A fully probabilistic framework for spatially distributed slope systems considering spatial cross-correlations of vector intensity measures

W. Du & G. Wang
The Hong Kong University of Science and Technology, Hong Kong, China

ABSTRACT: Earthquake-induced slope displacement is an important parameter for safety evaluation and earthquake design of slope systems. Traditional probabilistic seismic hazard analysis (PSHA) usually focuses on evaluating slope displacement at a particular location, and it is not suitable for spatially distributed slopes. This study proposes a computational efficient framework for fully probabilistic seismic displacement analysis of spatially distributed slope systems using spatially correlated vector intensity measures. First, a spatial cross-correlation model for three key ground motion intensity measures, i.e., peak ground acceleration (PGA), Arias intensity (Ia) and peak ground velocity (PGV) is developed. Monte Carlo simulation and data reduction techniques are utilized to generate spatially-correlated random fields for the vector intensity measures. The slope displacement hazards over the region are further quantified using empirical predictive equations. Finally, an illustrative example is presented to highlight the importance of the spatial correlation and the advantage of using spatially-correlated vector intensity measures.

1 INSTRUCTION

Estimating seismic displacement of natural slopes and earth structures is important for risk assessment of earthquake-induced landslides and performance-based evaluation of key infrastructures. In practice, the seismic slope displacement at a single site can be evaluated using a pseudo-probabilistic or a fully probabilistic approach based on hazard information derived from Probabilistic Seismic Hazard Analysis (PSHA) (Rathje & Saygili 2008). However, quantifying the seismic performance of a slope system over a spatially distributed region rather than at just a single site is critical for a variety of applications, including regional risk assessment of landslide and landslide-related damage to lifelines, road systems and portfolios of infrastructures in this region. Rigorous seismic analysis over a spatially distributed region is less straightforward than that for an individual site.

Two major issues need to be addressed in developing a rational analytical scheme for predicting earthquake-induced slope displacements in a regional scale. First, the spatial cross-correlations between important ground motion intensity measures (IMs) related to the estimation of seismic slope displacement have to be systematically studied. Recently the spatial correlations of some important IMs, such as the peak ground acceleration (PGA) and Arias intensity, have been developed by several researchers (e.g., Jayaram & Baker 2009, Du & Wang 2013). In addition, the cross-correlation between spectral accelerations at multiple periods is investigated (Loth & Baker 2013). Currently, there is no spatial cross-correlation study between PGA and other important parameters (Ia and PGV) available in the literature.

The second major issue is related with computational efficiency. For a fully probabilistic analysis of spatially distributed slopes, Monte Carlo-based Simulation (MCS) is the only feasible approach to rigorously treat all sources of uncertainties. In this study, we propose the following MCS-based framework: First, multiple magnitude-location earthquake scenarios are simulated with frequencies assigned according to recurrence relationships based on the seismicity of the source; Second, for each earthquake magnitude-location scenario, vector-IMs at all sites (called IM maps) will be randomly generated using GMPEs by incorporating inter-event variability and intra-event spatial variability (using the derived spatial cross-correlation matrix) in the process; Third, for each set of vector-IM map, multiple sliding displacement maps will be randomly generated using empirical displacement prediction equations and considering their corresponding uncertainties. In the end, the displacement hazard curve for the whole region can be further computed. The MCS approach would result in increasing computational demanding downstream in this process. Several recent studies have been devoted to

providing some suitable techniques to reduce the computational cost (e.g., Jayaram & Baker 2010, Han & Davidson 2012).

This paper aims at developing a framework for a fully probabilistic analysis of spatially distributed slopes. The spatial correlation between several key IMs most relevant to the prediction of seismic slope displacement, i.e., PGA, Ia and PGV is studied using eleven recent well-recorded earthquakes. A MCS-based computational framework is also developed to rigorously account for all sources of uncertainties. Three data reduction techniques are explored to reduce the computational cost. Following this framework, an illustrative example is also provided in the end.

2 SCALAR AND VECTOR IMS FOR PREDICTING SEISMIC SLOPE DISPLACEMENTS

A suitable prediction model is necessary to predict seismic slope displacement based on ground motion IMs. Since Newmark's pioneering work on the rigid sliding block method (Newmark 1965), the Newmark sliding displacement has important applications in evaluating natural slopes or earthquake-induced landslides. The Newmark displacement analysis assumes that the slope behaves as a rigid-plastic material, and the slope displacement is calculated by double integrating the part of the input acceleration that exceeds a critical value (a_c, which can be determined by the properties of slopes). It provides a simple index of seismic slope performance. In the past, a large number of empirical prediction equations have been proposed to predict the Newmark displacement based on a single (scalar) or multiple intensity measures (a vector IM). The PGA, Ia and PGV were often used as predictors for the Newmark displacement, (e.g., Saygili & Rathje 2008). As earthquake records are complex, transient time series, multiple ground motion IMs are necessary to represent different aspects of ground motion characteristics. The predictive models using a vector IM usually result in reduced aleatory variability (i.e., improved efficiency) and unbiased results for a wide range of earthquake scenarios.

In this study, four recently-developed Newmark displacement prediction equations are chosen based on a scalar IM (termed as PGA model), two-IMs (termed as (PGA, Ia) model and (PGA, PGV) model), as well as three-IMs (termed as (PGA, Ia, PGV) model) as follows (Saygili & Rathje 2008):

(1) PGA model:

$$\ln D = 5.52 - 4.43\left(\frac{a_c}{PGA}\right) - 20.39\left(\frac{a_c}{PGA}\right)^2 + 42.61\left(\frac{a_c}{PGA}\right)^3 - 28.74\left(\frac{a_c}{PGA}\right)^4 + 0.72\ln(PGA)$$

$$\sigma_{\ln D} = 1.13 \quad (1)$$

(2) (PGA, Ia) model:

$$\ln(D) = 2.39 - 5.24\left(\frac{a_c}{PGA}\right) - 18.78\left(\frac{a_c}{PGA}\right)^2 + 42.01\left(\frac{a_c}{PGA}\right)^3 - 29.15\left(\frac{a_c}{PGA}\right)^4 - 1.56\ln(PGA) + 1.38\ln(Ia)$$

$$\sigma_{\ln D} = 0.46 + 0.56\left(a_c / PGA\right) \quad (2)$$

(3) (PGA, PGV) model:

$$\ln D = -1.56 - 4.58\left(\frac{a_c}{PGA}\right) - 20.84\left(\frac{a_c}{PGA}\right)^2 + 44.75\left(\frac{a_c}{PGA}\right)^3 - 30.5\left(\frac{a_c}{PGA}\right)^4 - 0.64\ln(PGA) + 1.55\ln(PGV)$$

$$\sigma_{\ln D} = 0.41 + 0.52\left(a_c / PGA\right) \quad (3)$$

(4) (PGA, Ia, PGV) model:

$$\ln(D) = -0.74 - 4.93\left(\frac{a_c}{PGA}\right) - 19.91\left(\frac{a_c}{PGA}\right)^2 + 43.75\left(\frac{a_c}{PGA}\right)^3 - 30.12\left(\frac{a_c}{PGA}\right)^4 - 1.3\ln(PGA) + 1.04\ln(PGV) + 0.67\ln(Ia)$$

$$\sigma_{\ln D} = 0.2 + 0.79\left(a_c / PGA\right) \quad (4)$$

where D is the predicted sliding displacement in cm; a_c and PGA are in the unit of g; PGV is the peak ground velocity in cm/s; Ia is Arias intensity in the unit of m/s.

3 SPATIAL CROSS-CORRELATION FOR VECTOR IM [PGA, IA, PGV]

3.1 *Ground motion database for the spatial correlation*

A total of 2686 ground motion recordings from eleven earthquakes are compiled to develop the spatial cross-correlation models for PGA, Ia and PGV in this study. These earthquakes occurred in California (1994 Northridge earthquake, 2004 Parkfield earthquake, 2005 Anza earthquake, 2007 Alum Rock earthquake, 2008 Chino Hills earthquake), in Mexico (2010 EI Mayor Cucapah earthquake), in Japan (2000 Tottori earthquake, 2004 Niigata earthquake, 2007 Chuetsu earthquake and 2008 Iwate earthquake) and in Taiwan region (1999 Chi-Chi earthquake).

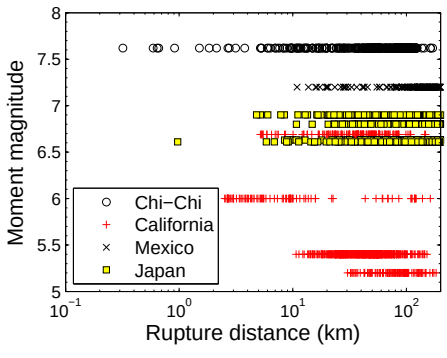


Figure 1. Magnitude and rupture distance distribution of records in the database.

The recorded time histories for these events are obtained from CESMD, CESMOS for US earthquakes and K-NET, Kik-Net for Japan earthquakes. The moment magnitude and rupture distance distribution of the data in the database is illustrated in Figure 1.

3.2 *Geostatistical analysis of intra-event residuals*

Based on ground motion prediction equations (GMPEs), the observed logarithmic IM, denoted as $\ln Y_{ij}$, at site j for an earthquake event i, can be written as follows:

$$\ln Y_{ij} = \overline{\ln Y_{ij}\left(M,R,\theta\right)} + \eta_i + \varepsilon_{ij} \tag{5}$$

where $\ln Y_{ij}(M,R,\theta)$ is the predicted median IM based on magnitude (M), rupture distance (R) and other variables (θ);η_i is the inter-event residuals with zero means and standard deviations of τ_i; and ε_{ij} denotes the intra-event residuals with zero means and standard deviations of σ_{ij}. Both η_i and ε_{ij} are assumed to be normally distributed independent random variables (Abrahamson & Youngs 1992). In this study, GMPEs developed by Campbell & Bozorgnia (2008, 2012) are used for PGA, PGV and Ia respectively.

For a vector **IM**=[PGA, Ia, PGV], the intra-event residuals can be assumed to follow a multivariate normal distribution. Under this assumption, the intra-event residuals $\boldsymbol{\varepsilon}_{ij} = \left(\varepsilon_{ij}^1,...,\varepsilon_{ij}^n\right)$ for n IMs for earthquake event i at site j can be fully obtained by their mean (zero vector in this case) and covariance matrix.

Semivariogram is a widely used statistical tool to estimate the spatial correlation of random variables. The semivariogram can be defined as measuring the average dissimilarity between two second-order stationary random variables Z_i and Z_j separated by a distance vector $\boldsymbol{h}$ as follows (Goovaerts 1997):

$$\gamma_{ij}\left(\boldsymbol{h}\right) = \frac{1}{2}E\left[\left(Z_i(u+\boldsymbol{h}) - Z_i(u)\right)\left(Z_j(u+\boldsymbol{h}) - Z_j(u)\right)\right] \tag{6}$$

where $Z_i(\boldsymbol{u})$ and $Z_i(\boldsymbol{u}+\boldsymbol{h})$ are variable Z_i evaluated at position $\boldsymbol{u}$ and at a position separated by a distance vector $\boldsymbol{h}$, respectively. In this study, Z_i and Z_j refer to the intra-event residuals of the i-th and the j-th component of the vector **IM**. Under the assumptions that the spatial field is isotropic and second-order stationary, a scalar variable $h = \left\|\boldsymbol{h}\right\|$ can be used in the formulation.

An exponential functional form can be used to fit the above empirical semivariogram data:

$$\tilde{\gamma}_{ij}\left(h\right) = a\left[1 - \exp(-3h/b)\right] \tag{7}$$

where a is the *sill* of the semivariogram, and b is the *range* of the semivariogram, defined as the separation distance h at which $\tilde{\gamma}_{ij}\left(h\right)$ equals 95% of the sill. The fitting parameters a and b can be obtained by least square or manual fit method. It is straightforward to show that the following relationship holds between the covariance function and the semivariogram function (Goovaerts 1997, p. 72-74):

$$C_{ij}(h) = \lim_{(h\to\infty)}\gamma_{ij}\left(h\right) - \gamma_{ij}\left(h\right) = C_{ij}(0) - \gamma_{ij}\left(h\right) \tag{8}$$

The unit-free correlation coefficient between two variables Z_i and Z_j is:

$$\rho_{ij}\left(h\right) = C_{ij}(h)\Big/\sqrt{C_{ii}(0)\cdot C_{jj}(0)} \tag{9}$$

Accordingly, the covariance matrix $\boldsymbol{C}(h)$ for the n-component vector **IM** is defined as follows:

$$\mathbf{C}(h) = \left[C_{ij}(h)\right] = \begin{bmatrix} C_{11}(h) & \cdots & C_{1n}(h) \\ \vdots & \ddots & \vdots \\ C_{n1}(h) & \cdots & C_{nn}(h) \end{bmatrix} \tag{10}$$

Therefore, the total covariance matrix $\boldsymbol{\Sigma}$(event i) can be implemented by submatrix $\boldsymbol{C}(h)$ in Equation (10) as follows:

$$\boldsymbol{\Sigma}(\text{event } i) = \begin{bmatrix} \boldsymbol{C}(0) & \cdots & \boldsymbol{C}(h_{1J}) \\ \vdots & \ddots & \vdots \\ \boldsymbol{C}(h_{J1}) & \cdots & \boldsymbol{C}(0) \end{bmatrix} \tag{11}$$

where h_{ij} represents the specific separation distance between site i and site j among a total of J sites (the separation distance is always zero for diagonal elements).

In summary, the total covariance matrix for the n-component vector **IM** at J sites can be obtained once the correlation range b in Equation (7) is obtained by semivariogram regression. The procedure (called "direct fit method") is straightforward and efficient. However, the total covariance matrix obtained by the direct fit method cannot guarantee the positive-definiteness, making it difficult to generate spatially correlated random field in application. Hence, a statistical approach termed as the linear model of coregionalization (LMC) will be subsequently adopted in this study to overcome the above limitation, such that the resulted total covariance matrix will guarantee to be positive-definite.

3.3 *Coregionalization matrix for vector IM [PGA, Ia, PGV]*

The linear model of coregionalization (LMC) can be used to decompose the semivariograms $\gamma_{ij}(h)$ as a linear combination of independent random functions. Accordingly, the semivariogram matrix $\boldsymbol{\Gamma}(h)$ can be decomposed as (Journel & Huijbregts 1978, p.171-173):

$$\boldsymbol{\Gamma}(h) = \sum_{l=1}^{L}\boldsymbol{B}^l g_l(h) \tag{12}$$

where $\mathbf{B}^l=[b_{ij}^l]$ is called the coregionalization matrix. It is to be noted that as long as the positive definiteness of matrix $\mathbf{B}^l$ is satisfied, the total covariance matrix is guaranteed to be positive definite regardless of the number of sites located in this region. This condition can be easily satisfied since $\mathbf{B}^l$ is just a $n\times n$ matrix (n is the number of IMs considered).

In this study, a short range (10 km) and a long range (60 km) exponential functions are used as the basic functions of $\boldsymbol{\Gamma}(h)$, h is in the unit of km:

$$\boldsymbol{\Gamma}(h) = \mathbf{B}^1\left(1-\exp\left(\frac{-3h}{10}\right)\right)+\mathbf{B}^2\left(1-\exp\left(\frac{-3h}{60}\right)\right) \quad (13)$$

Accordingly, the spatial correlation coefficient matrix can be obtained as:

$$\mathbf{R}(h) = \mathbf{P}^1\left(\exp\left(\frac{-3h}{10}\right)\right)+\mathbf{P}^2\left(\exp\left(\frac{-3h}{60}\right)\right) \quad (14)$$

where $\mathbf{R}(h)$ is the correlation matrix, $\mathbf{P}^1$ and $\mathbf{P}^2$ are standardized versions of $\mathbf{B}^1$ and $\mathbf{B}^2$.

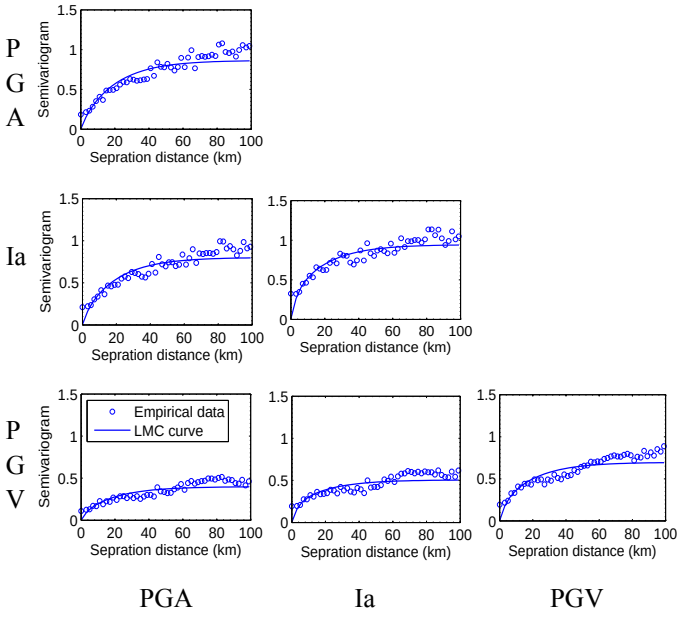


Figure 2. Cross-semivariograms and fitted LMS curves for the Chi-Chi earthquake.

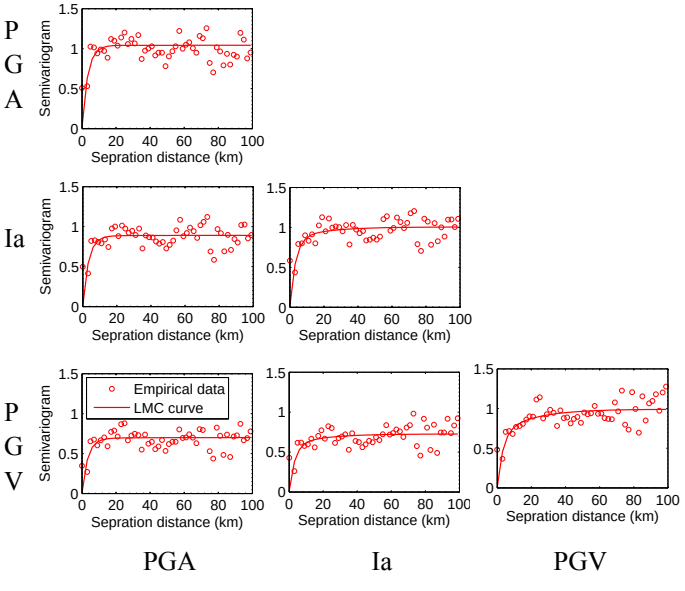


Figure 3. Cross-semivariograms and fitted LMS curves for the Northridge earthquake.

An iterative algorithm can be used to obtain a positive-definite matrix $\mathbf{B}^l$ efficiently for the LMC model. This method minimizes the weighted sum of squares of differences between the empirical semivariograms and the one estimated by the LMC model. The algorithm checks and ensures the positive-definiteness of $\mathbf{B}^l$ matrix during iterations. One can refer to (Goulard & Voltz 1992) for the details.

Following the above procedures, the cross-semivariograms with respect to separation distances for PGA, Ia and PGV, along with the fitted curves by the LMC method are illustrated in Figure 2 and Figure 3 for the Chi-Chi and the Northridge earthquake, respectively. All fitting curves approximate the empirical data reasonably well for each case. Similar results can be observed for other earthquakes considered in this study.

The final coregionalization matrix $\mathbf{P^1}$ and $\mathbf{P^2}$ for the vector **IM**=[PGA, Ia, PGV] can be obtained by averaging the coregionalization matrices obtained from each earthquake event as follows:

$$\mathbf{P}^1 = \begin{bmatrix} 0.61 & 0.57 & 0.39 \\ 0.57 & 0.67 & 0.46 \\ 0.39 & 0.46 & 0.50 \end{bmatrix} \quad \mathbf{P}^2 = \begin{bmatrix} 0.39 & 0.34 & 0.22 \\ 0.34 & 0.33 & 0.22 \\ 0.22 & 0.22 & 0.50 \end{bmatrix} \quad (15)$$

Both $\mathbf{P^1}$ and $\mathbf{P^2}$ are positive definite, which as we introduced earlier, is vitally important for stochastic simulation of spatially correlated fields. These proposed correlation matrices will be used to generate spatially correlated fields of PGA, Ia and PGV for calculating the Newmark displacement.

4 STOCHASTIC SIMULATION USING DATA REDUCTION TECHNIQUES

4.1 *Importance sampling (IS) technique*

Importance sampling is a widely used data reduction technique to sample earthquake scenarios. Generally speaking, earthquake magnitudes follow some recurrence relationships (e.g., Gutenberg-Richter law). Random sampling of earthquake scenario is inefficient because large magnitude events are infrequently sampled although they are more important in hazard analysis. Instead, the importance sampling technique preferentially samples the rare large events. The effects of IS technique are accounted for through assigning suitable weights to each sampling so that the occurrence rate of the earthquake scenarios can still be correctly represented. The procedure is introduced as follows:

(1) Rupture location is assumed to be uniformly distributed within each earthquake source zone, and a magnitude density function $f(m)$ is used to characterize each earthquake source. The range of magnitude (between a lower bound M_{min} and an upper bound M_{max}) is divided into n_m intervals. The interval can vary with magnitude (i.e., a smaller interval for larger magnitudes).

(2) A magnitude can be randomly selected within each interval [m_k, m_{k+1}], with an actual probability of $\int_{m_k}^{m_{k+1}} f(m)dm$. So a total of n_m earthquake events can be sampled for each rupture location. The sampling

probability is $1/n_m$ for each event.

(3) The importance sampling weight for each sampled scenario k is computed as:

$$IS_k = \int_{m_k}^{m_{k+1}} f(m)dm \Big/ 1/n_m \tag{16}$$

(4) If a total of N_M magnitude-location scenarios are generated from all earthquake sources, the actual annual occurrence probability for scenario j is assigned as:

$$P_j = \frac{IS_j}{\sum_{i=1}^{N_M} IS_i} \tag{17}$$

4.2 *Stratified sampling (SS) method*

In conventional MCS, a constant number of ground motion maps are generated for each earthquake scenario. Yet, the number of intensity maps can be reduced for some unimportant events (e.g., a small-magnitude far-distance event). Using the stratified sampling (SS) technique (Cochran, 1977), an optimal number of intensity maps can be assigned for each scenario. Let N_j denote the number of corresponding intensity maps for the j-th event. The total number of intensity maps is $N = \sum_{j=1}^{J} N_j$ for all events. The value of N_j for each event can be given as:

$$N_j = N \cdot \left[\frac{P_j \sqrt{P(y_{ij} \geq Y_{ir}) \cdot \left[1 - P(y_{ij} \geq Y_{ir})\right]}}{\sum_{j=1}^{J} P_j \sqrt{P(y_{ij} \geq Y_{ir}) \cdot \left[1 - P(y_{ij} \geq Y_{ir})\right]}} \right] \tag{18}$$

This method results in an optimal number of ground motion IM maps, N_j for each scenario j for each site i that minimizes the generated and the analytical hazard curves at a given return period r. We used four return periods (r =100, 475, 1000 and 2475 years) in this study. An averaged value of N_j over all sites, all return periods and all IMs is chosen as the final number of ground motion maps to be generated for each scenario. Finally, the annual occurrence probability for each ground motion intensity map is:

$$P_n = \frac{P_j}{N_j} \quad (n=1,\ldots,N) \tag{19}$$

4.3 *Summary of the fully probabilistic analysis procedure*

A fully probabilistic analysis procedure is summarized in the following steps:

Step 1: A set of earthquake magnitude-location scenarios can be simulated using stochastic method following magnitude-recurrence relationships. Important sampling technique is used in this step to reduce the number of samplings.

Step 2: The median predicted values of IMs and their corresponding standard deviation (τ_i and σ_{ij}) of the inter/intra-event residuals for each site are computed using Equation (5).

Step 3: The inter-event residuals (η_{ij}) are randomly generated following univariate normal distribution $\eta_{ij} = \mathcal{N}(0, \tau_i)$. The spatially-correlated intra-event residuals (ε_{ij}) are randomly generated for n IMs at all sites following multivariate normal distribution with a zero mean, standard deviation (σ_{ij}) and the total spatial correlation matrix **R**.

Step 4: Ground motion IM maps are calculated by combining the median, inter-event and intra-event residuals for each scenario. The SS method is used to determine the number of ground motion IM maps need to be generated, and P_n is calculated via Equation (19) for each IM map.

Step 5: The median predicted Newmark displacement and the standard deviation σ_D are computed for each IM map using predictive models. Displacement residuals are generated following univariate normal distribution $\mathcal{N}(0, \sigma_{\ln D})$.

Step 6: A total number of N_D displacement maps are retained using the SS method. For j-th displacement map, the corresponding probability is calculated as $P_{Dj} = P_n / N_n$, where N_n is the assigned number of displacement maps for the n-th IM map by SS method. Finally, the annual probability of exceedance λ_{D^*} for specific value D^* for site i can be computed as:

$$\lambda_{D^*} = \sum_{j=1}^{N_D} P_{Dj} \cdot P(D_{ij} \geq D^*) \tag{20}$$

where $P(D_{ij} \geq D^*)$ is again a binary function (equals 1 if 'true', and 0 otherwise).

5 AN ILLUSTRATIVE EXAMPLE

5.1 *Problem description*

In this section, a hypothetical area is investigated by the proposed fully probabilistic approach. The 30 km × 30 km area is divided into 900 sites separated by 1 km × 1 km in distance. A constant critical acceleration $a_c = 0.1g$ is assigned to all sites. A 5 km-long linear fault is located close to this area, shown in Figure 4(a). The following Gutenberg-Richter relationship is assumed to describe the seismicity of the source:

$$\log \lambda_m = 4.4 - 1.0 \times M_w \tag{21}$$

where λ_m is the mean annual rate of exceedance of the moment magnitude M_w. The minimum and maximum magnitudes are set as 4.4 and 7.5, respectively. The linear source is divided into five 1-km long segments. The location of rupture scenarios is randomly distributed within each segment. The GMPEs (Campbell & Bozorgnia 2008, 2012) are used to estimate the predicted median values for PGA, PGV and Ia, respectively. The four aforementioned models Equations (1)-(4) are adopted to compute the Newmark displacement.

The IS method is applied to stratify the range of

magnitudes. The partition interval is 0.3 for $4.4 \le M_w < 5.6$, 0.2 for $5.6 \le M_w \le 6.6$, and 0.1 for $M_w > 6.6$. 18 scenarios are sampled within each fault segment, resulting in a total of 90 scenarios considered. After applying the SS method, 400 maps are generated for each IM (PGA, Ia and PGV) in the "reduced set" with assigned probability P_n. One group of generated IM maps for PGA, Ia and PGV are shown in Figure 4 for demonstration. By applying SS method, a total of 3000 displacement maps are generated in the reduced set each with assigned probability P_D.

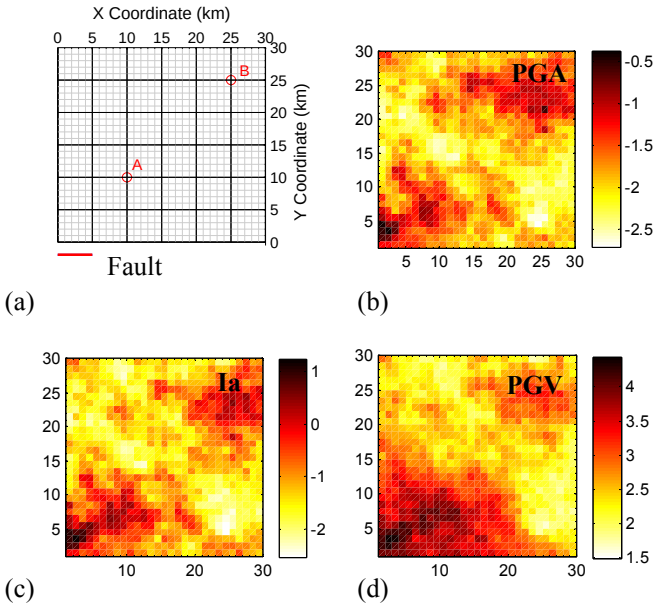


Figure 4. (a) A 30 km×30 km area divided into 1 km×1km grids showing the location of the fault trace. Site A and Site B are two selected sites for comparison. Examples of spatially-correlated vector IM fields for (b) PGA (in natural log scale, unit of g), (c) Ia (in natural log scale, unit of m/s) and (d) PGV (in natural log scale, unit of cm/s), respectively.

After the earthquake scenarios are sampled using IS, the conventional MCS method is used to generate 100 sets of inter and intra-event residuals for each intensity measure and each earthquake scenario for comparison (i.e., a total of 9000 IM maps, called the "large set"). For each group of intensity maps, the MCS is also carried out to get the 60 sets of displacement maps, resulting in a total number of 540000 displacement maps.

5.2 *Hazard consistency of the reduced sets*

The hazard consistency is checked by comparing the intensity and displacement hazard curves obtained from the reduced set and the large set of IM maps and displacement maps. The simulated intensity hazard curve can be obtained by calculating the probability of exceedance for each IM and each site as:

$$\lambda_{y^*} = \sum_{n=1}^{N} P_n \cdot P(y_{in} \ge y^*) \tag{22}$$

where P_n is the probability for the n-th IM map, y_{in} represents IM value at site i on the n-th IM map; $P(y_{in} \ge y^*)$ is binary function (equals 1 if $y_{in} \ge y^*$, 0 otherwise). On the other hand, the analytical ('true') intensity hazard curve for intensity measure Y at each site can be computed using PSHA approach:

$$\tilde{\lambda}_{y^*} = \iint P(Y \ge y^* | m,r) f(m) f(r) dm dr \tag{23}$$

where $P(Y \ge y^* | m,r)$ is computed using GMPEs by assuming lognormal distribution of IM.

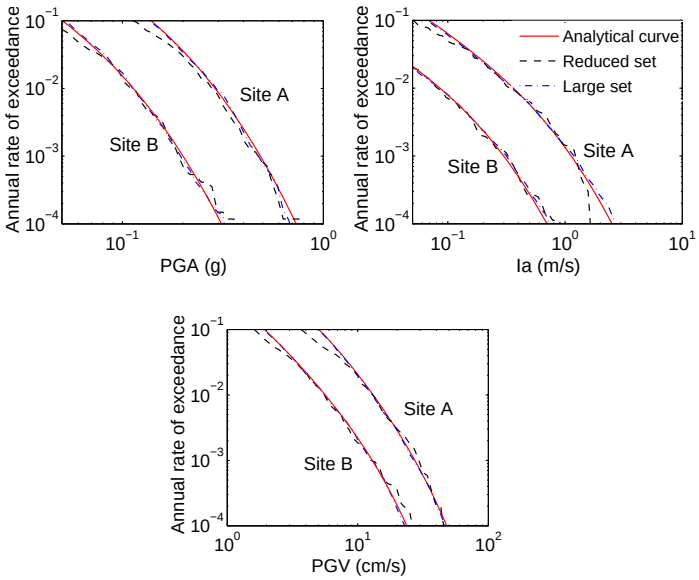
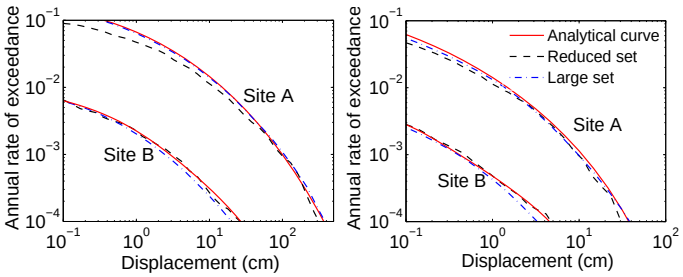


Figure 5. Comparison of seismic hazard curves obtained using the reduced IM maps (400 maps) and the large IM maps (9000 maps) method for PGA, PGV and Ia at site A and site B.



(a) PGA model (b) (PGA, Ia) model

Figure 6. Comparison of displacement hazard curves at two sites using the reduced set and the large set. (a) PGA model, and (b) (PGA, Ia) model.

Intensity hazard curves obtained from the reduce set and the large set are compared with the analytical hazard curves for PGA, Ia and PGV in Figure 5 for two representative sites A and B shown in Figure 4(a). Quite consistent results can be observed for both sites and all IMs. The displacement hazard

curves for the two sites are also compared in Figure 6 using the scalar (PGA) model and (PGA, Ia) model. It can be seen that the displacement hazard curves obtained by using data reduction technique are in reasonable agreement with these obtained using conventional MCS method, although the former only requires about one-180th number of realizations of the latter. Also, all these hazard curves obtained from sampling maps are consistent with the analytical hazard curves. By this example, it is demonstrated that the proposed computational framework can result in stable, fast and hazard-consistent results, and can be used to estimate the seismic risk over a large region.

5.3 *Importance of spatial correlation*

The importance of spatial correlation on the regional-scale hazard analysis is highlighted by considering several special cases. Given a specified value of D (denoted as D^*) and its exceedance area ratio AR^* (defined as the ratio of the areas where displacements exceed the specified D^* value against the total area of the region), the annual rate of exceedance (termed as "aggregated displacement hazard curve") can be computed as:

$$\lambda = \sum_{j=1}^{N_D} P_{Dj} P(D > D^* \,\&\, AR > AR^*) \qquad (24)$$

where P_{Dj} is the occurrence probability for displacement map j, and $P(D > D^* \,\&\, AR > AR^*)$ is a binary function (equals 1 when 'YES' and 0 otherwise).

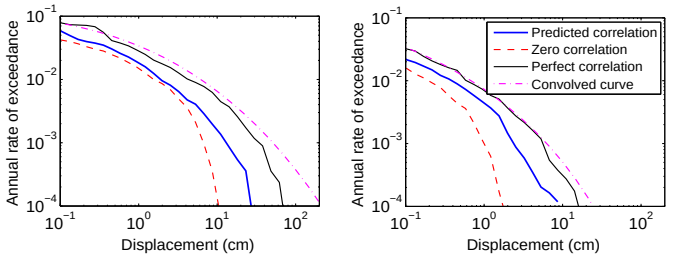


Figure 7. Displacement hazard curves for exceedance area ratio AR^* as 25% using (a) PGA model, and (b) (PGA, Ia) model

Figure 7 shows the aggregated displacement hazard curve using the scalar (PGA) model and the (PGA, Ia) model, by assuming that AR^* is 25%. The spatial correlation of IMs are assumed to follow (1) zero correlation (i.e., correlation range is zero), (2) predicted spatial correlation via Equations (14-15) and (3) perfect correlation (i.e., correlation range is infinite). The results demonstrate that, ignoring spatial correlation would yield an underestimated displacement hazard curve, especially for the rare cases. The case of perfect correlation, on the other hand, would overestimate the displacement hazard. For instance, given an annual rate of exceedance of 4×10^{-4}, the predicted displacement is 8.4 cm if the spatial correlation is zero; 21.7 cm if the predicted spatial correlation is considered and 47 cm for the perfectly-correlated case. In addition, the convolved analytical curve closely agrees with the perfectly-correlated case, indicating the sites are assumed to be perfectly correlated in this process. Hence, the convolved analytical solution will lead to overestimated risk estimate for spatially distributed slopes. Although only the scalar (PGA) model and the vector (PGA, Ia) model are used in Figure 7, similar conclusion can be drawn if other vector models are used.

5.4 *Influence of displacement prediction models*

The displacement hazard curves obtained using four different displacement prediction models are compared. Figure 8 shows the displacement hazard curves for individual sites (A and B), as well as for a given area ratio AR^* of 5% and 25%, respectively. The spatial correlation of vector IMs are computed using Equation (15). For all these cases, the displacement hazard curves obtained from the scalar PGA model are significantly higher than these obtained from vector models. The large discrepancy is not unexpected since a scalar displacement prediction model usually cannot satisfy the sufficiency requirement, e.g., the model exhibits systematic bias over earthquake magnitudes (Rathje and Saygili, 2008). On the other hand, rather consistent results are obtained using three vector-IM models, since the sufficiency requirement can be more easily satisfied using a vector model. The results clearly demonstrated the advantage of using vector IMs in displacement hazard analysis.

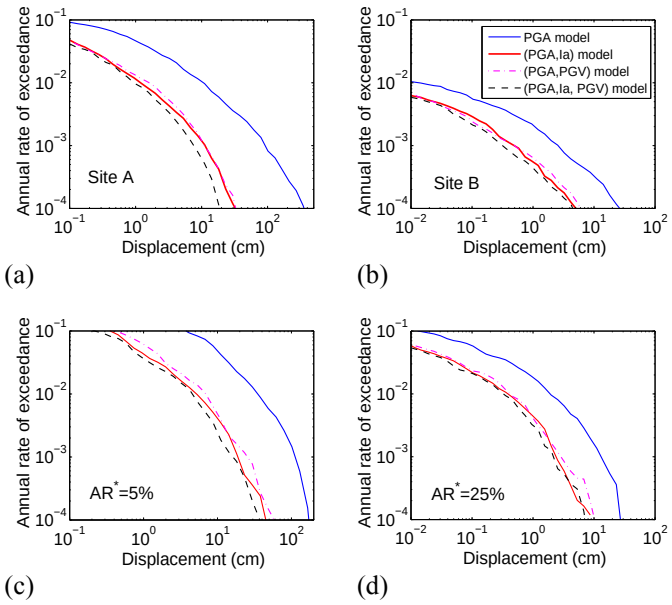


Figure 8. Displacement hazard curves using by different displacement prediction models for (a) site A, (b) site B, (c) exceedance area ratio AR^* as 5%, and (d) AR^* as 25%

6 CONCLUSIONS

This paper provided a computational efficient framework to conduct a fully probabilistic hazard analysis for spatially distributed slopes. The cross-correlations between the vector **IM** [PGA, Ia, PGV] are developed based on geostatistical modeling of strong motion data from eleven recent earthquakes. The coregionalization matrices provide a positive definite covariance matrix that enables generation of random fields of the vector IM preserving their spatial correlations.

The developed covariance model for the vector **IM** [PGA, Ia, PGV] is then used to quantify the spatial variability in a hypothetical region. The intensity maps and sliding displacement maps are generated by Monte Carlo method, and aleatory variability is incorporated in each step. To reduce the computational cost, several state-of-the-art data reduction techniques are also applied. The difference of hazard curves between 'reduced set' and 'large set' implies that results obtained from data reduction techniques can provide accurate results. The importance of spatial correlation is also emphasized using the example. Neglecting the spatial variability or using the convolved analytical solution would result in either underestimated or overestimated displacement hazard curves.

Finally, the displacement hazard curves computed by different predictive models (PGA model, (PGA, Ia) model, (PGA, PGV) model and (PGA, Ia, PGV) model) are compared. Except for the scaler (PGA) model, all vector models yield consistent displacement hazard curves for individual sites as well as for the whole region. The vector models demonstrate significant advantages over the scalar model, and underline the importance of using spatially-correlated vector **IM** in seismic hazard analysis of spatially distributed slopes.

ACKNOWLEGEMENTS

The authors acknowledge financial support from Hong Kong Research Grants Council Grant 620311, and DAG08/09.EG13, DAG11EG03G to conduct this study. Any opinions, findings, and conclusions or recommendations expressed in this material are ours and do not necessarily reflect those of the sponsor.

REFERENCES

Abrahamson N.A. & Youngs R.R. 1992. A stable algorithm for regressionanalyses using the random effects model. *Bulletin of the Seismological Society of America* 82:505-510.

Campbell K.W. & 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 0.1 to 10 s. *Earthquake Spectra* 24(1): 139–171.

Campbell K.W. & Bozorgnia Y. 2012. A comparison of ground motion prediction equations for Arias intensity and cumulative absolute velocity developed using a consistent database and functional form. *Earthquake Spectra* 28(3): 931-941.

Cochran W. 1977. Sampling Techniques (3rd edn). Wiley: New York.

Du W. & Wang G. 2013. Intra-event spatial correlations for cumulative absolute velocity, Arias intensity and spectral accelerations based on regional site conditions. *Bulletin of the Seismological Society of America* 103(2A): 1117-1129.

Goovaerts P. 1997. Geostatistics for natural resources evaluation. Oxford University Press: Oxford, New York.

Goulard M. & Voltz M. 1992. Linear coregionalization model: tools for estimation and choice of crossvariogram matrix". *Mathematical Geology* 24(3): 269-286.

Han Y. & Davidson R.A. 2012. Probabilistic seismic hazard analysis for spatially distributed infrastructure. *Earthquake Engineering & Structural Dynamics* 41(15): 2141-2158.

Jayaram N. & Baker J.W. 2009. Correlation model for spatially distributed ground-motion intensities. *Earthquake engineering and Structural Dynamics* 38:1687-1708.

Jayaram N. & Baker J.W. 2010. Efficient sampling and data reduction techniques for probabilistic seismic risk assessment. *Earthquake Engineering & Structural Dynamics* 39: 1109-1131.

Journel A.G. & Huijbregts C.J. 1978. Mining Geostatistics, Academic Press, London; 600 pp.

Loth C. & Baker J.W. 2013. A spatial cross-correlation model of ground motion spectral accelerations at multiple periods. *Earthquake Engineering & Structural Dynamics* 42(3): 397-417.

Newmark N.M. 1965. Effects of earthquakes on dams and embankments. *Geotechnique* 15(2): 139-160.

Rathje E.M. & Saygili G. 2008. Probabilistic seismic hazard analysis for the sliding displacement of slopes: scalar and vector approaches. *Journal of Geotechnical and Geoenvironmental Engineering* 134(6): 804-814.

Saygili G. & Rathje E.M. 2008. Empirical predictive models for earthquake-induced sliding displacements of slopes. *Journal of Geotechnical and Geoenvironmental Engineering* 134(6): 790-803.