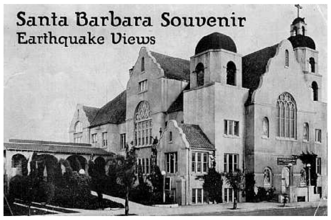

4^{th} *IASPEI / IAEE International Symposium:*

Effects of Surface Geology on Seismic Motion

August 23–26, 2011 · University of California Santa Barbara

A TERRAIN-BASED SITE CONDITIONS MAP OF CALIFORNIA WITH IMPLICATIONS FOR THE CONTIGUOUS UNITED STATES

Alan Yong
USGS
Pasadena, CA
USA

Susan E. Hough
USGS
Pasadena, CA
USA

Junko Iwahashi
GSI
Tsukuba, Ibaraki
Japan

Amy Braverman
JPL/Caltech
Pasadena, CA
USA

ABSTRACT

We present an approach based on geomorphometry to predict material properties and characterize site conditions using the V_{S30} (time-averaged shear-wave velocity to a depth of 30 meters) parameter. Our framework consists of an automated terrain classification scheme based on taxonomic criteria (slope gradient, local convexity and surface texture) that systematically identifies 16 terrain types from 1-km spatial resolution (30 arc sec) Shuttle Radar Topography Mission digital elevation models (SRTM DEMs). Using 853 V_{S30} values from California, we apply a simulation-based statistical method to determine the mean V_{S30} for each terrain type in California. We then compare the V_{S30} values to models based on individual proxies, such as mapped surface geology and topographic slope, and show that our systematic terrain-based approach consistently performs better than semi-empirical estimates based on individual proxies. To further evaluate our model, we apply our California-based estimates to terrains of the contiguous U.S. Comparisons of our estimates to 325 V_{S30} measurements outside of California, as well as estimates based on the topographic slope model, indicate our method to be statistically robust and more accurate. Our approach, thus, provides an objective and robust method for extending estimates of V_{S30} where in-situ measurements are sparse or not readily available.

INTRODUCTION

To assess site conditions, key factors including layer thickness and impedance (the product of the density of the material and the velocity of the propagating wave) are considered. Because shear-wave velocity (V_S) correlates with soil rigidity and is characterized by higher variability than density, V_S has generally been accepted as an appropriate measure of soil conditions (e.g., Borcherdt, 1970; Fumal, 1978; Aki, 1988; Borcherdt and Glassmoyer, 1992; Boore, 2006). Although variations in material properties at depths of tens to hundreds of meters below the Earth's surface are known to significantly influence ground motions, deeper variations are also important (e.g., Anderson *et al.*, 1996; Frankel *et al.*, 2002; Boore, 2004; Holzer *et al.*, 2005). Nevertheless, a number of studies (e.g., Borcherdt *et al.*, 1991; Boore *et al.*, 1993; Borcherdt, 1994; Dobry *et al.*, 2000; Boore *et al.*, 1994; 1997; Wills and Silva, 1998) have found good correlation between observed site amplification and V_{S30}, or the time-averaged shear-wave velocity to a depth of 30 meters below the surface. As a result, the V_{S30} parameter has since been widely adopted as the key parameter for site characterization.

When direct measurements are prohibitive, sparse or not readily available, the use of proxies, i.e., geologic properties and/or units, topographic slope, terrain types, etc., to infer V_S or V_{S30} is typically applied. For example, a number of proxy-based maps have been developed for all or parts of California (Joyner *et al.*, 1981; Tinsley and Fumal, 1985; VIC, 1993; Petersen *et al.*, 1997; Park and Elrick, 1998). More recently, Wills *et al.* (2000) presented the "first cut" of a geologic unit-based V_{S30} map for California that has found widespread usage in recent years. Further subdividing the original eight site classes used in the 2000 study, Wills and Clahan (2006) refined the urban areas (C. Wills, personal comm., 2006) of the earlier map by introducing 19 site classes based on a number of factors in addition to surficial geology—such as grain size, thickness of units, deep alluvium assumed for major basins, shallow depth for small and narrow basins, and degree of coarseness of sediments as a function of distance from mountain fronts.

While V_{S30} maps have traditionally been derived directly from geological maps, several studies (e.g., Romero and Rix, 2001; Wald and Allen, 2007; Yong *et al.*, 2008a; 2008b) have exploited newly available remote sensing data to develop site classification maps. As one of the earliest studies utilizing satellite imagery to account for site conditions, Romero and Rix (2001) used Landsat 7 Enhanced

Thematic Mapper Plus (ETM+) data to interpret the character of surficial sediments for estimating the potential of ground shaking in the Mississippi Embayment. Using Shuttle Radar Topography Mission digital elevation models (SRTM DEMs) (Farr *et al.*, 2007), Wald and Allen (2007) presented global V_{S30} maps based only on topographic slope. Most recently, Yong *et al.* (2008a; 2008b) performed semi-automated analyses of spectra (visible to thermal-infrared electro-magnetic frequency range) and 30-meter resolution DEM (stereoscopic correlation method) data produced from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) satellite recordings to interpret remotely-sensed geologic and topographic information. Subsequently, these results were assembled into terrain types to estimate V_{S30} in a 60 km by 60 km area surrounding Islamabad, Pakistan. This study demonstrated that satellite data can help circumvent the problems associated with inconsistencies found in traditional geologic mapping. They also showed that the combination of multiple parameters, such as geology and topography, yields a more robust estimate than uni-parametric models. However, results from earlier investigations (Yong *et al.*, 2005; 2006; 2007) indicated that a systematic and objective classification method that is both simple and yet robust for all terrain types is necessary for characterizing larger regions (e.g., Murray and Fonstad, 2007; Philips, 2007). To meet these requirements, we investigate the Iwahashi and Pike (2007) automated approach to develop a global terrain classification map that was also based on SRTM DEMs. As a basis for this study, we describe the details of our method where we explore the effectiveness of using a terrain-based model as the underlying framework for a V_{S30} proxy-based approach.

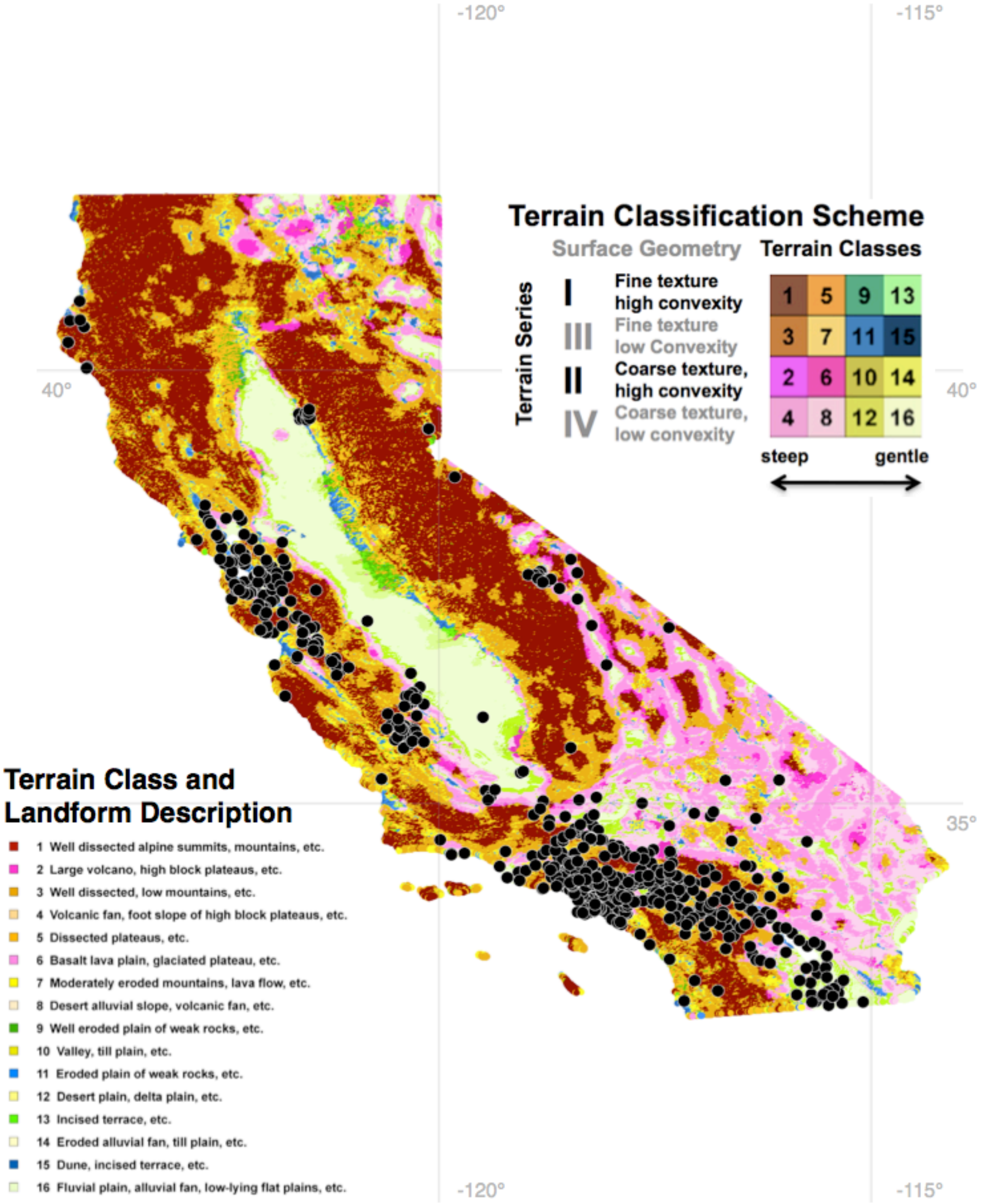

Fig. 1. Terrain map of California (Iwahashi and Pike, 2007) with locations (black circles) of 853 V_{S30} values.

DATA

Terrain-based Framework

Iwahashi and Pike (2007) (hereinafter as IP07) introduced an automated topography classification scheme to develop a globally consistent terrain classification map. The IP07 approach employed an algorithm that divides continuous topography (SRTM DEM) represented by elevations into a maximum possibility of 16 terrain classes through the use of a geometric signature consisting of three taxonomic criteria—slope gradient, local convexity and surface texture. The design of the topographic classification scheme is based on an unsupervised nested-means approach (Scripter, 1970) where the final output of terrain classes reflects the statistical properties of the input geometric variables rather than preset criteria. Unlike supervised algorithms, for which predetermined criteria for each output class do not adapt equally well to all locations and DEM resolutions, the unsupervised approach treats topography as a continuous random surface, independent of any spatial or morphological orderliness imposed by fluvial activity and other geomorphic processes. Since the resultant terrain classes are not determined *a priori* by training examples of target physiographic types, IP07 acknowledged that these classes are only approximately equivalent to such traditional terrain types as mountains, high hills, plains with hills, or tablelands in the manner of Hammond's (1964) taxonomy. Nevertheless, on the basis of subsequent map overlays and statistical analyses, IP07 showed that their U.S. terrain classification map closely resembles the classic physiographic divisions of Fenneman and Johnson (1946) and for more recent work, the geomorphic systems of Graf (1987). In our terrain framework, for example, the physiographic taxonomy established for the United States (Hammond, 1954; 1964) is used to relate the equivalent classes to the terrain category of California (Fig. 1). For the main part of this study, we use the same automated scheme and only adopt the California region out of the global dataset. In a later section, we discuss some results from preliminary tests, relating to use of an IP07 terrain map that covers the contiguous U.S.

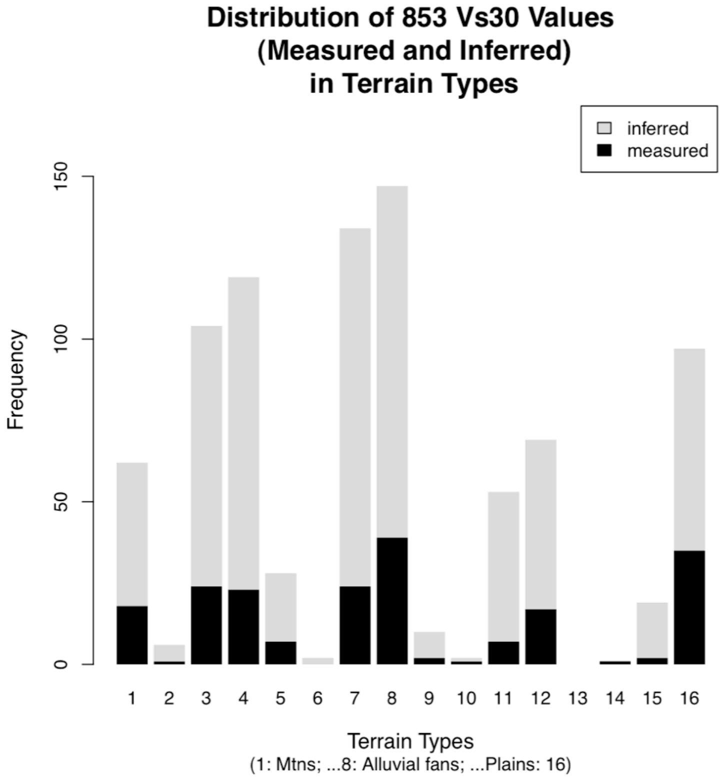

Fig. 2. Barplot describing the combined frequency distribution of measured (black) and inferred (grey) V_{S30} values.

V_{S30} values

In developing our terrain-based V_{S30} proxy model, we use shear-wave velocity data provided from Walt Silva (last revised July 17, 2006; written comm., 2007) of Pacific Engineering and Analysis (PE&A). As part of the NGA (Next Generation Attenuation) Strong-motion Database (Chiou *et al.*, 2008), the V_{S30} data consists of both measured and inferred values, i.e., measurements from recorded waveforms and extrapolations from proxy- and/or map- based correlations, respectively. Hence, we will continue to purposefully refer to this mix of V_{S30} data as "V_{S30} values" or "results" and not as "measurements" to avoid any misrepresentation about the nature of the source data. Of the 853 values in the database, 644 (75%) were inferred. Measured V_{S30} data were calculated from 201 (24%) V_S records with (minimum) 30 m depth profiles (Fig. 2). Other data used to infer V_{S30} include measurements are from six (0.7%) V_S profiles of less than 20 m and two (0.2%) borehole records extending to the 9 m depth. Furthermore, this sparse subset of measurement-derived V_{S30} values are incompatible due to the diversity of measurement methods used (Silva, personal comm., 2009). This limitation was previously noted by a number of authors (e.g., Wills, 1998; Boore, 2006; Moss, 2008). Biases between and imprecision associated with different methods control the overall epistemic uncertainties related to the results, and are directly attributable to the fundamental differences in principle that governs the respective geophysical method applied. For example, in a study comparing the variability of shear-wave velocity determinations from existing blind and comparative studies, Moss (2008) examined both the intramethod and intermethod variability and found a clear bias between invasive and noninvasive methods.

For V_{S30} values derived from proxy- and map-based inferences, concerns regarding data quality are generally related to the same issues as previously described for measured values. Despite these well-known issues, many of these data were used by Wills *et al.* (2000) to develop the Site Conditions Map of California and later by Wald and Allen (2007). In effect, the Silva (personal comm., 2007) V_{S30} database used in this study is an updated version of the V_{S30} values used by both Wills *et al.* (2000) and Wald and Allen (2007). Wills *et al*. (2000) used 556 V_{S30} values; for their California map Wald and Allen (2007) used an updated database of 767 values as well as other values in active tectonic regions (e.g., Utah [204], Taiwan [387] and Italy [43]).

We use all 853 V_{S30} results from the updated (2007) Silva V_{S30} database to develop our terrain-based V_{S30} proxy model. To maintain a framework consistent with the Wills *et al.* (2000) and Wald and Allen (2007) approaches, we also do not appraise the quality of the V_{S30} values used in our study other than acknowledging the lack of quality assurance in the V_{S30} data, and addressing the limitations of V_{S30} as a measure of site conditions later.

METHODS

Developing the Terrain-based V_{S30} Proxy Model

After establishing our terrain-based framework, we compare available V_{S30} values with our terrain classification map. Using the spatial locations for the 853 V_{S30} values, we determine the corresponding terrain-type class (Fig. 1). No V_{S30} values are associated with areas identified as our terrain category of Class 13, which fit the description of incised terraces (Fig. 2). Such a result might be expected because relatively few V_{S30} values are available at hard rock sites. V_{S30} values generally tend to be in major urban/metropolitan areas that are typically located on soft basin-type terrains (Fig. 2). Also, in the case of Class 13, the terrace terrain-type has the second-lowest ranking number (2,106 pixel cells) of occurrences in California representing <1% of the total cells (574,049) that make up the regionally extensive and morphologically diverse state (Yong *et al.*, 2010).

Following the association of V_{S30} values to the 15 site classes, we randomly select 556 values (same number used by Wills *et al.*, 2000) and calculate the mean V_{S30} values for each of the 15 terrain classes. Here, the portion of V_{S30} values assigned within each terrain class is based on the distribution of the original 853 values we find in each of the 15 terrain categories (Fig. 2). We use a standard cross-validation (non-replacement) method to compare our prediction to the remaining 297 V_{S30} measurements by examining its mean-squared prediction error (MSPE):

$$\Delta = \frac{1}{N}\sum_{i=1}^{N}(V_{S30_i} - prediction_i)^2 \tag{1}$$

The routine, from random selection through cross-validation, is iterated one thousand times. At the end of these runs, we select the model with the lowest MSPE (14,414) as our best predictor of V_{S30} for each terrain type (Table 1).

Terrain Type	Description	Freq. in CA (%)	Predicted V_{S30} (m/s)
1	Well dissected alpine summits, mountains, etc.	21	519
2	Large volcanoes, high block plateaus, etc.	2	393
3	Well dissected, low mountains, etc.	14	547
4	Volcanic fans, foot slopes of high block plateaus, etc.	19	459
5	Dissected plateaus, etc.	6	402
6	Basalt lava plains, glaciated plateaus, etc.	1	345
7	Moderately eroded plains of weak rocks, etc.	7	388
8	Desert alluvial slopes, volcanic fans, etc.	12	374
9	Well eroded plains of weak rocks, etc.	1	497
10	Valleys, till plains, etc.	<1	349
11	Eroded plains of weak rocks, etc.	2	328
12	Desert plains, delta plains, etc.	4	297
13	Incised terraces, etc.	<1	—
14	Eroded alluvial fans, till plains, etc.	<1	209
15	Dunes, incised terraces, etc.	1	363
16	Fluvial plains, alluvial fans, low-lying flat plains, etc.	8	246

Table 1. Table describing terrain type, terrain description, percentage of V_{S30} values in California and the predicted V_{S30} values based on the mean V_{S30} value for each terrain type.

RESULTS

Evaluation of Model Performances

To assess our terrain-based V_{S30} proxy model beyond inferential statistics, we evaluate its performance compared to the geologic unit-based model (Wills *et al.*, 2000) and the topographic slope-based model (Wald and Allen, 2007). Since both of the latter approaches effectively employed the same V_{S30} data we use for the same area (state of California)—albeit, in varying quantities, i.e., Wills *et al.* (2000) and Wald and Allen (2007) relied on approximately 65% and 91% (respectively) of the values from the updated V_{S30} database—the commonality in data source reasonably supports the basis for our direct comparison of the different models. Using the "Spatial Join" function in the ESRI ArcInfo (version 9.3) application program, we associate the 853 sites to the seven categories of the Wills *et al.* (2000) model (data provided by C. Wills, written comm., 2006). On the basis of the site categories in the joined GIS-database tables, we translate every instance of their categories from class letter designations to corresponding mean V_{S30} values (Wills *et al.*, 2000). For the Wald and Allen (2007) model, which is available from the USGS (http://earthquake.usgs.gov/research/hazmaps/interactive/vs30/predefined.php; last accessed 18 July, 2008) as an evenly sampled xyz file, we develop a simple algorithm using the nearest neighbor approach (Clark and Evans, 1954) to associate their coordinate-and-V_{S30}-value relations with the 853 V_{S30} sites. To identify a straightforward and common parameter between the Wills *et al.* (2000) geologic unit-based categories and our own model, we use their reported mean V_{S30} values for each of their seven expanded NEHRP-based classes. For the Wald and Allen (2007) topographic slope-based categories, mean V_{S30} values were not reported (or readily available) so we use the assigned V_{S30} values in their downloadable text-file. We use the Warnes (2004) "bandplot" (locally-smoothed-mean and standard deviation) function and a customized subroutine—all in the open-source statistical environment "R" (last accessed 4 April, 2009; http://www.r-project.org/)—to consider the local mean and variance, and calculate the typical standard deviation (σ_{typical}) of each model (Fig. 3). As a result, we find the Wills *et al.* (2000) model has the highest σ_{typical}, approximately 195 m/s, while the Wald and Allen (2007) model and our model have approximately 142 m/s and 128 m/s, respectively. As part of our statistical analysis, we also calculate the median of the MSPE for each model and determine that all models perform significantly better than random.

As an extension of our efforts to develop a site characterization map for California, we explore the utility of our approach for a V_{S30} map of the contiguous U.S (Yong *et al.*, 2010). Like the California map (Fig. 1), the U.S. V_{S30} map (Fig. 4) is derived using the automated IP07 classification method that relies on the same taxonomic criteria (slope gradient, local convexity and surface texture) developed from geomorphometry to also identify 16 terrain types from the same 1-km spatial resolution (SRTM30) DEM. On the basis of the California model, we present our preliminary U.S. map by applying the same terrain classification approach and adopting the same V_{S30} relations determined for each terrain type in California. Although our database of 325 V_{S30} observations (for source, see Data and Resources Section) measured outside of California is currently too sparse to permit a rigorous validation of the terrain-based

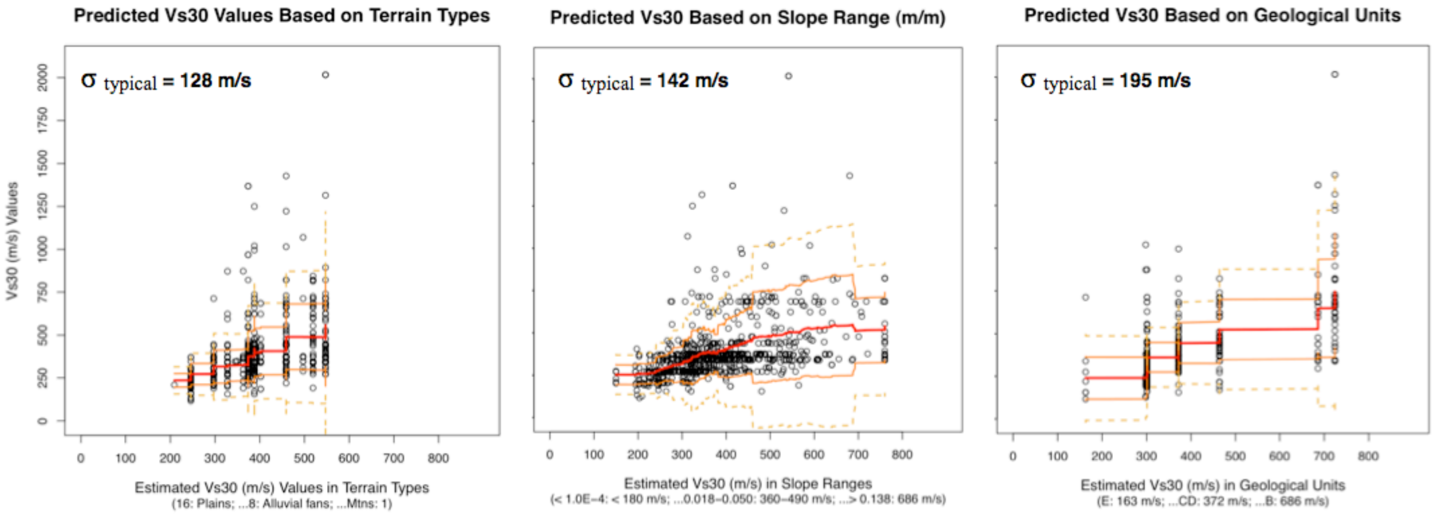

Fig. 3. Plots describing the smoothed mean and standard deviation for 853 V_{S30} values based on both measured and inferred data (California) against the terrain-based (left), topographic slope-based (middle), and geological unit-based (right) estimation models. Red lines indicate the smoothed mean values, orange dashes indicate the one standard deviation from the mean and yellow dash indicates two standard deviations from the mean.

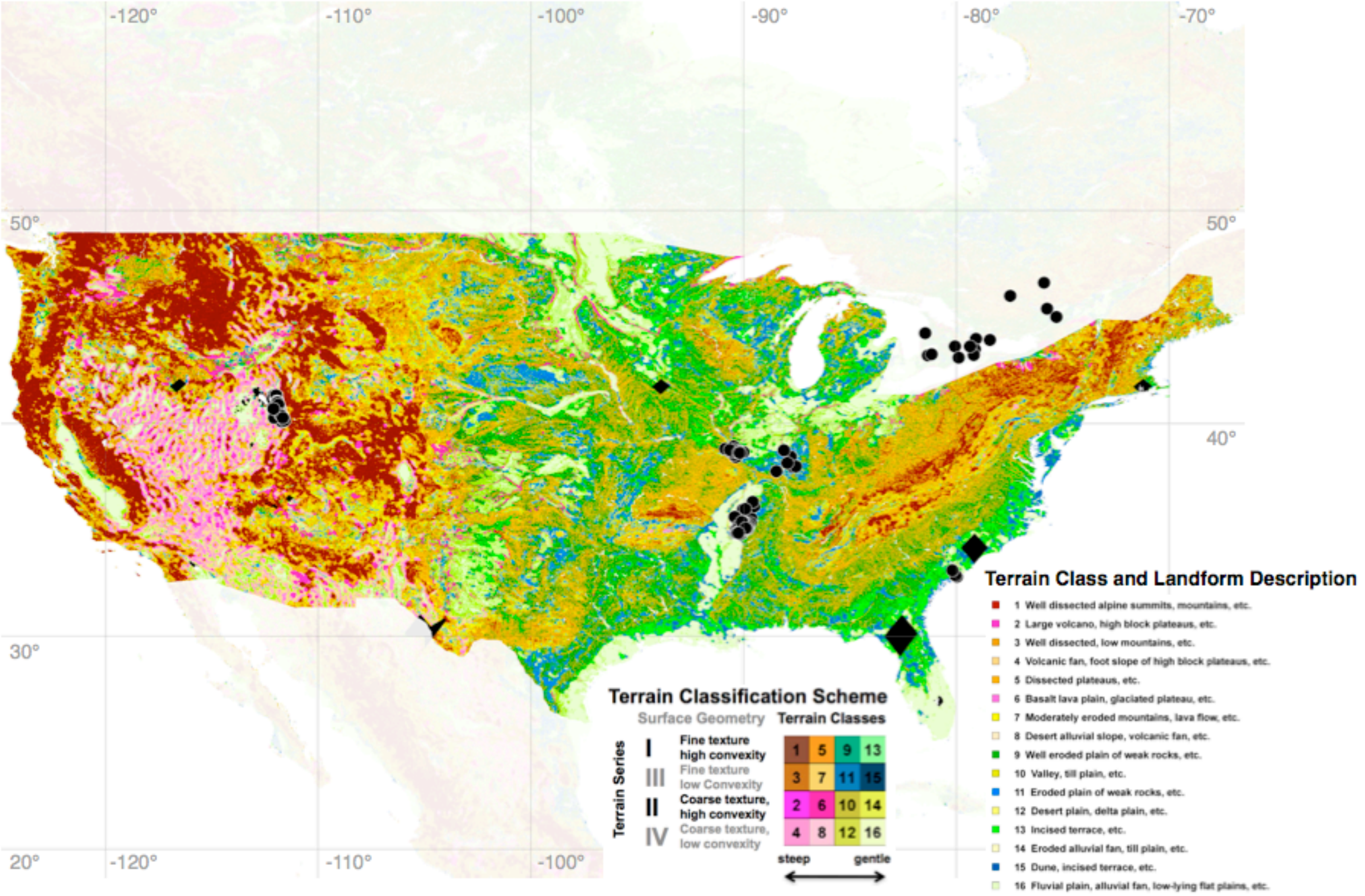

Fig. 4. Terrain map of the contiguous United States (Iwahashi and Pike, 2007) with locations (black circles) of 325 V_{S30} measurement-based values, including 14 records from Canada. Areas of poor DEM quality were left unclassified and designated by the authors as black polygons.

model, we plot the locally-smoothed mean and standard deviation (Fig. 5) of the terrain-based and topographic slope models for a preliminary assessment of their performances. At the time of our investigation, no geology-unit based V_{S30} estimation model of the contiguous U.S. was readily available. When comparing the σ_{typical} of each model for their respective performances in California and the U.S. (Fig. 3 and 5), we note the changes (Δ_{σ}) in their σ_{typical} and find that the terrain-based model ($\Delta_{\sigma} = +19$ m/s) is more stable than the topographic-slope based model ($\Delta_{\sigma} = +65$ m/s). We also find an appreciable improvement in the predictive performance of the terrain-based model ($\sigma_{\text{typical}} = 147$ m/s) over the topographic sloped-based model ($\sigma_{\text{typical}} = 213$ m/s).

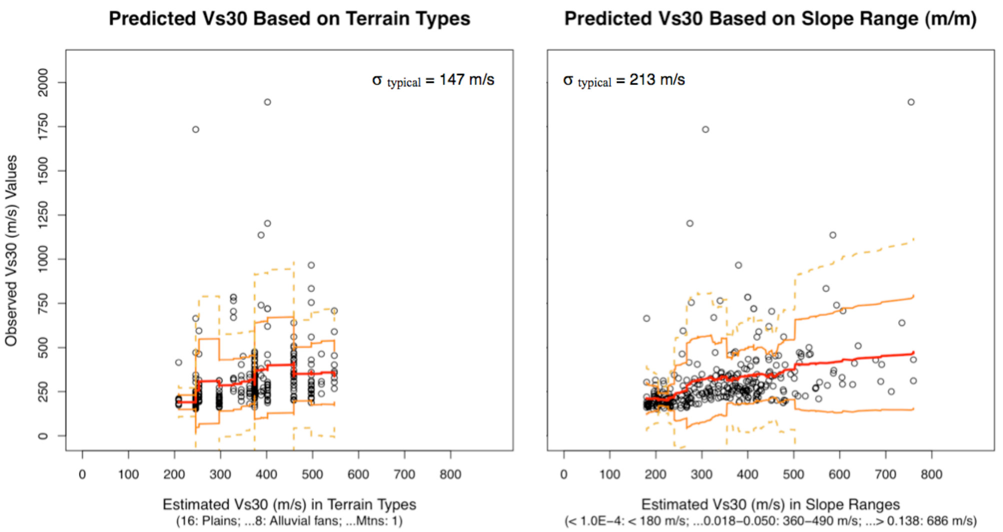

Fig. 5. Plots describing the smoothed mean and standard deviation for V_{S30} values based on 325 recorded measurements (contiguous U.S., outside of California) against the terrain-based (left) and topographic slope-based (right) estimation models. Red lines indicate the smoothed mean values, orange dashes indicate the one standard deviation from the mean and yellow dash indicates two standard deviations from the mean.

DISCUSSION

Our primary motives for developing the terrain-based model are effectively two-fold: to more fully exploit the nature of geomorphology as a proxy for site amplification, and to advance an approach that is completely objective with results firmly supported by robust statistical methods. We acknowledge that a significant portion (76%) of the data used to develop our model are based on inferred V_{S30} values and that such a strong dependence fundamentally contradicts the established principles of model development. However, it is necessary to depend on these proxy-based estimates because there are simply not enough openly-available measured values to meet standard statistical requirements. As previously indicated, the reliance on inferential data has precedent in state-of-art approaches, e.g., the geology-based and topographic slope-based models were also conditioned on effectively the same inference-based data. In the following section, we expand on these and other issues associated with our dependence on proxy-based models and the use of V_{S30} for estimating site amplification.

Although geomorphology is (in principle) an appropriate proxy for capturing the key characteristics of both geology and topography (Peterson, 1981; Harden, 1990; Ellis *et al.*, 1999; Saco *et al.*, 2006) and that morphometrics in terrain models meet the requirements for a consistent, objective, and quantitative framework that systematically encapsulates the parameters related to these near-surface properties (Oliver and Webster, 1986; Franklin, 1987; Dehn *et al.*, 2001; Minár and Evans, 2008; Pike *et al.*, 2009)—one initial concern is the role that terrain types play in our model. Because the proxies are by definition estimates of the predictors (i.e., terrain-type is the proxy for both geology and topography, which are in turn proxies for V_{S30}, which itself is a proxy for site amplification), this type of multi-tiered approach can inherently compound the uncertainties associated with each proxy. Despite these concerns, Iwahashi *et al.* (2010) examined results from multiple regression analyses of V_{S30} measurements at 1646 locations with topographic attributes (slope gradient, surface texture, and the logarithm of elevation) based on 50-m DEM and found that estimations of V_{S30} from a DEM are useful for earthquake vulnerability assessment.

One utility of the terrain-based approach is that material properties, which are not directly related to a single proxy, such as slope, can be inferred. This is especially true when estimating V_{S30} in flat or steep topographic areas. For such instances, the combination of terrain nomenclature (Hammond 1954; 1964) and the IP07 classification scheme provides a basis for differentiating material properties that do not always directly associate with individual proxies, such as topographic slope. For example, in a volcanic plain with flat topographic slope and high shear-wave velocity, topographic slope (alone) would predict high amplification (low V_{S30}). Furthermore, in active-tectonic regions, lower shear-wave velocity (high amplification) with steep topographic slope consisting of young sediments, a slope-based model would predict high V_{S30}. In both cases, geology plays the key role for distinguishing the exceptions in the relationship between the nature of topographic slope and V_{S30}. A manual analysis of the same regions (e.g., Modoc Plateau, Mendocino, etc.) in each model indicates significant differences in V_{S30} predicted (Fig. 1 and Wald and Allen, 2007). Actual observations of V_{S30} measurements from the regions are necessary to properly evaluate the performance of both models.

Clearly, the arbitrary nature of 30-meters as the key parameter for depth from the surface in the V_{S30} term is expected to have a strong first-order consequence on the effectiveness of the proxies. To address this issue, the use of more complicated parameterizations might be useful. Approaches worthy of further investigation, include the one-quarter wavelength method (Joyner *et al.*, 1981), or extra soil site considerations, e.g., sediment thickness coupled with shear-wave velocity (s_L) first proposed by Seed *et al.* (1976) and recently reviewed by Lee and Trifunac (2010). Another possible approach is to circumvent V_{S30} as a proxy by determining if the IP07 terrain scheme can be used to directly estimate site amplification; this will require the use of records from both permanent (e.g., California Integrated Seismographic Network) and temporary (e.g., USArray) seismographic station sites in order to meet minimum spatial sampling requirements. For now, in the absence of a more robust site parameter, we follow accepted convention and proceed with V_{S30} as the proxy for material induced amplification.

Considering the uncertainties associated with the V_{S30} term, in addition to other vexing factors such as artifacts associated with spatially referenced data, not to mention circularity issues relating to the co-mingling of measured- and inferred- (map-based) values found in the V_{S30} database, we recognize that these limitations will tend to confound our dependence on strict adherence to our statistically-rigorous systematic approach. Rather, it is in the face of these uncertainties that we consider an objective and systematic approach based firmly on statistics to be most warranted when investigating the predictive capability of the IP07 terrain model to estimate V_{S30}.

By statistically determining that each model performs significantly better than random, we are able to confirm that all of the proxies have a strong correlation to V_{S30} (Yong *et al*., 2010). However, when we consider standard regression analysis methods to evaluate the performance of the models by themselves and against each other, we encounter a fundamental statistic-related problem inherent in all three models. The respective proxies of the geologic unit-, topographic slope- and terrain-based models do not readily fit in any formalized description of a statistical variable—i.e., although the proxies are defined as names (nominal characteristic), at the same time their associations to V_{S30} are strongly dependent on numeric values tied to predefined (discrete variable) ranges on a common continuous scale with meaningful hierarchical order (ordinal characteristic). Nevertheless, it is in the utility of the standard deviation (σ) where we find a meaningful, yet simple, statistical measure for evaluating the performance of each model and the comparisons against each other. Here, based on the 853 values from the Silva database, we look for the typical amount of dispersion (σ_{typical}) in the V_{S30} values with respect to each model's proxy-based estimates. Because our terrain-based estimates depend on the average of the V_{S30} values for the respective terrain types, we choose the equivalent average V_{S30} values described by Will *et al.* (2000) for their eight respective classes to calculate the σ_{typical} of their geologic unit-based model. Although the topographic slope-based approach initially estimated V_{S30} on the basis of the median velocity of the subdivided NEHRP classifications, these results are not fully described in Wald and Allen (2007) and we find no means to calculate the average (nor median) of their V_{S30} estimates for each class because the complete data set is no longer available (Wald, written comm., 2009). In the face of these deficiencies, we instead rely on the latest and most-readily-available estimates from their USGS web-portal (http://earthquake.usgs.gov/research/hazmaps/interactive/vs30/predefined.php; last accessed 18 July, 2008) for comparison to the Silva V_{S30} data. Because we consider their estimates to possess the most statistically advantageous traits on the basis of the continuous nature and sheer abundance (1,401) of the V_{S30} data used in their analyses, we expect Wald and Allen (2007) estimations to yield the least σ_{typical}, hence rank as the best performing model when compared against the discretized proxies in geologic unit- and terrain-based models. Instead, as noted earlier, we find our terrain-based model to yield the lowest σ_{typical} (approximately 128 m/s), follow closely by the topographic slope-based (approximately 142 m/s) and the geologic unit-based models (approximately 195 m/s), respectively.

Regardless of the complexity and/or innovation associated with automated and systematic approaches, the accuracy of the derived model is inherently limited by the quality of the input data. For the topography of California, IP07 used DEMs of spatial resolution at 30-arc-sec (approximately 1-kilometer grid spacing or pixel size) derived from recordings measured by the Shuttle Radar Topography Mission (SRTM) (Farr *et al.*, 2007). Although higher resolutions of SRTM DEMs were also available (e.g., the SRTM recordings were sampled over a grid at spatial resolution of 1-arc-second by 1-arc-second or approximately 30 m by 30 m), we instead choose to adhere to the proven IP07 technique which produced reasonably accurate classifications (Iwahashi and Pike, 2007). At the same time, the use of the same 30-arc-sec topography employed by Wald and Allen (2007) allows us to directly compare the two methods. By comparison, the 1:250,000-scale maps used by Wills *et al.* (2000) have (in principle) the best-resolved details (pixel resolution of 125 m) with a detectable object-resolution of 250 m (Tobler, 1988).

CONCLUSION

To varying degrees of success, individual proxies such as mapped surface geology (Wills *et al.*, 2000) and topographic slope (Wald and Allen, 2007) have been used to develop V_{S30} (the average shear-wave velocity in the upper 30 meters) predictions for California. This type of broad-brush (single-parameter) approach to site characterization has been demonstrated to be useful for some purposes, especially in the absence of measured V_{S30} values (Frankel *et al.*, 2010; Thompson *et al.*, 2010). However, clearly no single proxy can fully account for variations in material properties that control V_{S30}. In an attempt to improve site characterization, we develop a method based on geomorphometry (slope gradient, local convexity and surface texture) to better predict their material properties and thus estimate V_{S30}. We use results from the automated terrain classification method based on taxonomic criteria derived from geomorphometry to systematically identify 16 terrain types from a 1-km spatial resolution (SRTM30 data) DEM of California (Iwahashi and Pike, 2007). Using 853 V_{S30} values, we assign V_{S30} values that are the average of values within each terrain type. We then compare these values to the single-proxy prediction models to determine relative performances. As expected, the overall results indicate that our approach performs better than the characterizations based on single proxies. In each case, we note that the reliance on inconsistent geologic data sets (Wills *et al.*, 2000), or a simple parameter, such as the 1-km resolution SRTM30 DEM derived slope gradient (Wald and Allen, 2007), are less discriminating in accounting for variations in site conditions. Furthermore, our preliminary studies also indicate the terrain-based framework to be more effective than topographic-slope alone when comparisons are extended to the contiguous U.S. (Yong *et al.*, 2010). Despite statistical tests that indicate our method to be, at worst case, a marginal improvement over the next-best method, we also note that by employing objective and systematic methods, as well as a multi-faceted framework that encapsulates geology and topographic-slope, our approach holds the most promise for improving the prediction of material (soil) properties when estimating site amplification. As new and spatially-distributed measurement-based V_{S30} data become available, we expect to be able to improve our predictive capabilities.

DATA AND RESOURCES

GIS data of world-wide automated terrain classification map from SRTM 30 can be downloaded at http://gisstar.gsi.go.jp/terrain/front_page.htm. The updated NGA Database Flat File (including measurement-based V_{S30} observations at 325 recording sites outside of California), referred in this article as the NGA (Next Generation Attenuation) Strong-motion Database, can be downloaded at http://peer.berkeley.edu/nga/flatfile.html.

ACKNOWLEDGEMENT

We are grateful to Chris Cramer, Jack Odum, Kris Pankow, Walt Silva and Rob Williams for providing the V_{S30} data. Special thanks to Walt Silva and Masashi Matsuoka, who provided many important insightful discussions about the role of V_{S30} in site amplification. We are also grateful to David Boore, Eric Thompson, Brad Aagaard, Christopher Stephens and Dominic Assimaki for constructive feedback that greatly improved our manuscript. Lisa Christiansen and Stan Schwartz were instrumental in preprocessing the GIS data. Lastly, we thank Richard J. Pike and Chris J. Wills for early discussions on the development of our project. Work done by Amy Braverman was performed at the Jet Propulsion Laboratory/California Institute of Technology, under contract to the National Aeronautics and Space Administration.

REFERENCES

Aki, K. [1988]. “Local site effects on strong ground motion”, in Proc. Earthq. Eng. and Soil Dyn. II − Recent Advances in Ground-motion Evaluation, New York, pp. 103−155.

Anderson, J. G., Y. Lee, Y. Zeng, and S. Day [1996]. “Control of strong motion by the upper 30 meters”, Bull. Seism. Soc. Am., Vol. 86, No. 6, pp. 1749-1759.

Boore, D. M. [2004]. “Can site response be predicted?”, J. Earthq. Eng., Vol. 8, No. 1, pp. 1-41.

Boore, D. M. [2006]. “Determining subsurface shear-wave velocities: a review”, in Proc. Third International Symposium on the Effects of Surface Geology on Seismic Motion, August 30- September 1, Grenoble, France, 21 pp.

Boore, D. M., W. B. Joyner, and T. E. Fumal [1993]. “Estimation of response spectra and peak accelerations from western North American earthquakes: an interim report”, U.S. Geol. Surv. Open-File Reprt. 93-509, 69 pp.

Borcherdt, R. D. [1970]. “Effects of local geology on ground motion near San Francisco Bay”, Bull. Seism. Soc. Am., Vol. 60, No. 1, pp. 29-61.

Borcherdt, R. D. [1994]. “Estimates of site-dependent response spectra for design (methodology and justification)”, Earthq. Spectra, Vol. 10, pp. 617-653.

Borcherdt, R. D., and G. Glassmoyer [1992]. “On the characteristics of local geology and their influence on ground motions generated by the Loma Prieta earthquake in the San Francisco bay region, California”, Bull. Seism. Soc. Am., Vol. 82, No. 2, pp. 603-641.

Borcherdt, R.D., C. M. Wentworth, A. Janssen, T. Fumal, and J. Gibbs [1991]. “Methodology for predictive GIS mapping of special study zones for strong ground shaking in the San Francisco bay region”, in Proc. Fourth Int. Cont. on Seismic Zonation, Earthquake Engineering Research Institute, Oakland, California, pp. 545-552.

Chiou, B., R. Darragh, N. Gregor, and W. Silva [2008]. “NGA project strong-motion database”, Earthq. Spectra, Vol. 24, No. 1, pp. 23-44.

Clark, P. J., and F. C. Evans [1954]. “Distance to nearest neighbor as a measure of spatial relations in populations”, Ecol., Vol. 35, pp. 445-453.

Dehn, M., H. Gärtner, and R. Dikau [1987]. “Principles of semantic modeling of landform structures”, Comp. Geosci., Vol. 27, pp. 1005-1010.

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”, Earthq. Spectra, Vol. 16, No. 1, pp. 41-67.

Ellis, M. A., Densmore, A. L., and Anderson, R. S. [1999]. “Development of mountainous topography in the Basin Ranges, USA”, Basin Res., Vol. 11, pp. 21-41.

Farr, T. G., P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriquez, L. Roth, D. Seal, S. Shaffer, J. Shimada, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf [2007]. “The shuttle radar topography mission”, Jet Propulsion Laboratory/California Institute of Technology, Pasadena, 43 pp.

Fenneman, N. M., and D. W. Johnson [1946]. “Physical division of the United States”, scale 1:7,000,000.

Frankel, A. D., D. L. Carver, and R. A. Williams [2002]. “Nonlinear and linear site response and basin effects in Seattle for the M 6.8 Nisqually, Washington, earthquake”, Bull. Seism. Soc. Am., Vol. 92, No. 6, pp. 2090-2109.

Frankel, A. D., S. Harmsen, C. Mueller, E. Calais, and J. Haase [1982]. “Documentation for initial seismic hazard maps for Haiti”, U.S. Geol. Surv. Open-File Reprt. 2010-1067, 12 pp.

Franklin, S. E. [1987]. “Geomorphometric processing of digital elevation models”, Comp. and Geosci., Vol. 13, No. 6, pp. 603-609.

Fumal, T. E. [1978]. “Correlations between seismic wave velocities and physical properties of near-surface geologic materials in the southern San Francisco Bay region, California”, U.S. Geol. Surv. Open-File Reprt. 78-1067, 114 pp.

Graf, W. L. E. [1987]. “Geomorphic Systems of North American; Centennial Special Volume 2”, Geological Society of America, Inc., Boulder, 643 pp.

Hammond, E. H. [1954]. “Small-scale continental landform maps”, Ann. Assoc. Am. Geogr, Vol. 44, pp. 33-42.

Hammond, E. H. [1964]. “Analysis of properties in land form geography: an application to broad-scale landform mapping”, Ann. Assoc. Am. Geogr. Vol. 54, pp. 11-19.

Harden, J. W. [1990]. “Soil development on stable landforms and implications for landscape studies”, Geomorphology, Vol. 3, pp. 391-398.

Holzer, T. L., M. J. Bennett, T. E. Noce, and J. C. Tinsley [2005]. “Shear-wave velocity of surficial geologic sediments in northern California: Statistical distributions and depth dependence”, Earthq. Spectra, Vol. 21, No., 1, pp. 161-177.

Iwahashi, J., and R. J. Pike [2007]. “Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature”, Geomorphology, Vol. 86, pp. 409-440.

Iwahashi, J., I. Kamiya and M. Matsuoka [2010]. “Regression analysis of Vs30 using topographic attributes from a 50-m DEM”, Geomorphology, Vol. 117, pp. 202-205.

Jensen, J. R. [2007]. “Introductory Digital Image Processing; A Remote Sensing Perspective” Prentice-Hall, Inc., Upper Saddle River, 526 pp.

Joyner, W. B., R. E. Warrick, and T. E. Fumal [1981]. The effect of Quaternary alluvium on strong ground motion in the Coyote Lake, California, earthquake of 1979, Bull. Seism. Soc. Am., Vol. 71, No. 4, pp. 1333-1350.

Lee, V. W., M. D. Trifunac, M. I. Todorovska, and E.I. Novikova [1995]. “Empirical equations describing attenuation of the peaks of strong ground motion, in terms of magnitude, distance, path effects and site conditions Los Angeles, California”, Reprt CE 95-02, 415 pp.

Lee, V. W., and M. D. Trifunac [2010]. “Should average shear-wave velocity in the top 30 m of soil be used to describe seismic amplification?”, Soil Dyn. Earthq. Eng., Vol. 30, pp. 1250-1258.

Martin, G. R., and R. Dobry [1994]. “Earthquake site response and seismic code provisions”, NCEER Bull., Vol. 8, pp. 1-6.

Minár, J. and I. S. Evans [2008]. “Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping”, Geomorphology, Vol. 95, pp. 236-259.

Moss, R. E. S. [2008]. “Quantifying measurement uncertainty of thirty-meter shear-wave velocity”, Bull. Seism. Soc. Am., Vol. 98, No. 3, pp. 1399-1411.

Murray, B. and M. A. Fonstad [2007]. “Preface: Complexity [and simplicity] in landscapes”, Geomorphology, Vol. 91, pp. 173-177.
Oliver, M. A. and R. Webster [1986]. “Semi-variograms for modeling the spatial pattern of landform and soil properties”, Earth Surface Processes and Landforms, Vol. 11, pp. 491-504.

Olsen, K. B., S. M. Day, J. B. Minster, Y. Cui, A. Chourasia, M. Faerman, R. Moore, P. Maechling, and T. Jordan [2006]. “Strong shaking in Los Angeles expected from southern San Andreas earthquake”, Geophys Res. Lett., Vol. 33, p. L07305.

Park, S., and S. Elrick [1998]. “Predictions of shear-wave velocities in southern California using surface geology”, Bull. Seism. Soc. Am., Vol. 88, No. 3, pp. 677-685.

Petersen, M., W. Bryant, C. Cramer, M. Reichle and C. Real [1997]. “Seismic ground-motion hazard mapping incorporating site effects for Los Angeles, Orange, and Ventura counties, California”, Bull. Seism. Soc. Am., Vol. 87, No. 1, pp. 249-255.

Peterson, F. F. [1981]. “Landforms of the basin and range province; defined for soil survey”, Nevada Agricultural Experiment Station- Max C. Fleischmann College of Agriculture- University of Nevada Technical Bulletin 28, 52 pp.

Philips, J. D. [2007]. “The perfect landscape”, Geomorphology, Vol. 84, pp. 159-169.

Pike, R. J., I. S. Evans, and T. Hengl [2009]. “Geomorphometry: a brief guide”, in Geomorphometry; Concepts, Software, Applications, Amsterdam, pp. 3-30.

Romero, S. M., and G. J. Rix [2001]. “Ground motion amplification of soils in the upper Mississippi embayment”, Mid-America Earthquake Center at School of Civil and Environmental Engineering at Georgia Institute of Technology Report No. GIT-CEE/GEO-01-1, pp. 46-112.

Saco, P. M., G. R. Willgoose, and G. R. Hancock [2006]. “Spatial organization of soil depths using a landform evolution model”, J. Geophys. Res., 111, F02016, doi:10.1029/2005JF000351.

Sánchez-Sesma, F. J. [1985]. “Diffraction of elastic SH waves by wedges”, Bull. Seism. Soc. Am., Vol. 75, No. 5, pp. 1435-1446.

Scripter, M. W. [1970]. “Nested-means map classes for statistical maps”, Ann. Assoc. Am. Geogr, Vol. 60, pp. 385-393.

Seed, H. B., C. Ugas, and J. Lysmer [1976]. “Site-dependent spectra for earthquake-resistant design”, Bull. Seism. Soc. Am., Vol. 66, No. 1, pp. 221-243.

Tobler, W.R., [1988], “Resolution, resampling and all that”, in Building Databases for Global Science, pp. 129-137.

Thompson, E. M., L. G. Baise, R. E. Kayen, Y. Tanaka, and H. Tanaka [2010]. “A geostatistical approach to mapping site response spectral amplifications”, Eng. Geol., Vol. 114, pp. 330-342.

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. I. Ziony [Editor], U.S. Geol. Surv. Profess. Pap. 1360, pp. 101-126.

VIC [1993]. “Shear wave velocity profiling of instrumented sites in the greater Los Angeles area using surface wave measurements”, in Report by Vibration Instruments Company, Tokyo, to the University of Southern California, Department of Civil Engineering.

Wald, D. J., and T. I. Allen [2007]. “Topographic slope as a proxy for seismic site conditions and amplification”, Bull. Seism. Soc. Am., Vol. 97, No. 5, pp. 1379-1395.

Warnes, G. R. [2004]. “Plot x-y points with locally smoothed mean and standard deviation”, http://pbil.univ-lyon1.fr/library/gregmisc/html/bandplot.html, (last accessed April 1, 2009).

Wills, C. J. [1998]. “Differences in shear-wave velocity due to measurement methods: a cautionary note”, Seism. Res. Lett., Vol. 69, No. 3, pp. 216-221.

Wills, C. J., and K. B. Clahan [2006]. “Developing a map of geological defined site-condition categories for California”, Bull. Seism. Soc. Am., Vol. 96, No. 4A, pp. 1483-1501.

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 the California based on geology and shear-wave velocity”, Bull. Seism. Soc. Am., Vol. 90, No. 6B, pp. S187-S208.

Wills, C. J., and W. Silva [1998]. “Shear-wave velocity characteristics of geologic units in California”, Earthq. Spectra, Vol. 14, No. 3, pp. 533-556.

Yong, A., S. Hough, K. Hudnut, S. Healy, J. Giberson, and G. Simila [2005]. “Geotechnical site characterization using remote sensing data: A feasibility study”, Seism. Res. Lett., Vol. 76, No. 2, p. 241 (Abstract).

Yong, A., S. Hough, H. Cox, K. Tiampo, A. Braverman, J. Harvey, S. Hook, K. Hudnut, and G. Simila [2006], “Balloons to satellites: A century of progress in geotechnical site characterization”, Seism. Res. Lett., Vol. 77, No. 2, p. 302 (Abstract).

Yong, A., S. E. Hough, C. Wills, and M. Abrams, [2007], “Site Characterization Using Satellite Imagery: A Progress Report”, Seism. Res. Lett., Vol. 78, No. 2, p. 272 (Abstract).

Yong, A., S. E. Hough, M. J. Abrams, H. M. Cox, C. J. Wills, and G. W. Simila [2008a]. “Site characterization using integrated imaging analysis methods on satellite data of the Islamabad, Pakistan, region”, Bull. Seism. Soc. Am., Vol. 98, No. 6, pp. 2679-2693.

Yong, A., S. E. Hough, M. J. Abrams, and C. J. Wills [2008b]. “Preliminary results for a semi-automated quantification of site effects using geomorphometry and ASTER satellite data for Mozambique, Pakistan and Turkey”, J. Earth Syst. Sci., Vol. 117, No. S2, pp. 797-808.

Yong, A., S. E. Hough, A. Braverman, and J. Iwahashi [2010]. “A terrain-based Vs30 estimation map of the contiguous United States”, Seism. Res. Lett., Vol. 81, No. 2, p. 294 (Abstract).