THE SEISMOLOGICAL SOCIETY OF AMERICA
400 Evelyn Ave., Suite 201
Albany, CA 94706-1375
(510) 525-5474; FAX (510) 525-7204
www.seismosoc.org

Bulletin of the Seismological Society of America, Vol. 101, No. 1, pp. 160–175, February 2011, doi: 10.1785/0120100038

Model Validations and Comparisons of the Next Generation Attenuation of Ground Motions (NGA–West) Project

by James Kaklamanos and Laurie G. Baise

Abstract The earthquake ground-motion prediction equations developed for the Next Generation Attenuation of Ground Motions (NGA–West) project in 2008 have established a new baseline for the estimation of ground-motion parameters, such as peak ground acceleration, peak ground velocity, and spectral acceleration, for shallow crustal earthquakes in active tectonic regions. We perform statistical goodness-of-fit analyses to quantitatively compare the predictive capabilities of the NGA models and their predecessors, using several testing subsets of the master database used in model development. In addition, we perform a blind comparison test using 1060 ground-motion records from seven recent earthquakes recorded in California: the 2003 **M** 6.5 San Simeon event, 2004 **M** 6.0 Parkfield event, 2005 **M** 5.2 Anza event, 2007 **M** 5.4 Alum Rock event, 2008 **M** 5.4 Chino Hills event, 2010 **M** 7.2 Baja event, and 2010 **M** 5.7 Ocotillo event. We assess how modeling decisions regarding the regression dataset, functional forms, input parameters, and model complexity influence the models' predictive capabilities. By comparing the performance of each model, we discuss various ground-motion modeling strategies and offer recommendations for model development. We find that increased model complexity does not necessarily lead to increased prediction accuracy, that the inclusion of aftershocks in regression datasets may result in decreased predictive capabilities for mainshocks, and that the use of measured site characteristics leads to greatly improved ground-motion predictions. A model validation framework is introduced to assess the prediction accuracy of ground-motion prediction equations and to aid in their future development.

Online Material: Additional information on testing datasets.

Introduction

The purpose of ground-motion prediction equations (also called ground-motion prediction relations or attenuation relationships) is to predict the ground motion at a given location as a function of earthquake magnitude, distance from the earthquake source, and other source, path, and site characteristics. Typical response variables in ground-motion prediction equations include peak ground acceleration (PGA), peak ground velocity (PGV), and 5%-damped pseudoabsolute response spectral acceleration (SA). A typical ground-motion prediction equation (GMPE) has the form

$$\ln Y = f(\mathbf{M}, R, \sum \text{Source}_i, \sum \text{Site}_i) + \varepsilon\sigma_T, \qquad (1)$$

where $\ln Y$ is the natural logarithm of the ground-motion parameter of interest, **M** is the magnitude of the earthquake, R is a measure of distance representing the path of seismic energy from the earthquake source to the site of interest, $\sum \text{Source}_i$ are other variables relating to the earthquake source (such as type of faulting, rupture width and depth, and fault dip), $\sum \text{Site}_i$ are variables relating to the site of interest (such as average shear wave velocity, geologic characteristics, and depth to bedrock), σ_T is the aleatory standard deviation (which includes interevent and intraevent components), and ε is the number of standard deviations to be considered in the calculations (Kramer, 1996; Douglas, 2003; Abrahamson *et al.*, 2008).

Peak values of ground-motion parameters are assumed to follow a lognormal distribution; as a result, regression is typically performed on the logarithm of the ground-motion parameter of interest. GMPEs are developed for specific tectonic environments using multivariate regression on ground-motion databases, and the relationships are updated as more earthquake data are obtained (Kramer, 1996; Abrahamson and Shedlock, 1997). It is important to note that many GMPEs are not purely statistical and that nonstatistical information, such as nonlinear site response, basin effects, oversaturation, and anelastic attenuation, is often used to constrain the models

(Abrahamson *et al.*, 2008). Over the past two decades, many GMPEs have become increasingly complex in terms of the number of explanatory variables and in the complexity of the functional forms. However, Douglas (2003) and Strasser *et al.* (2009) show that although GMPEs have become increasingly more complex over time, there has not been a marked improvement in the aleatory variability (sigma) of the ground-motion prediction estimates.

After a five-year effort, the Next Generation Attenuation of Ground Motions (NGA–West, or NGA) project was completed in 2008. The project, coordinated by the Pacific Earthquake Engineering Research Center, established five new GMPEs that predict ground-motion parameters for shallow crustal earthquakes in active tectonic regions (such as California). These models are the first large-scale update of GMPEs for this tectonic environment since 1997, when the previous generation of GMPEs was released. Table 1 lists the new NGA models and each model's predecessor, as well as the abbreviations for all models (Abrahamson and Silva, 1997; Abrahamson and Silva, 2008; Boore *et al.*, 1997; Boore and Atkinson, 2008; Campbell, 1997; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008a; Idriss, 1991; Idriss, 2008; Sadigh *et al.*, 1997).

A number of quantitative comparisons of the NGA models have been published. Ghasemi *et al.* (2008) and Shoja-Taheri *et al.* (2010) evaluate the NGA models for seismic hazard analysis in Iran. Campbell and Bozorgnia (2006), Scasserra *et al.* (2009), and Stafford *et al.* (2008) compare the NGA models with European models for seismic hazard analysis in Europe. Graves *et al.* (2008), Olsen *et al.* (2008, 2009), Star *et al.* (2008, 2010, 2011), and Stewart *et al.* (2008) compare the NGA models for simulated ground motions for various scenarios in southern California. Peruš and Fajfar (2010) employ a nonparametric approach to assess the influence of database and functional forms in GMPEs.

Many of the published comparisons of the NGA models have focused on specific regions or scenarios. In this paper, we perform a comprehensive comparison of the NGA models using an objective statistical validation framework. From the master database used to develop the five NGA models (the NGA flatfile), we develop testing subsets of ground-motion records that meet the requirements of the models. In addition, we implement the models on seven recent earthquakes that were not present in the flatfile; this represents a true blind test because the data were not used by the NGA developers. A blind comparison test is advantageous because model constraints outside of the regression are more fairly tested. For our blind comparisons, we compile a database of 1060 ground-motion records from the 2003 **M** 6.5 San Simeon earthquake, the 2004 **M** 6.0 Parkfield earthquake, the 2005 **M** 5.2 Anza earthquake, the 2007 **M** 5.4 Alum Rock earthquake, the 2008 **M** 5.4 Chino Hills earthquake, the 2010 **M** 7.2 Baja earthquake, and the 2010 **M** 5.7 Ocotillo earthquake. Characteristics of these earthquakes are presented in Table 2, and a map of the epicenters is shown in Figure 1. For the testing subsets derived from the flatfile and recent earthquakes, we test the models under different conditions, comparing their performances on mainshocks versus aftershocks, soil versus rock sites, measured versus inferred site characteristics, and various distance and magnitude ranges. To assess the level of improvement that has occurred over time, we compare the new NGA models with the previous generation of GMPEs from the 1990s. By calculating goodness-of-fit statistics, such as the Nash–Sutcliffe model efficiency coefficient (E) and the LH value, we compare the model predictions to the actual ground-motion records to assess the predictive capabilities of the models. Based on the suite of results, we discuss the manner in which model development decisions influence model performance. Ultimately, we present a model validation framework for assessing the prediction accuracy of GMPEs. Through detailed comparisons and analyses of the issues surrounding the NGA models, our suggestions may be useful in the development of future GMPEs.

Development of the NGA Models

In order to develop the NGA models, the research team compiled the NGA flatfile, an extensive dataset of 3551 ground-motion records from 173 shallow crustal earthquakes (Chiou *et al.*, 2008). In selecting the subportion of the NGA

Table 1
GMPEs Tested in This Study

NGA Models			Previous Models		
Team	Year	Abbreviation	Team	Year	Abbreviation
Abrahamson and Silva	2008	AS08	Abrahamson and Silva	1997	AS97
Boore and Atkinson	2008	BA08	Boore *et al.*	1997	BJF97
Campbell and Bozorgnia	2008	CB08	Campbell	1997	C97*
Chiou and Youngs	2008a	CY08	Sadigh *et al.*	1997	SCE97
Idriss	2008	I08	Idriss	1991	I91*

*In the early 2000s, Campbell and Bozorgnia released an update to Campbell's 1997 model (Campbell and Bozorgnia, 2003) and Idriss developed an unpublished update to his 1991 model (Idriss, 2002); however, to maintain consistency in the comparison between the sets of new and old models, we use all the GMPEs from the 1990s as the baseline for comparison.

Table 2
Recent Earthquakes in Blind Comparison Tests

Earthquake Name	Date (mm/dd/yyyy)	M	Hypocenter Latitude (°)	Hypocenter Longitude (°)	Hypocentral Depth (km)	No. of Records Used
San Simeon	12/22/2003	6.5	35.706	−121.102	4.7	30
Parkfield	09/28/2004	6.0	35.815	−120.374	7.9	94
Anza	06/12/2005	5.2	33.533	−116.578	13.5	126
Alum Rock	10/30/2007	5.4	37.432	−121.776	9.2	196
Chino Hills	07/29/2008	5.4	33.955	−117.765	13.6	377
Baja	04/04/2010	7.2	32.128	−115.303	10.0	141
Ocotillo	06/14/2010	5.7	32.698	−115.924	6.9	96

flatfile to use for regression, the model developers generally excluded records that were not representative of free-field conditions (e.g., records from tall structures, bridges, or dam crests), records from locations outside of the models' ranges of applicability, records with identified problems, and records lacking key information, although the specific decisions of each team varied (Power *et al.*, 2008). One of the most significant decisions for the researchers was whether to include aftershocks in their regression subsets. Three teams (AS08, CY08, and I08) opted to include aftershocks (Abrahamson *et al.*, 2008). Idriss (2008) only included rock sites (assumed to be sites with $V_{S30} \geq 450$ m/s); this significant difference isolates the I08 model from the others because it can only be applied to rock sites. As we will demonstrate, these dataset selection decisions greatly influence the models' prediction accuracy.

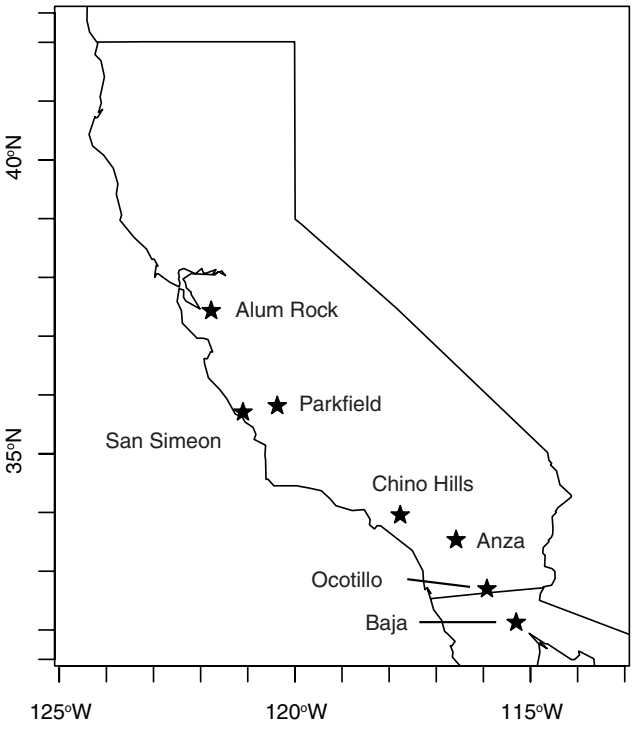

Figure 1. Map of earthquake epicenters for the seven earthquakes considered in the blind comparison tests.

Explanatory and Response Variables

Explanatory Variables

A summary of the explanatory variables used in the GMPEs is presented in Table 3. There is a wide range of model complexity, but, as a whole, the NGA models are much more complicated than their predecessors are. Based on the number of input parameters, I08 has the simplest formulation of the NGA models, followed by BA08, CB08, CY08, and AS08. Three of the NGA models (AS08, CB08, and CY08) include two or three different distance measures in the same model, whereas previous GMPEs incorporated just one per model. Four of the NGA models utilize the time-averaged shear wave velocity over the top 30 m of the subsurface (V_{S30}) as the primary site characteristic; Idriss (2008) does not quantitatively incorporate site effects, although separate coefficients for soft and hard rock sites are provided, assuming a boundary of $V_{S30} = 900$ m/s between soft and hard rock. Boore *et al.* (1997) were the first group to include V_{S30} in a GMPE. Other previous models utilize dummy variables (which take on a value of 1 if a condition is met and are 0 otherwise) to incorporate site conditions (Abrahamson and Silva, 1997; Campbell, 1997), or provide completely different sets of regression equations for soil and rock sites (Sadigh *et al.*, 1997). Furthermore, three of the NGA relations (AS08, CB08, and CY08) incorporate a depth parameter ($Z_{1.0}$ or $Z_{2.5}$) as a secondary site characteristic in addition to V_{S30}. One of the previous models (Campbell, 1997) includes a parameter D for depth to basement rock.

As illustrated in Table 3, an extensive set of explanatory variables is necessary to implement the ten GMPEs in this study. For the records in the NGA flatfile, many of the necessary parameters are explicitly included as columns, whereas others (such as the site coordinate, R_X) need to be calculated from the available information in the flatfile. For events without finite fault models, the unknown source and distance parameters were estimated using various recommendations described in Kaklamanos *et al.* (2010). Although estimates of V_{S30} are included for nearly every record in the flatfile, the depth parameters $Z_{1.0}$ (used in AS08 and CY08) and $Z_{2.5}$ (used in CB08) are not present for many records. The preferred method of determining the depth parameters is using a site-specific measured V_S profile that

Table 3
Explanatory Variables for Implementation of the GMPEs in This Study

Parameter	NGA Models					Previous Models				
	AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Source Parameters										
Moment magnitude, **M**	•	•	•	•	•	•	•	•	•	•
Depth to top of rupture, Z_{TOR}	•		•	•						
Down-dip rupture width, W	•									
Fault dip, δ	•		•	•						
Style-of-faulting flag (function of rake angle, λ)	•	•	•	•	•	•	•	•	•	•
Aftershock flag	•			•						
Path Parameters										
Closest distance to the rupture plane (rupture distance), R_{RUP}	•		•	•	•	•			•	•
Horizontal distance to the surface projection of the rupture (Joyner–Boore distance), R_{JB}	•	•	•	•			•			
Horizontal distance to top edge of rupture measured perpendicular to the strike (site coordinate), R_X	•			•						
Closest distance to the rupture plane within the zone of seismogenic rupture (seismogenic distance), R_{SEIS}								•		
Hypocentral distance, R_{HYP}										•
Hanging-wall flag	•			•		•				
Site Parameters										
Time-averaged shear wave velocity over the top 30 m of subsurface, V_{S30}	•	•	•	•			•			
Depth to bedrock or specific shear wave velocity horizon ($Z_{1.0}$, $Z_{2.5}$, or D) *	•		•	•				•		
Site conditions flag †						•		•	•	
PGA (or SA) on rock, as baseline for nonlinear site response	•	•	•	•		•				

*AS08 and CY08 use depth to $V_S = 1.0$ km/s ($Z_{1.0}$), CB08 uses depth to $V_S = 2.5$ km/s ($Z_{2.5}$), and C97 uses depth to basement rock (D).
†AS97 and SCE97 differentiate deep soil sites from sites composed of rock or shallow soil. C97 has separate categories for soil, soft rock, and hard rock.

extends to the 1.0 km/s and 2.5 km/s V_S horizons. However, only 54 sites in the NGA flatfile have measured V_S profiles that reach 1.0 km/s (Chiou and Youngs, 2008a), and even fewer reach 2.5 km/s. If a measured V_S profile or regional velocity model is unavailable, the depth parameters are determined by the recommendations of the model developers (Abrahamson and Silva, 2008; Chiou and Youngs, 2008a; Campbell and Bozorgnia, 2007). For example, Abrahamson and Silva (2008) recommend using the following median relationship to estimate $Z_{1.0}$ (m) from V_{S30} (m/s):

$$Z_{1.0} = \begin{cases} \exp(6.745) & \text{for } V_{S30} < 180 \text{ m/s} \\ \exp\left[6.745 - 1.35 \times \ln\left(\frac{V_{S30}}{180}\right)\right] & \text{for } 180 \le V_{S30} \le 500 \text{ m/s} \\ \exp\left[5.394 - 4.48 \times \ln\left(\frac{V_{S30}}{500}\right)\right] & \text{for } V_{S30} > 500 \text{ m/s}. \end{cases} \quad (2)$$

Chiou and Youngs (2008a) recommend using the following median relationship to estimate $Z_{1.0}$ from V_{S30}:

$$Z_{1.0} = \exp\left[28.5 - \frac{3.82}{8} \times \ln(V_{S30}^8 + 378.7^8)\right]. \quad (3)$$

Graphs of the two median relationships for $Z_{1.0}$ are presented in Figure 2, along with data from the 448 sites in the NGA flatfile with specified values of $Z_{1.0}$. Note that the majority of this flatfile data is from regional velocity models and not from boreholes. The large amount of scatter on this plot

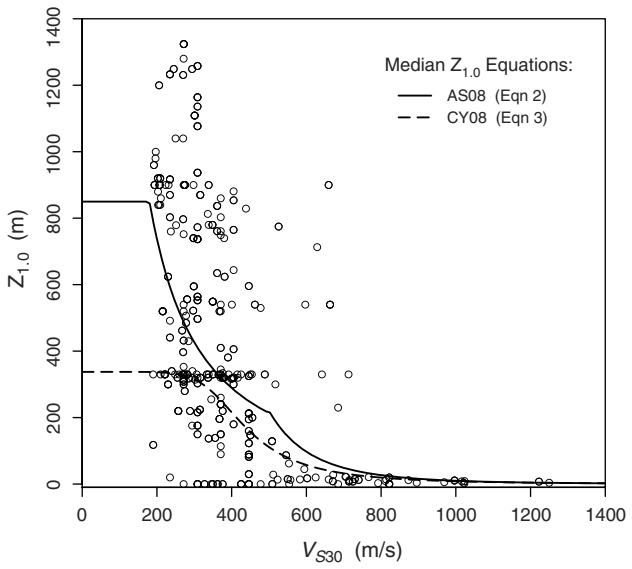

Figure 2. $Z_{1.0}$ versus V_{S30} for records in the NGA flatfile having specified values of $Z_{1.0}$. Also shown are the median equations employed by the AS08 and CY08 models for estimating $Z_{1.0}$ as a function of V_{S30}.

and its implications will be discussed in the Uncertainty of Site Parameters section.

For the seven recent earthquakes not present in the NGA flatfile, we determined the explanatory variables from alternate sources, discussed in the Data and Resources section. Ⓔ An extensive table of the explanatory and response variables that we gathered and calculated for our blind comparison tests is available as Table S1 in the electronic supplement to this paper, and may be thought of as a mini-flatfile for these recent events.

Response Variables

The response variables for the GMPEs are PGA, PGV, and 5%-damped pseudoabsolute response SA. All models have equations for PGA and SA, although the spectral periods with defined coefficients for SA vary from model to model, especially for the older models. In this study, we analyze PGA and SA for the six spectral periods represented most comprehensively in the U.S. Geological Survey (USGS) national seismic hazard maps: 0.1, 0.2, 0.3, 0.5, 1.0, and 2.0 s (Petersen *et al.*, 2008). Being represented in the national seismic hazard maps, these spectral periods have significant engineering consequences. Furthermore, all ten GMPEs have defined coefficients for these periods, so cross-comparisons can easily be made between the models. The previous GMPEs offer predictions for SA to maximum periods of 2–5 s, whereas the new GMPEs offer predictions up to 10 s. Although the new GMPEs can predict SA at long periods up to 10 s, the database of ground motions at long periods is small (Abrahamson and Silva, 2008), and the computed values of SA for long periods are more sensitive to noise (Boore and Atkinson, 2007). Accordingly, in this paper, we focus our analysis on spectral periods of 2 s and smaller.

The observed ground motions for records in the NGA database were obtained directly from the NGA flatfile. To calculate the observed ground motions for the recent earthquakes, which are not present in the flatfile, we first obtained the acceleration time histories and response spectra from various agencies with strong-motion stations (see Data and Resources section). The new and old GMPEs differ on how the two horizontal, orthogonal components of ground motion are combined to obtain a single value for a location. The previous GMPEs utilize the simple geometric mean of the as-recorded two horizontal components (GMxy), whereas the new GMPEs utilize GMRotI50, the geometric mean independent of the orientation of the instruments used to record the horizontal motion (Boore *et al.*, 2006). We utilized a FORTRAN procedure provided in Boore (2010) to compute GMRotI50 from the recorded acceleration time histories, and we utilized the conversions provided by Beyer and Bommer (2006) to convert between GMxy and GMRotI50 when necessary; fortunately, there is no systematic bias between GMxy and GMRotI50.

Development of Testing Subsets

Each model is applicable only within specific ranges of magnitude, distance, and V_{S30}. To meet the requirements of the five NGA models, a ground-motion record must have $5.0 \leq \mathbf{M} \leq 8.0$ (7.5 for normal-faulting earthquakes), R_{RUP} and R_{JB} no greater than 200 km, and $180 \leq V_{S30} \leq 1300$ m/s (if the I08 model is included, this requirement becomes $450 \leq V_{S30} \leq 1300$ m/s). In our tests, we only utilized ground-motion records that met these requirements. We also deleted records not representative of free-field conditions, records missing important information (such as SA), and records with identified problems. The most notable difference between the previous models and the new models is the smaller range of distances to which the previous models may be applied. Although some of the previous models do not specify maximum distances explicitly, 100 km is generally viewed as a reasonable limit (Campbell and Bozorgnia, 2003).

Figure 3 is a flowchart illustrating the testing subsets and the number of ground-motion records in each of the subsets. We assess the prediction accuracy of the models in various situations, including mainshocks versus aftershocks, soil versus rock sites, measured versus inferred site characteristics, and various distance and magnitude ranges. Because the BA08 and CB08 models were not developed for aftershocks, greater emphasis is placed on the mainshock subsets. In order to test the I08 model, we subdivide the datasets into soil sites ($180 \leq V_{S30} < 450$ m/s) and rock sites ($450 \leq V_{S30} \leq 1300$ m/s, where I08 is applicable). To compare the effect of distance on prediction accuracy, we utilize subdivisions of small ($R_{\mathrm{RUP}} \leq 10$ km), medium ($10 < R_{\mathrm{RUP}} \leq 100$ km), and large distances ($100 < R_{\mathrm{RUP}} \leq 200$ km).

As we will illustrate with the Parkfield earthquake, ground motion is often highly variable at distances less than 10 km (hence the first boundary), and the 100-km boundary separates the ranges of applicability of the previous and new models (we only test the previous models on the small-distance and medium-distance subsets). Finally, the NGA models are known to be biased at small-to-moderate magnitudes (Atkinson and Morrison, 2009; Chiou *et al.*, 2010), so we also subdivide the events by small-to-moderate magnitude ($\mathbf{M} \leq 6.5$) and large magnitude ($\mathbf{M} > 6.5$). In order to test all ten models simultaneously, we also create a special subset consisting of rock sites with small-to-medium distances. Ⓔ For details of the earthquakes and the corresponding numbers of records in each subset for NGA mainshocks, NGA aftershocks, and the recent earthquakes, please refer to Tables S2, S3, and S4, respectively, in the electronic supplement to this paper.

Goodness-of-Fit Measures

For each earthquake ground-motion record in each of the testing subsets, we computed median estimates of seven ground-motion parameters: PGA and SA at periods of 0.1, 0.2, 0.3, 0.5, 1.0, and 2.0 s. For details on our numerical implementation of the GMPEs, please see the Data and Resources section. Goodness-of-fit statistics are utilized to quantify the comparison of the model predictions with the observed ground-motion records. The first statistic we use as our basis of comparison is the Nash–Sutcliffe model efficiency coefficient (E), a commonly used statistic in hydrology (Nash and Sutcliffe, 1970). The coefficient of efficiency is calculated in logarithmic space by the equation

$$E = \left[1 - \frac{\sum_{i=1}^{N} (\ln Y_i - \ln \hat{Y}_i)^2}{\sum_{i=1}^{N} (\ln Y_i - \overline{\ln Y})^2}\right] \times 100\%, \qquad (4)$$

where N is the total number of ground-motion predictions, the observed values (PGA, SA, etc.) are denoted by Y_i, the predicted median values are denoted by $\hat{Y}_i$, and the mean of the logarithms of the observed values is denoted by $\overline{\ln Y}$. Because there are n ground-motion records and seven predicted ground-motion parameters per record, there are a total of $N = 7n$ ground-motion predictions in the summation. The value of E may vary between $-\infty$ and 100%; when E is less than zero, the arithmetic mean of the observed values has greater prediction accuracy than the model itself (Nash and Sutcliffe, 1970). The numerical values of E may be used to compare alternative models, with higher values indicating better agreement between observations and predictions.

The coefficient of efficiency has some important advantages over other commonly used goodness-of-fit statistics, such as the Pearson correlation coefficient (r). If a model is perfect, a plot of observed versus predicted values will follow a line with a unit slope. As an example, Figure 4 displays the observed and predicted PGA using the CB08 relationship on the aftershocks subset. Because the predicted values are generally larger than the observed values, the model is often overpredicting the low aftershock ground motions (note, however, that the CB08 model is not designed for aftershocks, and thus there is no expectation that the CB08 model should accurately predict the low ground motions generated by aftershocks). The correlation coefficient, which measures the dispersion about the least-squares regression line, will not detect this problem. However, the coefficient of efficiency, which measures the dispersion about the 1-to-1 line, will penalize the model for consistently overpredicting. Compared to many other goodness-of-fit statistics (such as r), the Nash–Sutcliffe model efficiency coefficient is more sensitive to additive and multiplicative differences between the model predictions and observations, and thus is a better indicator of goodness of fit (Legates and McCabe, 1999).

Although E adequately quantifies the prediction accuracy of the median relationships, it does not address the standard deviation relationships. We utilize the median LH value, introduced by Scherbaum *et al.* (2004), to test how well the aleatory variability (sigma) of the observations is predicted by the models. In a single number, the median LH value represents the degree to which a model's normalized residuals follow the assumption of zero mean and unit variance. As outlined by Scherbaum *et al.* (2004), the LH value for a single ground-motion prediction is given by

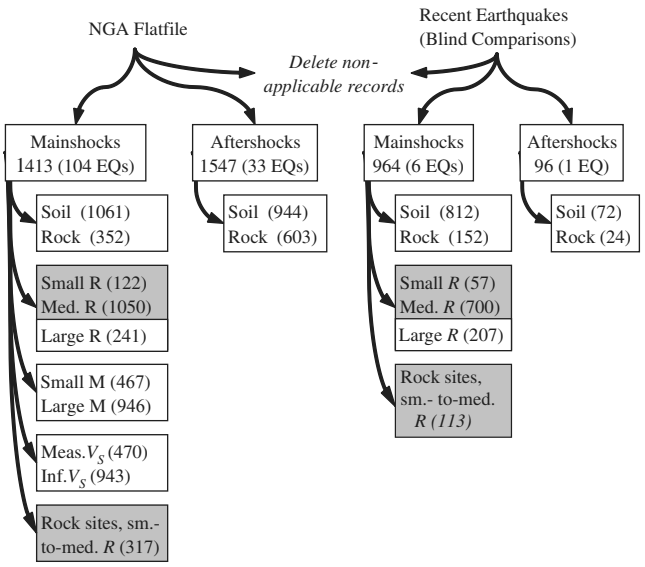

Figure 3. The subset delineation process, along with the number of records in each final subset. Datasets are divided into (a) soil sites ($180 \leq V_{S30} < 450$ m/s) and rock sites ($450 \leq V_{S30} \leq 1300$ m/s); (b) small distance ($R_{RUP} \leq 10$ km), medium distance ($10 < R_{RUP} \leq 100$ km), and large distance ($100 < R_{RUP} \leq 200$ km); (c) small magnitude ($5 \leq \mathbf{M} \leq 6.5$) and large magnitude ($6.5 < \mathbf{M} \leq 8$); (d) measured V_S and inferred V_S; and (e) a special subset consisting of rock sites with small-to-medium distances in order to test the ten models simultaneously. Because the previous models are only applicable for distances 100 km or less, they are only tested on the small-distance and medium-distance subsets (highlighted boxes).

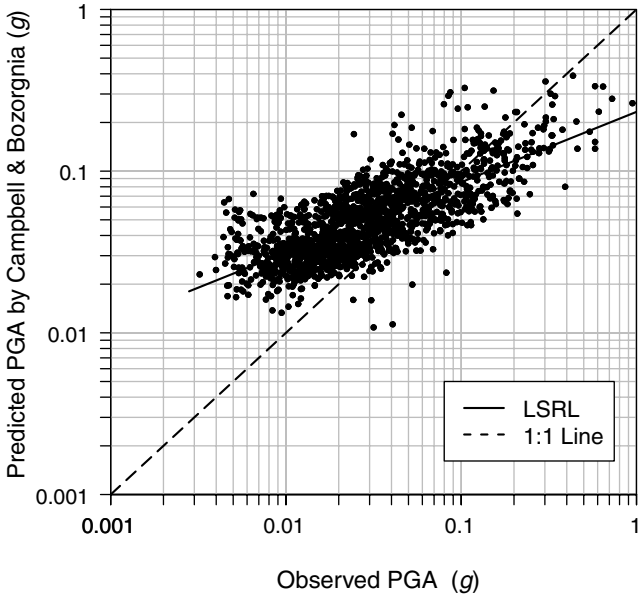

Figure 4. Advantages of the efficiency coefficient over the correlation coefficient. The observed versus predicted PGA using the CB08 relationship is plotted for the aftershocks dataset, along with the least-squares regression line (LSRL) and the ideal 1-to-1 line. The model is systematically overpredicting the actual ground motion, a phenomenon that the coefficient of efficiency—but not the coefficient of correlation—would capture. Note that aftershocks were not used to develop the CB08 model, and thus there is no expectation that the CB08 model should predict the low ground motions generated by aftershocks.

$$LH = \left[1 - \mathrm{Erf}\left(\frac{|z_0|}{\sqrt{2}}\right)\right] \times 100\%, \tag{5}$$

where Erf(•) is the error function, given by

$$\mathrm{Erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-\frac{t^2}{2}} dt, \tag{6}$$

and z_0 is the normalized model residual

$$z_0 = \frac{\ln Y - \ln \hat{Y}}{\sigma_T}. \tag{7}$$

If the model assumptions are matched exactly (that is, normalized residuals having zero mean and unit variance), the *LH* values for a set of predictions should be uniformly distributed between 0 and 100%. The median is used to summarize the distribution of *LH* values with a single number and should be equal to 50% if the model assumptions are correctly met. Thus, the median *LH* value may be used to compare alternate GMPEs; as the median *LH* value approaches 50% from either side, the goodness-of-fit increases, and the model assumptions are better met.

Results

NGA Flatfile Testing Subsets

The coefficients of efficiency and median *LH* values for the tests on the NGA flatfile are presented in Tables 4 and 5, respectively. The previous models are tested on the small-distance and medium-distance subsets, and the I08 and I91 models are only tested on the rock subsets. Comparing Tables 4 and 5, the model rankings are generally similar regardless of whether *E* or *LH* is used. For the total set of mainshocks, the values of *E* are near 70% for the four NGA models; the two models with the highest overall prediction accuracy, BA08 and CB08, are two of the simpler NGA models. The models' prediction accuracy for aftershocks is less than that for mainshocks, and there is a clear

Table 4
Coefficients of Efficiency (%) for NGA Testing Subsets

		NGA Models					Previous Models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Mainshocks (104 Events)											
Total set of mainshocks		68.8	69.0	71.2	66.9	-	-	-	-	-	-
Subdivision by site class	Soil	69.4	69.2	71.2	66.8	-	-	-	-	-	-
	Rock	67.4	68.6	71.1	67.2	64.4	-	-	-	-	-
Subdivision by distance	Small *R*	46.6	46.7	47.5	34.7	-	21.2	19.4	38.4	16.1	-
	Medium *R*	66.8	66.7	69.2	64.4	-	59.9	59.8	54.0	54.4	-
	Large *R*	26.1	29.7	35.1	30.4	-	-	-	-	-	-
Subdivision by magnitude	Small **M**	74.0	73.5	73.1	74.1	-	-	-	-	-	-
	Large **M**	62.7	63.6	68.0	59.0	-	-	-	-	-	-
Subdivision by measured/inferred V_s	Measured V_s	73.6	72.5	74.9	69.6	-	-	-	-	-	-
	Inferred V_s	64.7	65.8	67.9	63.9	-	-	-	-	-	-
Rock sites with small-to-medium *R*		70.3	70.5	73.4	67.3	66.5	59.3	61.9	61.0	57.2	54.7
Model rankings based on total *E*		3	2	1	4	-	-	-	-	-	-
Aftershocks (33 Events)											
Total set of aftershocks		59.6	47.8	44.7	60.7	-	-	-	-	-	-
Subdivision by site class	Soil	58.1	44.8	44.2	59.1	-	-	-	-	-	-
	Rock	59.9	49.7	43.0	61.2	60.5	-	-	-	-	-
Model rankings based on total *E*		2	3	4	1	-	-	-	-	-	-

Table 5
Median *LH* Values (%) for NGA Testing Subsets

		NGA Models					Previous Models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Mainshocks (104 Events)											
Total set of mainshocks		49.8	49.9	49.2	46.5	-	-	-	-	-	-
Subdivision by site class	Soil	51.9	51.2	50.3	47.5	-	-	-	-	-	-
	Rock	44.2	45.0	46.4	43.0	44.9	-	-	-	-	-
Subdivision by distance	Small *R*	47.6	52.5	50.7	43.3	-	42.4	35.2	39.1	37.7	-
	Medium *R*	50.1	49.9	48.8	47.1	-	44.0	38.4	37.2	36.2	-
	Large *R*	50.2	48.5	50.1	45.4	-	-	-	-	-	-
Subdivision by magnitude	Small **M**	52.3	47.3	43.9	49.6	-	-	-	-	-	-
	Large **M**	48.5	51.1	51.6	45.1	-	-	-	-	-	-
Subdivision by measured/inferred V_s	Measured V_s	50.3	50.9	51.1	46.0	-	-	-	-	-	-
	Inferred V_s	49.4	49.5	48.4	46.8	-	-	-	-	-	-
Rock sites with small-to-medium *R*		45.1	45.5	46.7	43.7	45.2	32.7	31.5	31.6	27.2	23.8
Model rankings based on total *LH*		2	1	3	4	-	-	-	-	-	-
Aftershocks (33 Events)											
Total set of aftershocks		46.7	36.2	30.7	49.0	-	-	-	-	-	-
Subdivision by site class	Soil	48.2	35.8	31.6	49.4	-	-	-	-	-	-
	Rock	44.7	36.7	29.2	48.4	50.0	-	-	-	-	-
Model rankings based on total *LH*		2	3	4	1	-	-	-	-	-	-

distinction between the models that were developed using aftershocks (AS08, CY08, and I08) and those that did not (BA08 and CB08). The models influenced by aftershocks have *E* values near 60% and median *LH* values between 45% and 50%; the models not designed for aftershocks have *E* values in the 40% range and median *LH* values in the high 30% range at best. Of the two models that did not use aftershocks, the BA08 model performs slightly better than the CB08 model, most likely because BA08 includes a severe anelastic attenuation term, which affects large-distance predictions for small-magnitude events; CB08 likely overestimates these ground motions because it does not include such a term.

When the NGA mainshock records are subdivided by site class (soil versus rock), the values of *E* are generally similar for both categories, although the median *LH* values fall slightly for rock sites. The prediction accuracy of the I08 model is last among the NGA models when compared using *E* and in the middle when compared using median *LH*. The probable reasons for the low prediction accuracy of the I08 model will be explained in the Incorporation of Aftershocks in Model Development section. When the NGA mainshock records are subdivided by magnitude, there is an improvement for small-magnitude events (values of *E* in the 70% range instead of the 60% range). The AS08 and CY08 models perform best for small-magnitude events, and the BA08 and CB08 models perform best for large-magnitude events, as we explain in the Discussion section. When the NGA mainshock records are subdivided according to whether V_{S30} is measured or inferred, there is a clear increase in prediction accuracy when the site has a measured V_S profile (values of *E* in the 70% range instead of the 60% range). In all cases, there is a clear difference in prediction accuracy between the NGA models and the previous generation of models; the NGA models show a significant improvement in both *E* and *LH*. Note, however, that the NGA models have an unfair advantage over their predecessors because the newer models had been exposed to a greater portion of the testing data during model development; this advantage is eliminated in the blind comparison tests on recent earthquakes.

When the NGA mainshock records are subdivided by distance, the GMPEs perform best at intermediate distances, where most data are available. In addition to the lack of data, ground-motion predictions at small distances are obscured by phenomena such as directivity and hanging-wall effects, which are magnified near the earthquake source. The NGA models incorporate hanging-wall effects but do not currently incorporate directivity effects; Somerville *et al.* (1997), Spudich and Chiou (2008), and Rowshandel (2010) discuss methods of incorporating directivity in GMPEs. The estimation of ground motion at large distances raises other complications, such as Moho bounce effects (Atkinson and Boore, 2006). Because they found a deficiency in the NGA flatfile, Abrahamson and Silva (2008) and Chiou and Youngs (2008a) used nonstatistical information to constrain their large-distance terms; this may explain why AS08 has the lowest *E* and CY08 has the lowest median *LH* in the large-distance subset. In the blind comparisons, which do not penalize models for having nonstatistical components, the performances of the AS08 and CY08 models improve in this category.

Blind Comparison Tests on Recent Earthquakes

In Tables 6 and 7, we present the coefficients of efficiency and median *LH* values for the blind comparison tests

Table 6
Coefficients of Efficiency (%) for Blind Comparison Tests on Recent Earthquakes

		NGA Models					Previous Models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Mainshocks (6 Events)											
Total set of mainshocks		71.9	69.9	71.802	71.799	-	-	-	-	-	-
Subdivision by site class	Soil	71.8	69.3	71.7	71.3	-	-	-	-	-	-
	Rock	70.7	70.8	70.6	72.2	66.8	-	-	-	-	-
Subdivision by distance	Small R	56.3	53.2	51.6	37.1	-	43.7	47.9	33.5	46.4	-
	Medium R	70.9	68.3	70.3	71.1	-	67.6	62.7	60.6	56.7	-
	Large R	46.0	45.2	50.0	49.1	-	-	-	-	-	-
Rock sites with small-to-medium R		76.8	75.7	75.3	76.3	74.3	72.7	66.3	65.2	71.7	74.4
Model rankings based on total E		1	4	2	3	-	-	-	-	-	-
Aftershocks (1 Event)											
Total set of aftershock records		55.5	68.6	74.0	46.0	-	-	-	-	-	-
Subdivision by site class	Soil	52.6	65.2	71.3	44.2	-	-	-	-	-	-
	Rock	59.6	75.0	79.1	46.1	73.5	-	-	-	-	-
Model rankings based on total E		3	2	1	4						

on 1060 ground-motion records from seven recent earthquakes recorded in California (Ⓔ compiled in Table S1 in the electronic supplement to this paper). Because the vast majority of events are of small magnitude and the vast majority of sites have inferred V_S profiles, we do not subdivide this dataset into categories by magnitude or by measured versus inferred V_S. The patterns of the results for the recent mainshocks are similar to the corresponding patterns for the NGA flatfile mainshocks; however, the coefficients of efficiency are generally slightly higher and the median LH values are generally slightly lower for the recent events. Depending on the model and subset, when Tables 4 and 6 are compared, the coefficients of efficiency increase for the blind comparisons up to 10% and increase even greater for the NGA models at large distances and the previous models at small distances. However, comparing Tables 5 and 7, the LH values generally decrease for the blind comparisons. These results show that the models can generate excellent median predictions in blind situations (shown by the slight increases in E) but that the aleatory variability may be more difficult to pinpoint (shown by the slight decreases in LH). In the subdivision by site class in Table 6, the BA08 model suffers a small decrease in E from rock to soil (1.5%, which is a greater decrease than that of the other NGA models). Of the four NGA models that incorporate site response, BA08 is the only model that does not include a depth parameter and perhaps loses some prediction accuracy by utilizing V_{S30} as the only site characteristic. In addition, now that the NGA models no longer have an advantage over the previous models, there is less of a difference in performance between the new and previous models. With the exception of E for the I91 and I08 models, however, each NGA model still outperforms its corresponding predecessor. Table 8 displays the NGA model performances in the blind comparison tests separated by

Table 7
Median LH Values (%) for Blind Comparison Tests on Recent Earthquakes

		NGA Models					Previous Models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Mainshocks (6 Events)											
Total set of mainshocks		50.4	41.7	42.0	46.4	-	-	-	-	-	-
Subdivision by site class	Soil	52.0	42.9	43.7	47.0	-	-	-	-	-	-
	Rock	43.6	34.4	33.1	43.3	42.9	-	-	-	-	-
Subdivision by distance	Small R	38.8	42.7	35.4	33.3	-	38.3	33.1	20.6	38.7	-
	Medium R	52.6	42.2	42.7	48.3	-	49.3	34.0	35.8	39.5	-
	Large R	45.9	39.1	41.2	42.5	-	-	-	-	-	-
Rock sites with small-to-medium R		47.3	37.9	34.9	45.7	45.5	41.8	21.8	29.7	41.7	44.1
Model rankings based on total LH		1	4	3	2	-	-	-	-	-	-
Aftershocks (1 Event)											
Total set of aftershock records		40.6	44.4	48.1	36.4	-	-	-	-	-	-
Subdivision by site class	Soil	38.5	44.4	49.9	35.7	-	-	-	-	-	-
	Rock	43.6	44.5	44.1	37.9	51.4	-	-	-	-	-
Model rankings based on total LH		3	2	1	4						

Table 8
NGA Model Performances in Blind Comparison Tests
Separated by Earthquake

	Coefficient of Efficiency, E (%)				Median LH Value (%)			
Earthquake	AS08	BA08	CB08	CY08	AS08	BA08	CB08	CY08
Parkfield	82.0	77.0	77.5	76.9	47.6	42.7	39.3	43.9
San Simeon	47.7	18.0	12.6	47.2	39.5	25.3	22.4	38.6
Anza	77.9	74.4	78.3	70.7	50.1	35.5	38.8	39.2
Alum Rock	42.9	41.0	41.2	50.5	40.9	29.9	26.8	41.1
Chino Hills	73.7	71.0	75.3	70.3	58.3	48.3	50.2	52.4
Baja	60.0	66.5	66.2	73.2	41.8	46.0	45.1	45.8
Ocotillo	55.5	68.6	74.0	46.0	40.6	44.4	48.1	36.4

earthquake. Interestingly, when the analysis is separated by earthquake, each NGA model offers the best predictions for at least one earthquake, and the worst predictions for at least one other.

The 2004 **M** 6.0 Parkfield earthquake is of special interest because it generated an unprecedented quantity of near-source ground-motion records. Of the 94 records we tested, 56 records (60%) are located within 10 km of the rupture plane. As noted by Shakal *et al.* (2005, 2006), the ground motions near the fault were highly variable during the Parkfield earthquake. Figure 5 is a plot of PGA versus rupture distance for the Parkfield dataset, along with the median relationships using the ten GMPEs in this study (drawn for the average value of V_{S30} across all sites the dataset, 399.3 m/s, which approximately corresponds to dense soil or soft rock). The increased ground-motion variability at rupture distances of less than 10 km is clearly observed, as manifested by the lower values of E and median LH for the models in the small-distance dataset.

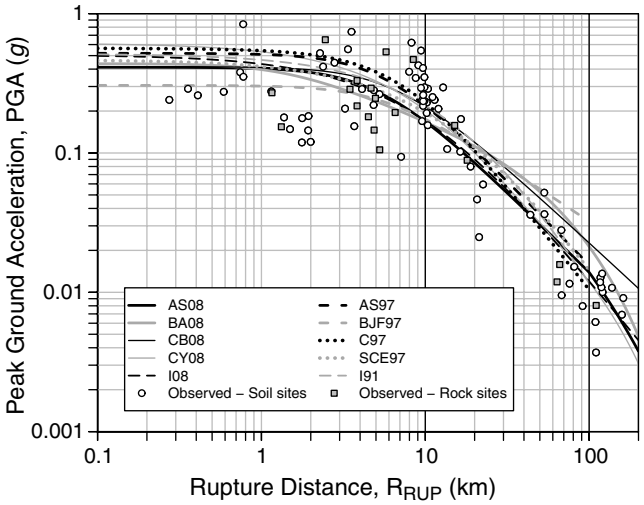

Figure 5. PGA versus R_{RUP} for the Parkfield dataset, along with the predicted ground motions using the average site characteristic among the 94 stations, $V_{S30} = 399.3$ m/s. The increase in ground-motion variability at distances less than 10 km is clearly observed. Points are separated into the categories of soil sites ($180 < V_{S30} < 450$ m/s) and rock sites ($450 \leq V_{S30} \leq 1300$ m/s), and the previous models are truncated at 100 km.

The results for the 14 June 2010 **M** 5.7 Ocotillo earthquake are particularly interesting. The event, considered to be an aftershock of the 4 April 2010 **M** 7.2 Baja earthquake, was recorded at many of the same stations as the Baja earthquake. The goodness-of-fit results for this aftershock are contrary to what we found for the NGA aftershocks: the BA08, CB08, and I08 models noticeably outperform the AS08 and CY08 models (which have an aftershock flag), both in terms of E and LH, as seen in Table 7. However, when the calculations are repeated assuming that the **M** 5.7 Ocotillo earthquake is a mainshock (i.e., by setting the aftershock flag in the AS08 and CY08 models to zero), the prediction accuracies of the AS08 and CY08 models reach a level similar to those of the other models, with E values in the 60% to 70% range and median LH values near 50%. These results suggest that the Ocotillo earthquake generated levels of ground motions more characteristic of a mainshock than an aftershock.

Discussion

Incorporation of Aftershocks in Model Development

As discussed earlier, each NGA modeling team made different decisions when selecting their regression datasets from the NGA flatfile, but one of the most significant decisions was whether to include aftershocks. The AS08, CY08, and I08 model teams included aftershocks in their regression datasets, most of which were aftershock records of the 1999 **M** 7.6 Chi-Chi, Taiwan, earthquake, which comprise 83% of the aftershock records in the flatfile. One potential problem with including such a high proportion of records from a single event is that the model may become overfit toward the characteristics of that event, and the model's ability to generalize to other situations is lowered. The AS08 and CY08 models outperform the BA08 and CB08 models on the small-magnitude NGA mainshock dataset but underperform on the large-magnitude NGA mainshock dataset. Because aftershocks tend to be small-magnitude events, the use of aftershock records in GMPE development may bias these models toward small-magnitude events.

The AS08, CY08, and I08 model teams included aftershocks in their regression datasets, but only the AS08 and

CY08 models include an aftershock dummy variable that reduces the ground-motion estimate when the earthquake is an aftershock. For a given magnitude, aftershocks tend to generate smaller ground motions than do mainshocks of the same magnitude, and there are differences in spectral scaling (Boore and Atkinson, 1989; Atkinson, 1993; Boore and Atkinson, 2008). Because aftershocks are associated with smaller ground motions than mainshocks for a given magnitude, the AS08 and CY08 teams utilize the aftershock dummy variable to alert the model to decrease its estimated ground motion when the GMPE is being used for an aftershock. The inclusion of aftershocks in the regression subset for the I08 model without an appropriate dummy variable effectively treats aftershocks as equivalent to mainshocks; thus, the ground-motion predictions for mainshocks are more prone to underprediction. This may explain why the I08 model ranks last among the NGA models in terms of E for all tests on mainshocks, as seen in Tables 4 and 6.

Because of the Poissonian assumption that events are independent, aftershocks are excluded from current probabilistic seismic hazard analyses (PSHAs); therefore, ground-motion predictions for mainshocks (especially large-magnitude mainshocks) have the most significant consequences. We do not intend to discount the importance of aftershocks, as aftershocks generate potentially devastating stresses and strains on already-fatigued systems. However, when aftershocks are included in the regression dataset for a model, there is a decrease in the prediction accuracy of large-magnitude mainshocks. Thus, like the results of Peruš and Fajfar (2010), our results suggest that aftershocks should either be used cautiously (or not used at all) in regression databases for GMPE development, especially in data-rich tectonically active regions. If model developers choose to include aftershocks in their regression datasets, we highly recommend that an aftershock dummy variable be included in the model. However, in areas of low-to-moderate seismicity or regions with limited ground-motion data, the elimination of aftershock records may not be so feasible; Douglas and Halldórsson (2010) suggest that aftershocks and mainshocks be treated equivalently in Europe, the Mediterranean, and the Middle East.

Uncertainty of Site Parameters

Of the model parameters, the greatest contribution to aleatory variability comes from V_{S30} (Abrahamson and Silva, 2008). One of the major problems of shear wave velocity data is that actual measurements are sparse. Only about 30% of the stations in the NGA database have measured values of V_{S30} (Power *et al.*, 2008); the remaining V_{S30} values are inferred using correlations of V_{S30} with surficial geology, such as those published by Wills and Clahan (2006). In Figure 6a, we investigate the accuracy of inferring V_{S30} from surficial geology. For the 258 California stations in the flatfile with measured V_S profiles, we utilize the correlations in Wills and Clahan (2006) to estimate V_{S30} from surficial geology.

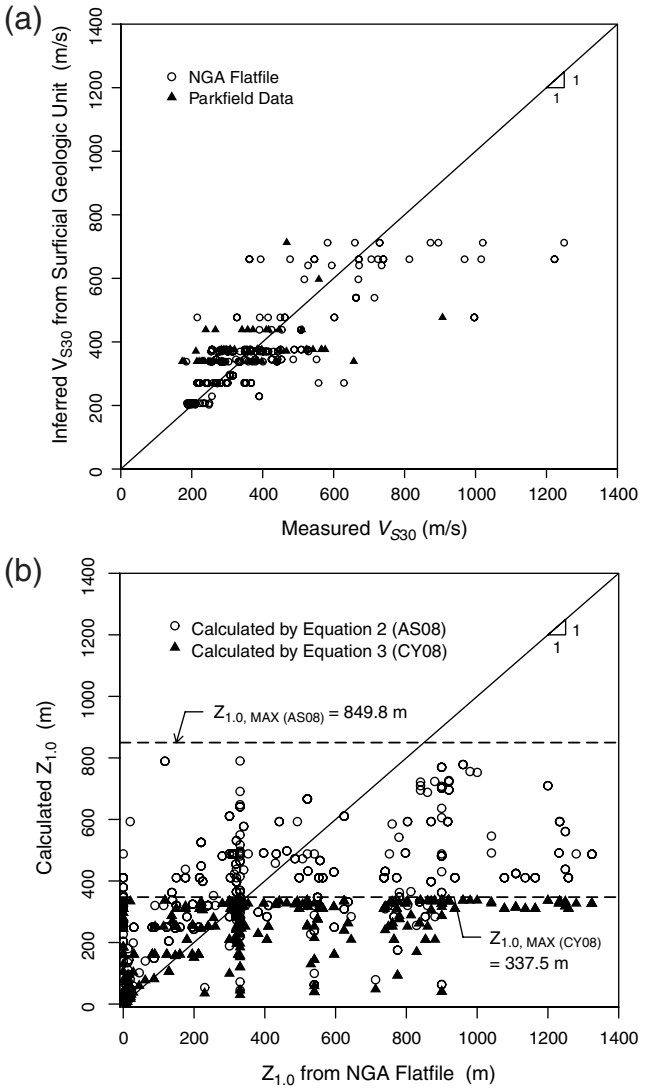

Figure 6. (a) Measured V_{S30} versus the corresponding values of V_{S30} inferred from surficial geology using Wills and Clahan (2006) for sites in California with measured V_S profiles; (b) comparison of $Z_{1.0}$ calculated from V_{S30} using the median equations versus the corresponding values of $Z_{1.0}$ from the flatfile, for equation 2 (AS08) and equation 3 (CY08).

In addition to the measured stations in the flatfile, we include 52 stations in the Parkfield, California, vicinity with measured V_S profiles provided by Thompson *et al.* (2010). We then plot the value of V_{S30} inferred from surficial geology (which would be used as an input parameter if there were no measurement at the site) against the known, measured value of V_{S30}. If the classification scheme based upon surficial geology were perfect, the plot would follow the 1-to-1 line; however, there is a substantial amount of scatter, and the discrete categories based on surficial geology are visible as horizontal groupings. The coefficient of efficiency for this V_{S30} estimation procedure is 55.6%. Although the NGA flatfile has estimated values of V_{S30} for almost every recording station, we suspect that many of the V_{S30} estimates based upon surficial geology are inaccurate.

Even more difficult to estimate than V_{S30} are the depth parameters, $Z_{1.0}$ and $Z_{2.5}$. In the Explanatory Variables section, we described the procedures for estimating the depth parameters when site-specific boreholes or regional velocity models are not available. In Figure 2, we presented a graph of the median relationships provided by the AS08 (equation 2) and CY08 (equation 3) models for estimating $Z_{1.0}$ from V_{S30}, along with data from the 448 sites in the NGA flatfile with specified values of $Z_{1.0}$. The considerable amount of scatter in Figure 2 is even more apparent when comparing plots of measured versus calculated $Z_{1.0}$, as seen in Figure 6b. For each of the 448 sites in the flatfile with specified values of $Z_{1.0}$, we calculate $Z_{1.0}$ from V_{S30} using equation (2) and equation (3) and then plot the calculated values of $Z_{1.0}$ versus the values of $Z_{1.0}$ specified in the NGA flatfile. The coefficients of efficiency are only 25.6% for the AS08 equation and −7.7% for the CY08 equation, indicating that the median equations suggested by Abrahamson and Silva (2008) and Chiou and Youngs (2008a) fare poorly when estimating $Z_{1.0}$ from the flatfile. Moreover, the $Z_{1.0}$ estimates are numerically bounded at maximums of 849.8 m for equation (2) and 337.5 m for equation (3). Abrahamson and Silva (2008) developed equation (2) from analytical site response models, and Chiou and Youngs (2008b) utilized an updated velocity model for southern California when they developed equation (3), which has smaller depth parameters than the previous velocity model reflected in the flatfile. Therefore, it is not surprising that there is some disagreement between equation (2) and equation (3) and the values in the flatfile, but the discrepancies demonstrate that the depth parameters and their methods of estimation are fraught with uncertainty.

When the depth parameters are estimated at unsampled locations, information about the ground surface (i.e., surficial geology) is used to estimate a parameter that involves 30 m of depth (V_{S30}), which in turn is used to estimate a parameter that typically involves depths much greater than 30 m ($Z_{1.0}$ or $Z_{2.5}$). The entire decision process is based upon weak correlations of average shear wave velocity with surficial geology. Scott *et al.* (2006) find that shear wave velocity correlates poorly with geologic units. Surficial geologic maps provide no information about how the geologic units change with depth and therefore are questionable for estimating V_{S30}. It is widely agreed that site characteristics should be incorporated into GMPEs, and Campbell (1989) finds that the inclusion of depth parameters in GMPEs can improve their predictive capabilities. As shown by the superior model performances for sites with measured V_S (as opposed to sites with inferred V_S) and as noted by Campbell and Bozorgnia (2008) and Chiou and Youngs (2008a), a greater emphasis on site-specific data collection would improve the prediction accuracy of GMPEs.

For developing more reliable ground-motion predictions, another issue (in addition to the lack of data) is determining which site parameters are the best predictors of amplification. When Boore *et al.* (1997) first introduced V_{S30} as an explanatory variable in a GMPE, they emphasized that the ideal averaging depth corresponds to one-quarter wavelength for the period of interest (Joyner and Fumal, 1984) and that 30 m is a choice limited by the lack of data at greater depths. Douglas *et al.* (2009) develop a new framework for incorporating site characteristics in GMPEs by utilizing the one-quarter wavelength velocity as the primary site characteristic. Rather than perpetuating the use of V_{S30} as the predominant site parameter, additional research into more effective site parameters may be beneficial for improvement of ground-motion predictions.

Improving Prediction Accuracy

Strasser *et al.* (2009) claim that reductions in uncertainty are not necessarily brought by increasing the number of explanatory variables in the models or by increasing the quantity of ground-motion records in the regression datasets, and Kuehn *et al.* (2009) warn about the dangers of overfitting using the current approaches to model development. Table 9 displays the NGA model performances on a superset of the NGA mainshocks and the recent earthquakes tested in this study. Based on E, the statistic more closely linked to the median predictions, the GMPE with the highest accuracy is CB08, which is one of the simpler GMPEs. The solid performance of the simpler NGA models in many of the subsets in this study lends credence to the suggestion that more complicated models do not necessarily offer more accurate predictions. The more complicated models with greater numbers of input parameters require more assumptions on variables that are difficult to predict before an earthquake actually happens. Site characteristics, however, can be measured prior to an earthquake, and a greater emphasis on site-specific V_S measurements may increase prediction accuracy, as supported by the superior performance of all models for sites with measured V_S over those with inferred V_S. A higher-quality regression dataset (but not necessarily higher-quantity) with greater measurements of site characteristics, coupled with simple functional forms in the GMPEs, may yield the best solution.

Conclusions

Using the Nash–Sutcliffe model efficiency coefficient (E) and the LH value as the primary goodness-of-fit statistics, we have compared the prediction accuracy of the five ground-motion prediction equations released as part of the

Table 9
NGA Model Performances for Superset of All NGA Mainshocks and Recent Earthquakes

	Model			
	AS08	BA08	CB08	CY08
Coefficient of efficiency, E (%)	75.2	74.9	76.7	74.0
Model rankings based on E	2	3	1	4
Median LH value (%)	49.7	46.7	46.6	46.1
Model rankings based on LH	1	2	3	4

Next Generation Attenuation of Ground Motions project. The focus of this paper is on ground-motion prediction equations for peak ground acceleration, peak ground velocity, and spectral acceleration because these are the ground-motion parameters addressed by the NGA project. Although we do not specifically analyze relationships for predicting other earthquake-related phenomena, we have implemented a statistical validation methodology that could be used as a framework for comparing alternative predictive relationships for other ground-motion parameters. First, we tested the NGA models and their predecessors on subsets of the database upon which they were developed. Then, we implemented a blind comparison test using 1060 ground-motion records from seven recent California earthquakes, which were not present in any of the databases used to develop the models. The newer models perform noticeably better than their previous counterparts; the five previous GMPEs should be considered obsolete.

We find that the decisions made by model developers when selecting their regression datasets, input parameters, and functional forms greatly influence the models' predictive capabilities. Including large numbers of records from single earthquakes (such as the high number of records from the Chi-Chi mainshock and aftershocks) can result in overfitting the data to those particular scenarios, thus reducing the models' ability to generalize to other situations. In addition, including aftershocks in the regression dataset may lead to unconservative ground-motion predictions if they are not properly handled by using an aftershock dummy variable. In general, however, aftershocks do not appear to help with model prediction accuracy and may bias models toward small-magnitude events. High model complexity, whether through large numbers of explanatory variables or convoluted functional forms, can also lead to overfitting. The results of this study show that the more complicated NGA models do not have a predictive advantage over the simpler NGA models. An increased emphasis on the measurement of site parameters would lead to a higher-quality regression dataset and better ground-motion predictions, as supported by the superior performances of the models at measured sites. The development of better site characteristics than V_{S30} may also improve the prediction accuracy of GMPEs.

Based on the results of this study, we do not wish to suggest that specific NGA models should be utilized more than others in general practice, as each model has positive and negative characteristics. In PSHA, it is advantageous to diversify the calculations and consider the estimates from multiple models when addressing epistemic uncertainty and selecting weights for logic trees. As we have seen in this study, when the analysis is broken into subsets and subdivisions, there are certain situations in which each NGA model performs superiorly. Rather than promoting the ubiquitous use of one model over another, we hope that the results of this study may be considered from a model development point of view. Especially important for prediction accuracy are decisions relating to model complexity, selection of input parameters, inclusion of aftershocks in the regression database, and allowing the regression database to be monopolized by single events. Nonetheless, the results of this study show that the NGA models are a significant improvement over their predecessors and have significant promise for ground-motion predictions in the future.

Data and Resources

To implement the GMPEs in this study, we utilized the open-source statistical language and environment R (R Development Core Team, 2010), and the *nga* R package (Kaklamanos and Thompson, 2010). Information about the NGA project, including a spreadsheet of the NGA flatfile and numerical implementations of the GMPEs in FORTRAN and Microsoft Excel, is publicly available on the Pacific Earthquake Engineering Research Next Generation Attenuation (NGA) project web site at http://peer.berkeley.edu/ngawest/index.html (last accessed June 2010). The classification of events in the NGA flatfile as either mainshocks or aftershocks was determined from information in Abrahamson and Silva (2008) and Boore and Atkinson (2007). Boore and Atkinson (2007) provide a useful record-by-record summary of reasons for excluding records from the NGA flatfile in their regression dataset.

For the blind comparison tests, source characteristics such as depth to top of rupture, down-dip rupture width, and fault dip were determined by selecting a finite fault model for the Parkfield earthquake (Dreger, 2004), San Simeon earthquake (Rolandone *et al.*, 2004), and Baja earthquake (Crowell and Bock, 2010). Finite fault models were not available for the smaller events, but the rake angle and fault dip were estimated by selecting the likely nodal plane from moment tensor solutions. The focal mechanisms were estimated for the Anza event by the Southern California Seismic Network at http://www.cisn.org/special/evt.05.06.12/14151344_email.txt, for Alum Rock by the Berkeley Seismological Laboratory at http://www.cisn.org/special/evt.07.10.30/AlumRock.htm, and for Chino Hills and Ocotillo by the Southern California Earthquake Center at http://www.data.scec.org/mtarchive/solution.jsp?eventid=14383980 and http://www.data.scec.org/MomentTensor/solutions/web_14745580/ci14745580_MT.html, respectively (all Web sites last accessed June 2010). The required distance measures were calculated from the source-to-site geometry for each location; a seismogenic depth of 3 km was assumed when calculating R_{SEIS} for the C97 model. When finite fault models were not available, R_{EPI} was used as a proxy for R_{JB}, and R_{HYP} was used as a proxy for R_{RUP}.

The site characteristics were obtained from a variety of sources, depending on the availability of information. First, some of the stations that recorded the recent earthquakes also recorded other earthquakes in the NGA flatfile; site characteristics for these stations were determined directly from the NGA flatfile. In addition, we utilized V_S profiles at 54 ground-motion recording stations in the Parkfield area that

were collected by Thompson *et al.* (2010) using the spectral analysis of surface waves test (Stokoe *et al.*, 1994), and two V_S profiles for the Turkey Flat Strong-Motion Array in Parkfield from Real (1988). For locations without measured V_S profiles, site characteristics were inferred from surficial geology using the classification scheme of Wills and Clahan (2006; Table 1), which the NGA research team also used to estimate V_{S30} for many records in the flatfile (Chiou *et al.*, 2008; Table 2). Geologic data were obtained from 1:250,000 scale regional maps (Jennings, 1958, 1959, 1962, 1967; Jennings and Strand, 1958, 1969; Matthews and Burnett, 1965; Rogers, 1965, 1967; Smith, 1964; Strand, 1962; Wagner *et al.*, 1991; available from the California Geological Survey at http://www.conservation.ca.gov/cgs/rghm/rgm/250k_index/Pages/250k_index.aspx; last accessed June 2010), and from Shakal *et al.* (2005), who used these maps for the Parkfield vicinity.

In addition to the aforementioned resources, location and site information were also obtained from the web sites of the seismic recording stations' networks: (1) the University of California Berkeley Digital Seismic Network (code: BK) at http://www.ncedc.org/bdsn/, (2) the California Geological Survey Strong Motion Instrumentation Program (CE) at http://www.conservation.ca.gov/cgs/smip/Pages/Index.aspx, (3) the Southern California Seismic Network (CI) at http://www.scsn.org/index.html (operated by the California Institute of Technology and U.S. Geological Survey [USGS]), (4) the Northern California Seismic Network (NC) at http://www.ncedc.org/ncsn/ (operated by USGS), (5) the National Strong-Motion Project (NP) at http://nsmp.wr.usgs.gov/ (operated by USGS), and (6) the California Department of Water Resources (WR) at http://wwwdwr.water.ca.gov/ (all were last accessed June 2010). Ground-motion records for the recent earthquakes were obtained from the following online databases: (1) the Center for Engineering Strong Motion Data at http://strongmotioncenter.org/, (2) the National Strong-Motion Project at http://nsmp.wr.usgs.gov/, and (3) the Consortium of Organizations for Strong-Motion Observation Systems at http://db.cosmos-eq.org (all were last accessed June 2010). Ⓔ A mini-flatfile for these recent earthquakes is presented as Table S1 in the electronic supplement to this paper. All other data used in this article came from the published sources listed in the references.

Acknowledgments

First, we would like to thank John Douglas, an anonymous reviewer, and associate editor Gail Atkinson for their incredibly detailed and insightful reviews and comments, which greatly improved the final manuscript. Additionally, we thank Eric Thompson and Richard Vogel for their useful advice that improved the quality and direction of this research. We would also like to acknowledge the researchers involved in the NGA project and their work that went into the creation of the NGA flatfile, the NGA GMPEs, and the supporting NGA projects and documentation. We commend them for their efforts in this area.

References

Abrahamson, N. A., and K. M. Shedlock (1997). Overview, *Seismol. Res. Lett.* **68,** 9–23.

Abrahamson, N. A., and W. J. Silva (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismol. Res. Lett.* **68,** 94–127.

Abrahamson, N. A., and W. J. Silva (2008). Summary of the Abrahamson and Silva NGA ground-motion relations, *Earthq. Spectra* **24,** 67–97.

Abrahamson, N. A., G. M. Atkinson, D. M. Boore, Y. Bozorgnia, K. W. Campbell, B. S.-J. Chiou, I. M. Idriss, W. J. Silva, and R. R. Youngs (2008). Comparisons of the NGA ground-motion relations, *Earthq. Spectra* **24,** 45–66.

Atkinson, G. M. (1993). Earthquake source spectra in eastern North America, *Bull. Seismol. Soc. Am.* **83,** 1778–1798.

Atkinson, G. M., and D. M. Boore (2006). Earthquake ground-motion prediction equations for eastern North America, *Bull. Seismol. Soc. Am.* **96,** 2181–2205.

Atkinson, G. M., and M. Morrison (2009). Observations on regional variability in ground-motion amplitudes for small-to-moderate earthquakes in North America, *Bull Seismol. Soc. Am.* **99,** 2393–2409.

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

Boore, D. M. (2010). TSPP—A collection of FORTRAN programs for processing and manipulating time series, version 2.1, *U.S. Geol. Surv. Open-File Rept. 2008-1111*, 56 pp.

Boore, D. M., and G. M. Atkinson (1989). Spectral scaling of the 1985 to 1988 Nahanni, Northwest Territories, earthquakes, *Bull. Seismol. Soc. Am.* **79,** 1736–1761.

Boore, D. M., and G. M. Atkinson (2007). Boore-Atkinson NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters, *PEER Report No. 2007/01*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, 242 pp.

Boore, D. M., and G. M. Atkinson (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, *Earthq. Spectra* **24,** 99–138.

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, *Seismol. Res. Lett.* **68,** 128–153.

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

Campbell, K. W. (1989). Empirical prediction of near-source ground motion for the Diablo Canyon power plant site, San Luis Obispo County, California, *U.S. Geol. Surv. Open-File Rept. 89–484*, 115 pp.

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, *Seismol. Res. Lett.* **68,** 154–179.

Campbell, K. W., and Y. Bozorgnia (2003). Updated near-source ground motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra, *Bull. Seismol. Soc. Am.* **93,** 314–331.

Campbell, K. W., and Y. Bozorgnia (2006). Next Generation Attenuation (NGA) empirical ground motion models: Can they be used in Europe?, *Proc., First European Conf. on Earthquake Engineering and Seismology*, Paper 458, Geneva, Switzerland, 3–8 September 2006.

Campbell, K. W., and Y. Bozorgnia (2007). Campbell-Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters, *PEER Report No. 2007/02*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, 246 pp.

Campbell, K. W., and Y. Bozorgnia (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s, *Earthq. Spectra* **24,** 139–171.

Chiou, B. S.-J., and R. R. Youngs (2008a). An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra* **24,** 173–215.

Chiou, B. S.-J., and R. R. Youngs (2008b). NGA model for the average horizontal component of peak ground motion and response spectra, *PEER Report No. 2008/09*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, 293 pp.

Chiou, B. S.-J., R. Darragh, N. Gregor, and W. Silva (2008). NGA project strong-motion database, *Earthq. Spectra* **24,** 23–44.

Chiou, B. S.-J., R. R. Youngs, N. A. Abrahamson, and K. Addo (2010). Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in California and its implications on regionalization of ground-motion prediction models, *Earthq. Spectra* **26,** 907–926.

Crowell, B., and Y. Bock (2010). Fault slip model for the Mayor-Cucapah earthquake, Scripps Institution of Oceanography, University of California, San Diego, http://geoapp03.ucsd.edu/gridsphere/gridsphere?cid = Sierra + El + Mayor + Earthquake.

Douglas, J. (2003). Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates, *Earth Sci. Rev.* **61,** 43–104.

Douglas, J., and B. Halldórsson (2010). On the use of aftershocks when deriving ground-motion prediction equations, in *Proc., 9th U. S. National and 10th Canadian Conf. on Earthquake Engineering*, Toronto, Canada, 25–29 July 2010.

Douglas, J., P. Gehl, L. F. Bonilla, O. Scotti, J. Régnier, A.-M. Duval, and E. Bertrand (2009). Making the most of available site information for empirical ground-motion prediction, *Bull. Seismol. Soc. Am.* **99,** 1502–1520.

Dreger, D. (2004). 09/28/2004 Preliminary slip model, Berkeley Seismological Laboratory, University of California, Berkeley, available at, *Finite-Source Rupture Model Database*, http://www.seismo.ethz.ch/srcmod/ (last accessed June 2010).

Ghasemi, H., M. Zare, and Y. Fukushima (2008). Ranking of several ground-motion models for seismic hazard analysis in Iran, *J. Geophys. Eng.* **5,** 301–310.

Graves, R. W., B. T. Aagaard, K. W. Hudnut, L. M. Star, J. P. Stewart, and T. H. Jordan (2008). Broadband simulations for M_w 7.8 southern San Andreas earthquakes: Ground motion sensitivity to rupture speed, *Geophys. Res. Lett.* **35,** L22302, doi 10.1029/2008GL035750.

Idriss, I. M. (1991). Selection of earthquake ground motions at rock sites, report prepared for the Structures Division, Building and Fire Research Laboratory, National Institute of Standards and Technology, Dept. of Civil Engineering, University of California, Davis, 35 pp.

Idriss, I. M. (2002). Attenuation relationship derived by I. M. Idriss in 2002, in *Empirical Model for Estimating the Average Horizontal Values of Pseudo-Absolute Spectral Accelerations Generated by Crustal Earthquakes (2007)*, interim report issued for USGS review, Pacific Earthquake Engineering Research Center, University of California, Berkeley, available at http://peer.berkeley.edu/ngawest/nga_models.html (last accessed June 2010).

Idriss, I. M. (2008). An NGA empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes, *Earthq. Spectra* **24,** 217–242.

Jennings, C. W. (1958). Geologic map of California: San Luis Obispo sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Jennings, C. W. (1959). Geologic map of California: Santa Maria sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Jennings, C. W. (1962). Geologic map of California: Long Beach sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Jennings, C. W. (1967). Geologic map of California: Salton Sea sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Jennings, C. W., and R. G. Strand (1958). Geologic map of California: Santa Cruz sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Jennings, C. W., and R. G. Strand (1969). Geologic map of California: Los Angeles sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Joyner, W. B., and T. E. Fumal (1984). Use of measured shear-wave velocity for predicting geologic site effects on strong ground motion, *Proc. of the 8th World Conf. on Earthquake Engineering* (San Francisco, California) **2,** 777–783.

Kaklamanos, J., and E. M. Thompson (2010). *nga*: NGA earthquake ground motion prediction equations, R package version 1.0, R Foundation for Statistical Computing, Vienna, Austria, http://cran.r-project.org/ (last accessed May 2010).

Kaklamanos, J., D. M. Boore, E. M. Thompson, and K. W. Campbell (2010). Implementation of the Next Generation Attenuation (NGA) ground-motion prediction equations in Fortran and R, *U.S. Geol. Surv. Open-File Rept. 2010–1296*, 47 pp.

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

Kuehn, N. M., F. Scherbaum, and C. Riggelsen (2009). Deriving empirical ground-motion models: Balancing data constraints and physical assumptions to optimize prediction capability, *Bull. Seismol. Soc. Am.* **99,** 2335–2347.

Legates, D. R., and G. J. McCabe Jr. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, *Water Resour. Res.* **35,** 233–241.

Matthews, R. A., and J. L. Burnett (1965). Geologic map of California: Fresno sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Nash, J. E., and J. V. Sutcliffe (1970). River flow forecasting through conceptual models: Part I, a discussion of principles, *J. Hydrol.* **10,** 282–290.

Olsen, K. B., S. M. Day, L. A. Dalguer, J. Mayhew, Y. Cui, J. Zhu, V. M. Cruz-Atienza, D. Roten, P. Maechling, T. H. Jordan, D. Okaya, and A. Chourasia (2009). ShakeOut-D: Ground motion estimates using an ensemble of large earthquakes on the southern San Andreas fault with spontaneous rupture propagation, *Geophys. Res. Lett.* **36,** L04303, doi 10.1029/2008GL036832.

Olsen, K. B., S. M. Day, J. B. Minster, Y. Cui, A. Chourasia, D. Okaya, P. Maechling, and T. Jordan (2008). TeraShake2: Spontaneous rupture simulations of M_w 7.7 earthquakes on the southern San Andreas fault, *Bull. Seismol. Soc. Am.* **98,** 1162–1185.

Petersen, M. D., A. D. Frankel, S. C. Harmsen, C. S. Mueller, K. M. Haller, R. L. Wheeler, R. L. Wesson, Y. Zeng, O. S. Boyd, D. M. Perkins, N. Luco, E. H. Field, C. J. Wills, and K. S. Rukstales (2008). Documentation for the 2008 update of the United States national seismic hazard maps, *U.S. Geol. Surv. Open-File Rept. 2008-1128*, 61 pp.

Peruš, I., and P. Fajfar (2010). Ground-motion prediction by a non-parametric approach, *Earthq. Eng. Struct. Dynam.* **39,** 1395–1416.

Power, M., B. Chiou, N. Abrahamson, Y. Bozorgnia, T. Shantz, and C. Roblee (2008). An overview of the NGA project, *Earthq. Spectra* **24,** 3–21.

R Development Core Team (2010). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, http://www.R-project.org (last accessed June 2010).

Real, C. R. (1988). Turkey Flat, USA, site effects test area, site characterization, *Technical Report No. 88-2*, Earthquake Shaking Assessment Unit, California Division of Mines and Geology, 252 pp.

Rogers, T. H. (1965). Geologic map of California, Santa Ana sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Rogers, T. H. (1967). Geologic map of California: San Bernardino sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Rolandone, F., R. Burgmann, D. Dreger, and M. Murray (2004). Coseismic slip distribution of the 22 December 2003 San Simeon earthquake, *Annual Report 2003–2004*, Berkeley Seismological Laboratory, University of California, Berkeley, http://seismo.berkeley.edu/annual_report/ar03_04/node19.html (last accessed June 2010).

Rowshandel, B. (2010). Directivity correction for the Next Generation Attenuation (NGA) relations, *Earthq. Spectra* **26,** 525–559.

Sadigh, K., C. Y. Chang, J. A. Egan, F. Makdisi, and R. R. Youngs (1997). Attenuation relationships for shallow crustal earthquakes based on California strong motion data, *Seismol. Res. Lett.* **68,** 180–189.

Scasserra, G., J. P. Stewart, P. Bazzurro, G. Lanzo, and F. Mollaioli (2009). A comparison of NGA ground motion prediction equations to Italian data, *Bull. Seismol. Soc. Am.* **99,** 2961–2978.

Scherbaum, F., F. Cotton, and P. Smit (2004). On the use of response spectral-reference data for the selection and ranking of ground-motion models for seismic-hazard analysis in regions of moderate seismicity: The case of rock motion, *Bull. Seismol. Soc. Am.* **94,** 2164–2185.

Scott, J. B., M. Clark, C. Lopez, A. Pancha, T. Rasmussen, S. B. Smith, W. Thelen, and J. N. Louie (2006). Three urban transects of shallow shear-velocity using the refraction microtremor method, *Proc. of the Managing Risk in Earthquake Country 100th Anniversary Earthquake Conf. Commemorating the 1906 San Francisco Earthquake*, San Francisco, 18–22 April 2006.

Shakal, A., V. Graizer, M. Huang, R. Borcherdt, H. Haddadi, K. Lin, C. Stephens, and P. Roffers (2005). Preliminary analysis of strong-motion recordings from the 28 September 2004 Parkfield, California earthquake, *Seismol. Res. Lett.* **76,** 27–39.

Shakal, A., H. Haddadi, V. Graizer, K. Lin, and M. Huang (2006). Some key features of the strong-motion data from the **M** 6.0 Parkfield, California, earthquake of 28 September 2004, *Bull. Seismol. Soc. Am.* **96,** S90–S118.

Shoja-Taheri, J., S. Naserieh, and H. Ghofrani (2010). A test of the applicability of NGA models to the strong ground motion data in the Iranian Plateau, *J. Earthq. Eng.* **14,** 278–292.

Smith, A. R. (1964). Geologic map of California: Bakersfield sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Somerville, P. G., N. F. Smith, R. W. Graves, and N. A. Abrahamson (1997). Modification of empirical strong ground motion attenuation relations to include amplitude and duration effects of rupture directivity, *Seismol. Res. Lett.* **68,** 199–222.

Spudich, P., and B. S.-J. Chiou (2008). Directivity in NGA earthquake ground motions: Analysis using isochrone theory, *Earthq. Spectra* **24,** 279–298.

Stafford, P. J., F. O. Strasser, and J. J. Bommer (2008). An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region, *Bull. Earthq. Eng.* **6,** 149–177.

Star, L. M., J. P. Stewart, and R. W. Graves (2010). Investigation of basin and directivity effects in broadband simulated ground motions, in *Proc., 9th U. S. National and 10th Canadian Conf. on Earthquake Engineering*, Toronto, Canada, 25–29 July 2010.

Star, L. M., J. P. Stewart, and R. W. Graves (2011). Comparison of ground motions from hybrid simulations to NGA prediction equations, *Earthq. Spectra* **27,** no. 2, in press.

Star, L. M., J. P. Stewart, R. W. Graves, and K. W. Hudnut (2008). Validation against NGA empirical model of simulated motions for **M** 7.8 rupture of San Andreas fault, in *Proc., 14th World Conf. on Earthquake Engineering*, Beijing, China, 12–17 October 2008.

Stewart, J. P., L. M. Star, and R. W. Graves (2008). Validation against NGA empirical model of simulated motions for **M** 7.15 rupture of Puente Hills fault, *Final Report*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, http://peer.berkeley.edu/tbi/pdf-ppt/PuenteHills_Synthetic_Ground_Motion_rpt3.pdf (last accessed June 2010).

Stokoe, K. H., II, S. G. Wright, J. A. Bay, and J. M. Roesset (1994). Characterization of geotechnical sites by SASW method, in *Geophysical Characterization of Sites*, ISSMFE Technical Committee #10, R. D. Woods (Editor), Oxford & IBH Publishing Co., New Delhi, India, 15–25.

Strand, R. G. (1962). Geologic map of California: San Diego—El Centro sheet, scale 1:250,000, *Calif. Div. Mines Geol.*

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: issues, insights, and challenges, *Seismol. Res. Lett.* **80,** 40–56.

Thompson, E. M., R. E. Kayen, B. Carkin, and H. Tanaka (2010). Surface wave site characterization at 52 strong motion recording stations affected by the Parkfield **M** 6.0 earthquake of 28 September 2004, *U.S. Geol. Surv. Open-File Rept. 2010-1168*, 117 pp.

Wagner, D. L., E. J. Bortugno, and R. D. McJunkin (1991). Geologic map of the San Francisco—San Jose Quadrangle, Regional Geologic Map 5A, scale 1:250,000, *Calif. Div. Mines Geol.*

Wills, C. J., and K. B. Clahan (2006). Developing a map of geologically defined site-condition categories for California, *Bull. Seismol. Soc. Am.* **96,** 1483–1501.

Department of Civil and Environmental Engineering
113 Anderson Hall
Tufts University
Medford, Massachusetts 02155
james.kaklamanos@tufts.edu
laurie.baise@tufts.edu

Manuscript received 9 February 2010