29 May 2010, 8:00 am - 9:30 am

Model Validation of Recent Ground Motion Prediction Relations for Shallow Crustal Earthquakes in Active Tectonic Regions

James Kaklamanos
Tufts University, Medford, MA

Laurie G. Baise
Tufts University, Medford, MA

Fifth International Conference on
Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics
and Symposium in Honor of Professor I.M. Idriss
May 24-29, 2010 • San Diego, California

MODEL VALIDATION OF RECENT GROUND MOTION PREDICTION RELATIONS FOR SHALLOW CRUSTAL EARTHQUAKES IN ACTIVE TECTONIC REGIONS

James Kaklamanos
Tufts University
Medford, Massachusetts-USA 02155

Laurie G. Baise
Tufts University
Medford, Massachusetts-USA 02155

ABSTRACT

Recent earthquake ground motion prediction relations, such as those developed from the Next Generation Attenuation of Ground Motions (NGA) project in 2008, have established a new baseline for the estimation of ground motion parameters such as peak ground acceleration (PGA), peak ground velocity (PGV), and spectral acceleration (*Sa*). When these models were published, very little was written about model validation or prediction accuracy. We perform statistical goodness-of-fit analyses to quantitatively compare the predictive abilities of these recent models. The prediction accuracy of the models is compared using several testing subsets of the master database used to develop the NGA models. In addition, we perform a blind comparison of the new models with previous simpler models, using ground motion records from the two most recent earthquakes of magnitude 6.0 or greater to strike mainland California: (1) the 2004 **M** 6.0 Parkfield earthquake, and (2) the 2003 **M** 6.5 San Simeon earthquake. By comparing the predictor variables and performance of different models, we discuss the sources of uncertainty in the estimates of ground motion parameters and offer recommendations for model development. This paper presents a model validation framework for assessing the prediction accuracy of ground motion prediction relations and aiding in their future development.

INTRODUCTION

The purpose of ground motion prediction equations (GMPEs; 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. The typical response variables in ground motion prediction relations are peak ground acceleration (PGA), peak ground velocity (PGV), and 5%-damped elastic pseudo-response spectral acceleration (*Sa*).

A typical ground motion prediction relation has the form

$$\ln \hat{Y} = f(\mathbf{M}, R, \sum Source_i, \sum Site_i), \qquad (1)$$

where $\ln \hat{Y}$ is the natural logarithm of the ground motion parameter of interest, **M** is the moment 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 (often the closest distance to the zone of rupture, R_{RUP}), $\sum Source_i$ are other variables relating to the earthquake source (such as type of faulting, rupture width and depth, and fault dip), and $\sum Site_i$ are variables relating to the site of interest (such as average shear wave velocity, geologic characteristics, or depth to bedrock) (Kramer, 1996; Abrahamson *et al*., 2008). Source parameters are generally constant for a given earthquake and do not vary from location to location; site parameters are generally constant for a given location and do not vary from earthquake to earthquake.

Peak values of ground motion parameters are assumed to follow a lognormal distribution; therefore, the logarithms of the ground motion parameters follow a normal distribution. To maintain normally-distributed residuals, regression is typically performed on the logarithm of the ground motion parameter of interest (Kramer, 1996). Ground motion prediction relations are developed for specific tectonic environments using multivariate regression on ground motion databases, and the relationships are updated as more earthquake data are obtained. 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 uncertainty of the ground motion prediction estimates.

After a five-year effort, the Next Generation Attenuation of Ground Motions (NGA) project was completed in 2008. The project, sponsored by the Pacific Earthquake Engineering Research Center (PEER), established five new GMPEs that

Table 1. GMPEs tested in this study

NGA models			Previous models		
Team	Year	Abbrev.	Team	Year	Abbrev.
Abrahamson and Silva	2008	AS08	Abrahamson and Silva	1997	AS97
Boore and Atkinson	2008	BA08	Boore, Joyner, and Fumal	1997	BJF97
Campbell and Bozorgnia	2008	CB08	Campbell	1997	C97*
Chiou and Youngs	2008	CY08	Sadigh, Chang, Egan, Makdisi, and Youngs	1997	SCE97
Idriss	2008	I08	Idriss	1991	I91*

NOTE:
* In the early 2000s, Campbell and Bozorgnia released an update to their 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 set of new and old models, we use all the GMPEs from the 1990s as the baseline for comparison.

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 (Abrahamson and Silva, 1997; Abrahamson and Silva, 2008; Boore *et al*., 1997; Boore and Atkinson, 2008; Campbell, 1997; Campbell and Bozorgnia, 2003; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008a; Idriss, 1991; Idriss, 2002; Idriss, 2008; Sadigh *et al.*, 1997).

The NGA models will have serious consequences, as they are beginning to serve as the basis for seismic hazard assessment in many applicable regions. In the most recent update to the national seismic hazard maps released by the United States Geological Survey (USGS), the NGA models are included in the hazard calculations (Petersen *et al.*, 2008). However, there have not been many published quantitative comparisons of the models. Ghasemi *et al.* (2008) compare and rank several GMPEs for seismic hazard analysis in Iran. Stafford *et al.* (2008) compare the NGA models with European models for seismic hazard analysis in the Euro-Mediterranean region. Star *et al.* (2008) and Stewart *et al*. (2008) compare the NGA models for simulated ground motions for various scenarios in southern California.

In this paper, we attempt to make a contribution to the geotechnical earthquake engineering literature by objectively comparing the models in a statistical validation framework. From the NGA flatfile, we develop testing databases of ground motion records that meet the requirements of the models. By calculating objective goodness-of-fit statistics, we compare the model predictions to the actual ground motion records to assess the predictive capabilities of the models. We test the models under different conditions, for soil and rock sites at various distances for both mainshocks and aftershocks. Based on the results, we discuss which GMPEs offer the most successful predictions in various situations, and the manner in which model development decisions influence model performance.

In addition, to assess the level of improvement that has occurred over time, we compare the new generation of models with the previous generation of models. We perform blind comparison tests by implementing the new and old models on recent earthquakes that were not present in any of the regression databases used for model development. In our comparisons, we utilize ground motion records from the two most recent earthquakes of magnitude 6.0 or greater to strike mainland California: (1) the 2004 **M** 6.0 Parkfield earthquake, and (2) the 2003 **M** 6.5 San Simeon earthquake. The results of these two tests are examples of how these GMPEs may perform when predicting ground motion for future earthquakes.

DATA

The NGA database

For use in developing the NGA models, researchers compiled an extensive data set of 3551 ground motion records from 173 shallow crustal earthquakes (Chiou *et al.*, 2008). A spreadsheet of the entire database, called the "NGA flatfile," is publicly available on the PEER NGA project web site (Pacific Earthquake Engineering Research Center, 2008). The researchers utilized subsets of this database in their regressions for model development. The research teams generally excluded records that were not representative of free-field conditions (e.g., records from basements, tall structures, or dam crests), records from locations not within the models' range of applicability, records lacking key information, and records with identified problems, although the specific decisions of each team varied (Power *et al.*, 2008).

The teams made various assumptions in selecting their final datasets for regression. One of the most significant decisions for the researchers was whether or not to include aftershocks in their regression subsets. Three teams (AS08, CY08, and I08) opted to include aftershocks (Abrahamson *et al*., 2008). Chiou and Youngs (2008a) only included records within 70 km of the earthquake source in order to avoid bias in the dataset, and they made assumptions on attenuation characteristics to extend their model to larger distances. Idriss (2008) only included rock sites (locations with $V_{S30} \geq 450$ m/s) in his model; 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.

Explanatory variables

A summary of the explanatory variables used in the GMPEs is presented in Table 2. There is a wide range of model complexity, but as a whole the NGA models are much more

complicated than their previous counterparts. 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 models generally incorporated just one per model. Four of the NGA models utilize the time-averaged shear wave velocity over the top 30 meters of the subsurface (V_{S30}) as the primary site characteristic; Idriss (2008) does not quantitatively incorporate site characteristics into his model, although he provides separate coefficients for soft and hard rock sites. The first researchers to quantitatively incorporate shear wave velocity in a GMPE were Boore *et al.* (1997). Other previous models utilize dummy variables (flags that take on a value of "0" or "1") 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}, where $Z_{1.0}$ is the depth to V_S = 1 km/sec, and $Z_{2.5}$ is the depth to V_S = 2.5 km/sec. One of the previous models (Campbell, 1997) includes a parameter D for depth to basement rock.

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) needed to be calculated from the available information in the flatfile. Although V_{S30} is included for nearly every record in the flatfile, the depth parameters $Z_{1.0}$ and $Z_{2.5}$ are not present for many records. When the values were absent, we estimated $Z_{1.0}$ or $Z_{2.5}$ from available information using recommendations from each of the research teams (Abrahamson and Silva, 2008; Chiou and Youngs, 2008a; Campbell and Bozorgnia, 2007). These procedures and their implications will be discussed later in this paper.

Table 2. Explanatory variables of the GMPEs in this study

		NGA models					Previous models				
	Parameter	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	•			•						
Distance and path parameters	Closest distance to the rupture plane, R_{RUP}	•		•	•	•	•			•	•
	Horizontal distance to the surface projection of the rupture (Joyner-Boore distance), R_{JB}	•	•	•	•			•			
	Horizontal distance to the top edge of the rupture measured perpendicular to the strike, 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 meters of the subsurface, V_{S30}	•	•	•	•			•			
	Depth to bedrock or specific shear wave velocity horizon ($Z_{1.0}$, $Z_{2.5}$, or D) [1]	•		•	•				•		
	Site conditions flag [2]						•		•	•	
	PGA (or Sa) on rock, as baseline for nonlinear site response	•	•	•	•		•				

NOTES:

1. AS08 and CY08 use depth to Vs = 1.0 km/s ($Z_{1.0}$), CB08 uses depth to Vs = 2.5 km/s ($Z_{2.5}$), and C97 uses depth to basement rock (D).
2. 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.

For the Parkfield and San Simeon earthquakes, which are not present in the NGA flatfile, we determined the explanatory variables from a variety of sources. 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 (Dreger, 2004) and San Simeon (Rolandone *et al.*, 2004) earthquakes. The distance measures were calculated directly from the source-to-site geometry for each location. Some of the stations that recorded the Parkfield and San Simeon earthquakes also recorded other earthquakes in the NGA flatfile; site characteristics for these stations were determined directly from the flatfile. For the other locations, site characteristics were determined from measured shear wave velocity data if available (Kayen, 2007; Real, 1988). For locations without measured V_S profiles, site characteristics were inferred from surficial geologic units (California Geological Survey, 2007; Shakal *et al.*, 2005) using the classification scheme of Wills and Clahan (2006), which the NGA research team also used to estimate V_{S30} for many records in the flatfile (Chiou *et al.*, 2008).

Response variables

The response variables for the GMPEs are peak ground acceleration (PGA), peak ground velocity (PGV), and 5%-damped elastic pseudo-response spectral acceleration (*Sa*). All models have equations for PGA and *Sa*, although the spectral periods with defined coefficients vary from model to model, especially for the older models. Four of the NGA models (AS08, BA08, CB08, and CY08) and one of the previous models (C97) have equations for PGV. In this study, we analyze PGA and *Sa* for the six spectral periods represented in the USGS national seismic hazard maps: 0.1, 0.2, 0.3, 0.5, 1.0, and 2.0 seconds (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. The previous GMPEs offer predictions for *Sa* to maximum periods of 2 to 5 seconds, whereas the new GMPEs offer predictions up to 10 seconds. Although the new GMPEs can predict *Sa* at long periods, 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 periods of 2 seconds and smaller.

The observed ground motions for records in the NGA database were obtained directly from the flatfile. To calculate the observed ground motions for the Parkfield and San Simeon 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 (California Geological Survey, 2004; United States Geological Survey, 2004). 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, 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 (2008) to compute GMRotI50 from the recorded acceleration time histories.

Ranges of applicability of the models

Each model is applicable only within specific ranges of magnitude, distance, and other variables. Table 3 presents the ranges of applicability of the five NGA models, which have specific requirements for magnitude, distance, and V_{S30}. Campbell and Bozorgnia (2008) specify some additional requirements regarding the depth parameter $Z_{2.5}$, depth to top of rupture Z_{TOR}, and fault dip δ. At the bottom of Table 3 is a summary of the requirements that an earthquake ground motion record must meet in order to be applicable to all five

Table 3. Ranges of applicability of the NGA models

Model	Magnitude	Distance [km]	V_{S30} [m/sec]	Additional requirements
AS08	$5.0 \le \mathbf{M} \le 8.5$	$R_{RUP} \le 200$	No specification	–
BA08	$5.0 \le \mathbf{M} \le 8.0$	$R_{JB} \le 200$	$180 \le V_{S30} \le 1300$	–
CB08	$4.0 \le \mathbf{M} \le 7.5$ (N*), 8.0 (R), 8.5 (SS)	$R_{RUP} \le 200$	$150 \le V_{S30} \le 1500$	$Z_{2.5} \le 10$ km $Z_{TOR} \le 15$ km $15° \le \delta \le 90°$
CY08	$4.0 \le \mathbf{M} \le 8.0$ (N and R), 8.5 (SS)	$R_{RUP} \le 200$	$150 \le V_{S30} \le 1500$	–
I08	$4.5 \le \mathbf{M} \le 8.0$	$R_{RUP} \le 200$	$V_{S30} \ge 450$	–
ALL	*Normal faulting:* $5.0 \le \mathbf{M} \le 7.5$ *Reverse and strike-slip faulting:* $5.0 \le \mathbf{M} \le 8.0$	$R_{RUP} \le 200$ and $R_{JB} \le 200$	*Excluding I08:* $180 \le V_{S30} \le 1300$ *Including I08:* $450 \le V_{S30} \le 1300$	$Z_{2.5} \le 10$ km $Z_{TOR} \le 15$ km $15° \le \delta \le 90°$

NOTE:
* For CB08 and CY08, different maximum magnitudes are specified for normal (N), reverse (R), and strike-slip (SS) faulting mechanisms.

Table 4. Ranges of applicability of the previous models

Model	Magnitude		Distance [km]		Spectral Period [sec]	
	Min	Max	Type	Max	Min	Max
AS97	–[1]	8.5	R_{RUP}	–	0.01	5.0
BJF97	5.5	7.5	R_{JB}	80	0.10	2.0
C97	5.0	–	R_{SEIS}	60	0.05	4.0
SCE97	4.0	–	R_{RUP}	100	0.07	4.0
I91	–	–	R_{RUP}, R_{HYP}[2]	–	0.03	5.0

NOTES:
1. An en dash (–) means that the model does not explicitly specify a minimum or maximum.
2. I91 specifies R_{HYP} for $\mathbf{M} \le 6$ and R_{RUP} for $\mathbf{M} > 6$.

models; we followed these requirements when developing our testing subsets.

Table 4 presents the ranges of applicability of the five previous models. In addition to the smaller range of spectral periods with defined coefficients, 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) suggest that even the models with specific distance limits (such as BJF97 and C97) may be reasonably extrapolated to 100 km.

Testing subsets

In order to perform a quantitative comparison of the predictive abilities of the NGA models, the first logical step is to test the models on the NGA database, upon which the new models were developed. The NGA database contains many records that do not meet the criteria specified in Table 3. As a result, we needed to generate testing subsets of the NGA database containing records that met all the requirements of the models. Figure 1 is a flowchart illustrating the testing subsets and the number of ground motion records in each of the subsets. We made a distinction between mainshocks and aftershocks because some of the NGA model developers (BA08 and CB08) did not include aftershocks in their regression databases. Thus, the testing of the BA08 and CB08 on the aftershocks subset serves as a blind test for these models. In order to test the I08 model, each subset was further subdivided into soil sites ($180 \leq V_{S30} < 450$ m/sec) and rock sites ($450 \leq V_{S30} \leq 1300$ m/sec, where I08 is applicable). Because of the differing assumptions made by the developers, the final testing subsets do not perfectly match the regression datasets of any of the developers, but they provide a useful basis for comparison.

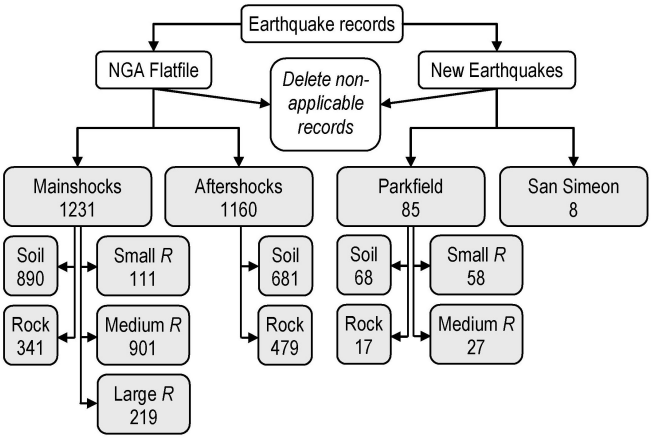

Fig 1. Flowchart of the subset delineation process, along with the sample size of each final subset.

We also compare the effect of distance on prediction accuracy, using subdivisions of small ($R_{RUP} \leq 10$ km), medium ($10 < R_{RUP} \leq 100$ km), and large distances ($100 < R_{RUP} \leq 200$ km) for subsets with sufficient data in each category. Ground motion often displays little attenuation 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. The previous models cannot be tested on the large-distance subset.

The blind comparison using the new and old models was performed using the database of new earthquakes. In order to test both the new and old models on the Parkfield and San Simeon earthquakes, we restricted records to distances no greater than 100 km from the earthquake source. Most of the records for the Parkfield earthquake are near-source, but most of the records for the San Simeon earthquake are at distances greater than 100 km. Because the database for the San Simeon earthquake was reduced to only 8 records, we did not subdivide this dataset. We did not compare the previous models with the new models using the NGA testing subsets. Because the NGA models were exposed to larger portions of these testing subsets during model development, the NGA models would have an unfair advantage.

In developing the final testing subsets, we deleted non-applicable records from the NGA flatfile if they did not meet the criteria in Table 3 (since only the new models are tested), and we deleted records from the database of new earthquakes if they did not meet the combined criteria in Tables 3 and 4 (since both the new and old models are tested). We also deleted records not representative of free-field conditions, records without finite fault models (which did not have values for R_{RUP}, R_{JB}, W, or Z_{TOR}), records missing other important information (such as V_{S30} or Sa), and records with identified problems. Boore and Atkinson (2007) provide a useful record-by-record summary of reasons for excluding records from the NGA flatfile.

METHODS

Implementation of the models

For each earthquake ground motion record in each of the testing subsets, we computed median estimates of PGA and Sa at periods 0.1, 0.2, 0.3, 0.5, 1.0, and 2.0 sec. We consider PGA to be part of the acceleration response spectrum with a period of 0.01 sec, because most of the models (CB08, CY08, and I08) have identical coefficients for PGA and Sa(0.01 sec). Idriss (2008) notes that the values of PGA and Sa(0.01 sec) in the flatfile are generally within 2% of each other.

A median prediction represents a model's best estimate of the expected ground motion at a site. Thus, a comparison of different models' median ground motion estimates with the observed values allows us to quantify the various models' goodness of fit. To perform our computations, we utilized the

open-source statistical language and environment **R** (R Core Development Team, 2009).

Goodness-of-fit measures

Goodness-of-fit statistics are utilized to quantify the comparison of the model predictions with the observed ground motion records. The primary 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 by the equation

$$E = \left[1 - \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} \left(Y_{ij} - \hat{Y}_{ij} \right)^2}{\sum_{j=1}^{m} \sum_{i=1}^{n} \left(Y_{ij} - \bar{Y} \right)^2} \right] \cdot 100\% , \qquad (2)$$

where m is the number of periods under consideration (for this study, $m = 7$), n is the number of ground motion records in the dataset, the observed values (PGA, *Sa*, etc.) are denoted by Y_{ij}, predicted values are denoted by $\hat{Y}_{ij}$, and the mean of the $n \times m$ matrix of observed values is denoted by $\bar{Y}$. This calculation weights contributions from each of the seven periods equally. The value of E may vary between $-\infty$ and 100%; when E is less than 0, the arithmetic mean of the observed values has greater prediction accuracy than the model itself. Compared to other goodness-of-fit statistics (such as the correlation coefficient, r), the coefficient of efficiency 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). We computed E using the observed and predicted values in real space (as opposed to logarithmic space, in which the models were developed), since seismic hazard assessment is concerned with the real values of the response variables (not the logarithms of the values).

RESULTS

Mainshocks

Table 5 presents the NGA models' prediction accuracy for PGA and *Sa*, when testing the models on mainshocks from the NGA database, the largest subset in this study. The coefficients of efficiency for the subdivisions of V_{S30} and R are presented, in addition to the total coefficient of efficiency on the entire mainshocks subset. In the total sense, the values of E are generally between 40 and 60 percent. All models perform more poorly at small and large distances than at intermediate distances. In two cases (CY08 for small distances and AS08 for large distances), we find E to be less than zero, indicating that the mean of the observed motions is a better predictor than the GMPEs in these cases. Quite interestingly, the two models with the highest prediction accuracy, BA08 and CB08, are two of the simpler NGA models. The more complicated models, AS08 and CY08, do not perform as well. When included in the testing subset for rock sites, the I08 model ranks fourth of five.

Table 5. Coefficients of efficiency for mainshocks in the NGA database

		NGA models				
		AS08	BA08	CB08	CY08	I08
Subdivision 1	Soil	57.7	59.5	60.4	53.7	–
	Rock	49.7	55.6	57.2	23.5	43.4
Subdivision 2	Small R	22.6	34.8	35.4	−11.8	–
	Medium R	45.5	45.4	47.2	36.7	–
	Large R	−6.5	15.3	23.8	3.5	–
Total E		54.8	58.1	59.3	42.7	–
Model rankings based on total E		3	2	1	4	–

Table 6. Coefficients of efficiency for aftershocks in the NGA database

		NGA models				
		AS08	BA08	CB08	CY08	I08
Subdivisions	Soil	51.2	49.8	44.6	45.8	–
	Rock	25.6	39.2	28.6	30.9	37.4
Total E		47.9	47.6	41.2	43.1	–
Model rankings based on total E		1	2	4	3	–

Aftershocks

The models' prediction accuracy for aftershocks is less than that of mainshocks, as seen in Table 6. The ranges of E are in the 40- to 50-percent range when the whole subset is analyzed. Most models suffer a considerable decrease in E from soil to rock, most noticeably the AS08 model. Only the AS08, CY08, and I08 model teams included aftershocks in their regression datasets. The CB08 model tends to over-predict ground motion for aftershocks, and thus has a lower coefficient of efficiency. However, the BA08 model, having not been influenced by any of the values in this subset during model development, performs surprisingly well.

Blind comparison test 1: Parkfield earthquake

The **M** 6.0 Parkfield earthquake of 2004 generated an unprecedented amount of near-source ground motion records. Of the 85 records with $R_{RUP} \leq 100$ km that we tested, 58 records (68%) are located within 10 km of the ruptured area. As noted by Shakal *et al*. (2005), the ground motions near the fault were highly variable in the Parkfield earthquake. The high near-source variability in the observed ground motions is manifested in the models' relatively low prediction accuracy for this earthquake, presented in Table 7. When the analysis is

separated into categories by distance, the coefficients of efficiency for the NGA models at small distances are similar to the values in mainshocks subset (E less than 30%). At medium distances, all models—new and old—have relatively high values of E (above 65%), with the CY08 model performing best. When the analysis is separated by V_{S30}, the models' performance at soil sites is better than their performance at rock sites.

For some—but not all—of the cases, the NGA models outperform their previous counterparts. When comparing the total dataset, the CB08 relationship has the highest coefficient of efficiency, followed by the BJF97 model, and then two of the NGA models (AS08 and BA08). Interestingly, the BJF97 model performs at a level similar to that of to its contemporary (BA08), and noticeably better than the other models in the previous generation. One difference between BJF97 and the other previous models is that Boore *et al*. (1997) were the first team to quantitatively incorporate site characteristics into a GMPE, perhaps giving the model greater prediction accuracy.

Blind comparison test 2: San Simeon earthquake

As Table 8 illustrates, the prediction accuracy of the models is much better for the San Simeon earthquake than for the Parkfield earthquake, most likely because highly variable near-source ground motions no longer dominate the database. The coefficients of efficiency are within a narrow range for the four NGA models (hovering near 70%), with the CY08 model performing best. Unlike the Parkfield earthquake, there is a clear difference between the new and previous models for the San Simeon earthquake. All four NGA models have higher values of E than their previous counterparts, with an average increase of 17.2%.

DISCUSSION

Incorporation of aftershocks in model development

As already explained, 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 or not to include aftershocks. The aftershock records of the 1999 **M** 7.6 Chi-Chi, Taiwan, earthquake, comprise 83% of the aftershock records in the flatfile and 97% of the aftershock records in our subset. (The reason why the proportion in our subset is so large is because finite fault models were only developed for 7 aftershocks in the NGA flatfile, four of which were aftershocks of the widely-recorded Chi-Chi earthquake.) The AS08, CY08, and I08 model teams included aftershocks in their regression datasets, most of which were from the Chi-Chi sequence. One potential problem with including such a high proportion of records from a single event is that the model may become over-fit toward the characteristics of that event, and the model's ability to generalize to other situations is lowered. This could be one potential reason why the BA08 and CB08 models outperformed the AS08, CY08, and I08 models on the mainshocks subset, which comprised more earthquakes (50) than any other subset in this study.

In assessing seismic hazards, the greatest hazard contribution comes from mainshocks, not aftershocks. For a given magnitude, aftershocks tend to generate smaller ground motions than mainshocks of same magnitude, and the spectral scaling is different (Boore and Atkinson, 2008; Boore and Atkinson, 1989; Atkinson 1993). Mainshocks are more likely to generate greater ground motions at a site. We do not intend to discount the importance of aftershocks, as aftershocks generate potentially devastating stresses and strains on already-fatigued systems, but we argue that mainshocks are more important from the point of view of seismic hazard map generation. Thus, the most important testing subset in this study is the mainshocks subset.

Table 7. Coefficients of efficiency for the Parkfield dataset

		NGA models					Previous models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Subdivision 1	Soil	36.6	34.7	42.0	24.3	–	32.4	40.0	32.7	31.6	–
	Rock	43.1	44.7	41.1	30.3	40.9	8.7	44.4	19.1	15.1	26.8
Subdivision 2	Small R	23.0	20.7	26.7	5.2	–	8.4	25.6	11.4	9.1	–
	Medium R	65.0	70.5	74.9	75.9	–	73.0	75.6	73.8	74.3	–
Total E		38.1	36.9	42.0	25.8	–	27.7	41.1	30.1	28.4	–
Model rankings based on total E		3	4	1	8	–	7	2	5	6	–

Table 8. Coefficients of efficiency for the San Simeon dataset

	NGA models					Previous models				
	AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
Total E	66.2	67.0	66.2	70.3	–	58.8	58.8	49.2	34.0	–
Model rankings based on total E	3 (tie)	2	3 (tie)	1	–	5 (tie)	5 (tie)	7	8	–

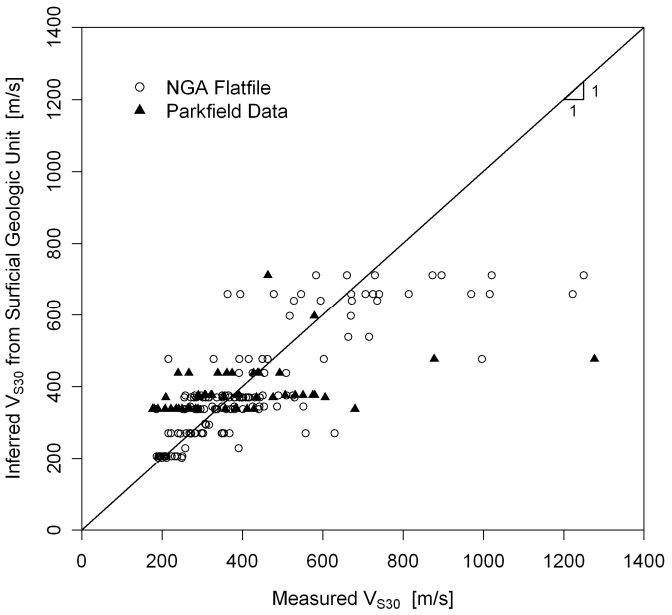

Fig. 2. Measured V_{S30} versus the corresponding values of V_{S30} inferred from surficial geology using Wills and Clahan (2006), for sites in California with measured V_S profiles.

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. 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, even though the ground motion and spectral scaling is known to differ. As a result, the ground motion predictions for mainshocks are more prone to underprediction. We recommend that model developers utilize aftershock dummy variables when they choose to include aftershocks in their regression datasets; however, our results suggest that aftershocks should not be included in model development.

Uncertainty of site parameters

Of the model parameters, the greatest contribution to epistemic uncertainty 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 remainder are inferred using correlations of V_{S30} with surficial geology, such as those published by Wills and Clahan (2006). However, Scott *et al.* (2004) find that shear wave velocity correlates poorly with geologic units. In Fig. 2, we explore 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. In addition to the measured stations in the flatfile, we include 55 stations in the Parkfield, California, vicinity with measured V_S profiles provided by Kayen (2007). There is a huge amount of scatter to this plot, and the discrete categories based on surficial geologic unit are clearly visible. The coefficient of efficiency for this V_{S30} estimation procedure is 52.9%. Although the NGA flatfile has estimated values of V_{S30} for almost every recording station, we suspect that many of the V_{S30} estimates are inaccurate.

Even more difficult to estimate than V_{S30} are the depth parameters, $Z_{1.0}$ (used in AS08 and CY08) and $Z_{2.5}$ (used in CB08). The preferred method of determining the depth parameters is using a site-specific measured V_S profile that extends to the 1.0 km/sec and 2.5 km/sec horizons. Unfortunately, only 54 sites in the NGA flatfile have measured V_S profiles that reach 1.0 km/sec (Chiou and Youngs, 2008a), and even fewer reach 2.5 km/sec. If the site is located in an area where a regional velocity model is available (such as San Francisco or Los Angeles), then the depth parameters may be determined from the regional velocity model. If a measured V_S profile or regional velocity model is unavailable (which is the most common case), the depth parameters are determined by the recommendations of the model developers.

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

$$Z_{1.0} = \begin{cases} \exp(6.745) & \text{for } V_{S30} < 180\ m/s \\ \exp\left[6.745 - 1.35 \cdot \ln\left(\frac{V_{S30}}{180}\right)\right] & \text{for } 180 \le V_{S30} \le 500\ m/s \\ \exp\left[5.394 - 4.48 \cdot \ln\left(\frac{V_{S30}}{500}\right)\right] & \text{for } V_{S30} > 500\ m/s\ . \end{cases} \quad (3)$$

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

$$Z_{1.0} = \exp\left[28.5 - \frac{3.82}{8} \cdot \ln\left({V_{S30}}^8 + 378.7^8\right)\right]. \quad (4)$$

In order to estimate $Z_{2.5}$, Campbell and Bozorgnia (2007) offer guidelines for extrapolating the estimates of $Z_{1.0}$ or $Z_{1.5}$ if these values are available. If neither $Z_{1.0}$ nor $Z_{1.5}$ is known, and if basin depth is not known to be particularly shallow or deep, then Campbell and Bozorgnia recommend assigning $Z_{2.5}$ to the "default value" of 2 km. In this case, V_{S30} will solely represent the site characteristics in the ground motion calculation. Campbell (2000) offers advice and clarifications for the depth to basement rock D found in his 1997 relationship.

Graphs of the two median relationships for $Z_{1.0}$ are presented in Fig. 3, along with data from the 448 sites in the NGA flatfile with specified values of $Z_{1.0}$. The considerable amount

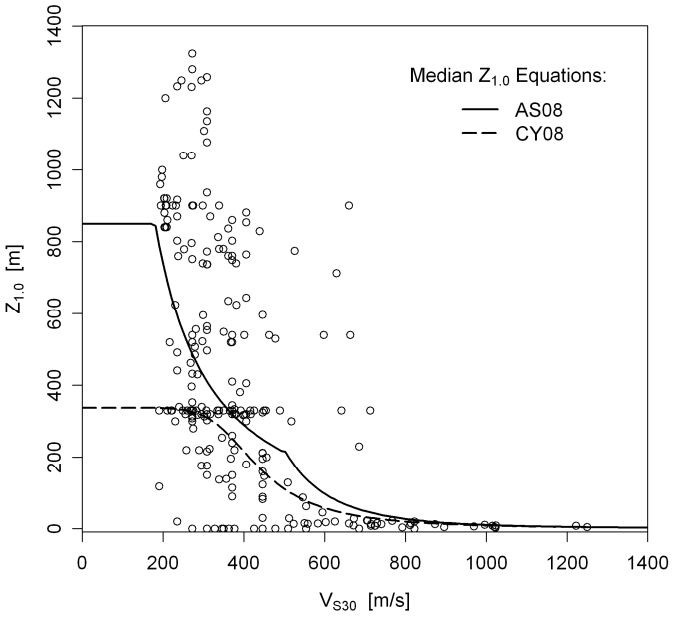

Fig. 3. $Z_{1.0}$ versus V_{S30} for records in the NGA flatfile having specified values of $Z_{1.0}$. Also shown are the median equations employed by the AS08 and CY08 models for estimating $Z_{1.0}$ as a function of V_{S30}.

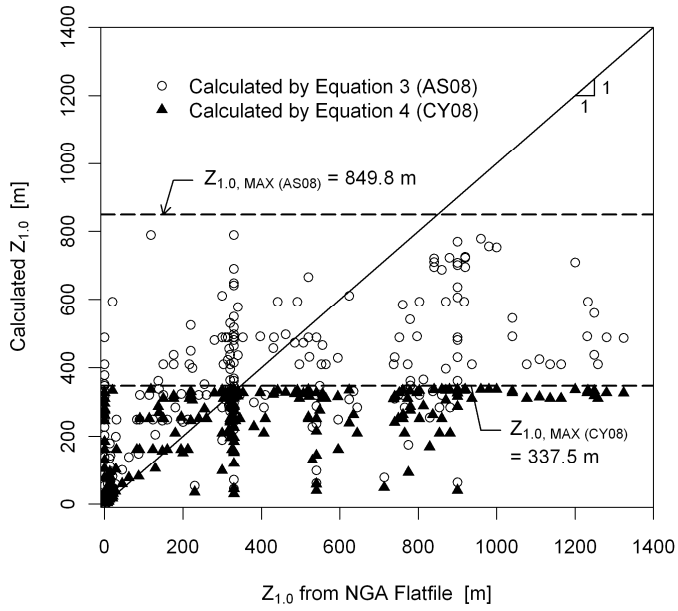

Fig. 4. Comparison of $Z_{1.0}$ calculated from V_{S30} using the median equations versus the corresponding values of $Z_{1.0}$ from the flatfile, for Equation 3 (AS08) and Equation 4 (CY08).

of scatter in Fig. 3 is even more apparent when comparing plots of measured versus calculated $Z_{1.0}$ (Fig 4). 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 Equations 3 and 4. 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 3 and 337.5 m for Equation 4. Abrahamson and Silva (2008) developed Equation 3 from analytical site response models. Chiou and Youngs (2008b) describe how they utilized an updated velocity model for southern California when they developed Equation 4, 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 Equations 3 and 4 and the values in the flatfile, but the discrepancies demonstrate that the depth parameters are fraught with uncertainty.

The key problem with site characteristics in GMPEs is a lack of measurements. For the majority of cases, site-specific V_{S30} measurements are unavailable; therefore, V_{S30} is inferred from surficial geology or site conditions. Then, the depth parameters are estimated from V_{S30}. When this happens, information about the ground surface (i.e., surficial geology) is being 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 $Z_{2.5}$). It is widely agreed that site characteristics should be incorporated into GMPEs. In our results, we find that models with quantitative site parameters (*i.e.*, AS08, BA08, CB08, CY08, BJF97, and C97) generally perform better than models that do not include site parameters, or models that only include a dummy variable. However, to improve the prediction accuracy of the models, we argue that there must be a greater emphasis on site-specific data collection. An increased database of site characteristics would reduce the scatter in Figs. 3 and 4, and would ultimately lead to more reliable ground motion predictions.

Effect of distance on prediction accuracy

As seen in Tables 5 and 7, the GMPEs perform best at intermediate distances, where the largest amount of data is available. In addition to the lack of data, ground motion is highly variable at small distances, and the estimation of ground motion at large distances raises other complications, such as Moho bounce effects. As a result, ground motions at small and large distances are more difficult to predict than ground motions at intermediate distances. The CY08 model performs poorly in several of the subsets, but performs superiorly in the blind comparison tests in the intermediate distance range. One of the key differences in model development is that Chiou and Youngs (2008a) only included sites within 70 km of rupture in their regression dataset, while the other NGA research teams included sites within 200 km. Perhaps the over-fitting of the CY08 model to intermediate distances gives it increased predictive capabilities within that range, and decreased predictive capabilities outside of that range.

Table 9. Summary of rankings

		NGA models					Previous models				
		AS08	BA08	CB08	CY08	I08	AS97	BJF97	C97	SCE97	I91
NGA flatfile	Mainshocks	3	2	1	4	–	–	–	–	–	–
	Aftershocks	1	2	4	3	–	–	–	–	–	–
Blind test	Parkfield	3	4	1	8	–	7	2	5	6	–
	San Simeon	3	2	3	1	–	5	5	7	8	–

Improving prediction accuracy

In this paper, we have focused our comparisons on the models' median estimates of PGA and *Sa*. The median equations are the models' best estimates of the expected ground motion, and they are the key focus of the model development efforts. However, we must emphasize the importance of the uncertainties of the median estimates. Ground motion variability has often been disregarded or discounted in the application of GMPEs (Bommer and Abrahamson, 2006). Strasser *et al.* (2009) present a good discussion of the contributions to uncertainty and the challenges in reducing them. They claim that reductions in uncertainty are not necessarily brought by increasing the number of explanatory variables in the models or increasing the quantity of ground motion records in the regression datasets. Kuehn *et al.* (2009) warn about the dangers of over-fitting using the current approaches to model development, and propose an alternative method for the development of GMPEs that uses generalization error minimization techniques.

Table 9 is a summary of the model rankings for all the testing subsets. The excellent performance of the simpler models (BA08 and CB08) on the mainshocks subset and their generally high rankings on the blind comparison subsets lends credence to the suggestion that more complicated models do not necessarily offer more accurate predictions. A higher-quality regression dataset (not necessarily higher-*quantity*) with greater measurements of site characteristics, coupled with simple functional forms in the GMPEs, may yield the best solution.

CONCLUSION

Using the Nash-Sutcliffe model efficiency coefficient (*E*) as the primary goodness-of-fit statistic, we have compared the prediction accuracy of the five ground motion prediction relations released as part of the NGA project. The coefficient of efficiency, commonly used in hydrology, is a superior goodness-of-fit statistic to many other statistics found in the literature, and it works well as a framework for validating alternative predictions. First, we tested the NGA models on subsets of the database upon which they were developed. Then, we compared the performance of the new models with the previous generation of models by implementing a blind comparison test on two recent California earthquakes (the 2004 **M** 6.0 Parkfield earthquake, and the 2003 **M** 6.5 San Simeon earthquake), which were not present in any of the databases used to develop the models. The newer models generally perform better than their previous counterparts in these blind tests, but all models had difficulty predicting the highly variable near-source ground motions of the Parkfield earthquake.

We find that the decisions that model developers make when selecting their regression datasets greatly influence the models' predictive capabilities. Allowing a model to be extrapolated to distances far beyond the range found in the regression dataset may be detrimental from a prediction standpoint. Especially on the blind comparison test, the CY08 model (developed only with data from distances no greater than 70 km from the earthquake source) performs relatively well at intermediate distances, but it performs more poorly than the other models at small and large distances. Including aftershocks in the regression dataset may lead to unconservative ground motion predictions if they are not properly dealt with through the use of an aftershock dummy variable, and in general do not appear to help in model prediction accuracy. 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 over-fitting the data to those particular scenarios, thus reducing the models' ability to generalize to other situations. High model complexity, whether through large numbers of explanatory variables or convoluted functional forms, can also lead to over-fitting. We find that two of the models with simpler functional forms and explanatory variables, BA08 and CB08, have the highest prediction accuracy when tested on a comprehensive subset of mainshocks from the NGA database, and also perform well in blind situations. These results suggest that increasing the complexity of GMPEs does not necessarily increase their prediction accuracy. Instead of increasing the models' complexity, an increased emphasis on the measurement of site parameters would lead to a higher-quality regression dataset, which would ultimately lead to better ground motion predictions in the future.

REFERENCES

Abrahamson, N., G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I. M. Idriss, W. Silva, R. Youngs [2008]. "Comparisons of the NGA Ground-Motion Relations," *Earthquake Spectra*, No. 24, pp. 45–66.

Abrahamson, N. A., and W. J. Silva [1997]. "Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes," *Seismol. Res. Lett.*, No. 68, pp. 94–127.

Abrahamson, N. A., and W. J. Silva [2008]. "Summary of the Abrahamson & Silva NGA Ground-Motion Relations," *Earthquake Spectra*, No. 24, pp. 67–97.

Atkinson, G. M. [1993]. "Earthquake Source Spectra in Eastern North America," *Bull. Seismol. Soc. Am.*, No. 83, pp. 1778–1798.

Bommer, J. J., and N. A. Abrahamson [2006]. "Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates?" *Bull. Seismol. Soc. Am.*, No. 96, pp. 1967–1977.

Boore, D. M. [2008]. "TSPP—A Collection of FORTRAN Programs for Processing and Manipulating Time Series," Version 1.5, *U.S. Geol. Surv. Open-File Rep. 2008–1111*.

Boore, D. M., and G. M. Atkinson [1989]. "Spectral Scaling of the 1985 to 1988 Nahanni, Northwest Territories, Earthquakes," *Bull. Seismol. Soc. Am.*, No. 79, pp. 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.

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," *Earthquake Spectra*, No. 24, pp. 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.*, No.68, pp. 128–153.

Boore, D. M., J. Watson-Lamprey, and N. A. Abrahamson [2006]. "Orientation-Independent Measures of Ground Motion," *Bull. Seismol. Soc. Am.*, No. 96, pp. 1502–1511.

California Geological Survey [2004]. Earthquake ground motion records for the Parkfield 2004-09-28 and San Simeon 2003-12-22 earthquakes. Accessed at *Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) Virtual Data Center*, <http://db.cosmos-eq.org>.

California Geological Survey [2007]. 1:250,000 Scale Regional Geologic Map Series; Monterrey, Fresno, San Luis Obispo, and Bakersfield quadrangles, <http://www.conservation.ca.gov/cgs/rghm/rgm/250k_index/Pages/250k_index.aspx>.

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.*, No. 68, pp. 154–179.

Campbell, K. W. [2000]. Erratum to "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.*, No. 71, pp. 352–354.

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.*, No. 93, pp. 314–331.

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.

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," *Earthquake Spectra*, No. 24, pp. 139–171.

Chiou, B. S., R. Darragh, N. Gregor, and W. Silva [2008]. "NGA Project Strong-Motion Database," *Earthquake Spectra*, No. 24, pp. 23–44.

Chiou, B. S., and R. R. Youngs [2008a]. "An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra," *Earthquake Spectra*, No. 24, pp. 173–215.

Chiou, B. S., and R. R. Youngs [2008b]. "An NGA Model for Average Horizontal Component of Peak Ground Motion and Response Spectra," *PEER Report No. 2008/09*, Pacific Earthquake Engineering Research Center, University of California, Berkeley.

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-Science Rev.*, No. 61, pp. 43–104.

Dreger, D. [2004]. "09/28/2004 Preliminary Slip Model," Berkeley Seismological Laboratory, University of California, Berkeley. Accessed at *Finite-Source Rupture Model Database*, M. Mai, <http://www.seismo.ethz.ch/srcmod/>.

Ghasemi, H., M. Zare, and Y. Fukushima [2008]. "Ranking of Several Ground-Motion Models for Seismic Hazard Analysis in Iran," *J. Geophys. and Engrg.*, No. 5, pp. 301–310.

Idriss, I. M. [1991]. "Procedures for Selecting Earthquake Ground Motions at Rock Sites," Report prepared for the Structures Division, Building and Fire Research Laboratory, National Institute of Standards and Technology. Department of Civil Engineering, University of California, Davis.

Idriss, I. M. [2002]. "Attenuation Relationship Derived by I. M. Idriss in 2002," reported in "Empirical Model for Estimating the Average Horizontal Values of Psuedo-Absolute Spectral Accelerations Generated by Crustal Earthquakes" [2007], Interim report issued for USGS review, <http://peer.berkeley.edu/products/nga_project.html>.

Idriss, I. M. [2008]. "An NGA Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes," *Earthquake Spectra*, No. 24, pp. 217–242.

Kayen, R. [2007]. Shear wave velocity data from the Parkfield, California, area. Personal communication.

Kramer, S. L. [1996]. *Geotechnical Earthquake Engineering*, Prentice Hall, Upper Saddle River, N.J.

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.*, No. 99, pp. 2335–2347.

Legates, D. R. and G. J. McCabe [1999]. "Evaluating the Use of Goodness-of-Fit Measures in Hydrologic and Hydroclimactic Model Validation." *Water Resources Res.*, No. 35, pp. 233–241.

Nash, J. E. and J. V. Sutcliffe [1970]. "River Flow Forecasting Through Conceptual Models. Part I – A Discussion of Principles." *J. Hydrology*, No. 10, pp. 282–290.

Pacific Earthquake Engineering Research Center [2008]. "Next Generation Attenuation of Ground Motions (NGA) Project," <http://peer.berkeley.edu/products/nga_project .html>.

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. Geological Survey Open-File Report 2008–1128*.

Power M., B. Chiou, N. Abrahamson, Y. Bozorgnia, T. Shantz, C. Roblee [2008]. "An Overview of the NGA Project," *Earthquake Spectra*, No. 24, pp. 3–21.

R Development Core Team [2009]. 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>.

Real, C. R. [1988]. "Turkey Flat, USA, Site Effects Test Area," *Technical Report No. 88-2*, California Department of Conservation, Division of Mines and Geology, Earthquake Shaking Assessment Unit.

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>.

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.* No. 68, pp. 180–189.

Scott, J. B., M. Clark, C. Lopez, A. Pancha, T. Rasmussen, S. B. Smith, W. Thelen, and J. N. Louie [2004]. "Three Urban Transects of Shallow Shear Wave Velocity Using the Refraction Microtremor Method," *Proceedings of the Managing Risk in Earthquake Country Conference Commemorating the 100th Anniversary of the 1906 Earthquake*, San Francisco, California.

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.*, No. 76, pp. 27–39.

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. Earthquake Engrg.*, No. 6, pp. 149–177.

Star, L. M., J. P. Stewart, R. W. Graves, K. W. Hudnut [2008]. "Validation Against NGA Empirical Model of Simulated Motions for M7.8 Rupture of San Andreas Fault," *The 14th World Conf. on Earthquake Engrg.*, Beijing, China.

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer [2009]. "Sigma: Issues, Insights, and Challenges," *Seismol. Res. Lett.*, No. 80, pp. 40–56.

Stewart, J. P., L. M. Star, and R. W. Graves [2008]. "Validation Against NGA Empirical Model of Simulated Motions for M7.15 Rupture of Puente Hills Fault," *Final Report*, Pacific Earthquake Engineering Research Center, University of California, Berkeley.

United States Geological Survey [2004]. Earthquake ground motion records for the Parkfield 2004-09-28 and San Simeon 2003-12-22 earthquakes. Accessed at *Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) Virtual Data Center*, <http://db.cosmos-eq.org>.

Wills, C. J., and K. B. Clahan [2006]. "Developing a Map of Geologically Defined Site-Condition Categories for California," *Bull. Seismol. Soc. Am.*, No. 96, pp. 1483–1501.