



Stochastic Simulation of Earthquake Ground Motion Footprints Constrained by Recorded Data and MMI Intensity Maps

M. Mahdyiar¹, B. Dodov¹, B. Shen-Tu¹, K. Shabestari¹, J. Guin¹, Y. Rong¹

ABSTRACT

Estimating the expected loss for an earthquake requires simulating sets of stochastic GM fields that reflect the inter-event and intra-site spatial correlation and at the same time are constrained by the available earthquake ground motion data. For many past earthquakes, the only available ground motion related data are the MMI maps. This paper presents a practical method for simulating stochastic ground motion fields that account for inter-event and intra-site correlation effects and are constrained by the recorded or inferred MMI-based ground motions.

Introduction

Estimating regional financial losses for earthquakes is important for both public safety and financial risk planning. A reliable regional loss estimate for an earthquake requires a comprehensive earthquake risk analysis model that is well calibrated against losses from past earthquakes. One of the most important components of the calibration process is the simulation of the ground motions for the calibrating earthquakes at the exposure locations. Ground motion (GM) simulation is also an important issue for real-time earthquake loss analysis, where GM footprints of an earthquake need to be constructed within a few hours of its occurrence. In general, the current practice for such analysis is to simulate regional GM footprints of earthquakes using the regional empirical attenuation equations and include some consideration for GM uncertainty.

For earthquake GM simulation, modelers often try to capture the regional characteristics of the earthquake by imposing GM constraints using the recorded ground motion data. However, for earthquakes that have no GM recordings, the available information is often limited to MMI intensity maps. In such cases the MMI data needs to be translated into GM values using empirical equations. There are large uncertainties in both the MMI contour maps and the MMI-PGA or MMI-Sa (spectral acceleration) empirical conversion equations, which need to be considered. The GM simulation becomes more complex as one also tries to account for the inter- and intra-event GM variability, and intra-station GM correlation. It is the objective of this study to investigate a cost-effective procedure for simulating stochastic GM footprints of earthquakes constrained by different types of GM data, obtained from direct recordings or inferred from MMI contour maps.

For earthquake loss analysis of portfolios, it is important to account for the GM uncertainty and spatial correlation. Even if one is only interested in estimating the expected loss and not the uncertainty in loss estimates, accounting for GM uncertainties is important because of the nonlinear damage response of buildings to earthquake- GM. Fig. 1, reproduced from a study by Mahdyiar et al. (2006), demonstrates this point. The figure shows a comparison between the losses for the 1906 San Francisco earthquake at different sites, calculated with and without taking into account the GM uncertainty. Buildings at sites close to the San Andreas fault

¹ AIR Worldwide, 131 Dartmouth Street, Boston, MA 02116

experience high levels of GM and, considering the typical shape of damage functions, mostly experience damage according to the linear parts of the damage functions. For these sites, there are good agreements between the mean losses calculated with and without considering GM uncertainty. This agreement, however, breaks down and the differences between these two estimates of the expected loss increase with increasing the source-site distances. At large source-site distances, buildings experience low levels of GM and are mostly damaged according to the nonlinear parts of the damage functions. For these sites, as is shown on Fig. 1, the expected loss, calculated by considering the GM uncertainty, is higher than the expected loss based on the median GM.

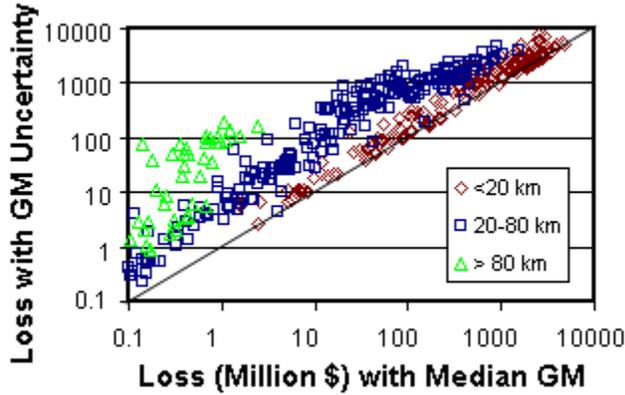


Figure 1. Expected loss values for the 1906 San Francisco earthquake at different sites calculated with and without considering the effects of GM uncertainty

Model Formulation

The most recent empirical attenuation equations formulate the ground motion uncertainties in terms of inter- and intra-event random effects. The inter-event variation reflects uncertainties in the source parameters, and the intra-event variation reflects the combined path, site, and other non-source related uncertainties (Abrahamson and Silva, 1997). A typical attenuation equation can be formulated as:

$$\ln(Y) = \overline{\ln(Y)} + \varepsilon_{Inter} + \varepsilon_{Intra} \quad (1)$$

where ε_{Inter} and ε_{Intra} are the inter- and intra-event random errors, assumed to be independent, and normally distributed with zero means and a standard deviation (SD) of σ_{Inter} and σ_{Intra} , respectively. It can be shown that such stochastic GM field produces the following spatial correlation for any two sites:

$$\rho_{GM} = \frac{\sigma_{Inter}^2}{\sigma_{Inter}^2 + \sigma_{Intra}^2} \quad (2)$$

However, the GM correlation for sites close to each other can be larger due to the intra-site correlation that depends on the distance between the sites. The intra-site correlation can make an important contribution to the regional loss distribution making the damage from some earthquakes very costly, (Bazzurro and Park, 2007). The overall GM correlation, after including the intra-site correlation, can be formulated as:

$$\rho_T = \frac{\sigma^2_{Inter} + \rho(d) * \sigma^2_{Intra}}{\sigma^2_{Inter} + \sigma^2_{Intra}} \quad (3)$$

where $\rho(d)$ is the intra-site correlation with d representing the distance between stations (Boore et al, 2003, and Bazzurro et al., 2007). Fig. 2 shows the distribution of ρ_T versus distance for typical values for σ_{Inter} , σ_{Intra} , and $\rho(d)$ (Lee et. al., 2005).

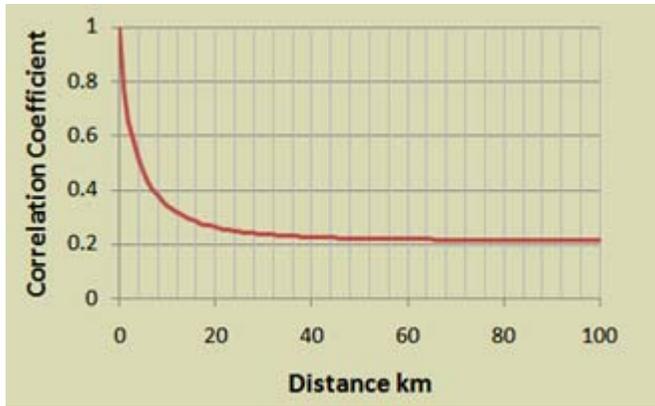


Figure 2. The integrated correlation coefficients for a pair of stations as a function of distance between the stations

Another important component of earthquake GM simulation involves capturing the earthquake's regional GM characteristics. For earthquakes with GM recordings, the GM information can be used to constrain the simulation at sites near the recording stations. For earthquakes with MMI data or contour maps, the MMI information can be translated into GM values and be used to constrain the simulation, while taking into consideration the uncertainty in the MMI-GM conversion. The task, therefore, is to formulate a stochastic GM simulation procedure that is based on the regional attenuation equations, capture the stochastic nature of the GM as described by Eq. 1, capture the intra-site correlation effects, and incorporate the constraints imposed by different kinds of observations.

A practical way to approach this problem, given the regional attenuation equations, is to simulate a set of stochastic GM residuals, with respect to the median GM, at sites of interest that conform to the imposed GM constraints. The ensemble of many sets of residual fields capture the realization of the regional inter-event, intra-event, and intra-site correlation effects. Given that the residuals are simulated with respect to the median GM for the reference site conditions, regional earthquake GM footprints can be constructed using the information on the shallow soil site conditions. Simulating the correlated random field with no intra-site correlation effects is

rather simple and can be achieved by a two-step simulation process based on Eq. 1. However, this approach creates ground motion fields with a constant site-to-site correlation that is independent of the distance between the sites.

There are different techniques available for simulating spatially correlated stochastic random fields, representing residuals, with variable intra-site correlation. Cholesky decomposition and simulation procedures have successfully been used for such purposes including the cases where constraints from observed data need to be imposed (Bazzurro et al. 2007). However, Cholesky decomposition can become time consuming if the size and the bandwidth of the covariance matrix are large. A cost-effective method for simulating a random field with isotropic spatial correlation is the spectral transformation and simulation technique (Keqiang and Xu, 2009). In this approach, the covariance matrix is transformed to the frequency domain using the Fast Fourier transformation (FFT). Imposing a random phase in the frequency domain, the correlated random field is simulated using the inverse FFT as follows:

$$S(k) = \frac{1}{2\pi} \int \int e^{ik\xi} \rho(\xi) d\xi \approx FFT(\rho(\xi)) \quad (4)$$

$$R(\text{Random Field}) = C * FFT^{-1}(\sqrt{S(k)} * A) \quad (5)$$

where $S(k)$ is the spectral density function, k is wave number, and A is a set of complex random numbers that capture the randomness of the field over the wave number domain. This procedure is a straightforward and relatively fast way of creating random fields over large scale grids with isotropic correlation effects. The values for A can be constructed by taking the FFT of a set of random numbers at grid points drawn from a Gaussian distribution. Multiple realizations of Eq. 5 will create a normally distributed set of random numbers at each grid point that conform to the GM residuals and the spatial correlation requirements. The GM at each site can be written as:

$$Y_i = \bar{Y}_i + R_i \quad (6)$$

The GM correlation coefficient for any two sites can be written as:

$$\rho_{ij} = \frac{E[(Y_i - \bar{Y}_i) * (Y_j - \bar{Y}_j)]}{\sigma_{Y_i} * \sigma_{Y_j}} = \frac{E[R_i * R_j]}{\sigma_i * \sigma_j} \quad (7)$$

where σ_{Y_i} and σ_{Y_j} are the SD of GM and σ_i and σ_j are the SD of the residuals. From Eq. 6, it is clear that SD of the GM and residuals are the same.

By adding these residuals to the median GM, and applying the shallow site response, a single stochastic realization of the GM footprint for the earthquake is created. However, as was discussed earlier, it is often necessary to constrain the simulation with the recorded GM data or inferred GM data from an MMI map. Let us consider a single site with the GM recordings of G_0 , which includes PGA and various spectral values. The objective is to capture the overall effect of the source, path, and site conditions that created the GM G_0 at the recording site, and

incorporate this effect in the GM simulation of sites nearby. Given the site condition of the recording site, the shallow soil site response can be removed from G_0 to calculate the GM at the site for the reference site conditions, and then the residuals, y_0 , with respect to the median GM of the attenuation equation. The residuals capture the overall effects of many complex phenomena that, in general, are formulated by the inter- and intra-event random effects shown in Eq. 2. Considering that the shallow site condition effects are removed, the residuals y_0 should have some prevailing local effects. For each grid point within some distance of the recording site we define a new random variable as a linear combination of y_0 and the simulated random field at the grid point as follows:

$$R_i^c = \alpha_i * y_0 + (1 - \alpha_i) * R_i \quad (8)$$

where α_i is a coefficient that controls the contribution of the constraining residual y_0 , relative to the contribution of the stochastically generated residuals at the grid point.

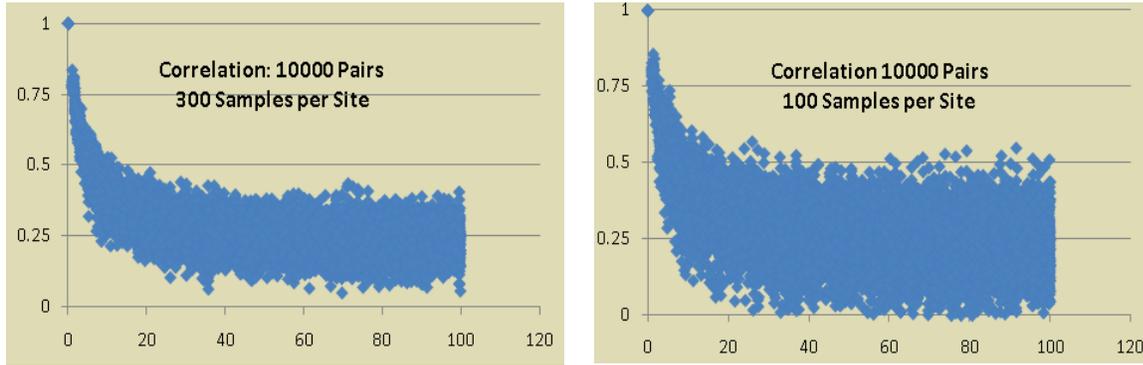


Figure 3. The distribution of correlation coefficients between simulated residuals at different pairs of sites as a function of the distance between the sites

The value of α_i is only a function of distance, and decreases as the distance between the grid and the recording site increases. It can be shown that the simulated residuals based on Eq. 8 produce a GM field with the same spatial correlation between any pair of grids as the simulated GM field based on the original unconstrained residuals. The simulated GM at each grid point based on Eq. 8 can be written as:

$$Y_i^c = \bar{Y}_i + R_i^c = \bar{Y}_i + \alpha_i * y_0 + (1 - \alpha_i) * R_i \quad (9)$$

The mean and the variance of this GM field are:

$$\bar{Y}_i^c = \bar{Y}_i + \alpha_i * y_0 \quad (10)$$

$$(\sigma_i^c)^2 = (1 - \alpha_i)^2 * E[(R_i)^2] = (1 - \alpha_i)^2 * \sigma_i^2 \quad (11)$$

The correlation coefficient between GMs at any two grid points can be written as:

$$\rho = \frac{E[(Y_i^c - \bar{Y}^c_i) * (Y_j^c - \bar{Y}^c_j)]}{\sigma_i^c * \sigma_j^c} = \frac{E[\{(1 - \alpha_i) * R_i\} * \{(1 - \alpha_j) * R_j\}]}{(1 - \alpha_i) * (1 - \alpha_j) * \sigma_i * \sigma_j} = \frac{E[(R_i * R_j)]}{\sigma_i * \sigma_j} \quad (12)$$

This is the same correlation coefficient as that of Eq. 7 for the unconstrained GM fields. This is an interesting result that validates the use of Eq. 8 for imposing constraints on randomly simulated fields of residuals. It is especially attractive since one can control the degree of contribution of the constraining GM data on the simulated random residuals. As will be discussed shortly, this becomes a useful procedure for imposing constraints from MMI contour maps to account for the large uncertainties that arise from inferring the GM from intensity. The formulation of Eq. 8 was presented for the case of imposing GM constraints from a single recording site. However, the formulation can be generalized to cases where the constraints are from multiple recording sites. In these cases, the modified random field can be written as:

$$R_i^c = K * \left(\sum_k \alpha_i^k * y_o^k + \sum_k (1 - \alpha_i^k) * R_i \right) \quad (13)$$

where K is the normalizing factor. It can be easily shown that this formulation maintains the same correlation coefficient as that of Eq. 7 for the unconstrained GM fields.

Fig. 3 shows the correlation coefficients for a randomly selected pair of sites, as a function of distance between the sites. The figure is based on samples of random numbers at each grid, as indicated. In both cases, the general pattern of the correlation coefficient shown on Fig. 2 is captured. However, as expected, the scatter of the correlation coefficients from the theoretical curve, Fig. 2, is larger for the case with the smaller number of samples. The results, presented in Fig. 3, are from a two-dimensional random field with a 1024x1024 grid points, at 1-km intervals. The simulation noise, (i.e., the scatter shown on Fig. 3) increases as the number of degrees of freedom increases. However, the pattern of scatter shown in the figure suggests that a 100-sample simulation can provide an unbiased estimate of the median field.

Ground Motion Simulation of Izmit 1999 Earthquake

We use the formulation of Eq. 13 to simulate a set of stochastic GM footprints for the M7.5 Izmit earthquake of 1999, constrained by both the observed GMs and the MMI contour map (Brian et. al., 2008 and Turkey General Directory of Disaster Affairs, 2009). Fig. 4 shows the earthquake rupture line and the locations of the recording sites. Using Fig. 2 as the basis for constructing the inter-event and intra-site correlation, we simulated 100 sets of random residual fields, normally distributed at each grid point and covering an area of 1024x1024 km² with 1-km

grid intervals. Each set of random fields, as was discussed, represents a single realization of the GM residuals.

Using Eq. 13, the random residuals within 10 km of each recording site were combined with the calculated residuals from the recorded data shown on Fig. 4. The values of α are designed in such a way that grids within 1 km of each recording site get 90% of the contribution from the observed GM data. The contribution from the observed data trends to zero at a distance of 10 km.

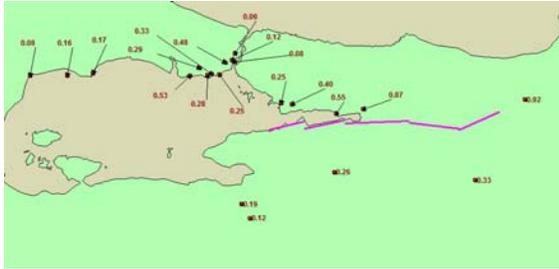


Figure 4. Surface rupture lines and the distribution of GM recording stations, 1999 Izmit earthquake

Fig. 5 shows the correlation coefficients for the simulated ground motions, with constraints, at randomly selected pairs of grid points. The correlation coefficients are shown as a function of the distance between the grid points for cases of using 100 and 300 samples at each point. In both cases, the general pattern of the correlation coefficients are very similar to Fig. 3, which indicates the success of using Eq. 13 for imposing GM constraints without disturbing the spatial correlation effects.

It should be mentioned that, before conducting the GM simulation, we examined the residuals of observed GMs to check for any systematic bias in the set of attenuation equations used in this study for calculating the median GM values. The residuals of the observed ground motions suggest some systematic bias beyond a distance of about 200 km, indicating that GM is overestimated at distant sites. This can be attributed to differences in the Q values for northwestern Turkey versus what might be the representative Q for the set of attenuation equations used in this analysis. We made a Q -related correction to adjust the observed bias. However, considering the low level of ground motions at distances beyond 200 km (where the bias is observed) further discussion of this issue is beyond the scope of this paper and would not affect the results presented here.

Fig. 6 shows four realizations of the 0.3 s S_a footprints for the 1999 Izmit earthquake, all constrained by the observed GM data. All simulations account for the shallow soil site response. The general GM footprints, and close evaluation of the areas near the recording sites (not shown), demonstrate the effectiveness of the methodology for creating a coherent stochastic GM random field. Obviously, more GM data for the earthquake will constrain the randomness of the simulated field and give a more realistic GM footprint for the earthquake. Realistically however, there is never enough data to fully constrain the high-frequency components of an earthquake's

GM field on a regional scale. Therefore, stochastic GM simulations, such as those shown on Fig. 6, are valuable for estimating regional loss distributions.

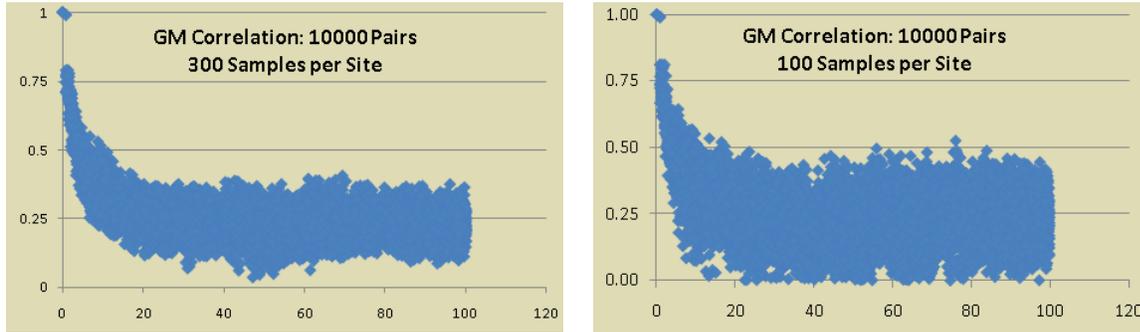


Figure 5. The distribution of the correlation coefficients between simulated GMs at different pairs of grids as a function of the distance between the grids. The simulated GMs include the constraining effects of 29 recorded GM values.

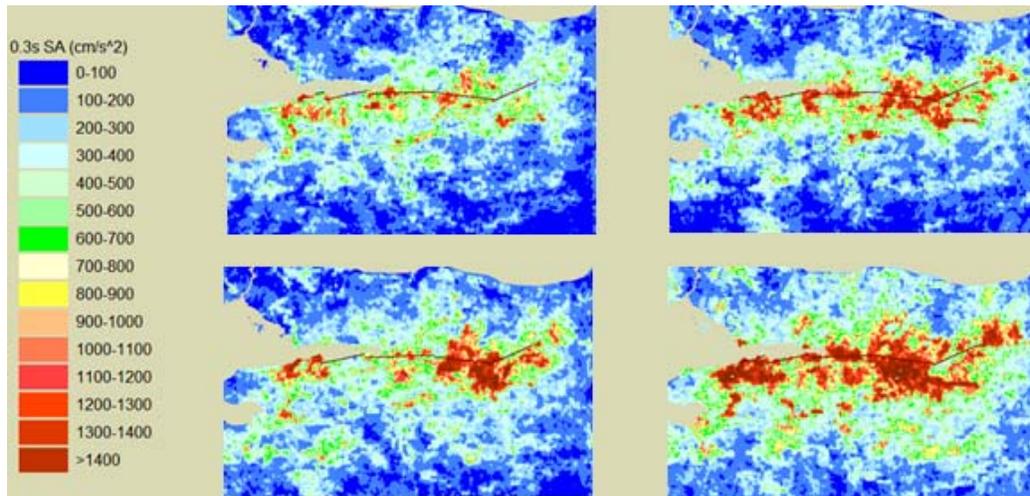


Figure 6. Four footprints of the M7.5 Izmit earthquake of 1999 (0.3 s SA)

For many past earthquakes, GM data is either very limited or not available at all. In such cases, it is desirable to use the earthquake’s MMI data, or contour map, to constrain the GM simulation. Obviously, the MMI data can be converted into GM values to constrain the GM simulation as described earlier. However, considering the qualitative nature of the MMI data and maps, there are large uncertainties in both the accuracy of the data and the MMI-GM empirical equation conversion.

The data integration formulation of Eq. 13 is well suited for accommodating the influence of uncertain GM data on the stochastic simulation. The parameter α in Eq. 13 is a relative measure of the contribution of the observed versus the simulated residuals. For the cases where real GM recordings are used, the value of α at grid points very close to recording sites should be close to one, forcing the simulated GM at these grid points to be very similar to the recorded GM at the sites. However, at sites with inferred GM data based on MMI, where there is uncertainty in

the inferred GM, the values of α at grid points close to the sites can be less than one, reflecting the modeler's confidence in the quality of the inferred data.

The attractiveness of the formulation of Eq. 13 is that it allows us to combine the effects of data that have varying degrees of accuracy. This is a practical way to constrain the GM simulation by mixing real GM recordings with inferred GM data based on MMI. Fig. 7a shows one stochastically simulated GM field, and the MMI contours, for the M7.5 Izmit earthquake of 1999, using one of the sets in Fig. 6. The GM footprint in this figure is constrained only by the 29 recorded GM data points.

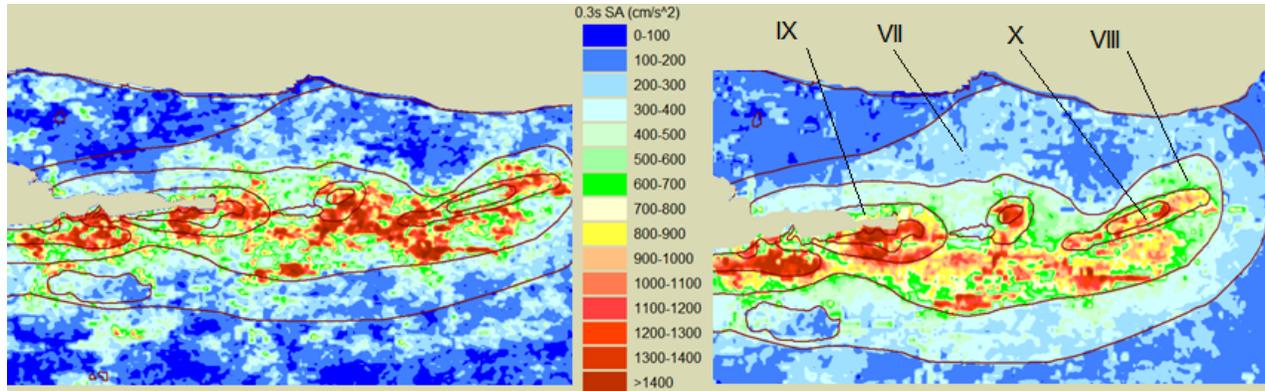


Figure 7. Stochastically simulated GM fields and MMI contours. Left, figure, Fig. 7a, shows a simulation that is constrained only by the 29 recorded GM data. Right, figure, Fig. 7b, shows the GM simulation that is constrained by the 29 recorded GM and inferred MMI-based GM at grid points. Both figures use the same set of random residuals for simulation.

Fig. 7b shows a reconstruction of the GM footprint shown on Fig. 7a but constrained by the inferred GM values from the MMI contour map. For this simulation, the contour map is digitized at 2 km intervals. Using the GM intensity tables (Atkinson and Kaka, 2007), the MMI at each grid point is translated to GM values using the medians for conversion. Using Eq. 13, the inferred, and the 29 recorded, GM residuals are used to constrain the GM simulation. The maximum α values of 0.7 and 0.9 are used in Eq. 13 for the inferred and recorded GM values, respectively. It should be stated that these values are selected for this presentation and can differ, based on modeler's judgment.

The GM footprint presented on Fig. 7b demonstrates the effectiveness of the methodology to stochastically capture the MMI values and regional pattern. Obviously, for the same set of random residual fields, different footprints reflecting different confidence levels on the MMI-based inferred GM data can be simulated. For this study, we sampled the entire MMI map. However, in reality, MMI maps are constructed based on data that are mostly concentrated around cities and towns with damage reports. Therefore, it makes sense to sample the MMI maps around the center of population.

Summary and Conclusion

Estimating the expected loss for an earthquake requires simulating different sets of stochastic GM fields that reflect the inter-event and intra-site spatial correlation, and are constrained by the available earthquake GM data. We presented a practical method for simulating such GM fields. The methodology presented here is versatile in that it allows one to incorporate into the simulation the effects of direct GM recordings as well as the less accurate MMI-based GM data. The data integration methodology is formulated to capture and reflect the modeler's confidence in the quality of GM data used in the simulation. The model presented here provides a practical approach for simulating realistic GM fields for past earthquakes with MMI maps using our present knowledge on the stochastic nature of GM fields and the related effects of spatial correlation.

References

- Bazzurro, P., and J. Park, 2007. "The effects of portfolio manipulation on earthquake portfolio loss estimates", Proceedings of ICASP10, Tokyo, Japan, July 31-August 4.
- Boore, D.M., J.F. Gibbs, W.B. Joyner, J.C. Tinsley, and D. J. Ponti, 2003. Estimated Ground Motion From the 1994 Northridge, California, Bull. Seism. Soc. Am. 93, 2737-2751
- Brian, C., R. Darragh, N. Gregor, and W. Silva, 2008. NGA project strong-motion database, Earthquake Spectra 24, 23-44.
- Keqiang, H. and X.F. Xu, 2009. Probability Upscaling of Material Failure Using Random Field Models – A Preliminary Investigation, Algorithms. 2, 750-763
- Lee, Y., J.G. Anderson, and Y. Zeng, 2005. Evaluation of Empirical Ground-Motion Relations in Southern California, Bulletin of the Seismological Society of America, Vol. 90, No. 6B, pp. S170-S186
- Park, J., P. Bazzurro, and J.W. Baker, 2007. Modeling Spatial Correlation of Ground Motion Intensity Measures for Regional Seismic Hazard and Portfolio Loss Estimation, Proceedings of ICASP10, July 31-August 3, Tokyo, Japan.
- Mahdyiar, M, J. Guin, T. Lai, K. Lum, and K. Shabestari, 2006. Uncertainty in Loss Estimation of a Repeat of 1906 Type Earthquake Due to the Uncertainties in Earthquake Source Parameters and Ground Motions, 8th National Conference on Earthquake Engineering, San Francisco
- Turkey General Directory of Disaster Affairs Earthquake Research Department website, 2009.