Impacts of Earthquake Hazard Uncertainties On Probabilistic Portfolio Loss Risk Assessment

G.L. Molas, H. Aslani, N. Shome, C. Cabrera, S. Taghavi, & J. Bryngelson
Risk Management Solutions,Inc. Newark, California, U.S.A.

SUMMARY:
The output of risk analyses is used by insurance and reinsurance companies for critical decision-making that can have a profound effect on the business. As more companies use and rely on the results of catastrophe risk analyses, it is important for companies to have a good understanding of the uncertainties associated with the risk loss metrics. Loss risk models are complex combinations of components, which in turn have uncertainties that are difficult to evaluate. This paper presents a simplified framework to incorporate uncertainties in the different components that can be combined using logic trees.

Specific cases for evaluating the effects of the seismic hazard uncertainties are given. The treatment of the epistemic and aleatory uncertainty of the ground motion models within the framework and the resulting uncertainties in the portfolio loss metrics are presented. Additional source of uncertainties can also be accommodated within the framework.

Keywords: Hazard Uncertainty; Probabilistic Portfolio Loss Estimates

1. INTRODUCTION

Large earthquakes such as the 2011 M_w6.3 Lyttelton earthquake in New Zealand and the M_w9.0 Tohoku earthquake in Japan were not included in the earthquake hazard models developed by the state agencies. This illustrates the importance of considering uncertainty in the earthquake source model for evaluating earthquake risks. The primary objective of modelling is to describe the behavior or relationship of a variable of interest with other variables that can be measured or assumed. The phenomena of earthquake, earthquake ground motion, and building response have inherent variabilities that are difficult to predict. Because of the complex nature of the damage or loss due to earthquakes, it is not easy to formulate equations that could capture the uncertainty due to different components of the loss model.

This study delves into the possible sources of the uncertainties that contribute to the total uncertainty of the insured losses due to earthquake shaking, and examines how these ultimately affect the variability of the insured loss. Although there are many possible sources for the uncertainty, this study focuses on a few sources that can have significant contributions to the loss uncertainty; in particular, the effect of the hazard uncertainties on the loss uncertainty. This paper also describes a simple methodology for quantifying the uncertainty in the loss results of a portfolio of buildings.

2. EARTHQUAKE RISK MODEL FRAMEWORK

The methodology for quantifying the earthquake risk for insurance and reinsurance companies uses a modular framework consisting of five basic modules (Figure 1). The Stochastic Event Module defines future events and their recurrence rates. The events are simulated from a source model that describes the types of earthquakes and their physical locations and dimensions. For a stochastic event, the expected ground motion at a given location of a building is calculated by the Hazard Module. The ground motion includes the effects of site amplification and other geotechnical properties provided by the Geocoding/Exposure Module. The Vulnerability Module uses the characteristics of the building(s) located at a site to select an appropriate damage curve to estimate the mean and coefficient of variation of the damage. Finally, the Financial Analysis module uses the financial terms of the insurance policies or contracts to calculate the loss for multiple stakeholders. This includes aggregating losses from multiple buildings if they are part of the policy.

Figure 1. Earthquake Risk Modelling Framework

The probabilistic framework used in this study is derived from the Performance-based Earthquake Engineering (PBEE) framework developed at the Pacific Earthquake Engineering Research (PEER) Center at the University of California, Berkeley (Deierlein, et al., 2003; Aslani and Miranda, 2005). This framework can be described as an expression of the total probability theorem as follows:

$$\nu(DV)=\iiint P(DV \mid DM)\left|dP(DM \mid EDP)\right|\left|dP(EDP \mid IM)\right|\left|d\nu(IM)\right| \tag{2.1}$$

where: *DV* is the decision variable (e.g., collapse), *DM* is the damage measure (e.g., mean loss curve), *EDP* is the engineering demand parameter (e.g., maximum inter-story drift), and *IM* is the intensity measure. In the context of this study, the decision variable is the financial loss and the intensity measure could be the pseudo-response acceleration or response displacement. The most common method to describe the distribution of IM is through the use of Ground Motion Predictive Equations (*GMPE*), which estimate the ground motion (intensity measure) from the event parameters (magnitude, depth), site parameters (distance, V_{S30}), etc. The relationship can also be expressed in the Probabilistic Seismic Hazard Analysis (PHSA) framework in discrete form as:

$$P(Y \geq y)=\sum_{i=1}^{R}\sum_{j=1}^{M} P(f(m,r) \geq y \mid m_j, r_i)\cdot P_M(m_j)\cdot P_R(r_i)\cdot \Delta m \cdot \Delta r \tag{2.2}$$

where: $f(m,r)$ represents a GMPE with associated uncertainty parameters, $P_M(m_j)$ is the probability distribution for Magnitude *j* for source *M*, and $P_R(r_i)$ is the probability distribution of the source-to-site distance.

The use of a probabilistic framework provides a unified way to incorporate the uncertainties of the components. It also allows consideration of the uncertainty of one of the components while keeping the others to their conditional mean values. This leads to a simple way of evaluating the sensitivity or uncertainty of the Decision Variable due to the uncertainty of the components in the framework.

3. METHODOLOGY

It can be very complex and cumbersome to determine the contribution of each of the sources of uncertainty in portfolio risk analyses. Each component in the risk analysis framework in Figure 1 has a number of subcomponents; thus, estimation of uncertainties of the components is very complex due to differences in the distribution characteristics of the subcomponents and correlation between them. The correlation in losses or decision variable due to the variation in different components of the model is also difficult to quantify. To improve our understanding of the effect of hazard uncertainties on the risk uncertainties, the individual effects of each of the components are evaluated. Deriving the relative sensitivity due to the variation in the different components of the model will be the initial step. This study focuses primarily on the hazard uncertainties; specifically, the uncertainties in the GMPEs.

3.1. Treatment of Hazard Uncertainty

3.1.1. Epistemic Uncertainty

Epistemic uncertainty arises from incomplete knowledge about either the methodologies or the modelling parameters. The epistemic uncertainty can be reduced by the analysis of more data and better understanding of the processes. The methodology for the generation of the 2008 National Seismic Hazard Maps (NSHM) for the United States developed by the United States Geological Survey (USGS) considers the epistemic uncertainty in the ground motion intensities by using a logic tree of multiple GMPEs to estimate the ground motion hazard at a given site (Petersen, et al., 2008). In addition, since the Next Generation Attenuation (NGA) West project had limited amount of data for large magnitude earthquakes and there was significant interaction among the developers with the possibility of influencing each other's decision, it was deemed necessary to specify additional variability for the Western United States. The amount of additional uncertainty ("dgnd" as specified by USGS) was based on the number of data used in the development of the GMPEs and given in Table 3.1.

In this study, a similar logic tree for multiple GMPEs is adopted to capture the GM uncertainties in earthquake hazard. The examples in this study for the Western U.S. also include the additional variability as specified in the USGS methodology. This study considers the three NGA relationships used in the NSHM, namely, Boore and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008). Different combinations of the GMPEs, with and without the additional variability, are considered to study the sensitivity of the estimated losses.

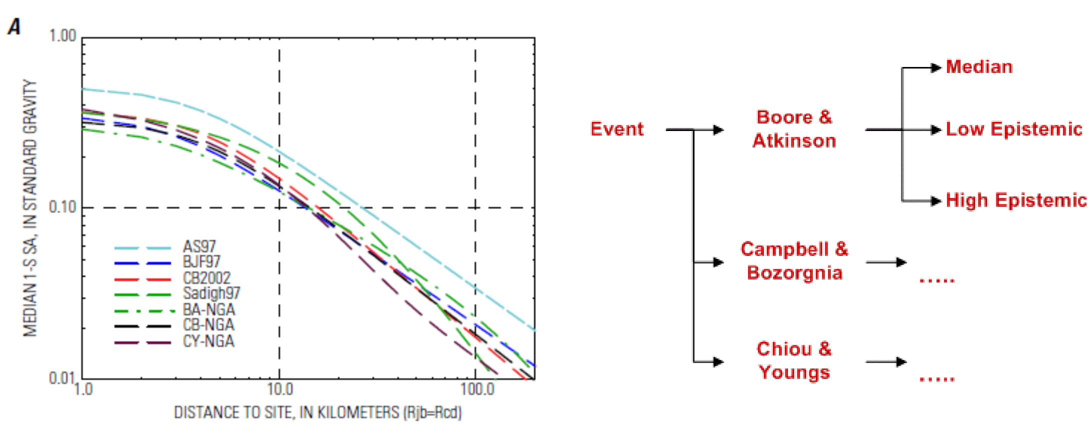

Figure 2. Epistemic uncertainty is considered by the use of multiple GMPEs and additional uncertainty (Petersen, et al., 2008). The models are organized as logic trees resulting in 9 individual models.

Table 3.1. Additional uncertainty applied to NGA relationships for the 2008 National Seismic Hazard Maps (Petersen, et al., 2008)

Magnitude	Rupture Distance (km)	Additional uncertainty term (dgnd)
5≤M<6	Rrup<10	0.375
5≤M<6	10≤Rrup<30	0.210
5≤M<6	Rrup≥30	0.245
6≤M<7	Rrup<10	0.230
6≤M<7	10≤Rrup<30	0.225
6≤M<7	Rrup≥30	0.230
M≥7	Rrup<10	0.400
M≥7	10≤Rrup<30	0.360
M≥7	Rrup≥30	0.310

3.1.2. Aleatory Uncertainty

Aleatory uncertainty is based on natural randomness and cannot be reduced by the inclusion of more data in the analysis. GMPEs specify the median and the standard deviations of GM intensities based on the observed records. In addition, the GM intensities are correlated spatially, such that the deviations from the mean at two nearby locations are more correlated than the deviations at locations that are farther apart. The effect of this spatial correlation on the uncertainty of the estimated insured loss was studied by Molas, et al. (2006) for a single large event. They assumed independence of the loss at different sites and found that the epistemic uncertainty in the ground motion model contributes significantly to the variability of the event loss estimate. This spatial correlation of GM intensities, however, is not considered in this study.

The ground motion uncertainty is convoluted with the damage uncertainty using Monte Carlo simulation to estimate its combined effect with the loss Coefficient of Variation (CV). This paper focuses only on the hazard uncertainties and does not show the results for all the different sources of uncertainties, which would have required the convolution process.

3.2. Measuring the Effect of Hazard Uncertainty on the Risk Uncertainty

Two commonly used loss metrics for the quantification of insurance risk are the Average Annual Loss (AAL), sometimes referred to as the pure premium, and the Loss Exceedance Probability (EP) curve. Given a set of stochastic events with a mean loss estimate, the AAL is the cross-product of the loss and the annual event rate. While the AAL is not affected by the aleatory uncertainty in the loss due to an event, it is affected by the epistemic variation in the mean loss, which is generated based on the GMPE logic tree. The loss (aleatory) uncertainty for an event, on the other hand, directly influences the EP curve calculations.

The EP curve estimates the annual probability of exceedance of loss. Assuming that each event is independent, the probability of exceeding a given loss threshold is the sum of the probability of exceeding the threshold given the event loss distributions. By calculating several loss thresholds, an exceedance curve is generated for the loss thresholds (Figure 3).

4. RESULTS AND DISCUSSION

4.1. Exposure for the Loss Calculations

The risk analysis methodology can be performed on a single location (for building-specific risk calculations) or on a number of locations (for portfolio risk calculations) over a region when the building characteristics are known. For the risk analysis of a portfolio of buildings, however, the distributions of the location and characteristics of the buildings exposed to the hazard play a crucial role.

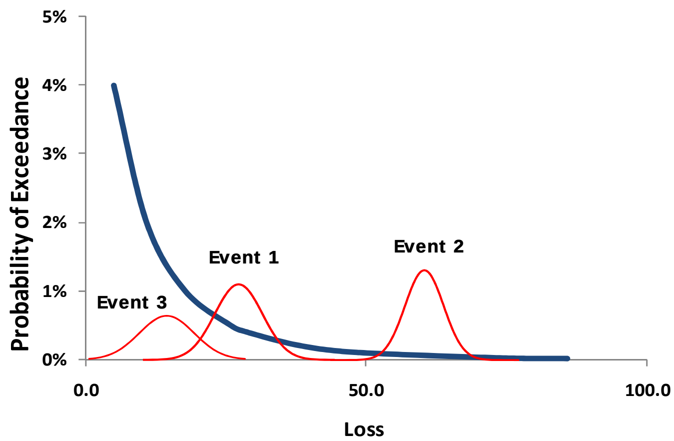

Figure 3. The probability of exceeding a loss threshold is calculated from the event loss distributions (in red). The exceedance probability (EP) curve (thick blue line) is generated when enough loss thresholds are analyzed.

Very high hazard in areas with very little exposure or number of buildings can be a low risk, while an infrequent event in high concentrations of exposures in downtown Memphis, for example, can be a significant contributor of risk to an insurer or reinsurer.

The exposure in this study uses the economic basis that represents the built environment, i.e., considers all the buildings in a region of interest. The portfolio of buildings in a built environment is represented by the economic exposure database (EED) and this is a modelled distribution of residential and non-residential buildings based on population, tax assessor's files, and other public and private sources of information. In this study, the distribution of building characteristics and occupancy types are based on multiple sources of information at different levels of resolution, such as the USGS 30-meter National Land Cover Database (NLCD), and the National Census data. The database used in this study is defined at the postcode level and includes building and content values.

4.2. Sensitivity of Event Losses

To show the range of losses related to the epistemic uncertainty from the choice of ground motion attenuation relationships, two scenarios, which are a repeat of the 1994 Northridge earthquake and of the 1906 Great San Francisco earthquake, are shown in Figures 4 and 5, respectively. The branches of the logic tree include the three NGA relationships used in the National Seismic Hazard maps as well as the additional epistemic uncertainties (lower and higher) given by the USGS methodology. In implementing a stochastic model, both the higher additional uncertainty and the lower additional uncertainty are considered equally. Since the relationship between ground motion and loss is non-linear, the inclusion of both will generally not cancel out in terms of losses.

Figure 4 shows that the loss estimates from the three NGA relationships are very similar, with approximately +/- 3% change from the reference model. However, when the additional epistemic uncertainty is considered, the change in loss becomes very significant; around -32% for the lower uncertainty and around +45% for the higher uncertainty. The effect of the additional uncertainty on the loss estimates among NGA relationships is similar since the same level of uncertainty is applied.

For stochastic modelling, combining the effects of both the high and low additional uncertainty produces very small changes in the loss estimates. When taken together, their effects mostly cancel each other, and the resulting effect of the additional epistemic uncertainty becomes minimal. This implies that the mean losses for stochastic modelling are not affected significantly by the additional uncertainty, but the estimation of the event loss standard deviation will be affected given the large percentage difference.

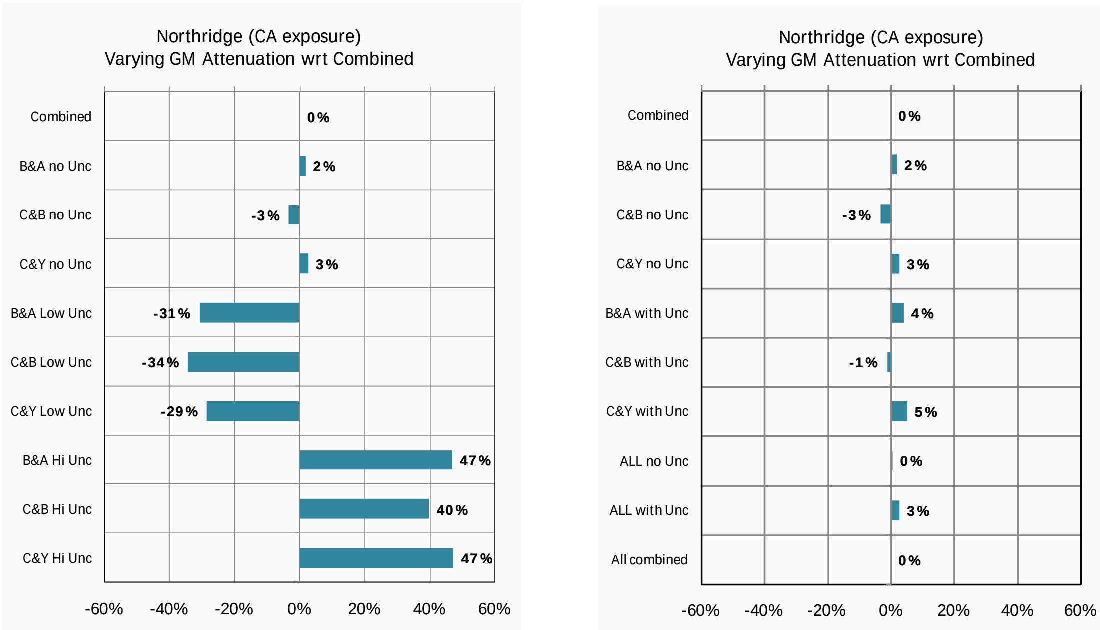

Figure 4. Sensitivity of loss estimates for a repeat of the 1994 Northridge earthquake in California. The uncertainty is from the ground motion predictive equations, with and without the additional epistemic uncertainty (left). Sensitivity to the loss estimates with the combined effect of the higher and lower additional epistemic uncertainty is also shown (right). Loss changes are relative to the loss estimated from combining the three NGA relationships with equal weight.

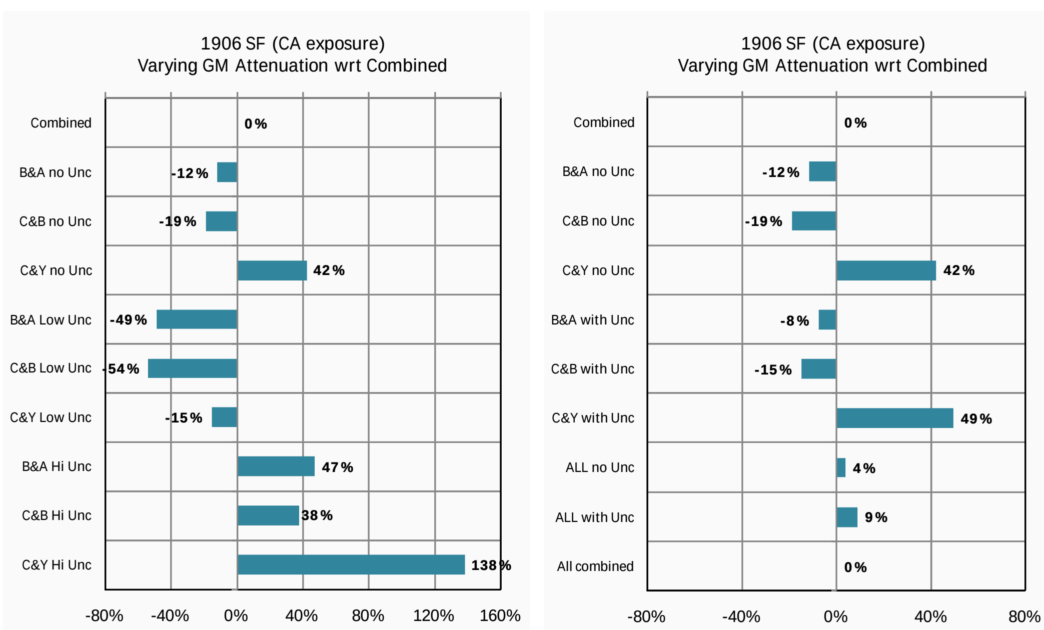

Figure 5. Sensitivity of loss estimates for a repeat of the 1906 San Francisco earthquake. The uncertainty is from the ground motion predictive equations, with and without the additional epistemic uncertainty (left). Sensitivity to the loss estimates with the combined effect of the higher and lower additional epistemic uncertainty is also shown (right).

Figure 5, however, shows much larger differences among the losses from the three NGA relationships for the repeat of the 1906 San Francisco earthquake. This may be due to the fact that the GMPEs are less constrained for the high magnitude events due to the limited data. The incremental effects of the additional higher and lower uncertainty are similar with the Northridge losses, but the total effects show very high sensitivity. As before, when the losses from the higher and lower uncertainties are combined, the changes in the mean loss results over the base NGA relationships are relatively smaller. These results show that the differences between the NGA relationships are dependent on the magnitude and other predictor variables, such as the distance and fault rupture type.

To better understand the effect of the additional uncertainty, the plot of Chiou and Youngs (2008) PGA attenuation for Magnitude 7.0, with and without the additional uncertainty, is given in Figure 6 (left). The range between the high and low estimates is significant, especially at small distances. Figure 6 (right) also shows the range of the Spectral acceleration (SA) for a distance = 10 km. It can also be seen that the range is significant, especially close to the peak of the spectra. The changes in the loss estimates are consistent with the ground motion values that are produced by the logic trees.

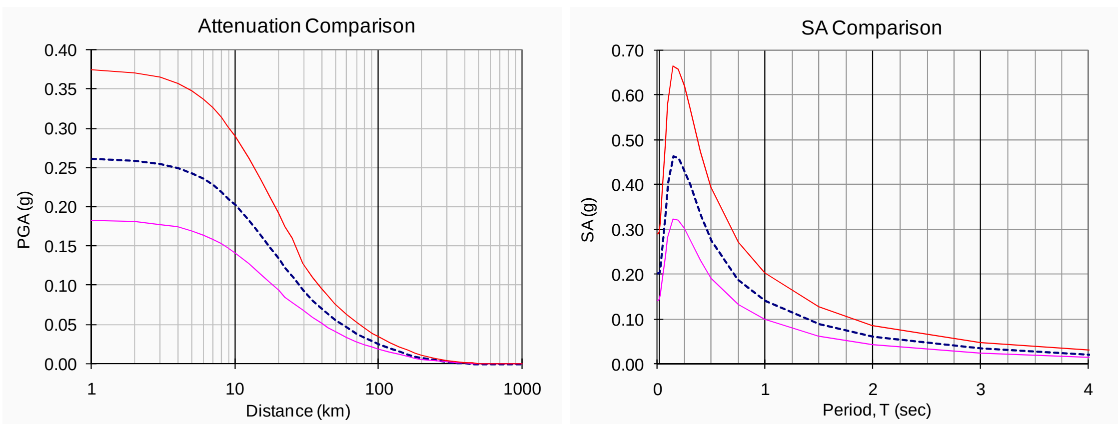

Figure 6. PGA attenuation (left) and SA (right) for Chiou and Youngs (2008) model with median and high and low additional uncertainty (Mw = 7.0, Vs30=760 m/s, SA plot for 10km distance)

The argument that the NGA relationships will not include a realistic epistemic uncertainty seems to hold in the 1994 Northridge scenario. It is noted that the number of observed ground motion data in NGA database for the magnitude range of the Northridge earthquake is quite numerous compared to the data in the higher magnitude range similar to the 1906 San Francisco event. Because of this, the GMPEs are well constrained for the Northridge event. For the 1906 San Francisco event (Figure 5), it seems that there is significant epistemic uncertainty in loss results represented by the three NGA relationships before applying the additional uncertainty. Similar studies for more scenarios would be useful for evaluating the epistemic uncertainty in the ground motion to the loss results.

4.3. Exceedance Probability Curves Generated from a Logic Tree

So far, sensitivity analysis in the loss results for a couple of historical events were performed and so the importance of each of the logic-tree branches has not played any role as yet. The relative weights of the logic-tree branches, however, are important for generating the exceedance probability curves, where each combination of branches in the logic tree will correspond to one EP curve. The weights become an important parameter when calculating the mean loss and confidence intervals around the mean for the EP curves.

For instance, an extreme value of a parameter may be specified in a logic tree, even though the probability of its occurrence may be small. This can give an indication of the range of losses due to uncertainty in the parameter, but the influence or weight of the logic-tree branch due to epistemic uncertainty in the parameter would be small to reflect the rarity of this particular outcome.

Given the difference in the results for the 1906 San Francisco and 1994 Northridge earthquake scenarios, the sensitivity of the loss metrics for the San Francisco and Los Angeles counties are investigated. Figure 7 shows the changes to the AAL and 250-year loss estimates due to uncertainty in the ground motion models.

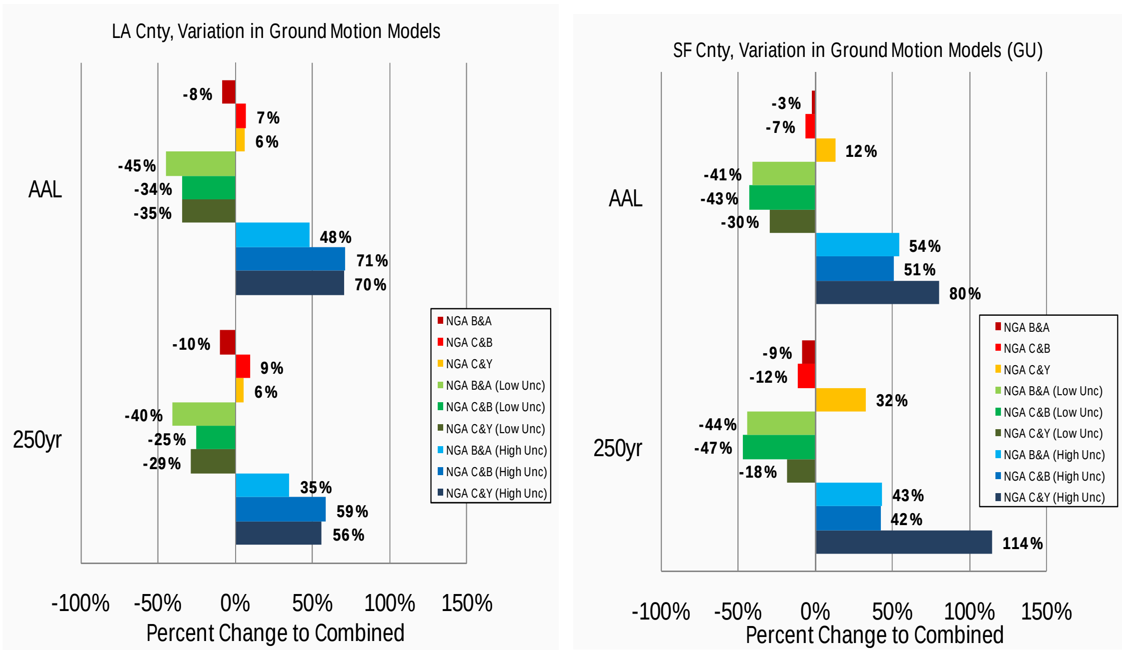

Figure 7. Sensitivity of the AAL and 250-year loss for L.A. County (left) and San Francisco County (right)

For Los Angeles County results, the AAL and 250-year loss results follow similar patterns as the changes for the Northridge event. Changes from the base NGA relationships are within 10%, while the changes for the models with additional epistemic uncertainty range from 35% to 70% with the higher uncertainty generating higher percentage changes due to the non-linearity of the loss with respect to the hazard. For San Francisco County, the patterns are similar, but with a bit more variability among the NGA models. Other than these general comments, it is difficult to find a specific pattern regarding the percentage changes for the AAL and 250-yr losses.

The changes for a smaller region like the Los Angeles and San Francisco counties can be affected by the types of events and types of exposures that dominate a given range of return periods. For the San Francisco region, the risk is dominated by large infrequent events along the San Andreas Fault in the same class as the 1906 San Francisco earthquake. The exposure is also characterized by a concentration of high rise buildings in the downtown area that are far from the major fault zones. For Los Angeles County, the risk is dominated by smaller thrust events that are nearer and affect exposure that is less concentrated and in a larger area. The effect of the Los Angeles basin on the loss estimates also plays an important role. This makes it important to perform the actual analysis, rather than try to find a way to model how uncertainties in the hazard will affect the uncertainty in the loss estimates.

The Loss EP curve represents the risk of a portfolio of buildings against earthquake loss. A single curve is associated with a specific model. When considering multiple models, as in the case of using logic trees, each model will result in a distinct EP curve and the distribution of the curve shows the range of the uncertainty effect on the risk. When the relative weights of the different models are considered, a distribution of the losses at different exceedance probabilities can be calculated. The confidence intervals can also be calculated by fitting a parametric model to this distribution (Figure 8).

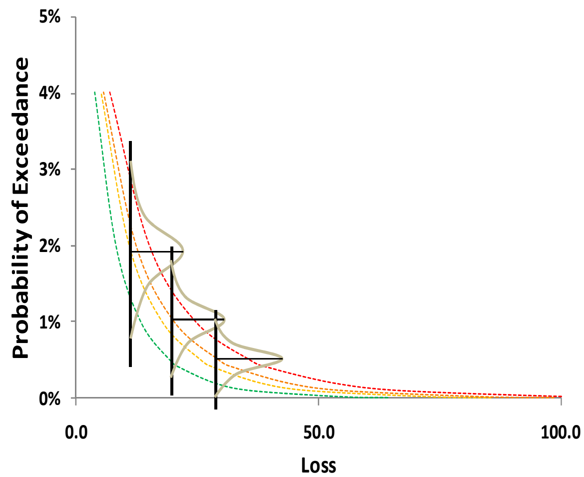

Figure 8. The mean and standard deviation of the exceedance probability can be evaluated given the relative weights of each EP curve. The distribution of exceedance probability is used for calculating the confidence intervals.

4.4. Incorporating Other Sources of Uncertainty

This study investigates how the hazard (limited to the ground motion predictive equations) uncertainty affects the portfolio loss uncertainty. There are many other possible sources of uncertainty within the hazard component (e.g., earthquake probability models) and in other components as well (e.g., vulnerability). The event rates as determined by the USGS National Seismic Hazard Map methodology involve logic trees that include multiple model choices for different components (e.g., fault model) resulting in a large number of branches. For the vulnerability component, the estimation of epistemic uncertainty due to different parameter assumptions (e.g., cost of repairing a damaged structure) is very complex due to difficulty in the estimation of different parameters of buildings and the correlation of those parameters for damage calculations (Jayaram, et al. 2012). The exposure used in the analysis has also significant uncertainty. Even when a single building is coded into a digital database, human errors can cause loss estimates to be very different, compared to one that is based on the most accurate information.

The framework described in this paper can be extended to allow for many more alternative models as long as these can be implemented as branches in a logic tree with appropriate weights. As discussed previously, each branch will generate one AAL, EP curve, and return period loss thresholds. The resulting number of EP curves can run into the hundreds, even thousands. The number of alternate models that can be analyzed is limited mostly by the available computing capacity. These results provide an estimate of the uncertainty of the loss metrics.

As more components are added in the framework, the range of exceedance probabilities for a given loss threshold can become much wider. The resulting mean and standard deviation curves can be easily calculated from the large number of EP curves. The ability to estimate the uncertainty of the loss metrics is very helpful in decision-making and risk management for insurance and reinsurance companies.

5. CONCLUDING REMARKS

This paper looks into the effects of the uncertainties in the choice of ground motion predictive equations on the uncertainty of the risk loss metrics for a portfolio. The changes in the losses are a valuable tool to evaluate the loss uncertainties. Incorporating the uncertainties of the event losses and the implementation of logic trees to consider the uncertainty of model parameters and methodologies in the risk framework generate metrics that can describe the level of uncertainty for a portfolio loss metrics. Cases presented in this paper illustrate the utility of the methodology.

The loss risk framework provides a unified way to combine the uncertainties of the different components of the stochastic risk model. The number of branches of the logic trees can be small or large depending on the understanding of the uncertainties. Branches can be added later, which will affect the relative weights of the existing alternative models. Since the EP curve for each model is not affected by the addition of logic tree branches, they need not be recalculated. Once the additional EP curves are available, only the mean and standard deviation of the exceedance probabilities have to be recalculated, which does not require much computational effort.

The presented framework, while useful and adaptable, is only the first step in the evaluation of risk uncertainties. More complicated simulation techniques are needed for a robust characterization of the secondary correlation and complex contract terms commonly found in insurance policies.

REFERENCES

Aslani H. and Miranda E. Probabilistic earthquake loss estimation and loss disaggregation in buildings. John A. Blume Earthquake Engineering Center, Report No. 157, Dept. of Civil Engineering, Stanford University, Stanford, California, 2005; 355 pages.

Boore, D. M. and Atkinson, G. M. (2008). Ground-motion prediction equations for the average horizontal 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. W. and Bozorgnia, Y. (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 10s. *Earthquake Spectra*. **24:1**, 139-171.

Chiou, B. S. J., and Youngs, R. R. (2008). An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*. **24:1**, 173-215.

Deierlein, G.G., Krawinkler, H, and Cornell, C.A. (2003). A Framework for Performance-Based Earthquake Engineering. *Pacific Conference on Earthquake Engineering*. Paper 140s. 8 p.

Jayaram, N., Shome, N. And Rahnama, M. (2012). Development of earthquake vulnerability functions for tall buildings. Earthquake Engineering and Structural Dynamics (to be published).

Molas, G.L., Anderson, R., Seneviratna, P. and Winkler, T. Jurukovski, D. and Rakicevic, Z**.** (2006). Uncertainty of Portfolio Loss Estimates for Large Earthquakes. *First European Conference on Earthquake Engineering and Seismology*. **Paper 1117**: 10 p.

Petersen, Mark D., Frankel, Arthur D., Harmsen, Stephen C., Mueller, Charles S., Haller, Kathleen M., Wheeler, Russell L., Wesson, Robert L., Zeng, Yuehua, Boyd, Oliver S., Perkins, David M., Luco, Nicolas, Field, Edward H., Wills, Chris J., and Rukstales, Kenneth S. (2008). Documentation for the 2008 Update of the United States National Seismic Hazard Maps: U.S. Geological Survey Open-File Report 2008–1128, 61 p.