Abstract. We consider two statistical regularities that were used to explain Omori’s law of the aftershock rate decay: the Lévy and Inverse Gaussian (IGD) distributions. These distributions are thought to describe stress behaviour influenced by various random factors: post-earthquake stress time history is described by a Brownian motion. Both distributions decay to zero for time intervals close to zero. But this feature contradicts the high immediate aftershock level according to Omori’s law. We propose that these statistical distributions are influenced by the power-law stress distribution near the earthquake focal zone and we derive new distributions as a mixture of power-law stress with the exponent $\psi$ and Lévy as well as IGD distributions. Such new distributions describe the resulting inter-earthquake time intervals and closely resemble Omori’s law. The new Lévy distribution has a pure power-law form with the exponent $-(1 + \psi/2)$ and the mixed IGD has two exponents: the same as Lévy for small time intervals and $-(1 + \psi)$ for longer times. For even longer time intervals this power-law behaviour should be replaced by a uniform seismicity rate corresponding to the long-term tectonic deformation. We compute these background rates using our former analysis of earthquake size distribution and its connection to plate tectonics. We analyze several earthquake catalogues to confirm and illustrate our theoretical results. Finally, we discuss how the parameters of random stress dynamics can be determined through a more
detailed statistical analysis of earthquake occurrence or by new laboratory experiments.

**Short running title:** RANDOM STRESS AND OMORI’S LAW

**Key words:** Omori’s law; Random stress; Aftershocks; Lévy distribution; Inverse Gaussian distribution; Seismicity as result of plate-tectonic deformation; Selection bias.

1 Introduction

Our aim is to investigate the connection between random stress and the Omori law of aftershock occurrence. Several attempts to explain Omori’s law have been published (see, for example, Kagan & Knopoff, 1987; Dieterich, 1994). Since stress in the Earth’s interiors cannot be easily measured or calculated, these studies usually consider stress as a scalar stochastic variable, ignoring for a while its tensorial nature. In this work we consider the stress as a scalar.

In analyzing earthquakes here we use the scalar seismic moment $M$ directly, but for easy comparison we convert it into an approximate moment magnitude using the relationship

$$m_W = \frac{2}{3} (\log_{10} M - 9.0),$$

(Hanks, 1992), where $M$ is measured in Newton m (Nm). Since we are using almost exclusively the moment magnitude, later we omit the subscript in $m_W$.

Kagan (1982) considered stress time history as a Brownian motion, i.e., randomly fluctuating stress acting on the stressed environment of an earthquake focal zone. Here the probability density for the time intervals between earthquake events has a Lévy distribution with a power-law tail having the exponent $-3/2$. Kagan & Knopoff (1987) added a tectonic drift component to the Brownian motion. For such a case, the distribution of inter-event times is the Inverse Gaussian distribution (IGD) which, depending on the value of the initial
stress drop and velocity of tectonic motion, can exhibit occurrence patterns ranging from
the Lévy distribution to a quasi-periodic occurrence.

Matthews et al. (2002) also proposed the IGD as a law for inter-earthquake intervals;
however, they considered only a limiting, quasi-periodic long-term distribution. The authors
suggest that “... [the Inverse Gaussian] distribution has the following noteworthy properties:
(1) the probability of immediate [large earthquake] rerupture is zero; (2) the hazard rate
increases steadily from zero at $t = 0$ to a finite maximum near the mean recurrence time ...
” and real earthquake occurrence follows the same pattern. For both practical (evaluation
of the seismic hazard) and theoretical reasons it is important to consider for and against
arguments for quasi-periodicity of earthquake occurrence.

Ellsworth (1995) and Ellsworth et al. (1999) present several sequences of characteristic,
quasi-periodic earthquakes mostly in western U.S., which, in their opinion, confirm a smaller
coefficient of variation than the Poisson process for earthquake recurrence intervals. Nadeau
et al. (1995) and Nadeau & Johnson (1998, and references therein) show a similar quasi-
periodic pattern for microearthquakes in the Parkfield area.

There is ample evidence that after almost any earthquake subsequent events follow
Omori’s law. This has been observed for smaller earthquakes (aftershocks) as well as large
events comparable in size to the original shock or even exceeding it (Kagan & Jackson, 1999;
Parsons, 2002). Omori’s law pattern of power-law rate decaying seismicity is observed at
decadal and century long scales (Utsu et al., 1995; Stein & Liu, 2009; Ebel, 2009).

Omori’s law aftershock behaviour is observed even for acoustic emission events (mi-
croearthquakes) (Ojala et al., 2004; Nechad et al., 2005) with the decay parameter value
similar to that determined for large and moderate earthquakes. Thus, the Omori law is a
general feature of earthquake occurrence observed for any brittle deformation of solids.

The presumed pattern of earthquake quasi-periodicity depends on the argument that
stress drops to zero or close to zero in the focal zone of an earthquake as suggested, for example, by Shimazaki & Nakata (1980). Thus, it would be necessary to wait for some time before a critical stress level is reached. Kagan & Jackson (1999) presented several examples of large earthquakes which followed after a short time interval (often just a few days) within a focal zone of similar large events. Such an inter-earthquake interval is clearly insufficient for stress to be replenished by tectonic motion. We present additional evidence for this pattern later in this paper.

Thus, for relatively short time intervals, earthquakes are clustered in time with the coefficient of variation of interevent times significantly greater than that for the Poisson process (Kagan & Jackson, 1991). Why is it widely believed that large earthquakes are quasi-periodic in time? Evidence mostly comes from specifically selected sequences (Bakun & Lindh, 1985; Ellsworth, 1995) or from paleo-seismic investigations (Ellsworth et al., 1999). However, paleo-seismic investigations have poor time resolution: two or more events occurring closely in time are likely to be identified as one event. Hence, their coefficient of variation estimates should be biased towards smaller values: a more periodic pattern. Some paleo-seismic investigations (Marco et al., 1996. Rockwell et al., 2000; Ken-Tor et al., 2001; Dawson et al., 2003; Kagan et al., 2011) also suggest that earthquakes follow a long-term clustering pattern at least in regions of slow tectonic deformation.

Additionally, paleo-seismicity studies the Earth’s displacement only at the surface. But temporal and spatial distributions of earthquakes rupturing the surface at one site substantially differ from general earthquake statistics (Kagan, 2005). The earthquake size distribution is also significantly different for site- and area-based statistics. Rigorously reliable statistical properties that are relevant for theoretical studies as well as for seismic hazard estimates can only be obtained by analyzing instrumental earthquake catalogues.

Measurements from instrumental earthquake catalogues indicate that short time intervals
between large earthquakes are much more frequent than even the Poisson model would sug-
gest (Kagan & Jackson, 1991; 1999). Moreover, the problem of biased selection is also hard
to avoid in historical and paleo-seismic data: sequences not displaying suggested patterns or
with a small number of events are less likely to be considered.

Sequences of microearthquakes in the Parkfield area (Nadeau et al., 1995; Nadeau &
Johnson, 1998) exhibit certain properties of characteristic quasi-periodic earthquakes: reg-
ularity of occurrence and nearly identical waveforms. They show many other interesting
features of earthquake occurrence. However, these micro-events are not characteristic earth-
quakes in a strict sense: they do not release almost all the tectonic deformation on a fault
segment while real characteristic earthquakes are assumed to do so.

As with large characteristic earthquakes, attempts using microearthquake quasi-
periodicity for their prediction have not yet succeeded. Zechar, J. D. (Southern California
Earthquake Center 2010 Annual meeting report, p. 106) attempted to forecast repeating
small earthquakes on the San Andreas fault near Parkfield, based on event regularity. While
retrospective tests indicated that the forecasts based on recurrence times were better than
random, the forward tests were unsuccessful.

Moreover, these micro-events apparently occur on isolated asperities surrounded by creep-
ing zones. Therefore, such an asperity exists as a secluded entity and may produce charac-
teristic quasi-periodic rupture events, similar to fracture in many laboratory experiments.
But tectonic earthquakes occur in an environment that is never isolated. This may explain
their ubiquitous clustering in time and space.

The quasi-periodic model for characteristic earthquakes (see for example, McCann et al.,
1979; Bakun & Lindh, 1985; Nishenko, 1991) has been put to a prospective test using new
data. Testing the validity of this hypothesis ended in failure (Rong et al., 2003 and references
therein; Bakun et al., 2005; Jackson & Kagan, 2006).
How can we investigate earthquake occurrence patterns in various tectonic zones? Our statistical studies of seismicity (see for example, Kagan, 1991, 2007; Kagan & Jackson, 1991; Kagan et al., 2010) are biased because earthquake rate is dominated by places with a high tectonic rate. Thus, the likelihood function maximum for our stochastic model (see more in Subsection 3.3) is mainly determined by these earthquakes. However, we should study earthquake patterns in continental areas (active and non-active) where the earthquake rate is low, but the vulnerable human population is large (see Stein & Liu, 2009; Parsons, 2009). A naive extrapolation of the aftershock sequence rates by assuming that the present rate would continue, may exaggerate seismic hazard. Instead, we need a convincing tool to produce a truer estimate.

Our recent results suggest that earthquake occurrence can be modeled by the Poisson cluster process: the combination of the long-term rate caused by tectonic strain with short-term Omori-type clustering. For large and medium strain rates, the model appears to work well (see again Kagan et al., 2010). The model is producing reasonable long- and short-term forecasts of earthquake occurrence both for local (Helmstetter et al., 2006) and global (Kagan & Jackson, 2011) seismicity. But to extrapolate such a model to small tectonic rates we would need additional arguments; we must extract the rates by comparing seismicity patterns with tectonic strain maps. Bird et al. (2010) discuss the dependence of the long-term inter-earthquake rate on strain. We need to show how the short-term properties of earthquakes behave in the regions of varying strain. As shown below, the transition time from Omori’s law decay to a quasi-stationary rate clearly increases with a decrease in the strain rate. As the result of these studies, we propose that the same statistical model can be applied for regions of both high and low tectonic strain.

In Section 2 below we consider two distributions to explain Omori’s law by random stress dynamics: the Lévy and Inverse Gaussian (IGD) distributions. We also derive a
modification of these distributions due to power-law character of the random stress in a focal area of an earthquake. The resulting distributions offer a better explanation of the observed Omori’s law decay in aftershock sequences. Section 3 discusses statistical analysis of earthquake occurrence with an emphasis on Omori’s decay of seismicity after a strong event. Two statistical techniques are used for the analysis: a second-order temporal moment of earthquake sequences (pair-wise interevent distribution) and the stochastic branching point processes. These two methods provide a quantitative description of the short-term behaviour of earthquake occurrence and its transition to the background Poisson rate. In Section 4, we review and update the earlier results (a) determining the background earthquake rate on the basis of the known tectonic deformation rate and (b) evaluating the size of the earthquake aftershock area for various focal mechanisms and in different tectonic zones. These results as well as the results of Section 3 allow us to calculate both short- and long-term seismicity rate for seismically active and less active areas. Subsection 4.3 illustrates an application of these techniques by calculating both rates for the New Madrid earthquake sequence.

2 Aftershock temporal distribution, theoretical analysis

2.1 Lévy distribution

Kagan (1982) proposed a heuristical model for an earthquake fracture as follows. At the moment that any given earthquake rupture stops, the stress near the edge of the rupture is lower than the critical breaking stress for extension of the fracture. The subsequent stress history near the earthquake rupture tip depends on other fractures in the neighborhood and additional random factors. In this case, the time history of the stress might resemble a
Brownian random walk. The stress-time function is thus given as a solution to the diffusion equation. When the stress reaches a critical threshold level, a new earthquake begins. The distribution density of the time intervals of the ‘first passage’ (Feller, 1966) is the Lévy distribution

\[ f(t) = \frac{\sigma}{\sqrt{2 \pi D t^3}} \exp \left( -\frac{\sigma^2}{2Dt} \right), \]

where \( D \) is the diffusion coefficient, \( t \) is the time, and \( \sigma \) is the threshold or barrier stress: the difference between initial breaking stress and that when rupture ceases. This distribution as a function of stress is the Rayleigh law; as a function of time it is the Lévy distribution (cf. Zolotarev, 1986). In this model, the stress is taken to be a scalar, which corresponds to the addition of perfectly aligned stress tensors; the interaction of misaligned tensors will be discussed in a later work. The Lévy distribution (2) was used by Kagan (1982, see Eqs. 8 and 9) to model the time dependent part of the space-time sequence of microearthquakes. Here we are address only the time sequence of events.

Fig. 1 displays several Lévy curves for various values of stress drop, \( \sigma \). The curve’s tail is a power-law similar to that exhibited by Omori’s law. However, for small values of time the curves decay to zero: depending on the stress drop, Brownian motion takes time to reach a critical stress level. This latter feature, though observed in aftershock sequences, is most likely caused by effects from coda waves of both the mainshock and stronger aftershocks that prevent identifying smaller aftershocks (Kagan, 2004). This opinion is supported by aftershock registration in the high-frequency domain (Enescu et al., 2009) where their time delay is substantially reduced. Measurements of total seismic moment rate decay in aftershock sequences (Kagan & Houston, 2005) suggest that aftershocks start right after the end (within one minute for a \( M6.7 \) earthquake) of a mainshock rupture. Therefore, the left-hand decay of the curves in Fig. 1 is not realistic.

There is another problem in comparison of theoretical distribution with aftershock ob-
servations. The Lévy curves may explain the inter-event time statistics only for the first generation (parent-child) of clustered earthquakes. Most observed aftershocks or clustered events are generally separated from their ‘parents’ by multiple generations. For example, the present aftershocks of the 1811-12 New Madrid earthquakes or the 1952 Kern County, California, earthquake are mostly distant relatives of their progenitors (mainshocks). In the Omori law we count all aftershocks, disregarding their parentage, whereas the Lévy law describes the time distribution for the next event: when the stress level reaches a critical value to trigger a new rupture. When we count all events, we decrease the exponent from 1.5 to 1.0 (Kagan & Knopoff, 1987).

We assume that the next earthquake may occur in any place in the focal zone of a previous event. Attempts to localize the initiation of strong earthquakes have not been successful; for example, the 2004 Parkfield earthquake occurred well outside the previously specified area (Jackson & Kagan, 2006).

We also assume that the stress drop after an earthquake has a power-law distribution: it is proportional to $\sigma^{-1-\psi}$, for $0 \leq \psi < 1.0$. Kagan (1994), Marsan (2005), Helmstetter et al. (2005), Lavallée (2008), Powers & Jordan (2010) support this statement.

Then for the Lévy distribution we obtain

$$\phi(t) \propto \int_0^\infty f(t) \sigma^{-1-\psi} d\sigma = \frac{1}{2 t \sqrt{\pi} (2 t D)^{\psi/2}} \Gamma\left(\frac{1-\psi}{2}\right),$$

where $\Gamma$ is a gamma function. The new modified Lévy distribution density function is a power-law with the exponent $-(1 + \psi/2)$. The density should be truncated at the left side, because aftershocks cannot be observed at short-time intervals close to a mainshock (Kagan, 2004). Normalizing the distribution would depend on this truncation. Kagan (1991, Eq. 8) introduced a minimum time ($t_M$), where $M$ is a seismic moment of a preceding event (see also Eq. 17 below). Before that time, the aftershock rate is measured with a substantial
undercount bias (Kagan, 2004).

2.2 Inverse Gaussian distribution

As Matthews et al. (2002) indicate, the Inverse Gaussian Distribution (IGD) has been known since 1915. However, this distribution only acquired its name in 1945. This is the name found in recent statistical encyclopedias and books (for example, Kotz et al., 2006; Seshadri, 1999; Chhikara & Folks, 1989), on the Wolfram website, and in Wikipedia, while the “Brownian passage-time” (BPT) is little known. Kagan & Knopoff (1987) proposed to use this distribution (without calling it IGD) to describe the inter-earthquake time distribution. Their major impetus was to explain aftershock statistics (Omori’s law) by stress dynamics.

Here we discuss the appropriate pairwise interval law for a model in which a steadily increasing source of stress, which we take to be due to plate tectonics, is added to or subtracted from a random or diffusion component, where the distribution (2) describes the density of earthquake recurrence times in the absence of tectonic loading. If the rate of tectonic loading is a constant $V$, the distribution density $f(t)$ is modified to (Feller, 1966, p. 368) become the Inverse Gaussian distribution

$$f(t) = \frac{\sigma}{\sqrt{2\pi D t^3}} \exp \left[ -\frac{(\sigma - Vt)^2}{2Dt} \right].$$ (4)

For tectonic loading velocity $V = 0$, this equation transforms to (2). Fig. 2 displays several IGD curves for various values of stress drop. For small values of $\sigma$ and $t$, the curves are similar to the values of Lévy distribution (2): if the stress drop is insignificant, or the tectonic loading influence is small in initial stages of stress behaviour.

Assuming again that the stress in a fault neighborhood is distributed according to a power-law, we obtain a new modified distribution of inter-earthquake time, based on the
\[\phi(t) \propto \int_0^\infty f(t) \sigma^{-1-\psi} d\sigma = \exp\left(-V^2 t/(2D)\right) \frac{\Gamma\left(1-\psi\right)}{2 t \sqrt{\pi} \left(2 f D\right)^\psi/2} \left[\Gamma\left(\frac{1-\psi}{2}\right) \right] _1 F_1 \left(\frac{1-\psi}{2}, \frac{1}{2}, \frac{V^2 t}{2D}\right) \]
\[+ V \sqrt{\frac{2t}{D}} \Gamma \left(1-\frac{\psi}{2}\right) _1 F_1 \left(1-\frac{\psi}{2}, \frac{3}{2}, \frac{V^2 t}{2D}\right), \tag{5}\]

where \(_1 F_1\) is a Kummer confluent hypergeometric function (Wolfram, 1999; Abramowitz & Stegun, 1972, p. 504). Another expression for the density is

\[\phi(t) \propto \exp\left(-V^2 t/(2D)\right) \frac{\Gamma\left(1-\psi\right)}{2 t \sqrt{\pi} \left(2 f D\right)^\psi/2} \left[\Gamma\left(\frac{1-\psi}{2}\right) \right] _1 F_1 \left(\frac{1-\psi}{2}, \frac{1}{2}, \frac{V^2 t}{2D}\right) U\left(\frac{1-\psi}{2}, \frac{1}{2}, \frac{V^2 t}{2D}\right), \tag{6}\]

where \(U\) a confluent hypergeometric function (ibid.). As in Eq. 3, both of the above distributions should be truncated as \(t \to 0\) and can be normalized after it.

For certain values of \(\psi\) the expressions (5 and 6) can be simplified. For example for \(\psi = 0\)

\[\phi(t) \propto \frac{1}{2t} \left[1 + \text{erf}\left(V \sqrt{\frac{t}{2D}}\right)\right]. \tag{7}\]

For \(\psi = 0.5\) we obtain two equations. For a positive \(V\), using Eq. 13.6.3 by Abramowitz & Stegun (1972), we transform (5) into

\[\phi(t) \propto \frac{1}{2t} \sqrt{\frac{\pi V}{4D}} \exp\left(-\frac{V^2 t}{4D}\right) \left[I_{-1/4}\left(\frac{V^2 t}{4D}\right) + I_{1/4}\left(\frac{V^2 t}{4D}\right)\right], \tag{8}\]

where \(I_{-1/4}\) and \(I_{1/4}\) are modified Bessel functions (Wolfram, 1999; Abramowitz & Stegun, 1972, p. 374).

For a negative \(V\) using Eq. 13.6.21 by Abramowitz & Stegun (1972), we transform (6) into

\[\phi(t) \propto \frac{1}{2t} \sqrt{-\frac{V}{\pi D}} \exp\left(-\frac{V^2 t}{4D}\right) K_{1/4}\left(\frac{V^2 t}{4D}\right), \tag{9}\]

where \(K_{1/4}\) is a modified Bessel function (Wolfram, 1999; Abramowitz & Stegun, 1972).

In Fig. 3 we display the new IGD curves for various values of the \(\psi\) parameter. Although the general behaviour of the curves remains power-law, the curves change their slope at the
time value of about 1.0. The curves with \( V = \pm \sqrt{D} \) show the distribution difference for tectonic loading sign; in the positive case tectonic movement is opposite to the fault displacement during an earthquake, whereas the negative sign corresponds to motion consistent with the earthquake mechanism. In the latter case, random fluctuations can bring the fault to rupture only during the early period of development.

To show the difference of curve slopes more clearly, the PDF values of Fig. 3 are multiplied by \( t^{1+\psi/2} \) in Fig. 4. For small time intervals the curves are horizontal, suggesting that the modified IGD is similar to the Lévy distribution in Eq. 3.

For small time values, the power-law exponents in Fig. 3 are essentially the same as for compounded Lévy law (3). This can be seen from modified Bessel function approximations for small values of the argument (Abramowitz & Stegun, 1972, Eqs. 9.6.7 and 9.6.9),

\[
I_{-1/4}(t) \propto \left(\frac{2}{t}\right)^{1/4} \quad \text{and} \quad K_{1/4}(t) \propto \left(\frac{2}{t}\right)^{1/4}.
\]  

In the general case the same result can be obtained with Eq. 13.5.5 by Abramowitz & Stegun (1972). Then for \( t \to 0 \) Eq. 5 transforms into

\[
\phi(t) \propto t^{-1-\psi/2}.
\]  

It is obvious from Figs. 3 and 4 that, except for the \( \psi = 0 \) curve, the slope of curves for large values of the time increases when compared with \( t \) close to zero. Using Eq. 13.1.27 of Abramowitz & Stegun (1972) we add an exponential term in (5) within the hypergeometric function \( _1F_1 \). Then for \( t \to \infty \) we obtain

\[
\phi(t) \propto t^{-1-\psi},
\]  

(Abramowitz & Stegun, 1972, Eq. 13.5.1). Therefore, in Eq. 8 the exponent would be \(-1.5\) (Abramowitz & Stegun, 1972, Eq. 9.7.1), whereas in Eq. 9 the power-law term (Abramowitz
& Stegun, 1972, Eq. 9.7.2) is multiplied by an exponential decay term

\[ \phi(t) \propto t^{-1.5} \exp \left( -\frac{V^2 t}{4D} \right). \]  

(13)

See Figs. 3 and 4.

3 Temporal distribution of aftershocks: Observations

Many observations of Omori’s law behaviour have been published (see Utsu et al., 1995 and references therein). There are some problems with these measurements. The standard interpretation of the Omori law is that all aftershocks are caused by a single mainshock. However, the mainshock is often followed by strong aftershocks and those are clearly accompanied by their own sequence of events, and so on. The second problem is that some earthquakes have very few or no aftershocks; such sequences cannot be included in the naive study of earthquake sequences, but can be analyzed as stochastic processes.

Thus, three techniques can be applied to study the temporal distribution of real earthquakes: 1) traditional, phenomenological techniques based on observing individual aftershock sequences; 2) using statistical moments of earthquake occurrence, considered as a point process; 3) applying stochastic process modeling to infer the parameter values of earthquake temporal interaction.

3.1 Aftershock sequences

Beginning with Omori (1894), the temporal distribution of aftershock numbers has been studied for more than one hundred years (Utsu et al., 1995). The aftershock rate decay is approximated as \( t^{-p} \) with the parameter \( p \) value close to 1.0 (ibid.; Kagan, 2004).

But the simple, superficial study of aftershock rate decay often encounters serious prob-
lems. First, only relatively rich aftershock sequences can be investigated by direct measurements; if there are too few aftershocks, their properties can be studied only by combining (stacking) many sequences. Second, to isolate individual sequences one should exclude any cases when one sequence is influenced by another, an arbitrary procedure which may introduce a selection bias. Third, an aftershock sequence often contains one or several large events which are clearly accompanied by a secondary aftershock sequence. Taking the influence of secondary earthquakes into account is not simple (see Section 3.3 for more detail). Fourth, some sequences start with a strong foreshock which is sometimes only slightly weaker than a mainshock. Again, handling this occurrence presents serious problem.

Therefore, strong bias may result from directly measuring Omori’s law exponents. Two other statistical methods considered below enable analysis of the whole earthquake occurrence as a point process, to minimize the problem of data and interpretation technique selection bias.

### 3.2 Temporal distribution for earthquake pairs

Kagan & Jackson (1991) investigated space-time pairwise earthquake correlation patterns in several earthquake catalogues. They showed that earthquake pairs follow a power-law distribution for small time and distance intervals. Kagan & Jackson (1999) showed that contrary to the seismic gap model (McCann et al., 1979; Nishenko, 1991; Matthews et al., 2002) the clustering pattern continues for strong ($m \geq 7.5$) earthquakes. Large shallow earthquakes can re-occur after a small time interval and follow the Omori-type temporal distribution.

Table 1 shows the location and focal mechanism difference for $m \geq 7.5$ global shallow earthquakes in the GCMT catalogue (Ekström et al., 2005; Ekström, 2007) from 1976-2010. The table format is similar to Table 1 in Kagan & Jackson (1999). However, here we kept
only those pairs in the table for which their focal zone overlap (the \( \eta \)-parameter) is greater than 1.0:

\[
\eta = \frac{L_1 + L_2}{2R},
\]

where \( L_1 \) and \( L_2 \) are the respective rupture lengths for the first and second earthquakes in the pair (see Eq. 3 in Kagan & Jackson, 1999) and \( R \) is the distance between the centroids. Therefore, if \( \eta \geq 1.0 \) the earthquake focal zones would intersect. For several doublets \( \eta \geq 2 \), implying that the smaller event should be largely within the focal zone of the larger earthquake. Inspecting the time difference and the 3-D rotation angle between focal mechanisms suggests that these high \( \eta \) pairs may occur after very short time intervals and have very similar double-couple mechanisms.

All earthquakes in the table occur in subduction zones as defined in Kagan et al. (2010). However, even with relatively high deformation velocity at these plate boundaries, the inter-earthquake time is in most cases substantially lower than the time necessary for tectonic motion to restore the critical stress conditions by the occurrence time of the second earthquake (see the last column in Table 1 by Kagan & Jackson, 1999).

Fig. 5 shows how the normalized number of \( m \geq 6.5 \) shallow earthquake pairs depends on the tectonic deformation rate as defined by Bird et al. (2010). Three curves are shown: all earthquakes from the GCMT catalogue, earthquakes from subduction zones (Kagan et al., 2010), and events from active continental zones. As expected, earthquakes in the trench zones occur in higher rate zones compared to all earthquakes, while most (about 80%) of the earthquake pairs in active continents are concentrated in zones with the deformation rate less than \( 20 \times 10^{-20} \) events/m\(^2\)/s.

In the following displays we would like to show the distribution of temporal intervals between earthquakes when measured in a catalogue of a limited duration, \( T \). Fig. 6 shows the integration domain for calculating earthquake pair rates. The diagram is a square with a
side length equal to a catalogue duration, \( T \); since the plot is symmetric, only the lower-right portion of the square is shown. The first event shown as a filled circle, is supposed to be at the square diagonal, the second one at the end of a hatched area. We assume that the time difference between earthquakes cannot be less than \( t_0 \) (similar to \( t_M \) in Eq. 17).

For the Poisson process, the interval pair density is uniform (see Kagan & Jackson, 1991, Eq. 1). Thus, the rate is proportional to the hatched area. For the normalized survival rate

\[
   n_p = \left( \frac{T - t}{T - t_0} \right)^2, \tag{15}
\]

where \( t_0 \) is the minimum time interval, \( T \) is the catalogue duration and \( t \) is the inter-earthquake time interval.

For the power-law time distribution with distribution density \( \phi(t) \propto t^{-1-\theta} \), we obtain the normalized survival rate by integrating over the domain shown in Fig. 6:

\[
   n_p = \frac{(t^{-\theta} - T^{-\theta})}{(t_0^{-\theta} - T^{-\theta})}. \tag{16}
\]

Unfortunately, the number of \( m \geq 7.5 \) earthquakes shown in Table 1 is too small to find the variation of the earthquake pair time differences in various tectonic zones. Such an analysis can be performed only for \( m \geq 6.5 \) earthquakes. For such events, their spatial separation becomes comparable with the location errors in the GCMT catalogue (Kagan & Jackson, 1999; Kagan et al., 2010). The results are less reliable when rupture areas for earthquake pairs can be superimposed due to the location errors.

However, even for \( m \geq 6.5 \) events the pair numbers are usually too small, especially for longer time intervals, to carry out rigorous statistical procedures. A simulation or a bootstrap procedure can be used as an alternative method to evaluate uncertainties. However, our model of earthquake occurrence is a branching process governed by fractal distributions. A simulation has not yet been developed in such a situation (see more discussion in Kagan,
2007; 2010, as well as in description of Fig. 12 below). Thus, we use a largely qualitative analysis of the pair temporal distributions.

In Figs. 7-10 the temporal distribution of inter-earthquake times for all \( m \geq 6.5 \) event pairs in the GCMT catalogue is shown as it depends on earthquake rate determined by Bird et al. (2010). Several normalized approximations for pair intervals are also given: the Poisson distribution of earthquakes (Eq. 15) in the time span 1976-2010 and several power-law interval (Eq. 16) dependencies.

The distribution curves consist of two parts: for small time intervals, they follow a power-law and for larger intervals the distribution is parallel to the Poisson rate. Similar results for smaller earthquakes are obtained by Touati et al. (2009). The transition from one behaviour to another occurs sooner for zones with a higher tectonic deformation rate: by the curves inspection we find that in Fig. 7 the transition is observed for the time of about 6,000 days; in Fig. 8 it is about 4,000-5,000 days, and in Fig. 9 it is less than 3,000 days.

Fig. 10 shows the time interval distribution for active continental areas. In this case a visual inspection suggests that the best approximation is the power-law; no transition to the Poisson rate is observable. This absence can be explained by a low deformation rate in these zones (see Fig. 5). As was observed for many aftershock sequences in continental and slowly deforming areas, aftershock sequences continue according to Omori’s law for decades and even centuries (Utsu et al., 1995; Stein & Liu, 2009; Ebel, 2009). Here the span of 34 years covered by the GCMT catalogue is likely insufficient to demonstrate the transition from an aftershock sequence to a background, Poisson rate.

We compute an average recurrence time (\( \bar{t} \)) for earthquakes in Figs. 7-10. The smallest value for \( \bar{t} \) is observed for distributions with a significant component of the power-law: Figs. 7 and 10. We also calculate the coefficient of variation (\( C_v \)) of earthquake inter-occurrence time as a ratio of the standard deviation (\( \sigma_t \)) to the average time \( \bar{t} \) (Kagan
& Jackson, 1991, Fig. 1). A completely random Poisson occurrence has the coefficient of variation equal to one, whereas quasi-periodicity yields a coefficient of less than one. For clustered earthquakes the coefficient is larger than one. Although the $C_v$ estimates are biased downwards when determined at a relatively short catalogue time span, their mutual relations are indicative of occurrence patterns. For Figs. 7-10, the $C_v$-values are 1.10, 0.925, 0.841, and 1.977, respectively. These evaluations again suggest that earthquakes in areas with a smaller tectonic rate become more clustered, and their Poisson component is diminished.

3.3 Stochastic branching processes

As we mentioned above, the mainshock is often followed by strong aftershocks and those are clearly accompanied by their own sequence of events, and so on. The patterns of multiple clustering have been described by stochastic branching processes (Hawkes & Adamopoulos, 1973; Kagan, 1991; Ogata, 1998). In this model, seismicity is approximated by a Poisson cluster process, in which clusters or sequences of earthquakes are statistically independent, although individual events in the cluster are triggered. However, Brémaud and Massoulié (2001) proposed a model in which all events belong to one branching cluster.

Kagan et al. (2010) applied the Critical Branching Models (CBM) to analyze earthquake occurrences statistically in several global and regional earthquake catalogues. Time intervals between earthquakes within a cluster (i.e., between any two dependent shocks) are assumed to be distributed according to a power-law

$$
\psi_{\Delta t}(\Delta t) = \theta t_M^{\theta}(\Delta t)^{-1-\theta}, \quad \Delta t \geq t_M.
$$

This is similar to Omori’s law. The parameter $\theta$ is an ‘earthquake memory’ factor, $t_M$ is the coda duration time of an earthquake with seismic moment $M_i$, and $\mu$ is the branching
(productivity) coefficient (Kagan et al., 2010, Eq. 9):

\[ \psi_M(M_i) = \mu \times \left( \frac{M_i}{M_t} \right)^{\delta}, \]

where the non-normalized function \( \psi_M(M_i) \) corresponds to the number of triggered shocks generated on the average by an earthquake with seismic moment \( M_i \) (Kagan, 1991) and \( M_t \) is the moment threshold: the smallest moment above which the catalogue can be considered to be complete.

During a likelihood search the \( \theta \)-values have been restricted within the interval \( 0.1 \leq \theta \leq 1.0 \); smaller or larger estimates are inadmissible because of physical considerations (Kagan, 1991; Kagan et al., 2010). These and similar constraints during the likelihood inversion make estimating parameter errors and the correlation between these estimates a difficult and unreliable procedure. Wang et al. (2010) carry out the statistical error analysis for the ETAS (Epidemic Type Aftershock Sequence) model that is similar to our CBM (see more in Section 6.2 by Kagan et al. 2010). They find significant correlations between statistical ETAS parameter estimates.

Therefore we investigate a correlation between estimates for two parameters \( \theta \) and \( \mu \) of the CBM, these parameters are most important for our discussion. In Fig. 11, most \( \theta \)-estimates are within the interval \( 0.1 \leq \theta \leq 0.5 \) and are negatively correlated with the estimates of the \( \mu \) coefficient. We also show the correlation coefficients of the parameter estimates for various sets of analyzed subcatalogues. The subscripts \( i \) at correlation coefficients (\( \rho_i \)) point to the table number in Kagan et al. (2010). These correlation coefficients also indicate a negative correlation of estimates.

Several determinations of the time decay exponent are carried out for the ETAS model. Ogata (1998, Tables 2-3) obtained the \( p \)-values (equivalent to our \( 1 + \theta \) coefficient), of the order 1.03–1.14. Ogata et al. (2003, Tables 1-2) estimates the \( p \)-values to vary within 1.05–
1.18. Ogata & Zhuang’s (2006, Tables 2-3) values are 1.02–1.05. Zhuang et al. (2005) obtained \(p\)-values of 1.14–1.15. Helmstetter et al. (2006, Table 1) used a branching model similar to the ETAS and calculated \(p = 1.18 - 1.20\).

To illustrate our fit of the temporal distribution of dependent earthquakes, average numbers of aftershocks following 15 \(m \geq 8.0\) GCMT earthquakes are displayed in Fig. 12 (similar to Fig. 13 in Kagan, 2004). We use the time period 1977-2003, so that all large earthquakes are approximately the same size (8.45 \(\geq m \geq 8.0\), i.e., excluding the 2004 Sumatra and its aftershocks). Since the GCMT catalogue has relatively few dependent events, we selected aftershocks from the the U.S. Geological Survey (2010) PDE (Preliminary Determinations of Epicenters) catalogue. The aftershock rate in the diagram is approximately constant above our estimate of the coda duration \(t_M\). For the logarithmic intervals, this corresponds to the standard form of the Omori law: the aftershock number \(n_a\) is proportional to \(1/t\). For the smaller time intervals, the aftershock numbers decline when compared to the Omori law prediction \((1/t)\). This decline is caused by several factors, the interference of mainshock coda waves being the most influential (Kagan, 2004). The decline is faster for weaker events \((ibid.)\).

For theoretical estimates, we used the results obtained during the likelihood function search (Kagan et al., 2010, Table 4) for the full PDE catalogue, \(m_t = 5.0\). The parameters values are: the branching coefficient \(\mu = 0.141\), the parent productivity exponent \(a_0 = 0.63\) \((a_0 = \delta \times 1.5\), see Eq.18\), and the time decay exponent \(\theta = 0.28\), i.e., in the middle of the scatterplot shown in Fig. 11. Theoretical estimates in Fig. 12 seem to be reasonably good at forecasting time intervals on the order of one day. For larger intervals, the expected aftershock numbers decrease as \(n_a \sim (\Delta t)^{-1.15}\): this is stronger than the regular Omori law would predict. As we suggested in Section 2.1 the Omori law assumes that all aftershocks are direct consequences of a mainshock, whereas a branching model regards any earthquake
as a possible progenitor of later events. Thus, later aftershocks are the combined offspring of a mainshock and all consequent earthquakes. With increase in time, the difference between Omori’s law and the CBM predictions would increase as well. Marsan & Lengliné (2008) show that in California catalogues due to cascading aftershock rates for direct and secondary triggering differ by a factor of 10 to about 50%. By numerical simulations of the ETAS model, Felzer et al., (2002) and Helmstetter & Sornette (2003) estimated that a substantial majority of aftershocks are indirectly triggered by the mainshock.

However, the simulations based on the ETAS model are not without problems. Sornette & Werner (2005, see their Fig. 2) argue that since events below the magnitude threshold are not included in the earthquake catalogues and in many simulations, the earthquake stochastic interrelation pattern is not fully reproduced. An earthquake which seems to be an independent event may be an offspring of a quake that is below threshold, similarly an apparent first generation aftershock may also depend on a preceding unobserved smaller event. All earthquake connections above the magnitude threshold are accounted for in the models based on branching along the magnitude/moment axis (Kagan, 2010, Fig. 1). Therefore, in Fig. 12 we simulate the aftershock distribution by the programs based on the branching along the magnitude axis (see, for example, Kagan & Knopoff, 1978).

The $p$-parameter in Omori’s law is often assumed to be 1.0. This $p$-value implies that the total aftershock number approaches infinity as the duration of the aftershock sequence increases. Figs. 3 and 4, however, suggest that in the presence of tectonic loading, the time exponent value should increase at longer time periods. Statistical determination of the exponent can usually be made for only short periods; thus, we do not yet have a good estimate of the $p$-value for the tail of an aftershock sequence.
4 Earthquake statistics and plate-tectonic deformation

4.1 Stationary earthquake rate due to tectonic deformation

The tapered Gutenberg-Richter (TGR) relation (Kagan, 2002b) has an exponential taper applied to the number of events with a large seismic moment. Its survivor function (1 − cumulative distribution) for the scalar seismic moment $M$ is

$$F(M) = (M_t/M)^\beta \exp \left( \frac{M_t - M}{M_c} \right) \quad \text{for} \quad M_t \leq M < \infty. \quad (19)$$

Here $M_c$ is the parameter controlling the distribution in the upper ranges of $M$ (‘the corner moment’), $M_t$ is the moment threshold (see Eq. 18); $\beta$ is the index parameter of the distribution; $\beta = \frac{2}{3}b$, $b$ is a familiar $b$-value of the Gutenberg-Richter distribution (G-R, Gutenberg and Richter, 1944).

By evaluating the first moment of the distribution (19), we can obtain a theoretical estimate of the seismic moment flux (Kagan, 2002b)

$$\dot{M}_s = \frac{\alpha_0 M_0}{1 - \beta} M_c^{1-\beta} \Gamma(2 - \beta) \exp(M_0/M_c), \quad (20)$$

where $\Gamma$ is a gamma function and $\alpha_0$ is the seismic activity level (occurrence rate) for earthquakes with moment $M_0$ and greater.

We compare the seismic moment rate with the tectonic moment rate ($\dot{M}_T$):

$$\dot{M}_T = \mu W \int_A |\dot{\epsilon}| \, dA = \dot{M}_s / \chi, \quad (21)$$

where $\chi$ is the seismic coupling (or seismic efficiency) coefficient, $\mu$ is the elastic shear modulus, $W$ is the seismogenic width of the lithosphere, $\dot{\epsilon}$ is the average long-term horizontal strain rate, and $A$ is the area under consideration. At present, some variables in the equation cannot be evaluated with great accuracy; to overcome this difficulty we calculate a product of
these variables: the ‘effective width’ of seismogenic zone $W_e$ or coupled seismogenic thickness (Bird & Kagan, 2004):

$$W_e = W \chi .$$ (22)

Bird & Kagan (2004, Eq. 11) propose another, more exact formula for calculating the tectonic moment rate appropriate to a plate boundary zone.

In regions of high seismicity, instead of Eqs. 20–22 we can use measured long-term seismic activity to infer the earthquake rate by extrapolating (19) to any moment level (see Bird & Liu, 2007, Eqs. 4 and 5). Bird et al. (2010) presented an algorithm and tables for a long-term world-wide forecast of shallow seismicity based on the Global Strain Rate Map (GSRM) by Kreemer et al. (2003). Because GSRM does not estimate tectonic strain-rates of stable plate interiors, a simple empirical-averaging method has been used. Thus, the seismicity in plate interiors is represented by a uniform rate.

Since the seismicity level in plate interiors may vary by orders of magnitude, the uniform rate may strongly under- or over-estimate the seismicity rate. Therefore, we apply Eqs. 20–22 to evaluate first the tectonic moment rate and then a long-term forecast for these regions

$$\alpha_0 = \frac{\dot{M}_T \chi (1 - \beta) \exp (-M_0/M_c)}{M_0^\beta M_c^{1-\beta} \Gamma(2 - \beta)} .$$ (23)

By calculating $\alpha_0$ for a particular choice of $M_0$, we may re-normalize Eq. 19 and obtain earthquake size distribution for any region with a known strain rate and corner moment.

4.2 Length of aftershock zone

Eq. 23 above can be used to evaluate the background seismicity level for an aftershock zone. Such a calculation would use the area of an earthquake focal zone. This area can be estimated by a dimension of aftershock zone for each event. Kagan (2002a) evaluated how the aftershock zone size for mainshocks $m \geq 7.0$ depends upon on the earthquake magnitude by
approximating aftershock epicentre maps through a two-dimensional Gaussian distribution. The major ellipse axis is taken as a quantitative measure of the mainshock focal zone size.

In Fig. 13 we display the regression curves for GCMT/PDE earthquakes: all earthquakes for three choices of focal mechanisms (updated Fig. 6a by Kagan, 2002a). In regression curves we use $m = 8.25$ as a reference point. For example, for the quadratic regression

$$L = \log_{10} \ell = a_0 + a_1(m - 8.25) + a_2(m - 8.25)^2,$$

(24)

where $\ell$ is the length of the aftershock zone in km. For the linear regression we set $a_2 = 0$. Fig. 14 displays the regression in a similar format for active continental tectonic zones (Kagan et al., 2010).

Table 2 summarizes the results of regression analysis for all global earthquakes, as well as for events in subduction zones (trenches) and on active continental zones. Earthquakes are also subdivided by their focal mechanism. Other tectonic zones lack a sufficient number of $m \geq 7.0$ mainshocks to carry out this statistical analysis.

The following conclusions can be made from Table 2: (a) aftershock zones exhibit similar scaling; (b) zone length ($\ell$) on average is proportional to moment $M^{1/3}$; and (c) the value of $a_0$ parameter (zone length for the $m8.25$ earthquake) is close to $10^{2.5} (316)$ km for all cases. Normal earthquakes (rows 5-6) exhibit slightly different scaling; zone length ($\ell$) for the linear regression is proportional to moment $M^{1/2.8}$. Scaling for strike-slip earthquakes (rows 7-8) also differs a little from average: zone length ($\ell$) is proportional to moment $M^{1/3.5}$. However, the earthquake numbers in these subsets are small, thus it is possible that these variations are due to random fluctuations.

Only three subsets show a substantial nonlinearity: (a) trenches with strike-slip focal mechanisms (rows 15-16); (b) continents with all focal mechanisms (rows 17-18); and (c) continents with strike-slip focal mechanisms (rows 21-22). However, the earthquake numbers
are small in all these plots, and although (b) and (c) aftershocks display zone lengths which increase strongly for the largest earthquakes (the feature often quoted in other studies of length scaling – see, for instance, Wells, & Coppersmith, 1994, and subsequent publications citing that paper), (a) earthquakes exhibit an opposite behaviour.

In all diagrams the standard errors ($\sigma$) are almost the same for the linear and quadratic regression. The maximum errors ($\epsilon_{\text{max}}$) follow the same pattern. This pattern suggest that linear regression is sufficient to approximate the data. Although the quadratic regression fit yields no statistically significant improvement in almost any diagram, the sign of the quadratic correction term is negative for most cases. The negative value of the $a_2$ regression coefficient means that increase in the aftershock zone length is weaker for the largest earthquakes. This feature contradicts those often quoted in other studies of length scaling (see Kagan, 2002a for details). Thus, the slope of the regression curve is either stable or decreases at the high magnitude end. No saturation effect for large earthquakes occurs in the data. Results in Table 2 imply that the major ellipse axis $a$ (length) of an earthquake focal zone can be approximated by

$$a = 316 \times 10^{(m-8.25)/2} \text{ km}.$$  \hspace{1cm} (25)

We conclude that earthquake rupture length is proportional to the cube root of moment, which implies that width and slip should scale the same way. Otherwise, one of them increases less strongly with moment and the other more strongly. For either that would pose the problem of “inverse saturation.”

We assume that the majority of aftershocks are concentrated within an ellipse having 2-$a$ major axis. The probability that a point lies inside a 2-$a$ ellipse is shown in Eq. 5 by Kagan (2002a). If we know the length of an earthquake focal zone, we can calculate its area. We assume, for example, that the ratio of the major ellipse axis to the minor axis is 1/4; then
area $S$ of the focal zone is

$$S = \pi a^2.$$  \hfill (26)

### 4.3 Example: New Madrid earthquake sequence of 1811-12

To illustrate the arguments and results of the previous sections, we calculate seismicity parameters of the New Madrid earthquake sequence (1811-12) and its consequences. There is substantial literature on this sequence (Hough, 2004; Stein & Liu, 2009; Calais et al., 2010, and references therein).

Three or four large earthquakes with magnitudes on the order of 7.3–7.8 occurred over a few months of 1811-12 in the New Madrid area; aftershocks of these events are still registered. As an illustration, we would assume that only one $m8$ event occurred at that time. If in reality earthquakes were smaller than such an event, their total focal zone and combined aftershock sequence at the present time would be equivalent to about one $m8$ mainshock.

The size of the focal zone can be evaluated by using regression equations in Figs. 13 and 14. The first plot contains many earthquakes but most of these events are in subduction zones. The second diagram uses earthquakes in active continental zones; a focal size of these earthquakes is likely to resemble the New Madrid area which can be classified as plate-interior (Kagan et al., 2010). Too few large earthquakes are available within plate-interior to obtain their features. The difference between regression parameters in Figs. 13 and 14 is small; therefore the size of earthquake focal zones either does not change in various tectonic zones or changes slightly.

For an $m8$ earthquake, calculations yield 227 km and 259 km as the length of the focal zone, defined as the $4\sigma$ major axis of an ellipse comprising a majority of aftershocks (Kagan, 2002a). The linear regression is used in both cases: the former value corresponds to Fig. 13 and the latter to Fig. 14. The two estimates are similar and roughly correspond to the
size of the present aftershock zone, as shown, for example, in Calais et al. (2010). For the Cottonwood fault that is about 110-120 km length (ibid., Fig. 1), the magnitude estimate from Fig. 14 is \( m \sim 7.2 - 7.4 \).

To calculate the surface area of an aftershock zone, we assume that the minor axis of the ellipse is 1/4 of the major axis, taken as 240 km. Then we obtain the \( m8 \) earthquake focal area as 11300 km\(^2\). Taking the stationary strain rate as \( \dot{\epsilon} = 10^{-9} \) yr\(^{-1} \) (Calais et al., 2006), we compute the tectonic moment rate (21): \( 3.4 \times 10^{15} \) Nm/year. Assuming that 50% of the tectonic rate is released seismically (Bird & Kagan, 2004), we obtain the background rate \( \alpha_0 = 1.27 \times 10^{-3} \) m \( \geq \) 5 earthquakes per year [we take in (23) \( M_c = 10^{21} \) Nm, \( W = 20 \) km, \( \beta = 2/3 \); then \( \Gamma(4/3) = 0.893 \)]. We use (19) to calculate the recurrence time for an \( m \geq 8 \) earthquake in the focal zone of the New Madrid events: more than two million years. In this computation any \( m8 \) earthquake with an epicentre or centroid in the focal zone counts: in Eq. (23) we do not request that the entire rupture of such an earthquake be contained in the zone. The recurrence time is an average value; even for events as large as \( m \geq 8 \) the earthquake occurrence is clustered (see Section 3.2). Thus, a new large earthquake can follow after a relatively short time period, as exemplified by the 1811-12 New Madrid sequence.

A similar calculation of the \( m8 \) earthquake rate could be carried out by using the results of Bird et al. (2010, p. 188). They calculate an estimate of the mean intraplate seismicity rate of \( 4.27 \times 10^{-22} \) m\(^{-2} \) s\(^{-1} \) for \( m_0 = 5.66 \) earthquakes. Bird et al. use a slightly different formula for the conversion of the scalar moment to magnitude. They use 9.05 instead of our 9.0 coefficient in (1). Thus their estimate would yield the rate for the New Madrid area \( \alpha_0 = 7.5 \times 10^{-4} \) m \( \geq \) 5 earthquakes per year, the value of the same order of magnitude as computed above.

We would like to calculate the duration of an aftershock sequence up until the aftershock rate decays to the background level. The results in Fig. 12 can be applied to this purpose.
However, we need to make a correction for the mainshock and aftershock magnitudes ($m_8$ instead of $m_8.15$ and $m \geq 5.0$ instead of $m_b \geq 4.9$ in the plot, respectively). In the diagram the aftershock rate per one interval (the intervals increase consequently by a factor of two) is 7 events. This translates into $4.55 \cdot 5$ events per interval for our choice of magnitudes (Kagan et al., 2010). After comparing the background and aftershock rates (we take $4.55 \cdot 5$ aftershocks per the first day, decaying according to Omori’s law, with $1/t$ rate with time), we discover that the aftershock sequence would reach the background rate in about 3,600 years. This duration estimate agree roughly with Stein & Liu’s (2009) value. In these calculations, we presume that no independent large earthquake clusters would occur during the aftershock sequence. The possible occurrence of spontaneous events makes any evaluation of aftershock sequence duration largely approximate.

Stein & Liu (2009) obtained aftershock duration values for several sequences using Eq. 14 from Dieterich (1994). This equation employs parameters whose values for actual earthquake focal zones are not known. Generally, the parameters have been back adjusted based on the statistics of earthquake occurrence. This may explain apparently reasonable fit of Dieterich’s formula to aftershock sequences.

In contrast, we obtain the aftershock sequence duration by extending the tapered Gutenberg-Richter and Omori’s laws and using their well-known properties and measured geometrical features of tectonic deformation. Moreover, according to Stein & Liu (2009, Fig. 1c) the New Madrid aftershock rate for the last 50 years was about $0.5 \cdot 4$ events per year. Computations based on Omori’s law similar to those shown above, yield the rate 175 years after the mainshock occurrence of about $0.26 \cdot 4$ events per year. This number is close to that shown above.
5 Discussion

Two classical statistical earthquake distributions largely governed our approach to analyzing seismicity: Omori’s law and the Gutenberg-Richter relation. As explained above, recent developments in earthquake size statistics considerably improved our understanding of earthquake occurrence and could lead to significantly better estimates of seismic hazard. For the G-R law the earthquake temporal distribution is mostly irrelevant, since size distribution of clustered events is largely independent of their history. Similar progress in understanding earthquake time statistics is much more difficult to achieve. We cannot ignore spatial variables and the available data are not as extensive, so the task is more complex. Only by applying rigorous methods, by analyzing carefully systematic and random effects, and by critical testing of models and hypotheses we will be able to advance in solving this problem.

In previous sections we derived the time distribution for earthquake occurrence; the aftershock distribution is shown to be controlled by power-laws. How can the parameters of these distributions be determined?

If one excludes the interiors of plates, the tectonic deformation rate $V$ is reasonably well known for plate boundaries and for active continents (Kagan et al., 2010; Bird et al., 2010). The diffusion rate $D$ is presently unknown. If we could obtain the earthquake temporal distribution as shown in Figs. 1 and 2, the $D$ evaluation would be easy. These distributions are derived for a particular area within an earthquake fault zone. However, if the stress in the focal zone of an earthquake is distributed according to the power-law with an exponent $\psi$ (see Eqs. 3 and 5), the problem becomes more acute.

Figs. 3 and 4 suggest that the distribution temporal behaviour changes when $V = \sqrt{D}$. This change relates to the first generation of offspring. Thus, we should not be able to see it in regular Omori plots which combine many generations of aftershocks. The inversion of earthquake occurrence parameters based on stochastic branching processes yields needed
first generation effects. However, in present models (both CBM and ETAS) as discussed in Section 3.3, temporal dependence is parameterized by just one exponent. These models should in principle demonstrate changes in the temporal pattern, if more complicated temporal function is applied. However, results from statistical analysis are very uncertain even for one-parameter time decay. Given the contemporary quantity and quality of earthquake catalogues, it is unlikely more complicated models would be effective in resolving this issue.

Perhaps new laboratory experiments (Zaiser, 2006) may help solve the problem of diffusion rate evaluation, but it is not clear whether such measurements are possible. The acoustic emission event rate exhibits fore- and aftershock sequences associated with dynamic failure of the test specimen (see, for example, Ojala et al., 2004). These and similar tests can be used to infer the dependence of the Omori law parameters on spatial scale and stress diffusion rate.

Results from statistical analysis of earthquake occurrence in our previous publications (Kagan, 2002b; Bird & Kagan, 2004; Kagan et al., 2010), as well as the results reported above, suggest that the earthquake process in all tectonic provinces can be described by the same model. We advocate the Poisson cluster process with clusters controlled by a critical branching process and a power-law time dependence. Combined with the earthquake size distribution approximation by the tapered Gutenberg-Richter (TGR) law, such a model allows a quantitative forecast of a spatially variable, time-independent (long-term) earthquake rate. It will optimally smooth the seismicity record (Kagan & Jackson, 2011) or translate the plate-tectonic and geodetic rate into a seismic hazard estimate (Bird & Liu, 2007; Bird et al., 2010).

A short-term forecast can be performed by using the temporal properties of earthquake clusters, an extrapolation which uses a variant of Omori’s law to estimate future earthquake rate (Kagan & Jackson, 2011). In Section 4.3 we presented an example of such calculations.
6 Conclusion

• 1. Two statistical temporal distributions – the Lévy and Inverse Gaussian – are used to explain Omori’s law of the aftershock rate decay.

• 2. Since these distributions have an unrealistic feature of decaying to zero at small time intervals, we consider their modification, by mixing them with a power-law stress distribution in the earthquake focal zone. The resulting distributions have a power-law behaviour for the duration of the aftershock sequence.

• 3. For longer time intervals earthquake occurrence is dominated by independent events which follow the Poisson law. The statistical distribution of interevent time for large earthquakes in various tectonic zones is analyzed to demonstrate this feature and infer the properties of the earthquake occurrence. As another method of a quantitative statistical analysis we apply stochastic branching process and show the results for the approximation of the temporal behaviour for earthquake occurrence.

• 4. We discuss how the Poisson rate of background seismicity can be calculated knowing the average tectonic strain rate and maximum (corner) magnitude. The latter variable can be estimated by using our earlier results.

• 5. In order to use our results for earthquake hazard calculation the size of earthquake rupture needs to be known. We estimate the length of the aftershock zone for various focal mechanisms and different tectonic environments, demonstrating that the length is either independent or only slightly dependent on these features.

• 6. To demonstrate the application of our results to earthquake hazard computations, we calculate the background rates for the New Madrid region, as well as the aftershock rates for the sequence of 1811-12 events.

• 7. In conclusion, this paper suggests a method for calculating long- and short-term
seismicity estimates, based on a theoretical inference about classical, statistical earthquake distributions: the Omori law and the G-R relation. A statistical analysis of the earthquake occurrence carried out in our previous papers and in this work make such a seismic hazard evaluation more reliable and accurate.

Acknowledgments

I am grateful to Dave Jackson, Paul Davis, and Peter Bird for useful discussion and suggestions. Per Jögi helped with the Mathematica computations. Comments by three anonymous reviewers and by the Associate Editor Massimo Cocco have been very helpful in revising the manuscript. I thank Kathleen Jackson for significant improvements in the text. The author appreciates support from the National Science Foundation through grants EAR-0711515 and EAR-0944218, as well as from the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. Publication 1493, SCEC.


34


Marsan, D., 2005. The role of small earthquakes in redistributing crustal elastic stress,


Ogata, Y., & J. C. Zhuang, 2006. Space-time ETAS models and an improved extension,


dence, R.I., pp. 284; Russian original 1983.
Table 1: Pairs of shallow earthquakes \( m \geq 7.5 \)

<table>
<thead>
<tr>
<th>No</th>
<th>First Event</th>
<th></th>
<th>Second Event</th>
<th></th>
<th>Difference</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Date</td>
<td>Coord.</td>
<td>Date</td>
<td>Coord.</td>
<td></td>
<td>( R )</td>
<td>( \Phi )</td>
<td>( \Delta t )</td>
</tr>
<tr>
<td></td>
<td>Lat.</td>
<td>Long.</td>
<td>Lat.</td>
<td>Long.</td>
<td>km</td>
<td>°</td>
<td>day</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1977/06/22</td>
<td>-22.9</td>
<td>-174.9</td>
<td>2009/03/19</td>
<td>-23.1</td>
<td>75</td>
<td>55</td>
<td>11593.26</td>
</tr>
<tr>
<td>2</td>
<td>1978/03/23</td>
<td>44.1</td>
<td>149.3</td>
<td>1978/03/24</td>
<td>44.2</td>
<td>7</td>
<td>25</td>
<td>1.69</td>
</tr>
<tr>
<td>3</td>
<td>1980/07/08</td>
<td>-12.9</td>
<td>166.2</td>
<td>1980/07/17</td>
<td>-12.4</td>
<td>62</td>
<td>18</td>
<td>8.85</td>
</tr>
<tr>
<td>4</td>
<td>1980/07/08</td>
<td>-12.9</td>
<td>166.2</td>
<td>1997/04/21</td>
<td>-13.2</td>
<td>33</td>
<td>42</td>
<td>6130.53</td>
</tr>
<tr>
<td>5</td>
<td>1980/07/08</td>
<td>-12.9</td>
<td>166.2</td>
<td>2009/10/07</td>
<td>-12.6</td>
<td>37</td>
<td>13</td>
<td>10682.95</td>
</tr>
<tr>
<td>6</td>
<td>1980/07/17</td>
<td>-12.4</td>
<td>165.9</td>
<td>2009/10/07</td>
<td>-12.6</td>
<td>41</td>
<td>14</td>
<td>10674.10</td>
</tr>
<tr>
<td>7</td>
<td>1980/07/17</td>
<td>-12.4</td>
<td>165.9</td>
<td>2009/10/07</td>
<td>-11.9</td>
<td>65</td>
<td>12</td>
<td>10674.11</td>
</tr>
<tr>
<td>8</td>
<td>1983/03/18</td>
<td>-4.9</td>
<td>153.3</td>
<td>2000/11/16</td>
<td>-4.6</td>
<td>83</td>
<td>72</td>
<td>6452.83</td>
</tr>
<tr>
<td>9</td>
<td>1983/03/18</td>
<td>-4.9</td>
<td>153.3</td>
<td>2000/11/16</td>
<td>-5.0</td>
<td>47</td>
<td>91</td>
<td>6452.94</td>
</tr>
<tr>
<td>10</td>
<td>1985/09/19</td>
<td>17.9</td>
<td>-102.0</td>
<td>1985/09/21</td>
<td>17.6</td>
<td>71</td>
<td>14</td>
<td>1.51</td>
</tr>
<tr>
<td>11</td>
<td>1987/03/05</td>
<td>-24.4</td>
<td>-70.9</td>
<td>1995/07/30</td>
<td>-24.2</td>
<td>33</td>
<td>7</td>
<td>3068.83</td>
</tr>
<tr>
<td>12</td>
<td>1990/04/18</td>
<td>1.3</td>
<td>123.3</td>
<td>1991/06/20</td>
<td>1.0</td>
<td>37</td>
<td>29</td>
<td>427.65</td>
</tr>
<tr>
<td>13</td>
<td>1995/08/16</td>
<td>-5.5</td>
<td>153.6</td>
<td>2000/11/16</td>
<td>-5.0</td>
<td>76</td>
<td>74</td>
<td>1918.89</td>
</tr>
<tr>
<td>14</td>
<td>1997/04/21</td>
<td>-13.2</td>
<td>166.2</td>
<td>2009/10/07</td>
<td>-12.6</td>
<td>70</td>
<td>30</td>
<td>4552.42</td>
</tr>
<tr>
<td>15</td>
<td>2000/06/04</td>
<td>-4.7</td>
<td>101.9</td>
<td>2007/09/12</td>
<td>-3.8</td>
<td>150</td>
<td>85</td>
<td>2655.78</td>
</tr>
<tr>
<td>16</td>
<td>2000/11/16</td>
<td>-4.6</td>
<td>152.8</td>
<td>2000/11/16</td>
<td>-5.0</td>
<td>67</td>
<td>89</td>
<td>0.12</td>
</tr>
<tr>
<td>17</td>
<td>2000/11/16</td>
<td>-4.6</td>
<td>152.8</td>
<td>2000/11/17</td>
<td>-5.3</td>
<td>93</td>
<td>88</td>
<td>1.67</td>
</tr>
<tr>
<td>18</td>
<td>2001/06/23</td>
<td>-17.3</td>
<td>-72.7</td>
<td>2001/07/07</td>
<td>-17.5</td>
<td>34</td>
<td>8</td>
<td>13.54</td>
</tr>
<tr>
<td>19</td>
<td>2005/03/28</td>
<td>1.7</td>
<td>97.1</td>
<td>2010/04/06</td>
<td>2.0</td>
<td>58</td>
<td>7</td>
<td>1835.25</td>
</tr>
<tr>
<td>20</td>
<td>2006/11/15</td>
<td>46.7</td>
<td>154.3</td>
<td>2007/01/13</td>
<td>46.2</td>
<td>70</td>
<td>82</td>
<td>58.71</td>
</tr>
<tr>
<td>21</td>
<td>2007/09/12</td>
<td>-3.8</td>
<td>101.0</td>
<td>2007/09/12</td>
<td>-2.5</td>
<td>176</td>
<td>11</td>
<td>0.53</td>
</tr>
<tr>
<td>22</td>
<td>2007/09/12</td>
<td>-3.8</td>
<td>101.0</td>
<td>2010/10/25</td>
<td>-3.7</td>
<td>189</td>
<td>8</td>
<td>1139.15</td>
</tr>
</tbody>
</table>

\( R \) – centroid distance, \( \Phi \) – 3-D rotation angle between focal mechanisms, \( \Delta t \) – time interval between events, \( \eta \) – degree of zone overlap, the ratio of earthquake focal zone sizes to twice their distance, see Equations (2,3) in Kagan & Jackson (1999). The total earthquake number with magnitude \( m \geq 7.50 \) for the period 1976/1/1–2010/10/25 is 121. The maximum epicentroid distance is 250.00 km.
Table 2: Aftershock zone log length vs mainshock moment magnitude $m$

<table>
<thead>
<tr>
<th>#</th>
<th>Tectonic zone</th>
<th>Focal mech.</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$\sigma$</th>
<th>$\epsilon_{\text{max}}$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>All</td>
<td>All</td>
<td>2.48</td>
<td>0.492</td>
<td>–</td>
<td>0.134</td>
<td>0.468</td>
<td>160</td>
</tr>
<tr>
<td>2</td>
<td>All</td>
<td>All</td>
<td>2.48</td>
<td>0.493</td>
<td>0.0013</td>
<td>0.134</td>
<td>0.468</td>
<td>160</td>
</tr>
<tr>
<td>3</td>
<td>All</td>
<td>Thrust</td>
<td>2.48</td>
<td>0.501</td>
<td>–</td>
<td>0.132</td>
<td>0.457</td>
<td>115</td>
</tr>
<tr>
<td>4</td>
<td>All</td>
<td>Thrust</td>
<td>2.48</td>
<td>0.499</td>
<td>0.0022</td>
<td>0.132</td>
<td>0.458</td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>All</td>
<td>Normal</td>
<td>2.47</td>
<td>0.532</td>
<td>–</td>
<td>0.076</td>
<td>0.132</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>All</td>
<td>Normal</td>
<td>2.46</td>
<td>0.427</td>
<td>–0.0884</td>
<td>0.075</td>
<td>0.137</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>All</td>
<td>Str.-Slip</td>
<td>2.47</td>
<td>0.437</td>
<td>–</td>
<td>0.153</td>
<td>0.276</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>All</td>
<td>Str.-Slip</td>
<td>2.49</td>
<td>0.490</td>
<td>0.0038</td>
<td>0.153</td>
<td>0.278</td>
<td>30</td>
</tr>
<tr>
<td>9</td>
<td>Trench</td>
<td>All</td>
<td>2.47</td>
<td>0.499</td>
<td>–</td>
<td>0.131</td>
<td>0.454</td>
<td>129</td>
</tr>
<tr>
<td>10</td>
<td>Trench</td>
<td>All</td>
<td>2.47</td>
<td>0.482</td>
<td>–0.0177</td>
<td>0.131</td>
<td>0.449</td>
<td>129</td>
</tr>
<tr>
<td>11</td>
<td>Trench</td>
<td>Thrust</td>
<td>2.48</td>
<td>0.500</td>
<td>–</td>
<td>0.135</td>
<td>0.460</td>
<td>104</td>
</tr>
<tr>
<td>12</td>
<td>Trench</td>
<td>Thrust</td>
<td>2.48</td>
<td>0.488</td>
<td>–0.0142</td>
<td>0.135</td>
<td>0.455</td>
<td>104</td>
</tr>
<tr>
<td>13</td>
<td>Trench</td>
<td>Normal</td>
<td>2.47</td>
<td>0.543</td>
<td>–</td>
<td>0.067</td>
<td>0.120</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>Trench</td>
<td>Normal</td>
<td>2.46</td>
<td>0.440</td>
<td>–0.0915</td>
<td>0.065</td>
<td>0.114</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>Trench</td>
<td>Str.-Slip</td>
<td>2.38</td>
<td>0.409</td>
<td>–</td>
<td>0.146</td>
<td>0.302</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>Trench</td>
<td>Str.-Slip</td>
<td>2.29</td>
<td>0.034</td>
<td>–0.2790</td>
<td>0.142</td>
<td>0.268</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>Active Cont.</td>
<td>All</td>
<td>2.52</td>
<td>0.504</td>
<td>–</td>
<td>0.138</td>
<td>0.256</td>
<td>27</td>
</tr>
<tr>
<td>18</td>
<td>Active Cont.</td>
<td>All</td>
<td>2.74</td>
<td>1.180</td>
<td>0.4250</td>
<td>0.135</td>
<td>0.220</td>
<td>27</td>
</tr>
<tr>
<td>19</td>
<td>Active Cont.</td>
<td>Thrust</td>
<td>2.44</td>
<td>0.480</td>
<td>–</td>
<td>0.112</td>
<td>0.176</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>Active Cont.</td>
<td>Thrust</td>
<td>3.00</td>
<td>2.210</td>
<td>1.0800</td>
<td>0.064</td>
<td>0.130</td>
<td>10</td>
</tr>
<tr>
<td>21</td>
<td>Active Cont.</td>
<td>Str.-Slip</td>
<td>2.50</td>
<td>0.419</td>
<td>–</td>
<td>0.145</td>
<td>0.314</td>
<td>14</td>
</tr>
<tr>
<td>22</td>
<td>Active Cont.</td>
<td>Str.-Slip</td>
<td>2.65</td>
<td>0.889</td>
<td>0.2990</td>
<td>0.143</td>
<td>0.295</td>
<td>14</td>
</tr>
</tbody>
</table>

$a_0$, $a_1$, $a_2$ are regression coefficients in Eq. 24; $\sigma$ is the standard uncertainty; $\epsilon_{\text{max}}$ is the maximum error, i.e., the largest difference between the data and the fit shown in Figs. 13,14; $n$ is the number of aftershock sequences. The catalogues time interval is 1977/1/1–2010/09/21.
Figure 1:

Plot of PDFs for the Lévy distribution (2), $D = 1.0$; From left to right $\sigma = 0.001$ (yellow), 0.01 (green), 0.1 (cyan), 1.0 (red), 10.0 (blue), 100.0 (magenta).
Figure 2:

Plot of PDFs for the IGD distribution (4), $V = \sqrt{D}$; $D = V^2$; From left to right $\sigma = 0.001$ (yellow), 0.01 (green), 0.1 (cyan), 1.0 (red), 10.0 (blue), 100.0 (magenta), 1000.0 (black).
Figure 3:

Plot of PDFs for the IGD distribution (Eqs. 5 – 9).

Curve with squares (7): $V = \sqrt{D}$, $\psi = 0.0$;
curve with diamonds (8): $V = \sqrt{D}$, $\psi = 0.5$;
curve with crosses (9): $V = -\sqrt{D}$, $\psi = 0.5$;
curve with circles (5): $V = \sqrt{D}$, $\psi = 0.9$. 
Plot of PDFs for the IGD distribution (Eqs. 5 – 9), multiplied by $t^{1+\psi/2}$, $D = 1.0$.

Curve with squares (7): $V = \sqrt{D}$, $\psi = 0.0$;
curve with diamonds (8): $V = \sqrt{D}$, $\psi = 0.5$;
curve with crosses (9): $V = -\sqrt{D}$, $\psi = 0.5$;
curve with circles (5): $V = \sqrt{D}$, $\psi = 0.9$. 
Figure 5:

Normalized cumulative distribution of earthquake rate for $m \geq 6.5$ shallow earthquake pairs in all zones, trench (subduction) zones, and active continental zones. The pair numbers $N$, average rate $\langle \lambda \rangle$ and its standard deviation $\sigma_\lambda$ are also shown. Earthquake rate is taken from a table by Bird et al. (2010) for a magnitude threshold $m_t = 5.66$. 

48
Integration domain for calculating inter-earthquake rates in a catalogue of duration $T$. The minimum time interval, $t_0$, corresponds to the coda duration time of the first event.
Figure 7:

Earthquake pair numbers in all zones with rates $0 - 26 \times 10^{-20}$ events/m$^2$/s. Earthquake rate is taken from a table by Bird et al. (2010) for a magnitude threshold $m_t = 5.66$. Pair number is 240. Solid curve is earthquake pair numbers; dashed curve is Poisson approximation (15), dash-dotted curves are for power-law approximations (16), the $\theta$-value is 0.5, 0.65, 0.75, 0.85, 0.925, and 0.99 from top to bottom. The catalogues time interval is 1976/1/1–2010/11/14. Average recurrence time ($\bar{t}$) for earthquakes is $\bar{t} = 3272 \pm 3596$ days.
Earthquake pair numbers in all zones with rates $26 - 45 \times 10^{-20}$ events/m$^2$/s. Pair number is 333. Average recurrence time ($\bar{t}$) for earthquakes is $\bar{t} = 3507 \pm 3242$ days. For notation see Fig. 7.
Figure 9:

Earthquake pair numbers in all zones with rates \( \geq 45 \times 10^{-20} \) events/m\(^2\)/s. Pair number is 264. Average recurrence time \( \bar{t} \) for earthquakes is \( \bar{t} = 3947 \pm 3319 \) days. For notation see Fig. 7.
Figure 10:

Earthquake pair numbers in active continental zones. Pair number is 37. Average recurrence time ($\bar{t}$) for earthquakes is $\bar{t} = 1304 \pm 2578$ days. For notation see Fig. 7.
Figure 11:

Correlation between maximum likelihood estimates of $\mu$ and $\theta$ parameters in several earthquake catalogues (Kagan et al., 2010).
Figure 12:

The average aftershock numbers $n_a$ from the PDE catalogue in logarithmic time intervals following $m \geq 8.0$ GCMT earthquakes during 1977-2003. Line with plus signs shows $m_b \geq 4.9$ aftershocks; dashed lines are theoretical estimates for the first generation aftershocks. Two curves display results of earthquake catalogue simulation: line with circles shows first generation aftershocks, line with squares indicates total aftershock numbers.
Plot of log aftershock zone length ($L$) against moment magnitude ($m$). Rupture length is determined using a 1-day aftershock pattern. Values of the correlation coefficient ($\rho$), coefficients for linear (dashed line) and quadratic (solid line) regression, standard ($\sigma$) and maximum ($\epsilon_{\text{max}}$) errors, and the total number ($n$) of aftershock sequences are shown in the diagram. The both lines for the linear regression and quadratic approximation practically overlap in the plot.

Circles – thrust mainshocks;
Stars – normal mainshocks;
Pluses – strike-slip mainshocks.
Figure 14:

Plot of log aftershock zone length ($L$) against moment magnitude ($m$) for earthquakes in active continental zones (Kagan et al., 2010). For notation see Fig. 13.

$$L = 2.74 + 1.18 (m_w - 8.25) + 0.425 (m_w - 8.25)^2$$

$$\rho = 0.75, \ L = 2.52 + 0.504 (m_w - 8.25)$$

$$\sigma = 0.138, \ \varepsilon_{\text{max}} = 0.256, \ n = 27$$

$$\sigma = 0.135, \ \varepsilon_{\text{max}} = 0.22$$