Bulletin of the Seismological Society of America, Vol. 94, No. 2, pp. 591–597, April 2004

Estimating $\overline{V}_s(30)$ (or NEHRP Site Classes) from Shallow Velocity Models (Depths < 30 m)

by David M. Boore

Abstract The average velocity to 30 m [$\overline{V}_s(30)$] is a widely used parameter for classifying sites to predict their potential to amplify seismic shaking. In many cases, however, models of shallow shear-wave velocities, from which $\overline{V}_s(30)$ can be computed, do not extend to 30 m. If the data for these cases are to be used, some method of extrapolating the velocities must be devised. Four methods for doing this are described here and are illustrated using data from 135 boreholes in California for which the velocity model extends to at least 30 m. Methods using correlations between shallow velocity and $\overline{V}_s(30)$ result in significantly less bias for shallow models than the simplest method of assuming that the lowermost velocity extends to 30 m. In addition, for all methods the percent of sites misclassified is generally less than 10% and falls to negligible values for velocity models extending to at least 25 m. Although the methods using correlations do a better job on average of estimating $\overline{V}_s(30)$, the simplest method will generally result in a lower value of $\overline{V}_s(30)$ and thus yield a more conservative estimate of ground motion [which generally increases as $\overline{V}_s(30)$ decreases].

Introduction

The average shear-wave velocity of the top 30 m of the Earth [$\overline{V}_s(30)$, which is computed by dividing 30 m by the travel time from the surface to 30 m] is an important parameter used in classifying sites in recent building codes (e.g., Dobry *et al.*, 2000; BSSC, 2001) and in loss estimation. The site classes estimated from shallow shear-wave velocity models are also important in deriving strong-motion prediction equations (e.g., Boore *et al.*, 1997), in construction of maps of National Earthquake Hazard Reduction Program (NEHRP) site classes (e.g., Wills *et al.*, 2000), and in applications of building codes to specific sites. Many measurements of near-surface shear-wave velocity, however, do not reach 30 m. For example, with one exception the shear-wave velocities at the Kyoshin Network (K-NET) stations in Japan (data source: www.k-net.bosai.go.jp/) are between 10 and 20 m (the one exception is K-NET station AKT019, with a depth to bottom of 5.04 m). Another example comes from a recent compilation of shear-wave velocities from 277 boreholes in California, more than half (142) of which are shallower than 30 m (Boore, 2003). A histogram of the depth to the deepest measurement for boreholes in that compilation is given in Figure 1, from which it can be seen that most of the shallow holes were drilled and logged before 1990. Although most of the holes are near 30 m, 31 have values less than 25 m. A third example is seismic cone penetrometer measurements made in the Oakland–Alameda area of California (Holzer *et al.*, 2002, 2004), where 193 out of 202 soundings are less than 30 m, with 146 of these being less than 20 m.

This note compares several ways of estimating $\overline{V}_s(30)$ from velocity models that do not reach 30 m (the models could be determined from either invasive or noninvasive methods). The simplest method assumes that the lowermost velocity of the model extends to 30 m; the other methods use correlations of shallow velocities and $\overline{V}_s(30)$.

Methods of Extrapolation

In the methods discussed here, a key quantity is the time-averaged velocity $\overline{V}_s(d)$ to a depth d. Generally d is the depth to the bottom of the velocity model, which is not necessarily the depth to the bottom of the borehole or the depth of the deepest measurement in a borehole if the velocity model was determined from borehole logging. I refrain from using the phrase "depth to bottom of borehole," which is meaningless for velocity models determined from noninvasive methods. The time-averaged velocity is computed from the equation

$$\overline{V}_s(d) = d/tt(d), \tag{1}$$

where the travel time tt(d) to depth d is given by

$$tt(d) = \int_0^d \frac{dz}{V_s(z)}. \tag{2}$$

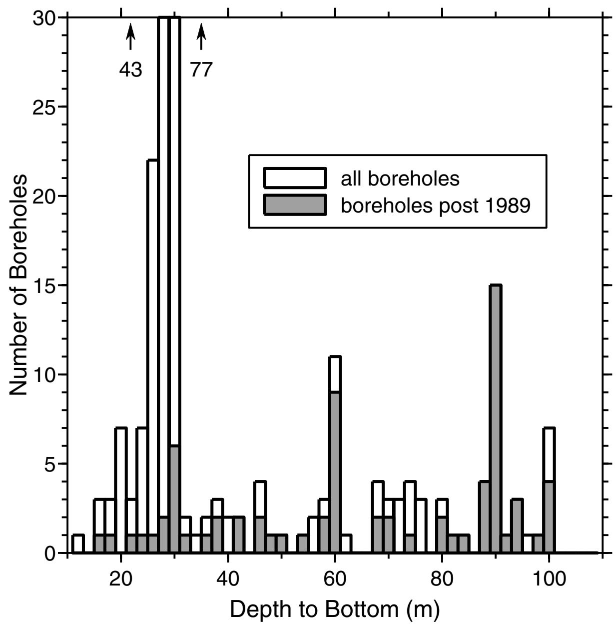

Figure 1. Distribution of the depths to the bottom of boreholes from California tabulated by Boore (2003). Note that most of the holes have depths near 30 m (the two highest bars have been capped so as to show better the distribution of other depths; the numbers indicate the height of the bars).

Table 1
Definition of NEHRP Site Classes in Terms of $\overline{V}_s(30)$, the Average Shear-Wave Velocity to 30 m

Site class	Range of $\overline{V}_s(30)$ (m/sec)
A	$1500 < \overline{V}_s(30)$
B	$760 < \overline{V}_s(30) \leq 1500$
C	$360 < \overline{V}_s(30) \leq 760$
D	$180 \leq \overline{V}_s(30) \leq 360$
E	$\overline{V}_s(30) < 180$

Constructed from information in BSSC, 2001.

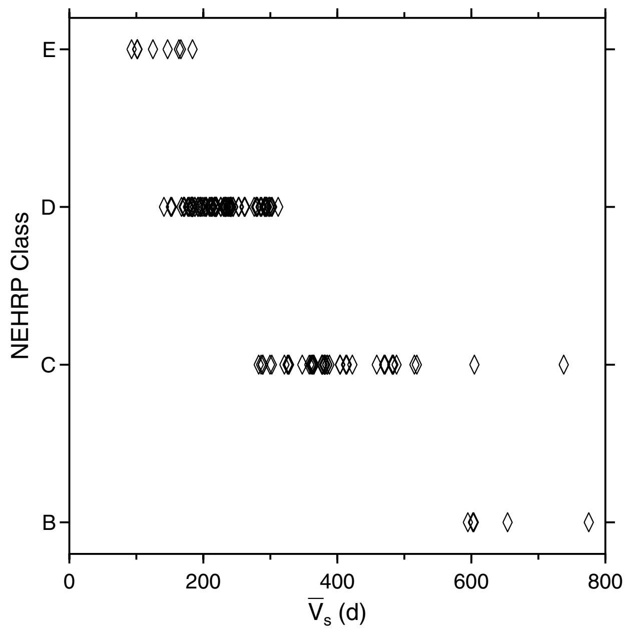

Figure 2. An example of NEHRP class as a function of $\overline{V}_s(d)$ for 135 boreholes. In this example, $d = 16$ m.

In equation (2), $V_s(z)$ is the depth-dependent velocity model. NEHRP classes are determined by $\overline{V}_s(30)$, as indicated in Table 1. The purpose of this note is to investigate ways of approximating $\overline{V}_s(30)$ if $d < 30$ m. If all that is desired is the NEHRP class, a simple method is to use correlations between the NEHRP class and $\overline{V}_s(d)$. I studied the correlation from the 135 California boreholes that extended to at least 30 m, for various assumed depths, and found that there was some overlap in site class for a given value of $\overline{V}_s(d)$ (an example is shown in Fig. 2). For this reason, and because $\overline{V}_s(30)$ is useful as a continuous parameter for characterizing site response in regression equations (e.g., Boore *et al.*, 1997), I investigate here some methods that make more use of the travel-time information from boreholes.

Extrapolation Assuming Constant Velocity

If the velocity model is available only to depth d, an assumption about the velocity between d and 30 m can be used to compute an estimate of $\overline{V}_s(30)$ using the following equation:

$$\overline{V}_s(30) = 30/(\mathrm{tt}(d) + (30 - d)/V_{\mathrm{eff}}), \tag{3}$$

where V_{eff} is the assumed effective velocity from depth d to 30 m. The simplest assumption is that V_{eff} equals the velocity at the bottom of the velocity model:

$$V_{\mathrm{eff}} = V_s(d). \tag{4}$$

Because the velocity in general increases with depth for both geological and geotechnical reasons, however, this method for determining V_{eff} will usually lead to an underestimate of $\overline{V}_s(30)$ and therefore to site classes that may be biased toward larger letter values (the problem should be worse the shallower the model). This drawback to the simplest method led to the methods described next.

Extrapolation Using the Correlation between $\overline{V}_s(30)$ and $\overline{V}_s(d)$

Another method uses the correlation between $\overline{V}_s(30)$ and $\overline{V}_s(d)$. Using the same set of 135 boreholes for which the actual depths reached or exceeded 30 m, I found that plots of $\overline{V}_s(30)$ against $\overline{V}_s(d)$ for a series of assumed depths d could be fit by a straight line. The scatter in these plots, however, increased with $\overline{V}_s(d)$. For this reason, a power-law relation between $\overline{V}_s(30)$ and $\overline{V}_s(d)$ was assumed, for which a straight line can be fit to the logarithms of the quantities. Figure 3 shows examples for four assumed depths. The correlation is good, even for the shallowest depth considered here (10 m). Table 2 gives the regression coefficients for the equation

$$\log \overline{V}_s(30) = a + b \log \overline{V}_s(d) \tag{5}$$

for depths ranging from 10 to 29 m. The velocity at the bottom of the model [$V_s(d)$] rather than the average velocity $\overline{V}_s(d)$ was also considered as the predictor variable, but the scatter was worse. $V_s(d)$ has the advantage, however, that it can be used if shallower parts of the velocity model are missing, as is often the case with suspension log results; I discuss later a method for determining NEHRP class that uses $V_s(d)$.

Equation (5) was used in two ways to determine $\overline{V}_s(30)$ and thus NEHRP class. The first was simply to insert the value of $\overline{V}_s(d)$ computed from the velocity model into equation (5). A second, somewhat more elaborate method used an estimate of $\log \overline{V}_s(30)$ randomly drawn from a Gaussian distribution with mean given by equation (5) and standard deviation given in Table 2 (the contribution to the variance due to uncertainty in the slope of the line is insignificant).

Extrapolation Based on Velocity Statistics to Determine Site Class

Another way to account for the general increase of velocity in a statistical way is discussed briefly in the appendix to Atkinson and Boore (2003) and is elaborated on here. For each borehole, I computed the ratio of $V_s(d)$ to the effective velocity (V_{eff}) needed to raise the site class to the next stiffer class than given by the simple extrapolation using equations (3) and (4), for a series of depths ranging in 1-m increments from 10 to 29 m. For each depth, the values of $V_{\mathrm{eff}}/V_s(d)$ were tabulated in increments of 0.1 units, starting from 0.4,

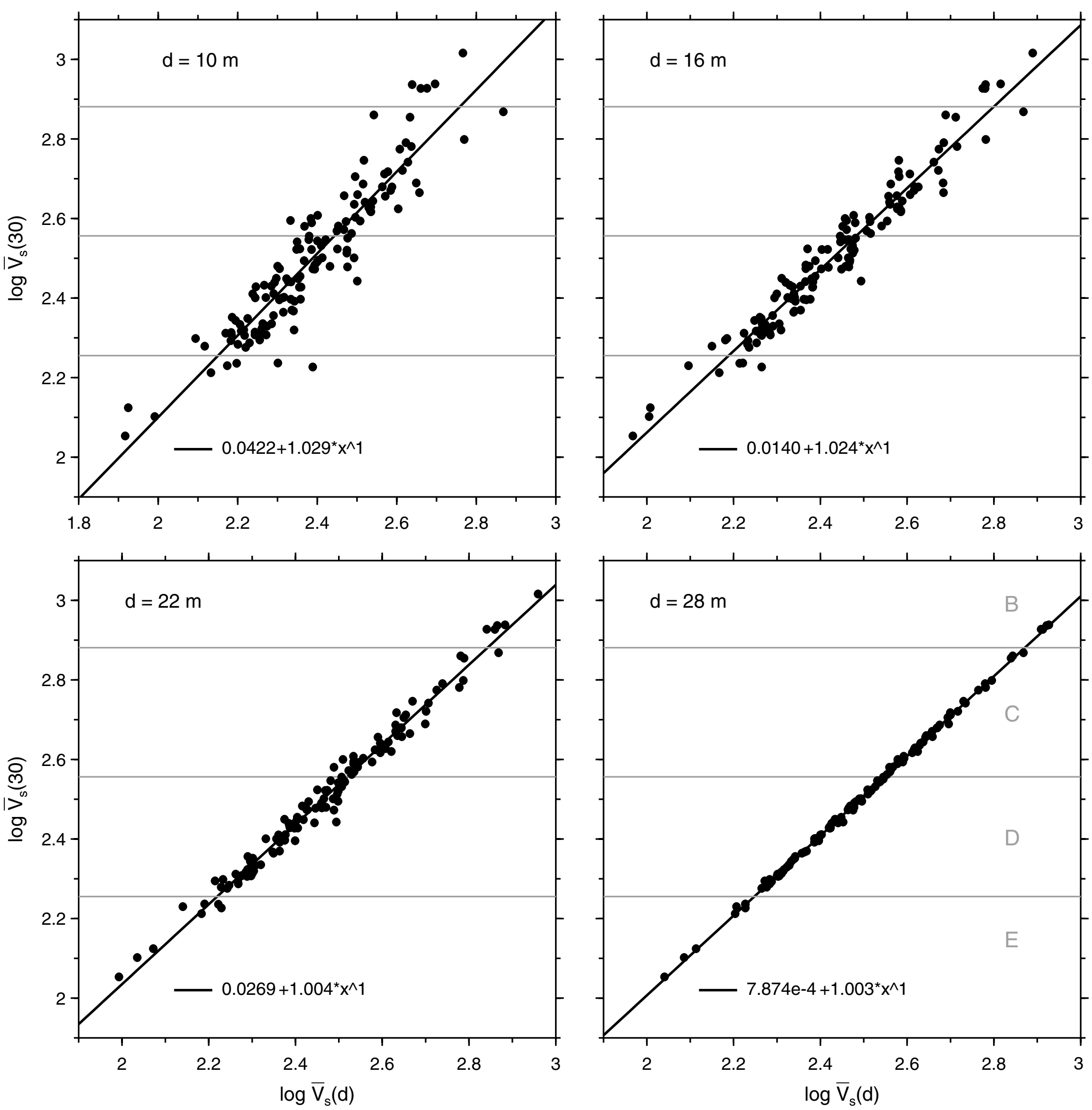

Figure 3. Fit of straight line to $\log \overline{V}_s(30)$ as a function of $\log \overline{V}_s(d)$, for $d = 10$, 16, 22, and 28 m. Velocities in meter per second. The gray lines show the boundaries between NEHRP site classes; the classes are shown in the lower right-hand graph.

Table 2
Coefficients of the Equation $\log \overline{V}_s(30) = a + b \log \overline{V}_s(d)$

d	a	b	σ
10	4.2062E − 02	1.0292E + 00	7.1260E − 02
11	2.2140E − 02	1.0341E + 00	6.4722E − 02
12	1.2571E − 02	1.0352E + 00	5.9353E − 02
13	1.4186E − 02	1.0318E + 00	5.4754E − 02
14	1.2300E − 02	1.0297E + 00	5.0086E − 02
15	1.3795E − 02	1.0263E + 00	4.5925E − 02
16	1.3893E − 02	1.0237E + 00	4.2219E − 02
17	1.9565E − 02	1.0190E + 00	3.9422E − 02
18	2.4879E − 02	1.0144E + 00	3.6365E − 02
19	2.5614E − 02	1.0117E + 00	3.3233E − 02
20	2.5439E − 02	1.0095E + 00	3.0181E − 02
21	2.5311E − 02	1.0072E + 00	2.7001E − 02
22	2.6900E − 02	1.0044E + 00	2.4087E − 02
23	2.2207E − 02	1.0042E + 00	2.0826E − 02
24	1.6891E − 02	1.0043E + 00	1.7676E − 02
25	1.1483E − 02	1.0045E + 00	1.4691E − 02
26	6.5646E − 03	1.0045E + 00	1.1452E − 02
27	2.5190E − 03	1.0043E + 00	8.3871E − 03
28	7.7322E − 04	1.0031E + 00	5.5264E − 03
29	4.3143E − 04	1.0015E + 00	2.7355E − 03

σ is the standard deviation of the residuals about the fitted line; velocities in meters per second.

and a complementary cumulative distribution in terms of percent (with a maximum value of 100) was computed and plotted. A power law of the form

$$P(\xi > V_{\text{eff}}/V_s(d)) = a(V_{\text{eff}}/V_s(d))^b \quad (6)$$

was then fit to the empirical distribution in a selected range of $V_{\text{eff}}/V_s(d)$. A variety of functions were tried, but the power law was simple and gave a good fit in most cases. Where the power law failed to provide a good fit [usually at small values of $V_{\text{eff}}/V_s(d)$ and deeper depths], P was given a value of 100 for arguments less than a subjectively chosen value. Figure 4 gives four examples of the fit at selected depths, and Table 3 gives values of the coefficients for all depths considered.

Equation (6) gives the probability $P(\xi > V_{\text{eff}}/V_s(d))$ of exceeding $V_{\text{eff}}/V_s(d)$ and can be used to decide if the site class should be changed, following this procedure:

1. Compute a provisional site class based on the simple extrapolation (equations 3 and 4).
2. Use equation (3) to solve for the $V_{\text{eff}}/V_s(d)$ needed to move to the next stiffer site class (softer site classes are not considered because in only a few cases from the set of 135 boreholes would the site class actually become softer; this example may not be universally applicable, for there may be regions in which velocities might generally decrease with depth in the upper tens of meters).
3. Evaluate equation (6) for this value of $V_{\text{eff}}/V_s(d)$ (paying attention to the values below which P is taken to be 100; see Table 3). If P equals 100, then the provisional site class is changed.
4. If P from the previous step does not equal 100, then a random number r uniformly distributed between 0 and 100 is generated. If $r \leq P$, then the provisional site class is changed to the next stiffer class (e.g., D to C), and if $r > P$ the class is set to the provisional site class (no change in class). If the procedure is applied many times, this step guarantees that the number of class changes will agree with the number expected from the probability value P that a change should occur.

Illustration of the Methods

The methods for estimating $\overline{V}_s(30)$ are illustrated using a subset of the 277 borehole models compiled by Boore (2003). The subset consists of 135 boreholes whose velocity models extend to at least 30 m. All of the boreholes are from California, with 5 from northern California, 38 from the San Francisco Bay area, 26 from central California, and 66 from southern California. Of those from southern California, about 36 are from the greater Los Angeles area and 18 are from the Imperial Valley. The formations near the surface are mostly Holocene in age, followed in number by those of Pleistocene age. A few boreholes are located at sites with highly weathered Mesozoic rocks at the surface. Except for the simplest method (equations 3 and 4), the use of this subset of boreholes does not constitute a true test of the methods because the same dataset used to derive the statistical quantities used in the methods is used for the application of the methods. A random subset of the 135 holes could have been selected for determination of the quantities in Tables 2 and 3 and the methods then tested against the rest of the holes. In view of the rather limited number of boreholes, I opted instead for a better determination of the statistical quantities. Thus the comparisons between actual and derived site classes are more a consistency check and an illustration of the methods than a test of the methods.

Depths ranging in 1-m increments from 10 to 29 m were chosen, assuming for each depth that the deeper part of the velocity model was unknown. Examples of the simple method and the method based on the statistics of $V_{\text{eff}}/V_s(d)$ are given in Table 4 for an assumed depth to bottom of 10 m. The entries in the columns will help in understanding the method. For example, for hole a, $V_{\text{eff}}/V_s(d) = 1.31$ is required to move the class based on simple extrapolation (D) to the next stiffer class (C); from equation (6) and Table 3, the probability P of such a value is 31%, but the random probability r is 79%, so no change was made in the class (the value of r depends on the particular seed used in the random number generator; other seeds might result in a change in site class). The examples in Table 4 were chosen to illustrate four cases: (1) both methods gave the same site class as the actual site class; (2) the simple extrapolation gave the wrong class, but the probabilty-based method gave the correct class; (3) the simple extrapolation gave the correct class, but the probabilty-based method gave the wrong

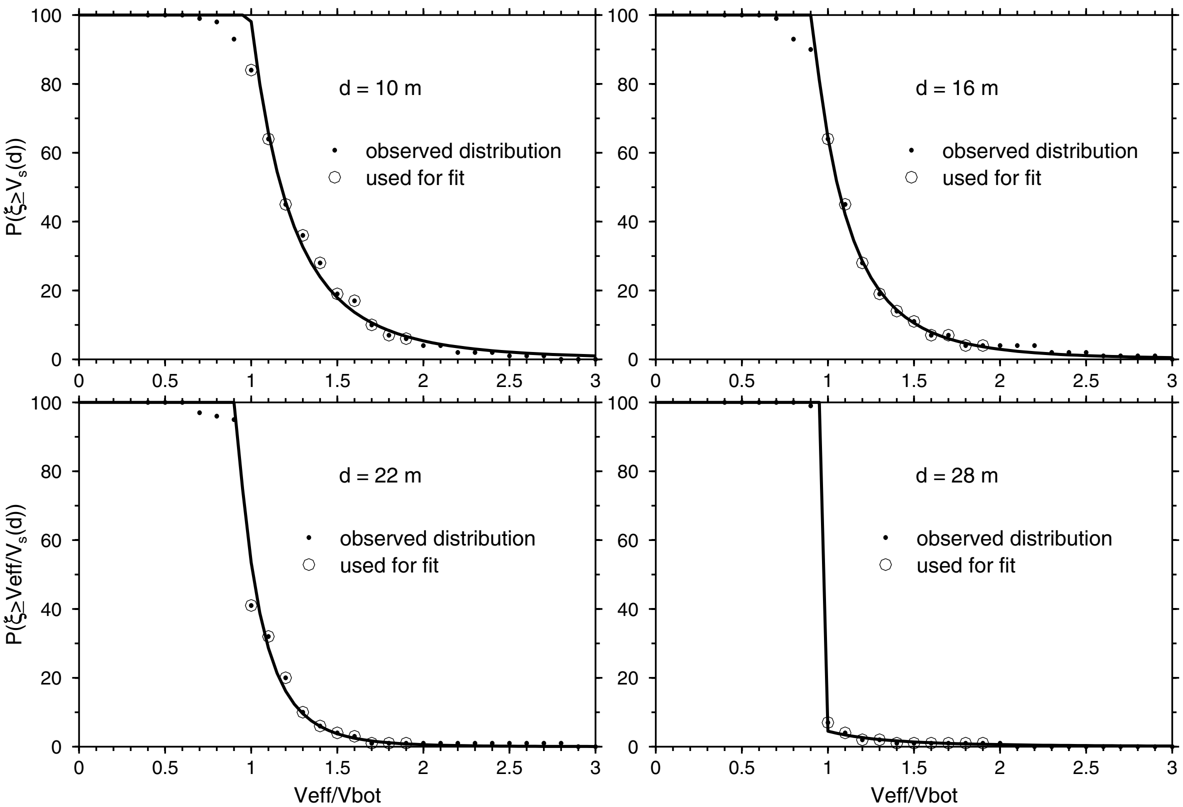

Figure 4. The complement of the cumulative probability of the ratio $V_{eff}/V_s(d)$ (see text) from 135 boreholes in California, for bottom depths of 10, 16, 22, and 28 m. Also shown are the power-law fits to the indicated observations (open circles), as used in the analysis [with P set equal to 100 for $V_{eff}/V_s(d)$ less than the values given in the fourth column of Table 3].

class; and (4) both extrapolation methods gave the wrong class.

For each method, an error count was kept for each assumed depth to bottom, with separate counts being kept for no change, changes to a stiffer class, and changes to a softer class, according to how far apart the reassigned site classes were from the actual site classes (in practice, this never exceeded one site class, so the count arrays were only incremented by unity). For example, if the real class were C but the reassigned class were D, then the array for erroneous changes to a softer site would be incremented by 1 (the array for erroneous changes to stiffer sites was incremented by −1). For each assumed depth, the cumulative number of erroneous changes was converted to a percentage of all boreholes with an erroneous change of site class. The results for all methods, plotted as a function of depth, are shown in Figure 5. Also included are the results of adding the percent of erroneous changes to stiffer and softer sites; this can be thought of as the total bias for each procedure. Figure 5 shows that the simple method gives a result biased toward softer sites, as expected. The biases for the other methods are significantly smaller for all depths. The version of the method using a Gaussian distribution about the regression relation (equation 5) gives somewhat smaller bias than using equation (5) by itself. The more fully probabilistic method using equation (6) gives better results than using the linear regression methods [but that method only yields an estimate of site class and not the continuous variable $\overline{V}_s(30)$].

Overall, the chance of an error in site classification is less than 10% for all but the shallowest depths; for the borehole dataset used in this article, with most holes closer to or deeper than 30 m (Fig. 1), the chance of an error is negligible using any method.

Discussion and Conclusions

Using 135 boreholes for which the velocity model extends to at least 30 m, several methods for extrapolating V_s to 30 m were investigated. Methods using the statistical properties of the relation between $\overline{V}_s(30)$ and shallower velocities resulted in significantly less bias for shallow models than the simple method of assuming that the lowermost velocity extends to 30 m. In addition, for all methods the percent of sites misclassified is generally less than 10% and falls to negligible values for velocity models extending to at least 25 m. The results suggest that determination of site classes for the USGS dataset of Boore (2003) will have few errors using any method. The results, however, might prove useful

Table 3
Coefficients of Power-Law Fit (Equation 6) to the Complementary Cumulative Distribution of Values of $V_{eff}/V_s(d)$

d(m)	a	b	ξ_{100}:use	ξ_{100}:pwr
10	98.053	−4.193	1.00	0.995
11	89.217	−4.461	0.97	0.975
12	91.365	−4.389	0.98	0.980
13	74.125	−3.773	0.92	0.924
14	63.179	−3.957	0.89	0.890
15	60.873	−4.090	0.89	0.886
16	64.418	−4.473	0.91	0.906
17	64.626	−4.499	0.91	0.908
18	52.342	−4.581	0.87	0.868
19	52.367	−4.129	0.85	0.855
20	54.560	−4.864	0.88	0.883
21	47.235	−6.291	0.89	0.888
22	53.445	−6.558	0.91	0.909
23	43.609	−7.170	0.89	0.891
24	35.723	−5.885	1.00	0.840
25	29.602	−5.314	1.00	0.795
26	13.790	−5.885	1.00	0.714
27	11.280	−4.416	1.00	0.610
28	4.488	−2.931	1.00	0.347
29	2.168	−3.165	1.00	0.298

The column headed ξ_{100}:use gives the values of $V_{eff}/V_s(d)$ below which P is set to 100 in use, and the column headed ξ_{100}:pwr gives the values of $V_{eff}/V_s(d)$ that yield $P = 100$ when inserted into equation (6).

Table 4
Examples of Determination of NEHRP Site Class using Simple Extrapolation and Probability-Based Extrapolation

Hole	$\overline{V}_s(30)$:a	$\overline{V}_s(30)$:x	$V_{eff}/V_s(d)$	P	r	Class:a	Class:x	Class:p
a	331	303	1.31	31	79	D	D	D
b	392	351	1.04	82	21	C	D	C
c	297	308	1.33	30	15	D	D	C
d	203	160	1.20	45	85	D	E	E

The values are for a velocity model stopping at 10 m. Velocities are in meters per second. In the column labels, ":a," ":x," and ":p" stand for actual values, extrapolated based on the simple model, and extrapolated values based on the probability model, respectively. The value of $V_{eff}/V_s(d)$ is that required to change to a stiffer class than given by the simple extrapolation. The actual boreholes are not identified because the procedure is not intended to make the proper prediction on a hole-by-hole basis; it should make the proper predictions on the average when applied to a large set of boreholes.

for other datasets, such as the K-NET dataset or other sites for which shear-wave velocity models might extend to depths significantly less than 30 m (such as many seismic cone penetrometer soundings).

An important caveat: none of the methods are likely to give a correct value of $\overline{V}_s(30)$ for a specific site. The procedures are inherently statistical, and over many sites the values should be correct on the average; the procedures only make sense when the data from many stations are being used in a statistical way, such as a regression analysis, and $\overline{V}_s(30)$ or site class for any particular site is not important.

The results here could be extended to determine the correlation between $\overline{V}_s(30)$ and the average velocity between a depth range that does not reach the surface (i.e., using a nonzero lower limit in equation 2). This would be useful for application to suspension log results, which never reach the surface. The results could also consider geological information, such as whether the site is expected to be underlain to 30 m by Holocene materials alone, Pleistocene materials alone, or some combination. As it is, the analysis is dominated by class NEHRP C and D sites, and the methods might not work as well for NEHRP class E or class B sites. A way of incorporating possible overall differences in the velocity models due to different types of sites is to use some measure of the gradient of the velocity model as a basis for extrapolation. This would also help deal with situations where the velocity might not generally increase with depth, as it does for the California boreholes used to illustrate the methods in this article.

Acknowledgments

I thank Tom Holzer, Chris Stephens, and two anonymous reviewers for comments that significantly improved the article. I also thank Tom Holzer and Mike Bennett for providing the seismic cone penetrometer data for the Oakland–Alameda region.

References

Atkinson, G. M., and D. M. Boore (2003). Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions, *Bull. Seism. Soc. Am.* **93,** 1703–1729.

Boore, D. M. (2003). A compendium of *P*- and *S*-wave velocities from surface-to-borehole logging: summary and reanalysis of previously published data and analysis of unpublished data, *U.S. Geol. Surv. Open-File Rept. 03-191*, 13 pp.; http://quake.usgs.gov/~boore (last accessed February 2004).

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.

Building Seismic Safety Council (BSSC) (2001). NEHRP recommended provisions for seismic regulations for new buildings and other structures, 2000 Edition, Part 1: Provisions, prepared by the Building Seismic Safety Council for the Federal Emergency Management Agency (Report FEMA 368), Washington, D.C.

Dobry, R., R. D. Borcherdt, C. B. Crouse, I. M. Idriss, W. B. Joyner, G. R. Martin, M. S. Power, E. E. Rinne, and R. B. Seed (2000). New site coefficients and site classification system used in recent building seismic code provisions, *Earthquake Spectra* **16,** 41–67.

Holzer, T. L., M. J. Bennett, T. E. Noce, A. C. Padovani, and J. C. Tinsley III (2002). Liquefaction hazard and shaking amplification maps of Alameda, Berkeley, Emeryville, Oakland, and Piedmont, California: a digital database, *U.S. Geol. Surv. Open-File Rept. 02-296*; http://geopubs.wr.usgs.gov/open-file/of02-296/ (last accessed February 2004).

Holzer, T. L., M. J. Bennett, T. E. Noce, A. C. Padovani, and J. C. Tinsley III (2004). Shear-wave velocity of surficial geologic sediments: statistical distributions and depth dependence, *Earthquake Spectra* **20,** (in press).

Wills, C. J., M. Petersen, W. A. Bryant, M. Reichle, G. J. 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,** S187–S208.

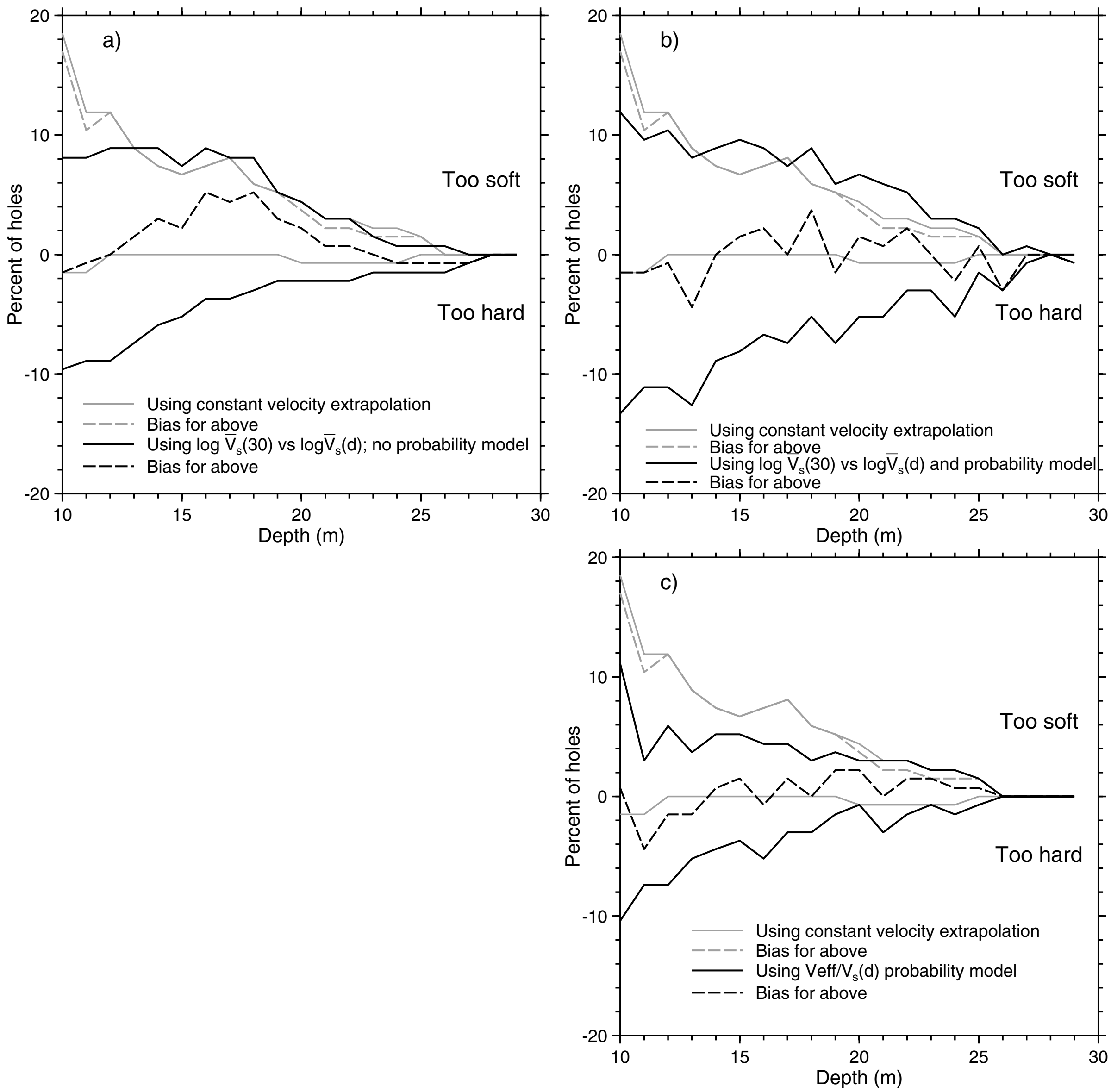

Figure 5. Summary of changes to the NEHRP classes using various methods to obtain $\overline{V}_s(30)$: (a) using the regression fit of log $\overline{V}_s(30)$ as a function of log $\overline{V}_s(d)$, without accounting for the scatter about the line; (b) using the regression fit and the scatter about the line; and (c) using the probability distribution of $V_{eff}/V_s(d)$. For comparison, the gray line in each graph shows the result of using simple extrapolation assuming a constant velocity from the bottom depth to 30 m. Nonzero values indicate incorrect class changes, as the percent of holes, with positive and negative values indicating a new class that is too soft and too hard compared to the actual class, respectively. The dashed lines are the overall bias, obtained by adding the number of sites misclassified as being too hard and being too soft.

U.S. Geological Survey
345 Middlefield Rd., MS 977
Menlo Park, California 94025
boore@usgs.gov

Manuscript received 29 May 2003.