Bulletin of the Seismological Society of America, 90, 6B, pp. S209–S221, December 2000

A Modified Ground-Motion Attenuation Relationship for Southern California that Accounts for Detailed Site Classification and a Basin-Depth Effect

by Edward H. Field*

Abstract The attenuation relationship presented by Boore *et al.* (1997) has been evaluated and customized with respect to southern California strong-motion data (for peak ground acceleration (PGA) and 0.3-, 1.0-, and 3.0-sec period spectral acceleration). This study was motivated by the recent availability of a new site-classification map by Wills *et al.* (2000), which distinguishes seven different site categories for California based on the 1994 NEHRP classification. With few exceptions, each of the five site types represented in the southern California strong-motion database exhibit distinct amplification factors, supporting use of the Wills *et al.* (2000) map for microzonation purposes. Following other studies, a basin-depth term was also found to be significant and therefore added to the relationship. Sites near the center of the LA Basin exhibit shaking levels up to a factor of 2 greater, on average, than otherwise equivalent sites near the edge. Relative to Boore *et al.* (1997), the other primary difference here is that PGA exhibits less variation among the Wills *et al.* (2000) site types. In fact, the PGA amplification implied by the basin-depth effect is greater than that implied by site classification. The model does not explicitly account for nonlinear sediment effects, which, if important, will most likely influence rock-site PGA predictions the most. Evidence for a magnitude-dependent variability, or prediction uncertainty, is also found and included as an option.

Introduction

This study is part of the Southern California Earthquake Center (SCEC) Phase III effort to determine the extent to which site effects can be accounted for in probabilistic seismic hazard analysis [see SCEC Phase III working group (2000) for an overview]. One question addressed by the working group is whether a site-classification scheme more detailed than rock versus soil is warranted. Two articles, one by Steidl (2000) and one by Lee and Anderson (2000), evaluated whether the detailed geology map of Tinsley and Fumal (1985), which divides the Quaternary into eight subunits, could be reliably used for microzonation purposes. By making a statistical comparison of empirical amplification factors for different categories, neither study could find compelling support for a subclassification beyond rock versus soil. Both studies did, however, see a significant basin-depth effect (discussed more later in this article).

Following completion of the studies by Steidl (2000) and Lee and Anderson (2000), a new site-classification map was made available by Wills *et al.* (2000). The new map uses the 1994 NEHRP classification scheme adopted in the 1997 Uniform Building Code, which defines site types A, B, C, D, and E based on the average shear-wave velocity in the upper 30 meters ($\bar{V}_s$). The map was compiled for the entire state of California by correlating detailed geology, in 1:250,000 scale maps, with *in situ* shear-wave velocity measurements. Because many geological units fell near boundaries of the NEHRP categories, Wills *et al.* (2000) included intermediate categories as well (BC, CD, and DE). The complete list of site types and associated shear-wave velocities is given in Table 1.

The purpose of this study is to evaluate whether microzonation based on the Wills *et al.* (2000) map is supported by southern California strong-motion data. The investigation benefits extensively from the work of Steidl (2000) and Lee and Anderson (2000), and in many ways mimics their analyses. As in those studies, amplification factors are estimated from differences between observed and predicted ground-motion levels (residuals). The analysis is conducted for peak acceleration (PGA) and 5% damped response spectral acceleration (SA) at 0.3-, 1.0-, and 3.0-sec periods. The Boore *et al.* (1997) attenuation relationship (hereafter referred to as

*Present address: U.S. Geological Survey, 525 South Wilson Ave., Pasadena, California 91106, email: *field@usgs.gov*

Table 1

The site-type categories from Wills et al. (2000), the 1994 NEHPR range of 30-meter shear wave velocities, and the shear-wave velocity assigned to each category in this study.

Wills *et al.* Category*	V_s (m/sec) NEHRP	V_s (m/sec) Used Here
	760–1,500	1000
	NA	760
	360–760	560
	NA	360
	180–360	270
	NA	180
	<180	NA

*There are no NEHRP A categories in California, and no E categories in southern California

BJF97) is used for the predictions. This differs from Lee and Anderson (2000) who used the Abrahamson and Silva (1997) attenuation relationship, and Steidl who used the Sadigh *et al.* (1997) relationship for rock. BJF97 is applied here because the site effect is parameterized explicitly with V_s (the average 30-m shear-wave velocity), making it the most natural choice with respect to the Wills *et al.* (2000) classification scheme.

The observed data are from the SCEC Phase III strong-motion database (Steidl and Lee, 2000). The earthquakes are listed in Table 2, along with the number of observations for each site type. In addition to comparing data to the default BJF97 predictions, the attenuation-relationship parameters have also been optimized with respect to the southern California observations. Finally, based on the work of Trifunac and Lee (e.g., 1978), Campbell (e.g., 1997), Lee and Trifunac (1995), Steidl (2000), and Lee and Anderson (2000), a basin-depth effect is evaluated and found to be significant.

There are varying degrees of statistical sophistication that can be applied in developing an attenuation relationship, and it is acknowledged that some studies have gone beyond that represented here (some have been less rigorous). The level chosen for this study is believed to be commensurate with presently available data and with the inherent limitations of any model that attempts to quantify earthquake ground motion with only a few parameters. Attenuation relationships are constantly being updated, particularly now given the recent spate of $M \geq 7$ events. These include the Kocaeli earthquake in Turkey (e.g., Toksöz et al., 1999), the Hector Mine earthquake in southern California (e.g., Scientists from the United States Geological Survey, the Southern California Earthquake Center, and California Division of Mines and Geology, 2000), and the Chi-Chi earthquake in Taiwan (Shin *et al.*, 2000). It is hoped that this study will guide future developments, and provide an interim solution until updated models are available. In terms of probabilistic seismic hazard analysis, the implications of this and other attenuation relationships is evaluated in another SCEC Phase III article (Field and Petersen, 2000).

Analysis Methodology

The BJF97 attenuation relationship has the following form:

$$\mu(M, r_{jb}, V_s) = b_1 + b_2(M - 6) + b_3(M - 6)^2 + b_5 \ln((r_{jb}^2 + h^2)^{0.5}) + b_v \ln(V_s/V_a) \quad (1)$$

where μ is the predicted natural logarithm of the ground-motion parameter (e.g., ln(PGA)), M is moment magnitude, r_{jb} is the closest distance to the vertical projection of rupture (in km), and V_s is the average 30-meter shear-wave velocity at the site (m/sec). The other parameters (b_1, b_2, b_3, b_5, h, b_v, and V_a) are solved for in the regression with empirical data, with b_1 optionally being separated into strike-slip (b_{1ss}) versus reverse-slip (b_{1rs}) components. V_a is an arbitrary reference site velocity (arbitrary because it trades off with b_1), and h is a fictitious depth needed to prevent zero distance. Note that the site-effect term is linear in that it is independent of magnitude, distance, and ground-motion level.

The regression is carried out in the context of the random-effects model described by Abrahamson and Youngs (1992). That is, not only is there a random component between observations of a given earthquake, but there is also a random component between earthquakes (for example, to account for variations in dynamic stress drop). This is written as:

$$y_{ij} = \mu_{ij} + \eta_i + \varepsilon_{ij} \quad (2)$$

where y_{ij} is the natural logarithm of the observed ground-motion parameter for event i and site j, μ_{ij} is the predicted value given by equation (1), η_i is the random effect for the i^{th} event (inter-event term), and ε_{ij} represents the intraevent variation. The terms η_i and ε_{ij} are assumed to be independent and normally distributed with variances τ^2 and σ^2, respectively.

Recall that the goal is to infer site effects from differences (or residuals) between observed and predicted values ($y_{ij} - \mu_{ij}$). That is, if any site type has a statistically significant bias with respect to the predictions, then we are warranted in applying a site correction. Note in Table 2 that the data are "unbalanced" in that each earthquake has a different number of recordings for each site type. As such, it is important to correct the observations for the interevent terms (η_i) before inferring site effects from residuals. Otherwise we risk mapping the ground motion from any unusual earthquake into site effects at the particular locations where event was recorded. Again, this is done in the context of the random-effects methodology outlined by Abrahamson and Youngs (1992). The exact procedure used here is as follows:

Step 1: Choose a trial value of the fictitious depth h, thereby linearizing equation (1), and set all interevent terms equal to zero ($\eta_i = 0$).

Step 2: After correcting the observed data for the present

Table 2
The Earthquakes in the SCEC Phase III Database, Their Moment Magnitude (mag), and the Number of Observations Available for each Site Type*

	Event Date and Name	mag	B	BC	C	CD	D	DE
			0	0	0	0		0
			0	0	0	0		0
			0	0	0	0		0
			0	0	0	0		0
			0	0	0	0	2	0
			0	0	0	0	1	0
			0	0	0	0	1	0
			0	0	0	0	1	0
			0	0	0	0	1	0
10			0	0	0	0	1	0
11			0	0	0	2	1	0
12			0	2	0	3	1	0
13			0	7	4	5	4	1
14			0	0	0	0	1	0
15			0	0	0	2	1	0
16			1	0	0	0	24	0
17			1	0	0	0	5	0
18			6	2	0	7	16	0
19			0	6	4	46	48	1
20			0	1	0	7	1	0
21			0	0	0			0
22			0	0	0	0	1	0
23			0	2	0	5	0	0
24			0	0	0	1	4	0
25			0	0	0	3	2	0
26			4		0	14	24	0
27			0	15	7	70	68	0
28			0	0	1	1	2	0
			12	36	16	166	215	2
			(10)	(27)	(13)	(119)	(187)	(1)

*(Horizontal-component pairs are counted as one observation). The bottom row gives the total number for each site category. The numbers for 3.0-sec SA are somewhat lower (listed in parentheses only for the totals).

interevent term estimates ($y_{ij} - \eta_i$), solve for the attenuation relationship parameters (b_{1ss}, b_{1rv}, b_2, b_3, b_5, b_v) using a standard, linear inversion technique (e.g., singular value decomposition).

Step 3: Solve for the interevent and intraevent variances (τ^2 and σ^2, respectively) by maximizing the likelihood equation given by Abrahamson and Youngs (1992)

$$\ln L = -\frac{1}{2} N \ln(2\pi) - \frac{1}{2}(N - M)\ln(\sigma^2) - \frac{1}{2}\sum_{i=1}^{M}(\sigma^2 + n_i\tau^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{M}\sum_{j=1}^{n_i}(y_{ij} - \mu_{ij})^2 - \frac{1}{2}\frac{\tau^2}{\sigma^2}\sum_{i=1}^{M}\frac{\left(\sum_{j=1}^{n_i} r_{ij}\right)}{(\sigma^2 + n_i\tau^2)} \quad (3)$$

where $r_{ij} = (y_{ij} - \mu_{ij})$, N is the total number of records, M is the number of events, and n_i is the number of records for the ith event. This equation must be solved numerically; here, a grid-search is conducted over τ and σ between 0.01 and 1, with a step-size of 0.01. Once σ and τ are thereby found, the interevent terms are obtained as:

$$\eta_i = \frac{\tau^2 \sum_{j=1}^{n_i} r_{ij}}{n_i\tau^2 + \sigma^2}$$

Step 4: Repeat steps 2 and 3 until the solution converges (typically 30 iterations).

Step 5: Repeat steps 1–4 for the entire range of reasonable fictitious depths, h, and choose the solution that maximizes equation (3).

Once all parameters have been obtained, then site-effect terms can be obtained by averaging the interevent-corrected residuals ($y_{ij} - \eta_i - \mu_{ij}$) using whatever classification scheme one chooses.

Default Parameter Results

BJF97 provided parameter coefficients for shallow, crustal earthquakes in active tectonic regions (see their Table

8). Those for PGA and 5% damped response-spectral acceleration at 0.3- and 1.0-sec periods are reproduced in Table 3 (BJF97 did not provide values for 3.0-sec SA). Note that these values, and all others discussed in this article, are for the natural logarithm of the average horizontal component (in units of G). This section evaluates how well these "default" values predict the SCEC Phase-III strong-motion data for southern California (Steidl and Lee, 2000).

It should be noted that there is some overlap between the database used by BJF97 and that applied here. However, there are also some important differences. The SCEC Phase III data are restricted to southern California earthquakes, whereas BJF97 used all available data for western North America. Even with the geographical restriction, the Phase III database is larger than that used by BJF97 (447 PGA and SA observations from 28 earthquakes, compared to their 271 PGA observations from 19 events, and 112 SA observations from 14 events). The difference is largely due to the 160 observations from the 1994 Northridge earthquake that were not available to BJF97 (see Steidl and Lee [2000] for details).

The SCEC Phase III database lists one of three different focal mechanisms for each earthquake; b_1 in equation (1) has been set as follows for each: b_{1ss} for strike-slip events, b_{1rv} for reverse-slip events, and $0.5(b_{1ss} + b_{1rv})$ for oblique-slip events. This convention follows that of Lee *et al.* (2000), and differs from BJF97 in that they included oblique-slip events in b_{1rv}. The two horizontal components have been geometrically averaged (unless only one was available) to be consistent with the model predictions. Note that when solving for the interevent terms in this section, all coefficients (including h) were constrained to the default BJF97 values (in step 2 above). Only a single overall bias was solved for, which was less than 4% in all cases, and this was added to b_{1ss} and b_{1rv} as a correction. The site class for each observation was set according to the Wills *et al.* (2000) map, and each class was assigned the associated V_s value listed in Table 1 (no site-specific shear-wave velocity measurements were directly used).

The average interevent corrected residuals for each site type, relative to BJF97 predictions for class BC, are plotted in Figure 1. Also shown is the trend predicted by the default BJF97 model (solid line). The overall agreement in Figure 1 is somewhat surprising in light of the results of Steidl (2000) and Lee and Anderson (2000). The observed trends for 0.3- and 1.0-sec SA agree well with that predicted by BJF97, although there seems to be a small systematic bias for the 0.3-sec SA case. For the 1.0-sec case, all site classes except BC are significantly different from one another at the one-sigma level. This implies that the classification is indeed warranted. The residuals for PGA, on the other hand, clearly show less variation than predicted by the default BJF97 parameters. This suggests that some fine-tuning is in order with respect to southern California data.

Custom Fit Parameters

Based on the systematic difference in site-class residuals seen in the previous section (Fig. 1), the BJF97 model coefficients were custom fit to the southern California data. Following the five-step procedure outlined previously, I obtained the values listed in Table 4. Note that the arbitrary reference velocity, V_a, has been set to 760 m/sec in all cases, which should be considered when comparing b_1 values between Tables 3 and 4. Note also that 3.0-sec SA parameter coefficients are included here. For 1.0- and 3.0-second SA, b_3 initially turned out positive (but was within one sigma of zero). Following BJF97, these cases were redone with it set to zero (effectively imposing a negativity constraint on b_3).

The average residuals for each site type, relative to BC, for the custom-fit parameters in Table 4 are shown in Figure 2. Also plotted are the residuals predicted by this model (solid line). The difference seen between site types generally supports the classification scheme. For 1.0-sec SA, the model implies that D sites are amplified by ~22% relative to CD sites, and by a factor of 2.5 relative to B sites (amplification = $(V_s/V_a)^{b_v}$). In terms of the response spectral values, the only distinction among site types that seems questionable is C versus BC, but neither are inconsistent with the model either. There is only one case where the 68% confidence limits fall outside the model (type C for 3.0-sec SA), which is not unexpected since the one-sigma level has been plotted. PGA exhibits less of a difference between site types, with perhaps only two classifications being justified (D-CD versus C-BC-C).

Basin Depth Effect

Following the work of others (e.g., Trifunac and Lee, 1978; Campbell, 1997; Lee and Trifunac, 1995; Lee and

Table 3
Default BJF97 Parameter Values for PGA and 0.3- and 1.0-sec Response SA

	b_{1ss}	b_{1rv}	b_2	b_3	b_5	b_v	V_s	h	σ_{tot}†
PGA	−0.313		0.527	0.000	−0.778	−0.371	1396	5.57	
0.3-sec SA	0.598		0.769	−0.161	−0.893	−0.401	2133	5.94	
1.0-sec SA	−1.133		1.036	−0.032	−0.798	−0.698	1406	2.9	

*No 3.0-sec SA values are provided by BJF97.

†These values are for the average horizontal component (σ_1 and σ_2 in their Table 8 added in quadrature), not the random horizontal component.

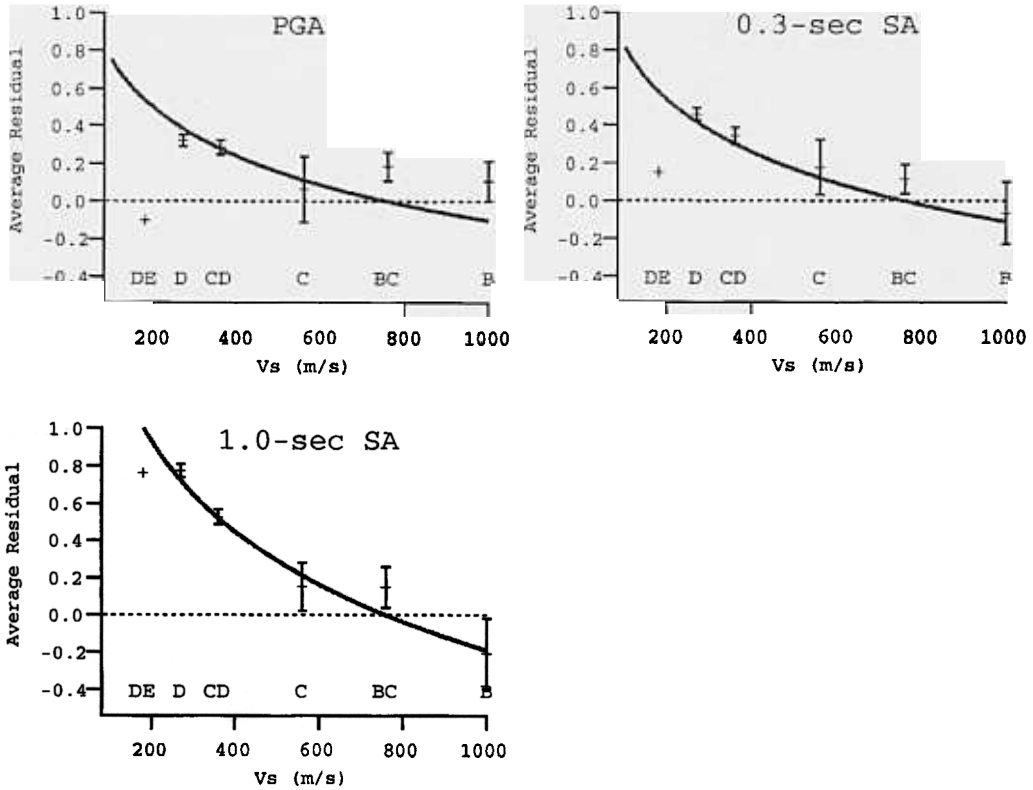

Figure 1. Average residuals (plus and minus one standard deviation of the mean) for each site category listed in Table 1. The residuals are plotted with respect to the default BJF97 prediction for the BC class (V_s = 760). The solid line is the BJF97-predicted residual ($b_v \ln(V_s/760)$). Note that the DE residuals have no uncertainty shown because only two observations were available.

Table 4
The BJF97 Model Coefficients Custom Fit to the Southern California Data*

	b1ss	b1rv	b2	b3	b5	bv	h	σ†	τ‡
PGA	0.853	0.872	0.442	−0.067	−0.960	−0.154	8.90	0.47	0.23
	(0.28)	(0.27)	(0.15)	(0.16)	(0.07)	(0.14)	(NA)	(0.02)	(NA)
0.3-sec	0.995	1.096	0.501	−0.112	−0.841	−0.350	7.20	0.53	0.26
SA	(0.27)	(0.25)	(0.15)	(0.16)	(0.07)	(0.14)	(NA)	(0.02)	(NA)
1.0-sec	−0.164	−0.267	0.903	0.0	−0.914	−0.704	6.20	0.53	0.22
SA	(0.26)	(0.24)	(0.11)	(NA)	(0.06)	(0.14)	(NA)	(0.02)	(NA)
3.0-sec	−2.267	−2.681	1.083	0.0	−0.720	−0.674	3.00	0.52	0.30
SA	(0.26)	(0.24)	(0.13)	(NA)	(0.06)	(0.15)	(NA)	(0.02)	(NA)

*The reference velocity, V_a, is 760 m/sec in all cases. Note that only 357 observations were available for 3.0-sec SA (as opposed to 447 for the other ground-motion parameters). The values in parentheses are one-sigma uncertainty estimates from the Singular Value Decomposition (assuming h is known exactly).

†The σ uncertainties are based on the approximation $\sigma/(2N)^{0.5}$ (Spiegel, 1975, pg 162), but are also consistent with the expression given by Abrahamson and Youngs (1992, n. 508).

‡The τ uncertainty estimates given by Abrahamson and Youngs (1992, p. 508) were deemed unreliable due to the limited number of earthquakes.

Anderson, 2000; and Steidl, 2000), this section examines the dependence of interevent corrected residuals with respect to basin depth. The SCEC Phase III strong-motion database (Steidl and Lee, 2000) provides estimates of basin depth at some of the observation sites. These are defined relative to the 2.5 km/sec shear-wave velocity isosurface in the SCEC 3D velocity model (Magistrale *et al.*, 2000). Of the 447 observations in the database, 197 have basin-depth estimates (151 for 3-sec SA).

Interevent corrected residuals are plotted versus basin

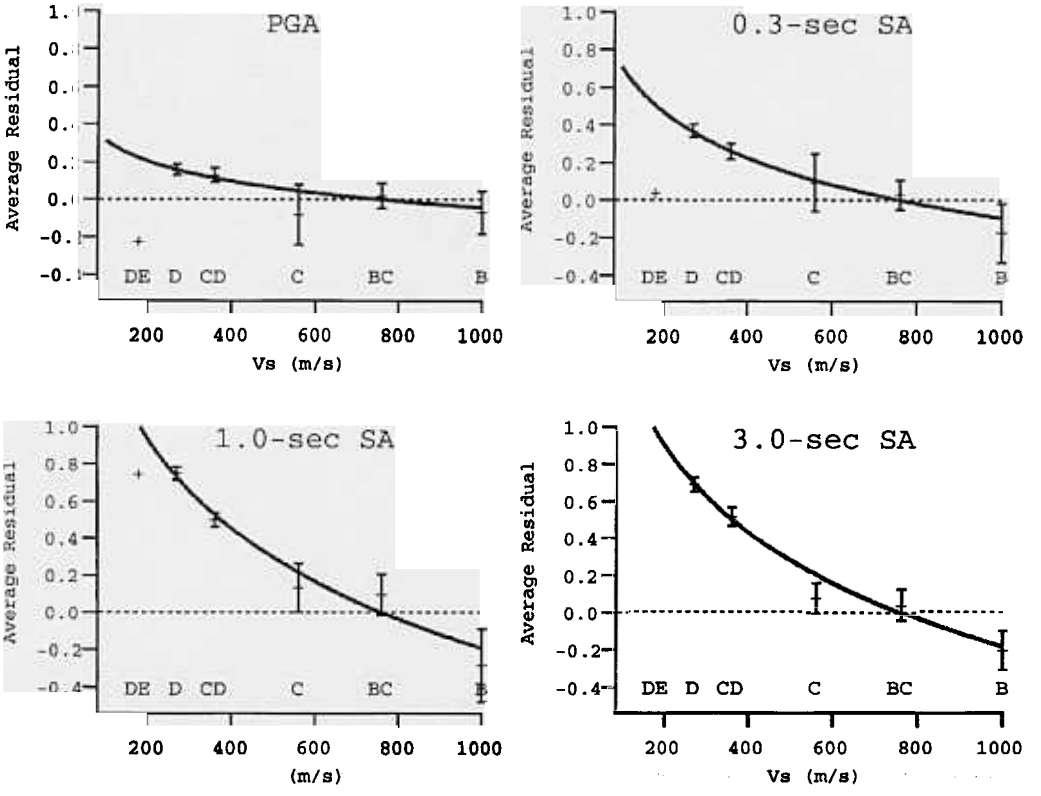

Figure 2. Average residuals relative to BC, as in Figure 1, but for the custom-fit coefficients listed in Table 4. The solid line is the residual predicted using the new value of b_v (from Table 4). The 3.0-sec SA residual for DE is not shown because it has a value of 1.5 (off the scale) and is based on only one observation.

depth in Figure 3. These residuals are determined using the custom fit coefficients listed in Table 4, and relative to the prediction appropriate to each site class (not relative to BC predictions as plotted in Figures 1 and 2). Listed in each plot is the best-fit slope and intercept along with their uncertainties in parentheses. The trends are also listed in Table 5. In all cases the slope is significantly different from zero by more than two sigma. The best-fit straight line for 1.0 second SA implies an amplification factor of ~2 in going from zero to 6 km depth (the latter approximately being the maximum depth for the LA Basin). That trend for PGA implies up to 50% amplification for the LA basin, which is more than the influence of the detailed site classification inferred above (Fig. 2).

There is a potential problem in using only a subset of data. Applying the custom-fit coefficients (from Table 4) ensures that there will be no significant trends when plotting all residuals against any of the independent variables: distance, magnitude, site type, or style of faulting. However, the same may not be true for the data subset. This is exemplified for 3.0-sec SA in Figure 4, which shows a significant trend with respect to magnitude for the subset of residuals where depth is known. In fact, the difference in ground motion implied by the magnitude trend is greater than that implied by the basin effect. Although Figure 4 was the most dramatic case the author could find, it nevertheless raises the question of whether the regionally fit coefficients are applicable to the subregion, and whether the basin-depth effect seen in Figure 3 is an artifact of something else.

To further investigate this possibility, a set of parameter coefficients was obtained for the subset of data where depth is known. The resulting basin-depth trends (not shown) were very close to those in Figure 3; that is, the differences were a small fraction of the uncertainties. Therefore, the basin-depth effect is apparently not an artifact of data selection. The question remains, however, as to which set of model coefficients should be applied. Most parameter coefficients were not significantly different between the two cases. In fact, the only significant differences were for b_2 (the linear magnitude dependence) at 1.0- and 3.0-sec SA. Furthermore, these were significant only at the one-sigma level and not at two-sigma. Therefore, it is concluded that the subset-data coefficients should not be used in lieu of the full-database coefficients (which is why they have not been listed). Nevertheless, this issue highlights the possibility that some subregions might indeed have significantly different coefficient values.

In general, the coefficient for any new independent variable (including basin depth) should be solved for simultaneously with the other coefficients in the five-step procedure

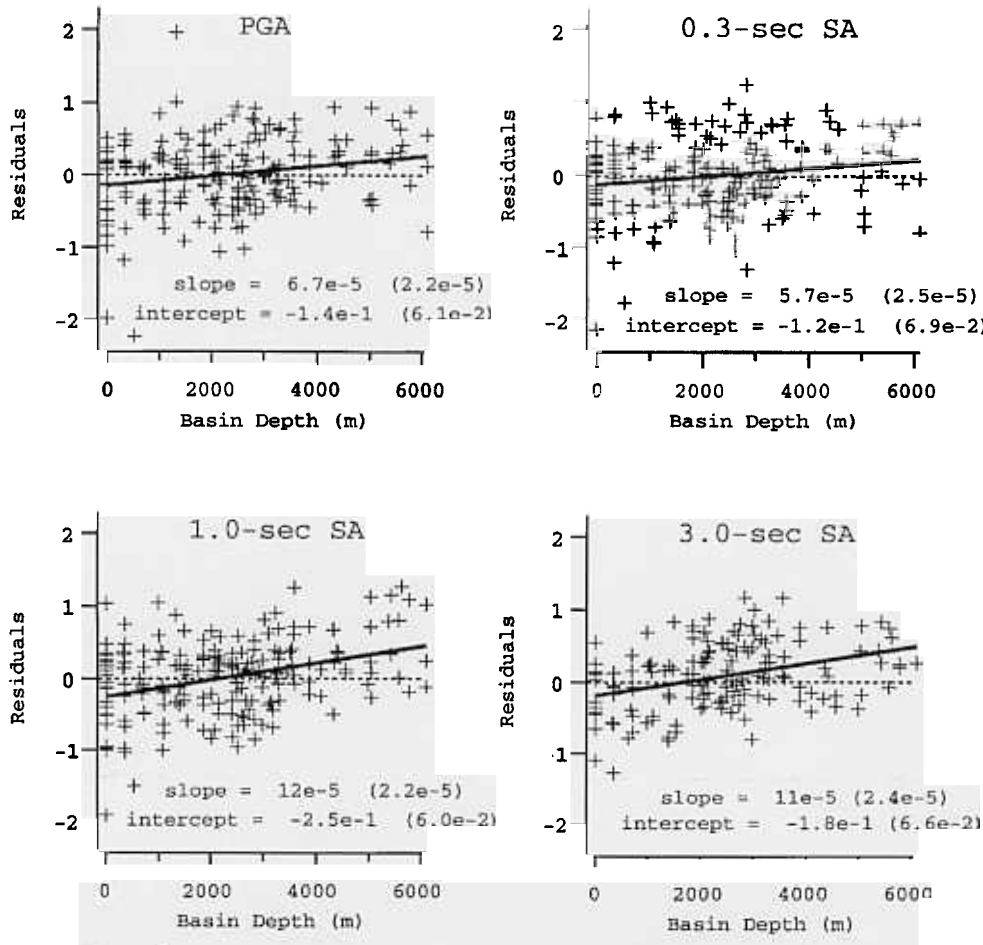

Figure 3. Interevent corrected residuals versus basin depth (defined with respect to the 2.5 km/s shear-wave velocity isosurface). The values listed in parentheses are the one-sigma uncertainties.

Table 5
The Basin-Depth Effect Inferred from Figure 3*

Ground Motion Parameter	Trend
PGA	6.7e-5depth − 0.14
0.3-sec SA	5.7e-5depth − 0.12
1.0-sec SA	12e-5depth − 0.25
3.0-sec SA	11e-5depth − 0.18

*Depth is measured in meters and is defined by the 2.5 km/sec shear-wave velocity isosurface. The trend can be added to equation (1) as a correction factor.

outlined above. However, because basin depth is known at only a subset of sites, its contribution had to be determined after solving for the others. Future studies should include the basin-depth parameter explicitly to the extent possible.

Nonlinear Site Response

The BJF97 model does not explicitly account for nonlinear site response in terms of having site-amplification factors that depend on the ground-motion level (see Beresnev and Wen (1996) or Field *et al.* (1998) for a general discussion of this topic). This is in contrast to the model by Abrahamson and Silva (1997), for example, which includes nonlinearity by virtue of having amplification factors that depend on rock-site PGA. The BJF97 model has a single amplification factor regardless of the shaking level, which represents the average for all data used in the regression. Recall that the custom-fit b_v value for PGA found here (Table 4) is lower than the original BJF97 value (Table 3). This suggests that there is less contrast between rock and sediment PGA in southern California, which could be attributable to more nonlinearity in the data used here (perhaps from the 1994 Northridge earthquake).

We can test for nonlinearity by looking for trends in residuals plotted versus the predicted ground motion for rock (e.g., Steidl, 2000; Joyner and Boore, written comm.). This is done in Figure 5, where residuals for site-class D are plotted against the values predicted for site-class BC. Also shown is the best-fit line for each case. The slope for PGA is indeed significant (although just barely), and the sign is consistent with the kind of nonlinearity expected on the basis

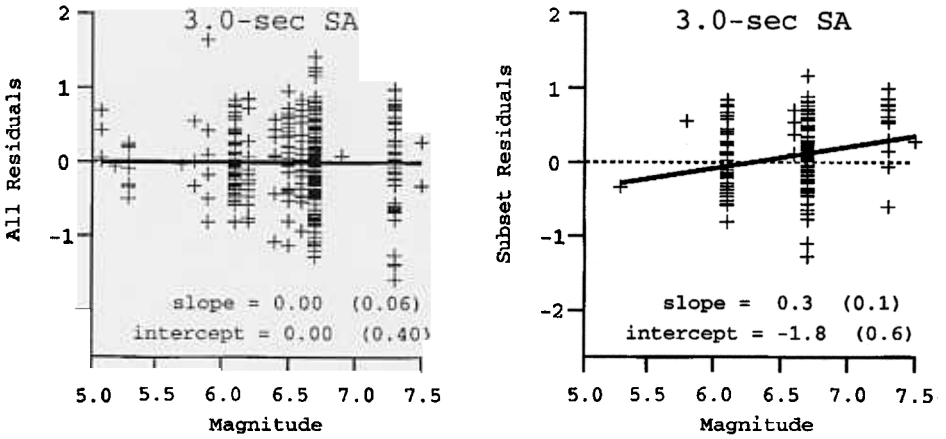

Figure 4. Interevent corrected residuals for all the southern California data (left) and for the subset where basin depth is known (right) plotted versus earthquake magnitude.

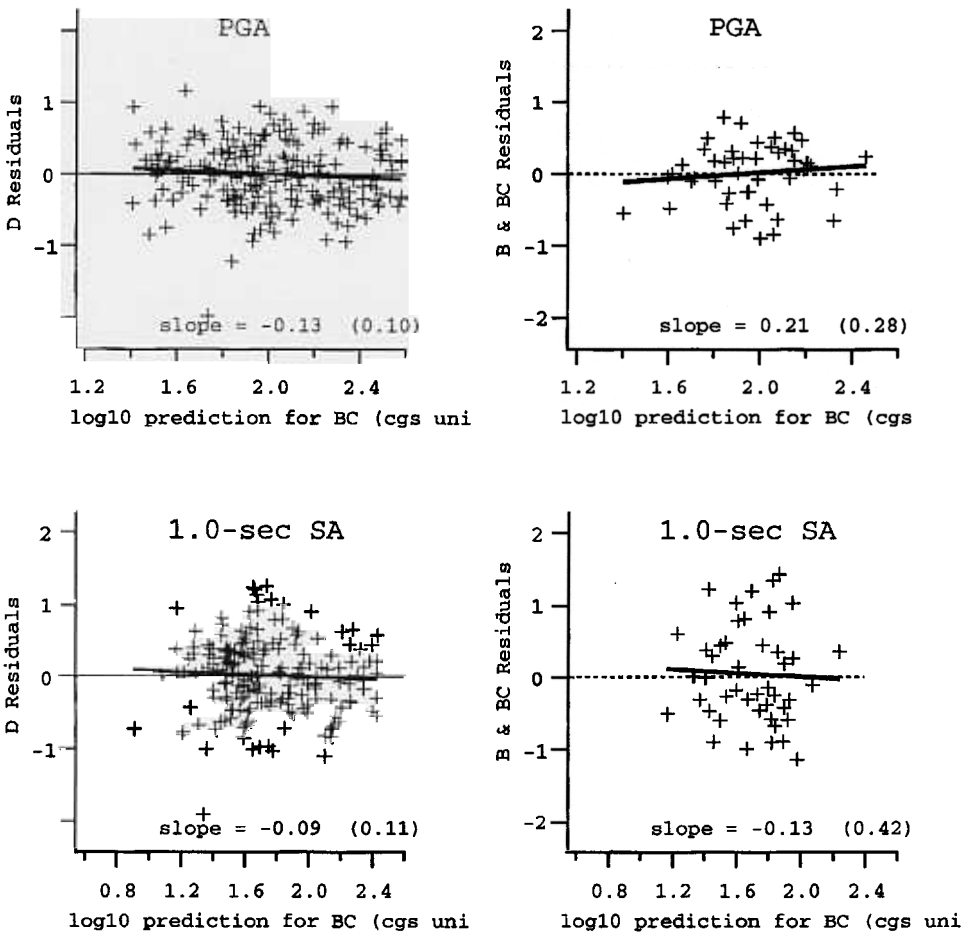

Figure 5. Interevent corrected residuals for site class D (left), and for site classes B and BC combined (right), plotted against the ground motion level predicted for class BC. Similar plots for 0.3- and 3.0-sec SA reveal no significant slopes at the 68% confidence level.

of geotechnical laboratory testing. That is, the model tends to overpredict D-class PGA by about 7% at the highest level of shaking and underpredict it by about 7% for the weakest level. The slopes for the other ground-motion parameters are not significant, implying a lack of evidence for nonlinearity.

Note that a non-zero slope in D-class residuals, as plotted in Figure 5, is not the only possible manifestation of nonlinearity. For example, it is clear from Table 2 that D sites dominate the strong-motion database (215 out of 447 observations). Therefore, it is possible that nonlinearity has been folded into the magnitude, distance, and/or interevent terms in the regression, leaving less of an apparent trend with respect to sediment residuals plotted against the shaking level. If this were the case, we might expect to see a significant and opposite trend in rock-site residuals. That is, if nonlinearity is present and somehow accounted for at sediment sites, the model would tend to underpredict rock-site PGA at high levels of shaking and overpredict at the lowest levels. To investigate this possibility, Figure 5 includes B- and BC-class residuals (combined), plotted versus the class predictions. Again, the rock-site residuals for PGA show the trend expected from nonlinearity (a positive slope). However, none of the B/BC slopes, including that for PGA, are significantly different from zero at the 68% confidence level. The uncertainties on the rock-site slopes in Figure 5 are much greater than those for the D-class sites, which reflects a relatively limited number of rock-site observations available (there are only 12 B- and 36 BC-site observations).

The southern California data are clearly insufficient to resolve the significance of nonlinearity with the approach taken here. More recordings are needed, especially at rock sites and over as wide a range of ground-motion levels as possible. Steidl (2000) reached a similar conclusion. We must now either wait for more earthquakes in southern California, or include data from other regions as well. In fact, Joyner and Boore have already taken this latter approach (written comm.), finding significant nonlinear effects at short periods (0.2-sec SA), but not at longer periods (1.0-sec SA).

With respect to the model presented in this article, it is concluded that nonlinearity may effect PGA (and perhaps 0.3-sec SA), but that it probably manifests more as an underprediction at rock sites than an overprediction at sediment sites at higher levels of shaking. No strong evidence for nonlinearity is seen for 0.3-, 1.0-, and 3.0-sec SA. Finally, it is important to note that class types E and DE, which have the highest likelihood of exhibiting nonlinear behavior, have not been addressed here because they are not adequately represented in the southern California strong-motion database.

Interevent Variability (τ)

An important aspect of any ground-motion model is the prediction uncertainty, part of which comes from the interevent variability, τ, introduced earlier. Abrahamson and Youngs (1992) provide approximate uncertainty estimates for both τ and σ (the latter being the intraevent variability). As discussed in the next section, their expression for σ was found to be reliable. However, their expression for τ was not, perhaps because there are too few events in the database to make the approximation valid. To investigate the uncertainty of this parameter, and to test the regression scheme as a whole, a set of Monte Carlo simulations was carried out. 250 synthetic databases were created, each with the same distribution of independent variables as represented in the actual data set. The parameters listed in Table 4 were used to create the synthetic data, and interevent and intraevent terms were added by sampling values from a zero-mean, Gaussian distribution with a standard deviation appropriate to σ and τ, respectively (also from Table 4). The coefficients were then determined from each synthetic database using the five-step procedure outlined above. The distribution of τ values obtained from the 250 simulations is shown in Figure 6 for PGA and 3.0-sec SA (the results for 0.3- and 1.0-sec SA are similar). Also plotted is a Gaussian distribution (see caption for details). It is not believed that the τ estimates actually follow this distribution; in fact they cannot because negative values are impossible. The Gaussian is shown simply for illustrative purposes.

There are three issues to note from Figure 6. First, in both cases, the average τ value recovered from the Monte Carlo simulations (obtained as the square root of the average variance, τ^2) is about 0.05 lower than that used to make the synthetic data. This suggests a bias in the estimation methodology. Second, the standard deviation of recovered values is roughly 0.07. Third, there is a conspicuous peak for the bin at 0.01 (the lowest value tested). This, as well as the bias just noted, suggests a potential problem with the five-step analysis methodology, or perhaps an error in its implementation here. However, all other parameters (b_{1ss}, b_{1rv}, b_2, b_3, b_5, b_v, and σ), and their uncertainties were nicely recovered in the Monte Carlo simulations. Thus, the problem seems to exist for τ alone.

Faced with a potential quagmire of statistical approximations, the author chose not to delve further into this issue here. Rather, it is concluded that τ is uncertain by about 0.1 (and possibly biased as well). Refinements beyond this are left to future studies.

Intra-Event Variability (σ)

Another source of uncertainty comes from the intraevent variability (σ). The values determined in the five-step procedure (listed in Table 4) are equivalent to taking the square root of the average squared interevent corrected residuals (ignoring lost degrees of freedom which make a negligible change compared to the uncertainties). Some previous studies have found that σ varies with magnitude (e.g., Youngs *et al.*, 1995) and/or site type (e.g., Sadigh *et al.*, 1997). Such dependencies are examined here by looking for trends between residuals squared and the independent variable in question.

Figure 7 presents the results with respect to magnitude.

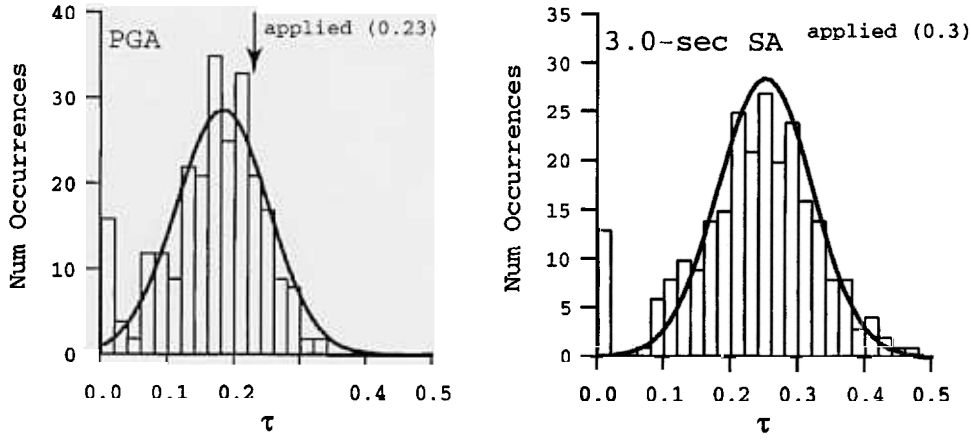

Figure 6. The distribution (histogram) of τ values obtained from 250 synthetic data sets in the Monte Carlo simulations. The arrow shows the value of τ applied in generating the synthetic data, and the solid line is a Gaussian distribution with the same mean (actually, the square root of the average variance [τ^2]) and a standard deviation of 0.07 (not actually fit to the data, but chosen simply for illustrative purposes).

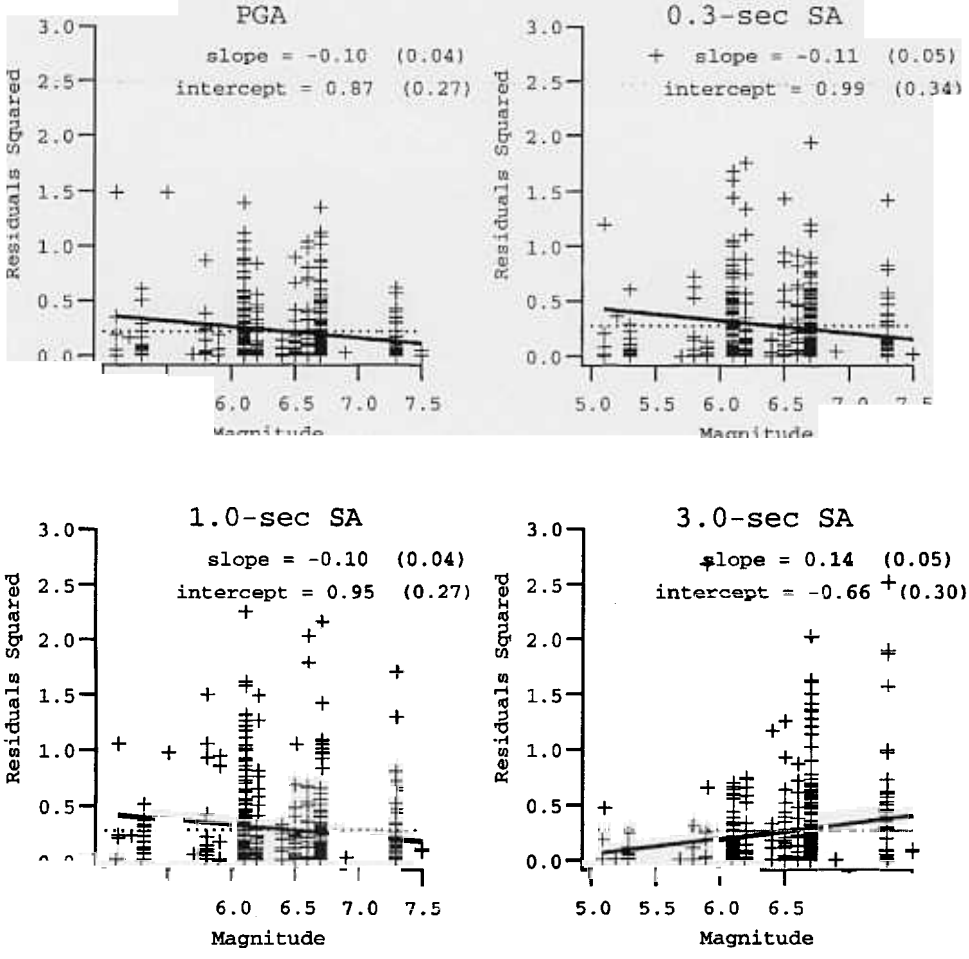

Figure 7. Squared residuals plotted versus magnitude. The dotted values represent the average (equivalent to σ in Table 4 squared), and the solid line is the best fit found by least squares. The values written in parentheses are the one-sigma uncertainties for the slope and intercept.

The dotted horizontal line shows the average of all residuals squared (equivalent to the σ value listed in Table 4 squared). The least-squares, straight-line fit is shown in black, and the slope and intercept are also listed along with their uncertainties. All four cases in Figure 7 show a significant trend with magnitude. PGA, 0.3 and 1.0-sec SA exhibit a decrease in σ with increasing magnitude, whereas the 3.0-sec case exhibits an increase. It is acknowledged that this approach may be statistically dubious, as the residuals squared are not Gaussian distributed. However, statistically equivalent results were obtained by computing σ over separate magnitude bins (e.g., Steidl, 2000), or for each earthquake separately (e.g., Lee and Anderson, 2000), and then computing a weighted least squares line.

Plots similar to those in Figure 7 also were made with respect to the logarithm of the average 30-m shear-wave velocity for each site category. No significant trends are observed for PGA and 0.3-sec SA. However, σ for 1.0-sec SA increases, and that for 3.0-sec SA decreases with increasing site velocity. Because the effect is smaller than that seen with respect to magnitude, it is not considered further here.

Again, it is acknowledged that there are more statistically sophisticated ways of testing for such trends (e.g., Youngs *et al.,* 1995). However, the analysis here seems adequate for present purposes. Finally, it should also be noted that the residuals plotted in Figure 7 have not been corrected for the basin-depth effect because it is known only at a subset of sites. Comparing the subset-data residuals before and after the depth correction implies a reduction in σ of less than ~5 percent, or about the same as the uncertainty in σ itself.

Total Variability (σ_{tot})

The total uncertainty (or variability) for the attenuation relationship is given as:

$$\sigma_{tot} = (\sigma^2 + \tau^2)^{0.5}$$

With results from the previous sections, both magnitude-dependent and magnitude independent σ_{tot} values are listed in Table 6. Based on the aforementioned Monte Carlo simulations, the uncertainty of σ_{tot} values in Table 6 is roughly 0.05 (a more quantitative statement does not seem warranted given the potential bias in τ discussed above, and the fact that the true distribution is unknown). Comparing the magnitude-independent values listed in Table 6 to the BJF97 default values listed in Table 3, it is clear the latter are somewhat smaller (especially for 0.3-second SA).

Tests of Normality

A fundamental assumption in all attenuation relationships is that residuals have normal or Gaussian distribution. This null hypothesis could not be rejected for any of the ground-motion parameters examined here, at any reasonable significance level, using the Kolmogorov-Smirnov test (e.g., Scheaffer and McClave, 1982). Theoretical quantile-quantile plots were also generated, but are not shown because nothing notable was observed.

Table 6
Total Uncertainties σ_{tot}*

	Magnitude Dependent $M \leq 7$	Magnitude Dependent $M > 7$	Magnitude Independent
PGA	$(0.93 - 0.10M)^{0.5}$	0.48	
0.3-sec SA	$(1.06 - 0.11M)^{0.5}$	0.54	
1.0-sec SA	$(1.00 - 0.10M)^{0.5}$	0.55	
3.0-sec SA	$(-0.57 + 0.14M)^{0.5}$	0.64	

*The magnitude-independent values are σ and τ from Table 4 added in quadrature. The magnitude-dependent values were obtained by squaring τ from Table 4, adding it to the best fit lines in Figure 7, and then taking the square root of the entire expression. The magnitude-dependent values above M = 7 are arbitrarily set to the value at that magnitude.

Discussion

A basic assumption made here (and in other SCEC Phase III studies) is that the southern California database is an adequate representation of all possible earthquakes. The hope has been that focussing on this subset would tell us something unique about the region, and that we have not been led astray by any atypical events in the database. Future studies will hopefully reveal the validity of this assumption.

The detailed site-classification map presented by Wills *et al.* (2000), paired with the BJF97 attenuation model and custom-fit parameters obtained here, is justified by the southern California data. Perhaps this should not be surprising in that there is certainly some overlap in the shear-wave velocity data used by BJF97 and Wills *et al.* (2000). However, even though the studies may not be completely independent, it is still satisfying to see significant differences in ground motion between the categories. This is especially good news in light of the opposite conclusions of Steidl (2000) and Lee and Anderson (2000) with respect to the Tinsley and Fumal (1985) detailed-geology map.

In addition to the site classification, there is a significant basin-depth effect (Figure 3). In fact, for PGA the influence of basin depth is greater than that of detailed geology. Although basin effects have been recognized in attenuation relationships previously for longer-period motions (e.g., Trifunac and Lee, 1978; Campbell, 1997; Lee and Trifunac, 1995), to our knowledge the SCEC Phase III articles are the first to recognize the influence for PGA as well (Steidl, 2000; Lee and Anderson, 2000; this study), The basin effect inferred and applied here is linear with respect to depth (Figure 3). This is supported by the average trend found in finite-difference simulations for nine different earthquake scenarios (Olsen, 2000, also part of the SCEC Phase III effort). However, some other depth dependence (e.g., Campbell, 1997) cannot be ruled out.

The physical mechanism for the basin effect is not yet

clear, as depth may be a proxy for something else. For example, the Phase III article by Joyner (2000) argues that distance from the basin edge, which generally correlates with depth, is the relevant parameter. In fact, the actual physical attribute may differ between longer and shorter period ground motion. The fact that PGA is amplified at deeper basin sites, rather than reduced by anelastic attenuation, is particularly surprising. Perhaps general focusing from basin concavity dominates the influence of anelasticity. Alternatively, it is possible that the basin effect is a complete artifact of something else. However, it cannot be a trade-off between the Wills *et al.* (2000) categories and basin depth because the basin effect was inferred after the influence of site type was effectively maximized. It could, however, reflect variations within a site category. For example, perhaps V_s within class D correlates with basin depth. If basin depth is such a proxy, the relevant question is whether there are any conditions where applying it anyway will lead to misleading results. Field and Petersen (2000) discuss this issue in the context of probabilistic hazard analysis.

Both magnitude-dependent and independent values for the total variability, σ_{tot}, are given in Table 6. The trends toward lower values at higher magnitudes are consistent with previous findings (e.g., Youngs *et al.*, 1995). However, to the author's knowledge, the opposite trend for 3.0-sec SA, seen also by Lee and Anderson (2000), is new to the Phase III studies. With respect to PGA, Youngs *et al.* (1995) discuss how soil nonlinearity may reduce amplification factors for stronger shaking, and thereby lower σ with increasing magnitude. Nonlinearity may have the opposite effect for long-period motion, where amplification factors increase with stronger shaking as site resonances shift to lower frequencies (e.g., Beresnev and Wen, 1996; Field *et al.*, 1998). This could explain the increase in σ_{tot} with magnitude at 3.0-sec SA. Finite-source effects, which increase with period and event size (e.g., Abrahamson and Somerville, 1996), could also be responsible for the observed trend. For now, choosing between the magnitude dependent and independent values listed in Table 6 is left to user discretion. The implications of this choice with respect to probabilistic hazard are discussed by Field and Petersen (2000).

Although the attenuation model presented here is presumably state of the art, its potential inadequacies should be noted. For one, the BJF97 model does not explicitly include nonlinear site effects in the form of amplification factors that depend on the level of ground motion. Even so, there is no strong indication of this inadequacy in the data residuals (Fig. 5). Either nonlinearity has been accounted for implicitly (masquerading as a magnitude and/or distance dependence), the data do not span a wide enough range of ground-motion levels to make it apparent, or both. If nonlinear effects are significant at higher levels of shaking, they will most likely manifest as an underprediction of rock-site PGA rather than an overprediction of sediment-site PGA. Another potential problem with the BJF97 model, as suggested by Anderson (2000), is that the distance-dependence does not vary with magnitude (although BJF97 saw no evidence for such an effect in their data). Even with these potential inadequacies, however, the BJF97 model predicts existing southern California data as well as other competing models (Lee *et al.*, 2000). The question remains as to how well it and other attenuation relationships will do with respect to conditions not represented in the database; e.g., near-field ground motion for large earthquakes.

Conclusions

This article has modified the BJF97 attenuation relationship in accordance with southern California data, the Wills *et al.* (2000) detailed site-classification map, and a basin-depth term. The final model for the four ground-motion parameters is given by equation (1) with the coefficients listed in Table 4. If depth to the 2.5 km/sec shear-wave velocity isosurface is known, then the appropriate basin-depth term listed in Table 5 can be added to equation (1). The total variability, σ_{tot}, can be set with either the magnitude dependent or independent case listed in Table 6 (the choice is left to the user).

The results here justify applying the site categories of Wills *et al.* (2000), and support previous findings regarding basin-depth effects. In fact, basin depth has a stronger influence on PGA than does site category. The level of statistical sophistication applied here is thought to be appropriate for present purposes, especially since more data will soon be forthcoming. These results can both serve as a guide to future attenuation-relationship developments and provide an interim model for use in southern California. The model presented here may not be applicable to NEHRP class A or E sites.

Acknowledgments

The author would like to thank William Joyner, Kenneth Campbell, and Mike Fehler for thoughtful review comments, and Norman Abrahamson, Yajie Lee, David Boore, and Jamison Steidl for many lengthy conversations. This research was supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-8920136 and USGS Cooperative Agreements 14-08-0001-A0899 and 1434-HQ-97AG01718. The SCEC contribution number for this paper is 542.

References

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

Abrahamson, N. A., and P. G. Somerville (1996). Effects of the hanging wall and footwall on ground motions recorded during the Northridge earthquake, *Bull. Seism. Soc. Am.* **86,** S93–S99.

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

Anderson, J. G. (2000). Expected shape of regressions for ground-motion parameters on rock, *Bull. Seism. Soc. Am.* **90,** no. 6B, S43–S52.

Beresnev, I. A., and K.-L. Wen (1996). Nonlinear soil response: a reality?, *Bull. Seism. Soc. Am.* **86,** 1964–1978.

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

Campbell, K. W. (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra, *Seism. Res. Lett.* **68,** 154–179.

Field, E. H., and M. D. Petersen (2000). A test of various site-effect parameterizations in probabilistic seismic hazard analyses of southern California, *Bull. Seism. Soc. Am.* **90,** no. 6B, S222–S244.

Field, E. H., S. Kramer, A.-W. Elgamal, J. D. Bray, N. Matasovic, P. A. Johnson, C. Cramer, C. Roblee, D. J. Wald, L. F. Bonilla, P. P. Dimitriu, and J. G. Anderson (1998). Nonlinear site response: where we're at (a report from a SCEC/PEER seminar and workshop), *Seism. Res. Lett.* **69,** 230–234.

Field, E. H., and the SCEC Phase III working group (2000). Accounting for site effects in probabilistic seismic hazard analyses of southern California: overview of the SCEC Phase III report, *Bull. Seism. Soc. Am.* **90,** no. 6B, S1–S31.

Joyner, W. B. (2000). Strong motion from surface waves in deep sedimentary basins, *Bull. Seism. Soc. Am.* **90,** no. 6B, S95–S112.

Lee, Y., and J. G. Anderson (2000). Potential for improving ground-motion relations in southern California by incorporating various site parameters, *Bull. Seism. Soc. Am.* **90,** no. 6B, S170–S186.

Lee, V. W., and M. D. Trifunac (1995). Frequency dependent attenuation functions and Fourier amplitude spectra of horizontal strong earthquake ground motion in California: final task report on the characteristics of earthquake ground motions for seismic design, (submitted to State of California Department of Transportation, Los Angeles County Department of Public Works, City of Los Angeles Department of Public Works).

Lee, Y., J. G. Anderson, and Y. Zeng (2000). Evaluation of empirical ground-motion relations in southern California, *Bull. Seism. Soc. Am.* **90,** no. 6B, S136–S148.

Magistrale, H., S. Day, R. W. Clayton, and R. Graves (2000). The SCEC southern California reference three-dimensional seismic velocity model version 2, *Bull. Seism. Soc. Am.* **90,** no. 6B, S65–S76.

Olsen, K. B. (2000). Site amplification in the Los Angeles basin from three-dimensional modeling of ground motion, *Bull. Seism. Soc. Am.* **90,** no. 6B, S77–S94.

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. *Seism. Res. Lett.* **68,** 180–189.

Scheaffer, R. L., and J. T. McClave (1982). Statistics for Engineers, Duxbury Press, Boston, Massachusetts.

Scientists from the U.S. Geological Survey, Southern California Earthquake Center, and California Division of Mines and Geology (2000). Preliminary Report on the 16 October 1997 M 7.1 Hector Mine, California, Earthquake, *Seism. Res. Lett.* **71,** 11–23.

Shin, T. C., K. W. Kuo, W. H. K. Lee, T. L. Teng, and Y. B. Tsai (2000). A preliminary report on the 1999 Chi-Chi (Taiwan) earthquake, *Seism. Res. Lett.* **71,** 24–30.

Spiegel, M. R. (1975). Theory and problems of probability and statistics, *Schaum's Outline* Series, McGraw-Hill, New York.

Steidl, J. H. (2000). Site response in southern California for probabilistic seismic hazard analysis, *Bull. Seism. Soc. Am.* **90,** no. 6B, S149–S169.

Steidl, J. H., and Y. Lee (2000). The SCEC Phase III strong-motion database, *Bull. Seism. Soc. Am.* **90,** no. 6B, S113–S135.

Tinsley, J. C., and T. E. Fumal (1985). Mapping Quaternary sedimentary deposits for areal variations in shaking response, in *Evaluating Earthquake Hazards in the Los Angeles Region: An Earth Science Perspective,* J. E. Ziony (Editor), *U.S. Geol. Surv. Prof. Paper 1360,* 101–126.

Toksöz, M. N., R. E. Reilinger, C. G. Doll, A. A. Barka, and N. Yalcin (1999). Izmit (Turkey) earthquake of 17 August 1999: first report, *Seism. Res. Lett.* **70,** 669–679.

Trifunac, M. D., and V. W. Lee (1978). Dependence of the Fourier amplitude spectra of strong motion acceleration on the depth of sedimentary deposits, University of Southern California, Department of Civil Engineering, Report No. CE 78-14, Los Angeles.

Wills, C. J., M. Petersen, W. A. Bryant, M. Reichle, G. I. Saucedo, S. Tan, G. Taylor, and J. Treiman (2000). A site-conditions map for California based on geology and shear-wave velocity, *Bull. Seism. Soc. Am.* **90,** no. 6B, S187–S208.

Youngs, R. R., N. Abrahamson, F. I. Makdisi, and K. Sadigh (1995). Magnitude-dependent variance of peak ground acceleration, *Bull. Seism. Soc. Am.* **85,** 1161–1176.

Department of Earthquake Sciences
University of Southern California
Los Angeles, California 90089-0740

Manuscript received 1 July 2000.