Stochastic processes and earthquake occurrence models

Kagan (2006) summarized the discussion in this chapter. Most statistical 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. In Section 5.5 we demonstrate that the focal zone of an earthquake, especially a large one, cannot be regarded as a point. Figure 6.1, for example, shows that earthquake rupture duration needs to be taken into account when very small time intervals are considered. In Fig. 7.7 we show the influence of inter-earthquake time intervals on the spatial structure of earthquake distribution.

A more basic way to study the multidimensional structure of earthquake occurrence 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 process were made by Vere-Jones (1970), Kagan (1973), and Ogata (1988). Many researchers (Console et al. 2003b; Helmstetter and Sornette 2003, 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 the application of statistical methods for earthquake prediction, both long- and short-term. Below we briefly review the available methods for earthquake occurrence analysis and their application in earthquake forecasting. We then discuss how these methods can be improved.

In Section 3.3 we discuss the continuum branching model of earthquake rupture: a 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.

3.1 Earthquake clustering and branching processes

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 the 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 $m_8$. 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 difficulties.

Figure 3.1 displays a sketch of earthquake catalog data and their models 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 (Schorlemmer et al. 2007). 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 (see

![Fig. 3.1](image)

**Fig. 3.1** Earthquake branching models. Open circles indicate (1) unobserved or modeled events; filled circles (2) observed earthquakes. 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. The diagonal solid lines connecting the events represent hypothesized triggering influences. Arrows (5) indicate the direction of the branching process: down magnitude axis in (b) and along time axis in (c). (a) Observational data. (b) Branching-in-moment (magnitude) model. (c) Branching-in-time model. Source: Kagan (2006), Fig 17.
Chapter 4). We must use quantitative statistical methods to describe earthquake clustering. Only an application of stochastic point process theory can provide a robust approach to the problem (see below in this section).

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 almost all are based on the theory of branching processes (Harris 1963; Athreya and Ney 1972). 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 (see also Fig. 3.2). The branching continues for each next generation of events. These multidimensional models are:

(A) Point process branching along the magnitude axis, introduced by Kagan (1973) and shown in Fig. 3.1b.

(B) Point process branching along the time axis (Hawkes 1971; Hawkes and Adamopoulos 1973; Hawkes and Oakes 1974) – called the Hawkes self-exciting process. In earthquake occurrence studies the process is called the CBM (Critical Branching Model), see Kagan and Knopoff (1987b) or the ETAS (Epidemic Type Aftershock Sequence) model (Ogata 1988), respectively (see Fig. 3.1c). Hawkes and Tukey (see discussion section in Kagan 1973) debate the difference between branching in earthquake size and in time.

(C) Continuum-state critical branching process which develops along the time axis (Kagan and Knopoff 1981; Kagan 1982; see Section 3.3).

The first two models (A and B) use the Poisson cluster process to approximate the earthquake occurrence. In these models, earthquake clusters are assumed to follow the Poisson occurrence on temporal or magnitude axes. 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 1973; Kagan and Knopoff 1987b; Ogata 1988, 2004). Bremaud and Massoulie (2001) recently proposed a variant of

Fig. 3.2 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; for the critical branching process the mean number is equal to one. The development of any particle is independent of all the other particles in this or previous generations. Simulations of time, position, and orientation of descendant offspring are shown in Fig. 3.3. Source: Kagan (2006), Fig. 18.
Hawkes’ process with no independent events (immigrants). However, in earthquake catalogs limited in time-span, we need to introduce independent events.

Models (A and B) 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. Math-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 3.3). As we discuss below, other models may have certain advantages in earthquake forecasting and the representation of seismicity. But the 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 search (Kagan 1991b; Ogata 1988, 2004; Chapter 9). 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 by a special procedure.

Most of the statistical models for earthquake occurrence (Kagan and Knopoff 1987b; Ogata 1988; Kagan 1991b) treat earthquake catalogs as a population set, with earthquakes considered as individual entities. As we discuss in Chapter 6, “an individual earthquake” is not a physical entity. Instead it is the result of the interpretation of seismograms and the selection by catalog compilers. Thus, extrapolations of observed features to smaller inter-earthquake time intervals, smaller size earthquakes, etc., may see a model breakdown. Such an approximation of deterioration is caused not by the physical properties of earthquake occurrence, but by the peculiarities of earthquake identification technique and catalogs. Why is this?

### 3.2 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, Chapter 9, in particular Section 9.5.2). 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 random and systematic errors in the earthquake catalogs needs to be performed before each statistical analysis. Otherwise, unless the effect being studied is very strong, the analysis 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 the 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 (see, for instance, Figs. 10.10, 10.11). \( \log(1000) = 6.9 \); 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 the wake of strong events (Chapter 6). They are also incomplete generally for small earthquakes (Section 5.2). 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 sufficiently investigated: 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 the 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 (1991a) 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. They conjectured that short-term clustering results
from stress redistribution in a brittle crust; long-term clustering is most likely due to space-temporal irregularities of 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 occurrence maps, a point forecast needs to be supplemented by an extended source model. In contrast to models (A) and (B), model (C) is in principle defined in a continuum which can simulate realistic, complex rupture processes extended in time, space, and fault orientation.

3.3 Critical continuum-state branching model of earthquake rupture

3.3.1 Time-magnitude simulation

Kagan and Knopoff (1981) proposed the model of the earthquake occurrence based on the continuum-state critical branching random process. This model, which uses only a few parameters, reproduces the most important time-magnitude statistical properties of the earthquake occurrence. Kagan (1982) expanded the model to include a full geometrical description of earthquake occurrence so that the resulting design presents a complete kinematic picture of an earthquake process (Subsection 3.3.2). 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. 3.2. 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. 3.3a), with the probability density function (PDF) similar to Omori’s law

\[ X(t) \propto t^{-1-u}. \]  

(3.1)

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

Kagan and Knopoff (1987a) show that the distribution (3.1) 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 (see more in Chapter 6 and Subsection 5.3.3). 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 the PDF (cf. Eq. 6.10 below)

\[ f_{1/2}(t) = \frac{1}{t} \sqrt{2\pi t} \exp \left( -\frac{1}{2t} \right). \]  

(3.2)
**Fig. 3.3** 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 rotational (angular) Cauchy distribution (Eq. 8.54). The secondary dislocations can produce new events according to the same law (see Fig. 3.2). Source: Kagan (2006), Fig. 19.

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 seismic moment release, so that cumulative plots can be interpreted as analogs to the cumulative moment-release plots used in discussing real earthquake seismograms (Fig. 3.4). However, in Section 5.3 we show that due to the varying orientation of earthquake sources, the cumulative moment may differ from a scalar sum of elementary events.

The intense clustering of the critical process results in this cumulative plot taking on a self-similar, quasi-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. By applying similar criteria to those used to identify real seismic events, Kagan and Knopoff (1981) were able from the time series record to list simulated “events,” each with its own seismic moment or magnitude.

More realistic synthetic seismograms can be obtained by applying the Green’s functions that account for varying position and focal mechanism orientations of each elementary dislocation (see the next subsection). The individual events may then be selected from such seismograms, each event having occurrence time, hypocenter and centroid position and focal mechanism defined by averaging the locations and focal mechanisms of the elementary dislocations comprising the event. In this way a synthetic seismic catalog 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 3.1, Chapter 9) yields similar values of basic parameters describing an earthquake occurrence.

3.3.2 Space-focal mechanism simulation

In the third stage of modeling, the spatial coordinates (location of the dislocation disc center, its orientation, and slip direction) are simulated according to Fig. 3.3b. This model of earthquake rupture incorporated results of 2-, 3-, and 4-point spatial statistical moment studies (Kagan and Knopoff 1980; Kagan 1981a,b, see also Chapter 7) and reproduced the inferred geometrical properties of hypocenter distributions. 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, implying infinitesimal seismic moment value for the loop. 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 (see Eq. 8.54). Most rotations are infinitesimal, though in rare cases large rotations give rise to fault branching seen by inspection of synthetic ruptures (Kagan 1982; see Fig. 3.5). As we explained above, the 3-D rotations are described by the group \( \text{SO}(3) \); 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 dislocation discs with a fixed plane.

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 8.1, this distribution also has a simple, physical explanation:
theoretical arguments (Zolotarev 1986; Kagan 1990) and simulations demonstrate that the stress tensor in the medium with defects follows Cauchy distribution. This stress distribution should result in the rotational Cauchy distribution of earthquake focal mechanisms (ibid.).

The obtained distribution of fault traces looks like actual earthquake fault maps. The spatial statistical moment functions are also qualitatively similar to those in Fig. 7.1: the 1/D and 1/V behaviors have been reproduced, although no rigorous tests have been attempted (Kagan and Vere-Jones 1996).

For extended rupture, like that shown in Fig. 3.5, we can calculate the seismic moment tensor of an earthquake or earthquake sequence

\[ \mathbf{M} = \sum_{i=0}^{N} \mu \left( \prod_{j} \mathbf{q}_{i} \times \cdots \times \mathbf{q}_{j} \times \cdots \times \mathbf{q}_{0} \right) , \]  

(3.3)

where \( \mathbf{q}_{0} \) is a quaternion (see Chapter 8) corresponding to the initial dislocation as in Fig. 3.2. The quaternion product, ending with \( \mathbf{q}_{i} \), describes a combination of

---

**Fig. 3.5** Stages in the evolution of an episode in the branching simulation model: intersection with a fixed XY plane of the synthetic dislocation discs. The range of dislocation numbers for one random realization is shown in each frame. Source: Kagan (1982), Fig. 7.
3-D rotations at the path $\xi_i$ in a branching process leading to the $i$-th dislocation. In a branching process such a path is unique. Each of the quaternion product $q_j$ components follows the rotational Cauchy distribution Eq. (8.54). Thus, the quaternion product in the formula represents the orientation of the $i$-th dislocation. The operator $\mu(\cdot)$ converts the orientation (quaternion) into the seismic moment tensor (Kagan and Jackson 1994, their Appendix; Eq. 8.37).

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

Despite the interest in the model, and the many questions which it raises, it has not (to our knowledge) been the subject of any significant analytical investigations. In particular, it is not clear how key aspects, such as its evident spatial and temporal self-similarity, or the clustering of events subsequently identified as earthquakes, can be deduced from its defining elements. The main purposes of this section are to briefly summarize the model structure, and to indicate some possible directions for pursuing these questions (Kagan and Vere-Jones 1996).