Applications of Statistics and Probability in Civil Engineering – Kanda, Takada & Furuta (eds)
© 2007 Taylor & Francis Group, London, ISBN 978-0-415-45211-3

Modeling spatial correlation of ground motion Intensity Measures for regional seismic hazard and portfolio loss estimation

J. Park & P. Bazzurro
AIR Worldwide Corporation, San Francisco, USA

J.W. Baker
Stanford University, Stanford, USA

ABSTRACT: The probabilistic assessment of ground motion intensity measures (IMs; e.g., peak ground acceleration) at an individual site is a standard practice. Less attention has been devoted to estimating the statistical dependence of IMs from a single event at multiple sites due to common source and wave traveling paths and to similar distance to fault asperities. This paper explores the site-to-site correlation of IMs and demonstrate its use in seismic hazard and loss estimation. The use of a spatial correlation function will be shown in two applications: 1) evaluation of the effect of modeling spatially correlated ground motion fields on loss estimates for portfolios of structures with different spacing patterns; and 2) simulation of U.S. Geological Survey (USGS) ground motion ShakeMaps consistent both with the recorded IM values at each station and with the IM correlation structure. This last case could be extremely valuable in USGS-sponsored efforts such as PAGER and ShakeCast.

1 INTRODUCTION

Many private and public stakeholders are strongly affected by the impact of earthquakes on a regional basis rather than on a single property at a specific site. The stakeholders include government and relief organizations that need to prepare for future events and manage emergency response, and private organizations that have spatially distributed assets. Whether for mitigating future seismic risk or managing response after an earthquake, assessment of regional earthquake impact requires a probabilistic description of the ground motion field that an event could or has generated. With knowledge, albeit probabilistic, of the level of ground shaking at a regional level one could more accurately estimate, for example, (1) the monetary losses caused to specific structures or portfolio of structures owned by a corporation or insured by an insurance company, (2) the number of people killed, injured, or displaced in a given region; (3) whether the access to certain critical buildings, such as hospitals, might be restricted due to yellow or red tagging; and (4) the probability that distributed lifeline networks for power, water, and transportation may be interrupted.

The probabilistic assessment of ground motion intensity measures, IMs (e.g., peak ground acceleration or spectral quantities) at an individual site based on the event magnitude, the source-to-site distance, and the local soil conditions is a standard practice that started in the late 1960s. Much less attention has been devoted, however, to estimating the statistical dependence of ground motion IMs from a single event at multiple sites. This paper provides a contribution to filling this research gap. Furthermore, we intend here to explore the effects of site-to-site correlation of ground motion IMs in more depth, and demonstrate their use in seismic loss estimation of spatially extended portfolios of structures.

2 ESTIMATION OF SPATIAL CORRELATION

In general, three effects contribute to the correlation of ground motion IMs at two sites: (a) they have been generated by the same earthquake (e.g., a high stress-drop earthquake may generate ground motions in the region that are, on average, higher at all sites than the median values from events of the same magnitude); (b) the seismic waves travel over a similar path from source to closely spaced; sites and (c) if the dimensions of the fault rupture are large compared to the distance from the sites to the source, adjacent locations may be located close to or far from the asperities on the fault rupture.

Modern ground motion prediction equations implicitly recognize the first cause of dependency via a specific inter-event error term, η_i, as follows:

$$\ln Y_{i,j} = \overline{\ln Y_{i,j}} + \tau\eta_i + \sigma\varepsilon_{i,j} \quad (1)$$

where $Y_{i,j}$ is the ground motion IM of interest (e.g., $S_a(T_1)$), $\overline{\ln Y_{i,j}}$ is the median value of the log of Y predicted by the attenuation equation at site j given the magnitude and distance of earthquake i and the local site conditions, η_i is the aforementioned inter-event standard normal error term, $\varepsilon_{i,j}$ is the site-to-site intra-event standard normal error term, and τ and σ are the corresponding standard deviations of the two error terms, or "residuals."

An alternative formulation for Equation 1, which was common in older prediction equations, is given by

$$\ln Y_{i,j} = \overline{\ln Y_{i,j}} + \tilde{\sigma}\tilde{\varepsilon}_{i,j} \tag{2}$$

where $\tilde{\varepsilon}_{i,j}$ is a random variable representing both the inter-event and intra-event variation at site j from earthquake i. By comparing Equations 1 and 2, it is can be seen that $\tilde{\sigma}$ must equal $\sqrt{\sigma^2 + \tau^2}$ for the variances of the two equations to be equal, and that

$$\tilde{\varepsilon}_{i,j} = \frac{\tau\eta_i + \sigma\varepsilon_{i,j}}{\tilde{\sigma}} \tag{3}$$

In the context of assessing site-to-site correlation of ground motion IMs, it is convenient to use the model in Equation 2 for at least three reasons: (a) there is now only one residual term for each observation ($\overline{\ln Y_{i,j}}$ and $\tilde{\sigma}$ are provided by ground motion prediction equations, and $Y_{i,j}$ is observed, so $\tilde{\varepsilon}_{i,j}$ can be computed directly); (b) the residual term is easy to compute (the values of $\eta_i, i = 1, \ldots, N$, for all the N earthquakes and the frequency-dependent values of τ are usually not included by the developers of ground motion prediction equations in their publications); and (c) Equation 2 is also the form commonly used in probabilistic seismic hazard analysis (PSHA) computer programs, so spatial correlation models in this format can be more easily incorporated into existing software.

An example of observed $\tilde{\varepsilon}_{i,j}$ residuals for peak ground acceleration (PGA) from the 1999 Chi-Chi, Taiwan, earthquake is shown in Figure 1; these residuals, whose value is indicated by the color scale, are the combined residuals from Equations 2 and 3 and represent both the inter- and intra-event error terms.

The intra-event site-to-site ground motion correlation, namely the correlation between the two random variables $\varepsilon_{i,j}$ and $\varepsilon_{i,k}$ at two different sites j and k, due to commonality of wave paths and of distance from asperities has not yet been extensively investigated. Spatial dependence can be observed in Figure 1, by noting that residuals at nearby locations have similar values. This intra-event site-to-site correlation, which is of course not addressed in attenuation relationships for single sites, is crucial for the spatially distributed applications mentioned above.

The limited research on this topic to date indicates that the correlation of peak ground acceleration

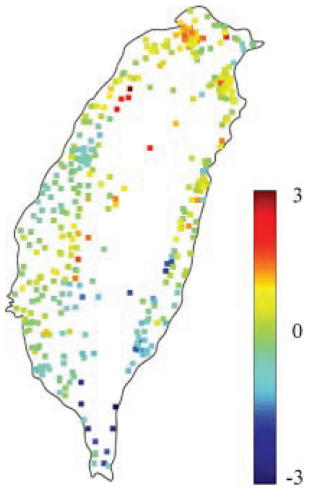

Figure 1. Observed attenuation residuals from the 1999 Chi-Chi, Taiwan, earthquake.

or velocity values decreases with increasing spacing between two sites. This correlation can be estimated by computing empirical correlation coefficients for $\varepsilon_{i,j}$ values at a site separation distance h (plus or minus some tolerance). Because the η_i value is fixed for each ith earthquake, it is effectively a constant when empirical correlation coefficients are estimated from a single earthquake. Thus, correlation coefficients obtained from $\varepsilon_{i,j}$ values or $\tilde{\varepsilon}_{i,j}$ values will be identical, but these correlation coefficients only represent the correlation in the $\varepsilon_{i,j}$ values. To obtain correlation coefficients for the $\tilde{\varepsilon}_{i,j}$ values, one must add the effect of the η_i random variable, which is perfectly correlated at all distances but cannot be detected from the previous empirical correlation coefficients. The total correlation in $\tilde{\varepsilon}_{i,j}$ values is thus

$$\tilde{\rho}(h) = \frac{\tau^2 + \sigma^2\rho(h)}{\tilde{\sigma}^2} \tag{4}$$

where $\rho(h) = \rho_{\varepsilon_{i,j,1},\varepsilon_{i,j,2}}(h)$ is the empirical correlation coefficient calculated for intra-event $\varepsilon_{i,j}$ values separated by a distance h, and $\tilde{\rho}(h) = \rho_{\tilde{\varepsilon}_{i,j,1},\tilde{\varepsilon}_{i,j,2}}(h)$ is the correlation coefficient for the total $\tilde{\varepsilon}_{i,j}$ values defined in Equation 3. Note that for very close sites (i.e., $h \to 0$) the correlation $\tilde{\rho}(h)$ of IMs, of course, tends toward one, whereas for very distant sites (i.e., $h \to \infty$) it is simply given by the ratio of the inter-term-variance to the total variance, as expected. The variance of the difference of the same IM quantity at two sites, k and l, separated by a distance h is simply

$$VAR[X_k - X_l] = 2\sigma^2(1 - \rho(h)) \tag{5}$$

Several models currently exist for these correlation functions (e.g., Boore *et al.*, 2003; Kawakami

and Mogi; 2003; Wang and Takada, 2005; Jeon and O'Rourke, 2005). The models differ in their results, potentially due to differences in the IM parameter (e.g., peak ground acceleration versus peak ground velocity) or in the earthquake events being studied, and to differences in estimation methodologies. Using a comprehensive study of many IM parameters within the framework described here, the authors expect to identify and resolve these differences shortly.

In the simulation steps that follow below, the correlation coefficient $\tilde{\rho}(h)$ in Equation 4, which incorporates correlation due to common inter-event residuals and spatial correlation of intra-event residuals, will be used.

3 SIMULATION OF GROUND MOTION IMS FOR LOSS ESTIMATION

It is a standard practice to model a ground motion IM, $Y_{i,j}$ (e.g., the spectral acceleration $S_a(T_1)$, at period T_1), generated by earthquake i at a site j as a log-normally distributed random variable (RV). In the notation of Equation 2 this implies that its natural logarithm is a Gaussian RV, i.e., that $\ln Y_{i,j} \sim N(\overline{\ln Y_{i,j}}; \tilde{\sigma}^2)$. We now make the reasonable assumption that the $\ln Y_{i,j}$ for all $j = 1, \ldots, M$ sites affected by earthquake i are not only marginally normally distributed, but also jointly normally distributed RV's. Under this assumption, the generation of a M-site random field of ground motion IM values that are spatially correlated involves the generation of a multivariate Gaussian vector of correlated, standard normal random variables, $\mathbf{X} = [X_1, X_2, \ldots, X_M]$, with a (symmetric, positive definite) covariance matrix Σ, or in shorthand $\mathbf{X} \sim N_M(\mathbf{0}, \Sigma)$. Here, $X_j = \tilde{\varepsilon}_{i,j}$ is the total residual term in the model for the computation of $\ln Y_{i,j}$ described by Equations 2 and 3. Note that, because $\tilde{\varepsilon}_{i,j}$ has a standard deviation of one due to normalization, the element of Σ in the k^{th} row and l^{th} column is computed by simply evaluating Equation 4, using the distance between site k and site l to evaluate the correlation function. Earthquake i can be either a realization of a *future event* or of a *past event* for which empirical observations of ground motion IMs at sites $k = 1, 2, \ldots, K$ are available.

In the future event case when no empirical observations of ground motions are available, the generation of $\mathbf{X}$ can be done simply by recognizing that $\mathbf{X} = \mathbf{AU}$ where $\mathbf{U} = [\mathbf{U_1}, \mathbf{U_2}, \ldots, \mathbf{U_M}]$ is a vector of independent standard normal variates and the matrix $\mathbf{A}$ is the Cholesky decomposition (i.e, matrix square root) of Σ, that is the unique lower triangular matrix $\mathbf{A}$ such that $\mathbf{AA^T} = \Sigma$ (e.g., Melchers, 1999).

These $\tilde{\varepsilon}_{i,j}$ values (see Figure 2 for an example based on the correlation function developed by Boore *et al.*, 2003) can then be added to the mean ground motion term, $\overline{\ln Y_{i,j}}$, in Equation 2 to obtain realizations of spatially correlated ground motion IMs. This is precisely what has been done to produce the map of 5%-damped S_a(0.3s) shown in Figure 3, which is based on the $\tilde{\varepsilon}_{i,j}$ field in Figure 2 and is compatible with the rupture and magnitude of the 1994, Northridge earthquake. For the generation of this map we used the digital soil map of Southern California (http://www.consrv.ca.gov/), and the ground motion prediction equation of Abrahamson and Silva (1997). Note that in generating Figure 3, we assumed that the correlation function for PGA developed by Boore *et al.* (2003) is also applicable to the spatial correlation of S_a(0.3s). The map in Figure 3, like the ShakeMap produced by USGS (http://earthquake.usgs.gov/eqcenter/shakemap/sc/shake/Northridge#0.3_sec_Period) is consistent with the same rupture scenario but, unlike the "real" ShakeMap, it does not preserve by design

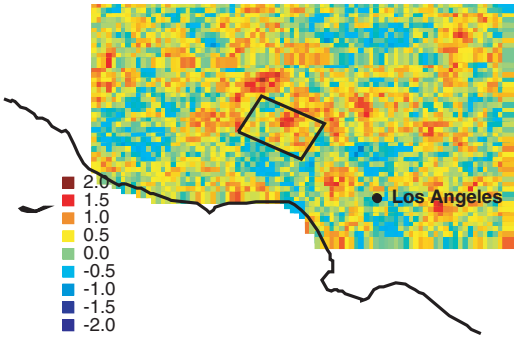

Figure 2. One realization of an $\tilde{\varepsilon}_{i,j}$ random field whose spatial correlation is modeled after Boore et al. (2003). The rectangle is the projection of the 1994 Northridge earthquake rupture in southern California.

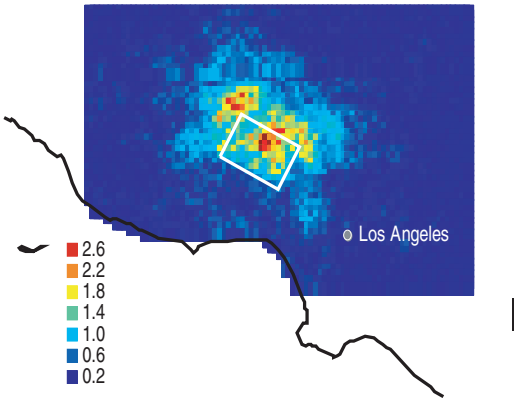

Figure 3. Map of spatially correlated 5% damped S_a(0.3 s) compatible with the rupture (white rectangle) of the 1994 Northridge earthquake. This map was generated using the map of $\tilde{\varepsilon}_{i,j}$ shown in Figure 2.

any of the recordings from the real event and is produced for the geometric mean, not the maximum value, of S_a(0.3s) of the two horizontal components.

Section 5 presents an alternative method for producing maps of correlated ground motion IMs for past events that are consistent with ground motion measurements at the recording stations.

4 PORTFOLIO LOSS ESTIMATION RESULTS

Using the simulation methodology in Section 3, we have studied the seismic risk of two portfolios of structures in the San Francisco Bay Area in Northern California

- **Portfolio 1**: 1,090 properties (Figure 3) in the same county.
- **Portfolio 2**: 133 properties (inset of Figure 3) in the same postal code area.

Both portfolios are made of residential 1- to 3-story woodframe buildings with an assumed replacement value of \$0.5 M each. We will refer to Portfolio 1 as the large portfolio and to Portfolio 2 as the small portfolio. Note that here the adjectives "large" and "small" denote the spatial size of the footprint of the portfolios and not the number of properties contained in them. Similar qualitative conclusions as those presented below would have been found even if the spatially confined Portfolio 2 contained more structures than in the widely spread Portfolio 1.

As with conventional PSHA, the two portfolios were subjected to a large catalog of future earthquake events to estimate the likelihood of losses due to structural and non-structural damage. To study the effects of spatial ground motion correlation modeling on portfolio loss estimates, we generated a random field of S_a(0.3 s) in the affected region for each event in the catalog according to the six approaches listed below. In this loss estimation exercise, the random fields were generated for S_a(0.3 s) because this ground motion IM is a good predictor of the response of stiff woodframe buildings. Note that in all six cases, all the RV's were simulated according to a Gaussian distribution truncated to $\pm 3\tilde{\sigma}$.

1. Independent ground motions at each site (i.e., $\tilde{\varepsilon}_{i,j}$ and $\tilde{\varepsilon}_{i,k}$ are independent RV's if $j \neq k$, and $\rho(h) = 0$).
2. Spatially correlated ground motion field with reduced correlation $\tilde{\rho}(h) = \tilde{\rho} = \tau^2 / (\sigma^2 + \tau^2)$. This expression is derived from Equation 4 after zeroing the $\rho(h)$ term. The spatial correlation modeled in this case is only due to the common value of the inter-event term η_i (see Equation 1) at all sites. The values of $\sigma^2 = 0.66\tilde{\sigma}^2$ and $\tau^2 = 0.34\tilde{\sigma}^2$ for S_a(0.3 s) were taken from Lee *et al.* (2000).

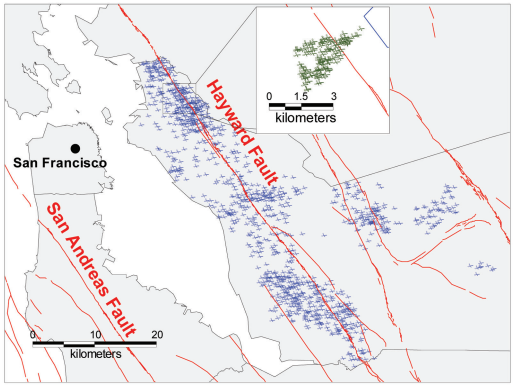

Figure 4. Location of the large and small (inset) portfolios considered for the loss estimation analysis.

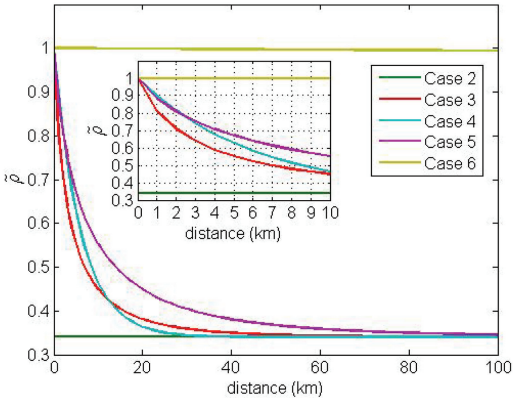

Figure 5. Four ground motion correlation models for $\tilde{\rho}(h)$ as a function of inter-site distance, h, used in the loss estimation analysis.

3. Spatially correlated ground motion field, where the intra-event correlation term $\rho(h) = 1 - \left[1 - e^{-\sqrt{0.6h}}\right]^2$ in Equation 4 is obtained according to the function by Boore *et al.* (2003).
4. Spatially correlated ground motion field with a preliminary alternative correlation function $\rho(h) = e^{-(h/6)}$ developed by Baker (2006) based on an analysis of the residuals for the Chi-Chi earthquake only.
5. As in Case 3 but with the separation distance at which the correlation coefficient has decreased to $\rho = e^{-1}$ (Vanmarcke 1983) suggested by Boore *et al.* doubled (i.e., 0.6 is replaced by 0.3 in the equation above). This case is considered as an upper bound of ground motion correlation.
6. $\tilde{\varepsilon}_{i,j}$ and $\tilde{\varepsilon}_{i,k}$ are assumed to be perfectly positively correlated RV's.

The trend of $\tilde{\rho}(h)$ with site-to-site distance, h, for Cases 2–6 is shown in Figure 5. Cases 3–5 are fully

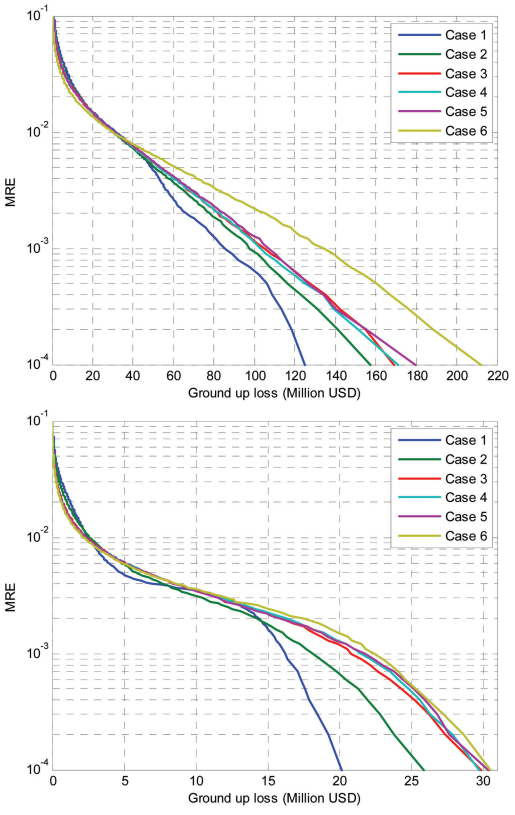

Figure 6. Mean rate of exceedance curves for losses to the large portfolio (top) and small portfolio (bottom) computed using the six proposed models for simulating random fields of ground motion IMs for each earthquake.

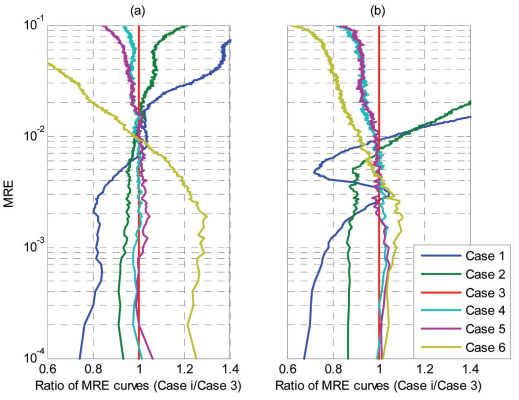

Figure 7. Ratio of MRE curves from Cases 1–6 to MRE curve for Case 3, where Boore *et al*. (2003) correlation function is used. The left panel refers to the large portfolio and the right panel to the small portfolio.

consistent with the methodology in Section 3, Case 2 is considered because, for its simplicity, it has been used to model ground motion correlation in many applications. Case 1 has a place in here because in the overwhelming majority of cases ground motion spatial correlation is simply disregarded. Finally, Case 6 is an unrealistic, extreme case included for comparison purposes only.

The losses expressed as a fraction of the replacement cost, called damage ratio (DR) experienced by a building at given site were simulated from a so-called "damage function" (DF), which is not shown here. A DF is a probabilistic relationship that presents the DR as a function of the severity of the ground motion IM, here Sa(0.3 s), observed at the site. In all six cases the portfolio losses for each event were obtained by simply adding the losses predicted at each site, if any. The losses caused by all the events in the catalog and the annual rate of occurrence of each earthquake were then assembled to produce a curve showing the annual mean rate of exceedance (MRE) of portfolio losses of different amounts. The MRE curves for the six cases above are shown in Figure 6 for both the large and the small portfolios displayed in Figure 4.

The loss exceedance curves in Figure 6 show that the trend of the MRE curve is systematically distorted if the spatial correlation of ground motion IMs is ignored (Case 1). The rare losses are consistently underestimated and the frequent losses are overestimated compared to those produced by Cases 3 and 4, which we consider to be the benchmarks for this exercise. This assertion becomes even more evident when considering the ratio of the MRE curves from Cases 1-6 to the MRE curve of the reference Case 3 as a function of MRE (Figure 7). For the large and small portfolios, the underestimations occur for losses corresponding to MREs less than or equal to 7×10^{-3} and 1×10^{-2}, or mean return periods (MRPs) of about 150 or 100 years and longer, respectively.

As expected, neglecting the intra event site-to-site correlation of IMs, namely setting $\rho(h) = 0$ in Equation 4 (Case 2), produces loss estimates that are less biased than those from Case 1. The underestimation/overestimation of losses is still significant and begins at about the same MRE (and MRP) values as in Case 1. As expected, the bias introduced by neglecting or underestimating the spatial correlation of IMs is more severe for small portfolios than it is for large ones. Neglecting ground motion correlation makes the occurrence of extremely high (or low) ground motion values everywhere in the footprint of the portfolio essentially an impossible occurrence. When a portfolio footprint (here about 4 km for the small portfolio) is within the separation distance of the correlation model (4 km for Case 3 and 6 km for Case 4), then a scenario with significantly higher (lower) than expected ground motions at all sites in the portfolio is a much

more likely event than it is for portfolios with large footprints. At an MRP of 1,000 years, for example, the large portfolio losses are underestimated by 19% in Case 1 and by 7% in Case 2. For the small portfolio the bias in percentage terms increases to 22% for Case 1 and 13% for Case 2.

The loss exceedance curves for Cases 3, 4, and 5, namely those that include more realistic models of the intra-event spatial correlation of IMs, are quite interestingly fairly similar for both portfolios. The losses are within ±10% for all MRE values even for Case 5 where the spatial correlation is, by design, overestimated. Additional MRE curves (not shown here) computed for a portfolio of about 25 structures confined within a 1 km-diameter circle confirm these findings (i.e., ±20% in that case for all MRE values of practical interest).

The similarity of MRE curves produced by different models for spatial correlation of IMs implies that, at least for this application and these two portfolios, the details of the correlation function do not have a significant impact on the losses of portfolios of structures with footprints ranging from one to more than one hundred kilometers. For widely spread portfolios, a significant departure from this cluster of consistent MRE curves can only be obtained either when the site-to-site spatial correlation of IMs is disregarded (Case 1), as discussed above, or when it is unrealistically overestimated (Figure 7 left), such as in Case 6 where the ground motion IMs at different sites are considered perfectly positively correlated RV's regardless of inter-site distance. For portfolios with small footprints, on the other hand, assuming perfect positive correlation of IMs appears to produce a less significant bias in the loss estimates

5 CHARACTERIZATION OF UNCERTAINTY IN SHAKEMAPS

The tools described above can also be modified to simulate spatially correlated ground motions that are also consistent with observed ground motion IMs at individual sites with recording instruments. This approach is useful for rapid post-earthquake ground motion estimation tools such as ShakeMap. By more completely characterizing the ground motion IMs that may have occurred at locations with no recordings, PAGER and other similar tools developed by USGS will be able to more accurately assess the uncertainty in post-earthquake rapid estimates of fatalities and losses.

In this case, the proposed alternative approach uses *conditional* simulation of $\tilde{\varepsilon}_{i,j}$ values at individual locations, *given* observations from recording stations and previously simulated locations. These $\tilde{\varepsilon}_{i,j}$ values can then be added back to mean attenuation predictions in Equation 2 to obtain realizations of ground motion IMs.

The joint realizations of $\tilde{\varepsilon}_{i,j}$ values in a region of interest can be generated using a conditional simulation approach. To do this, the approach described in Section 3 is modified as follows. The vector **X** is still used to describe the joint Gaussian distribution of $\tilde{\varepsilon}_{i,j}$ values at the M sites of interest (the sites where properties are located as well as the sites where recordings were obtained). We now partition this vector into two sub-vectors **X'** (containing the sites of interest for loss analysis) and $\mathbf{X}_{\text{obs}}$ (consisting of the sites with observed ground motion intensities).

By re-ordering the elements of **X** and partitioning the relevant matrices, the distribution of the vector can be denoted

$$\left[\begin{array}{c}\mathbf{X}_{obs}\\ \hline \mathbf{X'}\end{array}\right] \sim N_M\left(\left[\begin{array}{c}\mathbf{0}\\ \hline \mathbf{0}\end{array}\right], \left[\begin{array}{c|c}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{array}\right]\right) \tag{6}$$

where, as before, the elements of Σ are evaluated using Equation 4 using the separation distance between the two associated sites. Nothing in this equation requires knowledge of observed ground motion intensities – the ordering is merely for notation and convenience at the next step.

The observed values of $\tilde{\varepsilon}_{i,j}$ are incorporated into the model by noting that the values of $\mathbf{X}_{\text{obs}}$, are now known. With this information, the conditional distribution of **X'**, given knowledge that the vector of $\tilde{\varepsilon}_{i,j}$ values at $\mathbf{X}_{\text{obs}}$, are equal to **x** is given by

$$\left[\mathbf{X'} \mid \mathbf{X}_{obs} = \mathbf{x}\right] \sim N_M\left(\left[\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{x}\right], \left[\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\right]\right) \tag{7}$$

where $[A|B]$ is used to denote the distribution of A given B. Whereas the original (marginal) mean and covariance of **X'** were **0** and Σ_{22}, respectively, those have been updated based on the observed values. It is interesting to note that with this multivariate normal model, the conditional covariance of **X'**, $\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$, is always less than Σ_{22}, and that it is independent of the observed values **x**.

Once the mean and standard deviation of this conditional distribution have been computed using Equation 7, the Cholesky decomposition and simulation approach described above can proceed as before. Simulated values at the sites with recordings will always equal the recorded value (because the conditional mean is equal to the observed value and the conditional standard deviation is equal to zero). Conditional simulations at sites nearby the recordings will also have reduced standard deviations and adjusted means in accordance with the specified spatial correlation.

The joint realizations of $\tilde{\varepsilon}_{i,j}$ values in a region of interest can also be generated using a series of successive conditional simulations (Goovaerts 1997). First, the conditional distribution at an arbitrary unsampled location, $\mathbf{X'}_1$, is determined, conditional upon values

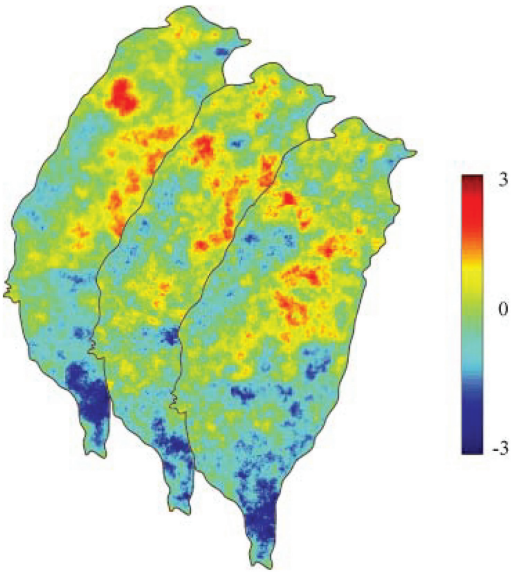

Figure 8. Conditional simulations of $\tilde{\varepsilon}_{i,j}$ from the Chi-Chi earthquake, conditional upon the observed values in Figure 1.

of the originally sampled data points. This conditional distribution is obtained using Equation 7, but now **X'** is only a scalar rather than a vector. A simulated value of $\tilde{\varepsilon}_{i,j}$ is generated using this conditional distribution, and this simulated value is fixed while additional $\tilde{\varepsilon}_{i,j}$ values are simulated conditional upon this observed value. The final simulations are mathematically identical to the one-step simulation approach described in more detail here, but the conditional simulation approach sometimes has computational advantages because the conditioning covariance matrices can be reduced to account for only locations nearby the individual location of interest, thus reducing the size of the Σ_{11} matrix that must be inverted.

Three example $\tilde{\varepsilon}_{i,j}$ simulated random fields that are consistent with observed $\tilde{\varepsilon}_{i,j}$ values from Figure 1 are shown in Figure 8. At every point in the simulations of Figure 8 where an observation is available in Figure 1, the simulations are equal to the observed values. At locations near to recordings, the variance of $\tilde{\varepsilon}_{i,j}$ values from the simulations is reduced because of the correlation with the nearby observed value.

The simulations of Figure 8 were produced using an implementation of Equation 7 in the open-source GSLIB geostatistics software package (Deutsch and Journel 1997). These $\tilde{\varepsilon}_{i,j}$ simulations can be incorporated into Equation 2 to produce simulations of ground motion IMs that are spatially dependent and also consistent with observed intensities at the recording station locations. The simulations can then be used to quantify uncertainty in ShakeMap-type ground motion intensity maps, or to perform post-earthquake portfolio loss analysis that incorporates measured ground motion intensities.

6 CONCLUSIONS AND RECOMMENDATIONS

This paper has shown how empirically-calibrated spatially correlated random fields of ground motion IMs (e.g., PGA or spectral accelerations), can be developed and utilized for hazard analysis and loss estimation.

The sources of correlation in ground motion intensities were discussed, and treatment of these uncertainties in current ground motion prediction models was described. Within this treatment approach, a model for correlation on inter- and intra-event ground motion intensity residuals was described, and an approach for modeling these correlations was presented. Inter-event residuals are perfectly correlated within a specific earthquake, while intra-event residuals are only partially correlated, with a correlation model that must be determined empirically.

Once the correlation model for these residuals was determined, an approach for simulating random fields of correlated ground motion IMs was described. Under the reasonable assumption that the IM residuals have a joint normal distribution, and by utilizing the Cholesky decomposition of the covariance matrix for this joint normal random vector, it is possible to efficiently simulate realizations of this spatially correlated IM field. The correlation modeling and ground motion simulation results were then effectively used in two important applications:

- Generation of spatially correlated random fields of given IMs (similar to the USGS-sponsored ShakeMaps) for both future scenario earthquakes and past events.
- Loss estimation for spatially extended portfolios of structures.

As an example of the first application, we generated maps of S_a(0.3 s) consistent with the Northridge 1994 earthquake fault rupture but not with the Northridge available recordings and others consistent with both the rupture and the recordings of the Chi-Chi, Taiwan, 1999 earthquake. In the latter case the simulated IM maps at sites nearby the recording stations also have reduced standard deviations and adjusted means in accordance with the specified spatial correlation. A set of such maps could be generated in a post-earthquake environment as alternative realizations of the USGS ShakeMaps, which away from the recording stations rely on median values of IMs.

The second application has shown that only by modeling the spatial correlation of IMs can one avoid introducing a bias in portfolio loss estimates. Neglecting or

underestimating the spatial correlation of IMs tends to overestimate frequent losses and underestimate rare ones. The opposite holds if the spatial correlation of IMs is severely overestimated with respect to empirical observations. This study has shown, however, that at least for the cases considered here the portfolio loss estimates seem to be fairly insensitive to the details of the mathematical model adopted to describe the spatial correlation of IMs. This suggests that the choice of the specific correlation model for simulation and loss estimation studies may not be critical, but it does not imply that such correlations can be completely neglected. Note that these findings may not hold true in all cases. More research is needed to evaluate their generality.

REFERENCES

Abrahamson, N. and W. Silva (1997). "Equations for estimating horizontal response spectra and peak acceleration from West North American Earthquakes: A summary of recent work", *Seismological Research Letters*, **68**, 94–127.

Baker J.W. 2006. Personal Communication, April.

Boore D.M., Gibbs J.F., Joyner W.B., Tinsley J.C., and Ponti D.J., 2003. Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California, *Bulletin of the Seismological Society of America,* **93** (6), 2737–2751.

Deutsch, C.V., and Journel, A.G., 1997. *GSLIB geostatistical software library and user's guide*, Oxford University Press, New York.

Goovaerts P., 1997. *Geostatistics for natural resources evaluation*, Oxford University Press, New York, xiv, 483 pp.

Jeon S.-S. and O'Rourke T.D., 2005. Northridge earthquake effects on pipelines and residential buildings, *Bulletin of the Seismological Society of America*, **95** (1), 294–318.

Kawakami, H., and Mogi, H., 2003. "Analyzing Spatial Intraevent Variability of Peak Ground Accelerations as a Function of Separation Distance." *Bulletin of the Seismological Society of America*, 93(3), 1079–1090.

Lee, Y., Anderson, J.G., and Zeng, Y. (2005). "Evaluation of Empirical Ground-Motion Relations in Southern California", *Bulletin of the Seismological Society of America*, Vol. 90, No. 6B, pp. S170–S186.

Melchers, R.E. (1999). "Structural Reliability Analysis and Prediction", 2nd Edition, John Wiley & Sons, Chichester, U.K.

Stanford Center for Reservoir Forcasting, 2006. *The Stanford Geostatistical Modeling Software (S-GEMS)*. http://sgems.sourceforge.net/.

Vanmarcke E., 1983. *Random fields, analysis and synthesis*, MIT Press, Cambridge, Mass., XII, 382 p.

Wang M. and Takada T., 2005. Macrospatial correlation model of seismic ground motions, *Earthquake Spectra*, **21** (4), 1137–1156.