A review of earthquake occurrence models for seismic hazard analysis

Thalia Anagnos

Department of Civil Engineering, San Jose State University, San Jose, CA 95192, USA

Anne S. Kiremidjian

Department of Civil Engineering, Stanford University, Stanford, CA 94305, USA

A large number of probabilistic earthquake occurrence models are currently available for seismic hazard assessment. This paper reviews the basic assumptions of the various models, summarizes their stochastic representations and discusses the parameters needed for applications. While the Poisson model is one of the most commonly used in practice it is limited in its representation of the physical earthquake driving mechanism and in its characterization of distinct seismicity patterns. From comparisons of the various models, it is observed that while the Poisson model may apply to regions characterized by moderate frequent earthquakes, other stochastic representations such as the Markov and semi-Markov models describe the sequences of events more adequately at regions with large infrequent earthquakes. Regions that have unique seismicity patterns such as clustering foreshock-mainshock-aftershock sequences are better represented by other stochastic models. It is found, however, that some of these models are difficult to implement and rather restrictive primarily because they require a considerable amount of additional data for model parameter estimation.

INTRODUCTION

Reliable estimation of the seismic hazard in a region requires the prediction of the size, location and magnitude of future earthquake events. An incomplete understanding of the earthquake phenomenon, however, has led to the development of primarily long-term hazard assessment tools relying on statistical averages of earthquake occurrences without considerations of specific patterns. As knowledge of the geophysical mechanisms that drive earthquake events has increased, so have the corresponding mathematical representations. Over the past two decades, numerous probabilistic models have been developed to depict various aspects of seismic occurrence patterns. The trend has been to introduce models that are specific to a particular region or fault. Some models reflect an apparent memoryless property, while others describe energy release sequences that are time and size dependent. Yet other models account for clustering, cyclicity, aftershock sequences and other patterns in the occurrence data.

At present, the number of probabilistic earthquake occurrence models is so overwhelming that a need exists to examine them and to assess their usefulness and applicability in various regions. The purpose of this paper is to review existing stochastic earthquake occurrence models and discuss their application to seismic hazard analysis. The underlying geophysical and modelling assumptions, the critical parameters and data needed for their determination, and the limitations of the various models are summarized and whenever possible critiqued for their applicability in regional earthquake hazard analyses. There is no simple answer to the question which is the best model. Ultimately a great deal of engineering judgement is involved in decisions about which models provide the best assessment of seismic hazard for a particular region. Table 1 provides a summary of the earthquake occurence models found in the literature and gives a brief comment about unique characteristics of each. While an effort is made to include most available models, the list is not intended to be comprehensive.

Accepted December 1987. Discussion closes May 1988.

STOCHASTIC MODEL FORMULATIONS

The objective in seismic hazard modelling is to obtain long term predictions of the occurrences of seismic events. Most often the prediction is expressed in the form of probabilities of exceedence of a specified earthquake magnitude over a period of time t or as the expected number of such events. Thus, if $N(t)$ represents the number of events in time $(0, t)$ and M defines the size of the events, then that probability is expressed as $P\{N(t)>0$ and $M \geqslant m, (0, t)\}$ and the expected number of events are $E[N(t)>0$ and $M \geqslant m, (0, t)]$. More recently, attention has also focused on the time dependence of earthquakes and representation of that time dependence. For example, information on seismic gaps, characteristic events, time of occurrence and magnitude of the last seismic event can play an important role in hazard computations. Thus, the probabilities of occurrence of at least one event of size M or greater in time (t_1, t_1+t) given that the last event was of

Table 1. Summary of occurrence models used in seismic hazard analyses[a]

Reference	Comments
POISSON MODELS	
(2) Cornell, 1968	Homogeneous Poisson, point-source model, log-linear magnitude-frequency relation, extreme type I distribution for largest annnual event.
(3) Cornell and Vanmarke, 1969	Homogeneous Poisson, point-source model, log-linear magnitude-frequency relation with upper limit on magnitude.
(33) Esteva, 1969	Homogeneous Poisson, point-source model, Bayesian parameters, gamma prior and posterior on occurrence rate.
(4) Milne and Davenport, 1969	Homogeneous Poisson, point-source model. Seismic hazard maps for Canada developed.
(61) Liu and Fagel, 1972	Homogeneous Poisson earthquake occurrence, stochastic model for earthquake ground motion.
(13) Merz and Cornell, 1973a	Homogeneous Poisson, point-source model, quadratic magnitude-frequency relation, extreme type II distribution for largest annual event.
(10) Shah *et al.*, 1975	Homogeneous Poisson, point-source model, bi-linear magnitude-frequency relation.
(5) Der Kiureghian and Ang, 1977	Homogeneous Poisson, fault-rupture model.
(6) Douglas and Ryall, 1977	Homogeneous Poisson, fault-rupture model.
(29) Mortgat and Shah, 1979	Homogeneous Poisson, fault-rupture model, Bayesian parameters.
(7) Blume and Kiremidjian, 1979	Homogeneous Poisson, fault-rupture model, occurrence rate computed from fault dislocation and plate boundary activity.
(62) Guagenti and Scirocco, 1980	Homogeneous Poisson, Bayesian updating to include information from observed precursors.
(15) Kijko and Sellevoll, 1981	Homogeneous Poisson, triple exponential distribution for largest annual event.
(8) Mohammadi and Ang, 1982	Homogeneous Poisson, fault-rupture model, includes probability of fault rupture strike on lifeline link.
(22) Vere-Jones and Ozaki, 1982	Compound Poisson process, cyclic rate for occurrence of earthquake clusters, independent distribution (e.g., geometric) for cluster size.
(11) Araya and Der Kiureghian, 1986	Fault-rupture model including directivity effects.
MARKOV MODELS	
(38) Vere-Jones, 1966	Continuous-time, continuous-state Markov process, aftershocks modelled as sequence of events of decreasing frequency and magnitude.
(39) Knopoff, 1971	Stationary continuous-time, continuous-state Markov process, models stored elastic energy of deformation, main events and aftershocks.
(37) Vagliente, 1973	Two-state Markov chain, states defined as success or failure (occurrence or nonoccurrence of earthquake in specified time interval).
(40) Veneziano and Cornell, 1974	Simulation and Markov Model, temporal and spatial dependence, earthquake occurrence when shear stress equals static friction stress, stress redistribution causes rupture propagation.
(42) Lomnitz-Adler, 1983	Simulation of Markov model to give a simplified representation of the spatial distribution of earthquakes on adjacent faults.
SEMI-MARKOV MODEL	
(43) Patwardhan *et al.*, 1980	Discrete-state semi-Markov process, time between events depends on magnitude of previous and next events.
(46) Anagnos and Kiremidjian, 1984	Semi-Markov model for time predictable earthquake sequences with application to characteristic earthquakes in the Parkfield region.
(48) Anagnos and Kiremidjian, 1985	Discrete-state time-predictable stochastic model with spatial dependence, Weibull distributed interarrival times.
(47) Guagenti and Molina, 1984	Semi-Markov model for time and slip predictable earthquake sequences.
(68) Cornell and Winterstein, 1986	Semi-Markov model for combined time and slip predictable model.
RENEWAL MODELS	
(25) Bender, 1984	Poisson model with two possible occurrence rates depending upon whether fault is in active or inactive cycle. Probabilities of moving from active to inactive or vice versa are constant.
(63) Hagiwara, 1974; (64) Rikitake, 1975	Weibull interarrival times, distribution parameters estimated from strain data, magnitude not considered.
(49) Kameda and Ozaki, 1979	'Double Poisson' renewal model, exponential interarrival times with rate γ in $(0, t_0)$, after t_0, if no earthquake occurs, rate increases to $\nu > \gamma$.
(51) Savy *et al.*, 1980	Weibull interarrival times, fault rupture modelled as series of coherent patches.
(50) Kameda and Takagi, 1981	Renewal process for major faults combined with nonstationary Poisson process for secondary sources, Markov chain for migration between major faults.

Table 1 cont.

Table 1 (cont.)

Reference	Comments
(52) Kiremidjian and Anagnos, 1984	Slip-predictable recurrence, Markov renewal model, interarrival times Weibull distributed.
(65) Grandori *et al.*, 1984	Probability densities of interevent times of aftershocks and longer recurrence events are combined to form the overall interarrival time distribution for the process.
TRIGGER MODELS	
(55) Vere-Jones and Davies, 1966	Homogeneous Poisson trigger events, decay function describes probability of shock occurring in time t after trigger event, magnitude not included.
(56) Shlien and Toksöz, 1970	Compound Poisson, z distribution for cluster size, magnitude not included in model.
(57) Merz and Cornell, 1973b	Homogeneous Poisson trigger events, nonhomogeneous Poisson aftershocks occurrences including spatial distribution of aftershock locations.
(66) Lai, 1977	Homogeneous Poisson trigger events and aftershocks, models magnitudes and occurrence times.
(58,59,60,67) Kagan and Knopoff, 1976, 1977, 1980, 1984	Independent main events, branching along energy axis, defines temporal and spatial distribution of triggered events, extensions include Bayesian analysis and depth dependence.

[a] Revised from Anagnos and Kiremidjian, 1985

size M_0 and there were no events in time $(0, t_1)$, expressed as $P\{N(t)>0 \text{ and } M \geqslant m, (t_1, t_1+t) \mid N(t_1)=0, M_0, (0, t_1)\}$ are also being sought.

In order to evaluate these probabilities various stochastic formulations have been used to model the occurrence process $\{N(t),\ t \geqslant 0\}$ and the associated earthquake size. Many of the current models are based on similar assumptions with some variations in the form of application. In the following discussion, models with similar assumptions are grouped together as either Poisson, Markov, semi-Markov, renewal or trigger models. Example references listed in Table 1 are cited within each group of processes. It should be noted that not all models listed in the table are explicitly referred to in the text primarily because they do not fall within one of the categories discussed in this paper. However, a brief description for each reference is provided in the table.

Poisson models

Earthquake events have long been assumed to occur randomly in time, space and magnitude. A terse look at the plot of earthquake epicenters in any seismic region of the world would reveal a great scatter even when tectonic features are relatively well known. This initial observation has led to the assumption that earthquakes form a stochastically independent sequence of events in time and space. The Poisson process satisfies this independence assumption and, as such, has been used extensively in seismic hazard analysis. The sequence of events forms a memoryless process where the occurrence of a subsequent event does not depend on the time, size or location of the last or any of the preceeding events. Defining again $N(t)$ as the number of events in the interval $(0, t)$ the counting sequence $\{N(t),\ t \geqslant 0\}$ is Poisson provided that

$$P\{N(t+s)-N(t)=k\} = \frac{e^{-\lambda s}(\lambda s)^k}{k!} \quad \text{for } k=0,1,2,\ldots;\ \lambda>0 \tag{1}$$

where λ is the rate of occurrence of events. The interarrival times for the Poisson process $\{T_1, T_2, T_3, \ldots, T_n\}$ must be exponentially distributed with probability density function $f_T(t)$ and cumulative distribution $F_T(t)$ given by

$$f_T(t) = \lambda e^{-\lambda t} \qquad t \geqslant 0,\ \lambda > 0 \tag{2}$$

$$F_T(t) = 1 - e^{-\lambda t} \qquad t \geqslant 0,\ \lambda > 0 \tag{3}$$

The hazard function, or failure rate, is defined as

$$r(t) = \frac{f_T(t)}{[1-F_T(t)]} = \frac{-d}{dt} \log[1-F_T(t)] \tag{4}$$

The hazard rate of the Poisson process is equal to the constant implying that the probability of occurrence of an earthquake in a future small increment of time, Δt, remains constant regardless of the size of the last event or the elapsed time since its occurrence. In physical terms this means that the energy release in a large earthquake does not affect the reservoir of stored energy that is available to produce subsequent earthquakes[1]. Consequently, if a large earthquake occurs at a point in time, the likelihood of another large event occurring in the near future is not changed.

Despite this counter-intuitive model characteristic, the Poisson model is used most widely at the present time. The primary reasons for its popularity are simplicity of model formulation, the small number of parameters to be estimated, the diversity of regions where it can be applied and the relative ease with which hazard due to several sources can be combined. Table 1 identifies the seismic hazard analysis models which are based on the Poisson assumption of earthquake occurrences. The simplest of these models assumes that earthquakes occur independently and release all of their energy at a point. The hazard at a site due to seismicity in a generalized area source or along a line source corresponding to a known fault is obtained by summing contributions from individual points over these sources[2-4].

With the recognition of the importance of the rupture length associated with large earthquakes and its effect on the intensity of ground motion felt at the site, the elementary point source models were modified to include rupture length in the calculation of source-to-site distance[5-7]. Other considerations, such as the intersection of a fault with a lifeline link[8] and the amount of differential displacement along the rupture length of a fault due to a given size event[9] have also been included. Applications to different regions in the world led to the modelling of dipping planes with varying slopes[10]. Most recently, the directivity effects of source rupture have been represented by Araya and Der Kiureghian[11].

In all of the above models the rate of occurrence, λ, is related to the Gutenberg–Richter equation describing the number of earthquake events, $N(m)$, of a given magnitude, m, or greater in time t given by

$$\ln N(m) = \alpha - \beta m \tag{5}$$

where α and β are empirical constants. Thus for a given magnitude of exceedance the frequency of occurrences is

$$\lambda(t) = e^{\alpha - \beta m} \quad \text{for } m_0 \leqslant m \leqslant m_{\max} \tag{6}$$

where m_0 and $m_{\max}$ are the lower and upper bound magnitudes respectively[3,5]. The probability density function of magnitudes, $f_M(m)$, is most often assumed to be a truncated exponential given by

$$f_M(m) = \frac{\beta\, e^{-\beta m}}{[e^{-\beta m_0} - e^{-\beta m_{\max}})]} \tag{7}$$

The lower magnitude arises from practical considerations. For example, earthquakes with Richter magnitudes smaller than 3.0 are not known to cause any structural damage and are thus excluded from the analysis. The upper magnitude is governed by the maximum earthquake capacity of a fault. A considerable amount of controversy exists, however, regarding both the validity of the Gutenberg–Richter relationship and the selection of the upper bound magnitude.

Investigations of earthquake occurrence data for various regions of the world have shown that the log-linear frequency law fits the data relatively well in the mid-size magnitude range and usually very poorly in the upper and lower magnitude tails. Explanations can be offered for the lack of fit in the lower magnitude range. Records may be missing from the data since instruments have been placed at many locations only within the last few decades. This presents the additional problems of nonhomogeneous distribution of data over time and the presence of intervals with zero observations in the large magnitude ranges. The former problem has been treated in various ways including the approach proposed by Weichert[12] which treats the unequal observation periods for different magnitudes separately when deriving the frequency law. In order to overcome the latter problem of overestimating the frequency of large magnitude events and to represent more realistically the nonlinear character of the cumulative frequency *versus* magnitude data, the following were proposed: a quadratic relationship[13], a bi-linear one[10], and a modified form of the Gutenberg–Richter relation based on seismic moment and fault slip[14]. The third formulation produces a triple exponential function of the largest annual event[15]. Another form of the recurrence relationship is based on the hypothesis that individual faults tend to generate earthquakes of approximately the same size within half a magnitude. The frequency distribution of these 'characteristic events' is nonlinear, consisting of an exponential region and a region with zero slope[16]. Weichert's approach and the bi-linear function do not change the overall model formulation with the bi-linear function being applied extensively.

Difficulties with estimating the upper magnitude bound has led to considerable discussion among scientists and engineers. Most often the upper bound magnitude is arrived at by a combination of methods[17]. One method is to review past seismicity to determine the largest earthquake event ever recorded in the study region. If available, empirical magnitude *versus* rupture length relationships (or magnitude *versus* fault rupture area relationships) may be combined with information about fault type and fault dimensions, thus essentially representing the physical limitations of the fault zone in the estimate of the maximum magnitude. Most often these relationships are represented by log–linear equations. In addition, geologic evidence of large magnitude events from trenching and carbon dating can be combined with relationships of displacement *versus* magnitude. Finally, the slip rate may be correlated with the maximum magnitude on a fault. Since none of these methods are reliable, the uncertainty associated with the upper bound magnitude can be considerable. Depending on the assumptions that are made, hazard probabilities may be very sensitive to the upper bound magnitude. In addition, the larger size events can be the most significant contributors to the seismic hazard in a region, especially when the source is a considerable distance away from the study site[16]. One approach to this problem has been to account for the uncertainty of the upper bound magnitude by treating it as a random variable[18-20].

Validation of the Poisson assumptions is in general very difficult due to the sparsity of data in most regions of the world and the lack of understanding of the geophysical earthquake generating processes. The representation of the sequence of events as a memoryless process has been shown to be adequate for sequences of main events in a certain catalog in Southern California[21]. However, this may not be a good representation of data for other geographical locations. In addition, foreshock and aftershock sequences cannot be represented by the homogeneous Poisson model because they appear as clusters in the data. A compound Poisson process may be more suitable for representing clustering of events[22]. In such a model, clusters containing Y_n events are assumed to be independent identically distributed and to occur as Poisson sequences. The number of events in a cluster, Y_n, are independent identically distributed random variables and are independent of the Poisson process. Then the total number of earthquake events in time $(0, t)$ is described by the compound Poisson process, $\{X(t), t \geqslant 0\}$ given by

$$X(t) = \sum_{n=0}^{n=N(t)} Y_n \tag{8}$$

This model, however, requires that the distribution of Y_n be known in addition to the Poisson process parameters.

The assumption of a constant rate of earthquake occurrence, and hence the use of the homogeneous

Poisson model, has often been questioned. An examination of earthquakes in China since 1177 B.C.[23,24] showed the rate of occurrence to increase and decrease periodically in a cycle of about 300 years. It was also shown that while the stationary Poisson model can still be used for seismic hazard estimation when there is periodicity in the seismicity rate, the occurrence rate for the model should be estimated with caution, taking into account recent trends. Bender[25] attempts to represent the cyclic pattern of earthquake activity by the use of a simple renewal model in which the rate of the Poisson process alternates between two values. However, the periodic effect associated with certain earthquake catalogs has been shown not to be significant when clustering is taken properly in account[22].

Estimation of the occurrence rate of the Poisson model using historic catalogues presents another difficulty because of incompleteness and biases in the data. Various methods have been employed to obtain estimates of the occurrence rate incorporating instrumental, historical, geological and subjective information. The occurrence rate, which is related to the Gutenberg–Richter relationship (equation (5)), is most often obtained by simple linear regression techniques using instrumentally recorded data[2,3,10]. Other methods for estimating the parameters of equation (5) include the maximim likelihood[12,22,26] and maximum entropy[27,28]. Often, instrumentally recorded data are augmented by historic accounts[23,29] or geologic information such as slip rate or moment rate[7,30–32]. Bayesian methods are frequently used to combine the results of several of these techniques or to incorporate subjective information[29,32–33].

Despite its many limitations, the Poisson model is extensively used in seismic hazard assessment and the development of seismic hazard maps. Methods for improving the parameter estimates and to address parameter uncertainty have also been proposed. However, one is still faced with the problem of describing specific patterns that are apparent in earthquake catalogues and geophysical earthquake generating mechanisms. As will be discussed in the next section, the consequence of representing earthquakes as a memoryless sequence of events is that the hazard at a site can be overestimated or underestimated when occurrences are time, magnitude or location dependent.

Markov and semi-Markov models

Modelling of faults in laboratory experiments[35,36] have shown that as two sides of a fault move in opposite directions they remain locked until sufficient shear stress builds up, then slip occurs and the fault subsequently locks again. Thus, a sequence of earthquakes can be represented by a process of strain accumulation interrupted by sudden releases. This laboratory representation of the elastic rebound theory suggests that the times of occurrence and magnitudes of a sequence of earthquakes on a given source may not be stochastically independent. In the attempt to overcome the modelling problems associated with the memoryless property of the Poisson process, other stochastic models have been considered. Markov models and semi-Markov models are useful in describing a unique type of dependence in a sequence of events. For these models a state space $E=\{1,2,3,\ldots,N\}$ is defined such that the states may correspond to various fault stress levels, energy release levels or magnitudes of earthquake events. The process $\{X(t), t\geqslant 0\}$ describes the visits to these states and is said to be a Markov process provided that

$$P\{X(t+s)=j|X(h)=i, 0<h\leqslant t\}=P\{X(t+s)=j|X(t)=i\}$$
$$\text{for } t,\ s>0 \text{ and all } i,j\in E \quad (9)$$

Thus the probability of being in some state j at a future time $t+s$ is deduced from knowledge of the state i at an earlier time t and is independent of the history of the process up to time t. The transition probabilities, $P\{X(t+s)=j\,|\,X(t)=i\}$, completely determine the Markov process. The Markov process assumes that the holding time, $h_{ij}(t)$, defined as the probability that the process stays in state i for a time period t before it moves to state j is exponentially distributed with parameters conditional on state i. In comparison, the semi-Markov model is not restricted to exponentially distributed holding times. In addition, for the semi-Markov process the holding times in a given state are identically distributed conditional on both the current state and the next state thus providing greater flexibility in modelling. In most semi-Markov earthquake occurrence models, parameters and distributions have been chosen to assure increasing hazard rates for the holding time distributions (e.g., Weibull, gamma) implying that the probability of an earthquake occurring in the near future increases with the time since the last event. For example, the hazard for the Weibull distribution given as $r(t)=\lambda v t^{v-1}\exp\{-\lambda t^{v}\}$ has an increasing hazard rate for parameters $v>1$ and $\lambda>0$. The increasing hazard rate captures some of the characteristics of stress build-up and release. A disadvantage of these models is the large amount of data needed for estimating parameters. However, a major advantage is that information on seismic gaps can be included and hazard forecasts can be updated to reflect the occurrence of the most recent event. The following are some examples of Markovian models used in seismic hazard computations.

Vagliente[37] represents the seismic process as a two state Markov chain with the states defined as occurrence and nonoccurrence of earthquakes in a specified time interval. By modelling the energy or stress accumulation and release, Markov models have also been used to describe aftershock sequences[38] as well as sequences of main events followed by aftershocks[39,40]. Veneziano and Cornell[40] consider the spatial redistribution of stress and consequently the spatial dependence between seismic events. Uribe-Carvajal and Nyland[41] have extended Knopoff's model[39] by including a finite element model of the study region to describe the spatial distribution of events. Lomnitz-Adler[42] simulates what could be interpreted as a Markovian model to represent a simplified spatial distribution of earthquakes on a series of faults by including the accumulation and release of stress on adjacent blocks. In many of these formulations the energy or stress levels constitute the states of the process. Visits from one state to another represent the occurrence of earthquakes and are described by the transition probabilities. These probabilities, however, are difficult to obtain from the very limited data and thus the models have not been applied to any particular region.

Semi-Markov models have been used to represent the sequence of large magnitude events and to characterize spatial and temporal seismic gaps found in the earthquake occurrence catalogues[43,44]. In order to develop the holding time distributions and state transition probabilities, Patwardhan *et al.*[43] and Cluff *et al.*[44] use historical and geologic data. However, they rely mostly on subjective input to develop these probabilities to complement the small amount of available data. Bayesian techniques are used to combine the various sources of information. Several other semi-Markov models based on the time- and slip-predictable hypotheses[45] have been developed[46,47]. Anagnos and Kiremidjian[46] model the mechanism of strain accumulation and release on a specified seismic source. Conditional holding times for this model are considered to be Weibull distributed with an increasing hazard rate in the range of values of interest. The state transition probabilities are derived by considering the stress accumulation (or slip) rate and the amount of stress release (or seismic slip) associated with various magnitude levels. The conditional holding time distributions and the state transition probabilities are developed from historical data as well as from geophysical considerations of the fault mechanism not employed in previous models.

This particular model has been extended to include spatial patterns of earthquake occurrence[48]. The state space of the spatial-temporal model is a set $E=\{1,2,\ldots,N\}$ describing the stress level and location on a fault. The set $\{Y_n: n\geqslant 0\}$ are independent identically distributed random variables assuming values in E and $\{T_n: n\geqslant 0\}$ are nonnegative random variables such that $0<T_1<T_2,\ldots$ The stochastic process $\{(Y_n, T_n): n\geqslant 0\}$ is a Markov renewal process. For this model Y_n represents a pair (S_n, L_n) where S_n is the stress level at the epicenter immediately after the earthquake event and L_n is the location of the epicenter that event along the fault. The process also keeps track of the maximum stress level immediately after an event denoted by S_n^+. A number of simplifying assumptions are needed for the spatially dependent model in order to develop the transition probabilities and the holding time distribution. Further detail on the development of this model is given in Ref. 48. Simulations of earthquake sequences using this model have demonstrated the effect of spatial dependence. The relatively large number of states needed to represent the spatial extent of the seismic source and the magnitude range significant for hazard assessment may pose, however, computational difficulties.

From applications of the time-predictable formulation to the plate boundary along the San Andreas fault it was observed that hazard forecasts could differ significantly from hazard estimates obtained by the commonly used Poisson model when a long time has elapsed since the last major sesmic event. The observation can have serious implications for the estimation of seismic design parameters particularly for critical facilities. These models are applicable primarily to plate boundaries and regions characterized by large infrequent events. A difficulty with all of these models can be in defining an initial condition, that is knowing the size and the time of occurrence of the last seismic event. Verification of the Markov and semi-Markov models temporal and spatial dependence has been limited to testing by simulation[40,48].

Renewal models

Another group of models that attempt to represent specific patterns in the earthquake catalogues or physical characteristics of a region are those based on renewal theory. In these types of models, the process restarts after the occurrence of each event. Therefore, the interarrival times are independent identically distributed random variables. The Poisson process is one of the simplest examples of a renewal process. Kameda and Ozaki[49] describe a particular type of temporal dependence, in which the rate of a Poisson process can have two distinct values depending upon the time since the last event. Kameda and Takagi[50] combine a renewal model for major offshore seismic sources with a nonstationary Poisson-type model for secondary sources to represent interaction between faults. Another, Poisson-type renewal model is developed by Savy *et al.*[51] to reflect strain build-up that takes place in the time interval between earthquake occurrences. In their model the occurrence rate increases with time and returns to its original value after each earthquake. Many of these models have a limited use because they represent a specific catalogue or they are idealizations of particular patterns of earthquake sequences. Their verification has been performed only for the earthquake catalogues specfic to the given geographic location.

Kiremidjian and Anagnos[52] use a Markov renewal model based on the slip-predictable hypothesis to describe the dependence of the size of future earthquakes on the elapsed time since the last event. Interarrival times are assumed to be Weibull distributed with an increasing hazard rate for the range of parameter values of interest in applications. The advantages of this model similar to the semi-Markovian models are that temporal seismic gaps are well represented and the additional information needed for parameter estimation can be obtained from considerations of the geophysical process at a fault. Applications of the model, however, are usually limited to regions characterized by large magnitude infrequent events or interplate boundary faults. Parameter estimation can be obtained from considerations of the geophysical process at a fault[53,54].

Trigger and branching models

As was discussed earlier, it is difficult to either accept or reject the Poisson assumptions for modelling earthquake occurrences. Since statistical analysis of earthquake frequencies have revealed correlations between the number of events in successive time intervals and values of the Poisson index, trigger models were proposed to account for variations from Poisson-like behaviour[55-57]. In these types of models the trigger events (or events initiating the group of shocks) occur at the instants of s simple Poisson processes with a constant rate. The decay of events within a cluster is described by the function $\mu(t)$ with the properties given by Ref. 55 as

$$\mu(t)=0 \qquad \text{for } t<0 \tag{10a}$$

$$\mu(t)\geqslant 0 \qquad \text{for } t\geqslant 0 \tag{10b}$$

$$\int_0^{\infty} \mu(t)\,\mathrm{d}t=1 \tag{10c}$$

The decay functions, $\mu(t)$, are most often considered to be exponential or inverse power functions of time. The conditional probability that an event will occur in a small interval of time $(t+s, t+s+\mathrm{d}s)$, given that a trigger event occurred at time t, is assumed to be independent of t and equal to $A\mu(s)$. The quantity A characterizes the size of the group. It is assumed to be a random variable with a distribution $F_A(a)$ having a finite mean and variance. The separate trigger events generate independent sequences of shocks and the random variables A, associated with distinct trigger events, are also independent[55]. The distribution of the number of shocks in a single trigger event is given by

$$p_n = \int_0^\infty \alpha^n \, e^{-\alpha} \, dF_A(\alpha)/n! \qquad (11)$$

This type of model is used to represent catalogues of earthquakes as well as main events followed by aftershocks. Trigger processes have been employed[55,56] to include foreshock and aftershock clusters in the hazard analysis estimates. Aftershock clusters in the Merz and Cornell model[57] occur as nonhomogeneous Poisson processes. The spatial distribution of aftershock epicenters is considered in this model by confining them within a limited area around the main shock. It is important to understand the value of these types of models because aftershocks may be a significant contributor to the hazard at a site.

Kagan and Knopoff[58-60] propose another approach for representing the space, time and magnitude relationship among earthquakes through the use of a branching renewal model. The particular structure of these models resembles that of the trigger model. In addition, in their formulation a statistical method which couples the maximum likelihood method and second order moment calculations has been used to determine the space-time-magnitude correlation. From the analysis of worldwide data they have demonstrated the migration of epicenters, the existence of gaps following large events and the occurrence of foreshocks and aftershocks associated with a main event.

Estimation of the parameters of trigger and branching models is even more difficult than for many of the models discussed earlier because of the greater detail needed in the earthquake data sequences. Their verification has also been limited by the lack of sufficient information. Due to their modelling limitations, these models have not been applied to any particular region.

SUMMARY OF RESEARCH NEEDS

The multitude of models described in this paper were grouped according to their stochastic formulations. The main distinction, however, which separates these models is their potential for widespread application. While the homogenous Poisson models have been the most popular for seismic hazard estimation and have been extensively applied to various regions in the world, they have been shown to be limited in their representation of the geophysical earthquake driving mechanisms. Their basic assumptions of temporal and spatial independence and homogeneity have been shown to be a major limitation when characterizing large infrequent events or when modelling seismic hazard in regions with specific occurrence patterns. Their validity has been demonstrated for only a few regions. The popularity of the Poisson model is due to its simplicity and relative ease of application. Because these types of models have been studied so extensively, however, the main difficulties with their parameter estimation and sensitivity to parameter uncertainties have been revealed and alternate solutions have been proposed. With the incorporation of these addended formulations, the simplicity of the Poisson model has been somewhat lost leading to rather involved techniques for parameter estimation or to more complex forms of the model itself.

Most of the remaining models attempt to represent either a specific pattern observed in a catalogue or a pattern that is built on the geophysical assumptions of the earthquake occurrence mechanism. Among these, the models based on geophysical considerations appear to be the most promising. For example, temporal variations in the occurrence sequences such as those associated with seismic gaps are best represented through time dependent models. It has been shown that the time and space dependent models would provide very different hazard estimates from the Poisson model depending, for example, on the duration of the gap (i.e., the elapsed time since the last event) and the average interarrival time. In particular, if the elapsed time since the last major earthquake is of the order of the mean interarrival time, the time-dependent models will predict a considerably greater hazard than the Poisson model. This observation contradicts the common belief that the Poisson model provides the most conservative estimates of the hazard. Contrary to many of the other non-Poissonian models, estimation of the parameters of the stochastic time- and slip-predictable models do not rely exclusively on earthquake occurrence data. The type of information needed for these models is most often obtained from geophysical considerations. For example, average strain rates or moment rates can be used to estimate the stress accumulation rate along a fault. The time dependent models are still untested for their sensitivity to variabilities in the data and in general are restricted to large infrequent events.

The remaining stochastic earthquake occurrence models are found to be very limited in their applications. The large amount of data needed for the development of their parameters as well as the constraints of the underlying assumptions, the complexity of model formulations and the extensive computational effort needed, make them too restrictive for widespread applications.

In the dilemma as to which model should be used for hazard estimation in a particular region, it appears from this review that initial estimates of that hazard are best obtained by the homogenous Poisson model. If, however, specific features can be identified for the region and the type of information is available to apply one of the other models that model should be further explored and used for seismic hazard analysis.

New models will inevitably be developed as the geophysical earthquake generating mechanism becomes better understood. Future research effort should be concentrated on developing models which attempt to reflect the physical mechanism of seismic events even if sufficient data for developing the parameters of these models are not available. The uncertainty in these models will reflect the systematic error due to the lack of data. This uncertainty can be reduced over time as these data become more available. In comparison, errors resulting from physically vague models are difficult to assess and reduce. Particular attention should be given to the various rupture processes along a seismic fault and the transfer of energy from the rupture zone to adjacent locations. Stochastic models representing these mechanical models can then be developed to represent the sequence of events along a seismotectonic feature. It should be recognized, however, that the simple models, such as the Poisson models should not be abandoned since they have been shown to be sufficiently good predictors of smaller size earthquakes. Thus, a combination of models, one for large events and one for smaller size events may be appropriate. The specific division is difficult to generalize. However, for a given fault the division into large and small events for model selection purposes would depend on the physical conditions. For example, for events which do not have associated significant fault displacements or rupture zones a Poisson model would seem appropriate. For the events with identifiable rupture zones a model reflecting this rupture potential should be sought.

The greatest difficulty in applying all of the discussed models is the lack of sufficient data. However, with the increased understanding of the physical process and the development of physically more accurate models, the types of data that are needed will also be identified. Only after such data are properly characterized can they be gathered and analysed. The degree of sophistication built into a model and the amount of effort put into collecting the appropriate amount of data, however, should always be consistent with any subsequent analysis and use of results.

REFERENCES

1 Aki, K. Some Problems in Statistical Seismology, *Zisin*, 1956, **8**, 205–228

2 Cornell, C. A. Engineering Seismic Risk Analysis, *Bull. Seism. Soc. Amer.*, 1968, **58**, 1583–1606

3 Cornell, C. A. and Vanmarke, E. H. The Major Influences on Seismic Risk, *Proc. World Conf. Eq. Eng., 4th*, Santiago, Chile, 1969, 69–83

4 Milne, W. G. and Davenport, A. G. Distribution of Earthquake Risk in Canada, *Bull. Seism. Soc. Amer.*, 1969, **59**, 729–754

5 Der Kiureghian, A. and Ang, A. H.-S. A Fault Rupture Model for Seismic Risk Analysis, *Bull. Seism. Soc. Amer.*, 1977, **67**, 1173–1194

6 Douglas, B. M. and Ryall, A. Seismic Risk in Linear Source Regions, with Applications to the San Andreas Fault, *Bull. Seism. Soc. Amer.*, 1977, **67**, 233–241

7 Blume, J. A. and Kiremidjian, A. S. Probabilistic Procedures for Peak Ground Motions, *J. Struc. Div., ASCE*, 1979, **105**, 2293–2311

8 Mohammadi, J. and Ang, A. H.-S. Seismic Hazard Analysis of Lifelines, *J. Struc. Div., ASCE*, 1982, **108**, 1232–2349

9 Kiremidjian, A. S. Reliability of Structures Subjected to Differential Fault Slip, *Int. J. Eq. Engr. Str. Dyn.*, 1984, **12**(5), 603–618

10 Shah, H. C., Mortgat, C. P., Kiremidjian, A. and Zsutty, T. C. A Study of Seismic Risk for Nicaragua, Report 11, John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA, 1975

11 Araya, R. and Der Kiureghian, A. Seismic Hazard Analysis Including Source Directivity Effects, *Proc. Third US Nat. Conf. Eq. Eng.*, Charleston, South Carolina, 1986, **1**, 269–280

12 Weichert, D. Estimation of the Earthquake Recurrence Parameters for Unequal Observation Periods for Different Magnitudes, *Bull. Seism. Soc. Am.*, 1980, **70**(4), 1337–1347

13 Merz, H. A. and Cornell, C. A. Seismic Risk Analysis Based on a Quadratic Magnitude-Frequency Law, *Bull. Seism. Soc. Amer.*, 1973a, **63**, 1999–2006

14 Lomnitz-Adler, J. and Lomnitz, C. A Modified Form of the Gutenberg–Richter Magnitude-Frequency Relation, *Bull. Seism. Soc. Amer.*, 1979, **69**, 1209–1214

15 Kijko, A. and Sellevoll, M. A. Triple Exponential Distribution, A Modified Model for the Occurrence of Large Earthquakes, *Bull. Seism. Soc. Amer.*, 1981, **71**, 2097–2101

16 Youngs, R. R. and Coppersmith, K. J. Fault Slip Rates and Earthquake Occurrence Relationships, *Bull. Seis. Soc. Amer.*, 1985, **75**(4), 939–964

17 Schwartz, D. P., Coppersmith, K. J. and Swan III, F. H. Methods for Estimating Maximum Earthquake Magnitude, *Eighth World Conf. Eg. Eng.*, 1984, **I**, 279–285

18 McGuire, R. K. Effects of Uncertainty in Seismicity on Estimates of Seismic Hazard for the East Coast of the United States, *Bull. Seism. Soc. Amer.*, 1977, **67**, 827–848

19 Kijko, A. Is it Necessary to Construct Empirical Distribution of Maximum Earthquake Magnitude?, *BSSA*, 1984, **74**, 339–347

20 Kulkarni, R. B., Youngs, R. K. and Coppersmith, K. J. Assessment of Confidence Intervals for Results of Seismic Hazard Analysis, *Proc. World Conf. Eq. Eng., 8th*, San Francisco, California, 1984, **1**, 263–270

21 Gardner, J. K. and Knopoff, L. Is the Sequence of Earthquakes in Southern California, with Aftershocks Removed, Poissonian?, *Bull. Seism. Soc. Amer.*, 1974, **64**, 1363–1367

22 Vere-Jones, D. and Ozaki, T. Some Examples of Statistical Estimation Applied to Earthquake Data, Cyclic Poisson and Self-Exciting Models, *Ann. Inst. Stat. Math.*, 1982, **34**, Part B, 189–207

23 McGuire, R. K. Adequacy of Simple Probability Models for Calculating Felt-Shaking Hazard, Using the Chinese Earthquake Catalog, *Bull. Seism. Soc. Amer.*, 1979, **69**, 877–892

24 McGuire, R. K. and Barnhard, T. P. Effects of Temporal Variations in Seismicity on Seismic Hazard, *Bull. Seism. Soc. Amer.*, 1981, **71**, 321–334

25 Bender, B. A Two-State Poisson Model for Seismic Hazard Estimation, *Bull. Seism. Soc. Amer.*, 1984, **74**, 1463–1468

26 Bender, B. Maximum Likelihood Estimation of b Values for Magnitude Grouped Data, *Bull. Seism. Soc. Amer.*, 1983, **73**, 831–851

27 Berrill, J. B. and Davis, R. O. Maximum Entropy and the Magnitude Distribution, *Bull. Seism. Soc. Amer.*, 1980, **70**, 1823–1831

28 Dong, W. M., Bao, A. B. and Shah, H. C. Use of Maximum Entropy Principle in Earthquake Recurrence Relationships, *Bull. Seism. Soc. Amer.*, 1984a, **74**(2), 725–737

29 Mortgat, C. P. and Shah, H. C. A Bayesian Model for Seismic Hazard Mapping, *Bull. Seism. Soc. Amer.*, 1979, **69**, 1237–1251

30 Anderson, J. G. Estimating the Seismicity from Geological Structure for Seismic Risk Studies, *Bull. Seism. Soc. Am.*, 1979, **69**, 135–158

31 Molnar, P. Earthquake Recurrence Intervals and Plate Tectonics, *Bull. Seism. Soc. Amer.*, 1979, **69**, 115–133

32 Dong, W., Shah, H. C., Bao, A. and Mortgat, C. P. Utilization of Geophysical Information in Bayesian Seismic Hazard Model, *Int. J. Soil Dyn. Eq. Eng.*, 1984b, **3**, 103–111

33 Esteva, L. Seismicity Prediction: A Bayesian Approach, *Proc. World Conf. Eq. Eng., 4th*, Santiago, Chile, 1969, **1**, 172–184

34 Campbell, K. W. Bayesian Analysis of Extreme Earthquake Occurrences, Part 1. Probabilistic Hazard Model, *Bull. Seism. Soc. Amer.*, 1982, **72**, 1689–1705

35 Byerlee, J. D. and Brace, W. F. Stick-Slip Stable Sliding and Earthquakes – Effect of Rock Type, Pressure, Strain Rate, and Stiffness, *J. Geoph. Res.*, 1968, **73**, 6031–6037

36 Brune, J. N. Earthquake Modeling by Stick-Slip Along Precut Surfaces in Stressed Foam Rubber, *Bull. Seism. Soc. Amer.*, 1973, **63**, 2105–2119

37 Vagliente, V. N. Forecasting Risk Inherent in Earthquake Resistant Design, Technical Report 174, Department of Civil Engineering, Stanford University, Stanford, CA, 1973

38 Vere-Jones, D. A Markov Model for Aftershock Occurrence, *Pure Appl. Geoph.*, 1966, **64**, 31–42

39 Knopoff, L. A Stochastic Model for the Occurrence of Main-Sequence Earthquakes, *Rev. Geop. Space Phys.*, 1971, **9**, 175–188

40 Veneziano, D. and Cornell, C. A. Earthquake Models with Spatial and Temporal Memory for Engineering Seismic Risk Analysis, Report R74-18, Department of Civil Engineering, Massachusetts Ins. of Tech., Cambridge, MA, 1974

41 Uribe-Carvajal, A. and Nyland, E. Stochastic Finite Element Calculation of Seismic Risk, *Tectonophysics*, 1985, **117**, 193–203

42 Lomitz-Adler, J. A Statistical Model of the Earthquake Process, *Bull. Seism. Soc. Amer.*, 1983, **73**, 853–862

43 Patwardhan, A. S., Kulkarni, R. B. and Tocher, D. A Semi-Markov Model for Characterizing Recurrence of Great

Earthquakes, *Bull. Seism. Soc. Amer.*, 1980, **70**, 323–347

44 Cluff, L. S., Patwardhan, A. S. and Coppersmith, K. J. Estimating the Probability of Occurrences of Surface Faulting Earthquakes on the Wasatch Fault Zone, Utah, *Bull. Seism. Soc. Amer.*, 1980, **70**, 1463–1478

45 Shimazaki, K. and Nakata, T. Time-Predictable Recurrence Model for Large Earthquakes, *Geoph. Res. Lett.*, 1980, **7**, 279–282

46 Anagnos, T. and Kiremidjian, A. S. Stochastic Time-Predictable Model for Earthquake Occurrences, *Bull. Seism. Soc. Am.*, 1984, **74**, 2593–2611

47 Guagenti-Grandori, E. and Molina, C. Semi-Markov Processes in Seismic Risk Analysis, Int. Symp. on Semi-Markov Proc. and Their Appls., Brussels Belgium, 1984

48 Anagnos, T. and Kiremidjian, A. S. A Stochastic Earthquake Recurrence Model with Temporal and Spatial Dependence, Report No. 76, John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA, 1985

49 Kameda, H. and Ozaki, Y. A Renewal Process Model for Use in Seismic Risk Analysis, *Memoirs of Faculty of Engineering*, Kyoto University, 1979, **XLI**, 11–35

50 Kameda, H. and Takagi, H. Seismic Hazard Estimation Based on Non-Poisson Earthquake Occurrences, Rev. Meeting US-Japan Cooperative Research Seis. Risk Anal. and Appl. to Reliability-Based Des. Lifeline Sys., 1981, 56 pp.

51 Savy, J. B., Shah, H. C. and Boore, D. Nonstationary Risk Model with Geophysical Input, *J. Struc. Div., ASCE*, 1980, **106**, 145–163

52 Kiremidjian, A. and Anagnos, T. Stochastic Slip-Predictable Model for Earthquake Occurrences, *Bull. Seism. Soc. Amer.*, 1984, **74**, 739–755

53 Kiremidjian, A. S. and Suzuki, S. A Stochastic Model for Site Ground Motions from Temporally Dependent Earthquakes, *Bull. Seism. Soc. Am.*, 1987, **77**

54 Suzuki, S. and Kiremidjian, A. S. Risk Consistent Design Spectra from Theoretical Earthquake Ground Motions, paper submitted to *Int. J. Earth. Enq. and Struct. Dyn.*, November 1987

55 Vere-Jones, D. and Davies, R. B. Statistical Survey of Earthquakes in the Main Seismic Region of New Zealand, Part 2: Time Series Analysis, *NZ J. Geol. Geoph.*, 1966, **9**, 251–284

56 Shlien, S. and Toksöz, M. N. A Clustering Model for Earthquake Occurrences, *Bull. Seism. Soc. Amer.*, 1970, **60**, 1765–1787

57 Merz, H. A. and Cornell, C. A. Aftershocks in Engineering Seismic Risk Analysis, Report R73-25, Department of Civil Engineering, Massachusetts Inst. of Tech., Cambridge, MA, 1973b

58 Kagan, Y. and Knopoff, L. Statistical Search for Non-Random Features of the Seismicity of Strong Earthquakes, *Phys. Earth Plant. Int.*, 1976, **12**, 291–318

59 Kagan, Y. and Knopoff, L. Earthquake Risk Prediction as a Stochastic Process, *Phys. Earth Planet. Int.*, 1977, **14**, 97–108

60 Kagan, Y. and Knopoff, L. Dependence of Seismicity on Depth, *Bull. Seism. Soc. Amer.*, 1980, **70**, 1811–1822

61 Liu, S. C. and Fagel, L. W. Earthquake Environment for Physical Design: A Statistical Analysis, *Bell Sys. Tech. Journal*, 1972, **51**, 1957–1982

62 Guagenti-Grandori, E. and Scirocco, F. A Discussion of Seismic Risk Including Precursors, *Bull. Seism. Soc. Amer.*, 1980, **70**, 2245–2251

63 Hagiwara, Y. Probability of Earthquake Occurrence as Obtained from a Weibull Distribution Analysis of Crustal Strain, *Tectonophysics*, 1974, **23**, 313–318

64 Rikitake, T. Statistics of Ultimate Strain of the Earth's Crust and Probability of Earthquake Occurrence, *Tectonophysics*, 1975, **26**, 1–21

65 Grandori, G., Guagenti, E. and Petrini, V. On the Use of Renewal Processes in Seismic Hazard Analysis, *Proc. World Conf. Eq. Eng., 8th*, San Francisco, California, 1984, 287–294

66 Lai, C. D. A Two-Dimensional 'Immigration-Branching' Model with Application to Earthquake Occurrence Times and Energies, *J. Appl. Prob.*, 1977, **14**, 464–474

67 Kagan, Y. and Knopoff, L. The Two Point Correlation Function of the Seismic Moment Tensor, Institute of Geophysics, University of California, Los Angeles, 1984

68 Cornell, C. A. and Winterstein, S. R. Applicability of the Poisson Earthquake Occurrence Model, *Seismic Hazard Methodology* for the *Central* and *Eastern United States*, EPRI Research Report, 1986, p. 101