Conditional Spectrum Computation Incorporating Multiple Causal Earthquakes and Ground-Motion Prediction Models

by Ting Lin, Stephen C. Harmsen, Jack W. Baker, and Nicolas Luco

Abstract  The conditional spectrum (CS) is a target spectrum (with conditional mean and conditional standard deviation) that links seismic hazard information with ground-motion selection for nonlinear dynamic analysis. Probabilistic seismic hazard analysis (PSHA) estimates the ground-motion hazard by incorporating the aleatory uncertainties in all earthquake scenarios and resulting ground motions, as well as the epistemic uncertainties in ground-motion prediction models (GMPMs) and seismic source models. Typical CS calculations to date are produced for a single earthquake scenario using a single GMPM, but more precise use requires consideration of at least multiple causal earthquakes and multiple GMPMs that are often considered in a PSHA computation. This paper presents the mathematics underlying these more precise CS calculations. Despite requiring more effort to compute than approximate calculations using a single causal earthquake and GMPM, the proposed approach produces an exact output that has a theoretical basis. To demonstrate the results of this approach and compare the exact and approximate calculations, several example calculations are performed for real sites in the western United States. The results also provide some insights regarding the circumstances under which approximate results are likely to closely match more exact results. To facilitate these more precise calculations for real applications, the exact CS calculations can now be performed for real sites in the United States using new deaggregation features in the U.S. Geological Survey hazard mapping tools. Details regarding this implementation are discussed in this paper.

Introduction

Ground-motion selection for structural and geotechnical system analysis is often associated with a target response spectrum that is derived from probabilistic seismic hazard analysis (PSHA) results. The conditional spectrum (CS) is one such target spectrum that estimates the distribution (with mean and standard deviation) of the response spectrum, conditioned on the occurrence of a target spectral acceleration value at the period of interest. As this CS concept is considered for practical use, several common approximations need to be further explored. Typical CS calculations to date are produced for a single earthquake ground-motion scenario (i.e., magnitude, distance, and ground-motion intensity of interest) and computed using a single ground-motion prediction model (GMPM, previously known as an attenuation relation, a ground-motion prediction equation, a ground-motion model, or a ground-motion relation). The scenario is generally determined from PSHA deaggregation, but PSHA deaggregation calculations for real sites often show that multiple earthquake scenarios contribute to the occurrence of a given ground-motion intensity. Additionally, modern PSHA calculations are performed with multiple GMPMs using a logic tree that also includes seismic source models. Incorporating those features is thus necessary to compute a CS that is fully consistent with the PSHA calculations upon which it is based.

This paper presents the methodology for performing refined CS computations that precisely incorporate the aleatory uncertainties (which are inherently random) in earthquake events with all possible magnitudes and distances, as well as the epistemic uncertainties (which are due to limited knowledge) from multiple GMPMs and seismic source models. Three approximate calculation approaches and the exact calculation approach are presented, with increasing levels of complexity and accuracy. To demonstrate, several example calculations are performed for representative sites with different surrounding seismic sources: Stanford in northern California, Bissell in southern California, and Seattle in the Pacific Northwest. The results evaluate the exact and approximate calculations, and analyze factors that contribute to the differences in accuracy. Note that while the exact approach is more cumbersome, it does not need to be computed by the user, because these exact CS calculations have been implemented in the U.S. Geological Survey (USGS) seismic hazard mapping tools, and could be incorporated into other PSHA software as well. Details regarding this new tool
and issues related to implementation of these concepts are provided.

Basic Conditional Spectrum Computation

A wide variety of techniques has been developed in the past to select ground-motion inputs for nonlinear dynamic analysis (Haselton et al., 2009; Katsanos et al., 2010). One common approach involves selecting ground motions with response spectra that match the target spectrum (Watson-Lamprey and Abrahamson, 2006; Beyer and Bommer, 2007; American Society of Civil Engineers, 2010; Applied Technology Council, 2011). The conditional mean spectrum (CMS) is one such spectrum that incorporates correlation across periods to estimate the expected $Sa$ values at all periods $T_i$ [Sa($T_i$)] given the target $Sa$ value at the period of interest $T^*$ [Sa($T^*$)] (Baker and Cornell, 2006; Somerville and Hamburger, 2009; Abrahamson and Al Atik, 2010; Baker, 2011; Gulerce and Abrahamson, 2011; Somerville and Thio, 2011).

The basic CMS computation procedure is as follows. First, we obtain the target spectral acceleration at period $T^*$, $Sa(T^*)$ from PSAH and its associated mean causal earthquake magnitude ($M$), distance ($R$), and other parameters ($\theta$), from deaggregation. Next, we use a GMPM to obtain the logarithmic mean and standard deviation of $Sa$ at all periods $T_i$, denoted as $\mu_{ln Sa}(M, R, \theta, T_i)$ and $\sigma_{ln Sa}(M, \theta, T_i)$. For any $Sa(T_i)$ value, we compute the $\varepsilon(T_i)$, the number of standard deviations by which $ln Sa(T_i)$ differs from the mean spectral ordinate predicted by a given GMPM, $\mu_{ln Sa}(M, R, \theta, T_i)$, at $T_i$:

$$\varepsilon(T_i) = \frac{ln Sa(T_i) - \mu_{ln Sa}(M, R, \theta, T_i)}{\sigma_{ln Sa}(M, \theta, T_i)}.$$  

(1)

The target $\varepsilon(T^*)$, for the target $Sa(T^*)$ value, can also be computed using equation (1). We can then compute the conditional mean spectral acceleration at other periods $T_i$, $\mu_{ln Sa(T_i)|ln Sa(T^*)}$, using the correlation coefficient between pairs of $\varepsilon$ values at two periods, $\rho(\varepsilon(T_i), \varepsilon(T^*))$, hereafter referred to as $\rho(T_i, T^*)$ (Baker and Jayaram, 2008):

$$\mu_{ln Sa(T_i)|ln Sa(T^*)} = \mu_{ln Sa}(M, R, \theta, T_i) + \rho(T_i, T^*)\sigma_{ln Sa}(M, \theta, T_i)\varepsilon(T^*).$$  

(2)

The spectrum defined by $\mu_{ln Sa(T_i)|ln Sa(T^*)}$ in equation (2) has been termed the CMS, because it specifies the mean values of $ln Sa(T_i)$, the exponentials of which are equivalent to the median values of $Sa(T_i)$ if it is lognormally distributed, conditional on the value of $ln Sa(T^*)$ (Baker, 2011).

Similarly, we can compute $\sigma_{ln Sa(T_i)|ln Sa(T^*)}$, the conditional standard deviation of spectral acceleration at period $T_i$, conditioned on the value of $Sa$ at $T^*$:

$$\sigma_{ln Sa(T_i)|ln Sa(T^*)} = \sigma_{ln Sa}(M, \theta, T_i)\sqrt{1 - \rho^2(T_i, T^*)}.$$  

(3)

The conditional standard deviation $\sigma_{ln Sa(T_i)|ln Sa(T^*)}$ from equation (3), when combined with the conditional mean value $\mu_{ln Sa(T_i)|ln Sa(T^*)}$ from equation (2), specifies a complete distribution of logarithmic spectral acceleration values at all periods (where the distribution at a given period is Gaussian, as justified by Jayaram and Baker, 2008). We term the resulting spectrum distribution as a CS, to be distinguished from the CMS that does not consider the variability specified by equation (3). It is noteworthy that as an extension of the CMS approach, Bradley (2010) proposed a generalized conditional intensity measure (GCIM) approach that considers the complete distribution of conditional intensity measures other than $Sa$.

The input earthquake parameters required for the previous CS calculations are those required by the GMPM and can include magnitude ($M$), distance ($R$), and other parameters such as rupture mechanism and site conditions ($\theta$). In this paper, we will sometimes use $M/R$ in the text as an abbreviation for $M/R/\theta$ (although $\theta$ will be included in equations where appropriate). To implement the CS in practice, we need to determine which $M/R$ and GMPM to use. Common approximations to compute the CS include using the mean magnitude and distance from deaggregation, along with a single GMPM (Baker, 2011). As explained more in the next section, these approximations need to be evaluated for the practical implementation of CS as a target spectrum for selecting ground motions.

Deaggregation

Computing CS at real sites requires us to consider the following two factors: First, deaggregation will produce multiple causal earthquake magnitude and distance values for a given $Sa(T^*)$ amplitude, as illustrated in the USGS deaggregation plots in Figure 1, where the height of each column represents the percentage contribution from each $M/R$ combination. Second, PSHA calculations use multiple GMPMs and seismic source models to compute seismic hazard at a site through a logic tree (Kramer, 1996; McGuire, 2004; Scherbaum et al., 2005; Bommer and Scherbaum, 2008; Petersen et al., 2008). A refined CS computation, therefore, needs to consider multiple causal earthquakes and GMPMs, if not multiple seismic source models.

Probability seismic hazard analysis accounts for the aleatory uncertainties in earthquake events by combining the frequencies of occurrence of all earthquake scenarios with different magnitudes and distances with predictions of resulting ground-motion intensity, in order to compute seismic hazard at a site (Kramer, 1996; McGuire, 2004). Probability seismic hazard analysis also incorporates the epistemic uncertainties in the seismic source models and ground-motion predictions by considering multiple models in a logic tree. For instance, the USGS utilizes three GMPMs (Boore and Atkinson, 2008; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008) for crustal seismic sources in the western United States (Petersen et al., 2008). Traditional PSHA deaggregation (McGuire, 1995; Bazzurro and Cornell, 1999; Harmsen, 2001), however, only reports distributions of
causal $M/R/\varepsilon$ values given an $Sa$ amplitude. The McGuire (1995) deaggregation is conditional on $Sa$ that equals a target value, termed $Sa$ occurrence, while the Bazzurro and Cornell (1999) deaggregation is conditional on $Sa$ that exceeds a target value, termed $Sa$ exceedance. Depending on the $Sa$ values of interest, either deaggregation approach can be used. Such deaggregation can be extended to include distributions of the logic-tree branches, such as GMPMs, that contribute to predictions of $Sa$ occurrence (or exceedance).

Just as the deaggregation of magnitude and distance identifies the relative contribution of each earthquake scenario to $Sa$ occurrence (or exceedance), the deaggregation of GMPMs tells us the probability that the occurrence (or exceedance) of that $Sa$ level is predicted by a specific GMPM. Note that the GMPM deaggregation weights differ from the logic-tree weights; in decision analysis (Benjamin and Cornell, 1970), the logic-tree weights are equivalent to prior weights, whereas the deaggregation weights may be interpreted as posterior weights given the occurrence (or exceedance) of the ground-motion amplitude of interest. Additional details on GMPM deaggregation are provided in Lin and Baker (2011). For the purpose of response spectrum predictions, the key elements are the GMPMs and their input earthquake parameters (e.g., $M$, $R$, $\varepsilon$). Hence, the focus of deaggregation here will be on these two. The other portions of the logic tree (e.g., recurrence type, rates, maximum magnitude) do not influence spectrum predictions and so can be grouped for the purpose of these calculations. The USGS has recently begun providing GMPM deaggregation outputs (see Data and Resources). These outputs facilitate the exact calculations of CS described next.

Conditional Spectrum Calculation Approaches

We now consider several options for computing CS that incorporate consideration of multiple causal earthquake
scenarios and multiple GMPMs, as well as multiple seismic source models and their logic-tree branches. We introduce several approximate calculation approaches with increasing complexity but also with increasing accuracy, followed by the exact calculation. Differences between the approaches are highlighted, and these approaches are evaluated later to determine the accuracy of the approximate approaches.

Method 1: Approximate CS Using Mean M/R and a Single GMPM

The most basic method for computing an approximate CS is to utilize a single earthquake scenario and single GMPM, so that equations (2) and (3) can be used directly. In current practice, this is done by taking the mean value of the causal magnitudes and distances from deaggregation, denoted here as \( M \) and \( R \) (Baker, 2011). Similarly, the mean value of other causal parameters (\( \theta \)) can be obtained or inferred from deaggregation. These mean values can then be used with a single GMPM (even though the underlying hazard analysis utilized several GMPMs). The resulting CS calculations are given in the following equations, utilizing the subscript \( k \) to denote that the calculations are based on a single GMPM indexed by \( k \). The equations are also denoted as being approximately equal to the true CS values, given the simplifications made here.

\[
\mu_{\ln S_{\text{eq}}(T)} \approx \mu_{\ln S_{\text{eq}}(\tilde{M}, \tilde{R}, \tilde{\theta}, T_i)} + \rho(T_i, T^*) \sigma_{\ln S_{\text{eq}}(\tilde{M}, \tilde{R}, \tilde{\theta}, T_i)} \varepsilon(T^*), \tag{4}
\]

and

\[
\sigma_{\ln S_{\text{eq}}(T)} \approx \sigma_{\ln S_{\text{eq}}(\tilde{M}, \tilde{R}, \tilde{\theta}, T_i)} \sqrt{1 - \rho^2(T_i, T^*)}, \tag{5}
\]

where the mean and standard deviation predicted by GMPM \( k \) are denoted \( \mu_{\ln S_{\text{eq}}(T)} \) and \( \sigma_{\ln S_{\text{eq}}(T)} \), and the CS computed for GMPM \( k \) is denoted \( \mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \) and \( \sigma_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \). Note that the correlation coefficient, \( \rho \), is assumed to be constant for each GMPM; this could be revised if desired.

Method 2: Approximate CS Using Mean M/R and GMPMs with Logic-Tree Weights

We can refine method 1 by considering all GMPMs used in the PSHA computation. The PSHA logic tree weights each model (these weights can be equal or unequal), and here we denote the weight for model \( k \) as \( p_k^l \), where the superscript \( l \) refers to logic tree. To obtain an approximate CS using all of these GMPMs, we repeat the single-GMPM calculation (equations 4 and 5) to obtain \( \mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \) and \( \sigma_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \) for each GMPM. We then sum up the resulting mean spectra \( [\mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*)] \), weighted by the logic-tree weights, to get a mean spectrum

\[
\mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \approx \sum_k p_k^l \mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*). \tag{6}
\]

In equations (6) and (7) we no longer have a subscript \( k \) on the resulting \( \mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \) and \( \sigma_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \), because the results are no longer specific to a single GMPM but incorporate multiple GMPMs.

Although the use of logic-tree weights is not rigorously correct, we introduce it here as a convenient approximation because these equations do not require GMPM deaggregation outputs, which are not currently available from many PSHA software tools.

Method 3: Approximate CS Using GMPM-Specific Mean M/R and GMPMs with Deaggregation Weights

In this section, we further refine the CS calculations by taking advantage of GMPM deaggregation if it is available (e.g., as it is from the new USGS tools). First, GMPM deaggregation will provide separate \( M/R \) deaggregation for each prediction model. Here we will use the mean \( M \) and \( R \) values for each model, denoted as \( \tilde{M}_k \) and \( \tilde{R}_k \) for GMPM \( k \). Using these values, means and standard deviations of the CS can be computed for GMPM \( k \) as follows:

\[
\mu_{\ln S_{\text{eq}}(T)} \mid \ln S_{\text{eq}}(T^*) \approx \mu_{\ln S_{\text{eq}}(\tilde{M}_k, \tilde{R}_k, \tilde{\theta}_k, T_i)} + \rho(T_i, T^*) \sigma_{\ln S_{\text{eq}}(\tilde{M}_k, \tilde{R}_k, \tilde{\theta}_k, T_i)} \varepsilon_k(T^*), \tag{8}
\]

and

\[
\sigma_{\ln S_{\text{eq}}(T)} \approx \sigma_{\ln S_{\text{eq}}(\tilde{M}_k, \tilde{R}_k, \tilde{\theta}_k, T_i)} \sqrt{1 - \rho^2(T_i, T^*)}. \tag{9}
\]

Note that the GMPM-specific \( \tilde{M}_k \) and \( \tilde{R}_k \) (in equations 8 and 9) are different from the \( M \) and \( R \) with respect to all GMPMs (in equations 4 and 5). As an intermediate step (e.g., between \( \tilde{R} \) and \( \tilde{R}_k \)), the concept of conditional deaggregation given GMPM can also be extended to compute \( \tilde{M}_k, \tilde{R}_k \mid \{M_k, \tilde{R}_k\} \) in a cascading or Rosenblatt-distribution manner.
The second GMPM deaggregation output used in this method is the probability that GMPM k predicted occurrence (or exceedance) of the $S_a$. These deaggregation probabilities are denoted as $p_k^d$, where the superscript $d$ refers to deaggregation, and they are generally not equal to the PSHA logic-tree weights that have been denoted as $p_k^s$. Utilizing these weights, along with the GMPM-specific $\mu_{\ln S_a(T_i)}|\ln S_a(T^*)$ and $\sigma_{\ln S_a(T_i)}|\ln S_a(T^*)$ from equations 8 and 9, a composite CS can be computed:

$$\mu_{\ln S_a(T_i)}|\ln S_a(T^*) \approx \sum_k p_k^d \mu_{\ln S_a(T_i)}|\ln S_a(T^*).$$

(10)

and

$$\sigma_{\ln S_a(T_i)}|\ln S_a(T^*) \approx \sqrt{\sum_k p_k^d [\sigma_{\ln S_a(T_i)}|\ln S_a(T^*)]^2 + (\mu_{\ln S_a(T_i)}|\ln S_a(T^*) - \mu_{\ln S_a(T_i)}|\ln S_a(T^*))^2].$$

(11)

As with method 2, the mean spectrum is the mean (over GMPMs) of the GMPM-specific means, except that here we have utilized the more appropriate deaggregation weights $p_k^d$. The standard deviation of the spectrum again contains contributions from the individual GMPM conditional standard deviations, plus the uncertainty from the variations in mean spectra across GMPMs. While this method incorporates more exact information than methods 1 or 2, it is still approximate because it utilizes only the mean earthquake scenario for a given GMPM. Method 4 will resolve that final approximation.

Method 4: “Exact” CS Using Multiple Causal Earthquake $M/R$ and GMPMs with Deaggregation Weights

With this final method we now account exactly for, when we compute the CS, the contribution that each earthquake magnitude/distance and each GMPM makes to the seismic hazard. For each causal earthquake combination $M_j/R_j$ and GMPM $k$, we can obtain a corresponding mean and standard deviation of the CS, denoted $\mu_{\ln S_a(T_i)}|\ln S_a(T^*)$ and $\sigma_{\ln S_a(T_i)}|\ln S_a(T^*)$, as

$$\mu_{\ln S_a(T_i)}|\ln S_a(T^*) = \mu_{\ln S_a}(M_j, R_j, \theta_j, T_i) + \rho(T^*, T_i) \varepsilon_j(T^*) \times \sigma_{\ln S_a}(M_j, \theta_j, T_i).$$

(12)

and

$$\sigma_{\ln S_a(T_i)}|\ln S_a(T^*) = \sigma_{\ln S_a}(M_j, \theta_j, T_i) \sqrt{1 - \rho^2(T_i, T^*)}. $$

(13)

A PSHA deaggregation that includes GMPM deaggregation will provide the weights, $p_{j,k}^d$, that indicate the contribution of each $M_j/R_j$ and GMPM $k$ to occurrence (or exceedance) of the $S_a$ of interest. Note that here we are considering the contributions of individual $M_j/R_j$ rather than approximating all contributing earthquake scenarios by simply a mean $M$ and $R$, as we did in equations (8) and (9). The exact CS incorporating multiple $M/R$ and GMPMs can then be evaluated by combining these individual

$$\mu_{\ln S_a(T_i)}|\ln S_a(T^*) \text{ and } \sigma_{\ln S_a(T_i)}|\ln S_a(T^*) \text{ with their corresponding weights, } p_{j,k}^d,$$

(14)

and

$$\mu_{\ln S_a(T_i)}|\ln S_a(T^*) \approx \sum_k \sum_j p_{j,k}^d \mu_{\ln S_a(T_i)}|\ln S_a(T^*).$$

(15)

Equations (14) and (15) are similar to those in method 3, except that there is now a second set of summations in the equations to account for the effect of multiple $M_j/R_j$ values instead of the single-mean values $M_k$ and $R_k$ from method 3. These equations provide an exact answer to the question, what is the mean and standard deviation of the response spectra associated with ground motions having a target $S_a(T^*)$, when that $S_a(T^*)$ could potentially result from multiple earthquake scenarios, and the $S_a$ predictions come from a logic tree with multiple GMPMs? It requires more effort to compute than the approximate approaches commonly used today, and it requires detailed deaggregation information including hazard contributions of GMPMs and other parameters, $\theta$, that may not be available.

Alternatively, an exact CS can be computed directly from the earthquake parameters and GMPMs that were used in PSHA computation to aggregate the hazard, as described in the next section. Although treatments of aleatory and epistemic uncertainties are often separated in PSHA, a single seismic hazard curve is typically derived with the consideration of aleatory uncertainties from multiple causal earthquakes and epistemic uncertainties from multiple GMPMs and
seismic source models (Petersen et al., 2008). Here we take a similar approach to compute a single CS that combines these aleatory and epistemic uncertainties.

**Aggregation Approach to Method 4 (Exact CS)**

Although the GMPSM deaggregations used in methods 3 and 4 are now available from the USGS, it would be impractical to provide deaggregations with respect to all of the other branches of the PSHA logic tree (e.g., for alternative moment-area equations in California and for body wave to moment magnitude equations in the central and eastern United States). Likewise, it would be cumbersome to provide deaggregations with respect to all of the other GMPSM input parameters besides $M$ and $R$ (e.g., rupture mechanism and hanging/footwall indicators), which were previously denoted as $\theta$. To account for these other branches and parameters in method 4 without additional deaggregation results, a CS can be calculated during the PSHA computation (using equations 12 and 13) for each and every earthquake source and logic-tree branch. These numerous CS results can then be combined using equations (14) and (15), but now with $j$ representing all the earthquake sources and $k$ representing all the logic-tree branches. In this case, the deaggregation weights in the equations are simply taken from the PSHA contribution (mean annual exceedance frequency) for each earthquake source and logic-tree branch, normalized by the total aggregated hazard.

An advantage of this aggregation approach to method 4 is that the other GMPSM input parameters $\theta$ no longer need to be inferred in calculating a CS. This is because we can first calculate a CS for each and every earthquake source (not to mention logic-tree branch) using the same corresponding GMPSM input parameters used in the PSHA computation. Based on the examples for three sites presented next, any numerical differences between the aggregation approach and method 4 are not expected to be significant for the CMS (or mean CS) at a site. Also based on the examples, however, the standard deviation of a CS (via equation 15) using the results for every earthquake source and logic-tree branch would be expected to yield differences, although the practical significance of the differences is not known. The standard deviation from the aggregation approach would capture all of the uncertainties considered in the PSHA computation.

An implementation of this aggregation approach to calculating an exact CS is now provided as part of USGS online hazard tools. Details and limitations of this implementation are discussed in the final section, Conditional Spectrum Calculation Tools from USGS.

**Example Calculations for Three Sites**

To demonstrate the numerical results that are now available using the previous equations and the USGS online hazard tools, this section provides a set of example calculations to determine whether using the “exact” method 4 provides results that have practical differences from those obtained using the simpler approximate methods 1 to 3.

To evaluate the accuracy of the approximate methods, CS for three locations are computed using the methods described previously. Target spectral accelerations are obtained for $Sa(0.2s)$ with 10% probability of exceedance in 50 years (i.e., a return period of 475 years), and for $Sa(1s)$ with 2% probability of exceedance in 50 years (i.e., a return period of 2475 years).

**Description of Example Sites and GMPSMs**

We consider three locations in the western United States with relatively high hazard but differing surrounding seismic sources: Stanford, Bissell, and Seattle. Deaggregation results for these three sites are shown in Figure 1. Ground-motion hazard at Stanford, located in northern California, is dominated by a single, shallow crustal earthquake source, the San Andreas fault zone. Ground-motion hazard at Bissell, located in southern California, has contributions from multiple earthquake sources, but they are all shallow crustal sources. Ground-motion hazard at Seattle has contributions from multiple earthquake sources of different types: shallow crustal, subduction zone interface, and intraplate (each of which has its own set of GMPSMs). All three sites are assumed to have a time-averaged shear-wave velocity in the top 30 meters of the soil ($V_{50}$) of 760 m/s.

The hazard calculations and deaggregation results all come from the USGS models and hazard mapping tools. The USGS model (Petersen et al., 2008) assigns equal logic-tree weights to GMPSMs near Stanford and Bissell, but unequal logic-tree weights to GMPSMs near Seattle. The USGS model uses three GMPSMs for the crustal sources near Stanford, Bissell, and Seattle (Boore and Atkinson, 2008; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008), three GMPSMs for the subduction zone interface sources (Youngs et al., 1997; Atkinson and Boore, 2003; Zhao et al., 2006), and three GMPSMs for the intraplate sources near Seattle (Youngs et al., 1997; Atkinson and Boore, 2003, which provides models that can be used in both the Global and Cascadia regions). The Youngs et al. (1997), Atkinson and Boore (2003), and Zhao et al. (2006) models are developed for both subduction zone interface and intraplate sources. The GMPSMs vary in their required input variables. The mean $\ln Sa$ predictions depend on $M$, $R$, and other source and site characteristics ($\theta$) such as depth to top of rupture, faulting mechanism, and hanging-wall effect. Some of the GMPSMs predict standard deviations of $\ln Sa$ that depend only on the period of interest (Boore and Atkinson, 2008), while others are magnitude dependent (Chiou and Youngs, 2008).

**Deaggregation Information**

The targets $Sa(0.2s)$ with 10% probability of exceedance in 50 years and $Sa(1s)$ with 2% probability of exceedance in 50 years are obtained from PSHA for the three example sites considered. The associated causal earthquake
magnitudes and distances are obtained from USGS deaggregation (see Data and Resources), and their deaggregation plots are shown in Figure 1. Other parameters, θ, that are associated with each causal earthquake M/R combination can be obtained directly from the parameters that were used for PSHA computation (i.e., via the aggregation approach to method 4), from deaggregation outputs if available, or inferred from the characteristics of contributing earthquake sources. With the target Sa(T*) and \( \mu_{ln,Sa} (M, R, \theta, T^*) \) and \( \sigma_{ln,Sa} (M, \theta, T^*) \) predictions for each causal M/R/θ, we can then back calculate ε(T*) using equation (1). The logic-tree weights for each GMPM (\( p_k^d \)) are obtained from the USGS documentation (Petersen et al., 2008), and the deaggregation weights (\( p_k^d/\mu_k \)) are obtained from the USGS deaggregation tools (see Data and Resources).

### Conditional Spectra Results

Conditional spectra results can be computed for each example case using the approximate or exact methods described previously. Depending on the level of approximation, the CS for a GMPM k can be computed using method 1 (equations 4 and 5) for approximate CS with mean M/R (M and R), method 3 (equations 8 and 9) for approximate CS with GMPM-specific mean M/R (M_k and R_k), or method 4 (equations 12 and 13) for all contributing CS with GMPM-specific causal earthquake M/R (M_j and R_j). The composite CS can then be computed with the corresponding weights using method 2 (equations 6 and 7) for approximate CS with GMPM logic-tree weights (\( p_k^d/\mu_k \)), method 3 (equations 10 and 11) for approximate CS with GMPM deaggregation weights (\( p_k^d/\mu_k \)), or method 4 (equations 14 and 15) for exact CS with deaggregation weights associated with each causal earthquake and GMPM (\( p_k^d/\mu_k \)). The resulting CS obtained from these four methods are plotted in Figure 2. Also plotted in the figure are the CMS (but not conditional standard deviations) resulting from the aggregation approach to method 4, which are nearly identical to the method 4 results despite the inclusion of additional logic-tree branches and earthquake sources. These CMS are obtained from the USGS calculation tools described in the final section.

We can make several observations from the results of Figure 2. For the Stanford and Bissell sites, CS computed using method 1 are very similar to results from methods 2 and 3, but they differ more for Seattle. This is because the GMPMs used in Seattle, some of which are for subduction zone events, result in more varied predictions; using a mean M/R that represents a variety of earthquake sources with a GMPM appropriate for a single-source type could be anticipated to produce these varied predictions. For Seattle, methods 2 and 3, which more carefully address the contributions of multiple GMPMs, do a better job of approximating the exact results from method 4.

For all three sites, the approximate methods 2 and 3 work better for conditional mean estimation than for conditional standard deviation estimation compared with the exact method 4, because method 4 produces higher conditional standard deviations in every case. A closer examination of equation (15) reveals two components of contribution to the exact conditional standard deviations: first, a contribution from \( \sigma_{ln,Sa} \), that is, variance in \( \ln Sa \) for a given \( M_j/R_j \) and GMPM_k; second, a contribution from \( \mu_{ln,Sa} \), due to variation in \( M_j/R_j \) and GMPM_k. In other words, the total variance comes from the expectation of the variance, which is the first term, and the variance of the expectation, which is the second term (Ditlevsen, 1981). The individual contributions from these two terms are plotted in Figure 3a-c. Methods 2 and 3 very closely approximate the contribution from \( \sigma_{ln,Sa} \) for the Stanford and Bissell sites, although the match is not as good for Seattle. Method 4 also includes the contribution from \( \mu_{ln,Sa} \), which is, however, not well captured by methods 2 and 3. Figure 3d-f shows the \( \mu_{ln,Sa} \) predicted for each \( M_j/R_j \) and GMPM_k, and it is the weighted variance of these spectra that creates the contribution to the overall standard deviation from \( \mu_{ln,Sa} \). Thus, we see that this variance of expectations contribution may not be negligible; the approximate methods do not capture this well, because they consider only mean M/R values and thus cannot identify the contribution to uncertainty from multiple M/R contributions to ground-motion hazard. Similarly, we anticipate that the deaggregation approach to the exact method 4 underestimates the conditional standard deviation, relative to the aggregation approach that accounts for every logic-tree branch and earthquake source used in the PSHA computation. We do not yet know whether this underestimation is practically significant.

Figure 4 shows the target CS computed using methods 2 and 4 for all three sites and both target \( Sa \) values: \( Sa(0.2s) \) with 10% probability of exceedance in 50 years and \( Sa(1s) \) with 2% probability of exceedance in 50 years. Method 2 is chosen as the reference approximate method here because it has an appealing combination of incorporating multiple GMPMs but not requiring any GMPM deaggregation information. For all three sites, the mean estimates of the approximate and exact methods are in close agreement, while the approximate standard deviations underestimate the exact result; this effect is most pronounced at periods far from the conditioning period (i.e., 0.2 s or 1 s).

The method 2 approximation appears to work best for sites with a single earthquake source (e.g., Stanford), followed by sites with multiple earthquakes sources of the same type (e.g., Bissell) and sites with multiple differing earthquake source types (e.g., Seattle). This is because there are several contributing factors to the accuracy of the approximation: (1) the input causal earthquake parameters, (2) the GMPMs used, (3) the GMPM deaggregation weights.

First, the importance of considering multiple causal M/R values depends upon how many M/R values contribute significantly to the hazard. In cases where all contributions to hazard come from a narrow range of magnitudes and distances (e.g., Stanford), the mean M/R is representative of the most important individual contributing M/R, so computations based only on the mean M/R are very precise.
However, in cases where hazard contributions come from a broader range of magnitudes and distances (especially in the case of Seattle), the mean $M/R$ deviates from any individual contributing $M/R$, so computations based on the mean $M/R$ only may result in a slight shift in conditional mean estimates in addition to reduced conditional standard deviations. Similarly, for the common situation of low- or moderate-seismicity sites in the central and eastern United States, for example, those that are located several hundred kilometers from the New Madrid seismic region where the low-frequency deaggregation can be strongly bimodal, computations based on the mean $M/R$ only may not be precise. Furthermore, variations in other parameters besides $M$ and $R$, such as depth to the top of rupture for different source types, can also affect the accuracy of the approximation because they contribute to the $S_a$ prediction.

Second, the similarity of the GMPMs affects approximations in their treatment. For cases where the ground-motion predictions for a given $M/R$ vary significantly between models (e.g., Seattle, where the subduction zone and crustal prediction models vary significantly), approximate treatment of GMPMs is less effective, and so a CS using a single GMPM (or using approximate weights on multiple models) may produce inaccurate results. Similar inaccuracies can be expected in the central and eastern United States. On the other hand, the three GMPMs used at Stanford and Bissell all tend to produce similar predictions, so for those cases the choice of the GMPM used to compute CS may not be as critical.

Figure 2. CMS at Stanford, Bissell, and Seattle (a, b, and c, respectively) and conditional standard deviation spectra at Stanford, Bissell, and Seattle (d, e, and f, respectively) using methods 1 to 4 for $S_a(0.2s)$ with 10% probability of exceedance in 50 years. The color version of this figure is available only in the electronic edition.
Third, as the GMPM deaggregation weights differ more from the GMPM logic-tree weights, approximate treatment of those weights works less effectively. For instance, in the (not too uncommon) case where one GMPM strongly dominates the deaggregation, the approximation that assumes GMPM logic-tree weights is expected to deviate more substantially from the computation that utilizes GMPM deaggregation weights. The GMPM deaggregation weights vary with the period of interest \( T \), the target \( Sa(T) \) amplitude of interest, and the location, so it is difficult to develop simple rules for when the approximations work well. But if the GMPM predictions are similar to each other for the \( M/R \) values contributing significantly to hazard, then the deaggregation weights are often similar to the logic-tree weights, and the predictions are also in good agreement (as noted in the previous paragraph), so the approximations are generally good in those cases.

### Impact of Approximations on Ground-Motion Selection

The importance of approximations in the CS computations will depend upon how the results affect any engineering decisions that may be made. The most common use for a CS is as a target response spectrum for ground-motion selection and scaling (Baker and Cornell, 2006; Baker, 2011).
Approximations to the CS might have an influence on ground motions selected from a database (because the selected ground motions match the target response spectrum), and that could affect nonlinear dynamic analysis results. On the other hand, the finite size of recorded ground-motion databases means that minor changes in the spectrum target may not result in substantially different ground motions being selected; in that case, the approximations would not have an appreciable impact on structural analysis results.

To illustrate, the target CS mean and variance are computed using both the exact (method 4) and approximate (method 2) approaches at the Bissell site for the $S_a(0.2\, s)$ with 10% probability of exceedance in 50 years. Ground motions can be selected to match these target spectrum mean and variance using the procedure of Jayaram et al. (2011), which assumes a Gaussian distribution. In general, the conditional logarithmic $S_a$ distribution is not Gaussian when multiple causal earthquakes and/or multiple GMPMs are considered, and hence the mean and variance alone may not describe the entire distribution; they still provide useful insights, however. Alternatively, Bradley (2010) considers multiple causal earthquake sources and the complete distribution of the conditional ground-motion intensity measure. The target response spectra and corresponding ground motions selected via Jayaram et al. (2011) are shown in Figures 4b and 5. The means of the CS using both methods are in close agreement, but the standard deviation of the CS using method 4 is higher than that using method 2, especially at periods farther away from the conditioning period of 0.2 s. Consequently, the spectra of the ground motions selected

Figure 4. CS computed using methods 2 and 4 for $S_a(0.2\, s)$ with 10% probability of exceedance in 50 years at Stanford, Bissell, and Seattle (a, b, and c, respectively) and $S_a(1\, s)$ with 2% probability of exceedance in 50 years at Stanford, Bissell, and Seattle (d, e, and f, respectively). The color version of this figure is available only in the electronic edition.
Individual CS can be constructed for each earthquake source, and reflect contributions from different earthquake sources, including those that are small and not significant in a separate set of ground motions selected to match both the approximate and exact CS, described previously and are provided by the tool, issues with the current implementation, and future features. One such future feature is providing the standard deviations of CS in addition to the means.

Description of Calculation Tool

As shown in Figure 6, the USGS 2008 Interactive Deaggregation web tool provides CMS for a user-specified:

- location (address or latitude and longitude) anywhere in the continental United States;
- mean exceedance probabilities of 1%, 2%, 5%, 10%, 20%, or 50% in 30, 50, 75, 100, or 200 years;
- spectral acceleration period of 0 s (corresponding to peak ground acceleration), 0.1 s, 0.2 s, 0.3 s, 0.5 s, 1 s, or 2 s, or additionally 3 s, 4 s, or 5 s for locations in the western United States (west of −115° longitude);
- Vₛ₃₀ of soil between 180 and 1300 m/s for locations in the central and eastern United States (east of −100° longitude), or of 760 m/s for locations in between.

As output, the tool currently provides graphs, tables, and text files for four different types of CMS, all calculated according to the Aggregation Approach to Method 4 (Exact CS) section. The four different types are:

- an overall CMS that accounts for all of the earthquake sources and logic-tree branches considered in the USGS PSHA computation (with a few exceptions described in the next section);
- a CMS for each GMPM that accounts for all of the earthquake sources and logic-tree branches related to the particular GMPM;
- a CMS for each M/R/ε bin that contribute most to the total aggregated hazard, accounting for all of the earthquake sources and logic-tree branches within that bin;
- a CMS for each M/R/ε bin and GMPM (currently only in text output), accounting for all of the earthquake sources and logic-tree branches within the bin that are related to the particular GMPM.

Conditional Spectrum Calculation Tools from USGS

The USGS National Seismic Hazard Mapping Project website now provides an option for CMS, as part of the 2008 Interactive Deaggregations web tool (see Data and Resources). The spectra are consistent with the 2008 national seismic hazard maps for the continental United States and are computed via the Aggregation Approach to Method 4 (Exact CS) section. Described next are the inputs to and outputs from the tool, issues with the current implementation, and future features. One such future feature is providing the standard deviations of CS in addition to the means.
The CMS for each GMPM demonstrates the effect of using only a single GMPM. As mentioned previously, this effect can be particularly significant in cases when more than one earthquake source type (e.g., subduction zone, shallow crustal) contributes significantly to the total aggregated hazard. In such cases, the CMS for two or more \( M/R/\varepsilon \) bins and/or GMPMs, each corresponding to a different source type, may be more useful than a single overall CMS.

Current Implementation Issues

While the web tool implementation described previously accounts for practically all of the earthquake sources and logic-tree branches considered in computing the USGS 2008 national seismic hazard maps, there are some exceptions. The least significant of these is that CMS calculations are not carried out for earthquake sources that do not contribute appreciably to the total aggregated hazard (i.e., those with mean annual exceedance frequency less than \( 10^{-6} \)). Also for the sake of limiting computation time, CMS are not calculated for the two USGS logic-tree branches that quantify additional epistemic uncertainty among the GMPMs for shallow crustal earthquakes in the western United States (see Petersen et al., 2008 for details). The logic-tree branches corresponding to the GMPMs themselves are fully accounted for, however. In the central and eastern United States, the USGS logic-tree branch for temporal clustering of New Madrid seismic zone earthquakes is not yet incorporated into the 2008 Interactive Deaggregations web tool, and hence is not accounted for in the CMS calculations. The numerical impacts of these exceptions on the calculated CMS have not yet been quantified, but they are anticipated to be relatively insignificant.

In being consistent with the 2008 national seismic hazard maps, the USGS web tool utilizes several different GMPMs for shallow crustal, subduction zone, and stable continental earthquake sources. However, for the correlation coefficients needed to calculate CS, the tool currently only uses the Baker and Jayaram (2008) model, which was developed with ground-motion data exclusively from shallow crustal earthquakes. A recent study of subduction zone ground motions from Japan suggests that this correlation model is also a reasonable representation for subduction zone earthquake sources (Jayaram et al., 2011). For stable continental earthquake sources, little data either confirm or contradict this model. In general, studies of correlation models have shown them to be relatively insensitive to the particular GMPM, earthquake magnitude, distance, and rupture mechanism (Baker, 2005; Jayaram et al., 2011). Thus, the numerical impact of using the single correlation model is anticipated to be relatively insignificant. Additional and/or updated correlation models corresponding to different earthquake sources and/or GMPMs could be incorporated into the web tool as they become available.

Future Features

As mentioned previously, in the future the USGS web tool will provide CS, not just CMS. The conditional standard deviations will likewise be calculated according to the aggregation approach to computing an exact CS. Whereas the current implementation for the conditional means has been seen to match results from the deaggregation approach of method 4, the standard deviations of a CS calculated by the web tool are anticipated to be higher and more inclusive of all the uncertainties accounted for in the USGS national seismic hazard maps.

Lastly, the current web tool calculates CMS with weights that are for exceedance of the \( Sa \) value specified by a user (via a selected mean exceedance probability). A future tool will provide weights for occurrence of an \( Sa \) value. In order to do so, the tool will optionally allow a user to specify an \( Sa \) value of interest, for example, a risk-targeted maximum considered earthquake (MCE) ground-motion value from the American Society of Civil Engineers (2010).
Conclusions

Approximate and exact computations of CS were proposed and used for example calculations for Stanford, Bissell, and Seattle. Exact CS mean and standard deviation calculations can incorporate multiple GMPMs and causal earthquake $M/R/\theta$ combinations, as well as multiple seismic source models and their logic-tree branches. Varying levels of approximations were also considered, that replaced multiple $M/R$ combinations with simply the mean $M/R$ from deaggregation, and either considered only a single GMPM or performed an approximate weighting of several GMPMs. These approximations are potentially appealing because of their ease of computation and because they do not require deaggregation of GMPM weights, a result that is not yet widely available in conventional PSHA software.

The approximate CS calculations appear to be more accurate for conditional mean estimation than for conditional standard deviation estimation. The exact conditional standard deviation is always higher than approximate results because of the additional contribution from the variance in mean logarithmic spectral accelerations due to variation in causal earthquakes and GMPMs. The input causal earthquake parameters and the GMPMs used along with the corresponding weights affect the accuracy of the approximation in CS computation. Exact calculation methods may be needed for locations with hazard contributions from multiple earthquake sources, where errors from approximations are higher as a result of multiple contributing earthquake magnitudes and distances, and variation in predictions from the input GMPMs.

The exact CS calculations require extension of traditional PSHA deaggregation, which considers only magnitude, distance, and $\epsilon$, to deaggregation of GMPMs. This additional deaggregation output is now available as part of the hazard results provided by the USGS. Further, the USGS now provides CMS results using the exact calculation aggregation approach described here. These new calculation tools should be useful in facilitating hazard-consistent ground-motion analysis and will allow for exact spectra to be computed without requiring cumbersome calculations by users.

Data and Resources

Deaggregation and CMS data used in this paper can be obtained from the USGS National Seismic Hazard Mapping Project website (http://earthquake.usgs.gov/hazards, last accessed June 2012), as part of the 2008 Interactive Deaggregations web tool at http://geohazards.usgs.gov/deaggint/2008/ (last accessed June 2012).

Acknowledgments

The authors thank Brendon Bradley, Gabriel Toro, Ned Field, Eduardo Miranda, and Ivan Wong for their helpful reviews of the paper. Also, thanks to Eric Martinez for his assistance on the USGS implementation. This work was supported by the USGS external grants program, under Awards 07HQAG0129 and 08HQAG0115. Any opinions, findings, conclusions, or recommendations presented in this material are those of the authors, and do not necessarily reflect those of the funding agency.

References


Campbell, K. W., and Y. Bozorgnia (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5%–damped linear elastic response spectra for periods ranging from 0.01 to 10 s, Earthquake Spectra 24, no. 1, 139–171.


Department of Civil and Environmental Engineering
Stanford University
Stanford, California 94305-4020
tinglin@stanford.edu
bakerjw@stanford.edu
(T.L., J.W.B.)

United States Geological Survey
Denver Federal Center
MS 966, Box 25046
Denver, Colorado 80225
harmesen@usgs.gov
nluco@usgs.gov
(S.C.H., N.L.)

Manuscript received 10 October 2011