A Mixed Model for Earthquake Interevent Times

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 ; abdelhak_t@yahoo.fr

Abstract A mixed model is proposed to fit earthquake interevent time distribution. In this model, the whole distribution is constructed by mixing the distribution of clustered seismicity, with a suitable distribution of background seismicity. Namely, the fit is tested assuming a clustered seismicity component modeled by a non-homogeneous Poisson process, and a background component modeled using different hypothetical models (Exponential, Gamma and Weibull). For southern California, Japan and Turkey, the best fit is found when a Weibull distribution is implemented as a model for background seismicity. Our study uses Earthquake Random Sampling (ERS) method we introduced recently. It is performed here to account for space-time clustering of earthquakes at different distances from a given source, and to increase the number of samples used to estimate earthquake interevent time distribution and its power law scaling. For Japan, the contribution of clustered pairs of events to the whole distribution is analyzed for different magnitude cutoffs m_c and different time periods. The results show that power laws are mainly produced by the dominance of correlated pairs at small and long time ranges. In particular, both power laws, observed at short and long time ranges, can be attributed to time-space clustering revealed by the standard Gardner and Knopoff's declustering windows.

Keywords Recurrence times · Distribution fit · Mixed model · Palm-Khintchine equations · Power laws · Scaling

1 Introduction

Seismicity temporal patterns remain one of the hottest topics discussed in earth sciences (e.g., Utsu 2002). Nowadays, characteristic signatures of these patterns are resumed in a series of scaling laws (or power laws) depicting pertinent scale invariant behaviors in seismicity. Typical examples, dating back to the last century, are the classical Gutenberg-Richter, Omori, and Båth laws (Omori 1894; Gutenberg and Richter 1942; Utsu 1961; Båth 1965). More recent examples are induced by the fractal nature of plate tectonics,

faults, and hypocenter distribution (e.g., Kagan and Knopoff 1980; Sornette and Pisarenko 2003).

Recently, Bak et al. (2002) and Christensen et al. (2002) used aggregated seismicity from different locations in Southern California to introduce a new scaling law (called hereafter "Bak's scaling law") for earthquake interevent times. The law describing two power law (PL) regimes separated by a characteristic kink, has been generalized and extended to the case of disaggregated seismicity within a single region (Corral 2003, 2004a,b, 2005). However, the scaling law for a single region was different and presented a steeper decrease at large interevent time comparing to Bak's scaling law. These two versions of earthquake interevent time distribution (EITD) are sometimes called local and mixed EITD, in reference to the distribution for a single region and Bak's scaling law which mixes interevent times from different locations. Nowadays, extensive research is focused on the study of EITD and its scaling for broad areas (e.g., Saichev and Sornette 2006, 2007; Corral 2007; Molchan and Kronrod 2007; Talbi and Yamazaki 2009). In particular, many studies reported departures and discrepancies in the claimed scaling laws and their parameterization (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, Talbi and Yamazaki 2009).

Despite the above critics and the broad use of local and global catalogs from all over the world to establish both EITD scaling laws, the fluctuations of their parameterization had rarely been studied (Talbi and Yamazaki 2009). Former studies were restricted to testing the scaling accuracy and in the affirmative to depicting the potential new interesting features it implies, using the nonhomogeneous Poisson process (e.g., Lindman et al. 2005; Molchan 2005) and the epidemic type aftershock sequence (ETAS) (Ogata 1988) model (e.g., Hainzl et al. 2006; Saichev and Sornette 2006, 2007). Early arguments against EITD scaling had been evoked by Molchan (2005) who urged the exponential up tail of the distribution (in sharp contrast with observations) under some general assumptions. The study was taken back by Saichev and Sornette (2006, 2007) who reviewed extensively the scaling issue and showed that the presence of Omori-Utsu clusters destroy the exact unified scaling law while giving an approximate unified law fitting the real data.

Whereas for a single region the local EITD has been fitted with a Gamma distribution since its establishment (Corral, 2003), Bak's scaling law is relatively difficult to fit because of the slow decrease at long interevent times. This former yields a controversial non-exponential decrease, inconsistent with a plausible Poisson model at this range. The first attempt to fit Bak's scaling law, which is mainly motivated by the problematic non-exponential decrease, has been challenged by Saichev and Sornette (2007). Despite their original approach and detailed synthesis, their analysis is restricted to the ETAS model and uses sophisticated mathematical background which is not accessible to non-specialists. Empirical tools combined to nonparametric models could be useful to simplify their view and extrapolate their conclusions to the observed "model-free" seismicity structure. An attempt in this way is given here using a simple model which combines empirical observations with some standard hypothetical distributions to fit the data.

This study is a continuation of our former work Talbi and Yamazaki (2009) mainly based on the study of EITD deviation from the PL behavior. It extrapolates some of our

former results and methodology, to fit EITD assuming the decomposition of seismicity into background and clustered (or dependent) components. We proceed in three steps. First, our data base is filtered using classical declustering algorithms to assess artificially the role of spatio-temporal correlation in the observed scaling. Secondly, we develop analytically our mixed model using point processes Palm theory. Finally we carry on the fit issue in application using several implemented background candidate distributions. As a relevant result, our model is found to fit the data exceptionally well when a Weibull distribution is implemented as interevent time distribution for the background seismicity process.

2 Earthquake catalog preparation and model selection

Main earthquake data files used in this study are described in Talbi and Yamazaki (2009). For Japan, the file includes events from the JMA catalog, and a compiled version of Utsu catalog for the historical period. For southern California, files covering the time period 1932-2005 were compiled from the Southern California National Network, together with Kagan catalog (Kagan et al. 2006). Further details about these data sources and their compilation can be found in Talbi and Yamazaki (2009). Turkey data covering the period 550 BC-2004 were downloaded from Kandilli Observatory website at the link http://www.koeri.boun.edu.tr/sismo/veri_bank/ mainw.htm. Fig. 1a, b, c show the selected areas and the epicenter distribution of the earthquake data files used in this study.

(a)

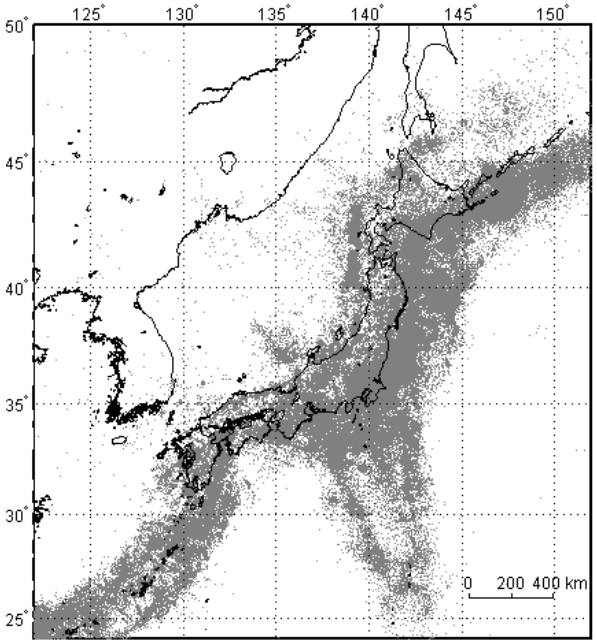
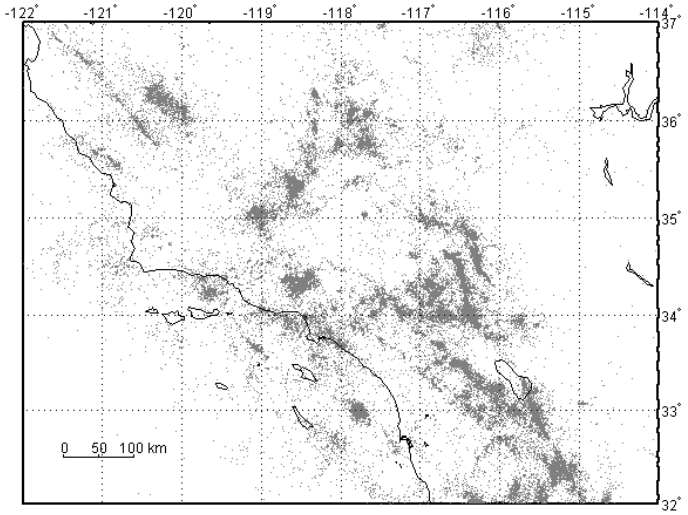
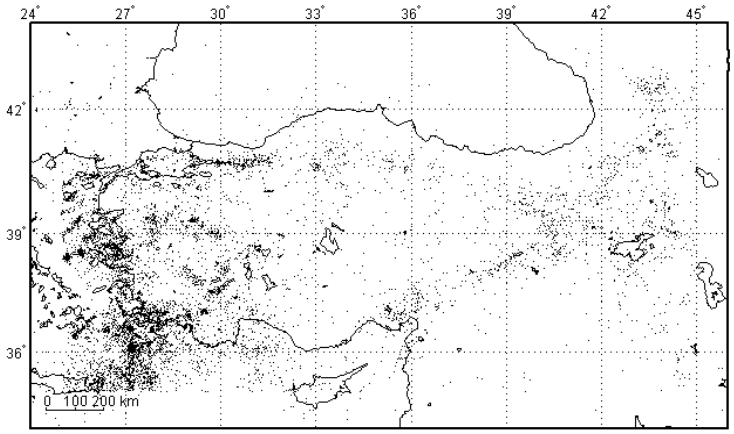


Fig. 1 Epicenter distribution of earthquakes in **(a)** Japan between 679 and 2005 as compiled from the JMA and Utsu catalogs, **(b)** Southern California between 1932 and 2005 and **(c)** Turkey between 550 BC and 2004.

(b)



(c)



Former study of completeness magnitudes by Talbi and Yamazaki (2009) led to four schemes for Japan and three schemes for southern California. The same study was carried out for Turkey and resulted in three schemes. For each region, these schemes cover different time periods and are complete above some threshold magnitudes m_c, as summarized in Table 1 below,

Table 1

Parameters of the sampling schemes used in this study.

m_c, N and R denote the magnitude of completeness, the number of events with magnitude $M \geq m_c$, and the ERS sampling radius, respectively.

Scheme	Time period	m_c	N	R [km]
Southern California				
1	1990-2005	2.5	21257	50
2	1947-2005	3.5	6006	50
3	1932-2005	4.7	556	50
Japan				
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
Turkey				
1	1988-2004	3.5	4472	50
2	1960-2004	4.5	1537	100
3	1960-2004	5.5	127	200

In all the following sections, only Japan data are used extensively. The use of Southern California and Turkey data is restricted to the tests of our mixed model in section 6.

3 Declustering the earthquake catalog

The identification and the remove of aftershocks in earthquake sequences, commonly called ‘declustering’ is not an exact science and depend more or less on some subjective parameters (controlling the time-space widows) predefined in the approach used. Several declustering algorithms were proposed, among which the most standard are Gardner and Knopoff (1974) and Reasenberg (1985). In fact, declustering can only be used for qualitative preliminary studies; quantitative studies are affected by artifacts created by the subjectivity in the methodology. Being aware of this shortcoming and after intensive tests on the catalog used, we adopted Gardner and Knopoff (1974) approach. 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. In this method, space-time windows are considered around each event in the catalog. The size and the duration of each window vary with the magnitude of the potential mainshock. The largest event in each window is identified as mainshock, while the others are removed as either foreshocks or aftershocks. Fig. 2 shows the seismicity rate in Japan during the period 1990-2005 before and after declustering. Main fluctuations are removed by the declustering algorithm, making the activity rate close to stationarity.

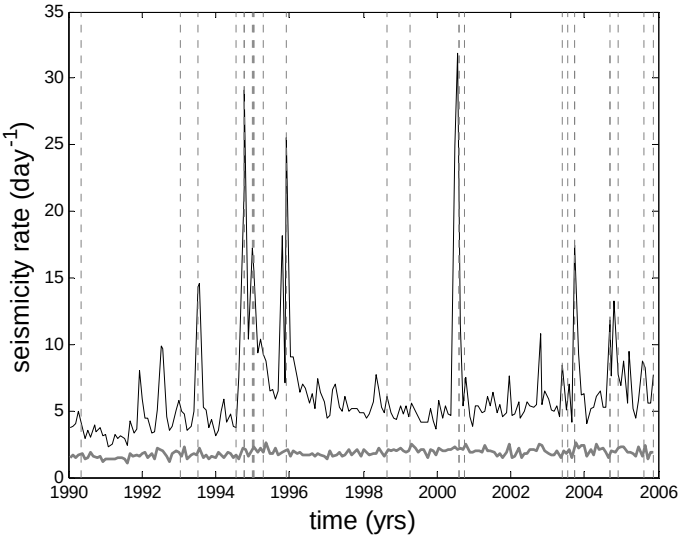


Fig. 2 Seismicity rate during the period 1990-2005 for events with magnitude M≥3.5 in the original JMA catalog (*black line*) and in the Gardner and Knopoff declustered catalog (*gray line*). Large fluctuations of seismicity are removed as shown from the resulting rate after declustering (*gray line*). Vertical dashed lines show the time of each M≥7.0 earthquake.

In Section 5, the results of declustering are used carefully in the qualitative preliminary assessment of the contribution of clustered pairs of events to the observed PL scaling. The other sections are independently carrying out the parameter estimation either analytically or based on empirical results.

4 Earthquake Random Sampling – ERS

The ERS algorithm was introduced by Talbi and Yamazaki (2009). A description of its main steps can be found in the same reference. Interevent time samples are drawn many times from a random collection of disks with a predefined radius R to produce a mean estimate of EITD. The main steps can be resumed as follows.

Step1. A random epicenter $x^{(0)}$ (identified by its longitude and latitude) is selected from the epicenter distribution map.

Step 2. Assuming a given target radius $r>0$, a second epicenter $x^{(1)}$ is selected as the nearest neighbor of $x^{(0)}$ at a distance greater or equal to r from $x^{(0)}$.

Step 3. $x^{(2)}$ is selected as the nearest neighbor of $x^{(1)}$ at a distance greater or equal to r from $x^{(0)}$ and $x^{(1)}$

Steps 4, 5,…, i defines $x^{(3)}$, $x^{(4)}$,…, $x^{(i-2)}$, $x^{(i-1)}$ in a similar way. For example, $x^{(i-1)}$ is defined in the step i as follows:

Step i. $x^{(i-1)}$ is the nearest neighbor of $x^{(i-2)}$ at a distance greater or equal to r from $x^{(0)}$, $x^{(1)}$,…, $x^{(i-2)}$.

If we deal with a finite number of epicenters and at least one of them is situated outside the disk centered on $x^{(0)}$ with radius r, then our algorithm stop in the step m+2 with $m\geq 0$, for which the nearest neighbors of $x^{(m)}$ (say $x^{(m+1)}$) is at a distance smaller than r from one of the above selected epicenters $x^{(0)}$, $x^{(1)}$,…, $x^{(m)}$. This step exclude $x^{(m+1)}$ which is therefore not selected.

The set of points $(x^{(i)})_{i=0,1,\ldots,m}$ constructed earlier are called *r-target points* whereas the family of the above processes, defined for all randomly chosen origin $x^{(0)}$, is called the family of *r-target processes*.

ERS is used iteratively to provide samples required in EITD estimation. In each iteration, a set of *r*-target points $(x^{(i)})_{i=0,1,\ldots,m}$, is selected. Then, for each *sampling disk* $D(x^{(i)}, R)$ centered on $x^{(i)}$ with radius $R = r/2$, the corresponding interevent times $\tau_j^{(i)}$, $j = 1, 2, \ldots, n_i$, are calculated (n_i is the number of interevent times). The resulting m samples $\left(\tau_j^{(i)}\right)_{j=1,\ldots,n_i}^{i=0,1,\ldots,m}$ are mixed together to estimate EITD. After several ERS runs, the mean distribution is derived. In general, *Earthquake Random Sampling* with target radius r and sampling radius R is noted $ERS(r,R)$, whereas the case $R = r/2$, considered in this study, is noted $ERS(R)$.

Fig. 3a displays the ERS(50) strategy map for Japan scheme 2 of Table 1, with the corresponding sampling *R*-disks and target *r*-disks (R=50 km, r=100 km). In this case, the interevent times have been calculated for events occurred within 50 km from targets, and using only sampling disks including more than 50 events. Fig. 3b shows the extent of the sampling regions after 10 ERS runs for Japan scheme 1 of Table 1.

(a)

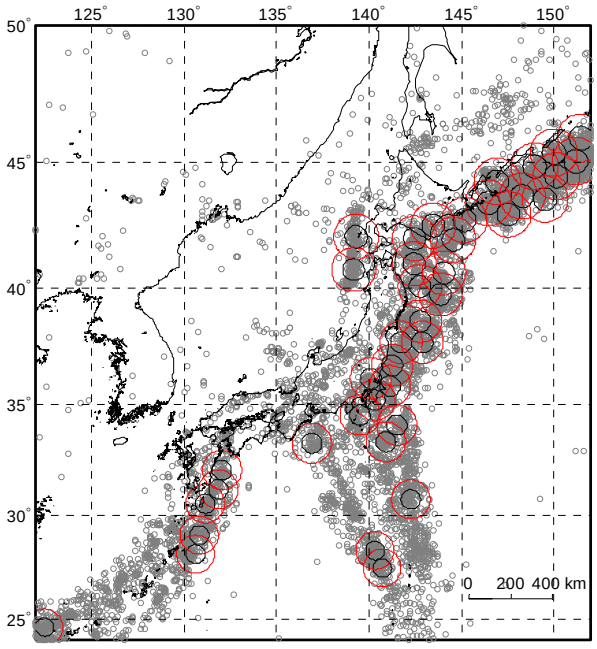
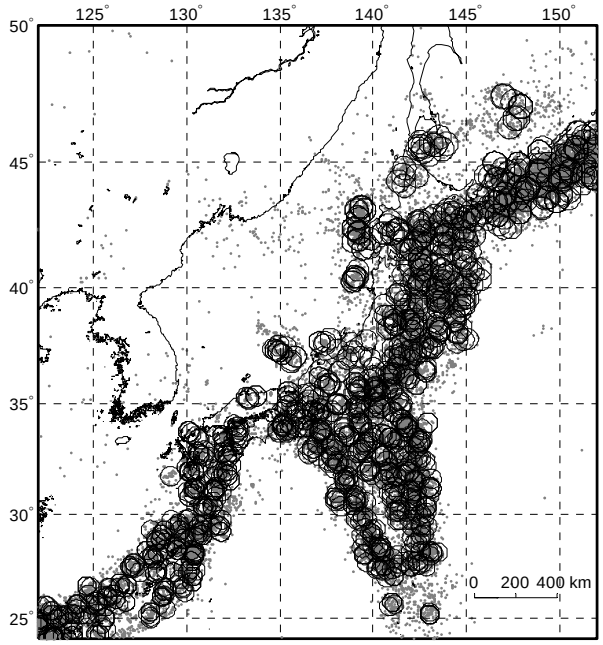


Fig. 3 (a) One run ERS(50) sampling strategy map, showing sampling disks with R=50km and target disks with r=100km, for Japan scheme 2 of Table 1. **(b)** Stacked sampling disks defining the sampling region extent corresponding to ten ERS(50) runs of Japan scheme 1.

(b)



5 Analysis of earthquake interevent time distribution

In this section, we focus on the analysis of the EITD as mixed from different locations using ERS. Following Talbi and Yamazaki (2009), the distribution is scaled by the inverse of the mean interevent time $\bar{\tau}$ of the mixed sample $\left(\tau_j^{(i)}\right)_{j=1,\ldots,n_i}^{i=0,1,\ldots,m}$ as follows,

$$D(\tau)=\frac{1}{\bar{\tau}}\tilde{F}\left(\frac{\tau}{\bar{\tau}}\right), \tag{1}$$

with,

$$\bar{\tau}=\frac{\sum_{i=0}^{m}\sum_{j=1}^{n_i}\tau_j^{(i)}}{\sum_{i=0}^{m}n_i} \tag{2}$$

Note that $\bar{\tau}$ could be replaced by the inverse of the total mean rate of seismicity used by Corral (2003, 2004a) (Talbi and Yamazaki, 2009).

Since the ERS procedure select randomly the first target point $x^{(0)}$, different EITD estimates are obtained when running ERS many times. The resulting averaged EITD is particularly consistent in case of poor data. For example, in Fig. 4 below, we were able to reproduce the shape of the mean EITD using 10^3 ERS(200) runs for M≥6.5 events corresponding to Japan scheme 4.

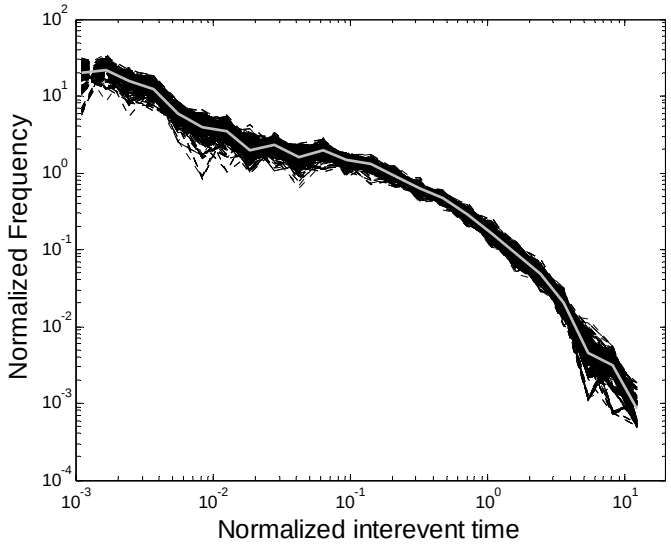
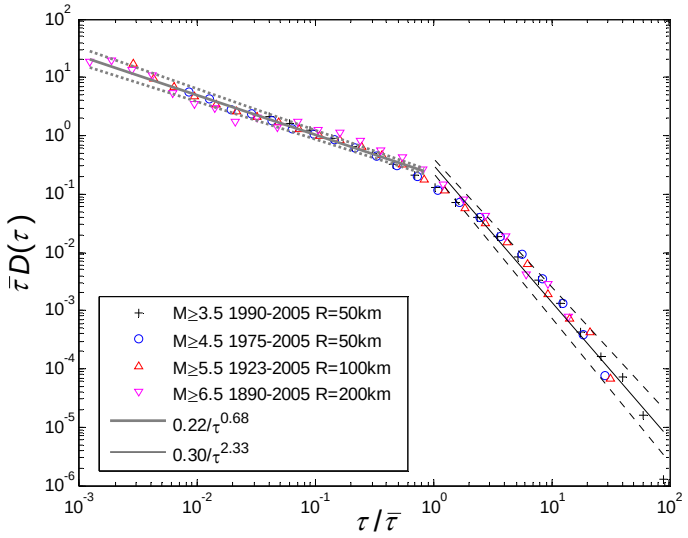


Fig. 4 Mean EITD (Eq. 1) as estimated for Japan scheme 4 using 10^3 ERS(200) runs. *Dashed lines* correspond to different ERS runs whereas the mean distribution is shown in grey.

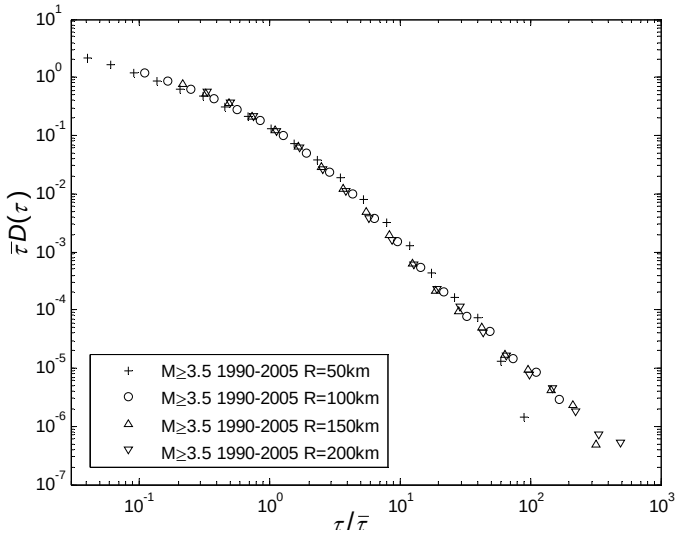
In order to estimate EITD for different magnitude thresholds, we considered completeness magnitudes m_c from the schemes already selected in Table 1. EITD for Japan has been estimated using the schemes 1, 2, 3 and 4 with sampling radiuses R equal to 50, 50, 100 and 200 km, respectively. We performed 10, 50, 100 and 500 ERS runs for the schemes 1, 2, 3 and 4 respectively. Fig. 5a shows the mean EITD for the whole catalog with its corresponding PL scaling. More details about the calculation of the mean distribution and the PL regressions can be found in Talbi and Yamazaki (2009). Note that the choice of the sampling radius R does not affect in any way the shape of the distribution D (e.g., Fig. 5b below and Fig. 6 of Talbi and Yamazaki (2009)). R in Table 1 is selected in order to guarantee a sufficient number of events in each sampling disk.

Fig. 5. **(a)** Stacked mean EITDs for Japan using the schemes in Table 1. *Solid lines* show the power law tendency whereas *dashed lines* show the 95% confidence limits. **(b)** Stack of the EITDs for Japan scheme 1 (Table 1), obtained using ERS(R) with different sampling radiuses R = 50, 100, 150, 200 km. **(c)** EITDs for Japan estimated using Gardner and Knopoff declustered catalog. The s*olid curve* shows the exponential fit with maximum likelihood estimate of the mean λ = 0.73. *Dotted curves* correspond to the exponential distributions obtained using the 95% confidence limits of λ (λ_1=0.49 and λ_2=1.19). There is no trace of the first power law observed for the whole catalog in Fig 5a. The main body of the distributions is close to the corresponding exponential distribution at 95%, apart from the far up tail at standardized interevent times exceeding 10. The deviation corresponding to high magnitude cutoffs (M≥5.5 and M≥6.5) observed for standardized interevent times below 0.01, is attributed to the poor data at this time range. The decrease at the upper tail is still consistent with a power law (*dashed line*).

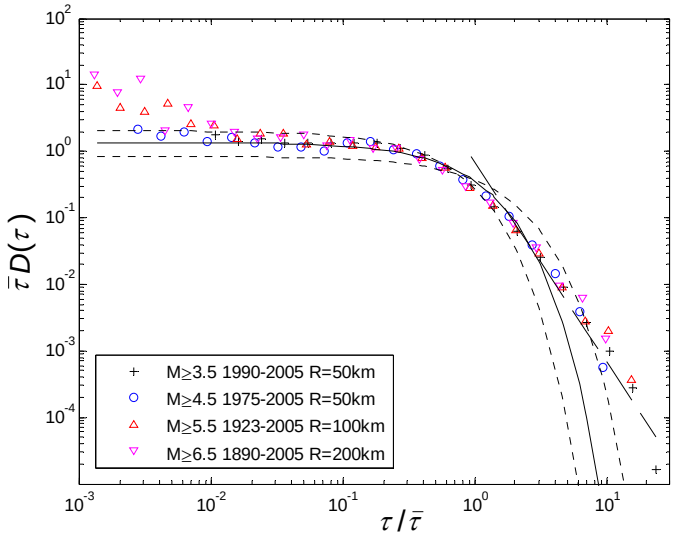
(a)



(b)



(c)



A simple way to test the contribution of spatio-temporal clustering to the observed scaling is to compute the same distributions after removing correlated events in the sense of Gardner and Knopoff (1974) windowing algorithm. Such distributions are plotted in Fig. 5c. As suspected, the first PL is completely disappearing, showing that the filtered clustering structure is responsible of the PL tendency at short time scales. To ensure that this conclusion is not an artifact of the specific space-time windows used, we checked in addition the stability of our results using the windowing suggested by Uhrhammer (1986) and Knopoff (2000). In all cases, the first PL is destroyed. This conclusion generalizes the former results by Saichev and Sornette (2007) showing that in the context of ETAS model, the first PL is aftershock made with exponent close to the Omori *p*-value. At the opposite, the second PL does not disappear totally, but drop close to an exponential decrease especially for scheme 2 (Fig. 5c). The drop is consistent with a steep long range PL. This situation can be produced by the deficiency of the time-space windowing used to remove aftershocks. Indeed the declustering algorithm may remove also some background seismicity identified as clustered. A long range clustering is also possible (e.g., Kagan and Jackson, 1991; Corral 2004b; Lombardi and Marzocchi, 2007; Lennartz et al, 2008). At long range, the distributions reflect the behavior of interevent times mainly for low seismic rates which induces the observed PL for interevent times (Corral 2003; Saichev and Sornette, 2006, 2007).

In order to analyze the contribution of dependent events on the whole interevent time scale, the whole and background distributions D and D^B are estimated using the whole and the declustered catalogs respectively; then, the following deviation d is computed,

$$d(\tau) = \left| D(\tau) - D^B(\tau) \right| \tag{3}$$

Note in this case the following decomposition of the distribution D,

$$D(\tau) = D^B(\tau) \pm d(\tau) \tag{4}$$

The deviation *d* for Japan scheme 1, 2 and 3 (Table 1) is plotted in Fig. 6,

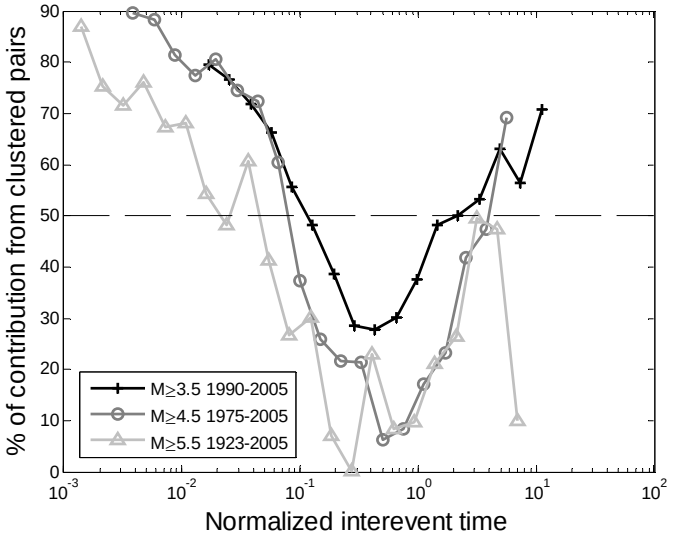


Fig. 6 Percentage of contribution of the clustered pairs of events to the whole distribution, calculated using Gardner and Knopoff (1974) declustering windows. The contribution is dominant for short interevent times and decrease gradually. Uncorrelated pairs are dominant around the change point between the two power laws. However, the contribution from clustered pairs starts to increase again after the change point. Their dominance at large interevent times is supposed to reflect long range correlation or failure of the declustering algorithm at long time range. Note that the power laws in Fig 5a are observed when the clustered pairs are dominant.

As expected, the clustered pairs of events are most active at short time ranges, with a maximal contribution (about 80-90%). This marks the dominance of the correlated process at this time range. The contribution drops gradually (to a level below 27%) around the kink at $\bar{\tau} = 1$, marking the dominance of the uncorrelated pairs of events. But amazingly starts to increase again (to reach eventually 50-70%) for relatively large time ranges. This increase in the contribution from correlated pairs at long ranges is supposed to be in part responsible of the long range PL of *D*. Since some residual catalogs could exhibit long term clustering (Kagan and Jackson, 1991), this behavior can be produced by correlated events that are possibly acting at long ranges producing the observed PL. Another possibility is that some independent events are identified as clustered events by the declustering windows. Note that the declustering algorithm removes all earthquakes within the influence volume of mainshocks, including eventual independent events, especially for low magnitude schemes with events more close in space and time. However, this hypothesis can not be checked since the definition of clustered events in any declustering algorithm is subjective.

The difficulties faced in the former analysis and the subjectivity in the definition of earthquake clusters motivated us to look for a theoretical seismicity model that accounts systematically for both seismicity components with no need to separate between them in practice (even though modeled separately in theory). Next section uses Palm theory to introduce analytically a mixed seismic model for EITD, in which background and clustered seismicity components contribute with proper weights. The parameters are fitted directly to the whole EITD although the model is tested using different background distributions.

6 Mixed models to fit EITD

For a stationary point process, the distribution of the forward recurrence time *F* is a weight biased form of the interevent time distribution *D* (e.g., Cox and Isham 1980, Daley and Vere-Jones 1988),

$$\frac{dF(x)}{dx} = \frac{1}{\mu}(1 - D(x)); \quad x > 0 \tag{5}$$

where μ is the mean interevent time and *d* the differential operator.
In point processes Palm theory, the former equation is commonly known as Palm-Khintchine equation (e.g., Cox and Isham 1980, Daley and Vere-Jones 1988).

Let us note τ the random variable measuring the interevent time between two successive events and τ_1 the time from an arbitrary origin to the first event (forward recurrence time). Then, if we note *P* the probability associated to the distributions *D* and *F*, Eq. 5 can be written equivalently in the following form,

$$-\frac{dP(\tau_1 > x)}{dx} = \frac{1}{\mu}(1 - D(x)) = \frac{1}{\mu}P(\tau > x) \tag{6}$$

If we assume that seismicity can be shared into a background and clustered components, both described by some corresponding independent processes, then a combination of the former equations for both components lead to the following interevent time probability density function (pdf) *f*,

$$f(x) = \frac{\mu}{\mu_B}P(\tau_1^D > x)\, f^B(x) + \frac{\mu}{\mu_D}P(\tau_1^B > x)\, f^D(x) + 2\frac{\mu}{\mu_B \mu_D}(1 - D^B(x))(1 - D^D(x)) \tag{7}$$

where D^B, f^B, μ_B, $\tau_1{}^B$ and D^D, f^D, μ_D, $\tau_1{}^D$ are the cumulative distribution, the density, the mean interevent time and the forward recurrence time for the background and dependent seismicity components respectively.

The derivation of Eq. 7 is developed in the Appendix. This equation, which is the culmination of the present work, defines a basic characteristic decomposition of EITD. It allows us in particular to calculate the pdf *f* explicitly, when models for background and clustered seismicity are substituted.

Let us put,

$$\begin{aligned} w^B(x) &= \frac{\mu}{\mu_B}P(\tau_1^D > x) \\ w^D(x) &= \frac{\mu}{\mu_D}P(\tau_1^B > x) \\ R(x) &= 2\frac{\mu}{\mu_B \mu_D}(1 - D^B(x))(1 - D^D(x)) \end{aligned} \tag{8}$$

The former decomposition reads simply,

$$f(x) = w^B(x) f^B(x) + w^D(x) f^D(x) + R(x) \quad (9)$$

So that the density of mixed interevent times is viewed as the weighted mixture of the background and clustered seismicity densities f^B and f^D plus a correction term R. For very small x values (x<<1), f can be approximated by,

$$f(x) \approx \frac{\mu}{\mu_B} f^B(x) + \frac{\mu}{\mu_D} f^D(x) + 2\frac{\mu}{\mu_B \mu_D} \quad (10)$$

The calculation of the pdf f using the characteristic equation Eq. 7 requires the preliminary knowledge of distributions f^B and f^D. In this study, we assume dependent events to occur as a non homogeneous Poisson process, so that their interevent time distribution obeys the following PL (Utsu et al. 1995; Yakovlev et al. 2005; Shcherbakov et al. 2005, 2006)

$$f^D(x) = \frac{K}{(x+c)^{2-1/p}} \quad (11)$$

where K, c are constants and p the Omori-Utsu p-value. In empirical studies, typical c values are found to range between 0.5 to 20 hours (0.02 to 1day) (Utsu 1961; Reasenberg and Jones 1989, 1994; Utsu et al. 1995). In the following, this parameter is set to the interevent time cutoff τ_c = 0.2 used by Talbi and Yamazaki (2009), whereas p is supposed to vary between 1 and 1.5.

The normalization of f^D together with the calculation of D^D, μ_D and $P(\tau_1^D > x)$ are developed in the Appendix. For both seismicity components, these distributions and parameters are required to carry out the calculation of the mixed pdf f from Eq. 7. For background seismicity, D^B, μ_B and $P(\tau_1^B > x)$ take different forms according to the selected distribution f^B (Appendix).

In this study, the pdf f is examined for different hypothetical background distributions (exponential, Gamma and Weibull). The calculation of f is developed in the Appendix. The following analytic forms of the pdf f are found for Exponential, Gamma and Weibull background distributions respectively.

$$f(x) = \mu\, \varphi(x) e^{-x/\mu_B} \quad (12)$$

$$f(x) = \mu \left[P_1(x)\, x^{k-1} e^{-x/\theta} + P_2(x) \frac{1}{(x+c)^{1+\alpha}} + P_3(x) \frac{1}{(x+c)^{\alpha}} \right] \quad (13)$$

$$f(x) = \mu \left[\frac{1}{\theta \Gamma\left(1+\frac{1}{k}\right)} Q_1(x) e^{-(x/\theta)^k} + \frac{1}{\mu_D} Q_2(x) \frac{1}{(x+c)^{1+\alpha}} \right] \quad (14)$$

In the former equations α is linked to the Omori-Utsu p-value, $\alpha = 1\text{-}1/p$, $0<\alpha<1$, whereas k and θ are the parameters of the Gamma and Weibull distributions in Eq. 13 and Eq. 14 respectively. The functions φ, Q_1, Q_2, P_i , $i = 1, 2, 3$ depend on the parameters of background and dependent events distributions as defined in the Appendix. The Exponential case considered in Eq. 12 is a special case of the Gamma case in Eq. 13 for k=1, so that only the distributions 13 and 14 are fitted to the data. The parameterization 13-A33 defines a mixed distribution noted MixGamma(c,p,k,θ) which is illustrated for Japan in Fig. 7 with the parameters $c = 0.2$, $p = 1.5$, $k = 0.2$ and $\theta = 5.5$.

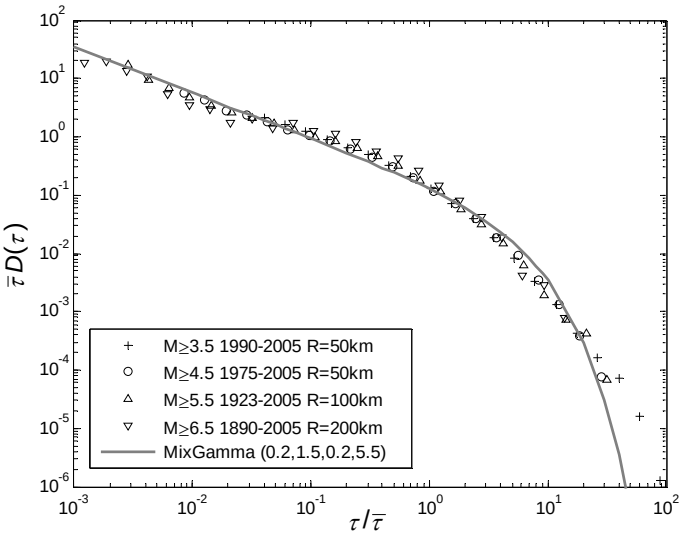


Fig. 7 EITDs fitted using the mixed Gamma model MixGamma (c,p,k,θ) corresponding to the parameterization 13-A33, with $c = 0.2$, $p = 1.5$, $k = 0.2$ and $\theta = 5.5$. Deviation is observed at all scales especially at the up tail where the theoretical fit fall down more rapidly than observed.

It is characterized by a PL like decrease followed by a rapid drop at large interevent times. Although it succeeds to catch the overall shape of the distribution, some deviation is observed at all scales especially at the long time range tail where the theoretical fit fall down more rapidly than observed. The parameterization 14-A40 defines a mixed distribution noted MixWeibull(c,p,k,θ). This former is illustrated for Japan in Fig. 8 for $c = 0.2$, $p = 1.5$, $k = 0.5$ and $\theta = 0.7$ (solid curve). The parameterization 14-A41 is also illustrated for $K = 0.07$, $p = 1$, $k = 0.50$ and $\theta = 0.70$ and an upper interevent time cutoff $\tau_{max} = 1000$ (dotted curve).

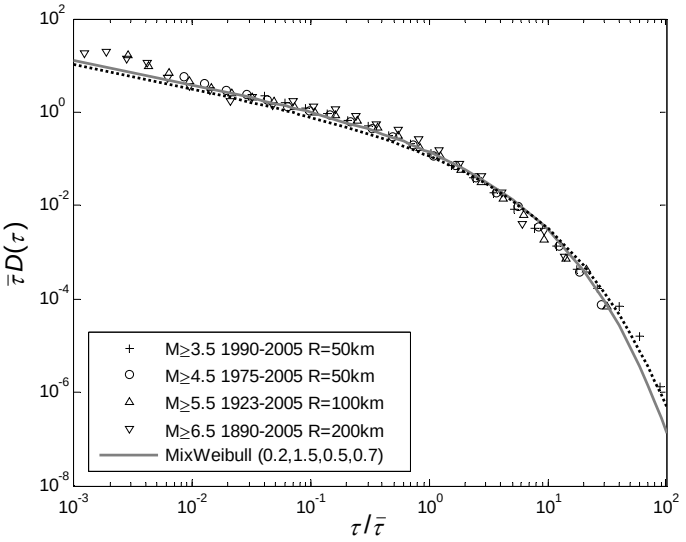


Fig. 8 EITDs for Japan, fitted using the mixed Weibull model MixWeibull(c,p,k,θ) corresponding to the parameterization 14-A40, with c = 0.2, p = 1.5, k = 0.5 and θ = 0.7 (*Solid curve*). *Dotted curve* shows the fit corresponding to the parameterization 14-A41, with K=0.07, p=1, k=0.50 and θ=0.70 and τ_{max} = 1000

The theoretical fit embraces quite well the observed structure of the distribution and succeeds to fit even the observed up tail with an acceptable deviation. To show the superiority of the mixed Weibull fit in Fig 8 comparing to the mixed Gamma fit in Fig 7, we plotted in Fig. 9 the residual mean squares $RMS(\tau > z\bar{\tau})$ for each model. These residuals were calculated for different standardized interevent time cutoffs z according to the following relation,

$$RMS(\tau > z\bar{\tau}) = \sum_{i=1}^{n}\left(f_i - \hat{f}_i\right)^2 \Big/ (n-q) \qquad (15)$$

where f_i, $\hat{f}_i$, n, q are the observed frequency, the predicted or model frequency, the sample size and the number of free parameters in the considered model, respectively. In the former equation, q can be reduced to two for all models compared in Fig. 9. Indeed, c and p are found quite stable and can be assumed constant (c = 0.2 and p = 1.5) in the mixed Weibull and Gamma models, whereas the PL fit is found quite stable at large interevent times (Fig 9, 10 of Talbi and Yamazaki (2009)), so that only the first PL parameters (say c_1 and p_1) are considered. The free parameters k and θ in 13-A33 and 14-A40 are found relatively close to the maximum likelihood estimates corresponding to the background distribution calculated using the declustered catalog. They can be assessed easily by minimizing the residuals in Eq 15. The PL free parameters c_1 and p_1 have been already estimated for Japan and southern California in Talbi and Yamazaki (2009).

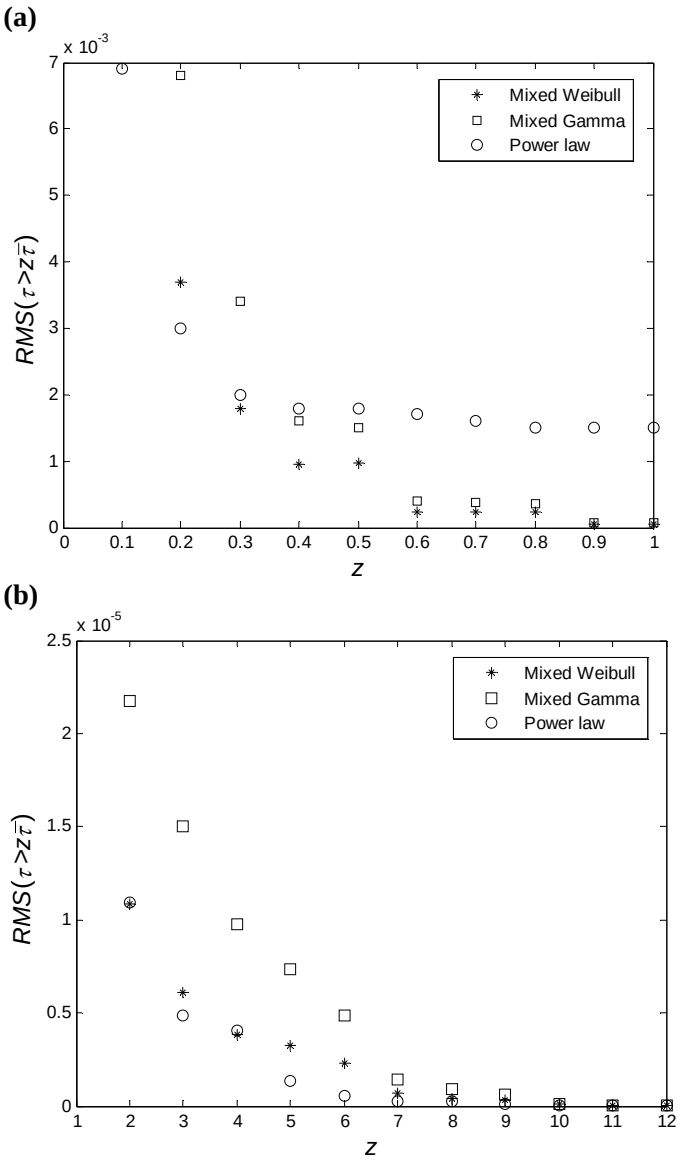


Fig. 9 Residual mean squares for Japan, calculated from Eq. 15 for normalized recurrence times $\tau/\bar{\tau}$ greater than *z*. Because of the large difference in scales, the residuals for $z \leq 1$ and $z > 1$ are separately plotted in **(a)** and **(b),** respectively. The mixed Gamma and mixed Weibull fits in Fig. 7 and Fig. 8 (*Solid curve*), and the doubly power law fit in Fig. 5a, are compared. We can opt for the mixed Weibull model starting from about $z = 0.3$.

Fig. 9 shows that the mixed Weibull model provides the best fit for standardized recurrence times *z* exceeding 0.3, comparing with the former mixed Gamma model in Fig. 7 and the two PLs fit in Fig. 5a. The mixed Weibull model is especially superior at both short and long range tails of the distribution.

Fig. 10a show the EITDs fit for southern California using our both mixed models, corresponding to the parameterization 14-A40, with $c = 0.2$, $p = 1.5$, $k = 0.4$ and $\theta = 0.7$ (*Solid curve*), and the parameterization 13-A33, with $c = 0.2$, $p = 1.5$, $k = 0.2$ and $\theta = 5.5$ (*Dotted curve*). The mixed Gamma model clearly deviates at long time ranges. Fig. 10b shows the residual mean squares corresponding to Fig. 10a, and those from the PL fit of Fig. 7a in Talbi and Yamazaki (2009).

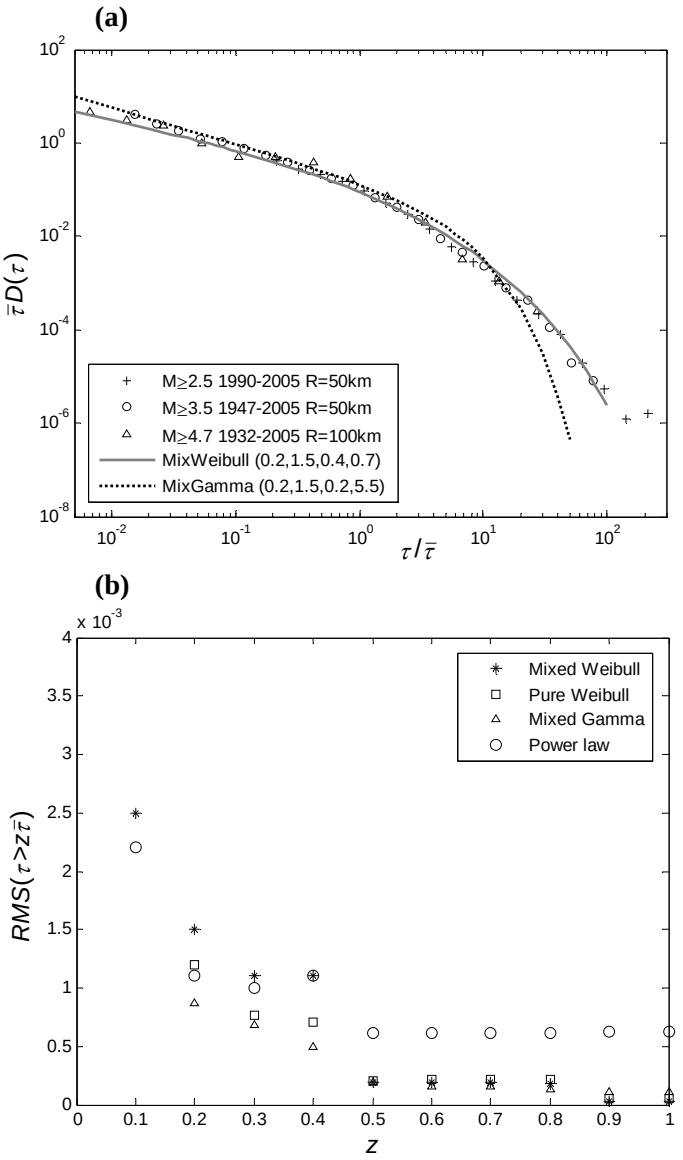


Fig. 10 (a) EITDs for southern California, fitted using the mixed Weibull model MixWeibull(c,p,k,θ) corresponding to the parameterization 14-A40 with c = 0.2, p = 1.5, k = 0.4 and θ = 0.7 (*Solid curve*), and the mixed Gamma model MixGamma(c,p,k,θ) corresponding to the parameterization 13-A33, with c = 0.2, p = 1.5, k = 0.2 and θ = 5.5 (*Dotted curve*) **(b)** Residual mean squares for southern California calculated from Eq. 15 for normalized recurrence times $\tau/\bar{\tau}$ greater than z, corresponding to the mixed Gamma and the mixed Weibull fits in the former Fig. 10a, the doubly power law and the pure Weibull fits in Fig. 7a and Fig. 12a of Talbi and Yamazaki (2009) respectively. The mixed Weibull model performs well starting from about z = 0.5.

In this case, we can opt for our mixed Weibull model for $z > 0.5$, even if the mixed Gamma model is the best fit for z values between 0.2 and 0.4. Note that for z values exceeding 1, the mixed Weibull is preferred. The residuals for $z > 1$ are not shown here because the comparison at this range is clear from Fig 10a. Finally, in case of Turkey, the mixed Weibull model is the best fit on the whole range of z values, namely exceeding 0.1 (Fig. 11a, b). As in the case of Japan, it is especially superior at both short and long range tails of the distribution (Fig. 11a, b). The power laws fit in this case is shown in Fig. 11c.

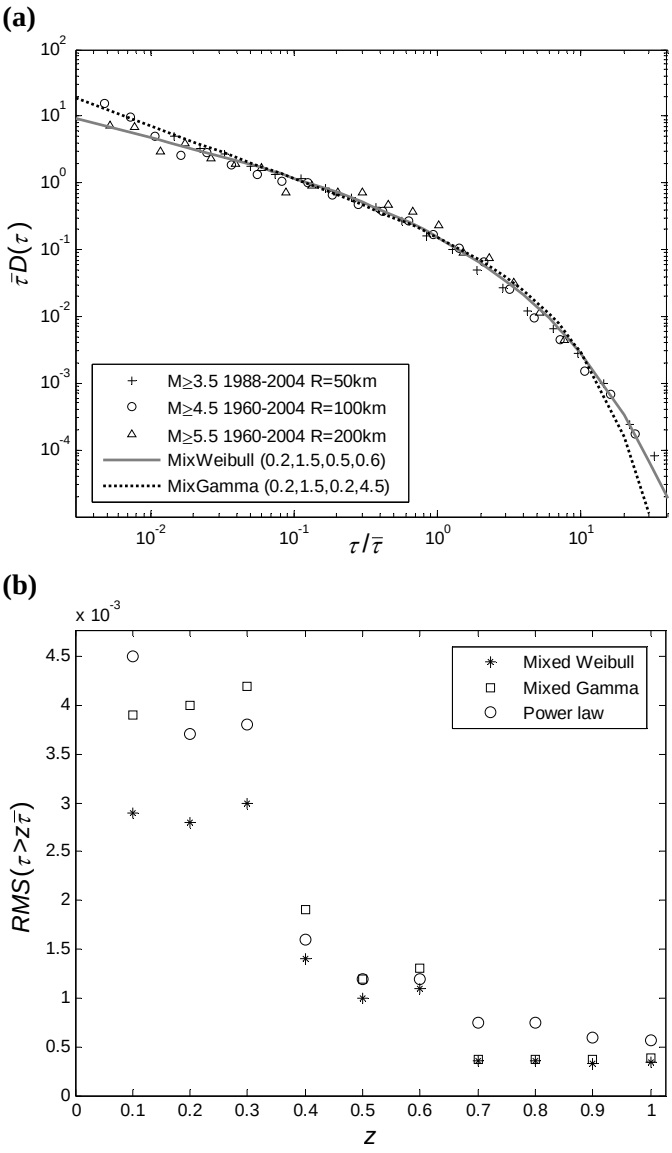
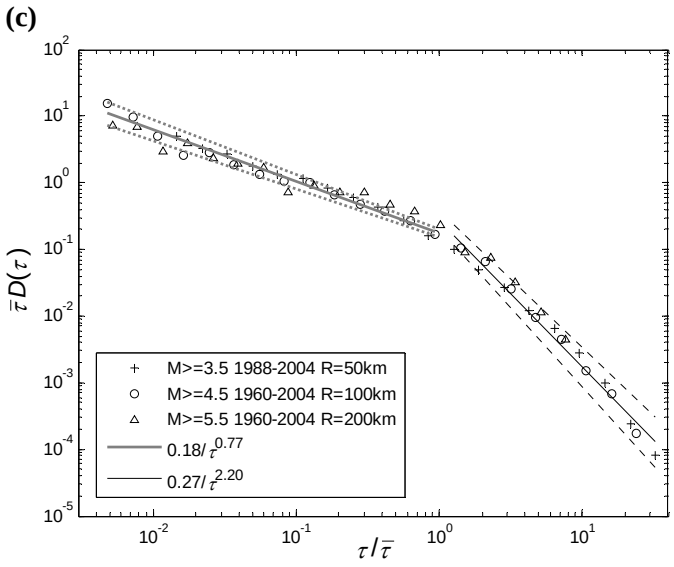


Fig. 11 (a) EITDs for Turkey, fitted using the mixed Weibull model MixWeibull(c,p,k,θ) corresponding to the parameterization 14-A40 with c = 0.2, p = 1.5, k = 0.5 and θ = 0.6 (*Solid curve*), and the mixed Gamma model MixGamma(c,p,k,θ) corresponding to the parameterization 13-A33, with c = 0.2, p = 1.5, k = 0.2 and θ = 4.5 (*Dotted curve*). **(b)** Residual mean squares for Turkey calculated from Eq. 15 for normalized recurrence times $\tau/\bar{\tau}$ greater than z. The mixed Gamma and the mixed Weibull fits in Fig. 11a, and the doubly power law fit in Fig. 11c, are compared. The mixed Weibull model provides the best fit. **(c)** The doubly power law fit for EITDs in Turkey.



In conclusion, the former results elect generally the mixed Weibull model as the best fit for EITD in the studied regions. Furthermore, the selected model occurs with quite stable parameterization, namely with c = 0.2, p = 1.5, k and θ estimated roughly in the ranges 0.4-0.5 and 0.6-0.7, respectively.

7 Conclusion

We have proposed a general model to describe the statistics of interevent times, averaged over multiple regions. This work was motivated by our former study proposing a Weibull distribution to fit southern California data, alternatively to the power laws (Talbi and Yamazaki 2009). Our model succeeded to describe quite well seismicity in Japan, southern California and Turkey, as a mixture of clustered seismicity modeled by a non-homogeneous Poisson process (inducing a power law interevent time distribution), and background seismicity with interevent times at best described by a stretched exponential (Weibull) distribution. This view shows that the classical description of earthquakes as mixture of background and dependent events can still account for and simply explain recent empirical observations of EITD. This conclusion is reached by a combination of analysis, which starts from the derivation of the so-called Palm equations for both seismicity components, to go to the combination of the derived equations in a single model, and end with the derivation of the analytic expression of the EITD for different hypothetical background distributions.

In the preliminary analysis, a usual filtering of the JMA catalog using standard space-time windows showed that both power laws, observed at short and long time ranges, can be attributed to the observed correlation (or clustering) structure within the filtering windows. In particular, the scaling and the power laws of EITDs reflect the observed correlation between earthquakes in space and time, as captured by the standard declustering windowing algorithms.

These results suggest a careful interpretation of the apparent scaling of EITD. In particular, a power law at short time ranges and a stretched exponential tail at long time ranges can constitute a better alternative fit breaking the power law at long time range. Our model deserves further development and tests, in particular using an inverse Gaussian background distribution.

Appendix

A1. Characteristic decomposition of EITD

Let us assume the seismicity shared into its background and clustered components both described by corresponding independent processes. In this case, the probability that the time to the next event from a fixed origin 0 exceeds a given time $x > 0$, say $P(\tau_1 > x)$, is the product of the probabilities for the two components,

$$P(\tau_1 > x) = P(\tau_1^B > x)P(\tau_1^D > x) \qquad \text{(A1)}$$

where τ_1^B and τ_1^D are the forward recurrence times corresponding to background and clustered seismicity respectively. Let us apply the derivation operator to both sides of the former equation,

$$\frac{dP(\tau_1 > x)}{dx} = P(\tau_1^B > x)\frac{dP(\tau_1^D > x)}{dx} + P(\tau_1^D > x)\frac{dP(\tau_1^B > x)}{dx} \qquad \text{(A2)}$$

The derivatives in the former equation can be obtained from Eq. 6. Indeed, this former holds also for the background and clustered components as follows,

$$\frac{dP(\tau_1^B > x)}{dx} = -\frac{1 - D^B(x)}{\mu_B} \qquad \text{(A3)}$$

$$\frac{dP(\tau_1^D > x)}{dx} = -\frac{1 - D^D(x)}{\mu_D} \qquad \text{(A3')}$$

The substitution of the derivatives defined in Eqs. 6, A3 and A3' into Eq. A2 yields,

$$\frac{1 - D(x)}{\mu} = P(\tau_1^B > x)\frac{1 - D^D(x)}{\mu_D} + P(\tau_1^D > x)\frac{1 - D^B(x)}{\mu_B} \qquad \text{(A4)}$$

where D^B, μ_B, D^D, and μ_D are the distribution and the mean interevent time of the background and dependent series of events, respectively. Finally, the following analytic expression of the distribution D is obtained,

$$D(x) = 1 - \left[\frac{\mu}{\mu_D}P(\tau_1^B > x)(1 - D^D(x)) + \frac{\mu}{\mu_B}P(\tau_1^D > x)(1 - D^B(x))\right] \qquad \text{(A5)}$$

The pdf f below follows by derivation,

$$f(x) = -\frac{\mu}{\mu_D}\left[\frac{dP(\tau_1^B > x)}{dx}(1 - D^D(x)) - P(\tau_1^B > x) f^D(x)\right] - \cdots$$
$$\frac{\mu}{\mu_B}\left[\frac{dP(\tau_1^D > x)}{dx}(1 - D^B(x)) - P(\tau_1^D > x) f^B(x)\right] \quad \text{(A6)}$$

In the former equation, f^B and f^D are the pdfs of the background and clustered seismicity components respectively.

Replacing the derivatives A3, A3' into Eq. A6, we finally get the characteristic decomposition of f as in Eq. 7,

A2 Distribution of dependent events

To normalize f^D as a pdf, we distinguish two cases, p>1 and p=1.

Case p>1

In this case f^D is normalized to,

$$f^D(x) = \frac{\alpha c^\alpha}{(x+c)^{1+\alpha}} \quad \text{(A7)}$$

with α =1-1/p, 0<α <1

The corresponding cumulative distribution is,

$$D^D(x) = 1 - \frac{c^\alpha}{(x+c)^\alpha} \quad \text{(A8)}$$

The substitution of D^D into Eq. A3' and the integration of both sides of the equation between 0 and x yields,

$$P(\tau_1^D > x) = 1 - \frac{c^\alpha}{(1-\alpha)\mu_D}\left((x+c)^{1-\alpha} - c^{1-\alpha}\right) \quad \text{(A9)}$$

Let us apply an upper cut off τ_{max}>0 to account for temporal dependency and avoid an infinite mean. In this case, the mean μ_D of the former distribution is by definition,

$$\mu_D = \int_0^{\tau_{\max}} x\, f^D(x)\, dx \quad \text{(A10)}$$

Replacing f^D from Eq. A7 we obtain,

$$\mu_D = \alpha c^\alpha \int_0^{\tau_{\max}} \frac{x}{(x+c)^{1+\alpha}}\, dx \quad \text{(A11)}$$

The former integration can be performed using the following decomposition,

$$\mu_D = \alpha c^\alpha \left(\int_0^{\tau_{\max}} \frac{1}{(x+c)^\alpha}\, dx - \int_0^{\tau_{\max}} \frac{c}{(x+c)^{1+\alpha}}\, dx\right) \quad \text{(A12)}$$

Finally, after a simple calculation we get the following mean value μ_D,

$$\mu_D = \frac{\alpha c^{\alpha}}{1-\alpha}(\tau_{\max}+c)^{1-\alpha} + \frac{c^{1+\alpha}}{(\tau_{\max}+c)^{\alpha}} - \frac{c^{1-\alpha}}{\alpha(1-\alpha)} \quad \text{(A13)}$$

Case *p*=1

This corresponds to the case α=0 in which f^D, D^D and $P(\tau_1^D>x)$ take the following forms,

$$f^D(x) = \frac{K}{x+c} \quad \text{(A14)}$$

$$D^D(x) = K\log\left(\frac{x+c}{c}\right) \quad \text{(A15)}$$

$$P(\tau_1^D > x) = 1 - \frac{K}{\mu_D} c\log c + \frac{K}{\mu_D}(x+c)\log(x+c) - \left(\frac{K + K\log c + 1}{\mu_D}\right)x \quad \text{(A16)}$$

Similarly to the former case, the mean μ_D is calculated after applying an upper cutoff τ_{max}. In this case, the normalization condition for D^D implies,

$$D^D(\tau_{\max}) = K\log\left(\frac{\tau_{\max}+c}{c}\right) = 1 \quad \text{(A17)}$$

The former equation links τ_{max} to the distribution parameters c and K as,

$$\tau_{\max} = c\left(e^{1/K} - 1\right) \quad \text{(A18)}$$

The mean μ_D is obtained from the following integration,

$$\mu_D = \int_0^{\tau_{\max}} \frac{x}{x+c}dx = \int_0^{\tau_{\max}} dx - \int_0^{\tau_{\max}} \frac{c}{x+c}dx \quad \text{(A19)}$$

After a simple calculation, μ_D takes the following form,

$$\mu_D = K\left(\tau_{\max} - c\,(1/K + \log c) + c\log c\right) \quad \text{(A20)}$$

Finally, substituting τ_{max} from Eq. A18, we get the following mean μ_D,

$$\mu_D = Kc(e^{1/K} - 1 - 1/K) \quad \text{(A21)}$$

A3 Background events distribution

In the following, the pdf f is examined for three hypothetical background distributions: Exponential, Gamma and Weibull.

A3.1 Exponential background distribution

This case corresponds to Poisson background occurrences with a mean μ_B and the following density and cumulative distributions,

$$f^B(x) = \frac{1}{\mu_B}e^{-x/\mu_B} \quad \text{(A22)}$$

$$D^B(x) = 1 - e^{-x/\mu_B} \tag{A23}$$

The probability $P(\tau_1^B > x)$ can be calculated by integration from Eq. A3,

$$P(\tau_1^B > x) = e^{-x/\mu_B} \tag{A24}$$

Case *p*>1

In this case, *f* takes the analytic form in Eq. 12 after the substitution of Eqs. A7, A8, A9, A22, A23 and A24, into the characteristic decomposition 7. The function φ is defined as,

$$\varphi(x) = a + b(x+c)^{1-\alpha} + \frac{d}{(x+c)^{1+\alpha}} + \frac{h}{(x+c)^{\alpha}} \tag{A25}$$

with,

$$\begin{aligned} a &= \frac{1}{\mu_B^{\,2}}\left(1 + \frac{c}{(1-\alpha)\mu_D}\right) \\ b &= -\left(\frac{c^{\alpha}}{(1-\alpha)}\right)\frac{1}{\mu_B^{\,2}\mu_D} \\ d &= \frac{\alpha c^{\alpha}}{\mu_D} \\ h &= \frac{2c^{\alpha}}{\mu_B \mu_D} \end{aligned} \tag{A26}$$

As suspected, for small *x* values ($x \ll 1$), *f* decreases as a PL with exponent $1+\alpha$, whereas it decreases rapidly for large *x* values ($x \gg 1$). Note that to allow a simple form of the pdf *f*, μ_D is not replaced here by its value from Eq. A13. However, in application, this former was used with a given upper cut off τ_{max}.

Case *p*=1

After the substitution of Eqs. A14, A15, A16, A22, A23 and A24 into the characteristic decomposition 7, *f* is obtained as in Eq. 12 with a function φ,

$$\varphi(x) = c' + a'\,x + b'(x+c)\log(x+c) + \frac{d'}{x+c} + h'\log(x+c) \tag{A27}$$

Here,

$$
\begin{aligned}
c' &= \frac{1}{\mu_B{}^2} - \frac{Kc}{\mu_B{}^2 \mu_D}\log c + \frac{2}{\mu_B \mu_D} + \frac{2K}{\mu_B \mu_D}\log c \\
a' &= -\frac{K + K\log c + 1}{\mu_B{}^2 \mu_D} \\
b' &= \frac{K}{\mu_B{}^2 \mu_D} \\
d' &= \frac{K}{\mu_D} \\
h' &= -\frac{2K}{\mu_B \mu_D}
\end{aligned}
\qquad \text{(A28)}
$$

Note that for simplicity μ_D obtained from Eq. A21 is not replaced here although it is used in application.

A3.2 Gamma background distribution

This choice is motivated by the view of the distribution as the sum of independent exponential components within each local sampling disk. The induced Gamma distribution has the following form,

$$f^B(x) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-x/\theta}; \quad x \geq 0,\ k, \theta > 0 \qquad \text{(A29)}$$

The cumulative distribution D^B follows by integration,

$$D^B(x) = \frac{\gamma(k, x/\theta)}{\Gamma(k)} \qquad \text{(A30)}$$

where $\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt$ and $\gamma(k, x/\theta) = \int_0^{x/\theta} t^{k-1} e^{-t} dt$ are the Gamma and the incomplete Gamma functions, respectively. Note that the former Exponential distribution Eq. A22 is embedded in the Gamma distribution Eq. A29 for $k = 1$.

Using Eq. A3 and replacing the Gamma mean $\mu_B = k\theta$ we obtain the following probability for the forward recurrence time,

$$P(\tau_1^B > x) = 1 - \frac{x}{k\theta} + \frac{1}{k\theta\Gamma(k)} I_{k,\theta}(x) \qquad \text{(A31)}$$

with,

$$I_{k,\theta}(x) = \int_0^x \gamma(k, u/\theta)\, du = \int_0^x \int_0^{u/\theta} t^{k-1} e^{-t} dt\, du \qquad \text{(A32)}$$

Case *p*>1

After the substitution of Eqs. A7, A8, A9, A29, A30 and A31 into Eq. 7, the pdf *f* takes the analytic form in Eq. 13. In this case, the functions P_1, P_2 and P_3 are defined as follows,

$$P_1(x) = \frac{1}{k\theta^{k+1}\Gamma(k)}\left(1 + \frac{c}{(1-\alpha)\mu_D} - \frac{c^\alpha}{(1-\alpha)\mu_D}(x+c)^{1-\alpha}\right)$$
$$P_2(x) = \frac{\alpha c^\alpha}{\mu_D}\left(1 + \frac{c}{k\theta} + \frac{I_{k,\theta}(x)}{k\theta\Gamma(k)}\right) \qquad \text{(A33)}$$
$$P_3(x) = 2\frac{c^\alpha}{k\theta\mu_D}\left(1 - \frac{\alpha}{2} - \frac{\gamma(k, x/\theta)}{\Gamma(k)}\right)$$

Case *p*=1

Similarly to the former case, *f* takes the analytic form in Eq. 13 after replacing Eqs. A14, A15, A16, A29, A30 and A31 in the characteristic decomposition 7. In this case, the functions P_1, P_2 and P_3 become,

$$P_1(x) = \frac{1}{k\theta^{k+1}\Gamma(k)}\left(1 - \frac{K}{\mu_D}c\log c + \frac{K}{\mu_D}(x+c)\log(x+c) - \frac{K + K\log c + 1}{\mu_D}x\right)$$
$$P_2(x) = \frac{K}{\mu_D}\left(1 - \frac{x}{k\theta} + \frac{I_{k,\theta}(x)}{k\theta\Gamma(k)}\right)$$
$$P_3(x) = \frac{2}{k\theta\mu_D}\left(1 + K\log c - \frac{\gamma(k, x/\theta)}{\Gamma(k)} + \left(\frac{\gamma(k, x/\theta)}{\Gamma(k)} - K\right)\log(x+c)\right)$$

(A34)

A3.3 Weibull background distribution

Recently, the Weibull distribution (Weibull 1951) has been extensively used 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). In addition, we showed that it presents the best fit for intermediate and long time ranges of southern California EITD (Talbi and Yamazaki 2009). These considerations motivated us to use it here as a model for background seismicity, especially to catch the proprieties of EITD at long time ranges.

The Weibull distribution with parameters *k*, θ takes the following form,

$$f^B(x) = k\theta^{-k}x^{k-1}e^{-(x/\theta)^k}; \quad x \geq 0,\ k, \theta > 0 \qquad \text{(A35)}$$

Its cumulative distribution writes,

$$D^B(x) = 1 - e^{-(x/\theta)^k} \qquad \text{(A36)}$$

The forward interevent time probability $P(\tau_1^B > x)$ can be calculated from Eq. A3 by replacing the following mean of the Weibull distribution,

$$\mu_B = \theta\Gamma\left(1+\frac{1}{k}\right) \tag{A37}$$

It follows that,

$$P(\tau_1^B > x) = 1 - \frac{1}{\theta\Gamma\left(1+\frac{1}{k}\right)} J_{k,\theta}(x) \tag{A38}$$

Where $J_{k,\theta}(x)$ is the following integral function,

$$J_{k,\theta}(x) = \int_0^x e^{-(t/\theta)^k} dt \tag{A39}$$

Case *p*>1

f takes the analytic form in Eq. 14 after replacing Eqs. A7, A8, A9, A35, A36 and A38 in the characteristic decomposition 7. The functions Q_1 and Q_1 are defined as follows,

$$\begin{aligned} Q_1(x) &= \left(1 - \frac{c^\alpha}{(1-\alpha)\mu_D}(x+c)^{1-\alpha} + \frac{c}{(1-\alpha)\mu_D}\right) k\theta^{-k} x^{k-1} + \frac{2c^\alpha}{\mu_D}\frac{1}{(x+c)^\alpha} \\ Q_2(x) &= \alpha c^\alpha \left(1 - \frac{1}{\theta\Gamma\left(1+\frac{1}{k}\right)} J_{k,\theta}(x)\right) \end{aligned} \tag{A40}$$

Case *p*=1

In the same way, *f* takes the analytic form in Eq. 14 after replacing Eqs. A14, A15, A16, A35, A36 and A38 in the characteristic decomposition 7. The functions Q_1 and Q_2 are defined as follows,

$$Q_1(x) = \left(1 - \frac{K}{\mu_D} c \log c + \frac{K}{\mu_D}(x+c)\log(x+c) - \frac{K + K\log c + 1}{\mu_D} x\right) k\theta^{-k} x^{k-1} + \cdots$$
$$\frac{2}{\mu_D}\left(1 - K\log\left(\frac{x+c}{c}\right)\right)$$
$$Q_2(x) = K\left(1 - \frac{1}{\theta\Gamma\left(1+\frac{1}{k}\right)} J_{k,\theta}(x)\right)$$
(A41)

Acknowledgements

This work has been supported, in part, by the scholarship from Ministry of Education, Culture, Sport, Science and Technology (Monbukagakusho), the government of Japan, and by the Algerian Centre de Recherche en Astronomie Astrophysique and Geophysique (CRAAG). The authors are grateful to the Japanese Meteorological Agency (JMA) for providing catalog data. We thank M. Hamdache from CRAAG for carefully reading the original version of the manuscript and his constructive criticisms. David Vere-Jones is acknowledged for his thoughtful comments and ideas that inspired much of the mathematical developments in this paper. The authors thank the editor Torsten Dahm and two anonymous reviewers for their comments that improved significantly an earlier version of the manuscript.

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

Båth M (1965) Lateral inhomogeneities in the upper mantle, Tectonophysics 2: 483–514

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 (2005) Mixing of rescaled data and Bayesian inference for earthquake recurrence times. Nonlin Processes Geophys 12**:**89–100

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

Gutenberg B, Richter CF (1942) Earthquake magnitude, intensity, energy and acceleration. Bull Seismol Soc Am 32**:**162–191

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

Kagan YY, Jackson DD (1991) Long-term earthquake clustering. Geophys J Int 104(1):117–134 doi:10.1111/j.1365-246X.1991.tb02498.x

Kagan YY, Knopoff L (1980) Spatial distribution of earthquakes: the two-point correlation function. Geophys J R Astr Soc 62**:**303–320

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

Lennartz S, Livina VN, Bunde A, Havlin S (2008) Long-term memory in earthquakes and the distribution of interoccurrence times. Europhysics Letters 81 doi:10.1209/0295-5075/81/69001

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

Lombardi AM, Marzocchi W (2007) Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes. J Geophys Res 112**(**B02303) doi:10.1029/2006JB004568

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

Omori F (1894) On the after-shocks of earthquakes. J Coll Sci Tokyo Imp Univ 7: 111–200

Reasenberg P (1985) Second-order moment of central California seismicity, 1969–1982. J Geophys Res 90:5479–5495

Reasenberg PA, Jones MJ (1989) Earthquake hazard after a mainshock in California. Science 243(4895) doi:10.1126/science.243.4895.1173

Reasenberg PA, Jones MJ (1994) Earthquake aftershocks: update. Science 265(5176) doi:10.1126/science.265.5176.1251

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

Shcherbakov R, Turcotte DL, Rundle JB (2006) Scaling proprieties of the Parkfield aftershock sequence. Bull Seismol Soc Am 96(4B):S376–S384 doi:10.1785/0120050815

Shcherbakov R, Yakovlev G, Turcotte DL, Rundle JB (2005) A model for the distribution of aftershock waiting times. Phys Rev Lett 95 doi:10.1103/PhysRevLett.95.218501

Sornette D, Pisarenko V (2003) Fractal plate tectonics. Geophys Res Lett 30**:**1105 doi:10.1029/2002GL015043

Talbi A, Yamazaki F (2009) Sensitivity analysis of the parameters of earthquake recurrence time power law scaling. J Seismol 13:53–72 doi:10.1007/s10950-008-9115-1

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

Uhrhammer R (1986) Characteristics of northern and southern California seismicity. Earthquake Notes 57**(**21)

Utsu T (1961) A statistical study on occurrence of aftershocks. Geophys Mag 30**:**521–605

Utsu T (2002) Statistical features of seismicity, International handbook of earthquake and engineering seismology. Part B:719–732

Utsu T, Ogata Y, Matsu'ura RS (1995) The centenary of the Omori formula for a decay law of aftershocks activity. J Phys Earth 43**:**1−33

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

Yakovlev G, Rundle JB, Shcherbakov R, Turcotte DL (2005) Inter-arrival time distribution for the non-homogeneous Poisson process. arXiv:cond-mat/0507657v1

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