Crespo, M. J., Benjumea, B., Moratalla, J. M., Lacoma, L., Macau, A., González, Á., Gutiérrez, F., Stafford, P. J. (2022): A proxy-based model for estimating V30 in the Iberian Peninsula. - Soil Dynamics and Earthquake Engineering, 155, 107165.

https://doi.org/10.1016/j.soildyn.2022.107165

A proxy-based model for estimating V_{S30} in the Iberian Peninsula

Maria J. Crespo[1,2] ORCID ID 0000-0002-0648-2253
Beatriz Benjumea[3,*] ORCID ID 0000-0002-0673-3411
José M. Moratalla[4] ORCID ID 0000-0001-8413-2606
Luis Lacoma[1]
Albert Macau[3] ORCID ID 0000-0001-8315-9831
Álvaro González[5,6] ORCID ID 0000-0002-9300-4283
Francisco Gutiérrez[7]
Peter J. Stafford[8] ORCID ID 0000-0003-0988-8934

[1] Principia Ingenieros Consultores, Velázquez 94, 28006 Madrid, Spain
[2] Universidad Politécnica de Madrid - E.T.S.I. Minas y Energía, Rios Rosas, 21, 28003 Madrid, Spain
[3] Institut Cartogràfic i Geològic de Catalunya, Barcelona, Spain
[4] Department of Natural Hazards and risks, GNS Science, Lower Hutt, New Zealand
[5] Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, 08193 Bellaterra, Spain
[6] GFZ German Research Centre for Geosciences, Telegrafenberg, 14467 Potsdam, Germany
[7] Departamento de Ciencias de la Tierra, Universidad de Zaragoza, Campus Plaza San Francisco, 50009 Zaragoza, Spain
[8] Department of Civil & Environmental Engineering, Imperial College London, London SW7 2AZ, UK
[*] Present address: CN Instituto Geológico y Minero CSIC. La Calera, 1. 28760 Tres Cantos, Madrid, Spain

Abstract

The time-averaged shear-wave velocity in the upper 30 meters of the ground, V_{S30}, is a key soil descriptor for estimating site response despite its recognized limitations. It is employed in both, site-specific probabilistic hazard assessments (PSHAs) and regional seismic codes.

This work presents a model for estimating V_{S30} in the Iberian Peninsula as a function of three proxies: topographic slope, geological age and lithology at each site. Tasks accomplished include: 1) gathering existing V_s profiles and calculating their V_{S30}; 2) defining an adequate set of representative age and lithological groups; 3) classifying the available V_S profiles according to these groups; and 4) carrying out a regression analysis between V_{S30}, slopes, age and lithological groups.

Based on the regression analysis and the dependency on the slope, some of the initially proposed groups were amalgamated, before proposing the final model. This model considers topographic slope values extracted from a digital elevation model (DEM) with 200 m horizontal resolution, plus six geological age groups and four lithological groups. It provides an estimate of the mean and standard deviation of $\log V_{S30}$ (and hence V_{S30}), which can be used for sites without direct estimates of velocity profiles (and V_{S30}) in the Iberian Peninsula.

Keywords: V_{S30}, shear wave velocity, proxy, topographic slope, geological age, lithology, ground motion, Iberian Peninsula

1 Introduction

For many applications in the fields of earthquake engineering and engineering seismology, it is useful to have information on how the shear waves propagate in the vicinity of the recording stations or at another site of interest. One of the parameters correlating with ground-motion amplification is the time-averaged shear-wave velocity in the upper 30 m of the ground, the so called V_{S30}. It is widely recognized that the V_{S30} has limitations and is far from presenting a complete description of the soil behavior (Boore et al., 2011; Seyhan et al., 2014), but in practical terms it is still employed in many design codes as a representative descriptor and it is foreseen that the V_{S30} will still be employed in seismic codes for some years. However, shear-wave velocity measurements are not usually carried out in conventional geotechnical site characterizations that focus upon static attributes or they are not available at the initial stages of a project. In many other applications, such as computing ShakeMaps, performing regional earthquake loss estimation, or portfolio risk assessments, it is practically infeasible to conduct field investigations to infer V_{S30} values at all sites of interest. Simple methods for estimating V_{S30} without undertaking site-specific investigations can be valuable in such situations. In the absence of in-situ measured V_S data, it is very common to estimate V_{s30} as a function of indirect descriptors (proxies) of the site. A typical application of this procedure is characterizing seismic station sites in networks, for instance, in France (Hollender et al., 2018), Italy (Foti et al., 2011), and Switzerland (Michel et al, 2014), the development of ground motion prediction equations (Seyhan et al., 2014), a study of its performance by comparison with new measurements (Savvaidis et al., 2018) or the construction of V_{S30} maps (Heath et al., 2020). Hence both, practitioners and researchers can benefit from a proxy-based V_{S30} model. Ideally, the proxy-based prediction model should be constructed with data that share certain characteristics with the sites where such estimation is needed.

Several proxy-based models have been developed in the last two decades for predicting the V_{S30} value as a function of several terrain descriptors. Some global models have been proposed. The most commonly adopted models is that of Wald & Allen (2007), which was applied to the entire globe by Allen & Wald (2007). The results from that application are available through the USGS web page. The method of Wald & Allen (2007) is based on the topographic slope from a digital elevation model (DEM), in particular, the DEM derived from the Shuttle Radar Topography Mission at 30 arcsec resolution (SRTM30). Allen & Wald (2007) assigned a fixed value of 180 m/s to all sites with a slope below a certain value, and of 760 m/s to all sites with a slope above another threshold. This upper boundary is a potential limitation for use in Spain where relatively hard rock conditions exist, and the restriction of 760m/s doesn't allow for discrimination in site resposne over these stiffer sites. Moreover, although applying a global model is theoretically feasible (since it is based on a globally available proxy), in practice it usually needs to be adapted to the specific region where it will be applied. This adaptation is frequently accomplished by combining the slope with other descriptors such as geological or geomorphological units, lithology, or elevation. One of the first works which used geology-based proxies was that of Tinsley & Fumal (1985), who showed a strong correlation between Quaternary deposits and variation in ground motion amplification. A good literature summary of proxy-based methods for V_{S30} estimation was presented by Ahdi et al. (2017) in their Table 1.

For the particular case of the Iberian Peninsula, there are three works related to the prediction of V_{S30}:

– Núñez et al. (2012) presented a map for the Iberian Peninsula and the Balearic Islands with the six NEHRP site classes (each having a representative V_{S30}) defined by the Building Seismic Safety Council (1998). The class assignments were inferred from the geological units depicted in a 1:1,000,000 scale map (IGME, 1994). This work could have provided a very interesting reference for comparison with the results from the present study. Unfortunately, however, the Núñez et al. (2012) work adopted an ambiguous map projection that prevents direct comparisons from being made (although broad qualitative comparisons remain possible). Núñez et al. (2012) noted that the calculated isoseismals for several past earthquakes were more realistic if the site effects resulting from their map were taken into account, but they did not calibrate this map directly with V_{S30}. Moreover, the discrepancies between the contacts of the geological units depicted in the geological map by IGME (1994) and the most recent and detailed one (IGME, 2015) are on the order of 1 km to 3 km. So, a direct calibration with measured values of V_{S30} and the use of newer geological maps, especially with higher resolution, should yield more robust results and avoid geological misclassifications of the sites.

– The work by Vilanova et al. (2018) covers just the mainland Portuguese territory, which only represents around 16% of the Iberian Peninsula (92,000 km^2 out of 582,000 km^2) and has a significantly more restricted geological diversity. They proposed a geology-based proxy that finally considers three lithological groups (igneous, metamorphic and old sedimentary rocks plus Neogene and Quaternary formations). This geological classification combines lithology and age.

– The work by Sá et al. (2020) requires the estimation of the V_{S30} for the SW of the Iberian Peninsula in a study of loss assessment and, in the absence of a local V_{S30} model, they employ two generic proxy-based models based on geology (Wills & Clahan, 2006) and on topographic gradient (Wald & Allen, 2007), finding a better performance of the geologically based model.

The goal of the present work is to develop a proxy-based V_{S30} model for the Iberian Peninsula, overcoming the mentioned limitations of previous works and being more accurate than global methods. This model is calibrated using a database of measured V_S profiles and/or V_{S30} values covering an ample range of geological domains of the Iberian Peninsula. This is the first compilation of its kind for mainland Spain. Figure 1 shows the location of the sites with measured V_{S30} used as input in this work, plotted on a map of the geological domains of the Iberian Peninsula and including the measurements presented in Vilanova et al. (2018). Most of the measurements located in Portugal correspond to three domains: the Lower Tagus Basin, the Iberian Massif (Variscan Orogen) and the Western Margin, the remaining measurements are located in the Cenozoic Belt and Tertiary and Quaternary surficial sediments of southern Portugal. As shown in Figure 1, the Spanish portion of the Iberian Peninsula includes five large Cenozoic sedimentary basins and several Alpine mountain ranges.

In order to fulfil the above objective, the following tasks have been undertaken:

a) Compilation of V_S profiles. In general, the documentation identified for a particular site includes one or more V_S profiles obtained using either a single or several shear-wave velocity measurement techniques. In some cases, the source includes just the V_{S30} value.
b) Compilation, including calculation when needed, of V_{S30} values.

c) Identification of candidate proxies and acquisition of spatially-distributed data covering the whole Iberian Peninsula.
d) Establishment of an adequate classification (categorization) for the geological/lithological proxies in accordance with the number of sites for which V_S profiles are available.
e) Numerical analysis to establish correlations between the measured V_{S30} and the identified proxies.
f) Check circumstances affecting points that present high residuals or that deviate excessively from the general tendency of the sample.
g) Propose the most suitable proxy-based model.

2 V_{S30} Database

2.1 V_S profile collection

A total of 580 sites were identified for which one or more in-situ measurements of the shear-wave velocity (V_S) are available. The type of work that motivated those measurements can be divided into the following categories:

- PhD Theses. This category includes measurements compiled for investigations conducted within four PhD Thesis (Clavero, 2014; Feriche, 2013; Olona, 2014; Pérez, 2011), including a total of 29 measurements.
- Public Databases. This category includes measurements compiled and provided by the ICGC (Institut Cartogràphic i Geològic de Catalunya) with a total of 85 sites.
- Journals and Conferences. This category compiles measurements used in works described in papers published in journals and/or conference proceedings and not belonging to the previous category. A total of 322 sites coming from eight publications have been included (García-Fernández & Jiménez, 2012; Carvalho et al, 2013; Alguacil et al, 2014; Dias et al, 2014; Navarro et al, 2014a,b; Rueda et al, 2015; Vilanova et al., 2018;). From the work by Vilanova et al. (2018), only the measurements not already included in the publications of Carvalho et al. (2013) and Dias et al. (2014) were added to avoid duplication.
- Private Projects. These are geophysical campaigns generally conducted for civil works with stringent requirements in terms of some type of dynamic load. This type of geophysical campaign is mainly carried out for the nuclear and gas industry, as well as for the construction of the Spanish high-speed railway. This category also includes 52 unpublished measurements contributed by M. Navarro (*pers. comm.*). The total number of sites coming from private projects is 144.

In some cases, only the V_{S30} has been provided by the authors, and not the complete V_S profile as a function of depth. Namely, for the data supplied by M. Navarro (*pers. comm.*), and the data from García-Fernández & Jiménez (2012), Carvalho et al. (2013), Dias et al. (2014), Rueda et al. (2015) and Vilanova et al. (2018).

The origin of the sites composing the database is distributed as indicated in Table 1.

The authors acknowledge the limitations of the database, in particular its small size considering the area covered. To our knowledge, this is, as of today, the largest compilation of publicly availabe V_S profiles in the Iberian Peninsula, including also some privately owned ones (Table 1). The goal is to develop the best possible proxy-based model with the existing information, always improving the already existing models; the database is intended to be in continuous enlargement with new measurements so as to conveniently update the model resulting from this work as and when the updated database allows a significant improvement.

Table 1. Sources of the V_{S30} measurements from the 580 sites included in the database.

Source of V_{S30} data	Percentage
Journals and conferences	55.7%
Private projects	24.8%
Public database	14.5%
PhD theses	5.0%

2.2 General characteristics of the database

The geographical and geological distribution of the sites is illustrated in Figure 1, where the 580 sites are plotted. There are 160 sites in Portugal (about 27% of the total), with all others in Spain. This proportion is about 1.7 times that of the mentioned proportion (16%) of the area of the Iberian Peninsula covered by Portugal.

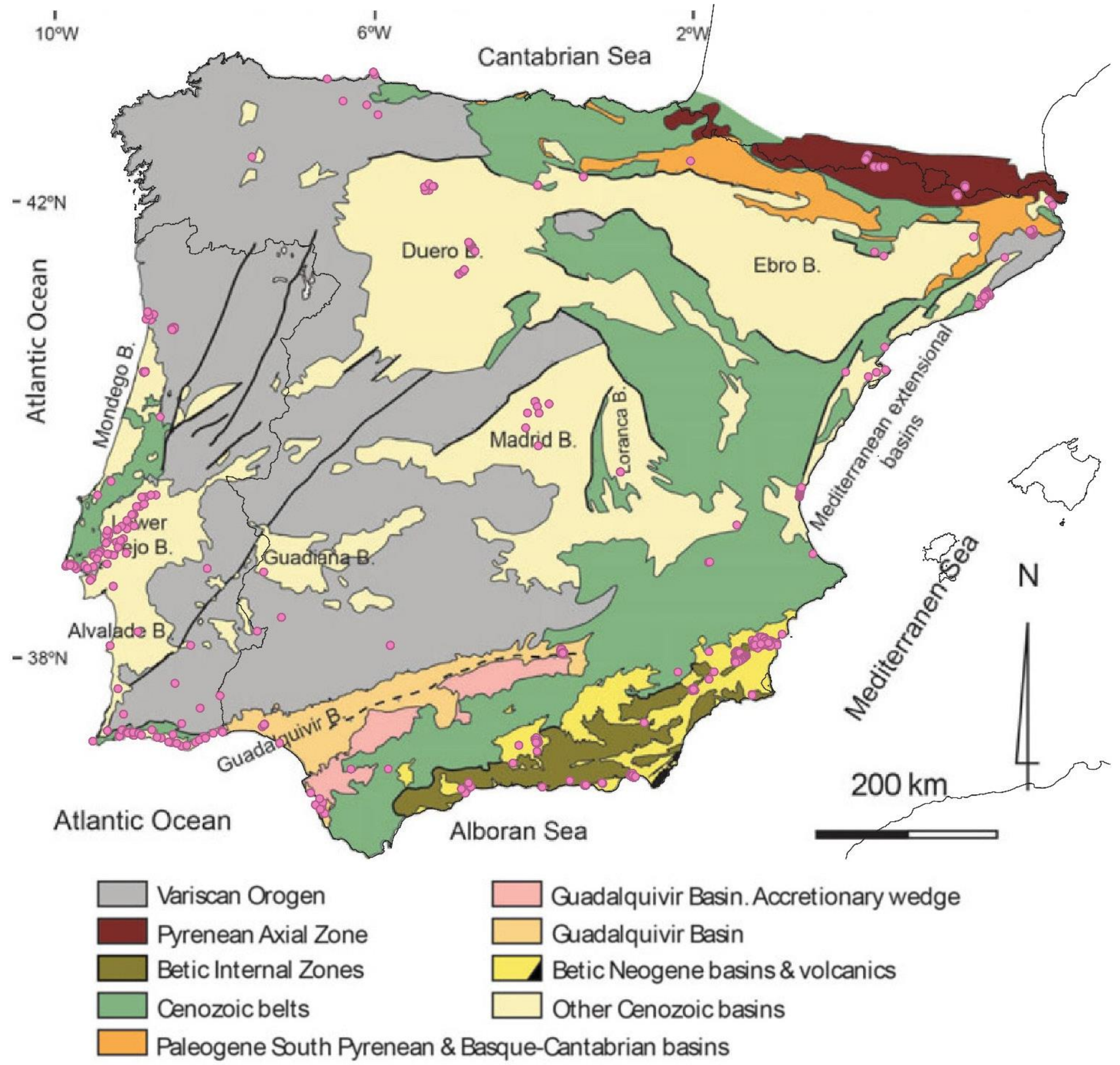

Figure 1. Distribution of the final 580 sites (pink points) with V_S profiles represented on a map depicting the main geological domains of the Iberian Peninsula (Braga & Cunha, 2019). Basins are generally related to the presence of unconsolidated sediments at surface.

Velocity measurements of different types were acquired. For each site there may be just one measurement or several measurements either of the same type or of different types. In the case of sites with several measurements, and lacking judgment from the original author(s) about the quality of the data, if all measurements exceed 30 m depth, the order of preference established in this work has been: down-hole, cross-hole and then borehole logging (concurrence of several measurements only occurs for these three types). When the measurements are of the same type, the deepest one has been given priority. There are other measurements, in particular those published or owned by M. Navarro, for which several types of non-invasive techniques were available, but the owner has explicitly indicated the order of preference.

The distribution of the sites in the zones identified by Braga & Cunha (2019) (Figure 1) can be seen in Table 2. As reflected in the figure and verified in the table, most of the sites fall in the Betic basins and in other Cenozoic basins (category which includes the Ebro, Duero, Madrid, Loranca, Guadiana, Mondregu and Tagus Basins). And the least populated categories are the two ones related to the Guadalquivir Basin with 6 sites each, which is about 1% of the total number of sites.

The type of measurements selected for the 580 sites can be classified as follows according to the method applied:

- Cross-hole: There are four sites for which this type of V_S profile was adopted.
- Down-hole: There are 20 sites with a down-hole measurement as a preferred V_S profile.
- Passive MASW (Multichannel Analysis of Surface Waves): There are 138 sites with passive MASW, in all cases it is the Refraction Microtremor technique (ReMi).
- Seismic Cone Penetration Test (SCPTu): There are 4 sites with this type of measurement as the only available one. These 4 sites come from the work of Vilanova et al. (2018).
- Active MASW: The database includes 122 measurements of this type. Many of them have been provided by the ICGC, and some come from privately-owned databases. The work of Vilanova et al. (2018) includes some 60 measurements of this type.
- 2D Passive array: There is a total of 103 sites with this type of measurement coming from PhD Theses, the ICGC database, published papers (mainly with M. Navarro as author), and privately owned.
- Shear-wave refraction: These sites, which contribute 85 measurements, come from the works of Carvalho et al. (2013), Dias et al. (2014) and Vilanova et al. (2018).
- H/V Inversion: The 54 sites of this type are derived from the work of García-Fernández & Jiménez (2012).
- Mini-Array: This is a novel technique (Cho et al., 2013) being applied for the first time in Spain by M. Navarro. At present there are 50 sites with this type of measurement (Candela-Medel et al., 2018).

The distribution of the sites in terms of the type of selected measurement can be seen in Table 3.

Table 2. Distribution of sites in the different zones identified by Braga & Cunha (2019).

Geological zone	Number of sites	Percentage
Variscan Orogen	66	11.4%
Pyrenean Axial Zone	14	2.41%
Betic Internal Zones	10	1.72%
Cenozoic Belts	46	7.93%
Paleogene South Pyrenean & Basque Cantabrian Basins	13	2.24%
Guadalquivir Basin.	6	1.03%

Accretionary Wedge.		
Guadalquivir Basin	74	12.76%
Betic Neogene Basins & Volcanics	170	29.31%
Other Cenozoic Basins	181	31.20%

Table 3. Methods used for the retrieval of V_S information included in the database.

Method for V_S retrieval	Percentage
Active MASW	21.0%
2D Passive array	17.7%
Passive MASW	23.8%
S-wave Refraction	14.7%
H/V Inversion	9.3%
Mini-Array	8.6%
Down-hole logging	3.5%
Cross-hole logging	0.7%
SCPTu	0.7%

The database we compiled for each site includes the following entries: latitude, longitude, elevation, data source, provider, measurement technique, investigation depth, V_{S30} (calculated or reported).

2.3 V_{S30} calculation

There are 362 sites, all 160 sites in Portugal and 202 in Spain (those from García-Fernández & Jiménez, 2012, Rueda et al., 2015, and those provided by M. Navarro), for which the V_S profile has not been made available to us, and just the V_{S30} has been provided by the authors. When the profile did not reach 30 m depth, M. Navarro assumed a constant extrapolation from the deepest measured value down to 30 m. In the Portuguese database there are two assumptions, one of them being equivalent to the criteria indicated by M. Navarro. The type of assumption made in this respect is not explicit for the dataset by García-Fernández & Jiménez (2012). An alternative approach could have been to adopt Boore (2004) approach for inferring V_{S30} from velocity profiles shallower than 30 m, however, since for a significant number of places we only had the V_{S30} and the predominant criterion was the one indicated above, we adopted it for the remaining sites that needed such an extrapolation.

For the sites with a measured V_S profile, the time averaged V_{S30} has been calculated as follows:

$$V_{S30} = \frac{30\ m}{\sum t_i}$$

$$t_i = \frac{\mathrm{z_i}}{V_{\mathrm{si}}}$$

where V_{Si} is the shear-wave velocity for each of the i different soil layers and z_{i} the respective thicknesses bounded to cover exactly the first 30 m.

The distribution of V_{S30} values and that for their decimal logarithm for the 580 sites is presented in Figure 2. The log V_{S30} histogram shows that this measurement follows a normal distribution, which is consistent with earlier reported findings (Boore et al, 2011; Vilanova et al, 2018). The decimal logarithm, referred to simply with "log" in the formulae, will be employed in this study.

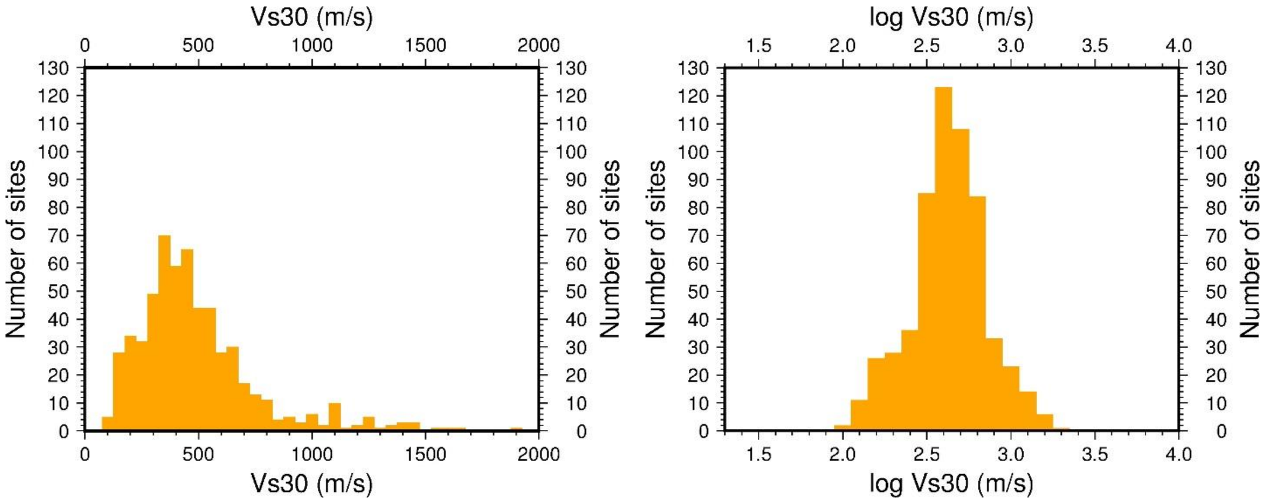

Figure 2. Histograms of the V_{S30} and log V_{S30} for all sites in the database.

3 Candidate proxies. Metadata compilation.

3.1 Gradient (slope)

For calculating the slope gradient, a digital elevation model (DEM) is needed. Three DEMs have been provided by the Spanish National Geographic Institute (IGN) with three different horizontal resolutions: 500 m, 200 m and 20 m. A new 1000 m resolution DEM was constructed out of the 500 m one by resampling values with the nearest neighbour algorithm. This 1000 m DEM has the purpose of approaching the 30 arcsec resolution employed by several authors, in particular Wald & Allen (2007). Some authors indicate that working with resolutions lower than 3 arcsec (~90 m) reduces performance (Stewart et al, 2014; Vilanova et al, 2018), which is also consistent with what was already indicated by Allen & Wald (2009).

The three DEM models retained for further work are the 1000 m, 500 m and 200 m ones. A digital slope model (DSM) has been constructed from each DEM based on the first-order derivative estimation. Figure 3 shows one of these three DSM, namely the 200 m one.

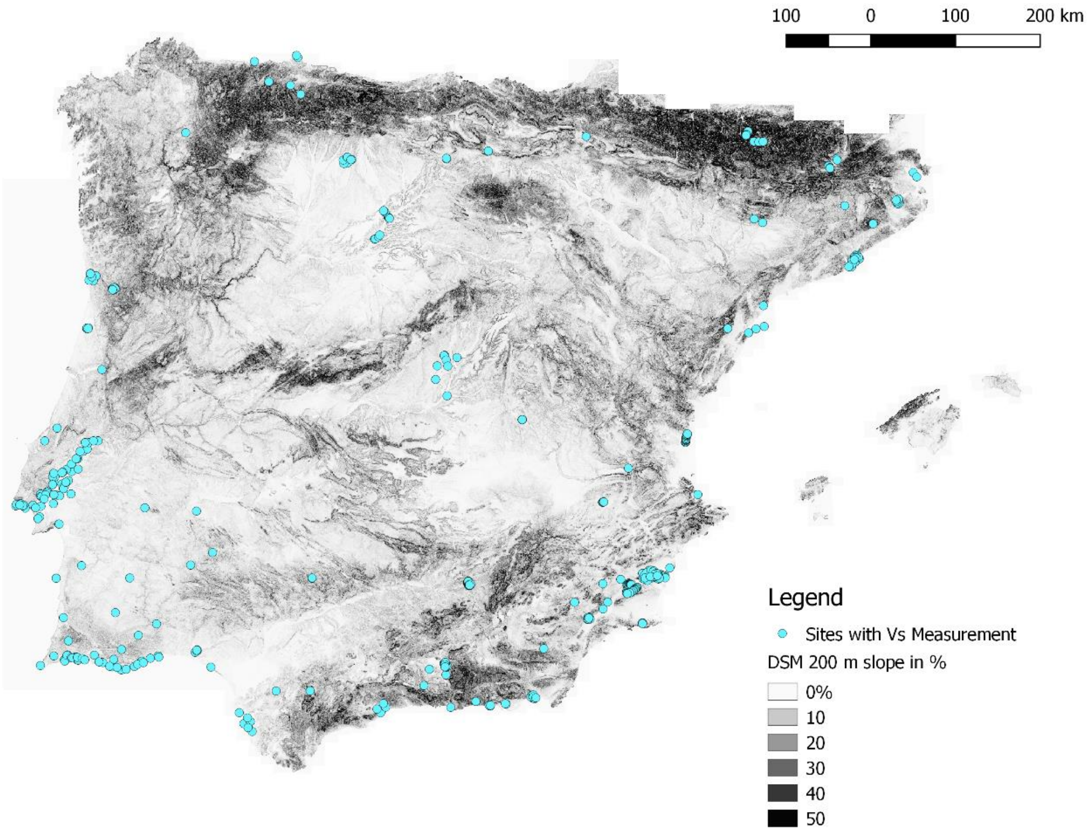

Figure 3. Digital slope model at 200 m resolution (constructed out of the 200 m DEM provided by IGN) with location of sites with V_S measurements in the data base.

The distributions of slopes obtained for the 580 sites in the database and from the DEMs with the three different resolutions are presented in Figure 4. A slight shift towards higher slopes is observed in the data obtained from the higher resolution models, which is expected due to the effect of smoothing in the lower resolutions.

Having the slopes and the V_{S30} values (section 2.3), it is possible to plot the V_{S30} values versus the slope. The three resulting plots, considering the three DSM with different resolution, are presented in Figure 5. The result of the least squares linear fit in log-log scale is very similar in the three cases, with a slight shift towards higher slopes and a better coefficient of determination, as expected, as the resolution increases. Based on these results and the considerations indicated at the beginning of this section, the 200 m DEM has been retained for the subsequent calculations.

It is also interesting to compare these regressions with the model proposed by Wald & Allen (2007) based on a 30 arcsec resolution map (ca. 1000 m). Figure 6 allows comparing our data and its fit for the slope derived from the 1000 m resolution map with the model proposed by Wald & Allen (2007), both for active continental regions (ACR) and stable continental regions (SCR). Considering the regression obtained with our data as reference, the ACR model tends to underestimate V_{S30} in the range with most data (log of slope between -1 and 1), while the SCR model overestimates it above V_{S30} ~ 300 m/s. This comparison, which reflects that a generic proxy like the one by Wald & Allen (2007) is not satisfactory, motivated going ahead with this work in the search of an improved model.

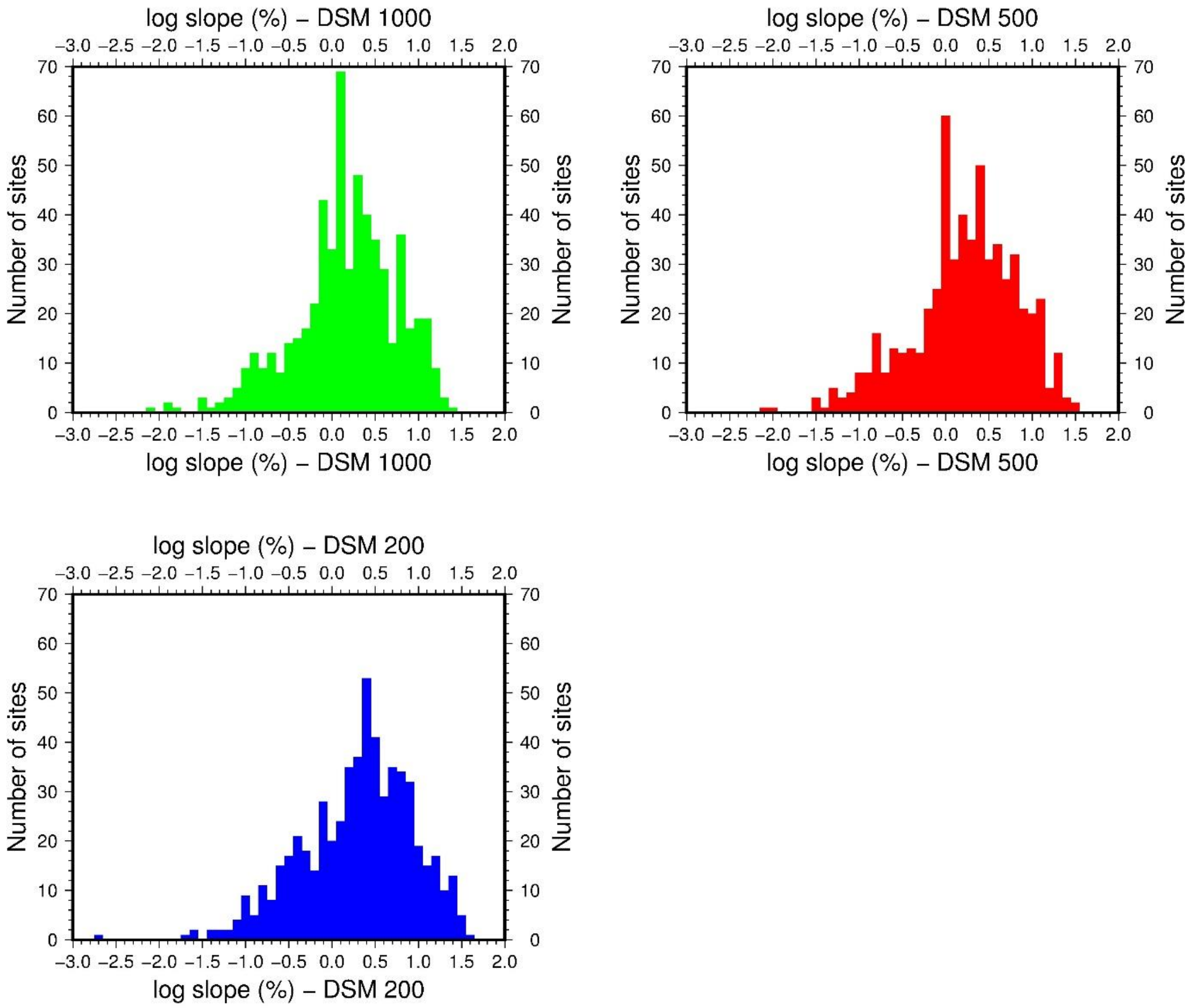

Figure 4. Distribution of slopes for the 580 sites in the database derived from the three DSMs with different resolutions.

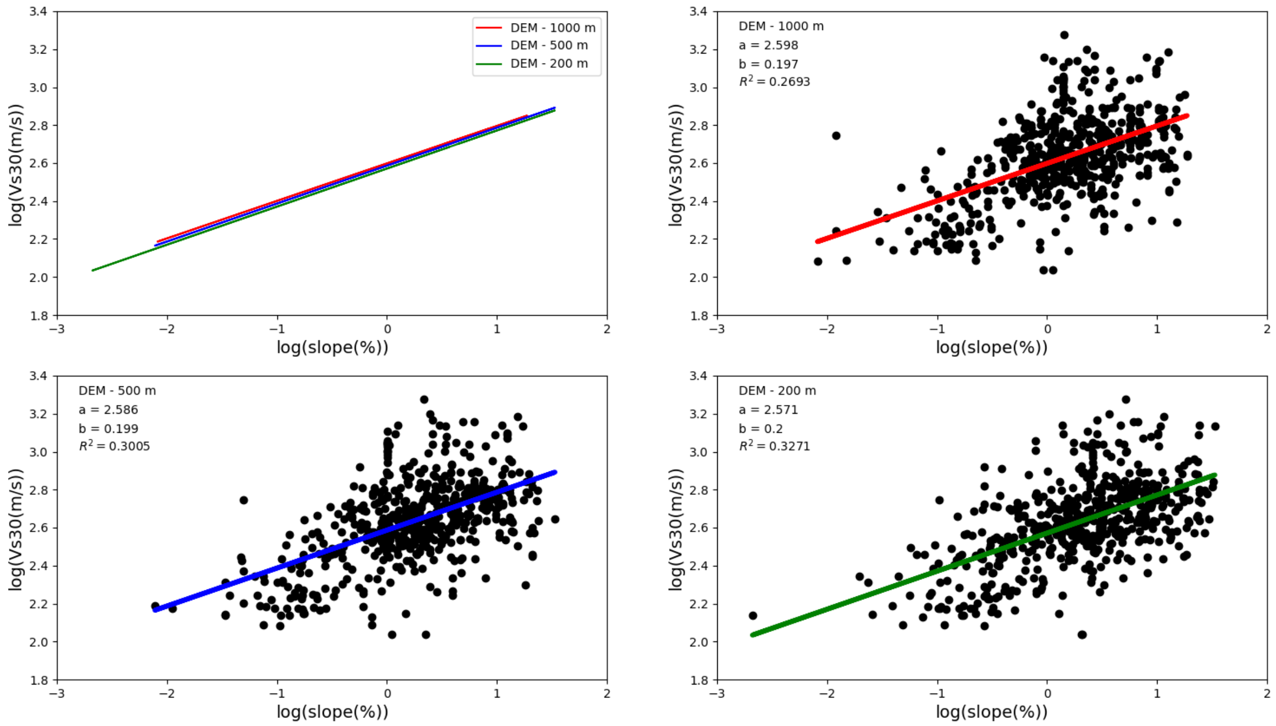

Figure 5. V_{S30} dependency on topographical slopes for the 580 sites in the database and for the three DEM resolutions. Solid lines represent least squares fits for all points and three different DEM resolutions; the top left plot shows the same three fits together.

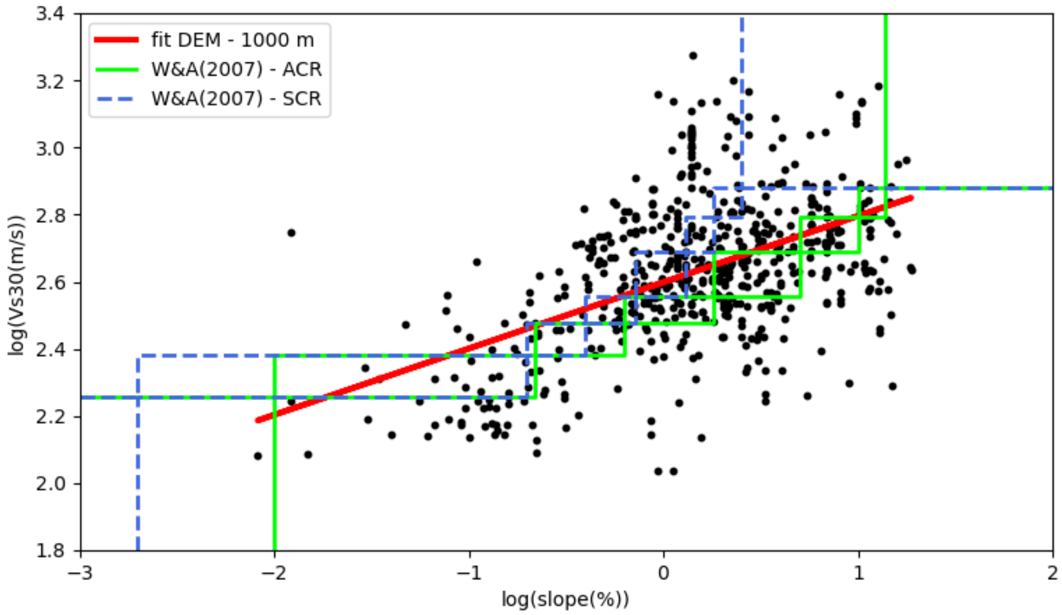

Figure 6. Model proposed by Wald & Allen (2007) (green and blue lines) compared to the data of the Iberian Peninsula (red line) (this work), both using slope data derived from 30 arcsec resolution DEMs.

3.2 Lithology and geological age. Sources of information

The official geological cartography of Spain is compiled and published at different scales by the Spanish Geological Survey (Instituto Geológico y Minero de España, IGME). We decided to use the 1:50,000 scale geological maps, since these are the ones with the highest resolution which cover all the Spanish territory and because they specify, for each geological unit with a given age, the corresponding lithology (one or several kinds of rocks or surficial deposits, in order of relative frequency, for each unit). This allows identifying the nature of the material outcroping the sites with higher confidence than in other maps.

The IGME has published two series of 1:50,000 scale maps, MAGNA and GEODE (IGME, 2020). The latter one, resulting from the digitisation, homogenisation and refinement of the MAGNA series, was the preferred one. The Catalonian region is not included in the GEODE map. However, there is a 1:50,000 scale map for Catalonia, the Geoíndex, uniform all across Catalonia, maintained by the ICGC, which provides similar information to that in the GEODE. Both the GEODE and the Geoíndex are vectorized and available online through a Web Map Service (WMS).

Neither MAGNA nor GEODE cover the Portuguese region. The Portuguese Geological Survey (Laboratório Nacional de Energia e Geologia, LNEG) has several geological maps, including a 1:50,000 scale series, but it does not cover the entire Portuguese territory. This map is also available through a WMS, although it consists of a non-georeferenced scanned copy (raster format), so the geological/lithological assignments cannot be carried out automatically.

The retrieval of lithologies from the GEODE and Geoíndex maps was performed with a Python script. For the Portuguese sites, S. Vilanova provided the assignment made for their work (Vilanova et al, 2018), which included most of the sites they had used. A manual assignment was performed for the remaining sites. Most of the sites fall on a 1:50,000 scale map and the 1:200,000 scale maps were employed for a small fraction of sites.

Table 4 presents the number of sites for the three sources of information, GEODE, Geoíndex and LNEG maps, and the number of individual different lithological descriptions for the sites in each case. The number in brackets shows the decomposition for the Portuguese sites between descriptions in the 1:50,000 scale map, which is the case for most sites, and the 1:200,000 scale map.

Table 4. Number of sites falling on each map and resulting number of lithological descriptions.

Source	# of sites	# of descriptions
GEODE	367	72
Geoíndex	53	30
LNEG	160	79 (68 + 11)
Total	580	181

4 Proxy Development Method

4.1 Groups of geological age

The first step in the grouping process is to use geological age as a potential attribute that helps to obtain distinctive group characteristics such as mean or standard deviation. Figure 7 shows the geological map of Spain with a 1:1,000,000 scale with the distribution of available V_S measurements over the geological age layer. This map is presented only for illustrative purposes since larger scale geological maps have been used for the proxy's analysis. Unfortunately, a continuous geological age map of the whole Iberian Peninsula is not available, and we show the Spanish portion since it includes a larger variability in geological units than the Portuguese portion.

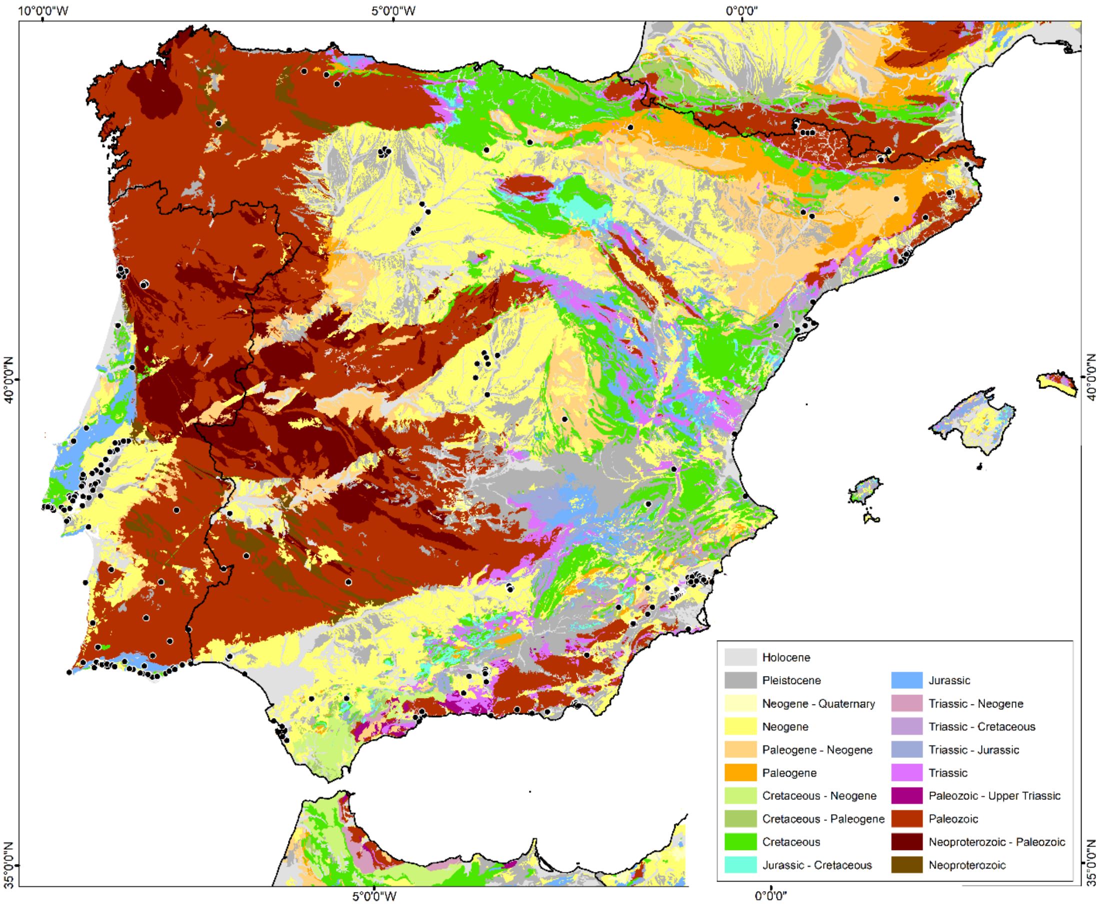

Figure 7. Geological map of Spain, originally with a scale of 1:1,000,000, showing the geological age layer (www.igme.es). Black dots show the locations of shear-wave velocity profiles. For the geological and lithological classification of the sites, higher resolution maps were used.

Table 5 shows the number of sites distributed by the main geological periods. This information has been retrieved from the 1:50,000 geological maps where available or from the 1:200,000 one.

This previous classification by age has been simplified to ensure well populated groups (see right part of Table 5); the grouping is inspired in similar classifications conducted by other authors with the same purpose (Stewart et al, 2014; Vilanova et al, 2018).

Table 5. Number of sites for each geological age.

Geological age	# of sites	%	Geological age (simplified)	# of sites	%
Precambrian	1	0.2	Paleozoic and older	69	11.9
Paleozoic	68	11.7			
Triassic	2	0.3	Mesozoic	25	4.3
Jurassic	4	0.7			
Cretaceous	19	3.3			
Paleogene	2	0.3	Tertiary	165	28.4
Neogene	163	28.1			
Pleistocene	89	15.3	Pleistocene	89	15.3
Holocene	232	40.0	Holocene	232	40.0

Although this grouping mixes eras (Paleozoic and Mesozoic) with periods (Tertiary) and epochs (Pleistocene and Holocene), it does cover the geological time scale and the final grouping on the basis of V_{S30} performance is given preference. Most of the sites in the Tertiary are located on Neogene deposits, and the only site corresponding to Precambrian age has been grouped with those on the Paleozoic.

Figure 8 shows the histograms of log V_{S30} for each age group. Except for the Mesozoic, all classes can be respectively fitted with a normal distribution, for which the average and standard deviation are reported. The standard deviations are rather large, especially for the Paleozoic and Holocene groups, suggesting that attributes other than the geological age may contribute to the observed dispersion. Hence, a separate analysis of each group is required (Section 4.3). The assumption of lognormal distribution will be considered in the statistical treatment of the sample.

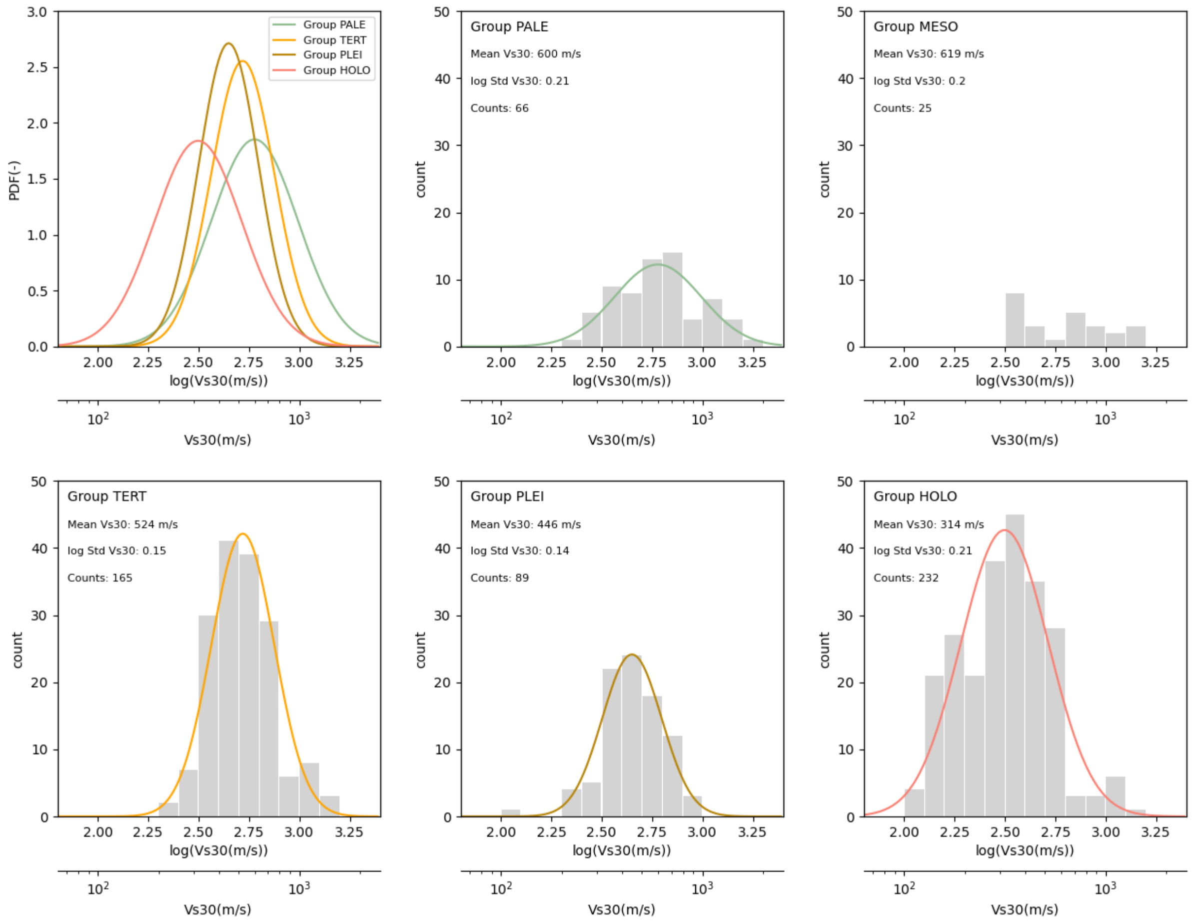

Figure 8. Histograms of log V_{S30} corresponding to each age group. In each plot, the continuous line shows the fit to a normal distribution. The corresponding mean and standard deviation of log V_{S30} are also displayed, except for the Mesozoic group.

4.2 Groups of lithologies

A total of 181 different lithological descriptions were retrieved for the geological units associated with the 580 sites from the sources of information indicated in section 3.2. For constructing the proxy-based model, initially, these were grouped (Table 6) according to a set of six lithological groups, ordered, in broad terms, in decreasing order of mechanical resistance, which are then expected to be also ordered, consistently, from highest to lowest V_{S30}:

1. Igneous and metamorphic rocks (e.g., granite, basalt, gneiss) characterized by a relatively low density of discontinuity planes.
2. Metamorphic rocks with a high density of discontinuity planes (e.g., fissility: slate, phyllite).
3. Carbonate rocks (e.g., limestone, dolomite).
4. Detritic, coarse-grained rocks (e.g., conglomerate, sandstone).
5. Detritic, fine-grained rocks (e.g., argillite, shale, marl, limolite).
6. Unconsolidated deposits.

Table 6. Number of sites ascribed to each of the six preliminary lithological groups.

	# of sites	%
1. Igneous and metamorphic rocks	40	6.9
2. Metamorphic rocks with abundant fissility	22	3.8
3. Carbonate rocks	13	2.2
4. Detritic rocks; coarse-grained	142	24.5
5. Detritic rocks; fine-grained	43	7.4
6. Unconsolidated deposits	320	55.2

Group 6 is the most populated, and ideally should be divided into coarse-grained vs. fine-grained unconsolidated deposits. However, this distinction could not be adopted because the descriptors available in the 1:50,000 geological maps enable making this distinction in less than half of the cases.

When the lithological description contains more than one lithology, the site has been assigned to the group corresponding to the first one mentioned in the description, which according to international normal practice corresponds to the predominant one. This criterion may be uncertain, especially for the detritic rocks of groups 4 (coarse-grained) and 5 (fine-grained); in numerous cases the description includes three or more lithologies mixed from these two groups (i.e not all of them are coarse-grained or fine-grained), hence the overall predominant group (be it Group 4 or 5) might not necessarily be the one indicated by the first descriptor.

The distribution of log V_{S30} values over the six groups described above is presented in Figure 9, together with the mean and standard deviation for each group. As can be seen in the probability density function (PDF) plots, Group 1 presents a significantly lower mean V_{S30} than Groups 2 and 3, as opposed to what was expected when the groups were first defined. This will be discussed in the following section. Group 2 has the same mean V_{S30} as Group 3, but a lower standard deviation.

The slight increase in V_{S30} observed from Group 4 to 5 (PDF plot in Figure 9), which is related with the distribution of Group 5 deviating from a normal one, can be explained by the mentioned frequent mixing of fine- and coarse-grained detritic rocks in the same description. In particular, there are 12 points that have been included in Group 5 with a V_{S30} higher than 1000 m/s (see the histogram in Figure 9), but for which the lithological descriptions are quite long, including several lithologies corresponding to coarse-grained detritic rocks and even a carbonate rock in one case. Both groups, 4 and 5, have similar ranges of V_{S30}. A specific study to explore the statistical dependency of these two groups was conducted, in particular an analysis of variance (ANOVA) test and an F-test. The result in terms of the p-value for the F-statistic was p = 0.27 for the ANOVA, which is above the most common significance levels; for the F-test the result was p = 0.03. According to these results these two groups are not statistically different at the commonly adopted 95% confidence level. It should be noted that Group 5 deviates from the normal distribution, hence the application of ANOVA should be interpreted with care. In any case, results are in agreement with the comparison of the means and the histograms.

With the exceptions mentioned above, the overall tendency observed in the means is consistent with the original expectations when designing the groups, namely a decreasing V_{S30} in the order the groups have been numbered.

The standard deviations found are between 0.1 and 0.2, equivalent to 0.25-0.45 when employing natural logarithms. This range is consistent with the values found by Seyhan et al. (2014) or Ahdi et al. (2017), and somewhat below that reported by Vilanova et al. (2018). It should be recalled that the database used by Vilanova et al. (2018) constitutes a subset of that compiled here.

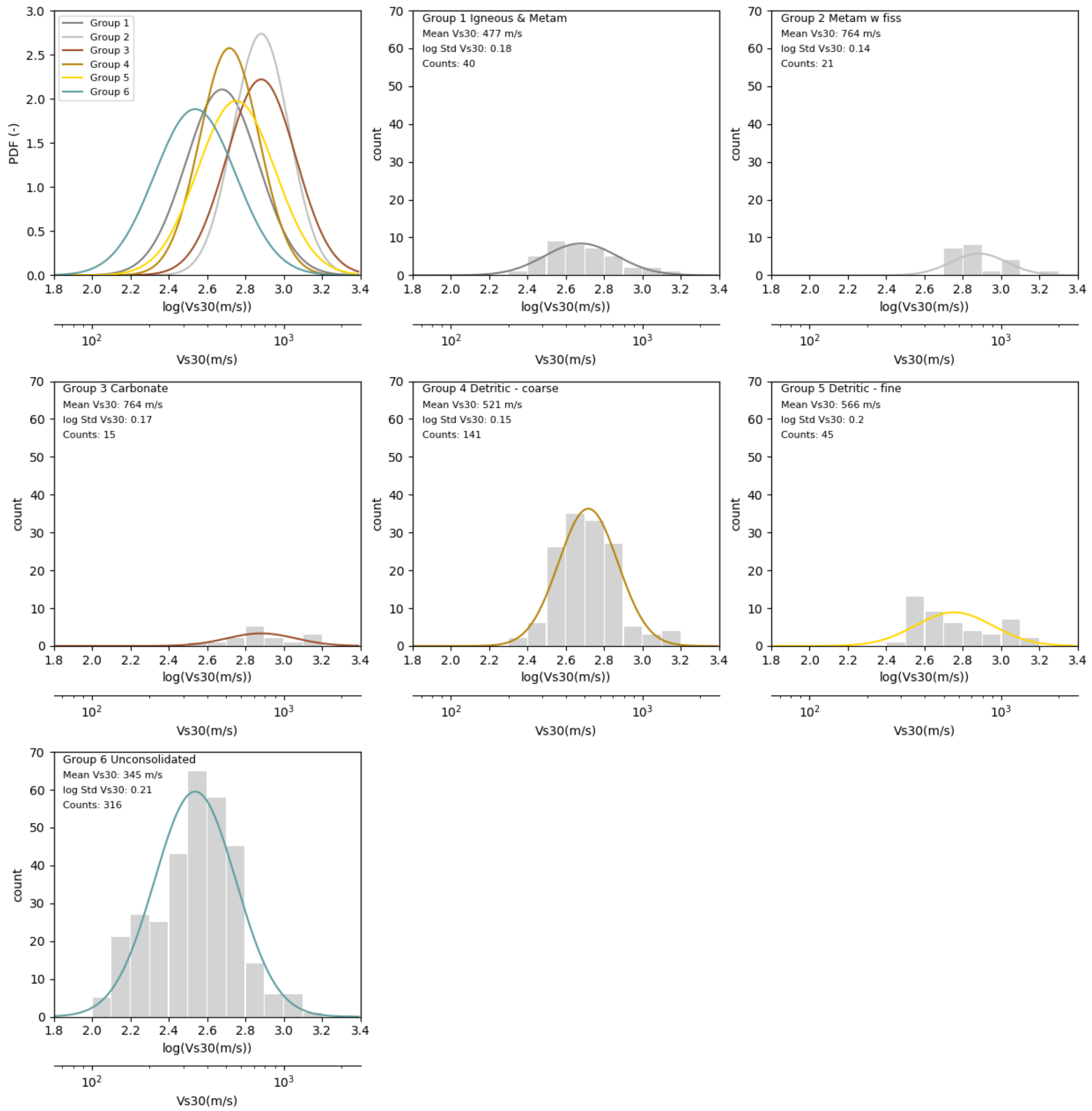

Figure 9. Distributions of log V_{S30} in the six lithological groups. The corresponding lognormal distribution for V_{S30} is also plotted over each histogram and all of them together in the top left plot.

4.3 Analysis of groups

4.3.1 Cross correlation of groups and geographic distribution

In this section, we will analyze the statistical results of the grouping made previously by age and lithology to understand the constraints of the V_{S30} database, considering the hypothesis of lognormal distribution already introduced in section 4.1. In addition, we will cross-correlate groups by age and lithology, following geological and statistical criteria, in order to obtain correlations between V_{S30} and slope and then a model with a higher predictive potential.

First, we calculated the cross distribution of sites between the grouping by age and the grouping by lithologies, shown in Table 7. There are age groups dominated by a given lithology: the Paleozoic group by igneous and metamorphic rocks (lithological Groups 1 and 2); the Tertiary group by detritic rocks, and the Pleistocene and Holocene groups by non-consolidated deposits.

If the cross distribution is looked at from the lithological groups side, in general lithological groups are concentrated in one or two geological ages. The exception is for the carbonate rocks in Group 3, which have measurements from all ages. However, this is also the least populated group.

Table 7. Cross distribution between grouping by geological age and grouping by lithology.

	Group 1 Igneous & Metamorphic	Group 2 Metamorphic w. fissility	Group 3 Carbonate	Group 4 Detritic – coarse	Group 5 Detritic – fine	Group 6 Unconsolidated deposits
I. Paleozoic	38	20	1	7	0	0
II. Mesozoic	1	0	8	16	0	0
III. Tertiary	0	0	4	118	43	0
IV. Pleistocene	1	0	1	0	2	85
V. Holocene	0	0	1	0	0	231

One of the attributes that could influence the deviation from a normal distribution is the regional influence. Figure 10a shows the distribution of the V_{S30} values among regions, color coded by geological age. Regarding the Mesozoic group, which, as previously noted (Figure 8), suggests deviations away from a normal distribution, most sites correspond to the Cenozoic belts. A number of these points correspond to a sector of the Valencia region where the geological map describes the presence of a particular Lower Cretaceous facies known as Utrillas (Rodríguez-López et al., 2009). This consists of alternating siliciclastic sediments (sands/gravels, low-cemented sandstones and clays). Therefore, lower shear-wave velocity is expected for these sites.

Sites included in the Paleozoic group are mostly located in the Iberian Massif, identified in Figure 10 as Variscan Orogen, both in the Spanish and Portuguese sides, and are rarely located in the other regions. Lithological Groups 1 and 2 have a similar situation, as observed in Figure 10b, which shows the distribution of the V_{S30} values among regions, color coded by lithology.

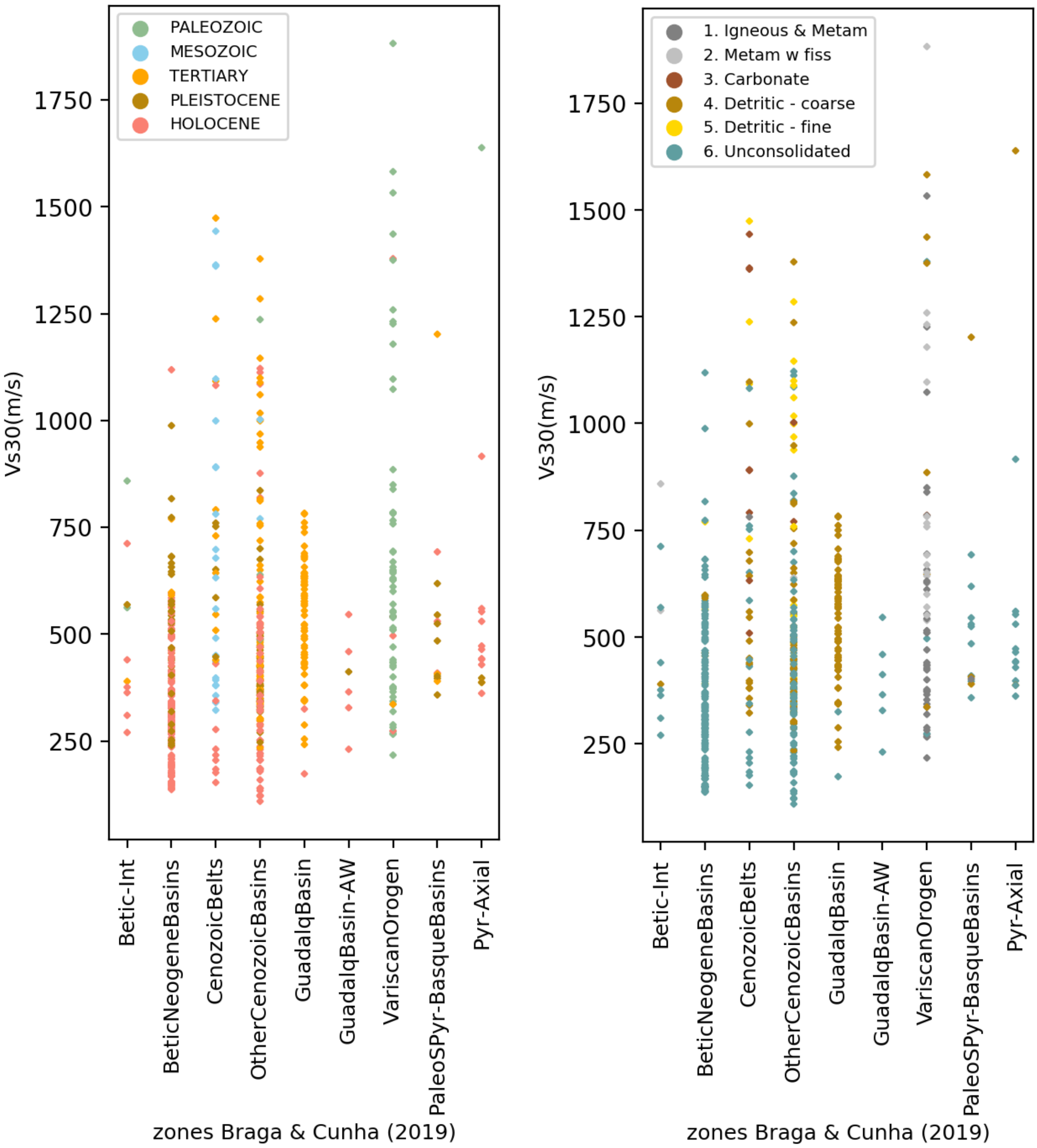

Figure 10. a) V_{S30} distribution by geologic regions of the Iberian Peninsula, shown in Figure 1 (Braga and Cunha 2019). Color code indicates the geological age. b) As in a), but with the grouping by lithology.

Figure 11a shows the histogram of V_{S30} for all the sites of the Paleozoic group (which is the same histogram as in Figure 8 for the Paleozoic). The standard deviation is high, whereas the mean value, 600 m/s, is low with respect to what would be expected. Figure 11c also shows this characteristic with the mean V_{S30} on the West part of the Iberian Peninsula (Portugal side) for Paleozoic sites as low as 546 m/s. Vilanova et al. (2018) noted that the mean for this group in their compilation (which included igneous and metamorphic rocks) was also low. Six of these sites, located by the Duero river shoreline, have a V_{S30} of less than 300 m/s despite pertaining to the Paleozoic group; these velocity values seem too low for lying on this age. These data come from the work of Santos (2011, included in the compilation by Vilanova et al., 2018), who intentionally selected the sites with the purpose of identifying the contact between extremely weathered, degraded rock and fresh rock. As a result, these estimates may be biased to low V_{S30} values and do not constitute a random representative sample. Removing these sites from the Paleozoic group leads to a mean V_{S30} of 888 m/s which is closer to the expected value for this geological age (Figure 11b).The authors of Vilanova et al. (2018) pointed out (pers. comm.) that, moreover, other sites in the Portuguese region are also located on a highly weathered soil profiles overlying intact rocks.

The Paleozoic group is mainly composed by lithological groups 1 and 2 and the same fact explained above also applies to lithological Group 1. The overall sample contains 40 sites falling in Group 1, and 35 belong to the Portuguese territory for which only the V_{S30} (and not the full profile) was available (section 2.3). The mean of all these 40 sites is 477 m/s, and the mean for the sites in the Spanish territory is 780 m/s, although the latter value derives just from a sample of 5 elements.

This situation of finding crystalline rock in the geological map which in reality corresponds to the presence of highly weathered strata, which can be quite deep, over intact rocks, yielding low V_{S30} values, is in fact a characteristic circumstance of the west part of the Iberian Peninsula. It seems appropriate to differentiate the sample between the West part of the Iberian Peninsula, where there is a predominance of weathered strata, and the rest with a relatively fresh rock. Additionally, this type of circumstances reinforces the importance of a local model.

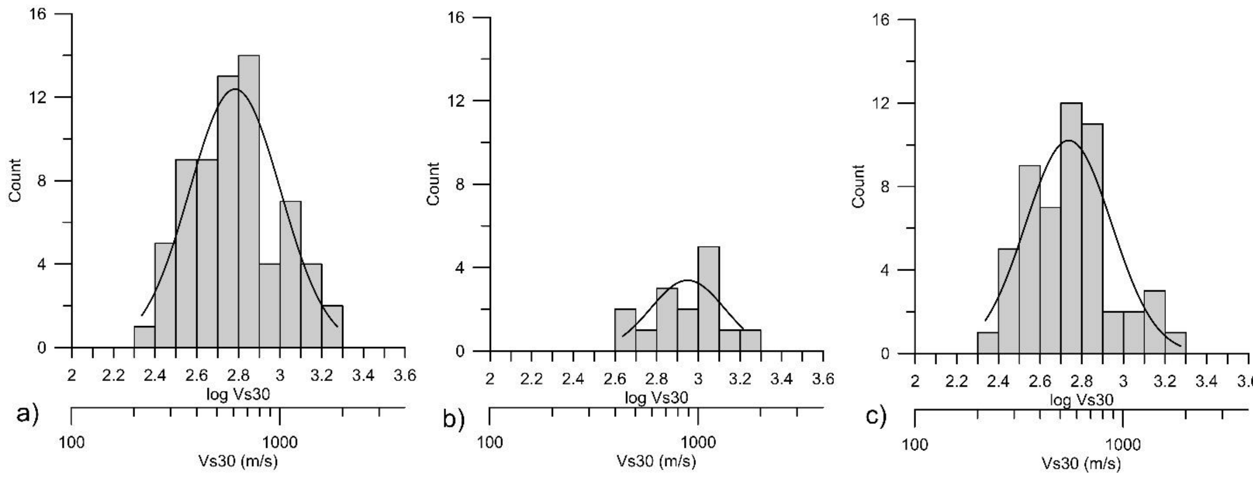

Figure 11. Histogram of the log V_{S30} corresponding to a) Paleozoic group of the whole Iberian Peninsula b) Paleozoic group of the Spanish territory where rock is relatively fresh c) Paleozoic group of the West part of the Iberian Peninsula (Portuguese area) where weathered strata predominate. The black lines correspond to the fits to normal distributions.

4.3.2 Dependency on slope

Next, we introduce the slope as another proxy that could improve the performance of the model. The dependency of V_{S30} with respect to the topographic slopes from the 200 m resolution DEM has been represented for each group of age and lithology. The mean variation of V_{S30} with the topographic slope is evaluated as:

$$\overline{\log V_{S30}} = a + b \log(s)$$

Where V_{S30} is the time averaged shear-wave velocity in m/s, s is the slope in % from the 200 m resolution DEM and a and b are regression coefficients.

Figure 12 presents least-squares linear regressions for the geological age groups. The type of lithology within each group is represented with color coded dots, in order to visualize the information provided in Table 7. The regressions are calculated based on binned means with equal weights, which is a methodology already employed by other authors in similar works (i.e., Stewart et al., 2014). Employing binned means reduces the influence of the clustered data. Other authors, like Vilanova et al. (2018) have directly applied a geographical declustering to the data. The V_{S30} proxy model obtained by Stewart et al. (2014) has been used for comparison in Figure 12 since their approach is based on geology and their studied area is also located in the South Europe region.

Similarly, Figure 13 presents analogous plots for the lithological groups. Linear regressions (calculated again with the binned means) are shown for each lithological group, while dots are color-coded according to the geological age groups.

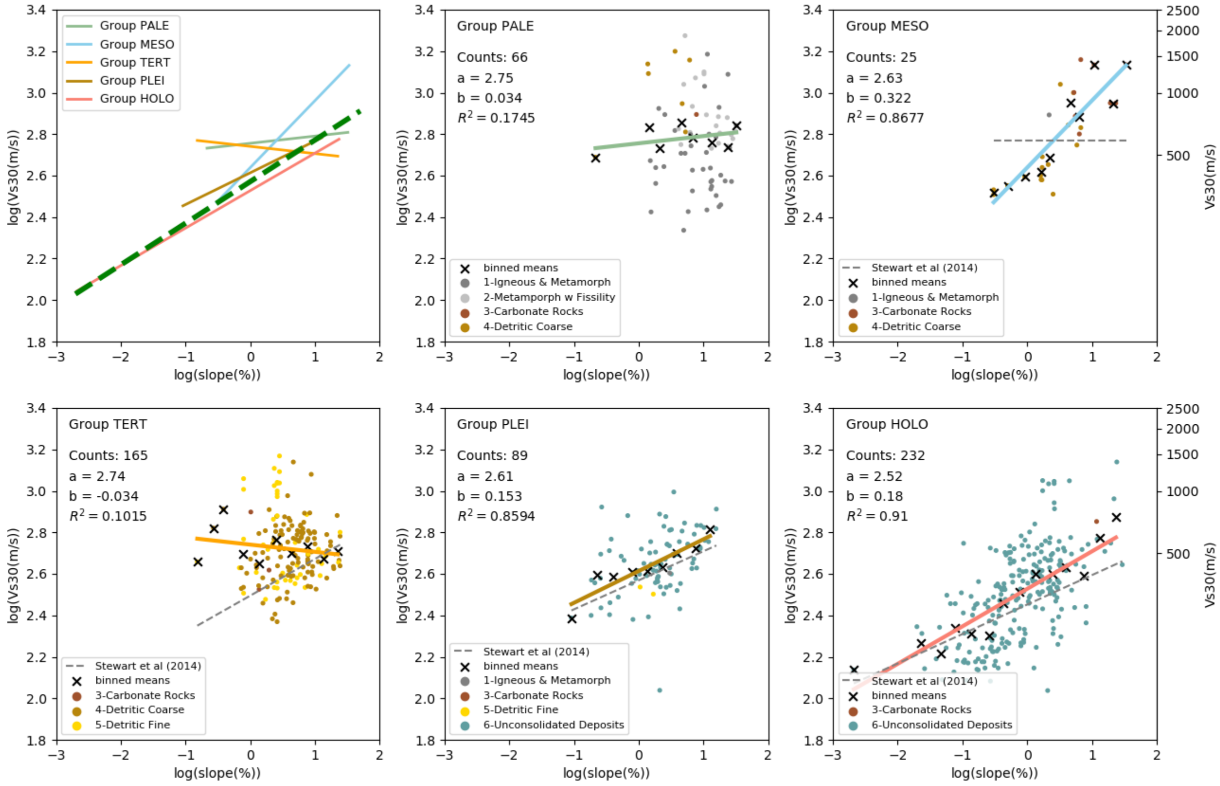

Figure 12 V_{S30} dependency on the topographic slope for the five geological groups and for the whole dataset. Dependency on lithology is presented by color-coded points. An independent fit is presented for each geological group. Results by Stewart et al. (2014) are included in grey dashed lines. The top left plot presents the five independent adjustments with binned means together with the adjustment considering the complete dataset (green dashed line) which was already presented in Figure 5. For the Tertiary the sample adjustment issues a negative slope with both fits.

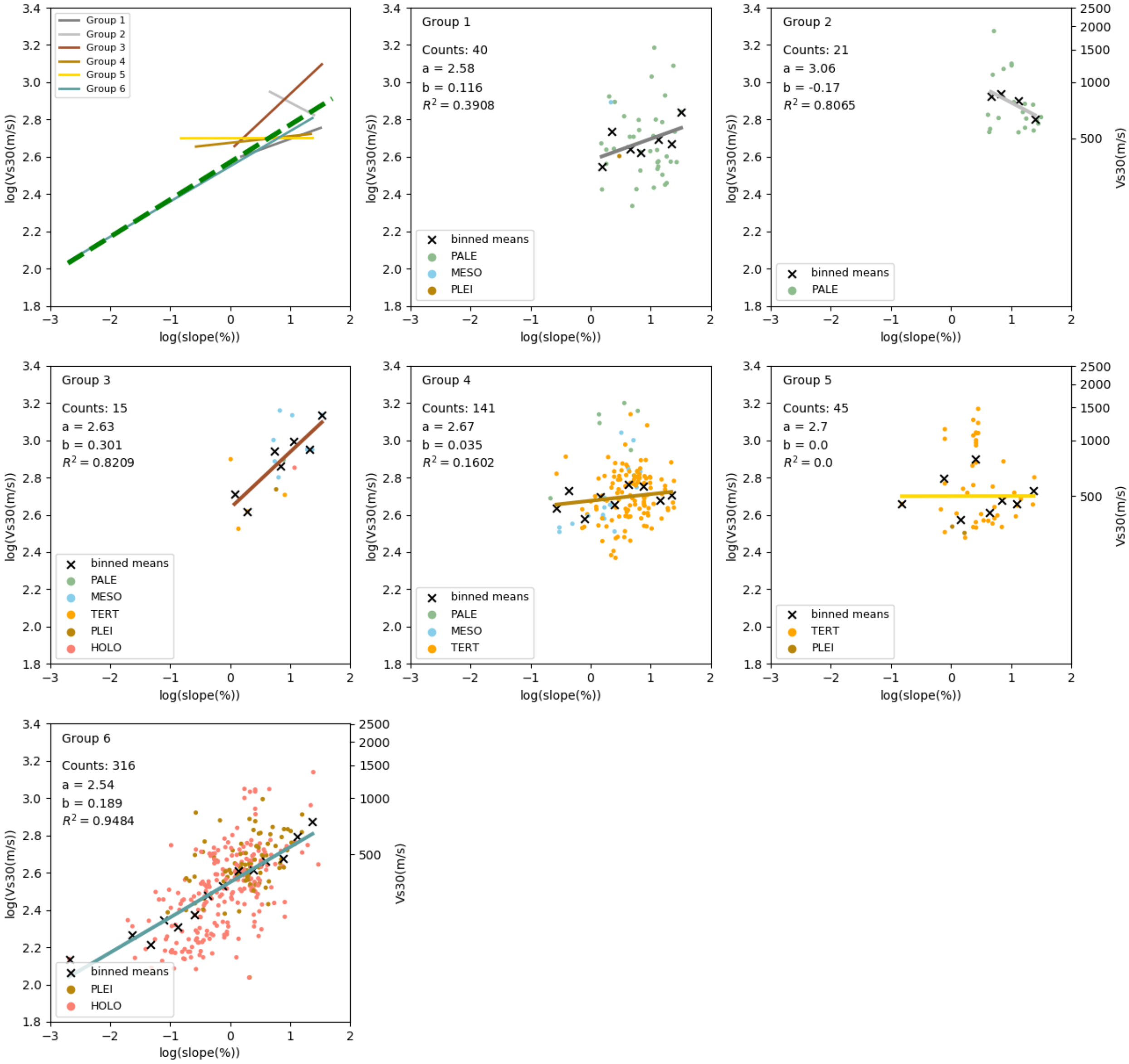

Figure 13. V_{S30} dependency on the topographic gradient for the six lithological groups and for the whole dataset. An independent fit is presented for each lithological group. Dependency on geological age is presented by color-coded points. The top left plot presents the six independent fits with binned means together with the fit considering the complete dataset (green dashed line) which was already presented in Figure 5.

In the regressions, positive slopes of the linear fitting, which could indicate a good correlation of V_{S30} vs. slope, are obtained in all cases, except for the Tertiary and the lithological Group 2. For groups having a negative slope in the regression analysis, or only a weak trend with the topographic gradient, the mean and standard deviation of the sample are the only results to be considered, and no dependence will be established with respect to the topographic gradient. A similar criterion was employed by Ahdi et al. (2017) and Stewart et al. (2014).

The fit for the Paleozoic group (Figure 12) has a low correlation coefficient. Considering the discussions presented above regarding the Paleozoic as well as the regressions in Figure 12, it seems appropriate to have the Paleozoic group represented solely by its mean value, including as additional information the different means obtained for Spain and Portugal.

A similar situation is found for Groups 1 and 2, which are the two main lithologies present in the Paleozoic. Given the negative slope obtained for Group 2 and the not so high correlation coefficient for Group 1, it seems appropriate to merge these two lithological groups into a single one that includes all igneous and metamorphic rocks, regardless the presence or not of fissility planes. Future enlargements of the database, with more points of the samples falling in these groups, could allow further refinements and a split for the Paleozoic based on lithological information.

The Mesozoic group presents a good correlation coefficient. The sample is not very large, with a total of 25 elements, being 16 coarse detritic rocks (Group 4), 8 carbonate rocks (Group 3), and one site belonging to Group 1. The variability introduced by the presence of different lithologies is considered to be well captured by the dependency on the slope, being probably in the case classified as “erosional areas” by Wills et al. (2015). The 16 points that belong to Group 4 all correspond to Utrillas facies, which presents a particular kind of lithology in the Valencia area as it has already been discussed above. This fact is also considered to be appropriately represented with the dependency on the slope. Stewart et al. (2014) used a smaller sample size for the Mesozoic, and did not find a significant trend, therefore for this age they proposed a constant value coincident with the mean (included in Figure 12 for comparison).

Group 3 (carbonate rocks) is the only lithology present in all geological ages (Table 7). The fit (Figure 13) presents a good correlation coefficient and, similarly to what it has been indicated for the Mesozoic, it is considered that the dependency on this age heterogeneity is well represented by the dependency on the gradient.

The Tertiary group is composed mainly by detritic rocks, both coarse (75%) and fine (25%) grained. The fit for the Tertiary issues a negative slope (Figure 12). For this reason, it seems appropriate to propose a constant value based on the mean.

Similarly, the trends observed for the detritic groups is not very significant (even null for the case of fine-grained ones) and with low correlation coefficients. This fact, together with the discussion in section 4.2, about the criterion followed for assigning lithologies to groups, suggests that detritic rocks should be all in the same group.

Pleistocene and Holocene are mainly composed of unconsolidated deposits; they both have a highly populated sample, especially the Holocene with 232 points. Both fits (Figure 12) present high correlation coefficients, and the agreement with the proposal by Stewart et al. (2014) for Greece is better than in other ages, being particularly good for the Pleistocene.

Group 6 (unconsolidated deposits) is solely composed by Holocene and Pleistocene ages (**Error! Reference source not found.**Table 7) and its fit (Figure 13) presents a high correlation coefficient. Inversely to what has happened in the previous cases, in this case, it is the geological age that can be employed for refining the lithological group, differentiating between Holocene or Pleistocene. Each fit for the Holocene and Pleistocene falls outside the 95%

confidence intervals of the other one, respectively (Figure 14), except for a small range of slopes above 10%. Noting this statistically significant difference, it has been decided to keep both age groups as independent classifications.

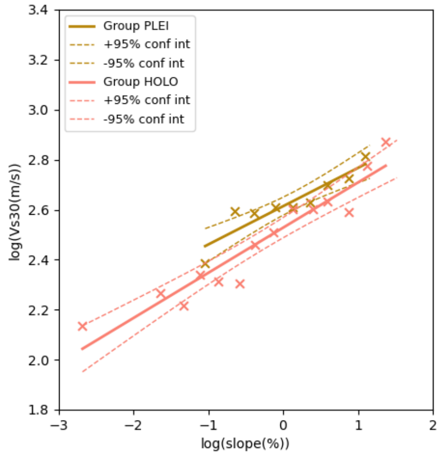

Figure 14. Confidence intervals for the fits of the Pleistocene and Holocene groups.

5 Proposed model and proxy performance

Considering the analysis presented in section 4 after the information provided by cross correlation of groups, geographic distribution, histograms, and correlations of V_{S30} with slope, a proxy model has been proposed.

The final model gives the option of relying on the geological age (introducing a dependency on the slope for certain ages); or alternatively a proxy model based on four lithological groups (named L1, L2… for differentiating them from the six initial ones), including the dependency on the slope when possible and refined with the geological age for one of the groups.

After the grouping argued in section 4.3.2, the final lithological groups, compared to the initial ones in Table 6, are:

- Group L1: Igneous and metamorphic rocks, regardless of fissility (groups 1 and 2).
- Group L2: Carbonate rocks (group 3).
- Group L3: Detritic rocks (coarse- or fine-grained, groups 4 and 5).
- Group L4: Unconsolidated deposits (group 6).

Figure 15 and Table 8 present the model based on geological ages. For the Mesozoic, Pleistocene and Holocene, the model introduces a dependency on the slope. For the Paleozoic and Tertiary, the model provides just a constant V_{S30}, independent of the topographic gradient; additionally for the Paleozoic it has been considered appropriate to introduce a regional dependency distinguishing between the western (Portuguese) part of the Iberian Peninsula where weathered strata predominates and the rest of the territory where the rock is usually relatively fresh.

Table 8 presents values *a* and *b* for each group together with their standard errors in the cases there is a dependency on the topographic slope, and for all cases the global standard deviation. For the groups with a model depending on the topographic slope, this standard deviation is the one resulting from the distribution of residuals.

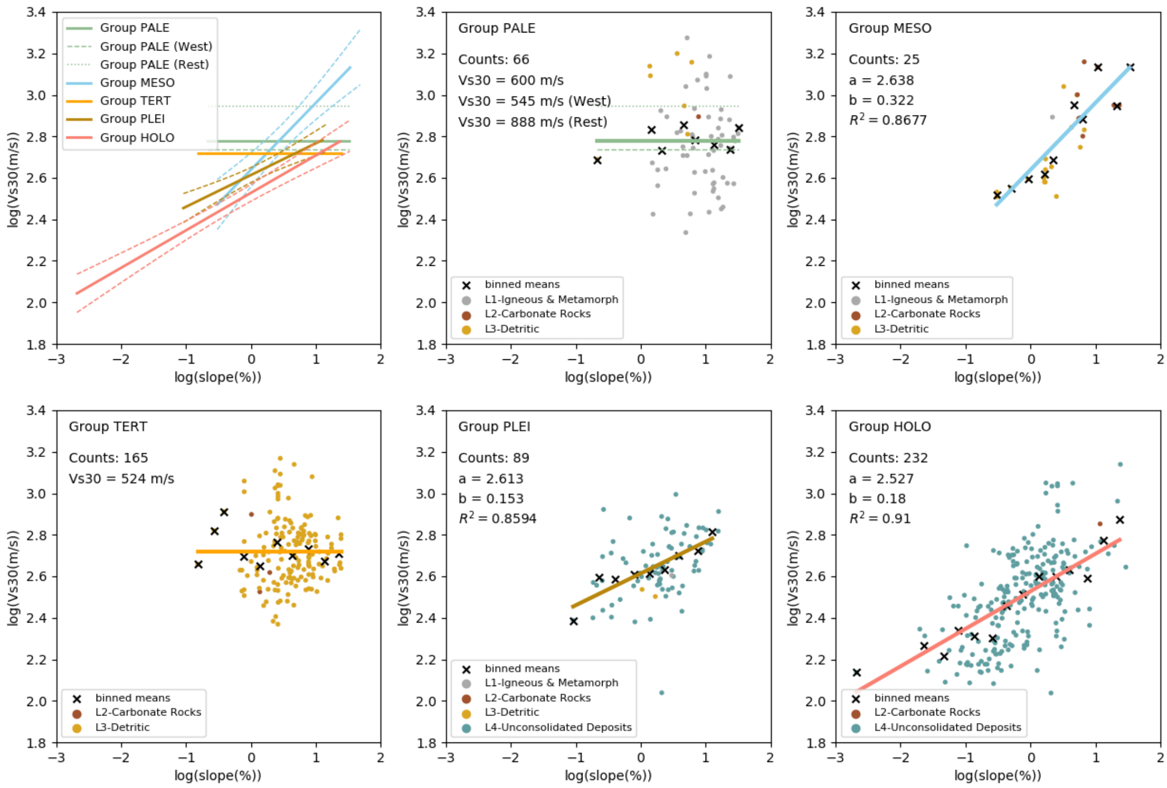

Figure 15 Final V_{S30} proxy model, based on geological age and topographic slope. For the Paleozoic and Tertiary ages only a mean V_{S30} is specified, regardless of the topographic slope. The dependency on the final lithologic groups is presented, just for information, by color-coded points. The top left plot presents the five fits together.

Table 8. Proposed model for the definition of the mean V_{S30} as a function of the geological age and the topographic slope from the 200 m resolution DEM, regardless of lithology.

$\overline{\log V_{S30}} = a + b \log(s)$				
		a	b	Standard deviation
Paleozoic	All	2.783	0.0	0.219
	Weathered	2.737	0.0	0.176
	Relatively fresh	2.948	0.0	0.207
Mesozoic		2.638 +/- 0.0445	0.322	0.125
Tertiary		2.719	0.0	0.150
Pleistocene		2.613 +/- 0.0234	0.153 +/- 0.0159	0.134
Holocene		2.527 +/- 0.0164	0.180 +/- 0.0189	0.174

Figure 16 and Table 9 present the model based on lithologies. It has four groups, and includes a dependency on slope for carbonate rocks (L2) and unconsolidated deposits (L4) and provides just a constant V_{S30} for igneous and metamorphic rocks (L1) and for detritic rocks (L3). The information in Table 9 is analogous to the one described for Table 8: values *a* and *b* are indicated together with their standard errors in the cases there is a dependency on the topographic slope, and the global standard deviation for all cases.

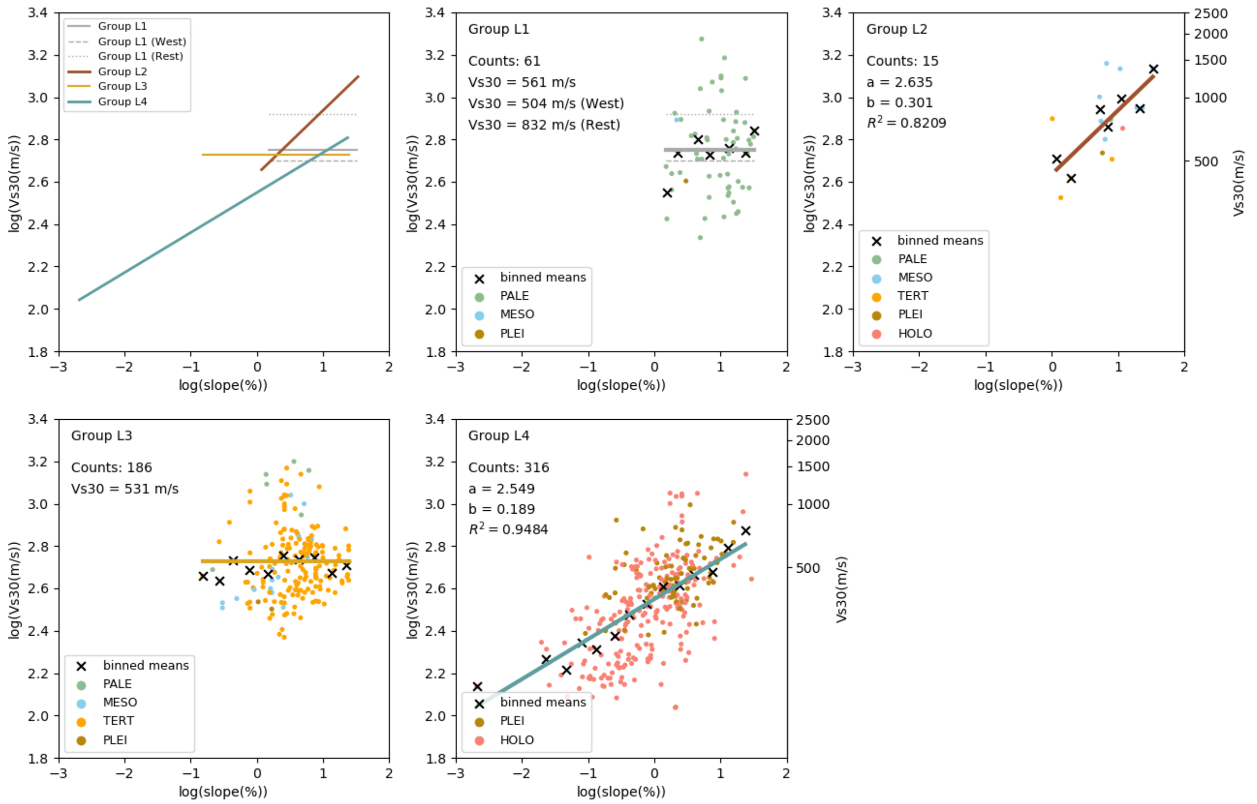

Figure 16. V_{S30} proxy model based on lithology, topographic slope and geological age. Lithological groups are reduced to four (L1 to L4, see text). V_{S30} for L1 and L3 are only based on the mean. Cross dependency on geological age is shown by color-coded dots, and is considered to provide different fits for L4 (seeTable 9). The top left plot presents the four fits together.

Figure 17 presents the residuals for the model presented in Figure 15 and Table 8, and Figure 18 shows the residuals for the model presented in Figure 16 and Table 9. For the cases with no dependency on the slope, a histogram is presented whereas for those cases that include a dependency on the slope this dependency is also shown for the residuals. The scatter plots of the residuals, in general, show a fairly random pattern around the x axis and the histograms approach a reasonably normal distribution.

Table 9. Proposed model for the definition of the mean V_{S30} as a function of lithology, geologic age and topographic slope from the 200 m resolution DEM.

$\overline{\log V_{S30}} = a + b\log(s)$				
		a	b	Standard deviation
L1. Igneous and metamorphic rocks	All	2.749	0.0	0.200
	Weathered	2.702	0.0	0.177
	Relatively fresh	2.920	0.0	0.187
L2. Carbonate rocks.		2.635 +/- 0.0630	0.301 +/- 0.0609	0.146
L3. Detritic rocks		2.725	0.0	0.219
L4. Unconsolidated deposits	All	2.549 +/- 0.0127	0.189 +/- 0.0146	0.169
	Pleistocene	2.613 +/- 0.0234	0.153 +/- 0.0159	0.134
	Holocene	2.527 +/- 0.0164	0.180 +/- 0.0189	0.174

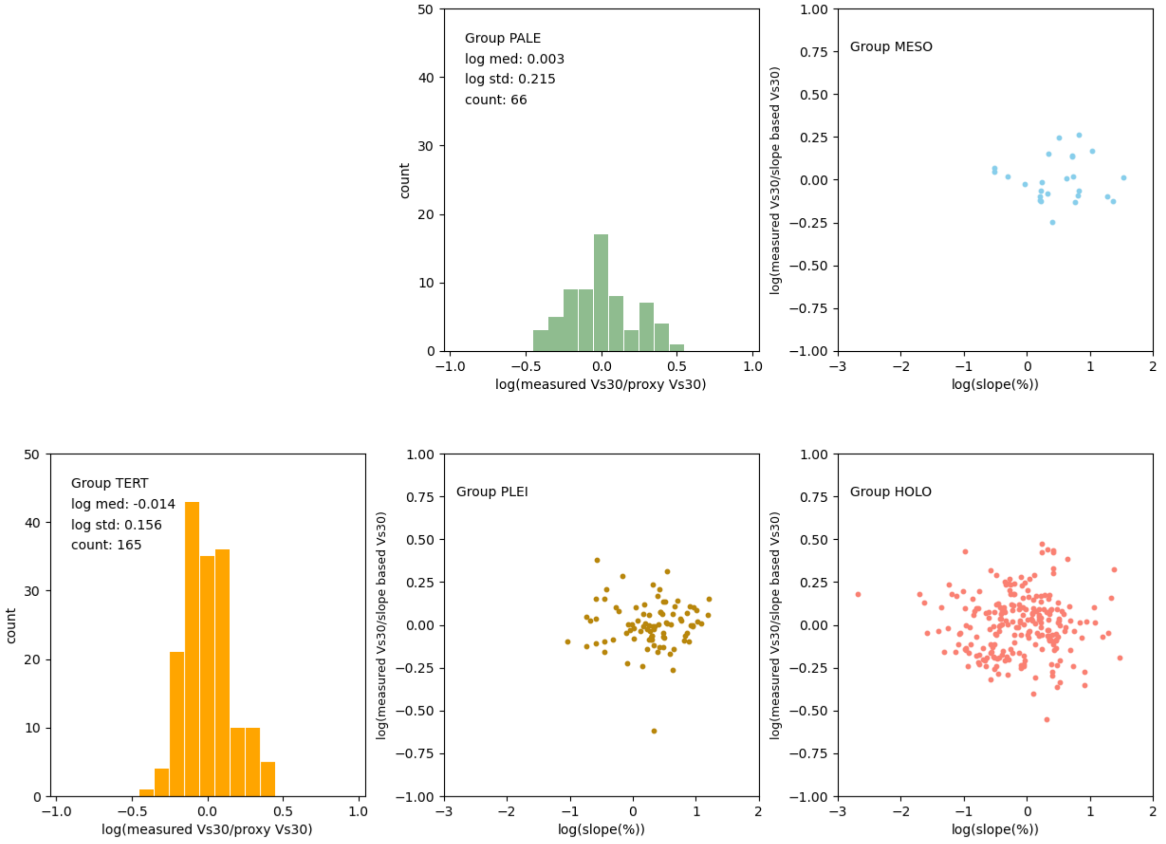

Figure 17. Residuals for the V_{S30} proxy model based on geological age groups (Figure 15 and Table 8).

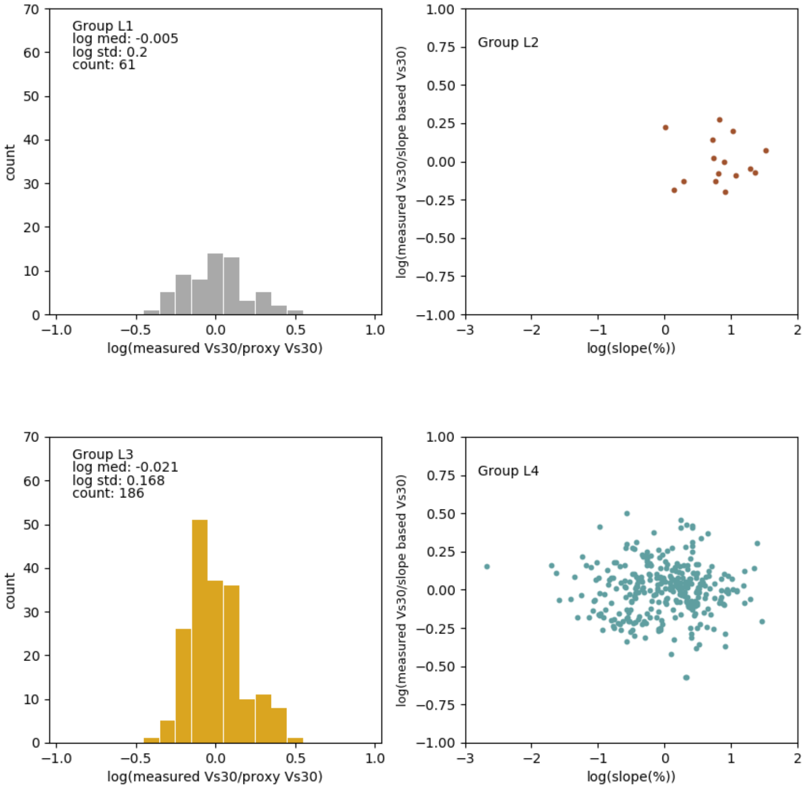

Figure 18. Residuals for the V_{S30} proxy model based on lithological groups (Figure 16 and Table 9).

6 Summary and Conclusions

A proxy-based model has been developed for estimating the average V_{S30} as a function of the lithology, the geological age and the topographic slope.

A V_{S30} database has been created with information distributed across the entire Iberian Peninsula. The database previously compiled for a similar work in Portugal has been incorporated, representing nearly 30% of the total number of sites. The proportion of Portuguese sites in the database is about double of the proportion of the area of Portugal in the Iberian Peninsula.

For the Portuguese sites, as well as for some of the Spanish ones, only the V_{S30} is available, while for the rest of the sites an in-situ measured V_S profile exists and is employed for calculating the V_{S30}. Consistency has been ensured in the way the V_{S30} has been calculated from the profiles provided by the different contributors to the database.

The available V_{S30} data showed a clear correlation with the topographic gradient, although an initial comparison with the widely known gradient-based model by Wald & Allen (2007) was not fully satisfactory, so it was concluded that developing a particular model for the Iberian Peninsula was justified.

The topographic slope, geological age and lithology have been chosen as descriptors for developing the Iberian model. The topography is provided by DEMs of the IGN with different resolutions. The lithology/geology comes from geological maps published by three different geological surveys; their scale was 1:50,000, except for some sites in Portugal, where it was 1:200,000.

After a preliminary study, the 200 m resolution topography and the derived digital slope model was chosen as the most suitable one for conducting the present study.

The V_{S30} proxy model proposed gives the option of relying on the geological age or lithology, depending on the convenience of the end user. A cross study between age and lithological groups shows that in most cases each geological age has one or two predominant lithologies. The final model includes six age groups and four lithological groups. Initially we attempted to divide certain groups of geological age considering their predominant lithologies, but this has not been finally possible; in some cases due to the too small sample size, and in the particular case of the Tertiary group (mainly composed of detritic rocks), because it was not possible to determine from the maps the actual fraction of fine- and coarse-grained rocks at the sites. However, for the lithological group of unconsolidated deposits, represented by an extensive sample of V_{S30} sites, the distinction between the two predominant ages (Pleistocene and Holocene) has been implemented in the model.

A dependency on the slope is introduced for the groups where the data evidences a good correlation between the V_{S30} and the topographic gradient, being the case for four age groups and two lithological groups. This dependency is supported by geological or lithological criteria, linked to erosional areas (such as the ones included in Mesozoic groups) or to depositional areas (like it is observed for Holocene and Pleistocene groups).

Future hazard analysis and risk assessments in the Iberian Peninsula can benefit from this proxy-based model. In this type of analyses knowing the site V_{S30} is essential, either for developing a site-specific ground motion model or for a precise use of an applicable one. With a proxy-based model like the one presented here, this estimation can be easily obtained with information which is public and regionally available.

Finally, the V_{S30} database is in continuous development with additions of new sites; this will allow a revision of the model presented in this study, that overcomes the limitations discussed above. This increase of the V_{S30} information is of special importance in an area like the Iberian Peninsula where there is a high geological and lithological variability.

7 Acknowledgements

The authors wish to thank IGN for the assistance with the DEMs, IGME for its support when accessing the Geode WMS, ICGC for providing the existing measurements in their database, Manuel Navarro for providing the unpublished Mini-Array measurements and Raquel Martín-Banda for her help in the presentation of the geological maps. We are also grateful to Julian Bommer, Luis Cabañas, Adrián Rodríguez-Marek, and Bob Youngs for their useful comments and observations. Á.G. acknowledges funding from the Spanish Ministry of Science and Innovation (grants “Juan de la Cierva” FJCI-2016-29307 and “José Castillejo” CAS19/00298). Finally, we want to thank Jonathan Stewart, and an anonymous reviewer for constructive comments and observations on the manuscript that helped us to improve the presentation of our work.

8 References

Ahdi, S.K., J.P. Stewart, T.D. Ancheta, D.Y. Kwak & D. Mitra (2017). Development of V_S profile database and proxy-based models for V_{S30} prediction in the Pacific Northwest region of North America. *Bulletin of the Seismological Society of America* **107**(4), 1781-1801.

Alguacil, G., F. Vidal, M. Navarro, A. García-Jerez & J. Pérez-Muelas (2014). Characterization of earthquake shaking severity in the town of Lorca during the May 11, 2011 event. *Bulletin of Earthquake Engineering* **12**, 1889-1908.

Allen, T.I. & D.J. Wald (2007). *Topographic slope as a proxy for seismic site-conditions (V_{S30}) and amplification around the globe.* USGS Open-File Report 2007-1357, US Geological Survey, Reston, Virginia.

Allen, T.I. & D.J. Wald (2009). On the use of high-resolution topographic data as a proxy for seismic site conditions (V_{S30}). *Bulletin of the Seismological Society of America* **99**(2A), 935-943.

Boore, D. M. (2004). Estimating V_{S30} (or NEHRP site classes) from Shallow velocity models (depth < 30 m). *Bulletin of the Seismological Society of America*, **98**(2), 591–597.

Boore, D. M., E. M. Thompson & H. Cadet (2011). Regional correlations of V_{S30} and velocities averaged over depths less than and greater than 30 meters. *Bulletin of the Seismological Society of America*, **101**(6), 3046–3059.

Braga, J.C. & Cunha, P.P. (2019) *Introduction in The Geology of Iberia: A Geodynamic Approach*. C. Quesada and J. T. Oliveira (eds.), Regional Geology Reviews. Springer, Cham.

Building Seismic Safety Council (1998) *1997 Edition NEHRP Recommended provisions for the development of seismic regulations for new buildings and other structures*. FEMA 302. Building Seismic Safety Council, Washington.

Candela-Medel, R., Oda, Y., Navarro, M., Enomoto, T. and García-Jerez, A. (2018). V_{S30} structure of Murcia city (southeast of Spain) from mini-array observations and HSR measurements. *Near Surface Geoscience Conference & Exhibition 2018*, 9-12 September 2018, Porto, Portugal.

Carvalho, J., R. Dias, C. Pinto, T. Cunha, J. Leote, S. Vilanova, J. Narciso, J. Borges, R. Ghose (2013). Earthquake mitigation in the Lisbon and Lower Tagus Valley area, Portugal, *Thirteenth International Congress of the Brazilian Geophysical Society*, 26-29 September, Rio de Janeiro (Brazil), 4201.

Clavero, D. (2014). *Microzonación sísmica de la ciudad de Málaga: Aproximación teórica y empírica*, PhD Thesis, Universidad de Granada, Spain.

Cho, I., Senna, S., & Fujiwara, H. (2013). Miniature array analysis of microtremors. *Geophysics*, *78*(1), KS13-KS23.

Dias, R., J. Carvalho, C. Pinto, J. Leote, S.P. Vilanova, J. Narciso & R. Ghose (2014). Site effect studies in the Lower Tagus region, *Comunicações Geológicas* 101, Especial II, 893-896.

Feriche, M. (2013). *Elaboración de escenarios de daños sísmicos en la ciudad de Granada*. PhD Thesis, Universidad de Granada, Spain.

Foti S, Parolai S, Bergamo P, Di Giulio G, Maraschini M, Milana G, Picozzi M, Puglia R (2011). Surface wave surveys for seismic site characterization of accelerometric stations in ITACA. *Bulletin of Earthquake Engineering* 9:1797–1820. doi:10.1007/s10518-011-9306-y.

García-Fernández, M. & M.J. Jiménez (2012) Site characterization in the Vega Baja, SE Spain, using ambient-noise H/V analysis. *Bulletin of Earthquake Engineering* **10**, 1163-1191.

Heath, D.C., Wald, D.J., Worden, C.B., Thompson, E.M. and Smockzyk, G.M. (2020). A Global Hybrid VS30 map with a topographic slope-based default and regional map insets. *Earthquake Spectra* **36**(3), 1570-1584.

Hollender, F., Cornou, C., Dechamp, A., Oghalaei, K., Renalier, F., Maufroy, E., Burnouf, C., Thomassin, S., Wathelet, M., Bard, P., Boutin, V., Desbordes, C., Douste-Bacqué, I., Foundotos, L., Guyonnet-Benaize, C., Perron, V., Régnier, J., Roullé, A., Langlais, M. and Sicilia, D. (2018) Characterization of site conditions (soil class, V_{S30}, velocity profiles) for 33 stations from the French permanent accelerometric network (RAP) using surface-wave methods. *Bulletin of Earthquake Engineering* **16**:6, 2337-2365.

ICGC (2020) Geoíndex – Cartografia Geologica http://siurana.icgc.cat/arcgis/services/geologic/icgc_mg50m/MapServer/WMSServer?, Last Access, September.

IGME (1994). *Mapa geológico de España a escala 1:1,000,000*. Instituto Geológico y Minero de España, Madrid.

IGME (2015). *Mapa geológico de la Península Ibérica, Baleares y Canarias a escala 1:1,000,000, edición 2015*. Instituto Geológico y Minero de España, Madrid.

IGME (2020) GEODE - Cartografia geológica digital continua a escala 1:50.000, http://mapas.igme.es/gis/services/Cartografia_Geologica/IGME_Geode_50/MapServer/WMSServer?request=getcapabilities&service=wms&version=1.3.0 , Last Access, September.

Michel, C., Edwards, B., Poggi, V., Burjanek, J., Roten, D., Cauzzi, C. and Fah, D. (2014). Assessment of site effects in alpine regions through systematic site characterization of seismic stations. *Bulletin of the Seismological Society of America* **104**:2809–2826. doi:10.1785/0120140097.

Navarro, M., P. Martínez-Pagán, A. García-Jerez, J. Pérez-Cuevas, L. González-García, F. Vidal & T. Enomoto (2014a). Comparative study of the SPAC and MASW methods to obtain the V_{S30} structure for the seismic site effect evaluation in Almeria town, SE Spain. *2nd European Conference on Earthquake Engineering and Seismology*, Istambul, August 25-29.

Navarro, M., García-Jerez, A., Alcalá, F. J., Vidal, F., & Enomoto, T. (2014b). Local site effect microzonation of Lorca town (SE Spain). *Bulletin of Earthquake Engineering*, **12**(5), 1933-1959.

Núñez, A., J. Rueda & J. Mezcua (2012). A site amplification factor map of the Iberian Peninsula and the Balearic Islands. *Natural Hazards* **65**, 461-476.

Olona, J. (2014) *Integración de metodologías geofísicas para la caracterización geológico-geotécnica del terreno,* PhD Thesis, Universidad de Oviedo, Spain.

Pérez, I. (2011) *Caracterización geotécnica de los suelos de Madrid mediante la técnica REMI. Aplicaciones en la ingeniería civil*. PhD Thesis. Universidad Complutense de Madrid, Spain.

Rodríguez López, J. P., Meléndez Hevia, N., Soria, A. R., & De Boer, P. L. (2009). Reinterpretación estratigráfica y sedimentológica de las formaciones Escucha y Utrillas de la Cordillera Ibérica. *Revista de la Sociedad Geológica de España*, **22**(3-4), 163-219.

Rueda, J., Mezcua, J., García-Blanco, R.M., Núñez, A. and Fernández de Villalta, M. (2015). Seismic scenario including site-effect determination in Torreperogil and Sabiote, Jaén (Spain), after the 2013 earthquake sequence. *Natural Hazards* **79**, 675-697.

Sá, L.F., Morales-Esteban, A. and Neyra, P.D. (2020). Regional correlations for estimating seismic amplification. Implications for loss assessment in SW Iberia. *Soil Dynamics and Earthquake Engineering* **130**, p.105993.

Santos, P. (2011). *Cartografia de espessura de alteração numa zona piloto da margem do Douro através de métodos sismicos: Implicações para o ordenamento do território*, Ph.D. Thesis, Faculdade de Ciências da Universidade do Porto, Portugal.

Savvaidis, A., Makra, K., Klimis, N., Zargli, E., Kiratzi, A. and Theodoulidis, N. (2018). Comparison of V_{S30} using measured, assigned and proxy values in three cities of Northern Greece. Engineering Geology **238**, 63-78.

Seyhan, E., J.P. Stewart, T.D. Ancheta, R.B. Darragh & R.W. Graves (2014). NGA-West2 site database. *Earthquake Spectra* **30**(3), 1007-1024.

Stewart, J. P., Klimis, N. Savvaidis, A.,Theodoulidis, N., Zargli, E., Athanasopoulos, G., Pelekis, P., Mylonakis, G. and Margaris, B. (2014). Compilation of a local V_S profile database and its application for inference of V_{S30} from geologic- and terrain-based proxies, *Bulletin of the Seismological Society of America* **104**, 2827–2841.

Tinsley, J. C., & Fumal, T. E. (1985). Mapping Quaternary sedimentary deposits for areal variations in shaking response. In: *Evaluating Earthquake Hazards in the Los Angeles Region - An Earth Science Perspective*, USGS Professional Paper 1360, 101-126.

Vilanova, V., J. Narciso, J.P. Carvalho, I. Lopes, M. Quinta-Ferreira, C.C. Pinto, R. Moura, J. Borges & E.S. Nemser (2018). Developing a geologically based V_{S30} site-condition model for Portugal: Methodology and assessment of the performance of proxies. *Bulletin of the Seismological Society of America* **108**(1), 322-337.

Wald, D.J & T.I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America* **97**(5), 1379-1395.

Wills, C.J. & Clahan, K.B. (2006). Developing a map of geologically defined site-condition categories for California. *Bulletin of the Seismological Society of America* **96**(4A), pp.1483-1501.

Wills, C. J., Gutierrez, C. I., Perez, F. G., & Branum, D. M. (2015). A Next Generation VS30 Map for California Based on Geology and Topography. *Bulletin of the Seismological Society of America* **105**(6), 3083-3091.