**First European Conference on Earthquake Engineering and Seismology**
(a joint event of the 13th ECEE & 30th General Assembly of the ESC)
**Geneva, Switzerland, 3-8 September 2006**
**Paper Number: keynote**

# SEISMIC HAZARD ASSESSMENT: PROBLEMS WITH CURRENT PRACTICE AND FUTURE DEVELOPMENTS

**Norman ABRAHAMSON[1]**

## SUMMARY

This paper addresses issues with current practice in seismic hazard analysis and gives some recommendations for improvements. A review of probabilistic and deterministic approaches is given. The remainder of the paper is focused on PSHA. There continue to be hazard studies that do not properly treat the ground motion variability, leading to systematic underestimations of the hazard. Next, some of the short-comings in current PSHA practice that can be improved without major revisions to the overall methodology are discussed. These include selecting the bin size for the deaggregation, use of Uniform Hazard Spectra, estimation of scenario spectra, estimation of epistemic uncertainty, degree of spatial smoothing of seismicity, and the use of a strict lower bound magnitude. Finally, the future directions of PHSA are addressed. These include incorporating site-specific site response into PSHA and taking project specific structure response into account in conducting PHSA.

## 1. INTRODUCTION

Any seismic hazard analysis involves the consideration of at least four basic parameters: the earthquake magnitude, the distance from the earthquake to the site, the site condition, and the ground motion for the given magnitude, distance and site condition. While the site condition can be determined with measurements, the earthquake magnitude, distance, and ground motion can only be characterized by probability distributions with randomness (aleatory variability). The distance is a computed parameter that depends on the earthquake location and rupture dimension,which in turn have randomness.

### 1.1 Aleatory Variability and Epistemic Uncertainty

Aleatory variability and epistemic uncertainty are terms used in seismic hazard analysis that are not commonly used in other fields, but the concepts are well known. Aleatory variability is the natural randomness in a process. It is a result of our simplified modeling of a complex process and is parameterized by probability density functions (pdf). Epistemic uncertainty is the scientific uncertainty in the simplified model of the process and is characterized by alternative models (e.g. alternative pdfs).

The terms randomness and uncertainty have also been used for aleatory variability and epistemic uncertainty, respectively; however, these terms are commonly used in generic ways. As a result, they are often mixed up when discussing hazard analysis. The terms “aleatory variability” and “epistemic uncertainty” provide an unambiguous terminology.

[1] Pacific Gas & Electric Company, 245 Market Street, San Francisco, CA , 94105, USA
Email : naa2@pge.com

## 1.2 Deterministic vs Probabilistic Approaches

Deterministic and probabilistic approaches have both been used in practice for over 30 years, but there remains widespread misunderstandings of the two methods by engineering and earth science professionals practicing in earthquake engineering. This section provides a brief review of these two approaches.

In the deterministic approach, individual earthquake scenarios (earthquake magnitude and location) are developed for each relevant seismic source and a specified ground motion probability level is selected (by tradition, it is usually either 0 or 1 standard deviation above the median). Based on the earthquake location, the distance to the site is computed. Given the magnitude, distance, and number of standard deviations for the ground motion, the ground motion is then computed for each earthquake scenario using a ground motion model (attenuation relation) that is based on either empirical ground motions or numerical simulations of ground motions. The largest ground motion from any of the considered scenarios is used for the design ground motion. The approach is "deterministic" in that single values of the parameters (magnitude, distance, and number of standard deviations for the ground motion) are selected for each scenario.

In the probabilistic approach, all possible and relevant deterministic earthquake scenarios (all possible magnitude and location combinations) are considered as well as all possible ground motion probability levels (a range of the number of standard deviations above or below the median). The set of scenarios should only include physically possible earthquakes. The scenarios from a deterministic analysis are all included in the full set of scenarios from the probabilistic analysis. For each earthquake scenario, the ground motions are computed for each possible value of the number of standard deviations above or below the median ground motion.

Up to this point, the probabilistic analysis is just a large number of deterministic analyses. Given this large suite of deterministic ground motions, which one do we select for design? One approach would be to select the largest ground motion from any of the scenarios. That is, select the scenario with the worst-case ground motion. The problem with that approach is that the worst-case ground motion will usually be very large with a large cost impact on design, and they are so rare that the high cost is not justified. If there is not a large impact on cost or they are not too rare, then the worst-case ground motion may be appropriate for design. In some low seismic areas, the cost may not be too high, but in most projects, this will not be the case.

I believe that much of the confusion regarding probabilistic seismic hazard analyses comes down to the difference between the occurrence of earthquake scenarios (magnitude and location) and the occurrence of earthquake shaking at a site. We design structures for a ground motion at a site, not for the earthquake magnitude and distance. This may seem circular, in that the ground motion at the site is computed from the earthquake scenario, but a key element is the variability of the ground motion for a given earthquake scenario. This is an important distinction that is often lost. The largest ground motions are controlled by the number of standard deviations above the median ground motion. As noted above, the deterministic approach traditionally uses at most 1 standard deviation above the median for the ground motion, but in the probabilistic approach, larger values of the number of standard deviations above the median ground motion are considered. As a result, the worst-case ground motions will be much larger then the 84th percentile deterministic ground motions.

I am not aware of any projects advocating the use of worst-case ground motions for design with the possible exception of the Yucca Mtn Project. Therefore, whether a deterministic or probabilistic approach is used, there will be a finite chance of the design ground motions being exceeded. That is, we accept that there is a non-zero chance of failure of the structure. In a deterministic approach, the chance of exceeding the design ground motion is not addressed directly. The hope is that using a deterministic approach, the chance of failure of a structure will be small enough. In a probabilistic approach, we address this chance of failure by estimating the chance of exceeding the design ground motion and in some cases by estimating the probability of failure of the structure using probabilistic risk analysis.

Both deterministic and probabilistic approaches result in probabilistic statements about the design ground motion. In the deterministic approach, the ground motion has a probability of being exceeded given that the scenario earthquake has occurred. In the probabilistic approach, the ground motion has a specified probability of being exceeded in a given time period. In either case, the ground motions are, in general, not worst-case ground motions. The selection of an appropriate ground motion that is lower than the worst-case ground motion comes down to defining what is "reasonable". The main purpose of a PSHA is to provide a method for selecting the deterministic scenarios that are "reasonable" from the large suite of possible scenarios (magnitude, distance, and number of standard deviations).

In summary, a deterministic analysis is just a simplified PSHA in which a limited set of candidate scenarios are considered. If we use a consistent definition of what is “reasonable”, then we will get consistent results from deterministic and probabilistic hazard studies. The two approaches lead to different results if they have inconsistent definitions of what is “reasonable”.

In a deterministic approach, we may consider that the median or 84th percentile ground motion is reasonable because we are using a "conservative" earthquake magnitude at a short distance. One common view is that using a larger number of standard deviation would be "piling conservatism on top of conservatism". The basis for this statement is that the selected earthquake is already rare so the ground motion should not also have to be rare (e.g. a large number of standard deviations). To address this view, consider two faults in California: the San Jacinto fault and the Bear Mountain fault. The Imperial segment of the San Jacinto fault generates a magnitude 6.5 or larger earthquake about every 80 years. The Bear Mountain fault generates a magnitude 6.5 or larger earthquake about every 10,000 years. Given that the Bear Mountain fault generates large earthquakes so infrequently, using a median ground motion may be reasonable since the combined annual chance of the earthquake occurring and the ground motion being greater than the median is 1/20,000. In contrast, for the Imperial segment, the 84th percentile ground motion has a 1/480 annual chance of being exceeded. The selection of a "reasonable" number of standard deviations of the ground motion must consider how often the earthquakes occur. In some cases, ground motions above the 84th percentile deterministic ground motion may be reasonable.

## 2. ASPECTS OF POOR PRACTICE

The most common continuing problems with PSHA practice are related to the incorrect treatment of the ground motion variability. As describe above, there are three main random (aleatory) variables that must be considered in a PSHA: earthquake magnitude, site to source distance (e.g. earthquake location), and the number of standard deviations of the ground motion (epsilon). The early development of PSHA did not account for ground motion variability [e.g. Cornell, 1968], but this short-coming was correctly quickly [Esteva, 1970]. Even though the need to consider the variability in all three of these parameters is well established, there continue to be PSHA studies that exclude or put tight limits on the ground motion variability. This practice leads to a systematic underestimation of the hazard {Bommer and Abrahamson, 2006].

### 2.1 Excluding Ground Motion Variability

One of the common aspect of poor practice is the exclusion of the ground motion variability from the hazard calculation. The usual justification given for this approach is that the hazard analyst wants to use the “best estimate” ground motion and not introduce conservatism. What is misunderstood is that the ground motion models give a distribution of the ground motion for a given magnitude and distance rather than a single value. The width of the distribution is an integral part of the ground motion model and cannot be ignored. Accounting for the ground motion variability is not a conservatism that can be included or excluded based on engineering judgment. A PSHA that excludes the ground motion variability is simply wrong. If the design ground motions resulting from a PSHA with the proper treatment of the ground motion variability are too large to use, then the proper response is to either modify the design or accept more risk (e.g. design for a shorter return period). Using an incorrect PSHA calculation to reduce the design values is not the answer.

### 2.2 Severe Truncation of the Ground Motion Variability

A second aspect of poor practice problem is the use of a severe truncation on the ground motion variability. The ground motion variability is generally modelled by a lognormal distribution and a common practice in PSHA is to truncate the log normal distribution at some maximum number of standard deviations. Typically, the lognormal distribution is truncated at between 2 and 3 standard deviations above the median, but in some projects, it is truncated at just 1 standard deviation. Severe truncation (at 1 standard deviation above the median) is not appropriate and leads to an understimation of the hazard.

The basis for truncating the ground motion distribution is that it is thought that the lognormal distribution may not be an accurate model for the upper tail of the distribution. When the empirical ground motion models were based on data sets with a few hundred recordings, there was not enough data to test this part of the distribution. For example, to test the lognormal distribution at 3 standard deviations requires 1000s of recordings. With modern strong ground motion data sets, there are 1000s of data available at the lower levels of shaking. Analyses of the residuals from modern attenuation relations has shown that the ground motion variability is consistent with a lognormal distribution up to at least 3 standard deviations [EPRI, 2006a]. This study concluded that there is no statistical basis for truncating the ground motion distribution below 3 standard deviations. There could still be a truncation of the ground motion distribution based on physical limits of the material underlying the site (e.g. at some maximum ground motion value) such as used in the PEGASOS project [Abrahamson et al, 2002] and the recent updates to the Yucca Mtn Project.

## 3. ASPECTS OF CURRENT PRACTICE THAT NEED IMPROVEMENT

In this section, I describe some of the short-comings in current PSHA practice that can be improved without major revisions to the overall methodology. These include selecting the bin size for the deaggregation, use of Uniform Hazard Spectra, estimation of scenario spectra, estimation of epistemic uncertainty, degree of spatial smoothing of seismicity, and the use of a strict lower bound magnitude.

### 3.1 Deaggregation Bins

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 considered together, it is difficult to understand of what is controlling the hazard from the hazard curve by itself. To provide insight into what events are the most important for the hazard, the hazard at a given ground motion level is broken down into its contributions from different earthquake scenarios. This process is called deaggregation [e.g. Bazzurro and Cornell, 1999]. In a deaggregation, the fractional contribution of different scenario groups to the total hazard is computed. The most common form is a two-dimensional deaggregation in magnitude and distance bins. This approach allows the dominant scenario earthquakes (magnitude and distance pair) to be identified.

As an example, consider the seismic sources shown in Figure 1. This example has two nearby faults with low activity rates, a distant fault with a high activity rate, and a background zone. A typical deaggregation is shown in Figure 2. This is based on equal spacing of the magnitude and distance bins. The deaggregation shows that the hazard is dominated by both the nearby faults (M6.5-7.0, 0-10 km) and distant faults (M7.5-8.0, 110-120 km). The dominant scenario is usually determined by either the mean or the mode of the deaggregation.

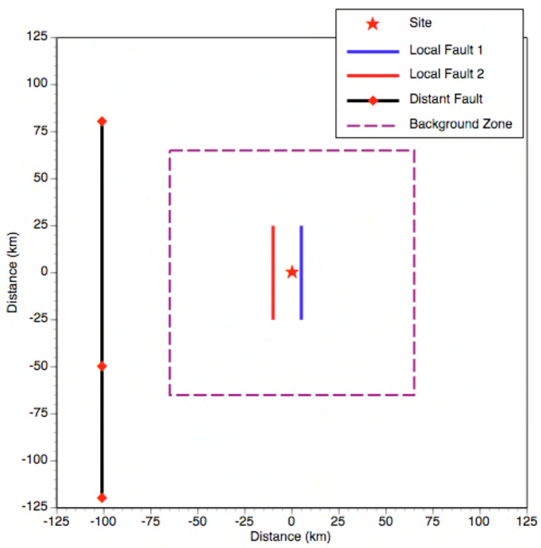


**Figure 1: Sources used in the hazard example. The diamonds on the distant fault represent segmentation points.**

The mean has advantages in that it is defined unambiguously and does not depend of the bin size. The disadvantage is that it may correspond to a scenario that is not realistic if the there are two or more sources with significant contribution to the hazard. In the example shown in Figure 2, the mean magnitude is 7.1 and the mean distance is 50 km. This scenario does not correspond to any of the fault sources.

The mode is the most likely scenario group. It corresponds to the scenario group that has the largest deaggregation value. The mode has the advantage it will always correspond to a realistic source. The disadvantage is that the mode depends on the grouping of the scenarios, so it is not robust. In the example, the mode from Figure 2 corresponds to the local faulting (M6.5-7.0, 0-10 km), but the mode changes if the bin sizes are changed. For example, if we used finer distance bins at short distances and broader distance bins at larger distances, then the mode would change to the large distant earthquake (Figure 3). While the two deaggregations show the same result in that they both show two scenarios dominate the hazard, the best single dominate event (as often used in national hazard maps) can be sensitive to the deaggregation bin size.

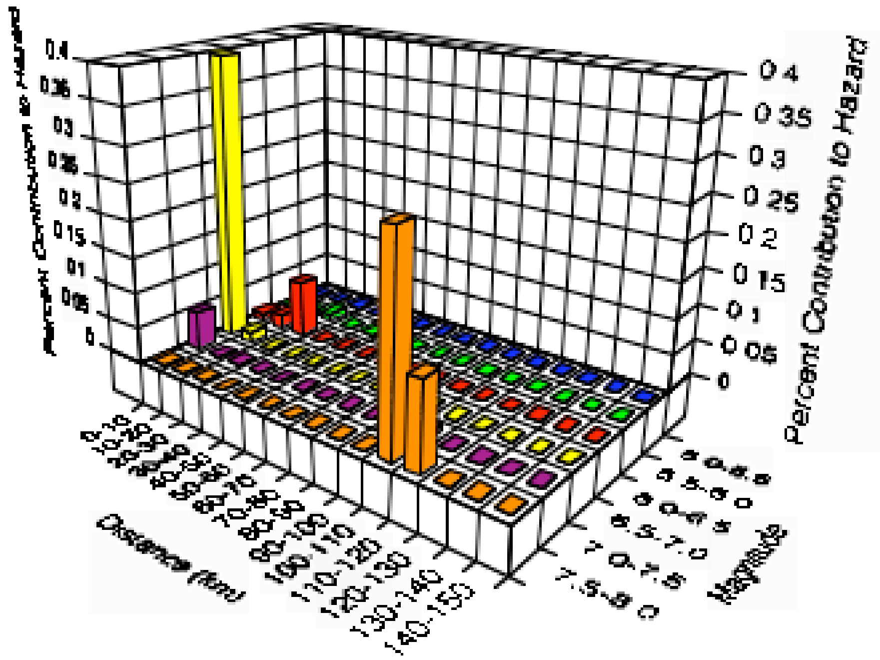


**Figure 2: Example deaggregation for T=1 sec, and a return period of 500 years using equal bins in distance.**

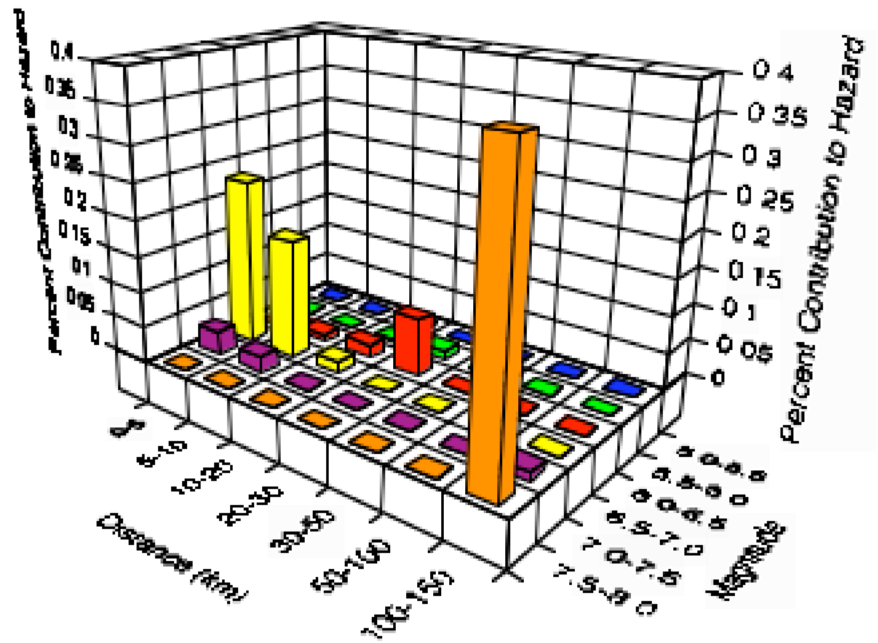


**Figure 3: Example deaggregation for T=1 sec, and a return period of 500 years using increasing distance bin size.**

In most hazard studies, little thought is given to the selection of the magnitude and distance bins. The appropriate bin size depends on how the results are to be used. One common use of deaggregation is to define scenarios for selecting time histories to be used in more detailed analysis of structures. In this case, the appropriate bin size depends on how the characteristics of the time histories (e.g. duration, near fault pulses, spectral shape) change with magnitude and distance. For example, if the differences in the time history characteristics at distances of 100 km and 110 km are less than the differences at distances of 10 to 20 km, then wider distance bins should be used at larger distances.

The selection of appropriate bin sizes for deaggregation requires feedback on the sensitivity of the structural response to the magnitude and distances. With this information, the bin size could be selected so that the each bin size corresponds to the same change in the expected response. This implies that the selection of the optimal deaggregation bins will be project-specific. For example, the optimal deaggregation bins for liquefaction analysis may not be the optimal deaggregation bins for evaluation of a building.

An alternative approach is use fine deaggregation bins, identify the controlling scenarios, associate the bins with the identified scenarios, and then compute the mean magnitude and mean distance for each controlling scenario. This approach is robust in that it will not be sensitive to the bin size but it can get complicated if the scenarios overlap.

The key issue here is that if the mode of the deaggregation is used to define the scenario, as is standard practice, then the results are sensitive to the bin size. There is not a single optimum bin size for all projects, so the hazard analyst needs to consider how the deaggregation is to be used by the project in the selection of the bin sizes.

### 3.2 Uniform Hazard Spectra

A common method for developing design spectra based on the probabilistic approach is to use the uniform hazard spectrum (UHS). The UHS is developed by first computing the hazard at a suite of spectral periods. Next, for a given return period, the ground motion is measured from the hazard curves for each spectral period. Finally, these ground motions are then plotted at their respective spectral periods. The term "uniform hazard spectrum" is used 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, a uniform 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. The UHS is the only step in a PSHA that combines the ground motions from different earthquakes.

The "mixing" of earthquakes in the UHS is often cited as a disadvantage of PSHA. There is nothing in the PSHA method that requires using a UHS. Based on the deaggregation, multiple spectra (for each important source) can be developed. The reason for using a UHS rather than using multiple spectra for the individual scenarios is to reduce the number of engineering analyses required. A deterministic analysis has the same issue. If one deterministic scenario leads to the largest spectral values for long spectral periods and a different deterministic scenario leads to the largest spectral values for short spectral periods, it is common practice to develop a single design spectrum that envelopes the two deterministic spectra. In this case, the deterministic design spectrum also does not represent a single earthquake.

The choice of using a UHS rather than multiple spectra for the different scenarios should be the decision of the engineering analyst, not the hazard analyst. The project engineers need to decide if it is worth the additional analysis costs of using multiple spectra to avoid exciting a broad period range in a single evaluation.

In practice, the hazard analysts often only provide the UHS in their hazard report. Without the scenario spectra also given in the hazard report, the project engineer often incorrectly assumes that they must use the UHS. To improve practice, hazard report should always include a comparison of the UHS with the spectra from the individual representative events identified by the deaggregation. This gives the engineer the information needed to make a decision whether to evaluate multiple scenarios one at a time or to envelope the spectra from the multiple scenarios to reduce the number of analyses required.

In the nuclear industry in the U.S., the standard practice is to break the UHS down into at least two controlling events: one for the high frequency (5-10 Hz), and one for the low frequencies (1-2.5 Hz). In this approach, the controlling earthquake for each frequency range is determined from the deaggregation and the median spectral shape (epsilon=0) for each controlling earthquakes are scaled up to the UHS value over the appropriate frequency range. This gives the scenario spectra that are used to develop design time histories.

### 3.3 Spectra for Controlling Scenarios

A drawback with using the median spectral shapes to define the spectra for the controlling scenarios is that the resulting spectrum may not be representative of the spectrum that is expected if the scenario earthquake occurs and the spectral value at the specified spectral period equals the UHS spectrum at that period.

For longer return periods, UHS will correspond to ground motions greater than the median ground motion (e.g. epsilon > 0). Recorded ground motions with positive epsilons at a particular period will tend to have peaks in the spectra at that period. (Part of the reason that the recorded ground motion had a positive epsilon was because it had a peak at the spectral period of interest). The spectral values at other spectral periods will, on average, not also have as strong of a peak. Using the median spectral shape corresponds to assuming that the other spectral values at all of the periods are also peaks.

Rather than using a median spectral shape for this event, the expected spectral shape given the epsilon at the long periods can be used [Baker and Cornel, 2006]. Based on evaluations of the residuals from attenuation relations, the expected epsilon at spectral period T can be computed given the epsilon at some reference spectral period $T_o$.

$$\varepsilon(T) = \varepsilon(T_o)c \qquad (1)$$

Examples of the values of coefficient c computed from the Abrahamson and Silva [2006] NGA relation using strong motion data for M>6.5 for rock sites are listed in Table 1 for reference periods of 0.2, 1.0, and 2.0 sec. To apply this model, the value of epsilon at the reference spectral period is needed. The reference epsilon is found by finding the epsilon needed to scale the ground motion from the controlling scenarios to the UHS. The expected spectrum is given by

$$Sa_{Expected}(T) = Sa_{med}(T) \exp(\varepsilon(T)\sigma(T)) \qquad (2)$$

**Table 1: Slope of the Relation Between Epsilons for the Expected Spectral Shape**

| Period (Sec) | $T_o$=0.2 | $T_o$=1.0 | $T_o$=2.0 |
|---|---|---|---|
| 0.0 | 0.91 | 0.68 | 0.43 |
| 0.075 | 0.91 | 0.54 | 0.31 |
| 0.1 | 0.91 | 0.50 | 0.27 |
| 0.2 | 1.00 | 0.48 | 0.26 |
| 0.3 | 0.93 | 0.63 | 0.39 |
| 0.4 | 0.84 | 0.71 | 0.45 |
| 0.5 | 0.71 | 0.77 | 0.52 |
| 0.75 | 0.62 | 0.92 | 0.66 |
| 1.0 | 0.45 | 1.00 | 0.76 |
| 1.5 | 0.37 | 0.87 | 0.85 |
| 2.0 | 0.26 | 0.81 | 1.00 |
| 3.0 | 0.24 | 0.77 | 0.94 |

Figures 4 and 5 shows examples of the T=2 second scenario spectra for return periods of 500 years and 10,000 years, respectively for distant fault from Figure 1 (M7.7, R=100 km). This scenario only applies to the long period part of the UHS. For the 500 year return period, the reference epsilon is 0.83 and for the 10,000 return period, the reference epsilon is 2.5. The expected spectrum is compared to the spectrum based on the median spectral shape. As the epsilon values become larger (long return period) the differences between the expected spectral shape and the median shape become larger. This occurs because it becomes less likely that the spectrum at the periods other than T=2 sec will have as strong of a peak as at T=2 sec as the reference epsilon value increases. This indicates that if we get an earthquake from the distant fault with the very high T=2 sec content, it will likely have relatively low spectral content at other frequencies. That is, if the T=2 sec spectral acceleration for an earthquake is unusually large (e.g. an outlier) than it is not likely to also be an outlier at the short spectral periods. This affects the response of the structure if the structure has more than one significant model (for buildings) or is sensitive to both the short and long periods as for slope stability.

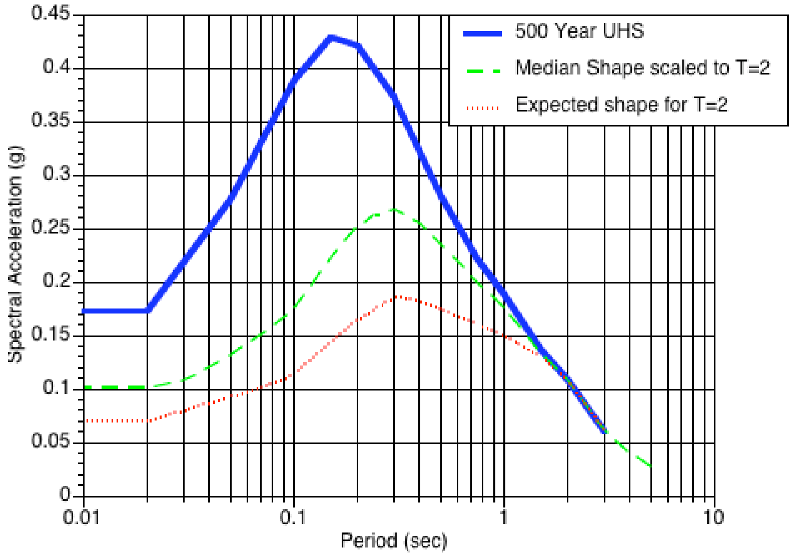


**Figure 4. Comparison of the median spectral shape scaled to the UHS at T=2 sec with the expected shape for the 500 year return period for T=2 sec.**

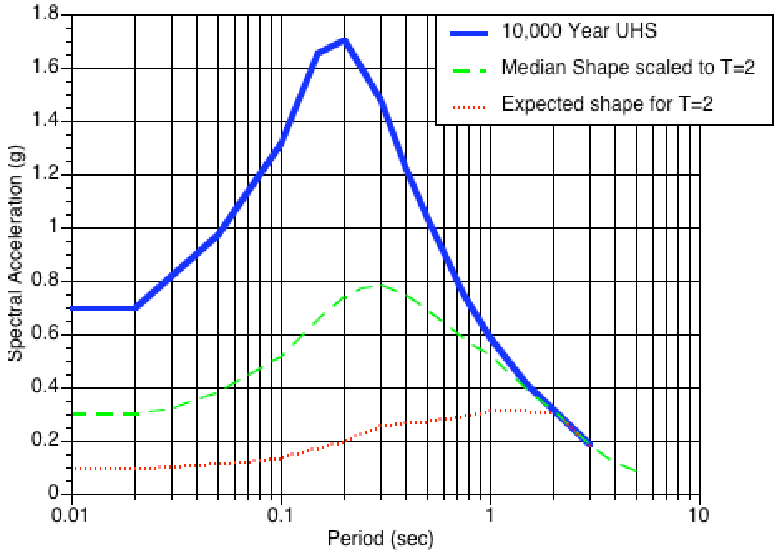


**Figure 5: Comparison of the median spectral shape scaled to the UHS at T=2 sec with the expected shape for the 10,000 year return period for T=2 sec.**

### 1.4 Underestimation of Epistemic Uncertainty

Epistemic uncertainty is considered by using alternative models and/or parameter values for the source characterization and ground motion attenuation relation. For each combination of alternative models, the hazard is recomputed resulting in a suite of alternative hazard curves. In seismic hazard analyses, it is common to use logic trees to handle the epistemic uncertainty.

A logic tree consists of a series of branches that describe the alternative models and/or parameter values. At each branch, there is a set of branch tips that represent the alternative credible models or parameter. The weights on the branch tips represent the judgment about the relative credibility (or merit) of the alternative models [Abrahamson and Bommer, 2005]. The branch tip weights typically sum to unity at each branch point. Only epistemic uncertainty should be on the logic tree. A common error in seismic hazard analyses is to put aleatory variability on some of the branches.

Prior to the use of logic tress, the approach was to develop the single best model. In controversial projects, there would be disagreement between the project sponsors, regulators, and interveners as to which model was the best model. Logic trees were used to allow multiple models to be considered with weights that reflected the degree of belief of the scientific community (or at least the seismic hazard analyst) in the alternative models. In this way, all proposed models that were credible could be considered without having to select a single best model.

As the use of logic trees has increased, the logic trees are commonly interpreted to represent the scientific uncertainty in the source characterization and ground motion attenuation; however, in practice, logic trees represent the range of available alternative models. In many cases, the range of available models will not cover the scientific uncertainty. In developing the epistemic uncertainty, the guiding concept should be that less data means larger uncertainty. This seems like a simple concept, yet it is often not followed.

Consider two faults, one that has had many studies and another that has had only a single study. In current practice, the logic tree will typically consider only available models (sometime this is further restricted to models published in referred journals). So for the well-studied fault, there will be several alternative models available, but for the poorly studied fault there will be only a single model available. By considering only the one available model for the poorly studied fault, 100% weight is given to that model, implying that there is no uncertainty in the model. In contrast, for the well-studied fault, there will be several alternative models available over which the weights will be spread. As a result, the computed epistemic uncertainty will be larger for the fault with more data. Another consequence is that as additional research is conducted, additional models will be developed, leading to more branches on the logic tree and larger uncertainty in the computed hazard. Over time this is what has happened. Our estimates of the epistemic uncertainty have increased, not decreased, as additional data have been collected and new models developed.

Clearly, using only available models will underestimate the epistemic uncertainty for sources and regions with little or no data, but how can we develop uncertainty estimates with no data? Our guiding concept is that less data means larger uncertainty. Regions with more data and more models can be used as a lower bound of the uncertainty for regions with little or no data. This does not require that we apply complex models to regions with no constraints on the inputs, but rather what is needed is a set of generic uncertainties that can be used with simple models to capture the possible range of behaviors of more complex models for regions with little or no data.

## 1.5 Spatial Smoothing of Seismicity

Most PSHA studies will include areal source zones. Typically, the historical seismicity is assumed to be uniformly distributed over specified source zones or the historical seismicity is smoothed spatially with some smoothing kernel of a selected width in distance. The degree of spatial smoothing to use is a recurring issue in PSHA. In some studies, very little smoothing is applied and the model of the seismicity closely approximates the historical seismicity. A drawback of this approach is that it does not allow for future earthquakes to occur in regions without historical seismicity. In other studies, the seismicity is smoothed over large regions which accomodates the possible occurrence of earthquakes in regions without historical seismicity. A drawback of this approach is that it may put too high a rate of seismicity in the regions with little or no historical earthquakes.

A common problem is that there is too much smoothing of the seismicity. As an example, consider a dam site located in the westen Foothills in California. In addition to the faults, background sources were used in the PSHA to account for the observed seismicity not associated with known mapped faults. The background seismicity is shown in Figure 6. An PSHA was conducted assuming that the seismicity within 50 km was uniformly distributed over the zone. The deaggregation for a return period of 1500 years showed that the controlling source was a magnitude 5.5 earthquake at a distance of 10-20 km; however, the plotted seismicity shown in Figure 6 shows that while there have been many historical earthquakes within 50 km of the site, there have been no historical earthquakes with magnitudes greater than 3.0 located within 17 km of the site. So the hazard result is not consistent with the observed seismicity.

A simple evaluation of the recorded seismicity indicates that the rate of earthquakes within 50 km of the site is not uniform. There have been 40 independent earthquakes (excluding aftershocks) within 50 km of the site with magnitude greater than 3.0 in the last 24 years. If the seismicity was uniformly distributed, then the chance of having 40 events occurring within 50 km of the site, but with zero occurring within 17 km of the site is given by

$$P(\text{zero eqk with } R<17 \text{ km}) = (1-(17/50)^2)^{40} = 0.007 \qquad (3)$$

Therefore, there is less than 1% chance that the seismicity is uniformly distributed.

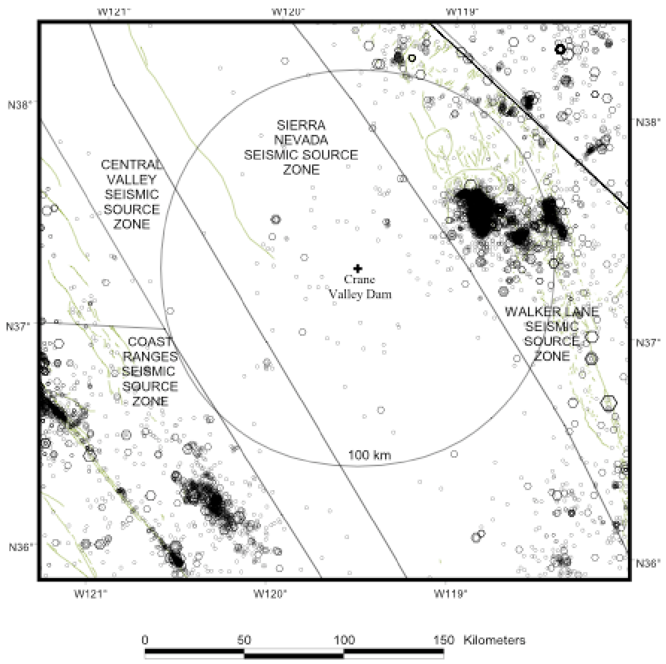


**Figure 6: Example of background seismicity for a site in the Western Foothills in California.**

To account for this non-uniform spatial distribution of seismicity within this areal source zone, the seismicity within 50 km can be modeled using two source zones: one given by a circle with a radius of 17 km centered on the site and one given by a donut shaped region with an outer radius of 50 km and an inner radius of 17 km centered on the dam. The seismicity for the donut shaped region can be directly estimated from the observed seismicity, since there are 40 earthquakes in the region, but the rate cannot be directly estimated for the small circle with zero earthquakes. One approach for estimating the seismicity rate for the zones with not historical seismicity is to set the rate such that there is a specified probability, P, of not observing any earthquakes during the time period, $N_{yr}$, in which the catalog is complete. Assuming a Poisson model for the earthquake recurrence, the rate is given by

$$rate = -\ln(1-P)/N_{yr} \tag{4}$$

If we set P=0.5 and use $N_{yr}$=24 years, then the rate is 0.15 eqk/yr for the 17 km radius zone. This corresponds to a rate of $1.6 \times 10^{-5}$ eqk/yr/km$^2$ which is much less than the rate of $2.4 \times 10^{-4}$ eqk/yr/km$^2$ for the donut-shaped region. Other values of the P can be selected, but this gives a method to model the lack of observed seismicity very close to the site. Using this new source model, the deaggregation from a revised PSHA showed that the controlling earthquake was a magnitude 6.0 earthquake at a distance of 20-30 km which is consistent with the observed. seismicity.

A more general method to determine the optimal spatial smoothing is needed. Musson [2004] evaluated the spatial smoothing used in seismic source models by comparing the predicted seismicity rates from the smoothed model with the observed seismicity. Using Monte Carlo methods, multiple realization of seismicity over the length of the historical catalog are developed. For each realization, the rates within specified sub-regions are computed. The distribution of number of earthquakes within each small region is then compared to the observed distribution for the same sub-regions to check if the observed seisimcity pattern is consistent with the smoothed model.

This approach can be used to select the optimum smoothing. Starting with a very broad spatial smoothing, the observed seismicity distributions can be compared to the simulated distributions. With very broad smoothing, the oberved seismicity will likely be inconsistent with the model. The extent of the spatial smoothing can be reduced until there is at least a 10% chance, for example, that the observations are consistent with the model. This method would not work if we start with very little smoothing, because this will always be consistent with the observations.

## 1.6 Lower Bound Magnitude

Probabilistic seismic hazard analysis (PSHA) for a site combines the hazard from all possible earthquakes in the site region that are potentially damaging. In current practice, non-damaging earthquakes are those with magnitudes below a selected lower bound earthquake magnitude. Earthquakes above the minimum magnitude are considered to be potentially damaging, and earthquakes below the minimum magnitude are not potentially damaging. Typically, the lower bound magnitude is set at M5.0 because well-engineered structures have not been observed to be damaged in earthquakes less than M5.0, but in some PSHAs, smaller lower bound magnitudes are used (e.g. M4.0 or 4.5)

The lower bound magnitude approach used in PSHA is equivalent to assuming that the probability of an earthquake being potentially damaging is a step function. For example, if the minimum moment magnitude is 5.0, then a moment magnitude 5.01 has a probability of 1.0 of being potentially damaging, whereas a moment magnitude 4.99 has a probability of 0.0 of being potentially damaging. Clearly, the step function is not realistic and does not properly represent the potential for damage as a function of earthquake magnitude. The transition from not potentially damaging to potentially damaging should be a smoother distribution on magnitude.

A drawback of using a fixed lower bound magnitude to remove not potentially damaging earthquakes from the hazard analysis is that in some cases, the results (both the hazard curve and the deaggregation) can be sensitive to the selection of the lower bound magnitude. In these cases, the deaggregation will show a significant contribution to the hazard from earthquakes just above the lower bound magnitude. This is typical for sites in which a nearby background source zone has a large contribution to the hazard. As an example, the deaggregation of the 10 Hz spectral acceleration hazard (in magnitude only) is shown in Figure 7 for a site located in the Eastern United States and a return period of 10,000 years using minimum magnitudes of 4.6. The deaggregation shows that there is significant contribution to the hazard from small magnitude earthquakes. If this lower bound magnitude was raised to 5.5, then there would be a significant change on the hazard (reduced by a factor of 2) and the deaggregation (shifted to larger magnitudes).

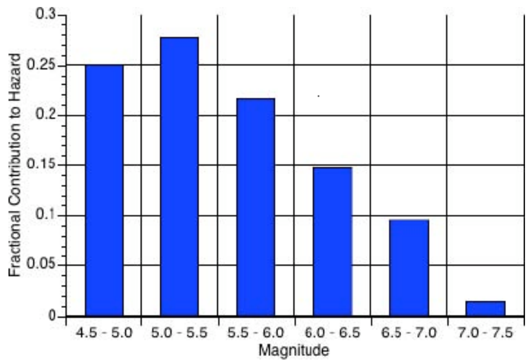


**Figure 7: Example of the deaggregation of the 20 Hz spectral acceleration for a site in the eastern United States for a lower bound magnitude of 4.6.**

This sensitivity to the lower bound magnitude could be avoided by applying a smooth transition from not potentially damaging to potentially damaging. One option is to set the potentially damaging in terms of some additional ground motion parameter. For nuclear power plants, Reed and Kennedy [1985] suggested using the "cumulative absolute velocity" (CAV) to identify earthquakes that are not potentially damaging regardless of the response spectral values. The CAV is defined as the average value of the absolute value of acceleration during 1 sec time windows that include an acceleration of 0.025g or larger, multiplied by the total duration of the 1-sec time windows. They recommended that if the CAV value were less than 0.016g-sec, then the ground motion would not be potentially damaging to well-engineered structures.

The standard hazard is written as

$$v(Sa > z) = \sum_{i=1}^{N_{source}} N_i(M > M_{\min}) \int_{M_{\min}}^{M_{\max i}} \int_{r=0}^{\infty} f_{mi}(M) f_{ri}(r,M)\, P(Sa > z \mid M, r)\, dr\, dM \tag{5}$$

The hazard can be modified to remove earthquakes with CAV<$CAV_{min}$ from the hazard calculation [EPRI, 2006a].

$$\nu(Sa > z, CAV > CAV_{\min}) = \sum_{i=1}^{N_{source}} N_i(M > M_{\min}) \int_{M_{\min}}^{M_{\max i}} \int_{r=0}^{\infty} f_{mi}(M) f_{ri}(r,M)\, P(Sa > z, CAV > CAV_{\min} \mid M, r)\, dr\, dM \tag{6}$$

The difference in the hazard integral is that instead of the probability of Sa>z for a given M and R, we have the joint probability of Sa>z and CAV>$CAV_{min}$ for a given M and R. In eq 7, we have the joint probability of Sa>z and CAV>$CAV_{min}$. The Sa and CAV are not independent so we need to compute the conditional probability of CAV >$CAV_{min}$ given Sa>z. If the spectral acceleration is determined by the M and R through the ground motion attenuation relation and the CAV is determined by M and R, then it is not apparent why are we concerned with the dependence of CAV on Sa. It would appear that the dependence is already accommodated through the M and R variables. The reason we need to consider the dependence is that there is aleatory variability of the ground motion and CAV for a given M and R. The ground motion attenuation relation gives the median and standard deviation of ln(Sa) for a given M and R. So even if the M and R are known values, there is a large range of ground motions and CAV values that could occur. Intuitively, if we have a higher than average ground motion, then we expect to also have a higher than average CAV; conversely, if we have a lower than average ground motion, then we expect a lower than average CAV. Therefore, we expect the CAV variability to be correlated with the Sa variability.

To account for the correlation, the CAV model is developed so that it depends on M, R, and Sa, not just on M and R. This avoids the need to develop a new ground motion attenuation relation that is dependent on CAV.

$$\nu(Sa > z, CAV > CAV_{\min}) = \sum_{i=1}^{N_{source}} N_i(M > M_{\min}) \int_{M_{\min}}^{M_{\max i}} \int_{r=0}^{\infty} f_{mi}(M) f_{ri}(r,M)\, P(Sa > z \mid M, R)\, P(CAV > CAV_{\min} \mid Sa > z, M, r)\, dr\, dM \tag{7}$$

This form of the hazard integral has an implicit integration over the ground motion variability. This hazard integral can be re-written to explicitly integrate over the ground motion variability:

$$\nu(Sa > z, CAV > CAV_{\min}) = \sum_{i=1}^{N_{source}} N_i(M > M_{\min}) \int_{M_{\min}}^{M_{\max i}} \int_{r=0}^{\infty} \int_{\varepsilon} f_{mi}(M) f_{ri}(r,M)\, P(Sa > z \mid M, R, \varepsilon)\, P(CAV > CAV_{\min} \mid Sa, M, r)\, dr\, dM d\varepsilon \tag{8}$$

where ε is the number of standard deviations of the ground motion. The advantage of the form of eq. (8) is that the CAV model is dependent on the Sa value rather than Sa>z. It is easier to develop a model for CAV given Sa, than it is to develop a model for CAV given Sa>z. The approach shown here for CAV will be applicable to any other additional parameter that may be considered in defining potentially damaging and non-potentially damaging ground motions.

Figure 8 shows an example of computing the hazard for 20 Hz spectral acceleration with the additional constraint that the CAV>0.16g-sec. The hazard curves becomes flat at low ground motion levels because it is the rate of the PGA exceeding 0.025g (since the CAV is defined to be zero for PGA<0.025g). At very low probability levels, the use of CAV does not have a strong effect because the CAV>0.16g-sec for large spectral acceleration even for small magnitude earthquakes. This may indicate that a different parameter should be considered that is not so strongly correlated with Sa.

The deaggregation with and without the CAV filtering is shown in Figure 9. Using CAV filtering removes many of the smaller magnitude events and leads to a smooth decay in the contribution of smaller magnitudes.

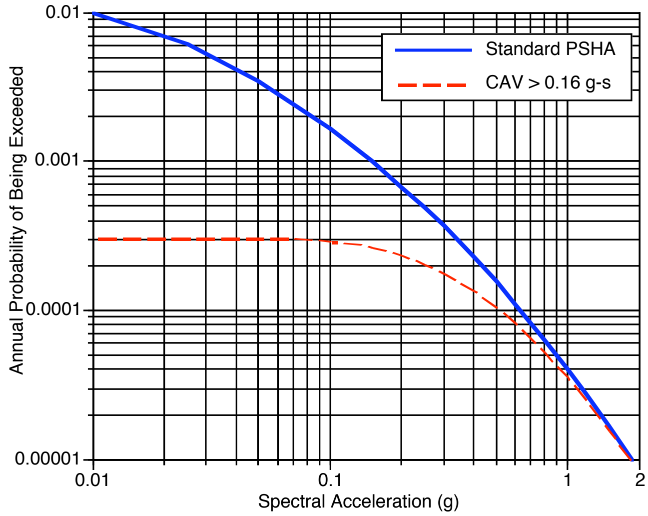


**Figure 8: 20 Hz spectral acceleration for a return period of 10,000 years for a site in the Eastern U.S.**

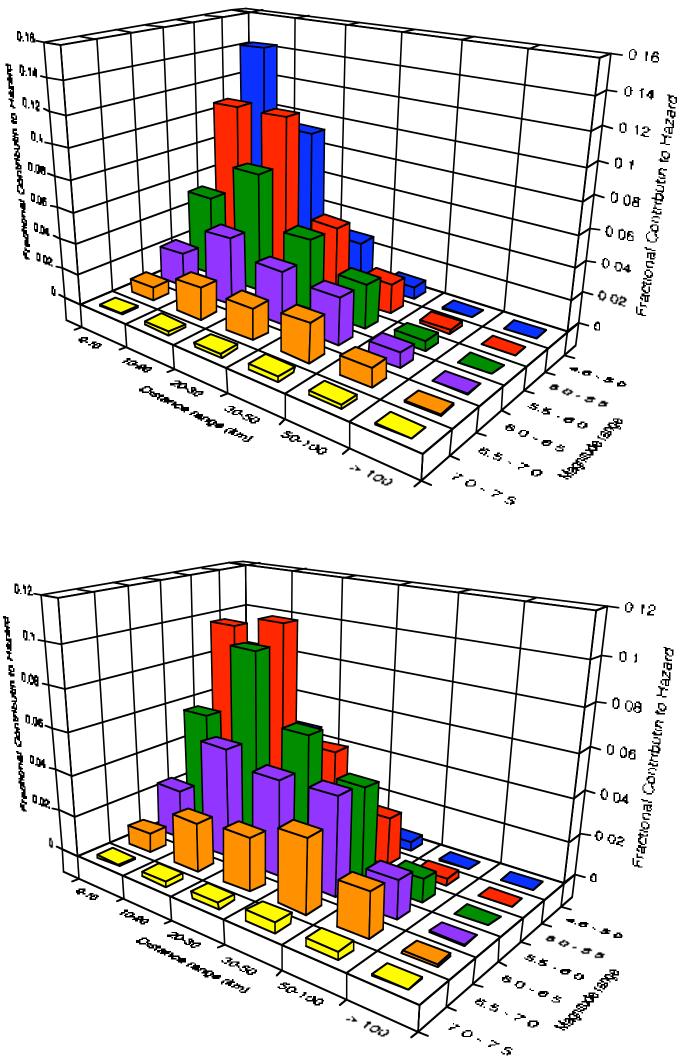


**Figure 9: Deaggregation of 20 Hz spectral acceleration for a return period of 10,000 years for a site in the Eastern U.S. The upper plot is the standard PSHA and the lower plot is the CAV filtered deaggregation.**

## 4. FUTURE DEVELOPMENTS

The previous section addressed potential improvements in PSHA but using the current general appraoch. Major advances in PSHA methodology require much greater communication between the hazard analyst and the project engineers.

### 4.1 Incorporating Site-Specific Site Response in to PSHA

For developing probabilistically based design ground motions for soil sites, there are two main approaches that are commonly used: (1) conduct the PSHA for a rock site condition, develop a rock site design spectrum, and then conduct a deterministic site response analysis (either site-specific or based on code based amplification factors), or (2) conduct the PSHA for the soil site condition directly using generic ground motion models that are considered to be applicable to site condition. These two approaches often lead to signficantly different UHS.

One of the reasons for the differences is that the generic attenuations often correspond to linear site response in terms of estimating the median ground motion and/or the variability of the ground motion. Most empirical ground motion models include the site effect as a simple scale factor that applies to the rock motion [e.g. BJF, 1997] which implies linear site-response. Other models, such as Abrahamson and Silva [1997], include a non-linear site response term in the model for the median ground motion, but use a constant standard deviation about the median. Since the ground motion is modeled as lognormal (see section 1), using a constant value of the standard deviation for soil sites is only appropirate for linear site response. So models such as Abrahamson and Silva [1997] have a mixture of non-linear and linear site response aspects.

McGuire et al. [2001] present several alternative methods for incorporating site response into a PSHA. The most direct method is to incorporate the site response directly into the hazard calculation. In this approach, a large suite of site response analyses is conducted for a range of input rock ground motion levels. A model is then developed for the median amplification and the aleatory variability of the amplification as a function of the rock motion level. The soil hazard curve is then computed by integrating the hazard over the rock motion ground motions as shown below.

The standard rock hazard can be written as

$$v(Sa_{rock}(T) > z) = \sum_{i=1}^{N_{source}} N_i(M_{\min}) \int_M \int_R f_m(M)\, f_R(R,M)\; P(Sa_{rock}(T) > z \mid \hat{S}a_{rock}(M,R), \sigma_{rock})\, dM\, dR \tag{9}$$

The site-specific soil hazard can be written as

$$v(Sa_{soil}(T) > z) = \sum_{i=1}^{N_{source}} N_i(M_{\min}) \int_M \int_R \int_\varepsilon f_m(M)\, f_R(R,M) f_\varepsilon(\varepsilon_{rock})\; P(Sa_{soil}(T) > z \mid \hat{S}a_{soil}(M,R,\varepsilon), \sigma_{amp}(Sa_{rock}))\, dM dR d\varepsilon \tag{10}$$

where $\varepsilon$ is the number of standard deviation of the spectral acceleration on rock and

$$\hat{S}a_{soil}(M,R,\varepsilon) = Sa_{rock}(M,R,\varepsilon)\, Amp(Sa_{rock}) \tag{11}$$

While computationally straightforward, this approach requires a large site-response effort because the amplification and the variability of the amplification need to be computed for the full range of rock ground motions levels, earthquake magnitudes, and input time histories. To be practical, this approach requires that an efficient method to conduct a large number of site response analyses is needed. Site response that uses random vibration theory is well suited for this type of problem [Rathje and Ozbey, 2006]. I expect that the use of site-specific amplifications models in PSHA will become more common over the next 10 years.

## 4.2 Using Structure-Specific Information in the Hazard Analysis

One common view in earthquake engineering is that the hazard analysis for a site should be done independent of the properties of the structure. The argument in favor of this approach is that the site does not know about the structure, so the free-field hazard should be conducted without regard to the properties of the structure (SSI effects are a different issue). If engineer does a complete evaluation of the structure for each scenario in the hazard (e.g. for each magnitude, distance and ground motion level), then this approach would work well; however, most projects only evaluate the structure for a small number of scenarios. If only a small number of scenarios are considered, then the properties of the structure should be considered so that the appropriate scenarios and ground motoin levels are selected for analysis.

To demonstrate this, consider an example in which the ground motion with a 500 year return period is developed. In typical practice, a PSHA is conducted, the ground motion with a 500 year return period is determined from the hazard curve, and the controlling earthquake is taken as the mode of the magnitude-distance deaggregation. If a time history is needed for structural analysis, then a time history, consistent with the controlling magnitude and distance, is selected and scaled to the target ground motion.

In the PSHA, the rates of the ground motion that exceeded a specific threshold were summed from all earthquakes scenarios. It make sense to sum the rates of exceeding a ground motion level from the different scenarios if the scenarios all have the same impact on the structure for the same ground motion level. This may not be the case. For example, a small magnitude earthquake with the same spectral acceleration at a given spectral period may not have the same impact on the structure as a large magnitude earthquake with the same spectral acceleration. This issue has been considered for years in liquefaction assessment through the use of magnitude weighting factors (MWF). In liquefaction assessments, the PGA and the number of cycles are both needed. Typically, the earthquake magnitude is used as a proxy for the number of cycles and the MWF are used to normalize the PGA to the PGA for an equivalent M7.5 earthquake. For this case, the PSHA would be conducted by summing the rates of the magnitude-weighted PGA exceeding a specified value. The magnitude-weighted PGA with a 500-year return period would be used with a M7.5 time history. In this liquefaction assessment example, it does not make sense to use the PGA value with a 500-year return period and a single magnitude from the controlling source. This is one example of why the structure response should be considered in the PSHA.

Next, consider a more complicated case that came up on a project in California. I've simplified this example for illustration. The site is located between two faults with similar size earthquakes. The seismic issue for this site is a slope failure and a key parameter is the peak ground velocity (PGV). Following standard practice, a PSHA was conducted and resulted in the PGV hazard curve shown in Figure 10. The deaggregation for this example showed that a magnitude 7.0-7.5 earthquake was found to be the controlling earthquake.

Non-linear dynamic analyses of the slope showed that the slope deformation depended on the direction that the waves approached the slope. In this case, ground motions from fault 1 with a specified PGV lead to twice the slope deformation as ground motions from fault 2 with the same PGV.

Follwong standard practice, design time histories were developed by selecting recordings from the appropriate magnitude and distance range and then scaled them to be consistent with the PGV with a 1000 year return period. The problem with this approach is that the PGVs from the two faults don't have the same effect on the structure even though they have similar magnitudes. Just as with the liquefaction assessment, the PGVs need to ne normalized to a selected direction of propagation (e.g. we need to apply "propagation direction weighting factors" similar to the MWF used in liquefaction assessments. Then the selected time history can be scaled to the normalized PGV and applied to the reference direction of propagation. Using the more severe direction of propagation as the reference, the modified PGV hazard is shown in Figure 10. This figure shows that the time history should be scaled to 125 cm/s rather than 150 cm/s for the 1000 year return period if the wave propagation is assumed to be from fault 1.

In both of these examples, the response of the structure affects how the PSHA is conducted. These issues can be avoided if a risk calculation is computed for the key structural response parameter, rather than just conducting the hazard for the ground motion parameter. This gets back to the introductory discussion of the probabilistic approaches. Using a PSHA to select ground motion levels based on a set return period is just a simplified risk calculation. The further we can push the calculation to risk directly (for example using inelastic response spectral values rather than elastic spectral values), we will reduce the inconsistencies that come up in both deterministic and probabilistic approaches.

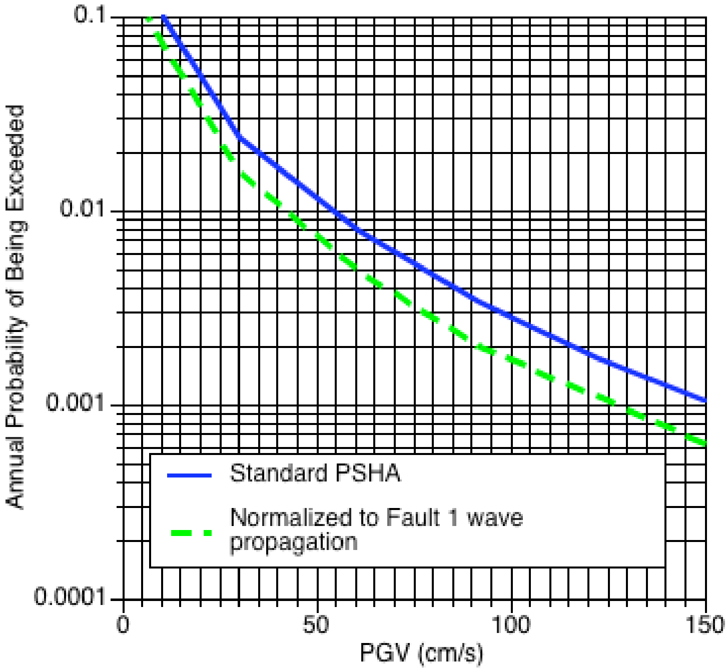


**Figure 10. Impact of accounting for structure specific effects on the PGV hazard**

### 4.3 Vector Hazard

In a standard PSHA, the hazard is computed for scalar ground motion parameters, such as spectral acceleration at a single period and damping. In vector hazard, the PSHA is conducted for two or more scalar ground motion parameters. Since the response of structures can depend on more than one ground motion parameter, using vector hazard allows for more accurate predictions of the response by using more complete descriptions of the scenarios. For example, for buildings, knowing the spectral acceleration at both the fundamental period and twice the fundamental period improves the prediction of the response of structures as they become more non-linear.

Conducting a vector PSHA is straight-forward but it requires additional ground motion models. In addition to the standard "attenuation relations" for the ground motion parameters of interest, the covariance of the ground motion parameters is also needed. This is computed from the correlation of the residuals of the two (or more) ground motion parameters.

The main difficulty with a vector PSHA is the implementation of the results. There is not a single parameter with a return period that can be read from a hazard curve. Instead, a vector PSHA leads to a set of ground motions parameters with a rate of ocurrence of the vector of parameters. For the case of a two parameter vector hazard, this would be the rate of occurrence of pairs of the two parameters (e.g. $Sa(T_1)$ and $Sa(T_2)$ pairs). To use this information requires a seismic risk analyis be conducted for some structual response parameter. This requires that models of the structural response be developed as a function of the vector of ground motion parameters. These structural response models can be developed in academic studies that can conduct large numbers of evaluations of the models, but they are beyond the scope of most projects.

In my opinion, the future of PSHA will be in vector hazard analysis, but before that happens, we need to develop simplifed methods to approximate the structural response as a function of a vector of key ground motion parameters. With these simplfied models that are feasible within the budget of typical projects, vector hazard analysis will have a much broader application. This will also help to focus the ground motion studies on developing models that better represent realistic earthquake loading conditions.

## 5. REFERENCES


Abrahamson, N. A. and J. J. Bommer (2005). Probability and uncertainty in seismic hazard analysis, Earthquake Spectra, 21, 603-607.

Abrahamson, N. A. and W. Silva (1997). Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes, *Seis. Res. Letters*, Vol. **68**, No. 1, p. 94-127.

Abrahamson, N.A., Birkhauser P., Koller, M., Mayer-Rosa, D. Smit, P., Sprecher, C., Tinic, S. and Graf, R. (2002). PEGASOS – a comprehensive probabilistic seismic hazard assessment for nuclear power plants in Switzerland. *12th European Conference on Earthquake Engineering*, London, Paper No. 633.

Abrahamson, N. A. and W. J. Silva (2006). Update of the Abrahamson and Silva attenuation relation for shallow crustal earthquakes in active tectonic regions, Pacific Earthquake Engineering Research Center, Berkeley, California.

Bommer, J, J, and N. A. Abrahamson (2006). Why do modern seismic hazard analyses often lead to increased hazard estiamtes?, Bull. Seism. Soc. Am., in Press.

Boore, D. M., W. B. Joyner, and T. E. Fumal (1997). Equations for estimating horizontal respones spectra and peak accelereation from western North American earthquakes: a summary of recent work, Seism. Res. Let., **68**, 128-153.

Baker, J. and C. A. Cornell (2006). Spectral shape, epsilon, and record selection, Earth. Eng. Struct. Dyn, **35**, 1077-1095.

Bazzurro, P. and C.A. Cornell (1999). Disaggregation of seismic hazard, Bull. Seism. Soc. Am., **89**, 501-520.

Cornell, C. A. (1968). Engineering siemsic risk analysis, Bull. Seism. Soc. Am., **58**, 1583-1606. Erratum: **59**, 1733.

EPRI (2006a). Program on Technology Innovation: Truncation of the lognormal distribution and value of the standard deviation for ground motoin models in the Central and eastern United States, EPRI, Palo Alto, CA and the U.S. Department of Energy: 2006, Report #1013105.

EPRI (2006b). Use of CAV in determining effects of small magnitude earhtquakes on seismic hazard analyses, EPRI, Palo Alto, CA and the U.S. Department of Energy: 2006. Report #1012965

Esteva, L. (1970). Seismic risk and seismic design descitions, In Seismic Desgin for Nuclear Power Plants, Ed. R. J. Hansen, MIT Press, Cambridge Mass, 142-182.

McGuire, R. K, W. J. Silva, and C. J. Consantino (2001). Technical basis for revision of regulatory guidence on design gorund motions: Hazard- and risk-consistent ground motion spectra guidelines, US Nuclear Regulatory Comm., NUREG/CR-6728.

McGuire, R. K. (2004). Analysis of Seismic Hazard and Risk, EERI Monograph, Earhtquake Engineering research Center, Oakland, California.

Musson, R. M. W. (2004). Objective validation of seismic hazard source models, Proc. 13WCEE, Vancouver.

PEER (2005) NGA Strong Motion Database. http://peer.Berkeley.edu/NGA, 2005.

Rathje, E.M., and Ozbey, M.C. (2006). Validation of Random Vibration Theory-Based Site Response Analysis, submitted to ASCE Journal of Geotechnical and Geoenvironmental Engineering.

Reed, J.W. and R.P. Kennedy (1988). A Criterion for Determining Exceedance of the Operating Basis Earthquake, EPRI Report NP-5939, July 1988.