is written as,

$$\bar{\tau}\, D(\tau) \approx \phi(\tau/\bar{\tau}) \qquad (3)$$

where $\phi$ is the scaling (or the unified) law. Note that $1/\overline{\tau}$ is basically the mean seismic rate $\overline{R}_\tau$ used by Corral (2003, 2004a). Indeed, consider the mixed sample $\left(\tau_j^{(i)}\right)_{j=1,\ldots,n_i}^{i=0,1,\ldots,m}$, we have on one hand,

$$\overline{\tau}^{-1} = \frac{\sum_{i=0}^{m} n_i}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} \tag{4}$$

where $m$ is the number of sampling disks, $n_i$ the number of recurrence times issued from the $i^{\text{th}}$ sampling disk and $\tau_j^{(i)}$ the $j^{\text{th}}$ recurrence time issued from the $i^{\text{th}}$ sampling disk.

On the another hand, $\overline{R}_\tau$ is written,

$$\overline{R}_\tau = \frac{\sum_{i=0}^{m} (n_i + 1)}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} = \frac{\sum_{i=0}^{m} n_i}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} + \frac{m}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} \tag{5}$$

Since each target disk contains at least 50 recurrence times, the second term at the right hand of the former equation is at least 50 times less than the first one. Therefore, it is neglected,

$$\frac{m}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} \leq \frac{1}{50} \frac{\sum_{i=0}^{m} n_i}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} \tag{6}$$

Finally, we obtain the following approximation:

$$\overline{R}_\tau \approx \frac{\sum_{i=0}^{m} n_i}{\sum_{i=0}^{m}\sum_{j=1}^{n_i} \tau_j^{(i)}} \tag{7}$$

From (4) and (7), it follows that,

$$\overline{\tau}^{-1} \approx \overline{R}_\tau \tag{8}$$

The main assumption we try to discuss here is the doubly PL behavior of $D$ holding with a kink around the mean recurrence time $\overline{\tau}$, that is $D$ scales as in (3) with

$$\phi(\tau/\bar{\tau}) = \begin{cases} c_1/(\tau/\bar{\tau})^{p_1} & 0 < \tau < \bar{\tau} - \varepsilon \\ \bar{\tau}\, D(\tau) & \bar{\tau} - \varepsilon \le \tau < \bar{\tau} + \varepsilon \\ c_2/(\tau/\bar{\tau})^{p_2} & \tau \ge \bar{\tau} + \varepsilon \end{cases} \quad (9)$$

in which $c_1$, $c_2$, $p_1$, $p_2$>0 are the unknown PLs parameters and $\varepsilon$ is a small positive constant. Note that to model the position of the kink between the two PLs which may be observed at different positions for different regions (Corral, 2004a, in his Figure 2), the analytic form of $\phi$ is not specified for $\tau \in [\bar{\tau} - \varepsilon,\ \bar{\tau} + \varepsilon[$.

Since $D$ is a probability density function (pdf), its sum over $[0,\ +\infty[$ should be equal to 1. This normalization condition leads, after some calculations, to the following characteristic approximation:

$$\frac{c_1}{1-p_1} + \frac{c_2}{p_2 - 1} + (c_1 p_1 - c_2 p_2)\left(\frac{\varepsilon}{\bar{\tau}}\right)^2 \approx 1 \quad (10)$$

The derivation of the expression (10) is developed in the Appendix C. For $\varepsilon << \bar{\tau}$, the first two terms in this expression prevail and we obtain,

$$\psi(c_1, c_2 p_1, p_2) = \frac{c_1}{1-p_1} + \frac{c_2}{p_2 - 1} \approx 1 \quad (11)$$

This approximation describes the link between the PL exponent at short range $p_1$ and the one at long range $p_2$. Here, $c_1$ and $c_2$ are linked to the probability at the distribution tails. Indeed, from equations (C8) in the appendix C, it follows that $c_1 \approx P(\tau \le \tau_0)$ with $\tau_0 = \bar{\tau}(1-p_1)^{\frac{1}{1-p_1}}$ and similarly from (C6) we deduce that $c_2 \approx P(\tau > \tau_1)$ with $\tau_1 = \bar{\tau}(p_2 - 1)^{\frac{1}{1-p_2}}$. Note that the approximation (11) could be derived simply from two PLs considered for the whole time scale with a fixed kink at $\bar{\tau}$. However, in the expression (10), the term $(c_1 p_1 - c_2 p_2)\left(\frac{\varepsilon}{\bar{\tau}}\right)^2$ results from modeling more reasonable situation in which the kink is variable or difficult to locate. In addition, the approximation (11) provides a straightforward way for testing the stability of the scaling defined in (3, 9). Indeed, for a given set of the parameters $c_1$, $c_2$, $p_1$ and $p_2$, any significant deviation of the statistic $\psi(c_1, c_2, p_1, p_2)$ from its theoretical value 1 will call for the rejection of the PL scaling defined in (3, 9). In the following section, the mean estimates of the parameters $c_1$, $c_2$, $p_1$ and $p_2$ are computed and discussed for different regions and different magnitude thresholds. Based on the obtained results, empirical estimates of $\psi$ are computed and used to discuss the scaling (3, 9).

## 3 Results and Discussion

In their pioneering work, Bak et al. (2002) discarded empirically interevent times less than 40 seconds from their analysis to avoid missing events at short time scales. However, recent studies show that 40 seconds is not sufficient and many events are missed above this level. Indeed, the magnitude of completeness significantly increases after large events and many short-term events are missed from the catalog in the early part of the aftershock sequence at a level exceeding 40 seconds (Kagan 2004; Helmstetter et al. 2006). For this reason and given that censuring a large portion of the data will lead to poor statistics, we adopted the conservative interevent time cutoff $\tau_c$ = 0.2 days obtained for $m_0$ = 7.5 and $m_c$ = 3.5 from the following reversed version of the relation proposed by Helmstetter et al (2006).

$$\log_{10}(\tau_c) = \frac{m_0 - m_c - 4.5}{0.75} \tag{12}$$

where $\tau_c$ is the recurrence time cutoff, $m_0$ the mainshock magnitude and $m_c$ the threshold magnitude.

Interevent times were sampled using ERS with the scheme parameters shown in Table 1. Note that the choice of $R$ does not affect the shape of the distribution $D$ (e.g., Figure 6) and that the radiuses in Table 1, has been selected to guarantee a sufficient number of events in each target disk.

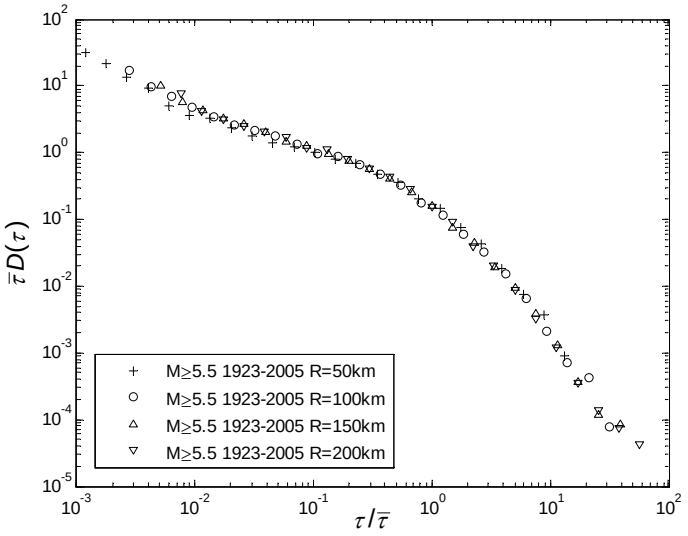



**Fig. 6** Scaled interevent time distributions for $m_c$ = 5.5 obtained using ERS($R$) with different sampling radiuses $R$ = 50, 100, 150, 200 km.

We performed 1, 10, 50 and 100 ERS runs for magnitude thresholds $m_c$ = 2.5, 3.5, 4.5(or 4.7) and 5.5, respectively. Sampling disks containing less than 50 events and interevent times below $\tau_c$ = 0.2 days were not considered in the analysis. The distribution $D$ has been estimated for Southern California and Japan data using the usual binning technique (Bak et al. 2002). First, the distribution $D$ was estimated for each magnitude threshold $m_c$. Then, all the distributions were stacked together and fitted using the least square linear regression on the log-log scale. The cases $\tau/\bar{\tau} < 1$ and $\tau/\bar{\tau} > 1$ were

considered separately to estimate the first and the second PL, respectively. The resulting PLs explain more than 97% of the data dispersion in all the cases as shown in Figure 7.

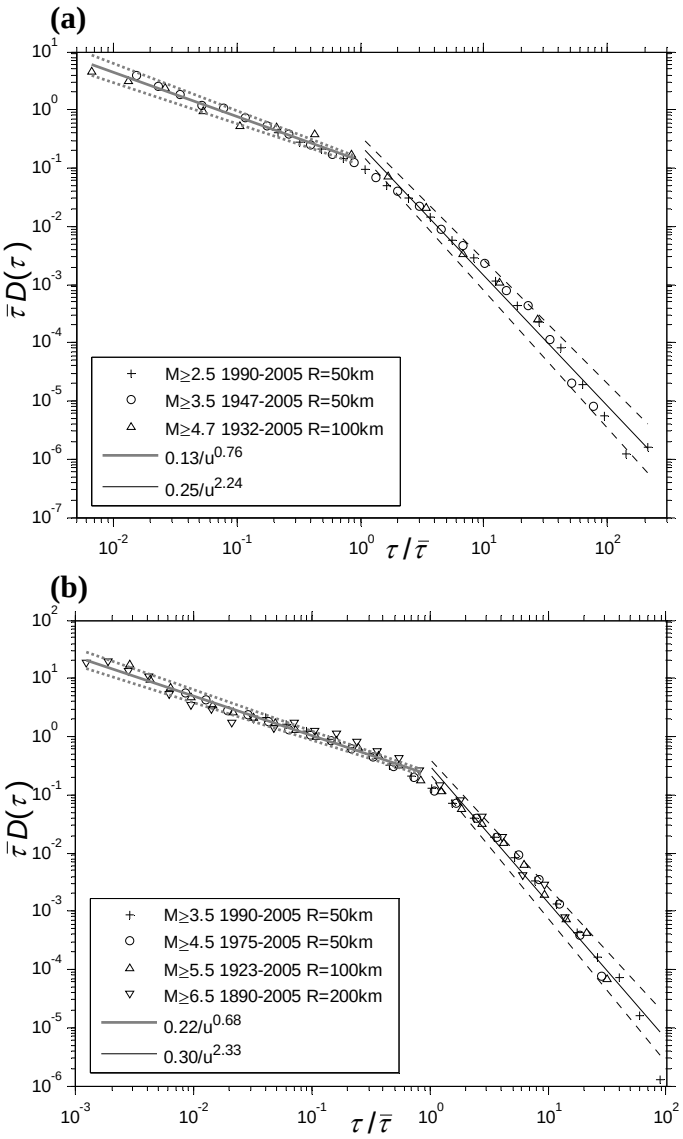



**Fig. 7** Scaled interevent time distribution obtained using ERS($R$) for **(a)** Southern California and **(b)** Japan. Solid lines show tendency to power law whereas dashed lines show the 95% confidence limit of the linear least square adjustment

The solid lines shown on Figure 7 are the regression lines of the first and second PL. Dashed lines correspond to the 95% confidence limits from the confidence intervals of the regression parameters.

The PL parameter estimates and their confidence intervals are summarized in Figure 8. The parameter estimates found by Corral (2004a, his Figure 2) for several regions in the world are also shown in the figure. Clearly, the first PL parameters, $c_1$ and $p_1$, are not stable since the rectangular areas I($c_1$)×I($p_1$) (where I($c_1$) is the confidence interval of $c_1$ and I($p_1$) is that of $p_1$) corresponding to Japan and Southern California in Figure 8 do not intersect. This is further confirmed from Corral's results showing a critical estimate of $p_1$ (around 1). On the other hand, the second PL parameters $c_2$ and $p_2$

are largely variable and their estimates are located roughly near the borders of the intersection zone between the rectangular areas I($c_2$)×I($p_2$) (where I($c_2$) is the confidence interval of $c_2$ and I($p_2$) is that of $p_2$). Corral's second PL estimate is located exactly on the border of the intersection zone I($c_2$)×I($p_2$) and is likely to be significantly different from the estimate for Japan. The situation can be clearer if the precision of Corral's results (confidence limits) was given. Yet, in our study the confidence limits of the PL exponents are assessed for the first time. Discrepancies in the second PL are difficult to evaluate significantly and need more investigation because of the deficit of large interevent times. As a general note, we highlight the slower decay and the steeper one predicted for short and large time ranges respectively in case of Japan (Fig. 8). These arguments do not support a unique scaling law for all the regions, and hence we conclude that in general the scaling parameters are more likely to be dependent on the region of study.

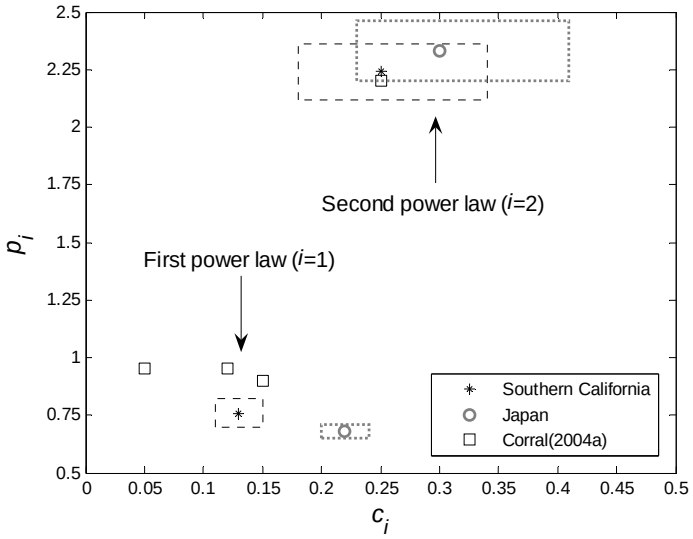



**Fig. 8** Variation of the mean power law parameter estimates for different regions. $c_1$ and $c_2$ are shown in the abscises whereas $p_1$ and $p_2$ are shown in the ordinates. Dashed lines are the 95% confidence rectangles

To test the sensitivity of the PL parameters to different threshold magnitudes, their mean values for each magnitude scale were computed after several ERS runs. One hundred (100) ERS runs were performed for events exceeding magnitudes 2.5, 3.5, 4.5 and 5.5, whereas for events exceeding the magnitude 6.5, 500 runs were carried out to achieve stable estimates of the PL parameters. These mean PL parameters were sorted in ascending order and their 95% confidence intervals estimated nonparametrically. For example, if $p_{(1)}, p_{(2)}, \ldots, p_{(100)}$ is the ordered sample of $p$, the lower and upper confidence interval bounds of the mean value $\overline{p} = \frac{1}{100}\sum_{i=1}^{100} p_i$ are $(p_{(5)} + p_{(6)})/2$ and $(p_{(95)} + p_{(96)})/2$, respectively. The obtained estimates are summarized in Figures 9 and 10, together with their 95% confidence intervals.

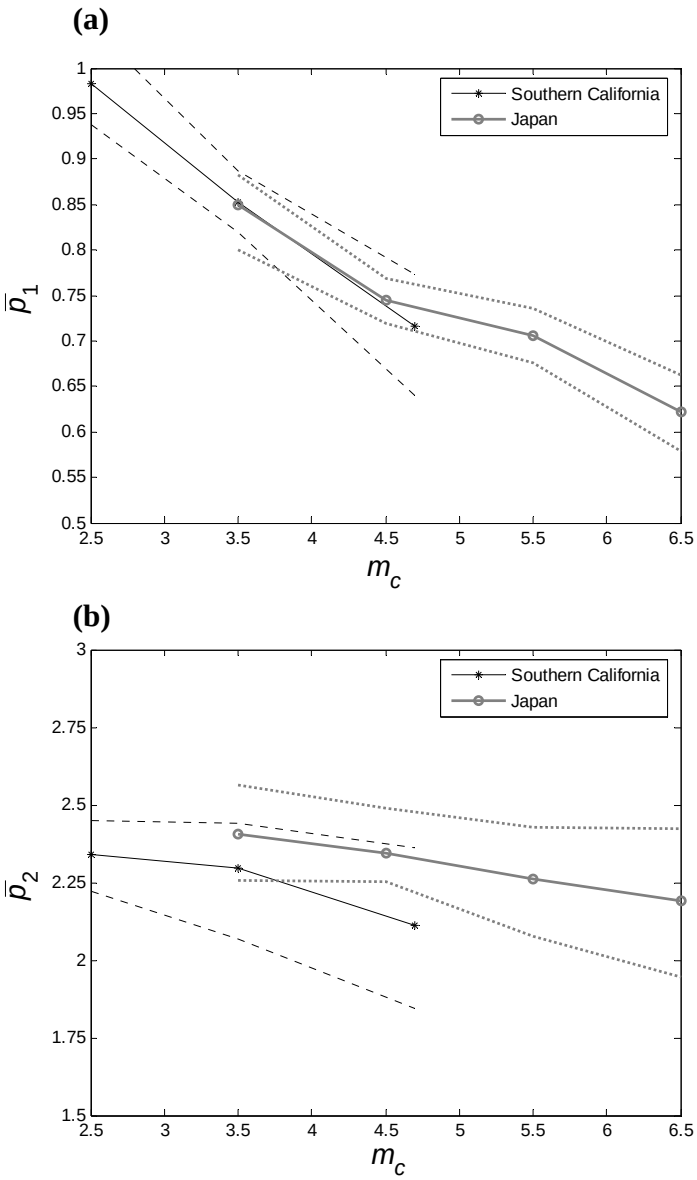



**Fig. 9** Variation of the mean estimates of **(a)** $p_1$ and **(b)** $p_2$ for different magnitude thresholds. Linear trends are found for $\overline{p}_1$ in case of Southern California and $\overline{p}_2$ in case of Japan. *Dashed lines* show the 95% confidence limits of the parameter estimates

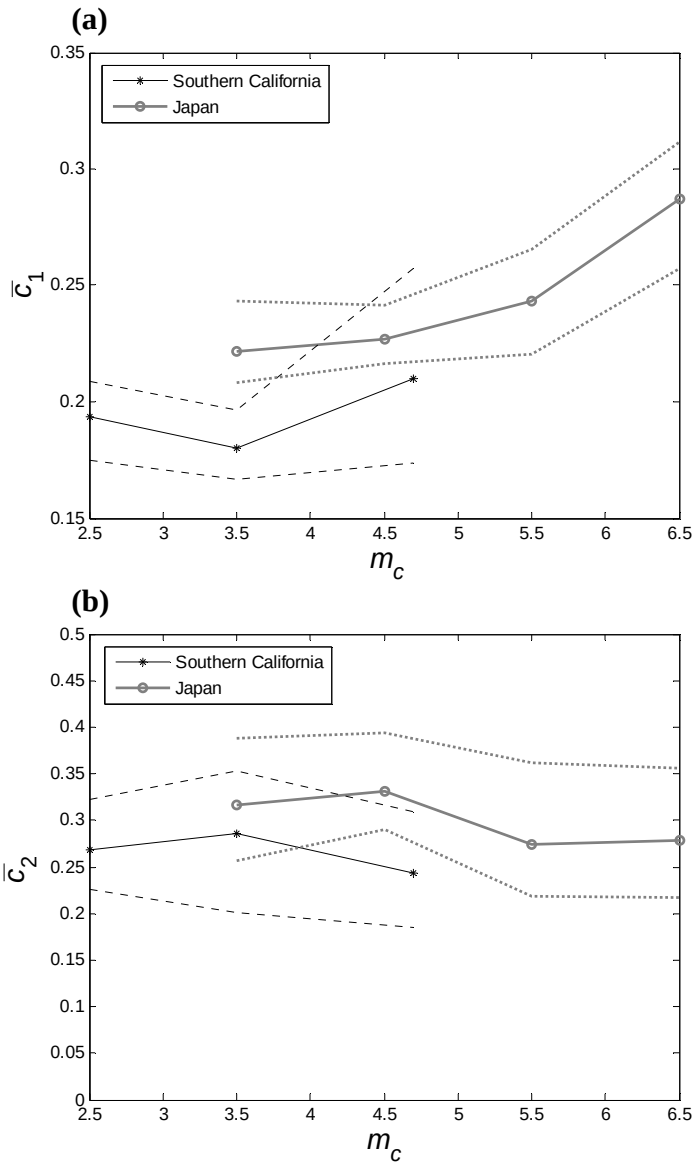



**Fig. 10** Variation of the mean estimates of **(a)** $c_1$ and **(b)** $c_2$ for different magnitude thresholds. $c_1$ is generally increasing for intermediate and large magnitudes, whereas $c_2$ is relatively stable. Dashed lines show the 95% confidence limits of the parameter estimates

In both cases, $p_1$ decreases with increasing magnitudes from about 0.98 to 0.62. The decrease at short and intermediate magnitudes ($m_c$=2.5, 3.5 and 4.5) is linear with a slope of about -0.12 followed by much slower decrease at large magnitudes (Fig. 9a). It follows that $p_1$ is not unique for all the magnitude scales. On the other hand, $p_2$, in most cases, slightly decreases with increasing magnitudes (Fig. 9b). Strikingly, the decrease in the case of Japan is roughly described by the following equation:

$$p_2(m_c) = -0.07m_c + 2.7; \quad m_c \geq 3.5 \tag{13}$$

The fluctuations of the parameters $c_1$ and $c_2$ for different magnitudes are summarized in Figures 10. $c_1$ generally increases with increasing magnitudes (for moderate and large magnitudes). This trend may be explained by the decreasing number of events in the schemes with large magnitude cutoffs, thus resulting in a larger mean interevent time.

Consequently, the distribution shifts to the left and the probability at the left tail increases. $c_2$, which is supposed to decrease for the same reasons, shows quite stable estimates, around 0.30, because of the poor data around the upper tail (Fig. 10b).

Because the approximation (11) is a necessary condition for the scaling defined in (3, 9) to hold, it was used to test an eventual departure from such scaling for different magnitudes. Mean estimates of $\psi$ are shown for $m_c$=3.5, 4.5, 5.5 and 6.5 in Figure 11.

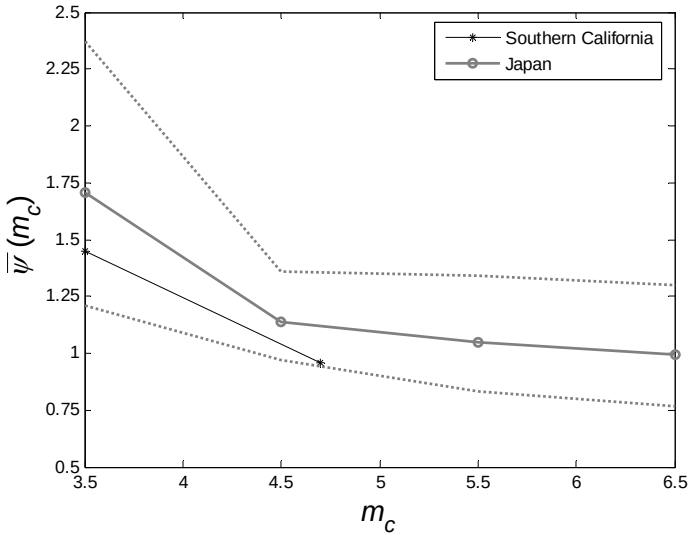



**Fig. 11** Mean estimates of $\psi$ for different magnitude thresholds. $\overline{\psi}$ is greater than one at small magnitudes while decreasing to one with increasing magnitudes

The results corresponding to small magnitudes cutoff $m_c$=2.5 were excluded because $p_1$ takes critical values close to 1 (Fig. 9a) and $\psi$ is unstable. The confidence intervals are shown only for Japan because of the high fluctuations otherwise. Typically, this statistics is significantly greater than 1 for small magnitudes while it decreases to 1 with increasing magnitudes. Note that the consideration of the approximation (10) will not change the results. Indeed using the parameter estimates for different magnitudes shown in Figures 9 and 10, we found in all the cases that,

$$\left|(c_1 p_1 - c_2 p_2)\left(\frac{\varepsilon}{\overline{\tau}}\right)^2\right| \le 2\times 10^{-4}\varepsilon^2; \quad m_c \ge 3.5 \qquad (14)$$

The upper bound in (14) is far below the deviation observed in Figure 11 for a large range of $\varepsilon$ values. It follows that the scaling (3, 9) do not hold for small magnitudes as a result of the deviation between $\psi$ and 1. Such deviation can be attributed to the high fluctuations in interevent times for small magnitude cutoffs. This conclusion reinforce the results by Saichev and Sornette (2007) showing that the PL scaling is broken by the Omori-Utsu clustering and do not hold strictly.

Alternatively to the PL fit, the Southern California data set was fitted by a Weibull distribution (Weibull, 1951), defined by the probability distribution function:

$$f_W(\tau;k,\lambda)=k\lambda^{-k}\tau^{k-1}e^{-\left(\frac{\tau}{\lambda}\right)^k};\quad \tau>0 \tag{15}$$

where $k$ and $\lambda$ are positive parameters. Weibull distribution attracts growing attention as a recurrence time model for large earthquakes (e.g. Newman et al. 2005; Yakovlev et al. 2006; Turcotte et al. 2007; Zoller and Hainzl 2007). Figure 12a shows the corresponding fit with maximum likelihood estimates of the parameters $k$ and $\lambda$ together with the PLs of Figure7a. Comparing with PLs, this distribution provides a better fit for large recurrence times $\tau > 0.3\times\bar{\tau}$ (Fig. 12b), with a residual sum of squares $RSS(\tau > 0.3\bar{\tau})$ equal to $7.7\times10^{-4}$ against about $10^{-3}$ for the PL fit of Figure 7a.

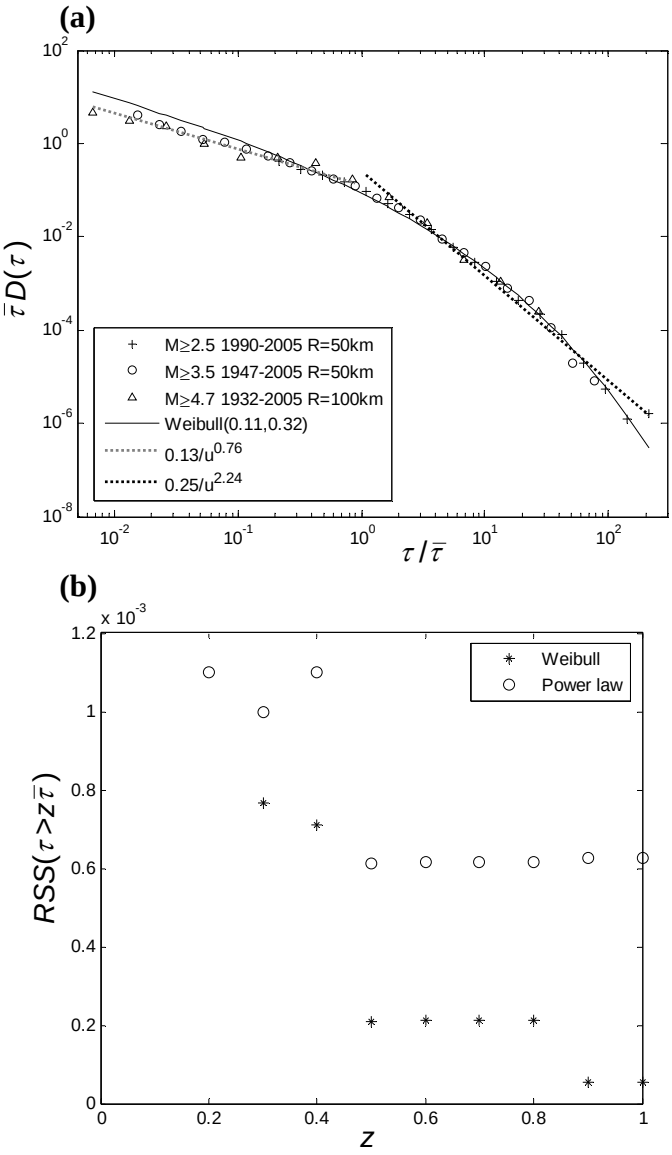



**Fig. 12 (a)** Weibull distribution with parameters $k$ = 0.11 and $\lambda$ = 0.32 and the PLs of Figure 7a fitted to Southern California data. **(b)** Residual sum of squares $RSS(\tau > z\bar{\tau})$ for normalized recurrence times $\tau/\bar{\tau}$ greater than $z$, corresponding to Weibull distribution in Figure 12a and the doubly power law fit in Figure 7a. Weibull distribution fits well the data starting from about 0.3 standardized interevent time

## 4 Conclusion

A robust estimation of the earthquake recurrence time distribution was carried out using a three step procedure. First, for the data set from Southern California and Japan, the magnitudes of completeness were assessed and a series of complete seismicity periods were selected. Then, a new sampling strategy for interevent times was applied to the data. Secondly, the distributions of the obtained interevent times were empirically assessed for tendency to PLs and their parameter for different regions were estimated. Finally, the parameterization of recurrence time distribution was discussed in two cases: when a doubly PL behavior is assumed, and in the case where PL parameters vary with different magnitude cutoffs.

The corresponding results showed that PL parameter estimates and their fluctuations are not consistent with a unique universal behavior. Significantly different estimates were found to govern the scaling for different magnitude cutoffs in these regions. In particular, the exponents of the claimed PL scaling generally decrease with increasing magnitudes. Also, most PL parameters are different from one region to another when averaged over all magnitude scales in a given region.

The hypothesis of a doubly PL scaling lead to a simple normalization condition that breaks for small magnitude thresholds, so that the scaling cannot hold for this magnitude range. It is also found that Weibull distribution fits the data quite well in case of Southern California, in particular for long time ranges where departure from the claimed PL scaling is difficult to evaluate. These results suggest that careful assessment of the magnitude of completeness and the interevent time cutoff should be taken into account before the interpretation of any empirical results concerning the recurrence time distribution.

## 5 Appendix

### Appendix A: Stepp's Approach (Stepp, 1971)

This method is based on the assumption of a stationary rate of the declustered catalog. After intensive tests on both catalogs, we adopted Gardner and Knopoff (1974) approach with the standard windows for Southern California. It is found more simple and stable than other methods and its windows parameters suitable for removing large fluctuations of seismic activity in space and time. Next, consider the magnitude classes $[m_i, m_{i+1}]$ with a fixed width ($\Delta m = m_{i+1}-m_i = 1$; $i = 1, 2, \ldots, I$). In that case, each magnitude class is assumed to be an independent point process in time following an homogeneous Poisson distribution with a constant rate $\lambda$,

$$P(T = k) = \frac{e^{-\lambda}\lambda^{k}}{k!} \qquad \text{(A1)}$$

where $T$ is the time period measured in years. For a time period $T = n$, if $k_1, k_2, \ldots, k_n$, are the mean number of earthquakes in each unit time interval of one year in the considered magnitude class. Then, an unbiased estimate of the mean rate of occurrence is given by:

$$\hat{\lambda} = \bar{k} = \frac{\sum_{i=1}^{T} k_i}{T} \qquad \text{(A2)}$$

with a standard deviation

$$\sigma\left(\hat{\lambda}\right) = \frac{\sqrt{\lambda}}{\sqrt{T}} \qquad \text{(A3)}$$

Assuming a stationary process and thus a constant rate λ, $\sigma\left(\hat{\lambda}\right)$ behaves as $\frac{1}{\sqrt{T}}$ for each magnitude class. Thus, its variation against time $T$ is linear on a log-log scale; in particular it is made with a parallel to the reference line

$$y = -\frac{1}{2}\log(x). \qquad \text{(A4)}$$

In practice, since recent time periods are more likely to be complete and with stationary rates, we start from the most recent time period of length $\Delta T$ (3 or 5 years for example) and step backward with a fixed lap $\Delta T$, measuring for each $T$ the standard deviation $\sigma\left(\hat{\lambda}\right)$. Figure 13 shows the plot of $\sigma\left(\hat{\lambda}\right)$ against $T$ with the corresponding empirical adjustment, for Japan seismicity with magnitude $m \in [3.5, 4.5]$.

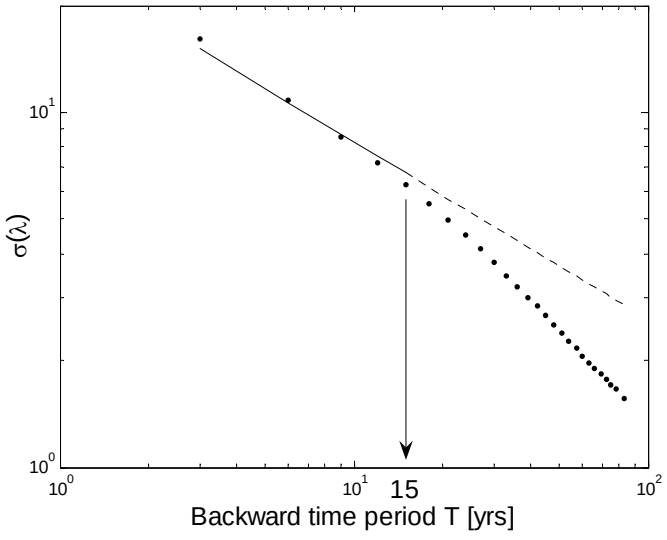



**Fig. 13** Variability of the occurrence rate $\sigma(\lambda)$ for events with magnitude $m \in [3.5\ 4.5]$, starting from the most recent period of length $\Delta T$ = 3 years, that is 2002-2005 and moving backward with a fixed step of ($T(i) = i\Delta T$). The linear adjustment explains quite well about five points. This corresponds to a period of $5\Delta T$ = 15 years, so that 1990-2005 is the maximal period during which earthquakes have been reported with a constant rate and hence completely reported.

## Appendix B: Earthquake Random Sampling

Lets assume that earthquakes occur as a point process $N$ in the three dimensional space of time-location (e.g., Cox and Isham, 1980; Daley and Vere-Jones, 1988), the magnitude and depth are disregarded for simplicity. Consider the occurrence of $N$ on the metric space $(E,d)$; where $E$ is a subset of $IR^2$ representing the set of epicenters identified by the pairs (longitude, latitude), and $d$ the Euclidian distance in $IR^2$. For a fixed real bound $r>0$, we construct a sub-process $\tilde{N}^{(r)}$ from the process $N$ as follows:

**1**. $X^{(0)}$ is a random point of the process $N$.

**2.** $X^{(1)}$ is the nearest neighbor of $X^{(0)}$ at a distance exceeding or equal to $r$ from $X^{(0)}$

i.e.,

$$d\left(X^{(0)},X^{(1)}\right)=\min\left\{d\left(X^{(0)},y\right);\, d\left(X^{(0)},y\right)\geq r,\, y\in N\subset IR^2\right\} \qquad \text{(B1)}$$

**3.** $X^{(2)}$ is the nearest neighbor of $X^{(1)}$ at a distance exceeding or equal to $r$ from $X^{(0)}$ and $X^{(1)}$ i.e.,

$$d\left(X^{(1)},X^{(2)}\right)=\min\left\{d\left(X^{(1)},y\right);\, y\notin D\left(X^{(0)},r\right)\cup D\left(X^{(1)},r\right),\, y\in N\subset IR^2\right\} \text{ (B2)}$$

$D\left(X^{(0)},r\right)$ and $D\left(X^{(1)},r\right)$ are the disks with radius $r$, centered on $X^{(0)}$ and $X^{(1)}$ respectively.

$X^{(3)}, X^{(4)},\ldots, X^{(i-2)}, X^{(i-1)},\ldots$ are defined in a similar way. For example, $X^{(i-1)}$ is defined in the step $i$ as follows:

***i.*** $X^{(i-1)}$ is the nearest neighbor of $X^{(i-2)}$ not belonging to the set of disks $\left\{D\left(X^{(j)},r\right)\right\}_{j=0,1,\ldots,i-2}$ (Figure 14).

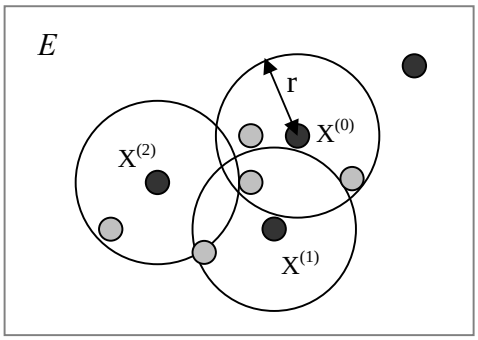



**Fig. 14** Schematic representation of target points construction. From all earthquake locations (balls), black balls are selected as target points.

For a finite point process $N$, if the disk $D\left(X^{(0)},r\right)$ do not include all the points, the algorithm stop in the step $m+2$ with $m\geq 0$, for which the nearest neighbors of $X^{(m)}$ belong to the set of disks $\left\{D\left(X^{(j)},r\right)\right\}_{j=0,1,\ldots,m}$. The set of points constructed earlier (defining the sub-process $\tilde{N}^{(r)}$) are called *r-target points*. The set of disks $\left\{D\left(X^{(j)},r\right)\right\}_{j=0,1,\ldots,m}$ are called *r-target disks*.

The ERS procedure has been built on as a tool for the estimation of earthquake recurrence time distribution $D$ in a given region. For this purpose, consider the set of $r$-target points $(x^{(i)})_{i=0,1,\ldots,m}$, then $x^{(0)}$ is a random epicenter corresponding to an event with magnitude $m_0$. Interevent times are calculated for each disk $D\left(x^{(i)}, R\right)$ centered on $x^{(i)}$ with radius $R=r/2$. Then, the $m$ samples $\left(\tau_j^{(i)}\right)_{j=1,\ldots,n_i}; i=0,1,\ldots,m$ obtained are mixed together to estimate $D$. We note $ERS(R)$ the *Epicenter Random Sampling* with sampling radius $R$. Note that resampling with ERS provides different interevent times in each run because of the randomly chosen starting target point $x^{(0)}$.

## Appendix C: Derivation of the approximation (10)

The normalization condition for $D$ reads,

$$\int_0^{\infty} D(s)ds = 1 \tag{C1}$$

In view of (3) and (9) and using the additive propriety of integration, the summation (C1) shares into,

$$\int_0^{\infty} D(s)ds = \int_0^{\bar{\tau}-\varepsilon} D(s)ds + \int_{\bar{\tau}-\varepsilon}^{\bar{\tau}+\varepsilon} D(s)ds + \int_{\bar{\tau}+\varepsilon}^{\infty} D(s)ds \tag{C2}$$

Let's note $I_1(D)$, $I_2(D)$ and $I_3(D)$ the first, the second and the third terms in the right hand of the equation (C2), respectively. Calculating these terms calls to the evaluation of $D(s)$ for $0 < s < \bar{\tau} - \varepsilon$, $\bar{\tau} - \varepsilon \le s < \bar{\tau} + \varepsilon$ and $s \ge \bar{\tau} + \varepsilon$. Let's consider first the case where $s \ge \bar{\tau} + \varepsilon$. From equations (3) and (9) we deduce that,

$$\bar{\tau}\, D(s) \approx \frac{c_2}{\left(s/\bar{\tau}\right)^{p_2}} \tag{C3}$$

Hence,

$$D(s) \approx \frac{c_2 \bar{\tau}^{\,p_2-1}}{s^{\,p_2}} \tag{C4}$$

By summing both terms in the former equation over $[s, +\infty[$, we obtain,

$$\int_s^{+\infty} D(u)du \approx \int_s^{+\infty} \frac{c_2 \bar{\tau}^{\,p_2-1}}{u^{\,p_2}} du \tag{C5}$$

It follows that,

$$1 - F_{\tau}(s) \approx c_2 \bar{\tau}^{\,p_2-1} \frac{s^{1-p_2}}{p_2 - 1} \tag{C6}$$

with $p_2>1$. $F_\tau$ is the cumulative probability distribution. $I_3(D)$ is calculated from equation (C6) by putting $s = \bar{\tau} + \varepsilon$,

$$I_3(D) = 1 - F_\tau(\bar{\tau} + \varepsilon) \approx \frac{c_2}{p_2 - 1} \bar{\tau}^{p_2 - 1} (\bar{\tau} + \varepsilon)^{1 - p_2} \qquad \text{(C7)}$$

where $p_2>1$.

In the same way, for the case where $0 < s < \bar{\tau} - \varepsilon$, summing $D$ between 0 and $\bar{\tau} - \varepsilon$ yields,

$$F_\tau(s) \approx c_1 \bar{\tau}^{p_1 - 1} \frac{s^{1 - p_1}}{1 - p_1} \qquad \text{(C8)}$$

with $p_1<1$. $I_1(D)$ is obtained from the former approximation for $s = \bar{\tau} - \varepsilon$,

$$I_1(D) = F_\tau(\bar{\tau} - \varepsilon) \approx \frac{c_1}{1 - p_1} \bar{\tau}^{p_1 - 1} (\bar{\tau} - \varepsilon)^{1 - p_1} \qquad \text{(C9)}$$

Finally we get from the summation of both hands in the approximations (C7) and (C9) that,

$$I_1(D) + I_3(D) \approx \frac{c_1}{1 - p_1} \left(1 - \frac{\varepsilon}{\bar{\tau}}\right)^{1 - p_1} + \frac{c_2}{p_2 - 1} \left(1 + \frac{\varepsilon}{\bar{\tau}}\right)^{1 - p_2} \qquad \text{(C10)}$$

On the other hand, if we suppose $D$ slightly varying (relatively constant) within the interval $[\bar{\tau} - \varepsilon,\ \bar{\tau} + \varepsilon]$, $I_2(D)$ could be approximated as follows,

$$I_2(D) = \int_{\bar{\tau} - \varepsilon}^{\bar{\tau} + \varepsilon} D(u)\, du \approx 2\varepsilon \hat{D}(\bar{\tau}) \qquad \text{(C11)}$$

where $\hat{D}(\bar{\tau})$ is the mean value of the distribution between $\bar{\tau} - \varepsilon$ and $\bar{\tau} + \varepsilon$, that is,

$$\hat{D}(\bar{\tau}) \approx \frac{1}{2} \left( \frac{c_1 \bar{\tau}^{p_1 - 1}}{(\bar{\tau} - \varepsilon)^{p_1}} + \frac{c_2 \bar{\tau}^{p_2 - 1}}{(\bar{\tau} + \varepsilon)^{p_2}} \right) \qquad \text{(C12)}$$

Suppose $\varepsilon << \bar{\tau}$, using the approximation,

$$(1 + x)^\alpha \approx 1 + \alpha x;\ x \to 0, \alpha \in IR - \{0\} \qquad \text{(C13)}$$

we get,

$$2\hat{D}(\bar{\tau}) = c_1 \bar{\tau}^{-1}\left(1 + p_1 \frac{\varepsilon}{\bar{\tau}}\right) + c_2 \bar{\tau}^{-1}\left(1 - p_2 \frac{\varepsilon}{\bar{\tau}}\right) \quad \text{(C14)}$$

Finally,

$$2\hat{D}(\bar{\tau}) = (c_1 + c_2)\frac{1}{\bar{\tau}} + (c_1 p_1 - c_2 p_2)\frac{\varepsilon}{\bar{\tau}^2} \quad \text{(C14')}$$

By substitution from (C14') and (C11), we obtain $I_2(D)$,

$$I_2(D) \approx (c_1 + c_2)\frac{\varepsilon}{\bar{\tau}} + (c_1 p_1 - c_2 p_2)\left(\frac{\varepsilon}{\bar{\tau}}\right)^2 \quad \text{(C15)}$$

Again using the approximation (C13) for $\varepsilon << \bar{\tau}$, the expression (C10) could be written as,

$$I_1(D) + I_3(D) \approx \frac{c_1}{1 - p_1}\left(1 - (1 - p_1)\frac{\varepsilon}{\bar{\tau}}\right) + \frac{c_2}{p_2 - 1}\left(1 + (1 - p_2)\frac{\varepsilon}{\bar{\tau}}\right) \quad \text{(C16)}$$

Finally, from (C1), (C2), (C15) and (C16), and after a simple calculation, the expression (10) is obtained.

## 6 Acknowledgments

This work has been supported by the scholarship from Ministry of Education, Culture, Sport, Science and Technology (Monbukagusho), the government of Japan, and by the Algerian Centre de Recherche en Astronomie Astrophysique et Geophysique (CRAAG). The authors are grateful to the Japanese Meteorological Agency (JMA) for providing the data used in this study. The authors thank two anonymous reviewers for their comments that improved an earlier version of the manuscript.

## 7 References

Bak P, Christensen K, Danon L, Scanlon T (2002) Unified scaling law for earthquakes. Phys Rev Lett 88, DOI 10.1103/PhysRevLett.88.178501

Bender B (1983) Maximum likelihood estimation of *b*-values for magnitude grouped data. Bull Seismol Soc Am 73(3):831–851.

Carbone V, Sorriso-Valvo L, Harabaglia P, Guerra I (2005) Unified scaling law for waiting times between seismic events. Europhys Lett: 71:1036–1042, DOI 10.1209/epl/i2005-10185-0

Christensen K, Danon L, Scanlon T, Bak P (2002) Unified scaling law for earthquakes. Proc Natl Acad Sci USA 99:2509–2513.

Corral A (2003) Local distributions and rate fluctuations in a unified scaling law for earthquakes. Phys Rev E 68, DOI 10.1103/PhysRevE.68.035102

Corral A (2004a) Universal local versus unified global scaling laws in the statistics of seismicity. Physica A 340:590–597. DOI 10.1016/j.physa.2004.05.010

Corral A (2004b) Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys Rev Lett 92**,** DOI 10.1103/PhysRevLett.92.108501

Corral A (2007) Statistical features of earthquake temporal occurrence. Lecture notes in physics 705:191–221, DOI 10.1007/3-540-35375-5_8, Springer, Berlin Heidelberg.

Corral A, Christensen K (2006) Comment on "Earthquakes descaled: on waiting time distributions and scaling laws". Phys Rev Lett 96, DOI 10.1103/ PhysRevLett.96. 109801

Cox DR, Isham V (1980) Point processes. Chapman and Hall, London

Daley DJ, Vere-Jones D (1988) An introduction to the theory of point processes. Springer-Verlag, New York

Davidsen J, Goltz C (2004) Are seismic waiting time distributions universal? Geophys Res Lett 31(21):L21612, DOI 10.1029/2004GL020892

Gardner, JK., Knopoff L (1974) Is the sequence of aftershocks in Southern California, with aftershocks removed, Poissonian? Bull Seismol Soc Am 64(5):1363–1367

Hainzl S, Scherbaum F, Beauval C (2006) Estimating background activity based on interevent-time distribution. Bull Seismol Soc Am 96(1):313–320

Helmstetter A, Kagan YY, Jackson DD (2006) Comparison of short-term and time-independent earthquake forecast models for Southern California. Bull Seismol Soc Am 96(1):90–106

Kagan YY (2004) Short-term proprieties of earthquake catalogs and models of earthquake source. Bull Seismol Soc Am 94(4):1207–1228

Kagan YY, Jackson DD, Rong YF (2006) A new catalog of Southern California earthquakes, 1800-2005. Seismol Res Lett 77(1):30–38

Knopoff L (2000) The magnitude distribution of declustered earthquakes in Southern California. Proc Natl Acad Sci USA 97(22):11880–11884

Lindman M, Jonsdottir K, Roberts R, Lund B, Bodvarsson R (2005) Earthquakes descaled: on waiting time distributions and scaling laws. Phys Rev Lett 94, DOI 10.1103/ PhysRevLett.94.108501

Lindman M, Jonsdottir K, Roberts R, Lund B, Bodvarsson R (2006) Earthquakes descaled: on waiting time distributions and scaling laws (Reply). Phys Rev Lett 96, DOI 10.1103/PhysRevLett.96.109802

Molchan G (2005) Interevent time distribution in seismicity: a theoretical approach. Pure Appl Geophys 162:1135–1150

Molchan G, Kronrod T (2007) Seismic interevent time: a spatial scaling and multifractality. Pure Appl Geophys 164:75–96, DOI 10.1007/s00024-006-0150-y

Newman W, Turcotte DL, Shcherbakov R, Rundle JB (2005) Why Weibull? In: Abstracts of the American Geophysical Union fall meeting, San Francisco, California, 5–9 December 2005

Ogata Y (1988) Statistical models for earthquakes occurrences and residual analysis for point processes. J Am Stat Assoc 83(401):9–27

Saichev A, Sornette D (2006) "Universal" distribution of interearthquake times explained. Phys Rev Lett 97, DOI 10.1103/PhysRevLett.97.078501

Saichev A, Sornette D (2007) Theory of earthquake recurrence times. J Geophys Res 112:B04313, DOI 10.1029/2006JB004536

Stepp JC (1971) An investigation of earthquake risk in the Puget Sound area by use of the of the type 1 distribution of largest extremes. Ph. D. Thesis, State University. of Pennsylvania

Stepp JC (1972) Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard. In: Proceedings of the International Conference on Microzonation, Seattle, WA, 1972, vol 2: 897–910

Talbi A, Yamazaki F (2007) Earthquake waiting time distribution: modeling and scaling law. In: Abstracts of Japan Geosciences Union meeting, Chiba, Japan, 19–24 May 2007

Turcotte DL, Abaimov SG, Shcherbakov R, Rundle JB (2007) Nonlinear dynamics of natural hazards. In: Tsonis, Anastasios A.; Elsner, James B. (eds.) Nonlinear Dynamics in Geosciences, Springer, New York, pp 557–580, DOI 10.1007/978-0-387-34918-3_30

Weibull W (1951) A statistical distribution of wide applicability. J Appl Mech 18(3):293–297

Weichert DH (1980) Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. Bull Seismol Soc Am 70(4):1337–1346

Wiemer S, Wyss M (2000) Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western Unites States, and Japan. Bull Seismol Soc Am 90(4):859–869

Yakovlev G, Turcotte DL, Rundle JB, Rundle PB (2006) Simulation-based distributions of earthquake recurrence times on the San Andreas fault system. Bull Seismol Soc Am 96(6):1995–2007

Zoller G, Hainzl S (2007) Recurrence time distributions of large earthquakes in a stochastic model for coupled fault systems: the role of fault interaction. Bull Seismol Soc Am 97(5):1679–1687