



HAL
open science

Modeling earthquake dynamics

Arthur Charpentier, Marilou Durand

► **To cite this version:**

Arthur Charpentier, Marilou Durand. Modeling earthquake dynamics. [Research Report] uqam. 2015. hal-00871883v2

HAL Id: hal-00871883

<https://hal.science/hal-00871883v2>

Submitted on 25 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MODELING EARTHQUAKE DYNAMICS

ARTHUR CHARPENTIER AND MARILOU DURAND

ABSTRACT. In this paper, we investigate questions arising in *Parsons and Geist* (2012). Pseudo causal models connecting magnitudes and waiting times are considered, through generalized regression. We do use conditional model (magnitude given previous waiting time, and conversely) as an extension to joint distribution model described in *Nikoloulopoulos and Karlis* (2008). On the one hand, we fit a Pareto distribution for earthquake magnitudes, where the tail index is a function of waiting time following previous earthquake; on the other hand, waiting times are modeled using a Gamma or a Weibull distribution, where parameters are function of the magnitude of the previous earthquake. We use those two models, alternatively, to generate the dynamics of earthquake occurrence, and to estimate the probability of occurrence of several earthquakes within a year, or a decade.

Arthur Charpentier, UQAM, 201, avenue du Président-Kennedy, Montréal (Québec), Canada H2X 3Y7 (corresponding author) charpentier.arthur@uqam.ca, and Marilou Durand, UQAM, 201, avenue du Président-Kennedy, Montréal (Québec), Canada H2X 3Y7.

1. INTRODUCTION AND MOTIVATION

1.1. Models for magnitude and occurrence process. The seismic gap hypothesis implies that earthquake hazard should be small immediately following the previous large earthquake, and should increase with time, since the last large event on certain fault or plate boundaries, see *Benioff* (1951), *Aki* (1956) or *McCann et al.* (1979) for early work, and more recently *Thatcher* (1989), *Kagan and Jackson* (1991a), *Nishenko and Sykes* (1993) or *Kagan and Jackson* (1995). The idea behind the seismic gap hypothesis can be found in much earlier work. For instance, *Gilbert* (2009) considered a ‘*rhythmic recurrence hypothesis*’ while *Reid* (1906) suggested that a large earthquake releases most of the stress in a given fault segment, and that further earthquakes there would unlikely occur until the stress is restored. For instance *Bakun and Lindh* (1985), *Sieh et al.* (1989), *Stein* (1995), *Sieh* (1996), *Murray and Segall* (2002), *Stein* (2002), *Kerr* (2004), *Weldon et al.* (2004), *Zhuang et al.* (2002), *Zhuang et al.* (2004) *Lombardi and Marzocchi* (2007), *Wang et al.* (2010), *Ogata and Katsura* (1993) and *Weldon et al.* (2004) mention temporal clustering of major earthquakes, with regular cycles, and characteristic earthquakes.

The idea that earthquake processes can be a *self-organized critical* phenomenon has been mentioned by several authors (see *Bak and Tang* (1989) or *Jensen* (1998)). In this paper, we propose to model dynamics of large earthquakes, following the work initiated in *Parsons and Geist* (2012). But since *large* earthquakes are rare, it is difficult to fit appropriate stochastic model. Our strategy is then to estimate models for *medium and large-size* earthquakes, linking magnitude of earthquakes and inter-occurrence durations, and then look at the implied dynamics on large ones.

1.2. Description of the dataset. In this paper, we use the ANSS Composite Earthquake Catalog, and selected data are in the Japanese area (see Figure 1), in the ranges of longitude $121^\circ - 155^\circ$ E, latitude $21^\circ - 48^\circ$ N, depth 0-100 km, from January 1st, 1965 till June, 30th, 2012. There are 20,445 events in this dataset, with magnitude higher than 2.7. Following *Christensen et al.* (2002), we shall keep *all* seismic events

from this dataset, since “*there seems to be no rigorous scientific method to distinguish whether an earthquake is a aftershock, because, according to this definition, the categorization depends on the lengths scale, magnitude, and time scale considered*” (Christensen *et al.* (2002), page 2509). Thus, all the observation will be considered, initially, in our study. But in order to have standard distributions for magnitude - for instance- *small* events will be removed from the dataset (and therefore, small aftershocks).

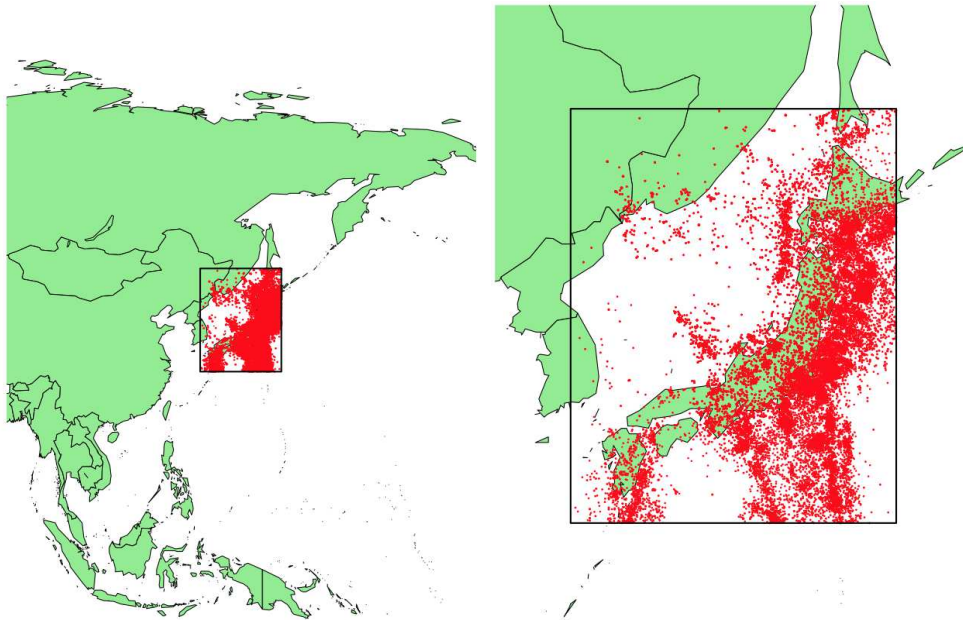


FIGURE 1. Location of earthquakes (overall dataset) from 1965 till 2012.

1.3. Structure of the paper. In section 2, we will look at standard distributions to model magnitude of earthquakes, or seismic moments (in section 2.1), and standard distributions to model occurrence of earthquakes, and inter-time distribution (in section 2.2). Those distributions have to be consistent with empirical observations from *Omori* (1984) or *Gutenberg and Richter* (1941). But assuming that there is no interaction might be a strong assumption. Even assuming a single correlation might be in contradiction with *Pacheco et al.* (1992), which suggested that from small to

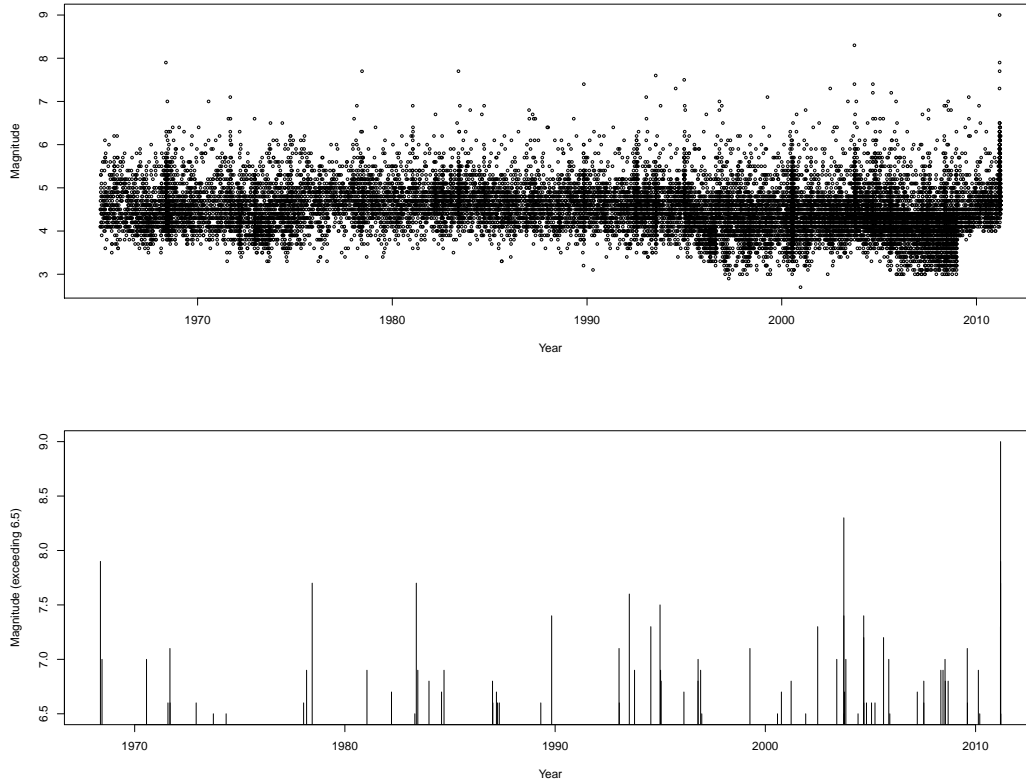


FIGURE 2. Series of earthquakes, from 1965 to 2012, with (on top) all earthquakes above Magnitude 3 in the database, and earthquakes with a Magnitude larger than 6.5 (below).

large earthquakes, the relationship frequency-size should change. Note that conditional distributions are considered here, which is different from *Nikoloulopoulos and Karlis* (2008) where a copula-approach is considered. The advantage is to focus on causal interactions, more than correlation or dependence aspects. Thus, in section 3, we will model magnitude of earthquakes, as a function of time elapses since the previous earthquake. Using a Pareto model, the tail index will be a function of time elapsed. Then in section 4, we model waiting time before the next earthquake as a function of the magnitude of the previous earthquake. Since there is no universal distribution used in literature to model inter-occurrence durations, three models will

be considered (Inverse Gaussian, Gamma and Weibull). Since we focus on events occurred on a small region, the causal relationship between time and magnitude, and then magnitude and time should be valid. Then, in Section 5 we mention how to integrate covariate to improve the models. Finally, in section 6, we use those models to generate scenarios of earthquakes over a century. From those scenarios, we focus only on *large* events, and describe the implied dynamics of earthquakes, from autocorrelations as in *Parsons and Geist* (2012) to distribution of the number of large events over a decade. Note that *Parsons and Geist* (2012) need to focus on worldwide earthquakes in order to have enough large earthquakes. In this paper, we focus on a smaller area (around Japan, see Figure 1), and consider *medium-size* earthquakes.

2. TEMPORAL AND SIZE DISTRIBUTIONS OF EARTHQUAKES

Statistical models for earthquakes are not new, since in the late XIXth century, *Omori* (1984) studied temporal distribution, and half a century later, *Gutenberg and Richter* (1941) studied size distribution. In this section, we will recall standard models used to model inter-earthquakes durations, and magnitude of earthquakes. Recently, ETAS model (combining the Gutenberg-Richter law and the Omori law) has been widely studied (*Gardner and Knopoff* (1974), *Udías and Rice* (1975), *Wang and Kuo* (1998), *Ogata* (1988a), *Ogata* (1988b), *Ogata* (1990)), but (only) one causal aspect is considered, where waiting times are functions of the magnitude of the previous earthquake. *Bak et al.* (2002) suggested a unified scaling law, followed by *Christensen et al.* (2002), *Corral* (2003). *Corral* (2004), *Corral and Christensen* (2006) or *Dauidsen and Goltz* (2005) among others.

2.1. Seismic moment and magnitude distribution. The common idea is that the distribution of seismic magnitude of earthquakes should be consistent with Gutenberg-Richter relation - from *Gutenberg and Richter* (1941) - stating that the expected number of earthquakes exceeding a magnitude $m \geq m_0$ should satisfy

$$\log_{10} N(m) = a - b \cdot m \tag{2.1}$$

where a is the logarithm of sample size for $m \geq m_0$ and b is the coefficient that relates small and large earthquakes, and which is empirically usually close to 1 (see e.g. *Lomnitz (1994)*, *Kagan (1993)*, *Kagan (2010)*, *Utsu (1999)* or *Godano and Pingue (2000)*). From this expression, it is possible to derive a relationship for the logarithm of the probability to exceed some given magnitude $m \geq m_0$, which will be an affine function in m . Let S denote the seismic moment, i.e.

$$M = \frac{2}{3} \log_{10} S - 6 \text{ or conversely } S = 10^{\frac{3}{2}(M+6)}$$

(the seismic moment is here measured in Newton meter, see *Kanamori (1977)* or *Rhoades (1996)*). The standard distribution used for seismic moment is then the Pareto distribution, since

$$\mathbb{P}(S > s) = \left(\frac{s}{s_0}\right)^{-\beta} \text{ for } s_0 \leq s \leq \infty, \quad (2.2)$$

where s_0 is a catalogue threshold (cutoff). Some authors called this a *power law* (see *Mega et al. (2003)*). The density of the seismic moment is then

$$f(s) = \beta s_0^\beta s^{-1-\beta} \text{ for } s \geq s_0.$$

Note that in that case, the average size of an earthquake (in excess of s_0) is

$$\mathbb{E}(S) = \frac{\beta \cdot s_0}{\beta - 1}.$$

Note that since the variable of interest is usually the magnitude, it might be interesting to observe that

$$\mathbb{E}(M) = \frac{2}{3} \mathbb{E}(\log_{10} S) - 6$$

thus

$$\mathbb{E}(\log_{10} S) = \int_0^\infty \mathbb{P}(\log_{10} S > x) dx = \log_{10}(s_0) + \int_{\log_{10}(s_0)}^\infty \left(\frac{10^x}{s_0}\right)^{-\beta} dx$$

i.e.

$$\mathbb{E}(M) = \frac{2}{3} \left(\log_{10}(s_0) + \frac{1}{\beta \log(10)} \right) - 6. \quad (2.3)$$

Note that a Pareto model on seismic moment means that the Magnitude has an exponential type distribution.

In this Pareto model, since s_0 (or equivalently m_0) is usually taken arbitrarily, one should denote in Equation 2.4 $\beta(m_0)$. The maximum likelihood estimator of the tail index is then

$$\hat{\beta}(m_0) = \frac{1}{n} \sum_{i=1}^n \log(s_i) - \log(s_0)$$

where n earthquakes had a seismic moment larger than s_0 . This estimator is popular in the literature on extremal events as Hill estimator, from ?).

On Figure 3 are plotted estimators of the tail index coefficient $\hat{\beta}$ for different values of m_0 . The plain horizon line is $\beta = 0.65$ that is mentioned in *Kagan (1991)* or *Kagan (2002a)*. Note that theoretical analysis of earthquake occurrence (*Vere-Jones (1976)* or *Vere-Jones (1977)*) suggests that, given its branching nature, the exponent β in the Pareto distribution of earthquake size distribution should be equal to $1/2$. The same values of power-law exponents are derived for percolation and self-organized criticality (SOC) processes in a high-dimensional space (*Kagan (1991)* or *Jensen (1998)*).

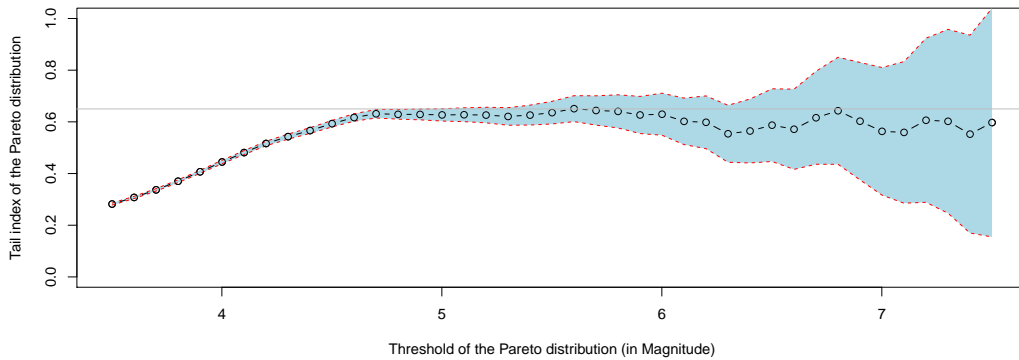


FIGURE 3. Evolution of the tail index of the Pareto distribution as a function of threshold, $\hat{\beta}(m_0)$, also called Hill plot.

Since seismic energy, or seismic moment flux has to be finite, one should introduce an upper limit of magnitude s_∞ . Then the distribution is a truncated Pareto one (see *Wyss (1973)*, *Knopoff and Kagan (1977)*, *Kijko (2004)* or *Holschneider (2011)*). In

that case, the exceedance probability function becomes

$$\mathbb{P}(S > s) = \frac{\left(\frac{s}{s_0}\right)^{-\beta} - \left(\frac{s_\infty}{s_0}\right)^{-\beta}}{1 - \left(\frac{s_\infty}{s_0}\right)^{-\beta}} \text{ for } s_0 \leq s \leq s_\infty, \quad (2.4)$$

It is possible to extend this distribution with an exponential taper applied to the cumulative number of events with seismic moment larger than s_0 (see e.g. *Vere-Jones et al.* (2001)), which yields a tapered Pareto distribution,

$$\mathbb{P}(S > s) = \left(\frac{s}{s_0}\right)^{-\beta} \exp\left[-\frac{s-s_0}{\alpha}\right] \text{ for } s_0 \leq s \leq \infty, \quad (2.5)$$

where α is called the upper corner moment. Note that such a distribution also appear in Pareto's work (see *Pareto* (1896)).

Finally, *Kagan* (1991) or *Main* (2000) used a gamma distribution, where the exceedance probability function is proportional to

$$\mathbb{P}(S > s) \propto \left(\frac{s}{s_0}\right)^{-\beta} \exp\left[-\frac{s-s_0}{\alpha}\right] \left(1 - \left(\frac{s}{\alpha}\right)^{-\beta} \exp\left(\frac{s}{\alpha}\right) \Gamma\left(1 - \beta, \frac{s}{\alpha}\right)\right) \quad (2.6)$$

where $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function $\Gamma(a, x) = \int_x^\infty e^{-t} t^{a-1} dt$.

2.2. Occurrence process and inter-time distribution. The standard model for occurrence processes is the Poisson process (see *Vere-Jones* (1970), *Vere-Jones* (1975) *Shlien and Toksöz* (1970), *Udías and Rice* (1975), *Leonard et al.* (2001), *Rundle et al.* (2005) *Langenbruch et al.* (2011)).

$$f(d) = \lambda \exp[-\lambda d] \text{ for } d \geq 0, \quad (2.7)$$

where $\lambda > 0$ is the rate parameter of this distribution. In that case, $\mathbb{E}(D) = \lambda^{-1}$.

But according to Omori's law - from *Omori* (1984) - the occurrence process should exhibit short-term time dependence that can be observe in empirical data (see *Kagan and Knopoff* (1987a), *Kagan and Knopoff* (1987b), *Davis et al.* (1989), *Davis*

and Frohlich (1991), Enescu et al. (2008) and reference therein). Thus, several authors have investigated laws for waiting times between earthquakes (or at least, seismic events), such as Parsons (2002), Carbone et al. (2005), Huc and Main (2003), Shcherbakov et al. (2005), Corral (2006), Corral (2007), Parsons (2008), Hainzl et al. (2006), Hainzl and Marsan (2008), Lennartz et al. (2008), Molchan (2005), Savage (1994) or Molchan and Kronrod (2007).

An idea is to use an Inverse Gaussian distribution, with density

$$f(d) = \left[\frac{\lambda}{2\pi d^3} \right]^{1/2} \exp \frac{-\lambda(d - \mu)^2}{2\mu^2 d} \text{ for } d \geq 0, \quad (2.8)$$

where $\mu > 0$ is the mean and $\lambda > 0$ is the shape parameter. In that case, $\mathbb{E}(D) = \mu$. Kagan and Knopoff (1987a) or Kagan (2011) suggested to use this distribution to describe the inter-earthquake time distribution. A motivation can be found also in Matthews et al. (2002), since a "Brownian passage-time" property suggests this distribution for inter-time distribution.

The Weibull distribution is the only distribution that has a scale-invariant hazard function, as motivated in Hagiwara (1974), Rikitake (1974), Newman et al. (2005), Hasumi et al. (2009a), Hasumi et al. (2009b), Hasumi et al. (2010) or Ramirez-Rojas et al. (2012), just to mention the most recent articles.

$$f(d) = \frac{k}{\lambda} \left(\frac{d}{\lambda} \right)^{k-1} e^{-(d/\lambda)^k} \text{ for } d \geq 0, \quad (2.9)$$

where $k > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter of the distribution. In this case

$$\mathbb{E}(D) = \lambda \Gamma \left(1 + \frac{1}{k} \right).$$

Note that if $k = 1$, D has an exponential distribution with mean λ^{-1} .

The Gamma distribution has been suggested in Utsu (1984)

$$f(d) = \frac{1}{\theta^k} \frac{1}{\Gamma(k)} d^{k-1} e^{-\frac{d}{\theta}} \text{ for } d \geq 0, \quad (2.10)$$

where $k > 0$ is the shape parameter and $\theta > 0$ is the scale parameter of the distribution. Here, $\mathbb{E}(D) = k\theta$. Note that if $k = 1$, D has an exponential distribution with

mean θ . For convenience, it is possible to write the Gamma density as a function of the exponential family, namely

$$f(d) = \exp\left(\frac{-y/\mu - \log \mu}{\phi} - \phi^{-1} \log \phi + (\phi^{-1} - 1) \log d - \log \Gamma(\phi^{-1})\right),$$

where $\mathbb{E}(D) = \mu$. Parameter ϕ is called dispersion.

3. MAGNITUDE AS A FUNCTION OF TIME BEFORE PREVIOUS EARTHQUAKE

In section 2.1, we have mentioned that the standard model for the seismic moment is a Pareto distribution. Consider now a Pareto generalized linear model, where exceedance probability of the seismic moment is given by

$$\mathbb{P}(S > s | D = d) = \left(\frac{s}{s_0}\right)^{-\beta(d)} = \left(\frac{s}{s_0}\right)^{-\exp[a_0(s_0) + a_1(s_0)d]} \quad \text{for } s_0 \leq s \leq \infty.$$

Since the model is fitted for earthquakes with moment magnitude exceeding s_0 , one should remember that parameters in the linear component of the regression are functions of s_0 (or m_0). It is possible to study significativity of D in the regression by plotting the Student-statistics associated with the test $H_0 : a_1(m_0) = 0$ (against $H_0 : a_1(m_0) \neq 0$), on Figure 4. Thus, if we keep earthquakes with magnitude exceeding 4.5, D could be significant in the regression.

In the Pareto model, with $m_0 = 4.5$, we obtain the following maximum likelihood estimators

$$\widehat{\beta}(d) = \exp\left[-\underset{(0.012250649)}{6.04090623} + \underset{(0.004220243)}{0.01501001} \cdot d\right]$$

and the predicted magnitude is here, according to Equation 2.3

$$\mathbb{E}(M | D = d) = \frac{2}{3} \left(\log_{10}(s_0) + \frac{1}{\widehat{\beta}(d) \log(10)} \right) - 6. \quad (\text{M})$$

(which is decreasing in d). The prediction $\widehat{\beta}(d)$ is plotted on Figure 5 (plain line), as well as a nonparametric spline regression (plain strong line), on top. On the bottom is plotted $\mathbb{E}(M | D = d)$ based on the two estimators (linear versus nonparametric).

The linear regression is statistically valid, but it might not be an appropriate model. Three different time ranges have to be considered, here (see Figure 6)

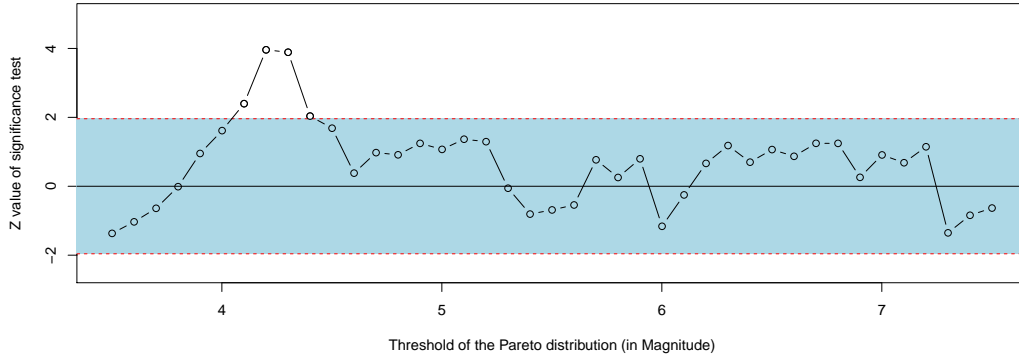


FIGURE 4. Z value of the test $H_0 : a_1(m_0) = 0$ (or equivalently $\beta(d)$ constant for all $d \geq 0$). The centered area is the acceptance region (non-significant coefficient).

- extremely small time d (a few hours), where earthquakes might be aftershocks,
- small time d (less than 3 days), where the expected magnitude is decreasing with time,
- larger time d (more than 3 days), where the expected magnitude is increasing with time.

Note that coefficients are significant for those three periods (and significantly different), see Figure 6.

Since non-parametric models cannot be used to predict values outside the range of observations, a parametric model based on three time ranges has been fitted,

4. INTER-EARTHQUAKE DURATION AS A FUNCTION OF PREVIOUS MAGNITUDE

Several distributions were mentioned in section 2.2 to model the inter-earthquake duration: Weibull (Equation (D₁)), Gamma (Equation (D₂)) and inverse-Gaussian (Equation (D₃)). For the three distributions, generalized linear models can be considered. Again, durations are here in days.

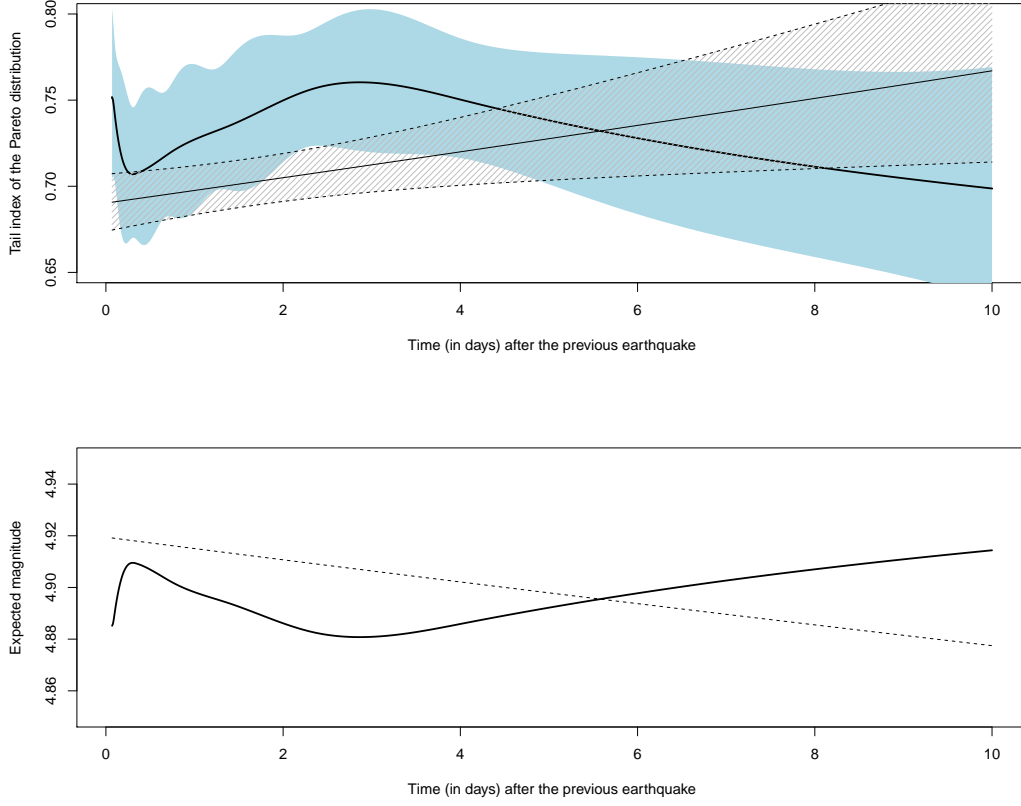


FIGURE 5. On top, $\hat{\beta}(d)$ (including pointwise confidence bands) and below $\mathbb{E}(M|D = d)$, as a function of the time d before the previous earthquake (in days).

In the Weibull model for the conditional duration, parameters k is function of m , but λ is not

$$k(m) = \exp\left[\begin{matrix} 0.49199 - 0.20775 \cdot m \\ (0.087217) \quad (0.017672) \end{matrix}\right] \text{ and } \lambda = \exp\left[\begin{matrix} 0.17745 \\ (0.017727) \end{matrix}\right] \quad (4.1)$$

and the predicted duration is here

$$\mathbb{E}(D|M = m) = \exp(0.17745) \cdot \Gamma\left(1 + \frac{1}{\exp(0.492 - 0.207m)}\right) \quad (D_1)$$

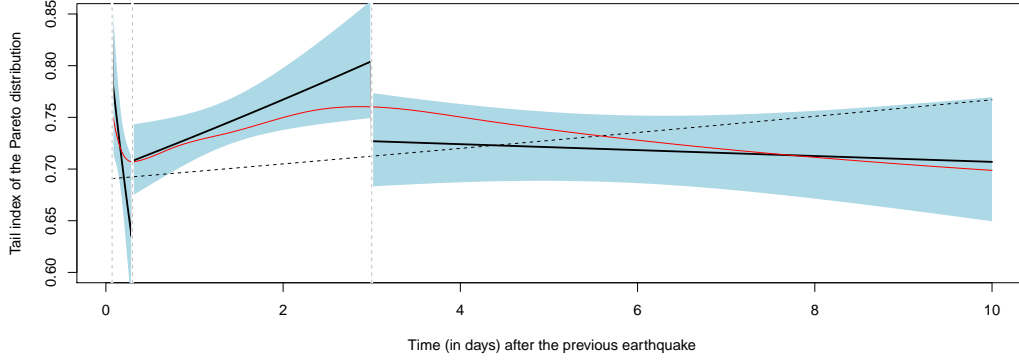


FIGURE 6. Evolution of $\widehat{\beta}(d)$ (including pointwise confidence band) as a function of the time d before the previous earthquake (in days).

which is increasing in m . Nonparametric spline regressions will also be considered, i.e. D given $M = m$ has distribution $W(\mu(m), \lambda)$ where $\mu(\cdot)$ is no longer the exponential of an affine function of m . The prediction can be visualized on Figure 7

In the Gamma model, maximum likelihood estimators of the two parameters are

$$\mu(m) = \exp\left[\underset{(0.16757)}{2.24487} - \underset{(0.03399)}{0.35262} \cdot m\right] \text{ and } \phi = 2.030214,$$

so that the predicted duration is here

$$\mathbb{E}(D|M = m) = \exp\left[\underset{(0.16757)}{2.24487} - \underset{(0.03399)}{0.35262} \cdot m\right], \quad (\text{D}_2)$$

which is decreasing in m . The prediction can be visualized on Figure 8

Finally, in the Inverse Gaussian model, the conditional distribution of D given $M = m$ is given by 2.8, where parameters μ is function of the magnitude of the previous earthquake (while λ is assumed to be constant). Using maximum likelihood techniques, and a log link function, we obtain the following parameters,

$$\mu(m) = \exp\left[\underset{(0.8184)}{2.4148} - \underset{(0.1631)}{0.3869}m\right]$$

so that the predicted duration is here

$$\mathbb{E}(D|M = m) = \mu(m) = \exp(2.4148 - 0.3869 \cdot m), \quad (\text{D}_3)$$

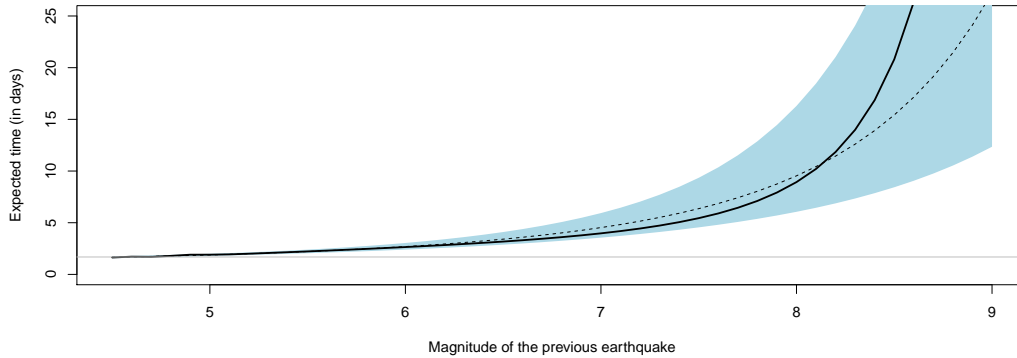


FIGURE 7. Evolution of $\mathbb{E}(D|M = m)$ using a *Weibull* model (including pointwise confidence band) as a function of the magnitude m of the previous earthquake (in days). The plain line is the nonparametric spline regression, while the dotted line is the model described above. The horizontal line is the average time in the dataset (for earthquakes exceeding 4.5), i.e. 1.68 days.

which is decreasing in m . The prediction is extremely close to the one obtained with a Gamma regression.

Clearly, the Weibull model and the Gamma-Inverse Gaussian regressions model very different patterns: On the one hand, with the Gamma-Inverse Gaussian regressions, after a very large earthquake, one should expect events soon in the future (within a few hours). Those are clearly aftershocks. On the other hand, with the Weibull regression, we observe the energy release described in *Bufe and Perkins* (2005) for instance: following a very large earthquake, one should wait a long time before observing another event. The predicted value is rather similar after a small earthquake (36 hours), but they are rather different for a large one: after an earthquake of size 8, the Weibull model suggest that the next event should occur in 10 days, while the Gamma model suggest 10 hours. To have a statistical measure of the goodness of fit, AIC and BIC are given in Table 1. The Inverse Gaussian model is rejected because

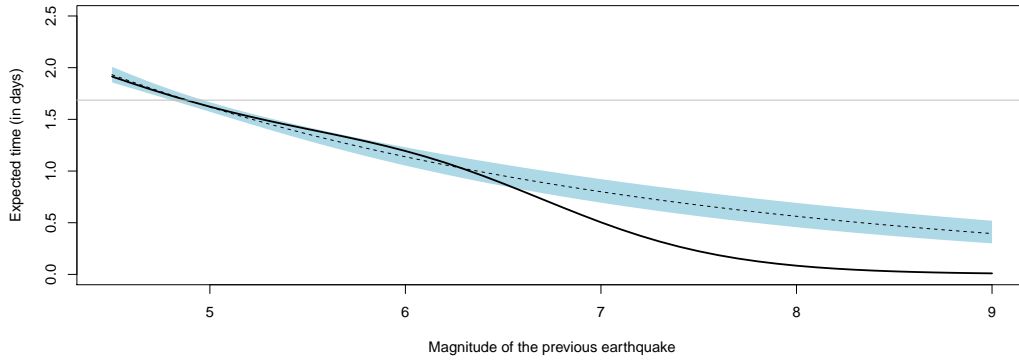


FIGURE 8. Evolution of $\mathbb{E}(D|M = m)$ using a *Gamma* model (including pointwise confidence band) as a function of the magnitude m of the previous earthquake (in days). The plain line is the nonparametric spline regression, while the dotted line is the model described above. The horizontal line is the average time in the dataset (for earthquakes exceeding 4.5), i.e. 1.68 days.

of the high values of AIC and BIC, but the other two models are - statistically - comparabl (even if the Gamma model provides a slightly better fit).

Model	Weibull (D ₁)	Gamma (D ₂)	Inverse Gaussian (D ₃)
AIC	24847.3	24584.6	38640.5
BIC	24868.9	24606.2	38662.2

TABLE 1. Akaike and Schwarz (Bayesian) Criteria for the three duration models.

5. MODELING COVARIATES

In the previous regression models, the only covariate that was considered was either the duration, or the magnitude, related to the previous quake (depending on the variable of interest). But several covariates can be used to improve the model (as

discussed in *Atkinson et al.* (1998) on geomorphology, or more recently, *Bell et al.* (2011) on volcanic eruptions, and references therein).

5.1. Using categorical covariates, e.g. the tectonic plate the previous earthquake was located on.

It is possible to use tectonic plates, for instance (see Figure 9) of the previous quake as a explanatory variable. The proportion of quakes in each plate is given in Table 3. If we use it as a covariate, the model is summarized in Table 2

Variable	Estimate	Std. Error	t value	p -value
plate A	2.48434	0.17096	14.53	<2e-16
plate B	2.51989	0.17458	14.43	<2e-16
plate C	2.66180	0.18423	14.45	<2e-16
plate D	2.73507	0.18025	15.17	<2e-16
plate E	2.20012	0.16980	12.96	<2e-16
Magnitude	-0.38537	0.03438	-11.21	<2e-16
dispersion φ	2.073606			

TABLE 2. Log-gamma regression of the time D as a function of the plate P and the magnitude (of the previous earthquake) M

distance from Tokyo	A	B	C	D	E
≤ 50 km	219	0	0	0	0
≤ 100 km	453	28	0	0	94
≤ 250 km	1,085	235	0	20	1,346
≤ 500 km	1,846	554	0	288	2,842
$\leq 1,000$ km	3,709	1,106	263	486	4,102

TABLE 3. Proportion of quakes in the 5 (primary and secondary) tectonic plates in the Japanese area, A: North American, B:Amur (Chinese) / Eurasian, C:Yangtze and Okinawa , D: Filipino, E: Pacific.

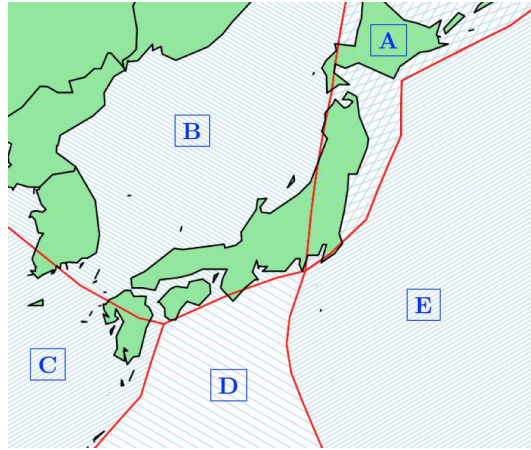


FIGURE 9. The 5 (primary and secondary) tectonic plates in the Japanese area, A: North American, B:Amur (Chinese) / Eurasian, C:Yangtze and Okinawa , D: Filipino, E: Pacific.

The prediction is given on Figure 11

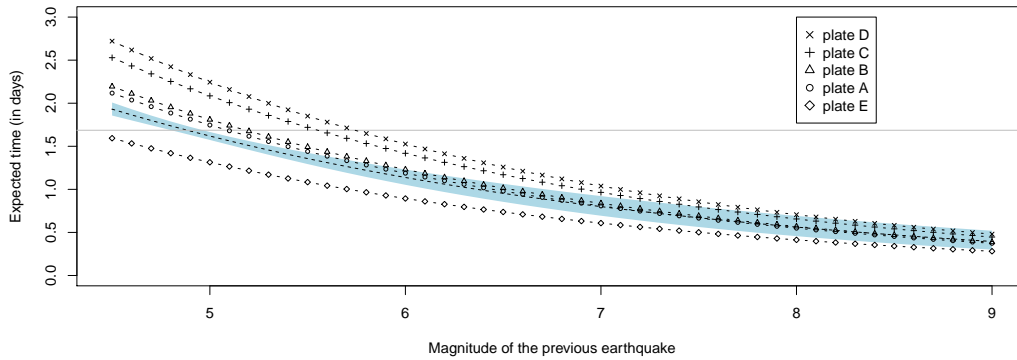


FIGURE 10. Evolution of $\mathbb{E}(D|M = m, P = p)$ using a *Gamma* model (including pointwise confidence band) as a function of the magnitude m of the previous earthquake (in days), and the location of the previous earthquake (the plate it was located on).

5.2. Using continuous covariates, e.g. the distance to a fault of the previous earthquake or the depth.

It is also possible to use other information related to the location of the previous earthquake, e.g. the distance to the border of the plate (which is usually a fault), or the depth of the earthquake.

Variable	Estimate	Std. Error	t value	p -value
(Intercept)	2.1799987	0.1714168	12.72	$< 2e-16$
Magnitude	-0.3855012	0.0345551	-11.16	$< 2e-16$
Depth	0.0019770	0.0001547	12.78	$< 2e-16$
Fault	0.0005866	0.0001340	4.38	$1.2e-05$
dispersion φ	2.096511			

TABLE 4. Log-gamma regression of the time D as a function of depth Y , the distance to a fault F and the magnitude (of the previous earthquake) M

The closer the fault, the more likely some aftershock will be observed.

6. IMPLIED DYNAMICS OF CAUSAL MODELING

In order to describe the dynamics implied by those models, 1,000 scenarios are generated, over a century. The iterative process is the following: we start from with average magnitude earthquake, then given that magnitude, we generate a waiting time, before the occurrence of the next event. Then, given the time elapsed, we generate a magnitude. And then, given that magnitude, we generate a waiting time, etc. Three models have been considered here,

- (B) as a benchmark, we generate independent events: independent magnitudes, and independent inter-occurrence durations, from an exponential distribution (occurrence is then driven by a standard homogeneous Poisson process)
- (W) one model, with Pareto seismic moments (with three ranges for durations), and Weibull inter-occurrence durations (linear in the previous magnitude),
- (G) one model, with Pareto seismic moments (with three ranges for durations), and Gamma inter-occurrence durations (linear in the previous magnitude),

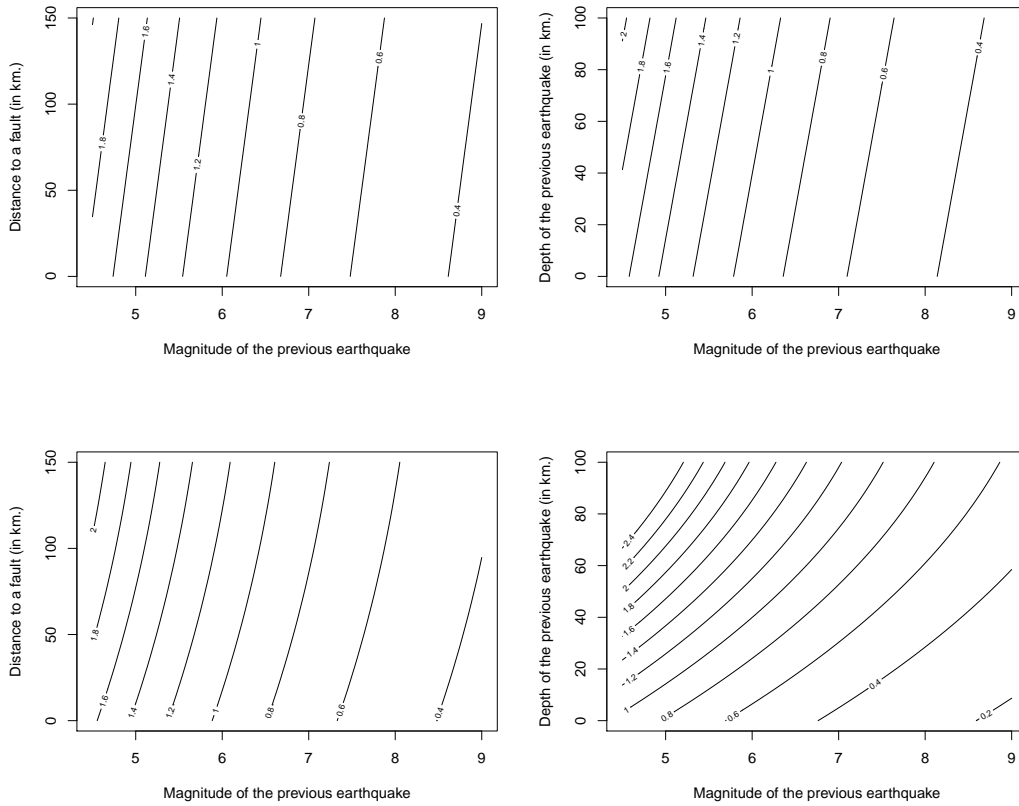


FIGURE 11. Level curves of $\mathbb{E}(D|M = m, F = f)$ using a *Gamma* model as a function of the magnitude m of the previous earthquake (in days), and the distance to the closest fault (in km.), on the left, and level curves of $\mathbb{E}(D|M = m, Y = y)$ as a function of the magnitude m of the previous earthquake (in days), and its depth (in km.), on the right. On top a standard log-Gamma additive model, and below when the covariate (either Y or F) is smoothed.

On Figure 12, one trajectory of those three models is plotted. In order to compare those models, the same set of uniform variates (used is the generation of conditional variables) is used.

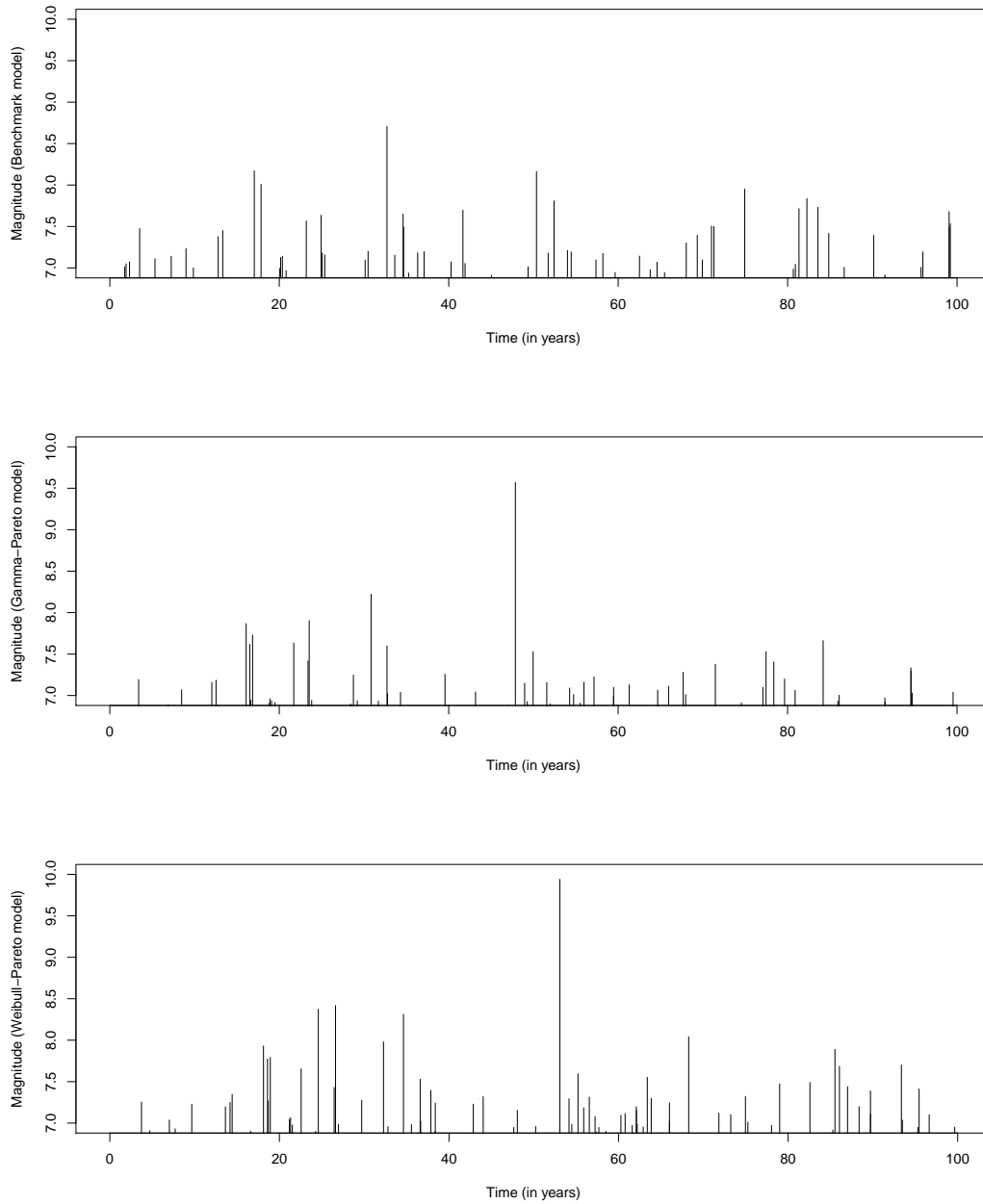


FIGURE 12. Earthquakes with magnitude exceeding 8, over a century, with the benchmark model (B) on top, then the Gamma-Pareto (G) below, and finally the Weibull-Pareto (W) in the bottom.

On Figure 13 are plotted the distribution of the number of large earthquakes, per year or per decade, where *large* is either larger than 7, 7.5 or 8.

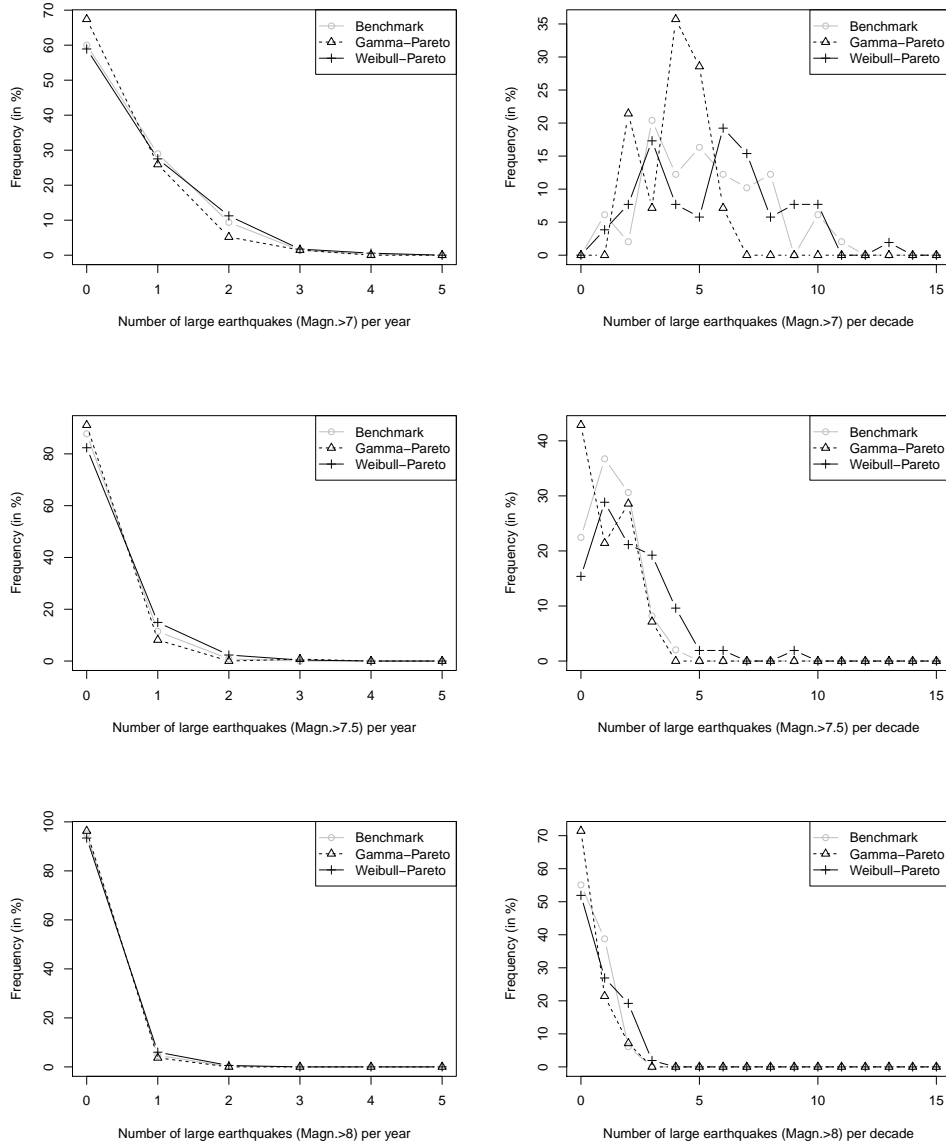


FIGURE 13. Number of large events, per year (on the left) or per decade (on the right), where *large* means (from top to bottom) a magnitude larger than 7, 7.5 or 8.

7. INFLUENCE OF THE AREA CONSIDERED

As a rule of thumb, $M=7$ events may rupture 50 to 80 km from the fault, whereas an $M=8$ might affect 150 to 400 km from the fault (see *Parsons (2008)*). Our main interest here is the region close to Tokyo, see Figure 14. Three different scenarios are considered, to compute the number of large events, per year or per decade,

- on Figure 15 are estimated the distribution of the number of big earthquakes, per year or per decade, within 1,000 km from Tokyo,
- on Figure 16 within 250 km,
- on Figure 17 within 100 km.

A summary can be found on Figure 18 where the probabilities to have - at least - two major earthquakes, closed to Tokyo, per year (on the left) or per decade (on the right) is plotted. As expected, the kind of dependence we consider will have a strong impact on the distribution of the number of large earthquakes. With the Weibull Pareto model, the probability of having two large earthquake is large, two times larger than a model with independence between duration and magnitude.

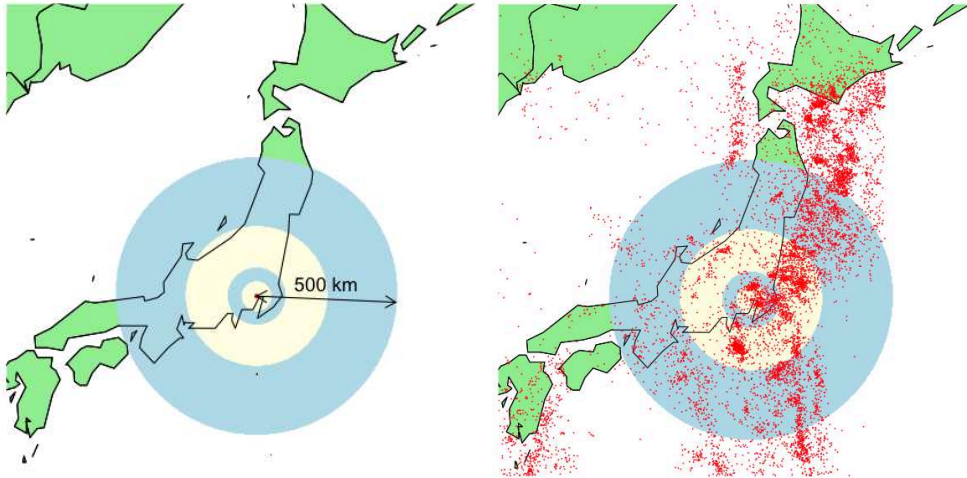


FIGURE 14. Four regions around Tokyo ($139^{\circ}45''$ E - $35^{\circ}40''$ N): less than 1,000km, less than 500 km, less than 250 km, 100 km and 50 km.

REFERENCES

- Aki, K. 1956. Some problems in statistical seismology. *Journal of Seismological Society of Japan*, **8**, 205-227.
- Atkinson, P., Jiskoot, J., Massari, R. & Murray, T. (1998). Generalized linear modelling in geomorphology. *Earth Surface Processes and Landforms*, **23**:13, 1185-1195.
- Bak, P., & Tang, C. 1989. Earthquakes as a self-organized critical phenomenon, *Journal of Geophysical Research*, *94*:15, 635-637.
- Bak, P., Christensen, K. , Danon, L. & Scanlon, T. 2002. Unified scaling law for earthquakes *Physics Review Letters*, **88**, 178501
- Bakun, W.H. & Lindh, A.G. 1985. The Parkfield, California, earthquake prediction experiment *Science*, **229**, 619-624
- Benioff, H. 1951. Earthquakes and rock creep. *Bulletin of the Seismological Society of America*, **41**:1, 31-62.
- Bell, A.F., Naylor, M., Heap, M.J. & Main, I.G. (2011). Forecasting volcanic eruptions and other material failure phenomena: An evaluation of the failure forecast method. *Geophysical Research Letters*, **38**, L15304
- Bufe, C. G. & Perkins, D.M. 2005. Evidence for a Global Seismic-Moment Release Sequence. *Bulletin of the Seismological Society of America*, **95**, 833-843.
- Carbone, V., Sorriso-Valvo, L., Harabaglia, P., & Guerra, I. 2005. Unified scaling law for waiting times between seismic events. *Europhysics Letters*, **71**:6, 1036-1042.
- Christensen, K. Danon, L. Scanlon, T. & Bak, P. (2002). Unified scaling law for earthquakes *Proceedings of the National Academia of Sciences U. S. A.*, **99**, 2509-2513.
- Corral, A. 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes *Physics Review, E*, **68**, 035102
- Corral, A. 2004. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes, *Physics Review Letters*, **92**.
- Corral, A. 2006. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics*, 424:177 - 193.

- Corral, A. 2007. Statistical features of earthquake temporal occurrence, Lecture Notes in Physics, **705**, 191-221.
- Corral, A. & Christensen, K. 2006. Comment on "Earthquakes Descaled: On Waiting Time Distributions and Scaling Laws" *Physics Review Letters*, **96**, 109801.
- Daividsen, J. & Goltz, C. 2005. Are seismic waiting time distributions universal? *Geophysics Research Letters*, **31**, L21612.
- Davis, P.M., Jackson, D.D. & Kagan, Y.Y. 1989. The longer it has been since the last earthquake, the longer the expected time till the next? *Bulletin of the Seismological Society of America*, **79**, 1439-1456.
- Davis, S.D. & Frohlich, C., 1991. Single-link cluster analysis of earthquake aftershocks – decay laws and regional variations, *Journal of Geophysical Research*, **96**, 6335-6350.
- Enescu, B., Struzik, Z. & Kiyono, K. 2008. On the recurrence time of earthquakes: insight from Vrancea (Romania) intermediate-depth events. *Geophysical Journal International*, **172**:1, 395-404.
- Gardner, J.K. & Knopoff, L. 1974. Is the sequence of earthquakes in Southern California, with aftershocks removed Poissonian? *Bulletin of the Seismological Society of America*, **64**, 1363-1367
- Gilbert, G. K. 1909. Earthquake forecasts, *Science*, **29**, 121-138.
- Godano, C. & Pingue F., 2000. Is the seismic moment-frequency relation universal?, *Geophysical Journal International* , **142**, 193-198.
- Gutenberg, B. & Richter, C.F., 1941. *Seismicity of the Earth*, *Geological Society of America, Special papers*, **34**, pp. 1-131.
- Hagiwara, Y. 1974. Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. *Tectonophysics*, **23**, 323-318
- Hainzl, S., Scherbaum, F. & Beauval, C. 2006. Estimating background activity based on interevent-time distribution *Bulletin of the Seismological Society of America*, **96**, 313-320
- Hainzl, S. & Marsan, D. 2008. Dependence of the Omori-Utsu law parameters on main shock magnitude: Observations and modeling, *Journal of Geophysical Research*,

113

- Hasumi, T., Akimoto, T. & Aizawa, Y. 2009. The Weibull - Log Weibull Distribution for Interoccurrence Times of Earthquakes *Physica A*, **388**, 491-498.
- Hasumi, T., Akimoto, T. & Aizawa, Y. 2009. The Weibull - Log Weibull transition of the interoccurrence time statistics in the two-dimensional Burridge-Knopoff earthquake model *Physica A*, **388**, 483.
- Hasumi, C., Chen, T., Akimoto & Aizawa, Y. 2010. The Weibull-log Weibull transition of interoccurrence time for synthetic and natural earthquakes. *Tectonophysics*, **485**, 9.
- Holschneider, M., Zöller, G., & Hainzl, S. 2011. Estimation of the Maximum Possible Magnitude in the Framework of a Doubly Truncated Gutenberg-Richter Model. *Bulletin of the Seismological Society of America*, **101**:4, 1649-1659.
- Huc, M. & Main, I.G. 2003. Anomalous stress diffusion in earthquake triggering: correlation length, time dependence, and directionality. *Journal of Geophysics Research*, **108**, 2324.
- Jensen, H.J. 1998. Self-organized criticality. Cambridge University Press.
- Kagan, Y. Y. 1991. Seismic moment distribution, *Geophysical Journal International*, **106**, 123-134.
- Kagan, Y.Y. 1993 Statistics of characteristic earthquakes, *Bulletin of the Seismological Society of America*, **83**, 7-24.
- Kagan, Y.Y., 2002a. Seismic moment distribution revisited: I. Statistical results *Geophysical Journal International* , **148**:3, 520-541.
- Kagan, Y.Y., 2010. Earthquake size distribution: Power-law with exponent $\beta=1/2$? *Tectonophysics*, **490**, 103-114.
- Kagan, Y. Y., 2011. Random stress and Omori's law, *Geophysical Journal International* , **186**:3, 1347-1364
- Kagan, Y.Y. & Jackson, D.D. 1991. Seismic Gap Hypothesis: Ten Years After. *Journal of Geophysical Research*, **96**:21, 419-421.
- Kagan, Y.Y. & Jackson, D.D., 1995. New seismic gap hypothesis: Five years after, *Journal of Geophysical Research*, **100**, 3943-3959.

- Kagan, Y.Y., & Knopoff, L. 1987a. Statistical short-term earthquake prediction, *Science*, **236**, 1563-1567.
- Kagan, Y.Y., & Knopoff, L. 1987b. Random stress and earthquake statistics: Time dependence. *Geophysical Journal. Royal Astronomical Society - A*, **88**, 723-731.
- Kanamori, H., 1977. The energy release in great earthquakes, *Journal of Geophysical Research*, **82**, 2981–2987.
- Kerr, R.A. 2004. Parkfield keeps secrets after a long-awaited quake. *Science*, **306**, 206-207.
- Kijko, A. 2004. Estimation of the Maximum Earthquake Magnitude, m_{\max} . *Pure and Applied Geophysics*, **161**, 1-27.
- Knopoff, L. & Kagan, Y.Y., 1977. Analysis of the theory of extremes as applied to earthquake problems, *Journal of Geophysical Research*, **82**, 5647-5657.
- Langenbruch, C., Dinske, C. & Shapiro, S. A. (2011). Inter event times of fluid induced earthquakes suggest their Poisson nature. *Geophysical Research Letters*, **38**.
- Lennartz S, Livina VN, Bunde A, & Havlin S. 2008. Long-term memory in earthquakes and the distribution of interoccurrence times. *Europhysics Letters*, **81**.
- Leonard, T., Papsouliotis O. & Main, I.G., 2001. A Poisson model for identifying characteristic size effects in frequency data: Application to frequency-size distributions for global earthquakes, “starquakes” and fault lengths, *Journal of Geophysical Research*, **106**:13, 473-484.
- Lombardi, A.M., Marzocchi, W. 2007. Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes. *Journal of Geophysics Research*, **112**.
- Lomnitz, C. 1994. Fundamentals of earthquakes prediction. John Wiley & Sons.
- Main, I., 2000. Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: fact or artifact?, *Bulletin of the Seismological Society of America*, **90**, 86-97.
- Matthews, M. V., W. L. Ellsworth, and P. A. Reasenberg, 2002. A Brownian model for recurrent earthquakes, *Bulletin of the Seismological Society of America*, **92**, 2233-2250.

- McCann, W.R., Nishenko, S.P., Sykes, I.R. & Krause, J. 1979. Seismic gaps and plate tectonics: seismic potential for major boundaries. *Pure and applied geophysics*, **117**, 1082-1147.
- Mega, M.S., Allegrini, P., Grigolini, P., Latora, V., Palatella, L., Rapisarda, A. & Vinciguerra, S. 2003. Power-law time distribution of large earthquakes. *Physics Review Letters* **90**:18, 188501.
- Molchan, G. 2005. Interevent time distribution in seismicity: a theoretical approach *Pure and Applied Geophysics*, **162**, 1135-1150.
- Molchan, G. & Kronrod, T. (2007) Seismic interevent time: a spatial scaling and multifractality. *Pure and Applied Geophysics*, **164**, 75-96.
- Murray, J. & Segall, P. 2002. Testing time-predictable earthquake recurrence by direct measurement of strain accumulation and release. *Nature*, **419**, 287-291.
- Newman, W., Turcotte, D. L., Shcherbakov, R. & Rundle, J. B. 2005. Why Weibull? *Proceedings of the American Geophysical Union*, Fall Meeting 2005, #S43D-07
- Nikolouloupoulos, A.K & Karlis, D. 2008. Fitting copulas to bivariate earthquake data: the seismic gap hypothesis revisited *Environmetrics*, **19**, 251–269.
- Nishenko, S.P. & Sykes, L.R. 1993. Comment on 'Seismic gap hypothesis: ten years after' by Y.Y. Kagan and D.D. Jackson. *Journal of Geophysical Research*, **98**, 9909-9916.
- Ogata, Y., 1988a. Statistical models for earthquake occurrence and residual analysis for point processes, *Journal of the American Statistical Association*, **83**, 9-27.
- Ogata, Y., 1998b. Space-time point-process models for earthquake occurrences, *Annals of the Institute of Statistical Mathematics*, **50**:2, 379-402.
- Ogata, Y., Seismicity analysis through point-process modeling: a review. *Pure and Applied Geophysics*, **155**, 471-507.
- Ogata, Y. & Katsura, K., 1993. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, *Geophysical Journal International* , **113**, 727-738.
- Omori, F. 1894. On aftershocks. *Journal of the College of Science, Imperial University of Tokyo*, **7**, 111-200.

- Pacheco, J.F., Scholz, C.H. & Sykes, L.R., 1992. Changes in frequency-size relationship from small to large earthquakes, *Nature*, **355**, 71-73.
- Pareto, V. 1896. Cours d'économie politique, tome 2. Lausanne, F. Rouge.
- Parsons, T. 2002. Global Omori law decay of triggered earthquakes: Large aftershocks outside the classical aftershock zone. *Journal of Geophysical Research*, **107**, 2199.
- Parsons, T. 2008. Earthquake recurrence on the south Hayward fault is most consistent with a time dependent, renewal process. *Geophysical Research Letters*, **35**.
- Parsons, T. & Geist, E.L. 2012. Were global $M \geq 8.3$ earthquake time intervals random between 1900-2011? *Bulletin of the Seismological Society of America*, **102**, 1-11.
- Ramírez-Rojas, A. Flores-Marquez , E.L. & Valverde-Esparza, S. 2012. Weibull Characterization of inter-event lags of Earthquakes occurred in the Pacific Coast of Mexico. *Geophysical Research Abstracts*, **14**, 13708.
- Reid, H. F. 1910. The Mechanics of the Earthquake. Vol. 2 of The California Earthquake of April 18, 1906. Report of the State Earthquake Investigation Commission (Carnegie Institution of Washington Publication 87).
- Rhoades, D.A. 1996 Estimation of the Gutenberg-Richter relation allowing for individual earthquake magnitude uncertainties *Tectonophysics*, **258**, 71-83.
- Rikitake, T. 1974. Probability of an earthquake occurrence as estimated from crustal strain, *Tectonophysics*, **23**, 299-312.
- Rundle, J.B., Rundle, P.B., Shcherbakov, R. & Turcotte, D.L. 2005. Earthquake waiting times: Theory and comparison with results from Virtual California simulations. *Geophysical Research Abstracts*, **7**, 2477.
- Savage, J.C. (1994). Empirical earthquake probabilities from observed recurrence intervals. *Bulletin of the Seismological Society of America*, **84**:1, 219-221.
- Shcherbakov, R., Yakovlev, G., Turcotte, D.L., & Rundle, J.B. 2005. Model for the distribution of aftershock interoccurrence times. *Physics Review Letters*, **95**, 218501.
- Shlien, S. & Toksöz,, M.N. 1970: A clustering model for earthquake occurrences. *Bulletin of the Seismological Society of America*, **60**, 1765-1787.

- Sieh, K. 1996. The repetition of large-earthquake ruptures. *Proceedings of the National Academy of Sciences U. S. A.*, **93**, 3764-3771.
- Sieh, K., Stuiver, M. & Brillinger, D. 1989. A more precise chronology of earthquakes produced by the San Andreas fault in Southern California *Journal of Geophysics Research*, **94**, 603-623.
- Stein, R.S. 1995. Characteristic or hazard? *Nature*, **370**, 443-444.
- Stein, R.S. 2002. Parkfield's unfulfilled promise. *Nature*, **419**, 257-258.
- Thatcher, W. 1989. Earthquake recurrence and risk assessment in circum-Pacific seismic gaps, *Nature*, **341**, 432-434.
- Udías, A. & Rice, J. 1975. Statistical analysis of microearthquake activity near San Andreas Geophysical Observatory, Hollister, California. *Bulletin of the Seismological Society of America*, **65**, 809-828.
- Utsu, T. 1984. Estimation of parameters for recurrence models of earthquakes. *Bulletin of the Earthquake Research Institute of the University of Tokyo*, **59**, 53- 66
- Utsu, T., 1999. Representation and analysis of the earthquake size distribution: A historical review and some new approaches, *Pure and Applied Geophysics*, **155**, 509-535.
- Vere-Jones, D., 1970. Stochastic models for earthquake occurrence (with discussion), *Journal of Royal Statistical Society, Series B*, **32**, 1-62.
- Vere-Jones, D., 1975. Stochastic models for earthquake sequences. *Geophysical Journal of the Royal Astronomical Society*, **42**, 811-826.
- Vere-Jones, D., 1976. A branching model for crack propagation, *Pure and Applied Geophysics*, **114**, 711-725.
- Vere-Jones, D., 1977. Statistical theories of crack propagation, *Mathematical Geology*, **9**, 455-481.
- Vere-Jones, D., Robinson, R. & Yang, W. 2001. Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation *Geophysical Journal International*, **144**:3, 517-531.
- Wang, J.H, & Kuo, C.-H. 1998. On the frequency distribution of interoccurrence times of earthquakes. *Journal of Seismology*, **2**, 351-358.

- Wang, Qi, David, Jackson, D. & Zhuang, J. 2010, Missing links in earthquake clustering models *Geophysics Research Letters*, **37**.
- Weldon, R. Scharer, K. Fumal, T. & Biasi, G. 2004. Wrightwood and the earthquake cycle: what a long recurrence record tells us about how faults work. *GSA Today*, **14:9**, 4-10
- Wyss, M. (1973), Towards a Physical Understanding of the Earthquake Frequency Distribution. *Geophysical Journal of the Royal Astronomical Society*, **31**, 341-359.
- Zhuang, J., Ogata, Y. & Vere-Jones, D. 2002. Stochastic declustering of space-time earthquake occurrences *Journal of the American Statistical Association*, **97**, 369-380
- Zhuang, J., Ogata, Y. & Vere-Jones, D. 2004 Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research*, **109**

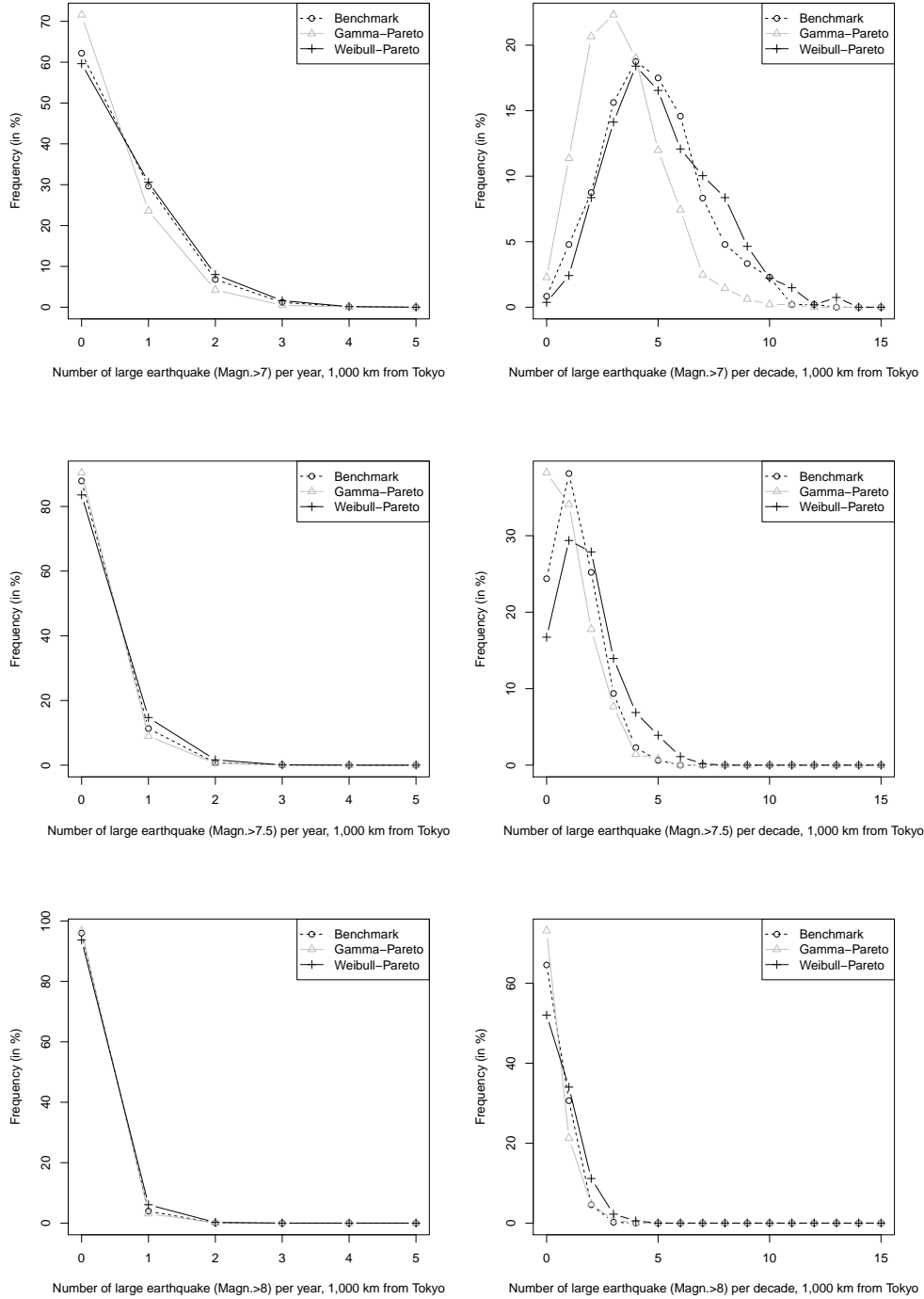


FIGURE 15. Number of large events, per year (on the left) or per decade (on the right), where *large* means (from top to bottom) a magnitude larger than 7, 7.5 or 8, within 1,000 km from Tokyo

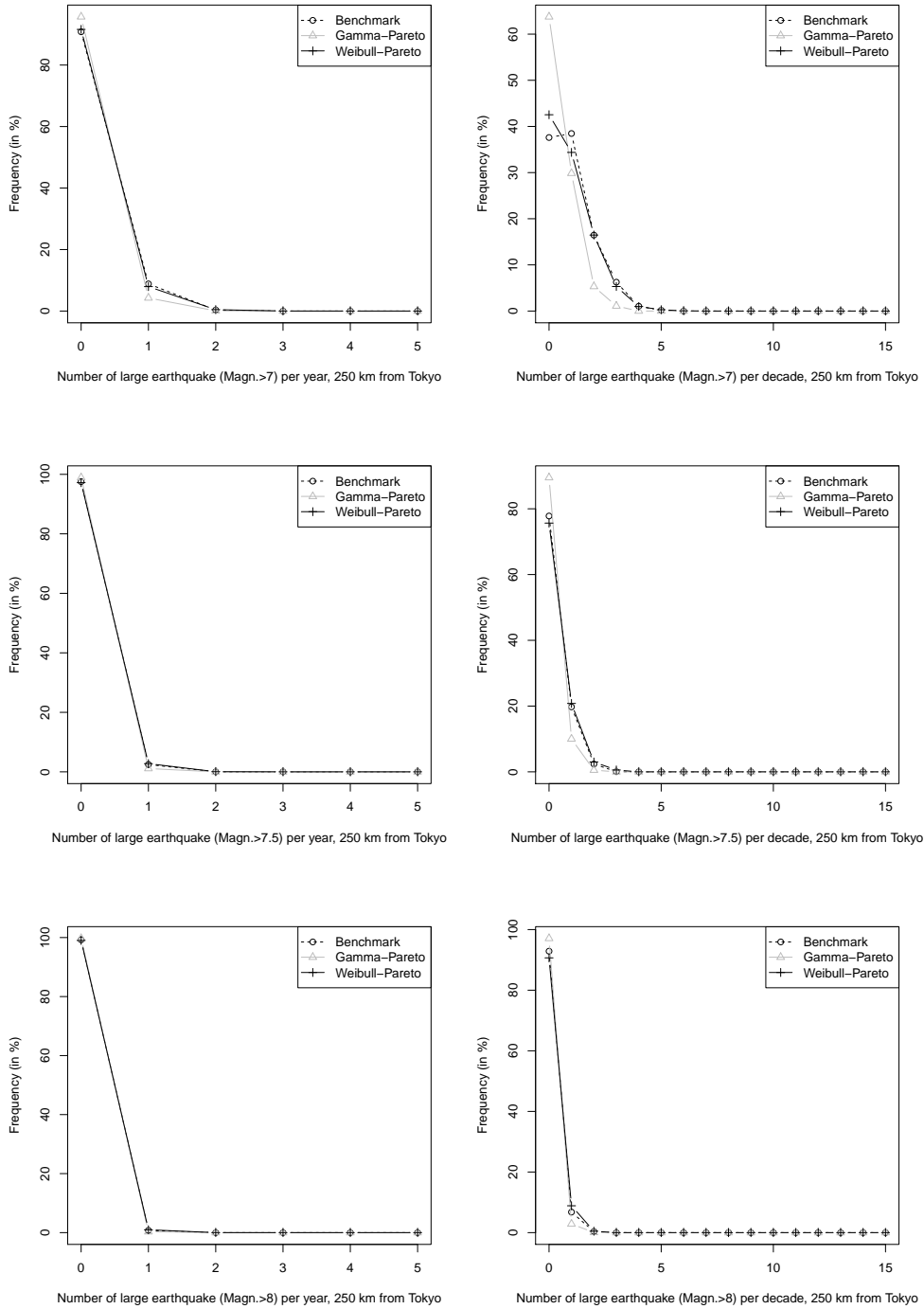


FIGURE 16. Number of large events, per year (on the left) or per decade (on the right), where *large* means (from top to bottom) a magnitude larger than 7, 7.5 or 8, within 250 km from Tokyo

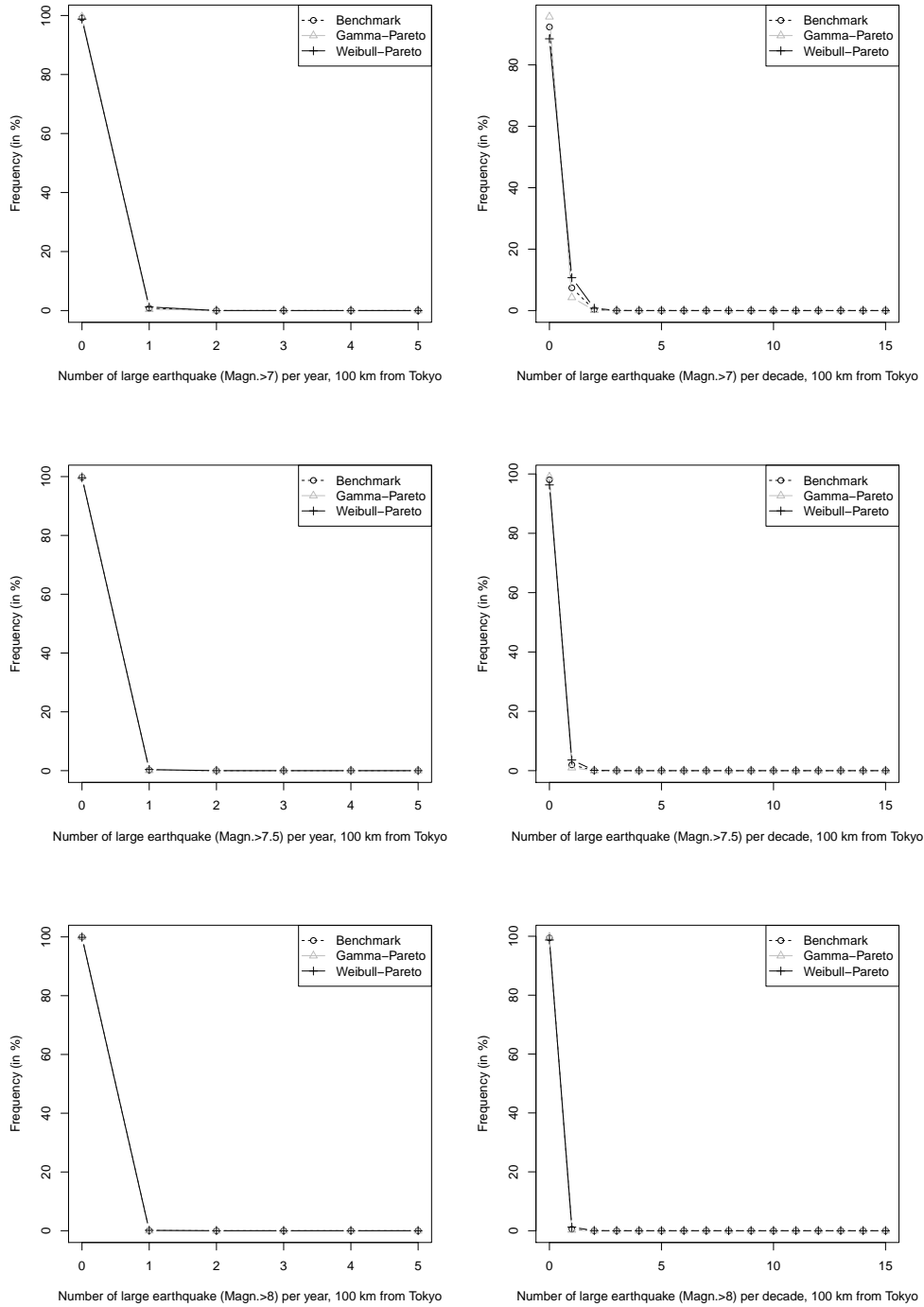


FIGURE 17. Number of large events, per year (on the left) or per decade (on the right), where *large* means (from top to bottom) a magnitude larger than 7, 7.5 or 8, within 100 km from Tokyo

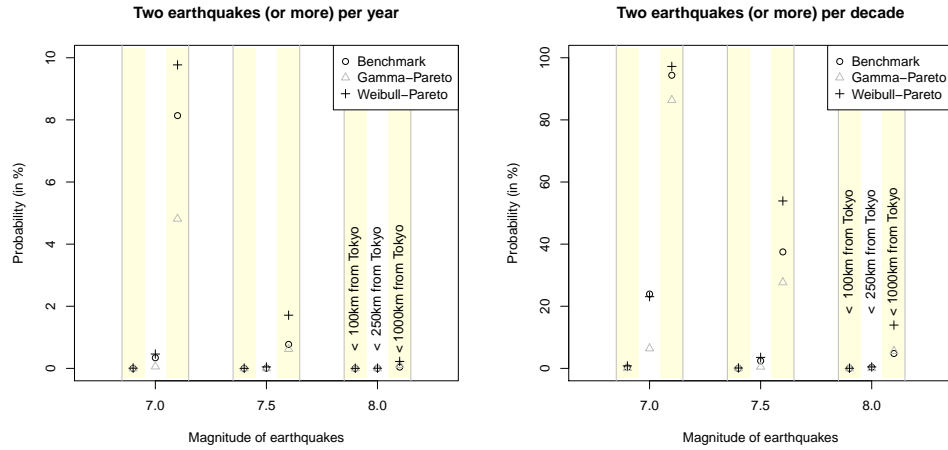


FIGURE 18. Probability to have two big earthquakes (or more) per year (on the left) or per decade (on the right), for different magnitudes (7, 7.5 and 8), within 100km from Tokyo, 250km or 1,000 km.