Title
Why does theoretical physics fail to explain and predict earthquake occurrence?

Journal
Lecture Notes in Physics, 705

Author
Kagan, Yan Y

Publication Date
2006

DOI
DOI 10.1007/3-540-35375-5_12

Peer reviewed

PostScript file created: December 7, 2006; time 1078 minutes

Why does theoretical physics fail to explain and predict earthquake occurrence?

Yan Y. Kagan[1]

[1] Department of Earth and Space Sciences, University of California, Los Angeles, California, USA

Abstract.

Several reasons for the failure can be proposed:

1. The multidimensional character of seismicity: time, space, and earthquake focal mechanism need to be modeled. The latter is a symmetric second-rank tensor of a special kind.

2. The intrinsic randomness of earthquake occurrence, necessitating the use of stochastic point processes and appropriate complex statistical techniques.

3. The scale-invariant or fractal properties of earthquake processes; the theory of random stable or heavy-tailed variables is significantly more difficult than that of Gaussian variables and is only now being developed. Earthquake process theory should be capable of being renormalized.

4. Statistical distributions of earthquake sizes, earthquake temporal interactions, spatial patterns and focal mechanisms are largely universal. The values of major parameters are similar for earthquakes in various tectonic zones. The universality of these distributions will enable a better foundation for earthquake process theory.

5. The quality of current earthquake data statistical analysis is low. Since little or no study of random and systematic errors is performed, most published statistical results are artifacts.

6. During earthquake rupture, propagation focal mechanisms sometimes undergo large 3-D rotations. These rotations require non-commutative algebra (e.g., quaternions and gauge theory) for accurate models of earthquake occurrence.

7. These phenomenological and theoretical difficulties are not limited to earthquakes: any fracture of brittle materials, tensile or shear, would encounter similar problems.

Index terms:

Keywords: Earthquakes; seismicity; earthquake size, time, space and focal mechanisms; scale-invariance; stochastic point process; statistical analysis; quaternions; earthquake forecasting

0148-0227/06/$9.00

1. Introduction

The difficulties of seismic analysis are obvious. Earthquake processes are inherently multidimensional (Kagan, 1994b; Kagan and Vere-Jones, 1996): in addition to the origin time, 3-D locations, and measures of size for each earthquake, the orientation of the rupture surface and its displacement requires for its representation either second-rank tensors or quaternions (see more below). Earthquake occurrence is characterized by extreme randomness; the stochastic nature of seismicity is not reducible by more numerous or more accurate measurements. Even a cursory inspection of seismological datasets suggests that earthquake occurrence as well as earthquake fault geometry are scale-invariant or fractal (Mandelbrot, 1983; Kagan, 1994b; Kagan and Vere-Jones, 1996; Turcotte, 1997; Sornette, 2003; see also http://www.esi-topics.com/earthquakes/interviews/YanYKagan.html).

Adequate mathematical and statistical techniques have only recently become available for analyzing fractal temporal, spatial, and tensor patterns of point process data generally and earthquake data in particular. Such methods are still in the development stage. Moreover, it is only in the past 25-30 years that the quality, precision and completeness of earthquake datasets and the processing power of modern computers have become sufficient to allow detailed, full-scale investigation of earthquake occurrence patterns.

After looking at recent publications on earthquake physics (for example, Lee *et al.*, 2002; Scholz, 2002; Kanamori and Brodsky, 2004), one gets the impression that knowledge of earthquake process is still at a rudimentary level. Why has progress in understanding earthquakes been so slow? Kagan (1992a) compared the seismicity description to another problem in physics: turbulence of fluids. Both phenomena are characterized by multidimensionality and stochasticity. Their major statistical ingredients are scale-invariant, and both have hierarchically organized structures. Moreover, the scale of self-similar structures in seismicity and turbulence extends over many orders of magnitude. The size of major structures which control deformation patterns in turbulence and brittle fracture is comparable to the maximum size of the region (see more in Kagan, 1994b).

Yaglom (2001, p. 4) commented that turbulence status differs from many other complex problems which twentieth century physics has solved or has considered.

> '[These problems] deal with some very special and complicated objects and processes relating to some extreme conditions which are very far from realities of the ordinary life... However, turbulence theory deals with the most ordinary and simple realities of the everyday life such as, e.g., the jet of water spurting from the kitchen tap. Therefore, the turbulence is well-deservedly often called "the last great unsolved problem of the classical physics."'

Although solving the Navier-Stokes equations, describing turbulent motion in fluids is one of the seven mathematical millennium problems for the 21st century (see http://www.claymath.org/millennium/), the turbulence problem is not among the ten millennium problems in physics presented by the University of Michigan, Ann Arbor (see http://feynman.physics.lsa.umich.edu/-strings2000/millennium.html), or among the 11 problems by the National Research Council's board on physics and astronomy (Haseltine, 2002). In his extensive and wide-ranging review of current theoretical physics, Penrose (2005) does not include the turbulence or Navier-Stokes equations in the book index.

Like fluid turbulence, the brittle fracture of solids is commonly encountered in everyday life, but so far there is no real theory explaining its properties or predicting outcomes of the simplest occurrences, such as a glass breaking. Although computer simulations of brittle fracture (for example, see O'Brien and Hodgins, 1999) are becoming more realistic, they cannot yet provide a scientifically faithful representation. Brittle fracture is a more difficult scientific problem than turbulence, and while the latter attracted first-class mathematicians and physicists, no such interest has been shown in the mathematical theory of fracture and large-scale deformation of solids.

One sees the degree of difficulty in assessing the effects of brittle fracture by looking at the investigation results of the space shuttle Columbia disaster (Columbia Accident Investigation Report, 2003, p. 83). The only way to test the possible cause of the accident – a breach on the edge of the shuttle wing – was to conduct a full-scale experiment. No realistic computation of the breach was possible.

Below we first review the seismological background information necessary for further discussion as well as basic models of earthquake occurrence (Section 2). Short Section 3 describes the available earthquake catalogs. In Sections 4-7 evidence for the scale-invariance of earthquake process is presented, in particular, marginal distributions for the multidimensional earthquake process. Fractal distributions of earthquake size, time intervals, spatial patterns, focal mechanism, and stress are discussed. Section 8 describes several multidimensional stochastic models used to approximate earthquake occurrence. They are all based on the theory of branching processes; in this case the multidimensional structure of earthquake occurrence is modeled. In Section 8.3 we discuss the branching model of earthquake rupture: a physical multidimensional model based on random stress interactions. The model uses very few free parameters and appears to reproduce all the fundamental statistical properties of earthquake occurrence. Section 8.4 briefly describes the application of statistical models to forecast an earthquake occurrence. The final discussion (Section 9) summarizes the results obtained thus far and discusses problems and challenges still facing seismologists.

2. Seismological background

2.1. Earthquakes

Since this paper is intended for seismologists, physicists, and mathematicians, we briefly describe earthquakes and earthquake catalogs as primary objects of the statistical study. A more complete discussion can be found in Kagan (1994b, pp. 162-165), Bolt (2003), Lee *et al.* (2002), Scholz (2002), Kanamori and Brodsky (2004). As a first approximation, an earthquake may be represented by a sudden shear failure – the appearance of a large quasi-planar dislocation loop (Aki and Richards, 2002) in rock material.

Fig. 1a shows a fault-plane trace on the surface of the Earth. Earthquake rupture starts on the fault-plane at a point called the 'hypocenter' (the 'epicenter' is a projection of the hypocenter on the Earth's surface), and propagates with a velocity close to that of shear waves (2.5–3.5 km/s). The 'centroid' is in the center of the ruptured area. Its position is determined by a seismic moment tensor inversion (Ekström *et al.*, 2005, and references therein). As a result of the rupture, two sides of the fault surface are displaced by a vector along the fault-plane. For large earthquakes, such displacement is on the order of a few meters.

The earthquake rupture excites seismic waves which are registered by seismographic stations. The seismograms are processed by computer programs to obtain a summary of the earthquake's properties. Routinely, these seismogram inversions characterize earthquakes by their origin times, hypocenter (centroid) positions, and second-rank symmetric seismic moment tensors.

Fig. 1c represents ('beachball') the quadrupolar radiation patterns of earthquakes. The focal plots involve painting on

a sphere the sense of the first motion of the primary, P-waves: solid for compressional motion and open for dilatational. Two orthogonal planes separating these areas are the fault and the auxiliary planes. During the routine determination of focal mechanisms, it is impossible to distinguish these planes. Their intersection is the null-axis (N-axis), the P-axis is in the middle of the open lune, and the T-axis in the middle of the closed lune. These three axes are called the 'principal axes of an earthquake focal mechanism,' and their orientation defines the mechanism.

In the system of coordinates of TPN axes, shown in Fig. 1c, the seismic moment tensor matrix is

$$\mathbf{M} = M \times \mathrm{diag}[1, -1, 0]\,, \qquad (1)$$

where M is a scalar seismic moment of an earthquake, measured in Newton-m (Nm). In an arbitrary system of coordinates all entries in 3×3 matrix (1) are non-zero. However, the tensor is always traceless, with a zero determinant. Hence it has only four degrees of freedom: one for the norm of the tensor (proportional to the scalar seismic moment) and three for orientation (they define the focal mechanism of an earthquake). Another equivalent representation of the earthquake focus is a quadrupole source of a particular type (Fig. 1b) known in seismology as a "double-couple" (Burridge and Knopoff, 1964; Aki and Richards, 2002; Kagan, 2005b). The three representations of focal mechanism shown in Fig. 1 as well as in (1) are mathematically equivalent; Kagan (2005b) discusses interrelations between these parameterizations.

2.2. Description of earthquake catalogs

Modern earthquake catalogs are collections of estimated earthquake origin times, hypocenter or centroid locations, measures of earthquake size (scalar seismic moment or appropriate magnitude), and earthquake focal mechanisms or seismic moment tensors (Aki and Richards, 2002). Such datasets in a certain sense fully describe each earthquake; for instance one can compute far-field, low-frequency seismic radiation using the above information. However, detailed studies of earthquake occurrences show that this description is far from complete, since each earthquake represents a process with moment tensor or focal mechanism varying in extended time-space. Moreover, because earthquakes have fractal features, even defining an 'individual' earthquake is problematic: earthquake catalog records are the result of a complex interaction of fault ruptures, seismographic recordings, and their interpretations (see Section 5).

Fig. 2 displays a map of the local catalog for southern California (Kagan *et al.*, 2006). Earthquake focal mechanisms are shown by a stereographic projection (Aki and Richards, 2002). The focal mechanisms can be characterized by a 3-D rotation from a fixed position; an alternative, more compact representation of each mechanism is a normalized quaternion (Kagan, 1991c; 2005b).

In Fig. 3 we display a map of earthquake centroids in the global Harvard CMT catalog (Ekström *et al.*, 2005, and references therein). Earthquakes are mostly concentrated at tectonic plate boundaries. Each earthquake in this catalog is characterized by a centroid moment tensor solution.

There are many other datasets which characterize earthquake processes, such as detailed investigations of earthquake rupture for particular events, or earthquake fault maps and descriptions of certain faults. The unique advantages of an earthquake catalog include relative completeness, uniformity of coverage, and quantitative estimates of errors. These properties make catalogs especially suitable for statistical analysis and modeling. The catalogs can be roughly subdivided into two categories: global or regional catalogs covering large areas (continents or their large parts), and local catalogs for particular areas such as southern California (Fig. 2) and still smaller areas. As we discuss below, each type has its own properties and problems (Kagan, 2006).

A collection of earthquake occurrences can be represented as a multidimensional stochastic process (Kagan, 1994b): $T \times \mathbf{R}^3 (= \mathbf{R}^2 \times H) \times M \times \mathbf{SO}(3) (= \Psi \times S^2)$ (time-space-size-orientation), where $\mathbf{R}^3$ or $\mathbf{R}^2$ is the Euclidian space, H is the depth dimension, M is the scalar seismic moment, and $\mathbf{SO}(3)$ is the 3-D special orthogonal (rotation) group. The latter may be represented as a rotation by the angle Ψ around a rotation pole distributed over the 2-D sphere S^2 (Kagan, 1991c). Multiple studies summarized in Kagan (1994b) and Kagan and Vere-Jones (1996) indicate that marginal earthquake distributions are scale-invariant for all the above variables. The fractal pattern breaks down for large size, distance or time intervals. We discuss this in other sections of the paper.

An important feature of the available earthquake catalogs is the range of the above variables related to the average error in estimating them. The ratio of the range to an error describes roughly the information one can obtain from a catalog. These ratios are only approximate to one order of magnitude (see more in Kagan, 2003): a summary is shown in Table 1.

From Table 1 we see that the temporal structure of earthquake occurrences can be detailed with great precision. The locations of earthquake foci are estimated relatively accurately in the horizontal plane, but vertical errors are often significantly larger. This effectively reduces available spatial information. The influence of location errors and other nuisance variables often extends well above a catalog's reported accuracy values (Kagan and Knopoff, 1980; Kagan, 1991a; 2003). Similarly, boundary effects can be observed at distances substantially smaller than a region's total size. Therefore, the scale-invariant range of the spatial distribution is likely to be smaller than the 10^2–10^3 shown in Table 1. Focal mechanisms, which have been reliably obtained in only the last 25 years, have large uncertainties also (Kagan, 2000; 2003).

Catalogs are a major source of information on earthquake occurrence. Since the late nineteenth century certain statistical features were established: Omori (1894) studied temporal distribution; Gutenberg and Richter (1944) investigated size distribution; quantitative investigations of spatial patterns started late (Kagan and Knopoff, 1980).

Kostrov (1974) proposed that earthquake displacement can be described by a second-rank symmetric tensor. Gilbert and Dziewonski (1975) were the first to obtain a

Table 1. Information available in earthquake catalogs

#	Variable	Accuracy (A)	Range (R)	R/A
1	Origin time, T	0.01–1 s	5–25 y	$10^9 - 10^{11}$
2	Horiz. space, R^2	3-10 km	3000 km	10^3
2′	Horiz. space, R^2	0.5 km	200 km	$10^{2.5}$
2″	Horiz. space, R^2	0.02 km	2-20 km	$10^2 - 10^3$
3	Vert. space, R	5-15 km	50 km	10
3″	Vert. space, R	0.1 km	10 km	10^2
4	Moment magn., m	0.07	6.0	10^2
5	Rot. angle, Ψ	10°	120°	10^3
6	Rot. pole, S^2		360°	

2, 3 – global catalogs (Kagan, 2003);
2′ – local catalogs;
2″, 3″ – wave correlation catalogs (e.g., Hauksson and Shearer, 2005; Shearer *et al.*, 2005).

tensor solution from seismograms. However, statistical investigations remained largely restricted to time-size-space regularities. Tensor or focal mechanism analysis is difficult because we lack appropriate statistical tools to analyze either second-rank tensors or quaternions which properly represent earthquake focal mechanisms (Kagan, 1991c; 2005b). For example, in two recent special issues of the geophysical journals dedicated to statistical analysis of seismicity (*Pure Appl. Geoph.*, **162**(6-7), 2005; *Tectonophysics*, **413**(1-2), 2006) no paper analyzed the statistics of the seismic moment tensor or earthquake focal mechanism. Kagan and Knopoff (1985a;b) and Kagan (1992b;c; 2000; 2005b) were the first to investigate the statistical properties of the seismic moment tensor (Section 7).

2.3. Earthquake temporal occurrence: quasi-periodic, Poisson, or clustered?

The periodic or quasi-periodic hypothesis of large earthquake occurrence has long been held by geoscientists. Similar hypotheses are called 'seismic gap' or 'seismic cycle' models. A seismic gap, according to such a model, is a fault or a plate segment for which the time since the previous large earthquake is long enough that stress builds up. Since earthquake occurrence is multidimensional and periodicity is a property of a one-dimensional process, the seismic record needs to be converted into a temporal series.

The characteristic hypothesis (Schwartz and Coppersmith, 1984) implies a sequence of recognizably similar events and provides the logical basis for discussing recurrence or quasi-periodicity. Recurrence intervals and their statistics are meaningless without a clear definition of the characteristic earthquake (Jackson and Kagan, 2006). A characteristic earthquake is assumed to release most of the tectonic deformation on a segment. Other earthquakes are significantly smaller than the characteristic one and hence can be ignored when moment release is calculated.

McCann *et al.* (1979) adopted the gap model and produced a colored map of "earthquake potential" for close to a hundred circum-Pacific zones. They assumed that the seismic potential increases with the absolute time since the last large earthquake. Nishenko (1991) refined the seismic gap model so that it could be more rigorously tested. He specified the geographical boundaries, characteristic magnitudes, and recurrence times for each segment. He used a quasi-periodic, characteristic recurrence model to estimate conditional earthquake probabilities for 125 plate boundary segments around the Pacific Rim.

Kagan and Jackson (1991) compared the model of McCann *et al.* (1979) against later earthquakes. They found that large earthquakes occurred more frequently in the very zones where McCann *et al.* had estimated low seismic potential. In other words, they found that large earthquakes are rather clustered in time. Kagan and Jackson (1995) also found that earthquakes after 1989 did not support Nishenko's (1991) gap model. Rong *et al.* (2003) concurred: both predictions were inconsistent with the later earthquake record.

Bakun and Lindh (1985) proposed that a magnitude 6 earthquake would occur at the Parkfield, California segment of the San Andreas fault with a 95% probability in the time window 1985-1993. The prediction model was based largely on the characteristic, quasi-periodic earthquake hypothesis. This was the only prediction reviewed and approved by the U.S. government. However, no such earthquake occurred till 28 September 2004, when an earthquake of magnitude 6.0 struck near Parkfield (Bakun *et al.*, 2005). Meanwhile, a complicated form of the seismic gap model was applied to estimate earthquakes probabilities in the San Francisco Bay region (Working Group, 2003). The Working Group concluded that "there is a 0.62 [0.38–0.85] probability of a major, damaging [$M \geq 6.7$] earthquake striking the greater San Francisco Bay Region over the next 30 years (2002–2031)."

Stark and Freedman (2003) argue that the probabilities defined in such a prediction are meaningless because they cannot be validated. They point out that in weather predictions, 50% probability of rain can be tested by counting the ratio of the number of rainy days to the total having this predictive value. No such possibility exists for the predictions concerning San Francisco Bay (Working Group, 2003) or Parkfield. Stark and Freedman (2003) finish their review of the San Francisco Bay earthquake prediction with the advice that readers "should largely ignore the USGS probability forecast."

This lack of falsifiability and the inability to construct an improved hypothesis both contradict the fundamental requirements of modern scientific method (Kuhn, 1965; Popper, 1980). After the gap model was formulated in its present form 20 or 30 years ago (McCann *et al.*, 1979; Nishenko, 1991; Schwartz and Coppersmith, 1984) the model proponents did not attempt to verify its fundamental assumptions with a critical test. The apparent failure of the predictions (see above) was not extensively analyzed and explained (see, for example, debate, in the Nature magazine, 1999 or Bakun *et al.*, 2005). Jackson and Kagan (2006) review the implications of the Parkfield 2004 event for earthquake prediction, the characteristic earthquake hypothesis, and the earthquake occurrence in general. They argue that a simpler *null* hypothesis based on the Gutenberg-Richter law (see Section 4) and Poisson time behavior better explains the Parkfield event sequence. Despite this breakdown of scientific methodology, the potentially incorrect model continued to be in use for seismic hazard assessment in the U.S. and many other countries (Jackson and Kagan, 2006).

How could this happen? Geosciences are largely observational and descriptive disciplines. Earth scientists are not trained to formulate falsifiable hypotheses, critically test them, systematically review possible sources of error, thoroughly rule out alternative explanations, and dismiss or improve the models thereafter. Oreskes (1999) discusses how American earth scientists summarily rejected the theory of continental drift for decades before the 1960s, though extensive evidence existed in support of it. Suppe (1998) analyzes logical structure of the arguments by one of the most influential papers supporting plate tectonics (Morgan, 1968) and considers critical methods to validate hypotheses. Kagan (1999) argues that the major challenge facing earthquake seismology is that new methods for hypothesis verification need to be developed. These methods should yield reproducible, objective results, and be as effective, for instance, as double-blind testing in medical research.

2.4. Earthquake faults: one fault, several faults, or an infinite number of faults?

In Fig. 4a we display the most commonly used geometry of an earthquake fault: a planar boundary between two rigid blocks. Other block boundaries are usually considered to be free. When Burridge and Knopoff proposed this model in 1967, it was the first mathematical treatment of earthquake rupture and a very important development. Since then, hundreds of papers have been published using this model or its variants. We show below why seismology needs a much more complicated geometrical model to represent brittle shear earthquake fracture:

- 1. The old model (Fig. 4a) is a closed, isolated system, whereas tectonic earthquakes occur in an open environment. This model justifies spurious quasi-periodicity, seismic gaps, and seismic cycle models (Section 2.3). No rigorous observational evidence exists for the presence of these features in earthquake records (see above).

• 2. An earthquake fault in the model (Fig. 4a) is a well-defined simple geometrical object – a planar surface with the dimension 2.0. In nature, an earthquake fault *system* is a fractal set. This set is not a surface; its dimension is about 2.2 (Section 6.2).

• 3. Two distinct scales are present in the diagram: irregularities of the planar surface and block size. Resistance to block motion due to breakage of microscopical surface inhomogeneities is described as a friction force. The friction is an appropriate model for man-made surfaces, where the scale of inhomogeneities is limited. In contrast, earthquakes are scale-invariant. The geometry and mechanical properties of an earthquake fault zone are the result of self-organization. They are fractal (Kagan, 1994b; 2006).

• 4. A displacement incompatibility problem is wrongly circumvented because of the flat plate boundaries. Real earthquake faults always contain triple junctions (see, for example, Figs. 2 and 3); further deformation is impossible without creating new fractures and rotational defects (disclinations).

• 5. Because the block boundary is planar, stress concentrations are practically absent after a major earthquake. Hence these models have few or no aftershocks.

• 6. All earthquakes in the model have the same focal mechanism. Any variations in mechanisms obvious during even a cursory inspection of maps (as in Fig. 2) are not taken into account.

King (1983), Turcotte (1986) and Andrews (1989) suggested that due to kinematic effects at fault junctions, the fault geometry of earthquakes may be represented as a scale-invariant matrix of faults. Gabrielov *et al.* (1996) developed a mathematical framework for calculating the kinematic and geometric incompatibility in a tectonic block system, both rigid and deformable. They concluded that due to geometric incompatibilities at fault junctions, new ruptures must be created to accommodate large plate tectonic deformations. Indeed, plate tectonic observations indicate that hundreds of km of deformation occur over the several million years of plate boundary existence (e.g. the San Andreas fault system).

Fig. 4b,c display a few alternative models of earthquake faults: a smooth surface boundary and a fractal surface. Unless the smooth boundary is a surface of rotation, no large scale block displacement is possible. Similarly, to move blocks along the fractal boundary, one needs to break the surface inhomogeneities. In contrast to the model of Fig. 4a, the largest inhomogeneities are comparable to the block size.

Obviously, if major faults comprising a plate boundary are not strictly parallel, fault junctions are unavoidable. The question is whether large deformations can be accommodated by a few faults delineating a few tectonic blocks (see, for example, Ben-Zion and Sammis, 2003), or whether an infinite number of faults must exist to account for such deformations.

The above considerations suggest again that the conventional models of tectonic block deformation need complete revision. If the number of faults and fault junctions is infinite, these junctions, as Gabrielov *et al.* (1996) suggest, constitute 'asperities' and 'barriers' for fault rupture propagation. These geometric complexities, not friction, should control the developing fault system and the stop-and-go feature of the earthquake rupture propagation. Kagan (1987) shows that when the earthquake rupture plane rotates, as in triple junctions, the third-rank seismic moment tensor, which can be identified with asperities or disclinations, becomes non-zero.

In Fig. 4d we show a picture of a fractal boundary zone between two rigid blocks. In this case, a complex fault pattern cannot be characterized as a surface: it is a fractal set of dislocations. In Sections 6 and 7 we attempt to characterize this pattern quantitatively.

2.5. Statistical and physical models of seismicity

As we mentioned above, hundreds of papers are scattered in geophysical and physical journals on the statistical properties of seismicity. They propose physical or phenomenological models of earthquake occurrence. We will describe the papers and their results briefly.

Several authors, starting with Bak *et al.* (2002), attempt to collapse the time-distance-size earthquake distribution into one plot (see also Corral, 2005; Baiesi and Paczuski, 2005). Such plots prominently demonstrate the scale-invariant structure of seismicity known previously from marginal distributions. However, as mentioned earlier, although temporal, spatial, and magnitude distributions are scale-invariant for small values of pertinent variables, for larger values the scale-invariant pattern is replaced by finite-scale effects (Sections 4–6). Moreover, even for small variable values, the distributions are influenced by various random and systematic effects. The study of such errors is difficult in a collapsed multidimensional plot.

There are several groups of physical seismicity models. Most of them employ the geometrical and mechanical scheme illustrated in Fig. 4a as their major paradigm: two blocks separated by a planar surface (Dieterich, 1994; Rice and Ben-Zion, 1996; Langer *et al.*, 1996, see also Kanamori and Brodsky, 2004). Our earlier discussion of this model is also valid for these attempts: they ignore the spatial and mechanical complexity of the earthquake fault zone. Consequently, the deficiencies listed in the previous section are present in these models as well. Moreover, since these paradigms describe only one boundary between blocks, they do not account for a complex interaction between other block boundaries and, in particular, triple junctions. Seismic maps (Figs. 2 and 3) convincingly demonstrate that earthquakes occur mostly at boundaries of relatively rigid blocks. This is a major idea of plate tectonics (Morgan, 1968; Oreskes, 1999). However, if blocks are rigid, stress concentrations at other block boundaries and a block's triple junctions should influence earthquake patterns at any boundary. Thus, even after a large earthquake, the stress on a particular boundary can be restored almost immediately due to the influence of the block's other boundaries and its junctions.

Lyakhovsky *et al.* (1997; 2005) base their seismicity model on the damage rheology theory. In this case, where the mechanical properties of the rock medium are modeled, even elementary geometrical properties of a fault system are not considered. As a result, the fault geometry and earthquake focal mechanism distribution fall outside their work.

As we mentioned in Section 2.3, theoretical developments need to be critically tested against observational evidence. Otherwise, they remain in the realm of speculation. At the present time, numerical earthquake models have shown no predictive capability exceeding or comparable to empirical predictions based on earthquake statistics. Even if a theoretical or physical model exhibits some predictive skill, we should always question whether the predictive power comes from a deeper theoretical understanding, or from the earthquake statistics results imbedded in the model.

The models described above have a large number of adjustable parameters, both obvious and hidden, to simulate a complicated pattern of seismic activity. Dyson (2004) says that Enrico Fermi advised him

> ... My friend Johnny von Neumann used to say, with four parameters I can fit an elephant ...

The observational evidence in support of these models generally consists of particular earthquake in specific regions. In a random process there is always the possibility

of using a large corpus of data to select a particular series of events which seem to agree with theoretical predictions. For the model confirmation to be credible, the criteria for data selection must be prescribed in advance (Jackson and Kagan, 2006).

Another possibility for perceived confirmation of theoretical models is publication bias. This bias is caused by the fact that research with currently fashionable results is potentially more likely to be submitted and published than work with less appealing outcomes (Duval and Tweedie, 2000).

Therefore, physical models may not have a theoretical predictive capability. As Section 8.4 reports, although several phenomenological models issue a quantitative prediction of future seismicity, no physical model has yet attempted to compete in such tests.

2.6. Laboratory and theoretical studies of fracture

In engineering science extensive work has been performed on the conditions of tensile crack initiation and propagation (e.g., Anderson, 2005). However, these efforts are concentrated on the problem of a single crack: the most important problem for engineers.

The problem of crack propagation and branching, far more relevant to earthquakes, has been recently addressed in several papers. In laboratory experiments, a crack develops instabilities which make its path and propagation velocity highly chaotic and unpredictable (Marder, 1998; Sharon and Fineberg, 1999; Bouchbinder *et al.*, 2005). These instabilities and a sensitive dependence on the initial conditions are due to crack propagation, especially at a speed close to the elastic wave velocity. Stress and fracture conditions in laboratory specimens differ significantly from those in earthquake fault zones: in the laboratory the boundary effects are controlled by the researcher. Therefore, fractures can self-organize only at spatial scales much smaller than those of the specimen. In fault zones, the stress, rock mechanical properties, and fault geometry are self-organized as large-scale self-similar patterns develop.

The calculations of molecular dynamics (Marder, 1998; Sharon and Fineberg, 1999; Buehler and Gao, 2006) demonstrate that basic properties of tensile fracture can be effectively derived from simple laws. Similarly, precise laboratory measurements of fault propagation demonstrate multiple branching of fault surfaces. These simulations reproduce the fractal character of a fracture. Moreover, calculating the total energy balance in laboratory fracture experiments (Sharon and Fineberg, 1996; 1999) demonstrates that almost all elastic energy goes into creating new surface. Although the conditions during tensile fracture differ from those of the shear failure in earthquakes, the above result may be significant for the problem of the heat paradox for earthquake faults (Scholz, 1996).

3. Modern earthquake catalogs

Detailed modern earthquake catalogs with estimates of focal mechanism and/or seismic moment tensor were compiled beginning in the 1970s. Several extensive catalog datasets are available at present. Frohlich and Davis (1999) and Kagan (2003) discuss the properties of global catalogs and their accuracy.

The global catalog of the centroid moment tensors (CMT) is compiled by the Harvard group (Ekström *et al.*, 2005). The catalog contains 22,476 solutions over a period from 1976/1/1 to 2004/12/31, see Fig. 3. The Harvard catalog includes seismic moment centroid times and locations as well as estimates of the seismic moment tensor components. Each tensor is constrained to have a zero trace (first invariant): no isotropic component. Double-couple (DC) solutions, or solutions with the tensor determinant equal to zero, are supplied as well. Almost all earthquake parameters are accompanied by internal estimates of error.

The PDE worldwide traditional catalog (Preliminary Determination of Epicenters, 1999, and references therein) is published by the USGS. The catalog measures earthquake size, using several magnitude scales. Body-wave (m_b) and surface-wave (M_S) magnitudes are provided for most moderate and large events since 1965 and 1968, respectively. The catalog contains more than 50,000 shallow earthquakes with $m_b \geq 5$ since 1965.

The problem for almost all catalogs, especially local and regional, is their inhomogeneity: since any local seismographic network is bounded, the accuracy and catalog completeness vary considerably within the catalog area. This inhomogeneity is especially strong for the seismographic networks concentrated on island chains. Where stations represent a quasi-linear array, the location accuracy differs strongly along the direction of the array as compared to the orthogonal direction.

4. Earthquake size distribution

The distribution of earthquake sizes is usually invoked as a first confirmation for virtually any model of seismicity. Moreover, this distribution is by far the most studied feature of statistical seismology. Starting with its first discussion by Ishimoto and Iida (1939) and then Gutenberg and Richter (1944), it has been established that earthquakes increase in number as a power-law as their sizes decrease. This dependence is usually referred to as the magnitude-frequency or the Gutenberg-Richter (G-R) relation, and its parameter (see Eq. 4 below) is commonly known as the '*b*-value'. A very large body of literature exists concerning the size distribution, its interpretation and possible correlation with geotectonics, stress, rock properties, etc. For example, a search of the ISI WEB OF KNOWLEDGE database (isiknowledge.com/) for keywords like **'earthquake* and b-value'** yields about 110 publications in the last four years. However, that proliferation has not led to a deeper understanding of earthquake generation.

4.1. Magnitude versus seismic moment

Magnitude is an empirical measure of earthquake size and many different magnitude scales are currently used (see Castellaro *et al.*, 2006). Several types of errors need to be investigated in earthquake size measurement. Some of them are known to be connected with earthquake magnitude determination: saturation of all magnitude scales (Kanamori, 1977), which is explained by the finite seismogram frequency for a seismographic network. Other types of errors are common to both magnitude and seismic moment determination (Kagan, 2003).

Relatively high-frequency seismic waves are used to determine magnitude, the effects of scattering, multipathing, focussing and unfocussing are stronger as the wave periods decrease. These effects cause great variations of wave amplitude which lead to larger uncertainties and biases in magnitude measurements (*cf.* Ekström and Dziewonski, 1988).

Seismographic networks are limited in detecting weak earthquakes and their essential parameters such as hypocenter location, origin time and magnitude. This results in another limitation of magnitude distributions: at the lower magnitude end, progressively larger number of events are missing from catalogs. Unfortunately, this lower magnitude cutoff is neither sharp nor uniform over time and space.

In this paper M denotes the scalar seismic moment, and m denotes the magnitude of an earthquake, b is the parameter for magnitude distribution and β is the corresponding

parameter for seismic moment distribution. Earthquake moment magnitude m_w is related to the scalar seismic moment M via (Kanamori, 1977; Hanks and Kanamori, 1979)

$$m_w = \frac{2}{3}\log_{10} M - C\,, \qquad (2)$$

where seismic moment M is measured in Newton-m, and C is usually taken to be between 6.0 and 6.1. Below we use $C = 6.0$. Eq. 2 allows us to use the moment magnitude as a proxy for a physical quantity: seismic moment.

Seismic moment is proportional to the amplitude of seismic waves at zero or close to zero frequency; hence its accuracy is higher than that using magnitudes. Kagan (2003) estimates that uncertainty in moment magnitude is on the order of 0.1-0.15 and is by a factor of 2 to 3 smaller than regular magnitude uncertainties.

Because of saturation and other systematic effects, each magnitude can only be evaluated over a limited range of earthquake size. Different magnitude scales are then compared by using a regression relation

$$m_1 = C_1 + C_2\, m_2\,, \qquad (3)$$

where m_i are magnitudes and C_j are coefficients for a linear regression. Although both magnitudes in (3) usually have errors of similar size, regular, not orthogonal regression, is commonly used (Castellaro *et al.*, 2006). This should cause a significant bias in converting one magnitude into another. Most earthquake catalogs initially use several magnitudes. To obtain a common magnitude value, catalog compilers transform various magnitudes, using variants of (3). Errors and systematic effects of such calculations should significantly shape the estimates of earthquake size.

Inspecting the value of the C_2 coefficient in (3) in various publications, one can see the degree of the problem in determining magnitudes. C_2 should be close to 1.0 at the range of earthquake size where both magnitudes are well-defined. This is rarely the case: the C_2-value often reaches 0.7 or 1.3 (Kagan, 2003, his Fig. 14). These fluctuations of the conversion coefficient may cause spurious variations of the b-value. In contrast, when the moment magnitude in different catalogs is compared, the C_2 coefficient is close to 1.0 (Kagan, 2003, his Fig. 12).

4.2. Seismic moment distribution

Gutenberg and Richter's (1944) magnitude-frequency relation is usually written as

$$\lg N(m) = a - b\, m\,, \qquad (4)$$

where $N(m)$ is the number of earthquakes with magnitude $\geq m$, and a and b are parameters: a characterizes seismic activity or earthquake productivity of a region and b parameter describes the relation between small and large earthquake numbers, $b \approx 1$. The expression (4) has been proposed in the above functional form by Vilfredo Pareto (1897, p. 305, his Eq. 1) for the financial income distribution.

The original G-R distribution (4) can be transformed into the Pareto distribution for the scalar seismic moment M

$$\phi(M) \;=\; \beta\, M_t^{\beta}\, M^{-1-\beta} \qquad \text{for} \quad M_t \leq M, \qquad (5)$$

where β is the index parameter of the distribution, $\beta = \frac{2}{3}b$ (see Eq. 2), and M_t is the observational threshold.

Simple consideration of the finiteness of seismic moment flux or the deformational energy, available for earthquake generation, requires that the Pareto relation (5) be modified at the large size end of the moment scale. The distribution density tail must have a decay stronger than $M^{-1-\beta}$ with $\beta > 1$. This problem is generally solved by introducing into the distribution an additional parameter called the *maximum* or *corner* moment (M_x or M_c).

The tapered G-R relation has an exponential taper applied to the cumulative number of events with the seismic moment larger than M (Vere-Jones *et al.*, 2001; Kagan, 2002a)

$$\Phi(M) \;= (M_t/M)^{\beta}\, \exp\left(\tfrac{M_t - M}{M_c}\right) \qquad \text{for} \quad M_t \leq M < \infty, \qquad (6)$$

here M_c is the parameter that controls the distribution in the upper ranges of M ('the corner moment'). Fig. 5 displays the seismic moment distribution and its approximation by (6) with a choice of several corner magnitudes (m_c) for the subduction zones (Bird and Kagan, 2004).

Many other displays of the moment-frequency relation for earthquakes before 2003 can be found at the power-point file http://moho.ess.ucla.edu/~kagan/Tokyo_Univ.ppt. Updates of the moment distribution parameters after the 2004 great Sumatra earthquake are presented at http://element.ess.ucla.edu/publications/2004_global_coupling-/2004_global_coupling.htm.

The corresponding probability density function is

$$\phi(M) = \left[\frac{\beta}{M} + \frac{1}{M_c}\right] (M_t/M)^{\beta}\, \exp\left(\frac{M_t - M}{M_c}\right). \qquad (7)$$

The above distribution in both expressions (Eqs. 6–7) was proposed by Pareto (1897, pp. 305-306, his Eqs. 2, 2*bis*, and 5).

In Fig. 6 we show the result of the maximum likelihood determination of the β-values for eight tectonic provinces (Bird and Kagan, 2004). All 95% confidence limits include $\beta \approx 2/3$ value. This can be considered a universal parameter of earthquake size distribution.

The next diagram (Fig. 7) displays the corner moment values evaluated for the same eight provinces. For convenience they are shown on the map of central America, where all the provinces are represented. In contrast to the β-value result, Bird and Kagan (2004) find that at least four distinct values of the corner magnitude seem to be required, based on the 95% confidence limits. These values include Oceanic Spreading Ridge (normal faulting, corner magnitude range, $m = 5.7 - 6.0$); Oceanic Transform Faults (medium and fast velocities, range, $6.4 - 7.0$); all the Continental zones, Oceanic Transform Faults and slow velocity/Oceanic Convergent Boundary (range, $7.4 - 8.7$); and Subduction zone boundaries (range, $9.1 - \infty$).

Using the earthquake size distribution (6, 7), we can calculate the seismic moment rate (Kagan, 2002c)

$$\dot{M}_s \;=\; \frac{\alpha_0\, M_0^{\beta}}{1-\beta}\, M_c^{1-\beta}\, \Gamma(2-\beta)\,, \qquad (8)$$

where α_0 is the annual rate of earthquakes with moment M_0 or greater and Γ is the gamma function. Subsequently we can compare it to the tectonic rate evaluated by plate motion or by geodesy (Kagan, 2002c, Bird and Kagan, 2004). Thus, plate tectonic predictions can be quantitatively related to seismic activity. Below we show that because of the power-law property of earthquake size distribution, any naive comparison of the cumulative seismic moment and tectonic deformation yields unreliable results in most cases.

4.3. Seismic moment sum distribution

The global distribution of the seismic moment is well approximated by a power-law (Pareto) distribution with index

$\beta \approx 2/3$ (5). This is a heavy-tailed distribution: it has the infinite mean and the standard deviation. Thus, if one uses the pure Pareto model, the Central Limit Theorem does not describe the distribution of the sum of seismic moments. The tapered (Eqs. 6–7) or truncated Pareto distribution appears to eliminate all summation problems. However, a detailed analysis shows that the Gaussian limit is reached for only a large number of observations. For a realistic number of events, the tapered or truncated Pareto still exhibits all the properties of a heavy-tailed distribution.

The cumulative seismic moment released in a region can be used as a proxy for total regional deformation of the Earth surface due to earthquakes. Formally, the strain rate for the volume of the deforming crust is proportional to the sum of the *tensor* moments of individual earthquakes (Kostrov, 1974). Thus, evaluating observed seismic moment rates is an important problem connected to the regional earthquake hazard assessment.

Let $X_i, i = 1, \ldots, n$ be independent identically distributed random variables with a common Pareto distribution (5), and let S_n denote their sum

$$S_n = \sum_{i=1}^{n} X_i \,. \tag{9}$$

If the exponent β of the power-law distributed variable is less than 2.0, the sum converges to a stable distribution (Samorodnitsky and Taqqu, 1994; Uchaikin and Zolotarev, 1999) with the probability density function

$$\phi(X, \beta, \gamma, \mu, \sigma) \,, \tag{10}$$

where γ is a symmetry parameter (for positive variables $\gamma = 1$, i.e., the sum is maximally-skewed), and μ, σ are shift and width parameters. For the Gaussian distribution (see Eq. 16 below) only the two last parameters are valid.

An arbitrary quantile z_q of the sum S_n can be approximated as (Zaliapin *et al.*, 2005)

$$z_q \approx z_q^{(1)} \equiv n^{1/\beta} x_q C_\beta + b_n, \tag{11}$$

where x_q solves the equation for the cumulative distribution F_β of the sum

$$F_\beta(x_q) = q, \tag{12}$$

and $z_q^{(1)}$ is a quantile for a maximally asymmetrical (maximally-skewed) stable distribution. For $\beta < 1$, $b_n = 0$ and

$$C_n \;=\; [\Gamma(1-\beta)\cos(\pi\beta/2)]^{1/\beta} \;. \tag{13}$$

Fig. 8 displays an example of simulated sums (S_n) for the Pareto distribution truncated at $y = M_x/M_t = 3.4 \times 10^4$ compared to the stable distribution quantiles. In this example we take the threshold moment $M_t = 10^{17}$ Nm or $m_t = 5.33$, the threshold of the recent Harvard catalog (Kagan, 2003) and the maximum magnitude $m_x = 8.35$ (Section 4.2).

According to (11), quantiles of the stable distribution increase as $n^{1/0.66}$, thus, for example, the median of the sum of 40 variables μ_{40} compared to μ_{20} is equal to

$$\mu_{40} \;\approx\; 2.86 \times \mu_{20} \quad \text{or} \quad \mu_{40} \;>\; \mu_{20} + \mu_{20} \,. \tag{14}$$

This behavior of the stable distribution sums may seem counter-intuitive, as is that of their other properties.

If the exponent β is less than 1.0, the sum of power-law distributed variables is comparable to the value of the largest observation M_n

$$E(S_n) = M_n/(1-\beta). \tag{15}$$

where E is a mathematical expectation sign (Feller, 1971).

Therefore, Eq. 14 means that in a sample of 40 earthquakes, there is a higher chance of encountering a large event which would significantly increase their sum than in a sample of only 20 earthquakes. Pisarenko (1998) as well as Huillet and Raynaud (2001) also note that for the heavy-tailed distributions, sum quantiles increase non-linearly with n.

The upper quantiles of the Pareto sum generally approach the stable distribution limit faster than do the lower quantiles (Zaliapin *et al.*, 2005). However, in Fig. 8, the upper quantiles depart from the theoretical curve for the stable distribution starting with $n = 2$ because of the upper limit truncation. The behavior of the lower quantile is essentially unaffected by the truncation until n exceeds 10^3.

When the number of summands is large, the truncation point y dominates the behavior of the quantiles. The sum is then distributed asymptotically according to the Gaussian law:

$$\lim_{n\to\infty} F_{S_n}(x) = \Phi\left(\frac{x \,-\, n\,\mu_y}{\sigma_y\,\sqrt{n}};\, 0,\, 1\right), \tag{16}$$

where Φ is the normal cumulative distribution (17), and the parameters μ_y and σ_y are given by (18), (19).

The standard Gaussian (normal) cdf with expectation μ and standard deviation σ is given by

$$\Phi\left(x; \mu, \sigma^2\right) = \frac{1}{\sigma\,\sqrt{2\,\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(y-\mu)^2}{2\,\sigma^2}\right) dy \,, \tag{17}$$

whereas

$$\mu_y = \frac{\beta}{1-\beta}\,\left(y^{1-\beta} - 1\right) / \left(1 - y^{-\beta}\right), \quad \beta \neq 1 \tag{18}$$

and

$$\sigma_y^2 = \frac{\beta}{2-\beta}\,\left(y^{2-\beta} - 1\right) / \left(1 - y^{-\beta}\right) - \mu_y^2, \quad \beta \neq 2 \tag{19}$$

are the conditional mean and variance of each summand (Kagan, 2002a; Zaliapin *et al.*, 2005), given the restriction on the maximum ($X < y$).

From the beginning of the plate tectonics hypothesis, it was assumed that earthquakes are due to plate boundary deformation. Calculations for global tectonics and large seismic regions justified such an approach. However, applying this assumption to smaller regions has usually been inconclusive, given the high variability of seismic moment sums.

Holt *et al.* (2000) compared the observed seismic moment release with the tectonic release inferred from the joint inversion of the GPS and the Quaternary rates of strain for south-east Asia. They also compared strain release with the earthquake record from 1965-1998. Fig. 9 shows the seismic coupling χ (the ratio between observed and predicted seismic moment) in 4 large regions and 42 subregions. The coupling is calculated as

$$\chi \;=\; \frac{1}{T}\sum_{i=1}^{n} M_i/\dot{M}_{\mathrm{tec}} \,, \tag{20}$$

where T is the total catalog time and $\dot{M}_{\mathrm{tec}}$ is the rate of tectonic deformation.

Three regimes are clearly seen, depending on the number of earthquakes in a region. These regimes are perfectly reproduced by the truncated Pareto model. Fig. 9 approximately displays the quantiles (upper, middle, lower) for the

ratio between the sum of n random variables and the corresponding mean (which is finite in the truncated model). Notice that we see the conventional Gaussian picture (the sum is proportional to the mean) only with a very large number of events exceeding m_t: $n > 1000$. Thus, the truncated Pareto model explains the non-linear behavior of the cumulative moment release for small to intermediate numbers of earthquakes as a transition between pure power and pure Gaussian approximations.

5. Temporal earthquake distribution

Omori (1894) showed that the aftershock rate decays approximately as

$$n(t) = \frac{K}{t+c}, \tag{21}$$

where K and c are coefficients, t is the time since the mainshock origin, and $n(t)$ is the aftershock frequency measured over a certain interval of time. Presently a more complicated equation is used to approximate the aftershock rate

$$n(t) = \frac{K}{(t+c)^p}. \tag{22}$$

This expression with the additional exponent parameter p is called the 'modified Omori formula' (Utsu *et al.*, 1995).

Kagan (2004) and Kagan and Houston (2005) argue that the observed saturation in the aftershock numbers described by the "time offset" parameter c in Omori's law is likely an artifact due to the under-reporting of small aftershocks. This under-reporting comes from the difficulty of detecting large numbers of small aftershocks in the mainshock coda, as well as other factors (Kagan, 2004). For even smaller time intervals, close to the rupture time of the mainshock and aftershocks, the point model of the earthquake process breaks down, so that Eqs. 21 and 22 are no longer valid.

Fig. 10 displays the aftershock distribution for the 2004 great Sumatra earthquake. The general time-magnitude aftershock pattern is seen in many other aftershock sequences (Kagan, 2004): larger aftershocks begin early in the sequence, and the occurrence rate is progressively delayed for weaker events. After the aftershocks start in any magnitude band, they seem to be almost uniformly distributed over the log time. This pattern would correspond to the aftershock rate's decay according to Omori's law (21).

Two displays in Fig. 10 exhibit an important property of earthquake catalogs: in the PDE catalog aftershocks start at about 10^{-2} days after the mainshock, whereas in the CMT catalog they start at about 10^{-1} days. The total number of aftershocks as well as their magnitude range also significantly differ in these two diagrams. The main reason is the frequency range of the seismograms used in compiling both catalogs: in the PDE catalog the aftershocks are determined using waves with 1 s period, whereas the CMT catalog uses low frequency (period 50 s and greater) waves. The magnitude estimates in the PDE catalog saturate at about $m_b = 6.0 - 6.5$ (Section 4.1): therefore, we see no large magnitude aftershocks in its display. On the other hand, long-period coda waves of the mainshock and large aftershocks in the CMT catalog extend over a longer time. They make it difficult to discern smaller events in the seismograms. Thus, Fig. 10 as well as the arguments in Kagan (2004) and Kagan and Houston (2005) demonstrate that the c-value depends on methods of seismogram interpretation. It is probably not a physical parameter.

Therefore, depending on the frequency characteristics of a seismographic network, the number of stations, and the seismogram processing technique, the same earthquake sequence could be variously identified. In one catalog it could be identified as one complex earthquake with some subevents, but in another as a foreshock/mainshock/aftershock sequence with many 'individual' events (Kagan, 2003; 2004). Thus 'an individual earthquake' results from interpretations and selections made by catalog compilers. It is not in itself a physical entity, as tacitly assumed in most statistical analyses of seismicity. It is a naming artifact.

Omori's law has been incorporated in many phenomenological and physical models of earthquake occurrence. Like aftershocks, foreshocks also follow the power-law rate increase before a mainshock (Kagan, 1991b; Ogata, 2004).

6. Earthquake location distribution

6.1. Multipoint spatial moments

In the early 1980s, Kagan and Knopoff (1980) and Kagan (1981a;b) investigated the spatial moment structure of earthquake hypo- and epicenters for global (PDE and others) catalogs and several local earthquake catalogs.

The two-, three-, and four-point moment functions were obtained and analyzed in these studies. The quality and quantity of the earthquake data were relatively poor at that time, and computer capacity was sufficient to study only a small subset of data at a time, especially when computing the higher moments. However, it became apparent that the spatial distribution of earthquakes has a scale-invariant pattern.

Spatial moment functions are basic to the investigation of hypocentral patterns. The principal quantities of the study are the proportions of k-tuples ($k = 2, 3, 4$) points from the catalog (Fig. 11) with the property that the maximum distance between any two points in the k-tuple does not exceed r, as a function of r, and the joint density function of the distances between the points forming such a k-tuple.

We write

$$q_k(r) = N_k(r)/N_k, \tag{23}$$

where $N_k(r)$ is the number of k-tuples with the stated property, and N_k is the total number of k-tuples from the catalog. The quantities $q_k(r)$ are computed first for the epicenters, as points in $\mathbf{R}^2$, and then for the hypocenters, as points in $\mathbf{R}^3$. This function can be interpreted as the average number of k-tuples within a distance r of an "average" point of the catalog.

To overcome the biases in such estimates, which arise from boundary effects, the ratios $q_k(r)$ may be compared to the corresponding values for simulated Poisson catalogs. The simulated catalogs are of the same size and extent as the original catalog, but their epicentral coordinates are uniformly distributed over the region, and the depth distribution is matched to that of the actual catalog.

This results in the ratios

$$Q_k(r) = q_k(r)/\tilde{q}_k(r), \tag{24}$$

where the tilde refers to the simulated catalog. Values of $Q_k(r)$ have been tabulated and graphed in various ways.

The graphs of the ratios $Q_k(r)$ against r typically display three ranges: the initial range, the middle range over which the $1/r$ behavior is observed, and the final range, in which the ratio approaches 1.0, as r approaches the diameter of the observed region. We interpret the first range as dominated by measurement errors. The second range illustrates self-similar behavior, and the third range is dominated by boundary effects. Kagan (2006) provides extensive analysis of various errors and biases in the 2-point moment evaluation.

Thus, our key results can be summarized as follows. The growth rates of the moment functions are consistent with a dimensional deficit of approximately 1.0. Within an order of magnitude over different radial and angular combinations:

- i) the distribution of pairs of points selected at random from the catalog is consistent with the density inversely proportional to the distance $1/D$ (Kagan and Knopoff, 1980; Kagan, 1991a; 2006);
- ii) the distribution of triplets of points selected at random from the catalog is consistent with the density inversely proportional to the area, $1/S$ (Kagan, 1981a);
- iii) the distribution of quadruplets of points selected at random from the catalog is (for the hypocenters only) consistent with the density inversely proportional to the volume, $1/V$ (Kagan, 1981b).

6.2. Correlation dimension

Kagan (1991a; 2006) revisited the two-point moment problem and was able to more accurately estimate the correlation dimension for shallow (0-70 km depth), intermediate (70-300 km), and deep (300-700 km) earthquakes. In these papers the dependence of the moment on the time interval between earthquakes was also investigated.

To demonstrate the influence of time limits on the correlation dimension, Fig. 12 shows the distribution of distances between accurately located hypocenters in southern California. The probability density function of distances $N_3(R)$ between these hypocenters, irrespective of the inter-earthquake time interval, is close to a power-law $N_3(R) \propto R^{\delta}$ in the range $0.1 \leq R \leq 5$ km, where $N_3(R)$ is the number of pairs in the 3-D space. The correlation fractal dimension (measured by the least-square linear regression of $\log(R)$ and $\log[N_3(R)]$ for $0.1 \leq R \leq 5$ km) is $\hat{\delta} \approx 1.5$ (black lines in Fig. 12, see also Helmstetter *et al.*, 2005). The faster decay for $R < 0.1$ km is due to location errors, and the roll-off for distances $R > 5$ km is due to the finite thickness of the seismogenic crust. For larger distances ($R > 50$ km), the $\hat{\delta}$ decrease is caused by catalog boundaries (Kagan, 2006).

For $N_3(R,t)$, the correlation dimension $\hat{\delta}$ increases between $\hat{\delta} \approx 0$ at times $t = 5$ minutes up to $\hat{\delta} \to 2$ for $t = 2500$ days. This maximum inter-event time of 2500 days is long enough that earthquake interactions are relatively small compared to the tectonic loading. This value $\hat{\delta} = 2$, measured for $t = 2500$ days, can thus be interpreted as approaching the fractal dimension of the active fault network.

In Fig. 13 we display epicentral and hypocentral moments for earthquakes in the PDE catalog at three depth intervals. We include all the earthquake pairs without taking the inter-earthquake time into account. The curves are normalized, so that the horizontal line corresponds to a self-similar distribution with $\delta = 2.0$. The curves below the horizontal line have $\delta \geq 2.0$ (the fractal dimension is equal to the tangent of the slope angle of the curve plus 2.0). Since the epicentral moments are defined in 2-D, the horizontal line corresponds to $\delta = 1.0$. To show their differences, we combine two types of curves in one plot: epicentral and hypocentral. As in Kagan and Knopoff (1980), epicentral moments yield a higher value of the exponent for distance ranges less than or comparable to the thickness of the appropriate layer. From simple geometrical arguments, the hypocentral curves are the preferred data input to calculate the fractal dimension (Kagan, 2006).

Comparing Figs. 12 and 13, we conclude that self-similarity of earthquake geometry is established up to the scale length of 0.1 km and less. Since the equations of elasticity lack any intrinsic scale, we expect that the property of self-similarity can be extended for the brittle fracture of disordered materials (rocks) up to the scale of a few millimeters: the size of rock grains. The upper cutoff for scale-invariance (2000 km) is connected to the size of major tectonic plates. The δ values in Fig. 13 demonstrate that the dimension decreases as the depth increases.

Kagan (2006) indicates that evaluating the fractal dimension for earthquake spatial patterns is difficult and prone to many errors and biases. This may explain the contrast with two other classical scale-invariant, universal exponents of earthquake distribution: unlike the G-R relation (Bird and Kagan 2004) and Omori's law (Kagan, 1991b; Kagan and Houston 2005), here the properties and value of the correlation dimension are not yet firmly established.

In most studies of earthquake spatial distribution, errors of location and other errors have not been properly considered (Kagan, 2006). This might explain the high values of fractal dimensions reported in many publications and the great variability of these values. Such findings may reflect not physical and geometrical properties of earthquake fracture, but rather various location and projection errors peculiar to the catalogs studied.

6.3. Spatial scaling

Kagan (2002b) investigated the distribution of aftershock zones for large earthquakes (scalar seismic moment $M \geq 10^{19.5}$ Nm, moment magnitude, $m \geq 7$) in global catalogs (CMT and PDE). The dependence of the aftershock zone length, l on the earthquake size was studied for three representative focal mechanisms: thrust, normal, and strike-slip.

The regression curves in Fig. 14 show that $M \propto l^3$ dependence continues up to $m = 9$ earthquakes. Estimated regression parameters for strike-slip and normal earthquakes are similar to those of thrust events, supporting the conjecture that the scaling relation is identical for earthquakes of various focal mechanisms. No observable scaling break or saturation occurs for the largest earthquakes ($M \geq 10^{21}$ Nm, $m \geq 8$). It is natural to assume that the aftershock zone length l is equal or proportional to the rupture length L. Thus, earthquake geometrical focal zone parameters are self-similar.

Using the derived scaling law and moment-frequency relation, we can derive the distribution of earthquake slip not only for a region, but also for a specific place on a fault (Kagan, 2005a). This distribution depends on the linear size of earthquake rupture. For example, if the rupture is relatively short, a particular spot on a fault would be ruptured less frequently but would have a larger slip.

7. Focal mechanism orientation and stress distribution

7.1. Focal mechanism distribution

It is difficult to measure the stress tensor itself in the deep interior of the Earth, but rotations of earthquake focal mechanisms may indicate the stress redistribution. Kagan (1982) introduced the rotational Cauchy distribution to represent rotations of focal mechanisms of micro-dislocations which comprise the focal zone of an earthquake. The rotational Cauchy distribution can be written as (Kagan 1990)

$$F(\Psi) = \frac{2}{\pi}\left[\arctan(A/\kappa) - \frac{A \times \kappa}{A^2 + \kappa^2}\right], \qquad (25)$$

where $A = \tan(\Psi/2)$ and Ψ is the rotation angle. The scale parameter κ of the Cauchy distribution represents the degree of *incoherence* or *complexity* of an earthquake fault.

An additional complication in studying the 3-D rotation of earthquake focal mechanisms is the symmetry of the source: the double-couple earthquake source has the rotational symmetry of a rectangular box with unequal sides.

Due to this symmetry, the maximum rotation angle for the earthquake source cannot exceed 120° (Kagan, 1990; 1991c).

Using the correspondence between the group **SO**(3) and the group of normalized quaternions, we solved an inverse problem of a 3-D rotation of double-couple earthquake sources. For each pair of focal mechanisms, we find a minimum 3-D rotation which transform one mechanism into another (Kagan, 1991c).

In Fig. 15 we display, as an example, the distributions of rotation angle Ψ for shallow earthquake pairs which are separated by a distance of less than 50 km and in a distance range of 400-500 km. We study whether the rotation of focal mechanisms depends on the location of the second earthquake in the pair with regard to the first event. Thus, we measure the rotation angle for hypocenters located in 30° cones around each principal axis (curves marked the T-, P-, and N-axes) of the first event (see Fig. 1). The curves in Fig. 15 for small distances are narrowly clustered, and are clearly well approximated by the rotational Cauchy distribution.

For large distances the curve corresponding to fault-planes (the N-axis) is clearly separated from the curves connected with the T- and P-axes. Although the rotation near the fault-plane is relatively small ($\kappa \approx 0.2$), the earthquakes situated in cones around the T- and P-axes have focal mechanisms essentially uncorrelated with the primary event: the curves are close to the curve corresponding to a completely random rotation of a double-couple (see formulas in Kagan, 2005b).

Fig. 16 displays a smoothed map of the average Ψ dependence on time and distance intervals for well-constrained earthquakes in the Harvard catalog (Frohlich and Davis, 1999; Kagan, 2000). The angle increases with distance between events. The increase with time interval (ΔT) is much less pronounced. For earthquake sequences clustered in time and space, the Ψ difference between focal mechanisms is small, on the order $10 - 15°$. These Ψ-values are close to the minimum uncertainty in Ψ evaluation (Kagan, 2003).

7.2. Random stress tensor

The aim of earthquake seismology is to rigorously describe the tensor stress field which triggers earthquakes. Until now, extensive attempts to study stress fields have been concentrated on stress tensor properties at particular points, especially at hypocentral locations of potential future earthquakes (e.g., Kagan, 1994a; Harris, 1998; Stein, 1999; Steacy *et al.*, 2005). However, if the earthquake spatial distribution is indeed fractal, the stress field must also be scale-invariant, representing an extremely complicated matrix with critical conditions for earthquake rupture satisfied in an infinite number of points. This would correspond to an infinite number of micro-earthquake occurrences, if one extrapolates the G-R law for earthquake size distribution (Section 4.2) toward earthquakes of size zero.

While it is apparent that earthquakes are triggered everywhere in seismic regions, the question remains unsolved why small earthquake ruptures develop into giant events which can cause massive destruction. Answering this question adequately will require a detailed description of the 3-D stress field geometry, including its singularities, limit cycles, and possible bifurcations (Gabrielov and Keilis-Borok, 1983). This is an extremely difficult and open problem: Gabrielov and Keilis-Borok (1983, p. 489) comment that "The [mathematical] problem of the complete description [of the topology of the field...] has not as yet been solved."

The Cauchy distribution is especially important for representing earthquake geometry. It can be shown by theoretical arguments (Zolotarev, 1986, pp. 45-46; Kagan, 1990) and by simulations (Kagan, 1990) that the stress tensor in the medium with defects follows this distribution. Kagan (1990) argues that the Cauchy distribution of the stress should produce the rotational Cauchy distribution of earthquake focal mechanisms.

For any point in an elastic medium which is surrounded by defects, the characteristic function for the random stress distribution can be written as

$$\log \phi(\zeta, \alpha) = \int_0^\infty [\exp(i\zeta\sigma r^{-3}) - 1]\nu(r) r^2 dr, \qquad (26)$$

where $\nu(r)$ is the density of defects which might depend on r, the distance of the defect from the reference (measurement) point, and σ is the normalized (for $r = 1$) stress Green function of an earthquake; stress decays with distance as r^{-3}. For the uniform 3-D distribution of defects, $\nu = \nu_0$. In this case (26) yields the Cauchy stress distribution.

Earthquake spatial distribution, as described in Section 6.2, is fractal. In (26) we should substitute the fractal distribution of sources $\nu = \nu_0 r^{\delta - D}$, where $D = 3$ is the Euclidean dimension of the space, and δ is a fractal correlation dimension of earthquake hypocenters. Then (*cf.* Zolotarev, 1986, eq. 1.1.16)

$$\begin{aligned} \log \phi(\zeta, \alpha) &= \nu_0 \int_0^\infty [\exp(i\zeta\sigma u) - 1] u^{(\delta/3)-1} du \\ &= \nu_0 \Gamma(-\alpha) |\zeta|^\alpha, \end{aligned} \qquad (27)$$

with $\alpha = \delta/3$. The above formula means that if $\delta = 3$, the resulting distribution is the Cauchy law (Zolotarev, 1986; Kagan, 1990), whereas for a fractal spatial distribution of earthquakes, $\alpha < 1$.

Analyzing seismic moment and stress tensors has been basic to earthquake seismology. Although tensors are fundamentally important, they have not been sufficiently investigated or interpreted from a statistical point of view in the earth sciences, with few exceptions. A linear error propagation was first independently proposed to derive the error estimate of the principal stresses and their orientations by Angelier *et al.* (1982) and Soler and van Gelder (1991). The correlation study on the invariant quantities of seismic moment tensors was investigated by Kagan and Knopoff (1985a;b). Kagan (1992b;c; 2000) further extended the correlation results on invariant quantities to analyze earthquake catalogs and interpret faulting geometry.

The study of random tensors has its root in nuclear physics (see e.g., Girko, 1990; Mehta, 1991) and multivariate statistics (see e.g., Anderson, 1958). For nuclear physics, a simple rotation-invariant distribution has been widely investigated. But in multivariate statistics, only a handful of large sample or asymptotic distribution results involving such distributions are available. These results, despite their significance, can not be applied directly to the Earth sciences, because the number of tensors derived from the same original source is generally small. In particular, the ratio of signal to noise is not large enough to neglect the effect of nonlinearity. More importantly, efforts in nuclear physics and statistical mathematics have largely been focused on the principal invariants, namely, the principal eigenvalues. Very little attention has been paid to random eigendirections, which are equally important in the Earth sciences. Moreover, the nonlinearity of the mapping onto the eigendirections and eigenvalues has been insufficiently studied. This nonlinearity could strongly affect the estimated eigenvalues and directions if the noise level is high. Xu (1999; 2002) and Cai *et al.* (2005, and references therein) have attempted to develop a probabilistic approach in dealing with random/stochastic tensors in geoscience. The main new results from such studies include exact distributions for the random eigenvalues and eigendirections. They also include accuracy estimates of a higher order and bias computations.

8. Stochastic processes and earthquake occurrence

Most distributions considered so far have been one-dimensional marginal distributions of the earthquake point process. Two enhancements of this picture need to be presented: multidimensional distributions are to be constructed and the point structure of the process needs revision. Fig. 10, for example, shows that earthquake rupture duration needs to be taken into account when very small time intervals are considered. In Fig. 12 we show the influence of inter-earthquake time intervals on the spatial structure of earthquake distribution. In Section 6.3 we show that the focal zone of an earthquake, especially a large one, cannot be regarded as a point.

A more basic way to study the multidimensional structure of earthquake process is to apply the theory of stochastic point processes (Daley and Vere-Jones, 2003), not ordinary statistical methods. The first applications of this theory to earthquake occurrence were made by Vere-Jones (1970), Kagan (1973a;b), and Ogata (1988). Many researchers (Console *et al.*, 2003; Helmstetter and Sornette, 2004, and others) have recently applied the theory of stochastic point processes to analyze earthquake occurrence and clustering. The major impetus for these investigations is application of statistical methods for earthquake forecasting, both long- and short-term. Below we briefly review the available methods for earthquake occurrence analysis and their application for earthquake forecasting. We then discuss how these methods can be improved.

8.1. Earthquake clustering

Almost any earthquake forecast requires proper accounting for earthquake clustering, mainly for aftershocks. If present, foreshocks may be used to calculate a mainshock probability. Even if we are mainly interested in a long-term earthquake forecast, the influence of short-term earthquake clustering on our results should be estimated. Moreover, a faithful modeling of the earthquake clustering is needed for any short-term forecast.

Clustering presents a special challenge since modern local catalogs have a magnitude range extending over several units: in California and Japan, the lower magnitude threshold is close to 1.0, whereas the largest earthquake may exceed 8.0. In such catalogs one should expect the aftershock numbers approaching or even exceeding millions after a very strong event. Handling these earthquakes and accounting for various systematic and random effects both present serious challenges.

Fig. 17 displays a sketch of earthquake catalog data in the magnitude-time format. The left part of all the diagrams is the past for which no information is available, and similarly for the right or future part. Some earthquakes are detected below the magnitude threshold, shown as a dashed line.

Aftershock sequences have traditionally been taken into account by catalog declustering. Declustering can be used only as a preliminary step in seismicity analysis: it is subjective; and many different techniques are available but they are not optimized and have not been rigorously tested. We must use quantitative statistical methods to rigorously describe earthquake clustering. Only an application of stochastic point process theory can provide a robust solution to the problem.

However, the multidimensional nature of earthquake occurrence, fractal or power-law properties of earthquake statistical distributions, and inhomogeneities of earthquake distributions all make it difficult to create and statistically analyze stochastic models. Over the years several such models of earthquake occurrence have been proposed and all are based on the theory of branching processes. Branching is expected to model the well-known property of primary and secondary clustering for aftershock sequences: a strong aftershock (or foreshock) tends to have its own sequence of dependent events. These multidimensional models are:

- (A) Point process branching along the magnitude axis, introduced by Kagan (1973a;b) and shown in Fig. 17b.
- (B) Point process branching along the time axis (Hawkes and Adamopoulos, 1973; Hawkes and Oakes, 1974; Kagan and Knopoff, 1987b; Ogata, 1988 – called Hawkes self-exciting process, see Fig. 17c). Hawkes and Tukey (see discussion section in Kagan, 1973b) debate the difference between branching in earthquake size and in time.
- (C) Continuum-state critical branching process develops along the time axis (Kagan and Knopoff, 1981; Kagan, 1982; see Section 8.3).

The first two models (A-B) use the Poisson cluster process to approximate the earthquake occurrence. In these models, earthquake clusters are assumed to follow the Poisson temporal occurrence. Earthquakes within a cluster are modeled by a multidimensional branching process which reproduces a temporal-spatial pattern of dependent events (mostly aftershocks) around the initial event of a sequence (Kagan, 1973a;b; Kagan and Knopoff, 1987b; Ogata, 1988; 2004).

These models employ in one form or another the classical statistical properties of earthquake occurrence: the G-R relation and Omori's law. Model (A) reproduces the G-R relation as the result of branching along the magnitude axis and uses Omori's law to describe earthquake clustering in time. Model (B) combines the G-R relation and Omori's law in a fairly empirical fashion to approximate seismicity. Physical model (C) yields the G-R law as the consequence of critical branching (Vere-Jones, 1976). It applies a version of Omori's law to the temporal distribution of micro-dislocations and simulates the position and orientation of dislocations to reproduce the entire earthquake process (Section 8.3). As we discuss below, other models may have certain advantages in earthquake forecasting and the representation of seismicity. But phenomenological model (B) is now almost exclusively used to statistically analyze and simulate earthquake occurrence (Kagan and Knopoff, 1987b; Kagan and Jackson, 2000; Ogata, 2004).

Models (A) and (B) can be parameterized to analyze earthquake catalogs. The optimal parameter values can then be found by the maximum likelihood method (Kagan, 1991b; Ogata, 1988; 2004). To account for earthquake clustering, one can put the obtained parameter values back into the model and find the probabilities for each event to be foreshock/mainshock/aftershock (Kagan and Knopoff, 1976; Zhuang *et al.*, 2004). If these probabilities are known, a catalog can be either declustered in an objective manner, or dependent events can be taken into account.

Most of the statistical models for earthquake occurrence (Ogata, 1988; Kagan, 1991b) treat earthquake catalogs as a population set, with earthquakes considered as individual entities. As we discuss in Section 5, 'an individual earthquake' is not a physical entity. Instead it is the result of interpretation and selection by catalog compilers. Thus, extrapolations of observed features toward smaller inter-earthquake time intervals, smaller size earthquakes, etc., may see a model breakdown. Such approximation deterioration is caused not by physical properties of earthquake occurrence, but by peculiarities of earthquake identification technique and catalogs. Why is this?

8.2. Several problems and challenges

- 1. Earthquake spatial distribution is very complex: the depth inhomogeneity, the fractal character of the spatial pattern, and various hypocenter location errors all make model parameterization difficult and create various biases in estimating parameters. Recent applications of stochastic point

processes for seismicity analysis often yield results which are incompatible or unstable: slight variations in the data, assumptions, or processing techniques yield significantly different parameter values (Kagan, 1991b). It is difficult to see whether these contradictions are caused by biases of analysis, data defects, or differences in parametrization.

• 2. A critical and careful analysis of errors in the earthquake catalogs needs to be performed before each statistical analysis. Otherwise, unless the effect being studied is very strong, the results are almost surely artifacts. The problem is that most errors in the earthquake data are caused by systematic effects, so they are more difficult to identify and to correct (Kagan, 2003).

• 3. There is no effective statistical tool to select proper models and check whether they fit the data. Likelihood methods and the 'Akaike Information Criterion' (AIC) dependent on them (see Ogata, 2004; Daley and Vere-Jones, 2004) apparently work only for regular processes: quasi-Gaussian in a continuous case and quasi-Poisson for discrete (point) processes. However, an earthquake occurrence is controlled by scale-invariant, fractal distributions, diverging to infinity. Although these infinities can be regularized by using renormalization procedures similar to techniques used in model (C), statistical tests applicable to such distributions have not been developed yet. Calculating the likelihood function for aftershock sequences illustrates this point: the rate of aftershock occurrence after a strong earthquake increases by a factor of thousands. Log(1000) = 6.9; hence, one close aftershock yields a contribution to the likelihood function analogous to about 7 free parameters.

• 4. What can be done in the present situation to obtain reliable statistical results? The model's number of degrees of freedom should be kept as small as possible: the new adjustable parameters are to be introduced only if they are critically tested against the data in various catalogs and against different tectonic environments.

• 5. Earthquake catalogs are incomplete in a wake of strong events (Section 5). They are also incomplete generally for small earthquakes (Section 4.2). Both of these effects need to be carefully accounted for (Kagan, 2004).

• 6. Until now, only worldwide seismicity or seismicity in certain seismic zones has been analyzed. Several tectonic provinces have not been investigated sufficiently: deep earthquakes, oceanic earthquakes, earthquakes in stable continental areas, and volcanic earthquakes. The dependence of earthquake clustering on the rate of tectonic deformation should also be investigated: for example, in continental areas (and specifically in California) aftershock sequences occur in zones of fast and slow deformation rate. Are the clustering properties of earthquakes the same in these conditions? A study of earthquake occurrence in these tectonic environments should yield important information on general properties of seismicity.

• 7. Apparently all the statistical models based on Omori's law fail to capture the properties of long-term earthquake clustering. Kagan and Jackson (1991) argued that, in addition to short-term clustering which manifests in foreshock/mainshock/aftershock shallow event sequences, long-term clustering also occurs. The latter phenomenon is common both to shallow and deep earthquakes. Kagan and Jackson (1991) conjectured that short-term clustering results from stress redistribution in a brittle crust; long-term clustering is most likely due to mantle convection.

• 8. Earthquake probabilities calculated using model (B) have a serious defect: if a strong event is preceded by a foreshock or a number of foreshocks, this large quake is considered dependent. Model (A) does not present this difficulty; the largest event in a cluster is always the mainshock.

• 9. Point models by definition provide only a point forecast. Each future earthquake is characterized by its location, magnitude, time, and possibly its focal mechanism. In reality, earthquakes are spatially extended and they are not instantaneous. This is especially important for large events. Therefore, to compute seismic hazard, a point forecast needs to be supplemented by an extended source model. In contrast with models (A) and (B), model (C) is in principle a continuum which can simulate realistic, complex rupture process extended in time, space, and fault orientation.

8.3. Critical continuum-state branching model of earthquake rupture

Kagan (1982) proposed a model of earthquake rupture which incorporated results of 2-, 3-, and 4-point moment studies (Kagan and Knopoff, 1980; Kagan, 1981a;b) and tried to reproduce the inferred geometrical properties of hypocenter distributions. The model was based on the propagation (governed by a critical branching process) of infinitesimal dislocation loops.

The simulation proceeds in three stages. In the first stage the branching family trees are started from a number of initial ancestors as in Fig. 18. The second stage of simulation involves adding time delays between the appearance of the parent and the offspring. The delay is power-law distributed (Fig. 19a)

$$X(t) \propto t^{-1-u}. \tag{28}$$

For shallow earthquakes Kagan and Knopoff (1981) find that $u \approx 1/2$.

Kagan and Knopoff (1987a) show that the distribution (28) may have a simple explanation: stresses at the end of an earthquake rupture are below the critical value and thereafter change randomly according to a one-dimensional Brownian motion. A new rupture starts when stress reaches a critical level. The level-set of this motion is a fractal set with a dimension $u = 0.5$ (Mandelbrot, 1983). The distribution of time intervals is Lévy type which has density (Zaliapin *et al.*, 2005)

$$f_{1/2}(x) = \frac{1}{x\sqrt{2\pi x}} \exp\left(-\frac{1}{2\,x}\right), \tag{29}$$

and cdf

$$F_{1/2}(x) = 2\left[1 - \Phi\left(\frac{1}{\sqrt{x}}; 0, 1\right)\right], \tag{30}$$

where Φ is the Gaussian distribution (17).

With this information available, a cumulative plot of the number of elementary events against time can be obtained. (In seismological terms, each elementary event is supposed to contribute a fixed amount to a scalar moment release, so that cumulative plots can be interpreted as analogues to the cumulative moment-release plots used in discussing real earthquake seismograms).

The intense clustering of the near critical process results in this cumulative plot taking on a self-similar, step-function appearance. By convoluting the derivative of this cumulative function with a suitably shaped Green's function, a record can be obtained which may be compared with the trace of a seismograph or its envelope in reality (Fig. 20). By applying similar criteria to those used to identify real, particular events, Kagan and Knopoff (1981) were able from the time series record to list simulated 'events,' each with its own 'magnitude'.

In the third stage of modeling, the spatial coordinates (location of disc center, orientation, and direction) are simulated according to Fig. 19b. Although in principle dislocations are infinitesimal, in practical simulations the dislocation loops are finite with a disc radius r_0. However, this

radius can be taken to be as small as possible. In such a case the critical branching process converts into a continuum-state process (Jirina, 1958).

The rotation of the focal mechanisms follows the 3-D rotational Cauchy distribution (25). Most rotations are infinitesimal, though in rare cases large rotations give rise to fault branching (Kagan, 1982). As we explained above, the 3-D rotations are described by the group $\mathbf{SO}(3)$; hence, our model is a random branching walk on non-commutative groups. From the results of this stage, it is possible to obtain a visual picture of the resulting "fractures" by plotting the intersection of the elementary discs with a fixed plane (Fig. 21).

It is partly from such pictures that the angular Cauchy distribution, rather than some other angular analogue of the stable distributions, has been chosen. As can be seen from Section 7.2, this distribution also has a simple, physical explanation. The obtained distribution of fault traces looks like actual earthquake fault maps. The spatial moment functions are also qualitatively similar to those in Fig. 11: the 1/D and 1/V behaviors have been reproduced, although no rigorous tests have been attempted (Kagan and Vere-Jones, 1996).

By averaging the locations of the elementary dislocations resulting in such an 'event', an approximate location for the 'hypocenter' of the event can be determined. The centroid, representing roughly the center of gravity of the locations of the elementary events contributing to the cluster (Fig. 1), can be determined as well. In this way a synthetic seismic catalogue can be produced in which events are listed in time sequence and associated with a hypocenter, magnitude and a focal mechanism. Processing the synthetic catalog through a maximum likelihood procedure similar to that used for real catalogs (Section 8.1) yields similar values of basic parameters describing an earthquake occurrence.

For extended rupture, like that shown in Fig. 21, we can calculate the seismic moment tensor of an earthquake or earthquake sequence

$$\mathbf{M} = \sum_{i=0}^{N} \mu \left(\prod_{\xi_i} q_i \times \ldots \times q_j \times \ldots \times q_0 \right) , \qquad (31)$$

where q_0 is a quaternion corresponding to the initial dislocation in Fig. 18. The quaternion product, ending with q_i, describes a combination of 3-D rotations at the path ξ_i in a branching process leading to the i-the dislocation. In a branching process such a path is unique. Each of the quaternion product q_j components follows the rotational Cauchy distribution (25). Thus, the quaternion product in the formula represents the orientation of the i-th dislocation. The operator $\mu(.)$ converts the orientation (quaternion) into the seismic moment tensor (Kagan and Jackson, 1994, their Appendix).

The quaternion multiplication is non-commutative, i.e., in general

$$q_1 \times q_2 \neq q_2 \times q_1 . \qquad (32)$$

Therefore, the resulting probability structure should be studied by a non-commutative probability theory (e.g., Voiculescu, 2000). Moreover, the moment tensor in (31) should have its smallest eigenvalue as non-zero. Hence the combined source would not be a double-couple (1), although it will likely only insignificantly differ from a double-couple (Kagan and Knopoff, 1985a). In this case normalized quaternions are insufficient to characterize complex moment tensors as in (31). These tensors, even after normalization, require four degrees of freedom for representation. Higher-order seismic moment tensors (Kagan, 1987) can be used to characterize the complex geometry of a fault rupture. The gauge theory of dislocations and disclinations in solids (Edelen and Lagoudas, 1988), most likely needs to be employed to describe complex earthquake rupture fully.

There has been recently renewed interest in the branching earthquake fault model. In a limited test, Libicki and Ben-Zion (2005) used a simplified procedure to reproduce some properties of the Kagan (1982) model.

8.4. Earthquake forecasting attempts

The fractal dimension of the earthquake process is lower than the embedding dimension: in time (1-D) the fractal dimension is 0.5 (28); the correlation dimension in 3-D space is 2.2 (Section 6.2). This allows us to forecast the spatial and temporal probability of earthquake occurrence.

8.4.1. Phenomenological branching models and earthquake hazard estimation

For the phenomenological forecast of seismic activity, we employed both methods shown in Figs. 17b;c: branching-in-magnitude and branching-in-time (see Kagan and Knopoff, 1977; Kagan and Knopoff, 1987b, respectively). It is not currently clear which technique would be more appropriate for earthquake forecasting. The advantages or drawbacks may depend on catalog properties or on goals of forecasting. With the first method (A), it is easier to calculate the earthquake rate at the detection threshold and extend the forecasts below the threshold. Forecasting in forward time would involve simulating various cluster probabilities (Kagan and Knopoff, 1977).

The second technique (B) is convenient for calculating the earthquake rate at the forward time boundary of an available catalog. This method is currently widely used in earthquake prediction efforts (see citations in the beginning of Section 8). However, to extend the forecast horizon into future, simulation is needed (Kagan and Jackson, 2000; Helmstetter and Sornette, 2004), since we need to consider earthquakes that occur between the end of catalog data and the prediction time. Similarly, if we want to take into account past seismicity (as shown in Figs. 17b;c, then simulation is needed in both models (A-B).

Since 1999 we have been running experimental short- and long-term forecasts of the west Pacific seismic activity (Kagan and Jackson, 2000). In Fig. 22 we display long-term forecast maps computed in 2000 for the north-west Pacific region.

We have tested the long-term forecast by a Monte-Carlo simulation (Kagan and Jackson 1994; 2000), see also http://scec.ess.ucla.edu/~ykagan/tests_index.html. The test involves comparing the forward prediction issued before the test period with a retrospective prediction optimized after 2002, when the earthquakes which occurred in 2000-2002 were known. If these two forecasts differ within the 95% confidence limit estimated by a simulation procedure, we consider the forward prediction successful. In effect, instead of competing against a null hypothesis which cannot be effectively defined for the inhomogeneous spatial distribution of seismicity, we test our results against the 'perfect', ideal model, specified on the basis of retroactively adjusting the model parameters. Using a similar technique, we produce a daily short- and long-term earthquake forecast for southern California (Helmstetter *et al.*, 2006).

Kagan and Knopoff (1987b) tested the short-term forecast by the maximum likelihood technique for a retrospective earthquake forecast at the San Andreas fault. Kossobokov (2006) tested our short-term western Pacific forecast and found that "... the achieved statistics are much better than random guessing." (However, see also our comment on his paper – Kagan and Jackson, 2006.)

8.4.2. Earthquake fault propagation modeling and earthquake hazard estimation

One way of using the continuum-state branching model of fault propagation (Section 8.3) is to apply it to maps of geologic faults and past earthquakes to predict the propagation of future earthquake ruptures. As we discussed earlier (Section 2.4), geometric compatibility conditions imply that faults must rupture virgin rock. The model which reproduces branching properties of real earthquake faults can be applied for extrapolation to fault data. Kagan and Knopoff (1984) made an early attempt to see how such a forecast could be formulated. They extrapolated the fault traces as the result of model simulation. In principle, if appropriate Green's functions are available, this model can generate a set of seismograms for each synthetic sequence.

9. Discussion

As we show in this paper, the earthquake process is controlled by scale-invariant, fractal distributions. Thus, mathematical "monsters" that were produced more than a hundred years ago (Mandelbrot, 1983; Feferman, 2000) are directly related to earthquake science and representations of seismicity.

Now we understand that we see and experience these monsters when, for example, we pour milk into tea or step over cracked pavement. Moreover, Section 8.3 proposes that these scale-invariant distributions can be at least generally explained by using simple assumptions on random stress behavior: (i) earthquake time behavior by the Brownian motion-like history of stress change and (ii) fault geometry by the Cauchy distribution of stress tensors due to randomly distributed defects in rock medium. This Cauchy distribution induces the Cauchy 3-D rotation of focal mechanisms (25). Such a physical and mathematical explanation is a relatively rare case in the study of fractal distributions (Mandelbrot, 1983).

However, it has not yet been explored whether simulated earthquakes (Section 8.3) are faithful representation of seismicity. Several reasons complicate the comparison: our observational data are not sufficiently detailed, especially with regard to spatial and angular resolution (see Table 1). The mathematical and logical structure of the stochastic model needs to be explored to see if it is consistent and can be extended to the continuum limit.

As we mentioned in several parts of this review paper, many of the mathematical techniques necessary to describe earthquake geometry and its occurrence are still being developed: (i) the theory of stable distributions and their statistics; (ii) statistics of 3-D rotations; (iii) random branching walk on non-commutative groups; (iv) the gauge theory of deformation in solids, etc.

But developing a comprehensible theory of earthquake rupture may encounter serious mathematical difficulties. Earthquake faults, as shown in Fig. 21, are stochastic fractal objects. The stress at the fractal boundary should be nowhere a differentiable function. Thus, it is possible that calculating earthquake rupture criteria for points close to a 'fault-tip' cannot be carried out effectively.

For example, for the deterministic Mandelbrot set (Mandelbrot, 1983, pp. 188-189) it has been shown (Blum *et al.*, 1998, p. 55) that even if we use *real-number* arithmetic operations, no algorithm can decide in a finite number of steps whether an arbitrary point in a complex plane is in the set. The reason for the 'undecidability' of the Mandelbrot set and many similar complex mathematical objects is that their boundary has a fractal Hausdorff dimension. Thus, it is possible that we cannot effectively calculate the boundary of earthquake rupture faults (Kagan, 1999).

Even if the above difficulties are resolved, more 'menacing' monsters are on the horizon: the Banach-Tarski theorem (French, 1988; Feferman, 2000) states that in a space of three and more dimensions a ball can be divided into several pieces and the pieces rearranged into two balls of the same size. This paradoxical result may mean that new ideas in the mathematics foundations may be needed to solve our problem.

In the introduction we mentioned the similarity of two major problems in classical physics: turbulence of fluids and fracture of solids (Kagan, 1992a) and an apparent lack of interest by theoretical physicists in solving the former problem. This, most likely, has a simple explanation. Goldstein (1969, p. 23) remarks:

> 'It was at a meeting of the British Association in London in 1932 that I remember that [Horace] Lamb remarked "I am an old man now, and when I die and go to Heaven there are two matters on which I hope for enlightenment. One is quantum electrodynamics, and the other is the turbulent motion of fluids. And about the former I am really rather optimistic."'

In other, apocryphal, more recent versions of the story Lamb is replaced by Einstein, von Neumann, Heisenberg, Feynman, and others (Yaglom, 2001; Gleick, 1987, pp. 121, 329). Does it reflect a general feeling among mathematical physicists that the turbulence problem may be unsolvable? The more complex problem of fracture in solids, including earthquake rupture process, may not be solved either.

This opinion starkly contradicts the optimism expressed by David Hilbert (1900) who said that any mathematical problem could be solved (see also comments by Feferman, 1994, p. 14):

> 'This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no **ignorabimus** [we shall not know].'

If we revisit our question in the title of this paper, the simple answer is that the theoretical explanation of earthquake occurrence is very difficult. It requires applying mathematical methods that are unfamiliar to geophysicists and physicists. For example, many papers and a few monographs (e.g., Fisher *et al.*, 1987) consider vector and axial statistics in 2-D and 3-D. However, there are almost no publications, except for those cited in Section 7, dealing with the statistics of 3-D rotations. Perhaps, a recent development of statistical theory for topological manifolds (Small 1996; Kendall *et al.* 1999) could be adapted for describing the complex geometry of earthquake faulting, including 3-D rotations of focal mechanisms.

What can be done? Clearly the level of mathematics employed in earthquake physics is inadequate. Presently, the mathematical tools used in seismological research go back to the 18th or to the first half of 19th century. As we explained above, the level of mathematics needs to be raised by the order of a magnitude. Results in the forefront of modern mathematical research should be employed to describe earthquake occurrence and the geometry of earthquake faults in particular. Mathematical disciplines, such as tensor analysis, matrix theory, group theory, topology, and theory of stochastic processes must be involved in the solution. Otherwise no significant progress is possible.

Although applied and pure mathematicians work in other geoscience disciplines like atmospheric and plasma physics or geodynamo theory, until now only statisticians have been studying earthquake occurrence problems. But if we look at the development of earthquake science in the U.S., no professional statisticians have been involved full time in the research. This situation contrasts with earthquake investigations in other countries: in Japan, Russia, and New Zealand statisticians have been involved. It is not surprising that earthquake prediction efforts in the U.S. have been particularly unsuccessful (see Section 2.3).

To summarize our discussion, we see that there are major, perhaps fundamental difficulties in creating a comprehensive physical/mathematical theory of brittle fracture and earthquake rupture process. On the other hand, developing quantitative models of earthquake occurrence needed to evaluate probabilistic seismic hazard is within our reach. It will require a combined effort of earth scientists, physicists, statisticians, and pure and applied mathematicians.

Acknowledgments. I appreciate partial support from the National Science Foundation through grants EAR 00-01128, EAR 04-09890, and DMS-0306526, as well as from the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The author thanks D. D. Jackson, P. Bird, F. Schoenberg and I. V. Zaliapin of UCLA, D. Vere-Jones of Wellington University and P. L. Xu of Kyoto University for very useful discussions. Invitation to the Workshop by Bikas K. Chakrabarti that led to this paper and his help with various problems in producing this paper is greatly appreciated. I am also very grateful to Kathleen Jackson who edited the final manuscript version. Publication 968, SCEC.

References

K. Aki, and P. Richards, *Quantitative Seismology*, 2nd ed., Sausalito, Calif., University Science Books (2002)

T. L. Anderson, *Fracture Mechanics: Fundamentals and Applications*, 3rd ed., Boca Raton, Taylor and Francis (2005)

T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York (1958)

D. J. Andrews, Mechanics of fault junctions, J. Geophys. Res., **94**, 9389-9397 (1989)

J. Angelier, A. Tarantola, B. Valette and S. Manoussis, Inversion of field data in fault tectonics to obtain regional stress - I. Single phase fault populations: a new method of computing stress tensor, Geophys. J. R. astr. Soc., **69**, 607-621 (1982)

P. Bak, K. Christensen, L. Danon and T. Scanlon, Unified scaling law for earthquakes, Phys. Rev. Lett., **88**, 178501(pp. 1-4) (2002)

W. H. Bakun and A. G. Lindh, The Parkfield, California, earthquake prediction experiment, Science, **229**, 619-624 (1985)

W. H. Bakun *et al.*, Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, Nature, **437**, 969-974 (2005)

M. Baiesi and M. Paczuski, Complex networks of earthquakes and aftershocks, Nonlinear Processes Geophys., **12**(1), 1-11 (2005)

Y. Ben-Zion and C. G. Sammis, Characterization of Fault Zones, Pure Appl. Geophys., **160**, 677-715 (2003)

P. Bhattacharyya, Of overlapping Cantor sets and earthquakes: analysis of the discrete Chakrabarti-Stinchcombe model, Physica A, **348**, 199-215 (2005)

P. Bhattacharyya, A. Chatterjee and B. K. Chakrabarti, A common mode of origin of power laws in models of market and earthquake, http://arxiv.org/abs/physics/0510038 (2005)

P. Bird and Y. Y. Kagan, Plate-tectonic analysis of shallow seismicity: apparent boundary width, beta, corner magnitude, coupled lithosphere thickness, and coupling in seven tectonic settings, Bull. Seismol. Soc. Amer., **94**(6), 2380-2399 (2004)

L. Blum, F. Cucker, M. Shub and S. Smale, *Complexity and Real Computation*, New York, Springer (1998)

B. A. Bolt, *Earthquakes*, 5th ed., New York, W. H. Freeman (2003)

E. Bouchbinder, I. Procaccia and S. Sela, Disentangling scaling properties in anisotropic fracture, Phys. Rev. Lett., **95**(25), 255503 (2005)

M. J. Buehler and H. J. Gao, Dynamical fracture instabilities due to local hyperelasticity at crack tips, Nature, **439**(7074), 307-310 (2006)

R. Burridge and L. Knopoff, Body force equivalents for seismic dislocations, Bull. Seismol. Soc. Amer., **54**, 1875-1888 (1964)

R. Burridge and L. Knopoff, Model and theoretical seismicity, Bull. Seismol. Soc. Amer., **57**, 341-371 (1967)

J. Q. Cai, E. W. Grafarend and B. Schaffrin, Statistical inference of the eigenspace components of a two-dimensional, symmetric rank-two random tensor, J. Geodesy, **78**(7-8), 425-436 (2005)

S. Castellaro, F. Mulargia and Y. Y. Kagan, Regression problems for magnitudes, Geophys. J. Int., accepted (2006) http://scec.ess.ucla.edu/~ykagan/silvia_index.html

B. K. Chakrabarti and R. B. Stinchcombe, Stick-slip statistics for two fractal surfaces: a model for earthquakes, Physica A, **270**(1-2), 27-34 (1999)

A. Corral, Renormalization-group transformations and correlations of seismicity, Phys. Rev. Lett., **95**(2), 028501 (2005)

Columbia Accident Investigation Board, Report, Volume I, NASA, U.S. government, Washington DC (2003)

R. Console, M. Murru and A. M. Lombardi, Refining earthquake clustering models, J. Geophys. Res., **108**(B10), Art. No. 2468 (2003)

D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer-Verlag, New York, 2-nd ed., Vol. 1 (2003)

D. J. Daley and D. Vere-Jones, Scoring probability forecasts for point processes: The entropy score and information gain, J. Applied Probability, **41A**, 297-312, (Sp. Iss.) (2004)

J. Dieterich, A constitutive law for rate of earthquake production and its application to earthquake clustering, J. Geophys. Res., **99**, 2601-2618 (1994)

S. Duval and R. Tweedie, Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis, Biometrics, **56**, 455-463 (2000)

F. Dyson, A meeting with Enrico Fermi – How one intuitive physicist rescued a team from fruitless research, Nature, **427**(6972), 297 (2004)

D. G. B. Edelen and D. C. Lagoudas, *Gauge Theory and Defects in Solids*, Amsterdam, North-Holland (1988)

G. Ekström and A. M. Dziewonski, Evidence of bias in estimation of earthquake size, Nature, **332**, 319-323 (1988)

G. Ekström, A. M. Dziewonski, N. N. Maternovskaya and M. Nettles, Global seismicity of 2003: Centroid-moment-tensor solutions for 1087 earthquakes, *Phys. Earth Planet. Inter.*, **148**(2-4), 327-351 (2005)

S. Feferman, Deciding the undecidable: wrestling with Hilbert's problems, math.stanford.edu/~feferman/papers/deciding.pdf (1994)

S. Feferman, Mathematical intuition vs. mathematical monsters, Synthese, **125**(3), 317-332 (2000)

W. Feller, *An Introduction to Probability Theory and its Applications*, **2**, 2-nd ed., J. Wiley, New York (1971)

N. I. Fisher, T. Lewis, B. J. J. Embleton, *Statistical Analysis of Spherical Data*, Cambridge, Cambridge University Press (1987)

R. M. French, The Banach-Tarski theorem, Math. Intelligencer, **10**(4), 21-28 (1988)

C. Frohlich and S. D. Davis, How well constrained are well-constrained T, B, and P axes in moment tensor catalogs?, J. Geophys. Res., **104**, 4901-4910 (1999)

A. M. Gabrielov and V. I. Keilis-Borok, Patterns of stress corrosion: geometry of the principal stresses, Pure Appl. Geophys., **121**, 477-494 (1983)

A. Gabrielov, V. Keilis-Borok and D. D. Jackson, Geometric incompatibility in a fault system, *P. Natl. Acad. Sci. USA* **93**, 3838-3842 (1996)

F. Gilbert and A. M. Dziewonski, An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, Phil. Trans. R. Soc. Lond. A, **278**, 187-269 (1975)

V. L. Girko, *Theory of Random Determinants*, Boston, Kluwer Academic Publishers (1990)

J. Gleick, *Chaos, Making a New Science*, Viking, New York (1987)

S. Goldstein, Fluid mechanics in the first half of this century, Annual Rev. Fluid Mech., **1**, 1-29 (1969)

B. Gutenberg and C. F. Richter, Frequency of earthquakes in California, Bull. Seism. Soc. Am., **34**, 185-188 (1944)

T. C. Hanks and H. Kanamori, A moment magnitude scale, J. Geophys. Res., **84**, 2348-2350 (1979)

R. A. Harris, Introduction to special section: Stress triggers, stress shadows, and implications for on seismic hazard, J. Geophys. Res., **103**, 24,347-24,358 (1998)

E. Haseltine, The 11 greatest unanswered questions of physics, Discover, **23**(2) (2002)

E. Hauksson and P. Shearer, Southern California hypocenter relocation with waveform cross-correlation, Part 1: Results using the double-difference method, Bull. seism. Soc. Am., **95**(3), 896-903 (2005)

A. G. Hawkes and L. Adamopoulos, Cluster models for earthquakes - Regional comparisons, Bull. Int. Statist. Inst., **45**(3), 454-461 (1973)

A. G. Hawkes and D. Oakes, A cluster process representation of a self-exciting process, J. Appl. Prob., **11**, 493-503 (1974)

A. Helmstetter and D. Sornette, Predictability in the epidemic-type aftershock sequence model of interacting triggered seismicity, J. Geophys. Res., **108**(B10), Art. No. 2482 (2004)

A. Helmstetter, Y. Y. Kagan and D. D. Jackson, Importance of small earthquakes for stress transfers and earthquake triggering, J. Geophys. Res., **110**(5), B05S08 (2005)

A. Helmstetter, Y. Y. Kagan and D. D. Jackson, Comparison of short-term and time-independent earthquake forecast models for southern California, Bull. Seismol. Soc. Amer., **96**(1), 90-106 (2006)

D. Hilbert, Mathematical Problems, Lecture at the International Congress of Mathematicians at Paris, English translation by M. W. Newson, Bull. Amer. Math. Soc., **8** (1902), 437-479 (1900)

W. E. Holt, N. Chamot-Rooke, X. Le Pichon, A. J. Haines, B. Shen-Tu and J. Ren, Velocity field in Asia inferred from Quaternary fault slip rates and Global Positioning System observations, J. Geophys. Res., **105**, 19,185-19,209 (2000)

T. Huillet and H. F. Raynaud, On rare and extreme events, Chaos, Solitons and Fractals, **12**, 823-844 (2001)

M. Ishimoto and K. Iida, Observations sur les seismes enregistres par le microsismographe construit dernierement (1), *Bull. Earthquake Res. Inst. Tokyo Univ.*, **17**, 443-478, (in Japanese) (1939)

D. D. Jackson and Y. Y. Kagan, The 2004 Parkfield earthquake, the 1985 prediction, and characteristic earthquakes: lessons for the future, Bull. Seismol. Soc. Amer., submitted (2006) http://scec.ess.ucla.edu/~ykagan/parkf2004_index.html

M. Jirina, Stochastic branching processes with continuous state space, Czech. Math. J., **8**, 292-313 (1958)

Y. Y. Kagan, A probabilistic description of the seismic regime, Izv. Acad. Sci. USSR, Phys. Solid Earth, 213-219, (English translation) (1973a)

Y. Y. Kagan, Statistical methods in the study of the seismic process (with discussion: comments by M. S. Bartlett, A. G. Hawkes, and J. W. Tukey), Bull. Int. Statist. Inst., **45(3)**, 437-453 (1973b)

Y. Y. Kagan, Spatial distribution of earthquakes: The three-point moment function, Geophys. J. Roy. astr. Soc., **67**, 697-717 (1981a)

Y. Y. Kagan, Spatial distribution of earthquakes: The four-point moment function, Geophys. J. Roy. astr. Soc., **67**, 719-733 (1981b)

Y. Y. Kagan, Stochastic model of earthquake fault geometry, Geophys. J. Roy. astr. Soc., **71**, 659-691 (1982)

Y. Y. Kagan, Point sources of elastic deformation: Elementary sources, static displacements, Geophys. J. Roy. astr. Soc., **90**, 1-34 (1987) (Errata, Geophys. J. R. Astron. Soc., **93**, 591, 1988.)

Y. Y. Kagan, Random stress and earthquake statistics: Spatial dependence, Geophys. J. Int., **102**, 573-583 (1990)

Y. Y. Kagan, Fractal dimension of brittle fracture, J. Nonlinear Sci., **1**, 1-16 (1991a)

Y. Y. Kagan, Likelihood analysis of earthquake catalogues, Geophys. J. Int., **106**, 135-148 (1991b)

Y. Y. Kagan, 3-D rotation of double-couple earthquake sources, Geophys. J. Int., **106**, 709-716 (1991c)

Y. Y. Kagan, Seismicity: Turbulence of solids, Nonlinear Sci. Today, **2**, 1-13 (1992a)

Y. Y. Kagan, On the geometry of an earthquake fault system, Phys. Earth Planet. Inter., **71**, 15-35 (1992b)

Y. Y. Kagan, Correlations of earthquake focal mechanisms, Geophys. J. Int., **110**, 305-320 (1992c)

Y. Y. Kagan, Incremental stress and earthquakes, Geophys. J. Int., **117**, 345-364 (1994a)

Y. Y. Kagan, Observational evidence for earthquakes as a nonlinear dynamic process, Physica D, **77**, 160-192 (1994b)

Y. Y. Kagan, Is earthquake seismology a hard, quantitative science?, Pure Appl. Geoph., **155**, 233-258 (1999)

Y. Y. Kagan, Temporal correlations of earthquake focal mechanisms, Geophys. J. Int., **143**, 881-897 (2000)

Y. Y. Kagan, Seismic moment distribution revisited: I. Statistical results, Geophys. J. Int., **148**, 520-541 (2002a)

Y. Y. Kagan, Aftershock zone scaling, Bull. Seismol. Soc. Amer., **92**(2), 641-655 (2002b)

Y. Y. Kagan, Seismic moment distribution revisited: II. Moment conservation principle, Geophys. J. Int., **149**, 731-754 (2002c)

Y. Y. Kagan, Accuracy of modern global earthquake catalogs, Phys. Earth Planet. Inter., **135**(2-3), 173-209 (2003)

Y. Y. Kagan, Short-term properties of earthquake catalogs and models of earthquake source, Bull. Seismol. Soc. Amer., **94**(4), 1207-1228 (2004)

Y. Y. Kagan, Earthquake slip distribution: a statistical model, J. Geophys. Res., **110**(5), B05S11 (2005a)

Y. Y. Kagan, Double-couple earthquake focal mechanism: random rotation and display, Geophys. J. Int., **163**(3), 1065-1072 (2005b)

Y. Y. Kagan, Earthquake spatial distribution: the correlation dimension, Geophys. J. Int., submitted (2006) http://scec.ess.ucla.edu/~ykagan/p2rev_index.html .

Y. Y. Kagan and H. Houston, Relation between mainshock rupture process and Omori's law for aftershock moment release rate, Geophys. J. Int., **163**(3), 1039-1048 (2005)

Y. Y. Kagan and D. D. Jackson, Long-term earthquake clustering, Geophys. J. Int., **104**, 117-133 (1991)

Y. Y. Kagan and D. D. Jackson, Probabilistic forecasting of earthquakes, Geophys. J. Int., **143**, 438-453 (2000)

Y. Y. Kagan and D. D. Jackson, Comment on 'Testing earthquake prediction methods: "The West Pacific short-term forecast of earthquakes with magnitude MwHRV $\geq$ 5.8"' by V. G. Kossobokov, Tectonophysics, **413**(1-2), 33-38 (2006)

Y. Y. Kagan D. D. Jackson and Y. F. Rong, A new catalog of southern California earthquakes, 1800-2005, Seism. Res. Lett., **77**(1), 30-38 (2006)

Y. Kagan and L. Knopoff, Statistical search for non-random features of the seismicity of strong earthquakes, Phys. Earth Planet. Inter., **12**(4), 291-318 (1976)

Y. Kagan and L. Knopoff, Earthquake risk prediction as a stochastic process, Phys. Earth Planet. Inter., **14**(2), 97-108 (1977)

Y. Y. Kagan and L. Knopoff, Spatial distribution of earthquakes: the two-point correlation function, Geophys. J. Roy. astr. Soc., **62**, 303-320 (1980)

Y. Y. Kagan and L. Knopoff, Stochastic synthesis of earthquake catalogs, J. Geophys. Res., **86**, 2853-2862 (1981)

Y. Y. Kagan and L. Knopoff, A stochastic model of earthquake occurrence, Proc. 8-th Int. Conf. Earthq. Eng., San Francisco, Calif., **1**, 295-302 (1984)

Y. Y. Kagan and L. Knopoff, The first-order statistical moment of the seismic moment tensor, Geophys. J. Roy. astr. Soc., **81**, 429-444 (1985a)

Y. Y. Kagan and L. Knopoff, The two-point correlation function of the seismic moment tensor, Geophys. J. Roy. astr. Soc., **83**, 637-656 (1985b)

Y. Y. Kagan and L. Knopoff, Random stress and earthquake statistics: Time dependence, Geophys. J. Roy. astr. Soc., **88**, 723-731 (1987a)

Y. Y. Kagan and L. Knopoff, Statistical short-term earthquake prediction, Science, **236**, 1563-1567 (1987b)

Y. Y. Kagan and D. Vere-Jones, Problems in the modelling and statistical analysis of earthquakes in: *Lecture Notes in Statistics (Athens Conference on Applied Probability and Time Series Analysis)*, **114**, C. C. Heyde, Yu. V. Prohorov, R. Pyke, and S. T. Rachev, eds., New York, Springer, pp. 398-425 (1996)

H. Kanamori, The energy release in great earthquakes, J. Geophys. Res., **82**, 2981-2987 (1977)

H. Kanamori and E. E. Brodsky, The physics of earthquakes, Rep. Prog. Phys., **67**, 1429-1496 (2004)

D. G. Kendall, D. Barden, T. K. Carne and H. Le, *Shape and Shape Theory*, New York, Wiley (1999)

G. King, The accommodation of large strains in the upper lithosphere of the Earth and other solids by self-similar fault systems: the geometrical origin of *b*-value, Pure Appl. Geophys., **121**, 761-815 (1983)

V. G. Kossobokov, Testing earthquake prediction methods: "The West Pacific short-term forecast of earthquakes with magnitude MwHRV $\geq$ 5.8", Tectonophysics, **413**(1-2), 25-31 (2006)

B. V. Kostrov, Seismic moment and energy of earthquakes, and seismic flow of rock, Izv. Acad. Sci. USSR, Phys. Solid Earth, January, 13-21 (1974)

T. S. Kuhn, Logic of discovery or psychology of research?, In *Criticism and the Growth of Knowledge*, eds. I. Lakatos and A. Musgrave, pp. 1-23, Cambr. Univ. Press, Cambrigde (1965)

J. S. Langer, J. M. Carlson, C. R. Myers and B. E. Shaw, Slip complexity in dynamic models of earthquake faults, Proc. Nat. Acad. Sci. USA, **93**, 3825-3829 (1996)

W. H. K. Lee, H. Kanamori, P. C. Jennings and C. Kisslinger, Eds, *IASPEI Handbook of Earthquake and Engineering Seismology, Part A*, Boston, Academic Press (2002)

E. Libicki and Y. Ben-Zion, Stochastic branching models of fault surfaces and estimated fractal dimension, Pure Appl. Geophys., **162**(6-7), 1077-1111 (2005)

V. Lyakhovsky, Y. Ben-Zion and A. Agnon, A viscoelastic damage rheology and rate- and state-dependent friction, Geophys. J. Int., **161**(1), 179-190; Correction, Geophys. J. Int., **161**(2), 420 (2005)

V. Lyakhovsky, Y. Ben-Zion and A. Agnon, Distributed damage, faulting, and friction, J. Geophys. Res., **102**, 27,635-27,649 (1997)

B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco, Calif., 2nd edition (1983)

M. Marder, Computational science – Unlocking dislocation secrets, Nature, **391**, 637-638 (1998)

W. R. McCann, S. P. Nishenko, L. R. Sykes and J. Krause, Seismic gaps and plate tectonics: seismic potential for major boundaries, Pure Appl. Geophys., **117**, 1082-1147 (1979)

M. L. Mehta, *Random Matrices*, 2nd ed., Boston, Academic Press (1991)

W. J. Morgan, Rises, trenches, great faults, and crustal blocks, J. Geophys. Res., **73**(6), 1959-1982 (1968)

Nature magazine, February-April of 1999. Debate on earthquake prediction, http://www.nature.com/nature/-debates/earthquake/equake_frameset.html

S. P. Nishenko, Circum-Pacific seismic potential – 1989-1999, Pure Appl. Geophys., **135**, 169-259 (1991)

J. F. O'Brien and J. K. Hodgins, Graphical modeling and animation of brittle fracture, Proceedings of Assoc. Computing Machinery (ACM) SIGGRAPH 99, 137-146 (1999)

A. Okabe, B. Boots, K. Sugihara and S. Chiu, *Spatial Tessellations*, 2nd ed. Wiley, Chichester (2000)

Y. Ogata, Statistical models for earthquake occurrence and residual analysis for point processes, J. Amer. Statist. Assoc., **83**, 9-27 (1988)

Y. Ogata, Space-time model for regional seismicity and detection of crustal stress changes, J. Geophys. Res., **109**(B3), Art. No. B03308. Correction J. Geophys. Res., **109**(B6), Art. No. B06308 (2004)

F. Omori, On the after-shocks of earthquakes, J. College Sci., Imp. Univ. Tokyo, **7**, 111-200 (with Plates IV-XIX) (1894)

N. Oreskes, *The Rejection of Continental Drift: Theory and Method in American Earth Science*, Oxford University Press, USA (1999)

V. Pareto, *Cours d'Économie Politique*, Tome Second, Lausanne, F. Rouge, see also Pareto, V., 1964. *Œuvres Complètes*, Publ. by de Giovanni Busino, Genève, Droz, v. II (1897)

R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*, New York, Knopf (2005)

V. F. Pisarenko, Non-linear growth of cumulative flood losses with time, Hydrological Processes, **12**, 461-470 (1998)

Preliminary Determination of Epicenters, (PDE), Monthly Listings, U.S. Dept. Interior/Geol. Survey, Nat. Earthquake Inform. Center, January, pp. 47 (1999) WEB: http://wwwneic.cr.-usgs.gov/neis/data_services/ftp_files.html.

K. R. Popper, *Logic of Scientific Discovery*, 2nd ed., London, Hutchinson (1980)

J. R. Rice and Y. Ben-Zion, Slip complexity in earthquake fault models, Proc. Nat. Acad. Sci. USA, **93**, 3811-3818 (1996)

Y.-F. Rong, D. D. Jackson and Y. Y. Kagan, Seismic gaps and earthquakes, J. Geophys. Res., **108**(B10), 2471, **ESE-6** (2003)

G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, New York, Chapman and Hall (1994)

C. H. Scholz, Faults without friction?, Nature, **381**, 556-557 (1996)

C. H. Scholz, *The Mechanics of Earthquakes and Faulting*, Cambr. Univ. Press, Cambridge, 2nd ed., (2002)

D. P. Schwartz and K. J. Coppersmith, Fault behavior and characteristic earthquakes: Examples from Wasatch and San Andreas fault zones, J. Geophys. Res., **89**, 5681-5698 (1984)

E. Sharon and J. Fineberg, Confirming the continuum theory of dynamic brittle fracture for fast cracks, Nature, **397**, 333-335 (1999)

E. Sharon and J. Fineberg, Microbranching instability and the dynamic fracture of brittle materials, Physical Review B, **54**, 7128-7139 (1996)

P. Shearer, E. Hauksson and G. Q. Lin, Southern California hypocenter relocation with waveform cross-correlation, part 2: Results using source-specific station terms and cluster analysis, Bull. seism. Soc. Am., **95**(3), 904-915 (2005)

C. G. Small, *The Statistical Theory of Shape*, New York, Springer (1996)

T. Soler and B. H. W. van Gelder, On covariances of eigenvalues and eigenvectors of second-rank symmetric tensors, Geophys. J. Int., **105**, 537-546 (1991)

D. Sornette, *Critical Phenomena in Natural Sciences (Chaos, Fractals, Self-organization, and Disorder: Concepts and Tools)*, New York, Springer, 2-nd ed. (2003)

P. B. Stark and D. A. Freedman, What is the chance of an earthquake?, Chapter 5.3 in *"EARTHQUAKE SCIENCE AND SEISMIC RISK REDUCTION"*, eds. F. Mulargia and R. J. Geller, pp. 201-213, Kluwer, Dordrecht (2003) preliminary draft of the document is available at http://oz.berkeley.edu/∼stark/Preprints/611.pdf .

S. Steacy, J. Gomberg and M. Cocco, Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard, J. Geophys. Res., **110**(B5), B05S01 (2005)

R. S. Stein, The role of stress transfer in earthquake occurrence, Nature, **402**, 605-609 (1999)

F. Suppe, The structure of a scientific paper, Phil. Sci., **65**(3), 381-405; also see pp. 417-424 (1998)

D. L. Turcotte, A fractal model for crustal deformation, Tectonophysics, **132**, 261-269 (1986)

D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics*, 2nd ed., Cambridge Univ. Press, Cambridge (1997)

V. V. Uchaikin and V. M. Zolotarev, *Chance and Stability: Stable Distributions and Their Applications*, Utrecht: VSP International Science Publishers (1999)

T. Utsu, Y. Ogata and R. S. Matsu'ura, The centenary of the Omori formula for a decay law of aftershock activity, J. Phys. Earth, **43**, 1-33 (1995)

D. Vere-Jones, Stochastic models for earthquake occurrence (with discussion), J. Roy. Stat. Soc., **B32**, 1-62 (1970)

D. Vere-Jones, A branching model for crack propagation, Pure Appl. Geophys., **114**, 711-725 (1976)

D. Vere-Jones, R. Robinson and W. Z. Yang, Remarks on the accelerated moment release model: problems of model formulation, simulation and estimation, Geophys. J. Int., **144**, 517-531 (2001)

D. Voiculescu, Lectures on free probability theory, in: *Lectures Notes in Mathematics, 1738*, pp. 283-349, ed. P. Bernard, Springer, Berlin (2000)

Working Group on California Earthquake Probabilities (WG02), *Earthquakes probabilities in the San Francisco Bay region: 2002 to 2031*, USGS, Open-file Rept. 03-214; (2003) http://pubs.usgs.gov/of/2003/of03-214 .

P. L. Xu, Isotropic probabilistic models for directions, planes, and referential systems, Proc. R. Soc. London, **458A**(2024), 2017-2038 (2002)

P. L. Xu, Spectral theory of constrained second-rank symmetric random tensors, Geophys. J. Int., **138**, 1-24 (1999)

A. Yaglom, The century of turbulence theory: The main achievements and unsolved problems, In: *New Trends in Turbulence*, eds. M. Lesieur, A. Yaglom, and F. David, pp. 1-52, NATO ASI, Session LXXIV, Springer, Berlin (2001)

I. V. Zaliapin, Y. Y. Kagan and F. Schoenberg, Approximating the distribution of Pareto sums, Pure Appl. Geoph., **162**(6-7), 1187-1228 (2005)

J. C. Zhuang, Y. Ogata and D. Vere-Jones, Analyzing earthquake clustering features by using stochastic reconstruction, J. Geophys. Res., **109**(B5), Art. No. B05301 (2004)

V. M. Zolotarev, *One-Dimensional Stable Distributions*, Amer. Math. Soc., Providence, R.I. (1986)

Yan Y. Kagan, Department of Earth and Space Sciences, University of California, Los Angeles, California, 90095-1567, USA; (e-mail: ykagan@ucla.edu)

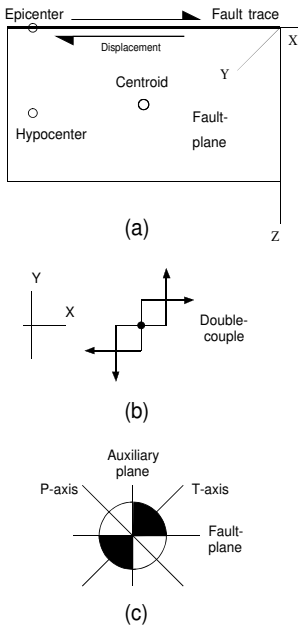

Figure 1. Schematic diagrams of earthquake focal mechanism.
(a) Fault-plane trace on the surface of the Earth. Earthquake rupture starts at the hypocenter (epicenter is the projection of a hypocenter on the Earth's surface), and propagates with velocity close to that of shear waves (2.5–3.5 km/s).
(b) Double-couple source, equivalent forces yield the same displacement as the extended fault (a) rupture in a far-field.
(c) Equal-area projection (Aki and Richards, 2002, p. 110) of quadrupole radiation patterns. The null (N) axis is orthogonal to the T- and P-axes, or it is located on the intersection of fault and auxiliary planes, i.e., perpendicular to the paper sheet in this display.

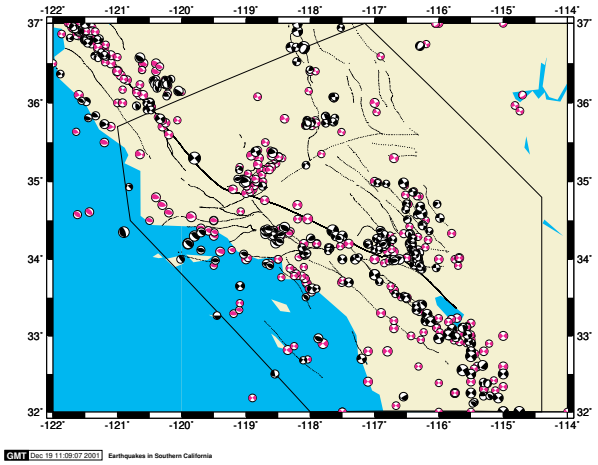

Figure 2. The southern California catalog and polygon: a region wherein the catalog is believed to be accurate and complete. Time period 1800–2002. Black beachballs – known solutions; orange beachballs – imputed solutions, obtained through interpolation from known focal mechanisms. Thin curved lines are active earthquake faults, including the San Andreas fault on which many earthquakes are concentrated.

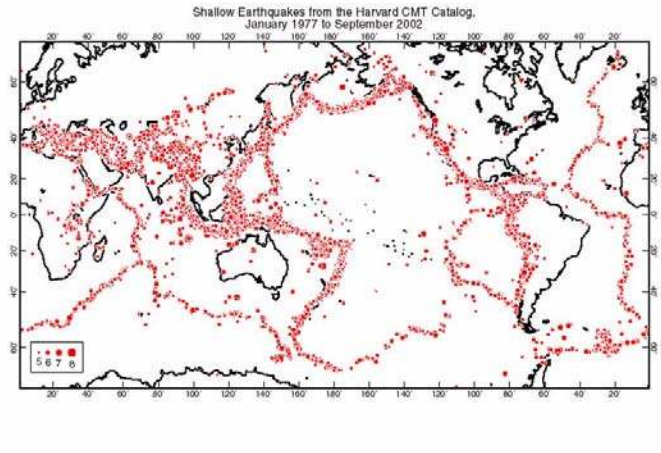

Figure 3. Location of shallow (depth 0-70 km) earthquakes in the Harvard CMT catalog, 1977/1/1–2002/9/30. Size of a symbol is proportional to earthquake magnitude.

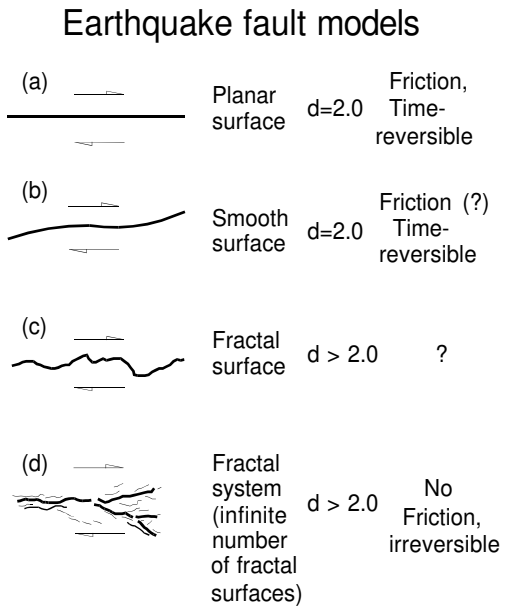

Figure 4. Earthquake fault models.

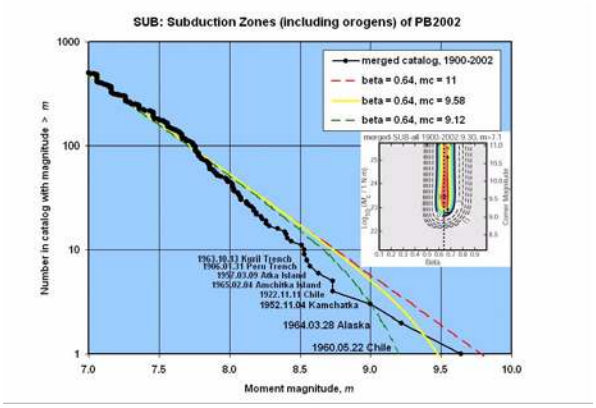

Figure 5. The size distribution for subduction earthquakes (Bird and Kagan, 2004). The inset shows the likelihood map which is used to estimate both the β and the corner moment (M_c) values.

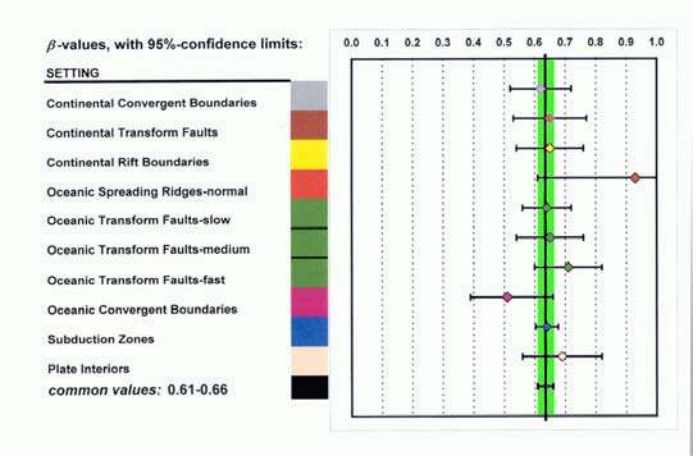

Figure 6. The β-value global distribution. The β-values ($\beta = \frac{2}{3}b$), determined by the maximum likelihood method for eight tectonic provinces (Bird and Kagan, 2004), are shown with their 95% confidence limits.

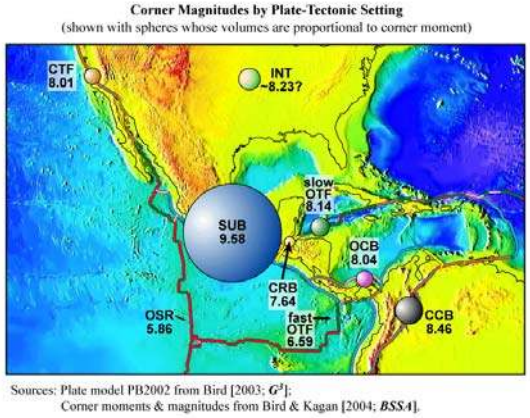

Figure 7. Corner moment (M_c) distribution for eight tectonic provinces (Bird and Kagan, 2004). In this figure the moment magnitude is calculated using $C = 6.05$ in (2). For plate boundary names and abbreviations see Fig. 6.

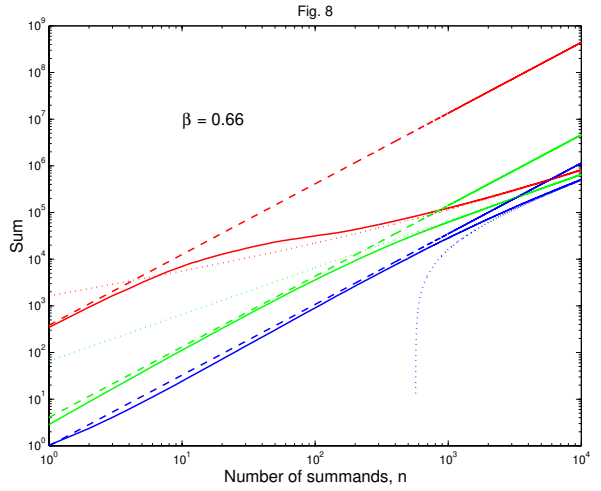

Figure 8. Quantiles for the sum S_n of truncated Pareto variables (upper limit $y = 3.4 \times 10^4$, Eq. 7) and their approximations as functions of the number of summands, n. Two approximations are considered: via the stable distribution, Eq. 11 (dashed lines) and Gaussian, Eq. 16 (dotted lines). Solid lines represent quantiles of simulated Pareto sums. Three thick upper curves are for the 0.98 quantile, three medium thickness middle curves are for the median, and three thin lower curves are for the 0.02 quantile.

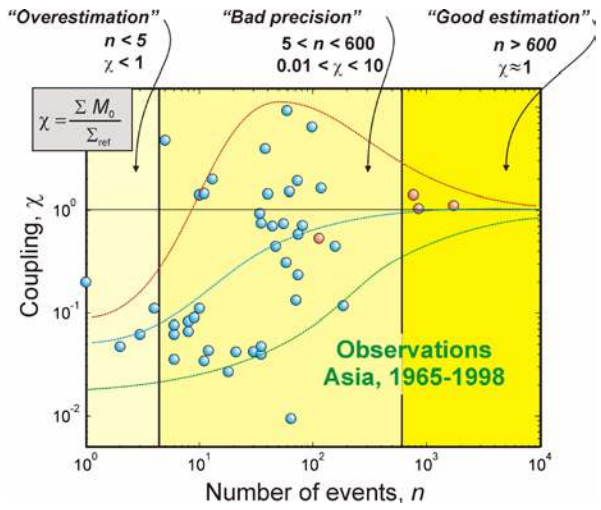

Figure 9. Coupling of tectonic deformation and seismic moment release for 42 subregions (grey disks) of southeast Asia (Holt *et al.*, 2000). Black disks are 4 large regions.

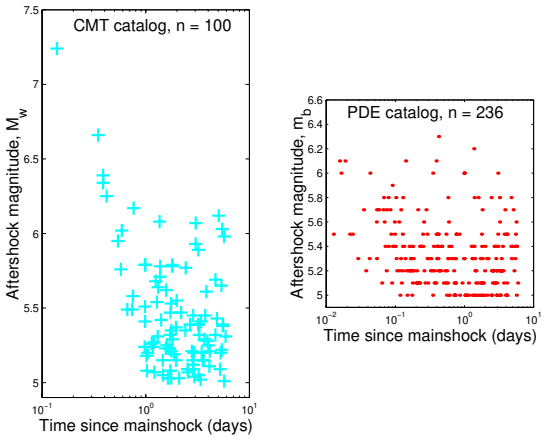

Figure 10. Time-magnitude distribution of the 2004/12/26 $M = 9.1$ Sumatra aftershocks; n is the aftershock number. The Harvard and the PDE earthquake catalogs are used. Events in the six days following the mainshock and between latitude 0°N and 15.0°N and longitude 90.0°E and 100.0°E were selected.

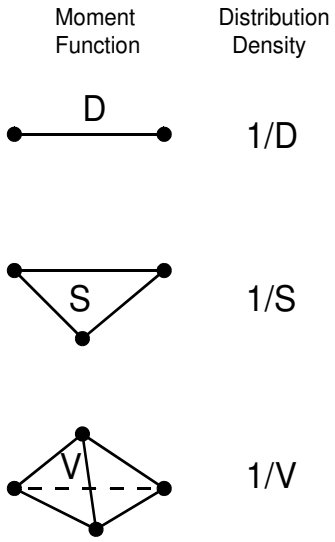

Figure 11. Schematic representation of 2-, 3-, and 4-point spatial moment functions and their suggested approximate densities.

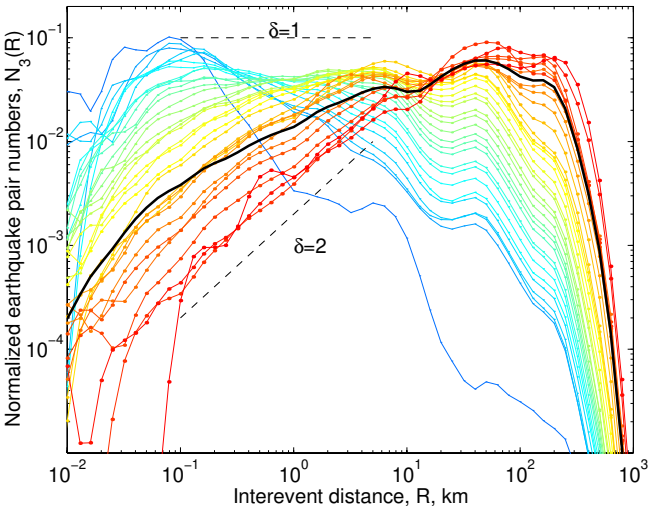

Figure 12. Distribution of distances between hypocenters $N_3(R,t)$ for the Hauksson and Shearer (2005) catalog, using only earthquake pairs with inter-event times in the range $[t,\ 1.25t]$. Time interval t increases from 1.4 minutes (blue curve) to 2500 days (red curve). We divide the earthquake pair number by R so that the horizontal line would correspond to $\delta = 1$. The black line is the function $N_3(R)$ measured for all earthquake pairs; it has a fractal dimension $\hat{\delta} \approx 1.5$ for $0.1 \leq R \leq 5$ km.

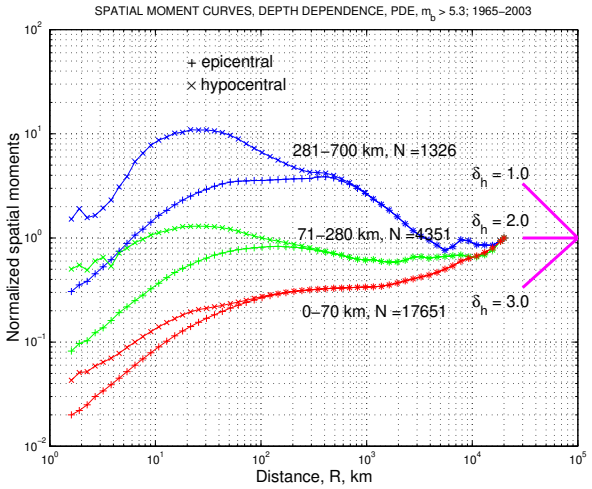

Figure 13. Hypocentral and epicentral spatial moment curves for various depth intervals. A complete PDE catalog (1965-2003) with $m_b \geq 5.3$ is used. In each of the coupled curves the upper curve is for the hypocentral moment and the lower curve for the epicentral. The two upper curves are for the depth interval 281-700 km, the middle ones are for the depth interval 71-280 km, and the lower ones are for the depth interval 0-70 km. Solid lines at the right show a slope of the curves corresponding to the integer values of the hypocentral correlation dimension of δ_h. For the epicentral moment $\delta_s = \delta_h - 1$.

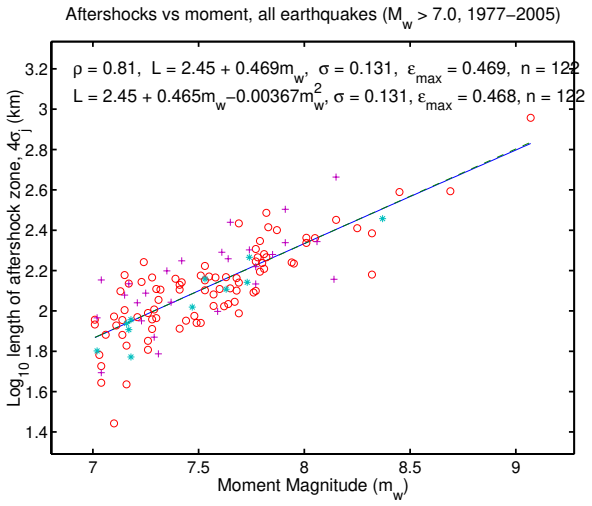

Figure 14. Plot of the log aftershock zone length (L) against moment magnitude (m). Magnitude values are shifted in formulas shown in the plot ($m_r = m - 8.25$). Rupture length is determined using a 1-day aftershock pattern. The diagram shows values of the correlation coefficient (ρ), coefficients for linear (dashed line) and quadratic (solid line) regression, standard (σ) and maximum (ϵ_{max}) errors, and the total number (n) of aftershock sequences. The dashed line is the linear regression; the solid line is a quadratic approximation.
Circle – thrust mainshocks;
Star – normal mainshocks;
Plus – strike-slip mainshocks.

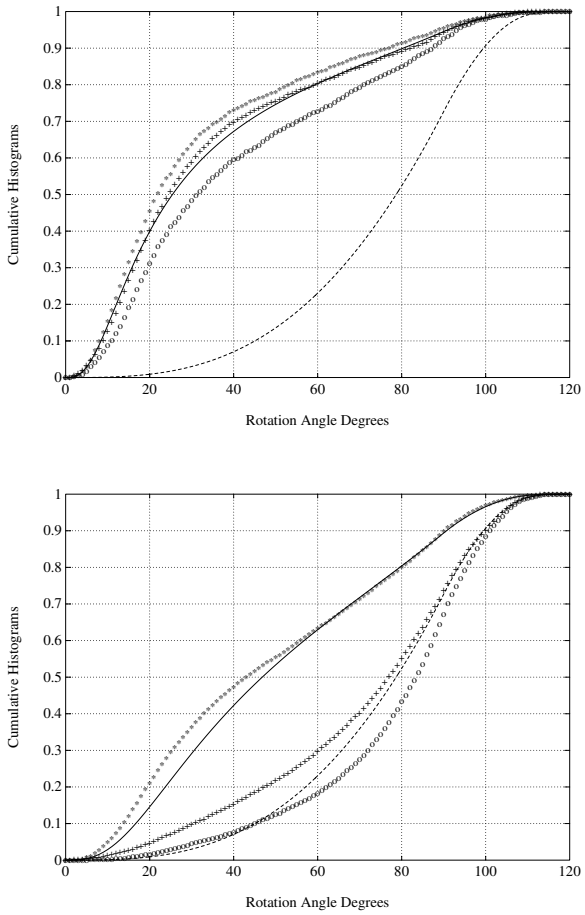

Figure 15. Distributions of rotation angles for pairs of focal mechanisms of shallow earthquakes from the Harvard catalog. Hypocenters are separated by distances: the upper diagram shows 0-50 km; the lower plots 400-500 km; circles indicate hypocenters in 30° cones around the T-axis; plusses hypocenters in 30° cones around the P-axis; stars hypocenters in 30° cones around the N-axis. The solid line is for the Cauchy rotation (25) with $\kappa = 0.1$ (upper plot); and $\kappa = 0.2$ (lower plot). The dashed line represents the random rotation.

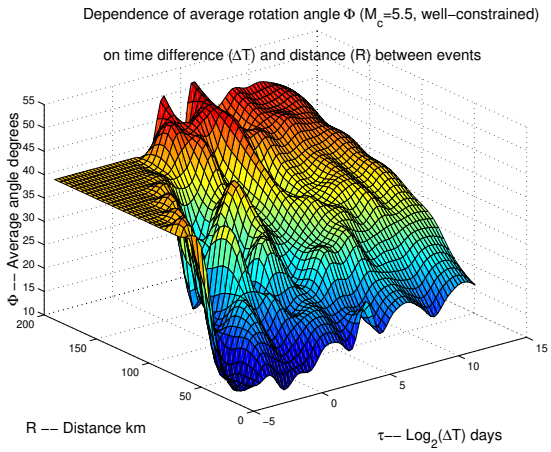

Figure 16. How the rotation angle Ψ depends on time difference and distance between two earthquakes for shallow well-constrained earthquakes with magnitude $m \geq 5.5$ registered in the time period 1977/1/1 – 1999/3/31 in the Harvard CMT catalog.

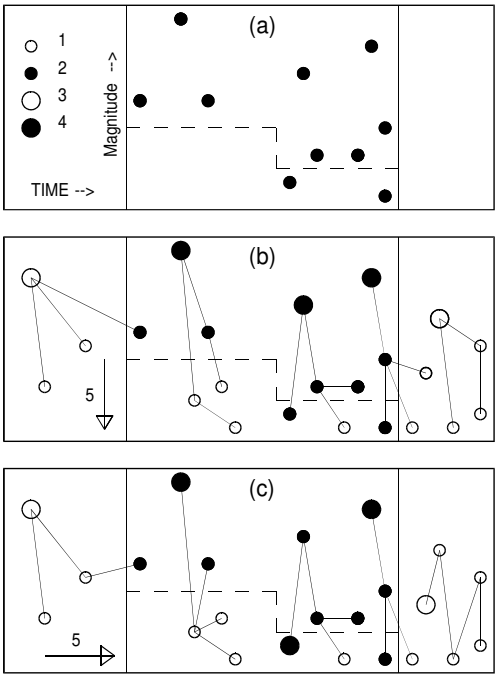

Figure 17. Earthquake branching models: Filled circles (2) indicate observed earthquakes; open circles (1) unobserved, modeled events. The dashed line represents observational magnitude threshold; the earthquake record above the threshold is complete. Many small events are not registered below this threshold. Large circles (3, 4) denote the initial (main) event of a cluster. Arrows (5) indicate the direction of the branching process: down magnitude axis in (b) and down time axis in (c).
(a) Observational data.
(b) Branching-in-moment (magnitude) model.
(c) Branching-in-time model.

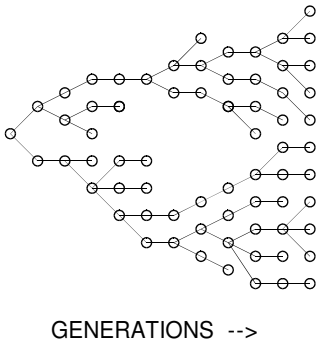

Figure 18. An example of a 'genealogical' tree of a critical branching process. The process starts with one 'particle' of zero generation. Each particle produces the Poissonian number of descendants with the mean equal to one. The development of any particle is independent of all other particles in this or previous generations. Time, position, and orientation of descendant offspring are shown in Fig. 19.

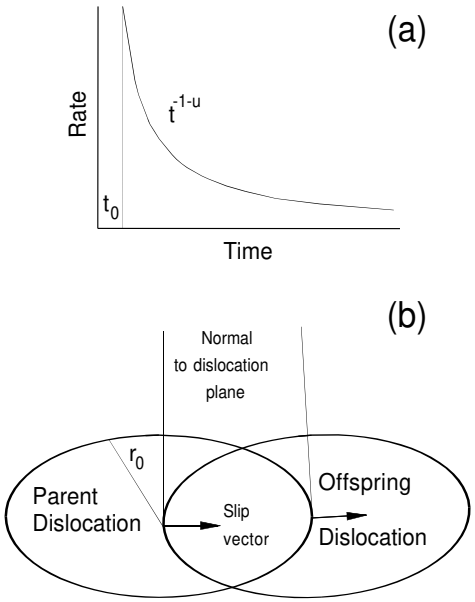

Figure 19. Schematic diagram of fault propagation. (a) Temporal rate of occurrence of dependent shocks. For shallow earthquakes $u \approx 0.5$, t_0 corresponds to rupture time for a dislocation disk in (b). (b) Spatial propagation of a synthetic fault. The initial infinitesimal circular dislocation of radius r_0 gives rise to a secondary event. The center of this dislocation is situated on the boundary of the initial dislocation loop. Solid lines indicate the vector that is normal to the fault plane in both dislocations; arrows show slip vectors. The fault-plane and slip vector of the secondary dislocation rotate following the Cauchy distribution (25). The secondary dislocations can produce new events according to the same law.

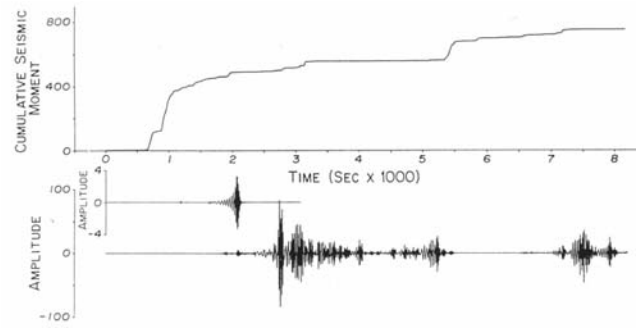

Figure 20. Cumulative event curve (interpreted here as "cumulative seismic moment" for a realisation of the branching process model with (below) an illustration of the filtered signal (using the theoretical seismogram in the middle) from which events and their "seismic moments" can be determined.

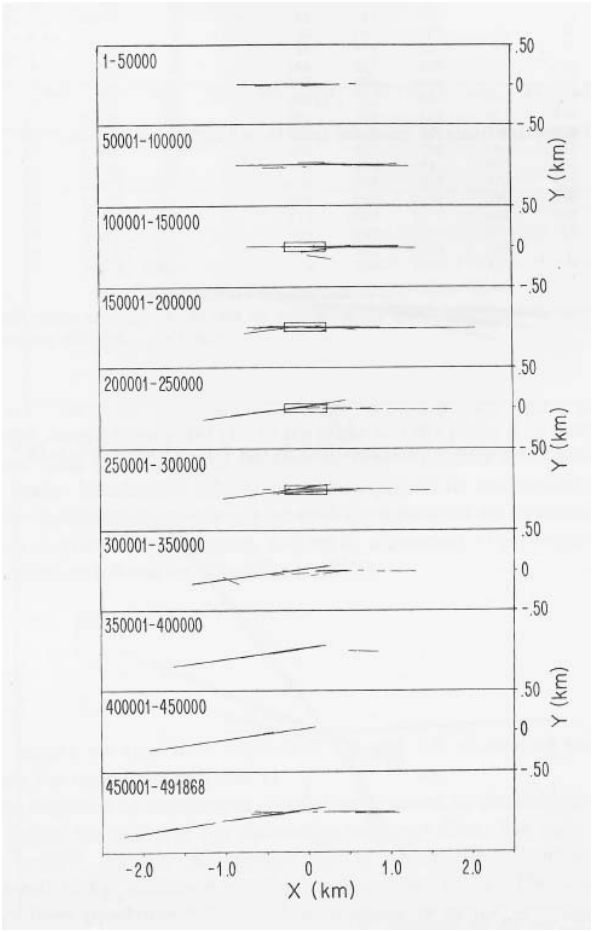

Figure 21. Stages in the evolution of an episode in the branching simulation model: intersection with a fixed plane of the dislocation discs.

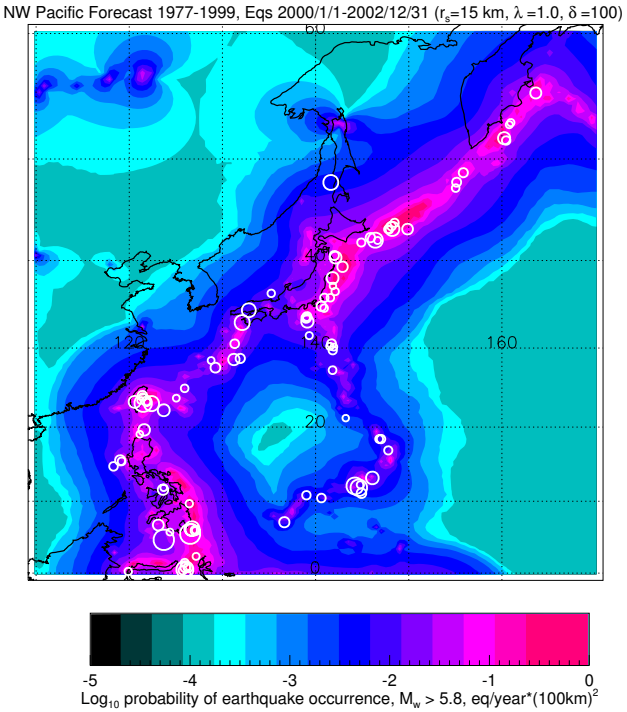

Figure 22. Northwest Pacific long-term seismicity forecast: latitude limits from 0.25°S to 60.25°N, longitude limits from 109.75°E to 170.25°E. Color scale tones show the probability of earthquake occurrence calculated using the Harvard 1977-1999 catalogue; earthquakes from 2000/1/1 to 2002/12/31 are shown in white. We demonstrate forecast effectiveness: displayed earthquakes occurred after a smoothed seismicity forecast was calculated.