Correlation of Spectral Acceleration Values from NGA Ground Motion Models

Jack W. Baker,[a] M.EERI, **and Nirmal Jayaram**[a]

Ground motion models (or "attenuation relationships") describe the probability distribution of spectral acceleration at an individual period, given a set of predictor variables such as magnitude and distance, but they do not address the correlations between spectral acceleration values at multiple periods or orientations. Those correlations are needed for several calculations related to seismic hazard analysis and ground motion selection. Four NGA models and the NGA ground motion database are used here to measure these correlations, and predictive equations are fit to the results. The equations are valid for periods from 0.01 seconds to 10 seconds, versus similar previous equations that were valid only between 0.05 and 5 seconds and produced unreasonable results if extrapolated. Use of the new NGA ground motion database also facilitates a first study of correlations from intra- and inter-event residuals. Observed correlations are not sensitive to the choice of accompanying ground motion model, and intra-event, inter-event, and total residuals all exhibit similar correlation structure. A single equation is thus applicable for a variety of correlation predictions. A simple example illustrates the use of the proposed equations for one hazard analysis application. [DOI: 10.1193/1.2857544]

INTRODUCTION

The utility of ground motion models (GMMs) can be extended if correlations of response spectral values at multiple periods or orientations (e.g., fault-normal/fault-parallel) are known. These correlations allow existing ground motion models to be adopted for predicting the joint distribution of spectral acceleration values at multiple periods, which is useful for vector-valued probabilistic seismic hazard analysis and generation of custom ground motion models. Predictions of these correlations have been previously proposed, but the NGA project's new GMMs and expanded ground motion library facilitate the development of an improved equation that is applicable over a wider range of periods. In this paper, the methodology for measuring and predicting these correlations is briefly outlined, a new correlation equation is developed, and the new results are examined and compared to previous comparable equations.

To describe precisely the correlations being studied, it is helpful to note that ground motion predictions take the following general form

[a] Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020

Earthquake Spectra, Volume 24, No. 1, pages 299–317, February 2008; © 2008, Earthquake Engineering Research Institute

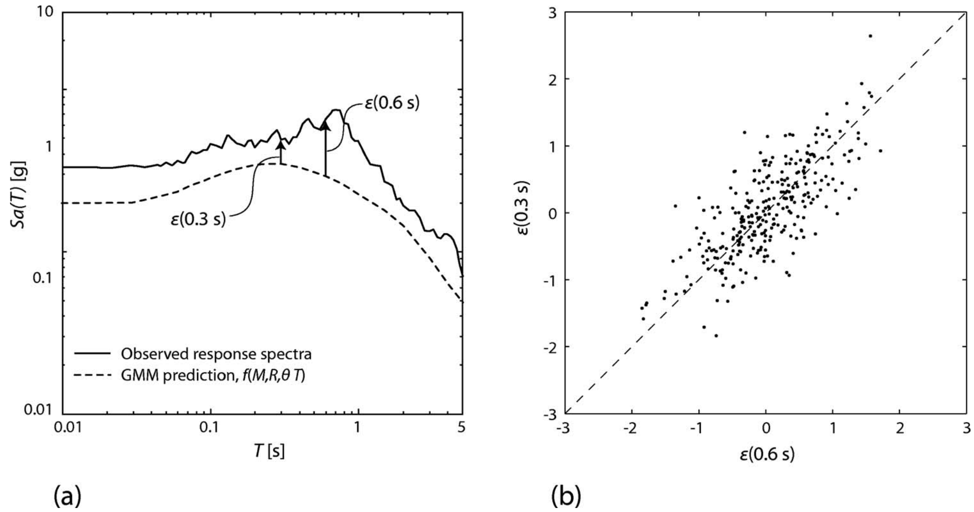

Figure 1. Illustration of the ε correlations being studied. (a) Predicted and observed response spectrum for one ground motion, with ε values at 0.3 s and 0.6 s highlighted. (b) Observed $\varepsilon(0.3\text{ s})$ and $\varepsilon(0.6\text{ s})$ values from a set of ground motions.

$$\ln Sa(T) = f(M, R, \theta, T) + \varepsilon(T) \tag{1}$$

where $f(M, R, \theta, T)$ is the predicted mean of the natural log of spectral acceleration (Sa) at a specified period (T), provided by the ground motion model. This predicted mean is a function of the earthquake magnitude (M), distance (R), and other parameters (θ) such as local site conditions and faulting mechanism. The term $\varepsilon(T)$ represents the difference between the actual logarithmic spectral acceleration, $\ln Sa(T)$, and its predicted mean value $f(M, R, \theta, T)$. For observed ground motions with known $Sa(T)$ and known M, R, etc., $\varepsilon(T)$ is a known number. For future ground motions, $\varepsilon(T)$ is represented by a random variable with a mean value of zero. The standard deviation of $\varepsilon(T)$ is estimated as part of standard ground motion models; this standard deviation is a function of the spectral acceleration period, and in some models is also a function of the earthquake magnitude. The symbol ε is sometimes used to describe the number of standard deviations between $f(M, R, \theta, T)$ and $\ln Sa(T)$, rather than the simple difference shown in Equation 1. Measured correlation coefficients are identical if either of these two ε definitions are used, however, so the results below are also applicable when using the "number of standard deviations" definition of ε.

The values of $\varepsilon(T)$ at differing periods are related probabilistically. For example, if a recorded spectral acceleration is stronger than expected (i.e., $\varepsilon(T)$ is greater than 0) at a given period, then it is likely to also be stronger than expected at adjacent periods. This relationship can be described probabilistically using correlation coefficients between ε's, as a function of the two periods of interest. To illustrate, an example response spectrum and predicted mean value are shown in Figure 1a. Epsilon values from a large set of ground motions are shown in Figure 1b, and they can be used to estimate an empirical correlation coefficient, as will be described below. The process will then be repeated for spectral acceleration values at many periods. For a ground motion with a given M, R, etc., $\varepsilon(T)$ is the only source of uncertainty in $\ln Sa$. Thus, the correlation of ε's completely describes the correlation of spectral acceleration values for the case of a single M, R, etc.

Note that there are several ways in which spectral acceleration might be defined in application, and $f(M, R, \theta, T)$ and the standard deviation of $\varepsilon(T)$ are affected by the choice of definition used (Boore et al. 2006). Results will be presented here for several definitions, but it will be seen that observed correlations are very similar in all cases. Comparisons will also be made to previous models that predicted correlations of this type (Abrahamson et al. 2003, Baker and Cornell 2006a, Inoue and Cornell 1990).

DEVELOPMENT OF CORRELATION EQUATIONS

The NGA ground motion library (http://peer.berkeley.edu/nga) was used to develop response spectra data for analysis, and four NGA ground motion models were used to compute predicted spectral acceleration values (Abrahamson and Silva 2008, Boore and Atkinson 2008, Campbell and Bozorgnia 2008, Chiou and Youngs 2008). For each record and period, the observed and predicted spectral acceleration values were used in Equation 1 to compute ε values. The usable period range differs for each record, so records are only used when both periods of interest fell within this range. Approximately 1000 to 2500 records are available at moderate periods, and approximately 300 to 500 records are available at 10 seconds (where the fewest records are available due to filtering of low-frequency signals in most records). The stated range of available records reflects variation in the number of ground motions used by the various ground motion models considered; only records used by the GMM authors were used to compute ε's associated with each model.

The correlation coefficient between two sets of observed ε values (e.g., Figure 1b) can be estimated using the maximum likelihood estimator (Kutner et al. 2004). Sometimes referred to as the Pearson product-moment correlation coefficient, it estimates the correlation coefficient between $\varepsilon(T_1)$ and $\varepsilon(T_2)$ as

$$\rho_{\varepsilon(T_1),\varepsilon(T_2)} = \frac{\sum_{i=1}^{n} (\varepsilon_i(T_1) - \overline{\varepsilon(T_1)})(\varepsilon_i(T_2) - \overline{\varepsilon(T_2)})}{\sqrt{\sum_{i=1}^{n} (\varepsilon_i(T_1) - \overline{\varepsilon(T_1)})^2 \sum_{i=1}^{n} (\varepsilon_i(T_2) - \overline{\varepsilon(T_2)})^2}} \tag{2}$$

where $\varepsilon_i(T_1)$ and $\varepsilon_i(T_2)$ are the ith observations of $\varepsilon(T_1)$ and $\varepsilon(T_2)$, $\overline{\varepsilon(T_1)}$ and $\overline{\varepsilon(T_2)}$ are their sample means, and n is the total number of observations (records). For example, a $\rho_{\varepsilon(T_1),\varepsilon(T_2)}$ of 0.74 is estimated from the data shown in Figure 1b. This calculation is repeated for each period pair of interest. The resulting correlations could be tabulated and used in a look-up table when needed, but the table would be difficult to transfer or reproduce in print. For this reason, the empirical correlation coefficients were fit with an analytical predictive equation that is easier to communicate.

The form of the predictive equation is chosen by judgment, and nonlinear least-squares regression is used to find the associated coefficients. The nonlinear least-squares algorithm works best when the errors for each observed value are of comparable size, but this is not the case here because correlation coefficients estimated in Equation 2 have non-constant standard errors. That is, one can have high confidence in observed correlations that are close to 1 or −1, but lower confidence in observed correlations that are close to 0. To correct for this, the Fisher z transformation (Kutner et al. 2004) was applied to the correlation coefficients

$$z = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right) \tag{3}$$

where ρ is the estimated correlation coefficient from Equation 2 and z is the transformed data. This "variance stabilizing transformation" leads to z values that have constant standard errors, so that the least-squares algorithm is optimal. Least-squares regression is then applied to these transformed values, rather than the original correlations, as follows

$$\min_{\boldsymbol{\beta}} \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{2}\ln\left(\frac{1+\rho_{i,j}}{1-\rho_{i,j}}\right) - \frac{1}{2}\ln\left(\frac{1+\hat{\rho}_{i,j}(\boldsymbol{\beta})}{1-\hat{\rho}_{i,j}(\boldsymbol{\beta})}\right)\right)^2 \tag{4}$$

where $\rho_{i,j}$ is the empirical correlation coefficient computed in Equation 2 at the period pair (T_i, T_j) and $\hat{\rho}_{i,j}(\boldsymbol{\beta})$ is the predicted correlation using a given analytical equation and a vector of associated coefficients $\boldsymbol{\beta}$. Results from this calculation are presented in the following section.

OBSERVED CORRELATIONS AND PREDICTIVE EQUATIONS

CORRELATIONS FOR GMRotI VALUES

The spectral acceleration definition used in the NGA ground motion models is typically "GMRotI50," also referred to as "GMRotI." This is the geometric mean of spectral accelerations of orthogonal horizontal components, after rotating the components to minimize the variation of the rotation-dependent geometric means over a range of periods (Boore et al. 2006). Correlation coefficients are first evaluated for this *Sa* definition.

Empirical correlation coefficients from four NGA GMMs (Abrahamson and Silva 2008, Boore and Atkinson 2008, Campbell and Bozorgnia 2008, Chiou and Youngs 2008), at a variety of period pairs, are shown in Figure 2 and Figure 3. Figure 2 shows correlation coefficients for a selected set of periods T_2, plotted versus T_1 values between 0.01 and 10 seconds. Figure 3 shows the same results, plotted using contours of correlation coefficients as a function of both T_1 and T_2. These results were used to fit a predictive equation, which is plotted in Figure 4. To evaluate this predictive equation, a series of initial calculations are first performed

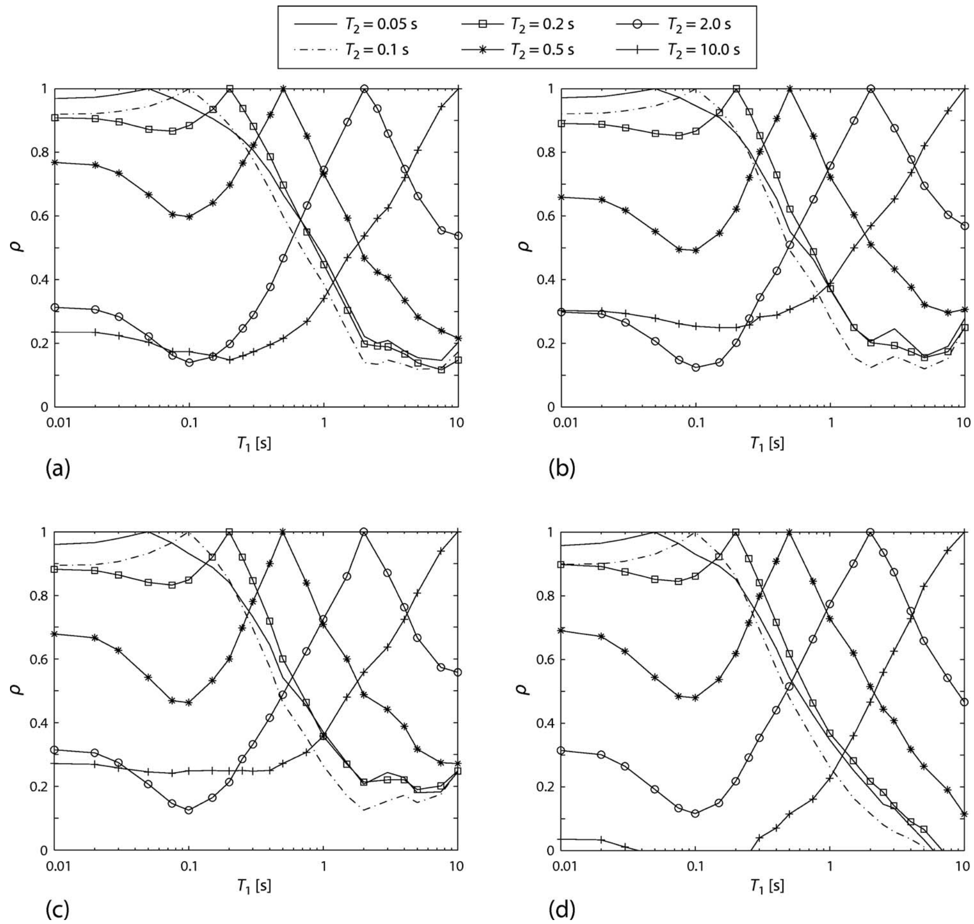

Figure 2. Plots of empirical correlation coefficients versus T_1, for several T_2 values. (a) Abrahamson and Silva (2008) model. (b) Boore and Atkinson (2008) model. (c) Campbell and Bozorgnia (2008) model. (d) Chiou and Youngs (2008) model.

$$C_1 = 1 - \cos\left(\frac{\pi}{2} - 0.366\ln\left(\frac{T_{\max}}{\max(T_{\min}, 0.109)}\right)\right)$$

$$C_2 = \begin{cases} 1 - 0.105\left(1 - \dfrac{1}{1 + e^{100T_{\max} - 5}}\right)\left(\dfrac{T_{\max} - T_{\min}}{T_{\max} - 0.0099}\right) & \text{if } T_{\max} < 0.2 \\ 0 & \text{otherwise} \end{cases}$$

$$C_3 = \begin{cases} C_2 & \text{if } T_{\max} < 0.109 \\ C_1 & \text{otherwise} \end{cases}$$

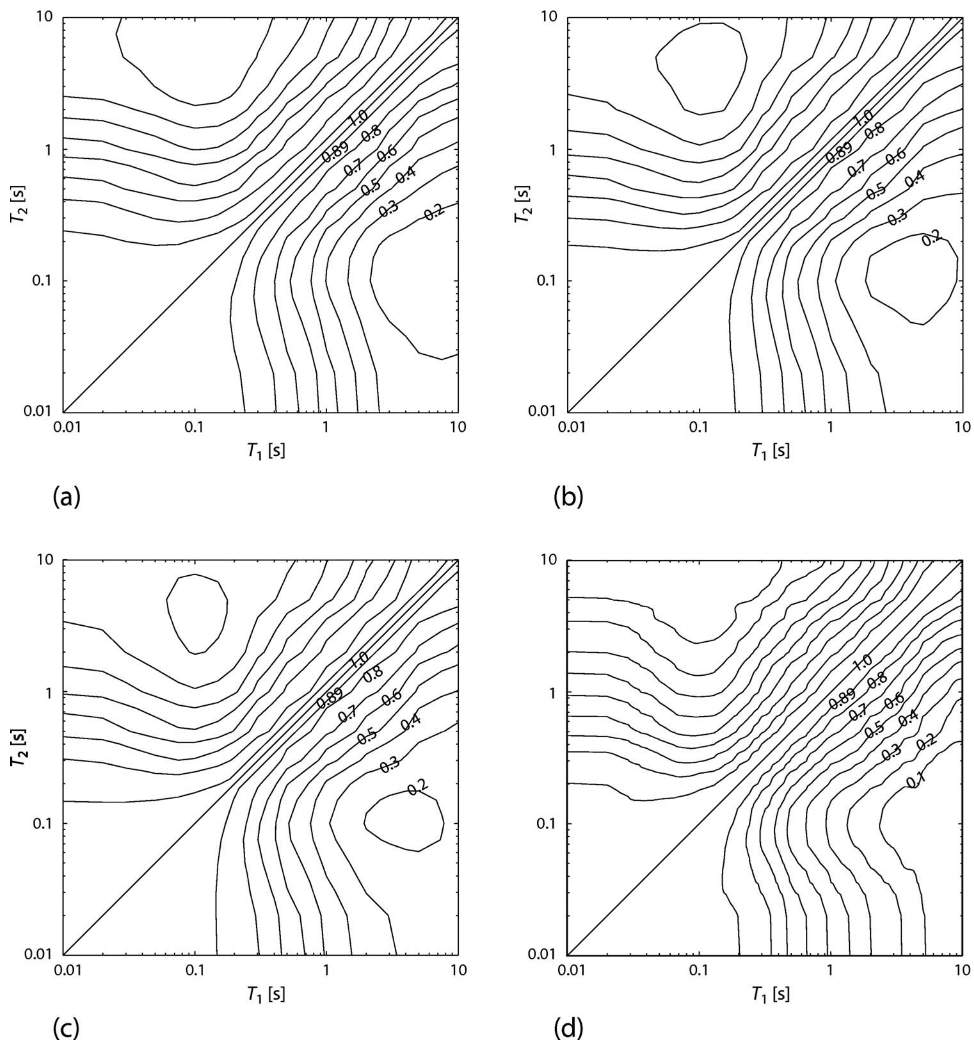

Figure 3. Contours of empirical correlation coefficients versus T_1 and T_2. Abrahamson and Silva (2008) model. (b) Boore and Atkinson (2008) model. (c) Campbell and Bozorgnia (2008) model. (d) Chiou and Youngs (2008) model.

$$C_4 = C_1 + 0.5(\sqrt{C_3} - C_3)\left(1 + \cos\left(\frac{\pi T_{\min}}{0.109}\right)\right) \tag{5}$$

where $T_{\min}=\min(T_1, T_2)$ and $T_{\max}=\max(T_1, T_2)$. The predicted correlation coefficient is then given by

$$\text{if } T_{\max} < 0.109 \quad \rho_{\varepsilon(T_1),\varepsilon(T_2)} = C_2$$

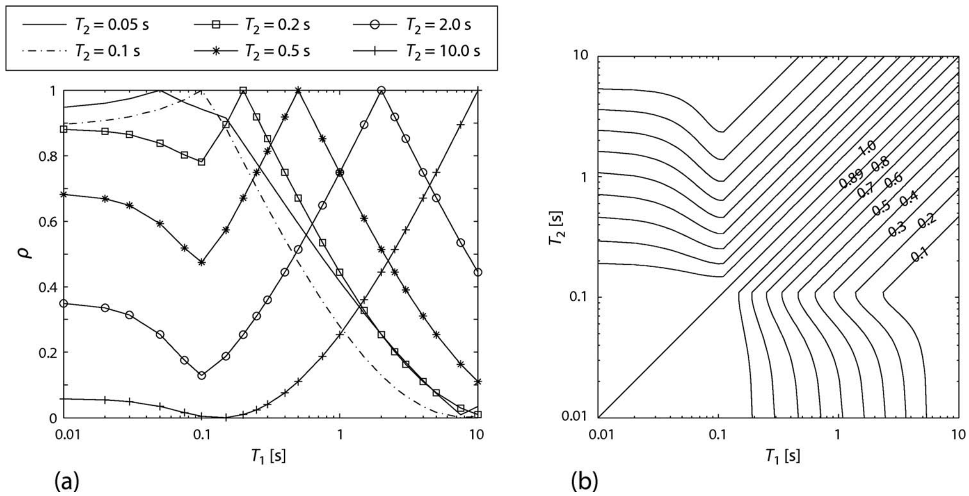

Figure 4. Plots of the predictions from Equation 6. (a) Plots of correlation coefficients versus T_1, for several T_2 values. (b) Contours of correlation coefficients versus T_1 andT_2.

$$\text{else if } T_{\min} > 0.109 \quad \rho_{\varepsilon(T_1),\varepsilon(T_2)} = C_1$$

$$\text{else if } T_{\max} < 0.2 \quad \rho_{\varepsilon(T_1),\varepsilon(T_2)} = \min(C_2, C_4)$$

$$\text{else} \quad \rho_{\varepsilon(T_1),\varepsilon(T_2)} = C_4 \tag{6}$$

The equation is valid when T_1 and T_2 are between 0.01 and 10 seconds. The form of the equation has no physical interpretation: it is simply a fit to observed data, and thus should not be extrapolated. Note that the numerical coefficients present in Equations 5 and 6 are the $\boldsymbol{\beta}$ coefficients obtained from Equation 4. This equation is slightly more complicated than previously proposed correlation equations, due to the need to fit observed data over a larger period range.

Data from the Chiou and Youngs (2008) model was used for the fitting, so the prediction matches that data most closely, although the agreement with data from the other models is good. At moderate periods, the close agreement of the various models (and their associated varying datasets) suggests that the computed correlations results are very stable and have little uncertainty. The only noticeable differences between ground motion models occur when T_1 and/or T_2 is larger than five seconds. Because the underlying GMMs and measured correlations are less precise at these long periods, due to the smaller number of ground motions with appropriate filter frequencies, creation of differing correlation predictions for each GMM was judged to be inappropriate. For example, if 300 observations are used to estimate a correlation coefficient of 0.1, the resulting 95% confidence interval for that estimate is −0.01 to 0.21 (Kutner et al. 2004), meaning that it is difficult to identify statistically significant differences among the low correlation values where results from the four models differ. (In contrast, if 300 obser-

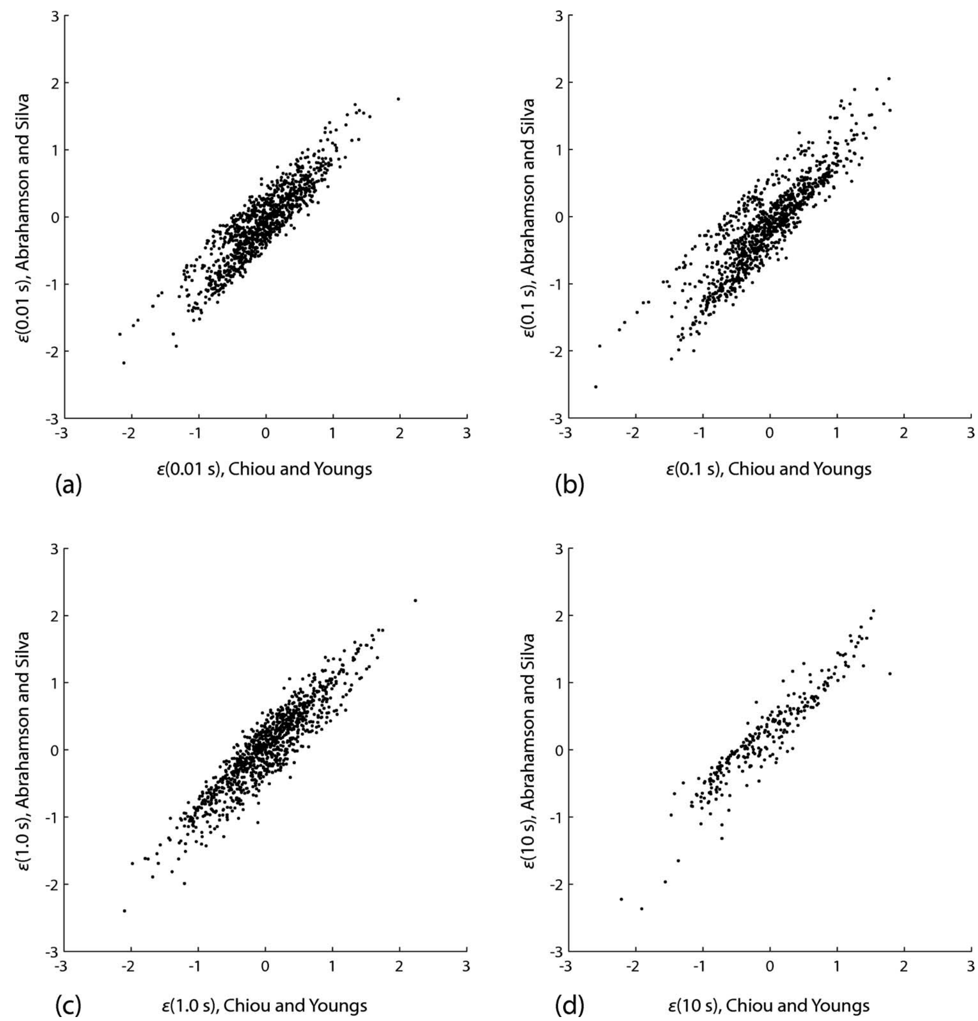

Figure 5. Scatter plots of ε values obtained from two ground motion prediction models. Values computed using the Abrahamson and Silva (2008) model are shown on the y axes, and values computed using the Chiou and Youngs (2008) model are shown on the x axes. (a) ε values at 0.01 s. (b) ε values at 0.1 s. (c) ε values at 1.0 s. (d) ε values at 10 s.

vations are used to estimate a correlation of 0.9, the 95% confidence interval is 0.88 to 0.92.) Perhaps as importantly, a difference of 0.1 or 0.2 in a low correlation value has very little practical impact on the joint distribution of two variables, which is the output of interest from this model.

Further support for using the same model with all ground motion models is provided by Figure 5, which shows scatter plots from observed ε's associated with two GMMs, at several periods. The correlation coefficients of these four data sets are between 0.88 and

0.9, indicating that the ε values associated with a given ground motion do not differ significantly when the underlying ground motion model is changed (similar results were observed when comparing ε values from other GMMs). Because of the similarity of observed ε values and observed correlation coefficients between models, Equation 6 is recommended for use with all four GMMs considered here. Note that in addition to reproducing empirical correlations, a predictive equation must also produce correlation matrices that are positive definite (in order to avoid predicting conditional standard deviations that are negative). The positive definiteness of predictions from Equation 6 has been verified.

CORRELATIONS FOR OTHER SA DEFINITIONS

Spectral acceleration definitions other than GMRotI may also be of interest to some users. The *Sa* of a single ground motion component may be of interest, or it may be useful to compute the geometric mean of *Sa*'s from two orthogonal ground motion components (Baker and Cornell, 2006b). It can be shown that the mean values of these spectral accelerations are nearly identical to the mean value of GMRotI (Beyer and Bommer, 2006), so the standard NGA ground motion models were used in Equation 1 to compute ε's and resulting correlations associated with these other *Sa* definitions. (Note that the standard deviation of $\ln Sa$ may vary by definition, but Equation 2 is not affected by the standard deviation, so correlation results do not require it to be known.) Figure 6 shows empirical correlations for GMRotI, geometric mean, and single-component spectral accelerations, using the Chiou and Youngs ground motion model. The similarity in results suggests that Equation 6 can be used for all of these *Sa* definitions. Note that for some widely spaced period pairs, negative correlations are observed; this may be because systematic differences between predicted and observed spectral shapes cause one period to be over-predicted while another is under-predicted. The negative correlations are small and not of any obvious practical concern in hazard calculations.

The results shown in Figure 6 are for specific orientations of the ground motion components (i.e., fault-normal and fault-parallel), rather than arbitrary orientations. Arbitrarily-oriented components also show the same correlation structure, which perhaps should not be surprising given the similarities in Figure 6. The strong empirical similarity is also supported by Baker and Cornell (2005, Appendix B), who showed analytically that correlations between geometric-mean and arbitrary-component *Sa*'s are nearly identical. Baker and Cornell also observed that correlations were independent of the ground motions' causal magnitudes and distances; that finding was assumed to hold here as well.

CORRELATIONS FOR SA VALUES OF ORTHOGONAL GROUND MOTION COMPONENTS

Correlations in spectral accelerations of orthogonal components of ground motions are also of interest (e.g., for determining the distribution of spectral amplitudes when analyzing 3-dimensional structures). Empirical correlations are shown in Figure 7 for *Sa* values of orthogonal ground motion components, when both *Sa* values have the same period. Results from four ground motion models are shown along with the correlation

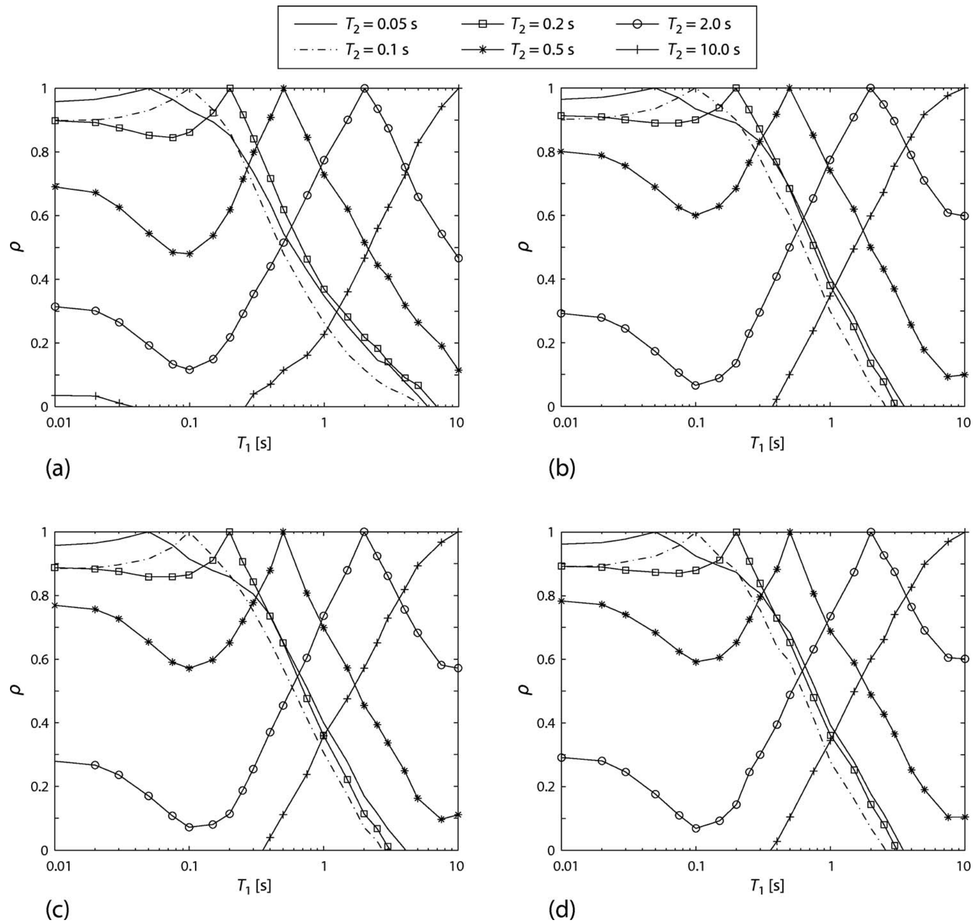

Figure 6. Empirical correlation coefficients from the Chiou and Youngs (2008) model versus T_1, for several T_2 values. (a) Correlations of GMRotI spectral accelerations. (b) Correlations of the geometric mean of the fault-normal and fault-parallel spectral accelerations. (c) Correlations of fault-normal spectral accelerations. (d) Correlations of fault-parallel spectral accelerations.

prediction from Baker and Cornell (2006a). The previous prediction closely matches the correlations observed here, even though the prediction was designed for a more restricted period range. That previous predictive equation is repeated here, given its continued validity

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

where $\rho_{\varepsilon_x(T),\varepsilon_y(T)}$ is used to denote the correlation between two epsilons, ε_x and ε_y, associated with orthogonal ground motion components at a given period T. Only the correlations from the Campbell and Bozorgnia model do not fit this predictive equation

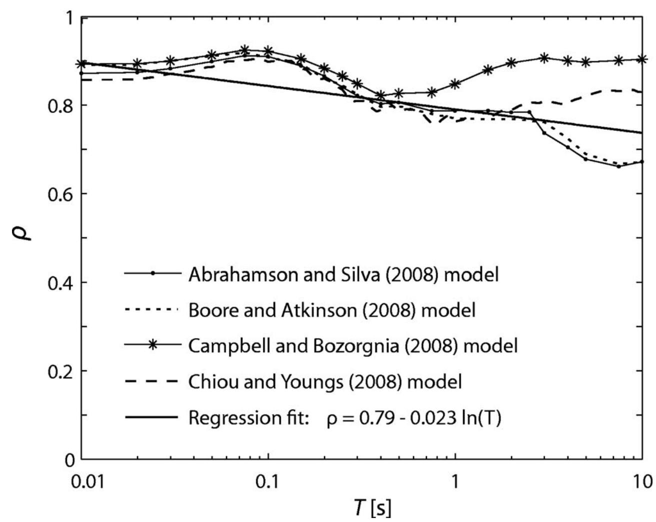

Figure 7. Observed and predicted correlations of ε values from orthogonal components of ground motion at a single period.

well, but the difference is again primarily at longer periods where limited data affects the accuracy of the models. For simplicity, and because of the potential effect of limited data at long periods, Equation 7 is recommended for use with all ground motion models.

In the case of spectral acceleration values in orthogonal directions with different periods, it appears that correlations can be estimated using a product of correlation predictions given in Equations 6 and 7 above. To evaluate Equation 7 in the case where two periods are of interest, the mean of the logarithmic periods is used, given the linearity of correlations as a function of $\ln(T)$. Figure 8 shows empirical correlations of orthogonal epsilon values at two periods, along with a correlation prediction given by

$$\rho_{\varepsilon_x(T_1),\varepsilon_y(T_2)} = \rho_{\varepsilon(T_1),\varepsilon(T_2)} \cdot (0.79 - 0.023 \cdot \ln(\sqrt{T_1 T_2})) \tag{8}$$

where $\rho_{\varepsilon(T_1),\varepsilon(T_2)}$ is given in Equation 6 and the term in parentheses is Equation 7 evaluated at the mean of the logarithmic periods (i.e., the log of the geometric mean of the periods). Because this prediction reasonably matches the empirical results shown in Figure 8, it is recommended as the predictive equation, rather than fitting a new equation. Estimating this correlation as the product of two other correlation coefficients corresponds to an assumption of Markovian dependence, where $\varepsilon_x(T_1)$ and $\varepsilon_y(T_2)$ are conditionally independent given either $\varepsilon_x(T_2)$ or $\varepsilon_y(T_1)$, and where an approximation is made that $\rho_{\varepsilon_x(T_1),\varepsilon_y(T_1)} \cong \rho_{\varepsilon_x(T_2),\varepsilon_y(T_2)} \cong \rho_{\varepsilon_x(\sqrt{T_1T_2}),\varepsilon_y(\sqrt{T_1T_2})}$ (Baker and Cornell 2006a). This product form also has the desirable property of matching Equation 7 in the special case where $T_1 = T_2$.

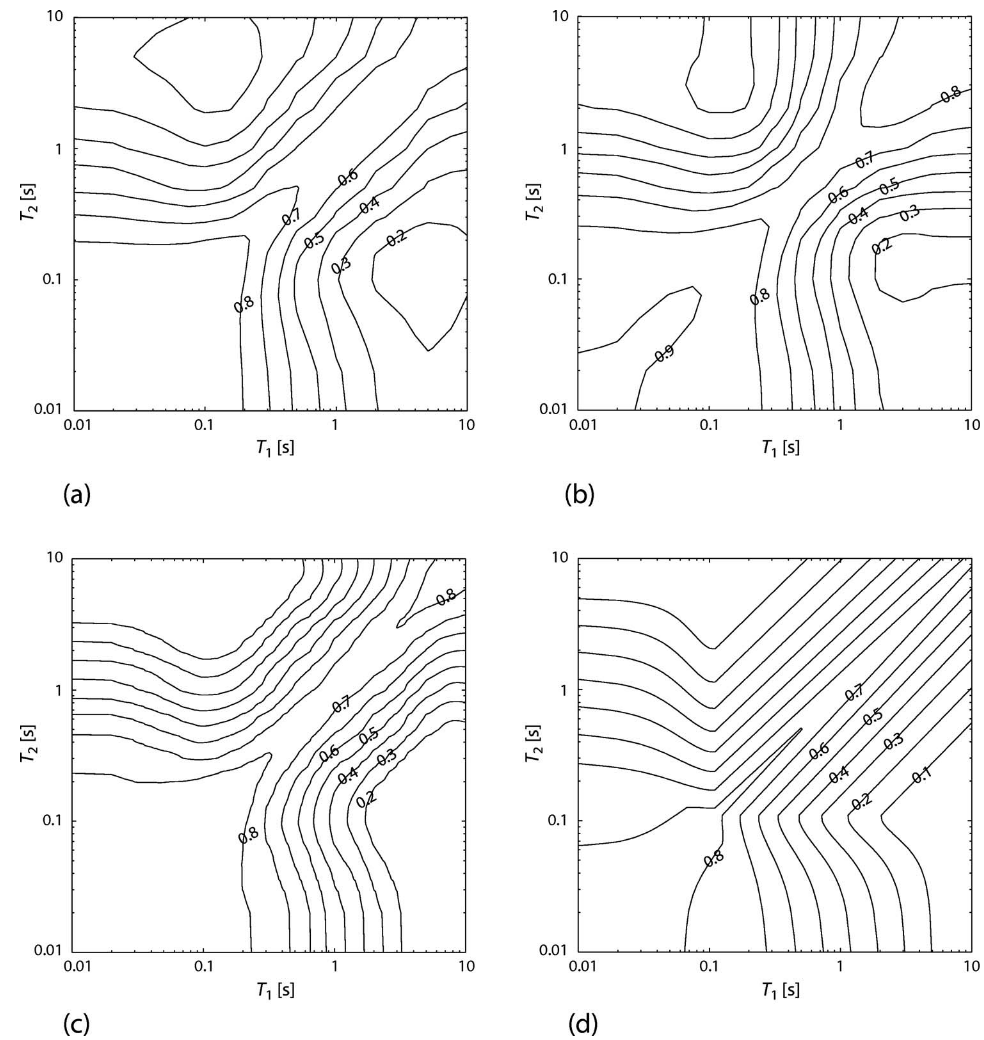

Figure 8. Contours of correlation coefficients versus T_1 andT_2, for orthogonal components of ground motion. (a) Empirical correlations from the Boore and Atkinson (2008) model. (b) Empirical correlations from the Campbell and Bozorgnia (2008) model. (c) Empirical correlations from the Chiou and Youngs (2008) model. (d) Predicted correlations using Equation 8.

CORRELATIONS FOR INTRA- AND INTER-EVENT RESIDUALS

Modern ground motion models refine Equation 1 by separating $\varepsilon(T)$ into two terms

$$\ln Sa(T) = f(M, R, \theta, T) + \eta(T) + \varepsilon'(T) \tag{9}$$

where $\eta(T)$ and $\varepsilon'(T)$ denote the inter-event and intra-event residual terms, respectively (Abrahamson and Youngs 1992). The inter-event term $\eta(T)$ is identical for all observed

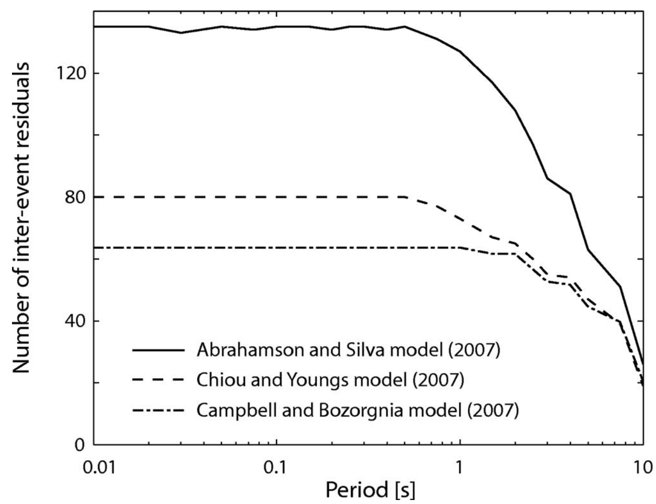

Figure 9. Number of observed inter-event residuals, by period and NGA model.

$Sa(T)$ values from a given earthquake event, while the intra-event term $\varepsilon'(T)$ is unique for each observed ground motion. Most users will need only correlations for the total $\varepsilon(T)$ as formulated in Equation 1, but the correlation coefficients for these inter- and intra-event epsilons may also be of interest.

By equating Equations 9 and 1, it is apparent that the total $\varepsilon(T)$ in Equation 1 is equal to $\varepsilon(T) = \eta(T) + \varepsilon'(T)$. The standard deviation of $\varepsilon'(T)$ is known to be significantly larger than the standard deviation of $\eta(T)$, so $\varepsilon'(T)$ is the dominant contributor to the total $\varepsilon(T)$ in Equation 1 and thus the correlations among intra-event residuals are well-modeled by Equation 6. This has been verified by observing that empirical correlation coefficients for $\varepsilon'(T)$ are nearly identical to those shown in Figure 2 and Figure 3, and has also been noted by Abrahamson and Silva (2007).

The NGA ground motion database is the first library with enough events to allow computation of a significant number of $\eta(T)$ values. A plot of the number of available $\eta(T)$ values is given in Figure 9; the number varies by period and by model, due to usable period limitations and differences in the events used by each author. The values of these residuals were obtained from the model authors, rather than being back-calculated, to ensure that no errors were introduced while calculating residuals. Empirical correlations of $\eta(T)$'s are shown in Figure 10 for three NGA models with available $\eta(T)$'s. While the limited number of observations causes more variability in these plots than in the comparable figures for total epsilons, the general trends are consistent with those observed for intra-event and total residual terms. (The limited number of observations also increases the uncertainty in the estimates, making it difficult to use statistical hypothesis testing to identify differences.) While it may be tempting to think of inter-event residuals as being uniformly positive or negative over large period ranges due to, for example, stress-drop variations between events, this result suggests that inter-event re-

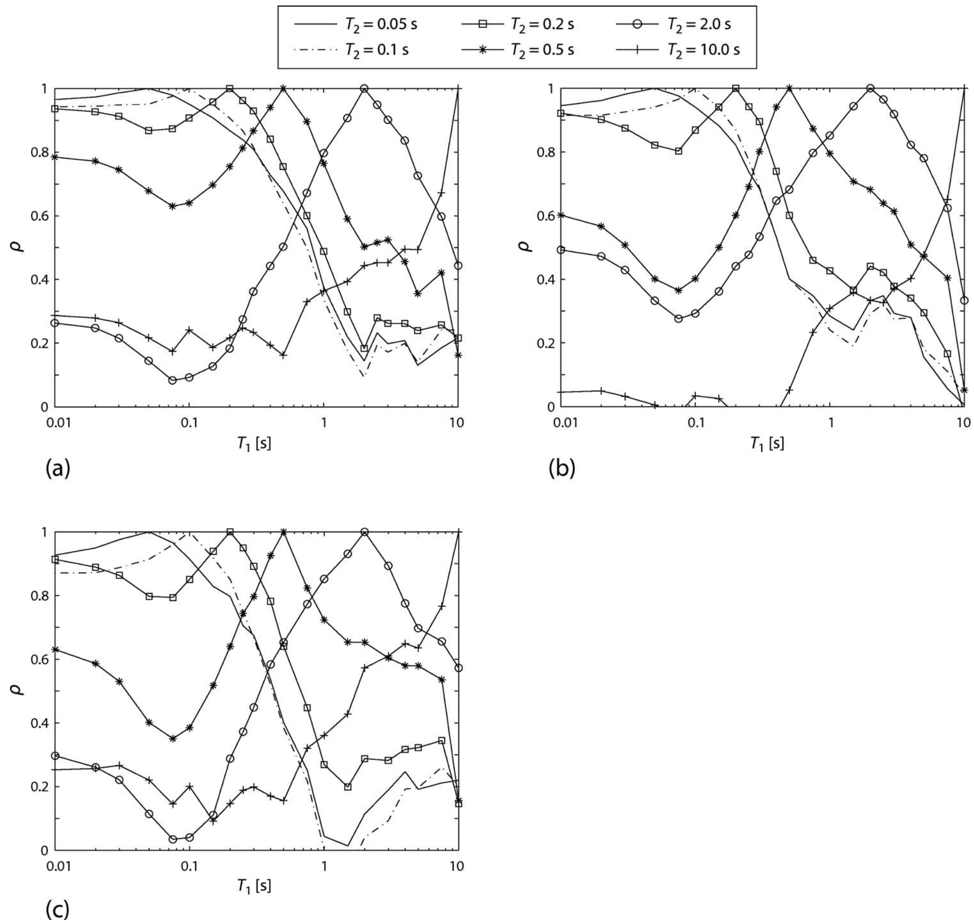

Figure 10. Plots of inter-event residual empirical correlation coefficients versus T_1, for several T_2 values. (a) Abrahamson and Silva (2008) model. (b) Chiou and Youngs (2008) model. (b) Campbell and Bozorgnia (2008) model.

siduals vary with period in the same way that intra-event residuals do. The use of Equation 6 to describe $\eta(T)$ correlations appears to be reasonable (note that Abrahamson and Silva 2007 also suggest that inter-event correlations be modeled using correlations from total epsilons).

COMPARISON TO PREVIOUS CORRELATION EQUATIONS

Correlation equations of the type described here have been proposed previously (Abrahamson et al. 2003, Baker and Cornell 2006a, Inoue and Cornell 1990). None of those equations, however, used a dataset as large as the NGA ground motion library, and they were valid over a smaller period range than the results given here. Those equations were also developed using ground motion models that will be superseded by the NGA

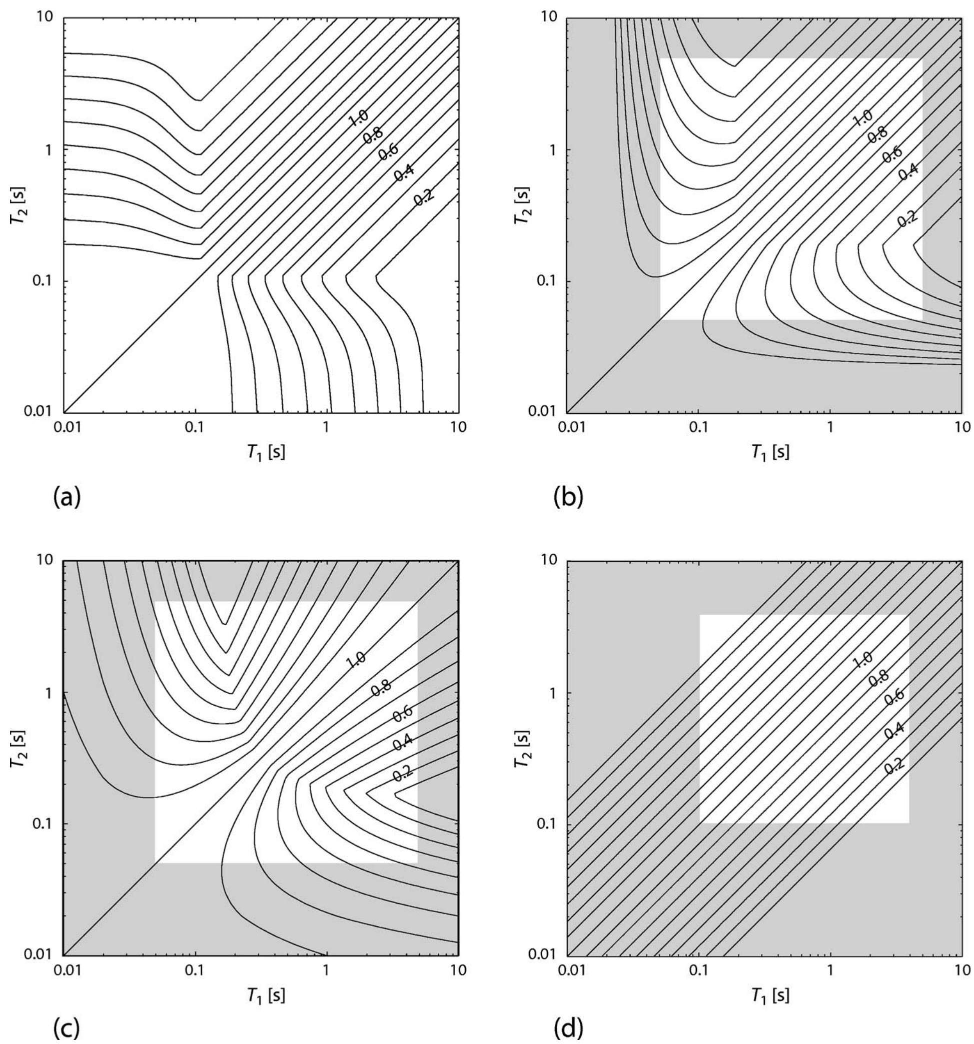

Figure 11. Contours of predicted correlation coefficients versus T_1 and T_2. Shaded regions indicate the period ranges for which the models are being extrapolated. (a) Predicted correlations using Equation 6. (b) Predicted correlations using Baker and Cornell (2006a). (c) Predicted correlations using Abrahamson et al. (2003). (d) Predicted correlations using Inoue and Cornell (1990).

models. Contours of predicted correlation coefficients are shown in Figure 11 for the three previously proposed correlation equations and the equation proposed here. Shaded regions correspond to extrapolations of the predictive equations, and it is clear that the extrapolations are not reasonable for covering an extended period range. Within their ranges of applicability, the four predictions produce similar results. It should be noted that Baker and Cornell (2006a) also predicted correlations involving response spectra of vertical ground motions, but those predictions are not revisited here, as the NGA ground motion models do not predict vertical *Sa* values.

EXAMPLE APPLICATION: GROUND MOTION MODEL FOR SA AVERAGED OVER A PERIOD BAND

To illustrate one use of the correlation equations presented above, a brief example application is presented here. Standard ground motion models provide the distribution of spectral acceleration values to be expected from an earthquake with a given magnitude, distance, etc., and for a given period T. When combined with the correlation equations presented above, a customized ground motion model can be created to predict the distribution of spectral acceleration averaged over any arbitrary period band. Consider the geometric mean of spectral acceleration values at a set of periods

$$Sa_{avg}(T_1, \ldots, T_n) = \left(\prod_{i=1}^{n} Sa(T_i) \right)^{1/n} \tag{10}$$

where $T_1, \ldots, T_n$ are n periods of interest, and $Sa_{avg}()$ denotes spectral acceleration averaged over a range of periods. This measure of ground motion intensity may be valuable if one desires to predict response of a structure that is sensitive to multiple periods of excitation. It is also useful for smoothing out peak-to-trough variability in response spectra, and thus providing a measure of average response spectra intensity. The product form in Equation 10 is chosen so that its logarithm is a simple summation

$$\ln Sa_{avg}(T_1, \ldots, T_n) = \frac{1}{n} \sum_{i=1}^{n} \ln Sa(T_i) \tag{11}$$

The $\ln Sa(T_i)$ terms have a joint normal distribution, so it follows that $\ln Sa_{avg}(T_1, \ldots, T_n)$ is also normally distributed (a property that would not hold if the geometric mean of Equation 10 was replaced with an arithmetic mean). Because the $\ln Sa(T_i)$ terms in Equation 11 are described by existing GMMs, the mean and standard deviation of $\ln Sa_{avg}(T_1, \ldots, T_n)$ can be easily computed as

$$E[\ln Sa_{avg}(T_1, \ldots, T_n)] = \frac{1}{n} \sum_{i=1}^{n} f(M, R, \theta, T_i) \tag{12}$$

$$Var[\ln Sa_{avg}(T_1, \ldots, T_n)] = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{\ln Sa(T_i), \ln Sa(T_j)} \sigma_{\ln Sa(T_i)} \sigma_{\ln Sa(T_j)} \tag{13}$$

where $\rho_{\ln Sa(T_i), \ln Sa(T_j)}$ is given in Equation 6, and $f(M, R, \theta, T_i)$ and $\sigma_{\ln Sa(T_i)}$ are the mean and standard deviation, respectively of $\ln Sa(T_i)$ as specified by standard ground motion models. The notation $E[]$ is used to denote the expected value, or mean, of $\ln Sa_{avg}$, and $Var[\ ]$ denotes its variance (i.e., the square of its standard deviation). Note that Equations 12 and 13 are the conditional logarithmic mean and variance for a given value of magnitude, distance, etc., as with standard ground motion prediction models for $\ln Sa(T)$. With the mean and standard deviation given by Equations 12 and 13, and with knowledge that $\ln Sa_{avg}$ is normally distributed, it is simple to perform probabilistic seismic hazard analysis in terms of $Sa_{avg}(T_1, \ldots, T_n)$ rather than the traditional Sa at a single period.

Note that the terms in the summation of Equation 11 could be weighted, and it would still be straightforward to compute the mean and variance of the summation. This might be useful if one wanted to create a structure-specific Sa_{avg} that weighted spectral accelerations at a structure's modal periods, with weights corresponding to the structure's modal participation factors. The formulation presented in this section has been previously used in various forms by several authors (Abrahamson et al. 2003, Cordova et al. 2001, Pacific Gas & Electric 1988; Shome and Cornell 1999).

In addition to the custom ground motion model described in this section, there exist several other potential applications of the correlation coefficient predictions. Vector-valued probabilistic seismic hazard analysis requires these correlations to compute the probability of joint occurrence of specified spectral acceleration values at multiple periods (Bazzurro and Cornell 2002). Correlations are also needed for ground motion selection procedures that account explicitly for the spectral shape of ground motions having "rare" or large spectral amplitudes at a specified period (Baker and Cornell 2006c).

CONCLUSIONS

Equations have been presented to predict correlations of spectral acceleration values, using the NGA ground motion library and the new NGA ground motion models (GMMs). A predictive equation was presented that provides correlations between logarithmic spectral accelerations at two periods. This equation was observed to be valid for a variety of definitions of spectral acceleration (i.e., spectral acceleration of an individual component, the geometric mean of spectral accelerations from two orthogonal components, and the orientation-independent GMRotI definition used by the NGA modelers). Additional equations were provided to predict correlations of spectral accelerations from orthogonally-oriented individual ground motion components. Correlations are modeled by computing correlations between prediction residuals, termed ε's. These residuals are defined as either the number of standard deviations between observed and predicted spectral acceleration values or simply the direct difference between observed and predicted values, but the computed correlations are the same for either definition.

The correlation equations are applicable for use with any of the NGA ground motion models, at periods between 0.01 and 10 seconds. When the periods of interest are less than 5 seconds, correlation coefficients from all considered models are essentially identical. If one period is greater than 5 seconds and the second period is significantly less than 5 seconds, correlations vary slightly among models. These variations are likely due to a lack of empirical data, and these widely-spaced period pairs are also of less engineering interest, so separate correlation equations were not developed for each model. The similarity of correlations from the various GMMs occurs because the correlations are dominated by the large record-to-record variability in observed spectral values from similar events. While slight differences in mean predicted values from the GMMs may be important for some applications, they do not affect computed ε values to a large enough extent that correlations change noticeably. This also explains why previous predictive equations fit to older GMMs produce similar correlations to those observed here (within their more limited period ranges).

In addition to providing an updated prediction for correlations of total residuals from spectral acceleration predictions, results were presented for inter-event and intra-event residuals. Intra-event residuals have essentially identical correlation structure to the total residuals, due to their dominant contribution to the total residual. Inter-event residuals also have a similar correlation structure to the total residuals, suggesting that their correlations can also be predicted using the equation developed for total residuals. Correlations for inter-event residuals have not been examined prior to the NGA project, due to the lack of a sufficiently large ground motion library.

An example application was presented to illustrate how the correlation equations can be combined with a standard ground motion model to produce a custom prediction of spectral acceleration averaged over a period range—a measure of ground motion intensity that may be useful for predicting response of structures affected by excitation at multiple periods or multiple orientations. Using the results of the example calculation, it is straightforward to perform seismic hazard analysis for this measure of ground motion intensity. Several other seismic hazard analysis and ground motion selection applications will also benefit from these updated equations.

ACKNOWLEDGMENTS

Thanks to Norm Abrahamson, David Boore, Yousef Bozorgnia, Ken Campbell, Brian Chiou, Bob Youngs and Jennie Watson-Lamprey for sharing some of the data used above. This work was partially supported 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.

REFERENCES

Abrahamson, N. A., and Youngs, R. R., 1992. A stable algorithm for regression analysis using the random effects model, *Bull. Seismol. Soc. Am.* **82**, 505–510.

Abrahamson, N. A., Kammerer, A., and Gregor, N., 2003. *Summary of scaling relations for spectral damping, peak velocity, and average spectral acceleration: Report for the PEGASOS project*, Personal communication.

Abrahamson, N. A., and Silva, W. J., 2007. *Abrahamson & Silva NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters*, Pacific Earthquake Engineering Research Center, University of California at Berkeley, Berkeley, California, 338 pp.

——, 2008. Summary of the Abrahamson & Silva NGA ground motion relations, *Earthquake Spectra* **24**, 67–97.

Baker, J. W., and Cornell, C. A., 2005. *Vector-valued ground motion intensity measures for probabilistic seismic demand analysis*, John A. Blume Earthquake Engineering Center, Report # *150*, Stanford, CA, 321 pp. http://blume.stanford.edu/Blume/Publications.htm

——, 2006a. Correlation of response spectral values for multi-component ground motions, *Bull. Seismol. Soc. Am.* **96**, 215–227.

——, 2006b. Which spectral acceleration are you using?, *Earthquake Spectra* **22**, 293–312.

——, 2006c. Spectral shape, epsilon and record selection, *Earthquake Eng. Struct. Dyn.* **35**, 1077–1095.

Bazzurro, P., and Cornell, C. A., 2002. Vector-valued probabilistic seismic hazard analysis, *7th U.S. National Conference on Earthquake Engineering*, Boston, MA, 10 pp.

Beyer, K., and Bommer, J. J., 2006. Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion, *Bull. Seismol. Soc. Am.* **96**, 1512–1522.

Boore, D. M., Watson-Lamprey, J., and Abrahamson, N. A., 2006. Orientation-independent measures of ground motion, *Bull. Seismol. Soc. Am.* **96**, 1502–1511.

Boore, D. M., and Atkinson, G. M., 2008. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthquake Spectra* **24**, 99–138.

Campbell, K. W., and Bozorgnia, Y., 2008. Campbell-Bozorgnia NGA horizontal ground motion model for PGA, PGV, PGD and 5% damped linear elastic response spectra, *Earthquake Spectra* **24**, 139–171.

Chiou, B., and Youngs, R. R., 2008. An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthquake Spectra* **24**, 173–215.

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

Inoue, T., and Cornell, C. A., 1990. *Seismic hazard analysis of multi-degree-of-freedom structures*, Reliability of Marine Structures, RMS-8, Stanford, CA, 70 pp.

Kutner, M. H., Nachtsheim, C., and Neter, J., 2004. *Applied Linear Regression Models*, McGraw-Hill/Irwin, New York, 701 pp.

Pacific Gas & Electric, 1988. *Final report of the Diablo Canyon long term seismic program*, US Nuclear Regulatory Commission, *Docket Numbers 50-275 and 50-323*.

Shome, N., and Cornell, C. A., 1999. *Probabilistic seismic demand analysis of nonlinear structures*, RMS Program, RMS-35, Stanford, CA, 320 pp. http://www.stanford.edu/group/rms/

(Received 29 June 2007; accepted 5 November 2007)