Bulletin of the Seismological Society of America, Vol. 96, No. 1, pp. 215–227, February 2006, doi: 10.1785/0120050060

Correlation of Response Spectral Values for Multicomponent Ground Motions

by Jack W. Baker and C. Allin Cornell

Abstract Ground-motion prediction (attenuation) models predict the probability distributions of spectral acceleration values for a specified earthquake event. These models provide only marginal distributions, however; they do not specify correlations among spectral accelerations with differing periods or orientations. In this article a large number of strong ground motions are used to empirically estimate these correlations, and nonlinear regression is used to develop approximate analytical equations for their evaluation. Because the correlations apply to residuals from a ground-motion prediction, they are in principle dependent on the ground-motion prediction model used. The observed correlations do not vary significantly when the underlying model is changed, however, suggesting that the predictions are applicable regardless of the model chosen by the analyst. The analytical correlation predictions improve upon previous predictions of correlations at differing periods in a randomly oriented horizontal ground-motion component. For correlations within a vertical ground motion or across orthogonal components of a ground motion, these results are believed to be the first of their kind.

The resulting correlation coefficient predictions are useful for a range of problems related to seismic hazard and the response of structures. Past uses of previous correlation predictions are described, and future applications of the new predictions are proposed. These applications will allow analysts to better understand the properties of single- and multicomponent earthquake ground motions.

Introduction

Spectral acceleration (*Sa*) values of earthquake ground motions are widely used in seismic hazard analysis and evaluation of structural response. Relatively little work has been done, however, to measure the joint distributions of multiple simultaneous spectral acceleration values. In this article, the authors measure correlation coefficients of spectral acceleration values to gain insight into these joint distributions and to facilitate other research. Correlations are presented for spectral acceleration values of a single ground-motion component at two differing periods and also for spectral accelerations of orthogonal components (horizontal/horizontal or horizontal/vertical) at two periods.

This analysis was performed in recognition of the many potential applications of the results. Several predictions have been developed previously for correlations of spectral acceleration values of a single ground-motion component, and past uses of those models are mentioned in the following text. The models for spectral accelerations of orthogonal components are new, and so several potential applications are described.

Motivation

Knowledge of correlation of spectral values has been used in several past studies to gain insight into seismic hazard and structural performance. Several past uses of the type of models presented here are discussed in this section, and potential future applications will be discussed later.

Conventional probabilistic seismic hazard analysis (PSHA) (Kramer, 1996) provides the mean annual rate of exceeding a specified value of a single ground-motion parameter, such as spectral acceleration at a given period. These hazard analyses can be repeated for spectral acceleration at several periods and presented simultaneously as uniform hazard spectra. But these uniform hazard spectra, being the locus of results from a suite of marginal hazard analyses for individual spectral values, should not be interpreted as providing any knowledge about the joint occurrence of spectral values at differing periods. To obtain knowledge about the joint or simultaneous occurrence of spectral acceleration at multiple periods, it is necessary to perform a vector-valued probabilistic seismic hazard analysis (VPSHA) (Bazzurro

and Cornell, 2002). This analysis is a direct extension of traditional PSHA, using the same information about the magnitudes, locations, and recurrence rates of earthquakes and the same ground-motion prediction (attenuation) models. The only additional information requirement is knowledge of the joint distribution of the spectral values for a given magnitude and distance. Logarithmic spectral acceleration values have been observed to be well represented by the normal distribution marginally, so the mild assumption that pairs of values are well represented by the joint normal distribution (and/or that the conditional distributions of one given the other are normal) is probably a reasonable one, but has not been investigated as yet to the authors' knowledge. Under this assumption, only correlation coefficients between spectral values at two periods are needed to define the joint distribution and proceed with VPSHA. The models presented in this article will provide improved predictions of these correlation coefficients, furthering the development of the vector-valued probabilistic seismic hazard analysis. Once a vector-valued PSHA has been performed, structural engineers can use this new information to improve the efficiency of probabilistic performance assessments of structures (e.g., Baker and Cornell 2004, 2005a). Because engineers are provided with more information about the spectral content of the ground motions occurring at a site, they are able to increase the precision of their structural response analyses.

Correlation of spectral acceleration values also arises implicitly in the development of ground-motion prediction models. The seismologists who develop these models often average the (log) spectral acceleration values of two perpendicular horizontal components of a ground motion and use these averaged data for fitting regression lines. This averaging decreases the noise in the data, allowing for more accurate estimates of the model parameters. Thus, PSHA calculations using these ground-motion prediction models provide mean exceedance rates for averaged spectral acceleration values, or more precisely for the geometric mean of the two horizontal components. But structural engineers often do not perform this averaging across components, which results in an inconsistency between PSHA and structural analysis. The work of the seismologist and engineer can be properly reconnected, however, once knowledge of correlations of spectral acceleration values in perpendicular ground-motion components is known (Baker and Cornell, 2006).

In addition to averaging across perpendicular components, averaging spectral accelerations across a range of periods is sometimes also performed, with a similar goal of reducing the variability of the resulting ground-motion intensity parameter for a given magnitude and distance (e.g., Pacific Gas & Electric, 1988; Shome and Cornell, 1999; N. A. Abrahamson, A. Kammerer, and N. Gregor, personal comm., 2003). With this work, knowledge of correlations of spectral accelerations is again needed, to quantify the effect of the averaging procedure.

A measure of ground-motion intensity proposed by Cordova *et al.* (2001) is a function of spectral acceleration at two periods. This measure was found to be useful for predicting response of the structure under consideration. A custom ground-motion prediction model was needed to complete the assessment of the structure, and was derived from an existing model by making use of an estimate of correlation between spectral acceleration values at the two periods of interest.

In addition to these past applications, there are several easily envisioned future applications that make use of the new information in this article regarding correlations across multiple components of ground motion. Examples are presented below, after the development of the predictive equations.

Note that there are other studies of earthquake ground motions that at first glance might appear similar to this work, but are in fact not related. For example, Penzien and Watabe (1975) observed that temporal cross-correlations of ground-motion accelerations at an instant in time are approximately zero along specified principle axes. That work does not imply anything about the phenomenon examined in this article: the correlation of peak spectral values (i.e., frequency content measures) at differing frequencies and orientations.

Analysis Procedure

Record Selection

The results presented in this study were derived empirically from a strong-motion data set based on worldwide recordings of shallow crustal earthquakes. The records for this study were taken from the PEER Strong Motion Database (2000). Records were selected based on the following criteria:

1. The site was classified as stiff soil: U.S. Geological Survey (USGS) class B-C or Geomatrix class B-D.
2. The recording was made in the free field or the first story of a structure.
3. All three components (two horizontal and one vertical) were available and had high-pass filter corner frequencies less than 0.2 Hz and low-pass filter corner frequencies greater than 18 Hz.
4. The earthquake magnitude was greater than 5.5.
5. The source-to-site distance was less than 100 km.

The recordings were left oriented as recorded rather than rotated into fault normal and fault parallel components, so they have effectively random orientations with respect to fault direction. This is analogous to the record orientations used to develop typical ground-motion prediction models. The Next Generation Attenuation project will produce a record library and predictive models that treat fault normal and fault parallel ground motions separately. When that project is completed it will be useful to compute correlations for fault normal and fault parallel ground motions separately,

but at present only randomly oriented ground motions are considered.

A total of 469 records from 31 earthquakes met the selection criteria, each consisting of three components of ground-motion recordings. Of the 469 records, 202 were from the 1999 Chi-Chi, Taiwan, earthquake. The Chi-Chi records were removed from the initial analysis to ensure that the results would not be excessively influenced by any peculiarities of the records from this single earthquake. The Chi-Chi records were later used to cross-validate the predictive equations.

Computation of Correlations

Using the 267 remaining three-component records, correlations were computed for the two horizontal and vertical components. At this point, the variable for which the correlation is estimated should be defined more clearly. The logarithmic spectral accelerations of the three ground-motion components can be represented by the following model:

$$\ln Sa_x(T) = f_H(M,R,T,\theta) + \sigma_H(M,T)\varepsilon_x(T) \quad (1)$$

$$\ln Sa_y(T) = f_H(M,R,T,\theta) + \sigma_H(M,T)\varepsilon_y(T) \quad (2)$$

$$\ln Sa_z(T) = f_V(M,R,T,\theta) + \sigma_V(M,T)\varepsilon_z(T), \quad (3)$$

where x and y are used to denote the two horizontal directions of the recording, and z is used to denote the vertical direction. The functions f_H (M, R, T, θ) and f_V (M, R, T, θ) are mean ground-motion predictions for the horizontal and vertical logarithmic response spectral values, respectively. These predictions are a function of the earthquake magnitude (M), distance (R), period (T), and other parameters (θ), such as the local soil conditions and faulting mechanism. These mean ground-motion predictions are deterministic, given the input parameters. The terms $\sigma_H(M, T)$ and $\sigma_V(M, T)$ account for the observed standard deviation of the logarithmic horizontal and vertical spectral accelerations, respectively. The standard deviations are observed to depend on the magnitude of the earthquake and the period of interest. Finally, the random variables $\varepsilon_x(T)$, $\varepsilon_y(T)$, and $\varepsilon_z(T)$ account for the randomness of the observations. Because the other terms in equations (1) through (3) have already accounted for the means and standard deviations of logarithmic spectral acceleration, the ε terms have means of zero and unit standard deviations.

The $f(M, R, T, \theta)$ and $\sigma(M, T)$ functions are completely defined by previously published ground-motion prediction models. What is not defined by standard ground-motion prediction models is the correlation between ε terms at different frequencies or for different components. For example, the correlation between the ε values of the two horizontal components of a given record is of interest: $\rho_{\varepsilon_x(T),\varepsilon_y(T)}$. The ε values of a record also vary as the period varies, and so the correlation of the ε values of a single component of a record at two periods is also of interest: $\rho_{\varepsilon_x(T_1),\varepsilon_x(T_2)}$. Finally, one might be interested in the correlations between two components at two differing periods: $\rho_{\varepsilon_x(T_1),\varepsilon_y(T_2)}$. These correlation values are estimated in this article.

According to equations (1) through (3), the computed ε values will vary somewhat for a given record depending on the ground-motion prediction model chosen. That is, ln $Sa(T)$ of a record is given and $f(M, R, T, \theta)$ and $\sigma(M, T)$ vary slightly among models, and so the $\varepsilon(T)$ value of a record must also vary among models to maintain equality. The correlations of ε values, however, were observed to be insensitive to the ground-motion prediction model considered. Correlations were computed here by using the model of Abrahamson and Silva (1997), but the results were found to be nearly identical when other models (specifically, Boore *et al.*, 1997; Campbell, 1997) were compared.

Once correlations of the ε values have been determined, we note that ln $Sa(T)$ is simply a linear function of $\varepsilon(T)$, with no other sources of uncertainty. Therefore, the correlation between, for example, ln $Sa_x(T)$ and ln $Sa_y(T)$ (for a given record) is equal to the correlation between $\varepsilon_x(T)$ and $\varepsilon_y(T)$. Thus, the procedure used here is to compute the ε values for all records, to remove the effect of magnitude, distance, etc., from the variation in observed spectral values. Correlations can be computed for these ε values, which are then appropriate to represent the correlations between ln Sa values for a given magnitude, distance, etc. For the same reason, the predictions can be used to represent logarithmic spectral velocity or spectral displacement. In general, the correlation of ln $Sa_x(T)$ and ln $Sa_y(T)$ is also a reasonable approximation for the correlation between $Sa_x(T)$ and $Sa_y(T)$ (Liu and Der Kiureghian, 1986). In applications such as vector-valued PSHA, however, it is often the logarithms of response spectral values that are used in the joint distributions, so the more precise correlation between ln $Sa_x(T)$ and ln $Sa_y(T)$ is sufficient in many cases.

To estimate correlation coefficients, we use the maximum likelihood estimator, sometimes referred to as the Pearson product-moment correlation coefficient (Neter *et al.*, 1996):

$$\hat{\rho}_{A,B} = \frac{\sum_{i=1}^{n}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^{n}(A_i - \bar{A})^2 \sum_{i=1}^{n}(B_i - \bar{B})^2}}, \quad (4)$$

where A and B are the random variables of interest (e.g., $\varepsilon_x(T)$ and $\varepsilon_y(T)$), $\bar{A}$ and $\bar{B}$ are their sample means, A_i is the ith observation of variable A, and n is the total number of observations (records). We perform this correlation computation for each pair of orientations (horizontal/horizontal in the same direction, horizontal/horizontal in perpendicular directions, vertical/vertical, and vertical/horizontal) and for each pair of periods of interest (75 periods between 0.05 and 5 sec). The matrix representing correlations for all combi-

nations of the 75 periods is most compactly displayed using a contour plot as a function of the periods T_1 and T_2 (e.g., Fig. 1).

Before performing further analysis, the correlation coefficients estimated using equation (4) were smoothed by using a simple averaging with correlation coefficients in a nearby neighborhood of periods to remove some of the noise in the estimates and make the underlying patterns in the correlation matrix clearer. A comparison of contours of the correlation matrix before and after smoothing is displayed in Figure 1.

Nonlinear Regression

Nonlinear least-squares regression was utilized to condense the data from a large empirical correlation matrix into a relatively simple predictive equation. Functional forms were chosen based on inspection of the correlation matrices, and coefficients for the functions were determined by using nonlinear regression. Correlation coefficients estimated from empirical data have nonconstant standard errors that depend on the true underlying correlation coefficient. For this reason, minimizing the squared error between the empirical correlation matrix and the predictive function would not be the optimal criteria for fitting the predictive function (i.e., fitting a correlation coefficient of 0.9 with an estimate of 0.8 is a worse error than fitting a correlation coefficient of 0.1 with an estimate of 0). For this reason the Fisher z transformation (Neter *et al.*, 1996) was applied to the correlation coefficients:

$$z = \frac{1}{2} \ln \left(\frac{1 + \rho}{1 - \rho} \right), \tag{5}$$

where ρ is an estimated correlation coefficient and z is the transformed data with a constant standard error. Simple least-squares regression could then be applied to these z values. The coefficients for the prediction equations were selected such that the squared prediction errors were minimized over the range of periods of interest:

$$\min_{\theta} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{1}{2} \ln \left(\frac{1 + \rho_{i,j}}{1 - \rho_{i,j}} \right) - \frac{1}{2} \ln \left(\frac{1 + \hat{\rho}_{i,j}(\boldsymbol{\theta})}{1 - \hat{\rho}_{i,j}(\boldsymbol{\theta})} \right) \right)^2, \tag{6}$$

where $\rho_{i,j}$ is the empirical correlation coefficient at the period pair (T_i, T_j) and $\hat{\rho}_{i,j}(\boldsymbol{\theta})$ is its predicted value using the functional forms shown below with a vector of coefficients $\boldsymbol{\theta}$. The resulting models are strictly empirical and thus should not be extrapolated beyond the range over which they were fit (periods between 0.05 and 5 sec, earthquake magnitudes between 5.5 and 7.6, and distances between 0 and 100 km).

Results

The results of these analyses are presented in the form of simple equations. They are broken into three separate cases, presented individually in the next sections.

Cases at a Single Period

Correlations between response spectral values with the same period but differing orientations are presented first. The correlation between horizontal orthogonal ε values at the period T is estimated by the equation:

$$\rho_{\varepsilon_x, \varepsilon_y} = 0.79 - 0.023 \cdot \ln (T). \tag{7}$$

The predictions from this function and the empirical correlations from the data set are displayed in Figure 2. By using the bootstrap to resample ground-motion records (Efron and Tibshirani, 1993), the slope of the regression line was found to be statistically significant, with a P-value of 0.001.

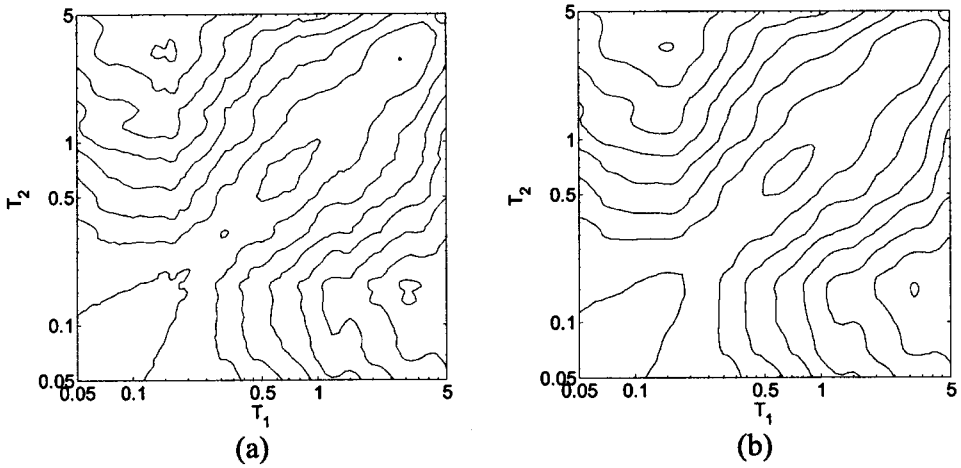

Figure 1. Effect of smoothing on the empirical correlation matrix for horizontal epsilons in perpendicular directions at two periods (T_1 and T_2). (a) Before smoothing. (b) After smoothing.

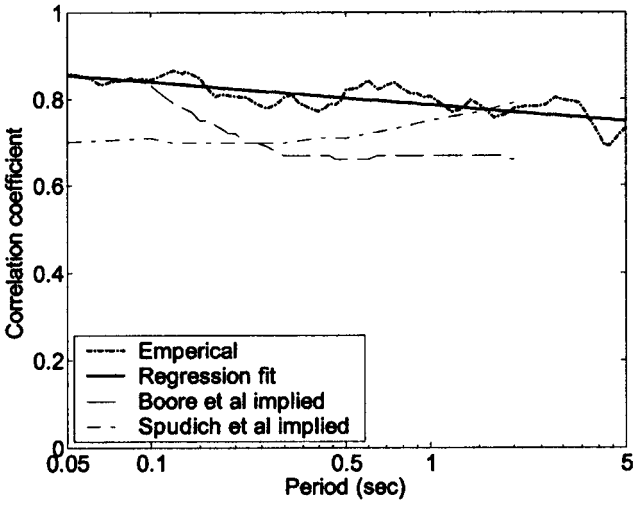

Figure 2. Correlation coefficients for perpendicular horizontal epsilons at the same period. Empirical results, the prediction from equation (7), and the correlations implied from the ratios of standard deviations in Boore *et al.* (1997) and Spudich *et al.* (1999).

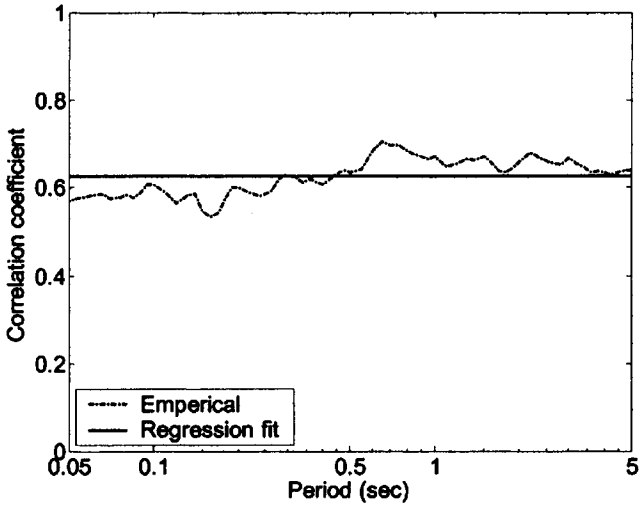

Figure 3. Correlation coefficients between horizontal epsilons and vertical epsilons at the same period. Empirical results and the prediction from equation (8).

The correlation between a horizontal ε value and a vertical ε value at the period T is estimated by the constant:

$$\rho_{\varepsilon_x,\varepsilon_z} = 0.63. \quad (8)$$

A regression fit was performed with a prediction as a function of T, but the slope was not statistically significant and so the prediction is a constant value for all periods. The empirical correlations from the data set and the predicted value are displayed in Figure 3.

Cases with Differing Periods but the Same Orientation

When the two periods of interest differ, more complex functional forms are needed. The correlation between the ε values of a single horizontal ground motion component at two differing periods is estimated by the function:

$$\rho_{\varepsilon_x,\varepsilon_x} = 1 - \cos\left(\frac{\pi}{2} - \left(0.359 + 0.163 I_{(T_{\min}<0.189)} \ln\frac{T_{\min}}{0.189}\right) \ln\frac{T_{\max}}{T_{\min}}\right), \quad (9)$$

where $I_{(T_{\min}<0.189)}$ is an indicator function equal to 1 if $T_{\min} < 0.189$ second and equal to 0 otherwise, implying that the form of the equation is simply $1 - \cos(a - b \ln(T_{\max}/T_{\min}))$ for periods larger than 0.189 sec. The variables $T_{\min}$ and $T_{\max}$ are used to denote the smaller and larger of the two periods of interest, respectively. The empirical correlations from the data set and the predictions from equation (9) are displayed in Figure 4a and b, respectively. Note that both the empirical data and this prediction imply that correlation does not always decrease with increasing separation of periods (e.g., the correlation between *Sa* values at $\{T_{\min} = 0.05$ sec, $T_{\max} = 1$ sec$\}$ is greater then the correlation at $\{T_{\min} = 0.2$ sec, $T_{\max} = 1$ sec$\}$). The authors see no obvious reason for this phenomenon, but it is seen clearly in the data.

Although equation (9) was fit for ε values of individual components, it is equally valid for ε values of geometric mean spectral acceleration values. This is shown both theoretically and empirically in Baker and Cornell (2005b, appendix B). The equation could also be used to approximately represent the correlation of inter-event ε values (which are of interest for modeling losses to portfolios of spatially distributed buildings), but the agreement is not as good in this situation. The definition of inter-event ε values can be found in Abrahamson and Silva (1997).

The correlation between the ε values of a vertical ground motion component at two different periods is estimated by the function:

$$\rho_{\varepsilon_z,\varepsilon_z} = 1 - 0.77 \ln\frac{T_{\max}}{T_{\min}} + 0.315 \left(\ln\frac{T_{\max}}{T_{\min}}\right)^{1.4}. \quad (10)$$

A comparison of this prediction with the empirical correlation coefficients is shown in Figure 5.

Cases with Differing Periods and Differing Orientations

For correlations of ε values between two horizontal components in perpendicular directions, it was hypothesized that perhaps the correlation coefficient could be represented as a product of the correlation due to perpendicular orientation and the correlation due to differing periods. The model of equation (7) was used to represent the perpendicular ori-

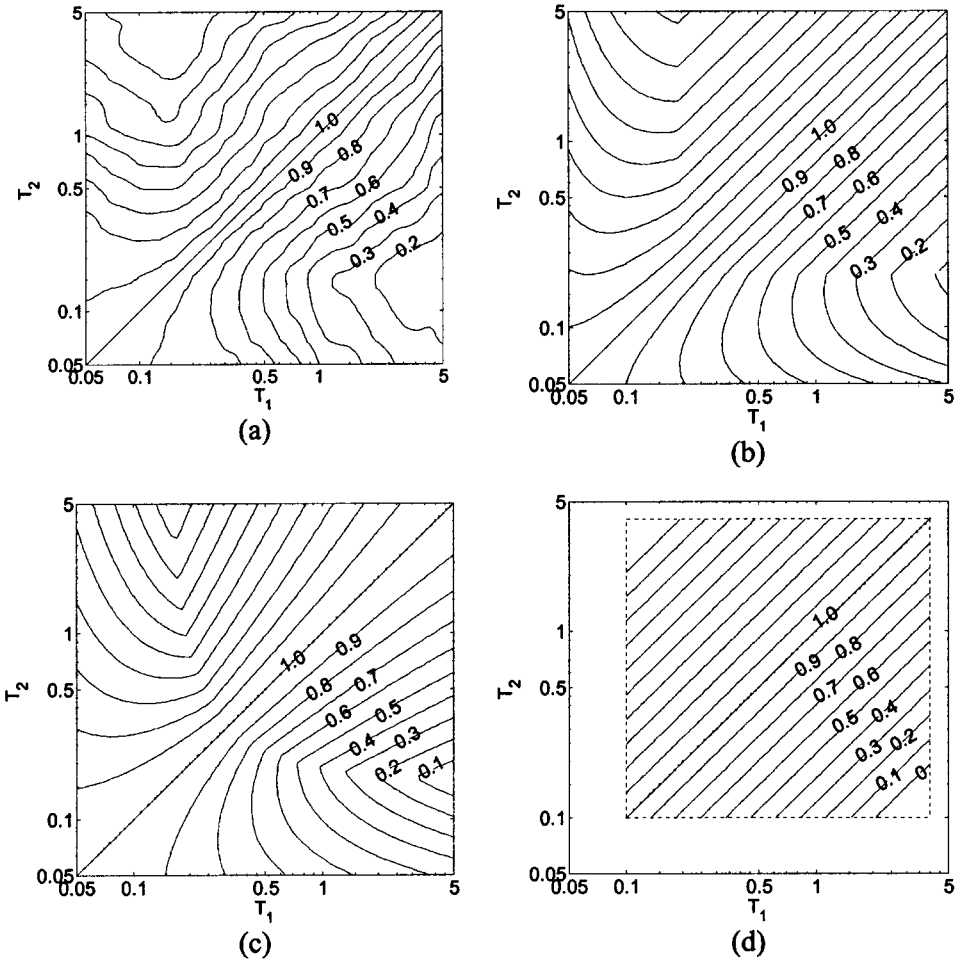

Figure 4. Correlation contours for horizontal epsilons in the same direction at two periods (T_1 and T_2). (a) Smoothed empirical results. (b) The prediction from equation (9). (c) The prediction from Abrahamson et al. (personal comm., 2003). (d) The prediction from Inoue and Cornell (1990).

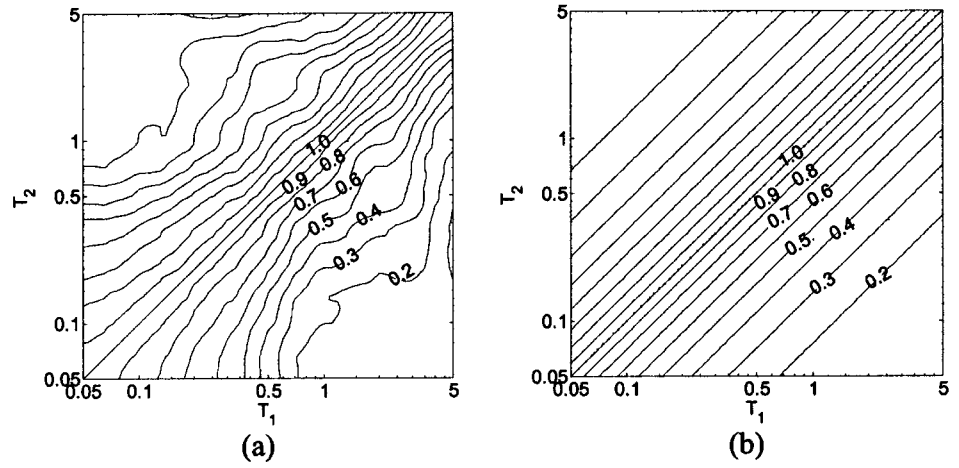

Figure 5. Correlation contours for vertical epsilons in the same direction at two periods (T_1 and T_2). (a) Smoothed empirical results. (b) The prediction from equation (10).

entations, evaluated at $T = \sqrt{T_{\min} T_{\max}}$, the geometric mean of the two periods of interest, and the model of equation (9) was used to represent the correlations at differing periods. The resulting product-form equation is:

$$\rho_{\varepsilon_x,\varepsilon_y} = (0.79 - 0.023 \cdot \ln \sqrt{T_{\min} T_{\max}}) \cdot \left(1 - \cos\left(\frac{\pi}{2} - \left(0.359 + 0.163 I_{(T_{\min}<0.189)} \ln \frac{T_{\min}}{0.189}\right) \ln \frac{T_{\max}}{T_{\min}}\right)\right). \quad (11)$$

A comparison of this prediction with the empirical correlation contours is shown in Figure 6. This form fits the empirical results well and agrees with equation (7) in the special case $T_{\min} = T_{\max}$. This "product of correlation coefficients" form implies that a Markov-process-like relationship exists, where $\varepsilon_x(T_1)$ and $\varepsilon_y(T_2)$ are conditionally (linearly) independent given either $\varepsilon_x(T_2)$ or $\varepsilon_y(T_1)$ (Ditlevsen, 1981, p. 339). Approximate conditional independence is observed in both the empirical data and in equation (11) (making the approximation $\rho_{\varepsilon_x(\sqrt{T_1T_2}),\varepsilon_y(\sqrt{T_1T_2})} \cong \rho_{\varepsilon_x(T_1),\varepsilon_y(T_1)} \cong \rho_{\varepsilon_x(T_2),\varepsilon_y(T_2)}$.

The same procedure was used to predict the correlations of epsilons between a horizontal and a vertical component of a ground motion at differing periods. The estimate of equation (8) at the geometric mean of the two periods was multiplied by a term with the functional form of equation (9). In this case, the coefficients of equation (9) did not provide a good fit to empirical results (which was expected because the coefficients are not associated with vertical motions), so two coefficients were re-estimated using nonlinear regression (the additional improvement from refitting all coefficients was negligible). The final estimate is given by:

$$\rho_{\varepsilon_x,\varepsilon_z} = (0.64 + 0.021 \cdot \ln \sqrt{T_{\min} T_{\max}}) \cdot \left(1 - \cos\left(\frac{\pi}{2} - \left(\ln \frac{T_{\max}}{T_{\min}}\right) \left(0.29 + 0.094 I_{(T_{\min}<0.189)} \ln \frac{T_{\min}}{0.189}\right)\right)\right). \quad (12)$$

This prediction is compared with empirical results in Figure 7. One notable feature of Figure 7a is that the contours are not symmetric about the line $T_1 = T_2$. That is, the correlation between *Sa* of the horizontal component at T_1 with the *Sa* of the vertical component at T_2 is not necessarily equal to the correlation of the *Sa* of the horizontal component at T_2 with the *Sa* of the vertical component at T_1. When the empirical correlations from the primary dataset were compared with the empirical correlations from the Chi-Chi dataset, however, variations between the two indicated that any asymmetries in the empirical correlations were likely due to sampling variability rather than to some underlying trend. For this reason, the form of equation (12) is left such that predictions are symmetric about the line $T_1 = T_2$.

Among the horizontal/vertical correlations, certain period pairs will be of more engineering interest than others. For instance, periods of vibration of buildings are typically much shorter in the vertical direction than the horizontal direction. Using a simple estimate of 0.1 sec for the vertical period of a typical building, it would be interesting to know the correlation of vertical *Sa* values at 0.1 sec with horizontal *Sa* values at a range of periods (corresponding to varying horizontal periods of vibration). This is shown in Figure 8 for both the empirical and predicted correlations.

For all of these predictions, several checks were performed. The positive definiteness of the predicted correlation matrices (when computed for a large array of periods simultaneously) was verified. A joint correlation matrix consisting of predictions in all three directions simultaneously was also found to be positive definite. This is a required property of a correlation matrix, and is necessary if one needs the joint distribution of *Sa* at many periods simultaneously (e.g., the "simulation of response spectra" application following). In addition, it was verified that the empirical correlations do not depend on magnitude or distance. This was done by taking windows of magnitude or distance values and comparing the computed correlation coefficients as the window moved to different magnitude or distance values. No trends were seen, and so the preceding models were left functionally independent of magnitude and distance. Supporting calculations for these conclusions can be found in Baker and Cornell (2005b, appendix B).

Some general observations can be made from the preceding data and analytical predictions. When both periods are the same, the correlation between the *Sa* values of two perpendicular horizontal components is roughly 0.8. For frame-type buildings, the first two periods of vibration in the same axis typically have a ratio of approximately 3:1. Spectral acceleration values at these two periods are often used by engineers (e.g., in response spectrum analysis; Chopra, 2001), and we see that if the periods are greater than 0.189 sec, the correlation coefficient between these two *Sa* values is approximately 0.6. When considering two periods with a ratio of 3:1 in orthogonal horizontal directions, we can use the Markov approximation and estimate the correlation coefficient as 0.8 * 0.6 = 0.48. When considering vertical ground motions, if we assume that the vertical period of interest is 0.1 sec and the horizontal period of interest is 0.5 to 1 sec (for midrise buildings) then we see that the correlation coefficient is approximately 0.3 to 0.4. If we define the correlation distance as the ratio of periods $T_{\max}/T_{\min}$ such that the correlation coefficient between the two is $e^{-1} = 0.37$, then the correlation distance is approximately 5 for vertical records and 6.5 for horizontal records (if $T_{\min} > 0.189$). These numbers may serve as useful rules-of-thumb for quick estimates.

Comparisons with Previous Work

The predictions provided by equations (7) through (12) are believed to be the first of their kind in most cases; except for two previous models analogous to equation (9). Inoue

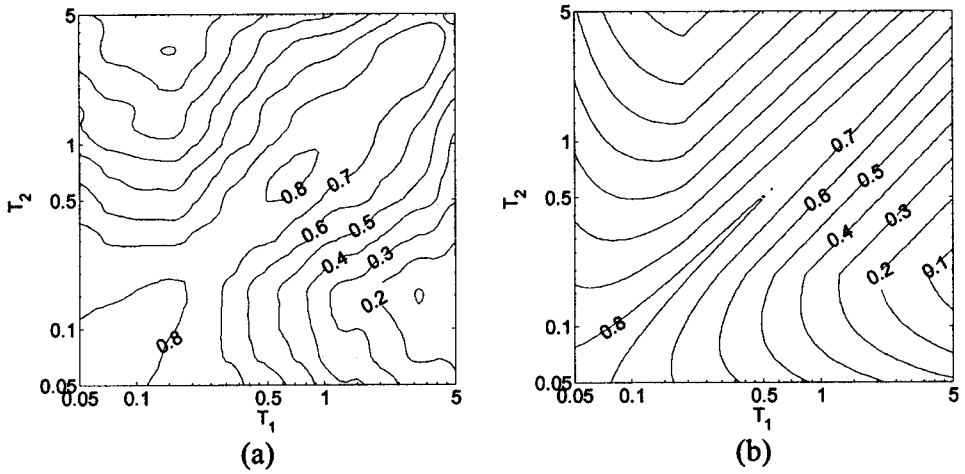

Figure 6. Correlation contours for horizontal epsilons in perpendicular directions at two periods (T_1 and T_2). (a) Smoothed empirical results. (b) The prediction from equation (11).

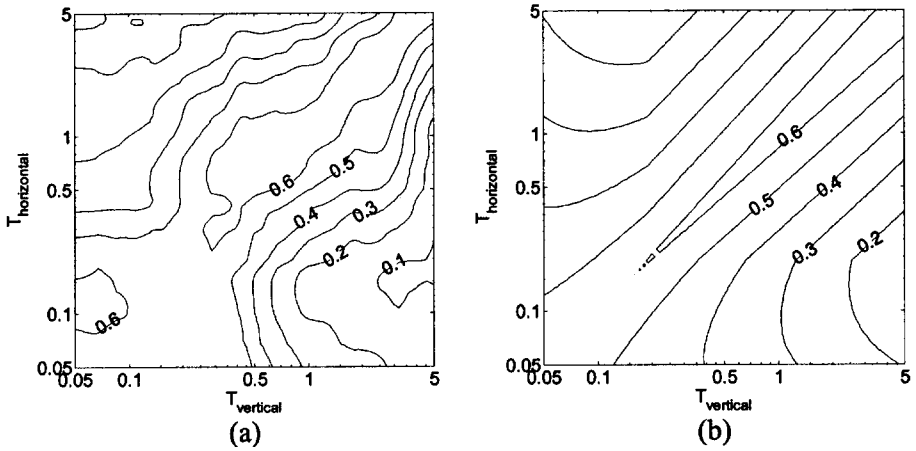

Figure 7. Correlation contours of vertical epsilons with horizontal epsilons at two periods (T_1 and T_2). (a) Smoothed empirical results. (b) The prediction from equation (12).

and Cornell (1990) proposed the following analogous model:

$$\rho_{\varepsilon_x,\varepsilon_x} = 1 - 0.33 \cdot \ln(T_{\max}/T_{\min}). \quad (13)$$

The function was fit over a period range between 0.1 sec and 4 sec, using 64 record components. Its contours are plotted in Figure 4c. This prediction agrees reasonably well with empirical results within the range over which it was fitted, but it would perform poorly if extrapolated to the larger period range used by the newer models.

Another model for the correlations of epsilons at two periods along the same component is provided by Abrahamson *et al.* (personal comm., 2003):

$$\rho_{\varepsilon_x,\varepsilon_x} = \begin{cases} \tanh(1 - 1.32 \cdot \log_{10}(X) + 0.072(\log_{10}(X))^2 - 0.05 - 0.5 \cdot \log_{10}(f_c)) & \text{for } f_c \le 3.16 \text{ Hz} \\ \tanh(1 - 1.32 \cdot \log_{10}(X) + 0.072(\log_{10}(X))^2 - 0.95 + 1.3 \cdot \log_{10}(f_c)) & \text{for } f_c > 3.16 \text{ Hz}, \end{cases} \quad (14)$$

where f_1 and f_2 are the larger and smaller frequencies, respectively, $f_c = 0.5(f_1 + f_2)$, and $X = \ln(f_1/f_2)$. The function was fit over a range of periods between 0.03 sec and 5 sec, using a record set similar to the one used here. The contours of this prediction are shown in Figure 4d. It is not dramatically different from that of equation (9), displayed in Figure 4b. Selected contours from Figure 4a, b, and c are overlaid in Figure 9 to aid comparisons of the predictions. Note that one desirable attribute missing from the Abrahamson *et al.* prediction is positive definiteness when the correlation matrix is computed for multiple periods simultaneously. This violates a required property of correlation matrices as discussed previously.

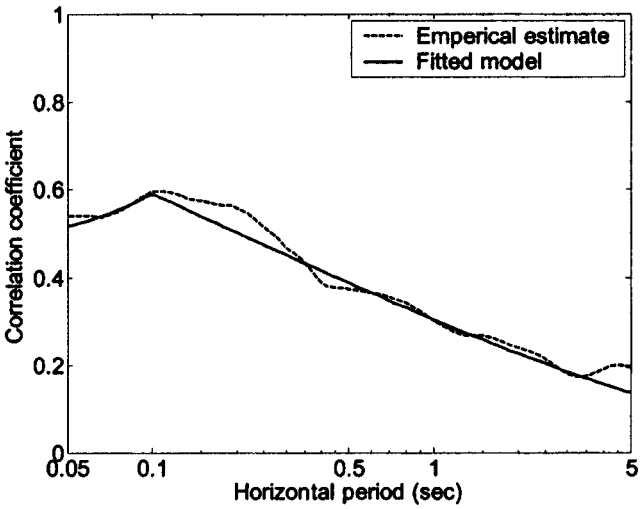

Figure 8. Correlation coefficient between vertical epsilons and horizontal epsilons when the period in the vertical direction is 0.1 sec.

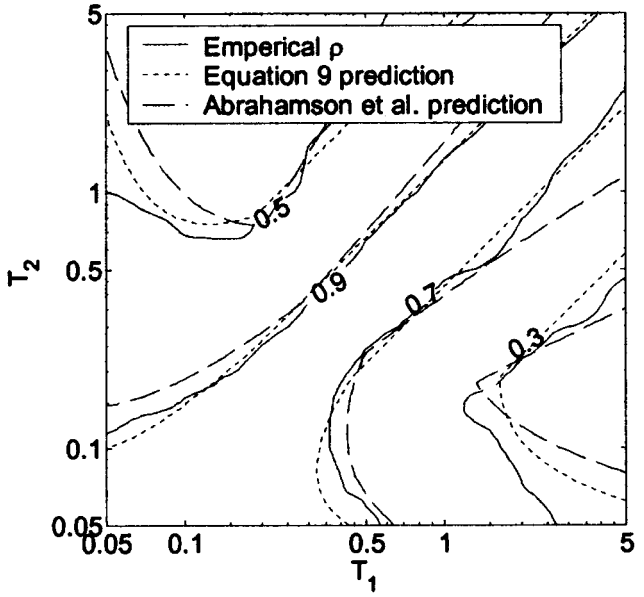

Figure 9. Overlaid contours at four correlation levels for the empirical correlations, the prediction from equation (9) and the prediction from Abrahamson *et al.* (personal comm., 2003).

In general, these previous models agree reasonably well with equation (9) proposed here. Equation (9) may be considered an improvement because of its increased range of periods as compared with the model of Inoue and Cornell and due to its positive definiteness property, which the Abrahamson *et al.* model does not possess. Equation (9) also produces smaller residuals than these two models when predicting correlations of either the primary record set above or the Chi-Chi record set (which was not used to fit any of the three models).

One additional study of spectral acceleration correlations used 30 records from the 1999 Chi-Chi, Taiwan, earthquake (Wang *et al.*, 2001). The results are not directly comparable with the work here, however, and so this work is not considered further.

Previous work also exists that can be indirectly compared with equation (7). Some ground-motion prediction models provide standard deviations of residuals for both the logarithmic spectral acceleration of a single horizontal component of a ground motion, or for the geometric mean of two orthogonal components (e.g., Boore *et al.*, 1997; Spudich *et al.*, 1999). By examining the ratios of the two standard deviations, one can back-calculate the implied correlation coefficient between the two components using the following equation:

$$\rho_{\varepsilon_x,\varepsilon_y} = 2\sigma^2_{\text{g.m.}}/\sigma^2_{\text{arb}} - 1, \tag{15}$$

where $\sigma_{\text{g.m.}}$ is the logarithmic standard deviation of the geometric mean of the two horizontal components, and σ_{arb} is the logarithmic standard deviation of an arbitrary component. The correlation coefficients implied by the models of Boore *et al.* (1997, 2005) and Spudich *et al.* (1999) are displayed in Figure 2. These models underestimate the correlation seen empirically in this study, but estimation of correlations was not a goal of these studies. The value of interest to these authors is $\sigma^2_{\text{arb}}/\sigma^2_{\text{g.m.}}$, but a slight change in this ratio can produce a large change in the correlation coefficient. Given that these authors were not interested in correlations and that the correlations calculated using equation (15) are very sensitive to the ratio $\sigma^2_{\text{arb}}/\sigma^2_{\text{g.m.}}$, it is perhaps not surprising that there is a slight discrepancy between the direct calculations of this article and the indirect back-calculations from previous work.

Applications

To demonstrate the usefulness of these new predictions and perhaps inspire new uses, several applications of the new correlation predictions are briefly described here.

Vector-Valued Hazard Analysis for Horizontal and Vertical Components of Ground Motion

The vector-valued hazard analysis methodology of Bazzurro and Cornell (2002) can now be easily applied to analysis of horizontal and vertical ground motions simultaneously. Consider a hypothetical two-dimensional building frame with a first-mode period of 1 sec in the horizontal direction and a first-mode period of 0.1 sec in the vertial direction. Using equation (12), we estimate a correlation coefficient between the *Sa* values at these two periods of 0.30. We assume that the building is located 8 km from a single fault which produces only (characteristic) magnitude 6.5 earthquakes with a mean return period of 500 years. Using the Abrahamson and Silva (1997) ground-motion prediction

model and the correlation coefficient predicted here, we can compute the joint distribution of horizontal and vertical spectral acceleration values at the specified first-mode periods. Contours of that hazard are displayed in Figure 10. Note that because of the low correlation, extreme values of *Sa* are unlikely to occur in the horizontal and vertical directions simultaneously. For example, we note that a horizontal *Sa* of 0.65*g* has a 2% probability of exceedance in 50 years, and a vertical *Sa* of 0.98*g* has a 2% probability of exceedance in 50 years. But the probability of exceeding both a horizontal *Sa* of 0.65*g* and a vertical *Sa* of 0.98*g* simultaneously is only 0.65% in 50 years. This suggests that designing for extreme ground motions in all directions simultaneously (e.g., by applying a horizontal Uniform Hazard Spectrum and a vertical Uniform Hazard Spectrum simultaneously), may be more conservative than intended. If one is primarily concerned with horizontal motions and thus uses the 2% in 50 years horizontal *Sa* value, the preferred vertical *Sa* design value would be that associated with the mean ln *Sa* in the vertical direction, given that *Sa* in the horizontal direction has exceeded 0.65*g*. This choice is consistent with load combination rules used in practice elsewhere (e.g., Norwegian Technology Standards Institution, 1999, pp. 17–18). For this example the design value of the vertical *Sa* was determined to be 0.66*g*, which is approximately 30% less than the *Sa* with a 2% probability of exceedance in 50 years.

Only a single magnitude/distance pair was used in this example for computational simplicity. Generalization of the analysis to incorporate multiple faults with multiple magnitudes and distances is merely a matter of coding the correlation prediction into a vector-valued hazard analysis program (Somerville and Thio, 2003). (As an approximation, one can also use the dominant value of ε obtained by disaggregation to obtain the associated ε values for other components.) No further mathematical developments are needed.

Ground-Motion Prediction Model for the Geometric Mean of Orthogonal Spectral Accelerations at Two Periods

The geometric mean of spectral acceleration values in two orthogonal horizontal directions is often computed in ground-motion prediction models. This quantity is useful for analyzing a three-dimensional structure subjected to ground motions in two horizontal directions, because it describes the intensity of ground motion in two directions using only a single parameter (Stewart *et al.*, 2001; Baker and Cornell, 2006). The geometric mean of spectral acceleration provided by ground-motion prediction models uses the same period of vibration in both directions, however, whereas a structure commonly has different periods of vibration in its two principal directions. Using the correlation models presented above, one can easily develop a "custom" correlation model incorporating the two periods of interest in a particular application.

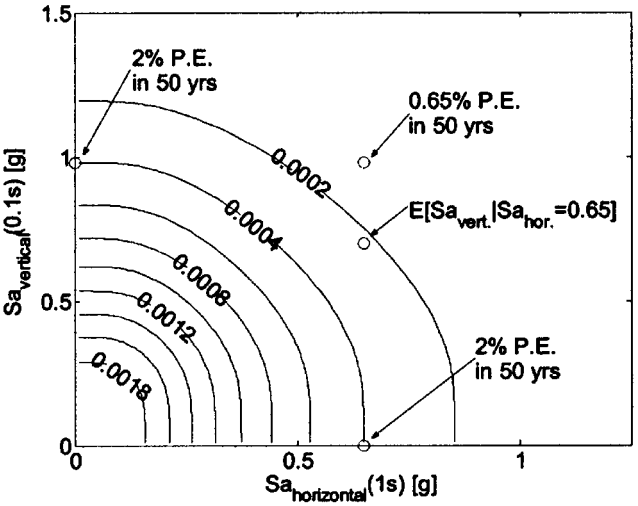

Figure 10. Contours of vector-valued probabilistic seismic-hazard analysis. The contours denote the mean annual rate of exceeding *both* the Sa_{vertical} *and* the $Sa_{\text{horizontal}}$ values. P.E., Probability of exceedance.

Ground-motion prediction models provide the $f_H(M, R, T, \theta)$ and $\sigma_H(M, T)$ terms for equation (1). We are interested in determining an analogous equation of the form:

$$\ln Sa_{\text{g.m.}}(T_1,T_2) = f_{\text{g.m.}}(M,R,T_1,T_2,\theta) + \sigma_{\text{g.m.}}(M,T_1,T_2)\varepsilon_{\text{g.m.}}(T_1,T_2), \quad (16)$$

where $Sa_{\text{g.m.}}(T_1, T_2)$ is the geometric mean of two orthogonal horizontal spectral accelerations at two periods T_1 and T_2. The function $f_{\text{g.m.}}(M, R, T_1, T_2, \theta)$ is the mean value of ln $Sa_{\text{g.m.}}(T_1, T_2)$ and $\sigma_{\text{g.m.}}(M, T_1, T_2)$ is the standard deviation. Recognizing that $\ln Sa_{\text{g.m.}}(T_1, T_2) = 1/2(\ln Sa_x(T_1) + \ln Sa_y(T_2))$, the mean and standard deviation terms can be derived from existing models and the correlation predictions presented above:

$$f_{\text{g.m.}}(M,R,T_1,T_2,\theta) = 1/2(f_H(M,R,T_1,\theta) + f_H(M,R,T_2,\theta)) \quad (17)$$

$$\sigma_{\text{g.m.}}(M,T_1,T_2) = \sqrt{\frac{\sigma_H^2(M,T_1)}{4} + \frac{\sigma_H^2(M,T_2)}{4} + \frac{\rho_{\varepsilon_x,\varepsilon_y}(T_1,T_2)\sigma_H(M,T_1)\sigma_H(M,T_2)}{2}}, \quad (18)$$

where $\rho_{\varepsilon_x,\varepsilon_y}(T_1,T_2)$ comes from equation (11). Remembering that many ground-motion prediction models provide the standard deviation of the geometric mean of two horizontal *Sa* values, rather than the standard deviation of a single component *Sa*, it may be first necessary to make the following conversion before calculating equation (18):

$$\sigma_H^2(M,T) = \frac{2\sigma_{\text{g.m.}}^2(M,T)}{1 + \rho_{\varepsilon_x,\varepsilon_y}}, \quad (19)$$

where $\rho_{\varepsilon_x,\varepsilon_y}$ comes from equation (7), $\sigma_{\text{g.m.}}(M, T)$ is the standard deviation of the geometric mean of two horizontal *Sa* values (the quantity presented in most ground-motion prediction models), and $\sigma_H(M, T)$ is the standard deviation of a single component *Sa* (the quantity used in equations 1, 2, and 18). The $f_H(M, R, T, \theta)$ term is unchanged regardless of whether the geometric mean or arbitrary component *Sa* definition is adopted, so it can be taken from the ground-motion prediction model without modification. Equations (17) through (19) can be easily implemented in a computer code alongside an existing ground-motion prediction, and the output used in the same way as the output from any other ground-motion prediction (e.g., for PSHA analysis). It is expected that this "custom" model and corresponding hazard analysis will allow an engineer to increase the precision of response analyses in cases where the structure of interest has different periods of vibration in its two principle directions.

Simulation of Response Spectra

The correlation predictions derived above can be used to simulate response spectra given an earthquake scenario. For a given earthquake magnitude (M), distance (R), and other parameters (θ), the distribution of horizontal ln *Sa* values at a range of periods ($T_1, T_2, \ldots, T_n$) can be obtained using the model of equation (1). The mean of ln $Sa(T_i)$ is equal to $f_H(M, R, T_i, \theta)$ and its standard deviation is equal to $\sigma_H(M, T_i)$. The covariance of ln $Sa(T_i)$ and ln $Sa(T_j)$ is equal to $\sigma_H(M, T_i)\sigma_H(M, T_j)\rho_{\varepsilon_x,\varepsilon_x}(T_i, T_j)$, where $\rho_{\varepsilon_x,\varepsilon_x}(T_i, T_j)$ is computed using equation (9). If we again assume a multivariate normal distribution for ln *Sa* values, then these means and covariances fully define the distribution that can be used for simulation. A comparison of empirical spectra and simulated spectra is shown in Figure 11. In Figure 11a and b, spectra are simulated using zero correlation and perfect correlation, respectively. In Figure 11c, spectra are simulated using the correlation prediction from equation (9). In Figure 11d, real spectra from recorded ground motions are shown for comparison. It is clear that the results in Figure 11a and b are not accurate representations of real record spectra, and so a model such as that presented here is needed for accurate simulation. Note that this simulation procedure can also be applied to simulation of multiple-component response spectra.

These simulated spectra may also be compared with synthetic ground motions to verify that the spectra of the synthetic motions show sufficient variability (or "roughness"). In cases where synthetic spectra are not "rough" enough corrective actions could be taken to increase the variability (e.g., "rough" spectra were generated in Sewell *et al.*, 1996).

Conclusions

Models have been presented for correlations of spectral response values of earthquake ground motions. The models include predictions of correlations for single-component ground motions measured at two differing periods, and also across orthogonal components of three-dimensional ground motions. The predictive models improve upon previous predictions of correlations at differing periods in a single horizontal ground motion. For correlations within a vertical ground motion or across orthogonal components of a ground motion, these predictions are believed to be the first of their kind. It is seen that correlations of *Sa* at differing periods across orthogonal horizontal components of a ground motion can be approximated as a product of the correlation of *Sa* values across differing components (at a single period) and the correlation at differing periods (in the same component).

The predictions are presented in the form of correlations of standardized residuals (epsilons) from established empirical ground-motion prediction models. These predictions provide information about the correlations of two logarithmic *Sa* values for a given magnitude and distance. Although the observed residuals in principle depend on the ground-motion prediction model chosen, the correlations do not vary significantly when the underlying model is changed. Thus the correlation predictions are applicable regardless of the ground-motion prediction model used by the analyst. This suggests that although one might repeat this exercise when the models from the current next generation attenuation project are released, the functional forms and even parameter values should not change appreciably, at least for vertical or randomly oriented horizontal components. (Correlations among or within fault normal and fault parallel components were not examined as part of this study, but it will be possible to examine them upon completion of the Next Generation Attenuation project.)

Several approximate "rule-of-thumb" correlation values can be determined from the preceding models. Correlation of orthogonal horizontal *Sa* values with the same period are comparatively highly correlated ($\rho \cong 0.8$), whereas horizontal and vertical *Sa* values at typical first-mode periods of midrise buildings are less correlated ($\rho \cong 0.3$–0.4). Several past and potential future applications are presented, illustrating that the correlations shown here are useful for a variety of earthquake hazard and engineering problems. Increased knowledge of response spectrum correlations will facilitate the further development of vector-valued probabilistic seismic hazard analysis and allow simple modification of existing ground-motion prediction models to develop custom predictions for any combination of periods and orientations. These applications will allow analysts to better understand the properties of multicomponent earthquake ground motions.

Data Source

The response spectra, magnitude, distance, and site condition data for all of the ground motions studied here came from the PEER Strong Motion Database (2000), http://peer.berkeley.edu/smcat/ (last accessed 29 June 2005).

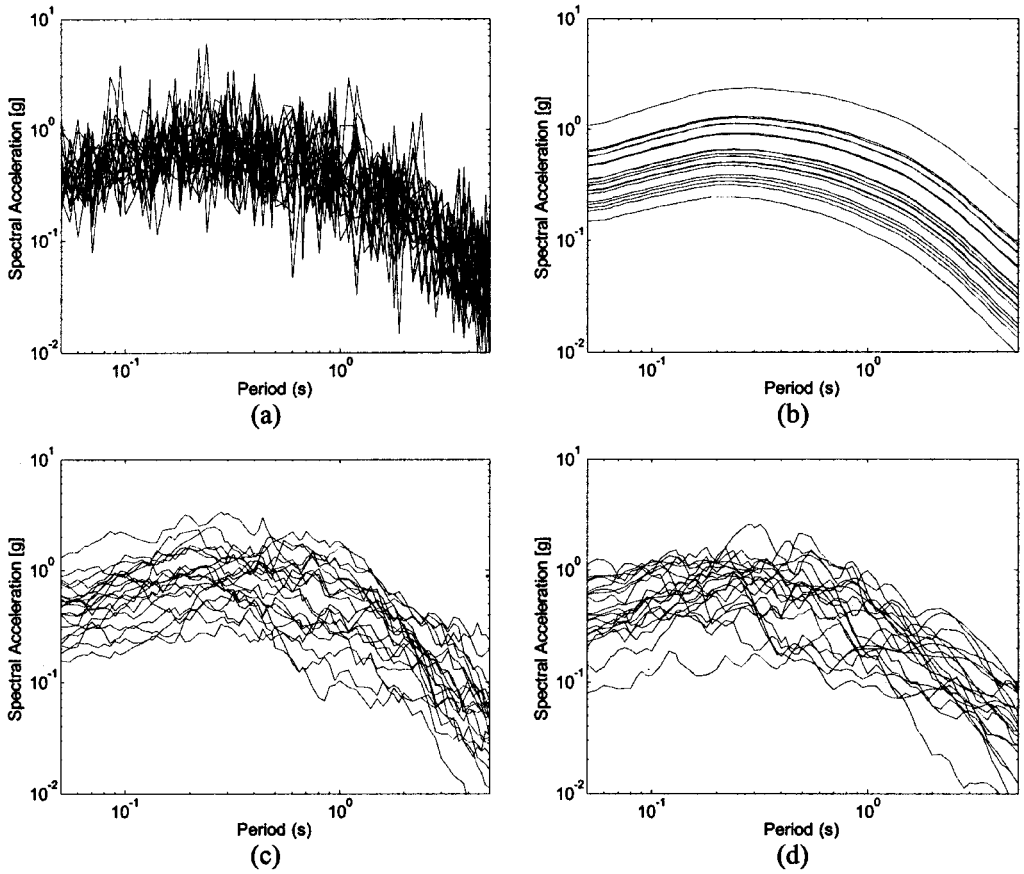

Figure 11. Samples of 20 response spectra from magnitude 6.5 earthquakes with a source-to-site distance of 8 km. The simulated spectra use means and variances from Abrahamson and Silva (1997). (a) Simulated spectra using correlation coefficients equal to zero between all periods. (b) Simulated spectra using correlation coefficients equal to one between all periods. (c) Simulated spectra using correlation coefficients from equation (9). (d) Real spectra from recorded ground motions with magnitude ≅6.5 and distance ≅ 8 km.

Acknowledgments

This work was supported primarily by the Earthquake Engineering Research Centers Program of the National Science Foundation, under Award Number EEC-9701568 through the Pacific Earthquake Engineering Research Center (PEER). Any opinions, findings and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation. The authors thank Paolo Bazzurro, Chris Cramer, and Gabriel Toro for their helpful reviews of this manuscript.

References

Abrahamson, N. A., and W. J. Silva (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seism. Res. Lett.* **68,** no. 1, 94–126.

Baker, J. W., and C. A. Cornell (2004). Choice of a vector of ground motion intensity measures for seismic demand hazard analysis, in *Proceedings, 13th World Conference on Earthquake Engineering*, Vancouver, Canada, 1–6 August 2004, 15 pp.

Baker, J. W., and C. A. Cornell (2005a). A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon, *Earthquake Eng. Struct. Dyn.* **34,** no. 10, 1193–1217.

Baker, J. W., and C. A. Cornell (2005b). *Vector-Valued Ground Motion Intensity Measures for Probabilistic Seismic Demand Analysis*, Stanford University, John A. Blume Earthquake Engineering Center Report 150, Stanford, California. http://blume.stanford.edu/Blume/Publications.htm (last accessed 29 June 2005)

Baker, J. W., and C. A. Cornell (2006). Which spectral acceleration are you using? *Earthquake Spectra* **22** (in press).

Bazzurro, P., and C. A. Cornell (2002). Vector-valued probabilistic seismic hazard analysis, in *7th U.S. National Conference on Earthquake Engineering*, Boston, Massachusetts, 21–25 July 2002, 10 pp.

Boore, D. M. (2005). Erratum: Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work, *Seismo. Res. Lett.* **76,** no. 3, 368–369.

Boore, D. M., W. B. Joyner, and T. E. Fumal (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work, *Seism. Res. Lett.* **68,** no. 1, 128–153.

Campbell, K. W. (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra, *Seism. Res. Lett.* **68,** no. 1, 154–179.

Chopra, A. K. (2001). *Dynamics of Structures: Theory and Applications to*

Earthquake Engineering, Prentice Hall, Upper Saddle River, New Jersey, 844 pp.

Cordova, P. P., G. G. Deierlein, S. S. F. Mehanny, and C. A. Cornell (2001). Development of a two-parameter seismic intensity measure and probabilistic assessment procedure, in *The Second U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures*, Sapporo, Hokkaido, 11–13 September 2000, 187–206.

Ditlevsen, O. (1981). *Uncertainty Modeling: With Applications to Multidimensional Civil Engineering Systems*, McGraw-Hill International Book Co., New York, 412 pp.

Efron, B., and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*, Chapman & Hall, New York, 436 pp.

Inoue, T., and C. A. Cornell (1990). *Seismic Hazard Analysis of Multi-Degree-of-Freedom Structures*, Reliability of Marine Structures, RMS-8, Stanford, California, 70 pp.

Kramer, S. L. (1996). *Geotechnical Earthquake Engineering*, Prentice Hall, Upper Saddle River, New Jersey, 653 pp.

Liu, P. L., and A. Der Kiureghian (1986). Multivariate distribution models with prescribed marginals and covariances, *Probab. Eng. Mech.* **1,** no. 2, 105–112.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman (1996). *Applied Linear Statistical Models*, McGraw-Hill, Boston, Massachusetts, 1408 pp.

Norwegian Technology Standards Institution (1999). *Actions and Action Effects*, N-003, Norwegian Technology Standards Institution, Oslo, Norway, 85 pp. www.standard.no/imaker.exe?id = 1399 (last accessed 13 March 2005).

Pacific Gas & Electric (1988). Final report of the Diablo Canyon long term seismic program, U.S. Nuclear Regulatory Commission, Docket Numbers 50-275 and 50-323, Washington, D.C.

PEER Strong Motion Database (2000). http://peer.berkeley.edu/smcat/ (last accessed 29 June 2005).

Penzien, J., and M. Watabe (1975). Characteristics of 3-D earthquake ground motions, *Earthquake Eng. Struct. Dyn.* **3,** 365–373.

Sewell, R. T., G. R. Toro, and R. K. McGuire (1996). Impact of ground motion characterization on conservatism and variability in seismic risk estimates, U.S. Nuclear Regulatory Commission, NUREG/CR-6467, Washington, D.C.

Shome, N., and C. A. Cornell (1999). Probabilistic seismic demand analysis of nonlinear structures, RMS Program, RMS-35, Stanford, California 320 pp. www.stanford.edu/group/rms/ (last accessed 14 March 2005).

Somerville, P. G., and H. K. Thio (2003). Probabilistic vector-valued ground motion intensity measures and engineering demand measures for the PEER Van Nuys Holiday Inn PBEE testbed, SCEC 2003 Project Report.

Spudich, P., W. B. Joyner, A. G. Lindh, D. M. Boore, B. M. Margaris, and J. B. Fletcher (1999). SEA99: a revised ground motion prediction relation for use in extensional tectonic regimes, *Bull. Seism. Soc. Am.* **89,** no. 5, 1156–1170.

Stewart, J. P., S.-J. Chiou, J. D. Bray, R. W. Graves, P. G. Somerville, and N. A. Abrahamson (2001). Ground motion evaluation procedures for performance-based design, Pacific Earthquake Engineering Research Center, University of California at Berkeley, PEER 2001-09, Berkeley, California, 229 pp.

Wang, G.-Q., X.-Y. Zhou, Z.-J. Ma, and P.-Z. Zhang (2001). A preliminary study on the randomness of response spectra of the 1999 Chi-Chi, Taiwan, earthquake, *Bull. Seism. Soc. Am.* **91,** no. 5, 1358–1369.

Department of Civil and Environmental Engineering
Stanford University
Stanford, California 94305-4020

Manuscript received 25 March 2005.