Precision of Seismic Hazard Evaluations in Central and Eastern North America

R. K. McGuire
Lettis Consultants International, Inc., Boulder, Colorado

SUMMARY
For critical facilities such as nuclear power plants, multiple teams of experts are often used to provide inputs for seismic source characteristics, ground motion estimation, and site response. Alternative, weighted interpretations of these inputs leads to multiple estimates of annual frequencies of exceedance that characterize the mean, uncertainty, and range of seismic hazard. This study examines and quantifies the imprecision in past seismic hazard results from major studies in intraplate regions. Uncertainties in mean seismic hazard estimates are quantified using the statistical bootstrap technique applied to assigned weights of alternative inputs. Results are that the total coefficient of variation (COV) in mean seismic hazard is a minimum of 0.25 at a hazard of 10^{-4}, 0.30 at a hazard of 10^{-5}, and 0.35 at a hazard of 10^{-6}. Uncertainty in site response is a small fraction of total COV, so these results apply to both rock and soil sites.

Keywords: seismic hazard, precision, bootstrap

1. INTRODUCTION

This study describes the level of precision that can be associated with seismic hazard estimates, with specific application to central and eastern North America (CENA). By “level of precision” is meant how well constrained the calculated mean hazard is, with respect to epistemic uncertainty. This precision reflects how much the seismic hazard estimates might change, if the analysis were to be repeated with independent experts who have access to the same basic information (geology, tectonics, seismicity, ground motion equations, site characterization). If a dataset or interpretation were to change, and that change causes a change in the assessed seismic hazard at a site, the level of precision can be used to judge whether that change in hazard is significant or insignificant. Thus the question of significance is closely linked to the level of precision with which we can assess seismic hazard.

There are three fundamental sets of information that contribute to the precision of seismic hazard estimates, as follows:

1. Seismic sources and parameters, which may be derived by teams of experts or by individuals.
2. Ground motion equations, which are generally derived by a single expert or team using available equations, but sometimes are derived by multiple experts.
3. Site response estimates, which are generally derived by a single expert, but sometimes are derived by multiple experts.

A realistic assumption can be made that, for seismic hazard analysis at a site, these information inputs are separate and independent. It is understood that ground motion equations are developed for a wide range of magnitudes and distances, and that site response estimates are developed for a wide range of input motions. Additionally, it is assumed that we are interested in the precision of the *mean* seismic hazard curves, rather than any particular fractile. The mean seismic hazard curve is recommended to derive seismic design levels for facilities (McGuire, Cornell, and Toro, 2005).

Estimates of the precision in mean hazard (mean annual frequency of exceedance) associated with each of these inputs can be made by examining existing seismic hazard results from published studies. Examples are given below for peak ground acceleration (PGA) and 1 Hz spectral acceleration; results for other spectral frequencies are reported in (CEUS-HS 2012).

2. CONCEPT

The underlying concept in calculating precision estimates is to examine the epistemic uncertainty in hazard caused by team-to-team variations or expert-to-expert variations in hazard from documented studies. For example, if 6 teams are used to derive seismic sources for a hazard estimate, there will be a distribution of total hazard (i.e., annual frequency of exceedance) for a given ground motion amplitude. This distribution of total hazard will have a standard deviation σ_{TH} caused by team-to-team variability, and this standard deviation can be calculated using the conditional total hazard curves for each team. The uncertainty in *overall mean* hazard σ_{MH} caused by the different seismic source interpretations is $\sigma_{MH} = \sigma_{TH}/\sqrt{6}$, assuming the teams' hazard estimates are uncorrelated. We put aside questions of team-to-team correlation that result from common data sets, availability of published papers, and similar items, because this correlation is a condition under which we are evaluating the precision of hazard. Similar "independent" teams would have access to the same data sets and published papers.

Because the absolute value of hazard varies over several orders of magnitude, we use as a measure of precision the coefficient of variation (COV) of the mean hazard. The COV is the calculated standard deviation of mean hazard (σ_{MH}) divided by the mean hazard. When used in this sense, the coefficient of variation is designated COV_{MH}.

3. UNCERTAINTIES FROM AREAL SEISMIC SOURCES

Figure 1 shows the calculated COV_{MH} as a function of mean hazard for PGA at 7 test sites studied in (CEUS 2012), using seismic sources documented in (EPRI 1989). These COV_{MH} were calculated at the 7 test sites using only hazard from the six team interpretations of seismic sources and do not including hazard from the New Madrid and Charleston sources of large earthquakes. At some sites (e.g. Manchester), area sources dominate the hazard. At other sites (e.g. Savannah), the hazard is dominated by the potential of large, nearby earthquakes, because the site lies very close to one of these zones (the Charleston seismic zone, in the case of Savannah) and the area sources contribute relatively less hazard. COV_{MH} tends to increase with decreasing annual frequency; between 10^{-4} and 10^{-6} (the mean hazard range of interest) it ranges from about 0.15 to 0.4.

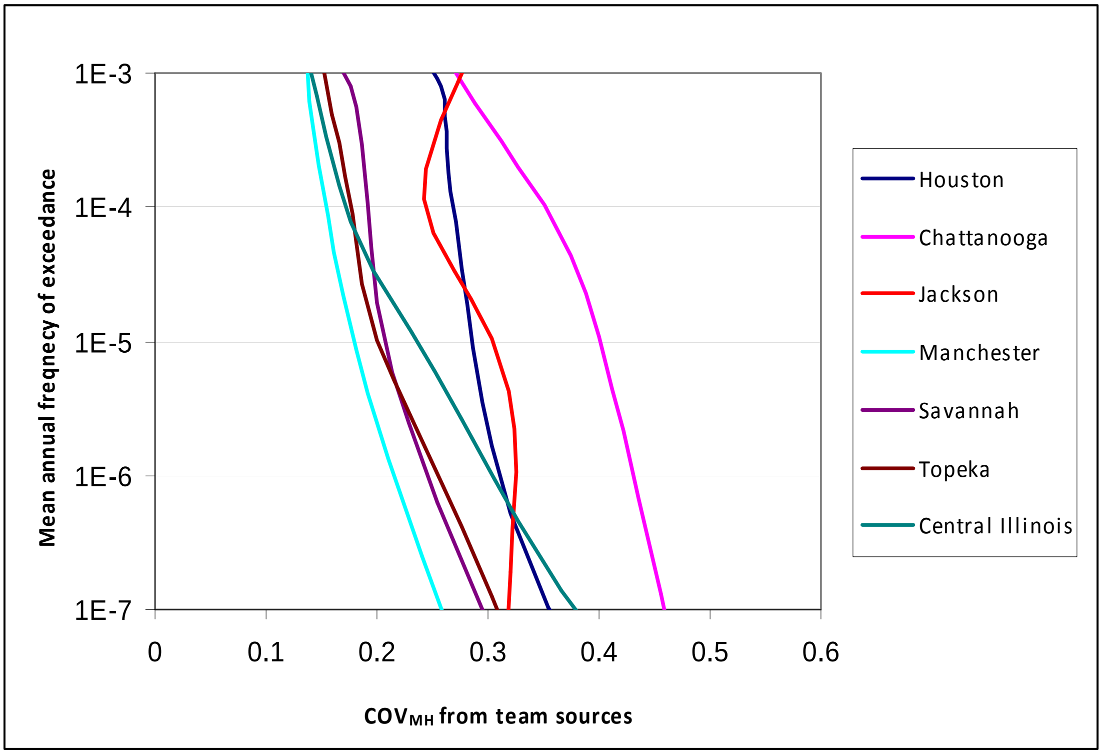

Figure 1. COV_{MH} from EPRI-1989 team sources vs. mean seismic hazard (i.e., mean annual frequency of exceedance) for seven test sites, for PGA.

Figure 2 shows COV_{MH} vs. mean hazard for PGA and 1 Hz spectral acceleration at four Swiss sites (i.e., Beznau, Goesgen, Liebstadt, and Muehlberg) studied during the PEGASOS project (NAGRA 2004). In this project, four experts developed areal seismic source interpretations, and Figure 2 plots COV_{MH}, calculated from the standard deviation of hazard σ_{MH} at each amplitude, as $\sigma_{MH} = \sigma_{TH}/\sqrt{4}$ because there were four experts who provided seismic source interpretations. For mean annual frequencies in the range 10^{-4} to 10^{-6}, COV_{MH} ranges from about 0.13 to 0.3, with one set of results (PGA for Goesgen) falling as low as 0.05 (see the solid blue curve in Figure 2).

The conclusion from Figures 1 and 2 regarding imprecision in seismic hazard estimates for area seismic sources is that typical COV_{MH} values will range from about 0.15 at a mean annual frequency of 10^{-4} to perhaps 0.3 at a mean annual frequency of 10^{-6}, with some sites (e.g. Chattanooga in Figure 1 and Goesgen PGA in Figure 2) falling outside that range.

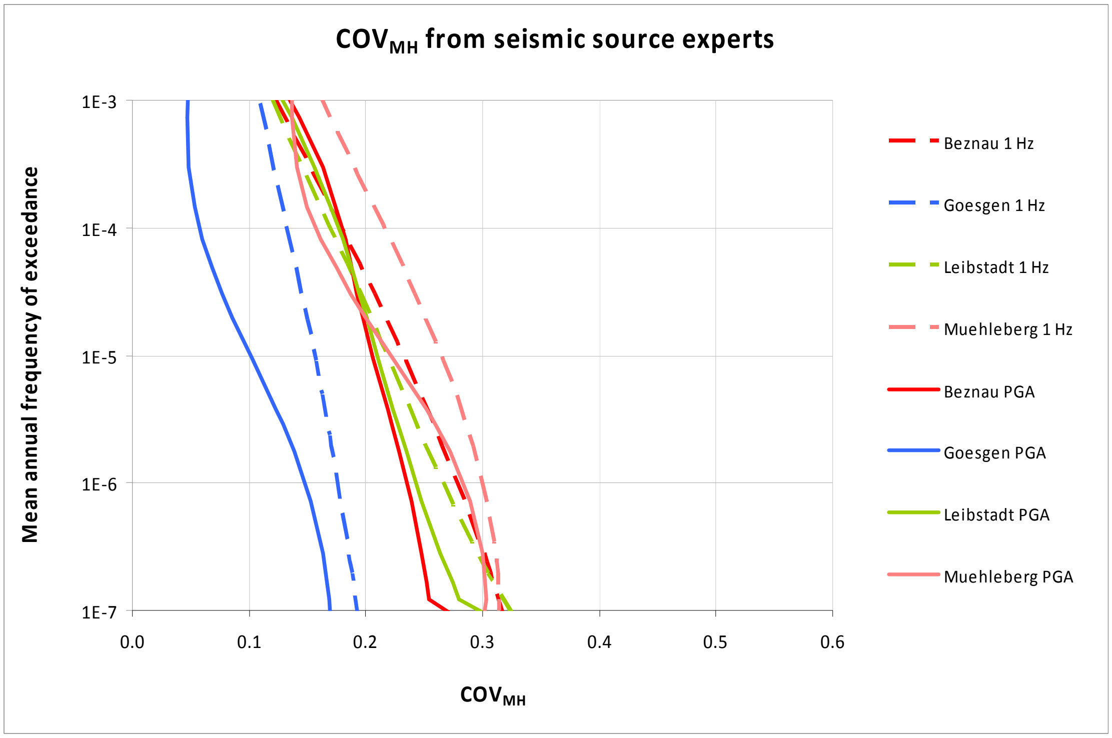

Figure 2. COV_{MH} from seismic source experts (PEGASOS project) vs. mean hazard for PGA and 1 Hz spectral acceleration.

4. UNCERTAINTIES FROM SOURCES OF LARGE EARTHQUAKES

For seismic hazard calculations in CENA, two sources of repeated large (**M** 7-7.5) earthquakes are the Charleston seismic zone and the New Madrid seismic zone. Herein this type of source is designated a "large EQ source." Nuclear plant seismic hazard studies have relied on two interpretations for these large EQ sources: The WLA model (Southern Nuclear Co, 2008) for the Charleston seismic zone and the Geomatrix model (Exelon Generation Co, 2003) for the New Madrid seismic zone. As an example, a general representation of the logic tree representing uncertainties in the Charleston seismic zone model is given in Table 4.1. For many sites in the southeastern US, seismic hazard is dominated by this source, rather than by local area sources. It is reasonable that there is uncertainty in the mean hazard coming from the source of large earthquakes, even though there is only one interpretation of epistemic uncertainty (e.g. Table 4.1).

Table 4.1. Summary of logic tree representing uncertainties for the Charleston seismic zone

Interpretation	**Alternatives**	**Weight on alternatives**	**Designation***
Geometry of source	4 geometries	0.7, 0.1, 0.1, 0.1	GEOM
Maximum magnitude	5 values	0.1, 0.25, 0.3, 0.25, 0.1	Mmax
Paleoseismic record length	2 periods	0.8, 0.2	SEIS
Activity rate given record	5 rates	0.1, 0.2, 0.4, 0.2, 0.1	RATE

*Designation of curves in Figures 3 and 4.

An independent evaluation by another investigator might assign somewhat different weights than those shown in Table 4.1. To determine what might be the effect of alternative weights, an adaptation of the statistical bootstrap technique (e. g. Efron, 1982) was used. This application has the underlying assumption that the weights given to alternative interpretations (e.g. in Table 4.1) are variables with distributions. It is reasonable that, to estimate a *minimum* variation on the weights given in Table 4.1, we should pick a COV_{WT} for the weights that corresponds to a change of 0.1 in the highest weight among the alternatives for each interpretation, because this is the precision with which weights were assigned. Designating this coefficient of variation COV_{WT}, we calculate the following values:

Source Geometry:	$COV_{WT} = 0.1/0.7 = 0.143$
Maximum magnitudes:	$COV_{WT} = 0.1/0.3 = 0.333$
Paleoseismic record length:	$COV_{WT} = 0.1/0.8 = 0.125$
Activity rate given record:	$COV_{WT} = 0.1/0.4 = 0.25$

The statistical bootstrap method consisted of generating random weights for the alternative interpretations given in Table 4.1, using the listed values as mean values and using the COV_{WT} given above to calculate standard deviations for the weights. A normal distribution for weights was assumed, truncated at 0 and 1. For each interpretation, the random weight for the alternative with the highest mean weight was generated first, and weights for the other alternatives followed. The values of these other weights are not independent, but depend on previously generated weights. In particular, they must sum to unity. Other dependencies are indicated in (CEUS-HS 2012).

The total mean hazard (annual frequency of exceedance) is the sum of weighted hazards from the available alternatives. For example, for the alternative geometries with 4 alternatives,

$$\text{mean}\ (H) = W_1\ H_1 + W_2\ H_2 + W_3\ H_3 + W_4\ H_4 \qquad (1)$$

where the H_i's are the mean hazard conditional on geometry i. In the current context, the H_i's are constant and the W_i's are random variables, so that,

$$\text{mean}\ (H) = \Sigma_i\ E[W_i]H_i \qquad (2)$$

(where E[.] indicates expectation) and

$$\sigma_k^2\ (H) = \Sigma\ \sigma_i^2\ H_i^2 + 2\ \Sigma_i\ \ \Sigma_{j>i}\ H_i\ H_j\ cov(W_i,\ W_j) \qquad (3)$$

where σ is standard deviation, *cov* is covariance, k indicates a specific interpretation from Table 4.1, and the σ_i's, H_i's, and W_i's are with respect to alternatives for that interpretation. The W_i's are correlated because, for example, a higher-than-mean value of W_1 will generally be associated with lower-than-mean values of the other W_i's, since they must sum to unity. The covariance of the W_i's can be estimated from samples generated using the bootstrap technique.

To calculate the total variance of the mean hazard (designated here as σ_{MH}^2), we assume that the contributions from the 4 alternatives in Table 4.1 are independent. This is an explicit assumption in the logic tree summarized in Table 4.1, e.g. the maximum magnitude alternatives and weights apply to all geometries. Under this independence assumption, COV_{MH} can be estimated as:

$$COV_{MH}^2 \simeq COV_{GEOM}^2 + COV_{Mmax}^2 + COV_{SEIS}^2 + COV_{RATE}^2 \qquad (4)$$

where Eq. (4) neglects cross-product terms involving the COVs that are small.

Figure 3 shows COV_K (where K represents GEOM, Mmax, SEIS, and RATE from Table 4.1) and COV_{MH} for PGA for the Charleston source, calculated for a site located at Columbia, South Carolina. It is evident that the alternative Mmax distribution dominates the uncertainty in mean hazard, except at low amplitudes (i.e., at high annual frequencies of exceedance).

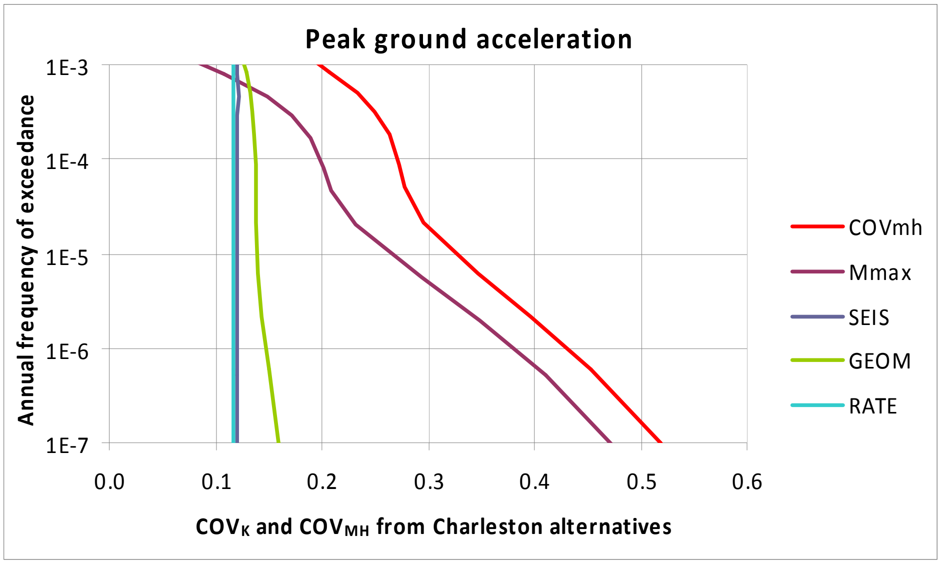

Figure 3. COV_K and COV_{MH} from Charleston alternatives for PGA, plotted vs. mean seismic hazard, for PGA at the Savannah site. COV_{MH} is the total COV of mean hazard, other labels for curves are given in Table 4.1.

Figure 4 shows a similar comparison of hazard sensitivity at the Jackson site to alternatives for the New Madrid seismic zone, which include Mmax, seismicity rate SEIS, and alternative geometries (indicated as "RFgeom," "NNgeom," and "NSgeom"). The latter are alternative geometries on the three faults in the New Madrid region, and a cluster model is used to calculate hazard.

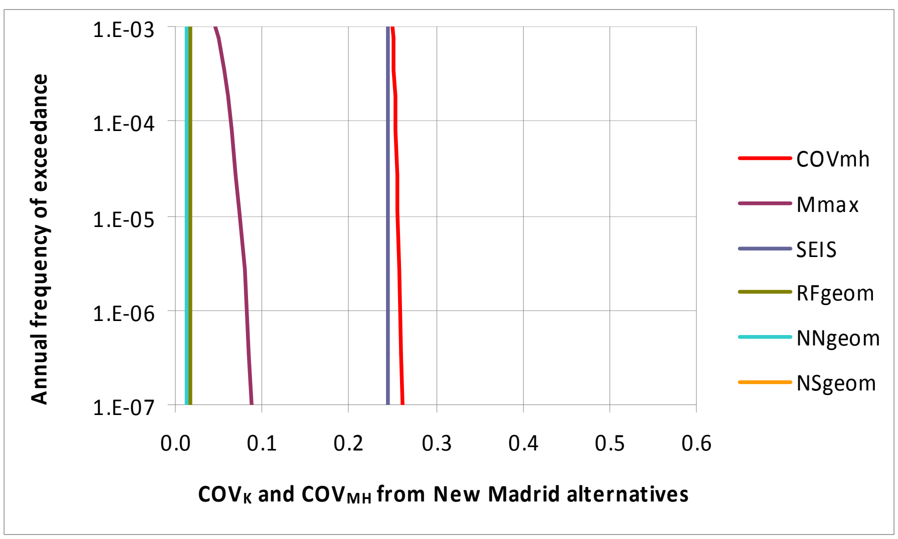

Figure 4. COV_K and COV_{MH} of total hazard from New Madrid for 1 Hz, plotted vs. mean hazard. COV_{MH} is the COV of mean hazard, see the text for other labels for curves.

Unlike the results for Charleston, the New Madrid model indicates that uncertainty in the rate of seismicity is the dominant contributor to uncertainty in hazard. The sensitivity to Mmax is low because, when one fault produces a high characteristic magnitude, other faults may produce a low characteristic magnitude during the cluster of earthquakes. COV_{MH} is about 0.25 for all amplitudes, and this result will be consistent across spectral frequencies because seismicity rate affects hazard equally across spectral frequencies.

5. UNCERTAINTIES FROM GROUND MOTION EQUATIONS

Several studies have results that can be used to estimate COV_{MH} that results from uncertainties in ground motion equations. EPRI (2005) and EPRI (2008) used ground motion equations published by EPRI (2004) for CENA, which consist of alternative equations with weights. The precision of hazard implied by these alternative models with weights was analyzed in a fashion similar to the Charleston seismic source, i.e. using an application of the statistical bootstrap technique. Weights given in EPRI (2004) for the various ground motion equations depend on whether ground motions come from an earthquake in an areal source or a large EQ source. Details of the bootstrap results are presented in (CEUS-HS 2012). Figure 5 summarizes a typical result from the Columbia site showing COV_{MH} vs. annual frequency of exceedance for PGA, for 3 values of the coefficient of variation on weights (COV_{WT}, with values of 0.3, 0.5, and 0.7) used in the bootstrap technique. COV_{WT}=0.5 is thought to be most representative, indicating COV_{MH} in the range of 0.15 to 0.3 for mean hazards in the range 10^{-4} to 10^{-6}.

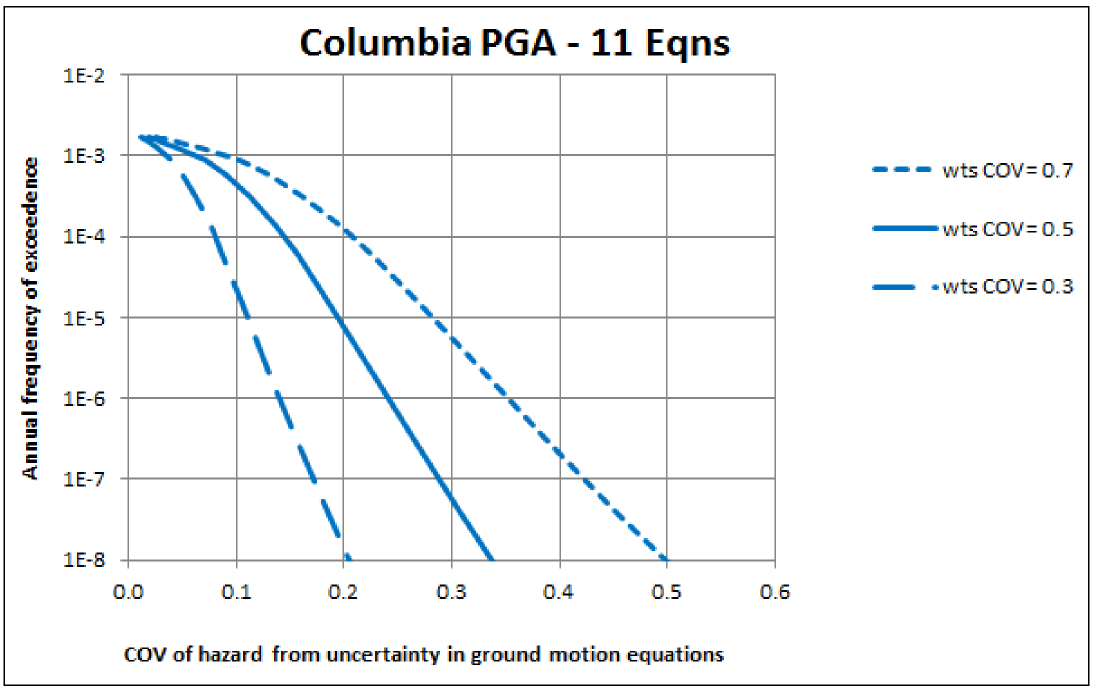

Figure 5. COV_{MH} vs. annual frequency of exceedance for the Columbia site from uncertainty in ground motion equation, for 3 values of COV_{WT}.

Hazard results that indicate hazard uncertainty from ground motion equations are also available from the PEGASOS study (NAGRA, 2004). In this study, five ground motion experts provided recommendations on sets of ground motion equations with weights, and hazard results are available at four Swiss nuclear power plant sites for PGA and 1 Hz SA conditional on each ground motion expert. The standard deviation of hazard σ_{MH} can be calculated for this set of conditional hazards, and COV_{MH} is taken as $\sigma_{MH}/\sqrt{5}$ divided by the overall mean hazard. Figure 6 shows COV_{MH} at the four sites, plotted vs. annual frequency of exceedance. For PGA the COV_{MH} exceeds 0.2, and for 1 Hz SA the COV_{MH} exceeds 0.3, for mean hazards in the range 10^{-4} to 10^{-6}. Note that these results do not include within-expert uncertainty, only uncertainty from expert-to-expert.

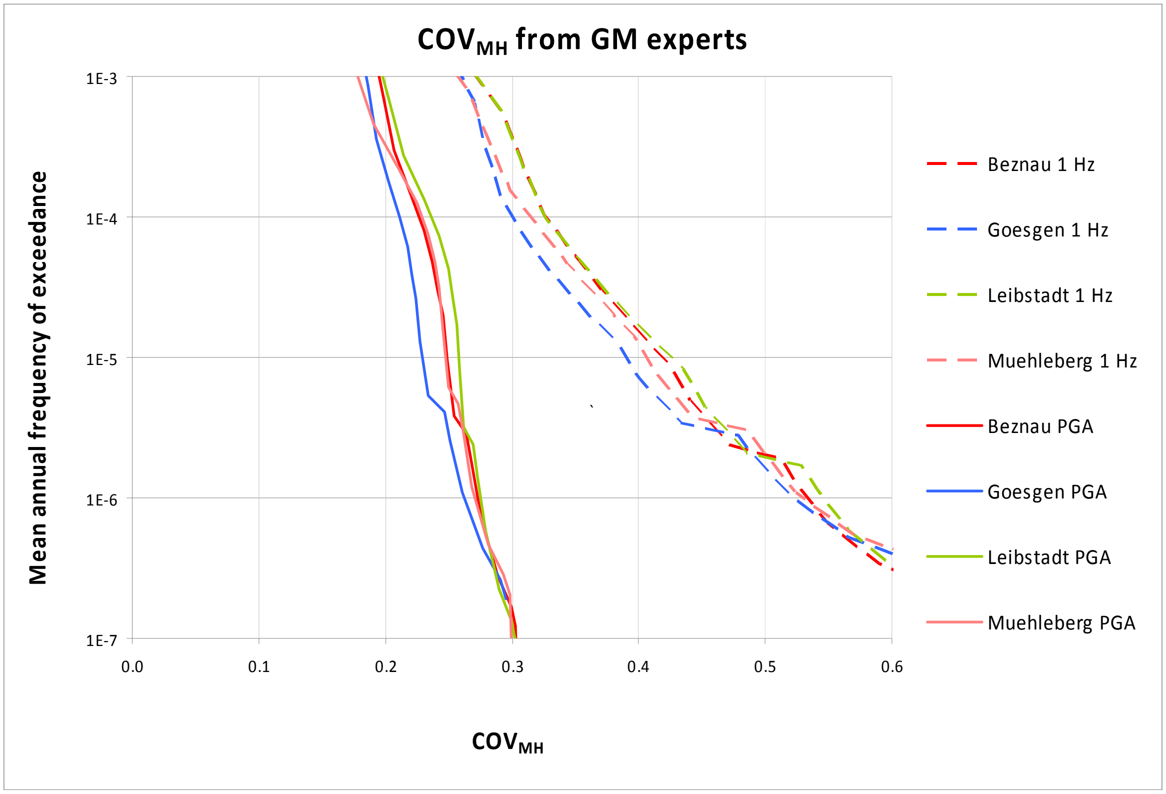

Figure 6. COV$_{MH}$ vs. annual frequency of exceedance for 4 sites from uncertainty in ground motion equation, from the PEGASOS study (NAGRA, 2004).

6. UNCERTAINTY IN SITE RESPONSE

Most sites in CENA are not classified as hard rock sites, and at these sites, uncertainty in site response plays a role in the uncertainty in site hazard calculations. Results from the PEGASOS project allow a direct estimate of the hazard uncertainty caused by uncertainty in site response calculations, because four site response experts provided recommendations on site response models, and hazard results are available at the 4 Swiss plant sites conditional on these 4 experts. The standard deviation of mean hazard σ_{MH} can be calculated for this set of conditional hazards, and COV$_{MH}$ is taken as $\sigma_{MH}/\sqrt{4}$ divided by the overall mean hazard. This analysis indicates that COV$_{MH}$ at the 4 sites for PGA and 1 Hz spectral acceleration (which are the only results available in this format) range from 0.05 for mean hazards of 10^{-4} to 0.4 for mean hazards of 10^{-6}. COV$_{MH}$ is generally lower for 1 Hz spectral acceleration than for PGA. The low estimate of COV$_{MH}$ for site response was also observed in the results of two EPRI-funded projects (2005, 2008) that calculated seismic hazard (including site response) at a group of nuclear power plants in CENA. Details are included in (CEUS-HS 2012).

7. CONCLUSIONS

Results are summarized in Table 7.1 as the minimum COV$_{MH}$ values observed in these sensitivity results. Some results were not used because they are not typical cases, or were down-weighted because the original source did not include within-expert variability. COV$_{MH}$ values are summarized by spectral frequency and annual frequency of exceedance, and results are given separately for area sources and large EQ sources. The last two columns represent the Total COV$_{MH}$, calculated as the square root of the sum of squares of the individual COV's for sites affected primarily by area sources and by large EQ sources.

Table 7.1. Minimum COV_{MH} values observed in seismic hazard from studies relevant to CENA

	Area sources	Large EQ sources	Ground motion (area sources)	Ground motion (large EQ sources)	Site response	Total COV_{MH}, site dominated by area sources	Total COV_{MH}, site dominated by large EQs
PGA, 1E-4	0.15	0.27	0.2	0.15	0.05	~0.25	~0.31
PGA, 1E-5	0.18	0.31	0.25	0.22	0.05	~0.31	~0.38
PGA, 1E-6	0.2	0.4	0.3	0.28	0.05	~0.36	~0.49
10 Hz, 1E-4	0.15	0.27	0.17	0.1	0.05	~0.23	~0.29
10 Hz, 1E-5	0.18	0.31	0.25	0.13	0.05	~0.31	~0.34
10 Hz, 1E-6	0.21	0.4	0.37	0.16	0.05	~0.43	~0.43
1 Hz, 1E-4	0.1	0.25	0.3	0.12	0.05	~0.32	~0.28
1 Hz, 1E-5	0.1	0.3	0.4	0.18	0.05	~0.42	~0.35
1 Hz, 1E-6	0.1	0.35	0.5	0.23	0.05	~0.51	~0.42

In general, Table 7.1 shows that minimum hazard uncertainties resulting from area source characteristics are smaller than minimum hazard uncertainties resulting from large EQ source characteristics. However, the reverse is true of uncertainties resulting from ground motion models, where minimum hazard uncertainties from area source ground motion models are larger than from large EQ ground motion models. These two effects compensate somewhat, so that total minimum uncertainties in hazard are comparable for the two types of sources. Uncertainty in site response contributes relatively little, at least for the example sites presented here from two major studies. As an overall conclusion, the minimum COV representing uncertainty in mean hazard over all spectral frequencies, and for annual mean hazards in the range 10^{-4} to 10^{-6}, can be taken to be about 0.25 for 10^{-4}, 0.3 for 10^{-5}, and 0.35 for 10^{-6}. Because the contribution of site response uncertainty is a small part of this total, this conclusion applies to both rock and soil sites.

For decisions regarding the significance of changes in seismic hazard, the above results should be interpreted as follows. If an alternative assumption or parameter is used in a seismic hazard study, and it potentially changes the calculated mean hazard (mean annual frequency of exceedance) by less than $\pm$25% for ground motions corresponding to 10^{-4} annual frequency of exceedance, and potentially changes the calculated hazard by less than $\pm$35% for ground motions corresponding to 10^{-6} annual frequency of exceedance, that potential change is less than the best (highest) level of precision with which we can calculate mean seismic hazard. Under these circumstances, the potential change could be deemed not significant. For many sites we cannot be this precise, and the uncertainty in mean hazard will be higher than this, but the above interpretation gives a reasonable lower-bound guideline with which to evaluate the significance of potential changes in mean hazard.

ACKNOWLEDGEMENT

This study was supported by a project funded by the U.S. Nuclear Regulatory Commission, the U.S. Department of Energy, and the Electric Power Research Institute. This support is gratefully acknowledged. No endorsement of these results is implied herein by these agencies, the conclusions are solely those of the author.

REFERENCES

CEUS (2012). *Central and Eastern United States Seismic Source Characterization for Nuclear Facilities*, Report published by NRC Report **NUREG-2115**, DOE Report **NE-0140**, EPRI Report **1021097**, 6 Volumes.

CEUS-HS (2012). *Hazard Significance,* Section 9.4 of *Central and Eastern United States Seismic Source Characterization for Nuclear Facilities*, Report published by NRC Report **NUREG-2115**, DOE Report **NE-0140**, EPRI Report **1021097**, Vol. 3, pp. 9-10 to 9-20.

Efron, Bradley (1982). *The jackknife, the bootstrap, and other resampling plans*, Soc. for Industrial and Applied Mathematics, Phil, PA, 92p.

EPRI (1989). Probabilistic Seismic Hazard Evaluations at Nuclear Plant Sites in the Central and Eastern United States: Resolution of the Charleston Earthquake Issue*, Elec. Power Res. Inst.*, **Rept. NP-6395-D**, Palo Alto, CA, Apr.

EPRI (2004). CEUS Ground Motion Project Final Report, *Elec. Power Res. Inst.,* **Rept. 1009684**, Palo Alto, CA, Dec.

EPRI (2005). Program on Technology Innovation: Assessment of a Performance-Based Approach for Determining Seismic Ground Motions for New Plant Sites, V1 and V2, *Elec. Power Res. Inst.,* **Rept. 1012044 and 1012045**, August.

EPRI (2008). Assessment of Seismic Hazard at 34 U.S. Nuclear Plant Sites, *Elec. Power Res. Inst.,* **Rept. 1016736,** August.

Exelon (2003). Submittal of Exelon Generation Company (EGC) application for an early site permit (ESP) for property co-located with existing Clinton Power Station (CPS) facility in Illinois, *US Nuc. Reg. Comm.* document accession no. **ML032721596**, Sept. 25.

McGuire, R.K., Cornell, C.A., and Toro, G.R. (2005), The Case for Using Mean Seismic Hazard, *Earthquake Spectra*, **21:3**, 879–886.

NAGRA (2004). Probabilistic Seismic Hazard Analysis for Swiss Nuclear Power Plant Sites (PEGASOS Project)*, Final Report to Unterausschuss Kernenergie (UAK) der Ueberlandwerke (UeW),* July 31.

Southern Nuclear Co (2008). Vogtle Early Site Permit Application, Revision 4, Part 2 – Site Safety Analysis Report, Chapter 2, “Site Characteristics,” see Updated Charleston Seismic Source (UCSS) model in Section 2.5.2.2.2.4, *US Nuc. Reg. Comm.* document accession no. **ML081020220**, March 28.