Bulletin of the Seismological Society of America, Vol. 98, No. 3, pp. 1304–1318, June 2008, doi: 10.1785/0120070199

Ground-Motion Prediction Equations for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty

by Gail M. Atkinson

Abstract I outline a referenced empirical approach to the development of ground-motion prediction equations (GMPEs). The technique is illustrated by using it to develop GMPEs for eastern North America (ENA). The approach combines the ENA ground-motion database with the empirical prediction equations of Boore and Atkinson (2008) for the reference region of western North America (WNA). The referenced empirical approach provides GMPEs for ENA that are in agreement with regional ground-motion observations, while being constrained to follow the overall scaling behavior of ground motion that is observed in better-instrumented active tectonic regions. They are presented as an alternative to the commonly used stochastic ground-motion relations for ENA. The motivation of the article is not to supplant stochastic GMPEs but is rather to consider other approaches that might shed light on their epistemic uncertainty.

Differences between the referenced empirical GMPEs of this study and the stochastic GMPEs of Atkinson and Boore (2006), along with inconsistencies between both of these studies and inferences based on intensity observations, suggest that uncertainty in median ENA GMPEs is about a factor of 1.5–2 for $\mathbf{M} \geq 5$ at distances from 10 to 70 km. Uncertainty is greater than a factor of 2 for large events ($\mathbf{M} \geq 7$) at distances within 10 km of the source. It may be that saturation effects not modeled in the stochastic predictions, but inferred from observations in other regions, cause overestimation of near-source amplitudes from large events in Atkinson and Boore (2006). On the other hand, these saturation effects cannot be directly verified in ENA data. Differences in predictions according to the approach taken are also significant at distances from 40 to 150 km, due to uncertainty in the shape of the attenuation function that will be realized in future earthquakes.

Introduction

Ground-motion prediction equations (GMPEs), providing estimates of peak ground motion and response spectral amplitudes as a function of earthquake magnitude and distance, are a key input to seismic hazard analysis. These equations allow us to estimate the average ground shaking effects for future earthquakes. In active tectonic regions, GMPEs are most often developed by empirical regression analysis of a database of recorded strong ground motions. An example is the prediction equations developed for shallow crustal earthquakes in active tectonic regions by the Pacific Earthquake Engineering Research Center–Next Generation Attenuation (PEER–NGA) project (Power *et al.*, 2006; Boore and Atkinson, 2008); there are many other such examples that have been commonly used in seismic hazard applications in western North America (WNA) (e.g., Abrahamson and Silva, 1997; Boore *et al.*, 1997; Sadigh *et al.*, 1997).

In eastern North America (ENA), the empirical database is not sufficient in the magnitude-distance range of engineering interest to allow direct development of GMPEs by empirical regression techniques. The problem is illustrated in Figure 1, which compares the ground-motion database available for active tectonic regions (as used by Boore and Atkinson [2008], in their PEER–NGA equations) with that available for ENA (from Atkinson and Boore [2006]). Observe that ENA has a significant body of ground-motion data available, but there is a paucity of recordings at close distances (< 50 km) and large magnitudes ($\mathbf{M} \geq 6$), which is what generally dominates seismic hazard. Because of this lack of critical data, GMPEs for ENA have been developed using a different approach. Almost all GMPEs in use for ENA today are based in whole or in part on a stochastic approach. This approach is model driven, though it is usually empirically calibrated. The most recent example is the ENA prediction equations developed by Atkinson and Boore (2006) (hereafter, AB06). They used a stochastic finite-fault model to simulate expected ground motions, for both hard-rock and

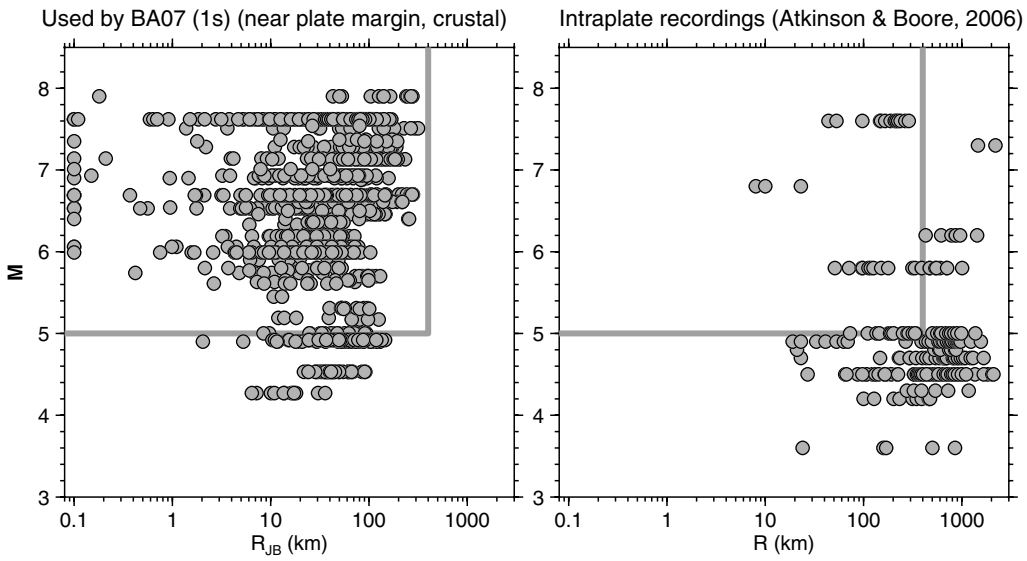

Figure 1. Distribution of response spectra data available for ground-motion studies in active tectonic regions, as used by Boore and Atkinson (2008), (left) and for ENA, as used by Atkinson and Boore (2006), (right).

for the National Earthquake Hazard Reduction Program (NEHRP) B/C boundary site conditions (average shear-wave velocity over the top 30 m, $V_{30} = 760$ m/sec) for earthquakes in ENA. The input to the simulations is a seismological model of the earthquake source and propagation processes, each component of which was derived from empirical data for ENA. The AB06 model thus has a substantial empirical foundation, but the scaling from that empirical foundation to the magnitude– ranges of most interest is inherently driven by the stochastic finite-fault model that underpins the equations. It was demonstrated in Atkinson and Boore (2006) that the AB06 equations provide a reasonable fit to the ENA ground-motion database of Figure 1 overall, but there are some notable exceptions, in particular for high frequencies at intermediate distances. This raises the question of how we may more fully utilize the ENA data in deriving GMPEs and/or in assessing their epistemic uncertainty.

In this article, I outline a new technique that I refer to as a referenced empirical approach to the development of GMPEs. I use the technique to develop GMPEs for ENA. The purpose of this exercise is not to replace the AB06 equations but rather is to provide an alternative model that may help to define their epistemic uncertainty, and to point to areas where the underlying seismological models may warrant further refinement to better fit observations. Alternative equations derived by different techniques provide a more realistic representation of epistemic uncertainty than do alternative equations that were all developed using a common technique. Although I have confidence in the stochastic approach to the development of GMPEs, I believe it is prudent to develop alternative equations using different approaches.

In the referenced empirical approach, the aim is to use empirical data to the largest extent possible. The ENA ground-motion database is used directly to develop suitable regional modifications to empirical ground-motion relations for active tectonic regions. The approach is similar in concept to the hybrid empirical approach of Campbell (2003) in that it is based on making adjustments to empirical equations from other regions. The rationale for such an approach is that empirical equations presumably capture important but complex source and distance scaling effects that are present in the data, but may be missing in a simple seismological model. By making regional adjustments to empirical GMPEs, I anchor my predictions to real experience from other more data-rich regions. The difference between the proposed referenced empirical approach and the hybrid empirical approach is that I do not use a stochastic model, or any other seismological model, to develop the regional adjustment factors. Instead, they are derived directly from the target-region ground-motion database.

The Referenced Empirical Approach

Basis of the Method

The idea behind the referenced empirical approach is to use the target-region ground-motion database in concert with GMPEs from active tectonic regions in order to make the best use of both region-specific data and global experience from better-instrumented regions. The key inputs are the target-region ground-motion database and a set of reference equations. In this study, the target region is ENA.

The most important underlying assumption is that the magnitude scaling of ground motions is the same for the target region (ENA) as that exhibited for shallow crustal earthquakes in WNA and other active tectonic regions, as is the overall near-source behavior with distance. Note that al-

though the scaling is assumed to be similar, no such assumption is made regarding overall levels. The scaling assumption is reasonable based on ground-motion scaling principles and can be verified to some extent over a limited magnitude range (at least from 4 to 6). In particular, the stress drop parameter that controls the high-frequency spectral level is generally considered to be approximately constant in both ENA and WNA over many orders of magnitude, although its average value may be different in the two regions (e.g., Kanamori and Anderson, 1975; Hanks and McGuire, 1981; Boore, 1983; Somerville *et al.*, 1987; Atkinson, 1993; Atkinson and Boore, 2006; Atkinson and Wald, 2007). In this case, we can use the empirical data for ENA to make modifications to GMPEs from WNA, by deriving appropriate adjustments to reference equations that will modify the overall level of the curves and their shape at larger distances. Adjustments to the overall level can accommodate such factors as regional variations in stress drop and event type, while adjustments to the distance coefficient accommodate regional attenuation. Thus I develop new GMPEs for ENA that are referenced to those that were derived from a larger database, richer in observations at large magnitudes and close distances.

The ENA ground-motion database used in this study is taken from Atkinson and Boore (2006) (the response spectral database provided in their electronic supplement); the data distribution in magnitude and distance is shown in Figure 1, while events are listed in Table 1. The AB06 database includes peak ground motions and response spectra from moment magnitude **M** 4.3–7.6, at distances from 10 to >1000 km, recorded on hard-rock sites (NEHRP A, with $V_{30} \sim 2000$ m/sec) in ENA. The Bhuj, India, data (**M** 7.6) are included, as it has been suggested they may be representative of the type of motions expected from large ENA events such as those in New Madrid. However, the applicability of these data is questionable, and the actual amplitudes and site conditions are also less reliable than for the other data sources. Similarly, the applicability of the Nahanni, Northwest Territory, data to ENA is questionable. These factors should be kept in mind when weighting any inferences that depend heavily on these two events. The data are horizontal component or equivalent. For stations where horizontal components were available, both horizontal components are listed. For stations where only the vertical component was available, it was converted to an equivalent horizontal component using a horizontal-to-vertical (H/V) ratio applicable for hard-rock sites (see AB06 for details). The database implicitly gives recorded horizontal components more weight, in that both horizontal components are listed in the database, whereas the converted-vertical component provides a single record per station. This is a simplistic way to make use of all the available data. I restrict the used distance range to <1000 km, as observations are sparse at greater distances.

I adjust the AB06 data amplitudes to provide equivalent values for NEHRP B/C conditions ($V_{30} = 760$ m/sec), as this is a standard reference site condition for more active regions. This adjustment is made using information in AB06. They provided GMPEs for both hard-rock (NEHRP A) and B/C site conditions. The ratio of their predictions for B/C conditions relative to those for A conditions is shown in Figure 2 for selected frequencies. The ratio is a weak function of magnitude and distance for each response spectra frequency (5%-damped pseudoacceleration [PSA]) but depends significantly on distance for peak ground acceleration (PGA) and peak ground velocity (PGV), owing to a dependence of these parameters on frequency content of the signal, which changes with distance. The genesis of these ratios is the difference in seismic impedance between NEHRP A and B/C sites, and its implications for frequency-dependent site amplification, as determined by quarter-wavelength estimates of the net amplification (e.g., Boore and Joyner, 1997). I multiplied each hard-rock ground-motion amplitude (Y_A) in the ENA database by the ratio of predicted B/C motions to predicted A motions (e.g., Y_{BC}/Y_A), as calculated from the AB06 equations, using the applicable magnitude and distance, for the appropriate frequency. This produced an ENA ground-motion database for B/C site conditions (available from the author on request).

It is acknowledged that the conversion of data (or, equivalently, GMPEs) from hard-rock to B/C boundary is a significant source of uncertainty. To maintain internal consistency, it is essential when using the results of this study that the same conversion factors be used. Explicitly, the GMPEs provided in the following discussion are applicable to the B/C boundary. If hard-rock equations for ENA are desired, they need to be generated from the B/C GMPEs by applying the same set of hard-rock to B/C boundary factors that were used to correct the data (as per Fig. 2). Provided this is done, the correct predictions for hard rock are ensured, within the framework of the method. However, the correctness of the predicted levels for all other site conditions (B/C and softer) is dependent on the correctness of the factors adopted to convert ground motions from hard-rock to B/C site conditions.

For most observations, the magnitudes are small enough, or distances large enough, that the choice of distance metric in the database is unimportant. AB06 used closest distance to the fault, whereas Boore and Atkinson (2008) used the closest distance to the surface projection of the fault (Joyner–Boore distance). I assume, with one exception, that all distances in the ENA database are equivalent to Joyner–Boore distance (R_{jb}). The exception is the Nahanni earthquake observations, for which I obtain the correct R_{jb} distances from the PEER–NGA database, for use as the distance metric (this event has near-source observations, so the distance metric is important). The first step of database preparation is equivalent to that used by Boore and Atkinson (2008), who applied empirical factors to reduce all observations from the PEER–NGA strong-motion database to the equivalent amplitudes for B/C site conditions prior to their regression analysis.

The reference GMPEs used in this study are the Boore and Atkinson (2008) (hereafter, BA08) relations for shallow

Table 1

Regression Coefficients of Residuals of ENA B/C Data with Respect to BA08 Predictions (Equations 1 and 2)

Event (mm/dd/yyyy)	M	Location	0.2 Hz	0.5	1	2	5	10	PGA	PGV
03/01/1925	6.4	Charlevoix, Que.	−0.446	−0.691	−0.415	−0.196				
01/11/1935	6.2	Timiskaming, Que.	−0.317	−0.828	−0.755	−0.634				
09/05/1944	5.8	Cornwall, Ont.	−0.483	−0.741	−0.523	−1.158				
01/19/1982	4.3	Gaza, NH			−0.666	−0.608	−0.299	0.085		−0.197
10/07/1983	5.0	Goodnow, NY			−0.230	−0.184	−0.125	0.060		0.038
12/23/1985	6.8	Nahanni, NWT		−0.246	−0.214	−0.148	−0.095	0.076	−0.015	−0.147
01/31/1986	4.8	Painsesville, OH			−0.091	−0.172	0.003	0.248		−0.024
07/12/1986	4.5	OH			−0.475	−0.440	−0.412	−0.117		−0.220
11/23/1988	4.3	Saguenay FS			−0.550	−0.648	−0.324	−0.001		−0.237
11/25/1988	5.8	Saguenay, Que.		0.062	0.089	0.255	0.445	0.652	0.617	0.478
10/19/1990	4.7	Mt. Laurier, Que.			−0.417	−0.403	−0.048	0.198		0.095
11/6/1997	4.5	Cap Rouge, Que	−0.542	−0.745	−0.718	−0.626	−0.385	−0.006		
09/25/1998	4.5	Cleveland, OH	0.360	−0.004	−0.140	−0.160	0.038	0.336		
03/16/1999	4.5	St.Anne, Que.	−0.182	−0.543	−0.507	−0.626	−0.363	−0.088		
01/01/2000	4.7	Kipawa, Ont.	−0.228	−0.636	−0.686	−0.520	−0.282	−0.046		
01/26/2001	7.6	Bhuj, India		0.338	0.291	0.480	0.583	0.621	0.527	
04/20/2002	5.0	Au Sable Forks, NY	−0.097	−0.346	−0.227	−0.325	−0.060	0.097		
03/06/2005	5.0	Riv. De Loup, Que.	−0.505	−0.651	−0.525	−0.446	−0.212	0.026	0.020	−0.045
c_1			−1.07E-03	5.20E-04	5.56E-04	1.13E-03	1.44E-03	1.24E-03	1.20E-03	−1.11E-03
c_2			1.49E-06	3.76E-07	7.44E-07	6.98E-07	1.27E-06	1.99E-06	2.30E-06	1.89E-06
c_0	Average		−0.271	−0.419	−0.376	−0.364	−0.102	0.143	0.287	−0.029
	Standard deviation		0.284	0.379	0.288	0.368	0.290	0.234	0.331	0.223
c_{0w}	Average		−0.319	−0.379	−0.404	−0.356	−0.155	0.093	0.163	0.047

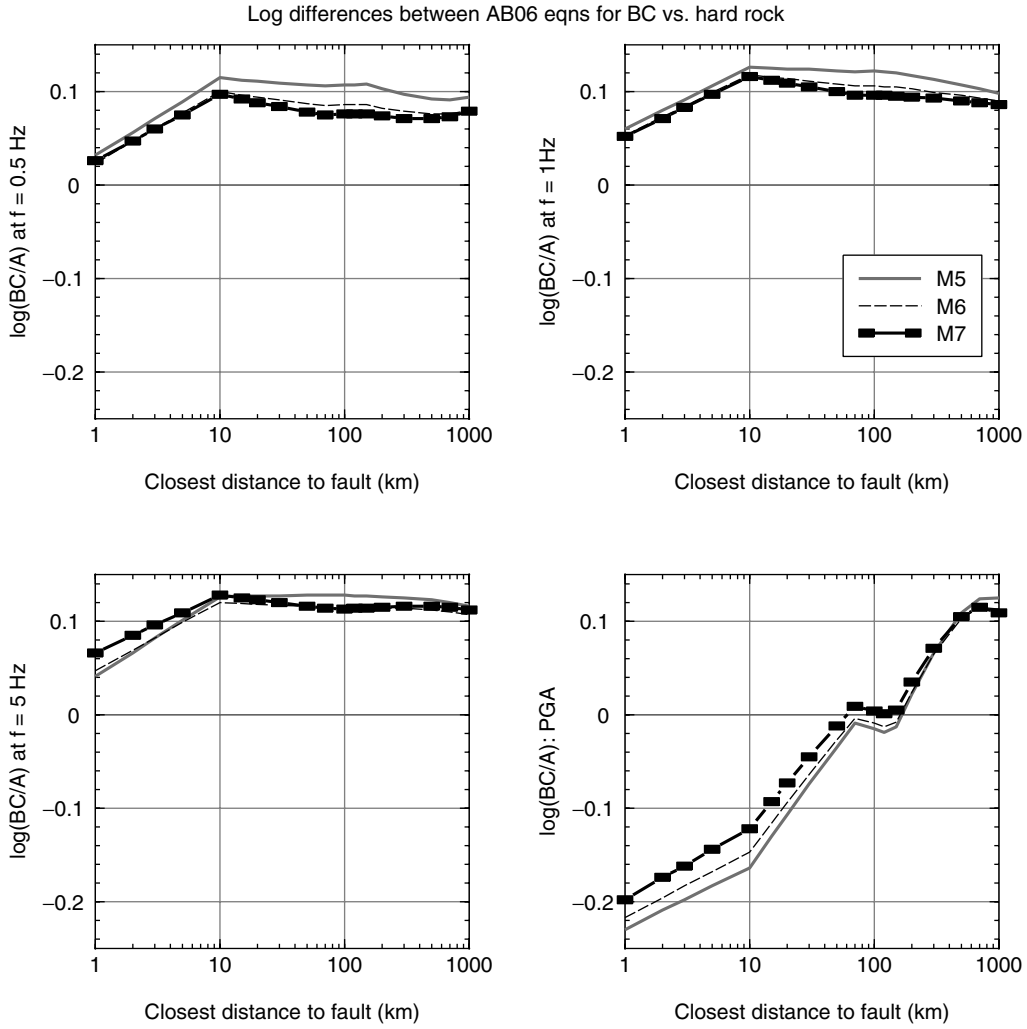

Figure 2. Ratio of predicted ground motions for sites on NEHRP B/C boundary site conditions ($V_{30} = 760$ m/sec) to motions on hard rock (NEHRP A, $V_{30} \sim 2000$ m/sec), according to GMPEs of Atkinson and Boore (2006), for PSA at 0.5, 1, and 5 Hz and PGA.

crustal earthquakes in active tectonic regions, developed in the PEER–NGA project and given for the average horizontal component (orientation-independent) for B/C site conditions. They are selected for their simplicity and convenience. Other reference relations from among the recent NGA equations could also be used.

The site-corrected ground motions for ENA (for B/C site conditions) are used with the reference BA08 equations (also for B/C conditions) to derive residuals, defined as the ratio of observed ENA motions to those predicted by BA08. Figure 3 plots the log (base 10) of the residuals for four selected frequencies as a function of Joyner–Boore distance (R_{jb}) (where log residual = log(observed ENA amplitude)− log(predicted amplitude from BA07). The residuals define a simple quadratic curve with distance (as shown by the line, determined in the next section). This curve can be fitted and used to define an adjustment factor to the BA08 GMPEs, the application of which will result in referenced empirical GMPEs for ENA. Other curve shapes were also considered, such as functions in $\log R$, or higher-order polynomials in R or $\log R$. A simple quadratic function in R provides the best fit; higher-order terms are not significant.

Fitting the Observed Residuals

In Figure 3, it is clear that the residuals from two events, the **M** 5.8 Saguenay earthquake and the **M** 7.6 Bhuj earthquake, show larger amplitudes relative to those of the other events, especially at higher frequencies (i.e., they are offset from the average curve). This suggests that to most accurately determine the overall shape of the residual trend with distance, we should allow an average-event-residual term to be defined to set the level for each event in the database, while retaining the quadratic shape with distance for all events. This is equivalent to stage 1 of a two-stage regression procedure (as per Joyner and Boore [1993]) and can be performed using a linear regression with dummy variables in order to extract the average-event-residual term of each earthquake and the distance scaling shape using

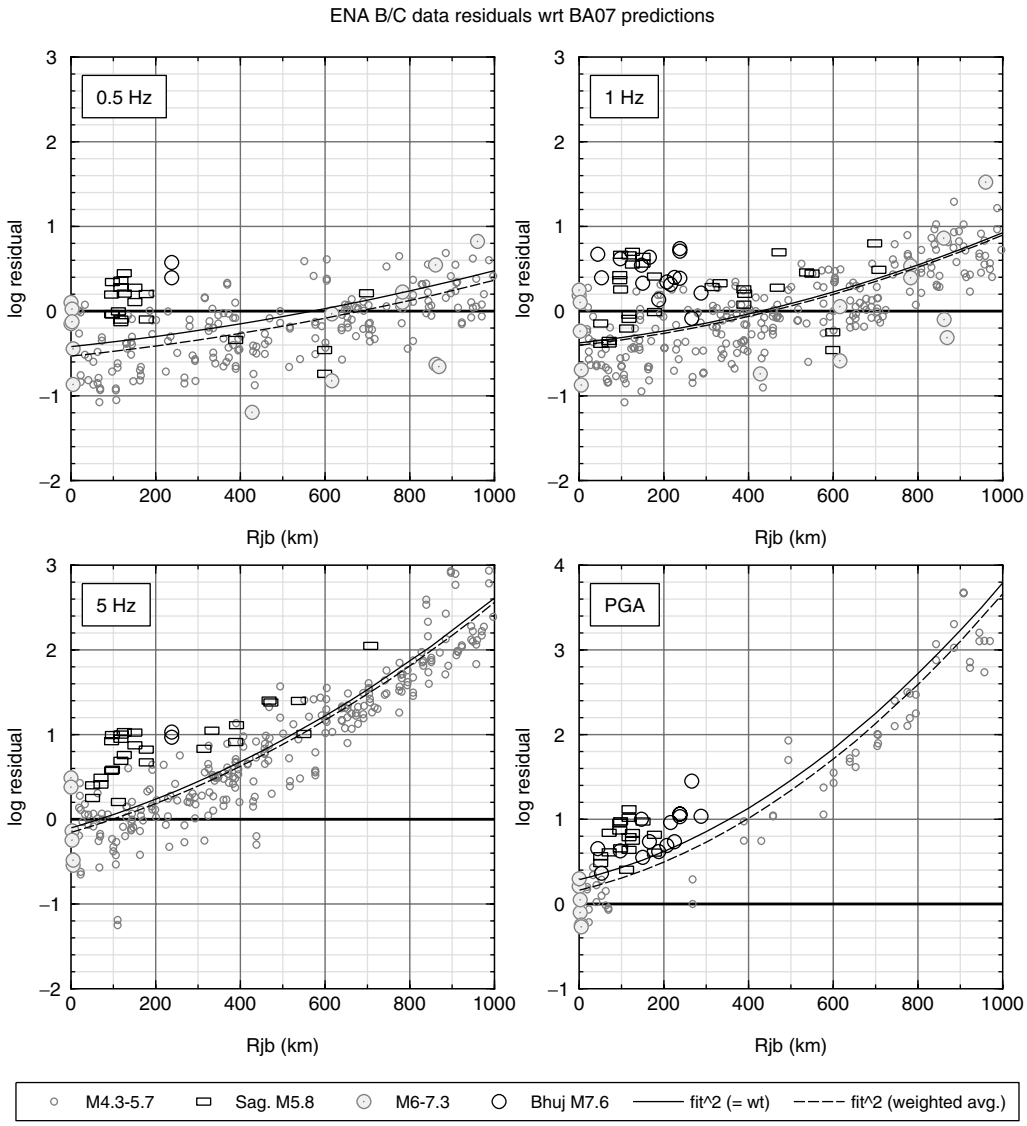

Figure 3. Residuals of ENA ground-motion data (B/C site conditions) with respect to the predictions of Boore and Atkinson (2008) for active tectonic regions, for PSA at 0.5, 1, and 5 and PGA. The solid line is the fit to the data for the average of the c_{0i} terms (equation 2, Table 1, c_0), while the dotted line is the fit using a weighted average of the c_{0i} terms (c_{ow}).

$$\log A_{ij} = \sum (c_0)_i E_i + c_1 (R_{\mathrm{jb}})_{ij} + c_2 (R_{\mathrm{jb}})_{ij}^2, \qquad (1)$$

where A_{ij} are the observed ground-motion amplitudes (earthquake i at station j) and $E_i = 1$ for earthquake i and 0 otherwise. Table 1 lists the constants (c_{0i}) for each event and the distance scaling terms c_1 and c_2. The c_{0i} terms are the average event residual for each event. Figure 4 plots the c_{0i} terms versus magnitude for four selected frequencies. There is no convincing trend with magnitude that persists across all frequencies. It may be argued that there is an apparent trend at 5 Hz, but the data are too sparse at larger magnitudes to be definitive, particularly in light of the questions regarding the applicability of the Bhuj datapoint at **M** 7.6. Overall, I conclude that the assumption of equivalent magnitude scaling in ENA and WNA was reasonable, though it cannot be entirely confirmed with existing data. Note the high average-event-residual terms for the Saguenay (**M** 5.8) and Bhuj (**M** 7.6) events. These events contribute heavily to the large standard deviation of average-event-residual terms (0.22–0.38 log units) as given in Table 1. They also contribute heavily to the adjustment factor for PGA, as there are few events in the ENA database for which PGA is available; the PGA adjustment factor is subject to greater uncertainty than that for the other ground-motion parameters for this reason.

The quadratic shape function in distance implicitly gives equal weight to all ground-motion observations, which is appropriate in a stage 1 regression versus distance. To define an average level correction (c_0) to use for ENA GMPEs, the overall level of the curve could be determined in two ways. One would be based on equal event weighting of the c_{0i} terms. This prevents undue bias in the equations from well-recorded

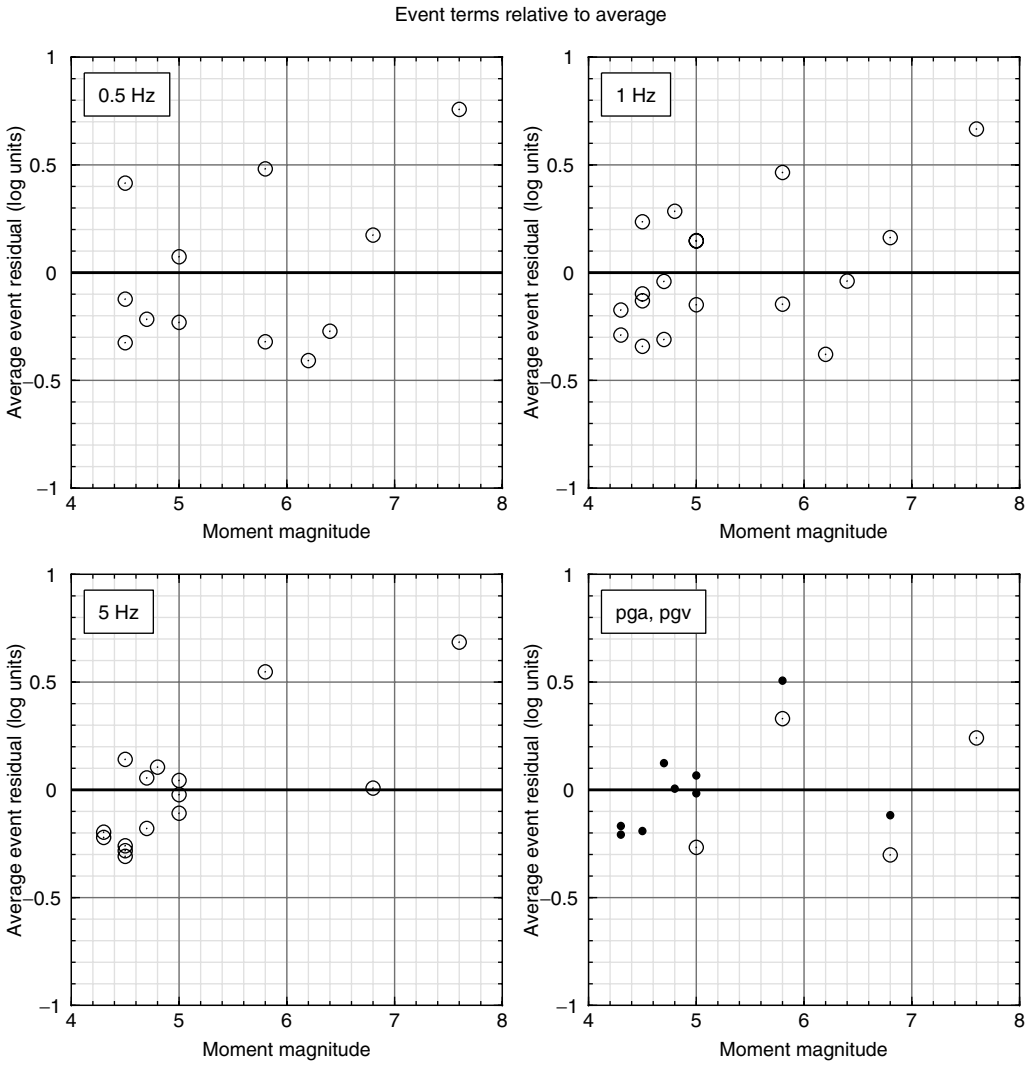

Figure 4. Average-event-residual terms (c_{0i}) from regression of residuals to equation (1), for PSA at 0.5, 1, and 5 Hz and PGA (open circles) and PGV (black dots).

events. On the other hand, to give more weight to events that are better-determined, the c_{0i} terms could be weighted by the number of observations for each event. This is equivalent to the approach taken in the two-stage regression procedure of Joyner and Boore (1993) and Boore and Atkinson (2008), in which equal-record weighting is used in a stage 1 regression to determine distance dependence and event terms, and a weighted fit is used in a stage 2 regression to determine magnitude scaling. In this case, however, there is no magnitude scaling in the stage 2 regression; we can use either a simple or a weighted average of the c_{0i} terms. Both of these approaches are taken, to explore the sensitivity of results.

An equal event weighting for the defined quadratic shape in distance is derived by calculating the average of the c_{0i} terms. This average ($= c_0$) and its standard deviation are listed in Table 1. All observations within 1000 km (approximately) are used in calculating the average event residuals; observations beyond 1000 km are too sparse to allow a robust determination of the residual function in distance and are therefore not used. The weighted average of the c_{0i} terms ($= c_{0w}$) is also calculated, where the weights are based on the number of observations per event at each frequency; these coefficients are also given in Table 1. The differences between the average-event-residual terms for these alternative approaches is relatively small (generally <0.1 log units); thus the weighting of the events is not a critical issue in determining the expected average ground motions.

The equal-event-weighted adjustment factor to predict ENA ground motions from the BA08 equations is given as

$$\log F = c_0 + c_1 R_{\mathrm{jb}} + c_2 R_{\mathrm{jb}}^2, \tag{2}$$

where c_0 may be replaced by c_{0w} if the weighted average is preferred. This function is plotted as the prediction line on Figure 3 (for both alternative values of c_0). To produce the referenced empirical ENA GMPEs, we simply add $\log F$ to the

log amplitude predictions of BA08, or equivalently multiply the BA08 predictions by the factor F. Thus the predicted ground-motion amplitude for ENA, for B/C conditions (Y_{ENA}), is just

$$Y_{\mathrm{ENA}} = FY_{\mathrm{BA07}}, \tag{3}$$

where Y_{BA07} is the predicted amplitude from BA08 for the corresponding magnitude, distance, and ground-motion parameter in active tectonic regions (e.g., PSA at a given frequency or PGA or PGV). Predictions can be made for an unknown fault mechanism or for the specific applicable mechanism if this is known. Predictions for other site conditions can be easily made using the site response factors of BA08, under the assumption that generic site response as a function of average shear-wave velocity is similar in ENA to that for active tectonic regions.

Evaluation of Results

The derived referenced empirical GMPEs for ENA (equations 2 and 3, with constants to derive F as given in Table 1) are plotted at four example frequencies, for **M** 5.5 and **M** 7.5, in Fig. 5, in comparison to the reference GMPEs for active tectonic regions of BA08 and the stochastic GMPEs for ENA of AB06 (all for B/C conditions). An unknown focal mechanism is assumed. The simple average c_0 values are used in these plots; if the weighted averages had been used, the referenced empirical GMPEs would plot very slightly lower (< 0.1 log units). Note that the AB06 distance metric

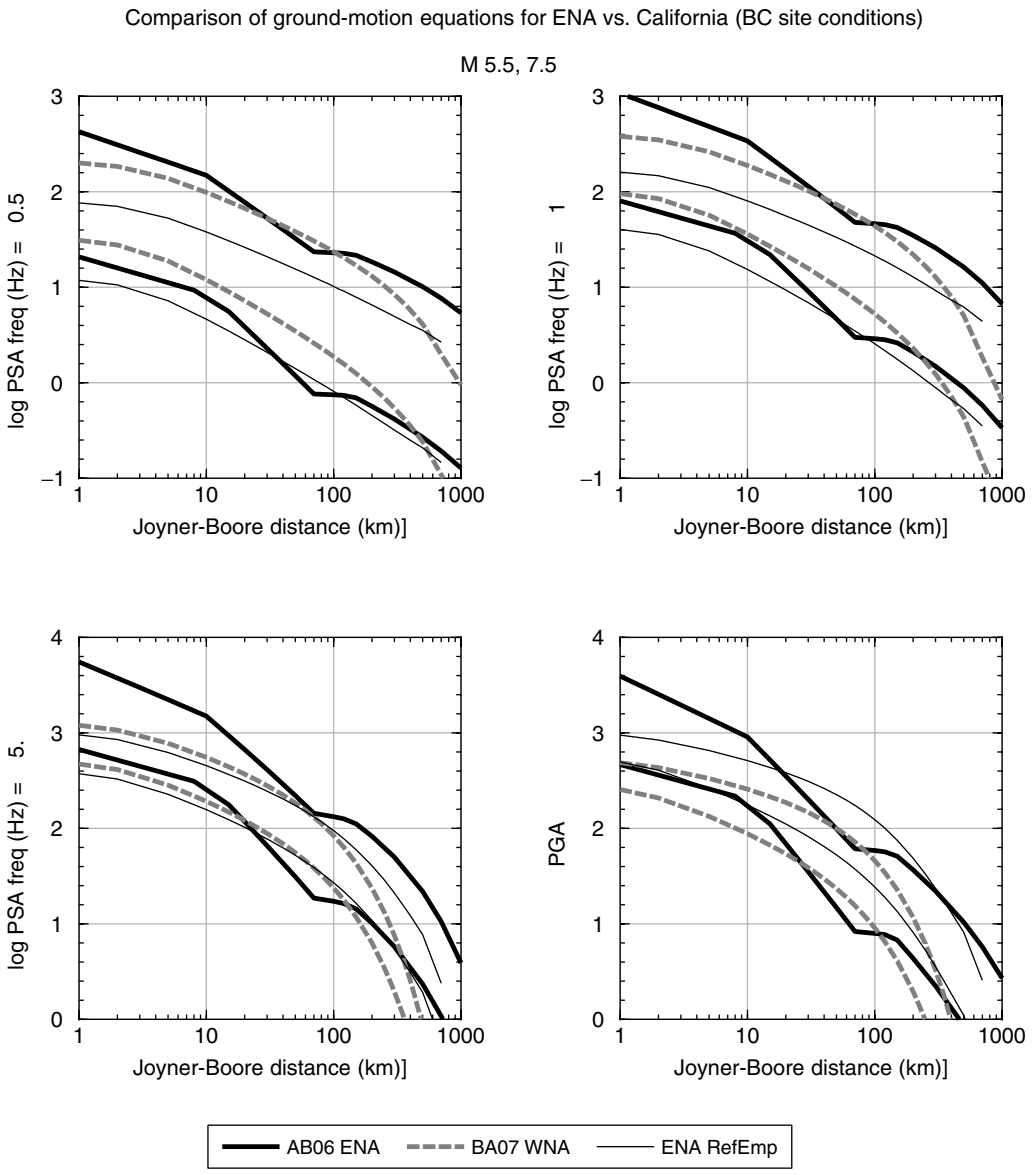

Figure 5. Comparison of GMPEs for ENA (B/C site conditions) according to Atkinson and Boore (2006) and this study (referenced empirical) with those for WNA according to Boore and Atkinson (2008), for **M** 5.5 and **M** 7.5, for PSA at 0.5, 1, and 5 Hz and PGA.

is the closest distance to fault. In this comparison, I assumed that the closest distance to the fault equals R_{jb} for **M** 7.5 (i.e., an **M** 7.5 would rupture the surface). For **M** 5.5 the closest distance to the fault should exceed R_{jb}, as small faults will not typically rupture the surface; to represent this in a simple way I assumed that the AB06 predictions for **M** 5.5 at a distance of 5 km from the fault correspond to the distance $R_{jb} = 1$ km. In other words, I assume that the top of the fault lies about 5-km beneath the surface; the exact distance is somewhat arbitrary as it depends on focal depth and could really be anywhere between $R_{jb} = 0$ and $R_{jb} = 10$. The reference empirical GMPEs are plotted only to a distance of 700 km, as they behave poorly beyond that distance—specifically, they begin to increase in amplitude with increasing distance, owing to the interaction of the distance function in BA08, which is not constrained by data at such distances, with the quadratic in distance of the adjustment function. Thus the reference empirical GMPEs should be truncated at a distance of 700 km.

In Figure 5 it appears that the reference empirical GMPEs are consistent with the AB06 GMPEs in a general sense in the distance range from 10 to 700 km, but they smooth out the pronounced trilinear shape function that is embedded in the AB06 model. The trilinear shape in AB06 models steep observed attenuation of spectral amplitudes in ENA at distances < 70 km, followed by a pronounced Moho bounce effect and a more gradual amplitude decay at larger distances. These trilinear effects are seen clearly in abundant ENA seismographic data compiled from smaller-magnitude earthquakes (**M** 3–5) (Atkinson, 2004); the response spectral data may be simply too sparse to show these effects, or the effects may be smoothed by the combination of events having observations in only some distance ranges.

Figure 5 also implies that the near-source amplitudes for large-magnitude events predicted by AB06 may be a gross overestimate, if we consider saturation effects observed near the source in active tectonic regions. Such effects are not modeled in the stochastic relations—although they could be modeled in future by introducing appropriate changes in the source scaling or near-source attenuation model to control amplitudes at near-fault distances (< 10 km). The differences in the reference empirical GMPEs and the AB06 GMPEs as seen in Figure 5 represent a good snapshot of epistemic uncertainty in ground-motion predictions in ENA.

In Figures 6–8, I compare the reference empirical GMPEs and the AB06 GMPEs to the ENA database (all for B/C conditions) for **M** 5, **M** 6, and **M** 7, respectively. In these comparisons, I assume that the top of the fault lies about 5-km beneath the surface for **M** 5 and **M** 6 (closest distance to the fault $= 5$ km maps to $R_{jb} = 1$ km). For **M** 7, two cases are shown: one for surface rupture and one for the top of the fault at 5-km depth. On balance, I conclude that the match of the data with the equations is comparable for the two alternative estimates. The distinct shape predicted by AB06 cannot be discerned in the data, but this may be a consequence of scatter, combined with the effects of plotting a one-unit magnitude range in each figure. Note that there are few to no data to distinguish between the very different predictions from the two sets of GMPEs at close distances and large magnitudes; note also the importance of the assumed closest distance of the fault from the surface in the AB06 predictions for large events (dashed versus solid AB06 predictions in Fig. 8). The major differences in predictions for large events at close distances may be a fair representation of epistemic uncertainty in this critical magnitude–distance range. In the stochastic model, the shape is driven by the attenuation model, which is not constrained by data at distances < 10 km; rather, the attenuation shape from 1 to 10 km is assumed in the stochastic model predictions to follow that observed from 10 to 70 km ($R^{-1.3}$ geometric attenuation). While this may be a reasonable assumption, it could potentially be modified by further data or modeling.

Examination of ENA Attenuation Shape

A major issue revealed by the comparison of the ground-motion predictions of the referenced empirical approach to that of AB06 involves the shape of the attenuation function, in particular, the attenuation rate at distances less than 70 km, and whether there is on average a pronounced Moho bounce effect. To examine the shape issue more closely, Figure 9 plots residuals from the ENA referenced empirical equations in the distance range to 200 km. Vertical bars show the hinge points of the AB06 GMPEs at 70 and 140 km. According to the AB06 shape, we would expect to see a negative trend in residuals in the first 70 km, followed by an offsetting positive trend from 70 to 140 km. There is a suggestion of such a trend in the figure (e.g., look at the data for events of **M** $<$ 5.8 at 1 Hz), but it is of marginal significance when examined statistically. Overall the trends are not statistically significant at $f \leq 2$ Hz and are barely significant at higher frequencies. In other words, when we fit a trend line to the data in specific distance ranges, such as 0–70 km or 70–140 km, the determined slope is not significantly different from zero. Furthermore, an attempt to draw out any such shape features by fitting the data of Figure 3 to a higher-order polynomial was unsuccessful (i.e., terms of higher order than the quadratic were not significant). As noted previously, the AB06 shape was derived from a larger database of smaller earthquakes (mostly **M** < 5), but based on Figure 9, this shape cannot be confirmed from the limited response spectral database for events of **M** ≥ 4.3. However, it should be kept in mind that the residuals at very short distances (< 10 km) are all from the Nahanni earthquake; because this event was recorded only at short distances, these residuals must be near zero by definition, and this may have the effect of biasing or obscuring apparent systematic trends of residuals with distance. It is also possible that there are attenuation trends that apply to the smaller-magnitude data that may not be applicable to larger events; this warrants further investigation in future studies.

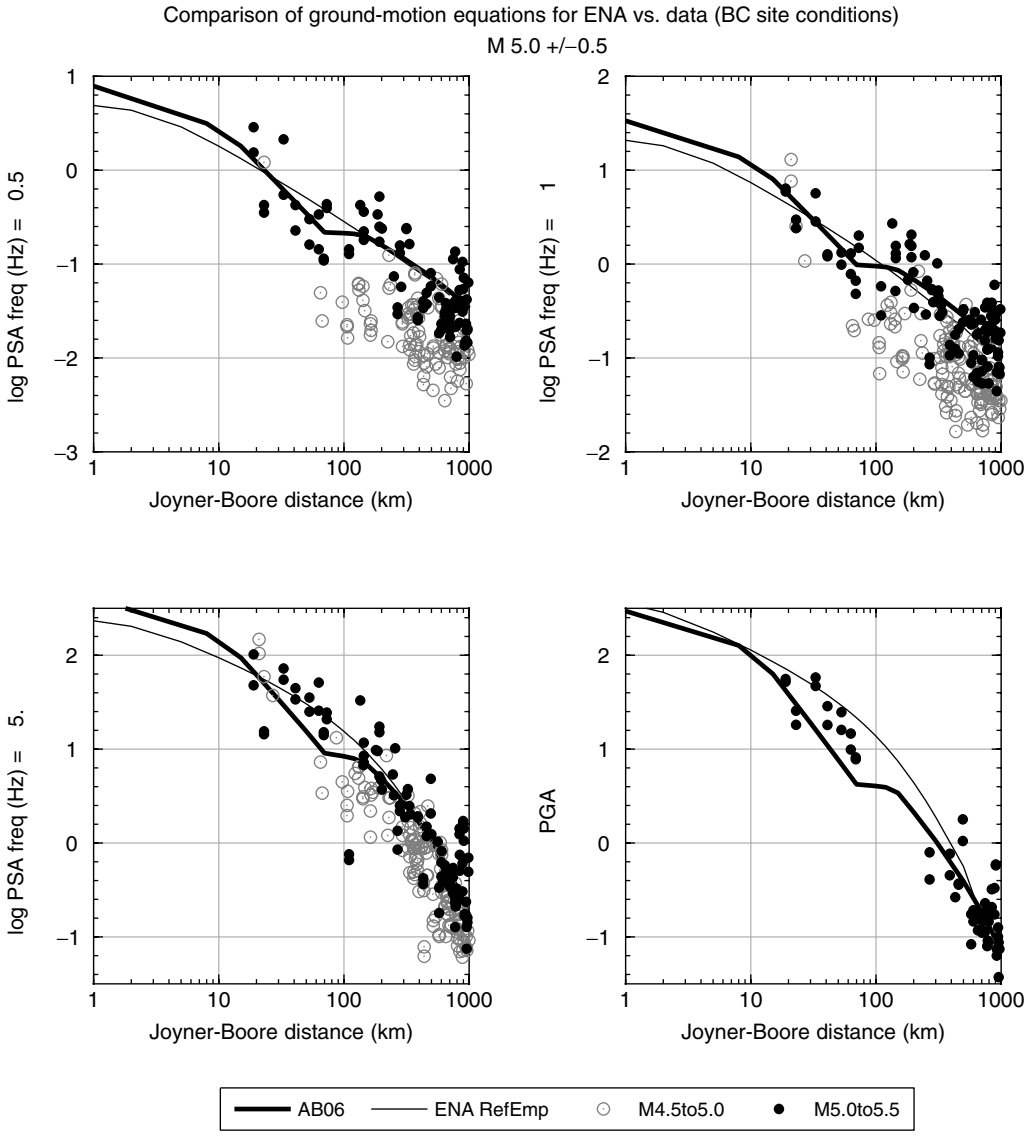

Figure 6. Comparison of ENA GMPEs of Atkinson and Boore (2006) and this study (referenced empirical) with ENA PSA data for 0.5, 1, and 5 Hz and PGA, all for B/C site conditions, for events of **M** 4.5–5.5.

Figure 10 shows the Fourier acceleration data used to derive the trilinear shape adopted by AB06, at one example frequency (1 Hz); it plots the Fourier amplitudes normalized by subtracting the event term (as determined by regression analysis in Atkinson, 2004) for each event. In these data, the steep decay ($R^{-1.3}$) in the AB06 model within 70 km is clearly supported, as is the pronounced Moho bounce. Only a handful of these data points, as indicated by the squares, are also present in the AB06 response spectral database, which focuses on larger events, and does not have good coverage with distance for individual events; for example, the response spectra data for Nahanni come only from near distances, while the Saguenay and Bhuj data come only from regional distances. It may be that the inadequacies of the data distribution in distance in the response spectral database are obscuring the actual attenuation shape.

Another potential source of information on attenuation shape, especially within the critical distance range of < 100 km, for which response spectral data are sparse, is intensity data. Atkinson and Wald (2007) demonstrate that these data contain important information on ground-motion characteristics and may be used to discern differences between eastern and western attenuation rates and relative ground-motion amplitudes. Atkinson and Wald (2007) derive a general attenuation function for the two regions based on intensity data collected by the Did You Feel It? web-based questionnaire of the U.S. Geological Survey (http://earthquake.usgs.gov/dyfi/ [last accessed March 2008]), which covers distances out to 1000 km. However the Atkinson and Wald attenuation function does not directly address regional differences within the critical distance range of contention here (from 10 to 70 km), as the adopted functional

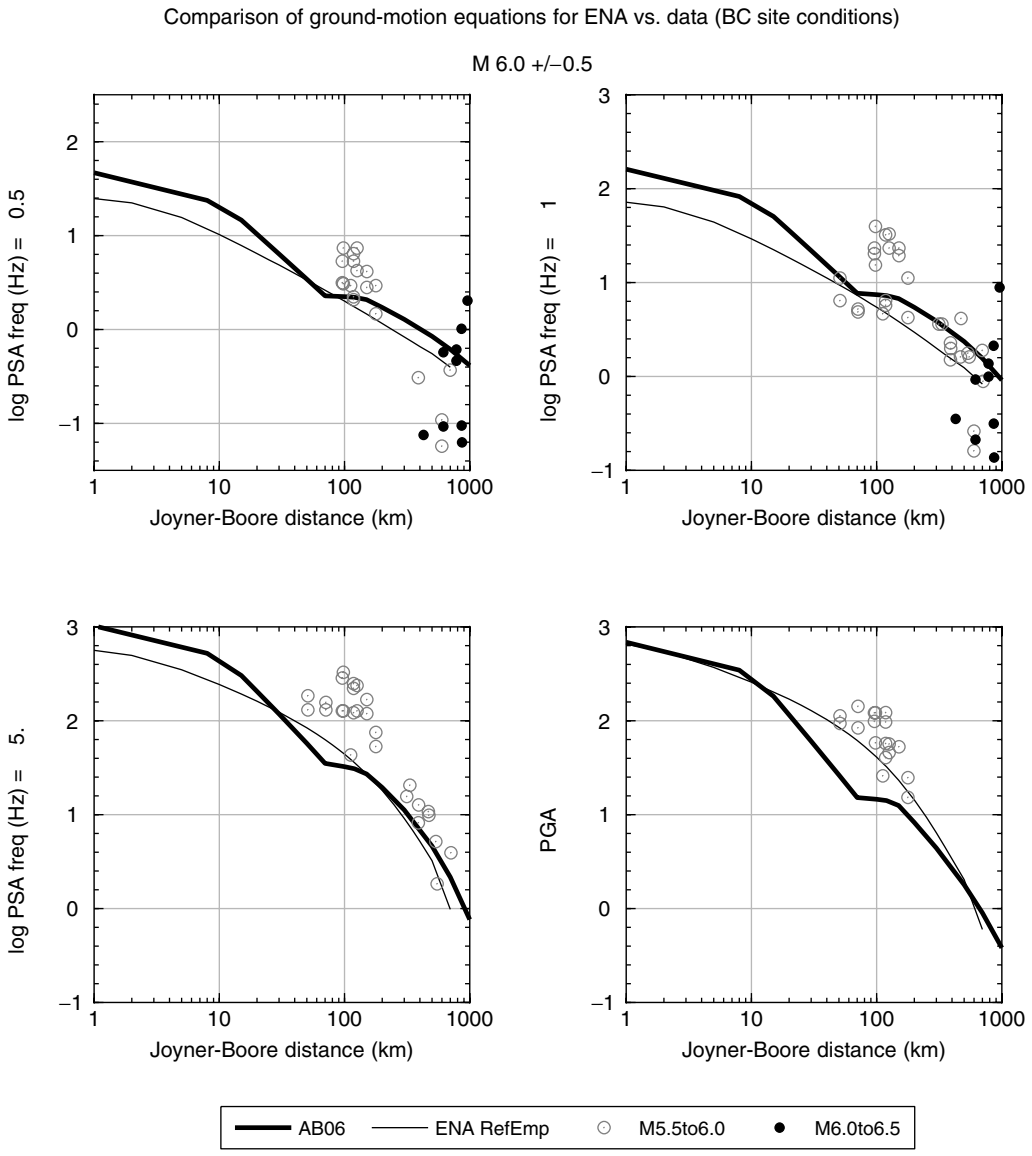

Figure 7. Comparison of ENA GMPEs of Atkinson and Boore (2006) and this study (referenced empirical) with ENA PSA data for 0.5, 1, and 5 Hz and PGA, all for B/C site conditions, for events of **M** 5.5–6.5. Note the data are dominated by observations from a single event, the Saguenay earthquake.

form smooths attenuation trends over a broader distance range. The relative near-distance attenuation rates in the two regions can be examined by analyzing the data compiled by Atkinson and Wald (2007) more closely. There are hundreds of observations within 70 km for each significant event in the database, enabling the attenuation slope of intensity in this distance range to be determined with reasonable confidence for individual events.

I use the intensity database compiled by Atkinson and Wald to determine the attenuation of intensity from an epicentral distance of approximately 5 km to approximately 65 km, for well-reported events of **M** 2.8–5.2 in both ENA and California. I consider only events for which there are ≥ 5 observations to constrain the average 5-km intensity value and ≥ 5 observations to constrain the average 65-km intensity value; there are 18 such events in ENA and 74 such events in California. The 5-km average comes from weighting, according to number of observations, the binned average intensity values at distances at 8 km or nearer, where distance bins ± 0.3 log units wide were used. The central values of the distance bins included in the 5-km average are 1, 2, 4, and 8 km (log distance values of 0., 0.3, 0.6, and 0.9); this average thus includes all data within 11 km of the epicenter (e.g., the largest distance is $8 \times 10^{0.15}$). The weighted central distance of the data included in the 5-km average is actually 6 km for the California dataset and 7 km for the ENA dataset. The 65-km value is the average binned intensity at 63 km (the log distance is 1.8) in both ENA and California, where

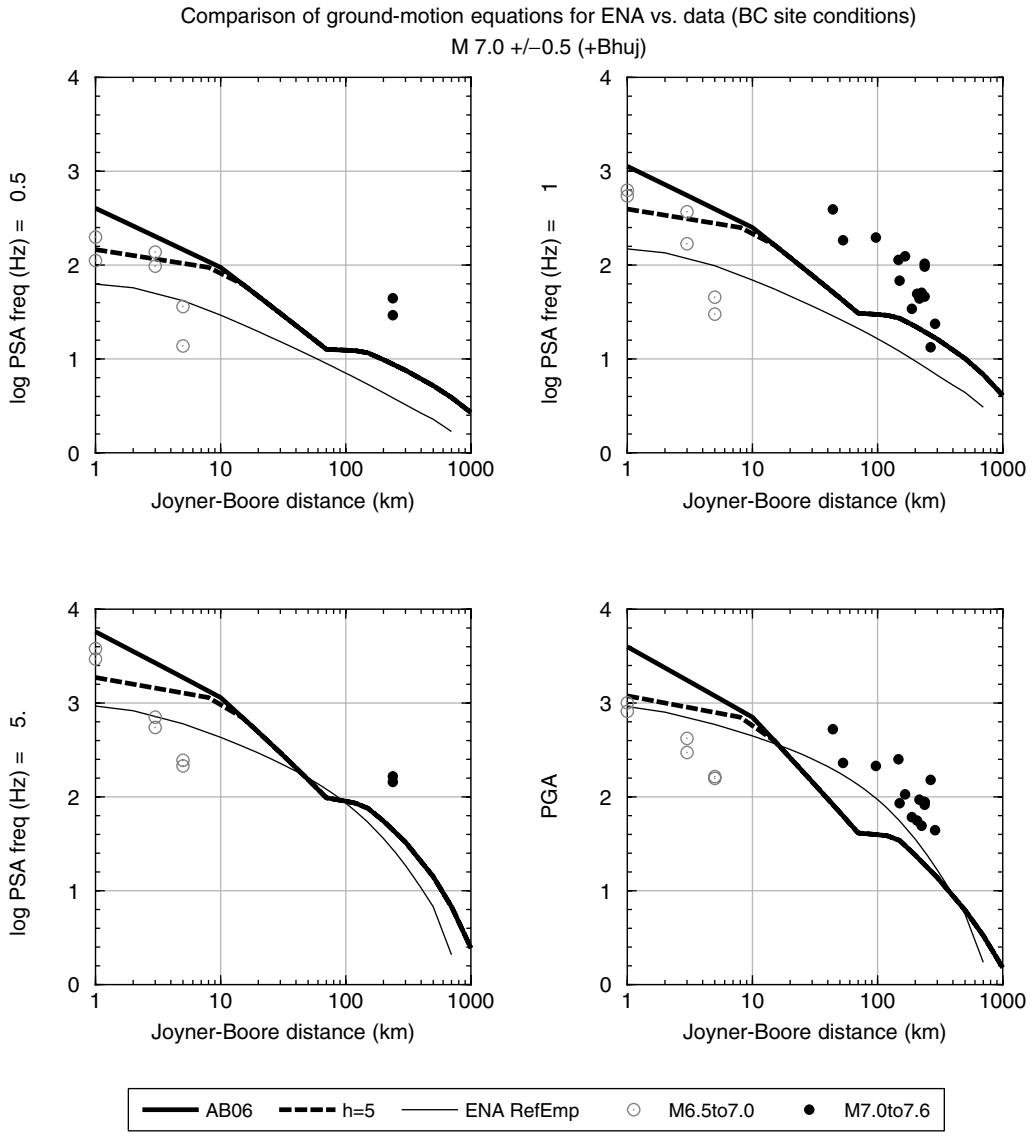

Figure 8. Comparison of ENA GMPEs of Atkinson and Boore (2006) and this study (referenced empirical) with ENA PSA data for 0.5, 1, and 5 Hz and PGA, all for B/C site conditions, for events of **M** 6.5–7.6. Note the observations at $\mathbf{M} > 7$ come entirely from the Bhuj earthquake.

the distance bin used in this case is 0.1 log units in width; thus the distance range of observations included in the bin is 56–71 km (the log distance is 1.75–1.85). It is acknowledged that the distances that apply to the determined did-you-feel-it (DYFI) intensity values cannot be precisely determined, as the DYFI measurements are averaged over a zip code area, where the distance is that to the center of the zip code. Uncertainty in the true distances that apply to the DYFI data averages may obscure attenuation behavior, especially near the source. This is an inherent limitation of using the DYFI data to infer attenuation.

Figure 11 plots the difference in intensity from the 5-km-distance bin to the 65-km-distance bin (I_5–I_{65}) for each event against magnitude. The mean drop in intensity from about 5 to 65 km, and its standard error, is 1.69 ± 0.19 in ENA versus 1.42 ± 0.07 in California. Thus the intensity data suggest a steeper attenuation of motion in ENA within 70 km than for California. However, the standard deviation of observations as observable on Figure 11 is large (0.8 for ENA, 0.6 for California), and the ENA average is significantly influenced by a few events with particularly large values of I_5–I_{65}. There are many ENA values for which I_5–I_{65} is comparable to the California average. The observed differences are only marginally significant.

On balance, I conclude that the weight of evidence suggests that the ENA attenuation within the first 70 km is steeper than that in active tectonic regions such as California. However, the fact remains that this is not directly observable in the ENA response spectral database that is available to date. Furthermore, the intensity data are somewhat ambigu-

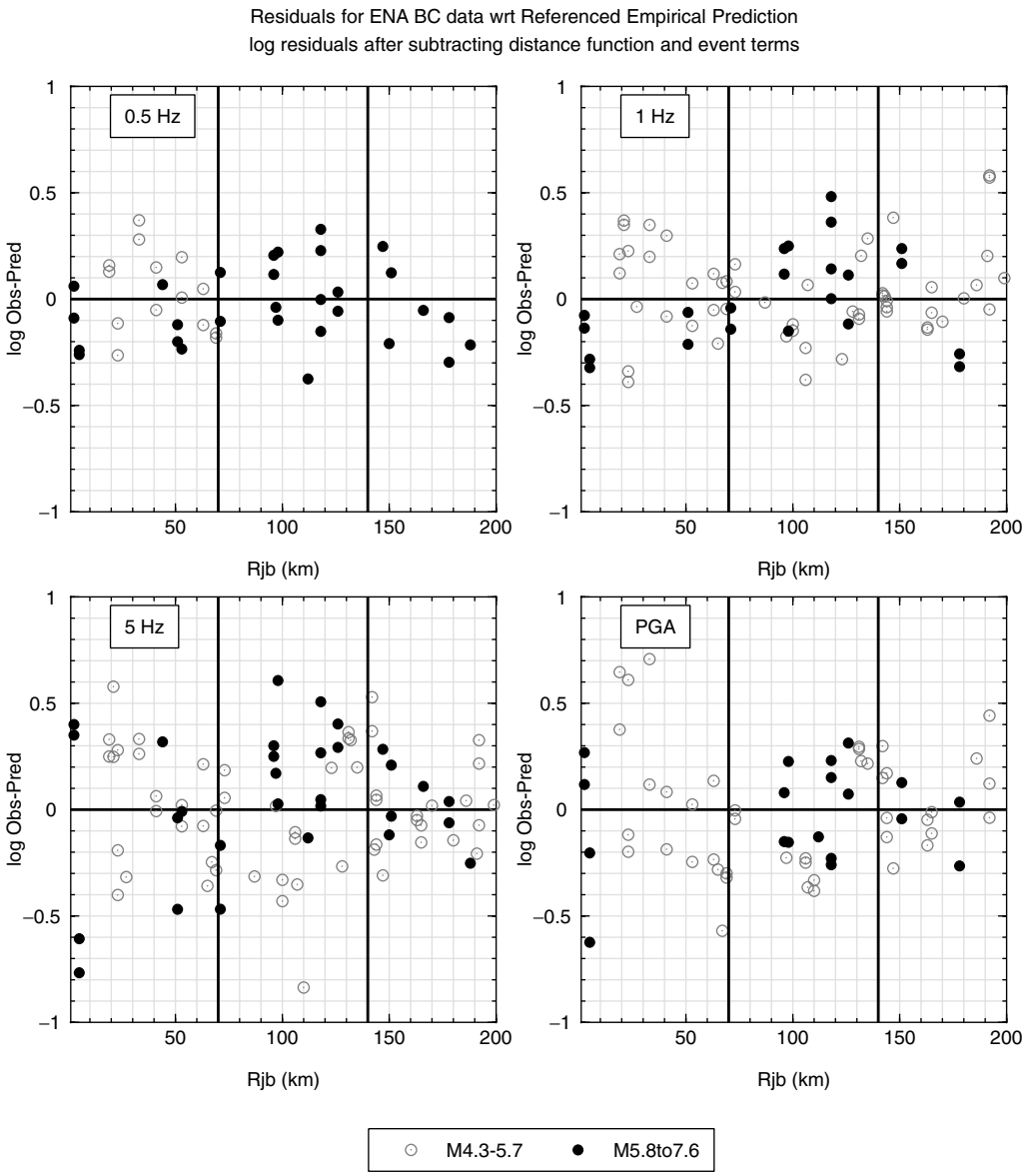

Figure 9. Residuals for ENA data with respect to the referenced empirical GMPEs of this study, for distances < 200 km, for PSA at 0.5, 1, and 5 Hz and PGA.

ous, indicating a higher average attenuation in ENA within 70 km, but not showing this consistently for all events. Whether this shape will be clearly observed in future earthquakes remains an area of epistemic uncertainty.

Discussion and Conclusions

The referenced empirical approach provides GMPEs for ENA that are in agreement with regional strong-motion observations, while being constrained to follow the overall scaling behavior of ground motion that is observed in better-instrumented active tectonic regions. They are presented as an alternative to the commonly used stochastic ground-motion relations for ENA, in order to partially express epistemic uncertainty in ENA GMPEs. The most important source of uncertainty is whether predicted attenuation shape effects in the AB06 model are in fact observable on average, or whether they are smoothed out as implied (but not proven) by the referenced empirical approach. Amplitude differences due to shape average out over a large distance range (10–500 km) but can be greater than a factor of 2 in some distance ranges, particularly from 40 to 150 km.

There are significant limitations to the referenced empirical approach as implemented here, which could map into sources of error. One potential problem is that fault dimensions for a given magnitude are likely to be smaller in ENA than in active tectonic regions (Somerville *et al.*, 2001). This could affect the relative relationships of near-source amplitudes and distance decay in ENA as compared to those implicit in the reference GMPEs for more active regions.

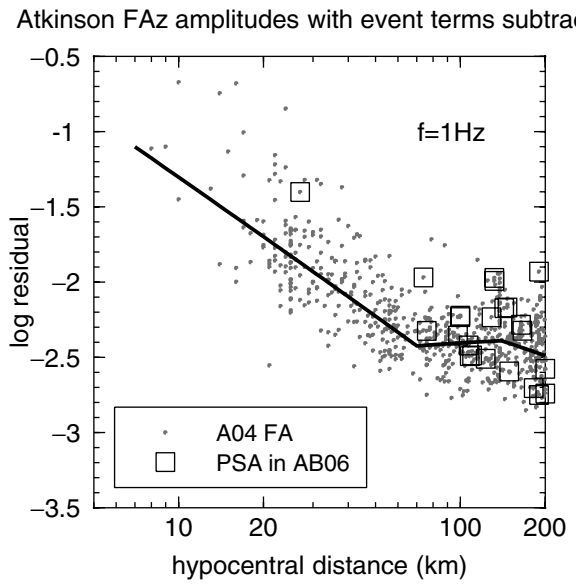

Figure 10. Fourier acceleration data at 1 Hz used in determining the attenuation shape of Atkinson (2004) (as adopted in AB06), normalized by subtracting the event term for each event (dots). Line shows trilinear attenuation function of Atkinson (2004). Squares show data points that are also in the AB06 response spectral database.

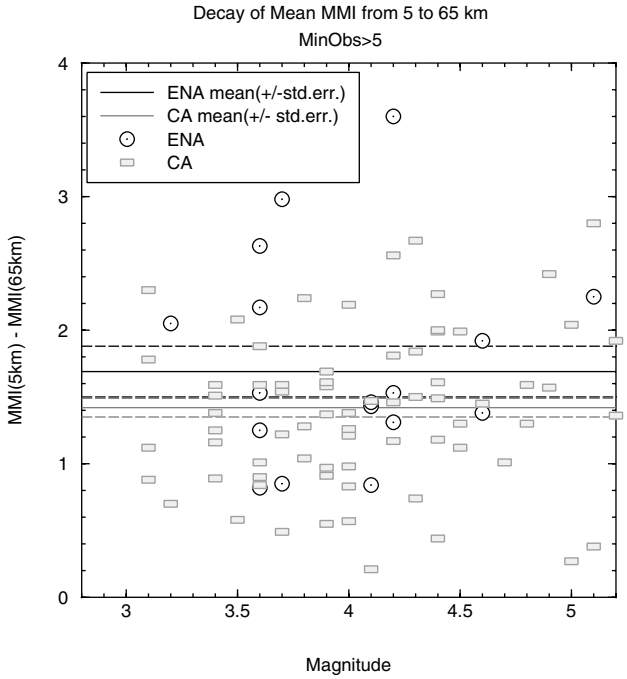

Figure 11. Difference in mean intensity of observations at about 5-km-epicentral distance (distance range of observations 1–11 km) and mean intensity at about 65-km distance (distance range of observations 56–71 km), for events having ≥5 observations within each distance bin. Data are from Atkinson and Wald (2007). Solid lines show mean values in ENA and California; dotted lines are standard error of means.

Furthermore, there may be a greater tendency of events in active regions to rupture the surface, which could affect ground motions—although the conclusion of Boore and Atkinson (2008) was that the significance of this effect is not clearly established.

Perhaps most significant among the limitations of the referenced empirical approach is that the ENA response spectra database may be simply insufficient to properly model all the effects that are important. In particular, the difference in the implied attenuation shape between the referenced empirical approach and AB06 is troubling. I believe that the AB06 attenuation model is more robust, as it takes advantage of the great wealth of low-magnitude data that are available; I would therefore give it more weight. The AB06 attenuation model is partially supported by intensity data, in terms of implied differences between ENA and California attenuation on average in the first 70 km. But the issue of relative differences in both amplitude levels and attenuation in ENA versus active tectonic regions clearly requires further study, preferably by directly comparing response spectra between regions in the same magnitude-distance range.

The referenced empirical GMPEs imply that ENA motions (on B/C site conditions) are actually less than those for active tectonic regions such as California at distances from 10 to 70 km at frequencies ≤2 Hz and are similar to California motions at 5 Hz. The ENA motions exceed the California motions only for higher frequencies and PGA and for distances >100 km. (Note: A similar conclusion is reached by comparing AB06 and BA08). This prediction is not in accord with the differences in intensity levels observed in ENA as compared to those in California (Atkinson and Wald, 2007). It is observed that intensity levels in this distance range are clearly about 1 unit higher in ENA than in California, for events of the same magnitude and distance (Atkinson and Wald, 2007). Based on typical empirical relationships between intensity and ground-motion amplitude (e.g., Atkinson and Kaka, 2007), this would imply larger ground-motion amplitudes in ENA than in California by at least a factor of 1.6. The intensity observations are consistent with the referenced empirical GMPEs (or with AB06) only if intensity is controlled by very high-frequency motions (10 Hz and PGA) or if site conditions are systematically different between the two regions, leading to larger average amplification of intensity observations in ENA. Thus there is an unresolved apparent inconsistency between intensity observations and ENA GMPEs.

Differences between the referenced empirical GMPEs of this study and the stochastic GMPEs of Atkinson and Boore (2006), along with inconsistencies between both of these studies and inferences based on intensity observations, suggests that uncertainty in median ENA GMPEs is at least a factor of 1.5–2 for $\mathbf{M} \geq 5$ at distances from 10 to 70 km. Uncertainty is greater than a factor of 2 for large events ($\mathbf{M} \geq 7$) at distances within 10 km of the source. It may be that saturation effects not modeled in the stochastic predictions, but inferred from observations in other regions, cause overestimation of near-source amplitudes from large events in AB06.

On the other hand, these saturation effects cannot be directly verified in ENA data. It is beyond the scope of this article to fully define the epistemic uncertainty in ENA ground-motion relations, but the factors estimated here provide some initial estimates of how large it may be—probably a factor of 1.5–2 at best.

Acknowledgments

Helpful comments and suggestions received from David Boore, Ken Campbell, Art Frankel, and Chris Cramer on a draft of this manuscript are gratefully acknowledged. Review comments of Jeff Munsey and an anonymous reviewer were very helpful in further improvements to the article.

References

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

Atkinson, G. (1993). Source spectra for earthquakes in eastern North America, *Bull. Seismol. Soc. Am.* **83,** 1778–1798.

Atkinson, G. (2004). Empirical attenuation of ground motion spectral amplitudes in southeastern Canada and the northeastern United States, *Bull. Seismol. Soc. Am.* **94,** 1079–1095.

Atkinson, G., and D. Boore (2006). Ground motion prediction equations for earthquakes in eastern North America, *Bull. Seismol. Soc. Am.* **96,** 2181–2205.

Atkinson, G., and S. Kaka (2007). Relationships between felt intensity and instrumental ground motions for earthquakes in the central United States and California, *Bull. Seismol. Soc. Am.* **97,** 497–510.

Atkinson, G., and D. Wald (2007). Modified Mercalli intensity: a surprisingly good measure of ground motion, *Seism. Res. Lett* **78,** 362–368.

Boore, D. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bull. Seismol. Soc. Am.* **73,** 1865–1894.

Boore, D., and G. Atkinson (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped SA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra* (in press): available from the Pacific Earthquake Engineering Research Center–Next Generation Attenuation project, http://peer.berkelely.edu (last accessed March 2008).

Boore, D., and W. Joyner (1997). Site amplifications for generic rock sites, *Bull. Seismol. Soc. Am.* **87,** 327–341.

Boore, D., W. Joyner, and T. Fumal (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work, *Seism. Res. Lett.* **68,** 128–153.

Campbell, K. (2003). Prediction of strong-ground motion using the hybrid-empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America, *Bull. Seismol. Soc. Am.* **93,** 1012–1033.

Hanks, T., and R. McGuire (1981). The character of high-frequency strong ground motion, *Bull. Seismol. Soc. Am.* **71,** 2071–2095.

Joyner, W., and D. Boore (1993). Methods for regression analysis of strong motion data, *Bull. Seismol. Soc. Am.* **83,** 469–487.

Kanamori, H., and D. Anderson (1975). Theoretical basis of some empirical relations in seismology, *Bull. Seismol. Soc. Am.* **65,** 1073–1095.

Power, M., B. Chiou, N. Abrahamson, and C. Roblee (2006). The "Next Generation of Ground Motion Attenuation Models" (NGA) project: an overview, in *Proceedings of the Eighth National Conference on Earthquake Engineering*, paper no. 2022.

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

Somerville, P., N. Collins, N. Abrahamson, R. Graves, and C. Saikia (2001). Ground motion attenuation relations for the central and eastern United States, a report to the U.S. Geological Survey, NEHRP External Research Program, award no. 99-HQ-GR-0098.

Somerville, P., J. McLaren, L. Lefevre, R. Burger, and D. Helmberger (1987). Comparison of source scaling relations of eastern and western North American earthquakes, *Bull. Seismiol. Soc. Am.* **77,** 322–346.

Department of Earth Sciences
University of Western Ontario
Arnprior, Ontario K7S 3T2 Canada
gmatkinson@aol.com

Manuscript received 7 August 2007