Bulletin of the Seismological Society of America, 89, 2, pp. 501–520, April 1999

Disaggregation of Seismic Hazard

by Paolo Bazzurro and C. Allin Cornell

Abstract Probabilistic seismic hazard analysis (PSHA) integrates over all potential earthquake occurrences and ground motions to estimate the mean frequency of exceedance of any given spectral acceleration at the site. For improved communication and insights, it is becoming common practice to display the relative contributions to that hazard from the range of values of magnitude, M, distance, R, and epsilon, ε, the number of standard deviations from the median ground motion as predicted by an attenuation equation.

The proposed disaggregation procedures, while conceptually similar, differ in several important points that are often not reported by the researchers and not appreciated by the users. We discuss here such issues, for example, definition of the probability distribution to be disaggregated, different disaggregation techniques, disaggregation of R versus ln R, and the effects of different binning strategies on the results. Misconception of these details may lead to unintended interpretations of the relative contributions to hazard.

Finally, we propose to improve the disaggregation process by displaying hazard contributions in terms of not R, but latitude, longitude, as well as M and ε. This permits a display directly on a typical map of the faults of the surrounding area and hence enables one to identify hazard-dominating scenario events and to associate them with one or more specific faults, rather than a given distance. This information makes it possible to account for other seismic source characteristics, such as rupture mechanism and near-source effects, during selection of scenario-based ground-motion time histories for structural analysis.

Introduction

As probabilistic seismic hazard analysis (PSHA) has become more realistic and comprehensive, it has become common practice to display the relative contributions to the hazard from the different values of the random components of the problem, specifically, the magnitude, M, the source-to-site distance, R, and, often, ε, a measure of the deviation of the ground motion from the predicted (median) value. The results, which are obtained separately for each fault and successively combined for all the faults in the region, are called the disaggregation of the PSHA. The reader should be aware that in a large body of literature, this technique is referred to as deaggregation. The disaggregation of hazard into relative contributions from different sources and earthquake events achieves an important twofold result: better insights and improved communication of the hazard, and a more informed characterization of the ground motion to be expected at the site.

The relevance of seismic hazard disaggregation was pointed out by the National Research Council (NRC) (1988). More recently, however, it has also been recognized by organizations, such as the U.S. Regulatory Commission (NRC) and the U.S. Department of Energy (DOE), which, for the nuclear and non-nuclear facilities under their jurisdiction, have required (see U.S. NRC, 1997; Benreuter *et al.*, 1996; Boissonnade *et al.*, 1995; U.S. DOE, 1996; Kimball and Chander, 1996) that the seismologists and engineers summarize the contributions to hazard by individual magnitude and distance ranges for the ground-motion levels corresponding to the reference probability prescribed by each agency. The SSHAC (1997) report, commissioned by NRC, DOE, and the Electric Power Research Institute (EPRI), has identified this PSHA disaggregation as one of the main elements of the seismic hazard documentation and provides guidance and recommendations on how to perform seismic hazard disaggregation and report its results. The basics of a disaggregation technique have also recently appeared in a textbook on earthquake engineering (Kramer, 1996).

Through the World Wide Web, organizations, such as the U.S. Geological Survey (USGS) and California Division of Mines and Geology (CDMG), have made available to the public different types of representations of disaggregated seismic hazard. USGS, for example, currently displays contributions of M and R ranges for several U.S. cities, and CDMG shows maps of the most probable seismic source dis-

tance and magnitude conditional on the exceedance of different levels of peak ground acceleration (PGA) and different spectral accelerations for counties in the Los Angeles area in southern California (see next section.)

The objectives of this study, as discussed more specifically later in this article, are a closer look at the details of PSHA disaggregation, to ensure that both analysts and users understand this useful tool and do not misinterpret the reported results, and the proposal of a new disaggregation technique. To understand and formally define hazard disaggregation, we must first review the fundamental aspects of PSHA itself.

Basics of PSHA

The probabilistic analysis procedure for the evaluation of the seismic hazard at a site has been long established (Cornell, 1968); it has been widely used and elaborated on by many. The conventional PSHA implies an integration of all the potential magnitudes and source distances to estimate the mean frequencies of earthquake ground motions occurring at the site in any given time period.

For the same local soil conditions, the intensity of the ground shaking at the site depends mainly on the magnitude, M, and source-to-site distance, R, of the causative event. For the same M and R values, however, empirical recordings have shown a great deal of scatter. Such variability is captured by a (standardized Gaussian) variable called *epsilon*, ε, which is defined here as the number of (logarithmic) standard deviations by which the (logarithmic) ground motion deviates from the median value predicted by an attenuation equation given M and R. As a measure for ground-motion intensity, in this work, we will consider linear spectral acceleration, S_a, which is, of course, oscillator frequency and damping dependent. To further clarify the definition of ε adopted here, a generic ground-motion attenuation equation for (logarithmic) S_a would read

$$\ln S_a = g(M, R, \boldsymbol{\theta}) + \varepsilon\sigma_{\ln S_a}, \tag{1}$$

where g represents the predictive functional form used during regression of the strong-motion database, and $\sigma_{\ln S_a}$ is the standard deviation of $\ln S_a$. In general, the empirical coefficients of g also depend on additional variables, $\boldsymbol{\theta}$, such as fault type and local soil conditions at the site.

Note that in the literature, the error term is often defined by a Gaussian variable that corresponds to the product $\eta = \varepsilon\sigma_{\ln S_a}$ here. We separate the error term into two parts in order to make the variable ε independent of M and R. The importance of this remark will be apparent later. Many of the early attenuation relationships, which used more limited ground-motion databases, considered $\sigma_{\ln S_a}$ as a constant. More recently, many researchers (e.g., Abrahamson and Silva, 1997) have found the value of $\sigma_{\ln S_a}$ to be dependent on the earthquake magnitude, M, while others (e.g., Campbell, 1997) prefer to account for the dependence of $\sigma_{\ln S_a}$ on the predicted value of the response parameter (e.g., S_a). In most empirical relationships, $\sigma_{\ln S_a}$ also decreases with increasing oscillator frequency.

More formally, the PSHA methodology allows computation of the mean annual frequency of exceedance, $\lambda_{S_a>x}$, at a site of a specified level x of S_a at an oscillatory frequency f and damping ξ based on the aggregated hazard from N sources located at different distances and capable of generating events of different magnitudes. Mathematically,

$$\lambda_{S_a>x} = \sum_{i=1}^{N} (\lambda_{S_a>x})_i = \sum_{i=1}^{N} \nu_i \left\{ \iiint I[S_a > x|m, r, \varepsilon] \, f_{M,R,\varepsilon}(m, r, \varepsilon) dm dr d\varepsilon \right\}_i, \tag{2}$$

where

- ν_i is the mean annual rate of occurrence of earthquakes generated by source i with magnitude greater than some specified lower bound (e.g., $m = 5.0$).
- $I[S_a > x|m, r, \varepsilon]$ is an indicator function for the S_a of a ground motion (generated by source i) of magnitude m, distance r, and ε standard deviations away from the median with respect to level x. This indicator function is equal to 1 if $\ln S_a(m, r, \varepsilon)$, as computed from equation (1), is greater than $\ln x$ and 0 otherwise.
- $f_{M,R,\varepsilon}(m, r, \varepsilon)$ is the joint probability density function of magnitude, M, distance R, and ε for source i. It should be observed that because ε is stochastically independent of M and R (although $\sigma_{\ln S_a}$ is not functionally so), then $f_{M,R,\varepsilon}(m, r, \varepsilon) = f_{M,R}(m, r)\, f_\varepsilon(\varepsilon)$, where $f_\varepsilon(\varepsilon) = (1/\sqrt{2\pi}) \exp(-\varepsilon^2/2)$ represents the standardized Gaussian distribution.

Modern PSHA allows also the explicit consideration of multiple hypotheses on the input assumptions of the seismicity model, such as, for example, the locations and other characteristics of seismic sources, the recurrence model of different earthquake sizes, the maximum magnitude of each source, and the median ground-motion attenuation with magnitude and distance (see equation 1).

This explicit treatment of model and parameter uncertainty (i.e., epistemic uncertainty) permits the evaluation of not only the mean but also of any desired fractile of the hazard estimate (e.g., the 15th, median, or 84th percentile estimates of the annual frequency of exceeding a specified S_a level). For example, the mean hazard of exceeding a level x of S_a is simply the weighted combinations of all the $\lambda_{S_a>x}$ values obtained from all the cases considered. Each weight expresses the degree of confidence in that particular realization of the seismicity model. A brief discussion on how epistemic uncertainty can be handled during hazard disaggregation is included in the next subsection.

Numerical integration is routinely used because in almost all realistic cases, the previous integrals cannot be solved analytically. The range of feasible values of M, R,

and ε for each source is divided into bins (or segments, or cells) of width Δm, Δr, and $\Delta\varepsilon$, respectively (not necessarily constant throughout the entire domain of values of each variable). The integrals in equation (2) are in practice replaced by summations. This operation implies that each source is capable of causing earthquakes of only a discrete number of magnitudes (usually assumed equal to the central value of each magnitude bin) at a discrete number of distances (usually, but not always, assumed to be equal to the central value of each distance bin), which, in turn, generate at the site ground motions of only a discrete number of standard deviations away from the predicted median motion, given the M and R pair. The only feasible values of ε are also routinely assumed to be coincident with the central values of each ε bin. The accuracy of this numerical procedure obviously increases as the sizes of the bins of all three variables decrease.

Definition of Seismic Hazard Disaggregation

From equation (2), it follows that the hazard is computed for a fixed intensity level, x, of spectral acceleration, S_a, at an oscillator frequency, f, and damping ratio, ξ. On the same lines, hazard disaggregation techniques are rigorously defined only for a given combination of the parameters x and $S_a(f, \xi)$. Unless stated otherwise, in this work, the disaggregation of hazard and, implicitly, any derived quantity (e.g., conditional probability density functions of M, R, and ε, or their summary statistics, such as means and modes) are always referred to such a condition. These quantities may very well be considerably different when one, or more, of the specified parameters is changed.

The disaggregation of the hazard from all N sources combined is usually obtained by accumulating in each 3D M, R, and ε bin the contribution to the global hazard, $\lambda_{S_a>x}$, during the numerical integration of equation (2) and, at the end of the calculations, by dividing the total contribution accumulated in each bin by the value of $\lambda_{S_a>x}$. Formally, therefore, this disaggregation represents the conditional probability distribution of M, R, and ε given the event that S_a exceeds x at the site. In different words, it is the sum of the ν_i-weighted integrands in equation (2), normalized to unit volume. If disaggregating the hazard from the ith fault alone is of interest, the M, R, and ε contributions to hazard from the ith fault only have to be normalized by $(\lambda_{S_a>x})_i$; the result is simply the ith integrand in equation (2), normalized to one.

The procedure previously described computes the relative contributions to the hazard originated by a specific characterization of the seismicity in the study region. When epistemic uncertainty is included in the PSHA calculations, in principle, one could disaggregate the hazard from each considered seismicity model. In realistic applications, this is highly impractical given the very large number of cases typically considered. However, the same procedure outlined earlier can still be applied to disaggregate, for example, the *mean* hazard. Computationally, this can be done by multiplying the hazard contribution in each bin by the weight assigned to the model under consideration. Because the sum of all the weights adds up to 1, the results of this operation provide the relative contributions to the mean hazard. All the disaggregation results for the Los Angeles case study to follow refer to the mean hazard. For brevity, hereafter, we will not distinguish between hazard and mean hazard when discussing disaggregation results.

The hazard can be simultaneously disaggregated in different types of bins, to be utilized and displayed later according to the particular application. We shall return to this matter later. For example, the contributions to hazard can be at the same time accumulated in 1D M bins, in 2D M-R bins, and in 3D M-R-ε bins. Throughout the study, we shall refer to these as 1D, 2D, and 3D hazard disaggregation techniques. In probability theory, these three different representations of the disaggregated hazard are called, respectively, the marginal probability mass function (PMF) of M, and the joint PMF of M-R and of M-R-ε, in all cases conditional on $S_a > x$ at the site. (In what follows, we shall frequently drop this conditional phrase for simplicity.)

For instance, Figure 1 shows the marginal PMFs of M and of R conditional on S_a exceeding 0.41 g at an f of 1 Hz and a ξ of 5% at a downtown Los Angeles site (this case study will be thoroughly discussed in the next section; see Fig. 2). For comparison, see also the joint PMF of M-R for the same site and same hazard level displayed in Figure 3. The joint M-R-ε is not shown here because it requires the equivalent of a 4D plot (e.g., see McGuire, 1995).

Sometimes, representations of the disaggregated hazard, such as those shown in the figures mentioned earlier, are exploited only to estimate the expected or most likely earthquake magnitude and source-to-site distance to cause the exceedance of the specified ground-motion parameter level at the site (i.e., $S_a \geq 0.41$ g in the previous example). In these cases, the results of the hazard disaggregation are interpreted and often condensed into central statistics (such as the means or the modes) with the sole intent of identifying such earthquake events or scenarios typically for developing design ground motions. These parameters are discussed in the next section.

Organization and Focus of This Study

Disaggregating the hazard in terms of M and R has lately become a routine practice in the seismic hazard evaluation community. What is not widely appreciated, however, is that the procedures that have been proposed in the past few years (see next section) for conducting seismic hazard disaggregation, while very similar in concept, are in fact rather heterogeneous in important details. Such methodologies, for example, neither use the same method for disaggregation nor disaggregate the same hazard (e.g., the mean versus the median hazard curve resulting from the multiple hypotheses/uncertainty analysis discussed in a previous subsection). Furthermore, different methods compute the contributions to hazard in terms of different quantities (M and R only or

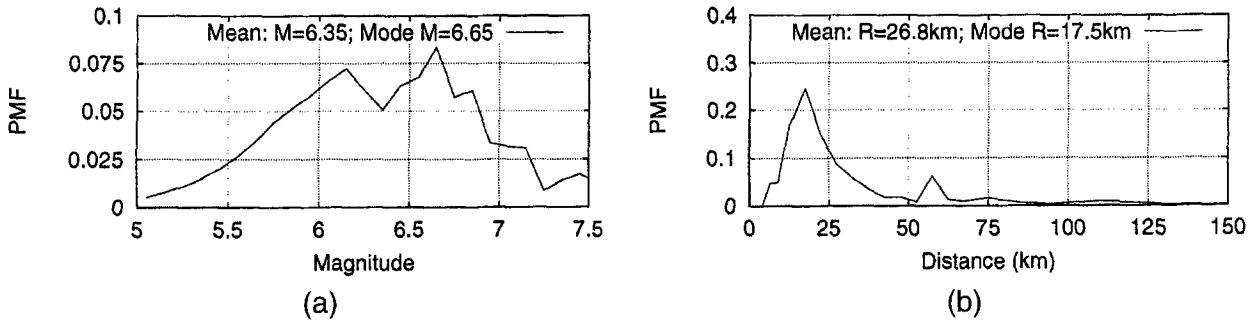


Figure 1. PMFs of (a) M and (b) R conditional on exceeding S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site.

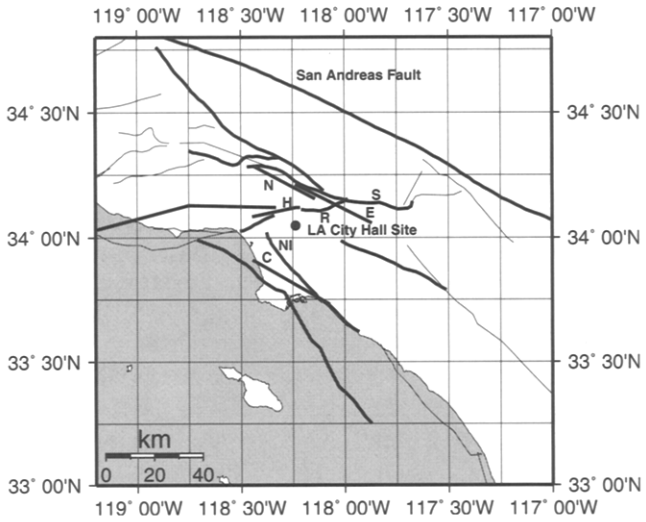


Figure 2. A map of all the major known southern California faults and the location of the Los Angeles City Hall site. Legend of the faults: S (Sierra Madre); N (Northridge); H (Hollywood); R (Raymond); E (Elysian Park); C (Compton Thrust); NI (Newport-Inglewood).

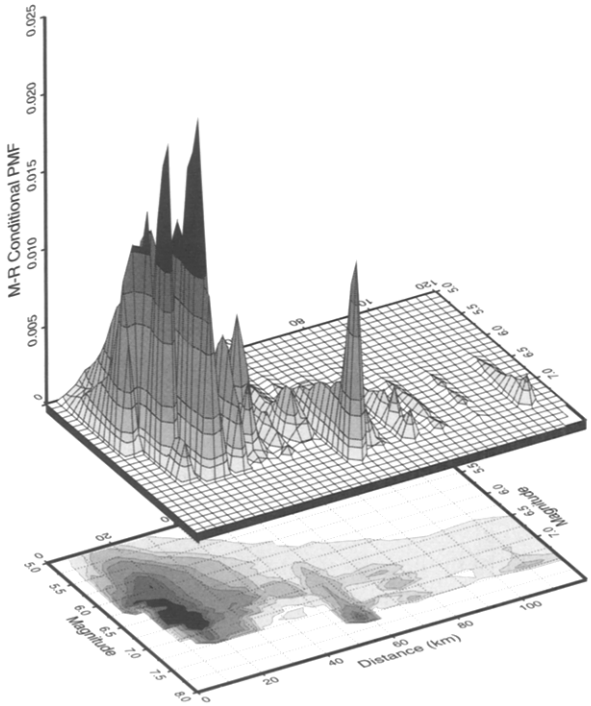


Figure 3. M-R PMF conditional on the exceedance of S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site.

ε as well), and the results are summarized and reported by using different central statistics (e.g., mean versus mode), some of which are more informative than others. Finally, some of the methods disaggregate the hazard conditional on exceeding the target value of the ground-motion parameter (e.g., S_a = 0.41 g) but accumulate the contributions only in M, R, and ε bins such that the target value is equaled when the bin values are included in a ground-motion predictive equation. The purpose of this different disaggregation technique is that it has the desirable property that the modal values of the joint M-R-ε (conditional) PMF, when substituted in the attenuation equation (such as equation 1), produce the exact target value, provided there is only a single such equation used in the analysis. McGuire (1995), who recognizes the common use of weighted, multiple, alternative attenuation equations, uses yet another scheme of disaggregation that has a similar objective (see next section).

While, for background, we call attention to these differences, the intent of this study is not a discussion of the relative merits of each proposed methodology for hazard disaggregation. Rather, the focus will be on the issues often hidden in the mathematical details that, if unstated by authors or unappreciated by users of disaggregated PSHA, can give rise to misleading or unintended results or interpretations. Issues such as disaggregation of PMF versus probability density function (PDF) of M, R (and, possibly, ε), disaggregation of R versus ln R, and the effect on the results of different binning sizes will be addressed in that section.

Finally, we will introduce a refinement to the current

state-of-the-art disaggregation techniques: the contributions to seismic hazard are computed in terms not of R but of latitude and longitude, thereby permitting a display directly on a typical map of the faults of the area surrounding the site. Such hazard contributions can be stored and displayed by means of the Geographic Information Systems tool. At each place of interest, the hazard can be further disaggregated in terms of the contributions by the other two variables, M and ε, due to each fault present at that location. We shall call this method 4D disaggregation. This proposed approach brings fresh perspectives to the understanding of the decomposition of the seismic hazard at a site.

Seismic Hazard Disaggregation Procedures

The early disaggregation studies that appeared in the literature did not compute the relative hazard contribution by different ranges of the three main variables in the PSHA, as the more modern methods do. Only M and R were considered, while the other important random variable (i.e., ε), which describes the departure of the ground motion from its median value (as predicted by an attenuation relation given M and R) was in early disaggregations almost always neglected. Regardless of whether ε was considered or not, these earlier procedures often did not explicitly compute or report the joint distribution (conditional on the exceedance of the target S_a level for a given f and ξ) of the basic PSHA variables. We consider them here because they share one main goal, namely, the identification of the seismic events dominating the hazard at the site, as derived from PSHA.

In the following overview, the focus is devoted mainly to outline similarities and differences of some of the proposed methods. The review starts from the earlier, less-refined methods and evolves to the current best state of practice.

Mean and Modal Values of M and R

Historically, mean values and modal values of M and R have been the two most popular contenders for the role of defining the dominant event. Events inducing the exceedance of any given level of ground-motion intensity (e.g., S_a) computed via PSHA were first summarized (McGuire and Shedlock, 1981) in terms of simply the mean values, $\bar{M}$ and $\bar{R}$, of magnitude and distance. For example, in the example considered in Figure 1, this approach would have reported $\bar{M} = 6.35$ and $\bar{R} = 26.8$ km. These mean values, globally evaluated for all the seismic sources around the site, were used there to further investigate the sensitivity of seismic hazard calculations to statistical uncertainties in models and parameters.

Today, modal values are preferred to means by many. The advantages of using mean values of M and R as final summary statistics are that they are simple to understand, to communicate, and to compute. In most cases, such values represent meaningful summaries, but, rigorously speaking, they do not describe the most likely magnitude or the most likely distance that may induce the specified (or larger) acceleration level at the site (see Fig. 1 and the next subsection for an example). In this respect, the two univariate modal values would be most called for (e.g., $M^* = 6.65$ and $R^* = 17.5$ km in the example in Fig. 1).

Moreover, it is also immediate to conceive of counter-examples where the (hazard-weighted) pair $\bar{M}$ and $\bar{R}$ do not represent a physically realizable earthquake, let alone one that is the dominating event. A site surrounded by two equally hazardous faults—one nearby capable of generating small-magnitude events and one much farther away causing characteristic earthquakes of much larger size—is the simplest example of such cases. The mean distance $\bar{R}$ (given $S_a \geq x$) will lie between the two faults, and the mean magnitude $\bar{M}$ will not be representative of events likely to occur at either fault.

Bivariate modal values (i.e., the peak, or the most probable $M - R$ pair, of the joint $M - R$ probability distribution) overcome this problem because they necessarily refer to an actual realizable source, at least within the resolution of the magnitude and distance binning required to estimate numerically the joint conditional distribution and compute its mode. However, the bookkeeping operation of accumulating the hazard in each single $M - R$ bin makes the computation of modal values more lengthy than that needed for assessing marginal means. It is interesting to note that the property of the bivariate mode (M^*; R^*) of describing a feasible event on a specific fault, in general, is not shared by the univariate modal values, M^* and R^*, of the two marginal distributions of M and R. It is sometimes forgotten, too, that bivariate and univariate modes (unlike means) are not necessarily coincident (see the next subsection for an example).

The use of mean values appears to be questionable also in those cases where different types of source-to-site distance definitions (e.g., closest distance to the rupture zone versus epicentral distance) are used in the same PSHA. This may occur, for example, when multiple attenuation laws are utilized.

A kind of mean value of M and of R was also adopted by Kameda and coworkers (see, e.g., Ishikawa and Kameda, 1988, 1991, 1993; Kameda *et al.*, 1994a, 1994b). They find the means of each of the *single*-fault disaggregations discussed in the previous section, giving as a result a collection of $\bar{M}$ and $\bar{R}$ pairs to be used as scenario events for subsequent use in analysis and design of structures. It should be noted that according to the definition of $\bar{M}$ and $\bar{R}$ proposed in this latter body of work, two different $\bar{M}$ and $\bar{R}$ pairs of values belonging to two area sources represent events that do not have the same target mean frequency of occurrence. They are defined conditional on that source producing a site $S_a \geq x$; this may be a much rarer event for a distant source, for example.

However defined, either globally or per source, $\bar{M}$ and $\bar{R}$ (and, incidentally, the medians too) are central statistics of the marginal distributions of M and R that do not capture any dependence between the two variables. In this respect

too, the use of the modal values of the joint *M-R* distribution conditional on exceeding the target S_a value appears, again, more appropriate.

Despite these several drawbacks, the use of the mean values of *M* and *R* has been adopted by U.S. NRC (1997) and by U.S. DOE (1996) in their recent guidelines for selecting controlling earthquake sizes and locations.

In light of these comments, the need for representing the seismic threat with a single set of ground-motion parameters appears better fulfilled by the bivariate mode (M^*; R^*) rather than by the marginal mean values, $\bar{M}$ and $\bar{R}$. The bivariate mode, however, coincides with the central value of one of the *M-R* bins (see Introduction). Hence, when modal values are reported, the binning sizes should be reported as well to provide a measure of their accuracy. Smaller, more accurate binning sizes may be used for these mean and mode computations than are displayed graphically, where coarse bar charts often replace figures such as Figures 1 and 3.

The use of modal values for the purpose of selecting hazard-dominating events was proposed in the literature by Stepp *et al.* (1993), McGuire (1995), and Chapman (1995). The procedure outlined by McGuire is reportedly at the basis of the predominant seismic magnitude and distance maps produced for southern California by Cramer and Petersen (1996) of CDMG.

Finally, it is worthwhile repeating that means and modes change, in general, with different levels of $S_a(f, \xi)$ or when the hazard is disaggregated for spectral acceleration at different oscillator frequencies and damping ratios.

Joint Distributions of *M*, *R*, and ε

As observed earlier, if for the sake of simplicity or whatever other reason the hazard need be described in terms of one event only, the most likely combination of ground-motion parameters to exceed the target S_a level at the site appears to be the most logical summary statistic. It should be kept in mind, however, that because of the nature of the PSHA approach, no single event will ever be able to fully describe the seismic threat at the site.

In many practical cases, the joint *M-R* distribution shows comparable contributions from more than one region of the *M* and *R* space. In such cases, considering only the mode may lead to underestimating the S_a level in some frequency range (see example later) and accounting for multiple dominant or controlling events, one for each peak of the joint distribution, seems to be more appropriate. This can be a major limitation of predominant *M* and *R* maps, such as those produced by Cramer and Petersen (1996), where the presence of multiple hazard-dominating events for the same site cannot be displayed.

Ground motions generated by earthquakes of *M* and *R* values very different from each other may show quite dissimilar characteristics (e.g., frequency content). Hence, when the hazard is dominated by multiple events, it is important that the different *M* and *R* values be reported, along with an estimation of their relative contributions and that these be properly accounted for during structural analysis and design.

One such case involves the Los Angeles City Hall site (Fig. 2), where the earthquake threat is posed both by several close-by buried and day-lighting thrust faults, which lay under the Los Angeles basin, and by the southern segments of the San Andreas Fault at a larger distance. The seismotectonic model of the region displayed in the figure was provided to us by Dr. Normal Abrahamson (1996). This model, which includes multiple assumptions of the seismicity parameters, was used to provide all the (mean) hazard results for the Los Angeles site. We utilized the modern attenuation law by Abrahamson and Silva (1997) with stiff-soil parameters for ground-motion prediction.

Figure 3 displays the site-specific *M-R* PMF conditional on exceeding an S_a of 0.41 *g* (approximately, the 100-yr mean return period value) at the frequency of 1 Hz and 5% damping. In this case, a constant binning of 0.1 of magnitude was employed. For distances below 10 km, the bins were chosen 2 to 3 km wide, from 10 to 70 km, the bin size was 5 km, and finally, a bin width of 10 km and increasingly larger was adopted for distances above 70 km. The contributions from distances larger than 120 km, however, are not displayed because they are negligible for this hazard level.

For the case in Figure 3, it is apparent that nearby-distance events of magnitude ranging from 5.5 to 7.0 dominate the hazard (more than 80% contribution). A more thorough examination of the raw PSHA results reveals that the main contribution to the hazard is caused by a group of 13 faults whose minimum distance values from the rupture area to the site range from 7.5 to 31 km, whose maximum magnitudes vary from 6.2 to 7.2, and whose types of rupture mechanisms include both reverse (the majority) and strike-slip faulting. The individual contributions of each close-by causative fault, however, are not immediately discernible from this plot. They will be much more evident later when the hazard is disaggregated versus latitude and longitude rather than *R*. The lower peaks in the distribution at distances of 55 km and larger are due to large-magnitude earthquakes that may be generated by different segments of the San Andreas Fault. The nearby faults that dominate the hazard and the San Andreas Fault are shown by heavy lines in Figure 2.

In a complicated case like the one under consideration, the use of single statistics, such as means or modes of *M* and *R* or $M - R$, is clearly not sufficient to describe the characteristics of the ground motions that are most likely to threaten the site. In this respect, the knowledge of the entire joint conditional distribution is necessary. In particular, the mean values, $\bar{M} = 6.35$ and $\bar{R} = 26.8$ km (Fig. 1), would suggest an event that is definitely not very likely to occur. Out of the 13 nearby faults, only four have a minimum distance above 20 km, and they are responsible for only 11% of the total hazard. The mode of the joint conditional PMF of *M* and *R* (see Fig. 3) is at (M^*; R^*) = (6.85; 17.5 km), and it is due (as will be more evident later) mainly to the Sierra Madre, the Newport–Inglewood, the Compton Thrust,

and the Elysian Park Faults. These faults contribute 40% of the total hazard. Sierra Madre is the single most hazardous seismogenic source for the site for this spectral acceleration level, but it contributes only 16% of the total hazard. More frequent, smaller magnitude events (say, M about 6 at R of 10 to 15 km) generated by the less active thrust faults, such as the Hollywood, Raymond, and Northridge Faults, should also be considered among the possible "controlling" events. Together they contribute another 28% of the total.

The importance of identifying multimodal contributions to the hazard can be appreciated by studying Figure 4. This figure displays the 5%-damped median horizontal spectra predicted for a stiff-soil site by the same attenuation law by Abrahamson and Silva (1997) for the three events identified previously. The parameters corresponding to a reverse fault and a site located on the footwall were used when computing the spectra. Notice also that, because here the spectral shapes rather than absolute amplitudes are of primary interest, the spectra have been scaled to have the same $S_a = 0.41\ g$ at 1 Hz. The dependence on magnitude of the frequency content of each ground motion is apparent in the figure. These differences may be important for computing the linear or nonlinear dynamic response of a 1-Hz multi-degree-of-freedom structure located at the site.

The disaggregation of the hazard in terms of the joint $M - R$ distribution has been recommended by many, among others Chapman (1995) and Kramer (1996). The studies of Stepp *et al.* (1993) and of McGuire (1995) include ε as well. It is worth noting that, rigorously speaking, the most likely event over the range of all feasible ones considered in the analysis can be found only by considering the triplet (M^*, R^*, ε^*) of values corresponding to the mode of the 3D joint M-R-ε distribution (conditional on exceeding the target S_a). This is because all three variables significantly affect the exceedance probability, and, in this case, the univariate M^* and R^* and the bivariate (M^*; R^*) are not necessarily preserved in the M and R modal values of the joint $M - R - \varepsilon$ distribution.

In the Los Angeles case considered so far, by coincidence, the M and R values of the 3D mode [i.e., (M^*; R^*; ε^*) = (6.85; 17.5 km; 0.81), where the ε value corresponds to the 79% fractile] of the joint M-R-ε PMF, and those of the 2D mode [i.e., (M^*; R^*) = (6.85; 17.5 km)] of the joint M-R PMF, both conditional on exceeding S_a (1 Hz, 5%) = 0.41 g, are the same. However, this is often not the case, especially if smaller bin sizes are used.

When the 3D modal values are substituted in the ground-motion attenuation relationship, the spectral acceleration value predicted for 1 Hz is equal to 0.46 g, approximately 12% larger than the target value. The discrepancy between the target spectral acceleration, whose exceedance defines the hazard disaggregation, and the spectral acceleration value predicted for the most likely event anticipated to exceed the target level is the topic of the following subsection.

Disaggregation and Target Spectral Acceleration

A very desirable property of any controlling event would be that, when the event parameters are substituted in the attenuation law used in the PSHA, the target spectral acceleration level would be recovered. As alluded to earlier, this property does not hold, and there is no theoretical reason why it should. However, in our experience, when the modal triplet (M^*; R^*; ε^*) of values is substituted into the attenuation relation, the difference (which, theoretically, is always in exceedance) is usually not very large (say, within 20%).

The issue of the difference between the target and the recovered S_a values has driven the already cited disaggregation procedure proposed by McGuire (Stepp *et al.*, 1993; McGuire, 1995). The main goal of this method is, in fact, the identification of a controlling event that matches the target uniform hazard spectrum (and not only the target spectral acceleration at the given frequency) to be used as a hazard-consistent scenario event for structural analyses.

It is worth remarking again, however, that any disaggregation method operates on the site-specific hazard of ex-

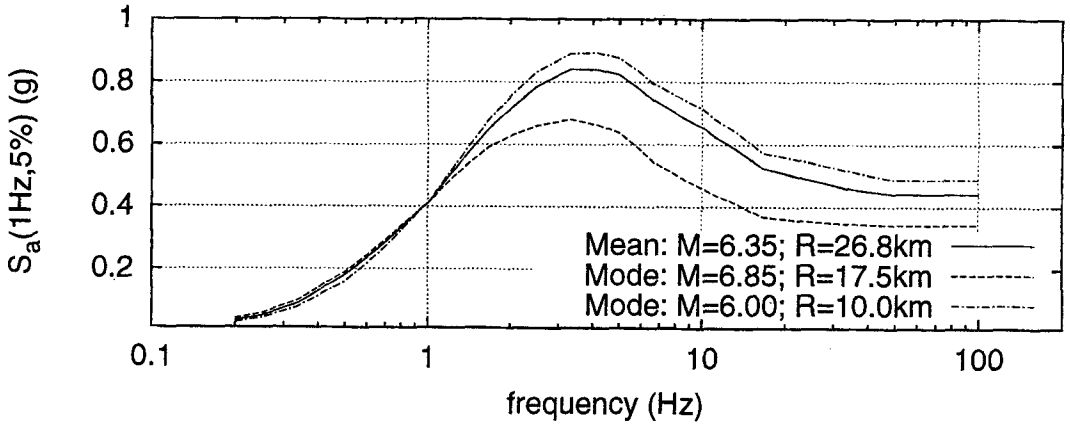


Figure 4. Median motion for three M-R pairs predicted using the Abrahamson and Silva (1997) attenuation law.

ceeding a fixed spectral acceleration level, x, at a specified oscillator frequency, f, and damping ξ. When matching of the entire response spectrum (or even a portion of it) by a single controlling event is desired, then any disaggregation procedure necessarily becomes less rigorous. As stated before, the hazards at different frequencies (even for the same mean return period) are often dominated by distinct events due to the different attenuation of the seismic waves at separate frequencies. Hence, for consistency, in the following, we concentrate our attention on the matching of a single spectral acceleration value at a specified frequency.

In order to achieve the matching, McGuire suggests disaggregating the probability of exceedance of the specified S_a level at the given frequency f by lumping the hazard contribution into the appropriate M, R, and ε bin such that the target value is equaled (not exceeded) when the values are substituted in the attenuation relation.

To make this matter clear, let us consider the following example. Assume that during the PSHA computations an event of large magnitude m at close distance r is being considered. For such an event, assume also that, according to the adopted ground-motion attenuation relation, the target $S_a = x$ level is equaled when the ground motion is one standard deviation below its median level (i.e., $\varepsilon = -1$). In other words, this event is so strong and close to the site that a (log) ground motion of average strength generated by this event will exceed the target S_a level by one standard deviation. With these numbers, on average, only 16 out of 100 ground motions generated by events of such m and r will produce an S_a value equal to the target value, x, or lower.

In this case, the procedure proposed by McGuire would assign the entire probability of exceedance of the acceleration x (i.e., 0.84 times the probability that an event of such intensity, m, would occur at that particular location, r) to the (cubic) bin that contains the values of m, r, and $\varepsilon = -1$. The bins containing m, r, and all the values of $\varepsilon > -1$ would not be assigned any contribution (see Fig. 5 for the 1D ε bins).

From the previous example, it becomes clear that this procedure, although perfectly reasonable given the matching purpose, disaggregates the hazard in such a way that the modal value of the conditional M-R-ε distribution is not guaranteed to describe the event that will most likely exceed the target $S_a = x$ level at the site. In this made-up example, in fact, the earthquake characterized by the m, r, and $\varepsilon = 0$ triplet would be a much more likely event to exceed x than the one selected by the foregoing procedure.

In the example just considered, if each single bin containing m, r, and the ε values above -1 had received its part of the hazard (proportional to probability mass in the ranges of the ε bins above -1), then the disaggregation procedure would have created the conventional joint conditional M-R-ε distribution, and its 3D mode would be the single, most likely event to exceed $S_a = x$ at the site. This alternative procedure was used in the Los Angeles City Hall case study reported in the previous subsection.

This latter hazard disaggregation method, however, does not guarantee any longer that the joint distribution modal value matches the target $S_a = x$ level. In this respect, recall that in the Los Angeles site example, the resulting modal event produced an acceleration S_a (1 Hz, 5%) = 0.46 g, which is larger than the target of 0.41 g. On the other hand, McGuire's procedure, which matches the target S_a value at a frequency of 1 Hz, would have identified a modal event at (M^*; R^*; ε^*) = (7.55; 57.5 km; 1.15) generated by the San Andreas Fault. This event does not coincide with the most likely event expected to exceed the target S_a at the site.

Therefore, it is clear that there is a trade-off between the desire of having the target spectral acceleration matched and the necessity of producing the proper joint conditional distribution, whose mode can be confidently said to be the event that will most likely generate at the site a ground motion exceeding the target S_a level at the specified frequency.

It should be recognized that these two methods, which disaggregate the hazard in M, R, and ε terms, are both perfectly legitimate; they simply respond to different specifications. Only the second one can be said to use the rigorously defined joint conditional PMF. It should be noted too that the first one treats the three variables unsymmetrically; only the probability mass of ε is redistributed to achieve the matching.

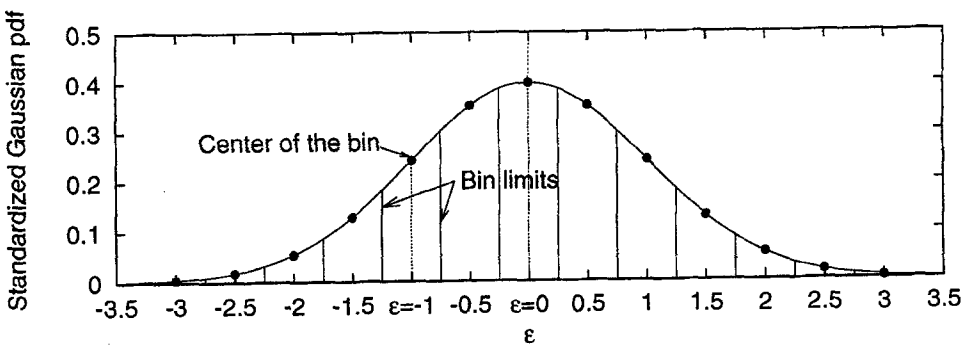


Figure 5. Univariate ε bins for the example considered in the text.

When the need to match the target S_a value at a given oscillator frequency, f, and damping, ξ, is felt to be of primary importance, a third method can be devised. This alternative method, which blends the characteristics of both disaggregation procedures, would take into consideration the M^* and R^* values of the most likely event [taken from the 3D (M^*-R^*-ε^*)], and then heuristically adjust the ε^* value to ε', the value necessary to recover the target $S_a(f, \xi)$ level when substituted in the attenuation law. This scaling procedure can be applied with accelerograms representative of the most likely event (i.e., of the previous modal values M^* and R^*) and with the median spectral shape as predicted for that event by an attenuation relationship. In the case of the Los Angeles City Hall site, this procedure would lead to a controlling event characterized by $M^* = 6.85$, $R^* = 17.5$ km, and $\varepsilon' = 0.63$ (smaller than the 3D modal value of $\varepsilon^* = 0.81$).

The justification for this proposal is that the spectral shape depends primarily on M and, secondarily, on R but does not significantly depend on ε. The mild dependence of the spectral shape on ε follows from the small difference between ε' and ε^* in most practical cases. This mixed approach would in fact ensure that the adopted controlling event has the most likely magnitude and distance. This somewhat heuristic method is similar to the one proposed by Chapman (1995), the only difference being that he proposes the use of (M^*; R^*) values taken from the M-R joint distribution instead of that from the full 3D M-R-ε joint distribution, as suggested here.

Different Distributions, Random Variables, and Binning Schemes

In the previous section, we discussed the several hazard disaggregation procedures proposed in the literature and introduced a technique for disaggregating the hazard in terms of M, R, and ε. These all will clearly produce somewhat different results. To date, little attention has been devoted, however, to uncovering the effects that apparently marginal details, such as the distribution that gets disaggregated in each bin, the precise definition of the random variables used for disaggregation, and the selection of bin sizes, have on the final perception of the hazard contributions and on the summary statistics employed to describe the dominating events.

Use of PMF and PDF in Hazard Disaggregation

As discussed earlier, disaggregating the hazard implies the computation of the relative contribution of each M, R, and ε bin to the probability of exceeding the target S_a level. The procedure of accumulating the hazard per bin is repeated for all possible earthquake locations and intensities compatible with the seismotectonic models of the region around the site.

Assume now that each bin is selected of constant width throughout the domain of each variable and, in particular, of width Δm in magnitude, Δr in distance, and $\Delta\varepsilon$ in ε. The relative hazard contribution is reported in terms of the probability that, given the exceedance of S_a estimated in accordance with a ground-motion attenuation relation, the event is of magnitude between $m \pm (\Delta m)/2$ at a distance included in $r \pm (\Delta r)/2$ and the (log) ground motion (given the magnitude and the distance) is between $\varepsilon \pm (\Delta\varepsilon)/2$ standard deviations away from the predicted (median) motion.

Although U.S. NRC (1997) recommends the magnitude and distance bins to be used for the facilities under its authority, the choice of the bin sizes for each variable is usually left to the discretion of the analyst (e.g., SSHAC, 1997). While it is obvious that the width of each bin should not be selected less than the step size used for each component in the PSHA numerical integration, there is a degree of arbitrariness regarding customary values to be used. Values of Δm reported in the literature range from 0.1 to 1.0, of Δr from 5 to 100 km (generally increasing with distance from the site), and of $\Delta\varepsilon$ from 0.1 to 1. For display purposes, the bin contribution to the hazard is usually assigned to the center point of the cell (at least when linear distance binning is used—see the next subsection).

As mentioned earlier, the integration in the PSHA is carried out in discrete steps, and the final M-R-ε joint distribution (conditional on the exceedance of S_a) is most naturally reported in terms of a probability mass function (PMF). A PMF was in fact displayed in Figure 3. This representation, however, is sensitive to the selection of the sizes of the cells; in this respect, it may be preferable to display the probability density function (PDF) instead.

The PDF representation, which is obtained by dividing the PMF contribution of each bin by the bin's size (i.e., in this 3D case, by the product $\Delta m\ \Delta r\ \Delta\varepsilon$), is independent of the bin selection, at least in the limit when bin sizes approach zero. Hence, whenever the binning selection is left to the analyst, showing the hazard contributions in terms of PDF would avoid a degree of arbitrariness that is present in the results when PMF is adopted. On the other hand, a PDF representation is sensitive to the precise definition of the variables that are disaggregated. We shall return to this matter in the next subsection when discussing the use of R versus $\ln R$ as a measure of distance.

If the bin sizes are selected to have the same width throughout the domain of all the variables, the PMF and PDF appear as scaled versions of one another when displayed as surfaces (as in Fig. 3 for M and R). However, as said before, this is seldom the case. While the bin size in M is usually kept constant (not always, see later in this section), the bins at small distances are typically selected to be much shorter than the bins at larger distances. For most practical purposes, events occurring at R values between 100 and 150 km are often times perceived as equally distant from the site, but not all the earthquakes between 0 and 50 km can be considered equally close. For example, U.S. NRC (1997) recommends the use of magnitude bins of 0.5 unit of width for all

magnitude values, while the distance bin sizes vary from 10 to over 100 km as the distance from the site increases.

When bins of uneven sizes are adopted, the conditional PMF and PDF graphical representations are not proportional any longer, and the display of the PMF may be misleading (see also the next subsection). Regions that appear to produce a high hazard contribution may, for instance, be caused by simply a larger bin size, while, at the same time, the probability density in the same region would be considerably less prominent. This phenomenon, obviously, may have also an effect on the location of the mode that is often used as the definition of the controlling event.

An example is shown in Figure 6 which displays the joint M-R PDF conditional on the exceedance of S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site. The binning sizes are the same employed to generate the corresponding PMF reported in Figure 3. From comparison of Figures 3 and 6, it is clear that the contributions to the hazard due to the San Andreas Fault segments above 70 km are made more apparent in the PMF by the increasingly larger bin size used for distance. The effect of uneven distance binning can also be noticed in the modal values of the PDF [i.e., $(M^*; R^*)$ = (6.15; 9 km)] that are different than those of the PMF [i.e., $(M^*; R^*)$ = (6.85; 17.5 km), see previous section]. Relative to the PMF, in the PDF representation, the hazard contributions at short distances are enhanced because of the smaller distance bin width below 10 km.

A clear advantage of the PMF representation, however, is that the ordinate corresponding to each cell can be immediately interpreted as the contribution to the hazard due to magnitude and distance (and, sometimes, ε) ranges of the bin itself. This property does not hold when PDF is adopted to display the hazard contributions, because in this case, it is the volume under the distribution that is unity and not the summation of the surface ordinates corresponding to the middle point of each bin. If a precise estimate of such contributions is of interest, one should either make use of a numerical algorithm to integrate the PDF or compute a PMF for a binning scheme in which the bin widths are constrained to be uniform.

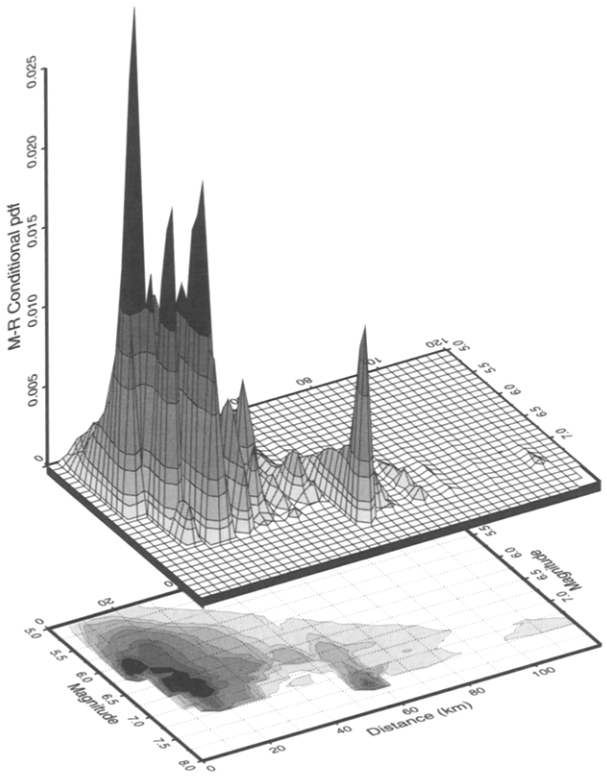


Figure 6. M-R PDF conditional on the exceedance of S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site. Compare Figure 3.

Uses of R Versus ln R

Some researchers prefer to evaluate the contributions to the hazard in terms not of the distance R but rather of the (natural) logarithm of R; that is, D = ln R. The regulations issued by U.S. NRC (1997), for example, require that the hazard be disaggregated in terms of ln R and that the contribution of each cell be assigned to the centroid of the ring area comprised by the lower and upper limit of each distance bin.

The reasoning is that the ln R form is more nearly symmetrical with M and $\eta = \varepsilon\sigma_{\ln S_a}$. In the typical regression for ln S_a (see equation 1), the dependence on M and ln R is approximately linear, as it is with η. (Secondary terms such as R, M^2, and RM may also appear, but they are comparatively weak terms in the regression.) For this reason, and also for saving computer time by giving somewhat less emphasis to large distances that contribute less to the hazard, several of the readily available PSHA codes use an integration step size that is constant in ln R rather than R.

The use of ln R in lieu of R, however, has some interesting effects on the hazard disaggregation mainly when the PDF representation of the hazard is used. We shall explore this issue with a simple example.

Let us consider a site located at the center of a circular area of 100-km radius where the seismicity is uniform. To concentrate our attention exclusively on distance, assume that only events of zero hypocenter depth and of magnitude 7 can occur in that area. We want to compute both PMF and PDF of R and D = ln R, conditional on the exceedance of S_a (1 Hz, 5%) = 0.3 g at the site when two linear and logarithmic distance binning schemes are used.

More precisely, the hazard, computed by the PSHA code using a constant step size of 1 km, is disaggregated in terms of R and D, and the contributions are accumulated both in 10 bins equally spaced in linear scale of R from 0 to 100 km (Fig. 7a) and in 10 bins equally spaced in (natural) logarithmic scale of R again from 0 to 100 km (Fig. 7b). In the linear scheme, each bin is, obviously, 10-km wide, while in the logarithmic scale, the width of each distance bin varies from approximately 1 km nearby the site to 37 km for the farthest

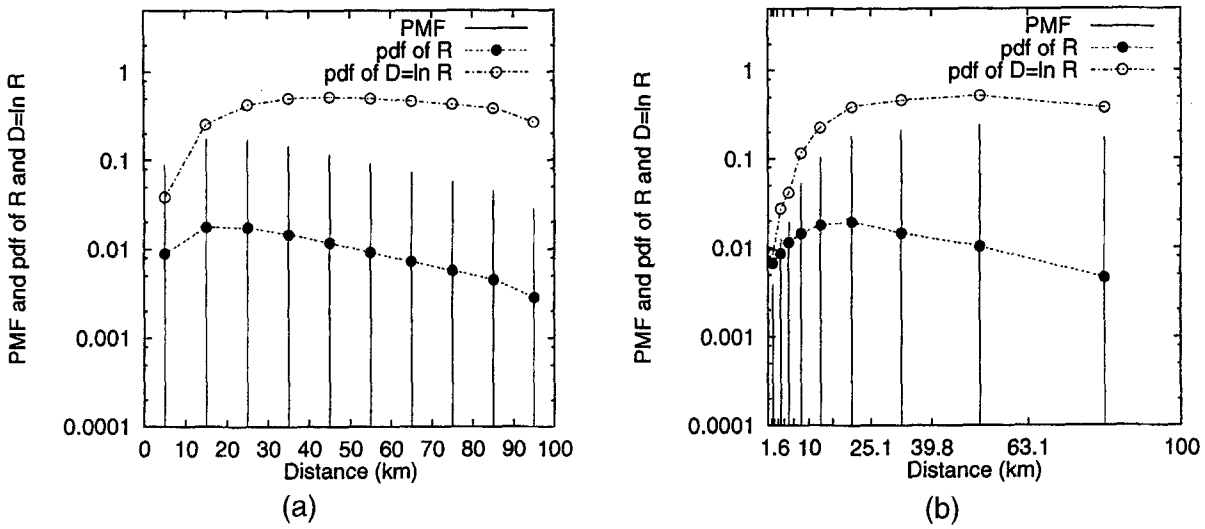


Figure 7. PMF and PDF of R and $D = \ln R$ for a circular area source example. The tic marks on the distance axis delimit the boundaries of the bins. (a) Regular binning in linear scale. (b) Regular binning in logarithmic scale.

bin spanning from 63.1 to 100 km. Since the differences are not critical when bin sizes are not wide, for simplicity, the contributions are assigned to the central value rather than to the centroid value of each distance cell.

The marginal PMF and PDF of R and D (conditional on exceeding the specified S_a value) are shown in Figures 7a and 7b. It is important to point out that only one PMF is shown in each figure because the two PMFs obtained when disaggregating in R and D coincide for the same binning strategy. (Notice that the PMF in this case is displayed as a series of vertical spikes.) As discussed in the previous subsection, the PDFs of R and D are insensitive to the binning scheme, unlike the PMFs, which show a very different shape. The PDFs of R in Figures 7a and 7b would exactly coincide if the binning sizes were chosen very small (approaching the step size used in PSHA computations and zero in the limit). A similar remark obviously holds for the PDFs of D.

Assume that the contributions to hazard are available for the two binning schemes in terms of PDF and PMF of R only. In this case, the analyst would conclude from the PDF of R that the events most likely to exceed the specified S_a value occur at 10 to 20 km from the site, regardless of the binning scheme adopted. On the other hand, a superficial review of the PMFs of R may lead to interpretations of the distance contributions dominating the hazard that are driven by the particular binning, that is, short distances for the linear binning scheme (Fig. 7a) and mid-range to high-range distances for the logarithmic strategy (Fig. 7b). The risk of possible misinterpretation when using the PMF is emphasized when, instead of spikes, a more customary line (or surface in more dimensions) representation is adopted for display. Of course, similar reasoning would follow if $D = \ln R$ were chosen.

Assume now that the analyst disaggregates the hazard in terms of PDF of both R and $D = \ln R$. In this case from either Figure 7a or 7b, one could deduce, when using R, that the contributions to hazard peak at a close distance ($R^* = 15$ km) and then fade away. On the other hand, distances from 30 to 80 km appear to dominate the hazard if D is preferred. The modal value in the latter case is shifted to a larger distance of 45 km. (Recall that both values have an accuracy of ± 5 km in Fig. 7a.)

In particular, it is interesting how the hazard contributions by D remain almost constant with distance, instead of exponentially decreasing as customarily observed. From a mathematical perspective, this behavior is easily explainable. In Figure 7a, for example, the bins are all $\Delta r = 10$ km wide throughout the entire domain when distance is measured in terms of R, but the bin width keeps decreasing when moving from 0 to 100 km (e.g., $\Delta d = 0.69$ for the 10- to 20-km bin and 0.11 for the 80- to 90-km bin) when D is used. Therefore, in the PDF of D, the relative contributions due to close distances are deamplified with respect to the contributions obtained from the PDF of R, and, conversely, those due to large distances are magnified.

The same pattern is observed also in the Los Angeles City Hall case study when the hazard of exceeding S_a (1 Hz, 5%) = 0.41 g is disaggregated in terms of D and represented by a PDF. From Figure 8 (to be compared with Figs. 3 and 6), it is evident how the contributions due to San Andreas Fault segments at 50 km and more from the site are greatly emphasized. In this case, the mode (M^*; D^*) = (7.55; 4.04) translates to an event, which might be interpreted as the dominating earthquake, which is located on the San Andreas Fault at 57 km from the site (much farther away than before when R was used).

We conclude that when the PMF is used to characterize the contributions to hazard, the binning scheme adopted is

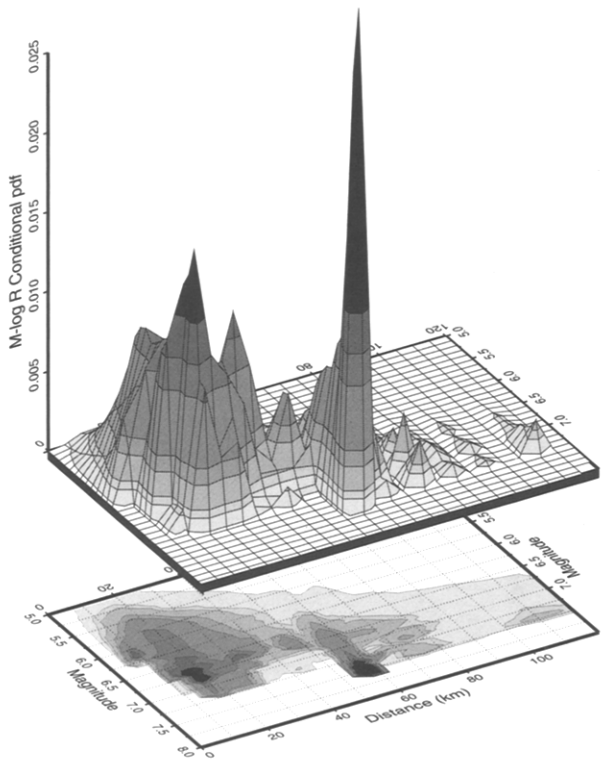


Figure 8. M-ln R PDF conditional on the exceedance of S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site. Compare Figures 3 and 6.

crucial, while the use of R or D is not. On the other hand, if the PDF is used, the use of R or D when disaggregating the hazard is critical, while the binning strategy becomes noninfluential (in the limit).

The authors' preference is to use the PDF of $M - R - \varepsilon$ or, alternatively, the PMF of $M - R - \varepsilon$ obtained by uniform R bins. We favor the use of R because it represents the more natural interpretation of distance. However, if a particular application suggests a binning strategy (see the next subsection for two examples), the PMF representation, unlike its PDF counterpart, is invariant to the selection of the distance measure (the PMFs in terms of R and D coincide). Hence, in such cases, it may be preferable to compute the contributions to hazard using PMF rather than PDF expressed in terms of either R or D.

Selection of the Binning Scheme

In the previous sections, we have implicitly assumed that the hazard is disaggregated for a generic application where the spectral acceleration S_a is the final response measure of interest to the analyst. Considerations about appropriate M and R bin dimensions for this generic case were given in a previous subsection. It should be emphasized again that, regardless of the variables used during hazard disaggregation (e.g., R or D = ln R), the selection of the binning scheme and even the dimension of the hazard disaggregation depend on the intended use. This concept is made clear by the following two examples.

A geotechnical engineer may want to assess the liquefaction resistance for a site by means of the simplified cyclic stress approach (e.g., Seed and Lee, 1966; Kramer, 1996). In this approach, the relationship between soil density, cyclic stress amplitude, and number of cycles to failure is defined in terms of the so-called cyclic stress ratio (CSR). The CSR is customarily compared to the CSR for earthquakes of magnitude 7.5, $CSR_{7.5}$ (Seed *et al.,* 1983), via the magnitude correction factor (MCF = $CSR/CSR_{7.5}$). The minimum CSR required to initiate liquefaction decreases with strong-motion duration and, hence, with magnitude. The relationship between MCF and M is displayed in Figure 9a (Kramer, 1996). In this case, the engineer may want to disaggregate the hazard for several values of the input ground motion (e.g., PGA) and investigate the contributions to hazard from a nonuniform binning scheme that preserves a constant MCF decay in each bin (see Fig. 9a).

In a different application, a structural engineer may be asked to assess the seismic performance of a tall, flexible high-rise building located close to a fault. He may be concerned that the strike-normal component of a ground motion at the site exceeds the design criterion. The difference between the strike-normal and the average horizontal S_a is in fact significant for short oscillator frequencies. For example, the average ratio between strike-normal and average horizontal S_a at the building fundamental frequency, f_1 = 0.25 Hz, for a distance of 1 km from the fault and a magnitude 6.0 event is just above 1.3 (Somerville *et al.,* 1997) (see also Fig. 9b). This ratio, called FN/AVG in the figure, is only mildly dependent on M but is strongly dependent on distance. The ratio decreases sharply with increasing distance from the fault. The engineer may then decide to disaggregate the hazard for the exceedance of S_a (0.25 Hz) equal to the original design level in 2D M-D bins. While the bins may be uniform in M, the widths of the D bins could be selected in such a way that the FN/AVG decreases constantly in adjacent bins. This criterion gives rise to the irregular distance binning shown in Figure 9b.

Disaggregation of Hazard in Latitude, Longitude, M, and ε

As seen in a previous section, in the best form to date, the seismic hazard is disaggregated in terms of the three main variables that appear in the PSHA calculations, that is, M, R, and ε. For complicated cases, such as the Los Angeles City Hall site considered in Figure 3, the M-R-ε contributions to hazard need a further interpretation to allow the identification of the causative faults. One might consider using the typically provided table of fault-by-fault contributions to the total hazard together with a complete set of the single-fault disaggregations (see Introduction). Even such an approach does not determine or display on a map the locations domi-

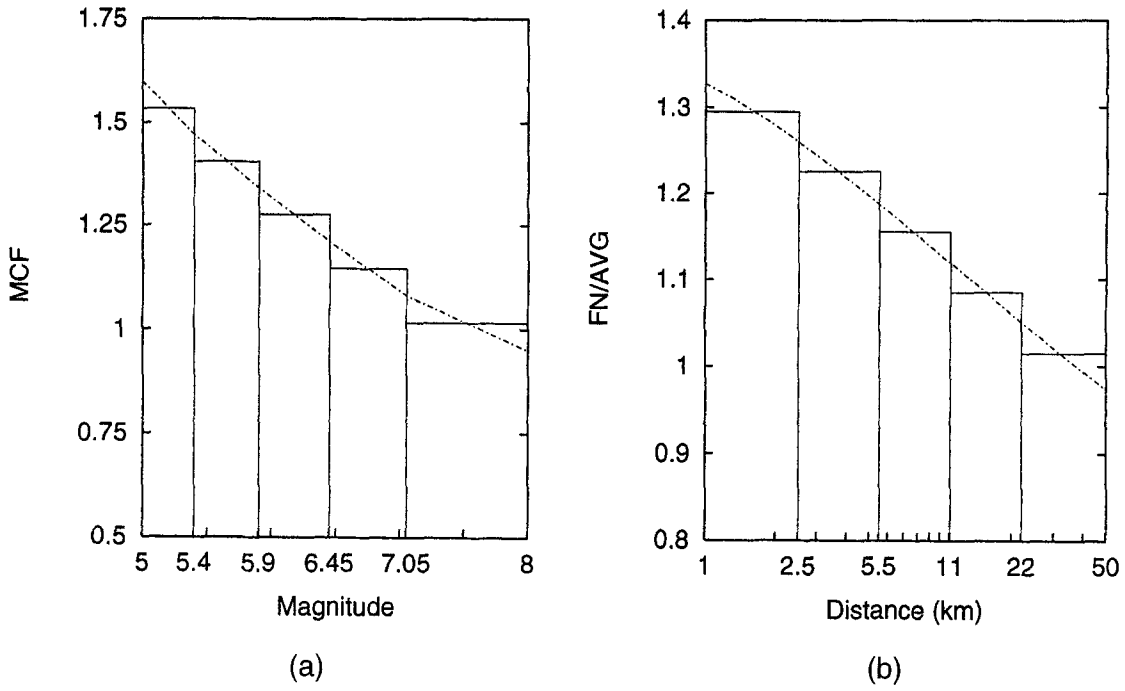


Figure 9. Binning schemes proposed for two different applications: (a) assessment of soil liquefaction resistance and (b) seismic performance evaluation of a building located close to a fault.

nating the hazard, a feature that would enhance the understanding and the ability to communicate the hazard.

The natural following step that we propose toward a further improvement of the hazard representation consists of displaying the contributions versus not three but four dimensions: latitude, longitude, M, and ε. It has been brought to our attention that a somewhat similar technique for disaggregating the hazard, versus latitude and longitude only, was independently proposed in the so-called gray literature (REI, 1989), but, to our knowledge, it did not have a wide circulation or subsequent application. This 4D spatial representation of the hazard, which is implemented here, has also been recently envisaged by Spudich (1997). Based on an early manuscript of this article, USGS has recently implemented a modified version of the 4D disaggregation that follows (see the World Wide Web site http://geohazards.cr.usgs.gov/eq/).

Disaggregation in 4D

In the present work, latitude and longitude values represent the coordinates of the surface projection of the closest point of the random rupture area. In general, however, latitude and longitude may be referred to the surface projection of any measure of the distance, R, from the source to the site, as defined in the ground-motion amplitude predictive relationships adopted in the PSHA. Note that if in the same PSHA one uses multiple attenuation laws requiring different source-to-site distance definitions for R (e.g., the closest distance to the rupture surface versus the hypocentral distance), the contributions to hazard disaggregated in latitude and longitude can still be combined. In contrast, in the 3D case, different definitions of R within the same disaggregation exercise may generate results that are difficult to interpret.

This disaggregation scheme permits the display of the hazard on a typical map of the faults surrounding the site, allowing an immediate identification of the locations on the faults dominating the hazard. Practically speaking, this formulation, along with the knowledge of the most likely magnitude, may be very helpful in establishing the specific earthquakes that present the greatest hazard to the site. The knowledge of the causative faults and of the most hazardous locations allows other seismic source characteristics, such as, for example, rupture mechanism, propagation path, and near-source effects, to be modeled. These characteristics have a direct impact on the severity and attributes of the motion to be expected at the site (e.g., spectral content, duration, degree of nonstationarity, critical pulses, etc.) that may be relevant in subsequent structural analyses.

An application is shown in Figure 10, where contours of hazard contributions for the same Los Angeles site and the same S_a level considered before are displayed in terms of latitude and longitude. For a correct interpretation of this figure, some comments are mandatory.

Hazard contributions appear to be arising from locations in between the surficial traces of different faults. During

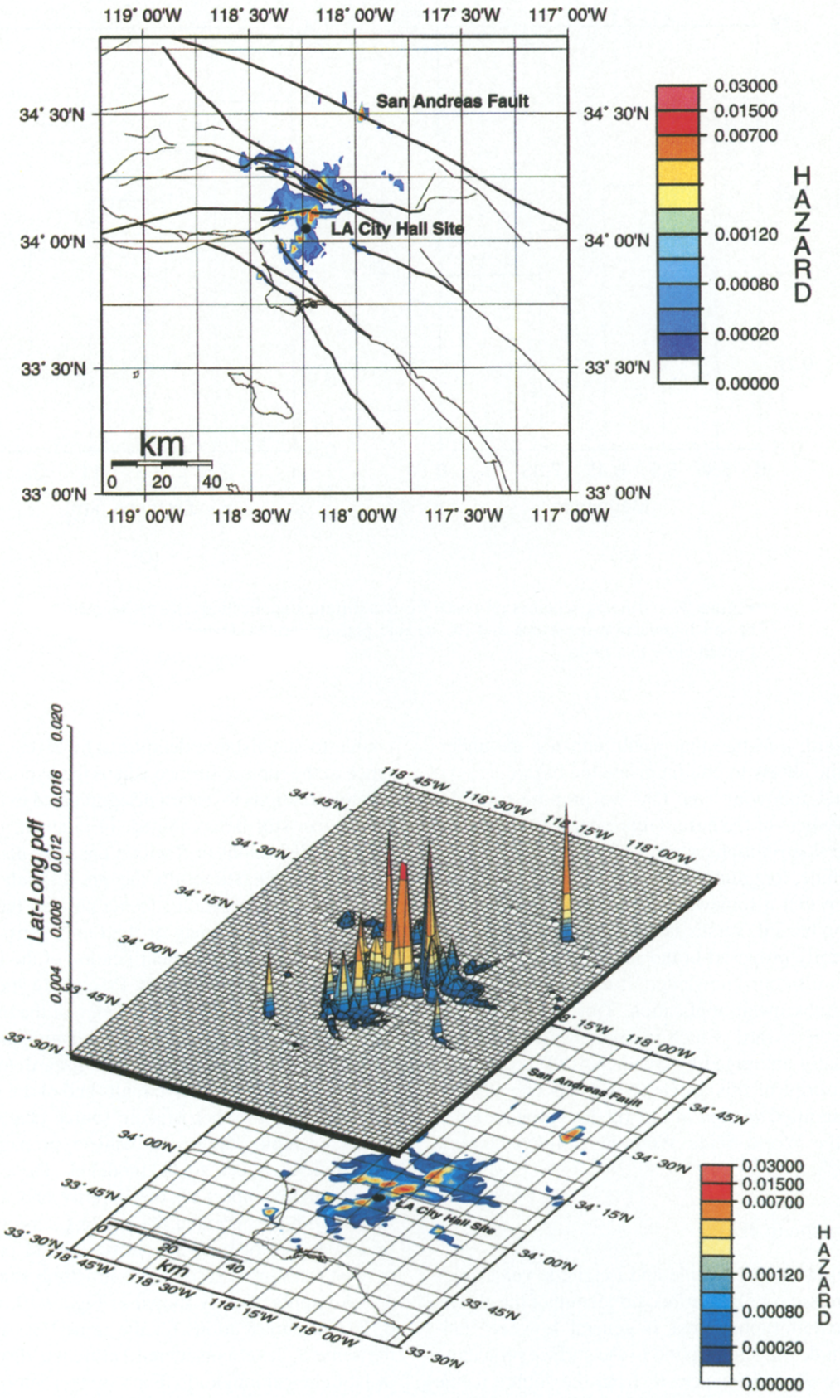


Figure 11. Three-dimensional view of contributions to hazard of exceeding S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site, disaggregated in latitude and longitude.

PSHA computations, the hypocenter is assumed to occur at random at different positions on the fault plane that, in some of the faults, are oriented downward with a small dip angle. Therefore, in some cases, the projection on the surface of the closest point on the rupture area (recall this is the measure of R adopted here) is not aligned with the faulting trace on the ground surface.

For instance, in our Los Angeles example, this is the case for the Northridge Fault that is positioned approximately 12 km NNE of the site and dips toward the site (see Fig. 2). This fault is responsible for the hazard contributions displayed as the light blue area SSW of the fault in Figure 10. Along the Northridge Fault plane, the point on the (random) rupture areas that is closest to the site lies underneath the surficial trace of the western tip of the Raymond Fault (in the middle of the orange-red zone in Fig. 10). Concentrated at this location is the closest of the four highest spikes of the latitude–longitude joint PDF (conditional on the exceedance of the specified spectral acceleration level) displayed in Figure 11. The contributions to hazard shown in Figure 10 are just the contour lines of this surface. For this particular S_a level at this frequency, the other two high concentrations of hazard close to the site are due to the Hollywood Fault (i.e., the spike farther west in Fig. 11) and to the Sierra Madre Fault (i.e., the tall spike located between the site and the San Andreas Fault in Fig. 11).

Notice also that the contribution due to some of the faults is not spread along the source but is mainly concentrated in a single spike located at the closest distance from the site. This occurs, for example, for the San Andreas Fault that is responsible for the isolated high spike on the right-hand side of Figure 11. This apparent concentration of hazard, again, is due to the distance measure adopted in the attenuation model used in the PSHA and, obviously, does not imply that that particular location is more active than any other along the fault. In the PSHA, earthquakes are in fact assumed to occur with equal likelihood at any location within the same source. Large events, however, tend to break at longer portions of the fault, making the location closest to the site more likely than any other to be the closest point on the rupture area for many different earthquakes. The contribution to hazard would have been more uniformly distributed along the fault plane if, for example, hypocentral distance had been used as distance parameter. The terminology *most hazardous location,* which is somewhat imprecise, should be considered and understood under this perspective.

With the aid of spatial disaggregation results such as those in Figures 10 and 11, the analyst can easily identify positions in space where he would like the other two dimensions displayed. At any specified location, the hazard can be further disaggregated in terms of M and ε, providing in this way information on the magnitudes of the events that are most likely to exceed the specified S_a level at the site, and on their relative strengths (in terms of ε) compared to median motion predicted for given magnitude and distance values. For example, in the Los Angeles case study, one may want to further disaggregate the hazard at the locations of the four highest spikes.

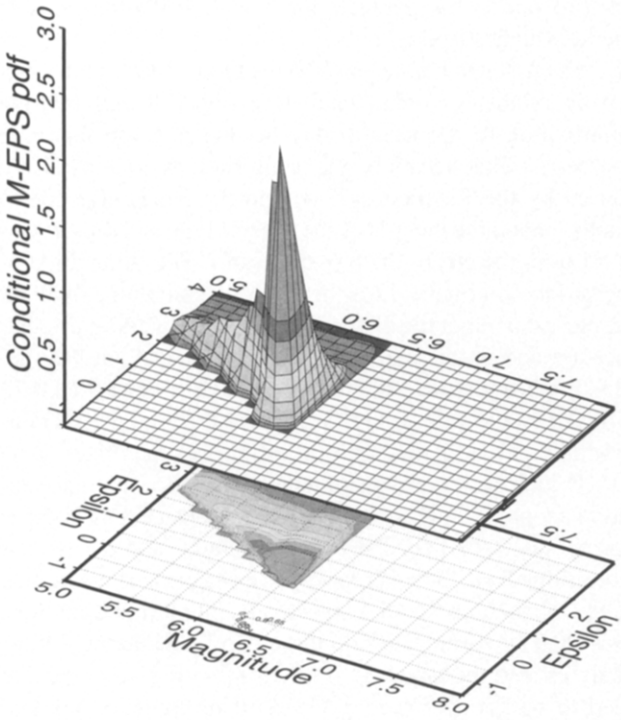
As we stated before, at some locations, such as the western tip of the Raymond Fault (see Figs. 10 and 11), the contribution to the hazard may be due to more than one source. At that location, the contributions to hazard are caused by the Raymond (77%) and the Northridge (23%) Faults (again, the latter is of the reverse type and dips downward from the NE to the SW direction). The summation of the surfaces in Figure 12 is the M-ε PDF conditional both on the closest point of the rupture surface being at this particular location and on the exceedance of the target S_a at the site. This implies that the volume below each surface (i.e., 0.77 in Fig. 12a and 0.23 in Fig. 12b) gives the total hazard contribution (at that particular location) due to each fault.

The characteristic magnitude-frequency recurrence relation adopted for these faults induces peaks near their upper magnitudes of 6.3 for the Raymond Fault, and 6.7 for the Northridge Fault. More precisely, the peaks are at magnitude values of 6.25 and 6.65, respectively. The ground motions generated by earthquakes on these two faults that will most likely exceed the target S_a (1 *Hz,* 5%) = 0.41 g at the site need to be stronger (i.e., $\varepsilon \geq 0.5$) than predicted for that distance and magnitude to exceed the target S_a. Note that in spite of the same surficial distance from the selected location to the site (about 7 km) and the larger maximum magnitude value for the Northridge Fault, the two contributions (i.e., Raymond versus Northridge) in Figures 12a and 12b peak at almost the same ε value of approximately 0.8. The reason is that the site is farther away from the Northridge Fault plane at that location (which, in our model with multiple hypotheses on the dip angle, is on average 12-km deep) than it is from the daylighting Raymond Fault.

The hazard contribution versus M and ε at the location of the spike in Figure 11 that lies on the San Andreas Fault is shown in Figure 13. It can be noticed that, given the larger distance from the site (circa 55 km), events of very large magnitude (from 7 to 7.6) causing ground motions of unusually high strength ($\varepsilon \geq 1$, which means the 84th percentile or higher) have to occur for the target acceleration level to be exceeded at the Los Angeles City Hall site. The M and ε contributions to hazard at the two locations on the Hollywood Fault and on the Sierra Madre Fault are similar to those displayed in Figures 12 and 13, and, therefore, they have been omitted here.

Most Likely Event at Most Likely Location

Seismicity Modeled by Faults. In the previous section, the most likely event to exceed the target S_a has been defined as being the one described by the modal values of the joint M-R-ε distribution. The natural extension would be to consider the mode of the 4D latitude-longitude-M-ε probability distribution (either PDF or PMF). This approach is not being proposed here for practical computational reasons only. In real cases, in fact, this 4D distribution may require large memory storage because, in order to keep accuracy in space,



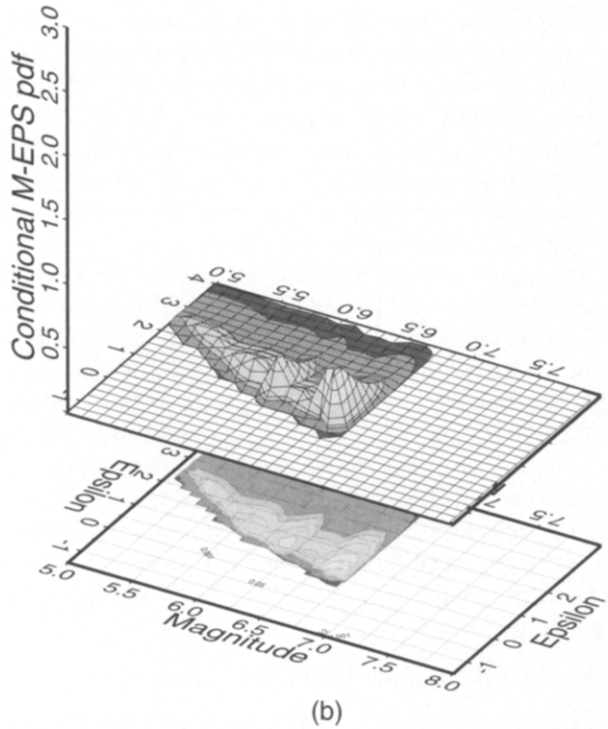


Figure 12. M and ε contributions to the hazard of exceeding S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site for one of the most likely locations. The contributions are presented for both causative faults: (a) the Raymond and (b) the Northridge Faults.

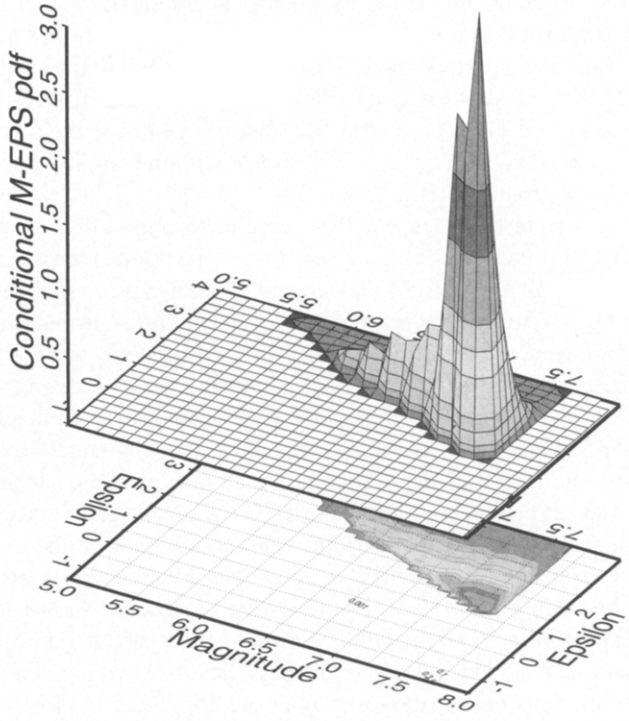


Figure 13. M and ε contributions to the exceedance of S_a (1 Hz, 5%) = 0.41 g at the Los Angeles City Hall site given that the earthquake is generated by the southern segment of the San Andreas Fault.

a limited bin size in latitude and longitude is required (e.g., in Figs. 10 and 11 in the proximity of the site, a bin size of 02′ 24″ was used in both directions). If feasible, however, this 4D distribution should be computed, and its mode (latitude*-longitude*-M*-ε*) should be used to rigorously identify the most likely event to exceed the target S_a at the site.

Alternatively, a less strict but more practical definition is the most probable event occurring at the most likely location, as identified via the hazard disaggregated in latitude and longitude. However, the mode of the bivariate (conditional) $M - \varepsilon$ distribution at the mode of the bivariate latitude–longitude distribution proposed here may or may not be the same as the mode of this 4D distribution.

Similarly, it is not necessarily the case that the distance of the most likely location, r_l (as identified by the largest contributions of all the latitude–longitude bins), is equal to the most likely distance, R* [as selected on the basis of the distance value of the (M*-R*-ε*) mode]. These two values, of course, are correlated. In our limited experience, the distance r_l (computed accounting also for the depth of the fault planes at that location) appears to be fairly close to R*, at least when the seismicity in the region is modeled by faults.

The Los Angeles case discussed so far is quite complicated because no one single location clearly dominates the hazard for the specified S_a level and oscillator frequency. The four highest peaks in Figure 11 have comparable heights and sizes. We have seen that the location where both the

Raymond and the Northridge Faults contribute to the hazard (which, strictly speaking, is the modal value of the contributions to hazard disaggregated in latitude and longitude) is approximately (on the ground surface) 7 km from the site. The (average, because multiple hypotheses on dip angles were used) hypocentral depth at this location is 12 km for Northridge and 2 km for Raymond. Hence, the value of r_l, recovered from this peak of the latitude–longitude disaggregation, appears to be 14 km, if the event is generated by Northridge, or 8 km, if the earthquake occurs on the Raymond Fault. If we consider instead the other two high spikes due to the Hollywood and the Sierra Madre Faults (Fig. 11), the values of r_l are 8 and 17 km, respectively. On the other hand, the spike on the San Andreas Fault is 55 km distant from the site. Other than the last, these distance values are not very different from the value of $R^* = 17.5$ km found by means of the 3D disaggregation scheme.

Again, in the Los Angeles case, the hazard for this 100-yr S_a level at the frequency of 1 Hz is dominated by events mainly from four different sites. The most likely event at the location where the Raymond and the Northridge Faults overlap is a magnitude 6.25 earthquake (at 8 km and with $\varepsilon \approx 0.8$) generated by the Raymond Fault (see Fig. 12a). The most likely events at the other three locations are (6.15*M*, 8 km, 0.8) caused by the Hollywood Fault, (7.15*M*, 17 km, 0.5) caused by the Sierra Madre Fault, and (7.50*M*, 55 km, 1.5) generated by the San Andreas Fault (Fig. 13). These events are different than the magnitude 6.85 event at 17.5 km (and $\varepsilon = 0.81$) identified in the previous section. Interestingly, all four events are predicted by the attenuation law to produce S_a (1 Hz, 5%) values at the site in excess of 0.41 *g* (more precisely, 0.49, 0.44, 0.46, and 0.48 *g*, respectively); the spectral shapes would also be somewhat different.

Theoretically, there is no reason to prefer one option over the other as the candidate for describing the hazard-dominating event. The event selected on the basis of the most likely location carries with it naturally the information about the causative fault and its style of faulting. This information can be exploited to derive an appropriate ground-motion spectral shape. Historically, however, engineers and seismologists are used to associating spectral shapes with distance values (plus magnitude). Hence, if this is the form of the specification, perhaps the most likely event identified by the mode of the 3D $M - R - \varepsilon$ PMF distribution should be preferred.

Seismicity Modeled by Area Sources. Regarding the most likely location issue, another interesting case worth discussing involves spatial disaggregation of hazard when the seismotectonic model of the region around the site comprises only area sources of uniform seismicity. This is often the case, for example, in the eastern United States where earthquake mechanisms are generally so poorly defined as to preclude distinction among individual faults.

As a limiting example, let us consider again the circular area of 100-km radius and zero seismogenic depth (see previous section). When the hazard is disaggregated in *R*, the contribution due to the site area source is reflected in a conditional PDF that, in general, starts from zero at zero distance, stays at zero up to the estimated depth of the seismogenic rupture surface (which is zero in this example), then starts climbing up to a peak, which is located closer to the site as the target acceleration level increases, and then slowly decreases to zero (see Fig. 7). This occurs because the distribution of *R* (or, similarly, of ln *R*) conditional on the exceedance of the S_a level is obtained by multiplying two factors:

1. the probability of occurrence, p_{occ}, which increases with distance because it is proportional to the ring area of width $r \pm (\Delta r)/2$ that increases with *r;* and
2. the probability of exceedance given occurrence, p_{exclocc}, which decreases with distance because of the attenuation of the seismic waves.

When the hazard is disaggregated in latitude and longitude, however, the p_{occ} remains spatially constant by definition of uniform seismicity. In this case, the distribution of *R* conditional on the exceedance of the target S_a reaches the peak at a distance equal to the seismogenic depth (or to the most likely depth value when multiple hypotheses are considered in the PSHA) and immediately starts decreasing.

Hence, the location of the highest peak in the latitude–longitude distribution is always coincident (at least within the bin resolution) with the site unless contributions from any other adjacent seismic sources are more prominent. Thus, the distance r_l is always shorter than the value of R^* obtained by disaggregating the hazard in $M - R - \varepsilon$ terms.

An example of hazard peaking at the site location is the Savannah River site in South Carolina. A map of the area with a model of source zones of uniform seismicity is shown in Figure 14 (Savy, 1994). The hazard corresponding to the exceedance of the 500-yr S_a (1 Hz, 5%) = 0.07 *g* at this site is disaggregated versus latitude and longitude. From Figure 15, it is evident that the contribution to hazard from the site area has a conical shape with peak at the site. At this spectral acceleration frequency and level, however, the contribution due to the Charleston area some 100 km away from the site is also significant. In this case, a double-event scenario would be most appropriate.

Conclusions and Recommendations

In this study, we have reviewed, to our knowledge, all the seismic hazard disaggregation procedures available in the literature at the time of writing and made an attempt to unveil the issues often hidden in mathematical details that may bear a considerable importance on the final perception of the hazard.

Among others, we examined a disaggregation procedure that lumps the hazard contributions only in those *M*, *R*, and ε bins that ensure that the target S_a value is equaled. The disaggregation results are affected by this matching require-

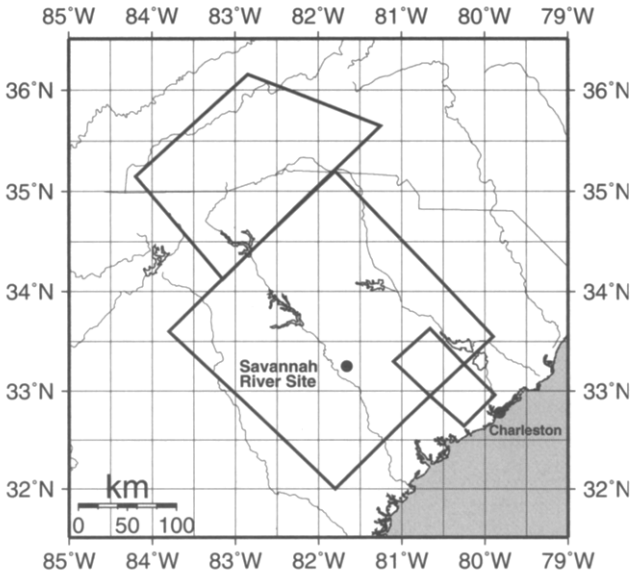


Figure 14. A PSHA source model of the region around the Savannah River site.

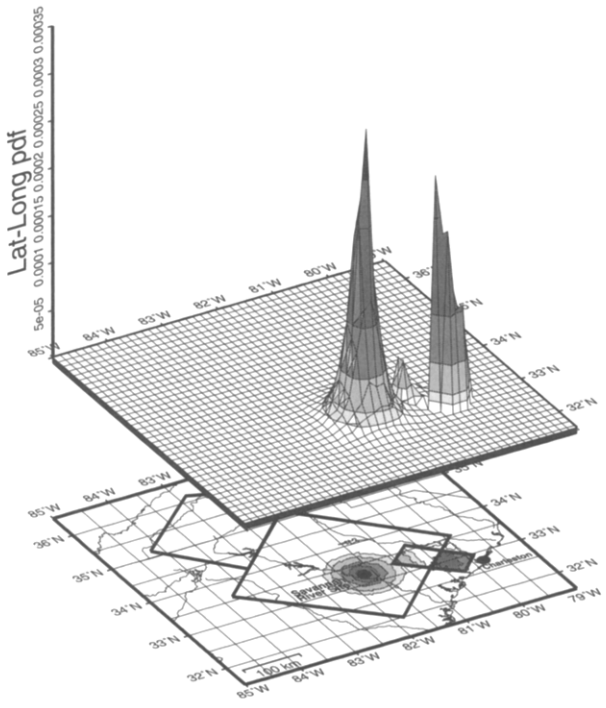


Figure 15. Three-dimensional view of contributions to hazard conditional on the exceedance of S_a (1 Hz, 5%) = 0.07 g at the Savannah River site, disaggregated in latitude and longitude. The surface shown in the figure is the latitude–longitude joint PDF.

ment. For example, the mode of the joint distribution of M, R, and ε conditional on $S_a \geqq x$ obtained using this technique does not necessarily identify the most likely event to cause $S_a \geqq x$ at the site. An alternative disaggregation approach that preserves this property is proposed.

We discussed also how multiple hypotheses on the input assumptions of the seismicity model can be considered in the disaggregation process. Contributions to the mean hazard from the ranges of the three basic variables in the PSHA were provided for a realistic case study for a site in downtown Los Angeles.

In particular, we have demonstrated how the hazard contributions may be significantly dependent on

- the distribution chosen for representing the relative contributions to hazard (PMF versus PDF);
- the variables used during disaggregation (different combinations of the basic variables M, R, and ε, but also different measures of distance, such as $D = \ln R$); and
- the binning scheme adopted.

In the PMF representation, which is in one sense the natural choice since the PSHA computations are carried out in a discrete (i.e., noncontinuous) way, the contribution of a bin represents directly the contribution to hazard from that bin. However, this value is dependent on the bin's size. This fact may lead to the undesirable results that two different analysts who use the same ground-motion predictive relation but adopt different binning schemes to disaggregate the hazard report different dominating earthquake events for the same site, the same spectral acceleration level at the same oscillator frequency, and damping.

In this regard, the PDF representation would achieve the result that this degree of arbitrariness be removed or, at least, be made less critical. The disadvantage, however, is that the fractional contribution to hazard of each magnitude, M, and distance, R (and, sometimes, ε), bin is not readily available when PDF is displayed in graphical form. If the PMF representation is adopted, we strongly recommend that details about the actual bin sizes used during computation be reported and displayed. This will make it possible for the reader both to interpret the figure properly and to estimate the accuracy of the derived statistics of interest, such as modes.

We also considered ln R rather than R as the distance variable utilized in hazard disaggregation. The use of ln R, preferred by some researchers and required by some organizations, tends to magnify the contributions from large distances and to deamplify those from short distances in comparison with the more commonly used R. We showed that disaggregating the hazard in terms of R or of ln R has an impact on the results only when a PDF representation is used. For the same binning scheme, if the relative contributions to hazard are computed by a PMF, the choice of the distance measure is irrelevant.

Because R is a more natural measure of distance than

$D = \ln R$, we favor the representation of the contributions to hazard in terms of joint PDF of $M - R - \varepsilon$ or, alternatively, of joint PMF of $M - R - \varepsilon$ computed using R bins of uniform width. The latter option is appropriate provided that the particular application does not call for a nonuniform binning strategy of R.

The issue regarding the selection of the binning scheme is practical more than conceptual and, of course, has a bearing only if a PMF representation is adopted. The choice of the bin sizes may be dictated by the specific application. If the application concerns a seismic hazard analysis conducted for generic engineering purposes, we recommend that M and ε be uniformly binned, unless particular reasons suggest otherwise. Regarding binning of distance, the choice depends on the geometry and location of the faults in the study region around the site. An increase in bin sizes for long distances is often preferable. For other applications, the binning selection criteria are of course different. For instance, a particular nonuniform binning in M is found to be appropriate for the purpose of assessing soil liquefaction resistance, while the seismic performance evaluation of a building located near a fault suggests R bins very refined at small distances.

Often times, disaggregation results are used to identify ground-motion accelerograms consistent with the earthquake events dominating the hazard. If summary statistics are to be used to determine such scenario earthquakes, in a generic application (i.e., when *a priori* no particular binning strategy is to be preferred), we prefer, again, the use of the 3D mode from the joint PDF distribution of $M - R - \varepsilon$ conditional on exceeding the target S_a.

The selection of scenario events on the basis of the modal values of M and R only (i.e., the modes of the marginal distributions instead of that of the full joint distribution) may identify earthquakes that are not the most likely events to exceed the hazard of interest at the site. We recognize, however, that there are specific cases involving, for example, the selection of appropriate near-field motions to represent forward rupture directivity effects, which may require *ad hoc* considerations.

Given the large variety of parameters that can have a bearing on the hazard contributions, at a minimum, we recommend that in the future, enough details of the disaggregation technique always be reported, such that the reader and user are sufficiently informed to make good inferences. Preferably, the graphical displays should also indicate explicitly bin sizes (if a PMF is shown) and the use of D ($= \ln R$) rather than R, if the former is used during hazard disaggregation (if a PDF is used).

It should be clearly emphasized that no one method is theoretically preferable over the others, but the potential users of the methods and of the results should be aware of the differences. We invite discussions of this study and statements of preferences by hazard analysts and users to begin a dialog that might lead to a healthy community consensus that would simplify future communication and comparisons.

Finally, we propose to supplement the disaggregation process by showing hazard contributions in terms of not distance but latitude, longitude, as well as M and ε. This permits a display directly on a typical map of the faults of the surrounding area and, hence, facilitates the identification and communication of the most hazardous earthquake locations. This information makes it easier to account for other seismic source characteristics, such as faulting style and near-source effects, during selection of scenario-based ground-motion time histories for structural analysis. The seismic hazard disaggregation in terms of latitude and longitude can be easily implemented in the Geographic Information Systems framework.

In summary, because such differences may affect users' interpretations and because there are no theoretical reasons to prefer one disaggregation option over another, we strongly encourage full reporting of the options selected. Finally, at this time, we recommend the following:

- If the application drives the binning strategy:
  - use the joint PMF of M-R-ε, and
  - report bin sizes and reasons for them.
- If the binning scheme is not clearly application driven:
  - use the joint PDF of M-R-ε or, alternatively, of M-ln R-ε, and
  - report the variable used to characterize distance (R or $D = \ln R$).
- To identify scenario earthquakes for selection of ground motions:
  - use most likely (modal) events from the joint M-R-ε or, alternatively, from the joint M-ln R-ε distribution (either PMF or PDF, according to the specific application as discussed earlier); and
  - use more than one event in multi-modal cases.
- To gain additional insights and improve communication of hazard:
  - use, additionally, 4D geographical disaggregation.

Acknowledgments

Funding for this research was provided by the Southern California Earthquake Center (SCEC) under Grant Number 699-718 and by the U.S. Nuclear Regulatory Commission through Contract NRC-04-95-075. SCEC is funded by NSF Cooperative Agreement EAR-8920136 and USGS Cooperative Agreements 14-08-0001-A0899 and 1434-HQ-97AG01718. SCEC Contribution Number 407. This support is gratefully acknowledged. We are also very thankful to Dr. Norman Abrahamson, who, besides insightful comments, provided the PSHA software package and the seismotectonic model of southern California used in this study. We also thank Dr. K. Campbell, Dr. B. Youngs, and Dr. J. Savy, whose comments improved the quality of this article. Finally, we wish to acknowledge the use of the Generic Mapping Tools software package by Wessel and Smith (1991) to produce most of the figures in this article.

References

Abrahamson, N. A. (1996). Personal communication, Pacific Gas & Electric Co., San Francisco, May.

Abrahamson, N. A. and W. J. Silva (1997). Empirical response spectra attenuation relations for shallow crustal earthquakes, *Seism. Res. Lett.* **68,** 94–127.

Benreuter, D. L., A. C. Boissonnade, and C. M. Short (1996). Investigation of techniques for the development of seismic design basis using probabilistic seismic hazard analysis, Lawrence Livermore National Laboratory, Rept. NUREG/CR-6606, prepared for U.S. Nuclear Regulatory Commission.

Boissonnade, A., N. Chokshi, D. Benreuter, and A. Murphy (1995). Determination of controlling earthquakes from probabilistic seismic hazard analysis for nuclear reactor sites, *Proc. of the Thirteenth International Conference on Structural Mechanics in Reactor Technology,* SMiRT 13, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, August 13–18, 771–776.

Campbell, K. C. (1997). Empirical near-source attenuation of horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra, *Seism. Res. Lett.* **68,** 154–179.

Chapman, M. C. (1995). A probabilistic approach to ground-motion selection for engineering design, *Bull. Seism. Soc. Am.* **85,** 937–942, June.

Cornell, C. A. (1968). Engineering seismic risk analysis, *Bull. Seism. Soc. Am.* **58,** 1583–1606.

Cramer, C. H. and M. D. Petersen (1996). Predominant seismic source distance and magnitude maps for Los Angeles, Orange, and Ventura Counties, California, *Bull. Seism. Soc. Am.* **86,** 1645–1649.

Ishikawa, Y. and H. Kameda (1988). Hazard-consistent magnitude and distance for extended seismic risk analysis, *Proc. of Ninth World Conference on Earthquake Engineering* **II,** 89–94.

Ishikawa, Y. and H. Kameda (1991). Probability-based determination of specific scenario earthquakes, *Proc. of Fourth International Conference on Seismic Zonation* **II,** 3–10.

Ishikawa, Y. and H. Kameda (1993). Scenario earthquakes vs. probabilistic seismic hazard, *Proc. of Fourth International Conference on Structural Safety and Reliability* **3,** 2139–2146, Innsbruck, Austria.

Kameda, H., C. H. Loh, and M. Nakajima (1994a). A comparative study in Japan and Taiwan by means of probabilistic scenario earthquakes, *Proc. of Fourth KAIST-NTU-KU Tri-Lateral Seminar/Workshop on Civil Engineering,* Kyoto, Japan, November 22–23, 27–37.

Kameda, H., Y. Ishikawa, and W. Li (1994b). Probability based determination of scenario earthquakes, in *Seismic Risk Assessment of Urban Facilities in a Sedimentary Region,* Chap. 4, Natural Hazard Reduction and Mitigation in the East Asia, Final Rept. Part 3, March, 67–85.

Kimball, J. K. and H. Chander (1996). Department of Energy seismic siting and design decisions: consistent use of seismic hazard analysis, *Proc. Twenty-Fourth Water Reactor Safety Information Meeting* **3,** 141–160, Bethesda, October 21–23, U.S. Nuclear Regulatory Commission, NUREG/CP-0157.

Kramer, S. L. (1996). *Geothecnical Earthquake Engineering,* Prentice-Hall International Series in Civil Engineering and Engineering Mechanics, Prentice-Hall, Upper Saddle River, New Jersey.

McGuire, R. K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop, *Bull. Seism. Soc. Am.* **85,** 1275–1284.

McGuire, R. K. and K. M. Shedlock (1981). Statistical uncertainties in seismic hazard evaluations in the United States, *Bull. Seism. Soc. Am.* **71,** 1287–1308.

National Research Council (NRC) (1988). Probabilistic seismic hazard analysis, Rept. of the Panel on Seismic Hazard Analysis, National Academy Press, Washington, D.C.

Risk Engineering, Inc. (REI) (1989). EQHAZARD primer, Special Rept. NP-6452-D, Research Project P101-46, prepared for Seismicity Owners Group and Electric Power Research Institute.

Savy, J. B. (1994). Seismic Hazard Characterization at the DOE Savannah River Site (SRS) Status Report, Lawrence Livermore National Laboratory, Rept. UCRL-ID-117452, June 24.

Seed, H. B. and K. L. Lee (1966). Liquefaction of saturated sands during cyclic loading, *J. Soil Mech. Found. Div. ASCE* **92,** 105–134.

Seed, H. B., I. M. Idriss, and I. Arango (1983). Evaluation of liquefaction potential using field performance data, *J. Geotech. Eng. ASCE* **109,** 458–482.

Senior Seismic Hazard Analysis Committee (SSHAC) (1997). Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts, NUREG/CR-6372, UCRL-ID-122160, Main Report 1, prepared for Lawrence Livermore National Laboratory, April.

Somerville, P., N. F. Smith, R. W. Graves, and N. A. Abrahamson (1997). Modification of empirical strong motion attenuation relations to include the amplitude and duration effects of rupture directivity, *Seism. Res. Lett.* **68,** 199–222.

Spudich, P. (1997). What seismology may be able to bring to future building codes, *Proc. of the ATC-35 Ground Motion Initiative Workshop,* Rancho Bernardo, California, July 30–31.

Stepp, J. C., W. J. Silva, R. K. McGuire, and R. T. Sewell (1993). Determination of earthquake design loads for high level nuclear waste repository facility, *Proc. of Fourth* DOE *Natural Phenomena Hazard Mitigation Conference* **II,** 651–657, Atlanta.

U.S. Nuclear Regulatory Commission (U.S. NRC) (1997). *Identification and Characterization of Seismic Sources and Determination of Safe Shutdown Earthquake Ground Motion,* Regulations 10 CFR Part 100, Regulatory Guide 1.165, Appendix C, Washington, D.C., 1.165-17/1.165-23.

U.S. Department of Energy (U.S. DOE) (1996). *Natural Phenomena Hazards Assessment Criteria,* DOE-STD-1023-96, Washington, D.C.

Wessel, P. and W. H. F. Smith (1991). GMT free software helps map and display data, *EOS* **72,** 441 and 445–446.

Department of Civil and Environmental Engineering
Stanford University
Stanford, California 94305-4020

Manuscript received 19 November 1997.