

A Generic Soil Velocity Model that Accounts for Near-Surface Conditions and Deeper Geologic Structure

Marafi, N.A.¹, Grant, A.², Maurer, B.W.³, Rateria, G.⁴, Eberhard, M.O.⁵, and Berman, J.W.⁶

Near-surface soil conditions can significantly alter the amplitude and frequency content of incoming ground motions – often with profound consequences for the built environment – and are thus important inputs to any ground-motion prediction. Previous soil-velocity models (SVM) have predicted shear-wave velocity profiles based on the time-averaged shear-wave velocity in the upper 30 m (V_{S30}). This article presents a generic soil-velocity model that accounts both for near-surface conditions (V_{S30}) and deeper geologic structure, as represented to the depth at which the profile reaches a velocity of 1.0 km/s ($Z_{1.0}$). To demonstrate the advantages of our new SVM, we apply it to the Cascadia Region of North America, where numerous geologic basins and glaciated landscapes give rise to a wide range of V_{S30} and $Z_{1.0}$ combinations. This soil velocity model yields good estimates of site response across all site conditions, and significantly improves upon a model calibrated using only V_{S30} data. In conjunction with existing models that describe the deep velocity structure of the region (e.g., Stephenson et al., 2017), the proposed model is particularly suited for use in regional-scale predictions of site response, liquefaction, landslides, infrastructure damage, and loss. The proposed methodology is broadly applicable to the development of SVMs elsewhere, and with improved understanding of near-surface and deep velocity structures, can facilitate more accurate ground-motion predictions globally.

Keywords: Cascadia, soil velocity, site response, ground motion prediction

¹Senior Modeler, Risk Management Solutions Inc., Newark, CA 94560; Affiliate Assistant Professor, Dept. of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195

²Research Civil Engineer, U.S. Geological Survey, Moffett Field, CA 94035

³Assistant Professor, Dept. of Civil and Environmental Eng., University of Washington, Seattle, WA 98195

⁴Graduate Research Assistant, Dept. of Civil and Environmental Eng., University of Washington, Seattle, WA 98195

⁵Professor, Dept. of Civil and Environmental Eng., University of Washington, Seattle, WA 98195

⁶Professor, Dept. of Civil and Environmental Eng., University of Washington, Seattle, WA 98195

INTRODUCTION

Subsurface seismic-wave velocity (i.e., site condition) can significantly affect ground-motion amplitude and frequency content. Reliable estimates of these velocities are thus needed to better predict ground motions and coseismic phenomena, such as liquefaction, landslides, infrastructure damage, and loss. For situations where detailed in-situ measurements are not available, such as regional-scale hazard assessments, estimates of subsurface conditions (e.g., via a soil-velocity model) are needed to approximate site response.

Accordingly, efforts have been made to better understand and predict seismic wave velocity profiles in the absence of direct measurement (e.g., among many, Boore and Joyner, 1997; Holzer et al., 2005; Allen and Wald, 2007; Castellaro et al., 2008; Boore et al., 2011; Boore, 2016; Parker et al., 2017). These efforts have typically focused more on predicting the time-averaged shear-wave velocity in the upper 30 m (V_{S30}), a required input to empirical ground-motion prediction equations, and less on explicitly predicting the variation of shear-wave velocity (V_S) with depth, a required input to ground-motion prediction methods based on wave propagation (e.g., Schnabel et al., 1972).

Boore and Joyner (1997) proposed a soil-velocity model (SVM) for the western U.S. that has been used in many applications (e.g., Frankel et al. 2018). More recently, Shi and Asimaki (2018) proposed a California-based SVM conditioned on V_{S30} and based on the functional form of Vrettos (1996). Provided a prediction of V_{S30} (e.g., using a proxy- or geology-based estimate), this model predicts V_S as a function of depth (z):

$$V_s(z) = \begin{cases} V_{S0} & , z < 2.5 \text{ m} \\ V_{S0}(1 + k(z - 2.5))^{\frac{1}{n}} & , z \geq 2.5 \text{ m} \end{cases} \quad (1)$$

where V_{S0} , k , and n are fitting parameters conditioned on V_{S30} and defined in Shi and Asimaki (2018). This model: (1) provides depth-continuous predictions of near-surface stratigraphy; (2) can be implemented at relatively fine spatial resolution (i.e., the resolution of the input V_{S30}); and (3) was fit to a large set of measured profiles distributed throughout California. However, having been based only on data from California, the applicability of the Shi and Asimaki (2018) model to other regions is unknown.

Soil velocity models are particularly important for physics-based simulations of earthquake ground motions and consequent impacts at a regional scale. Such simulations are especially useful

in regions with a paucity of historical records (e.g., where the earthquake return-period exceeds the observation interval), examples of which include the Alpine Fault Zone of New Zealand (Bradley et al., 2017), the New Madrid Seismic Zone of the Central U.S. (Ramirez-Guzman et al., 2015), and the Cascadia Subduction Zone of North America (Frankel et al., 2018; Wirth et al., 2018). In addition to populating gaps in ground-motion datasets, physics-based simulations help to elucidate and quantify complex ground-motion phenomena (e.g., the effects of directivity, basins, and topography) via explicit modeling of kinematic fault rupture, wave propagation, and the subsurface velocity structure. With respect to the latter, physics-based simulations often rely on so-called “Community Velocity Models” (CVMs), which map predictions of V_s versus depth across large areas. While numerous CVMs have been developed worldwide (e.g., among many, Magistrale et al., 2009; Cramer et al., 2016; Small et al., 2017; Stephenson et al., 2017), these models often: (1) are concerned more with predicting deep geologic structure, such as basin geometry, and less with near-surface conditions that can also significantly alter motions; and (2) have coarse spatial resolution. For example, Stephenson et al. (2017) provide V_s profiles for approximately 60 million hectares of the Cascadia Subduction Zone. While this model was recently utilized in physics-based simulations (e.g. Frankel et al. 2018, Wirth et al. 2018), it: (1) has a minimum V_s of 600 m/s, which corresponds to soft rock, and thus ignores the presence and influence of softer soils; and (2) has a minimum grid spacing at the ground surface of 200 m laterally and 100 m vertically, which is a coarser resolution than is typically used in geotechnical analysis.

In this paper, we propose a new soil-velocity model that accounts not only for V_{s30} , but also for the deeper velocity structure (which could be characterized by CVMs). Such a model would be particularly useful for regions in which the deeper velocity structure varies greatly. For example, the Cascadia region, considered herein to include Oregon, Washington, and southern British Columbia (Canada), is a geologically and topographically complex region. The Cascadia forearc is composed of multiple coastal mountain ranges, a north-south spanning lowland, and the Cascade volcanic arc. Deep sedimentary basins underly major population centers throughout the region (e.g., the Georgia, Everett, Seattle, Tacoma, and Tualatin basins) that significantly affect surface ground motions compared to equivalent non-basin sites (e.g., Marafi et al., 2017, Wirth et al., 2018; Roten et al., 2019). How shallow velocity profiles in basin and non-basin sites may differ is therefore of critical interest for the development of regional ground-motion predictions that incorporate near-surface soil conditions.

83 To provide a baseline with which to compare the performance of the proposed model, the
 84 performance of the Shi and Asimaki (2018) model, based solely on V_{S30} data, is herein evaluated
 85 for the Cascadia region. By accounting for the deeper geologic structure, our proposed model
 86 provides better estimates of shear-wave velocity with depth, and by corollary, better estimates of
 87 local site amplification for a wide range of site and geological conditions, as compared to Shi and
 88 Asimaki (2018).

89 Evaluation of Shi and Asimaki (2018) Model in the Cascadia Region

90 The performance of the Shi and Asimaki (2018) (SA18) SVM in Cascadia was evaluated using
 91 909 V_S profiles compiled by Ahdi et al. (2017) from Oregon, Washington, and British Columbia
 92 (Canada). These profiles were digitized from reports by state and national geologic surveys (e.g.,
 93 Washington Division of Natural Resources, U.S. Geological Survey). For this dataset, Figure 1
 94 quantifies the accuracies of SA18 predictions as a function of depth via the difference between the
 95 natural-log of the predicted V_S ($\ln V_{S,SA18}$) and the natural-log of the measured V_S ($\ln V_S$), termed
 96 the V_S residual. As shown in Figure 1, the medians of the V_S residuals are relatively small at depths
 97 less than 50 m, and again, for depths between 600 m and 1000 m. For example, the 50th percentile
 98 V_S residual reaches peak values of: +0.14 in the upper 50 m (~15% overprediction of V_S) and -0.19
 99 between depths of 600 m and 1000 m (~17% underprediction). In contrast, the velocity estimates
 100 exhibit strong bias at depths between 50 m and 600 m, in which the 50th percentile V_S residual
 101 reaches peak values of +0.84 (~132% overprediction). In addition, the variance of residuals
 102 indicates that even at depths where the median residual is zero, 32% of V_S profiles are either under-
 103 or overpredicted by at least 32%.

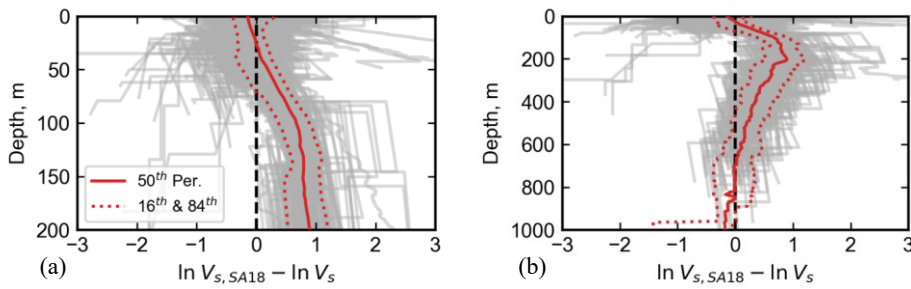


Figure 1. Shear-wave velocity (V_S) residuals (predictions from Shi and Asimaki, 2018) versus depth for 909 profiles compiled by Ahdi et al. (2017) for depths up to (a) 200 m; and (b) 1000 m.

Thus, when SA18 is applied to Cascadia, it: (1) tends to severely overpredict V_S in some settings, which will later be identified as thick soil deposits in basins; and (2) has large prediction variance, even when the median residual is zero. This suggests that motions predicted at the ground surface would, in some locations, be inaccurate for many spectral periods. Accordingly, we propose a new model that introduces conceptual improvements to the SA18 formulation by integrating information about the deeper geologic structure. A Cascadia-specific SVM is then developed using V_S profiles from the Cascadia region. Wave-propagation-based site-response analyses were used to quantify the performance of the new Cascadia SVM and identify the geologic conditions where these improvements are most critical to accurately predicting earthquake ground motions.

PROPOSED MODEL

A modified version of the SA18 model is proposed for predicting $V_S(z)$:

$$V_S(z) = \begin{cases} V_{S0} & , z < 2.5 \text{ m} \\ V_{S0} + 1000 \cdot \left(k \frac{z-2.5}{Z_{1.0}-2.5} \right)^{\frac{1}{n}} & , z \geq 2.5 \text{ m} \end{cases} \quad (2a)$$

where V_{S0} defines V_S at the surface; k controls the initial, near-surface rate-of-change in V_S ; n controls the rate-of-change in V_S at greater depths; $Z_{1.0}$ is the depth, in meters, where $V_S = 1$ km/s; V_S and V_{S0} have units of m/s; and z is depth in meters.

Prior to settling on the functional form of the proposed SVM, other forms such as polynomials were tested but deemed unfavorable because the equations: (a) highly deviated from the Shi and Asimaki (2018) model; (b) became too complex/non-unique to fit, with many regressed coefficients; (c) provided no additional accuracy; and (d) resulted in unrealistic predictions of soil profiles with input variables that were outside the parameter space of the dataset. The proposed functional form is believed to provide a “balanced” solution.

The proposed SVM builds conceptually on SA18 (Eq. 1) in two ways. First, the function defined in Eq. 1 is modified so that parameters k and V_{S0} have reduced interaction. This parameter interaction is not consistent with the Cascadia data and results in non-unique solutions, such that various model inputs result in a similar V_S profile, complicating model optimization. Second, to benefit from, and coalesce with, CVMs that map deep geologic structure, the predicted V_S is

anchored to 1000 m/s at depth $Z_{1.0}$, which roughly corresponds to the top of tertiary sedimentary rock. During model development, $Z_{1.0}$ was obtained from field measurements, whereas in forward use, it can be obtained from a regional CVM (e.g., Stephenson et al., 2017). Values of $n < 1$ result in convex profiles, such that the rate-of-change in V_s increases with depth; values of $n > 1$ result in concave profiles, such that the rate-of-change in V_s decreases with depth; and values of $n = 1$ result in constant rate-of-change in V_s . To anchor the predicted V_s at $Z_{1.0}$, the parameter k is defined as:

$$k = \left(\frac{1000 - V_{s0}}{1000} \right)^n \quad (2b)$$

Accordingly, to apply Eqs. 2a and 2b at any given location requires input variable $Z_{1.0}$ and model parameters V_{s0} and n . The prediction of these parameters will be discussed subsequently.

Shear Wave Velocity Profile Dataset

A subset of 218 V_s profiles from the Ahdi et al. (2017) compilation where $Z_{1.0}$ was directly measured from the velocity profile was used in this study. These data were compiled from multiple sources and methods (including downhole, seismic cone, P-S suspension, and various non-invasive geophysical methods) and detailed in Ahdi et al. (2017). Shown in Figures 2 and 3, respectively, are the V_{s30} - $Z_{1.0}$ parameter space and geographic locations for this subset. It can be seen in Figure 2 that the V_{s30} parameter space extends from 112 m/s to 871 m/s and is consistently populated, whereas the $Z_{1.0}$ parameter space extends from 5 m to 890 m but is unpopulated between ~100 m and 200 m. This gap in $Z_{1.0}$ corresponds to a geographic clustering, as mapped in Figure 3, such that profiles with $Z_{1.0} > 200$ m are mostly from sites located in the Fraser River Delta of British Columbia (180 profiles), whereas all profiles with $Z_{1.0} < 200$ m are from sites in Oregon and Washington (38 profiles). The profiles are sourced predominantly from surficial geologic units mapped as Quaternary sediment, including 70% in alluvium variants (e.g., loam, overbank, flood) and 15% in peat.

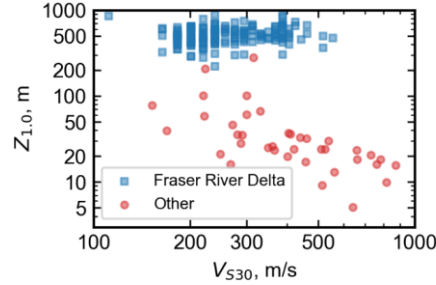


Figure 2. Distribution of V_{S30} and Z_{L0} from measured profiles in which V_s reaches 1000 m/s.

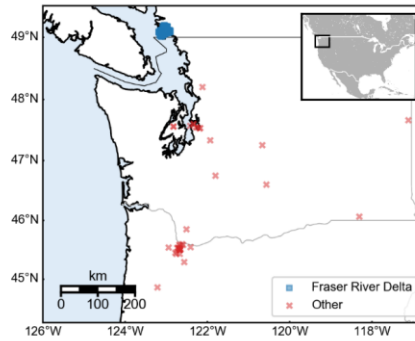


Figure 3. Map of measured profiles in which V_s reaches 1000 m/s.

Notably, a scarcity of Z_{L0} measurements eliminated 76% of the Ahdi et al. (2017) data from the dataset, since Z_{L0} is required in Eq. 2a. This scarcity is attributable to the difficulty and expense of deep profiling, which is often motivated in British Columbia by petroleum exploration (e.g., Hunter et al., 1998). In this regard, the lack of deep Z_{L0} measurements elsewhere should not be inferred as an absence of such conditions. To the contrary, profiles with $Z_{L0} > 200$ m are abundant in the Puget-Willamette lowland of Oregon and Washington (e.g., in the Everett, Seattle, Tacoma, Portland, and Tualatin sub-basins), where Z_{L0} may occasionally reach 1000 m (Yount et al., 1985; Myer et al., 2005; McPhee et al., 2014; Ahdi et al., 2017; Stephenson et al., 2017). Whereas V_s measurements are plentiful within these basins, the measurements rarely exceed a depth of 100 m.

Alternative depth to velocity (Z_x) inputs to Eq. 2 were explored (e.g., $Z_{0.6}$) to increase the available points in the dataset. However, such alternatives: (1) only marginally increased the size of the dataset; (2) did not measurably improve the accuracy of the final SVM; and (3) are not as

common as $Z_{1,0}$ in ground motion modeling, thereby potentially hindering transferability of the proposed model to other regions. Ultimately, as shown in Figure 2, there is a need for deep, continuous profiling of sedimentary basins in Washington and Oregon, which will confirm or update the model proposed herein.

Fitting Model Parameters

Using the subset of the Ahdi et al. dataset, model parameters V_{S0} and n were determined for each profile. This was achieved by minimizing the error between the measured V_S and that predicted by Eq. 2, with error for each profile computed as:

$$\varepsilon_{VS} = \frac{1}{N} \sum_{i=1}^N |\ln V_{s,predicted}(z_i) - \ln V_{s,measured}(z_i)| \quad (3)$$

In Eq. 3, each profile is discretized into 1-m thick layers, z_i is the depth from the surface to the midpoint of layer i , and N is the total number of layers for that profile. Only sediment layers with $V_s \leq 1000$ m/s were included in the error term, since the proposed model is intended to coalesce with a CVM at $Z_{1,0}$. Optimal parameters V_{S0} and n were found using the Nelder-Mead algorithm (Gao and Han, 2012, with a convergence tolerance of $1e-8$) to minimize ε_{VS} , which can be expressed as:

$$\{V_{S0}, n\} = \arg \min \varepsilon_{VS} \quad (4)$$

As examples of the results of this procedure, Figure 4 shows the fitted profiles from two sites: (a) a low- V_{S30} , high- $Z_{1,0}$ site; and (b) a high- V_{S30} , low- $Z_{1,0}$ site. In each case the optimal V_{S0} and n are shown, as is the prediction of V_S -versus-depth from SA18 (i.e., Eq. 1). It can be seen that SA18 is problematic at the high- $Z_{1,0}$ site, where V_S is difficult to predict in the absence of $Z_{1,0}$ information.

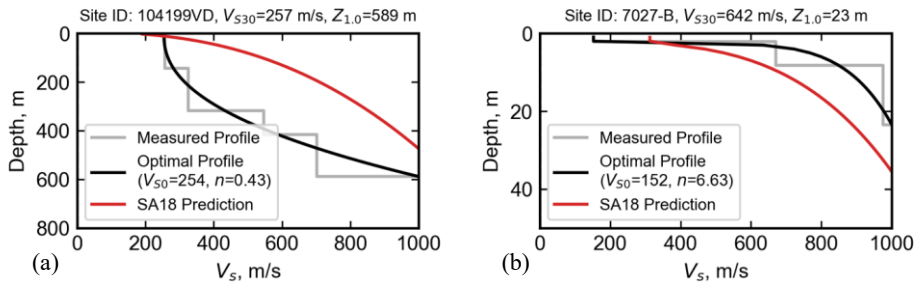


Figure 4. Shear-wave velocity (V_s) versus depth for (a) a low- V_{S30} , high- $Z_{1.0}$ site; and (b) a high- V_{S30} , low- $Z_{1.0}$ site.

Repeating this process for all profiles, Figure 5 summarizes the optimal, fitted model parameters, plotted with respect to each profile's measured V_{S30} and $Z_{1.0}$. As can be observed from the binned, average behavior in Figure 5a, V_{S0} tends to increase as V_{S30} increases. This was also observed in California for the SA18 model and is expected, given the link between V_s at the ground surface and V_s time-averaged over the top 30 m. In contrast, V_{S0} shows no correlation to $Z_{1.0}$ for the dataset (Figure 5b). While it is mathematically necessary that V_{S0} approach 1000 m/s as $Z_{1.0}$ approaches zero, the lack of any other trend is unsurprising, given that surficial soil deposits may be stiff or weak, and shallow or deep, independent of the rock that lies beneath.

Parameter n , which relates to the rate-of-change in V_s at greater depths, increases with increasing V_{S30} (Figure 5c) and decreases with increasing $Z_{1.0}$ (Figure 5d). This indicates that in profiles with high V_{S30} and low $Z_{1.0}$, V_s tends to increase rapidly at shallow depth and gradually thereafter, whereas in profiles with low V_{S30} and high $Z_{1.0}$, V_s tends to increase gradually at shallow depth, then more rapidly at depths approaching $Z_{1.0}$. While these behaviors were expected, the inclusion of $Z_{1.0}$ in Eq. 2, and preservation of reasonable V_{S30} - $Z_{1.0}$ behavior, represents an important conceptual advance. As will be demonstrated, this can result in significantly more accurate estimates of site-response relative to models that do not account for deeper geologic structure.

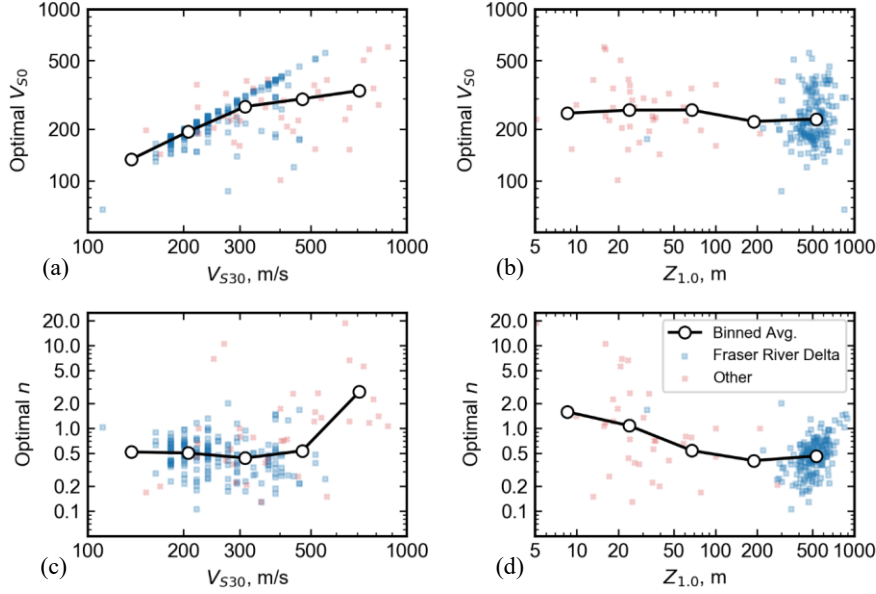


Figure 5. Optimal values of model parameter V_{S0} as a function of measured (a) V_{S30} and (b) $Z_{1.0}$; Optimal values of model parameter n as a function of measured (c) V_{S30} and (d) $Z_{1.0}$.

In addition to the trends investigated in Figure 5, correlations between the model parameters (i.e., n and V_{S0}) and other potential predictor variables were explored, including mapped glacial history, surficial geologic unit, and various geospatial parameters (e.g., surface mineralogy, roughness, and wetness; distance to, and elevation above, rivers, streams, and other water bodies; and compound-topographic-index). Some of these predictors showed promise for future study, particularly with a larger dataset, but they were not included in the current model. In this regard, future research could identify and include additional predictors of V_S as part of global modeling efforts. In the next section, the prediction of model parameters V_{S0} and n is explored using predictor variables V_{S30} and $Z_{1.0}$.

Predicting Model Parameters

Informed by the trends observed in Figure 5a, parameter V_{S0} is modelled using input variable V_{S30} and the following functional form:

$$V_{S0} = a_0 + a_1(V_{S30})^{a_2} \quad (5)$$

217 where a_0 is the y-axis intercept (i.e., V_{S0} when $V_{S30} = 0$); a_1 and a_2 controls the rate of variation of
 218 V_{S0} with V_{S30} . Coefficients a_0 , a_1 , and a_2 were determined by minimizing the prediction error of
 219 V_{S0} , expressed as:

$$220 \quad \{a_0, a_1, a_2\} = \arg \min \frac{1}{m} \sum_{j=1}^m |\ln V_{S0,j} - \ln V_{S0-optimal,j}| w_j \quad (6)$$

221 where V_{S0} is computed by Eq. 5; $V_{S0-optimal}$ is the optimal value of V_{S0} computed for each profile by
 222 Eq. 4; and w is a profile-specific weighting factor that mitigates bias from geographic clustering
 223 of the dataset. In particular, a majority of the profiles were sourced from the Fraser River Delta, as
 224 shown in Figures 2 and 3. Accordingly, w normalizes errors in proportion to the number of profiles
 225 from the Fraser Delta and elsewhere, such that both subsets have equal weighting. Ideally, the
 226 weighting function should be further disaggregated to account for error bias from other regions
 227 (i.e., Seattle and Portland), however, doing so would have resulted in an unreasonably small
 228 sample size by which to normalize errors. Following this approach, coefficients a_0 , a_1 , and
 229 a_2 were found using the Nelder-Mead algorithm (Gao and Han, 2012, with a convergence
 230 tolerance of $1e-8$), and are summarized in Table 1. These coefficients give physically reasonable
 231 predictions for all $V_{S30} > 100$ m/s, although the dataset is populated only from $112 \text{ m/s} < V_{S30} <$
 232 871 m/s . The resulting calibrated model for parameter V_{S0} is a function of V_{S30} is plotted in Figure
 233 6.

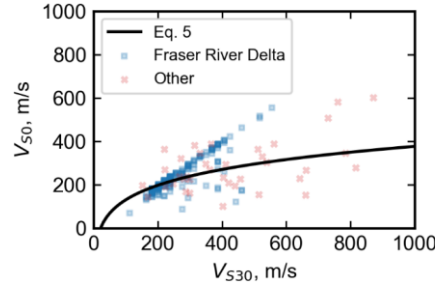
234

235

Table 1. Optimized Model Coefficients

Coefficient	Value	Coefficient	Value
a_0	-629	b_0	0.00912
a_1	434	b_1	0.646
a_2	0.122	b_2	-0.201
-	-	b_3	0.136

236

**Figure 6.** Model parameter V_{S0} predicted by Eq. 5 as a function of V_{S30} . Markers indicate the fitted, or optimal, V_{S0} values for each V_S profile.

237 Informed by the trends observed in Figures 5c and 5d, parameter n is predicted using input
 238 variables V_{S30} and $Z_{1.0}$ with the following functional form:

$$239 \quad n = b_0(V_{S30})^{b_1}(Z_{1.0})^{b_2}(V_{S30}Z_{1.0})^{b_3} \quad (7)$$

240 where b_0 is the initial slope, b_1 is the exponent on the V_{S30} term, b_2 is the exponent on the $Z_{1.0}$ term,
 241 and b_3 is the exponent on the $V_{S30}Z_{1.0}$ interaction term. Coefficients b_0 , b_1 , b_2 , and b_3 were
 242 determined by minimizing the prediction error of n , expressed as:

$$243 \quad \{b_0, b_1, b_2, b_3\} = \arg \min \frac{1}{m} \sum_{j=1}^m |\ln n_j - \ln n_{optimal,j}| w_j \quad (8)$$

244 where n is computed by Eq. 7; $n_{optimal}$ is the optimal n for each profile computed by Eq. 4; and w
 245 is as previously defined. Coefficients b_0 , b_1 , b_2 , and b_3 were again determined using the Nelder-
 246 Mead algorithm (Gao and Han, 2012, with a convergence tolerance of 1e-8) and are summarized
 247 in Table 1. These coefficients give physically reasonable predictions for all V_{S30} and $Z_{1.0}$.

248 The resulting predicted values for parameter n are plotted as a function of V_{S30} and $Z_{1.0}$ as a
 249 color contour in Figure 7a. Parameter k , which depends on predictions of both V_{S0} and n (see Eq.
 250 2b), is depicted similarly in Figure 7b. Consistent with trends previously observed in individual
 251 profiles, sites with low- V_{S30} and high- $Z_{1.0}$ have predicted $n < 1$ and predicted k approaching 1.0,
 252 meaning that the predicted V_S increases gradually at shallow depth, then more rapidly at depths
 253 approaching $Z_{1.0}$. In contrast, sites with high- V_{S30} and low- $Z_{1.0}$ have predicted $n > 1$ and predicted
 254 k approaching zero, meaning that the predicted V_S increases rapidly at shallow depth and gradually
 255 thereafter. Also shown in Figure 7 is the V_{S30} - $Z_{1.0}$ parameter space of the dataset (transparent
 256 circles). It should be noted that predictions beyond this space are extrapolations of the data.

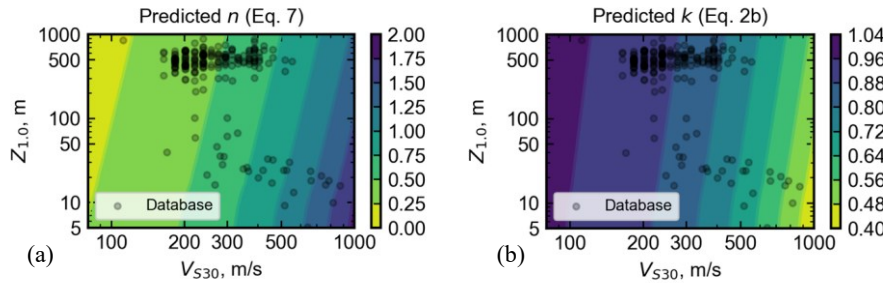


Figure 7. Model parameters (a) n and (b) k predicted by Eqs. 7 and 2b, respectively, as a function of V_{S30} and $Z_{1.0}$.

257 To assess whether these predictions exhibit bias for any ranges of V_{S30} and $Z_{1.0}$, prediction
 258 residuals are quantified in Figure 8 as the difference between the natural-log of the predicted
 259 parameter ($\ln V_{S0,pred}$ or $\ln n_{pred}$) and the natural-log of the optimal parameter ($\ln V_{S0,optimal}$ or $\ln n_{optimal}$). Figures 8a and 8b show the V_{S0} residuals, which average -0.023 and are nearly constant
 260 with respect to V_{S30} and $Z_{1.0}$. Similarly, Figures 8c and 8d show the n residuals, which average -
 261 0.013 but show significant negative bias for the highest- $Z_{1.0}$ bin in Figure 8d. This bin only contains
 262 five profiles, so it is possible that it is a statistical outlier, but it is also possible that the model
 263 underestimates the concavity of the deepest profiles. Accordingly, we conclude that V_{S0} - and n -
 264 residuals are acceptable with respect to the model's input variables, V_{S30} and $Z_{1.0}$, but also
 265 acknowledge that additional field data are needed for further model calibration.

267

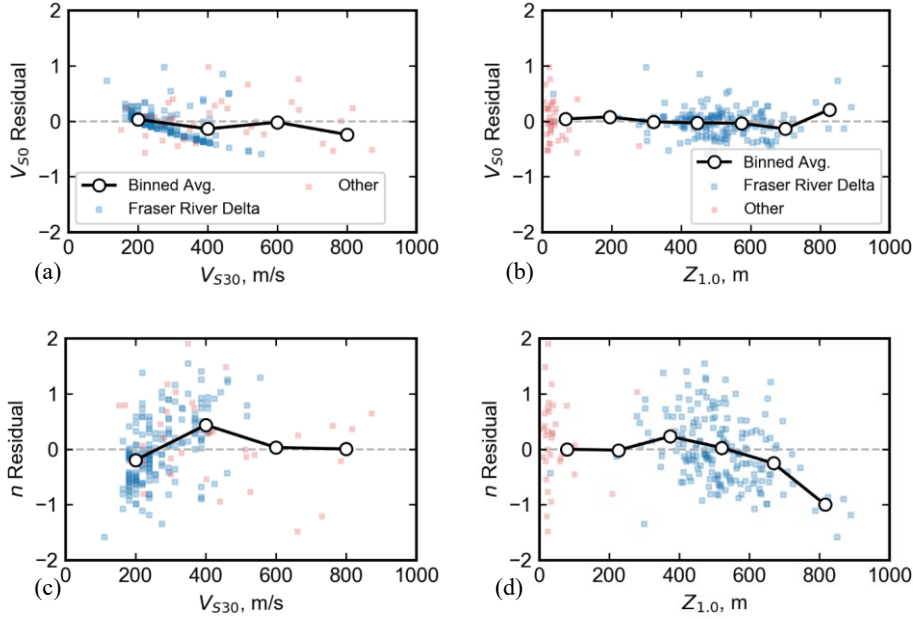


Figure 8. (a) V_{50} residuals with respect to V_{S30} ; (b) V_{50} residuals with respect to $Z_{1.0}$; (c) n residuals with respect to V_{S30} ; and (d) n residuals with respect to $Z_{1.0}$ (All residuals are in natural log units).

EVALUATION OF PROPOSED MODEL

The proposed SVM is defined by Eqs. 2, 5, and 7, along with the coefficients provided in Table 1. Figure 9 shows examples of V_s profiles predicted by the SVM, corresponding to $V_{S30} = 200$, 400, 600, and 800 m/s, and $Z_{1.0} = 50, 200$, and 400 m. These predictions are consistent with trends observed in the dataset. Extrapolations beyond the parameter space of the dataset are physically reasonable, though the SVM is intended to coalesce with a CVM at $V_s = 1000$ m/s and was not trained on measurements exceeding 1000 m/s. In the ensuing sections, we evaluate the performance of the proposed model in several ways.

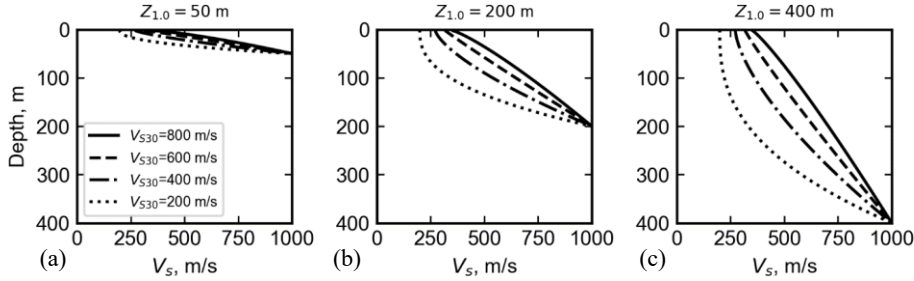


Figure 9. Shear-wave velocity with respect to depth for sites with various V_{S30} (in m/s) and (a) $Z_{1.0} = 50$ m; (b) $Z_{1.0} = 200$ m; and (c) $Z_{1.0} = 400$ m.

Predicted Profile V_{S30} Versus Input V_{S30}

Given that V_{S30} is an input to the SVM (and may be estimated using one of several proxy- or geology-based models), predicted profiles would ideally have a V_{S30} consistent with the assumed input. To evaluate this consistency, V_{S30} residuals are quantified in Figure 10a and defined as $\ln(V_{S30,\text{predicted}}) - \ln(V_{S30,\text{input}})$, where the latter is the input value and the former is V_{S30} computed from the resulting, predicted V_s profile (i.e., using Eq. 2). These residuals were computed for a wide range of values of V_{S30} and $Z_{1.0}$ and are depicted in Figure 10a as a color contour. An analogous analysis was performed for SA18 and is shown in Figure 10b. In each case, the V_{S30} - $Z_{1.0}$ parameter space of the Cascadia dataset is also plotted.

The computed values of V_{S30} for the SA18 profiles are consistent with the input V_{S30} , with 96% of dataset profiles having a V_{S30} residual less than ± 0.1 (i.e., $\pm 10\%$ error). However, SA18 gives undefined or implausible results for $V_{S30} < \sim 140$ m/s. Using the proposed Cascadia SVM, 53% of dataset profiles have a V_{S30} residual less than ± 0.1 and 73% have a residual less than ± 0.25 . It can also be seen in Figure 10a that some combinations of V_{S30} and $Z_{1.0}$ result in very large residuals. However, these combinations are generally unlikely to exist, low V_{S30} (< 200 m/s) and low $Z_{1.0}$ (< 40 m), or high V_{S30} (> 500 m/s) and high $Z_{1.0}$ (> 50 m), having no representation in the Ahdi et al. (2017) dataset. Lastly, while strict compliance with V_{S30} is ideal, such constraint can lead to large errors at greater depths. In this respect, the Cascadia SVM represents a compromise, reproducing V_{S30} adequately but not identically, in favor of better prediction of V_s at $z > 30$ m and overall improved site response estimation (subsequent section).

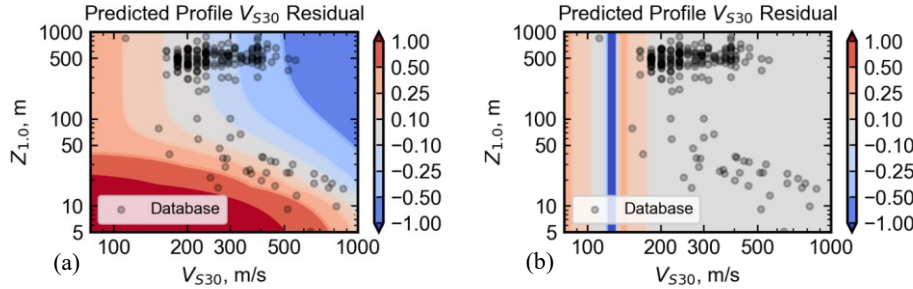


Figure 10. V_{S30} residual, as a function of input $Z_{1,0}$ and V_{S30} , for the (a) proposed Cascadia SVM; and (b) SA18 SVM.

Predicted Versus Measured V_S Profile

To evaluate the proposed SVM's performance at all depths, V_S residuals were computed in the same manner done for SA18 in Figure 1. However, to provide a consistent comparison, the evaluation of SA18 is repeated for the subset of profiles and is shown in Figure 11. These results are very similar to those in Figure 1, from which it may be inferred that the subset profiles are representative of the larger Ahdi et al. (2017) dataset. SA18 predictions are again relatively accurate for $z < 50$ m and $z > 600$ m but tend to have much larger errors at depths in between. Specifically, the median V_S residual reaches peak values of: -0.25 in the upper 50 m (~28% underprediction); -0.18 between depths of 600 m and 1000 m (~16% underprediction); and +0.85 between depths of 50 m and 600 m (~150% overprediction). The largest overpredictions generally correspond to soft, thick soil deposits in deep basins, which are underrepresented in the SA18 dataset.

Residuals for the Cascadia SVM, incorporating both V_{S30} and $Z_{1,0}$, are shown in Figure 12. Here, the median V_S residual reaches peak values of: -0.088 in the upper 50 m (~8% underprediction); -0.33 between depths of 600 m and 1000 m (~28% underprediction); and -0.13 between depths of 50 m and 600 m (~12% underprediction). Thus, the proposed SVM performs significantly better at depths up to 600 m, beyond which the model accuracies are similar. In addition, the variance of residuals is less in the upper 100 m using the proposed SVM and similar to SA18 at other depths.

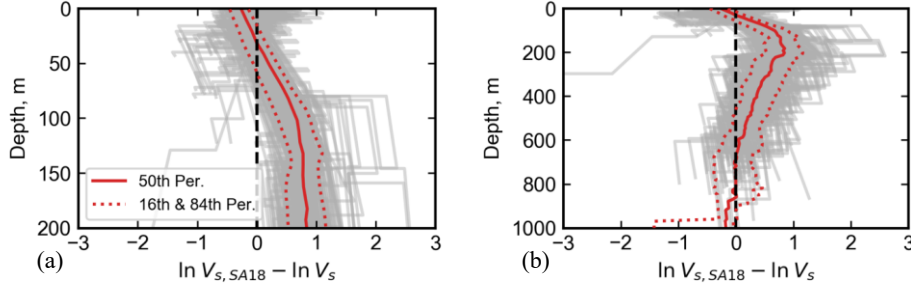


Figure 11. V_s residuals (predictions from SA18) for (a) depths up to 200 m; and (b) depths up to 1000 m, considering 218 profiles compiled by Ahdi et al. (2017). Positive residual indicate an overprediction by the model (i.e., SA18).

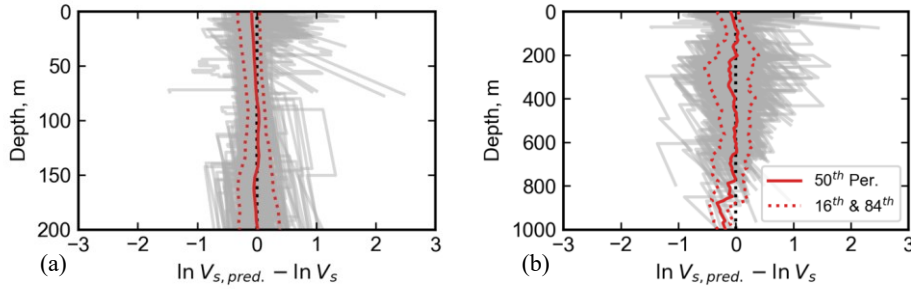


Figure 12. V_s residuals (predictions from our proposed Cascadia SVM) for depths: (a) up to 200 m; and (b) up to 1000 m, considering 218 profiles compiled by Ahdi et al. (2017). Positive residual indicate an overprediction by the model (i.e., our SVM).

To assess whether these predictions exhibit bias (i.e., consistently positive or negative residuals) for any ranges of V_{S30} and $Z_{1.0}$, the residuals from Figure 11 and 12 are parsed into four bins in Figure 13 and 14: (a) low V_{S30} and low $Z_{1.0}$, (b) low V_{S30} and high $Z_{1.0}$, (c) high V_{S30} and low $Z_{1.0}$; and (d) high V_{S30} and high $Z_{1.0}$. Considering depths up to 200 m, the average absolute values of the V_s residuals (using the proposed SVM) in these respective bins are 0.44, 0.15, 0.31, and 0.25 (Figure 14). By comparison, and as shown in Figure 13, SA18 produces analogous V_s residuals in these respective bins of 0.45, 0.89, 0.46, and 0.53. This suggests the proposed SVM better predicts V_s than SA18 for all bins. Notably, both SVMs resulted in larger residuals for profiles with low V_{S30} and low $Z_{1.0}$ (Figure 13a and Figure 14a). However, as previously discussed,

the dataset contains relatively fewer profiles with low $Z_{1.0}$. In this regard, a larger set of field measurements is needed to assess this behavior.

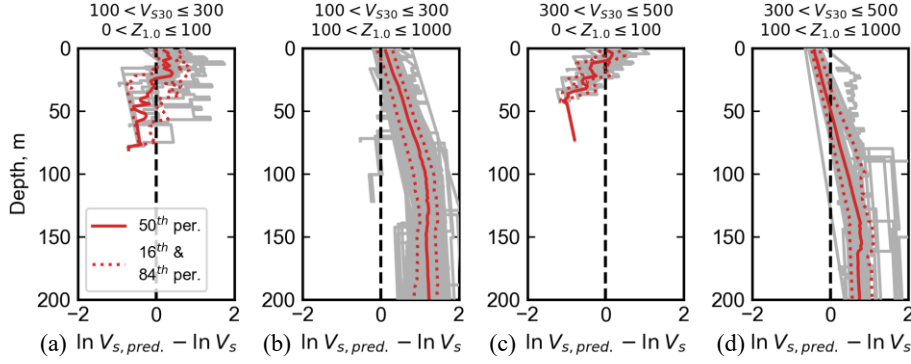


Figure 13. V_s residuals (predictions from SA18) from Figure 11 binned on: (a) low V_{S30} and low $Z_{1.0}$, (b) low V_{S30} and high $Z_{1.0}$, (c) high V_{S30} and low $Z_{1.0}$; and (d) high V_{S30} and high $Z_{1.0}$. Positive residual indicate an overprediction by the model (i.e., SA18).

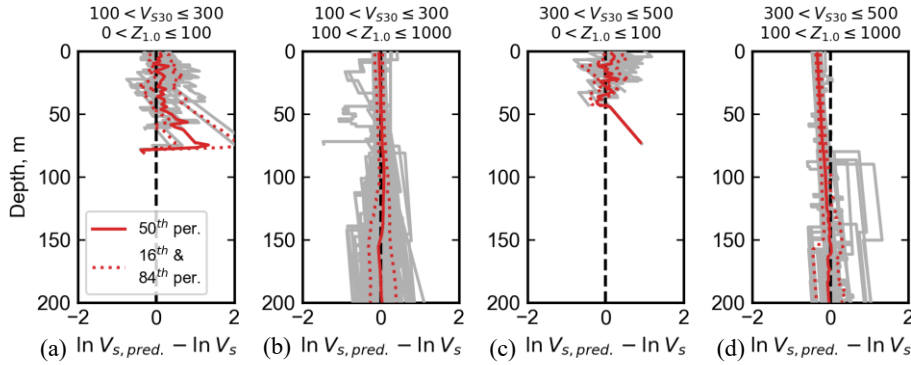


Figure 14. V_s residuals from Figure 12 binned on: (a) low V_{S30} and low $Z_{1.0}$, (b) low V_{S30} and high $Z_{1.0}$, (c) high V_{S30} and low $Z_{1.0}$; and (d) high V_{S30} and high $Z_{1.0}$. Positive residual indicate an overprediction by the model (i.e., our SVM).

Expected Site Response: Predicted Versus Measured V_s

While the preceding evaluation shows that the proposed SVM significantly improves predictions of V_s in Cascadia, it is important to evaluate the consequences for expected ground motions. That is, to what extent do inaccuracies in the predicted velocity structure result in erroneous estimates of site response? To assess the similarity of ground-motion predictions using measured versus predicted profiles, wave-propagation-based site response analyses were

performed on: (1) V_s profiles measured in the field; and (2) predictions of those V_s profiles made by SA18 and the SVM proposed herein. This was achieved via equivalent linear site-response analysis using *pysra* (Kottke, 2020), a Python implementation of the software *Strata* (Kottke and Rathje, 2008). Nonlinear material behavior was modelled using Darendeli (2001) with the following material assumptions: (1) plasticity index of 30; (2) soil density of 19.6 kN/m³; (3) ground water depth of 5 m; (4) at-rest earth pressure coefficient of 1.0; and (5) an over-consolidation ratio computed per Wair et al. (2012). Nonetheless, the most salient findings of these analyses were ultimately insensitive to the assumed material properties. Results are first presented in greater detail for two representative site conditions, considering two input motions. The results of replicate analyses on a large set of V_s profiles are then summarized.

First, the two Ahdi et al (2017) profiles shown in Figure 4 were selected to represent: (1) a soft site ($V_{S30} = 257$ m/s) with high $Z_{1.0}$ (589 m), and (2) a stiff site ($V_{S30} = 642$ m/s) with low $Z_{1.0}$ (23 m). Each profile was subjected to two horizontal input motions from the NGA-West-2 database (PEER, 2014): (1) a low-intensity record ($PGA = 0.06$ g) from the 1976 Mw5.91 Friuli earthquake on a site with $V_{S30} = 650$ m/s (Record Sequence Number [RSN] 133); and (2) a high-intensity record ($PGA = 0.65$ g) from the 2007 Mw6.8 Chuetsu-Oki earthquake on a site with $V_{S30} = 610$ m/s (RSN 4845). These motions were input at depths consistent with the near-surface velocities on which they were recorded.

The implications of V_s prediction for expected site response were studied via ratios of surface-to-input spectral acceleration (S_a), as shown in Figure 15 for the four representative combinations of site conditions and input motions. Such ratios were computed for each combination, namely using: (1) V_s as measured in the field; (2) V_s as predicted by SA18; and (3) V_s as predicted by the proposed SVM. As shown in Figures 15a and 15b, differences between expected motions on the site with high V_{S30} and low $Z_{1.0}$ are generally minor for measured versus predicted profiles. For these conditions, SA18 and the proposed SVM predict V_s profiles that are similar to one another, and to the measured profile, leading to similar estimates of site response. The differences that do arise from measured versus predicted profiles are partly attributable to the smooth gradient of each SVM, such that neither can predict strong impedance contrasts within a profile.

In contrast, as shown in Figures 15c and 15d, large differences in expected motions arise for the site with low V_{S30} and high $Z_{1.0}$. For these conditions, the proposed SVM performs significantly better, whereas SA18 predicts a soil profile that is stiffer and shorter than actual (i.e., it overpredicts

363 V_s at intermediate depths, as shown in Figure 11). This results in erroneous estimates of site
 364 response, such that S_a is overpredicted at shorter periods ($< \sim 0.5$ sec) and underpredicted at longer
 365 periods ($> \sim 2$ sec).

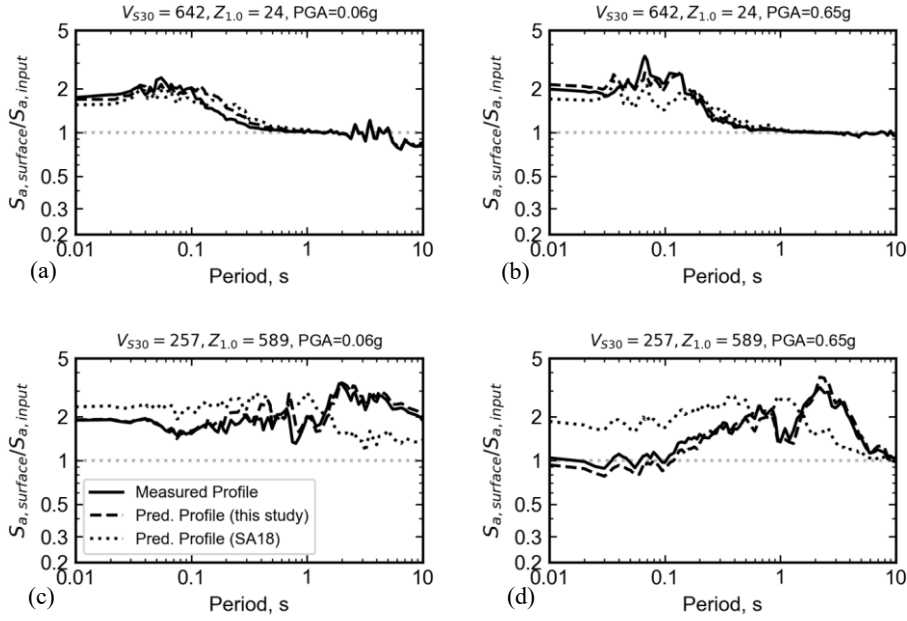


Figure 15. Ratios of surface-to-input acceleration response-spectra ($S_{a,surface} / S_{a,input}$) for measured versus predicted V_s profiles, considering representative site and loading conditions: (a) high V_{s30} , low $Z_{1,0}$, low input- PGA ; (b) high V_{s30} , low $Z_{1,0}$, high input- PGA ; (c) low V_{s30} , low $Z_{1,0}$, low input- PGA ; and (d) low V_{s30} , low $Z_{1,0}$, high input- PGA .

366 Next, this comparison was repeated for all 218 Ahdi et al. (2017) profiles with measured $Z_{1,0}$.
 367 These results are summarized in Figure 16 via the ratio of predicted-to-actual S_a amplification,
 368 where “predicted” and “actual” respectively refer to analyses using predicted and measured V_s . A
 369 ratio of 1.0 indicates that expected S_a values on predicted and measured profiles are identical.
 370 Further, these results are parsed by $Z_{1,0}$, such that ratios from sites with $Z_{1,0} < 100$ m are shown in
 371 Figures 16a and 16b while those from sites with $Z_{1,0} > 100$ m are shown in Figures 16c and 16d.
 372 These results generally mirror those presented above. Considering all profiles with $Z_{1,0} < 100$ m,
 373 the two SVMs result in estimates of site response that are similar. These estimates are also
 374 generally consistent with those derived from measured profiles, as indicated by average S_a -
 375 amplification ratios near unity. Conversely, considering all profiles with $Z_{1,0} > 100$ m, the two

376 SVMs result in drastically different estimates of site response. As shown in Figures 16c and 16d,
 377 the proposed Cascadia SVM results in ground motions that, on average, are very similar to those
 378 expected on measured profiles. In contrast, SA18 results in motions with considerably higher-than-
 379 actual S_a across shorter periods and considerably lower-than-actual S_a across longer periods. These
 380 mispredictions would have important ramifications for the expected, consequent impacts within
 381 Cascadia's basins. For example, in Figure 17, S_a -amplification ratios at natural periods (T_n) of 0.2
 382 sec, 1 sec, and 3 sec are extracted from Figure 16 for each profile and plotted with respect to $Z_{1.0}$.
 383 It is again clear that as $Z_{1.0}$ increases, the proposed SVM performs better, with SA18 tending to
 384 overpredict shorter-period motion and underpredict longer-period motion.

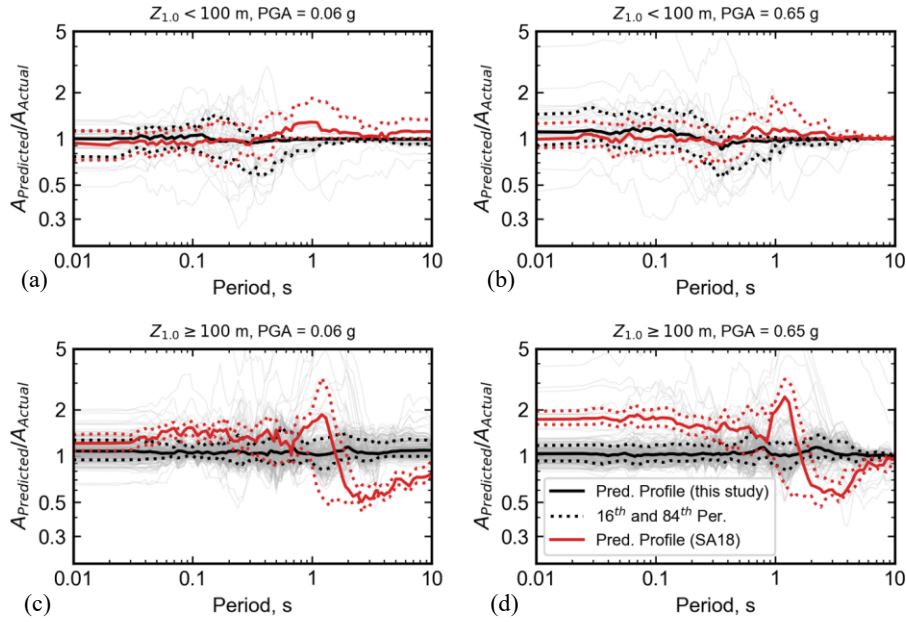


Figure 16. Predicted-to-actual S_a amplification (“predicted” and “actual” refer to analyses using predicted and measured V_s , respectively) for: (a) $Z_{1.0} < 100$ m, low input- PGA ; (b) $Z_{1.0} < 100$ m, high input- PGA ; (c) $Z_{1.0} > 100$ m, low input- PGA ; and (d) $Z_{1.0} > 100$ m, high input- PGA .

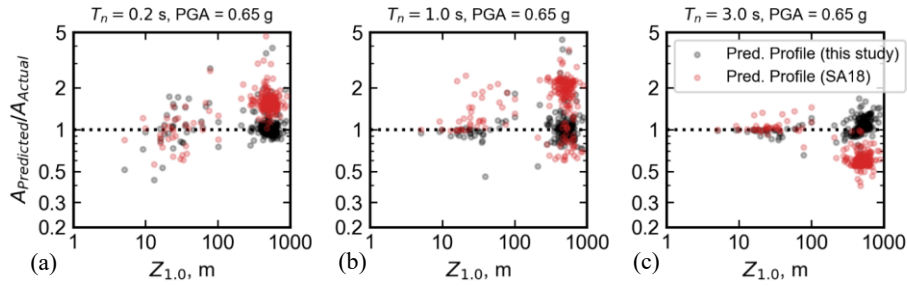


Figure 17. Predicted-to-actual S_a amplification, as a function of $Z_{1,0}$, at natural periods (T_n) of: (a) 0.2 sec; (b) 1.0 sec, and (c) 3.0 sec.

Conclusions

This study presents a new general soil velocity model (SVM) for prediction of V_s profiles. Such models are critical to the evaluation of local amplification of ground motions at sites where detailed, site-specific profiles are not available (e.g., regional assessments). The new model builds on previous work by accounting not only for near-surface soil conditions (V_{S30}), but also deeper geologic structure ($Z_{1,0}$). The model requires input variables V_{S30} and $Z_{1,0}$, which can be estimated without a detailed soil investigation. V_{S30} can be estimated using several proxy- or geology-based methods (e.g., Wald and Allen, 2007; Ahdi et al., 2017), and $Z_{1,0}$ can be estimated from Community Velocity Models (e.g., Stephenson et al., 2017), which provide detailed mapping of deeper geologic structure, but not of the near-surface.

The proposed SVM was calibrated with data from the Cascadia region and improves upon a California-specific model (Shi and Asimaki, 2018) based on V_{S30} alone, resulting in significantly more accurate estimates of site-response in Cascadia's deeper basins while still capturing shallow-site conditions. The new SVM gives physically reasonable predictions for $V_{S30} \geq 100$ m/s and $Z_{1,0} > 2.5$ m up to a depth of $Z_{1,0}$, where the model is intended to coalesce with a CVM. The model was not trained on V_s measurements exceeding 1000 m/s, so for depths beyond $Z_{1,0}$, it is expected that the profile would be given by the CVM. In addition, relatively few existing V_s measurements in the region reach 1000 m/s, limiting the dataset utilized herein. As more $Z_{1,0}$ data become available, the accuracy of this SVM will improve, in turn improving ground-motion predictions across the region. Additionally, it is assumed here that $Z_{1,0}$ scales with 'true' basin depth and that the underlying CVM accurately captures the velocity structure at large depths. Special cases where

large velocity inversions exist due to intermediate layers are beyond the scope of this work and would require site- and basin-specific modeling. Lastly, while this model was implemented for Cascadia, the modeling approach demonstrated herein is globally applicable and is particularly likely to be beneficial in regions with deep basins.

Acknowledgments

The authors would also like to thank Erin Wirth, Arthur Frankel, Annemarie Baltay, and three anonymous reviewers for their comments and suggestions which improved the quality of the manuscript. The presented study is based upon work supported by the U.S. Geological Survey (USGS) and National Science Foundation (NSF) under Grant Nos. G19AP00049 and CMMI-1751216, respectively. Computational infrastructure was provided by the Texas Advanced Computing Center (TACC) at the University of Texas at Austin and NSF Grant No. 1520817. However, any opinion, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF or TACC.

References

- Ahdi, S.K., Stewart, J.P., Ancheta, T.D., Kwak, D.Y., and Mitra, D. (2017). "Development of VS Profile Database and Proxy-Based Models for VS 30 Prediction in the Pacific Northwest Region of North America." *Bulletin of the Seismological Society of America*, 107(4), 1781-1801.
- Allen, T. I., & Wald, D. J. (2007). Topographic slope as a proxy for global seismic site conditions (VS30) and amplification around the globe: U.S. Geological Survey Open-File Report 2007-1357.
- Boore, D. M., Thompson, E. M., and Cadet, H. (2011). "Regional Correlations of VS30 and Velocities Averaged Over Depths Less Than and Greater Than 30 Meters." *Bulletin of the Seismological Society of America*, Seismological Society of America, 101(6), 3046-3059.
- Boore, D. M. (2016). "Determining Generic Velocity and Density Models for Crustal Amplification Calculations, with an Update of the Boore and Joyner (1997) Generic Site Amplification for $V_s(Z) = 760$ m/s." *Bulletin of the Seismological Society of America*, 106(1), 316-320.
- Boore, D. M., & Joyner, W. B. (1997). Site amplifications for generic rock sites. *Bulletin of the seismological society of America*, 87(2), 327-341.
- Bradley, B.A., Bae, S.E., Polak, V., Lee, R.L., Thomson, E.M. and Tarbali, K. (2017). "Ground motion simulations of great earthquakes on the Alpine Fault: effect of hypocenter location and comparison with empirical modelling." *New Zealand Journal of Geology and Geophysics*, 60(3), 188-198.
- Castellaro, S., Mulargia, F., and Rossi, P. L. (2008). "VS30: Proxy for seismic amplification?" *Seismological Research Letters*, 79(4), 540-543.

Formatted: Font: Bold

438 Cramer, C. H., Bauer, R. A., Chung, J. W., David Rogers, J., Pierce, L., Voigt, V., ... & Hempen, G. L. (2016).
 439 "St. Louis area earthquake hazards mapping project: Seismic and liquefaction hazard
 440 maps." *Seismological Research Letters*, 88(1), 206-223.

441 Darendeli, M. B. (2001). Development of a new family of normalized modulus reduction and material damping
 442 curves. University of Texas, Austin.

443 Frankel, A., Wirth, E., Marafi, N., Vidale, J., and Stephenson, W. (2018). "Broadband Synthetic Seismograms
 444 for Magnitude 9 Earthquakes on the Cascadia Megathrust Based On 3D Simulations and Stochastic
 445 Synthetics (Part 1): Methodology and Overall Results." *Bulletin of the Seismological Society of America*,
 446 108 (5A): 2347-2369.

447 Gao, F., & Han, L. (2012). "Implementing the Nelder-Mead simplex algorithm with adaptive
 448 parameters." *Computational Optimization and Applications*, 51(1), 259-277.

449 Holzer, T. L., Padovani, A. C., Bennett, M. J., Noce, T. E., and Tinsley, J. C. (2005). "Mapping NEHRP VS30
 450 site classes." *Earthquake Spectra*, 21(2), 353-370.

451 Hunter, J. A., R. A. Burns, R. L. Good, and C. F. Pelletier (1998). "A compilation of shear wave velocities and
 452 borehole geophysics logs in unconsolidated sediments of the Fraser River Delta, British Columbia." *Geol.*
 453 *Surv. Canada, Open-File Rept. 3622*, Ottawa, Canada.

454 Kottke, A. R. (2020). "Site Response Analysis for Python." <<https://github.com/arkottke/pysra>>. Date Accessed:
 455 Feb. 1st 2020. DOI: 10.5281/zenodo.3522104

456 Kottke, Albert R., and Ellen M. Rathje. (2008). "Technical manual for Strata." Report No.: 2008/10. Pacific
 457 Earthquake Engineering Research Center, University of California, Berkeley.

458 Magistrale, H., Pechmann, J., & Olsen, K. (2009). "The Wasatch Front, Utah, community seismic velocity
 459 model". *Seismol. Res. Lett.*, 80, 368.

460 Marafi, N. A., Eberhard, M. O., Berman, J. W., Wirth, E. A., & Frankel, A. D. (2017). Effects of Deep Basins
 461 on Structural Collapse during Large Subduction Earthquakes. *Earthquake Spectra*, 33(3), 963–997.
 462 <https://doi.org/10.1193/071916EQS114M>

463 McPhee, D. K., Langenheim, V. E., Wells, R. E., & Blakely, R. J. (2014). "Tectonic evolution of the Tualatin
 464 basin, northwest Oregon, as revealed by inversion of gravity data." *Geosphere*, 10(2), 264-275.

465 Myer, L., Downey, C., Clinkensbeard, J., Thomas, S., Stevens, S., Benson, S., Zheng, H., Herzog, H., and
 466 Biediger, B. (2005). "Preliminary geologic characterization of West Coast states for geologic
 467 sequestration." *DOE Contract No: DE-FC26-03NT41984*, California Energy Commission.

468 Parker, G. A., Harmon, J. A., Stewart, J. P., Hashash, Y. M., Kottke, A. R., Rathje, E. M., ... & Campbell, K.
 469 W. (2017). "Proxy-based VS 30 estimation in central and eastern North America." *Bulletin of the*
 470 *Seismological Society of America*, 107(1), 117-131.

471 PEER. (2014). "NGA West 2 Database." Retrieved April 24, 2018, from <http://peer.berkeley.edu/ngawest2/>.

472 Ramirez-Guzman, L., Graves, R. W., Olsen, K. B., Boyd, O. S., Cramer, C., Hartzell, S., ... & Zhong, J. (2015).
 473 Ground-motion simulations of 1811–1812 New Madrid earthquakes, central United States. *Bulletin of the*
 474 *Seismological Society of America*, 105(4), 1961-1988.

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

Formatted: Font: Bold

- Roten, D., Olsen, K. B., & Takedatsu, R. (2019). Numerical Simulation of M9 Megathrust Earthquakes in the Cascadia Subduction Zone. *Pure and Applied Geophysics*. <https://doi.org/10.1007/s00024-018-2085-5>
- Schnabel, P. B., Lysmer, J., & Seed, H. B. (1972). "SHAKE: A computer program for earthquake response analysis of horizontally layered sites." Report No. *EERC72-12*, *EERC*, 12.
- Shi, J., & Asimaki, D. (2018). A Generic Velocity Profile for Basin Sediments in California Conditioned on VS30. *Seismological Research Letters*, 89(4), 1397–1409. <https://doi.org/10.1785/0220170268>
- Small, P., Gill, D., Maechling, P. J., Taborda, R., Callaghan, S., Jordan, T. H., ... & Goulet, C. (2017). "The SCEC unified community velocity model software framework." *Seismological Research Letters*, 88(6), 1539-1552.
- Stephenson, W. J., Reitman, N. G., & Angster, S. J. (2017). P- and S-wave velocity models incorporating the Cascadia subduction zone for 3D earthquake ground motion simulations—Update for Open-File Report 2007–1348. In Open-File Report (Version 1.). <https://doi.org/10.3133/ofr20171152>.
- Vrettos, C. (1996). "Simple inversion procedure for shallow seismic refraction in continuously nonhomogeneous soils." *Soil Dynamics and Earthquake Engineering*, 15(6), 381-386.
- Wair, B. R., DeJong, J. T., & Shantz, T. (2012). "Guidelines for estimation of shear wave velocity profiles." Pacific Earthquake Engineering Research Center.
- Wirth, E. A., Frankel, A. D., Marafi, N., Vidale, J. E., and Stephenson, W. J. (2018). "Broadband Synthetic Seismograms for Magnitude 9 Earthquakes on the Cascadia Megathrust based on 3-D Simulations and Stochastic Synthetics (Part 2): Rupture Parameters and Variability." *Bulletin of the Seismological Society of America*, 108 (5A): 2370-2388.
- Yount, J.C., Dembroff, G. R., & Barats, G. M. (1985). "Map showing depth to bedrock in the Seattle 30' by 60' Quadrangle" U.S. Geological Survey Miscellaneous Field Studies Map 1692, scale 1:100,000, <https://doi.org/10.3133/mf1692>.

Formatted: Font: Bold

Formatted: Font: Bold