PROBABILISTIC SEISMIC HAZARD ANALYSIS METHOD FOR
MAPPING OF SPECTRAL AMPLITUDES AND OTHER DESIGN-
sPECIFIC QUANTITIES TO ESTIMATE THE EARTHQUAKE
EFFECTS ON MAN-MADE STRUCTURES

I.D. Gupta
Central Water and Power Research Station
Khadakwasla, Pune-411024

ABSTRACT

Along with the commonly used parameters like peak acceleration and response spectral amplitudes, knowledge of many other parameters like strong-motion duration, peak strains, likelihood for initiation of liquefaction, and permanent dislocations across faults is necessary to have a more comprehensive estimate of the earthquake effects on a variety of man-made structures. The present paper provides a concise but complete description of the probabilistic seismic hazard analysis (PSHA) method to map any of these quantities with a uniform probability of not being exceeded due to the total expected seismicity during a specified life period. Example results are presented to illustrate the application of the PSHA method in preparing the microzonation maps for several different hazard parameters. The paper also proposes simple practical solutions for some of the difficulties faced in implementing the existing PSHA method in real applications.

KEYWORDS: Probabilistic Seismic Hazard, Uncertainties, Hazard Parameters, Microzonation Maps

INTRODUCTION

The seismic hazard analysis is concerned with the evaluation of the levels of various natural effects of earthquakes, which may be of consequence for the safety of an existing or a proposed man-made structure at a site. Some important parameters used for characterization of seismic hazard can be listed as the peak ground acceleration (Cornell, 1968), Fourier and response spectrum amplitudes (McGuire, 1977; Anderson and Trifunac, 1977, 1978; Lee and Trifunac, 1985), strong motion duration (Papazachos et al., 1992), peak strains (Todorovska and Trifunac, 1996), surface faulting (Todorovska et al., 2005; Stepp et al., 2001), soil liquefaction (Todorovska and Trifunac, 1999), and landslides (Del Gaudio and Wasowski, 2004), out of which the response spectrum is the most widely used and extensively studied strong-motion functional. The concept of response spectrum method was introduced in early 1930s (Biot, 1932, 1933, 1934); and since 1970, it has become the principal tool in the design of earthquake-resistant structures (Trifunac, 2003), because of the simplicity and directness with which it relates the strong motion with the response of a structure.

The deterministic and the probabilistic are the two commonly used approaches for the seismic hazard analysis. In the deterministic approach, the value of a hazard parameter of interest is estimated for a specified earthquake magnitude assumed to occur at a fixed source-to-site distance (e.g., Reiter, 1990; Anderson, 1997; Krinitzsky, 2002). However, a single scenario earthquake is not able to provide a true picture of the seismic hazard at a site because different combinations of magnitude and distance contribute more significantly in different frequency bands. On the other hand, the probabilistic seismic hazard analysis (PSHA) approach takes into account the effects of all the earthquakes by considering the inherent random nature of earthquake magnitude, recurrence time, and epicentral location as well as that of the amplitude of the hazard parameter of interest (e.g., Cornell, 1968; McGuire, 1977; Anderson and Trifunac, 1978). The estimate of a hazard parameter by PSHA approach is thus not expected to be exceeded with a desired confidence level due to any of the earthquakes expected to occur during a given exposure period.

The PSHA formulation was first presented by Cornell (1968) for the peak ground acceleration. He modeled the randomness in magnitude by the Gutenberg-Richter’s frequency-magnitude relationship (Gutenberg and Richter, 1944), that in recurrence time by Poisson probability distribution, and that in
location by considering the epicenters to be confined to a point source, or distributed uniformly over a straight line fault or an annular area around the site. But the formulation of Cornell (1968) did not consider the random scattering in the amplitudes of the hazard parameter around the median attenuation relationship. Many other early studies (e.g., Milne and Davenport, 1969; Douglas and Ryall, 1975) as well as some later studies also (e.g., Kijko and Graham, 1999) have not considered the randomness in hazard parameter. Der Kiureghian (1977) showed that this randomness may be a significant source of uncertainty in the results of the hazard analysis. To have a uniformly conservative estimate of the hazard at all the frequencies, McGuire (1974, 1977) performed the PSHA for response spectrum amplitudes at several different frequencies, with the randomness in spectral amplitudes considered by a lognormal distribution. Anderson and Trifunac (1977, 1978) generalized the PSHA formulation by modeling the seismicity in a more realistic way and applied that to compute the Fourier amplitude spectra. They employed five different types of source: (a) a point source, (b) a line source (not necessarily straight), (c) an areal source with arbitrary boundary, (d) an arbitrarily dipping fault surface, and (e) a volume of arbitrary shape, to define the seismicity. Their formulation also included the effect of fault rupture dimensions, which may have significant effect on the hazard estimation (e.g., Ang, 1973; Der Kiureghian and Ang, 1975; Anderson and Trifunac, 1977).

Most of the recent developments in the PSHA approach have been primarily concerned with introducing different probabilistic models to describe the randomness in earthquake magnitude, recurrence time, and epicentral location to get more realistic descriptions for specific practical applications. However, due to inadequacy or lack of available data and incomplete understanding of the earthquake and ground-motion generating processes, it is generally difficult to specify the various input models and their parameters without any uncertainty. The current PSHA approach utilizes the logic-tree methodology (Kulkarni et al., 1984) to quantify the effect of these additional uncertainties, termed commonly as “epistemic” uncertainties. On the other hand, the basic PSHA approach considers only the inherent random uncertainties, which are termed as “aleatory” uncertainties. The logic-tree methodology provides a systematic graphical procedure for identifying all possible sets of input models and their parameters. An appropriate weight is assigned to each set of these inputs to the PSHA by assigning suitable weights to the various logic-tree branches for each input element. The basic PSHA is then performed for each set of inputs to get a complete picture of the effect of the epistemic uncertainties on the hazard estimation. However, there is no consensus on the way the uncertainties are to be assigned and on how to take the final decision with epistemic uncertainties (e.g., Klügel, 2005a, 2005b, 2005c; Musson et al., 2005; Budnitz et al., 2005).

If applied properly, the PSHA approach may prove a powerful method for estimation of site-specific design ground motions for practical applications (e.g., EPRI, 1986; Bernreuter et al., 1987; Todorovska et al., 1995; Senior Seismic Hazard Analysis Committee, 1997; USACE, 1999; Gupta, 2002a; McGuire, 2004). The results of PSHA may form a basis for earthquake-resistant design using both the simplified elastic analysis (e.g., BSSC, 1997) as well as more rigorous performance-based analysis (e.g., FEMA, 2000; Ellingwood, 2001; Bertero and Bertero, 2004). Another practical application of PSHA approach is in preparation of seismic zoning maps. Zoning may be done on a macro scale, such as those under GSHAP (1999) and several other studies (e.g., Frankel et al., 2002; Adams and Atkinson, 2003; Das et al., 2006), or on a micro scale including the regional and local site effects in a more detailed way (e.g., Lee and Trifunac, 1987; Trifunac, 1990a). The microzoning maps need not be limited only to the peak acceleration and the spectral amplitudes at selected frequencies. The recent developments have enabled to prepare the microzonation maps in terms of normalized peak strains (Todorovska and Trifunac, 1996), surface faulting (Todorovska et al., 2005), and liquefaction potential (Todorovska and Trifunac, 1999). The present paper provides an overview of the various aspects of the currently used PSHA approach with a number of illustrative example results. Many studies have proposed the deaggregation of probabilistic seismic hazard to represent the hazard equivalently by a single pair of earthquake magnitude and distance, which is considered necessary and useful in making certain engineering decisions (e.g., Chapman, 1995; Ishikawa and Kameda, 1988; Bazzurro and Cornell, 1999; McGuire, 1995). However, the use of PSHA to arrive at a single scenario earthquake by deaggregation has not been described in any detail.
THE PSHA FORMULATION

The PSHA formulation is fundamentally concerned with estimating the expected occurrence rate, \( \nu(Z > z) \), of exceeding a specified value, \( z \), of a random parameter, \( Z \), used for characterization of hazard at a site. For this purpose, the original formulation due to Cornell (1968) uses only those combinations of earthquake magnitude and distance, which may cause a specified mean or median estimate of \( Z \) to be exceeded. However, by considering the random scattering of the amplitudes of hazard parameter around the mean or median estimate, the occurrence rate can be defined using total probability theorem by the following generalized form of expression.

\[
\nu(Z > z) = \sum_{n=1}^{N} N_n(M_{\text{min}}) \int \int \delta(Z > z | M, R, \varepsilon) f_n(M) g_n(R) h(\varepsilon) dM dR d\varepsilon
\]

In this expression, \( N_n(M_{\text{min}}) \) represents the occurrence rate of earthquakes above a selected threshold magnitude \( M_{\text{min}} \) in the \( n \)th source zone, and the summation is taken over all the \( N \) number of source zones. Functions \( f_n(M) \) and \( g_n(R) \) are the probability density functions of magnitude and distance for this source. Further, the expression of Equation (1) is based on the assumption that the logarithm of the values of the hazard parameter for magnitude \( M \) and distance \( R \) follows a Gaussian distribution with mean value \( \mu(M, R) \) and standard deviation \( \sigma(M, R) \). The quantity \( \delta(Z > z | M, R, \varepsilon) \) as taken as 1.0 for \( \ln z \) equal to \( \mu(M, R) \) plus \( \varepsilon \) times \( \sigma(M, R) \) and zero otherwise, with \( h(\varepsilon) \) as the standard Gaussian distribution with zero mean and unit standard deviation. In practical applications, the probability distribution of the amplitudes of hazard parameter is usually truncated arbitrarily at two to three standard deviations, which cannot be considered appropriate. If at all, any upper limit on the hazard parameter has to be based on the physical grounds. However, it seems unlikely that this problem may be solved in the near future (Bommer et al., 2004).

Contrary to that assumed in Equation (1), the residuals of the hazard parameter need not necessarily be defined by a Gaussian density function (Trifunac and Lee, 1979). It will therefore be more generalized to replace the integral of the product of \( \delta(Z > z | M, R, \varepsilon) \) and \( h(\varepsilon) \) over \( \varepsilon \) by the probability of exceeding level \( z \) due to magnitude \( M \) at distance \( R \). Representing this probability by \( q(Z > z | M, R) \), the expression for the occurrence rate becomes

\[
\nu(Z > z) = \sum_{n=1}^{N} N_n(M_{\text{min}}) \int \int q(Z > z | M, R) f_n(M) g_n(R) dM dR
\]

By discretizing the magnitude and distance for the \( n \)th source zone into small intervals like \((M_j - \delta M_j, M_j + \delta M_j)\) and \((R_i - \delta R_i, R_i + \delta R_i)\) with central values \( M_j \) and \( R_i \), the occurrence rate of earthquakes in the \( j \)th magnitude and the \( i \)th distance interval can be defined as

\[
\lambda_n(M_j, R_i) = N_n(M_{\text{min}}) \int_{M_j - \delta M_j}^{M_j + \delta M_j} \int_{R_i - \delta R_i}^{R_i + \delta R_i} f_n(M) g_n(R) dM dR
\]

The expression of Equation (2) can thus be written in the following discrete form:

\[
\nu(Z > z) = \sum_{n=1}^{N} \sum_{j=1}^{J} \sum_{i=1}^{I} q(Z > z | M_j, R_i) \lambda_n(M_j, R_i)
\]

A total of \( J \) magnitude ranges and \( I \) distance ranges are considered for the summations in Equation (4). Further, if the same attenuation relation is applicable to all the seismic source zones, it is possible to use directly the total annual number, \( n(M_j, R_i) \), of earthquakes obtained by adding the numbers for all the source zones. The expression of Equation (4) thus becomes (Anderson and Trifunac, 1977, 1978)

\[
\nu(Z > z) = \sum_{j=1}^{J} \sum_{i=1}^{I} q(Z > z | M_j, R_i) n(M_j, R_i)
\]
In practical applications, the probabilistic hazard computation is commonly based on the expressions of Equations (4) and (5). By using the numbers, \( \lambda_n(I_0j, R_i) \) or \( n(I_0j, R_i) \), of earthquakes with epicentral intensity \( I_0j \) at distance \( R_i \), the probabilistic hazard can also be performed using the intensity data (Gupta, 1991). For this purpose, the probability, \( q(I > I_0j, R_i) \), of \( I > I_0 \) is obtained by summing over all the site intensities the product of the probability of exceeding value \( z \) due to a specified site intensity multiplied by the probability of occurrence of that site intensity due to the combination of \( I_0j \) and \( R_i \).

The reciprocal of \( u(I > I_0) \) gives the return period for the occurrence of an amplitude \( z \) or above of the hazard parameter. Assuming the occurrence rate \( \lambda_n(M_j, R_i) \) to follow a Poisson probability distribution, the occurrence rate \( u(I > I_0) \), which is a linear combination of \( \lambda_n(M_j, R_i) \), can also be described by a Poisson probability distribution. Thus, the probability of \( I > I_0 \) due to all the earthquakes in all the sources during an exposure period of \( Y \) years can be written as

\[
P(I > I_0|Y) = 1 - \exp \{-Y u(I > I_0)\}
\]  

(6)

From this, the return period for the occurrence of \( I > I_0 \) can be defined in terms of \( P(I > I_0|Y) \) as

\[
T(I > I_0) = \frac{1}{\ln(1 - P(I > I_0|Y))}
\]  

(7)

The plot of the probability \( P(I > I_0|Y) \) versus \( z \) is commonly known as the “hazard curve”. The most widely adopted practice is to plot the hazard curve in terms of the annual (\( Y = 1 \)) probability of exceedance. Assuming \( u(I > I_0) \) to be very small, the annual probability of exceedance is generally approximated by \( u(I > I_0) \). The hazard curves are sometimes also plotted as \( T(I > I_0) \) versus \( z \). The various representations of the hazard curve are shown schematically in Figure 1. It may be noted that though the PSHA can equivalently be described by any of the quantities \( u(I > I_0) \), \( T(I > I_0) \), \( P(I > I_0|Y) \), \( P(I > I_0|Y) \), which are interrelated by simple relations, the use of \( P(I > I_0|Y) \) provides a direct physical interpretation of the results of PSHA. If \( Z \) represents the Fourier or response spectral amplitudes at different periods, the hazard curves in terms of \( P(I > I_0|Y) \) can be used to obtain the complete spectrum with a constant probability of exceedance. A spectrum thus obtained is commonly termed as “uniform hazard spectrum”.

![Fig. 1 The various commonly used representations of the seismic hazard curves](image-url)
An alternative to the above analytical formulation for PSHA is the use of Monte Carlo simulation, in which a very long duration of earthquake catalog is generated from the probability density functions of magnitude, epicentral location, and the inter-event time for each source zone (e.g., Musson, 1999a; Smith, 2003). The amplitudes of the hazard parameter are then computed for all the earthquakes in the simulated catalog using a suitable probability density function for the residuals of the hazard parameter. The annual rate, $v(Z > z)$, is finally obtained by counting the number of years in which the maximum value of $Z$ exceeds a specified value $z$, and by dividing it by the total duration (in years) of the catalog. This procedure generally takes much more computational time without any apparent advantage for the case of Poisson occurrences of earthquakes. However, it may sometimes be more convenient to use the simulation to account for the epistemic type of uncertainties (e.g., Musson, 1999b; Smith, 2003).

1. PSHA with Non-Poisson Earthquake Occurrences

The foregoing hazard formulation is based on the Poisson assumption for the occurrences of earthquakes in a region, which may be violated in that the earthquakes may be characterized by long as well as short-term temporal correlations. Under the Poisson assumption, the inter-event times follow an exponential distribution with a constant rate of occurrence. However, very large magnitude events in seismically active areas may follow a long-term cyclic behavior with time-varying rate of occurrence. Such events are required to be described by a real-time renewal model, wherein the occurrence rate is small soon after a large earthquake and increases with the lapse of time since the last such event (e.g., Rikitake, 1976; Vere-Jones and Ozaki, 1982; Sykes and Nishenko, 1984; Thacher, 1984; Nishenko and Buland, 1987; Jara and Rosenblueth, 1988). Several studies have implemented the time-dependent renewal models in the PSHA approach (e.g., Kameda and Takagi, 1981; Kiremidjian and Suzuki, 1987; Cornell and Winterstein, 1988; Lee, 1992; Todorovska, 1994). A renewal process that satisfies all the Poisson assumptions except the constant occurrence rate is called as a “non-homogeneous Poisson process”. It is required to be defined by a time-dependent occurrence rate, which can be obtained from the hazard function based on the probability distribution of inter-event times. The PSHA formulation of Equation (6) for the stationary Poisson processes is applicable to such events also if their average occurrence rate is obtained using a time-dependent hazard function (Lee, 1992).

The expression of Equation (6) is, however, not applicable to the events like aftershocks and sequential earthquakes, which are characterized by strong spatio-temporal correlation among themselves as well as with the main shock. To include the effect of the aftershocks it is necessary to decluster the available earthquake catalog using a suitable algorithm (e.g., Keilis-Borok et al., 1972; Reasenberg, 1985; Maeda, 1996). Only the background seismicity is then described by the Poisson model, and the aftershocks by some other suitable model (e.g., Hagiwara, 1974; Utsu, 1984; Hong and Guo, 1995; Corral, 2004; Molchan, 2005). The aftershocks can also be described by a Poisson model with time-dependent occurrence rate defined by the modified Omori’s law (Utsu et al., 1995). If no standard distribution is found suitable, an actual probability density function can be obtained by summation of a suitable kernel function with the observed interevent times (Silverman, 1986). A large number of earthquake catalogs of $Y$ years duration are then simulated using Poisson distribution with constant occurrence rate for the mainshocks, and some of the above mentioned distributions for the aftershocks. Beauval et al. (2006) have proposed to simulate the combined seismicity using epidemic type aftershock sequence (Ogata, 1988). Next, the amplitudes of hazard parameters are simulated for all the earthquakes in these catalogs, from which the probability $P(Z > z|Y)$ is defined as the fraction of the total number of catalogs with the maximum value of the hazard parameter exceeding the value $z$.

A more efficient method to account for the effect of the aftershocks may perhaps be to generate only a single catalog of $Y$ years duration for the aftershocks, and assume them to occur in a literal way. If $\eta(M_l, R_k|Y)$ is the total number of aftershocks in $Y$ years in a small magnitude interval around central magnitude $M_l$ and in a small distance interval around central distance $R_k$, the probability of $Z > z$ due to these events to occur in a deterministic way can be defined as (Anderson and Trifunac, 1977)

$$P'(Z > z|Y) = 1 - \exp \left\{ \sum_{k=1}^{K} \sum_{l=1}^{L} \ln \left( 1 - q(Z > z|M_l, R_k) \right) \eta(M_l, R_k|Y) \right\}.$$  \hfill (8)
By carrying out the hazard analysis for the declustered catalog of the main shocks using the expression of Equation (6), the combined probability of $Z > z$ from both the main earthquakes and the aftershocks can be defined as

$$P^*(Z > z|Y) = 1 - \exp\left(-Y \nu(Z > z)\right) \left\{1 - P^*(Z > z|Y)\right\}$$

This expression is expected to provide adequately conservative estimate of the hazard for practical applications. Further, the effect of any other type of events occurring in a literal way (e.g., earthquake prediction) can also be included in $P^*(Z > z|Y)$ by including their numbers in $\eta(M_1, R_1|Y)$.

2. Steps Involved in PSHA Approach

The four basic steps involved in the implementation of the foregoing PSHA formulation are depicted schematically in Figure 2. The first step is to identify and demarcate the boundaries of the various seismic sources. Normally, the sources within about 300 to 400 km (depending on the tectonic region) of the site are sufficient for this purpose. Each of the sources is divided into a large number of small-size elements, and the expected seismicity in a source is distributed suitably among all the elements. The epicenters of all the earthquakes in an element are assumed to occur at the geometric center of the element. The probability distribution function $G(R)$ of the source-to-site distance $R$ is then defined using the distances to all the elements as illustrated in the top left panel in Figure 2.

To estimate the total number, $N_n(M_{\text{min}})$, of earthquakes with magnitude above $M_{\text{min}}$ in a source zone, the frequency-magnitude relationship due to Gutenberg and Richter (1944) is defined for each
A magnitude distribution function, $F(M)$, is also defined for each source to distribute these numbers among different magnitude intervals between a minimum magnitude, $M_{\text{min}}$, and a maximum magnitude, $M_{\text{max}}$. The exponentially decaying magnitude distribution is generally found suitable for area sources, whereas a characteristic earthquake model (Youngs and Coppersmith, 1985) is commonly used for individual faults. Alternatively, one may generate the synthetic catalogs for each source zone by estimating the parameters of the probability density functions for magnitude, occurrence time, and distance, defined from the available earthquake catalog.

A suitable attenuation relationship providing a probabilistic description of the amplitudes of the hazard parameter is required to be selected in Step 3. This should provide the mean or median estimate and the corresponding probability distribution of the residuals for specified earthquake magnitude, source-to-site distance, and site geologic and soil conditions. This is used to estimate the probability $q(Z > z|M, R)$ as illustrated in bottom left panel in Figure 2. A single attenuation relation may normally be applicable to all the source zones, but different relations may also be used, if necessary. For example, as in the northeast India, if a site is affected simultaneously by shallow crustal and deep subduction zone earthquakes, those are required to be described by different attenuation relations.

The fourth and the final step in the basic PSHA is to compute the hazard curves by integrating over all the magnitudes and distances in all the source zones. Several hazard curves are required to compute the uniform hazard spectra as shown in the bottom right panel in Figure 2 (Anderson and Trifunac, 1977). It may be noted that due to lack of exact scientific knowledge and inadequacy of available data, it may not be possible to establish the first three steps of PSHA in a unique way (Gupta, 2005). For example, there could be several possible choices for the definition of seismic source zones and distribution of distance, type of earthquake recurrence model and the maximum magnitude for each source, as well as for the attenuation relationship for the hazard parameter of interest. Due to these epistemic uncertainties, a large number of different sets of input with different weights may be possible in the PSHA, which can be identified by the logic-tree method (Kulkarni et al., 1984). A typical logic tree depicting the possible uncertainties in the various elements of the basic PSHA is shown in Figure 3. The basic principle to be followed in setting up a logic tree is that the branches emanating from a single node should cover only the physically realizable distinct possibilities, which may lead to significantly different estimate of the hazard.

In the logic tree of Figure 3, three sets of source zones with different weights may result from different interpretations and subjective judgments for a given database on seismotectonics and geological features in the region of interest. Two different sets with weights of 0.6 and 0.4 for the past earthquake catalog form the second element of the logic tree, which may result from the availability of several catalogs prepared by different organizations or from the use of different methods for homogenization of magnitudes in a given catalog. Two options with equal weights are shown for the two different types of recurrence relationships (to be explained in more detail later). Further, two different moment release rates
are considered in the recurrence relationship with constant moment rate. The next element in the logic tree is the maximum magnitude, for which three options as small, large and preferred with weights of 0.3, 0.3 and 0.4 are considered for each source zone. The spatial distribution of seismicity in a source zone is considered in two different ways as (a) uniform distribution, and (b) that based on spatially smoothed past seismicity. Finally, there are three different options for the ground motion attenuation model, with weights equal to 0.4, 0.3 and 0.3.

The example logic tree in Figure 3 has a total of 324 end branches, which are given by the product of the number of different options for each input element. The weight for an end branch is given by the product of the weights of all the intermediate branches leading to that branch. To account for the effect of the epistemic uncertainties, the basic PSHA is performed for all the combinations of the input leading to various end branches, and the resulting hazard curves are assigned the corresponding weights. These can be used to define the mean or the median hazard curve, as well as the hazard curves with desired confidence levels. However, at present, there is no widely accepted practice for the choice of the hazard curve for use in practical applications. The subsequent sections in the paper describe the first three steps providing the input for the fourth step in PSHA, and the possible epistemic uncertainties involved in each step. Illustrative example results are then presented for the uniform hazard Fourier and response spectra, and some other parameters of importance for characterizing the hazard.

**SEISMIC SOURCES AND DISTANCE DISTRIBUTION**

In an ideal situation, each source zone has to be an individual fault or fault segment. However, due to lack of knowledge about all the faults and wide dispersion of the epicenters of past earthquakes in relation to the known faults, broad area sources encompassing several faults are commonly used in real practice. Such seismic sources may be associated with the geological structures like uplifts, rifts, folds and volcanoes, which release the tectonic stresses and localize the seismic activity. Another type of seismic source used in practical applications is the “tectonic province”, which generally covers a large geographic area of diffused seismicity with no identifiable active faults or geological structures. The observed seismicity is sometimes seen to be highly concentrated in a very small area. This can be defined by a point source, if located far away from the site of interest. The source zones in a region are identified on the basis of some sort of geological, geophysical, geodetic and seismotectonic uniformity. The seismic potential of a source zone has to be distinctly different from the other adjacent sources. As the available data in most cases are not adequate, expert knowledge, detailed familiarity with the geology in the area, interpretation and judgment play important role in defining the seismic sources. The following four types of source zones can be considered sufficient for most of the practical PSHA applications.

**Point Source:** The seismicity in a point source is concentrated in a small area at very long distance from the site, and the fault with which it can be associated generally does not have to be identifiable. The geometric center of this small area is assumed to be the epicentral location for all the earthquakes expected to occur in the point source. Thus, the epicentral distance has a single fixed value in this case. However, the probability distribution of the closest distance to fault rupture can be defined assuming fault rupture to be a straight line and equally likely in all azimuthal directions (Anderson and Trifunac, 1977).

**Line Source:** In this type of source, the seismicity is related to a long fault and is usually, but not necessarily, assumed to be distributed uniformly over its entire length. Cornell (1968) considered a straight-line fault and provided an expression for the distance to a site from any point on the fault. Anderson and Trifunac (1977) included consideration of curved faults also. They divided the fault length \( L \) into \( N \) small elements of length \( \Delta L \) each, and assumed the midpoint of each element to be the location of the epicenters. Each fault element is normally assigned the same weight, but non-uniform weights may also be assigned if different segments of the fault are known to be characterized by different levels of seismicity. The distances to all the fault elements with corresponding weights can be used to find the probability distribution of the epicentral distance from the site selected for the estimation of hazard. Anderson and Trifunac (1977) also proposed to consider the effect of fault rupture. For this purpose, if \( l \) is the fault rupture length for a given earthquake magnitude, the epicenters are assumed to be located sequentially in any continuous fault segment of length \( (L-l) \). This obviates the need to account for the “unilateral” or “bilateral” nature of rupture propagation. The closest distance to the fault rupture is then estimated for each epicentral location for a given magnitude, and it is assigned a weight equal to that for
the corresponding epicentral location. All the closest distances with weights can be used to find the probability distribution of the closest distance.

**Dipping Plane Source:** Anderson and Trifunac (1977) introduced this type of source zone to describe the seismicity associated with a dipping fault plane. To find the probability distribution of the closest distance to fault rupture, let \( L \) be the total length and \( W \) the total width of the fault plane, and let \( l \) and \( w \) be the rupture length and width for a given earthquake magnitude. Then, similar to the line source, assuming the hypocenters to lie in any continuous area of length \((L-l)\) and width \((W-w)\) of the fault plane, one can find the closest distance to the fault rupture for each hypocentral location considered sequentially by dividing this area of the fault plane into small-size elements of length \( \Delta L \) and width \( \Delta W \). The probability distribution of the closest distance to fault rupture can be obtained by assigning suitable weights to each hypocentral location, which may be uniform or non-uniform.

**Area Source:** This is the most widely used type of source zone in the PSHA studies. One has to use gross area sources when the observed seismicity is associated with a localizing geologic structure or a tectonic province. Cornell (1968) considered the area type of source defined by an annular area around the site of interest, which was generalized by Anderson and Trifunac (1977) to be of any arbitrary shape and located anywhere with respect to the site. One may refer to Gupta (2006a) for a very comprehensive description and examples on defining the area type of source zones for India and surrounding areas. The probability density of epicentral distances for an area source can be obtained easily by dividing the source zone into small-size elements and by assuming the epicenters to lie at the geometric center of each element. For uniform distribution of seismicity, each epicentral location is assigned a weight in proportion to the area of the corresponding element. One can also consider non-uniform distribution on the basis of the past seismicity with proper spatial smoothing (e.g., Frankel, 1995; Woo, 1996; Kagan and Jackson, 2000).

Similar to the point source, the effect of fault rupture length may be accounted by assuming the rupture to be straight and equally likely in all the directions for each epicentral location (Anderson and Trifunac, 1977). If aftershocks are included separately in the analysis, those can be distributed around the main shocks according to an isotropic probability density function (e.g., Helmstetter et al., 2003; Zhuang et al., 2004).

To illustrate the application of the foregoing procedure for arriving at the probability distribution of the source-to-site distance, Figure 4 shows the hypothetical examples of two area sources A1 and A2 of diffused seismicity and two line sources L1 and L2 represented by vertical faults. To consider the effect of fault rupture, the rupture length and width are estimated using the empirical correlations due to Wells and Coppersmith (1994) for all fault types. Assuming the two faults to be characterized by the same seismic potential, the combined probability distribution function \( G(R) \) of the distances for them has been obtained as shown in Figure 5. In general, one can combine any number of faults in this way. The results in Figure 5 represent the probability distributions without the fault rupture as well as with the fault rupture for magnitudes equal to 5.5, 6.5 and 7.5. Similar results for the area source A1 with uniform spatial distribution of seismicity are shown in Figure 6, whereas Figure 7 shows the results with spatial distribution based on the spatially averaged past seismicity. Compared to the distances without fault rupture considered, the closest distances to fault rupture are seen to have increasingly smaller values with increase in the magnitude for both the line and the area sources. Further, the probability distributions based on the spatial distribution of past seismicity for the area source have shifted towards longer distances compared to those for the uniformly distributed seismicity. The probability density function, \( g(R) \), of source-to-site distance can be obtained by differentiating the distribution function \( G(R) \). Thus the value of the integral over distance in Equation (3) can directly be written as \( (G(R_i + \delta R_i) - G(R_i - \delta R_i)) \).

As mentioned before, in practical applications, the seismic sources cannot be defined in a unique way (Bender, 1986). Some subjectivity is inevitable due to inadequacy or non-availability of the required data, and also due to possible alternative interpretations of the available data. Borders between source regions are usually not sharp with respect to seismic activity. Furthermore, the complete understanding of the long-term tectonic processes is generally lacking in many cases. To account for the uncertainties in defining the source zones, more than one set of source zones are required to be used as indicated in the logic tree diagram of Figure 3. If considerable seismicity is known to have occurred in the region of interest, the spatial distribution can also be based purely on the past seismicity with a zoneless approach (e.g., Frankel, 1995; Woo, 1996; Das et al., 2006). However, ignoring the distinct geological and seismological knowledge may not always be appropriate.
Fig. 4 Typical examples of the area and line types of seismic sources

Fig. 5 Probability distribution of the source-to-site distance for the line sources in Figure 4 with uniform distribution of seismicity (the solid curve corresponds to the epicentral distance and the dashed curves to the closest distance to the fault rupture for different earthquake magnitudes)
Fig. 6  Probability distribution of the source-to-site distance for the area source A1 in Figure 4 with uniform distribution of seismicity (the solid curve corresponds to the epicentral distance and the dashed curves to the closest distance to the fault rupture for different earthquake magnitudes)

Fig. 7  Probability distribution of the source-to-site distance for the area source A1 in Figure 4 with non-uniform distribution of seismicity based on spatially averaged past seismicity with a correlation distance of 20 km
RECURRANCE RELATIONSHIP AND MAGNITUDE DISTRIBUTION

An earthquake recurrence relationship defines the annual occurrence rate, $N(M)$, of earthquakes with magnitude greater than or equal to $M$. Anagnos and Kiremidjian (1988) have reviewed the earthquake recurrence models for seismic hazard analysis. If $N(M_{\text{min}})$ is the total number of earthquakes above a selected threshold magnitude $M_{\text{min}}$, the number $N(M)$ can be written as a product of $N(M_{\text{min}})$ and the probability distribution function, $F(M)$, of the earthquake magnitude. The negative of the derivative of $F(M)$ gives the density function, $f(M)$, of the magnitude. Thus, the recurrence relationship can be used to obtain directly the value of the $N_n(M_{\text{min}})$ times the integral of $f(M)$ over the magnitude in Equation (3) as $\left( N_n(M_j - \delta M_j) - N_n(M_j + \delta M_j) \right)$. Thus, along with the probability distribution of source-to-site distance, the occurrence rate, $\lambda_n(M_j, R_i)$, in Equation (3) can be obtained as

$$\lambda_n(M_j, R_i) = \left[ N_n(M_j - \delta M_j) - N_n(M_j + \delta M_j) \right] \left[ G(R_i + \delta R_i) - G(R_i - \delta R_i) \right]$$ (10)

Gutenberg and Richter (1944) have defined a form of the earthquake recurrence relationship as

$$\log N(M) = a - bM$$ (11)

In this relation, $a$ and $b$ are the constants specific to a seismic source, which are commonly estimated using available data on past earthquakes. To evaluate $a$ and $b$, it is necessary to convert the available data into a common magnitude scale using suitable empirical conversion relations (e.g., Chung and Bernreuter, 1981; Utsu, 1982) and to remove the dependent events using an appropriate algorithm (e.g., Reasenberg, 1985; Maeda, 1996; Hainzl et al., 2006). It is also necessary to account for the incompleteness of lower magnitude earthquakes, for which several methods have been proposed by different investigators (e.g., Stepp, 1972; Lee and Brillinger, 1979; Tinti and Mulargia, 1985; Rydelek and Sacks, 1989; Weimer and Wyss, 2000; Albarello et al., 2001). However, the procedure due to Stepp (1972) can be considered quite suitable and convenient for the practical hazard analysis applications. Then, the parameters $a$ and $b$ in Equation (11) can be evaluated by least squares, maximum likelihood (e.g., Weicher, 1980; Bender, 1983), or the maximum entropy (Dong et al., 1984) method, but the maximum likelihood method is, in general, considered quite appropriate.

By defining $\beta = b \ln 10$, the relationship of Equation (11) can be expressed in terms of the total number, $N(M_{\text{min}})$, of earthquakes above a threshold magnitude $M_{\text{min}}$ and the probability distribution $F(M)$ as

$$N(M) = N(M_{\text{min}}) \exp\left(-\beta(M - M_{\text{min}})\right)$$ (12)

This relation does not impose an upper limit on the magnitude, whereas it is necessary to consider an upper bound magnitude, $M_{\text{max}}$, in the practical applications. An abrupt truncation of the relation of Equation (12) at magnitude $M_{\text{max}}$ is not considered appropriate, as it will result in an infinitely large value of the density function at magnitude $M_{\text{max}}$. This problem could be avoided if $N(M)$ tends asymptotically to zero as $M$ reaches $M_{\text{max}}$. For this purpose, the recurrence relationship has to be defined as (Page, 1968; Cornell and Vanmarcke, 1969)

$$N(M) = N(M_{\text{min}}) \frac{\exp\left(-\beta(M - M_{\text{min}})\right) - \exp\left(-\beta(M_{\text{max}} - M_{\text{min}})\right)}{1 - \exp\left(-\beta(M_{\text{max}} - M_{\text{min}})\right)}$$ (13)

Many investigators have suggested other alternative models with faster decay for larger magnitudes to avoid the estimation of the maximum magnitude, which generally suffers from large uncertainties (Bollinger et al., 1992; Kijko and Graham, 1998; Kijko, 2004). For example, Merz and Cornell (1973) used a quadratic model, and Lomnitz-Adler and Lomnitz (1979) suggested a double exponential model. Some more sophisticated models have been proposed by the later studies (e.g., Main and Burton, 1984; Main, 1996; Kagan, 1991, 1997; Burroughs and Tebbens, 2002).
The relation of Equation (13) is known as constant seismicity model, because it approaches zero asymptotically as \( M \) approaches \( M_{\text{max}} \), without altering the number of lower magnitude earthquakes. Thus, lowering of \( M_{\text{max}} \) will result in lower moment release rate, if it is not compensated by increasing the total number of earthquakes, \( N(M_{\text{min}}) \). By relating the seismic moment, \( M_0 \), to the earthquake magnitude, \( M \), with an empirical relation of the form, \( \log M_0(M) = c + dM \) (Hanks and Kanamori, 1979), the relationship of Equation (13) can be used to obtain the following relationship for the moment release rate:

\[
\dot{M}_0 = N(M_{\text{min}}) \frac{\exp(-\beta(M_{\text{max}} - M_{\text{min}}))}{1 - \exp(-\beta(M_{\text{max}} - M_{\text{min}}))} M_0(M_{\text{max}}) \frac{b}{d - b} \tag{14}
\]

For a given value of \( \dot{M}_0 \), the use of numbers \( N(M_{\text{min}}) \) obtained from this expression in the recurrence relation of Equation (13) will ensure the conservation of moment release for varying \( M_{\text{max}} \). The moment rate \( \dot{M}_0 \) can be obtained from \( \dot{M}_0 = \mu \dot{u} A \) (Brune, 1968), where \( \dot{u} \) is the geologically estimated long-term slip rate, \( A \) is the total fault rupture area, and \( \mu \) is the shear modulus of the rock mass at the fault. The constraint imposed by fault slip rate allows the development of fault-specific recurrence relationship in regions where the historical seismicity data are only sufficient to establish the regional recurrence rate for small-to-moderate size earthquakes. The expression of Equation (14) can also be used to determine the upper bound magnitude from knowledge of the fault slip rate for given values of \( N(M_{\text{min}}) \) and \( b \) from historical seismicity.

The exponentially decaying recurrence model of Equation (13) is able to describe the observed seismicity in non-fault-specific area type of sources. For many of the individual faults, the characteristic magnitude recurrence model due to Youngs and Coppersmith (1985) can describe better the behaviour of the observed seismicity. Certain faults are seen to generate repeatedly the maximum earthquakes in a narrow magnitude range with a much higher occurrence rate than that predicted by the recurrence relationship for smaller magnitudes on the same fault. This has given rise to the concept of characteristic earthquakes (Wesnousky et al., 1983; Schwartz and Coppersmith, 1984). The characteristic model assumes that more of the seismic energy is released by large magnitude events than that in the exponential model. The magnitude distribution for characteristic earthquakes is assumed to be uniform over the range \( M_c = (M_{\text{max}} - \Delta M_c) \) to \( M_{\text{max}} \). This is taken equal to the probability density at magnitude \( M' = (M_c - \Delta M') \), as defined by the exponential distribution of Equation (13) fitted to the earthquake data up to the magnitude \( M_c \). The characteristic recurrence model with \( N(M_{\text{min}}) \) as the rate of non-characteristic earthquakes, that is the total number of earthquakes in the magnitude range, \( M_{\text{min}} \) to \( M_c \), can thus be written as (Youngs and Coppersmith, 1985)

\[
N(M) = \begin{cases} 
N(M_{\text{min}}) \frac{\exp(-\beta(M - M_{\text{min}})) - \exp(-\beta(M_c - M_{\text{min}}))}{1 - \exp(-\beta(M_c - M_{\text{min}}))} + \dot{n}(M_c) \Delta M_c; & M_{\text{min}} \leq M < M_c \\
\dot{n}(M_c)(M_{\text{max}} - M); & M_c \leq M < M_{\text{max}} 
\end{cases}
\tag{15}
\]

In this expression, \( \dot{n}(M_c) \) is the probability density for the occurrence rate of the characteristic earthquakes, which is taken equal to the rate density at magnitude \( M' \), as given by the exponential distribution for magnitudes up to \( M_c \):

\[
\dot{n}(M_c) = N(M_{\text{min}}) \frac{\beta \exp(-\beta(M' - M_{\text{min}}))}{1 - \exp(-\beta(M_c - M_{\text{min}}))} \tag{16}
\]

Similar to that for Equation (14), the moment release rate for the recurrence relationship of Equation (15) can be obtained as
\[
\dot{M}_0 = N(M_{\text{min}}) \frac{\exp(-\beta(M_c - M_{\text{min}}))}{1 - \exp(-\beta(M_c - M_{\text{min}}))} M_0(M_{\text{max}}) \left[ \frac{b}{d-b} 10^{-dM_c} + \frac{b \exp(\beta)(1-10^{-dM_c})}{d} \right]
\]

The use of \(N(M_{\text{min}})\) obtained from this expression for a given \(\dot{M}_0\) into Equation (15) provides the characteristic recurrence model with constant moment release rate.

For the purpose of illustration, Figure 8 shows the abruptly truncated exponential model, asymptotically decaying exponential model, and the characteristic earthquake recurrence model for a moment release rate of \(\dot{M}_0 = 1.0 \times 10^{25}\) dyne-cm/year, \(M_{\text{min}} = 3.8\), and \(M_{\text{max}} = 8.0\). For the characteristic model, \(\Delta M_c\) and \(\Delta M'\) are both taken as 0.8. As mentioned before, for a fixed \(\dot{M}_0\), a change in \(M_{\text{max}}\) causes a change in the number of all the lower magnitude earthquakes in all the recurrence models. On the other hand, for a fixed total number \(N(M_{\text{min}})\) in the constant seismicity models, the change in \(M_{\text{max}}\) causes a change in the recurrence curves in the vicinity of magnitude \(M_{\text{max}}\) only. These characteristics are illustrated for the asymptotically decaying recurrence model in Figure 9, which shows the results for \(M_{\text{max}} = 6.0, 7.0\) and 8.0. For the constant seismicity case, the number \(N(M_{\text{min}})\) is kept fixed and equal to the number for \(M_{\text{max}} = 7.0\) in the constant moment case. The truncated exponential and the characteristic models will also show similar behaviour.

From the above description it is apparent that the specification of the magnitude-recurrence relation for a source zone may be associated with considerable epistemic uncertainties. There may be uncertainties in use of the exponential, characteristic, or some other model, and also as to the use of the constant-seismicity- or the constant-moment-release-rate model. The choice of the lower threshold and the maximum magnitudes, as well as the estimation of the moment release rate, may also be associated with some uncertainties. Further, depending on the empirical conversion relations used for homogenization of magnitude, criteria adopted for removal of dependent events, and the method used for identification of the
periods of completeness for different magnitude ranges, the recurrence parameters $a$ and $b$ may vary substantially. Lastly, all the other things being the same, the values of $a$ and $b$ may also vary with the method of estimation (e.g., least squares, maximum likelihood, or maximum entropy method). To account for the epistemic uncertainties in the recurrence relationship by the logic-tree approach, one may have to deal with several recurrence models with parameters varying over wide ranges as shown in Figure 3. However, rather than considering a large number of separate options, Lee and Trifunac (1985) have proposed to account for the random uncertainties in parameters $a$ and $b$ and that in $M_{\text{max}}$ by using a Bayesian estimate $N(M)$, obtained by multiplying the expected value of $N(M)$ with the probability of $M \leq M_{\text{max}}$.

Fig. 9 Comparison of the exponential recurrence model with constant-seismicity-rate (thin curves) and constant-moment-release-rate (thick curves) constraints

ATTENUATION AND SCALING RELATIONSHIPS

An attenuation or scaling relationship is required to obtain the probability, $q(z > z|M, R)$, of exceeding a specified value, $z$, of a hazard parameter, $Z$, due to an earthquake of magnitude $M$ at a source-to-site distance $R$. A median attenuation or scaling relationship is commonly developed by fitting a simple equation in terms of a limited number of earthquake and site parameters to the $z$-values observed during past earthquakes. For areas deficient in recorded data, simulated data using seismological source model approach have been also used to develop the attenuation relations for some of the hazard parameters (e.g., Huang and Huo, 1997; Gregor et al., 2002). A median attenuation relation is seen to be associated with large random uncertainties in that the observed or simulated $z$ values are generally scattered widely. This scattering can mainly be attributed to not considering the dependence on several parameters (e.g., stress-drop, radiation pattern), possible random errors in the values of the governing parameters (e.g., magnitude, distance, and site condition), and to the use of a simplified and idealized form for the attenuation equation. To quantify the random scattering in the data, the residuals between the observed values and the corresponding model predictions are defined by suitable probability distributions, due to which the attenuation relations become probabilistic in nature.
The random (aleatory) uncertainties in the attenuation and scaling relations can, in principle, be reduced to some extent by incorporating additional governing parameters in the model and by using more complicated functional forms for the attenuation equation. But in reality, it may not be possible to define accurately the values of the additional parameters and to get stable estimates of the added regression coefficients involved. Thus, the reduction in the aleatory uncertainties may be offset by an increase in the epistemic type of uncertainties in specifying the values of the input governing parameters and by inaccuracies in estimating the regression coefficients. Thus, unlike other input quantities to the PSHA, the classification of uncertainties as aleatory and epistemic in case of attenuation relationships is somewhat dubious (Atkinson and Boore, 1997; Toro et al., 1997). Therefore, only simple attenuation models with a limited number of parameters are used in practical applications. However, due to a limited database available in most real situations, the estimated mean or median relationship as well as the distribution of the residuals is generally associated with significant epistemic uncertainties. As these uncertainties cannot be defined directly from the database, several different attenuation relations with appropriate weights are used to account for their effects in practical applications (e.g., Sabetta et al., 2005; Bommer et al., 2005). This section presents a brief description of the attenuation and scaling relationships for several hazard parameters of importance to the safety of man-made structures.

1. Ground Motion Amplitudes

A complete description of the ground motion for earthquake engineering applications is provided by the acceleration time-histories for three orthogonal components of motion. However, it is not feasible to develop the attenuation relations directly for the acceleration time-histories. Therefore, the commonly used engineering practice is to synthesize the acceleration time-histories compatible to the response spectra (e.g., Seed et al., 1976; Mohraz, 1976; Newmark and Hall, 1982) by the peak ground acceleration (PGA), which is equivalent to the zero-period amplitude of the absolute acceleration normalized spectral shape (e.g., Seed et al., 1976; Mohraz, 1976; Newmark and Hall, 1982; Gupta and Joshi, 1993; Abrahamson, 1998). In this sense, the response spectrum is commonly considered to represent the intensity of ground motion, though in reality it represents the maximum response of a single-degree-of-freedom oscillator. An early approach to obtain the response spectra was to scale a normalized spectral shape (e.g., Seed et al., 1976; Mohraz, 1976; Newmark and Hall, 1982) by the peak ground acceleration (PGA), which is equivalent to the zero-period amplitude of the absolute acceleration spectrum. Most of the attenuation relations in the past were therefore developed for the PGA only. However, it is now well recognized that a normalized spectral shape is unable to represent the dependence on earthquake magnitude, distance, and site condition in a realistic way (Trifunac, 1992; Gupta, 2002b). A more appropriate method to estimate the response spectrum ordinates at different natural periods is to use the empirical attenuation relationships directly for the spectral amplitudes at each natural period (Trifunac, 1976b, 1978). Most of the recent attenuation relations have been, therefore, developed for both the peak acceleration and the response spectrum amplitudes at different natural periods or frequencies (Douglas, 2003). Alternatively, the acceleration time-histories can also be synthesized from the Fourier amplitude spectrum (Trifunac, 1971; Wong and Trifunac, 1979). Though some empirical attenuation relations are available for the Fourier spectrum amplitudes at different wave-periods or frequencies (e.g., Trifunac, 1976b, 1987, 1989; McGuire, 1978), the prediction of the Fourier amplitude spectra is more commonly based on the source-model approach (e.g., Petukhin et al., 1999; Sokolov et al., 2000).

The available studies on frequency-dependent attenuation relations have used widely differing functional forms, different types of earthquake magnitude (moment, surface-wave, or body-wave), and different measures of the source-to-site distance (epicentral, hypocentral, closest distance to the rupture surface, or closest distance to the surface projection of the rupture plane). Also, the site condition in different relations has been defined in widely varying ways, ranging from qualitative descriptions of the near-surface material to quantitative definitions based on shear-wave velocity. Nonlinear soil behaviour has been also accounted in some of the relations (Tsai, 2000; Atkinson and Boore, 2003). Following the work of Trifunac (1987) for the Fourier amplitude spectrum, Lee (1987) developed the attenuation relations for response spectrum amplitudes considering the effects of both local geological condition up to depths of a few kilometers and site soil condition up to 200 m depth. These relations have also accounted at each frequency the magnitude and distance saturation effects as well as the variation of geometrical spreading with distance, and they are thus considered to possess the properties desired on physical grounds. Many of the available relations lack in some or the other of these fundamental requirements, and hence the future developments are required to take these into account.

A site-specific estimation of design ground motion needs the attenuation relations based on the strong-motion data recorded in the target area of interest. However, the required data is either lacking or
inadequate for many parts of the world. It thus becomes necessary to use the relations based on the global data or those developed for some other regions. Due to strong regional dependence, the selection of suitable attenuation relations from the available relations for other host regions is not a straightforward task. Several different relations are thus required to be used with appropriate weights as indicated in the logic-tree diagram of Figure 3. The uncertainties arising due to the inability of defining the ground motion attenuation model for an area in a unique way is found to be a major source of uncertainty in the seismic hazard assessment (Stepp et al., 2001, Sabetta et al., 2005).

The initial selection of the ground motion relations is normally based on the geo-scientific criteria like similarities in the tectonic setting (e.g., compressional or extensional regime), source characteristics (e.g., stress drop), and the anelastic attenuation modeled by the $Q$-factor. As this selection may suffer from considerable personal judgment and biases, many investigators have proposed simple numerical criteria for updating and ranking the initial choice. The simplest update may be to adjust a selected attenuation relation by a constant scale factor to have closer fitting to the strong-motion data for the target region, if available. The hybrid empirical approach due to Campbell (2003, 2004) may provide a more comprehensive way for the purpose. Scherbaum et al. (2004) have proposed simple numerical criteria using available limited data to rank the selected and updated attenuation relations for their appropriateness for the target region, the application of which has been illustrated in some other studies (Cotton et al., 2006; Douglas et al., 2006). The ranking methodology has been also used to assign branch weights in the logic-tree for the ground attenuation model (Scherbaum et al., 2005; Bommer et al., 2005). However, the updating and ranking is generally based on very limited data from one or two earthquakes, which may sometimes lead to highly unrealistic results. Instead, one may impose higher confidence in a relation based on a very large worldwide database and accounting for the various dependencies in a physically realistic way.

In addition to updating for the fundamental differences between the target and the host regions, to combine several attenuation relations in a logic-tree, it is also necessary to make them uniform with respect to the definitions of the various governing parameters (Bommer et al., 2005). The effect of such conversions for the type of horizontal component of ground motion, magnitude scale, source-to-site distance, site condition, and the type of faulting, on the response spectral amplitudes computed from five typical attenuation relations (Abrahamson and Silva, 1997; Boore et al., 1997; Sabetta and Pugliese, 1996; Lussou et al., 2001; Berge-Thierry et al., 2003) is illustrated in Figure 10. The upper left panel in this figure shows the median spectra on the rock type of site condition as obtained by taking the distance as 5 km and magnitude as 5.0 in the original attenuation relations, with no regard to their compatibility. The lower left panel shows the corresponding spectra after converting all the relations into moment magnitude, shortest distance to the surface projection of fault rupture, and the geometric mean of the two horizontal components. For the relations that include the style of faulting as a predictor variable, the effect has been removed by assuming a reverse faulting with dip angle of 50°. The results in Figure 10 indicate that the homogenization of the attenuation equations may help in reducing the epistemic uncertainty to some extent. However, the reliability and applicability of such conversions to a target area of interest cannot generally be established. Further, Bommer et al. (2005) have proposed to carry across the random variability associated with the empirical conversion relations used for homogenization by enhancing the aleatory uncertainties in the original ground-motion relations, which cannot be considered appropriate on physical grounds. The upper and lower right panels in Figure 10 show the comparison between the original and the enhanced standard deviations for the adjusted attenuation equations.

On physical grounds, the conversion of the type of governing parameters in an attenuation equation should sometimes also help in reducing the errors. For example, the attenuation relationships in terms of the closest distance to fault rupture, $R_{\text{rup}}$, are expected to be characterized by smaller dispersion than those in terms of the epicentral distance, $R_{\text{epic}}$. Thus, enhancing the standard deviation for an equation adjusted from $R_{\text{epic}}$ to $R_{\text{rup}}$ is not appropriate. Similarly, an attenuation relationship for the random horizontal component is expected to show larger scattering than that in terms of the geometric mean of the two horizontal components, and enhancing the variability for such a conversion is also not reasonable. In addition, the transfer of aleatory component of the epistemic type of uncertainties into the aleatory type of uncertainties in the basic PSHA is not in order, because the effects of the epistemic and the aleatory uncertainties are accounted differently. Thus, it is proposed that the aleatory uncertainties in the conversion relations for various parameters in the attenuation relations be accounted by using the
expected estimate of the converted attenuation relationship without any change in the original aleatory uncertainties. However, if several different conversion relations are the likely candidates, one may consider additional branches in the logic tree for each set of conversion relations.

Fig. 10 Comparison of the original median acceleration response spectra and associated standard deviations obtained from five selected attenuation relations (upper panels) with those obtained after modifying the relations for the compatibility of various governing parameters (lower panels) (the results correspond to a magnitude 5.0 earthquake at a distance of 5 km and rock type of site condition; after Bommer et al., 2005)

2. Strong-Motion Duration

In addition to the amplitudes, it is also necessary to define the duration of strong motion to estimate the potential of an earthquake to cause damage to the structure at a site (Jeong and Iwan, 1988; Anderson and Bertero, 1991) and ground failure by liquefaction (Todorosvka and Trifunac, 1999). However, for use in different applications, the strong-motion duration is defined in several different ways (Theofanopulos and Watabe, 1989; Kawashima and Aizawa, 1989; Bommer and Martinez-Pereira, 1999). Also, duration depends on frequency (Bolt, 1973), and frequency-dependent duration is required for assigning the duration to generate the synthetic accelerograms (Wong and Trifunac, 1979; Gupta and Joshi, 1993). Some studies (Mohraz and Peng, 1989; Gupta and Trifunac, 1998) have also introduced the role of structural frequency and damping into the definition of duration. Most of the available studies are however concerned only with proposing new definitions of the duration, rather than developing prediction models needed for the PSHA studies.

The early studies on developing the prediction models for frequency-dependent duration are due to Trifunac and Westermo (1976, 1977, 1982). They defined the duration in several frequency bands as the sum of the separate strong-motion portions during which the Husid plot has the steepest slope and gains 90% of its final value. Novikova and Trifunac (1993, 1994a, 1994b, 1995) updated the early studies by using a much bigger database of uniformly processed accelerograms for the California region. They presented a family of prediction models, which can be used straightaway in seismic hazard analysis. Their
fundamental model for the durations of horizontal and vertical components is defined in terms of magnitude \( M \) and epicentral distance \( \Delta \), for a frequency band with central frequency \( f \) as

\[
dur^{(h/v)}(f) = a_i^{(h/v)}(f) + a_2(f)M + a_3(f)M^2 + a_4(f)\Delta
\]

(18)

In this expression, if \( M \) is less than \( M_{\text{min}}(f) = -a_2(f)/(2a_3(f)) \), it is replaced by \( M_{\text{min}}(f) \). To consider the effects of depth of sediments at the site, characteristic horizontal distance, \( R \), to the nearest rock outcrop capable of producing reflections, and the angle, \( \varphi \), subtended at the site by the surface of the reflecting rock, Novikova and Trifunac (1993) obtained some more comprehensive models by incorporating additional terms in the basic relationship of Equation (18).

Two relatively simple models also have been presented by considering the effect of only the depth of sediments at the site in one model, and that of only \( R \) and \( \varphi \) in the other model. Another simplified model, which may find wider practical applications, has been defined by adding the following two terms to characterize the site geology in a qualitative way:

\[
a_{13}(f)S^{(1)} + a_{14}(f)S^{(0)}
\]

(19)

where \( S^{(1)} \) and \( S^{(0)} \) are the index variables for the site geological condition defined as

\[
S^{(1)} = \begin{cases} 
1; & \text{Intermediate or undefinable} \\
0; & \text{Otherwise}
\end{cases} \quad \text{and} \quad S^{(0)} = \begin{cases} 
1; & \text{Sites over deep sediments} \\
0; & \text{Otherwise}
\end{cases}
\]

(20)

Another useful model considering the effects of both the local geological and site soil conditions is obtained by adding the following terms to the fundamental model of Equation (18):

\[
a_{15}(f)(2-s) + a_{11}(f)S^{(1)}_L + a_{12}(f)S^{(2)}_L
\]

(21)

In this expression, parameter \( s \) defines the site geological condition (\( s = 2 \) for basement rock, \( 0 \) for deep sediments, and \( 1 \) for intermediate or indefinable type of sites), and \( S^{(1)}_L \) and \( S^{(2)}_L \) are the index variables for the local soil condition defined as

\[
S^{(1)}_L = \begin{cases} 
1; & \text{Stiff soil sites} \\
0; & \text{Otherwise}
\end{cases} \quad \text{and} \quad S^{(2)}_L = \begin{cases} 
1; & \text{Deep soil sites} \\
0; & \text{Otherwise}
\end{cases}
\]

(22)

The values of the coefficients \( a_i^{(h)}(f) \), \( a_i^{(v)}(h) \), \( a_2(f) \), \( a_3(f) \), and \( a_4(f) \) in Equation (18), as well as the other coefficients in Equations (19)–(22), have been estimated by Novikova and Trifunac (1993) for 12-frequency bands with central frequencies ranging from 0.075 to 21.0 Hz. To define the probability of exceeding a specified value, \( \text{dur}(f) \), of the strong motion duration in a particular frequency band, Novikova and Trifunac (1993) have also presented probability density function for the relative residuals, \( \varepsilon = \text{dur}(f)/\text{dur}_\text{model}(f) \), where \( \text{dur}_\text{model}(f) \) is the duration estimated from a predictive model for specified \( M \), \( \Delta \), and site condition. This density function is given by

\[
p(\varepsilon) = \frac{1}{\eta} \frac{\varepsilon^b}{a + \varepsilon^c}; \quad \eta = a\frac{b+1}{c} \frac{\pi}{c} \left[ \sin \left( \frac{b+1)\pi}{c} \right) \right]^{-1}
\]

(23)

The coefficients \( a \), \( b \) and \( c \) in this relationship are also estimated by Novikova and Trifunac (1993) for the various predictive models by using the observed values of \( \varepsilon \) for each of the 12 frequency bands. The probability, \( q(\text{dur}(f) \mid M, \Delta) \), of exceeding a specified duration, \( \text{dur}(f) \), due to magnitude \( M \) at distance \( \Delta \) can thus be obtained using the density function of Equation (23) as

\[
q(\text{dur}(f) \mid M, \Delta) = 1 - \int_0^{\varepsilon} p(x)dx
\]

(24)

This can be used to carry out the PSHA studies for the strong motion duration in different frequency-bands for the same description of the seismicity as that used for the ground motion evaluation.
3. Critical SPT Value for Initiation of Liquefaction

Initiation of liquefaction in water-saturated cohesionless sands occurs when the effective stress in the ground is reduced to zero. The methodology of probabilistic seismic hazard analysis can also be used to estimate the likelihood of liquefaction at a site during specified exposure time. The basis for this is provided by the study of Trifunac (1995) on the occurrence or nonoccurrence of liquefaction due to specified earthquake magnitude and distance. Using 90 worldwide observations, Trifunac (1995) has proposed five different empirical models to obtain the standard penetration test (SPT) value, $\bar{N}$, corrected for the overburden pressure $\sigma_0$, that separates on average the observed cases of liquefaction from those of no liquefaction. These models are based on the seismic energy and are functions of $\sigma_0$ and the earthquake magnitude, distance and site geological condition. The model prediction can be viewed as a critical value, $N_{\text{crit}}$, of $N$ for liquefaction to occur under specified conditions. Liquefaction will occur at a site if the actual $N$ value is smaller than the estimated $N_{\text{crit}}$ value.

As the ground motion and the site characterization are associated with many inherent uncertainties, the observed data points are found scattered randomly around the model predictions. To consider the points in the database that violated the prediction, Trifunac (1995) defined the probability of liquefaction using a Gaussian probability distribution as the probability of $N_{\text{crit}}$ being greater than the actual $N$ value:

$$\text{Prob}\{N_{\text{crit}} > N\} = \frac{1}{\sqrt{2\pi}\sigma_N} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left(\frac{x - \mu_N}{\sigma_N}\right)^2\right) \, dx$$

In this equation, $\mu_N$ is the mean value and $\sigma_N$ the standard deviation of the corrected SPT value from the model. From the “$v^2 \cdot \text{dur}$” model of Trifunac (1995), $\mu_N$ can be obtained as

$$\mu_N = 87.3 \left(\frac{v_{\text{max}}^2 \cdot \text{dur}}{\sigma_0^{1/2}}\right)^{1/1.4} + 0.95$$

In this expression, $v_{\text{max}}$ is the peak ground velocity in cm/sec, which can be obtained from the following empirical relationship (Trifunac, 1976a):

$$\log_{10} v_{\text{max}} = \log_{10} A_0(R) + 3.059M - 0.201M^2 - 0.134s - 9.8135$$

Here, $\log_{10} A_0(R)$ is the Richter’s attenuation factor, and $s$ takes values of 0, 1 and 2 to define the local geological condition as explained earlier. Also, from the relationships due to Novikova and Trifunac (1993), the strong-motion duration $\text{dur}$ in Equation (27) can be defined as

$$\text{dur} = 7.8 - 3.86M + 0.57M^2 + 0.07R + 1.14 \left(-\frac{s}{2}\right)$$

For the model of Equation (26), the value of $\sigma_N$ is specified to be 5.5 (Trifunac, 1995). Thus, the probability of Equation (25) is equivalent to the conditional probability that liquefaction will occur at a site due to earthquake magnitude $M$ at distance $R$. This provides a basis to carry out the PSHA to evaluate the average return period of occurrence of liquefaction at a site with given $\bar{N}$ and $\sigma_0$ values. Equivalently, one can estimate the probability that liquefaction will occur during a specified exposure period. For a given value of $\sigma_0$ at the site, it is also possible to estimate with a specified confidence level the value of $\bar{N}$ for which liquefaction may occur during a given exposure period.

4. Permanent Fault Displacement

In addition to the ground motion amplitudes and the strong-motion duration, the estimation of hazard in terms of the permanent dislocations across faults may be required for situations like bridges, tunnels, aqueducts, and water and gas lines crossing over faults. For this purpose, it is necessary to predict the probability of exceeding a specified value of displacement due to a given magnitude of earthquake at a
given location on a fault plane. Unlike for other hazard parameters, the earthquakes on the fault of interest only are to be considered for assessing the hazard of permanent displacement. Further, one has to consider only those earthquakes on the fault, which will be able to cause fault rupture reaching the ground surface as well as the site on the fault. Thus, the conditional probability that the displacement at a site on the fault will exceed a specified value, \(d\), due to an earthquake of magnitude \(M\) at distance \(R\) on the fault can be defined as

\[
q(D > d | M, R) = \text{Prob}(D > d | M, R) \times \text{Prob( Rupture breaks the ground surface)} \times \text{Prob( Rupture extends horizontally to the site)}
\]

(29)

Depending on the assumptions regarding the distribution of earthquake locations on the fault plane and the direction of rupture (unilateral or bilateral), there may be several different ways to define the second and third probabilities on the right hand side of the above expression. For a fault plane with given length and width, Todorovska et al. (2007) have proposed a simple way to define these probabilities using the mean values and the standard deviations of rupture width and rupture length defined by empirical equations in terms of earthquake magnitude.

The mean value \(\mu\) and the standard deviation \(\sigma\) of the logarithm of the displacement \(d\) are defined by Todorovska et al. (2007) by the following empirical relationship:

\[
\mu = M - 2.2470 \log_{10}(\Delta) + 0.6489M + 0.0518\times 2 - 0.3407\nu - 2.9850 - 0.1369M^2 - 0.0306 + \log_{10}2 - 0.0090; \quad \sigma = 0.3975
\]

(30)

where \(M\) is the earthquake magnitude, \(\Delta\) is the representative source-to-station distance, \(L_R\) is the rupture length, and \(\nu\) represents the direction of motion (\(\nu = 0\) for horizontal and 1 for vertical component). The distance \(\Delta\) depends on both the physical distance and the size of the rupture as

\[
\Delta = S \left\{ \ln \frac{R^2 + H_R^2 + S^2}{R^2 + H_R^2 + S_0^2} \right\}^{-1/2}
\]

(31)

where \(H_R\) is the focal depth, \(S\) is the source dimension, and \(S_0\) is the source coherence radius. To estimate \(\mu\) from Equation (30), \(R\) is taken as zero, and \(H_R\) is taken as the depth to the center of the fault width. The source dimension, \(S\), has been defined by Todorovska et al. (2007) as

\[
S = \begin{cases} 0.0729(5.5 - M) \times 10^{0.5M} & ; \ M < 4.5 \\ -25.34 + 8.51M & ; \ 4.5 \leq M \leq 7.25 \end{cases}
\]

(32)

\(S_0\) is proposed to be taken as half of the smaller of \(S\) and \(S_f\), with \(S_f\) defined as

\[
S_f = \begin{cases} L_R(M) & ; \ M < 3.5 \\ L_R(M)/2.2 + W_R(M)/6 & ; \ 3.5 \leq M \leq 7.0 \\ L_R(M_{max})/2.2 + W_R(M_{max})/6 & ; \ M > M_{max} = 7.0 \end{cases}
\]

(33)

where \(L_R(M)\) and \(W_R(M)\) are the median values of the fault rupture length and width for magnitude \(M\). Using the \(\mu\) and \(\sigma\) values obtained as above, Todorovska et al. (2007) have approximated the probability, \(\text{Prob}(D > d | M, R)\), by a lognormal distribution as

\[
\text{Prob}(D > d | M, R) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\log d}^\infty \exp \left\{-\frac{1}{2} \left( \frac{\log x - \mu}{\sigma} \right)^2 \right\} \, dx
\]

(34)

It is thus possible to compute the probabilistic hazard of permanent dislocation at a site on a fault for a given description of the expected seismicity.
SENSITIVITY ANALYSIS OF PSHA

To illustrate the sensitivity of the PSHA results to the possible aleatory uncertainties in the various input models and their parameters, example results are computed for a hypothetical seismic source, i.e., a 200 km long straight line vertical fault, with the site located at a closest distance of 10 km from the midpoint of the fault trace. To define the seismicity of this fault, the preferred value of the moment release rate $M_0$ is assumed to be $1.0 \times 10^{25}$ dyne-cm/year, that of the $b$ value as 0.9, and that of the maximum magnitude $M_{\text{max}}$ as 8.0. To study the sensitivity of the hazard estimation, the values of the source parameters are varied around the preferred values as assumed. The hazard is evaluated in the form of uniform hazard Fourier spectra (UHFS) for an exposure period of 100 years, using the residual two-step mag-site-soil model due to Trifunac (1987). In this model, the Fourier amplitude spectrum, $FS(T)$, at each of several wave-periods, $T$, is defined by an empirical scaling relationship in terms of earthquake magnitude $M$, the representative source-to-site distance $\Delta$, site geologic condition defined by the parameter $s (= 0$ for deep sediments, 2 for basement rock, and 1 for difficult-to-classify or intermediate type of sites), and site soil condition defined by the parameter $S_l (= 0$ for rock, 1 for stiff, and 2 for deep soil sites). All the results are computed for a focal depth of 10 km. It may be mentioned that the generality of the results will not be affected due to these assumptions regarding the seismic source and its seismicity. Further, only the median estimates of UHFS are presented for the purpose of illustration, because the results for other confidence levels will qualitatively show similar behaviour.

First of all, Figure 11 presents the UHFS for the preferred values of all the source parameters and the seismicity corresponding to three different types of recurrence models, viz., the truncated, exponential, and characteristic models. For each of the models, the estimated seismicity is distributed uniformly over the complete fault length without considering the effect of fault rupture length for different magnitudes. The spectra for the truncated and the exponential models are seen to be quite close, and both are substantially higher than that for the characteristic model. Thus, the spectral amplitudes are seen to be dominated by the larger number of smaller magnitude earthquakes in the truncated and exponential models, and not by the higher moment release rate in the larger magnitude range in the characteristic model. This is further confirmed by the results in Figure 12, which shows the UHFS for the case of exponential recurrence model with three different values of $M_{\text{max}}$. In the low-period range, the spectrum for the lowest $M_{\text{max}}$ of 6.0 is seen to be the highest. Also, the spectrum for the $M_{\text{max}}$ value of 7.0 is throughout higher than that for $M_{\text{max}}$ of 8.0. The increase in the spectral amplitudes for lower values of $M_{\text{max}}$ is also due to the increase in the number of all the earthquakes up to the magnitude $M_{\text{max}}$. However, the increase in the spectral amplitudes is comparatively smaller in the longer period range, because larger magnitudes contribute more in the longer period range. The effect of the larger number of smaller magnitude earthquakes is further illustrated by the results in Figure 13 for varying value of $b$. The spectra for higher values of $b$ are seen to be higher, because the relative number of smaller magnitude earthquakes increases with increase in the value of $b$.

Further, the results in Figure 14 show the variation in the uniform hazard Fourier spectra with change in the moment release rate for a fixed $M_{\text{max}}$ of 8.0. Though a change in $M_0$ changes the number of earthquakes of all the magnitudes by a constant factor, the spectral amplitudes are not seen to change uniformly at all the wave periods. The middle curve in Figure 14 corresponds to the preferred value of $M_0$, whereas the upper and lower curves correspond respectively to twice and half of the preferred value. An increase in $M_0$ is seen to cause comparatively more increase in the longer period amplitudes. This is because larger magnitudes are characterized by higher contents of longer period waves, and also because the ground motion due to smaller magnitudes is not perceptible at larger distances.

The sensitivity of the uniform hazard Fourier spectra to the local soil and the site-geologic conditions has been studied the next. To illustrate only the effect of the local soil condition, Figure 15 presents the UHFS for various types of soil conditions overlying the basement rock. The spectral amplitudes on stiff and deep soil sites are seen to be amplified compared to those on rock sites for periods greater than about 0.34 s, whereas they are deamplified for the smaller periods. Compared to the spectrum for stiff soil site, the spectrum for deep soil condition is seen to have further amplification for the periods greater than...
about 1.6 s. The amplification of the longer period waves is due to the resonance of the soil layer, whereas the attenuation of low period waves is due to lower $Q$-value for the soil layer.

To illustrate the effect of the site geologic condition, Figure 11 presents the UHFS for a deep soil site lying over three different types of geologic site conditions. Compared to the deep soil site on basement rock, the spectrum for the deep soil site on deep sediments is seen to have significant amplification for the periods greater than about 0.11 s, and slight deamplification for the lower periods. The deamplification can be attributed to lower $Q$-value for sediments, whereas amplification is due to the resonance effect. On
the other hand, due to higher $Q$-value, there is slight amplification in the low-period range also; for the intermediate type of site geologic condition, due to smaller impedance contrast, the amplification in the longer period range is smaller than that for the deep sediments. All these observations can be explained on the basis of physical considerations that the amplification effect predominates the anelastic attenuation of the soil layer, and that the anelastic attenuation in the sediments is much less than that in the soil layer (Trifunac, 1990a; Gupta and Joshi, 1996). Thus the uniform hazard spectra are able to account for the dependence on the various governing parameters in a very realistic way.

![Fig. 13 Sensitivity of the UHFS to the $b$-values in the exponential recurrence model with a constant moment release rate](image)

![Fig. 14 Sensitivity of the UHFS to moment release rate in the exponential recurrence model with a constant maximum magnitude](image)
EXAMPLES OF PROBABILISTIC HAZARD MAPPING

The PSHA method can be used to prepare a microzonation map by estimating the values of a hazard parameter at a closely spaced grid of sites in the area of interest. Similar to that for a single site such maps are able to account, in a realistic way, for the effects of the level and distribution of seismicity in various earthquake sources as well as those of the soil and geological features in the area. Several typical examples of the microzonation maps for the Los Angeles metropolitan area, prepared by Lee and Trifunac (1987), Todorovska and Trifunac (1996, 1999), and Trifunac (1990b), are described in this section for the...
purpose of illustration. These studies for the Los Angeles area have considered 29 fault segments and a rectangular area of diffused seismicity as the seismic source zones, which are shown in Figure 17. Each fault segment in this figure is labeled by a serial number followed by two values within parentheses. The first value in the parentheses is the estimate of the moment release rate $M_0$ in dyne-cm/year, and the second value is the activity rate $a$ for an exposure period of 50 years. The value of $b$ for all the faults is taken as 0.86, and the values of $M_{\min}$ and $M_{\max}$ as 2.75 and 7.0, respectively. The expected number of earthquakes in 50 years for different values of central magnitude in the diffused rectangular source, as indicated in Figure 17, is assumed to occur uniformly over the entire source area. The major faults in the Los Angeles metropolitan area and the local geological condition in terms of the depth of sediments in kilometers at 5’×5’ grid points are shown in Figure 18, with grey areas indicating the rock outcrops (Lee and Trifunac, 1987; Todorovska and Trifunac, 1999).

![Diagram](image)

Fig. 17 Various faults and a rectangular area type of source in the California region (the seismic potential of each fault is specified by $M_0$ and activity rate $a$ in the parentheses, and that of the area source by the expected number of earthquakes in different magnitude intervals during a period of 50 years)

1. Microzoning in Terms of PSV Amplitudes

Lee and Trifunac (1987) and Trifunac (1990b) have prepared microzonation maps for the Los Angeles metropolitan area in terms of the pseudo relative velocity (PSV) spectrum amplitudes at different natural periods, using the empirical scaling relations due to Trifunac and Lee (1985). These relations define the spectral amplitude, $PSV(T)$, at period $T$, in terms of the earthquake magnitude, source-to-site distance, and the site geological condition defined by the depth of sediments in kilometers. Using the probability distribution of the residuals, these scaling relations can be used to obtain the conditional probability, $q(PSV(T) | M_j, R_i)$, that a spectral amplitude $PSV(T)$ will be exceeded at a site due to the earthquake magnitude $M_j$ at the distance $R_i$ from the site. Using these probabilities and the seismicity associated with the 29 fault segments and the diffused rectangular area source, PSHA has been carried out
to compute the $PSV(T)$ amplitudes with different confidence levels at each of the 1'×1' grid points. Some typical microzonation maps thus obtained for a confidence level of 0.50 and natural periods equal to 0.04, 0.34, 0.90 and 2.8 s are shown in Figure 19. The microzonation maps for many other natural periods and confidence levels are available in Lee and Trifunac (1987).

**Fig. 18** Thickness of sedimentary deposits in the Los Angeles Metropolitan area at a 5' grid of points, with the grey areas indicating the rock outcrops

**Fig. 19** Typical microzonation maps for the Los Angeles metropolitan area in terms of 5%-damped PSV spectrum amplitude for a confidence level of 0.50 and exposure period of 50 years for $T = 0.04, 0.34, 0.90$ and 2.8 s (after Lee and Trifunac, 1987)
The maps in Figure 19 correspond to the horizontal-component PSV spectrum with 5% damping. By reading the spectral amplitudes from such maps for a series of natural periods and for a particular set of confidence level and damping value, one can readily construct the uniform hazard response spectrum for any site in the area. Examples of such spectra for two typical sites, one at 33°45′N and 118°20′W on rock and the other at 33°45′N and 118°05′W on about 7.5 km thick sediments, are presented in Figure 20. It is seen that the difference between the two spectra is drastic, although both the sites are in the same metropolitan area. Similar situation may exist in the Delhi metropolitan area in India, where the seismicity can be associated with several fault segments and the site soil and geological conditions vary widely.

Fig. 20 Uniform risk spectra for two typical sites in the Los Angeles metropolitan area as obtained from the microzonation maps like those given in Figure 19

2. Microzonation for the Occurrence of Liquefaction

Using the formulation described earlier to evaluate the probability of occurrence of liquefaction at a site with a specified SPT value, $\overline{N}$, and the overburden pressure $\sigma_0$, due to an earthquake magnitude $M_j$ at distance $R$, Todorovska and Trifunac (1999) have prepared microzonation maps for the occurrence of liquefaction in the Los Angeles metropolitan area. Two typical maps showing the average return periods for the occurrence of liquefaction are shown in Figure 21 for two different $\overline{N}$ values and $\sigma_0$ taken as 40 kPa. If a site is characterized by the $\overline{N}$ and $\sigma_0$ values as specified in such a map, the corresponding return period gives the period for the liquefaction to occur at that site. By preparing such maps for a large number of $\overline{N}$ and $\sigma_0$ values, it is possible to identify the average recurrence period for the liquefaction to occur at any site in the area.

For a given value of the overburden pressure, the microzonation maps for the occurrence of liquefaction can also be prepared in terms of the distribution of the critical SPT values for which liquefaction may initiate at a site with a specified probability during a specified exposure period. A typical microzonation map of this type for the Los Angeles metropolitan area is shown in Figure 22. By preparing such maps for several different values of $\sigma_0$, the occurrence of liquefaction with desired probability and exposure period can be found readily at a site with the known $\overline{N}$ and $\sigma_0$ values. Such maps can therefore be considered more useful for practical engineering applications.
Fig. 21 Typical microzonation maps for the Los Angeles metropolitan area showing the average return period for initiation of liquefaction with overburden pressure of 40 kPa and the corrected SPT values of 10 (top) and 20 (bottom) (after Todorovska and Trifunac, 1999)

Fig. 22 Typical microzonation map for the Los Angeles metropolitan area showing the critical SPT values for the overburden pressure of 40 kPa, and probability to liquefy equal to 0.1 during the exposure of 50 years (the liquefaction may initiate at a site if the actual SPT value with the specified overburden pressure is less than that in the map; triangles in this and in the previous figures show the location of the King Harbor site, for which Todorovska and Trifunac (1999) showed the average return period, and liquefaction occurrence hazard versus $\frac{N}{\bar{N}}$)
SUMMARY AND CONCLUSIONS

The paper has presented a critical overview on the various aspects of the currently used probabilistic seismic hazard analysis (PSHA) formulation to compute the probability of exceeding a specified level of a hazard parameter at a site due to any of the earthquakes expected to occur during a given life-period. The PSHA formulation is fundamentally based on evaluating the annual frequency of occurrence for different levels of the hazard parameter due to the total expected seismicity. The probability of exceedance or the occurrence rate, plotted as a function of the level of hazard parameter, is known as the hazard curve. The inherent random uncertainties in the location, magnitude and inter-event times of an earthquake, and that in the observed amplitudes of the hazard parameter, are taken into account in defining the hazard curves. A hazard curve forms the basis for estimating the level of hazard with a desired confidence level or the annual frequency of occurrence.

To consider the randomness in earthquake location by defining the probability distribution of the source-to-site distance, the paper has described four different types of sources, covering the most practical conditions. These include a point source represented by highly concentrated seismicity at a long distance from the site, a line source corresponding to a vertical fault, a dipping fault plane source with specified location and orientation, and an arbitrarily shaped area source of diffused seismicity. For an assumed spatial distribution of the seismicity, it is straightforward to estimate any desired measure of the source-to-site distance and its probability distribution for a line or a dipping fault plane source (Gupta, 2006b). The point source is characterized by only a single fixed value of the distance, which can be defined easily when the fault rupture is not considered. However, for a point source with unknown fault, a fault rupture distance is proposed to be obtained by averaging over the distances corresponding to uniformly distributed fault length in all the directions. Similarly, the probability distribution of a measure of distance for the area source can be obtained by dividing the entire source area into a large number of small-size elements, and by assuming the geometric center of each element to be a point source with suitable weight. Recommendations have been made for using both uniform as well as non-uniform weights for this purpose.

The randomness in earthquake magnitude is considered by defining an appropriate recurrence relationship for each source. Though the exponential model with a lower and an upper bound magnitude is found to be suitable in most cases, the characteristic model may provide better description in the case of some fault-specific sources. Both these models can be defined with a constant seismicity or a constant-moment-release-rate constraint. If the available data on past earthquakes in a source does not fit any of the standard recurrence relationships well, it has been proposed to use the observed probability distribution obtained by summation of a suitable kernel function.

The randomness in the times of earthquake occurrence in a source is considered by defining a probability distribution function for the inter-event times. The PHSA formulation is most commonly based on the Poisson assumption, under which the inter-event times are described by an exponential distribution with a constant occurrence rate. This formulation is also applicable to a non-homogeneous Poisson process characterized by a time-dependent occurrence rate, such as large subduction zone earthquakes. However, the aftershocks and sequential events, which are characterized by strong spatio-temporal correlations, cannot be included in this category. The Monte-Carlo simulation method is commonly used to incorporate the effect of such dependent events into the PSHA formulation. To account for the effect of the dependent events in a more efficient and practically simple way with adequate conservatism, it has been proposed to model their occurrences in a literal way.

To consider the effect of the random scattering in the observed values of a parameter used to characterize the hazard, the PSHA formulation needs the probability distribution of the residuals between the observed and the mean or median estimates of the hazard parameter obtained by fitting an empirical attenuation relationship to the available data. Such probabilistic attenuation or scaling relations have to be specific to the region of an area selected for the mapping of hazard. In the absence of the region-specific relationships, suitably selected and updated relations from other regions are used commonly in practical applications. A larger number of such relations are known widely for the peak ground acceleration and the response spectrum amplitudes. The paper has also described the relations for many other parameters like the Fourier amplitude spectrum, frequency-dependent strong-motion duration, permanent displacement across a fault, and the critical SPT value for the initiation of liquefaction with a specified
overburden pressure at a site, which are not that well known. The mapping of these additional parameters is necessary to have a more realistic assessment of the seismic risk in an area.

Due to limited amount of the available data and incomplete scientific knowledge of the various physical phenomena involved in the generation of earthquakes and their various natural effects, it is generally difficult to define in a unique way the models and the parameters of a model used to describe the inherent randomness in the basic PSHA formulation. Based on subjective decisions by different experts or different interpretations by a single expert, in practical applications, several different alternatives are generally possible even with the same database and the state of knowledge. The inherent randomness and the inability to define this randomness in a unique way are commonly termed as aleatory and epistemic types of uncertainties, respectively. The paper gives a brief description of the logic-tree approach to consider the effects of the possible epistemic uncertainties in the various steps of the basic PSHA formulation. By assigning suitable weights that represent relative levels of confidence imposed in the various options for each of the input elements to the PSHA, the logic-tree approach is used to define all the different combinations of the input elements and the corresponding hazard curves with weights. It generally becomes very difficult to arrive at the final decision due to a wide dispersion in the hazard curves. By computing the example results for the Fourier amplitude spectra for a site, it has been illustrated that large epistemic uncertainties in some of the input model parameters may result in very large uncertainties in the hazard estimate. The effect of the epistemic uncertainties is, therefore, accounted commonly by using a higher level of hazard curve, leading to unduly large conservatism.

To rationalize the use of epistemic uncertainties, the paper has first proposed to reduce the number of hazard curves by treating the aleatory component of the epistemic uncertainty in defining the model parameters by using the Bayesian estimate of the parameters. Then, to the extent possible, a single consensus or a limited number of most appropriate models have been proposed to be used for each of the input elements. A model based purely on speculation and not supported by the available data or not having a plausible physical basis should not be included in the analysis. Finally, to avoid unreasonable conservatism, it has been recommended to use the hazard curve with the maximum weight to represent the decision with the highest level of confidence for the currently available data and knowledge. The use of this hazard curve with a suitably chosen confidence level is considered more rational in practical engineering applications, even if it differs from the reality to be known in future.

The PSHA formulation can thus be used to account for the effects of the random (aleatory) as well as the epistemic uncertainties in a reasonable way. However, perhaps with an intention to get an upper bound value of the hazard, the PSHA is sometimes used with extremely low annual frequencies of the order of $10^{-7}$ to $10^{-8}$ (Abrahamson et al., 2002; Stepp et al., 2001). If applied without imposing a truncation on the distribution of the residuals of the hazard parameter, this may result in unrealistically high estimate of the hazard, whereas truncation cannot be considered to provide a reliable estimate of the hazard for such small values of the annual frequency. It has been thus recommended that the PSHA be used without any truncation for only those applications where the safety requirements are within the practically realizable limits.

To illustrate the application of the PSHA method for hazard mapping, the paper has finally presented typical examples of microzonation maps for several parameters representing different natural effects of the earthquakes. These maps are able to exhibit in a balanced and physically realistic way the effects of the spatial distribution of the seismicity and the site condition in the area of interest. The implementation of the various proposals made in the paper is expected to make the PSHA method more amenable to practical applications for the hazard mapping to estimate the earthquake effects on man-made structures.

ACKNOWLEDGEMENTS

The author is very thankful to Mrs. V.M. Bendre, Director, CWPRS, for taking keen interest in the present study and for permitting to publish this paper.

REFERENCES


Probabilistic Seismic Hazard Analysis Method for Mapping of Spectral Amplitudes and Other Design-Specific Quantities to Estimate the Earthquake Effects on Man-Made Structures


