Application of Single-Station Sigma Ground Motion Prediction Equations in Practice

15 WCEE
LISBOA 2012

N.A. Abrahamson
Pacific Gas & Electric Company, San Francisco, CA

J.C. Hollenback
University of California Berkeley, United States

SUMMARY:
Traditionally, ground motion prediction equations have been built based on the ergodic assumption. Currently, there is a trend towards removing parts of the ergodic assumption from GMPEs. Application of partially ergodic GMPEs requires an estimate of the site term and quantification of its epistemic uncertainty. We provide guidance on how to apply these updated GMPEs in practice. An example application for Diablo Canyon Power Plant (DCPP) site is presented. Ground motion residuals can be used to estimate site terms. Two earthquakes observed at DCPP site, San Simeon M_w = 6.5 and Parkfield M_w = 6.0, were used to estimate site terms and associated epistemic uncertainty. Estimated site terms, in natural log units, range approximately from 0.55 at low frequencies to -0.24 at high frequencies. The standard deviation of epistemic uncertainty in site term estimates range from 0.2 - 0.23. The net effect is a significant change in the mean hazard.

Keywords: Seismic Hazard, Single-Station Sigma, Ground Motion Prediction

1. INTRODUCTION

Traditionally, empirically based ground motion prediction equations (GMPEs or ground motion models) have been built based on the ergodic assumption. An ergodic process is a random process whose distribution across space is identical to its distribution across time for a fixed location. To assume earthquake ground motions are an ergodic process is to assume that the spatial variability of ground motion at many sites is identical to the variability of ground motion at a specific site across time (Anderson and Brune 1999). The Next Generation Attenuation (NGA) Relationships (2008), for example, utilize the ergodic assumption. They are built with global data sets where ground motion data comes from many events, located in several different regions (e.g. Japan, Taiwan, Turkey, and California), where all events were recorded at multiple sites. In the development and application of these relationships the variability of ground motion across all different regions, wave propagation paths, and sites is assumed to represent the variability of future ground motion observations at a single site. Currently, there is a trend towards removing parts of the ergodic assumption from ground motion models. One variety of these partially non-ergodic models is known as a single-station sigma model. This paper examines the implications of the ergodic assumption, explains single-station sigma models, provides guidance on proper application of single-station sigma ground motion models, and demonstrates proper application with an example problem.

The initial efforts to remove parts of the ergodic assumption have focused on the contribution site response has on the ground motion model standard deviation (e.g. Atkinson 2006, Rodriguez-Marek et al 2011, Lin et al 2011). These new models are referred to as single-station sigma models and have a standard deviation that is more representative of the variability of ground motion observed at a single site. Table 1.1 lists some published values of single-station sigma for PGA.

Table 1.1. Published Values of Single-Station Sigma for PGA

Study	Rodriguez-Marek et al (2011)	Atkinson (2006)	Lin et al (2011)	Chen and Tsai (2002)
Ergodic Standard Deviation	0.799	0.711	0.680	0.731
Single-Station Sigma	0.672	0.617	0.619	0.631
% Reduction	16%	13%	9%	14%

1.1. The Ergodic Assumption & Single-Station Sigma

In general, the ergodic assumption is made when developing empirically based ground motion models to compensate for a lack of data. Anderson and Brune (1999) note that ergodic assumption for ground motion is incorrect. It is important to understand the two types of uncertainty considered in seismic hazard analysis, aleatory variability and epistemic uncertainty, to understand why the ergodic assumption is incorrect,. According to Toro et al. (1997), aleatory variability is "inherent to the unpredictable nature of future events" and "cannot be reduced by collection of additional information". They define epistemic uncertainty as "uncertainty that is due to incomplete knowledge and data about the physics of the earthquake process. In principle, epistemic uncertainty can be reduced by the collection of additional information." The main difference being aleatory variability cannot be reduced while epistemic uncertainty, in theory, can be reduced to zero with sufficient knowledge.

Empirical ground motion models developed using random effects or two-step regression methods will take the form:

$$\ln(Y_{es}) = f(X,\Theta) + \delta B_e + \delta W_{es} \tag{1.1}$$

where Y_{es} is the predicted ground motion parameter, subscripts *e* and *s* specify some given event and site respectively, $f(X,\Theta)$ is the ground motion model (median estimate of the natural log of the ground motion parameter Y), X is a vector of explanatory parameters (e.g. magnitude, distance, site, etc.), Θ is a vector of model coefficients, δB_e is a random variable that represents the between-event variability and δW_{es} is a random variable that represents the within-event variability. Between-event variability, sometimes referred to as inter-event variability, describes how ground motions generated from one event can be systematically higher or lower than model prediction, $f(X,\Theta)$, across all sites. Within-event variability, sometimes referred to as intra-event variability, describes how the ground motion observed at a site could be higher or lower than the event corrected model prediction. The standard deviations of δB_e and δW_{es} are τ and ϕ respectively. These terms are independent so the total standard deviation, σ, is given by.

$$\sigma = \sqrt{\tau^2 + \phi^2} \tag{1.2}$$

Embedded in a ground motion model is a median estimate of site response. This is based on a subset of the explanatory parameters, X_{site}, that describe the site (e.g. site class, depth to bed rock, etc.). The Abrahamson and Silva (2008) NGA model, for example, uses analytically based amplification functions to estimate site response based on a site's V_{S30}. Site response is a difficult phenomenon to predict, especially with simple explanatory parameters such as V_{S30}. As such, ground motion models will not be able to predict site response perfectly for any one site. Additionally, a group of sites that have the same X_{site} will not have identical site response, which is referred to as site-to-site variability.

When the ergodic assumption is made, site-to-site variability exists in any data set used to make a ground motion model, and contributes to the variability of the δW_{es} (i.e. the within-event aleatory variability described by the model). However, allowing site-to-site variability (variability across space) to contribute to σ violates the ergodic assumption because it does not describe the variability of ground motion at a given site for future ground motion observations. Rather, it describes the uncertainty in predicting site response when using simplistic models, such as those in ground motion models. Site response has systematic and repeatable effects on the ground motion observed at the ground surface. With increased knowledge such as a site's shear-wave velocity profile with depth and/or ground motion observations from past events, the uncertainty in the prediction of site response can be reduced. When the ergodic assumption is applied to ground motions both epistemic uncertainty, and aleatory variability are contributing to the standard deviation, σ, of a model. Ideally, the standard deviation should only represent aleatory variability of a ground motion model and any epistemic uncertainty should be captured using a logic tree approach. Aleatory variability has a great influence on predicted mean hazard (Abrahamson and Bommer 2006). Thus, failing to correctly partition these two types of uncertainty can lead to errors in the estimation of hazard.

Single-station sigma models attempt to remove the epistemic uncertainty associated with site response from the ground motion model standard deviation, σ. Writing Eqn. 1.1 more explicitly for site response we get:

$$\ln(Y_{es}) = f\left(X_{rock},\Theta\right) + \mathrm{Amp}\left(V_{S30}, f\left(X_{rock},\Theta\right)\right) + \delta B_e + \delta W_{es} \quad (1.3)$$

where $\mathrm{Amp}(V_{s30}, f(X_{rock},\Theta))$ is the amplification of the ground motion from the reference rock condition. The site response at a given location has a systematic, repeatable, and potentially knowable effect on the ground motion observed at a site. Because the amplification function doesn't model site response perfectly, embedded in the within-event residual, δW_{es}, is the systematic and repeatable effect site response has on the ground motion observation. This effect can be removed from the within-event residual as:

$$\ln(Y_{es}) = f\left(X_{rock},\Theta\right) + \mathrm{Amp}\left(V_{S30}, f\left(X_{rock},\Theta\right)\right) + \delta B_e + \delta WS_{es} + \widehat{\mathrm{S}}_s \quad (1.4)$$

where δWS_{es} is the single-station within-event residual and $\widehat{\mathrm{S}}_s$ is the site-specific site term given by:

$$\widehat{\mathrm{S}}_s = \frac{1}{NE_s}\sum_{e=1}^{NE_s} \delta W_{es} \quad (1.5)$$

Conceptually, the $\widehat{\mathrm{S}}_s$ represents the average difference in site response at site s from the model estimate, $\mathrm{Amp}(V_{s30}, f(X_{rock},\Theta))$, and has an uncertainty of ϕ_{S2S}, which is the quantification of the site-to-site variability (alluded to above). NE_s is the number of ground motion observations at site s. The single-station within-event residual, δWS_{es}, is what remains when $\widehat{\mathrm{S}}_s$ is removed from δW_{es} and is a random variable with a mean of zero and a standard deviation of ϕ_{ss}. The uncertainty in $\widehat{\mathrm{S}}_s$ can be reduced with information (e.g. ground motion observations or analytical site specific model). The site-to-site variability of the $\widehat{\mathrm{S}}_s$ term, ϕ_{S2S}, is related to ϕ_{ss} as below.

$$\phi_{S2S} = \sqrt{\phi^2 - \phi_{ss}^2} \quad (1.6)$$

The total standard deviation of a ground motion model that has the site-to-site variability removed is:

$$\sigma_{ss} = \sqrt{\tau^2 + \phi_{ss}^2} \quad (1.7)$$

This is known as single-station sigma. The site-to-site variability is not included because it is the epistemic uncertainty in the site-specific site term, $\widehat{\mathrm{S}}_s$. Single-station sigma models are only partially non-ergodic models. They have removed one component of the total ground motion variability that is not representative of the variability of future observations of ground motion at a single site from the model standard deviation. To get a fully non-ergodic model all of the components of the total ground motion variability that are not representative of the variability of future observations of ground motion at a single site must be removed. This paper focuses on single-station sigma models, where site response is the only systematic and repeatable effect removed from the standard deviation of the ground motion model. For a more detailed explanation of all the constituent parts of ground motion variability refer to Al Atik et al. (2010). Single-station sigma is a modest reduction compared with the fully ergodic ground motion standard deviation (10-15%); however, this reduction can have large effects on hazard at low probability levels. The median site-specific site terms, $\widehat{S}_s$, can also have a significant effect on the hazard.

2. APPLICATION OF SINGLE-STATION SIGMA MODELS

From a seismic hazard perspective, single-station sigma models are a welcome step forward. Purvance et al. (2008) point out examples of the over-prediction of ground motion estimates from a probabilistic seismic hazard analysis when compared to toppling accelerations of precariously balanced rocks. One explanation they offer is that the variability of ergodic ground motion models, σ, is too large, which leads to estimates of acceleration that are unrealistically high at low probability levels. Single-station sigma models are the first attempt at reducing the model variability to a more realistic level; however, the reduced variability cannot be used without the penalty of additional epistemic uncertainty in the site terms. There are two additional steps required to use single-station sigma models in practice: estimating the site term, $\widehat{S}_s$ for a site of interest, and quantifying the epistemic uncertainty in this estimate. In general there are two ways to estimate the site term. The first utilizes ground motion observations at a site of interest. The second involves using more sophisticated analytical site response models that are based on more descriptive data than that of the ground motion model (e.g. shear-wave velocity profile down to bedrock as apposed to site class or Vs_{30}). This paper will focus on the former method.

As shown in Eqn. 1.5, the site term, $\widehat{S}_s$, is the mean of the within-event residuals at a given site, δW_{es}. The within-event residual is given by:

$$\delta W_{es} = \ln(Y_{es}) - f(X,\Theta) - \delta B_e \qquad (2.1)$$

To quantify δW_{es} we need all three terms on the right hand side of Eqn. 2.1. One important issue is that the distance scaling in the data should be consistent with the distance scaling in the model so that the event term, δB_e, is applicable to the site, otherwise some path effects may be mapped into the site term.

The epistemic uncertainty (standard error of the mean) in the estimate of $\widehat{S}_s$ is:

$$SE\left[\widehat{S}_s\right] = \frac{\phi_{S2S}}{\sqrt{NE_s}} \qquad (2.2)$$

If there are no ground motion observations at site *s*, then the best estimate of $\widehat{S}_s$ is zero and the standard error is ϕ_{S2S}. If single-station sigma is used, then a branch must be added to the logic tree in a probabilistic seismic hazard analysis to capture the epistemic uncertainty in the $\widehat{S}_s$ term.

2.1. An example application of single-station sigma models

Here, an example application is provide to help illustrate how to use single-station sigma models in probabilistic seismic hazard analysis. The example site is Diablo Canyon Power Plant (DCPP) site, which is located on the central coast of California half way between Los Angeles and the San Francisco Bay Area.

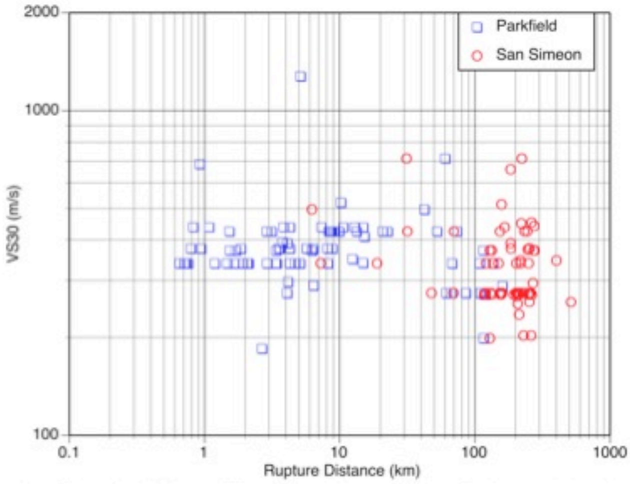

Figure 2.1. Distribution of Rupture distance and Vs_{30} for San Simeon and Parkfield events

DCPP has a free-field strong motion recording station on site. Two well-recorded earthquakes have been observed at this station: the 2003 San Simeon and 2004 Parkfield earthquakes. The distribution of rupture distances and Vs_{30} values of all sites that recorded these two earthquakes can be seen in Fig. 2.1. These events will be used to estimate the site term, $\widehat{S}_s$, for DCPP and to quantify the epistemic uncertainty in this estimate. DCPP is located at a rupture distance of 35km and 85km for the San Simeon and Parkfield events, respectively.

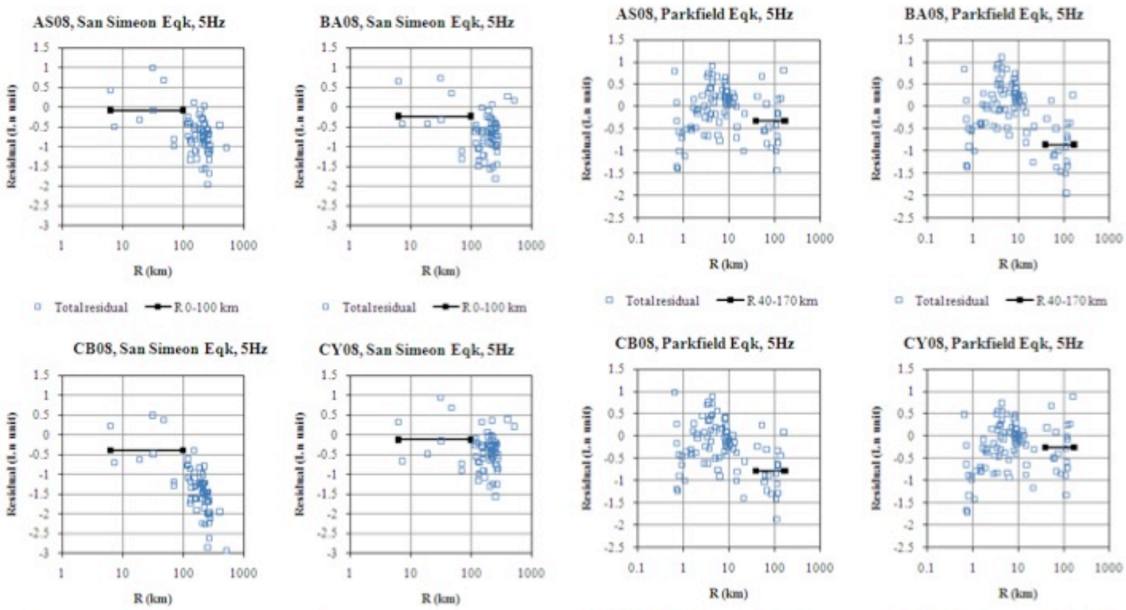

Figure 2.2. Total residuals for Sa(5hz) plotted against distance for San Simeon (left) and Parkfield (right) events

The site term was estimated using four out of the five NGA models. The Idriss model was omitted because the majority of sites that recorded these two events had Vs_{30} values of less than 450m/s and the Idriss model is for sites with Vs_{30}>450m/s so it could not be used to estimate the event terms. The first step is to estimate the event terms. The total residual, δ_{es}, is computed for all 4 NGA model used. Fig.2.2 shows the total residuals plotted against distance for 5Hz spectral acceleration, both the San Simeon and Parkfield events, and all 4 NGA models. For both events, there is a trend in the residuals with distance indicating that the large distance scaling in the NGA models is not applicable to these two earthquakes. This distance dependence is not related to the site term of interest. Therefore, we remove the effect of the distance trend on the event term by estimating the event term using a subset of the data near the distance to DCPP for each event.

For the San Simeon earthquake, the event term is taken as the average of the total residuals for a distance bandwidth of 0 to 100km. This distance bandwidth was selected because the rupture distance for this event was 35km and the residuals are nearly flat over 0 to 100 km. The black lines on the plots in Fig. 2.2 show the estimate of the event term for each respective model. The event terms for 5Hz spectral acceleration for San Simeon range from 0 to -0.5 ln units depending on the model.

For the Parkfield earthquake, the event term is taken as the average of the total residuals for a distance bandwidth of 40-170km. This distance bandwidth was selected because the rupture distance for this event was 85km and the residuals are nearly flat over the range of 40-170 km. The event terms for 5Hz spectral acceleration for Parkfield range from 0 to -1.0 ln units. This process was repeated for a sweet of spectral frequencies.

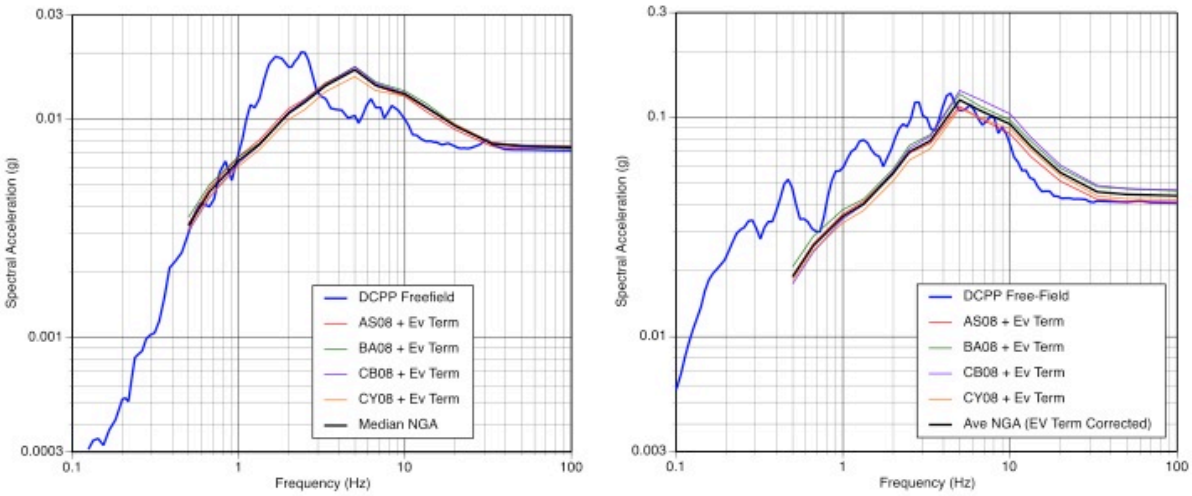

Figure 2.3. Comparison of response spectra of event corrected median NGA models with DCPP observation, San Simeon (left) Parkfield (right)

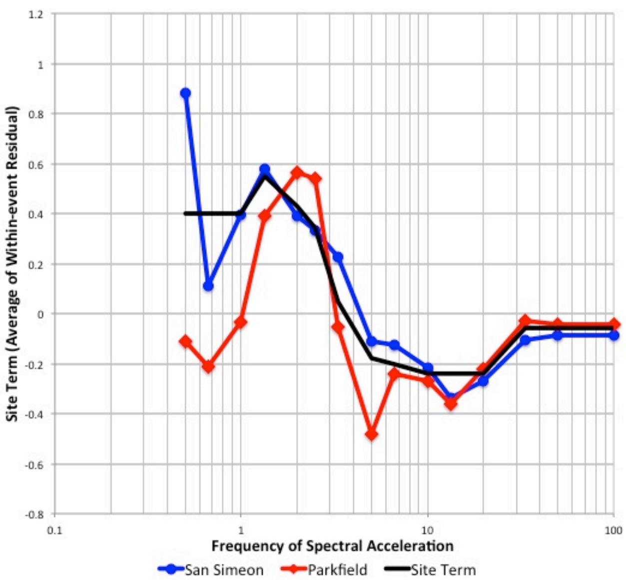

Figure 2.4. Model for site term at DCPP as a function of frequency of spectral acceleration

The GMPE-specific event term is then added back to the median ground motion model for each of the 4 NGA relationships and compared with the observed ground motion. This comparison is shown in

Fig. 2.3, both San Simeon and Parkfield. Eqn. 2.1 is used to calculate the within-event residual, δW_{es}, for each event. The average of δW_{es} for all 4 of the NGA models is plotted in Fig. 2.4 as a function of frequency for each event. This represents the average difference in site response at DCPP from the median estimate of the 4 NGA models. The site term for DCPP site at 5Hz spectral acceleration is estimated to be -0.18 ln units. Fig. 2.4 shows that the site term is a robust feature because the trend is the same for both events even though the two events had very different spectral shapes (as shown in Fig. 2.3). It also demonstrates that site response has a systematic, repeatable and potentially knowable effect on the ground motion observed at a site. The model for site term at DCPP suggests that this site amplifies low frequency motions, and reduces high frequencies.

With the estimates of the site terms, $\widehat{S}_s(f)$,for DCPP, the hazard can be computed using the single-station sigma in place of the traditional ergodic sigma. Here, we use the single-station sigma model developed for application to the NGA models (BCHydro, 2010):

$$\sigma_{ss} = (\text{T,M}) = \left(0.87 + 0.0037\ln(\text{T})\right)\sigma(\text{T,M}) \tag{2.3}$$

To capture the uncertainty in the site term estimate, we use ϕ_{S2S} = 0.30 which leads to an epistemic uncertainty in site term estimate of 0.21 natural log units. This is then used as input to the logic tree for the uncertainty in the ground motion model.

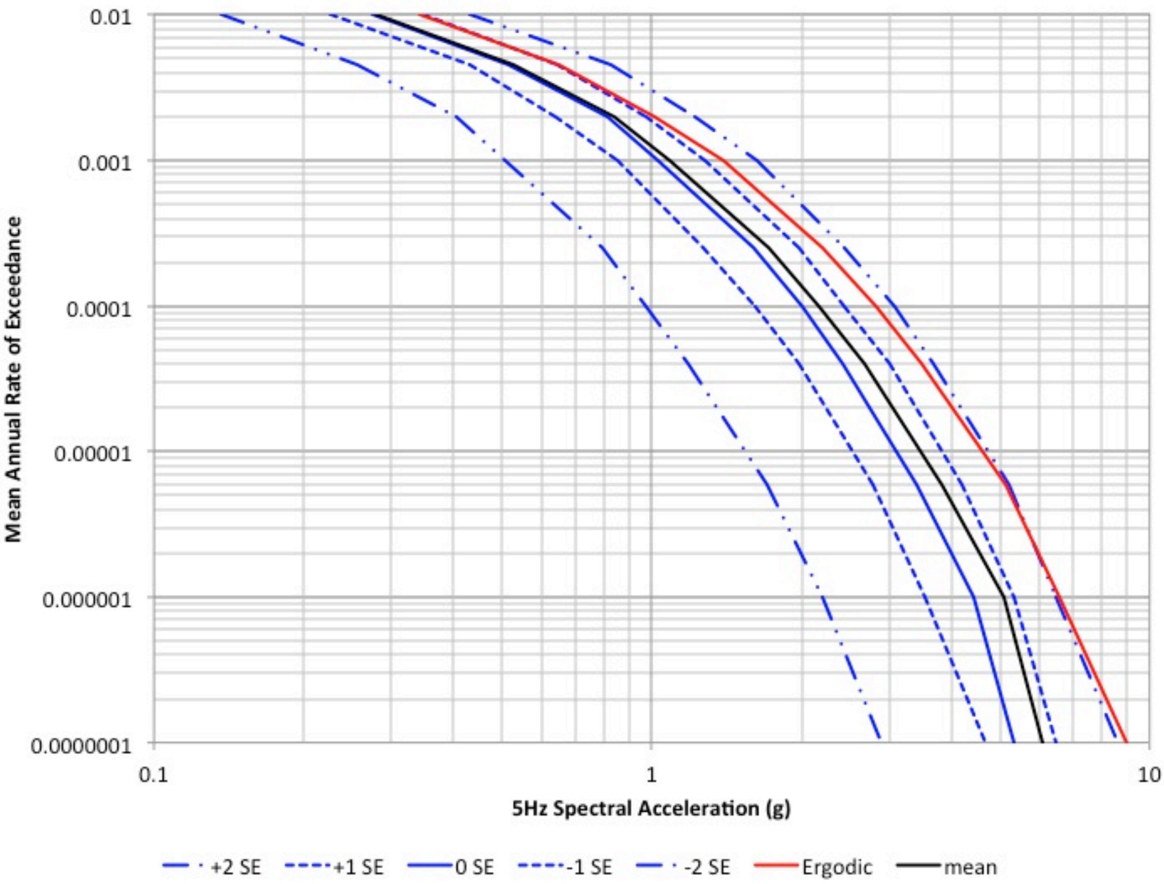

Figure 2.5. Hazard curves for 5Hz spectral acceleration showing ergodic curve (red), fractals of single-station sigma(blue), and mean single-station sigma (black)

As an example, a probabilistic seismic hazard analysis was run for the DCPP site for spectral acceleration at 5Hz using the source model described in PG&E (2011). Single-station sigma ground motion models were used with the above estimate of site term and its associated epistemic uncertainty. Fig. 2.5 shows the resulting hazard curves. The blue curves are different hazard fractals for single-station sigma ground motion models. The fractals reflect the epistemic uncertainty in the estimate of the site term with a range of $\pm 2\,SE\left[\widehat{S}_s\right]$. The black curve is the mean hazard using single-station sigma ground motion models. The red curve is the mean hazard using fully ergodic ground motion models. There is a modest decrease in the mean hazard using single-station sigma ground motion models.

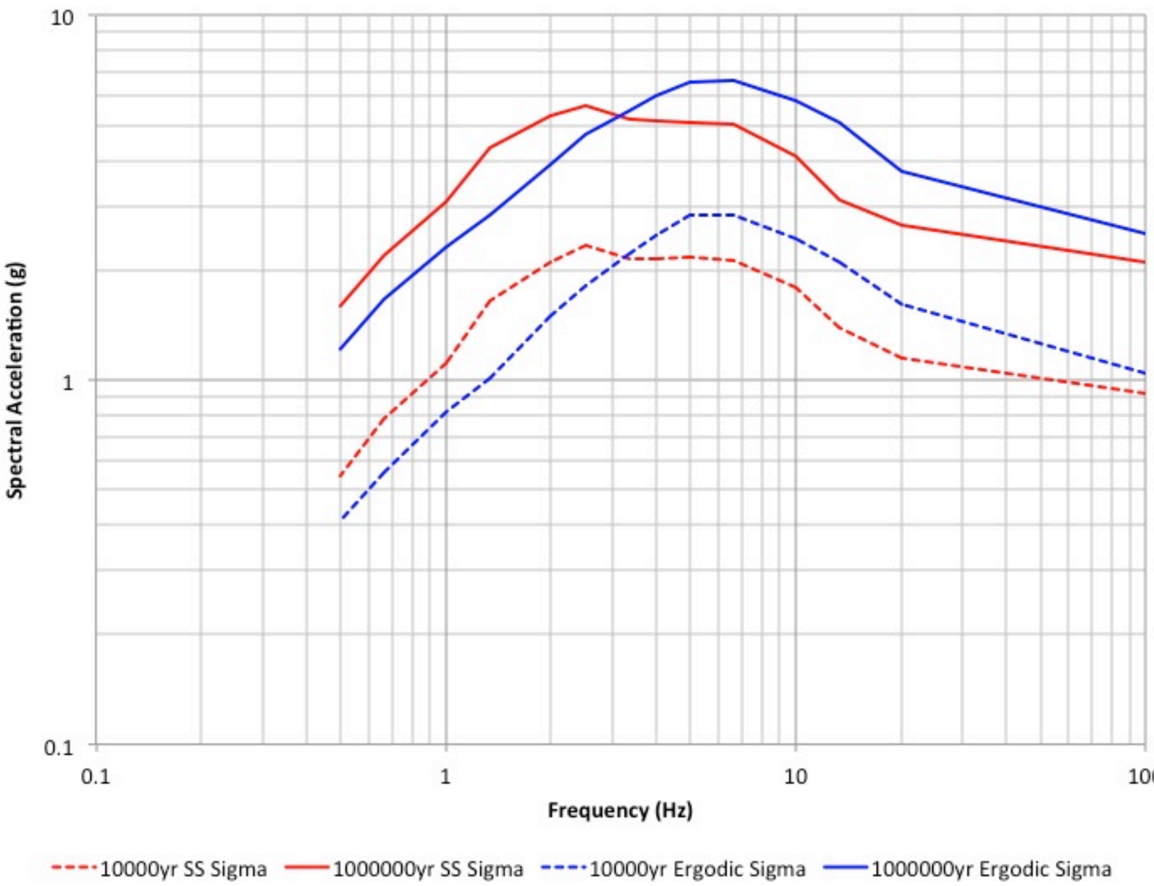

Figure 2.6. Uniform Hazard Spectra for both 10,000 and 1,000,000 year return period, Single-Station Sigma Mean Hazard (Red Curves) and Egrodic Mean Hazard (Blue Curves)

There are two main contributors to this decrease. First, at 5Hz spectral acceleration, DCPP site had a negative site term. This shifts the mean hazard curve to the left. Second, the reduced, partially non-ergodic, ground motion model standard deviation, σ_{ss}, has steepened the shape of the mean hazard curve at low probability levels. Fig. 2.6 shows the effects of the site term and single-station sigma across all frequencies. For the DCPP site there is an increase in the mean hazard for single-station sigma models at low frequencies. This is caused by a positive site term (see Fig. 2.4) that is dominating the effect of the decreased standard deviation. There is a decrease in the mean hazard for single-station sigma models at high frequencies that results from a negative site term and decreased model standard deviation.

3. DISCUSSION AND CONCLUSIONS

Single-station sigma is more representative of the variability of future ground motion observations observed at a single site. There are still systematic and repeatable path and source effects on ground motion that are contributing to single-station sigma; however it is a significant step towards a quantification of the true aleatory variability of the ground motion process. The use of single-station sigma ground motion models is encouraged. However, to properly use these models the site term, and its associated epistemic uncertainty, must be quantified. If the site term is unknown, then the epistemic uncertainty in the estimate of the site term is equal to the site-to-site variability ($SE\left[\widehat{\mathrm{S}}_s\right]$ = ϕ_{S2S}). This results in no change to the mean hazard when using a single-station sigma model but it is still preferred because it better represents the epistemic uncertainty in the hazard. Using the single-station sigma approach provides a framework to show how new site-specific data can be incorporated into the hazard estimate and clearly shows the benefit of operating accelerometers at sites to collect site-specific data.

REFERENCES

Abrahamson, A., Bommer, J. (2006). Why Do Modern Probabilistic Seismic-Hazard Analysis Often Lead to Increased Hazard Estimates? *Bulletin of the Seismological Society of America*. **96:6**,1967-1977.

Abrahamson, A., Silva, W. (2006). Summary of the Abrahamson & Silva NGA Ground-Motion Relations. *Earthquake Spectra*. **24:1**, 67-97.

Abrahamson, A., Youngs, R. (1992). A Stable Algorith for Regression Analyses Using The Random Effects Model. *Bulletin of the Seismological Society of America*. **82:1**, 505-510.

Atkinson. (2006). Single-Station Sigma. *Bulletin of the Seismological Society of America*. 96:2, 446-455.

Al Atik, L., Abrahamson, A., Bommer, J., Scherbaum, F., Cotton, F., Kuehn, N. (2010). The Variability of Ground-Motion Prediction Models and Its Components. *Seismological Research Letters*. **81:5**,794-801.

Anderson, J., Brune, J. (1999). Probabilistic Seismic Hazard Assessment Without the Ergodic Assumption. *Seismological Research Letters*. **70:1**,19-28.

BCHydro (2010). Probabilistic Seismic Hazard Analysis, Volume 3: Ground Motion Report, Draft Nov 3, 2010.

Boore, D., Atkinson, G. (2008). Ground-Motion Prediction Equations for the Average Horizantal component of PGA, PGV, and 5%-Damped PSA at Spectral Periods Between 0.01s and 10.0s. *Earthquake Spectra*. **24:1**,99-138.

Campbell, K., Bozorgnia, Y. (2008). Ground Motion Model for the Geometric Mean Horizantal component of PGA, PGV, and 5%-Damped Linear Elastic Response Spectra for Periods Ranging from 0.01s and 10.0s. *Earthquake Spectra*. **24:1**,139-172.

Chen, Y., Tsai, C. (2002). A stable algorithm for regression analyses using the random effects model, *Bulletin of the Seismological Society of America.* **92:1**,1984–1991.

Chiou, B., Youngs, R. (2008). An NGA Model for the Average Horizantal component of Peak Ground Motion and Response Spectra. *Earthquake Spectra*. **24:1**,173-215.

Lin, P., Chiou, B., Abrahamson, N., Walling, M., Lee, C., Cheng, C., (2011). Repeatable Source, Site, and Path Effects on the Standard Deviation for Empirical Ground-Motion Prediction Models, *Bulletin of the Seismological Society of America*. **101:5**,2281-2295

Pacific Gas and Electric Company (PG*&E) (2011).* REPORT ON THE ANALYSIS OF THE SHORELINE FAULT ZONE, CENTRAL COASTAL CALIFORNIA.

Purvance, D., Brune, J., Abrahamson, N., Anderson, J. (2008). Consistency of Precariously Balanced Rocks with Probabilistic Seismic Hazard Estimates in Southern California. *Bulletin of the Seismological Society of America.* **98:6**,2629–2640.

Rodriguez-Marek, A., Montalva, G., Cotton, F., Bonilla, F., (2011) Analysis of Single-Station Standard Deviation Using the KiK-net Data. *Bulletin of the Seismological Society of America.* **101:3**,1242–1258.

Toro, G., Abrahamson, N., Schneider J. (1997). Letter to the editor, *Seismological Research Letters.* **68:3**,481-482