STATE OF THE PRACTICE OF SEISMIC HAZARD EVALUATION

Norman A. Abrahamson[1]

INTRODUCTION

The main use of seismic hazard analyses is to develop rock outcrop or stiff soil ground motions for use in design. The quantitative descriptions of the ground motions can be in terms of simple scalar values (e.g. peak acceleration, peak velocity, peak displacement, response spectral values, or Arias intensity) or it can be in terms of time histories of acceleration, velocity, and displacement. Typically, the design ground motions are then used in geotechnical and/or structural engineering analyses. This paper does not address site-specific site response analyses. Site response is addressed in the accompanying theme lecture paper by Dobry.

The topic of this paper is the state-of-the-practice of seismic hazard analysis. The actual practice of seismic hazard analysis varies tremendously from poor to very good. The large variability in practice is not simply a reflection of project budgets; a large variation in practice exists for similar projects. Although the basic methodologies used in seismic hazard analysis are well established, as represented by short courses and seminars on seismic hazard analysis as well as numerous seismic hazard reports, these basic methodologies are generally not well understood.

Usually, the state-of-the-practice is defined as what the "average" practitioner would do. With large variations in the quality of seismic hazard evaluations in practice, it is difficult to determine what is "average". As a result, we have taken a different approach in this paper. Many of the poor aspects of the practice are due to widespread misunderstandings of the basics of seismic hazard analysis and misleading or inconsistent terminology that is used in practice. This applies to both those conducting the seismic hazard evaluation and those applying the results in an engineering analysis. In general, the high end of the practice is very good, but the results are often not well understood by the engineers who need to apply them. To address with this situation, we first discuss the some of the shortcomings of the current practice, and then discuss the state-of-the-practice at the high end. As a start, we describe many common misunderstandings regarding the development and interpretation of design ground motions that contribute to poor practice.

APPROACHES TO DEVELOPING DESIGN GROUND MOTIONS

There are two basic approaches to developing design ground motions that are commonly used in practice: deterministic and probabilistic. The selection of the probabilistic or the deterministic approach remains controversial. In our opinion, much of the controversy results from widespread misunderstandings of the two approaches.

First, there are different meanings of "deterministic" and "probabilistic" approaches to design ground motion in different countries. In the United States and most of the world, the deterministic approach selects individual earthquake scenarios (earthquake magnitude and distance) and specified ground motion probability level (by tradition, it is typically either 0 or 1 standard deviation above the median). The probabilistic approach considers all possible earthquake scenarios (all possible magnitude and distance combinations) as well as all possible ground motion probability levels (e.g. from -3 to 3 standard deviations from the median) along with their associated probabilities, and it computes the probability that any of the scenarios will produce a ground motion greater than the specific test value. The deterministic approach leads to a single ground motion for each scenario considered, whereas, the probabilistic approach leads to a hazard curve, giving the probability of exceeding various ground motion values

In contrast, in Japan, the terms "deterministic" and "probabilistic" usually refer only to the method used for estimating the ground motion from a specific earthquake scenario. In this case, "probabilistic" refers to using an empirical attenuation relation to predict the ground motion; whereas, "deterministic" refers to using a seismological model of the earthquake fault rupture, wave propagation (e.g. numerical simulation) for a specific fault geometry, seismic velocity structure, and seismological source properties

[1] Pacific Gas and Electric Company, Mail Code N4C, PO Box 770000, San Francisco, CA 94177 USA

(slip distribution, rise-time, rupture velocity, ect.) to estimate the ground motions. Both of these approaches would be called "deterministic" in the USA. This difference in terminology creates a lot of confusion. In this paper, we will be using the more widely used definition of deterministic and probabilistic as used in the USA.

COMMON MISUNDERSTANDINGS IN SEISMIC HAZARD ANALYSIS

Below we discuss common misunderstandings of the deterministic and probabilistic approaches to developing design ground motions. The list of common misunderstandings given in Table 1 was developed from experience giving short courses and seminars to practicing earthquake engineering professional on seismic hazard analysis and from experience reviewing seismic hazard reports in the USA and other countries.

Using the deterministic approach requires selecting a limited set of earthquake scenarios (magnitude, style-of-faulting, distance from the earthquake to the site), and the number of standard deviations from the median for the design ground motion level. Although the deterministic approach is very straightforward, the results are often misunderstood. Using the probabilistic approach requires defining a comprehensive set of scenarios as well as their probability of occurring. Probabilistic seismic hazard analysis (PSHA) is more complicated than deterministic analysis and is often seen as a "black box" by practicing engineers that use the results. While the mathematical formulation of PSHA as shown later can seem complex, in reality, a PSHA is just a large number of deterministic analysis with the added feature that it considers how likely the different deterministic scenarios are.

Many of the common misunderstandings in seismic hazard evaluations are a result of misleading terminology that is used in this field; for example, the term "maximum" usually does not refer to a true maximum. Other common misunderstandings are related to misinterpretations of attenuation relations. Some of the items listed in Table 1 may seem minor, but they contribute to the overall confusion in the field and to the large variability in the preparation of seismic hazard analyses.

TABLE 1: COMMON MISUNDERSTANDINGS AND SHORTCOMINGS IN CURRENT PRACTICE
1. The "maximum magnitude" is thought to be the largest magnitude that can occur on a fault.
2. Attenuation relations are assumed to always give the larger horizontal component.
3. The "distance" measures are used in various attenuation relations as though they were the same.
4. The terms "mean" and "median" for ground motion are used interchangeably.
5. Ground motions greater than one standard deviation above the median are thought to be unreasonable.
6. Deterministic ground motion estimates are thought not to include any probability
7. The ground motions from the Maximum Credible Earthquake are thought to be the worst case.
8. The terms "recurrence interval" and "return period" are used interchangeably.
9. Including the standard deviation of the ground motion in a PSHA is thought to be conservative.
10. The hazard curve is not well understood
11. The result of the probabilistic approach is thought to be unbounded ground motions.
12. Combining the hazard from multiple faults is thought to imply that the faults rupture at the same time.
13. In a PSHA, the ground motions for a given return period are incorrectly computed for each fault separately.
14. Probabilistic approach is thought to be unable to provide design earthquakes for the ground motion levels
15. The branches on logic trees for fault segmentation are often not set up correctly.

1. The "Maximum Magnitude" Is Thought To Be The Largest Magnitude That Can Occur On A Fault.

The "maximum magnitude" is often thought to be the largest magnitude that can occur on a fault. This is a case where the terminology is misleading. In either probabilistic or deterministic seismic hazard analyses, a magnitude is estimated for each fault source based on the fault dimension (area or length), or fault displacements. For example, a commonly used model relating magnitude to fault rupture area is given by Wells and Coppersmith (1994):

$$M = 0.98 \, \mathrm{Log}(A) + 4.07 \quad (1)$$

where M is the moment magnitude and A is the fault rupture area in km^2. This magnitude is often mislabeled as the "maximum magnitude". There is significant random (aleatory) variability associated with these scaling relations. For example, the magnitude from eq. (1) has a standard deviation of 0.24 magnitude units because not all earthquakes of a given magnitude have the same rupture area (e.g. variability in the stress-drop). The magnitude given by eq. (1) is the mean magnitude for the given fault rupture area. The true maximum magnitude is the magnitude at which the magnitude distribution is truncated (Figure 1). In Figure 1, the maximum magnitude is shown at 2 standard deviations above the mean. In practice, it is common to see the mean magnitude listed as the "maximum magnitude".

Some of the ideas for less confusing notation are awkward. For example, the term "mean maximum magnitude" could be used, but this is already used for describing the average "maximum magnitude" from alternative scaling relations (e.g. through logic trees).

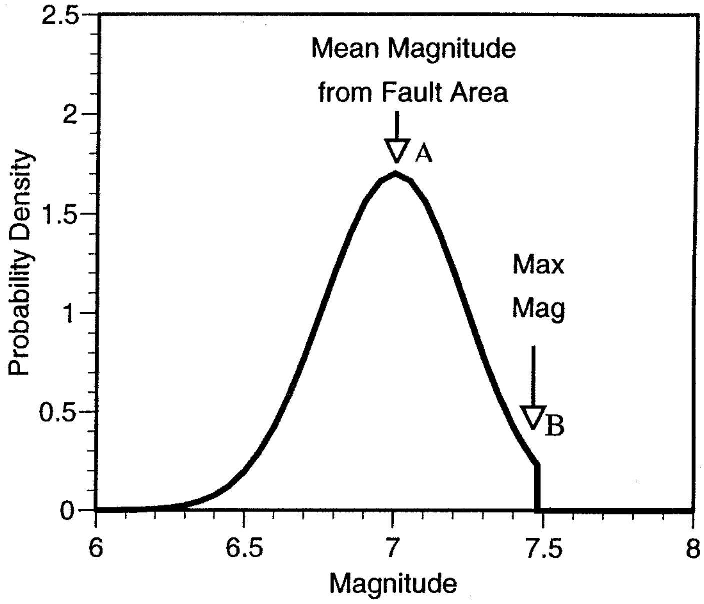

Figure 1. Comparison of mean magnitude and maximum magnitude for a given fault area based on truncation at 2 standard deviations above the mean.

With fault segmentation, the issue becomes more confused because the "maximum" magnitude could be based on rupture of a single segment or multiple segments. For fault sources, an alternative to the term "maximum magnitude" for point A is the term "mean characteristic earthquake magnitude" for a specified segment or set of segments, and reserve "maximum magnitude" for point B. Rather than struggling to define a different terminology, a better approach is to use clear language to describe how the value is obtained.

2. Attenuation Relations Are Assumed To Always Give The Larger Horizontal Component.

There is not a standard for treating the two horizontal components of ground motion in attenuation relations. Some attenuation relations use the larger of the two horizontal components, some use the arithmetic mean of the two components, some use the geometric mean of the two components, and some use both components (e.g. a random horizontal component). These differences are often ignored when applying the attenuation relations. Engineers using the design ground motions often assume that the design motions are for the larger horizontal component because most hazard reports are not specific about this aspect of the design ground motion.

Most attenuation relations developed in the United States are based on the average horizontal (geometric mean) component, but in other counties, the larger component is commonly used. For peak acceleration, the difference between the larger horizontal component and the average horizontal component is about 15%. This difference becomes larger at long periods ($T > 0.5$ seconds), particularly for near-fault ground motions.

Seismic hazard evaluations should clearly state what the horizontal design spectrum represents: an average horizontal component, a random horizontal component, or the larger of the two horizontal components.

3. The "Distance" Measures Are Used In Various Attenuation Relations As Though They Were The Same.

A single consistent definition of the site-to-source distance has not been widely adopted by different authors of attenuation relations. In applications of the attenuation relations, these differences in distance definitions are often ignored, but it is important to use the appropriate distance measure with each attenuation relation, particularly for short distances.

Commonly used distance measures include the following: r_{jb}, the closest horizontal distance to the vertical projection of the rupture (the "Joyner-Boore" distance); r_{rup}, the closest distance to the rupture surface (slant distance); r_{seis}, the closest distance to the seismogenic part of the rupture surface (assumes that near-surface rupture in sediments is non-seismogenic); r_{cent}, the centroid distance; r_{hyp}, the hypocentral distance; and r_{epi}, the epicentral distance. The first three distances measure some sort of closest distance to the rupture plane, whereas, the last three distances are point source measures. For large-magnitude earthquakes, the closest distance measures are generally preferred over the point source distances. Some of these different distance measures are shown graphically in Figure 2 for a vertical fault and for a dipping fault. The main differences are for sites located close to the fault. The appropriate distance measure should be used with each attenuation relation considered.

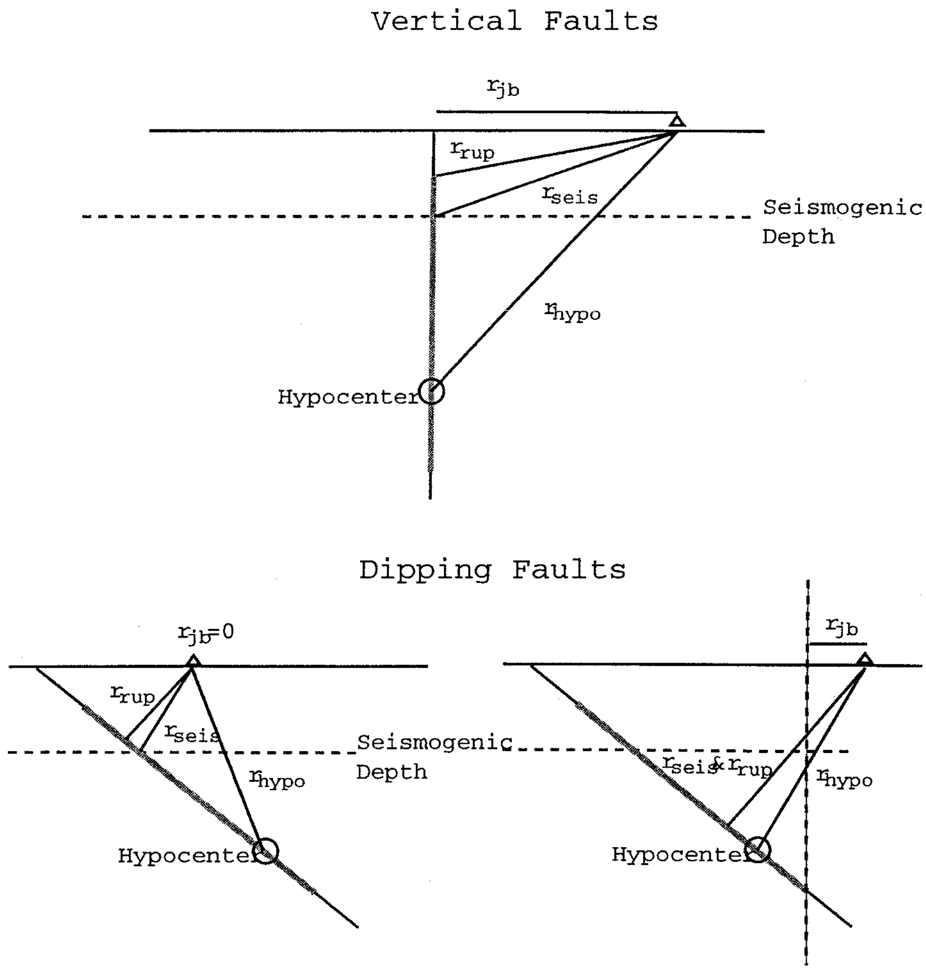

Figure 2. Various source to site distance measures for ground motion attenuation models shown in cross-section. The shaded region indicates the extent of the rupture.

4. The Terms "Mean" And "Median" For Ground Motion Are Used Interchangeably.

It is common to find the terms "mean" and "median" used interchangeably in descriptions of ground motions. Peak ground motion parameters, such as peak acceleration, peak velocity, peak displacement, and response spectral values are generally assumed to be lognormally distributed (Figure 3). For a log normal distribution, the mean and median are not the same. The mean is larger than the median due to the skewness of the lognormal distribution. Attenuation relations developed from statistical regression analyses usually give the median ground motion and the standard deviation of the logarithm of the ground motion. The relation between the median and the mean is given by:

$$\text{Mean} = \text{Median} * \exp(-\sigma^2/2) \quad (2)$$

where σ is the standard deviation in natural log units. For a typical standard deviation of about 0.6, the mean is about 20% larger than the median.

In most cases, when the term "mean" is used with respect to ground motion, it should have been "median". This misuse of terminology seems minor, but can create confusion. In some cases, practitioners have calculated the mean using eq. (2) to satisfy a "mean ground motion" design criterion, when the intent of the criteria was to use the median ground motion.

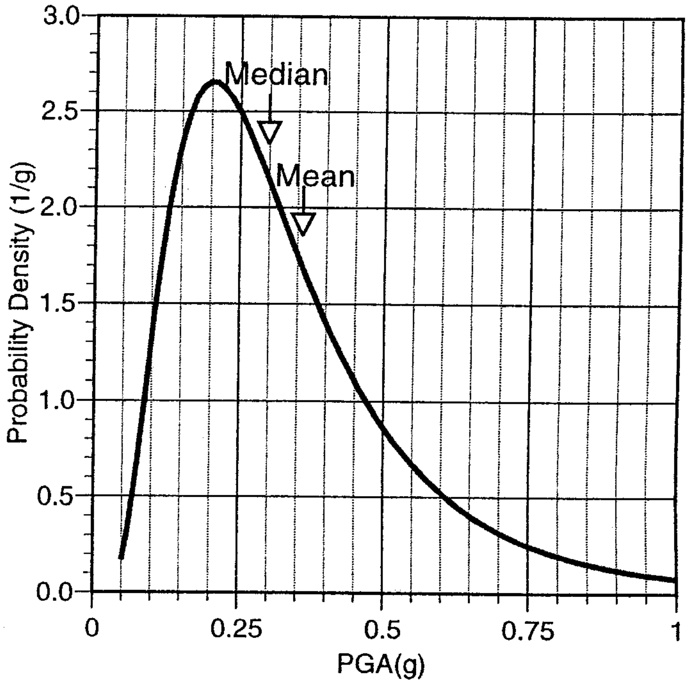

Figure 3. Comparison of the mean and median for a lognormal distribution with a standard deviation of 0.6.

5. Ground Motions Greater Than One Standard Deviation Above The Median Are Thought To Be Unreasonable.

A common belief is that that ground motions greater than the 84th percentile are "unreasonable

Since the attenuation relations are probabilistic descriptions of the ground motion, ground motions greater than 1 standard deviation above the median can and do occur, but just not very often. As the number of recordings in the strong motion data sets become larger, more observations exceeding one standard deviation are found. In the current strong motion data sets, there are hundreds of recorded ground motions that exceed one standard deviation above the median. Statistical evaluations of recorded ground motion data indicate that the data begin to deviate from a lognormal distribution above 2 standard deviations, indicating that the probability model for ground motion is less reliable above 2 standard deviations. An example of such an analysis is shown in Figure 4 which compares the observed distribution of 1080 log peak acceleration residuals to a lognormal distribution. In this type of quantile-quantile plot, the predicted number of standard deviations is on the x-axis and the observed number of standard deviations is on the y-axis. If the data points fall on the line, then the data follow the assumed distribution. The data (shown by the black dots) are close to the

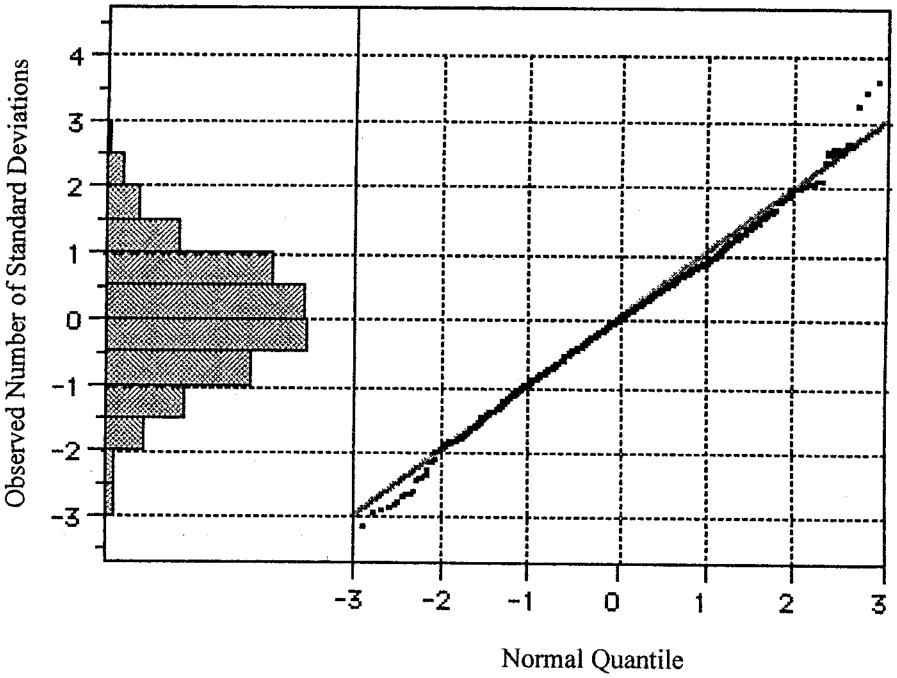

Figure 4. Normal quantile plot comparing the observed distribution of peak accelerations with the assumed lognomral distribution. If the data follow a lognormal distribution, the points would lie on the line.

predicted distribution (shown by the solid line) between –2 and 2 standard deviations; the data begin to deviate from the line above 2 standard deviations. This indicates that the computed probabilities of ground motions more than 2 standard deviations above the median, based on a lognormal distribution, may not be reliable. As a result, a common practice in probabilistic seismic hazard evaluations is to limit the number of standard deviations to be less than 3.

Another way to evaluate the distribution of ground motions is to tabulate the observed number of ground motion residuals that exceed a certain number of standard deviations and compare that to the predicted number. Table 2 compares the observed and predicted number of ground motions that would exceed 0 to 3 standard deviations above the median for a data set of 1050 recordings. This table also shows that ground motions closely follow the assumed lognormal distribution up to 2 standard deviations above median.

TABLE 2 : COMPARISON OF THE OBSERVED AND EXPECTED NUMBER OF POINTS EXCEEDING STANDARD DEVIATION LEVELS.

Standard Deviations	Number of Observations (out of 1080)	
	Expected	Observed
>0.0	540	547
>0.5	333	327
>1.0	171	143
>1.5	72	63
>2.0	25	26
>2.5	7	8
>3.0	1	3

In some seismic hazard evaluations, the ground motion distribution is truncated at 1 standard deviation above the median (implying that ground motion greater than 1σ above the median is unrealistic). This is an unconservative assumption. There is no empirical evidence to support truncation of the ground motion at less than 2σ. Doing so will result in an under prediction of the hazard.

Another contributing factor to the belief that ground motion greater than 1 sigma are unreasonable comes from the misconception that deterministic seismic hazard analysis is based on statistical, not probabilistic, descriptions of the ground motion. The thinking is that since statistics is used in developing the attenuation relations, then the attenuation relations must be statistical. The difference between statistics and probability may seem a minor detail, but it can affect the interpretation of the standard deviation of the ground motion. By considering the standard deviation to be statistical, it makes it easy to discount ground motions above 1 standard deviation as statistical fluctuations that don't need to be considered. In that sense, the median or 84th percentile values are considered to be just different "statistical levels".

In reality, attenuation relations consist of equations for the median ground motion, the standard deviation, and the form of the distribution (lognormal). The ground motion for a given magnitude and distance is described by a probability distribution. Statistics is the tool used to estimate the probability distribution of the ground motion. The statistical uncertainty leads to alternative estimates of the value of the median, the value of the standard deviation, and the form of the distribution. The standard deviation itself does not represent statistical uncertainty, but rather it represents randomness in the earthquake source process, wave propagation, and site response.

6. Deterministic Ground Motion Estimates Are Thought Not To Include Any Probability

A common belief is that probability is not considered in developing deterministic design ground motions. In reality, some aspects of probability are typically used in the development of deterministic ground motion, but they are not always acknowledged. As noted above, there is probability is involved in selecting the maximum magnitude for a source, the ground motion level (number of standard deviations), and, for areal sources, the distance from the earthquake to the site.

For the magnitude, probability is implicitly considered in terms of selecting an earthquake rupture scenario for segmented faults. The magnitude could be based on single segment ruptures or multiple segment ruptures. This selection depends on how likely it is for multiple segments to rupture in one earthquake.

The ground motion level is most often selected by the importance of the structure, but in some cases, the activity rate of the fault is used to help in the selection of either the median or 84th percentile ground motion. For example, for very low activity rate faults, a median ground motion may be appropriate, whereas, for high activity faults, an 84th percentile ground motion may be appropriate.

Using a deterministic approach does not avoid all aspects of probability. The concept of what deterministic ground motion is "reasonable" must address how likely the ground motion is to occur (at least in some rough sense). This misconception that deterministic approach does not include probability contributes to the confusion between deterministic and probabilistic approaches. In particular, the concept that the deterministic approach ensures a "safe" design, whereas, a probabilistic approach accepts some failures comes in part from this misunderstanding of the role of probability in selecting deterministic design ground motions.

7. The Ground Motions From The Maximum Credible Earthquake Are Thought To Be The Worst Case.

When the deterministic approach is used for design of important structures, the Maximum Credible Earthquake (MCE) is commonly used. The most common misunderstanding about the deterministic method is that the MCE is often interpreted to be a "worst-case" ground motion as implied by the term "maximum". As a minimum, there are three parameters needed to define the MCE ground motion: magnitude, distance, and number of standard deviations for the ground motion.

For faults, the magnitude of the MCE is usually the mean magnitude from a relation between the fault dimension (length or area) and the magnitude. As discussed above, this is not the maximum magnitude.

For faults, the distance from the fault to the site is typically the closest distance, so this is a worst case; however, for sources that are described by zones rather than specific faults, the distance to the site is typically not the closest distance. (For the source zone containing the site, the closest distance would be based on some depth directly below the site). For zones, the earthquake is put at some distance away from the site and at some depth. Just how far away to put it has always been an issue for the deterministic approach..

TABLE 3: COMPARISON OF DETERMINISTIC GROUND MOTIONS FOR T=1 SECOND SPECTRAL ACCELERATION (5% DAMPING)

Number of Std. Dev.	Mean Mag (M=7.0) D=10 KM	Maximum Mag (M=7.5) D=10 KM
-3.0	0.06	0.09
-2.5	0.08	0.12
-2.0	0.10	0.15
-1.5	0.14	0.19
-1.0	0.18	0.25
-0.5	0.24	0.33
0.0	**0.31**	0.42
0.5	0.41	0.55
1.0	**0.54**	0.71
1.5	0.71	0.92
2.0	0.94	1.20
2.5	1.24	1.55
3.0	1.63	**2.01**

The ground motion level is usually taken as either the median (50th percentile) or one standard deviation above the median (84th percentile). This is not the worst case ground motion; the worst case ground motion would be 2 to 3 standard deviations above the median. (See discussion of largest reasonable ground motions above)

Table 2 shows the T=1.0 second spectral accelerations at 5% damping for a range of number of standard deviations for the example case computed using the Sadigh et al (1997) attenuation relations for strike-slip earthquakes at a rock site. If we take the worst case to be 3 standard deviations above the median for ground motion and 2 standard deviations above the mean for the magnitude, then the worst case ground motion is 2.01g as compared to 0.54g for the 84th percentile from the MCE. The main factor in increasing the ground motion is the number of standard deviations. For this example, the increase in magnitude (from the mean to the maximum) causes about a 30% increase, whereas, the increase in the number of standard deviations from 1 to 3 causes an increase of about 200%.

The common response to such a discussion of worst case ground motions is that it is piling conservatism on top of conservatism and the end result is unreasonably conservative. In other words, the worst case ground motions are so rare that it would be unreasonable to design for them. This begs the question: what is reasonable? The typical practice does not use worst case ground motion; however, seismic hazard reports generally do nothing to dispel the misconception that MCE ground motions are maximum ground motions.

8. The Terms "Recurrence Interval" And "Return Period" Are Used Interchangeably.

The recurrence interval and return period have different meanings, but in practice, these terms are often used interchangeably causing confusion. The "recurrence interval" refers to the occurrence of earthquakes on a seismic source, not ground motion. The recurrence interval is the time interval between earthquakes of given magnitude or larger for a specified seismic source or set of sources. The "return period" refers to the occurrence of ground motion at a site. The return period is the one over the annual rate at which a ground motion level is exceeded at a site.

Return period is used instead of annual rate of exceedance because it is more convenient to say "500 years" than it is to say an " 0.002 per year"; however, this does not imply that the hazard evaluation is applicable over the time period of implied by the return period. In particular, for very small probabilities (e.g. return period of 10,000 years), the ground motion has a 1/10,000 annual chance of being exceeded. It may not represent what may happen over the next 10,000 years.

As an example, consider the simplest case in which there is just a single fault that generates only one magnitude earthquake. If the earthquake occurs every 300 years on average then the recurrence interval is 300 years. When the earthquake occurs, the median ground motion will be exceeded 50% of the time. Therefore, for this simplified example, the return period of the median ground motion is 600 years (300 divided by 0.5). Similarly, the return period of the 84th percentile ground motion is 3750 years (300 divided by 0.16).

9. Including The Standard Deviation Of The Ground Motion In A PSHA Is Thought To Be Conservative.

Some commercially available software for computing probabilistic seismic hazard includes an option of using a zero for the standard deviation of the ground motion attenuation relation. There is widespread misinterpretation of what this option does. It is commonly thought that including the standard deviation means that the hazard analysis is run using 84th percentile ground motions. As a result, the thought is that including the standard deviation is "more conservative" than using just the median ground motion. It is not "more conservative" to include the standard deviation in a PSHA; instead using a zero for the standard deviation significantly underestimates the hazard.

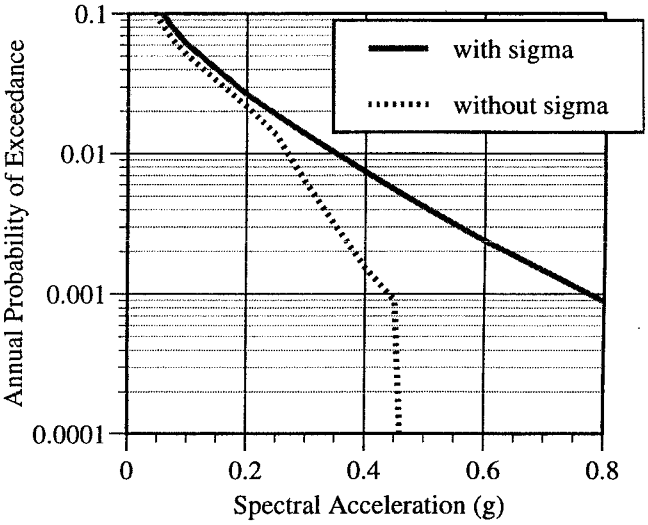

Figure 5. Example probabilistic hazard calculation showing the underprediction of the hazard caused by ignoring the standard deviation of the ground motion.

In PSHA, the standard deviation of the ground motion is part of the probabilistic description of the ground motion (see item 5 above). When probabilistic seismic hazard analysis was first developed by Cornell (1968), this probabilistic nature of the ground motion for a given magnitude and distance was not considered. Many of the early hazard studies followed this approach and set a precedent for ignoring the standard deviation. Part of the reason for not including the variability of the ground motion was that many early attenuation relations did not provide estimates of the standard deviation of the ground motion.

The importance of this missing source of variability was recognized in the 1980s. Some current seismic hazard evaluations continue to ignore the ground motion variability. This practice leads to underestimation of the hazard. As an example, the effect of ignoring the variability of the ground motion on the computed hazard is shown in Figure 5. The underestimation of the hazard becomes greater at low probabilities (long return periods).

10. The Hazard Curve Is Not Well Understood

Although probabilistic analysis is becoming common around the world and is an integral part of the 2000 International Building Code, a surprisingly large number of practicing engineers do not have a good understanding of what a hazard curve means. The hazard curves are simply used to read off the ground motion at a specified return period. This is an unhealthy state of practice. Without a good understanding of what hazard curves mean, there is not only a much greater chance that errors will be made in the hazard evaluation, but engineering judgement used in applying the design ground motions may also be affected.

One useful way to improve the understanding of a hazard curve is to break it down into scenario ground motions. By computing the difference in the hazard curve at closely spaced ground motions (e.g. every 0.1g), a histogram showing the rates of ground motions occurring at a site can be computed. An example is shown in Figure 6. In this example, there is a 0.005 annual probability of experiencing a peak acceleration between 0.1 and 0.2 g; and there is a 0.00006 annual probability of experiencing a peak acceleration between 0.4 and 0.5g. This type of ground motion histogram is much easier to understand because it deals with scenario ground motions. The current state-of-the-practice does not include ground motion histograms, but would be one way to greatly improve the general understanding of probabilistic ground motion evaluations.

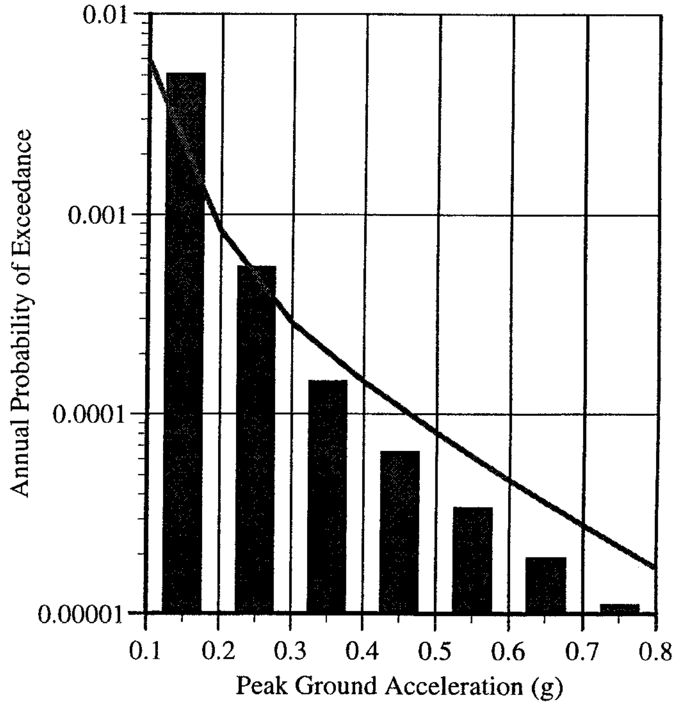

Figure 6. Example of a ground motion histogram computed from the hazard curve.

11. Probabilistically Based Design Ground Motions Are Thought To Be Unbounded.

The results of a probabilistic seismic hazard evaluation are typically presented in terms of the probability of exceedance. For example, a design ground motion may be defined as the ground motion with a 10% chance of being exceeded in 50 years. This terminology can be misleading. Saying that there is a 10% chance of exceeding a specified ground motion does not tell how large the ground motion could be. This terminology is misinterpreted to imply that the design ground motions are unbounded.

This misunderstanding could be addressed easily by considering the probability of non-exceedance. For example, the ground motion with 10% chance of being exceeded in 50 years is also the ground motion with 90% chance of not being exceeded in 50 years. A simple change in the terminology can remove this cause of confusion.

When ground motions are developed for a deterministic analysis, the terminology is probability of not being exceeded. For example, using the 84th percentile ground motion for the MCE is the ground motion with an 84 percent probability of not being exceeded if the postulated scenario earthquake were to occur. The deterministic ground motions are "unbounded" in the same way as the probabilistic ground motions. The difference in terminology for design motions adds to the confusion between deterministic and probabilistic approaches. Using the probability of not being exceeded makes more sense and should be used for the probabilistic analysis.

There is, however, a legitimate concern about how large the ground motion could be. This can be addressed using the ground motion histograms discussed in item 10 above. This figure shows how likely larger ground motions are. The flatter the slope of the hazard, the more likely it is to have larger ground motions.

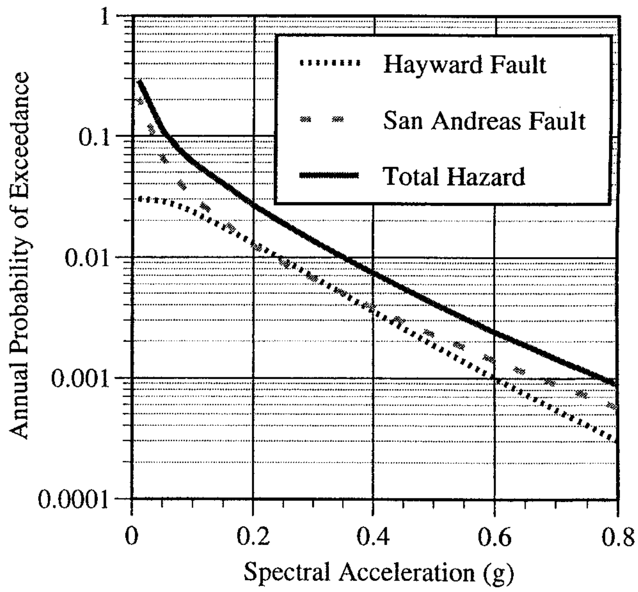

Figure 7. Example hazard calculation from multiple faults. The total hazard is the sum of the hazards from the individual faults.

12. Combining The Hazard From Multiple Faults Is Thought To Imply That The Faults Rupture At The Same Time.

Some of the basic concepts of probabilistic hazard analysis are often misunderstood. For example, the process of combining the hazard from multiple faults has been misinterpreted to imply that the faults rupture simultaneously (which would be very unlikely). Adding the hazard from multiple sources simply computes the chance that any of the sources generates a ground motion exceeding the test value.

For example, the T=1 second spectral acceleration hazard from the Hayward and San Andreas faults for a site in San Francisco is shown in Figure 7. The total hazard is the sum of the hazard from both fault (see eq. 5 below). This does not imply that the two faults rupture at the same time, rather, it considers that either fault is capable of generating design level ground motions at the site. For this example, a return period of 500 years corresponds to a spectral acceleration of 0.62g. At this return period, about 60% of the hazard is from the San Andreas fault and 40% of the hazard is from the Hayward fault.

13. In A PSHA, The Ground Motions For A Given Return Period Are Incorrectly Computed For Each Fault Separately.

Another source of misinterpretation related to the hazard curves from multiple faults is that the hazard from the individual faults is sometimes incorrectly treated separately. That is, for a given return period, the ground motion is estimated from the hazard curves from the individual faults rather than from the total hazard. For example, using the hazard shown in Figure 7, the spectral acceleration for a 1000 year return period is 0.78g, but some evaluations incorrectly specify the ground motion from each fault; in this case, 0.68g from the San Andreas and 0.60 g from the Hayward for a 1000 year return period. This is not correct. The ground motion should be estimated only from the total hazard curve, not from hazard curves from individual sources. The goal of a probabilistic analysis is to estimate the chance that a ground motion occurs at a site regardless of the source that generates the earthquake.

This incorrect interpretation of hazard curves stems from the need to define earthquake scenarios for the design ground motion. Associating scenario earthquakes to the design ground motion level is done through deaggregation (discussed later). The different sources will have different scenario earthquakes (in terms of magnitude, distance, and number of standard deviations of the ground motion), but the level the ground motion is the same for all faults (as given by the total hazard curve).

14. Probabilistic Approach Is Thought To Be Unable To Provide Design Earthquakes For The Ground Motion Levels

The deterministic approach results in an earthquake scenario (magnitude and distance) as well as the ground motion level. The probabilistic approach is often mistakenly thought to only provide a ground motion level. This misunderstanding results from probabilistic evaluations that provide only the hazard curves without a deaggregation. As discussed later under "Deaggragation", there is a well established procedure for determining earthquakes scenarios to associate with the probabilistic ground motion. Too often, however, deaggregations are not included in probabilistic seismic hazard evalauations in current practice. Without a deaggregation, the probabilistic evaluation gives the ground motion level without information on the earthquake scenarios that cause to the ground motion.

The state-of-the-practice for probabilistic seismic hazard evaluations needs to improve to always include deaggregations as standard practice. The process of deaggregating the hazard is not difficult or time consuming, but it adds a tremendous amount of information about the seismic hazard which greatly improves understanding of the PSHA results.

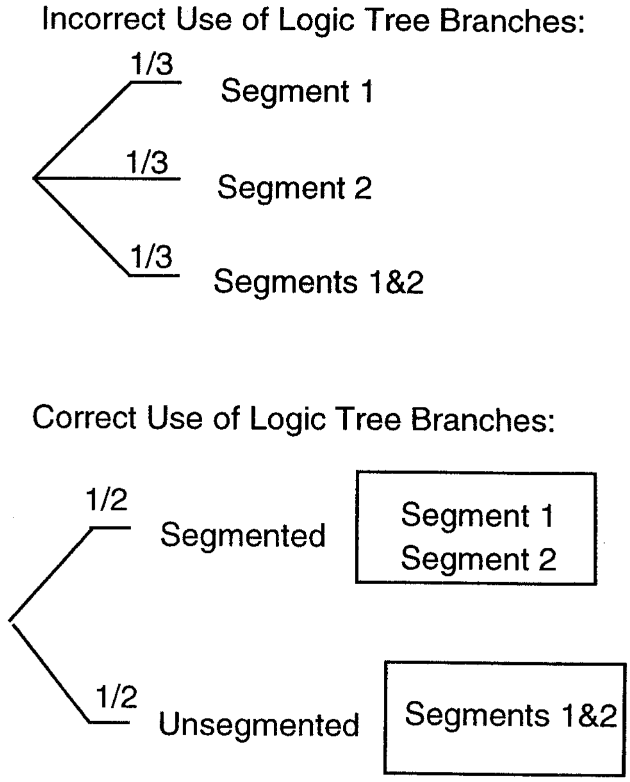

Figure 8. Example of misunderstanding of the branches in logic trees for fault segmentation.

15 The Branches On Logic Trees For Fault Segmentation Are Often Not Set Up Correctly.

The branches on logic trees should represent credible alternative models about the seismic source characterization and the ground motion attenuation. Although logic trees are usually only used in high end seismic hazard evaluations, often they are not set up correctly. A common problem with logic trees is that the alternative models are not complete. This is most common in the geologic input for the seismic source characterization. For example, consider the scientific uncertainty associated with fault segmentation. As a simple example, consider a fault with a postulated segmentation point. In many seismic hazard studies, branches in the logic tree are incorrectly set up for each potential rupture of the fault: segment 1 by itself, segment 2 by itself, and segments 1 and 2 together (Figure 8). The weights of the branches are then set to sum to unity. In this example, the weights of the segmented and unsegmented are intended to be equal.

The correct way to treat this uncertainty is to have a branch on whether the fault is segmented or not (Figure 8). One branch is that the fault is unsegmented; another branch is that the fault consists of two separate segments. Either branch in the logic tree needs to cover the entire fault length. In the second branch, the two segments are considered as separate sources. That is, the annual hazards from the two individual segments are added together in this branch. There should not be a branch of the logic tree in which just segment 1 ruptures by itself.

Incomplete branches on the logic tree will tend to underestimate the hazard. All branches should be checked to determine that they are complete.

ATTENUATION RELATIONS

Attenuation relations are probabilistic descriptions of the level of ground shaking as a function of the earthquake and site parameters. They are probabilistic in that they provide equations for the median ground motion and the standard deviation of the logarithm of the ground motion. The most common parameters that must be clearly defined in order to estimate ground motions using attenuation relations are: earthquake magnitude, type of faulting, site-to-source distance, and local site conditions (site classification). Moment magnitude is the preferred magnitude measure and has been adopted in nearly all recent attenuation relations. In contrast, a single consistent definition of the site-to-source distance (discussed earlier) and site classification has not been widely adopted by different authors of attenuation relations.

Source Type

Different tectonic regimes give rise to different ground motion attenuation relationships. Currently, three broad categories of tectonic regimes are typically used for attenuation relations used in seismic hazard assessments: shallow crustal earthquakes in active tectonic regions (e.g. California, Japan, New Zealand), shallow crustal earthquakes in stable continental regions (e.g. Australia, Eastern North America), and subduction zone earthquakes (e.g. Japan, Chile, Alaska, New Zealand). Many of the recent attenuation relations for these tectonic categories are given in a special issue of the Seismological Research Letters in February, 1997.

Most recent attenuation models distinguish between the ground motion from reverse and strike-slip earthquakes with the ground motion from reverse earthquakes being 20 to 40 percent larger than for strike-slip earthquakes. Due to the small number of normal faulting earthquakes in most strong motion data sets, the difference between ground motions for strike-slip and normal faulting earthquakes has not been included in most attenuation models. The standard practice is to use strike-slip attenuation relations to predict the ground motion from normal faults; however, recent evaluations of normal faulting earthquakes have found that the ground motions from normal faulting earthquakes are smaller than for strike-slip earthquakes.

Regionalization

As the number of recordings of strong ground motion increase, there has been a trend toward developing region-specific attenuation relations rather than just using the global average models developed for the broad tectonic categories. Often there is a tendency to overemphasize region specific data in developing region specific attenuation. Typically, there are not enough data in a specific region to completely determine the attenuation relation. In particular, usually there are not enough data close to the fault to constrain the behavior of the attenuation relation at short distances.

One way to address regionalization of attenuation relations is to only update parts of the global attenuation relations. For example, the simplest update is to estimate a constant scale factor to use to adjust a global attenuation model to a specific region. (This can reflect differences in the earthquake source or differences in the site categories.) If there are enough data over a range of distances, the slope of the attenuation could be updated while maintaining the magnitude scaling of the global model. An example of this type of regionalization of parts of the attenuation relation is the new attenuation relations developed for New Zealand (McVerry and Zhao, 1999). The peak acceleration attenuation relation from this region-specific model is compared to the global models in Figure 9.

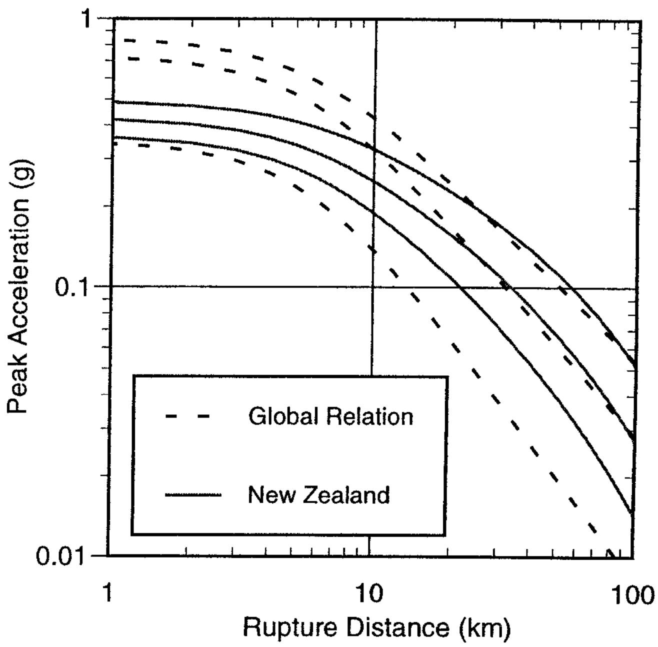

Figure 9. Comparision of regionalized attenuation relation for peak acceleration developed for New Zealand (strike-slip, weak rock) with the global model from Abrahamson and Silva (1997) for rock. The three curves in each set are for magnitude 5.5, 6.5, and 7.5.

Site Classifications

There are also several site classification schemes used in different attenuation relations, ranging from qualitative descriptions of the near-surface material to very quantitative definitions based on shear-wave velocity. Without consistent site classifications for the attenuation relations, it is often difficult to know how to apply the attenuation relations to a specific site. Site classifications can vary from country to country. The site classifications used in global models of the attenuation may not fit into the site classification system for a particular region. Site classifications based on the shear wave velocity, such as used in the IBC 2000, provide a basis for consistent site classifications around the world, but most existing attenuation relations do not use shear wave velocity for the site classification because the shear wave velocities are not widely available for strong motion station sites. Current efforts to collect detailed site information at strong motion sites should address this problem.

Even for "rock" site classifications, there is a range of what is classified as rock. In California, sites that are classified as "rock" often contain weathered rock and/or thin soil (< 20m thick). In practice, these "rock" ground motions are usually assumed to be outcrop rock (e.g. shear wave velocity of 1300 m/sec) for site response calculations. When used in a site response calculations, this outcrop motion should not include the effects of the weathered rock and thin soil. Idriss and Silva (1999) have evaluated suites of the recorded "rock" ground motions with measured shear-wave velocity profiles to determine the effect of the weathered rock and thin soil layers. They found that at high frequencies, the typical "rock" ground motions are 20 to 30 percent larger than true outcrop ground motions. It is common practice to ignore this difference between "rock" ground motions from attenuation relations and outcrop ground motions. As a result, the effects of the weathered rock/thin soil layers are often double counted in site response calculations.

All of the attenuation relations separate soft-soil sites from typical soil sites. Since the response of soft-soil sites is strongly site specific and there are few soft-soil strong motion recordings, response spectral attenuation relations have generally not been developed for soft-soil sites. This site condition is typically addressed with a site- specific site-response analysis.

Near Fault Directivity Effects

Existing attenuation relations used in practice do not explicitly include rupture directivity effects. Somerville et al. (1997) developed models of the fault rupture directivity effect that may be applied to any ground motion attenuation relationships for crustal earthquakes. The directivity model consists of two period dependent scale factors. One of the factors accounts for the change in shaking intensity in the average horizontal component of motion due to near-fault rupture directivity effects (higher ground motions for rupture toward the site and lower ground motions for rupture away from the site). The second factor reflects the directional nature of the shaking intensity using two ratios: fault normal (FN) and fault parallel (FP) versus the average (FA) component ratios. The two scaling factors depend on whether fault rupture is in the forward or backward direction, and also the length of fault rupturing toward the site.

There are several aspects of the empirical model for the average horizontal component scale factors developed by Somerville et al that needed to be modified to make the model applicable to a probabilistic hazard analysis. These modifications are given by Abrahamson (2000). The horizontal attenuation relations discussed earlier can be adjusted to account for near-fault directivity effects using these scale factors. Currently, models for the effects of directivity on the vertical component are not available. Directivity effects have been considered in some projects (important long period structures), but are generally not included in current seismic hazard analyses.

DETERMINISTIC SEISMIC HAZARD ANALYSIS

In a deterministic seismic hazard analysis, earthquake scenarios are evaluated separately. For each source, a scenario earthquake is defined (magnitude, distance, style-of-faulting and in some case rupture direction). The ground motion for the scenario earthquake is usually estimated using attenuation relations, but is sometimes estimated using seismological simulations of the ground motion.

Design Earthquakes

For faults, there are four attributes that need to be considered in developing a deterministic design earthquake: classification of active or inactive, the earthquake magnitude, the distance to the site, and the tectonic regime. These four attributes are discussed below.

Active vs Inactive

In a deterministic analysis, each source must be classified as either active or inactive. This classification can depend of the type of project. In some cases, only faults that have moved in the Holocene are considered active, whereas, in other cases, faults that have moved in the quaternary are considered active. Thus a fault may be considered active for critical structures but inactive for standard structures. The distinction between active and inactive can be unclear for low activity faults.

Scenario Earthquake Magnitude

The selection of the magnitude of the earthquake depends on the design approach that is being used. In some cases, two levels of earthquakes are considered. One is a upper level earthquake typically used for safety level evaluations and the other is a lower level earthquake typically used for functional level evaluations.

For fault sources, the magnitude of the upper level earthquake is for faults is usually estimated from the fault dimensions. The magnitude can be estimated from the fault length or from the fault area. In general, magnitude scaling relations based on the fault area are preferred over those based on the fault length. This is because the magnitude-area scaling relations are better constrained than the magnitude-length relations. Furthermore, since the seismic moment (given later in eq. 13) is proportional to the fault area scaling relations based on the fault length implicitly assume an average fault width. For crustal faults, fault width can vary greatly depending on the dip of the fault and the thickness of the crust. It is much better to estimate fault-specific fault widths than it is to assume an average fault width that may not apply to the specific fault in question. However, it remains common practice to estimate the magnitude from fault length.

Before fault segmentation concepts were developed, usually some fraction of the total fault dimension was used to estimate the magnitude of the design earthquake. For example, it was common to use 1/2 of the fault length. Fault segmentation studies have replaced this approach for well studied faults. Typically, the magnitude is estimated using the full dimension of the segment. The has brought up the issue of multiple segments rupturing at once. So selection of the earthquake scenario needs to consider how likely multi-segment ruptures are. Faults that have not been studied are often assumed to be unsegmented, however, this assumes knowledge about the fault. Assuming a fault is unsegmented is conservative for the deterministic approach, but it may be unconservative for the probabilistic approach (because it leads to longer recurrence intervals).

For lower level events, there is a wide range of approaches used to determine the magnitude. In some cases it is simply some level (e.g. 0.5 to 1.0 magnitude units) below the magnitude of the upper level earthquake. In other cases it may be based on the recurrence interval of earthquakes on each source. For example, the magnitude of the earthquake with a 100 year recurrence interval could be selected.

Scenario Earthquake Distance

For upper level events (e.g. MCE), it is common to assume that the fault ruptures the closest point on the fault to the site. This is a worst case distance, but it is reasonable because the large earthquake will likely rupture the fault closest to the site.

For areal source zones, the selection of the distance is more difficult. For the zone that contians the site, the worst case distance is zero because the earthquake could occur right under the site, but that would be very unlikely. One approach is to use the earthquake recurrence relation to determine an average distance for a given recurrence interval. This approach, sometimes called the semi-probabilistic approach, uses a probabilistic description of the occurrence of earthquakes to select deterministic earthquake scenarios. It requires selecting an appropriate recurrence interval. Alternatively, a common approach is to select the distance by judgement.

Tectonic Regime

The tectonic regime needs to be identified for each source so that the appropriate attenuation relation is used.

Ground Motion

To estimate the ground motion for the scenario earthquake, requires selecting the appropriate , site condition, ground motion model, number of standard deviations, and in some cases rupture directions.

Site Condition

Site effects can be considered in either a generic way or in a detailed site-specific analysis. If a generic approach is used, then the site category is identified and an appropriate attenuation relation for the site category is used. If a site-specific approach is used, then the a rock site attenuation relation is used to define the outcrop rock ground motion that is then used in a site response evaluation.

This difference between a gross site category and a site-specific site response applies to either a deterministic or probabilistic approach. However, to add to the confusion in the terminology in this field, there is a view by some geotechnical engineers that a "deterministic" approach ground motion is implies that a site-specific site response is conducted whereas a "probabilistic" approach implies that attenuation relations for broad site categories are used.

Ground Motion Model

The ground motion for the scenario earthquake can be computed from an appropriate attenuation relations or from seismological simulations of the source, wave propagation, and site response. Attenuation relations are much more commonly used than numerical simulations.

Number of Standard Deviations

In either case, the probability of the ground motion, given the scenario earthquake, needs to be specified (in terms of number of standard deviations above the median ground motion). In practice, usually 0 or 1 standard deviation (50th percentile or 84th percentile) is selected. This selection of the number of standard deviations usually considers the importance of the structure. Except for critical structures, the number of standard deviations to use is typically not specified in regulations; it is left up to the project to determine what to use.

Directivity

For important long period structures, directivity effects can be included using scale factors for the attenuation relation described earlier. To use these factors, the rupture direction needs to be selected. If directivity effects are included in a deterministic analysis, it is most common to assume rupture toward the site.

PROBABILISTIC SEISMIC HAZARD ANALYSIS

Modern probabilistic seismic hazard analyses follows the basic approach developed by Cornell (1968). The main changes from the original work is that the variability in ground motion (for a given magnitude and distance) is included. In addition, it has been generalized to fault (planar) sources. To handle planar sources, more source parameters are used (rupture dimensions, rupture location on the fault plane) and scientific (epistemic) uncertainty in the source characterization and ground motion attenuation is considered. The basic methodology involves computing how often a suite of specified levels of ground motion will be exceeded at the site. In this section, we describe the features of a PSHA study that represent the more experienced end of the state of practice.

The basic methodology involves computing how often a suite of specified levels of ground motion will be exceeded at the site. The hazard analysis computes the annual number of events that produce a ground motion parameter, A, that exceeds a specified level, z. This number of events per year, ν, is also called the "annual frequency of exceedance". The inverse of ν is called the "return period".

The calculation of the annual frequency of exceedance, ν, involves several probability distributions for each seismic source: the frequency of occurrence of earthquakes of various magnitudes, the rupture dimension and location of the earthquakes, and the attenuation of the ground motion from the earthquake rupture to the site. The occurrence rates of the earthquakes of various magnitudes are determined by the magnitude recurrence relations. The location of the earthquake depends on the geometry of the seismic source. The distance from the rupture to the site is computed from the earthquake location and rupture dimension. The ground motion at the site is determined from the attenuation relation.

For areal source zones, the earthquakes are usually assumed to be point sources. The rate at which the ground motion from the i^{th} source exceeds the test level z at the site, is given by

$$\nu_i(A > z) = N_i(M_{\min}) \int_{r=0}^{\infty} \int_{m=M_{\min}}^{M_{\max_i}} \int_{\varepsilon=\varepsilon_{\min}}^{\varepsilon_{\max}} f_{m_i}(m) f_{r_i}(r) f_\varepsilon(\varepsilon) P(A > z \mid m, r, \varepsilon)\, dr\, dm\, d\varepsilon \qquad (3)$$

where N_i (M_{min}) is the rate of earthquakes with magnitude greater than M_{min} from the i^{th} source, r is the distance measure, m is magnitude, ε is the number of standard deviations of the ground motion from the median ground motion, M_{max_i} is the maximum magnitude (for the i^{th} source), $f_m(m)$, and $f_r(r)$ are probability density functions for the magnitude, and distance which describe the relative likelihood of different earthquake scenarios, $f_\varepsilon(\varepsilon)$ is the probability density function for a ground motion variability, and $P(A>z|m,r,\varepsilon)$ is the probability that the ground motion exceeds the test level z for magnitude m, distance r, and number of standard deviations ε. Since the magnitude, distance and number of standard deviations fully describes the ground motion, $P(A>z|m,r,\varepsilon)$ is either 0 or 1. The formulation given in eq. (1) is different from the standard way of writing the hazard in that it explicitly denotes the variability of the ground motion. This was done to emphasize the role of the variability in the attenuation relation and to show that ground motion variability is treated in the same way as variability in the magnitude and distance of the earthquake scenario.

For point source models, the density function for closest distance, $f_r(r)$, is straightforward to compute; typically the hypocenters are uniformly distributed over the seismic source. For planar sources (e.g. known faults), we need to consider the finite dimension and location of the rupture in order to compute the closest distance. Specifically, we need to include the randomness in the rupture length, rupture width, rupture location along strike, and rupture location down dip. (Since the rupture width is often limited by the down-dip fault width, variability in the fault width is correlated to variability in the fault length. A common approach to the correlation is to compute the rupture area and rupture width and then back calculate the rupture length.) For planar sources eq. (3) becomes

$$\nu_i(A > z) = N_i(M_{\min}) \int_{W=0}^{\infty} \int_{RA=0}^{\infty} \int_{Ex=0}^{1} \int_{Ey}^{1} \int_{m=M_{\min}}^{M_{\max_i}} \int_{\varepsilon=\varepsilon_{\min}}^{\varepsilon_{\max}} f_{m_i}(m) f_{W_i}(m,W) f_{RA_i}(m,RA) f_{Ex_i}(x) \\ f_{Ey_i}(m,x) f_\varepsilon(\varepsilon) P(A > z \mid m, r_i(x,y,RA,W), \varepsilon)\, dW\, dRA\, dx\, dy\, dm\, d\varepsilon \qquad (4)$$

where $f_W(m,W)$, $f_{RA}(m,RA)$, $f_{Ex}(x)$, $f_{Ey}(y)$ are probability density functions for the rupture width, rupture area, location of the center of rupture along strike and location of the center of rupture down dip, respectively. In eq. (4), x and y give the location of the center of the rupture in terms of the fraction of the fault length and fault width, respectively (e.g. x=0 is one end of the fault and x=1.0 is the other end of the fault).

The hazard integral in eq. (4) appears complicated, but most of the computations are related to finding the distribution of the closest distance. Keep in mind that all that eq. (2) is doing is defining a complete set of possible earthquake scenarios (magnitude and distance combinations) with the full range of possible ground motions and keeping track of which scenarios lead to ground motions that exceed the test value z.

For multiple seismic sources, the total annual rate of events with ground motions that exceed z at the site is just the sum of the annual rate of events from the individual sources (assuming that the sources are independent).

$$\nu(A > z) = \sum_{i=1}^{Nsource} \nu_i(A > z) \qquad (5)$$

Where N_{source} is the total number of fault and areal sources. The rates for the individual sources, vi, generally represent the long term behavior of the source and do not include short term variations in the rate due to the time since the last large earthquake (e.g. periodicity effects) or fault interaction effects. These effects are discussed later in the time dependent hazard section.

Poisson Assumption

To convert the annual rate of events to a probability, we consider the probability that the ground motion exceeds test level z at least once during a specified time interval. A standard assumption is that the occurrence of earthquakes is a Poisson process. That is, there is no memory of past earthquakes, so the chance of an earthquake occurring in a given year does not depend on how long it has been since the last earthquake. If the occurrence of earthquakes is a Poisson process then the occurrence of peak ground motions is also a Poisson process. For a Poisson process, the probability of at least one occurrence of ground motion level z in t years is given by

$$P(A>z,t) = 1 - \exp(-\nu(A>z)t) \qquad (6)$$

For t=1 year, this probability is the annual hazard.

Time Dependent Hazard

While the most common assumption is that the occurrence of earthquakes is a Poisson process, an alternative model that is commonly used is the renewal model. In the renewal model, the occurrence of large earthquakes is assumed to have some periodicity. The conditional probability that an earthquake occurs in the next ΔT years given that it has not occurred in the last T years is given by

$$P(T,\Delta T) = \frac{\int_{T}^{T+\Delta T} f(t)\,dt}{\int_{T}^{\infty} f(t)\,dt}$$

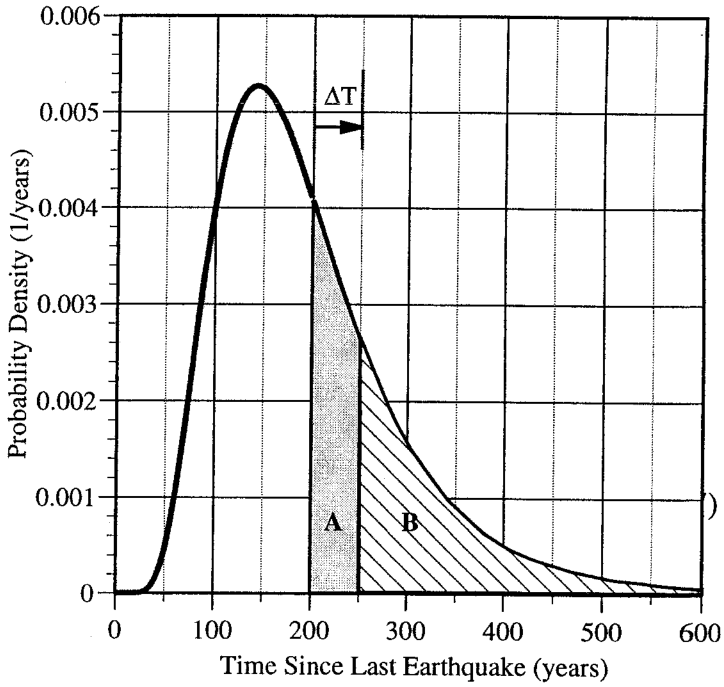

Figure 10. Computation of the conditional probability for the rewnel model.

where f(t) is the probability density function for the earthquake recurrence intervals. Several difference forms of the distribution of earthquake recurrence intervals have been used: normal, log-normal, Weibull, and Gamma. In engineering practice, the most commonly used distribution is the log-normal distribution. Although lognormal distributions are usually parameterized by the median and standard deviation, in renewal models, the usual approach is to parameterize the distribution by the mean and the coefficient of variation. For a log normal distribution, the relation between the mean and the median was given earlier by eq. (2).

This conditional probability is shown graphically in Figure 10. The probability distribution of earthquake recurrence intervals in this example is based on a lognormal distribution with a mean recurrence interval of 200 years and a coefficient of variation (cov) of 0.5. The conditional probability is computed for an exposure time of 50 years assuming that it has been 200 years since the last earthquake. Graphically, the conditional probability is given by the ratio of the area labeled "A" to the sum of the areas labeled "A" and "B". That is, P(T=200,ΔT=50)=A/(A+B).

An important parameter in the renewal model is the cov. The cov is a measure of the periodicity of the earthquake recurrence intervals. A small cov (e.g. cov < 0.3) indicates that the earthquakes are very periodic, whereas a large cov (e.g. cov>1) indicates that the earthquakes are not periodic. Early estimates of the cov found small covs of about 0.2 (e.g. Nishenko, 1982). More recent estimates of the cov are much larger, with cov values ranging from 0.3 to 0.7. In practice, the typical cov used in seismic hazard analysis between 0.4 and 0.6. The sensitivity of the conditional probability to the cov is shown in Figures 11 and 12 for a 50 and 5 year exposure periods, respectively. For comparison, the Poisson rate is also shown. These figures shows that the renewal model leads to higher probabilities once the elapse time since the last earthquake is greater than about one-half of the mean recurrence interval. In addition, these figures show that as the cov becomes larger, the conditional probability becomes closer to the Poisson probability.

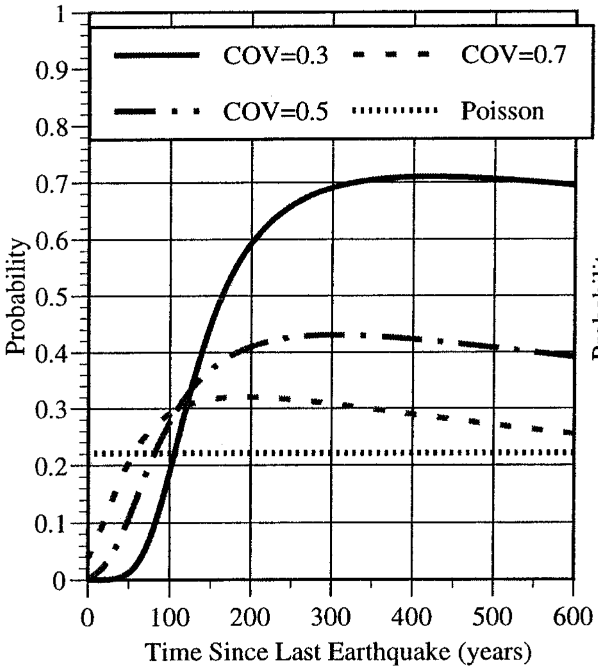

Figure 11. Sensitivity of the conditional probability for the renewala model for a 50 year exposure period.

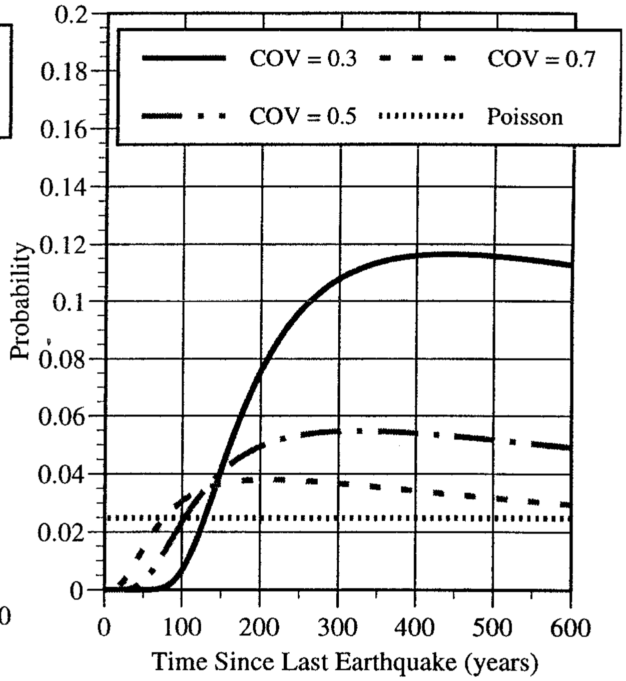

Figure 12. Sensitivity of the conditional probability for the renewala model for a 5 year exposure period.

Logic Trees

Scientific (epistemic) uncertainty is considered by using alternative models and/or parameter values for the probability density functions in eq. (4), the attenuation relation, and the activity rate. For each alternative model, the hazard is recomputed resulting in a suite of alternative hazard curves. Scientific uncertainty is typically handled using a logic tree approach for specifying the alternative models for the density functions, attenuation relations, and activity rates.

Branches on logic trees represent either-or branches. The branches represent alternative credible models. The weights on the branch represent the judgement about the credibility of the alternative models. These weights are often called probabilities, but they are better treated as weights. Branches in logic trees do not represent "sometimes" branches (e.g. randomness). For example, if a fault sometimes ruptures in individual segments and sometimes ruptures as multiple segments, then this variability is randomness that should be part of the probability density function in the hazard integral. The branches on the logic tree should reflect alternative estimates of the parameters and models included in the hazard integral.

To help keep track of the difference between what goes on the logic tree and what goes into the probability density functions, the terms "aleatory" variability and "epistemic" uncertainty are used. Aleatory variability is the randomness part and epistemic uncertainty is the uncertainty part. The reason for using aleatory and epistemic rather than "randomness" and "uncertainty" is that randomness and uncertainty are too common of terms that are often used interchangeably. Using the terms aleatory and epistemic leads to a more consistent use of terminology.

Source Characterization

In the 1970s and early 1980s, the seismic source characterization was typically based on historical seismicity data using seismic zones (called areal sources). In many parts of the world, this is still the standard of practice. In regions with geologic information on the faults (slip-rates or recurrence intervals) the geologic information is used to compute the activity rate. The slip-rate is converted to an earthquake activity rate by requiring the long-term seismic moment-rate on the fault to be in equilibrium. A key issue in this conversion is the distribution of sizes of earthquakes that release the seismic energy. For example, the seismic energy could be released in many moderate magnitude earthquakes or in a few large magnitude earthquakes. The relative rate of moderate to large magnitude earthquakes is described by a magnitude density function.

Together, the magnitude distribution and the activity rate are used to define the magnitude recurrence relation which describes how often each magnitude earthquake is expected to occur. This is computed by integrating the magnitude density function from right to left:

$$N(M)=N(M_{\min})\int_{m=M_{\max}}^{M} f_m(m)dm$$

Magnitude Distribution

The magnitude distribution, given by $f_m(m)$, describes the relative number of large magnitude and moderate magnitude events that occur on the seismic source. Two alternative magnitude density functions are usually considered in seismic hazard analyses: the truncated exponential model and the characteristic model.

The truncated exponential model is the standard Gutenberg-Richter model that is truncated at the minimum and maximum magnitudes and renormalized so that it integrates to unity. The density function for the truncated exponential (TE) model is given by

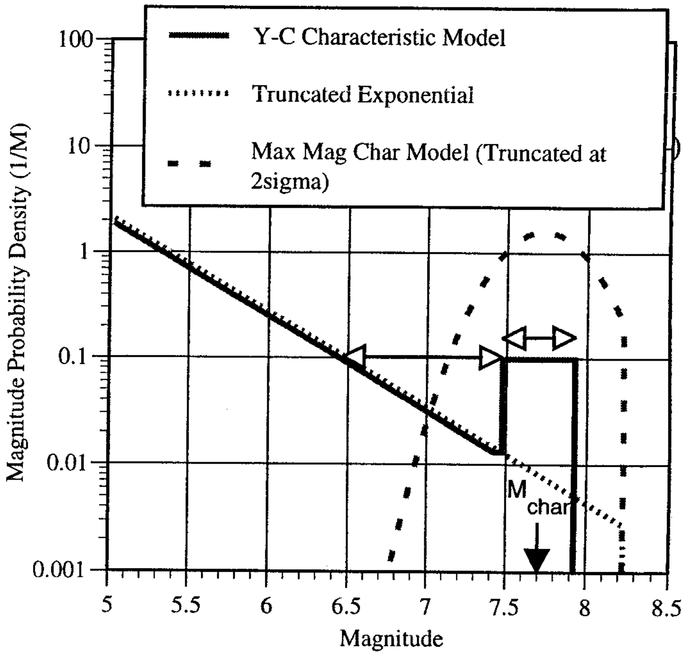

Figure 15. Comparison of commonly used magnitude probability density functions.

$$f_m^{TE}(m)=\frac{\beta\exp(-\beta(m-M_{\min})}{1-\beta\exp(-\beta(M_{\max}-M_{\min}))} \qquad (9)$$

where β is ln(10) times the b-value. Source-specific estimates of the b-value are usually used with this model. An example of the truncated exponential distribution is shown in Figure 15.

The characteristic model assumes that more of the seismic energy is released in large magnitude events than in the truncated exponential model. That is, there are fewer small magnitude events for every large magnitude event for the characteristic model than for the truncated exponential model. There are several alternative forms for the characteristic model. Two commonly used models are the Youngs and Coppersmith (1985) characteristic model and the "maximum magnitude" characteristic model.

The maximum magnitude characteristic model assumes that all of the seismic energy is released in characteristic earthquakes. The simplest form of this model uses a single magnitude for the characteristic earthquakes (CDMG, 1996). A more general form is a truncated normal distribution which allows a range of magnitudes for the characteristic earthquake consistent with the variability in the empirical data. Here the distribution is truncated at 2 standard deviations above the mean characteristic magnitude. The resulting density function for the truncated normal (TN) model is given by:

$$f_m^{TN}(m) = \frac{1}{0.977}\frac{1}{\sigma\sqrt{2\pi}}\exp\{\frac{(m-M_{char})^2}{2\sigma_m^2}\} \qquad for\ m \le \overline{M}_{char} + 2\sigma_m^2 \tag{10}$$

where M_{char} is the mean magnitude of the characteristic earthquake. This is also shown in Figure 15.

When these models are used to balance moment-rate on the fault (see below), then using a single magnitude vs using a symmetric distribution leads to different mean recurrence intervals of the characteristic earthquake due to the exponential scaling of seismic energy release as a function of earthquake magnitude. The truncated normal distribution better represents the randomness in the magnitude of the earthquake for a given rupture dimension (e.g. variability in the static stress-drop of the earthquake) than does the single value (delta function model).

The second commonly used characteristic earthquake model is the Youngs and Coppersmith (1985) model. In this model, the fault generates moderate magnitude earthquakes as well as characteristic earthquakes. The magnitude density function for this model is shown in Figure 15. This model has a boxcar distribution for the characteristic part and an exponential distribution for the moderate size earthquakes. The width of the boxcar is Δm_2 magnitude units. The relative number of earthquakes in these two parts are defined by setting the height of the boxcar to be equal to the value of the exponential distribution at Δm_1 magnitude units below the lower magnitude of the characteristic earthquake boxcar function.

The equations for the magnitude density function for the generalized form of the Youngs and Coppersmith characteristic model are given by

$$f_m^{YC}(m) = \begin{cases} \dfrac{1}{1+c}\,\dfrac{\beta\exp(-\beta(\overline{M}_{char}-M_{\min}-\Delta m_1-0.5\Delta m_2)}{1-\exp(-\beta(\overline{M}_{char}-M_{\min}-0.5\Delta m_2)} & for\ \overline{M}_{char}-\dfrac{\Delta m_2}{2} < M \le \overline{M}_{char}+\dfrac{\Delta m_2}{2} \\ \dfrac{1}{1+c}\,\dfrac{\beta\exp(-\beta(M-M_{\min})}{1-\exp(-\beta(\overline{M}_{char}-M_{\min}-0.5\Delta m_2)} & for\ M_{\min} \le M \le \overline{M}_{char}-\dfrac{\Delta m_2}{2} \end{cases} \tag{11}$$

where

$$c = \frac{\beta\exp(-\beta(\overline{M}_{char}-M_{\min}-\Delta m_1-0.5\Delta m_2)}{1-\exp(-\beta(M_{char}-M_{\min}-0.5\Delta m_2)} \tag{12}$$

In the Youngs and Coppersmith model, Δm_1=1.0 and Δm_2=0.5. This results in about 6% of the total seismic moment-rate being put into the exponential tail and 94% of the moment-rate being put into characteristic earthquakes.

Activity Rate

There are two common approaches used for estimating the fault activity rate: historical seismicity and geologic (and geodetic) information. If historical seismicity catalogs are used to estimate the activity rate, then the estimate of $N(M_{min})$ is usually based on fitting the truncated exponential model to the historical data. Maximum likelihood procedures are generally preferred over least-squares for estimating the activity rate and the b-value. When using geologic information on slip-rates of faults, the activity rate is computed by balancing the energy build-up estimated from geologic evidence with the total energy release of earthquakes. Knowing the dimension of the fault, the slip-rate, and the rigidity of the fault, we can balance the long term seismic moment can be set so that the fault is in equilibrium. (e.g. Youngs and Coppersmith, 1985).

The seismic energy release is balanced by requiring the build up of seismic moment to be equal to the release of seismic moment in earthquakes. The build up of seismic moment is computed from the long term slip-rate. The seismic moment, Mo (in dyne-cm), is given by

$$M_0 = \mu A D \tag{13}$$

where μ is the rigidity of the crust, A is the area of the fault (in cm^2), and D is the average displacement (slip) on the fault surface (in cm). The annual rate of build up of seismic moment is given by the time derivative of eq. (13):

$$\frac{dM_o}{dt} = \mu A S \tag{14}$$

where S is the slip-rate in cm/year. The seismic moment released during an earthquake is given by

$$\log 10\ M_0 = 1.5\ M + 16.05 \tag{15}$$

where M is the moment magnitude of the earthquake. To balance the moment build up and the moment release, the annual moment rate from the slip-rate is set equal to the sum of the moment released in all of the earthquakes that are expected to occur each year.

$$\mu AS = N(M_{\min}) \int_{m=M_{\min}}^{M_{\max}} f_m(m) 10^{(1.5m+16.05)}\, dm \tag{16}$$

Given the shear modulus (a typical value is 3×10^{11} dyne/cm^2), slip-rate, fault area, and magnitude density function, the activity rate, $N(M_{min})$ is given by

$$N(M_{\min}) = \frac{\mu AS}{\int_{m=M_{\min}}^{M_{\max}} f_m(m) 10^{(1.5m+16.05)}\, dm} \tag{17}$$

Magnitude Recurrence Relations

Although the density functions for the truncated exponential and Y&C characteristic models are similar at small magnitudes (Figure 15), if the geologic moment-rate is used to set the annual rate of events, $N(M_{min})$, then there is a large impact on the computed activity rate depending on the selection of the magnitude density function. Figure 16 shows the comparison of the magnitude recurrence relations for the alternative magnitude density functions when they are constrained to have the same total moment rate. The characteristic model has many fewer moderate magnitude events than the truncated exponential model (about a factor of 5 difference). The maximum magnitude model does not include moderate magnitude earthquakes. With this model, moderate magnitude earthquakes are generally considered using areal source zones.

The large difference in the recurrence rates of moderate magnitude earthquakes between the Y+C and truncated exponential models can be used to test the models against observations for some faults. The truncated exponential model significantly overestimates the number of moderate magnitude earthquakes. This discrepancy can be removed by increasing the maximum magnitude for the exponential model by about 1 magnitude unit. While this approach will satisfy the both the observed rates of moderate magnitude earthquakes and the geologically determined moment rate, it generally leads to unrealistically large maximum magnitudes for known fault segments (e.g. about 6 standard deviations above the mean from a magnitude-area scaling relation) or it requires combining segments of different faults into one huge rupture. Although the truncated exponential model does not work well for faults in which the geologic moment-rate is used to define the earthquake activity rate, in practice it is usually still included as a viable model in a logic tree because of it wide use in the past. Including the truncated exponential model is generally conservative for high frequency ground motion (f>5 hz) and unconservative for long period ground motions (T>2 seconds).

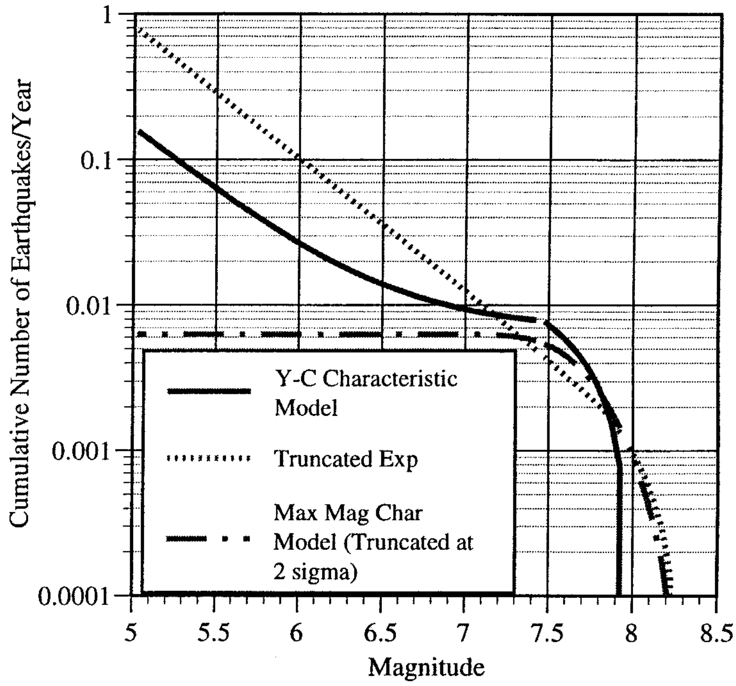

Figure 16. Comparison of commonly used magnitude recurrence relations with balanced moment-rate.

Rupture Dimension Distributions

Scaling relations between the rupture dimension and the magnitude are used in two different ways in probabilistic seismic hazard analyses. First, the rupture dimension is used to define the characteristic earthquake (and maximum magnitude). In this step, the magnitude is estimated from the fault dimension. In the hazard integral (eq. 4), a model is also needed to define the rupture area for a given magnitude. In this step, the rupture area is estimated from the magnitude.

Empirical relations between rupture area and magnitude are typically used; however, the regression results are not the same for a regression of magnitude given and area versus a regression of area given magnitude. In most hazard evaluations, different models are used for estimating M(A) versus A(M). The difference between these models due simply to the regression is shown in Figure 17. This figure show the commonly used relations developed by Wells and Coppersmith (1985) for M(A) and for A(M). These two models are similar, but they differ at larger magnitudes. While the application of these different models is consistent with the statistical derivation of the models, there is a problem of inconsistency when both models are used in the same source characterization.

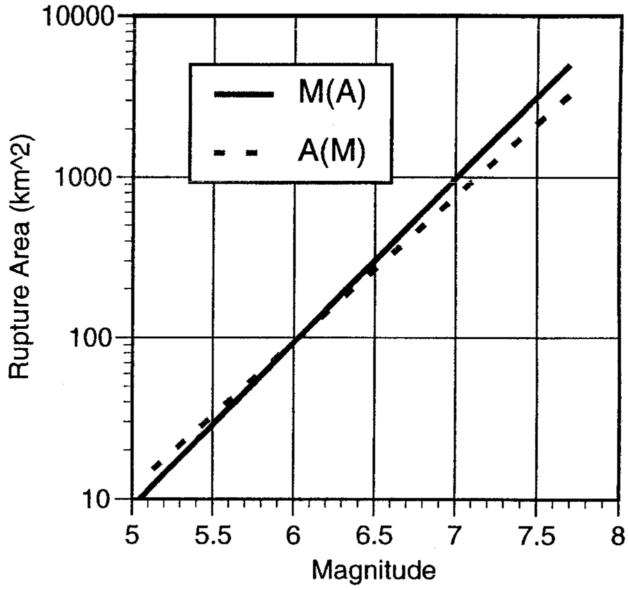

Figure 17. Comparison of empirical magnitude-area scaling relations. The solid line is for estimating mgnitude from the area and the dashed curve is for estimating the rupture area from the magnitude.

As an alternative, empirical models can be derived with a constraint that earthquakes have a constant stress-drop. This constraints the slope to be unity and leads to models that can be applied in either direction. The empirically derived slopes are close to unity, implying that such a constraint is consistent with the observations. Regional variations in the average stress-drop of earthquakes can be accommodated by different a constant in the scaling relation.

The hazard integral given in eq. (4) also requires a distribution for rupture width as a function of magnitude. Empirical models for this parameter are also available (e.g. Wells and Coppersmith, 1985). These models need to be truncated on fault specific basis to reflect variations in the down-dip fault width. For example, if the seismogenic crust has a thickness of 15 km and the fault has a dip of 45 degrees, then the maximum width is 21 km; however, if a fault in this same region has a dip of 90 degrees, then the maximum width is 15 km.

Rupture Location Density Functions

The hazard integral in eq. (4) also requires a model for the distribution of the ruptures on the fault plane (both along-strike and down-dip). For the rupture locations, a uniform distribution over the fault plane is commonly used. The resulting density functions for $f_{Ex}(x)$ and $f_{Ey}(y)$ are unity. In some applications, empirical distributions of the depths of earthquakes are used to define $f_{Ey}(y)$.

Equal Hazard Spectra

A common method for developing design spectra based on the probabilistic approach is equal hazard spectra (also called uniform hazard spectra). An equal hazard spectrum is developed by first computing the hazard at a suite of spectral periods using response spectral attenuation relations. That is, the hazard is computed independently for each spectral period. For a selected return period, the ground motion for each spectral period is measured from the hazard curves. These ground motions are then plotted at their respective spectral periods to form the equal hazard spectrum. This process is shown graphically in Figure 18.

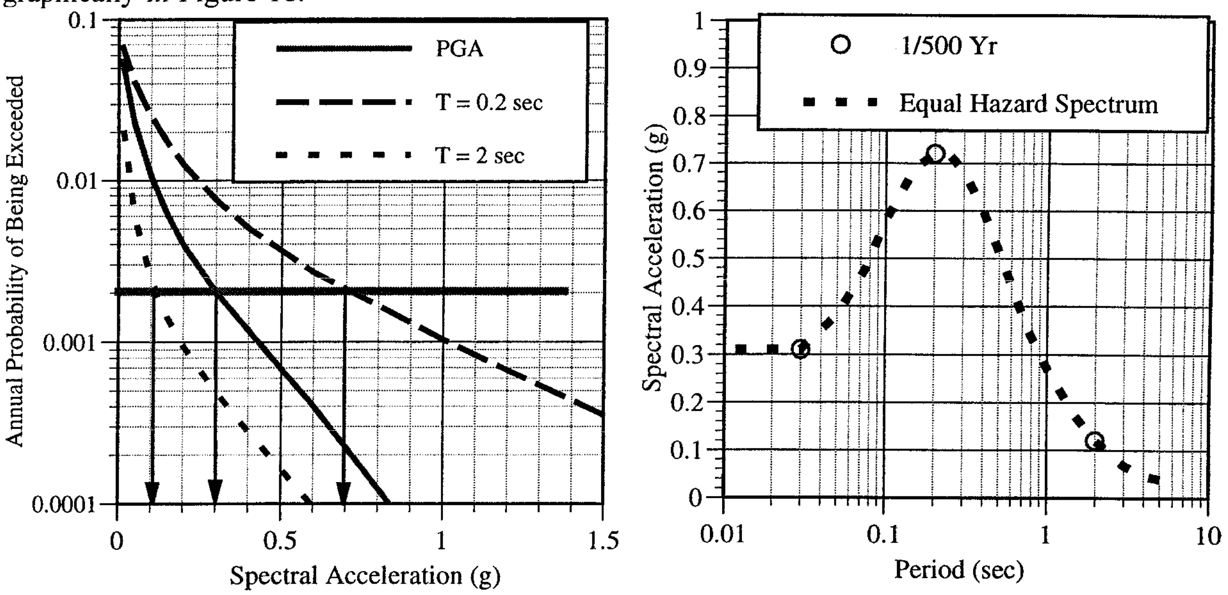

Figure 18. Procedure for developing equal hazard spectral. In this example, a return period of 500 years is used.

The term "equal hazard spectrum" because there is an equal probability of exceeding the ground motion at any period. Since the hazard is computed independently for each spectral period, in general, an equal hazard spectrum does not represent the spectrum of any single earthquake. It is common to find that the high frequency (f>5 Hz) ground motions are controlled by nearby moderate magnitude earthquakes, whereas, the long period (T>1 sec) ground motions are controlled by distant large magnitude earthquakes.

Deaggregation of Hazard

The hazard curve gives the combined effect of all magnitudes and distances on the probability of exceeding a given ground motion level. Since all of the sources, magnitudes, and distances are mixed together, it is difficult to get an intuitive understanding of what is controlling the hazard from the hazard curve by itself. A common practice is to break the hazard back down into its contributions from different magnitude and distance pairs to provide insight into what events are the most important for the hazard (at a given ground motion level). This process is called deaggregation (Bazzurro and Cornell, 1999).

With deaggregation, the fractional contribution of different subsets of the events to the total hazard is computed. The most common form of deaggregation is a two-dimensional deaggregation in magnitude and distance bins. The deaggregation by source allows the dominant seismic source to be identified. The deaggregation by magnitude and distance bins allows the dominant scenario earthquake (magnitude and distance pair) to be identified. The results of the deaggregation will be different for different probability levels (e.g. 100 yr vs 1000 yr return periods) and for different spectral periods.

Example Hazard Calculation

An example, the hazard is computed for the a site located 5 km from a low activity fault and 50 km from a large high activity fault (Figure 19). The hazard for peak acceleration and T=1 second spectral acceleration are shown in Figure 20 for a rock site condition. At short return periods (<200 years), the high activity, distant fault dominates that hazard at both PGA and T=1 second Sa. At long return periods (e.g. 10,000 years), the low activity local fault and the background zone dominate the PGA hazard, but the local fault and more distant fault contribute about equally to the long period hazard.

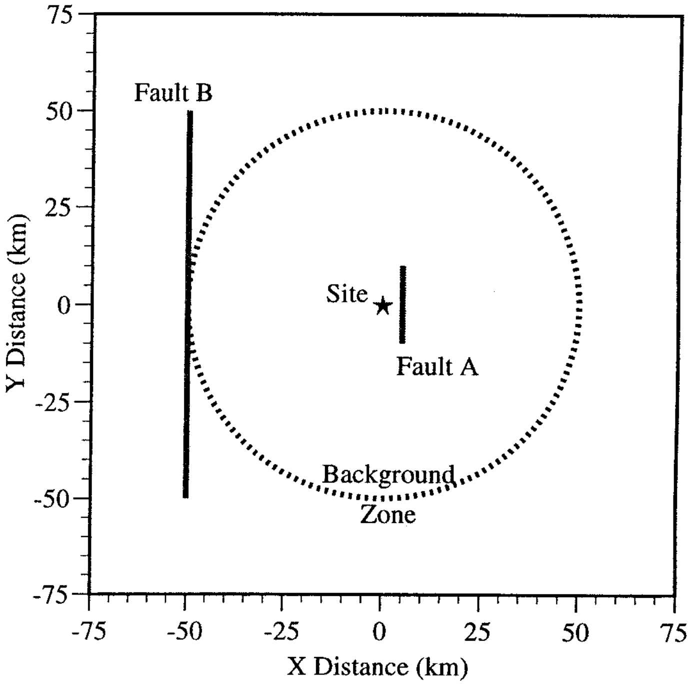

Figure 19. Source geometry used for the example.

The scientific (epistemic) uncertainty in the hazard curveis displayed in terms of fractiles from the suites of logic trees that result from the logic trees for the source characterization and ground motion attenuation. Figure 21 shows the uncertainty in the T=1 second spectral acceleration hazard. In addition to the mean hazard, the 5th , 50th, and 95th fractiles are shown. The range of probability shown in this example ranges from a factor of 2 to a factor of 10. This is a typical range of uncertainty for hazard curves. Note that the ground motion uncertainty for a given return period is a much smaller factor (about 1.5 for the 5 to 95 percentile range). Note that the median is in lower than the mean hazard. This is also a typical result. In practice, it is most common to use the mean hazard rather than the median hazard for selecting design events; however, higher fractiles could also be used depending on the importance of the project.

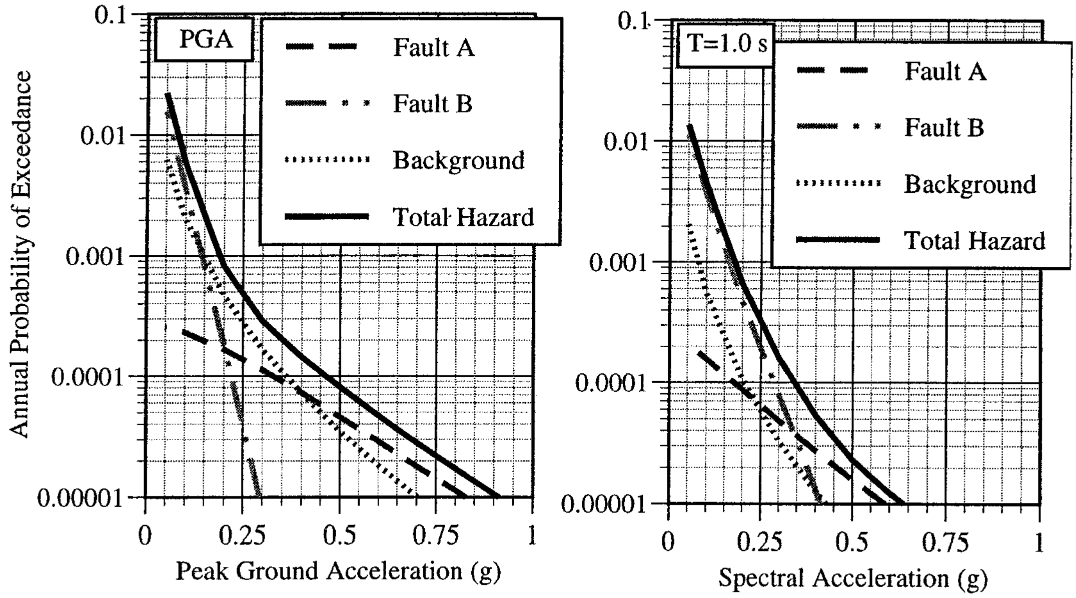

Figure 20. Mean hazard for PGA and T=1 sec spectral acceleration.

The results of the deaggregatization of the hazard are shown in Figure 22 and 23 at return periods of 500 years and 10,000 years for T=1 sec spectral acceleration. For the shorter return period, the deaggegation shows a dominant large magnitude, large distance event from fault B. At the long return period, both faults have a significant contribution to the hazard Typically the mode of the distribution is used rather than the mean. Using the mode ensures that the scenario is physically realizable. The difficulty with using the mode is that it can be sensitive to the bin size (Bazzurro and Cornell, 1999). Typically, equal size bins are used, but without much thought as to what the optimal bin size should be.

For bimodal distributions, such as the 10,000 year deaggregation, the best approach is to use multiple design events. The ground motion level is the same for the multiple events, but the characteristics of the time history (e.g. duration) would be different for the different events. In this example, time histories developed for this return period should consider both the local moderate magnitude event and the distant large magnitude event.

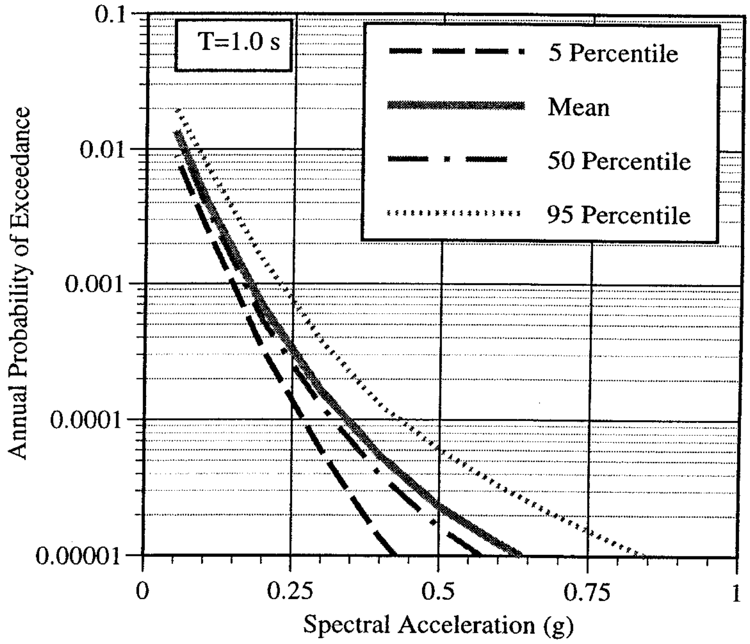

Figure 21. Example of the scientific (epistemci) uncertainty in the hazard as computed from the logic trees.

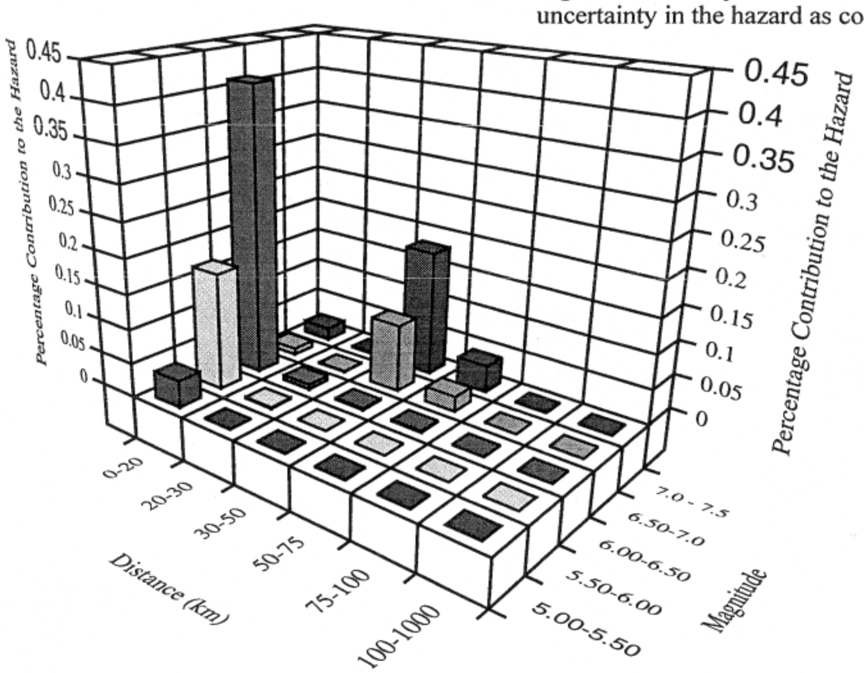

Figure 22. Deaggregation of the T=1 second spectral acceleration for a return period of 10,000 years.

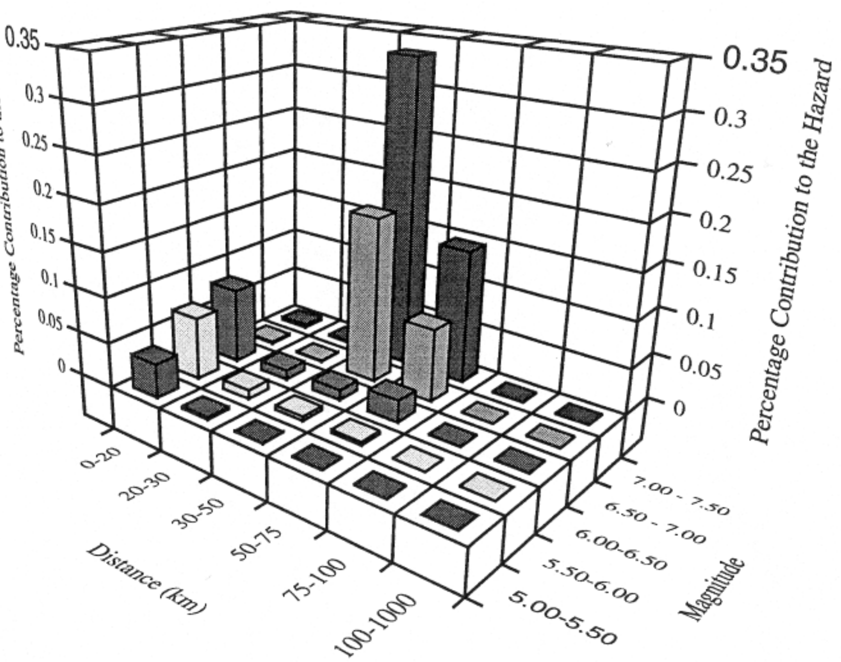

Figure 23. Deaggregation of T=1 sec spectral acceleration for a 500 year return period

COMPARISON OF THE DETERMINSTIC AND PROBABILISTIC APPROACHES

The deterministic and probabilistic approaches provide complementary information. The deterministic approach is based on a small subset of the large number scenarios considered in the probabilistic approach. Some of the advantages and disadvantages of the deterministic and probabilistic approaches are listed below.

Advantages And Disadvantages Of The Deterministic Approach

The main advantages of the deterministic approach are:

- There are simple rules for application
- It is easy to check the design earthquake and computed ground motion
- Does not require explicitly setting acceptable level of safety

The main disadvantages of the deterministic approach are:

- It is often mistakenly thought to lead to worst case ground motions
- It leads to a wide range of the probability of the design ground motion occurring.
- It can be sensitive to the active/not active determination for nearby faults

Advantages and Disadvantages of the Probabilistic Approach

The main advantages of the probabilistic approach are:

- It provides a consistent hazard level at all sites
- It considers the activity rate of faults (high activity faults are recognized to be more hazardous than low activity faults).
- It considers the total hazard from all faults (being close to many active faults is recognized as being more hazardous than being close to only one fault)
- It is not sensitive to active/inactive determinations for nearby faults
- It requires a clear quantitative description of what is "reasonable"

The main disadvantages of the probabilistic approach are:

- It is not well understood by the engineering community
- It is more difficult to check the computations
- It requires explicitly selecting a probability level to use

Discussion

A key issue in using the probabilistic approach is selecting the appropriate hazard level to use in defining the design ground motions (e.g. annual probability of 1/500, 1/1000, 1/2000, 1/10000, etc). That is, to use the probabilistic approach requires specifying the acceptable level of safety. This can be a difficult decision because it involves public policy; however, once the hazard level is selected, the probabilistic method provides an objective method for selecting appropriate design events. The probabilistic hazard can be deaggregated to determine the appropriate earthquake scenarios to use in a deterministic analysis (as shown above). This is important for defining the characteristics of the ground motion in terms of the non-stationary character of the time history (duration, pulse shapes, etc).

As discussed above, in most cases, a deterministic approach needs to back off from the worst case ground motion. The question is how far do we back off from the maximum? A rational approach is to reduce the ground motion below the maximum until there is a large enough chance that the ground motion will occur so that is should be considered in design. In other words, the appropriate deterministic ground motion should consider the probability of the ground motion occurring. In practice, this is not done explicitly. For example, using the MCE, the probability of the ground motion is not explicitly addressed but it sometimes considered implicitly in selecting the number of standard deviations to use for the ground motion (typically, the median or 84th percentile).

Since the deterministic ground motions are not worst case ground motions, there will be some significant non-zero probability of exceeding the MCE ground motion. The deterministic approach includes some risk, but it is not quantified. This unquantified hazard of exceeding the MCE ground motions can vary greatly from region to region. Using an MCE approach with a fixed number of standard deviations, this probability of exceeding the MCE ground motion can have a very wide range for different sites. For example, in California, the return period of the median peak acceleration for the MCE (from faults) varies from as little as 50 years to as much as 20,000 years in different parts of the state (Figure 24)

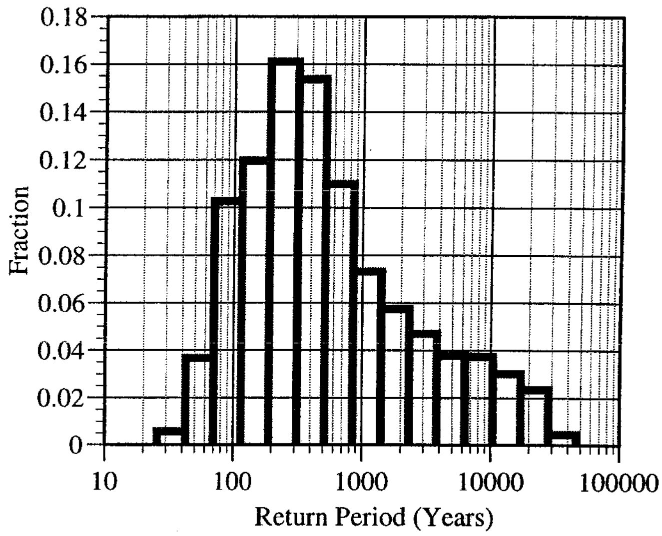

Figure 24. Distribution of the return period of the median ground motion for the MCE in California.

SUMMARY

The topic of this paper was the state-of-the-practice of seismic hazard evaluations. The variability in seismic hazard evaluations is so wide that there is no state-of-the-practice. This wide variability is in part due to inconsistent use of terminology, misleading terminology, and widespread misunderstandings of the basic methodologies used in seismic hazard evaluations. One source of the problems leading to the large variability in practice is the lack of well written, easy to understand papers or textbooks on the topic of seismic hazard analysis. In this paper we have highlighted some of the inconsistencies and misinterpretations that are common in seismic hazard evaluations and have presented what we believe the state-of-the-practice should strive to become.

REFERENCES

Abrahamson, N. A. and Shedlock, K. M. (1997). Overview, Seismological Research Letters, v. 68, No. 1, p. 9-23.

Abrahamson, N. A. and W. Silva, (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, Seismological Research Letters, v. 68, No. 1, p. 94-127.

Abrahamson, N. A. (2000). Effects of rupture directivity on probabilistic seismic hazard analysis, Proc. Sixth Int. Conf. On Seismic Zoneation, November 12-15, 2000, Pam Springs.

Bazzurro, P. and Cornell, C. A. (1999). Disaggregation of seismic hazard, Bulletin of the Seismological Society of America, v. 89., p. 335-341.

Cornell, C. A., 1968, Engineering seismic risk analysis, Bulletin of the Seismological Society of America, v. 58, p. 1583-1606.

California Division of Mines and Geology, 1996, Probabilistic seismic hazard assessment for the state of California, DMG Open-File Report 96-08, 33p.

Idriss, I. M. and Silva, W. (1999). Modification of Recorded Ground Motions to Rock Outcrop Ground Motion. Consultant's report to Pacific Gas and Electric Company.

McVerry, G. H. and J. X. Zhao (1999). A response spectrum model for New Zealand including high attenuation in the volcanic zone, Report prepared by the Institute of Geological and Nuclear Sciences Ltd.

Sadigh, K., Chang, C.-Y., Egan, J. A., Makdisi, F., and Youngs R. R. (1997). Attenuation relationships for shallow crustal earthquakes based on California strong motion data, Seismological Research Letters, v. 68, p. 180-189.

Somerville, P. G., Smith, N. F., Graves, R. W. and Abrahamson N. A. (1997). Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity, Seismological Research Letters, v. 68, p. 199-222.

Youngs, R. R. and Coppersmith, K. (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates, Bulletin of the Seismological Society of America, v. 75, p. 939-964.