https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Introducing inferred geomorphological sediment thickness as a new site proxy to predict ground-shaking amplification at regional scale. Application to Europe and Eastern Turkey

Karina Loviknes[1,2], Fabrice Cotton[1,2], Graeme Weatherill[1]

[1]Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Potsdam, 14467, Germany
[2]University of Potsdam, Potsdam, Potsdam, 14469, Germany

Correspondence to: Karina Loviknes (karinalo@gfz-potsdam.de)

Abstract. To test whether a globally inferred sediment thickness value from geomorphological studies can be used as a proxy to predict earthquake site amplification, we derive site amplification models from the relation between empirical amplification for sites in Europe and Turkey, and the geomorphological sediment thickness. The new site amplification predictions are then compared to predictions from site amplification models derived using the traditional site proxies, V_{S30} inferred from slope, slope itself, as well as geological era and slope combined. The ability of each proxy to capture the site amplification is evaluated based on the reduction in site-to-site variability caused by each proxy. The results show that the highest reduction is caused by geological era and slope combined, while the geomorphological sediment thickness show a slightly larger or equal reduction in site-to-site variability as inferred V_{S30} and slope. We therefore argue that including geology and geomorphology in site amplification modelling on regional scale can give an important added value and that globally or regionally inferred models for soil and sediment thickness from fields beyond engineering seismology can have a great potential in regional seismic hazard and risk assessments. Furthermore, the differences between the site amplification maps derived from different proxies capture the epistemic uncertainty of site amplification modelling. While, albeit the different proxies predict similar features on a large scale, local differences can be large. This shows that using only one proxy when predicting site amplification does not capture the full epistemic uncertainty, which is demonstrated by looking into detail on the site amplification maps predicted for Eastern Turkey and Syria, where the devastating Kahramanmaras Earthquake Sequence occurred in February 2023.

1 Introduction

Local geological features can have a strong impact on earthquake ground shaking. Especially at sites with mainly loose sediments, which have been observed to amplify the recorded ground motion. Knowing the soil and sediment composition of a site is therefore necessary for computing the possible earthquake site amplification for seismic hazard and risk assessments. For a single site and site-specific analysis, several site parameters for characterizing shallow site conditions (e.g., fundamental frequency f0, shear wave velocity profile, horizontal-to-vertical ratio HVSR, depth to bedrock etc.) can be obtained from seismic and geotechnical investigations and used to predict local site amplification (e.g., Bergamo et al., 2021; Cultrera et al.,

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

2021; Trifunac, 2016; Derras et al, 2017). For larger areas and regional site-amplification analysis, however, the site conditions must be derived from empirical relations between relevant proxies available through regional or global maps (e.g., Bergamo et al., 2022; Thompson et al., 2010). Currently, the common practice for characterizing site amplification in seismic hazard and risk assessment is using the average shear wave velocity of the upper 30 meters of the soil column (V_{S30}). For a single site the velocity profile and V_{S30} can be measured directly, but for larger areas and regions, however, V_{S30} must be inferred from other parameters. A much-used method to calculate V_{S30} is using slope from digital elevation models (DEMs), following Wald and Allen (2007). This method is based on the hypothesis that steep (high) slopes generally have less sediments and therefore higher shear-wave velocity (V_S), while flat (low) slopes are more likely to be basins filled with sediments and thus with lower V_S. Wald and Allen (2007) used measured V_{S30} to derive a relation between V_{S30} and slope for active and stable tectonic regions separately and provided a global map of predicted V_{S30} values. However, inferring V_{S30} based on slope has several limitations. As already stated by Wald and Allen (2007), the assumption of correlation between V_{S30} and slope breaks down for continental glaciated terrains and nominally flat volcanic plateaus. In addition, Lemoine et al. (2012) have shown that other geological conditions, in particular narrow sedimentary basins and small topographic heterogeneity, have a poor correlation with the V_{S30} model based on slope.

Since the Wald and Allen (2007) model, several V_{S30} maps based on new methods and other geological proxies in addition to slope has been made, both on local and national level (e.g., Thompson et al., 2014; Vilanova et al., 2018; Foster et al., 2019; Mori et al., 2020; Li et al., 2022). However, as also argued by Weatherill et al. (2020, 2023), the main purpose of V_{S30} is as a proxy to predict site amplification, and when inferring V_{S30} from other parameters, it thus becomes a proxy-of-a-proxy. In fact, site amplification predicted by V_{S30} based on slope show little improvement to the site-amplification models based directly on slope (Weatherill et al., 2020). Furthermore, it is important to keep in mind that inferred V_{S30} should not be used interchangeably with measured V_{S30} values without properly accounting for the additional uncertainty related to the V_{S30} calculations (Lemoine et al., 2012; Thompson and Wald, 2016; Weatherill et al., 2023). The variability of ground-motion predictions can have great impact on the resulting probabilistic seismic hazard and risk assessments and using inferred site proxies in place of measured site parameters results in an increase in uncertainty. To account for this increase in uncertainty, Weatherill et al. (2023) derived separate site-amplification models for measured and inferred proxies and compared their impact on the final hazard and risk calculation. It was found that, although the median amplification predicted using inferred V_{S30} were notably lower than the median predicted amplification using measured V_{S30}, the resulting seismic hazard and risk curves from the different approaches were within the same range. This emphasizes how seismic hazard is not only controlled by the median amplification, but also by the uncertainty and that the increase in uncertainty related to inferred proxies compensates for the change in predicted median amplification in a probabilistic hazard and risk context.

In this study, we follow the approach of Weatherill et al. (2020, 2023), to test the ability of new site proxies to predict site amplification. We skip the step of deriving a site-proxy ourselves and look beyond the field of engineering seismology for already available large-scale models of soil and sediment conditions that would allow for inference of soil amplification across a wide region. One of such models is the Pelletier et al. (2016) geomorphological model for sedimentary thickness. As the

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

thickness of soil and sediments down to bedrock is an important factor for modelling amplification of earthquake ground shaking, the thickness of porous weathered material above unweathered bedrock is necessary for land surface modelling of, for example, the water and carbon cycle (Pelletier et al., 2016). Pelletier et al. (2016) therefore developed a global data set of soil, intact regolith, and sedimentary deposit thicknesses intended as input for hydrology and ecosystem models. The model is based on a combination of data including slope, lithology and stratigraphy, and water table depth, all of which correlate with geotechnical soil conditions known to yield seismic amplification. Because it is based on more robust geomorphological theories than traditional inferred site proxies, like V_{S30} based on slope or geology, we acknowledge the potential value of the Pelletier et al. (2016) model and other similar large-scale models of soil thickness derived from other fields than our own, as possible input in site amplification modelling in large scale seismic hazard and risk modelling.

To test the suitability of the geomorphological model for ground-shaking prediction, we derive a simple site-amplification prediction model using empirical site amplification from the European Engineering Strong-Motion (ESM) dataset (Lanzano et al., 2019; Luzi et al., 2020). The empirical site amplification factors are derived as site-to-site residuals ($\delta S2S_s$) from a simple ground-motion model (GMM) following the method of Kotha et al. (2018, 2022). We compare the ability of the geomorphological model to predict site amplification, to site amplification models based on the traditional site proxies; V_{S30} derived from slope from Wald and Allen (2007), and slope alone, as well as a combination of slope and geological era. To better investigate the differences in the site-amplification prediction maps derived from the different proxies, we focus on Eastern Turkey and Syria where the recent February 2023 Kahramanmaras Earthquake Sequence occurred (Melgar et al., 2023; Petersen et al., 2023).

2 Site amplification factors

The empirical site amplification used in this study is the site-to-site residuals ($\delta S2S_s$), commonly called the "site term", derived from a simple GMM. We derive the GMM using a similar functional form as Kotha et al. (2018, 2020, 2022) and robust mixed-effects regression (rlmm, Koller 2016), where statistical outliers are down-weighted and hierarchical data is dealt with by distinguishing between fixed effects as explanatory variables and random effects as grouping factors (Bates et al., 2015).

A GMM is typically composed of three main explanatory variables describing the source, path and site effects of the ground motion. In its most basic form, magnitude and distance are used to describe the source and path, while V_{S30} is usually used to characterize the site effects. In the GMMs of Kotha et al. (2018, 2020, 2022), however, only the source and path effects are used as fixed effects, while the site is included as a random effect:

$$\ln(\mu) = f_{R,g}(R_{JB}) + f_{R,a}(R_{JB}) + f_M(M_W) + \delta B_e + \delta S2S_s + \delta WS_{e,s} \quad (1)$$

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

$$f_{R,g}(R_{JB}) = c_1 \ln \sqrt{\frac{R_{JB}^2 + h_D^2}{R_{ref}^2 + h_D^2}} \quad (2)$$

$$f_{R,a}(R_{JB}) = \frac{c_3}{100}\left(\sqrt{R_{JB}^2 + h_D^2} - \sqrt{R_{ref}^2 + h_D^2}\right) \quad (3)$$

$$f_M(M_W) = \begin{cases} b_1(M_W - M_h) + b_2(M_W - M_h)^2 & \text{if } M_W \leq M_h \\ b_3(M_W - M_h) & \text{if } M_W > M_h \end{cases} \quad (4)$$

Here, $\ln(\mu)$ is the median ground motion prediction and $f_{R,g}(R_{JB})$, $f_{R,a}(R_{JB})$ and $f_M(M_W)$ are the fixed effects capturing the scaling of the ground motion with geometric spreading, anelastic attenuation and magnitude for Joyner-Boore distance R_{JB}, hypocentral depth h_D and magnitude M_W. The between-event random effect δB_e and site-to-site random effect $\delta S2S_s$ represent the systematic deviation of recorded ground motions from the GMM median predictions related to an event e and a site s, respectively, and $\delta WS_{e,s}$ is the "remaining" record-to-record variability (Kotha et al., 2018; Loviknes et al., 2021). Not including a site-proxy dependent site term in the fixed effects ensures that the $\delta S2S_s$ captures all the site-specific response and thus can be used as an empirical site-amplification function describing the amplification of each station with respect to the median of all sites (Kotha et al., 2018). The $\delta S2S_s$ is assumed to follow a frequency-dependent normal distributions with standard deviation; φ_{s2s}:

$$\delta S2S_S = N(0, \varphi_{s2s}) \quad (5)$$

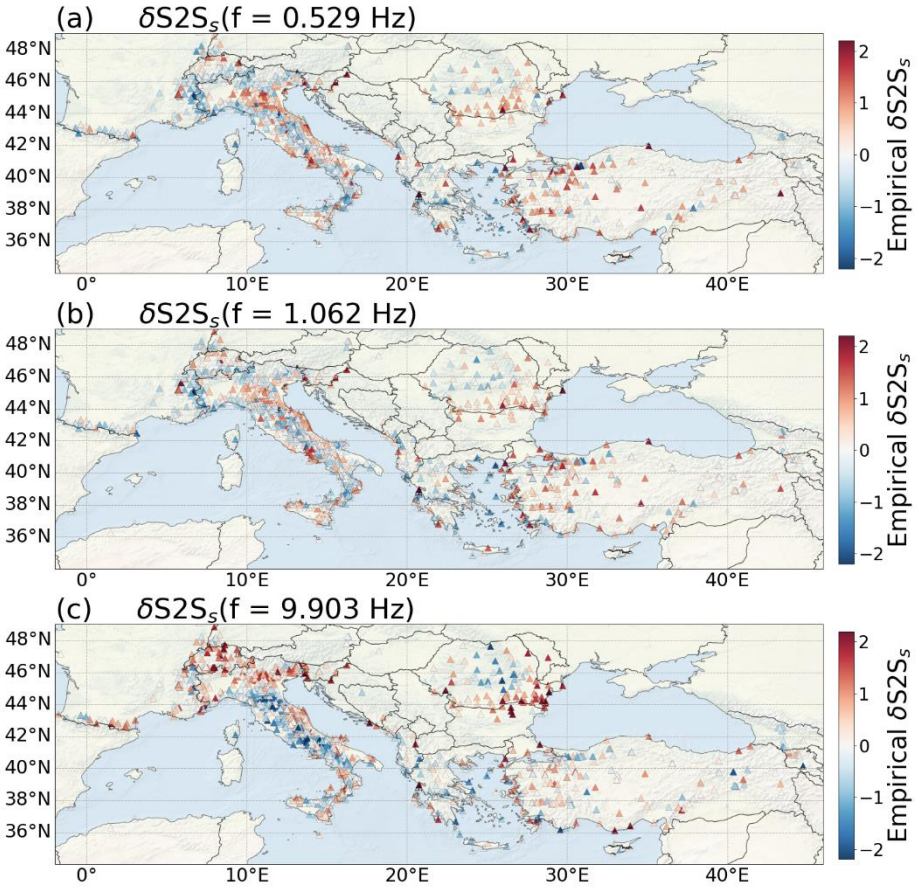
In this study the GMM and corresponding $\delta S2S_s$, is derived in the Fourier Amplitude Spectra (FAS) from the ESM dataset (Lanzano et al., 2019; Luzi et al., 2020). While most GMMs are derived for response spectral amplitudes (SA), representing the damped response of an elastic single-degree-of-freedom oscillator, we here derive the GMM and $\delta S2S_s$ in FAS to better capture the physical effects that can be masked in the response spectra, in particularly at high frequencies (Kotha et al., 2022; Bora et al., 2019; Bayless and Abrahamson, 2019).

To derive the GMM we use the same data selection criteria and similar functional form as Kotha et al. (2022). The GMM of Kotha et al. (2022) is a regionally adaptable models in FAS for shallow crustal earthquakes in Europe and Mediterranean regions. The rationalization in these models is represented by including an earthquake locality-to-locality variability term and an attenuation region-to-region variability term to the random effects. In the GMM derived for this study, these terms are not included and only event and site are used as random effects, this done to minimize the possibility that regional differences in site effects propagate into the region-to-region random-effect.

Unlike traditional site-amplification factors, $\delta S2S_s$ is not relative to a reference rock condition, but to $\delta S2S_s = 0$ which is the centre of the distribution, median, of all the station. The final $\delta S2S_s$ dataset contains site terms for 1680 stations in Europe and the Middle East, as shown on the map in Fig. 1 at f = 0.529, 1.062 and 9.903Hz. Although the site amplification shows a high

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

variability and are mainly dominated by very local effects, some regional effects can be observed, for example for Italy the amplification is mainly high (above the median, red) in the Po-Plains and low (below the median, blue) in the Alps.



125 **Figure 1: Map of the site-amplification factor $\delta S2S_s$ for (a) f = 0.529 Hz, (b) f = 1.062 Hz and (c) f = 9.903 Hz. The colour scale shows the amplification for each station, where red represent amplified ground motions with respect to the median of all the stations, and blue represents deamplified ground motions.**

3 Site proxies: inferred V_{S30}, slope, geomorphological sediment thickness and geological era

3.1 Inferred V_{S30} from slope

130 The V_{S30} dataset of Wald and Allen (2007) has had important implications for large scale seismic hazard and risk assessments and is arguably the most used inferred site proxy in seismic hazard and risk studies (Silva et al., 2020). Wald and Allen (2007) used measured V_{S30} from several location in United States, Taiwan, Italy and Australia to derive a relation between V_{S30} and slope, separating between active and stable tectonic regions. The slope was calculated from global 30 arc sec DEMs from the Shuttle Radar Topography Mission (SRTM30). Here we use the inferred V_{S30} values for Europe directly from the global map

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

published by Wald and Allen (2007). These values range from V_{S30} = 180 m/s to 900 m/s as shown in the map of Europe in Fig. 2a and the distribution plot in Fig. 3a.

3.2 Geomorphological Sediment Thickness (GST)

The Pelletier et al. (2016) model provides a gridded global dataset of soil, intact regolith, and sedimentary deposit thicknesses down to 50 meters. In their model, Pelletier et al. (2016) defines bedrock as the unweathered bedrock below unconsolidated material, which in lowlands are mainly considered sedimentary deposits, and high porosity material, which in uplands can be divided into regolith and soil where soil is the material that sustain life and regolith is the porous weathered material below soil (Pelletier et al. 2016; Holbrook et al., 2014). The model is therefore developed bypartitioning the Earth's surface into uplands and lowlands, which are then separated into hillslopes and valley bottoms. Uplands and lowlands are defined as areas undergoing net erosion and net deposition, respectively, over geological time scales and are distinguished using geological maps and topographic analysis. Hillslopes and valley bottoms are identified using topographic curvature from DEMs, where hillslopes are areas of unconfined surface water flow, while valley bottoms are areas of confined surface water flow. This distinction is particularly important for uplands, and the regolith, soil and sediment thickness values are derived separately for the three landform types: upland hillslopes, upland valley bottoms and lowlands. The values are calculated using mathematical formulas specific to each landform, based on World climate data, water table depths, soil thickness databases and depth-to-bedrock data, among others, as input. The final dataset provided by Pelletier et al. (2016) includes separate 30 arcsec pixel grids covering 60° S–90° N and 180° W–180° E for maximum upland regolith, average soil thickness for upland hillslope, average soil and sediment thickness for upland valley bottom and lowlands, and average soil and sediment thickness across all areas.

The Pelletier et al. (2016) model has been previously tested as a proxy for basin depth in Japan (Weatherill et al., 2020) and were included in the open-source site database of strong-motion stations in Japan by Zhu et al. (2021). However, both Weatherill et al. (2020) and Zhu et al. (2021) used the average soil and sediment thickness for upland valley bottom and lowlands value, while we in this study use the average soil and sediment thickness for all areas, from hereon referred to as geomorphological sedimentary thickness. The geomorphological sedimentary thickness ranges between 0 m and 50 m and are shown in the map of Europe in Fig. 2c and the distribution plot in Fig. 3c.

3.3 Geological era and slope for Europe

In the latest European Seismic Hazard and Risk model (ESHM20, ESRM20), geological era and slope are used to derive the site-response model (Crowley et al., 2021; Weatherill et al., 2023). This approach is based on Vilanova et al. (2018) who made a V_{S30} map for Portugal from geological maps, and Weatherill et al. (2020) who compared several approaches for deriving site amplification from inferred proxies in Japan, including geology and slope. The harmonized surface geology map of Europe is a combination of three geological maps for Europe and Iceland. Because several geological units, both following lithologic

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

(nature) and stratigraphic (age) classification, contain too few stations, the geological units were grouped into the following seven geological eras: Holocene, Pleistocene, Cenozoic, Cretaceous, Jurassic-Triassic, Pre-Cambrian, and Paleozoic. The map and station distribution of these eras are shown in Fig. 2d and 3d. The slope used in this model was calculated from the 2014 General Bathymetric Chart of the Oceans grid (GEBCO 2014, https://www.gebco.net/) and is shown in Fig. 2b and 3b. In this study we use the same slope and geological eras as Weatherill et al. (2023), which are available on the EFEHR seismic risk web-services (http://risk.efehr.org/site-model/).

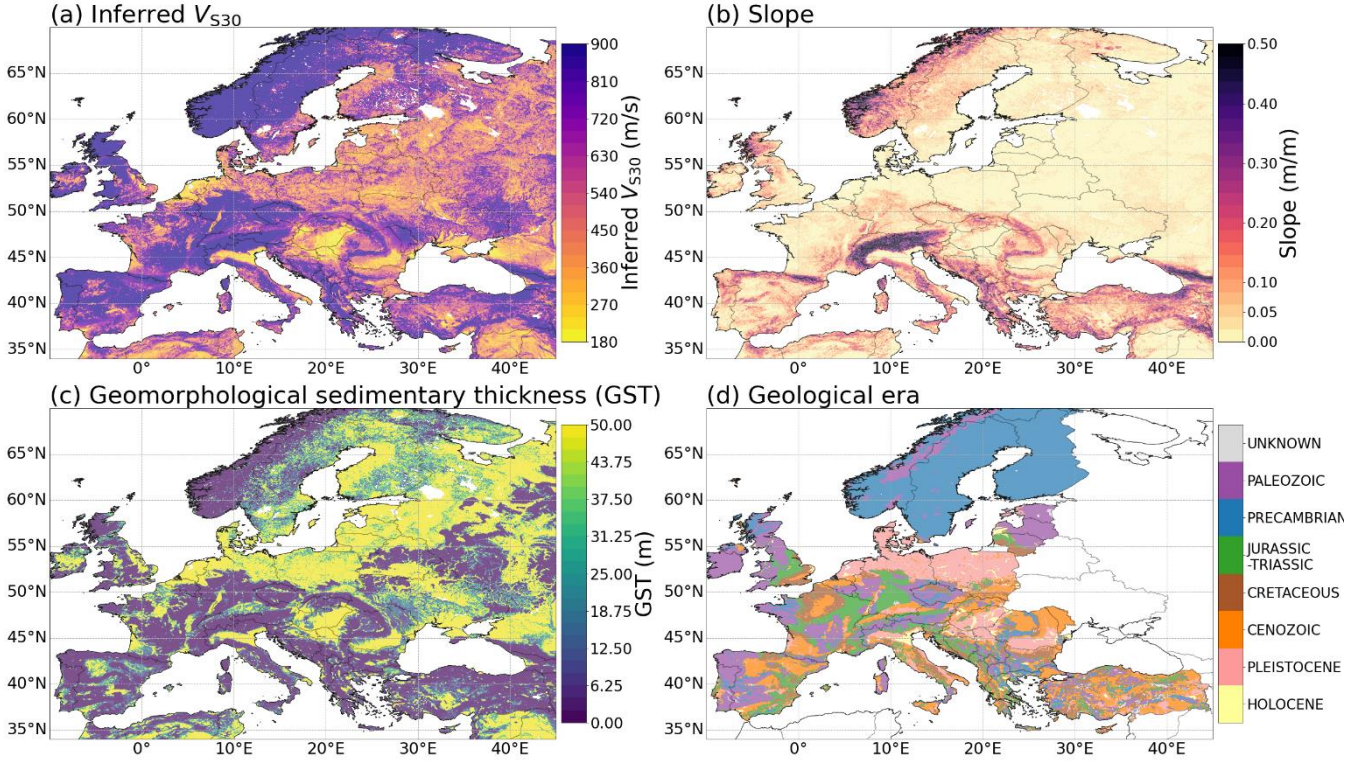


Figure 2: Map of the site-proxies to be tested in this study. (a) V_{S30} from slope by Wald and Allen (2007), (b) slope calculated from digital elevation models, (c) geomorphological sedimentary thickness by Pelletier et al. (2016), and (d) geological era used in the latest European Seismic Risk model (ESRM20, Crowley et al., 2021; Weatherill et al., 2023)

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

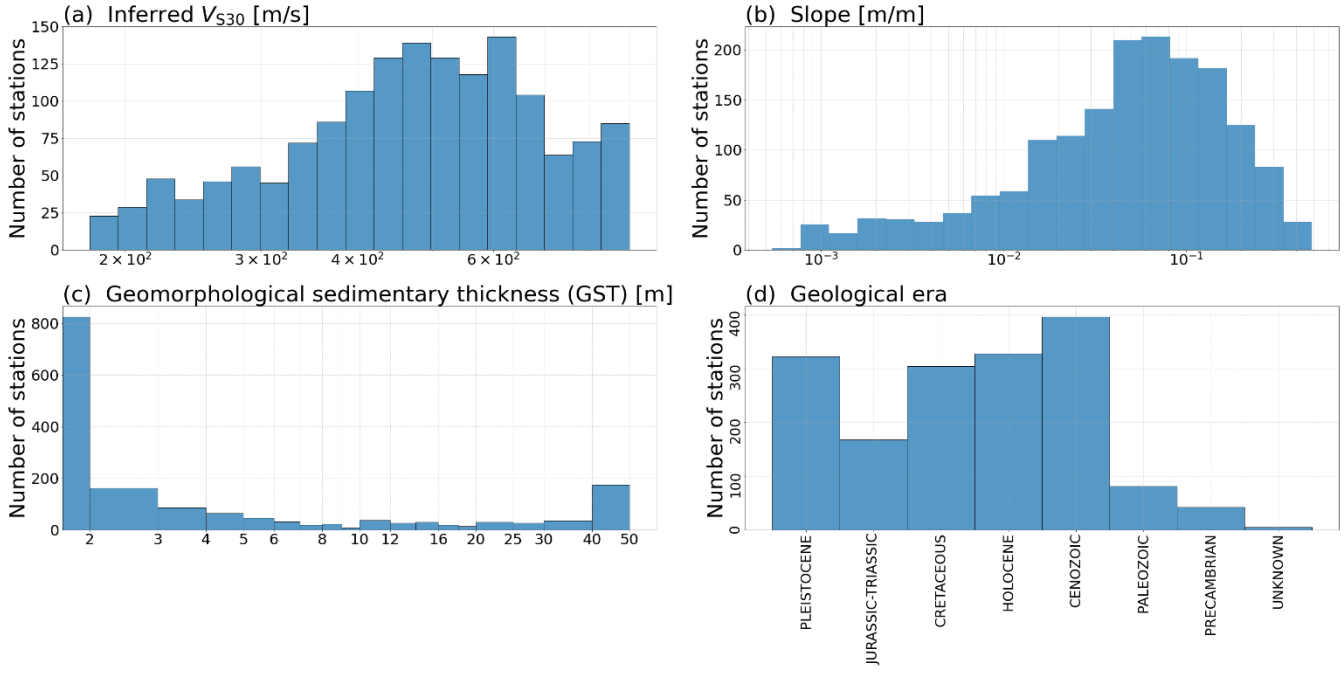


Figure 3: The distribution of the site-proxies to be tested in this study. (a) V_{S30} from slope by Wald and Allen (2007), (b) slope calculated from digital elevation models, (c) geomorphological sedimentary thickness by Pelletier et al. (2016), and (d) geological era used in the latest European Seismic Risk model (ESRM20, Crowley et al., 2021; Weatherill et al., 2023).

4 Amplifications predictions according to the different proxies

We evaluate the ability of the different proxies to predict site amplification by deriving a site-amplification model for each proxy. The relation between site amplification and measured V_{S30} is generally considered as log-linear, we therefore also assume a log-linear relation between the site amplification and the inferred site proxies and use linear regression to derive the site amplification models:

$$Y_S(f, Proxy) = a\, log(x_{\mathrm{Proxy}}) + b \qquad (6)$$

where $Y_s(f,Proxy)$ is the predicted site amplification for a site s at frequency f using a proxy x_{Proxy}, and a and b are the coefficients derived from the linear regression.

A log-normal distribution is generally assumed for V_{S30} and depth to bedrock. Figure 3 shows the distribution of the inferred proxies, and the log-normal assumption is not fully fulfilled for inferred V_{S30} and especially geomorphological sediment thickness. This is because the proxies are limited to a certain range during their calculation process. In the case of

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

geomorphological sediment thickness, Pelletier et al. (2016) set the maximum value to 50 m, effectively meaning the thickness of the sediment layer is 50 m *or more*.

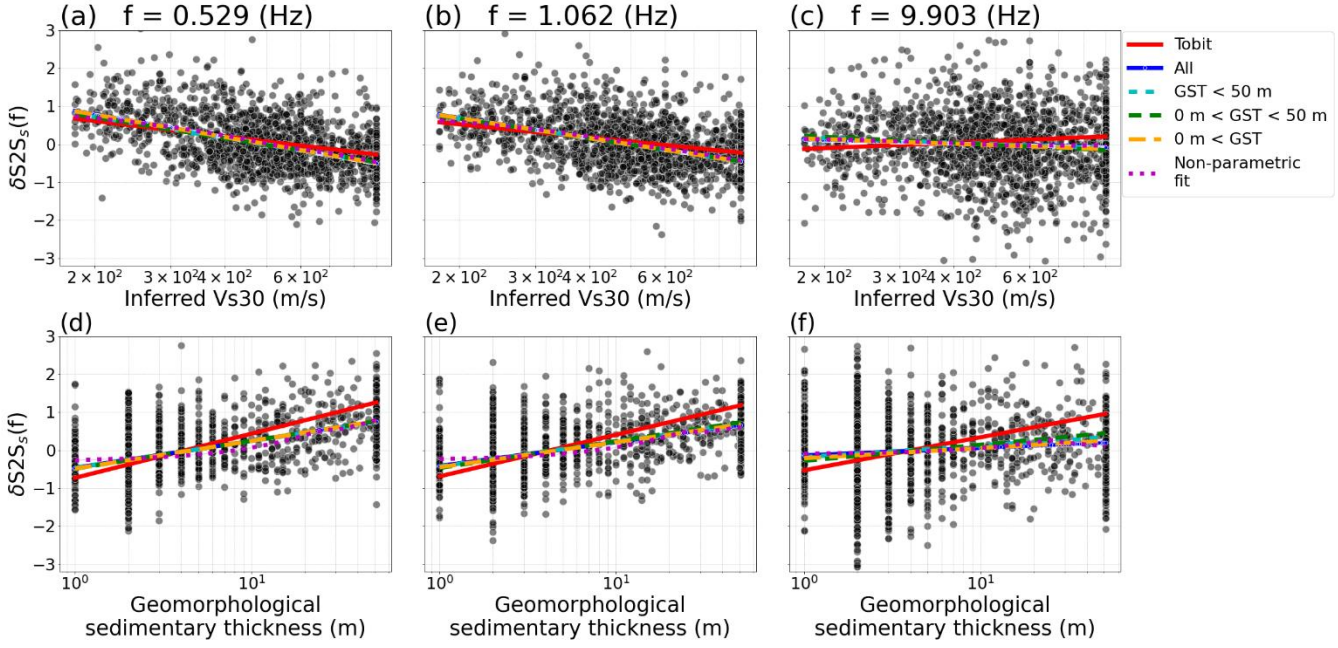


Figure 4: Inferred V_{S30} (a, b, c) and geomorphological sedimentary thickness (d, e, f) with the $\delta S2S_s$ of all the 1680 stations (black dots) in the ESM dataset for the frequencies f = 0.529 Hz, (a, d), f = 1.062 Hz (b, e) and f = 9.903 Hz (c, f). The regression lines are from the Tobit regression (solid red line), and from the linear regression on all the data (dashed blue line), on the selected dataset without the maximum geomorphological sediment thickness values (50 m, dashed cyan line), without the extreme values of geomorphological sediment thickness (0 and 50 m, dashed green line), and without the minimum geomorphological sediment thickness (0 m, dashed orange line). The general trend of the data is shown by a non-parametric fit (dotted magenta line).

When dealing with such so-called "censored data", Tobit regression (Tobin, 1958) is a possibility. The Tobit model is developed to estimate linear relationships when the dependent variable is censored and uses the likelihood function to deal with the uneven distribution (Amemiya, 1984). However, as can be seen in Fig. 4 (red line), the Tobit regression strongly overestimates the slope of the relation between the site proxies and site amplification, which is also demonstrated by the non-parametric fit (dotted magenta lines). The slope from the Tobit regression is especially overestimated at high frequencies, where a weak relation between the empirical site amplification and site proxies is expected due to the impact of small-scale heterogeneities at the site where even the location or housing of the strong-motion station can affect the amplification (Hollender et al., 2020) and the coarse spatial resolution (30 arc second) of the site amplification model.

Another alternative approach is to exclude the extreme values when running the regression, however, this step only excludes a high number of sites while not having a strong impact on the regression line (Fig. 4, blue, green, yellow, and cyan line).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Instead, we only omit sites with geomorphological sediment thickness = 0 m, which is the value that causes the highest unevenness in the distribution, as well as sites with missing values for any of the proxies, leaving us with 1508 sites for the regression. To evaluate the dependency of the regression on the selection of sites, we run a 10 fold cross validation test. In a 10-fold cross-validation test, the dataset is split into 10 equal parts and the model is derived on 10-1 parts and tested on the remaining 1 part of the dataset. This is done 10 times and for each run a different subset of the data is used for validation (Bishop and Nasrabadi, 2006). The distributions of the site proxies for each 10 fold cross validation iteration are shown in Fig. A1 in the Appendix.

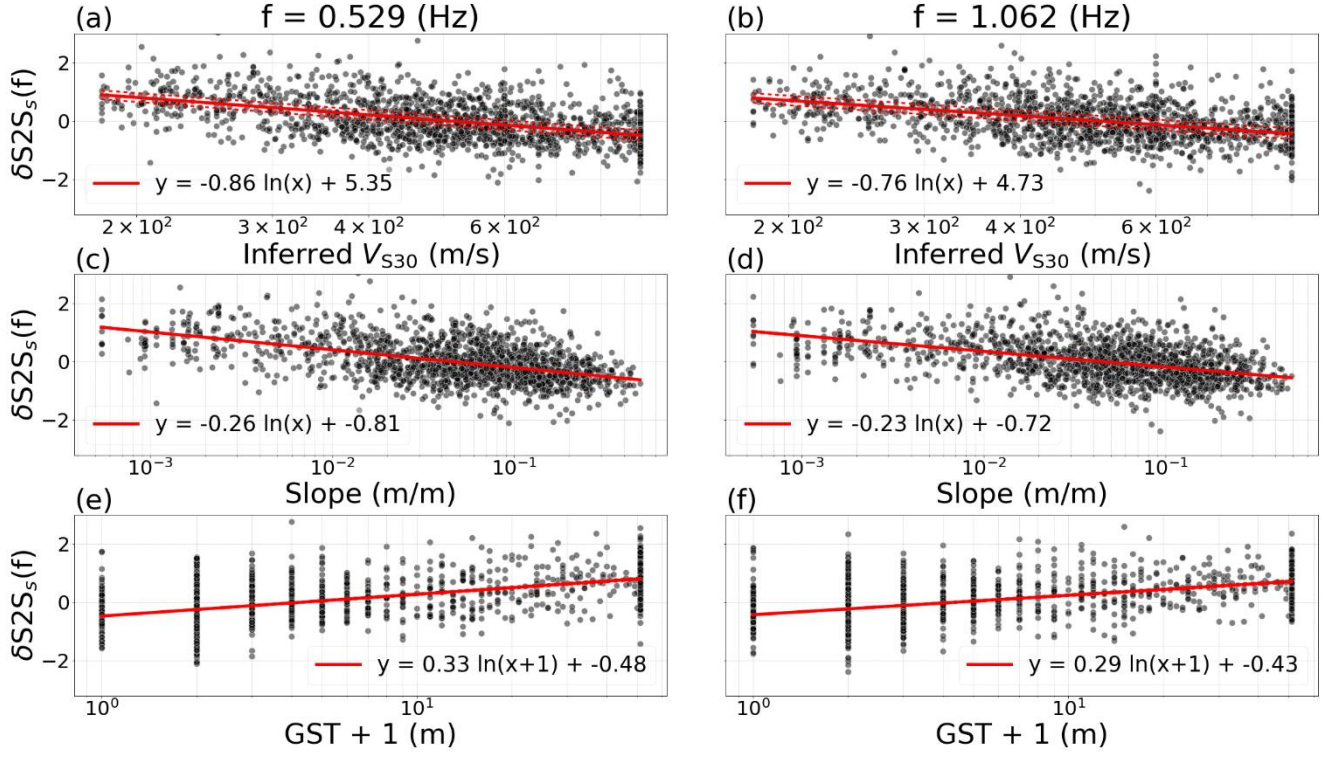


Figure 5: Linear regression (red lines) over the site proxies inferred V_{S30} (a, b), slope (c, d) and geomorphological sedimentary thickness (e, f), with the station $\delta S2S_s$ (black dots) for the frequencies f = 0.529 Hz, (left; a, c, e) and f = 1.062 Hz (left; b, d, f).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

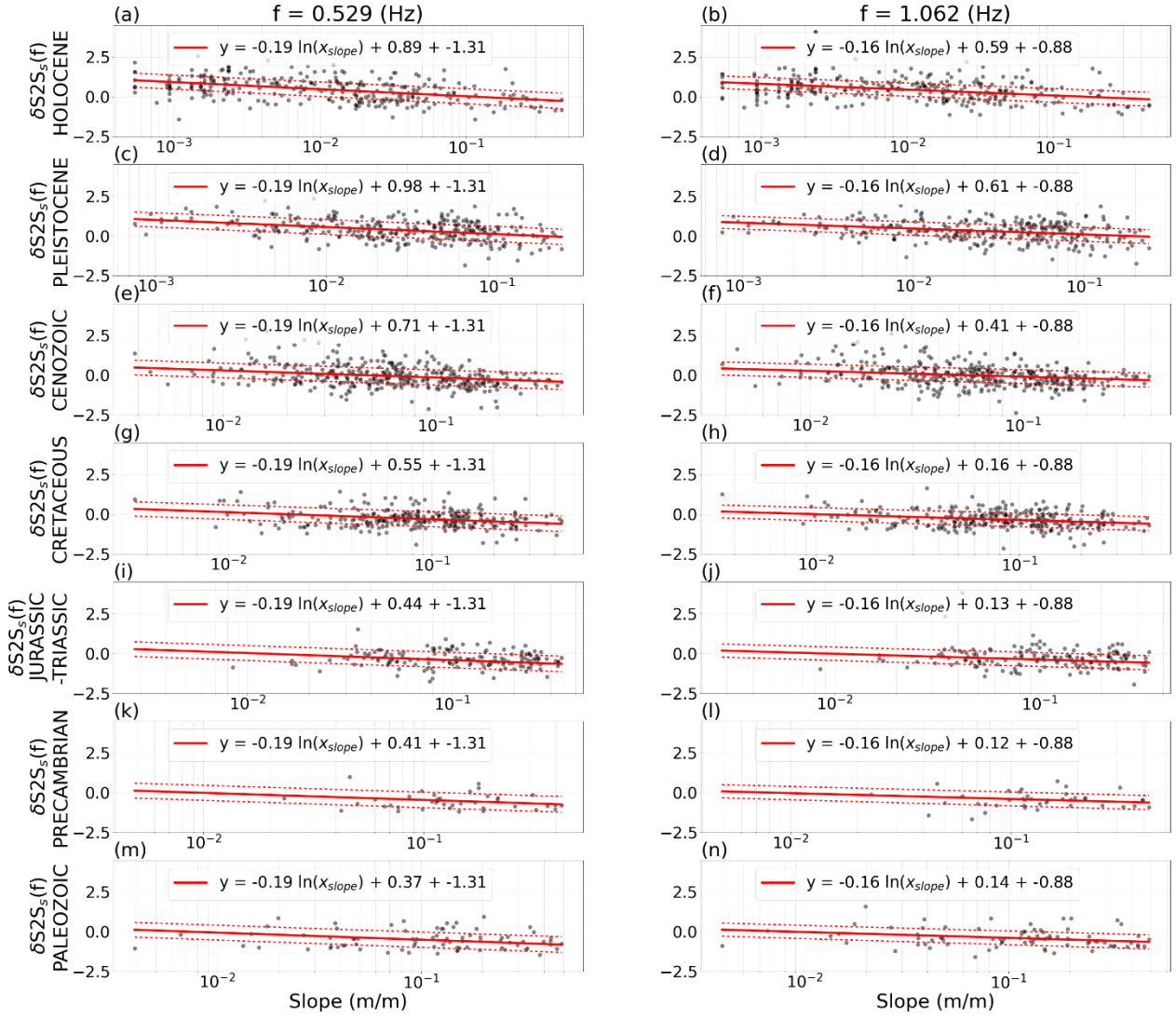


Figure 6: Linear regression (red lines) over slope and the geological eras; Holocene (a, b), Pleistocene (c, d), Cenozoic (e, f), Cretaceous (g, h), Jurassic-Triassic (i, j), Precambrian (k, l) and Palaeozoic (m, n), with the station $\delta S2S_s$ (black dots) for the frequencies f = 0.529 Hz, (left) and f = 1.062 Hz (right).

We derive site-amplification models from linear regression between $\delta S2S_s$ for the frequency range f = 0.460 − 9.903 Hz and the site proxies, V_{S30} from slope, slope alone, geomorphological sediment thickness and geological era combined with slope. The selected sites and the regression lines are shown for f = 0.529 Hz and f = 1.062 Hz in Fig. 5 for inferred V_{S30}, slope and geomorphological sediment thickness. Although the $\delta S2S_s$ shows a high scatter with the site proxies, a general trend of higher amplification for low V_{S30}, low slope and high sediment thickness and low amplification for vice versa, can be identified. For the regression over geological era and slope combined, we apply multiple linear regression, meaning with multiple independent

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

variables where the categorical predictor, geological era, is transformed to dummy variables for each era and the regression then derives a constant coefficient for slope with a different intercept for each geological era (Fig. 6). The variation in prediction coefficients due to the alteration of the training-set in the cross-validation process, shown as dotted red lines in Fig. 5 and 6, is small for all the proxies except inferred V_{S30} and geological era. This indicates that the site amplification models based on inferred V_{S30} and geological era and slope combined are more dependent on the data selection than the other proxies, causing a higher final uncertainty.

5 Reduction in site-to-site variability

After deriving an amplification model based on each proxy, we compare the predicted amplification $\delta S2S_s(f,Proxy)$ with the empirical amplification $\delta S2S_s(f)$ at a site s and frequency f. We measure the ability of each proxy to capture the site amplification using the reduction in site-to site variability φ_{s2s} as an indicator of the efficiency of each proxy to predict the amplification (Stewart et al., 2017; Zhu et al., 2022). We compute the corrected site term for each proxy-specific predicted amplification $\delta S2S_{s,cor.}(f,Proxy)$:

$$\delta S2S_{s,cor.}(f,Proxy) = \delta S2S_s(f) - \delta S2S_s(f,Proxy) \quad (7)$$

$$\delta S2S_{s,cor.}(f,Proxy) = N(0,\varphi_{s2s,cor.}(f,Proxy)) \quad (8)$$

Where $\varphi_{s2s,\,cor.}(f,Proxy)$ is the site-to-site variability of the sites corrected for the site-amplification prediction based on a proxy for frequency f, as shown in Fig. 7. The greater the reduction in variability, meaning a lower $\varphi_{s2s,\,cor.}(f,Proxy)$ the better the ability of the proxy to capture the site amplification. However, it is important to keep in mind that this measure and the correlation between $\delta S2S_s(f)$ and the site proxies are purely statistical and does not give any insight into what is causing the amplification and its variability.

As described in Section 4, the site-amplification models and corresponding reduction in φ_{s2s} were derived and calculated 10 times following the 10 fold cross validation technique. This means we are dealing with two sources of variability; the site-to-site variability φ_{s2s} and the variability related to running the regression on different subsets of the data in the 10 fold cross correlation, here called ε_{ee}. φ_{s2s} is a combination of the natural, random and irreducible (aleatory) variability of site response, and the epistemic uncertainty related the to the site proxies not being able to fully capture the site properties controlling the site amplification. ε_{ee} is fully epistemic as it is related to the difference between the site-amplification models derived using different datasets. Figure 7 shows the mean φ_{s2s} for all stations (black lines) and $\varphi_{s2s,\,cor.}(Proxy)$ for each proxy (coloured lines) from the 10 cross-validation iterations derived on the training-sets (Fig. 7a) and the validation-sets (Fig. 7b). The shaded areas around the means are the variance ε_{ee} related to the cross-validation process. ε_{ee} is, as expected, higher for the validation set (Fig. 7b), which for each iteration correspond to only 10% of the dataset, than for the training set (90% of the data, Fig.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

7a). Nonetheless, the general pattern is similar for both sets, where the highest reduction in site-to-site variability is caused by the site-amplification model based on geological era and slope. Inferred V_{S30}, slope and geomorphological sediment thickness show similar reduction in variability for both the training and validation set, with around 1% difference for the training-set. For the validation-set the reduction caused by Inferred V_{S30}, slope and geomorphological sediment thickness are within the same standard deviation ε_{ee} and are not clearly distinguishable (Fig. 7b). For both the training and validation set, none of the site-amplification models are distinguishable for frequencies above 3 Hz, which might be caused by the low resolution (30 arc seconds) of all the proxies considered in this study. The φ_{s2s} for each cross-validation iteration is shown in Fig. A2 in the Appendix.

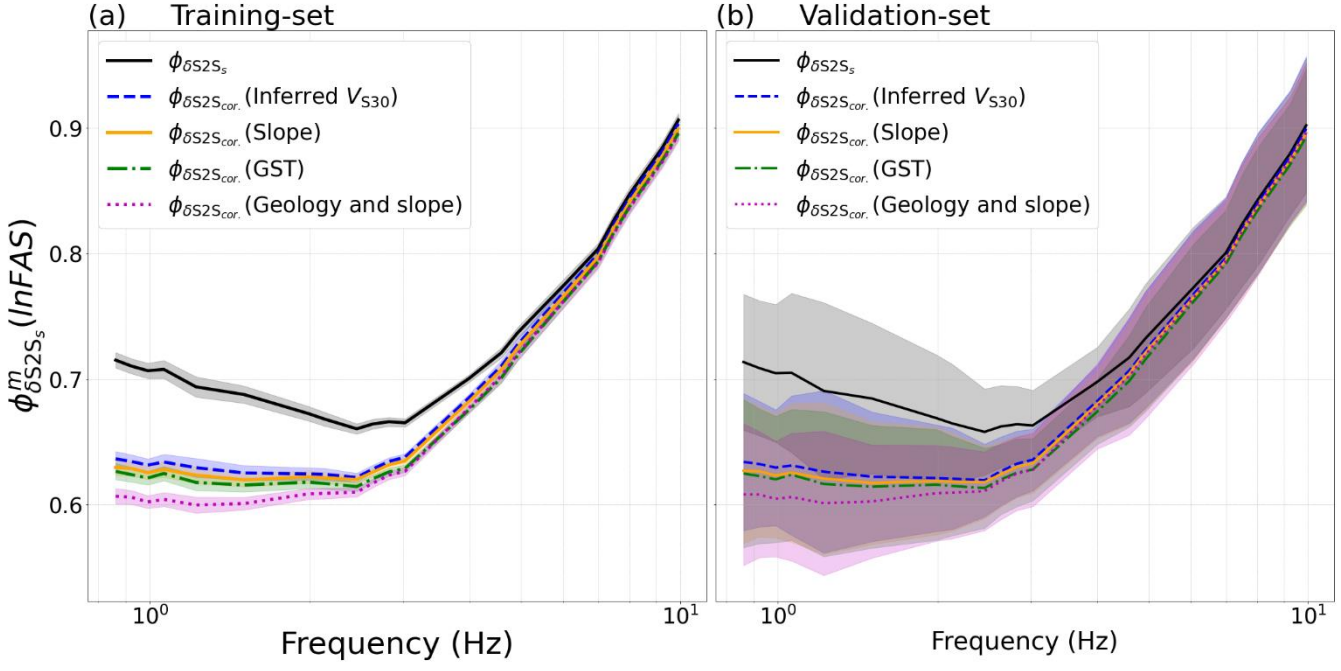


Figure 7: The site-to-site variability φ_{s2s} for all selected stations (solid black line) and the corrected site-to-site variability after subtracting the predicted site amplification using inferred V_{S30} ($\varphi_{s2s,cor.}(V_{S30})$, dashed blue lines), slope ($\varphi_{s2scor.}$(Slope), solid orange lines), geomorphological sediment thickness (GST) ($\varphi_{s2s,cor.}$(GST), dash-dotted green lines) and geology and slope ($\varphi_{s2scor.}$(Geology and slope), dotted magenta lines) from the empirical site amplification using (a) the 10-1 part training-set and (b) the 1 part validation-set.

The results shown in Fig. 7, indicate that the site-amplification model based on geological era and slope combined is better at capturing site amplification relative to the other proxies used in this study. These results are consistent with the findings of Weatherill et al. (2020), who also derived site-amplification models based on several inferred proxies for Japan. However, when applying the same model for Europe, Weatherill et al. (2023) found that the reduction from geological era and slope was not significantly lower than for inferred V_{S30} or slope alone. They speculate that this could be an artifact of the mixed-effects

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

regression used to derive the model, where geological era and slope were included as a random effect, meaning the coefficient for slope change with each geological era, which is different for the multiple linear regression applied in this study where the slope coefficient stays the same for each geological era. In addition, our study uses $\delta S2S(f)$ from FAS, which is likely to affect the site-to-site variability.

Nevertheless, the case that geological era and slope combined gives the highest reduction to the site-to-site variability shows the importance of including geology in site amplification modelling. Furthermore, the new proxy geomorphological sedimentary thickness shows similar or even slightly higher reduction in site-to-site variability as the traditional proxies inferred V_{S30} and slope, indicating the potential of geomorphological sedimentary thickness as an alternative site proxy for seismic hazard assessments for large areas or areas without measured site parameters.

6 Proxy based site amplification predictions and maps

To further evaluate the ability of the site proxies to predict site amplification we compare the predicted site amplification with empirical site amplification for the entire frequency range (f = 0.460 − 9.903 Hz, Fig. 8). The object of this study is to predict regional site amplification over a large area using regionally or globally available site proxies. It therefore does not make sense to compare our site-amplification predictions to empirical site amplification at a single site. Instead, we compare the predicted amplification to empirical amplification from a set of sites with similar measured V_{S30}. We use the measured V_{S30} provided in the ESM database and group the sites into three ranges representing soft soils (V_{S30} = 175 ± 25 m/s, Fig. 8a), stiff soils (V_{S30} = 375 ± 25 m/s, Fig. 8b) and rock sites (V_{S30} = 775 ± 25 m/s, Fig. 8c). Selecting corresponding ranges of site properties for predicting the site amplification using other proxies than inferred V_{S30}, requires some attention. The correlation between measured V_{S30} and these proxies are shown in Fig. 9 and show thatlow V_{S30} corresponds to low slope (Fig.9a, below orange line) and deeper sediments (Fig.9b, between orange lines). Although the relations show a high variability escepsially for high V_{S30}, we select low slope at 0.001 ± 0.001 m/m and high geomorphological sediment thickness at 40±10m for predicting soft soils. Equivalently we use slope at 0.02 ± 0.005 m/m (Fig.9a, between red lines) and 0.3 ± 0.005 m/m (Fig.9a, between purple lines), and geomorphological sediment thickness at 8 ± 2 m (Fig.9b, between red lines) and 1 ± 1 m (Fig.9b, below purple lines) for predicting stiff soil and rock, respectively. When selecting corresponding geological eras, we use Holocene for soft soil (Fig.9a, yellow scatter points), Pleistocene for medium soil (Fig.9a, pink scatter points) and the remaining geological eras (leaving out Cenozoic) Cretaceous, Jurassic-Triassic, Precambrian and Paleozoic are grouped together for rock (Fig.9c, brown, green, blue and purple scatter points). The selected ranges of site properties are given in Table 1.

Table 1: The ranges of site properties selected to correspond with the V_{S30}-ranges for soft soil (left column), stiff soil (centre column) and rock (right column).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Proxy	Soft soil	Stiff soil	Rock
Measured V_{S30}	175 ± 25 m/s	375 ± 25 m/s	775 ± 25 m/s
Inferred V_{S30}	175 ± 25 m/s	375 ± 25 m/s	775 ± 25 m/s
Slope	0.001 ± 0.001 m/m	0.02 ± 0.01 m/m	0.3 ± 0.1m/m
GST	40 ± 10 m	10 ± 2 m	1 ± 1 m
Geological era	Holocene	Pleistocene	Cretaceous, Jurassic-Triassic, Precambrian and Paleozoic

Figure 8 shows the mean of the empirical site amplification (solid black line) with the standard deviation (shaded black area) of the sites with similar measured V_{S30}. The variability of the empirical site amplification even within each site group (soft soil, stiff soil and rock) is very high, and the predicted site amplification using any of the proxies are within the standard deviation of the three V_{S30} ranges. For soft and stiff soils, the different proxies have similar predictions close to the mean of the empirical amplification. For rock, inferred, slope and geomorphological sediment thickness are in the upper range of the empirical site amplification standard deviation, but follow the shape of the mean empirical site amplification. The prediction based on geological era and slope, however, have a different shape and are slightly under-predicting for low frequencies and over-predicting for higher frequencies relative to the mean empirical site amplification. However, the differences between the predictions and empirical amplification could also be due to a poor correspondence between the selected ranges of inferred proxy values and the measured V_{S30}, in particularly for geological era.

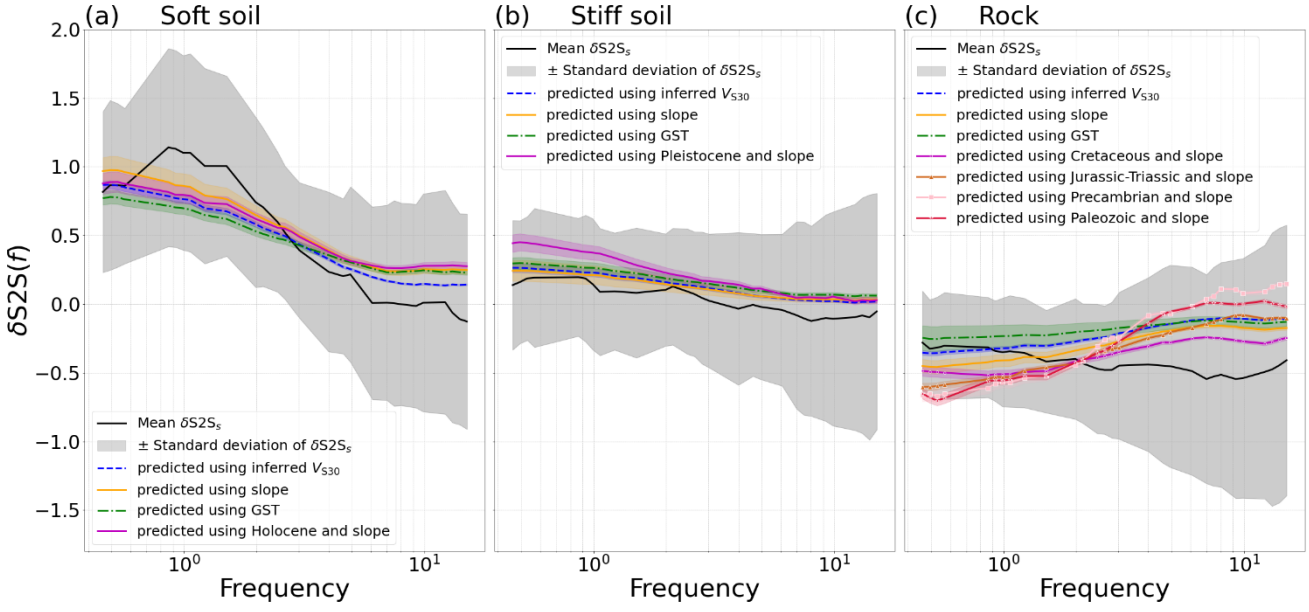
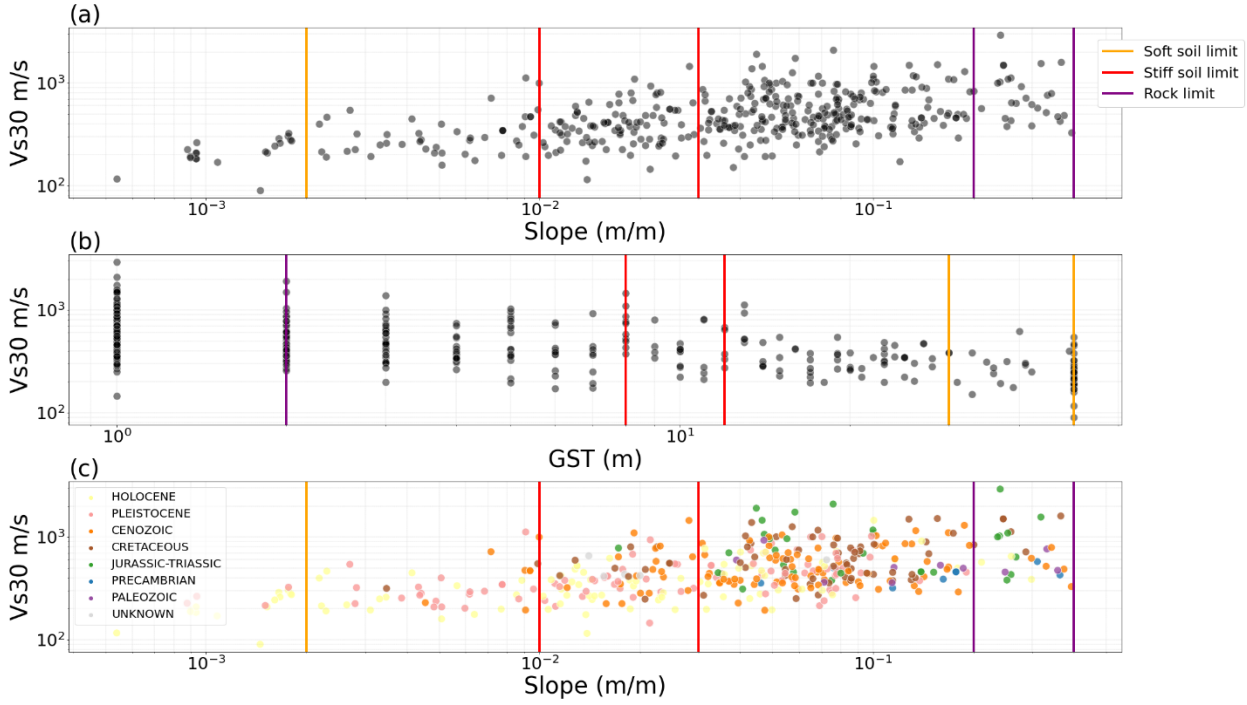


Figure 8: Empirical mean (solid black line) site amplification with standard deviation (shaded black area) compared to predicted site amplification for soft soil (a), stiff soil (b) and rock sites (c) using inferred V_{S30} (dashed blue lines), slope (solid orange lines), geomorphological sediment thickness (dash-dotted green lines) and geological era and slope (dotted magenta lines).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023



340 **Figure 9: The correlation between measured V_{S30} and (a) slope, (b) geomorphological sediment thickness (GST) and (c) slope and coloured by geological era. The yellow lines show the chosen ranges for soft soil (orange) stiff soil (red) and rock (purple).**

Finally, we use the proxy-based site-amplification models to predict the site amplification for entire Europe, as shown in Fig. 10 for $f = 1.062$ Hz. The amplification maps at $f = 0.529$ Hz and $f = 9.903$ Hz are shown in Figures A3 and A4 in the Appendix.
345 The site-amplification maps show the amplification (red) and deamplification (blue) relative to the median site amplification (white). This is different from conventional amplification maps used in PSHA, where the site amplification is relative to a rock reference. In this study we keep the amplification relative to the median to avoid biasing the result with poorly constrained rock properties. The maps all show similar features, for example high amplification in the deep-sediment basins of the Po-Plains in northern Italy, the Danube plains of Eastern Romania and the Great Hungarian plains, and strong deamplification in
350 the Alps, the Carpathian Mountains and Western Norway. However, clear differences are also evident in the different site-amplification maps, for example around the Rhine valley, in the Baltic and in Eastern Turkey.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

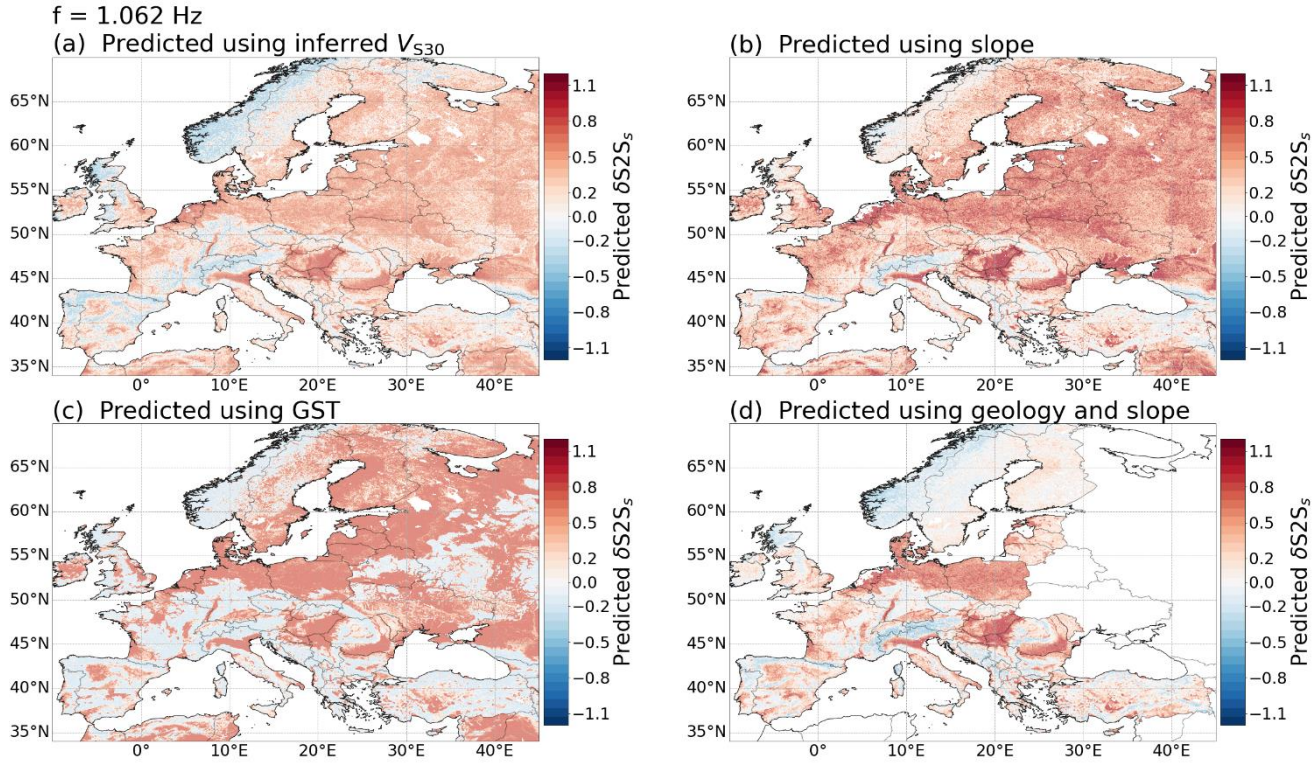


Figure 10: Predicted site amplification at f = 1.062 Hz for Europe using (a) inferred V_{S30} (b) slope, (c) geomorphological sedimentary thickness, and (d) geological era and slope.

355

6.1 Justifications of the differences in the site amplification predictions: A focus on Eastern Turkey

To investigate the differences in the predictions further, we zoom in on Eastern Turkey where the recent Kahramanmaras Earthquake Sequence of February 2023 occurred (Melgar et al., 2023; Petersen et al., 2023). Figures 11 and 12 show the site proxies and predicted site amplification at f = 1.062 Hz in Eastern Turkey and Syria. The main difference between the site
360 proxies and their corresponding site amplification is in the area south of the South-eastern Taurus Mountains, around the cities Gaziantep and Aleppo, and by the border between Turkey and Syria. In the map based on inferred V_{S30} (Fig. 11a) and slope (Fig. 11b), low values of V_{S30} and slope are present on both sides of the border and in the corresponding amplification maps (Fig. 12a and b), medium to high ($0.2 < \delta S2S_s < 0.6$) amplification is predicted around the border with gradual lower amplification towards the Taurus Mountains. The map of geomorphological sedimentary thickness (Fig. 11c) differs markedly
365 from that of inferred V_{S30} and slope north of the Turkey-Syria border with mainly shallow (< 10 m) sediments and low amplification ($-0.6 < \delta S2S_s < 0$) is predicted for that area (Fig. 12c). South of the Turkey-Syria border, however, the geomorphological sedimentary thickness map shows deep (> 30 m) sediments and predicts high amplification ($\delta S2S_s > 0.6$). This sharp difference at the Turkey-Syria border in the geomorphological sedimentary thickness is likely due to the fact that

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Pelletier et al. (2016), in the process of differentiating between upland and lowland, used separate geologic maps for Europe and the Arabian Peninsula. The site amplification map based on geological era and slope combined (Fig. 12d) predicts less high amplification than inferred V_{S30} and slope and mainly in concentrated areas north of the Turkey-Syria border (Fig. 12a and b). However, because the geological era maps were created for the European Seismic Hazard and Risk model, the maps based on geological era and slope end at the border of Turkey and does not include Syria in (Fig. 11d).

Because there are few stations in the area with the largest differences, it is difficult to make a solid argument for which proxy is more correct. For this purpose, a more detailed comparison with local values could be a possibility, but, as stated before, the aim of this study is not to re-create the exact site-specific amplification, but to predict the site amplification for larger areas using inferred proxies and capture, using several models, the epistemic uncertainty of such regional amplification prediction. Instead, we compare the empirical amplification from the few stations in the area with the predicted amplification in a similar way as in Fig. 8. We only use the 14 stations with measured V_{S30} values, these stations are marked in Fig. 12 and are described in Table 2. Because the measured V_{S30} of the considered stations only range between 300 and 700 m/s, it does not make sense to separate the stations into soft soil, stiff soil and rock V_{S30} ranges as done in Fig 8. Instead, the stations are separated into two ranges with for medium stiff soil sites (V_{S30} = 300−500m/s, Fig. 13a) and stiffer sites (V_{S30} = 500−700m/s, Fig. 13b). The comparison of empirical amplification and predicted amplification for the East-Turkish stations in the two V_{S30}-ranges shows that all the proxies are significantly under-predicting the site amplification for the medium stiff soil sites and over-predicting the stiffer sites. Figure 13 further shows that predicting site amplification using proxies cannot reproduce the full range of amplifications on a local level and that regionalized models might be necessary.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

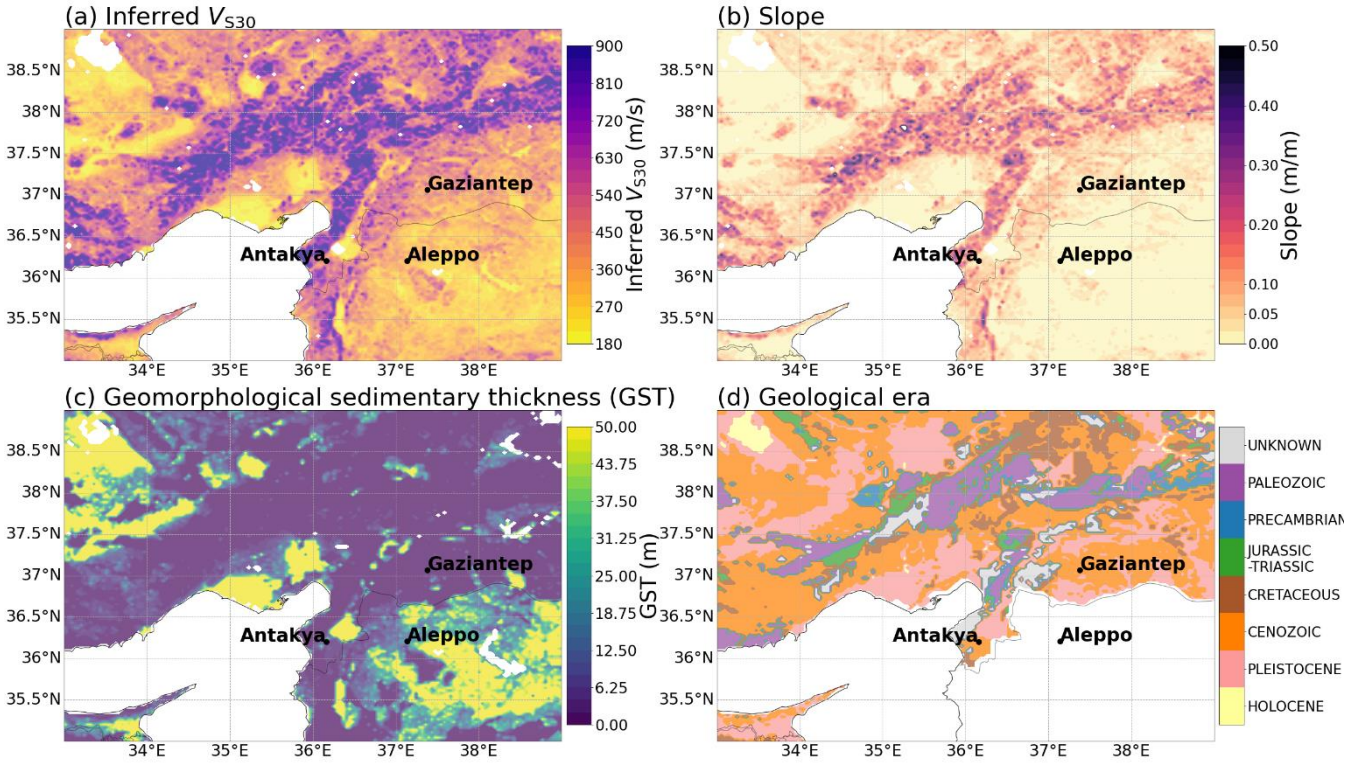


Figure 11: Map of Eastern Turkey and Syria coloured by the site-proxies (a) inferred V_{S30} (b) slope, (c) geomorphological sedimentary thickness, and (d) geological era. The cities Gaziantep, Antakya and Aleppo are indicated on the map.

Table 2: Station name, location, and site properties of the East Turkish stations with measured V_{S30} shown in Fig. 12.

Station code	Latitude	Longitude	Measured V_{S30} (m/s)	Inferred V_{S30} (m/s)	Slope (m/m)	GST (m)	Geological era
TK-4401	38.34962	38.34019	481.0	502.0	0.069	3.0	Cenozoic
TK-3102	36.21300	36.15900	469.0	340.0	0.009	27.0	Cenozoic
TK-4403	38.09616	37.88732	655.0	425.0	0.014	1.0	Unknown
TK-4605	38.20368	37.19771	315.0	313.0	0.003	38.0	Cretaceous
TK-4603	37.57998	36.93061	465.0	523.0	0.049	3.0	Pleistocene
TK-4604	37.57010	36.35737	613.0	625.0	0.064	2.0	Cenozoic
TK-3101	36.21423	36.15973	469.0	340.0	0.009	27.0	Cenozoic
TK-3103	36.11593	36.24722	344.0	512.0	0.044	4.0	Cenozoic
TK-3105	36.80262	36.51119	619.0	602.0	0.078	1.0	Pleistocene
TK-3111	36.37260	36.21973	338.0	619.0	0.051	4.0	Pleistocene
TK-2701	37.02546	36.63593	421.0	457.0	0.020	16.0	Unknown
TK-4607	37.48513	37.29775	672.0	501.0	0.059	-	Cenozoic
TK-2702	37.18430	36.73280	599.0	469.0	0.073	3.0	Unknown
TK-4601	37.53872	36.98187	345.0	349.0	0.017	22.0	Pleistocene

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

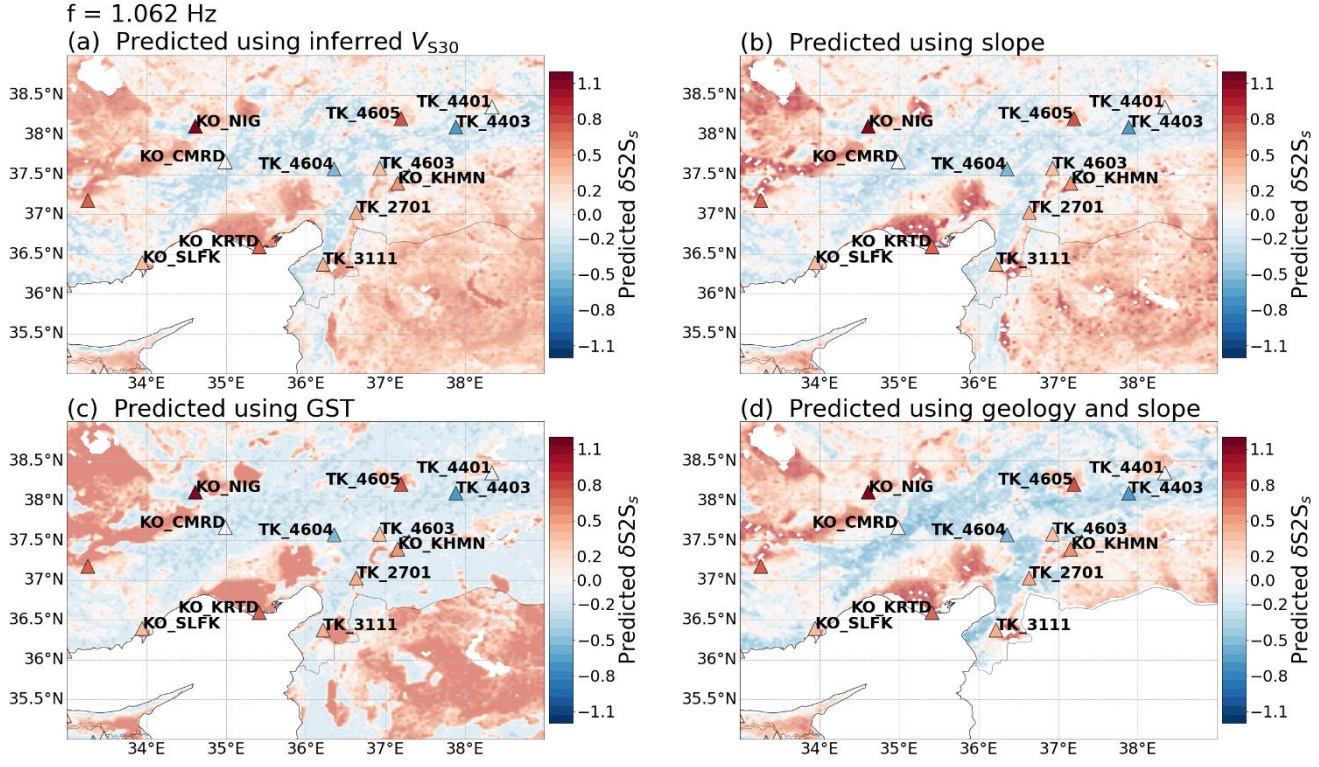


Figure 12: Predicted site amplification in Eastern Turkey and Syria at f = 1.062 Hz using (a) inferred V_{S30} (b) slope, (c) geomorphological sedimentary thickness, and (d) geological era and slope. The triangles represent the strong motion stations coloured by empirical site amplification δS2S$_s$ at f = 1.062 Hz. The stations with measured V_{S30}-values are indicated on the map

Another approach to investigate the variations in the amplification predictions and demonstrate the impact of the differences is to look at the site amplification predicted for a city. The ranges of the coordinates and site properties of the three cities Antakya, Aleppo and Gaziantep indicated in Fig. 11 are given in Table 3, and Fig. 14 shows the range of predicted amplification at f = 0.529 Hz and f = 1.062 Hz at these cities. Fig. 14 shows that even for relatively small areas and cities, where the accuracy of the expected amplification is the most important, using different proxies produce widely different amplification predictions. This is especially true for Gaziantep where geomorphological sedimentary thickness is shallow (between 0 and 2m) and predicts deamplification ($-0.1 < \delta S2S_s < -0.5$), while the other proxies predict the opposite ($0.2 < \delta S2S_s < 0.5$) and consist of slow V_{S30} (275 – 350 m/s), small slope (0.015 − 0.019 m/m) and young era (Cenozoic). Although this study focuses on large-scale areas, and more detailed and local site amplification studies are preferred for cities, such investigations are not always possible or widely available. It is therefore not uncommon to use inferred proxies like V_{S30} from slope or geology, even for small areas (e.g., Michelini et al., 2008; Thompson et al., 2014; Vilanova et al., 2018; Foster et al., 2019; Li et al., 2022). Our results are therefore still relevant on a city level because it shows the importance of capturing the epistemic uncertainty related

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

to site-amplification modelling when using inferred proxies and that using only inferred V_{S30}, which is often the standard procedure, is simply not enough.

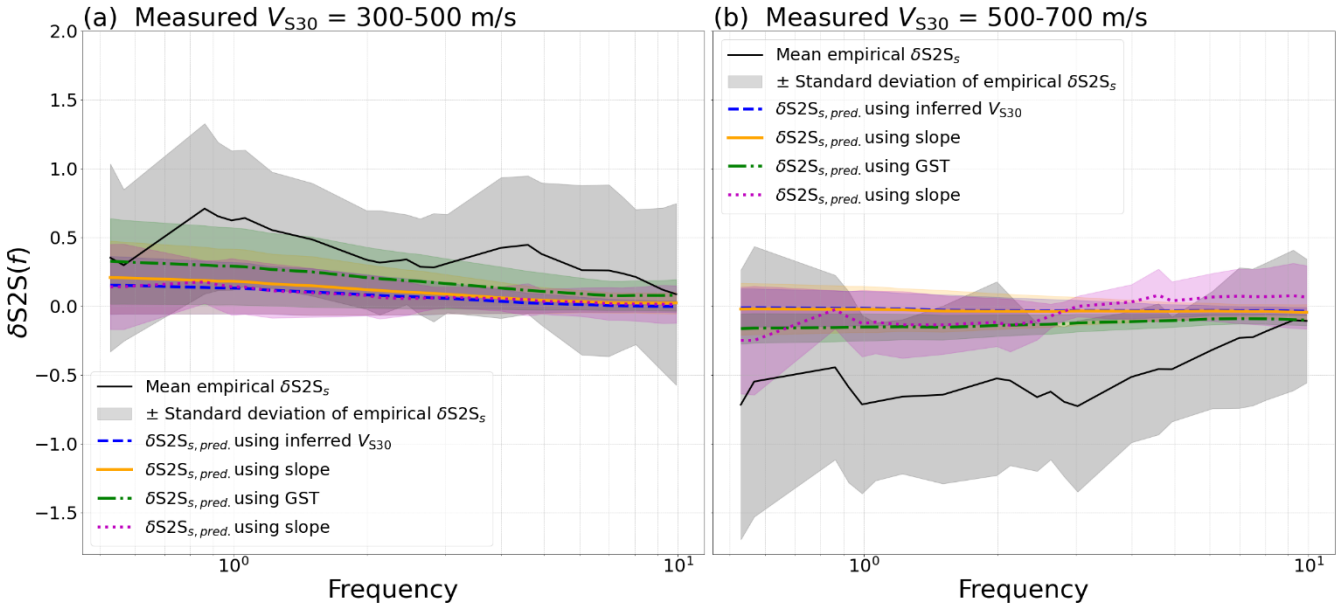


Figure 13: Empirical site amplification (solid black line) compared to predicted site amplification using the stations in Eastern-Turkey with measured V_{S30}-values, separated into medium stiff soil sites with (a) V_{S30} = 300 – 500 m/s and (b) stiff sites V_{S30} = 500 – 700 m/s, using inferred V_{S30} (dashed blue lines), slope (solid orange lines), geomorphological sediment thickness (dashed green lines) and geological era and slope (dotted magenta lines) at the stations.

Table 3: Range of location and site properties of the cities Antakya, Aleppo and Gaziantep shown in Fig. 11 and 14.

City	Longitude°	Latitude°	Inferred V_{S30} (m/s)	Slope (m/m)	GST (m)	Geological era
Antakya	36.14 - 36.18	36.19 - 36.22	290.0 - 873.0	0.0107 - 0.1694	2 - 16	Cenozoic
Aleppo	37.125 - 37.18	36.19 - 36.22	325.0 - 409.0	0.0109 - 0.0295	10 - 11	-
Gaziantep	37.36 - 37.4	37.05 - 37.08	275.0 - 350.0	0.0151 - 0.0192	0 – 2	Cenozoic

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

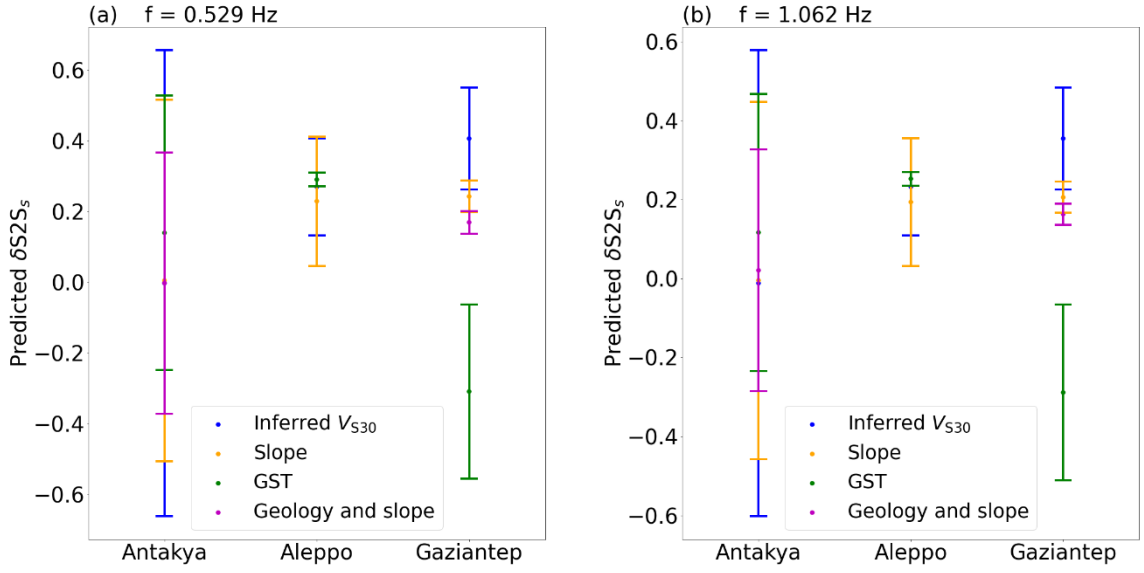


Figure 14: Predicted site-amplification ranges for the three cities Antakya, Aleppo and Gaziantep at (a) *f* = 0.529 Hz and (b) *f* = 1.062 Hz.

7 Conclusion

We have derived site-amplification models based on the geomorphological sediment thickness, as well as traditional site proxies like inferred V_{S30}, slope and geological era and slope combined, to test whether the geomorphological model for sediment thickness derived by Pelletier et al., (2016) can be used as an alternative site proxy. Although the predicted site-amplification maps based on the different proxies show similar trends, there are also noticeable differences. The differences in the site-amplification predictions based on the different proxies capture the epistemic uncertainty and emphasize the limitations of only using one proxy when modelling site amplification. This is especially important for smaller areas and on the level of cities where the impact of site amplification can be great. In this study we calculate the site amplification for Europe and the Middle-East but focus particularly on Eastern Turkey and Syria. The epistemic uncertainty of the predicted amplification in this area is large, especially for the city Gaziantep, which may impact the uncertainty of regional risk computations performed with proxy-based amplifications. As a measure of how well the site proxies capture the empirical site amplification we used the reduction in site-to-site variability. The results show that the site-amplification predictions based on geological era and slope combined cause the highest reduction while the prediction based on geomorphological sediment thickness causes a similar, but slightly larger, reduction in site-to-site variability than the traditional site proxies, inferred V_{S30} and slope. This result shows the value of including geology and geomorphology in prediction models for site amplification. Furthermore, the geological map used in this study is only available for Europe, while geomorphological sediment thickness is available globally and easily accessible. However, although the geomorphological sediment thickness has potential, further investigations and tests are needed before establishing it as an alternative to the much-used inferred V_{S30} model from Wald and

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Allen (2007), in particular in areas where inferred V_{S30} and slope are known to have a weak correlation with site amplification. Moreover, the correlation between the empirical site amplification and the site proxies all weakens above 3 Hz, which shows the need for models with higher resolution or including more local and shallow information. Our results therefore show the potential, of not only the geomorphological sediment thickness model, but also of other models for soil and sediment thickness from geomorphology and similar fields outside seismology and earthquake engineering. Finally, it is important to state that the site amplification models and maps developed in this study are made for the purpose of testing the different site proxies and showing the epistemic uncertainty related to using different proxies. Additionally, using inferred site proxies should only be done for regional seismic hazard studies of larger areas or when more detailed site parameters are missing.

Code and data availability

The empirical site amplification $\delta S2S_s$ and the coefficients for the proxy-based site amplification prediction models derived in this study, as well as a short code for computing and mapping the proxies and site-amplification predictions, are available on Zenodo: https://zenodo.org/record/8140143.

The V_{S30} dataset from Wald and Allen (2007) is available from the U.S. Geological Survey (USGS) and were downloaded from https://earthquake.usgs.gov/static/lfs/data/vs30/vs30.zip (last accessed 19.01.2023). The slope and geological map of Europe are available from the EFEHR seismic risk web-services (http://risk.efehr.org/site-model/) and can be downloaded from https://maps.eu-risk.eucentre.it/map/european-site-response-model-datasets/download (last accessed 23.09.2022), and https://nextcloud.gfz-potsdam.de/s/93ZR4ky8D4mDXb9 (last accessed 02.05.2022), respectively. The geomorphological sediment thickness from Pelletier et al. (2016) can be downloaded from https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1304 (last accessed 26.01.2021).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Appendix

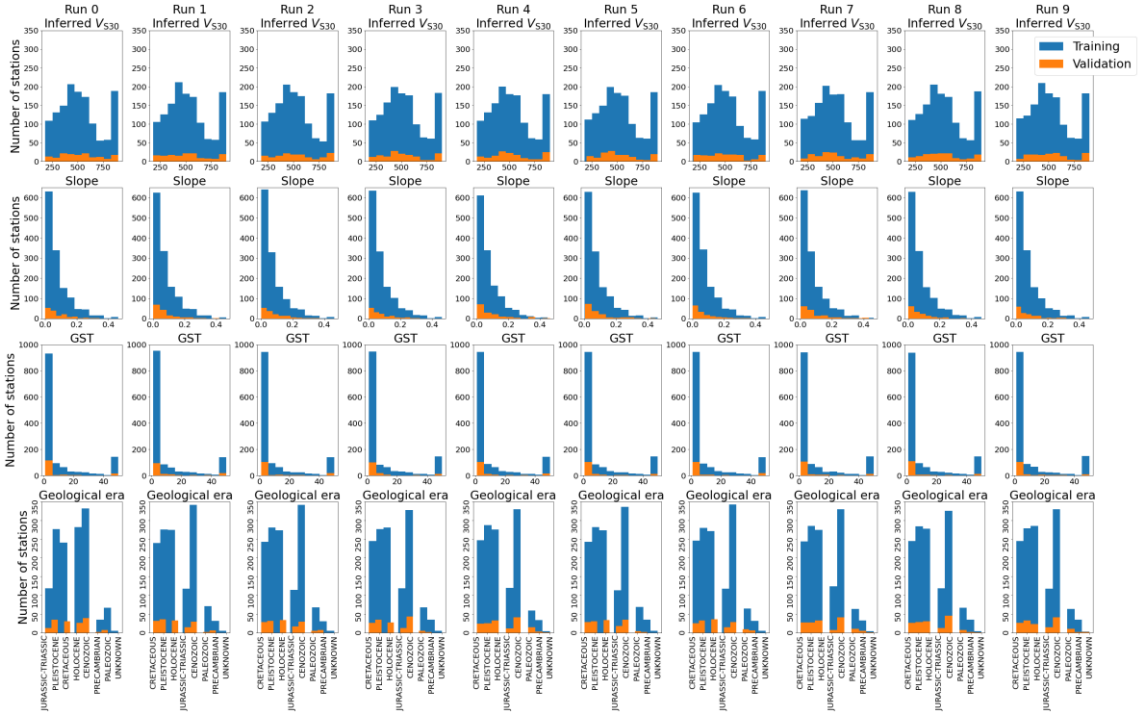


Figure A1: The distribution of inferred V_{S30} (top row), slope (second row), geomorphological sediment thickness (GST, third row) and geological era and slope (bottom row) for each 10 fold cross validation iteration using the 10-1 part training-set (blue) and the 1 part validation-set (orange).

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

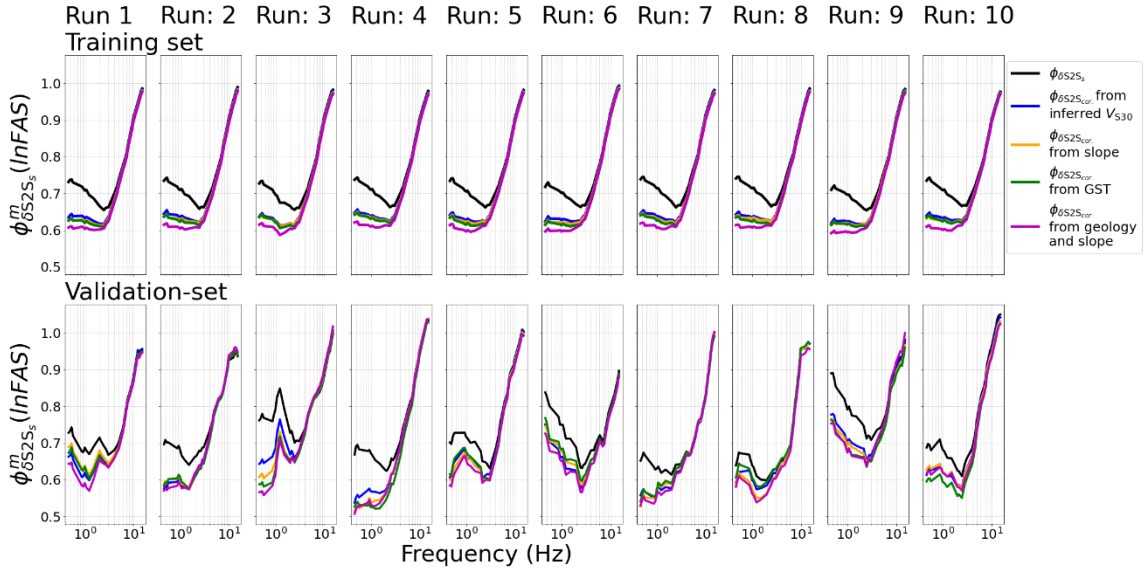


Figure A2: The site-to-site variability φ_{s2s} for all selected stations (black line) and the corrected site-to-site variability after subtracting the predicted site amplification using inferred V_{S30} ($\varphi_{s2s,cor.}(V_{S30})$, blue lines), slope ($\varphi_{s2s,cor.}(Slope)$, orange lines), geomorphological sediment thickness (GST) ($\varphi_{s2s,cor.}(GST)$, green lines) and geological era and slope ($\varphi_{s2scor.}(Geology\ and\ slope)$, magenta lines) from the empirical site amplification for each 10-fold cross validation iteration using the 10-1 part training-set (top row) and the 1 part validation-set (bottom row).

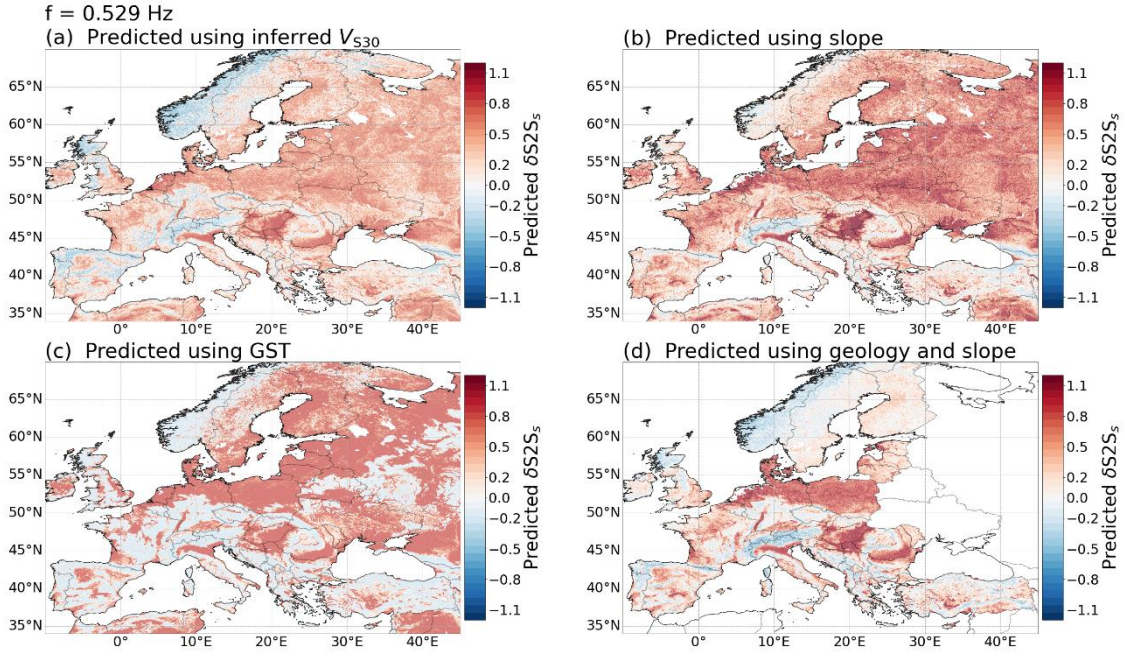


Figure A3: Predicted site amplification at f = 0.529 Hz for Europe using (a) inferred V_{S30} (b) slope, (c) geomorphological sedimentary thickness, and (d) geological era and slope.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

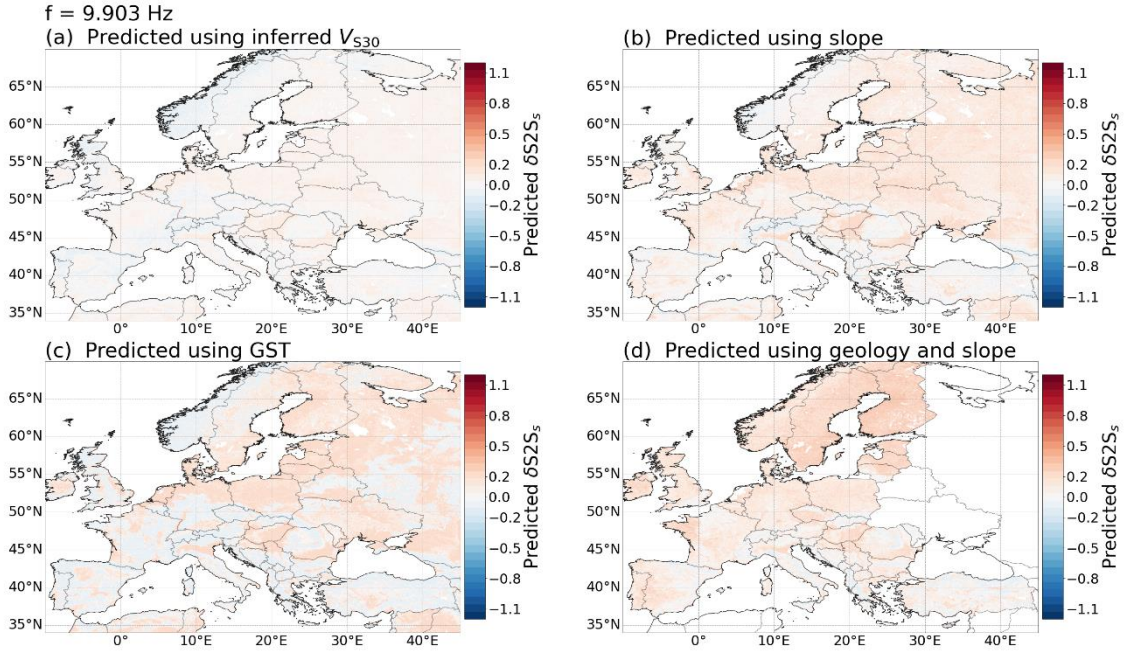


Figure A4: Predicted site amplification at f = 9.903 Hz for Europe using (a) inferred V_{S30} (b) slope, (c) geomorphological sedimentary thickness, and (d) geological era and slope.

485

Author contribution

Karina Loviknes: Conceptualization, methodology, data preparation, investigation, visualization, writing. Fabrice Cotton: Conceptualization, methodology, supervision, review, and editing. Graeme Weatherill: Conceptualization, data preparation, review, and editing.

490

Competing interests

Fabrice Cotton is a member of the editorial committee for this special issue of Natural Hazards and Earth System Sciences and therefore cannot be chosen as the editor in charge of the paper.

Acknowledgements

495 The authors are grateful to Jean Braun for valuable discussion and explanation on regolith. This research is funded by the European Commission, ITN-Marie Sklodowska-Curie New Challenges for Urban Engineering Seismology URBASIS-EU

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

project, under Grant Agreement 813137 and supported within the funding program "Open Access publication costs" Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 491075472.

References

Amemiya, T.: Tobit models: A survey, Journal of econometrics, 24, 3–61, https://doi.org/10.1016/0304-4076(84)90074-5, 1984.

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C.: Fitting linear mixed-effects models using lme4, J. Stat. Software, 67, 340 https://doi.org/10.18637/jss.v067.i01, 2015.

Bayless, J. and Abrahamson, N. A.: Summary of the BA18 ground-motion model for Fourier amplitude spectra for crustal earthquakes in California, Bulletin of the Seismological Society of America, 109, 2088–2105, https://doi.org/10.1785/0120190077, 2019.

Bergamo, P., Hammer, C., and Fäh, D.: On the relation between empirical amplification and proxies measured at Swiss and Japanese stations: Systematic regression analysis and neural network prediction of amplification, Bulletin of the Seismological Society of America, 111, 345 101–120, https://doi.org/10.1785/0120200228, 2021.

Bergamo, P., Hammer, C., and Fäh, D.: Correspondence between Site Amplification and Topographical, Geological Parameters: Collation of Data from Swiss and Japanese Stations, and Neural Networks-Based Prediction of Local Response, Bulletin of the Seismological Society of America, 112, 1008–1030, https://doi.org/10.1785/0120210225, 2022.

Bishop, C. M. and Nasrabadi, N. M.: Pattern recognition and machine learning, vol. 4, Springer, ISBN 978-0-387-31073-2, 2006.

Bora, S. S., Cotton, F., and Scherbaum, F.: NGA-West2 empirical Fourier and duration models to generate adjustable response spectra, Earthquake Spectra, 35, 61–93, https://doi.org/10.1193/110317EQS228M, 2019.

Crowley H, Dabbeek J, Despotaki V, Rodrigues D, Martins L, Silva V, Romão X, Pereira N, Weatherill G, Danciu L. European seismic risk model (ESRM20). EFEHR Technical Report. https://doi.org/10.7414/EUC-EFEHR-TR002-ESRM20, 2021

Cultrera, G., Cornou, C., Di Giulio, G., and Bard, P.-Y.: Indicators for site characterization at seismic station: recommendation from a dedicated survey, Bulletin of Earthquake Engineering, 19, 4171–4195, https://doi.org/10.1007/s10518-021-01136-7, 2021.

Derras, B., Bard, P. Y., and Cotton, F.: V_{S30}, slope, H_{800} and f_0: Performance of various site-condition proxies in reducing ground-motion aleatory variability and predicting nonlinear site response 4. Seismology, Earth, Planets and Space, 69, 133, https://doi.org/10.1186/s40623-017-0718-z, 2017.

Holbrook, W. S., Riebe, C. S., Elwaseif, M., L. Hayes, J., Basler-Reeder, K., L. Harry, D., Malazian, A., Dosseto, A., C. Hartsough, P., and W. Hopmans, J.: Geophysical constraints on deep weathering and water storage potential in the Southern Sierra Critical Zone Observatory, Earth Surface Processes and Landforms, 39, 366–380, https://doi.org/10.1002/esp.3502, 2014.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Hollender, F., Roumelioti, Z., Maufroy, E., Traversa, P., and Mariscal, A.: Can we trust high-frequency content in strong-motion database signals? Impact of housing, coupling, and installation depth of seismic sensors, Seismological Research Letters, 91, 2192–2205, https://doi.org/10.1785/0220190163, 2020.

Koller, M.: robustlmm: an R package for robust estimation of linear mixed-effects models, Journal of statistical software, 75, 1–24, 355 https://doi.org/10.18637/jss.v075.i06, 2016.

Kotha, S. R., Cotton, F., and Bindi, D.: A new approach to site classification: mixed-effects ground motion prediction equation with spectral clustering of site amplification functions, Soil Dynamics and Earthquake Engineering, 110, 318–329, https://doi.org/10.1016/j.soildyn.2018.01.051, 2018.

Kotha, S. R., Weatherill, G., Bindi, D., and Cotton, F.: A regionally-adaptable ground-motion model for shallow crustal earthquakes in 360 Europe, Bulletin of Earthquake Engineering, 18, 4091–4125, https://doi.org/10.1007/s10518-020-00869-1, 2020.

Kotha, S. R., Bindi, D., and Cotton, F.: A regionally adaptable ground-motion model for Fourier amplitude spectra of shallow crustal earthquakes in Europe, Bulletin of Earthquake Engineering, pp. 1–30, https://doi.org/10.1007/s10518-021-01255-1, 2022.

Lanzano, G., Sgobba, S., Luzi, L., Puglia, R., Pacor, F., Felicetta, C., D'Amico, M., Cotton, F., and Bindi, D.: The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics, Bulletin of Earthquake Engineering, 17, 561–582, 365, https://doi.org/10.1007/s10518-018-0480-z, 2019.

Lemoine, A., Douglas, J., and Cotton, F.: Testing the applicability of correlations between topographic slope and V_{S30} for Europe, Bulletin of the Seismological Society of America, 102, 2585–2599, https://doi.org/10.1785/0120110240, 2012.

Li, M., Rathje, E. M., Cox, B. R., and Yust, M.: A Texas-specific V_{S30} map incorporating geology and V_{S30} observations, Earthquake Spectra, 38, 521–542, https://doi.org/10.1177/87552930211033622, 2022.

Loviknes, K., Kotha, S. R., Cotton, F., and Schorlemmer, D.: Testing nonlinear amplification factors of ground-motion models, Bulletin of the Seismological Society of America, 111, 2121–2137, https://doi.org/10.1785/0120200386, 2021.

Luzi, L., Lanzano, G., Felicetta, C., D'Amico, M., Russo, E., Sgobba, S., and Pacor, F.: ORFEUS Working Group 5: Engineering Strong Motion Database (ESM)(Version 2.0), Istituto Nazionale di Geofisica e Vulcanologia (INGV) [data set], https://doi.org/10.13127/ESM.2, 2020.

Melgar, D., Taymaz, T., Ganas, A., Crowell, B. W., Öcalan, T., Kahraman, M., Tsironi, V., Yolsal-Çevikbil, S., Valkaniotis, S., Irmak, T. S., et al.: Sub-and super-shear ruptures during the 2023 Mw 7.8 and Mw 7.6 earthquake doublet in SE Türkiye, https://doi.org/10.31223/X52W9D, 2023.

Michelini, A., Faenza, L., Lauciani, V., and Malagnini, L.: ShakeMap implementation in Italy, Seismological Research Letters, 79, 688–697, https://doi.org/10.1785/gssrl.79.5.688, 2008.

Mori, F., Mendicelli, A., Moscatelli, M., Romagnoli, G., Peronace, E., and Naso, G.: A new V_{S30} map for Italy based on the seismic 380 microzonation dataset, Engineering Geology, 275, 105745, https://doi.org/10.17632/8458tgzc73.1, 2020.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Pelletier, J. D., Broxton, P. D., Hazenberg, P., Zeng, X., Troch, P. A., Niu, G.-Y., Williams, Z., Brunke, M. A., and Gochis, D.: A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling, Journal of Advances in Modeling Earth Systems, 8, 41–65, https://doi.org/10.1002/2015MS000526, 2016.

Petersen, G. M., Büyükakpinar, P., Vera Sanhueza, F. O., Metz, M., Cesca, S., Akbayram, K., Saul, J., and Dahm, T.: The 2023 Southeast Türkiye Seismic Sequence: Rupture of a Complex Fault Network, The Seismic Record, 3, 134–143, https://doi.org/10.1785/0320230008, 2023.

Silva, V., Amo-Oduro, D., Calderon, A., Costa, C., Dabbeek, J., Despotaki, V., Martins, L., Pagani, M., Rao, A., Simionato, M., et al.: Development of a global seismic risk model, Earthquake Spectra, 36, 372–394, https://doi.org/10.1177/8755293019899953, 2020.

Stewart, J. P., Afshari, K., and Goulet, C. A.: Non-ergodic site response in seismic hazard analysis, Earthquake Spectra, 33, 1385–1414, 390, https://doi.org/10.1193/081716eqs135m, 2017.

Thompson, E., Wald, D. J., and Worden, C.: A V_{S30} Map for California with Geologic and Topographic Constraints, Bulletin of the Seismological Society of America, 104, 2313–2321, https://doi.org/10.1785/0120130312, 2014.

Thompson, E. M. and Wald, D. J.: Uncertainty in V_{S30}-based site response, Bulletin of the Seismological Society of America, 106, 453–463, https://doi.org/10.1785/0120150214, 2016.

Thompson, E. M., Baise, L. G., Kayen, R. E., Tanaka, Y., and Tanaka, H.: A geostatistical approach to mapping site response spectral amplifications, Engineering geology, 114, 330–342, https://doi.org/10.1016/j.enggeo.2010.05.010, 2010.

Tobin, J.: Estimation of relationships for limited dependent variables, Econometrica: journal of the Econometric Society, pp. 24–36, https://doi.org/10.2307/1907382, 1958.

Trifunac, M. D.: Site conditions and earthquake ground motion–A review, Soil dynamics and earthquake engineering, 90, 88–100, 400, https://doi.org/10.1016/J.SOILDYN.2016.08.003, 2016.

Vilanova, S. P., Narciso, J., Carvalho, J. P., Lopes, I., Quinta-Ferreira, M., Pinto, C. C., Moura, R., Borges, J., and Nemser, E. S.: 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, 322–337, https://doi.org/10.1785/0120170213, 2018.

Wald, D. J. and Allen, T. I.: Topographic slope as a proxy for seismic site conditions and amplification, Bulletin of the Seismological Society of America, 97, 1379–1395, https://doi.org/10.1785/0120060267, 2007.

Weatherill, G., Kotha, S. R., and Cotton, F.: Re-thinking site amplification in regional seismic risk assessment, Earthquake Spectra, 36, 274–297, https://doi.org/10.1177/8755293019899956, 2020.

Weatherill, G., Crowley, H., Roullé, A., Tourlière, B., Lemoine, A., Gracianne, C., Kotha, S. R., and Cotton, F.: Modelling site response at regional scale for the 2020 European Seismic Risk Model (ESRM20), Bulletin of Earthquake Engineering, 21, 665–714, https://doi.org/10.1007/s10518-022-01526-5, 2023.

https://doi.org/10.5194/egusphere-2023-1370
Preprint. Discussion started: 18 July 2023

Zhu, C., Weatherill, G., Cotton, F., Pilz, M., Kwak, D. Y., and Kawase, H.: An open-source site database of strong-motion stations in Japan: K-NET and KiK-net (v1. 0.0), Earthquake Spectra, 37, 2126–2149, https://doi.org/10.1177/8755293020988028, 2021.

Zhu, C., Cotton, F., Kawase, H., Haendel, A., Pilz, M., and Nakano, K.: How well can we predict earthquake site response so far? Site-specific approaches, Earthquake Spectra, 38, 1047–1075, https://doi.org/10.1177/87552930211060859, 2022.