ORIGINAL ARTICLE

Modelling site response at regional scale for the 2020 European Seismic Risk Model (ESRM20)

Graeme Weatherill[1] · Helen Crowley[2] · Agathe Roullé[3] · Bruno Tourlière[4] ·
Anne Lemoine[3] · Cécile Gracianne[3] · Sreeram Reddy Kotha[5] ·
Fabrice Cotton[1]

Abstract

Quantitative estimation of seismic risk over a region requires both an underlying probabilistic seismic hazard model and a means to characterise shallow site response over a large scale. The 2020 European Seismic Risk Model (ESRM20) builds on the 2020 European Seismic Hazard Model (ESHM20), requiring additional information to firstly parameterise the local site condition across all of Europe, and subsequently determine its influence on the prediction of seismic ground motion. Initially, a harmonised digital geological database for Europe is compiled, alongside a model of topographic/bathymetric elevation and a database of 30 m averaged shearwave velocity measurements (V_{S30}), in order to produce separate 30 arc-second maps of inferred V_{S30} based on topography and on geology. We then capitalise on a large database of seismic recording stations in Europe for which site-to-site ground motion residuals ($\delta S2S_S$) have been determined with respect to the shallow crustal ground motion model used in the ESHM20. These residuals allow us to incorporate site amplification functions into the European GMM calibrated upon either observed or inferred V_{S30}, or on the European geology and topography models. We present the resulting pan-European seismic site amplification model and assess its impact on seismic hazard and risk compared against other approaches. The new site amplification model fulfils the requirements of the ESRM20 and, providing uncertainty is fully propagated, yields estimates of seismic hazard and risk at a large space scale that may be comparable to other methods often applied at local/urban scale where better-constrained site information is available.

Keywords Site response · Seismic risk · Geology · Aleatory/epistemic uncertainty · Europe · ESRM20

✉ Graeme Weatherill
graeme.weatherill@gfz-potsdam.de

Extended author information available on the last page of the article

1 Introduction

The characterisation of strong shaking and its likelihood of occurrence across a region forms the primary point of connection between the seismic hazard model (relating to the location, nature and probability of occurrence of strong shaking) and the risk model (addressing distribution of exposed building and the probability of observing given levels of damage and loss when subject to strong shaking). Building upon recent databases of ground motion in Europe (Lanzano et al. 2019) and current developments in the field of ground motion modelling, the 2020 European Seismic Hazard Model (ESHM20) has developed a comprehensive ground motion model (GMM) that characterises the expected level of shaking, its aleatory uncertainty and its regional variability (Kotha et al. 2020) and adopts a *scaled backbone* GMM logic tree to represent the epistemic uncertainty (Weatherill et al. 2020a; Weatherill and Cotton 2020).

The ESHM20 forms the seismic hazard input for the accompanying 2020 European Seismic Risk Model (ESRM20), which combines this information with comprehensive models of building exposure (residential, commercial and industrial) and vulnerability in order to deliver fully probabilistic assessments of seismic risk in terms of average annual loss (AAL) and loss exceedance curves and maps (Crowley et al. 2021). The ESRM20 sets a new objective for the seismic hazard calculation in addition to those of the ESHM20, and that relates to the role of seismic site condition. The key target of the ESHM20 is the definition of seismic hazard in terms of probability of exceedance of ground motion with respect to a Eurocode 8 Class A reference rock site, which in this case is assumed to have a 30 m averaged shearwave velocity, V_{S30}, of 800 m/s, and thus a depth to the 800 m/s V_S layer (H_{800}) of 0 m. Serving primarily the needs of defining seismic input in a manner that meets the requirements of Eurocode 8, the development of ground motion models for the ESHM20 have so far focussed on the influence of the seismic source and travel path on ground motion, in addition to their uncertainties and regional variabilities insofar as they influence the seismic hazard on the reference rock (Kotha et al. 2020; Weatherill et al. 2020a; Weatherill and Cotton 2020). The material and geometry of the superficial geological layers strongly modify the strong ground motion, however, both in terms of its amplitude, duration and wavefield composition, as can special topographic features. For the estimation of losses to buildings and/or infrastructure it is fundamental to account for these modifications, which can drastically increase the seismic hazard at the surface on which the structures are located. Site effects are therefore a key parameter in seismic risk estimation.

Though a substantial body of scientific literature can be found that addresses site-specific characterisation of ground motion and the optimal measures for defining the shallow site condition to improve prediction of local amplification (e.g. Derras et al. 2017; Bergamo et al. 2021; Zhu et al. 2022), the characterisation of seismic risk at national or regional scale presents a different set of challenges. In the current manuscript we detail the specific challenges for characterisation of such site response at regional scale, the additional uncertainties that must be accounted for and how these have been formally integrated into the approach for site characterisation within the ESRM20. In the following chapter we outline the strategies that have been adopted for regional scale characterisation of site response in previous applications of seismic risk at national scale, such those national and regional models comprising the recent global seismic risk model produced by the Global Earthquake Model (Silva et al. 2020 and references therein), and discuss whether these address the needs of the ESRM20. The subsequent chapters then focus on the development of the site characterisation model for Europe, beginning with the compilation and harmonisation

of relevant data sets before moving onto the construction of pan-European maps of inferred V_{S30}. The construction of the input data sets is only part of the complete process of site characterisation, however. The manuscript therefore proceeds to explain how the database of strong motion data are combined with the pan-European maps of site properties in order to calibrate the site amplification terms of the European ground motion models adopted within ESHM20. We outline a mixed data driven approach that uses the relations between observed station-to-station ground motion residuals ($\delta S2S_S$) and regional scale mappable site proxies to calibrate the model parameters, ensuring that the variability in the resulting surface ground motion accounts for additional uncertainty associated with the use of regionally mapped site properties in the amplification models rather than detailed site-specific studies. Using selected localities in Europe we then demonstrate how the regional scale site amplification approach adopted for the ESRM20 influences the resulting seismic hazard and risk curves, and how this compares with other approaches that might typically be applied when more detailed microzonation data is available. Finally, we outline the implications of our site response modelling strategy for future seismic hazard and risk analysis in Europe, including its adaptability and the role that improved site characterisation can play in reducing uncertainties.

2 Modelling site response at regional scale—current practices and limitations

As site effects are mainly driven by superficial geology and topography conditions, Digital Elevation Models (DEMs) and geological maps are fundamental data sets that have been widely used to assess site response mapping at regional scale (e.g., Wald and Allen 2007; Thompson et al. 2014; Forte et al. 2019; Chen et al. 2021; Silva et al. 2020). Recent papers have developed alternative methodologies based on both geological and DEM information to derive site effects proxies (mainly the V_{S30} parameter and/or the building code soil classes) at regional scale. Those methodologies consist of (1) defining geological and/or elevation data-based units (e.g., slope or lithology) (2) deriving the statistical distribution of site effects proxies (e.g., V_{S30}) for each unit from data sets of in situ observations, (3) building correlations between spatial units and site effects proxies to obtain a site condition model on the complete area.

A first category of site response models integrates topographic gradient (slope) or other DEM-derived elements (roughness, convexity, elevation, distance from mountain). These include the widely used Wald and Allen (2007) model, based only on slope criteria, or the Iwahashi and Pike (2007) model based on multiple geomorphological criteria. Site condition models from extrapolations of surface geology data (lithology and/or stratigraphy) were also proposed for different countries, such as that of Wills and Clahan (2006) in the United States, Lee and Tsai (2008) in Taiwan, McPherson and Hall (2013) in Australia, Foster et al. (2019) for New Zealand, Vilanova et al. (2018) for Portugal, Di Capua et al. (2016) and Forte et al. (2019) in Italy and Panzera et al. (2021) for Switzerland. Finally, some hybrid methods combine different information, such as terrain classes, lithostratigraphic criteria and DEM information (e.g., Stewart et al. 2014; Seyhan et al. 2014; or Ahdi et al. 2017). The reliability of such methodologies is strongly dependent on the correlations between classified units and site conditions proxies databases. In spite of the difficulties and limitations in predicting a local-scale effect at regional or continental scale (e.g., Lemoine et al. 2012; Seyhan et al. 2014), those methodologies are widely used to address

territorial planning and the generation of real-time ShakeMaps (Thompson et al. 2014), as well as previous regional and global scale seismic risk analyses (Silva et al. 2020).

Though the use of databases of inferred V_{S30} is commonplace, whether derived from topography, geology or some hybrid of the two, some perspective is needed as to what such models can deliver and what their implications are for characterising uncertainty in seismic risk analysis. When measured directly at a site from surface wave dispersion, seismic reflection or borehole logs, V_{S30} is merely a property of the upper layers of the soil. In reality, the local amplification of seismic waves at a site are influenced not only by the material properties of the upper 30 m, but rather from the entire soil column, and in some cases the surrounding subsurface topography and geology. For the purposes of empirical ground motion modelling, linear site amplification at a given spectral period of oscillation is related to V_{S30} primarily though correlation between the residual of the observed ground motion at a site, once corrected for the source and path effects, and the V_{S30} values of the sites in question. Though the correlations have a clear seismological and geotechnical basis, V_{S30} is itself a *proxy* for amplification. This has several consequences: the first is that there exists variability in the prediction of amplification that emerges from differences in the full site profile even among sites with similar, even identical, V_{S30}. In the case of ergodic seismic hazard analysis, this site-to-site variability is a constituent of the entire within event-variability and is modelled as a normal distribution with zero mean and a standard deviation of ϕ_{S2S}. The second consequence is that the degree of amplification at a given period resulting from the correlations with V_{S30} may not result just from the influence of shallow soil layers on the amplification but on the implicit correlations between V_{S30} and other properties of the soil column that may be present within the data set, such as basin depth or proximity to the basin edge (Weatherill et al. 2020b).

For modelling the influence of the local site conditions on ground motion in regional scale seismic risk, not only do we lack full profiles of site properties, we also lack observations of V_{S30} for all of the sites across the target region of interest. Instead, we have to rely on the values of V_{S30} inferred from the correlations between observed V_{S30} measurements and regionally mappable quantities (e.g., slope, geology etc.) described previously. It is important not to lose sight of the fact that our ultimate target for seismic risk analysis is not the prediction of V_{S30}, but rather the prediction of ground motion at the Earth's surface. In this sense, topographically and/or geologically *inferred* V_{S30} is *a proxy of a proxy* for local site amplification. Inferred V_{S30} values not only have a considerable amount of uncertainty, there are also many geological and/or geomorphological environments in which these correlations between V_{S30} and mappable proxy parameters break down entirely (Lemoine et al. 2012).

Why is this then a problem for seismic risk analysis? The limitations of the use of *inferred* V_{S30} arise not necessarily from the increased uncertainty in the resulting site amplification, but rather in the manner this uncertainty is treated in the seismic risk calculations. When developing the ground motion models, the databases of observed strong motions will likely contain a substantial proportion of records from stations for which the V_{S30} is measured directly, in some cases along with other parameters such as depth to bedrock, horizonal to vertical spectral ratio etc. The remainder of the V_{S30} values of the sites may come from inferred properties such as the topographic slope or seismic design code site class. The resulting amplification model, $f_{SITE}(T, V_{S30}, \theta)$ and its corresponding site-to-site variability, $\phi_{S2S}(T)$, will most likely reflect a predominantly *measured* V_{S30} site condition. As the majority of recent GMMs are designed with a view toward site-specific application, in which the V_{S30} (and other properties) of the target site are known, the ground motion modeller may even wish to limit the calibration of f_{SITE} exclusively to the subset

of records with measured V_{S30}. But in doing so there can be an inconsistency between the assumption of *measured* V_{S30} in the development of a GMM and its application in regional scale seismic risk analysis for which the V_{S30} values at the target sites are *inferred* from an uncertain proxy, as described in Weatherill et al. (2020b).

There are several possible ways to integrate the additional uncertainty from the use of *inferred* V_{S30} into the seismic risk calculation. One possibility is to address the issue in the process of GMM development using either a generalized maximum likelihood approach (Gehl et al. 2011) or Bayesian regression (Stafford 2014; Kuehn and Abrahamson 2018). These methods make the regression more robust to the uncertainty on the V_{S30} in the underlying data set and reducing the influence of the uncertain V_{S30} in the resulting ϕ_{S2S}. To date, however, very few GMMs have been developed in this fashion and though this assists in better calibrating the models when *inferred* values are mixed together with measured V_{S30} values, it does not explicitly propagate the additional uncertainty into the risk calculation.

Another possible approach would be to model this additional uncertainty as a conventional epistemic uncertainty, adding additional branches to the ground motion logic tree to account for the range of potential V_{S30} values for each site while retaining a ϕ_{S2S} in the ground motion model that reflects the observed V_{S30} case. The consequence of such a transfer from aleatory to epistemic uncertainty would be a reduction of the mean seismic hazard at a site (an outcome of a reduced ϕ_{S2S}), while at the same time ensuring that the quantiles of the hazard estimates from the logic tree reflect better the "true" epistemic uncertainty, which would be reduced over time with the acquisition of better site information. This approach is more aligned with the state-of-the-art in site-specific seismic hazard analysis, where it is often taken further by reducing the aleatory uncertainty in the ground motion model for the prediction of hazard on rock using partially or fully non-ergodic approaches that reduce or even remove the station-to-station variability ϕ_{S2S} (e.g., Stewart et al. 2017; Rodriguez-Marek et al. 2014; 2021).

For application across large regions, however, we encounter two problems. The first problem is one of computation, as moving this uncertainty into the ground motion logic tree requires another set of branches, thus increasing the calculation efforts. The second is that when applying the epistemic uncertainties in the site response to thousands of sites at the same time we encounter the problem of site-to-site correlation in epistemic uncertainty.

To illustrate these issues, consider an approach in which one adopts a GMM whose amplification is based on a measured V_{S30} and corresponding ϕ_{S2S} that reflects this the site-to-site variability based on exclusively measured V_{S30} values. To apply this at regional scale one would need to use regionally mappable proxies to infer a distribution of V_{S30} at each site, which could be characterised by lognormal distribution with a median ($\ln \overline{V_{S30}}$) and standard deviation ($\sigma_{\ln V_{S30}}$). To evaluate this epistemic uncertainty, we would add an additional set of branches into the GMM logic tree to approximate the distribution $\mathcal{N}\left(\ln \overline{V_{S30}}, \sigma_{\ln V_{S30}}\right)$ by N_{BR} branches and their respective weights (e.g., Miller and Rice 1983). To execute the logic tree in a seismic hazard or risk calculation, however, we would either have to treat each of the N_{SITES} independently, and thus generate an additional $N_{BR}^{N_{SITES}}$ end-branches to the logic tree (i.e. enumerating or sampling all permutations of the branches from all sites). Alternatively we would have to assume perfect correlation in the epistemic uncertainty for all sites, i.e. $V_{S30} = \ln \overline{V_{S30}} + \varepsilon_1 \cdot \sigma_{\ln V_{S30}}$, $\ln \overline{V_{S30}} + \varepsilon_2 \cdot \sigma_{\ln V_{S30}}, \ldots \ln \overline{V_{S30}} + \varepsilon_{N_{BR}} \cdot \sigma_{\ln V_{S30}}$, where the number of standard deviations $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{N_{BR}}$ is fixed for all N_{SITES} simultaneously. The first approach is not feasible to evaluate, while the second would effectively mean that for any given branch all sites would

be assumed to take a V_{S30} value corresponding to the same number of standard deviations above or below the median. This assigns disproportionately high weights to extreme outlier scenarios in which *all sites* would have higher or lower V_{S30} than the median. Perfect correlation in the site branches would vastly overestimate the resulting epistemic uncertainty distribution of the seismic risk curves, which in turn could affect the estimation of the mean risk.

In reality, neither true independence nor perfect correlation is correct, and there would likely be a degree of spatial correlation in the ε term that should decay with distance between sites. At present we have no such model of the spatial correlation in the $\varepsilon \cdot \sigma_{\ln V_{S30}}$ terms, nor any practical means to implement this into the seismic risk calculation. Therefore, we cannot simply expand the current state-of-the-art approaches for site-specific seismic hazard and risk analysis to apply to thousands of sites at the same time.

The third possible approach to managing uncertainty from site properties in regional scale risk calculations is to define separate calibrations of both f_{SITE} and ϕ_{S2S} terms of the ground motion model to reflect the case in which the site property is *observed* (via direct measurement) or *inferred* from a mappable proxy. In this case the expected amplification and the site-to-site variability reflect the ability of the proxy to predict the amplification observed in the ground motion data. This approach is not unprecedented in ground motion modelling, as the NGA West 2 GMMs of Abrahamson et al. (2014) and Chiou and Youngs (2014) also adapt their aleatory uncertainty term to reflect the increased uncertainty in the case when V_{S30} is inferred from proxy rather than measured directly. The effect of this approach is to transfer the epistemic uncertainty of the poorly known site properties into the aleatory variability of the ground motion model. This approach is not only practical, as it can be easily implemented in modern seismic hazard and risk calculation software and does not substantially increase the calculation time nor introduce artificial site-to-site correlations.

This brings us to the strategy for modelling site response in the seismic risk calculation at European scale. Initially we work with the V_{S30} proxy, widely used in the civil engineering community. To develop a pan-European model of V_{S30} we apply two models: a Digital Elevation Model (DEM) based methodology following the approach of Wald and Allen (2007), whose reliability was tested during the previous European Seventh Framework Program (FP7) project Seismic Hazard Harmonization in Europe (SHARE) (Lemoine et al. 2012), and a geology-based methodology built around the approach adopted in a recent V_{S30} model developed in Portugal (Vilanova et al. 2018). With the data sets established we then explore the observed station-to-station ground motion model residuals ($\delta S2S_S$) for more than a thousand stations across Europe and their correlation to *observed* $V_{S30,}$ *inferred* V_{S30} and the mapped information (slope and geology) that is used in their construction. These correlations enable us to calibrate f_{SITE} and ϕ_{S2S} reflecting either the *observed* (or *measured*) V_{S30} condition and the *inferred* condition. We can also go a step further, however, and define f_{SITE} and ϕ_{S2S} directly upon the slope and geology data and, in doing so, avoid invoking the proxy of *inferred* V_{S30} but rather allowing the scaling between slope and site amplification in the GMM to reflect differences in the geological setting of each site. The outcome of these combined activities is a model that allows us to predict ground motion at the surface at a 30 arc-second resolution that covers all target regions of the ESRM20 in Europe, taking into account the local site condition and propagating the uncertainty that emerges from the use of mappable data into the probabilistic seismic risk calculation.

3 Compilation and harmonisation of European datasets for site response modelling

3.1 Development of a harmonised surface geology map for Europe

A geological map informs on the nature (lithology) and age (stratigraphy) of outcropping formations, without specific information on their thickness. At continental scale, superficial and/or weathered formations are generally not reported, except for alluvial deposits. As there does not exist a harmonized geological map at European scale describing superficial formations for the full extent of the region required for the ESRM20 risk model, we created a composite map derived from three existing maps:

- The geological map at 1:1,500,000 from the European project ProMine (http://promine.gtk.fi/, Cassard et al. 2015). It was built from a compilation and homogenization of the national geological maps of involved countries (with scales varying from 1:500,000 to 1:1,000,000). One of the objectives of the ProMine project was to develop the first pan-European GIS-based database on mineral resources. In this framework, it provided a coherent litho-stratigraphic map at European scale based on the main geological events of the Caledonian, Varsican and Alpine orogenies. This map, originally built for mineral resource exploitation, was clearly focused on the ante-cenozoic formations and consequently missed details on the recent sedimentary deposits that cover a main part of North-Eastern Europe. This map covers the whole ESRM20 extent except Iceland, and the islands of Azores, Canaries, and Madeira.
- The geological map at 1:1,000,000 from OneGeologyEurope, available from the European Geological Data Infrastructure (EGDI) services (http://www.europe-geology.eu/). It provides a pan-European geological map of superficial formations. It was delivered from the geological surveys of 21 countries using INSPIRE data models focusing on outcropping superficial formations and giving information both on their lithology and stratigraphy. Unfortunately, it too only covers partially the ESRM20 extent since 15 countries are missing, amongst which are Austria, Greece, Switzerland, Iceland and most of the Balkans countries.
- The bedrock geological map of Iceland at 1:600,000 available from the Icelandic Institute of Natural History (Johannesson 2014). It provides a classification of bedrock on the basis of its age, type, and composition (https://en.ni.is/resources/publications/maps/geological-maps#&gid=null&pid=1) combined with the description information (in English) given by the open interactive geological map of Iceland available at: http://jardfraedikort.is

The harmonisation of geological data from the aforementioned sources followed five steps:

1. *Synthetic lithological coding of ProMine geological map:* all the lithologies described in ProMine dataset were reprocessed to produce a simplified lithological coding based on the nature of the geological formations. If two lithologies were present in the descriptions available in ProMine DB, only the first one, considered as dominant, was kept.
2. *Improvement of lithology coding through OneGeology data:* for the recent sedimentary deposits (represented by SDMT code), ProMine DB information was not sufficiently
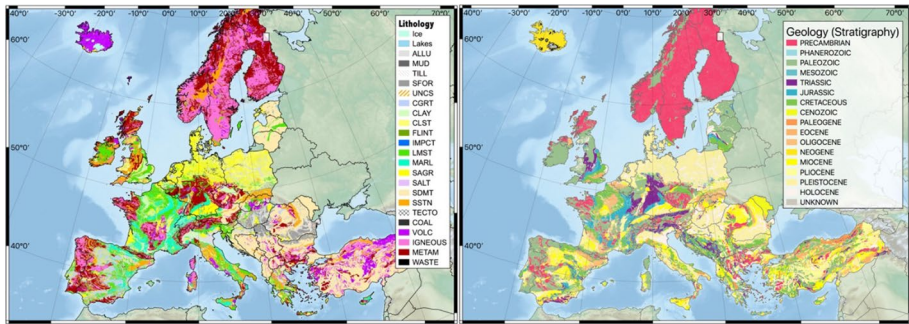
Fig. 1 Simplified lithological map (left) and stratigraphic map (right) of Europe. See "Appendices A" and "B" for the complete description of the lithological and stratigraphic codes

detailed. Instead, it was substituted by the OneGeology information, mainly in Spain and North-Eastern Europe. This step also permitted the addition of lithological information in the islands of Azores, Canary, and Madeira. The junctions between the two maps are satisfactory at 1:1,500,000 but could present some differences of interpretation locally.

3. *Completion of the lithological information for Iceland* through the Icelandic bedrock geological map described above. The final lithological information corresponds to 24 different codes (including lakes) detailed in "Appendix A" Table 3.

4. *Output of stratigraphic information for each lithological polygon.* We decided to keep the information on geological systems for Quaternary and Tertiary formations, which are less consolidated materials and therefore prone to produce site effects, and to keep only eras information for more ancient geological formations. The final stratigraphic information corresponds to 9 codes detailed in "Appendix A" Table 4.

5. *Production of the final litho-stratigraphic information* through the combination between the lithological codes and stratigraphic codes described above. The resulting harmonised geological database is shown in Fig. 1, both in terms of lithological and stratigraphic classifications.

3.2 Topography and bathymetry data

In order to be consistent with Wald and Allen (2007)'s method, we computed the topographic slopes with the same DEM resolution of 30 arc-seconds. Allen and Wald (2009) tested the use of higher resolution DEMs that require adapting the correlation between the slope distribution and site classification. For coherency with the continental scale of our work and with the resolution of European geological data (up to 1:1,500,000), however, we decided to keep the resolution of 30 arc-seconds. To avoid problems in the application of Wald and Allen (2007)'s method in coastal areas (highlighted in Lemoine et al. 2012), we computed topographic slopes from the 2014 General Bathymetric Chart of the Oceans grid (GEBCO_2014, https://www.gebco.net/). GEBCO_2014 is a continuous topographic and bathymetric model with spatial resolution 30 arc-seconds. Land data of GEBCO_2014 are 1 km average elevations determined from the SRTM30 gridded DEM, whereas the ocean portion of the grid was developed from various bathymetric compilations (Weatherall et al. 2015).

3.3 Database of V_{S30} observations

In order to build a European site amplification model, we collected as many publicly available data as possible on the site characterization of seismological stations (with measured V_{S30} values). The different contributions came from (see Fig. 2):

- The Engineering Strong Motion (ESM) database describing site characterization for European Strong Motion stations (Luzi et al. 2020 and references cited therein) with 469 measured V_{S30} values
- Database of V_{S30} measurements compiled during the SHARE project (Yenier et al. 2010; Lemoine et al. 2012) including specific site characterisation data from Switzerland [ETHZ] (Fäh and Huggenberger 2006; Fäh et al. 2007; Havenith et al. 2007; Poggi 2011; ETHZ, 2015) and France [labelled BRGM] (Roullé et al. 2010)
- Portuguese data used in the geologically based V_{S30} site-condition model of Vilanova et al. (2018)
- Dutch data from the site characterization of Groningen gas field (Noorlandt et al. 2018)
- Greek data used for inferring V_{S30} values from geologic- and terrain-base proxies (Stewart et al. 2014)
- Turkish data from the National Strong Motion project of Turkey and from a microzonation study on the city of Istanbul (Yilmaz et al. 2014)
- In addition to the site observations, V_{S30} from a high-resolution 3D velocity model of Cologne and surroundings from Pilz et al. (2021) is also included (not shown in Fig. 2)

The final database contains 1626 measured V_{S30} values across Europe (Fig. 2) and its corresponding distribution with respect to Eurocode 8 site class is shown Fig. 3. The V_S profile measurement methods used for estimating the V_{S30} value depend mainly on the operator responsible for the acquisition. The distribution of V_{S30} values in terms of acquisition methods shows that most of the data come from active seismic acquisition (e.g., Multi-Channel Analysis of Surface Waves [MASW], Spectral Analysis of Surface Waves [SASW], seismic refraction, etc.), which represents 44% of the dataset, and from borehole methods, which represent 33% of the dataset. The use of different acquisition methods for estimating V_S profiles can give rise to uncertainties, but the recent papers of Garofalo et al. (2016) showed that it does not induce significant uncertainty in V_{S30} estimation.

4 Development of Pan-European maps of V_{S30}

4.1 Development of V_{S30} maps following a topography-based approach

The GEBCO_2014 30 arc-second DEM was used to infer V_{S30} values for Europe from slope data using the Wald and Allen (2007) methodology. As discussed in Lemoine et al. (2012), topographic slopes derived from land terrain models only are associated with artifacts in coastal areas. We therefore decided to use the joint topographic/bathymetric DEM described in Sect. 3.2 and the topographic slope was calculated using the same algorithm as Wald and Allen (2007).

Ranges of V_{S30} values were calculated for both stable continental and active tectonic areas by applying the Wald and Allen (2007) correlations. The next step was to separate stable continent areas and active tectonic ones. The distinction was done using the SHARE

tectonic classification, modified from Delavaud et al. (2012), which is shown in "Appendix B". Turkey, which was not fully covered by the SHARE zonation, was considered as an active tectonic area for the whole country. A final merged map was built merging the information from the active and stable tectonic areas according to the SHARE classification (Delavaud et al. 2012), the result of which is a 30 arc-second resolution map providing V_{S30} ranges for the whole Europe including Iceland (Fig. 4).

4.2 Development of V_{S30} maps following a geology-based approach

The spatial distribution of the observed V_{S30} dataset is quite unbalanced, with more than 80% of the data coming from active tectonic areas (Italy, Greece, Turkey, the Alps and the Pyrenees) (Fig. 2). This poor spatial distribution of the V_{S30} data induces a bias in the analysis since some geological units will be underrepresented in the dataset as, for example, geologies from stable areas in Western and Northern Europe. However, the spatial analysis in Fig. 5 shows that an extrapolation of the V_{S30} information (coloured polygons) to all the geological polygons characterized by a similar lithology (grey polygons) leads to a rather good spatial coverage of V_{S30} information all over Europe.

We analysed the distributions of measured V_{S30} values for each of the lithological categories shown in Fig. 1 and "Appendix A". These values are assumed to be normally distributed in all cases. Figure 6 indicates that for soft soils, mean V_{S30} values are mostly between 250 and 450 m/s, corresponding to EC8 soil classes B to C, as expected from Fig. 3. On the contrary, results for stiff soils to rock are quite unexpected since they show low V_{S30}, with mean values between 324 and 587 m/s (except for MAFIC code) and mean values between 393 and 666 m/s, corresponding mainly to EC8 soil class B instead of the expected soil class A. This can be due to poor geological classification at station sites since geology is based on a 1:1,500,000 scale map, which is not precise enough for local analysis. It may also result from bias in V_{S30} sampling, since site characterizations of strong-motion stations are generally devoted to soft soils. Our observed V_{S30} dataset is limited in terms of rock V_{S30} values due to a lack of measurements from sites located on these types of geological units.

As the V_{S30} dataset collected for the ESRM20 was not sufficient to build our own geology-based model for estimating V_{S30}, we opted for a more pragmatic approach that uses an existing model, built on a regional scale, and extrapolated it to the rest of Europe. Several possible approaches adopted in various regions of continental Europe were considered. The first was the model developed by Stewart et al. (2014) in Greece, which combined terrain type, surface geology (age) and surface gradient information. This model is not easily reproducible at European scale, however, mainly due to the complexity of terrain classification. The second approach was the model developed by Di Capua et al. (2016) based on surface geology (lithology) and considering additional criteria such as geological age, consistency and terrain structure. As for the case of Greece, this model is not easily reproducible at European scale mainly due to a lack of harmonized surface geology at European scale in addition to the complexity of terrain classification. Finally, we considered the model developed by Vilanova et al. (2018) for Portugal, which is based mainly on a stratigraphic classification and was found to be easily reproducible at the European scale using the stratigraphic codes in "Appendix A". This approach was thus chosen and the mapping between the current stratigraphic classification and the classification of Vilanova et al. (2018) is given in Table 1.
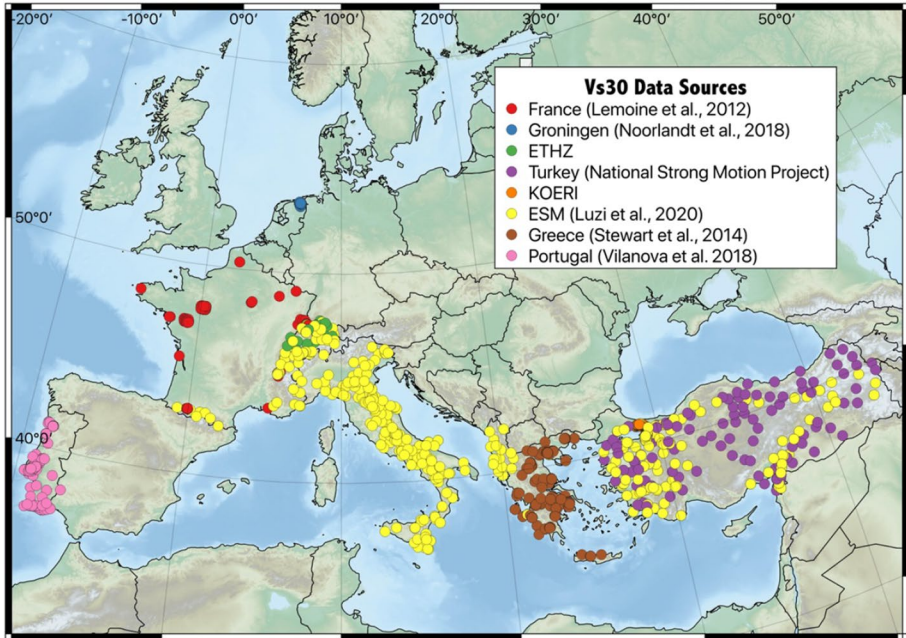
Fig. 2 Location of the measured V_{S30} values compiled in the database for inferring a European amplification model

Table 1 Correspondence between the stratigraphic classification of the geological data set and the classification proposed by Vilanova et al. (2018)

Synthetic system	Stratigraphic Code	Vilanova et al. (2018) classification
HOLOCENE	HC	F3 or F1 (only for volcanic and igneous formations: VOLC and IGENOUS codes)
NEOGENE	NG	F2 or F1 (only for metamorphic, volcanic and igneous formations: METAM, VOLC and IGENOUS codes)
PALEOGENE	PG	F2 or F1 (only for metamorphic, volcanic and igneous formations: METAM, VOLC and IGENOUS codes)
CENOZOIC	CN	F2
MESOZOIC	MS	F1
PALEOZOIC	PZ	F1
PHANEROZOIC	PH	F1
PRECAMBRIAN	PK	F1

The resulting map of the geologically based model from Vilanova et al. (2018) method is presented in Fig. 7. An important limitation of this kind of modelling is the processing of volcanic and weathered formations, which could present either consolidated or unconsolidated states (e.g., scoria cones, volcanic debris, slope colluvium, etc.). However, at

continental scale, the distinction between these two consolidation states is not available and would require a finer resolution analysis to be taken into account. Moreover, we could consider that it concerns mainly steep mountainous regions with few populated urban areas and therefore presents a low risk level for our purpose.

4.3 Comparison between inferred V_{S30} and observed V_{S30}

We compared the two V_{S30} models described above to the in-situ observations of V_{S30} from the compiled database shown in Fig. 2. The comparison was processed in terms of EC8 soil classes and the accuracy of the two different V_{S30} models to predict the correct soil class for the respective in-situ observations. From these comparisons we found that (i) the two models are better than random chance (accuracy > 25%), (ii) the topographically-inferred model using the approach of Wald and Allen (2007) yielded a greater accuracy (51%) than that of the geologically-inferred model using the approach of Vilanova et al. (2018) (38%), (iii) the accuracy of the two models are similar for the Eurocode 8 C class but the topographically inferred model improves on the geologically-inferred model for the A and B site classes. Though these comparisons give a general perspective on the relative performance of the two models, there are several caveats to be taken into account. Firstly, the in-situ observations are mainly located in active regions, with few tests on tectonically stable regions. Secondly, the Vilanova et al. (2018) model has been built primarily in a passive region and could potentially be improved by further adaptation for application in other contexts. Given the improved accuracy of the topographically inferred V_{S30} approach in predicting the observed V_{S30} when compared to the geologically inferred approach, the topographically inferred V_{S30} values are adopted in the subsequent analysis and European model. Both models are made available on the European Facility for Earthquake Hazard and Risk (EFEHR) website (http://risk.efehr.org/site-model/) (Table 2) .
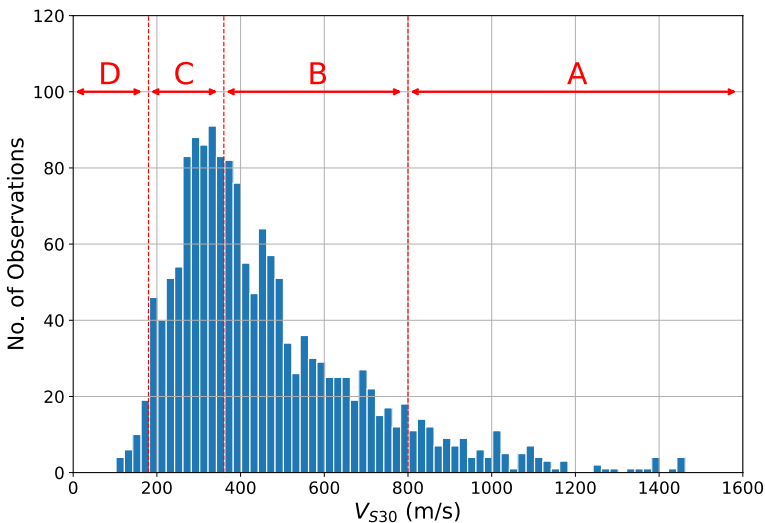

Fig. 3 Distribution of V_{S30} observations in the compiled database, with Eurocode 8 site class ranges indicated
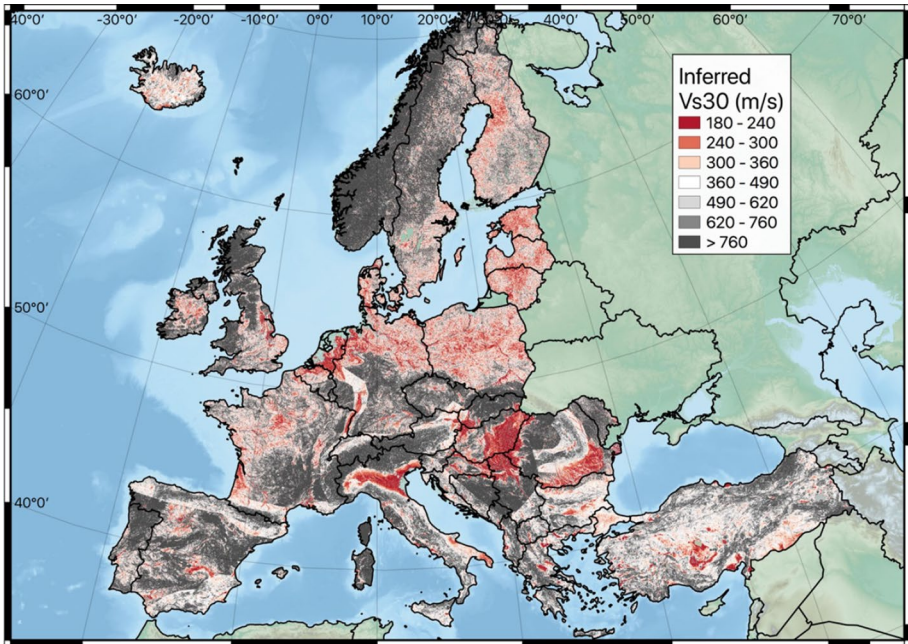
Fig. 4 Final ranges of V_{S30} over Europe inferred from Wald and Allen (2007) V_{S30}/slope correlations for active and stable areas with slope based on the GEBCO_2014 topography/bathymetry model
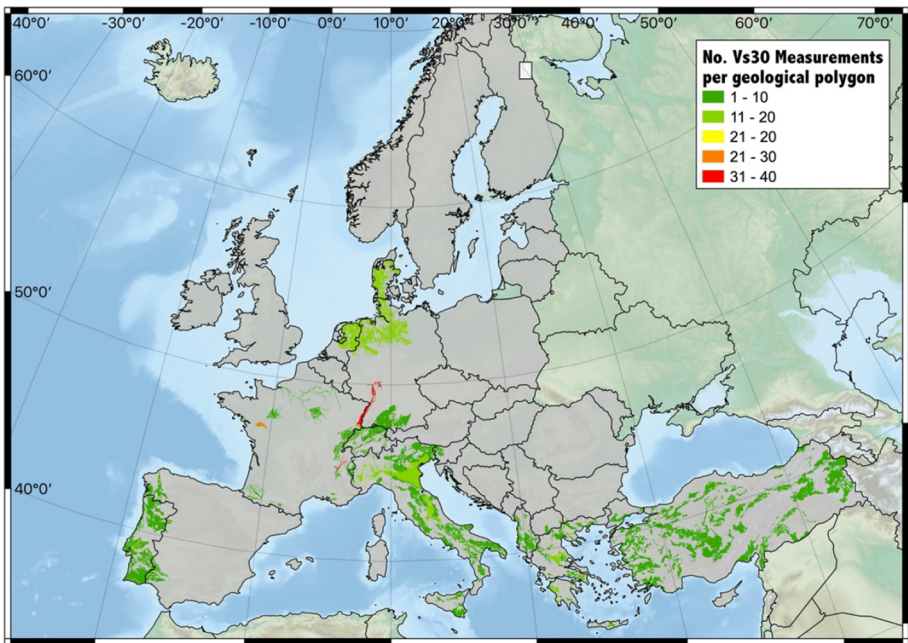

Fig. 5 Spatial distribution of V_{S30} measured data regarding lithological information. Coloured regions represent the number of V_{S30} data available in each polygon. Grey areas represent geological polygons where V_{S30} information is available by extrapolating the V_{S30} information available in similar lithology
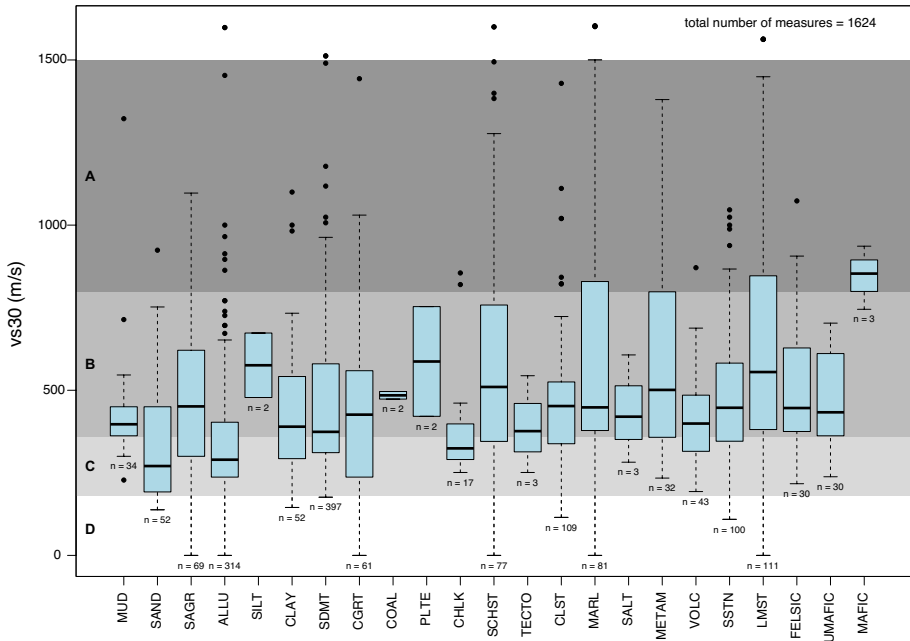
Fig. 6 Boxplots showing the measured V_{S30} distribution for each lithological code. The EC8 soil classes A, B, C and D are superimposed with grey levels. The EC8 class E was not considered since no information on the thickness on soft soil layers was available

Table 2 Accuracy of predicted V_{S30} against observed V_{S30} from the topographic approach of Wald and Allen (2007) and the geological approach of Vilanova et al. (2018)

		Predictions (% in parenthesis)								Total
		Topography (Wald and Allen 2007)				Geology (Vilanova et al. 2018)				
	N	A	B	C	D	A	B	C	D	
Observations	A	39 (2.40)	103 (6.34)	11 (0.68)	0 (0.00)	101 (6.22)	42 (2.58)	10 (0.62)	0 (0.00)	153 (9.42)
	B	93 (5.72)	517 (31.82)	183 (11.26)	0 (0.00)	324 (19.94)	291 (17.91)	178 (10.95)	0 (0.00)	793 (48.80)
	C	29 (1.78)	319 (19.63)	270 (16.62)	8 (0.49)	110 (6.77)	281 (17.29)	235 (14.46)	0 (0.00)	626 (38.52)
	D	5 (0.31)	10 (0.62)	27 (1.66)	11 (0.68)	5 (0.31)	23 (1.42)	25 (1.54)	0 (0.00)	53 (3.26)
	Total	166 (10.22)	949 (58.40)	491 (30.22)	19 (1.17)	540 (33.23)	637 (39.20)	448 (27.57)	0 (0.00)	1625 (100)

Total accuracy for the topographically inferred V_{S30} is 51% and for the geologically inferred approach 38%

5 Exploring the relation between the station-to-station ground motion site term ($\delta S2S_S$) and inferred site proxies

5.1 Structure of the ground motion model and its random effects

With the availability of the Engineering Strong Motion flatfile (Lanzano et al. 2019) the number of recording stations reporting multiple observations has increased by an order of magnitude in comparison to the preceding RESORCE strong motion database (Akkar et al. 2014b). In the construction of the ground motion model for application to shallow crustal seismicity in Europe, Kotha et al. (2020) identify 644 stations with more than five strong motion observations, and more than 1100 with three or more. Of the set of strong motion stations represented in the database used for the development of the Kotha et al. (2020) GMM, just over fifth of the strong motion stations are associated with a measurement of V_{S30}. For the remainder, V_{S30} values are defined from the topographically inferred model. Nevertheless, although the proportion of stations with observed V_{S30} values is small, we can however retrieve estimates of the direct ground motion amplification with far larger proportion of the data set from the repeated site-to-site variability ($\delta S2S_S$) inferred from the mixed effects regression (Bates et al. 2015).

The general structure of the ground motion model of Kotha et al. (2020) and its associated random effects from the mixed effects regression is given as:
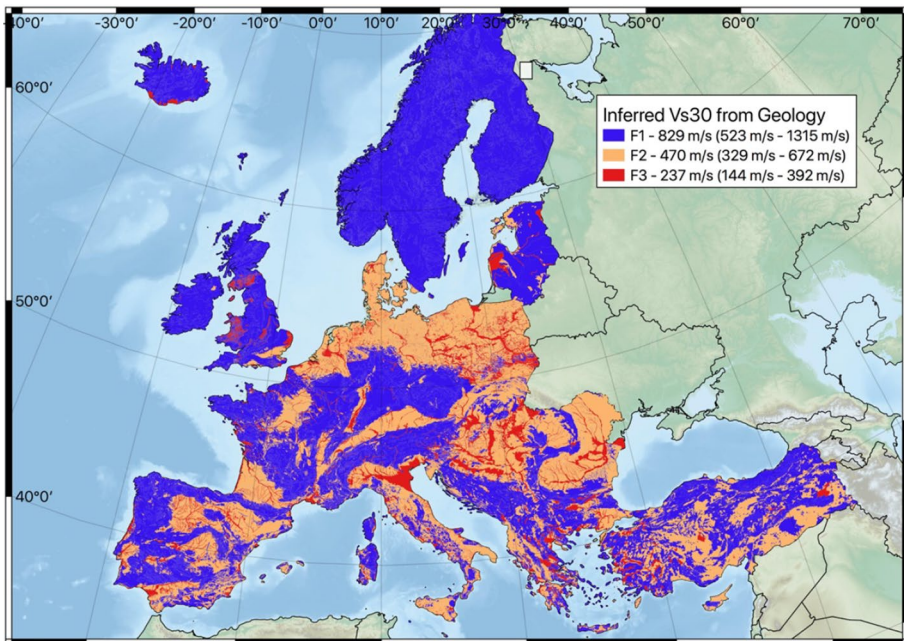
Fig. 7 Geology-based model of Vilanova et al. (2018) extrapolated to all of Europe. Mean V_{S30} given for each class along with -1σ to $+1\sigma$ range assuming a normal distribution

$$
\begin{aligned}
\ln Y_{ES}(T) =\, & e_1(T) + f_{r,g}\big(M_{W,E}, R_{JB,ES}, h_{D,E}, T\big) \\
& + f_{r,a}\big(R_{JB,ES}, T\big) + f_m\big(M_{W,E}, T\big) + \delta L2L_l(T) \\
& + \delta B_E^0(T) + \delta S2S_S(T) + \delta W_{ES}(T)
\end{aligned}
\tag{1}
$$

where Y_{ES} is the ground motion (PGA, PGV or spectral acceleration, Sa, at a period T) from event E, with magnitude $M_{W,E}$ and hypocentral depth $h_{D,E}$, reported at site S with a Joyner-Boore distance $(R_{JB,ES})$ from the source. $f_{r,g}\big(M_{W,E}, R_{JB,ES}, h_{D,E}, T\big)$, $f_{r,a}\big(R_{JB,ES}, T\big)$ and $f_m\big(M_{W,E}, T\big)$ are the geometric spreading, anelastic attenuation and magnitude scaling terms respectively, which are explained in further detail by Kotha et al. (2020). $\delta L2L_l(T)$ describes the source-region to source-region variability; effectively describing the extent to which earthquakes in a given region are more or less energetic with respect to the "average" of the data set. This is a property not of a particular event or station but of a particular region in which the event occurs, the regions being defined for this purpose using the TECTO large scale area source branch of the ESHM20 (Kotha et al. 2020). As described in further detail by Weatherill et al. (2020b), in the construction of the ground motion model logic tree for Europe, the $\delta L2L_l(T)$ term is treated as an epistemic uncertainty for the purpose of PSHA and mapped into the logic tree. Of the remaining random effects terms, $\delta B_E^0(T)$ describes the region-corrected event-to-event variability of the ground motion, $\delta S2S_S(T)$ the station-to-station variability and $\delta W_{ES}(T)$ the remaining event- and station-corrected within-event variability. The four random effects components described here are described by Gaussian distributions with means of zero and standard deviations $\tau_{L2L}(T)$, $\tau_E(T)$, $\phi_{S2S}(T)$ and $\phi_0(T)$ respectively.

The general functional form of the GMM assumes a relatively simple structure, containing only terms for the magnitude and distance scaling. This approach permits a greater exploration of the random effects to determine regional features that may be modelled directly in the GMM logic tree, as opposed to attempting to fit more complex models that may not be well constrained by the distribution of the data set (e.g., hanging wall effects, directivity, top of rupture depth scaling etc.). More importantly for the current analysis, it does not contain an explicit site amplification term as a function of V_{S30} or any specific parameter. As such, the resulting $\delta S2S_S(T)$ term refers to the relative station-to-station amplification *with respect to the centre of the data set*, rather than to a reference rock condition. This follows the approach of Kotha et al. (2018), who utilised cluster analysis on the resulting $\delta S2S_S(T)$ distribution to identify sites within the Japanese Kik-net database sharing common site characteristics, from which the definition reference rock *emerges* from the stations with similarly low $\delta S2S_S(T)$. The centre of the data set will, to some extent, reflect the underlying distribution of V_{S30} values (or other proxy) in the data, albeit that this does not refer to a single V_{S30} but rather a reference value that is dependent on period. Being centred on the middle of the data set, $\delta S2S_S(T)$ can take negative values, thus describing the de-amplification with respect to the "average" site condition implied by the data, as well as to positive amplification terms.

The complete data set of station-to-station terms contains estimates of the amplification factor $\delta S2S_S(T)$ for more than 1100 stations across Europe and the Middle East that are constrained by three or more observations. The corresponding spatial distribution can be seen for T = 0.2 s and T = 1.0 s in Fig. 8. Regional-scale trends in $\delta S2S_S$ are difficult to determine, though coherently higher $\delta S2S_S(T)$ values at longer periods can be seen in some of the more expansive low-lying basins and plains, such as the Po

Plain region of northern Italy and the Danube plains of Eastern Romania. Otherwise, the $\delta S2S_S$ values will likely reflect local scale features close to the station.

5.1.1 A general model for site amplification within the GMM

The complete set of station-to-station terms allows us to explore the possibility of identifying site properties that can be adopted as predictors of site response at a large geographical scale, as well as their impact on the uncertainties of the GMM. Of particular interest for this purpose is the distinction between *measured* site properties (i.e., those known by direct measurement at the site) and *inferred* site properties (i.e., those associated to the site by virtue of its location in a large-scale geographical data set, i.e. a proxy). For application in the ESHM20, the distinction between *measured* and *inferred* site parameters is of critical importance because of the two contexts being served, namely the engineering application (requiring the characterisation of the hazard on Eurocode 8 class A rock) and the seismic risk application (requiring the characterisation of the seismic ground motion for any location in the European exposure model according to the local soil condition). The critical assumption is made here that for application to the Eurocode 8 context the soil condition at the target site is *known*, and that the uncertainty in the ground motion model should reflect this case. For the ESRM20, the site condition for the majority of Europe may only be inferred from the topographic and/or geologic data described previously, and as such the uncertainty in the ground motion model should be adjusted accordingly.

The total aleatory uncertainty of the GMM in Eq. 1 is described by $\sigma_T = \sqrt{\tau_e^2 + \phi_0^2 + \phi_{S2S}^2}$, noting that for the current application τ_{L2L} is not included in the total uncertainty as it is treated as an epistemic uncertainty rather than aleatory (Weatherill et al. 2020a). In the case that $\delta S2S_S$ is well constrained for a site such that site-to-site variability can be removed from the GMM, the non-ergodic aleatory uncertainty can be used: $\sigma_T = \sqrt{\tau_e^2 + \phi_0^2}$. This might be considered the optimum reduction of site-to-site variability, but this is a property only of the site at which the measurement is taken (or an area surrounding it with limited radius). For a more general application, an explicit site amplification term $f_{site}(\boldsymbol{\theta_S}, T)$, is desired to reduce the total aleatory variability and this can be regressed from the $\delta S2S_S(T)$ data assuming the general form:
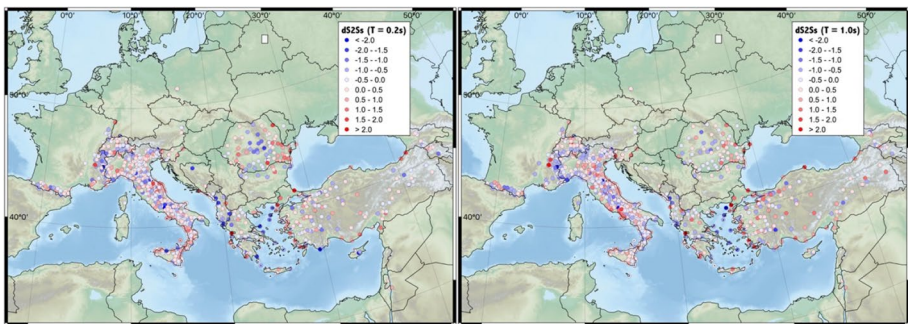

Fig. 8 Spatial distribution of station-to-station term, $\delta S2S_S$, across Europe for the period T = 0.2 s (left) and T = 1.0 s (right)
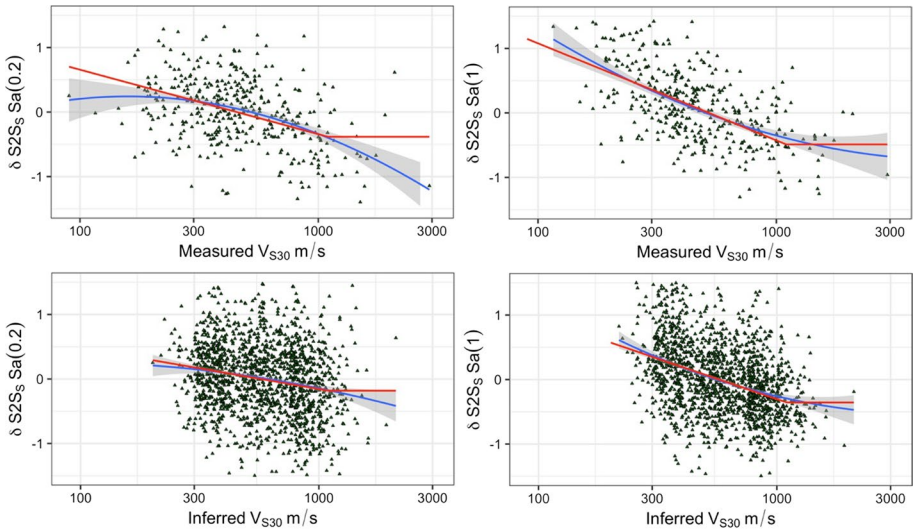
Fig. 9 Trends in $\delta S2S_S$ with respect to measured V_{S30} (top row) and topographically inferred V_{S30} (bottom row) for periods T = 0.2 s (left column) and T = 1.0 s (right column) respectively. Blue lines indicate the non-parametric LOESS fit to the data (and its uncertainty in shaded grey), while red lines indicate the fit of the f_{site} model from Eq. 3

$$
\begin{aligned}
\ln Y_{ij}(T) = & f_{R,g}(R_{JB}, M_W) + f_{R,a}(R_{JB}) + f_m(M_W) \\
& + f_{site}(\boldsymbol{\theta_S}, T) + \delta B_e^0(T) + \delta S2S_S^\theta(T) + \delta W_{ES}(T)
\end{aligned}
\tag{2}
$$

where $\boldsymbol{\theta_S}$ refers to one or more specific parameters of the site and $\delta S2S_S^\theta(T)$ the revised site-to-site variability, which is itself a Gaussian distribution with zero mean and a standard deviation of ϕ_{S2S}^θ. Kotha et al. (2020) adopt this approach in order to determine possible site amplification models dependent on V_{S30} (m/s) or topographic slope (m/m). In their model, a quadratic functional form is adopted and $f_{site}(V_{S30})$ fit only to the smaller subset of 419 stations with measured V_{S30}. Though the quadratic functional form may fit the dataset reasonably well, it lacks a physical basis and deviates from a simple linear model only at the extreme low and high V_{S30} or slope values, which are themselves poorly constrained by data. It also cannot be safely extrapolated to higher or lower V_{S30} values outside of the range of data, which may result in unexpectedly high or low amplifications at certain locations.

For the purposes of the ESHM20 and ESRM20 a different functional form is adopted, this time using a two segment piecewise linear model:

$$
f_{site}(\theta, T) = \begin{cases} g_1(T) \ln\left(\frac{\theta}{\theta_{ref}(T)}\right) & \text{for } \theta \leq \theta_c \\ g_1(T) \ln\left(\frac{\theta_c}{\theta_{ref}(T)}\right) & \text{for } \theta > \theta_C \end{cases} + \delta S2S_S^\theta(T)
\tag{3}
$$

where θ is the site property of interest, g_1 and θ_{ref} are period-dependent coefficients fit during the regression, and θ_C is a period independent cut-off parameter above which f_{site} is constant. $\delta S2S_S^\theta$ is the resulting site to site residual, which itself should be Gaussian distributed with a mean of zero and standard deviation of ϕ_{S2S}^θ. This general form applies for

all site properties considered in this analysis; hence the use of the superscript θ. Example correlations between $\delta S2S_S^{\theta}(T)$ and both measured V_{S30} and inferred V_{S30} from the 30 arc-second slope at a site are shown for the periods 0.2 s and 1.0 s respectively in Fig. 9.

There are several critical issues to be addressed in interpreting the plots shown in Fig. 8. Two trends are immediately obvious: the first is that a steeper slope (i.e., larger absolute g_1) can be seen for the longer period, Sa (1.0), than for the shorter period, and the second is that the scatter in the $f_{site}(\theta)$ is larger for the inferred V_{S30} case than the measured V_{S30} case such that $\phi_S^{INF} > \phi_S^{OBS}$. The latter observation is largely unsurprising and implies clearly that as a predictor of amplification the use of an indirect proxy for the shearwave velocity in the upper layers of the soil is poorer than a direct measurement of it. What is important here is that the degree to which this is the case, and therefore the degree to which the uncertainty on the prediction of the resulting surface ground motion should be increased.

The observation that the correlation between $\delta S2S_S$ and V_{S30} (both inferred and measured) is generally weaker for shorter period motion than longer period motion requires some careful interpretation. $\delta S2S_S$ is a *linear* indicator of the degree to which ground motions at the given site may be systematically greater or lower than the centre of the distribution of sites in general. By itself it does not reveal much about the specific factors that are causing it, nor the relative impact from multiple contributing factors. As such, in addition to the influence of the shallow soil conditions at a site, the $\delta S2S_S$ term at shorter periods may be controlled by factors such as soil nonlinearity or high frequency site κ_0 conditions. Similarly, at longer periods different compounding phenomena such as basin depth and 2D/3D resonance effects may come to dominate the $\delta S2S_S$. By basing f_{SITE} exclusively on the empirical site term we cannot define a predictive model that distinguishes between the various phenomena, but this does not mean that they are not represented in the overall distribution, ϕ_{S2S}.

In the case of the short period motion, the non-parametric locally estimated scatterplot smoothing (LOESS) regressions (indicated as blue lines and surrounding grey shaded regions) suggest that while a near linear trend of increasing $\delta S2S_S$ can be seen for much of the body of the range V_{S30} ($300 \leq V_{S30}$ (m/s) ≤ 1000), on the few soft soil sites with $V_{S30} < 250$ m/s the trend reverses (or at least flattens). The majority of the stations with measured $V_{S30} < 250$ m/s are located in the Po Plain, while the rest are located in mostly Quaternary basins or valleys within the Appenines and the Alpine foreland. When exploring the specific ground motions at those stations with both low V_{S30} and lower than expected $\delta S2S_S$, however, we find that in several cases the $\delta S2S_S$ term is constrained exclusively by low intensity shaking from smaller magnitude events (less than 0.02 g), which does not necessarily support the assumption that nonlinearity is playing the dominant role here.

We note here too that the location and housing of the strong motion sensor (e.g., free-field, in-building, close to large structures etc.) may contribute to the overall variability, especially at short periods. For 55% of the stations considered, no information is available in the ESM database regarding the location of installation and proximity to structures, while for those that are reported nearly 70% are identified as free-field or "free-field close to structure". The remainder are reported as located in basements or ground floors of buildings (predominantly one-storey masonry). While we cannot rule out the possibility of such site-specific factors influencing the $\delta S2S_S$ terms at short periods, we do not have enough information to identify specific trends or regional variation that could be attributed to these factors, and therefore cannot quantify the extent to which this contributes to the total aleatory variability. Though we hope that more information on housing and installation will be
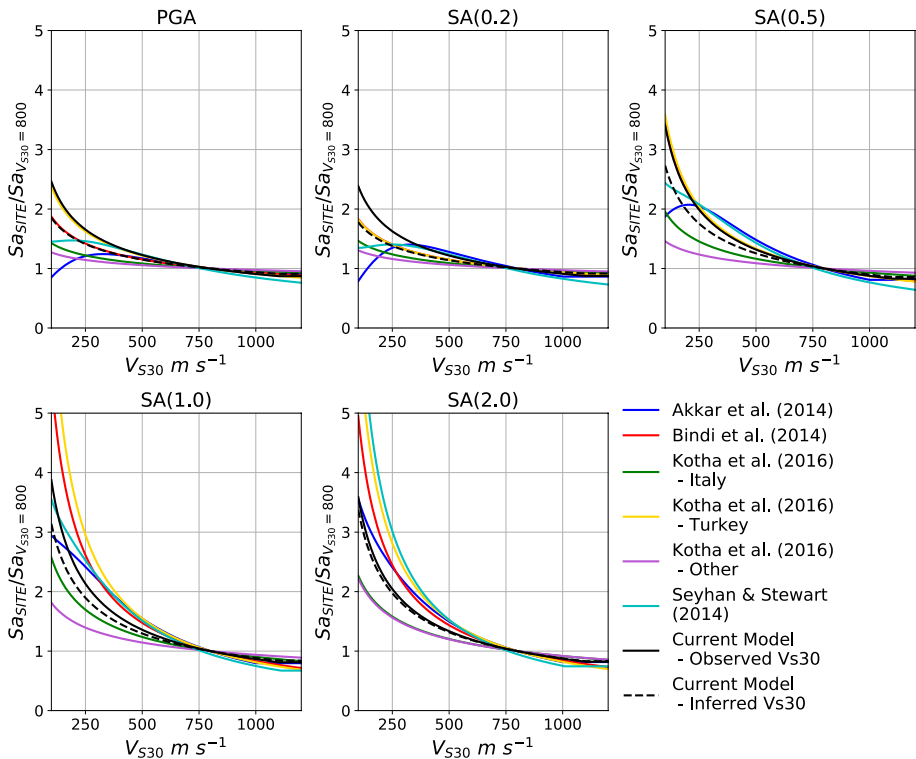
Fig. 10 Comparison of expected site amplification from different GMMs for a site located 15 km (Joyner-Boore distance) from a Mw 6.5 event, with respect to an equivalent site of V_{S30} 800 m/s. "Current Model" refers to the amplification functions presented in Sect. 5.2, separating two cases of measured V_{S30} (solid black line) and V_{S30} inferred from the topographic model presented in Sect. 4 (dashed black line)

included in future versions of the ESM database, we do not find any compelling reason to believe that these factors would impact significantly on the models calibrated here.

The site amplifications predicted by both the inferred and measured V_{S30} models here are compared in Fig. 10 against those implied by several existing European GMMs (Akkar et al. 2014a; Bindi et al. 2014; Kotha et al. 2016). These European models were all built using the European RESORCE strong motion data set (Akkar et al. 2014b), and as such share many of the same sites as this analysis. None of the preceding GMMs divide the sites between those using inferred V_{S30} and those using measured V_{S30}; however, each model differs slightly in terms of the modelling process or underlying assumptions. Akkar et al. (2014a), for example, is the only model to adopt a nonlinear site response factor, whereas Kotha et al. (2016) lack a term for soil nonlinearity but instead explore regional differences in the linear site amplification coefficient between Italy, Turkey and "other" regions in Europe. Also included is the site amplification model of Seyhan and Stewart (2014), which is based on the NGA West 2 data set and forms the basis for the site amplification models found within some of the resulting NGA West 2 GMMs (e.g., Boore et al. 2014). For much of the spectral period and V_{S30} range the current model predicts amplifications that are close to the middle of the spread of amplification factors predicted by the previous European models. Where the new model diverges from the existing models is in the case

of the measured V_{S30} on soft soil ($V_{S30} < 250$ m/s), where it predicts higher amplifications than all of the models except the Kotha et al. (2016) "Turkey" model. These comparisons would suggest that while the method adopted here for the calibration of a V_{S30} site amplification model may differ from those fit to previous European strong motion databases, the resulting level of amplification is mostly consistent with the range predicted by such models, even in the case of inferred V_{S30}.

5.1.2 Incorporating geology into the site amplification model

The dual objectives of the ESHM20 in terms of describing hazard on reference rock for engineering application and on soil for risk applications provide a basis for devising separate models for inferred and measured site condition. As the seismic risk application requires the definition of the site condition across the entire extent of the area covered within the exposure model, it is inevitable that V_{S30} inferred from proxy will come to play the dominant role. The topographically and geologically inferred V_{S30} maps shown in Sects. 4.1 and 4.2 form the primary data set for defining the site parameterisation for input into the seismic risk calculations. Adopting the inferred V_{S30} classification in the GMMs should ensure that an appropriate increase in uncertainty is assigned to the ground motion for using the proxy.

While the uncertainty in the amplification may be appropriately penalized in this framework, there still remain fundamental limitations to adopting the inferred V_{S30}. Principal amongst these is the demonstration that correlation between slope and V_{S30} is dependent not only on the tectonic classification of the region (active/stable) but on the lithology of the local geology itself. One of the rationales behind the Wald and Allen (2007) approach is not necessarily that the topographical slope itself is an explicit predictor of V_{S30}, but rather that the geomorphological environments that produce steep or shallow gradients are correlated to those where certain types of surface soil conditions are dominant. There are many specific geological and geomorphological environments where this is not necessarily the case, and these exceptions may be particularly pertinent when the objective is to use this information to model site amplification rather than simply predicting the V_{S30} itself.

Weatherill et al. (2020b) explored the relation between $\delta S2S_S$, as determined from a GMM fit using data from the Japanese Kik-net data set in a manner similar to that described here, and the properties of the site including those measured directly for the station and those inferred from topography and geology. They find that the correlation between $\delta S2S_S(T)$ and both measured and inferred site properties to be dependent upon the geology at the location of the station. Specifically, in regions where older surficial geological units of pre-Cenozoic are present there is little to no correlation between inferred V_{S30} and $\delta S2S_S$, while in Quaternary environments the correlations and their dependence on spectral period are clear. They also demonstrate that in terms of predicting site amplification, there is no discernible reduction in uncertainty for using V_{S30} inferred from topographic slope rather than using the topographic slope as the predictor variable in itself. The dependence of $f_{site}(\theta)$ on geology, whether θ in this case refers to measured V_{S30}, inferred V_{S30} or slope itself, was integrated into the amplification using a mixed effects regression process, which calibrated the coefficients of the model depending on the geological environment. This was utilised by Weatherill et al. (2020b) to produce a 30 arc-second site amplification model of Japan.

Given both the absolute number of stations in the ESM database for which $\delta S2S_S(T)$ has been determined, and the relatively low proportion with measured site properties, there is

a case for adopting the approach of Weatherill et al. (2020b) for application to Europe. The 30 arc-second topographic slope and geology are assigned to each station in the database for which $\delta S2S_S(T)$ is determined, using the topography and geology data sets described previously. To account for the influence of geology, the model for f_{site} is developed using robust linear mixed effects regression (Koller 2016) such that:

$$f_{site}(\theta, T) = \begin{cases} (g_1(T) \pm \delta g_1 | G, T) \ln\left(\frac{\theta}{(\theta_{ref} \pm \delta\theta_{ref}|G,T)}\right) & \text{for } \theta \leq \theta_c \\ (g_1(T) \pm \delta g_1 | G, T) \ln\left(\frac{\theta_c}{(\theta_{ref} \pm \delta\theta_{ref}|G,T)}\right) & \text{for } \theta > \theta_C \end{cases} + \delta S2S_S^\theta(T) \quad (4)$$

where $\delta g_1 | G$ and $\delta\theta_{ref} | G$ are the period-dependent, correlated random effects conditional on the geological category, G. As the subdivision of the data into the geological categories provides a smaller data for the fit of the random effects, the resulting random effect coefficients are smoothed to minimise the period-to-period variability. As was the case for Weatherill et al. (2020b), a critical balance has to be found between capturing as much of the relevant geological variability as possible, while still ensuring a sufficient volume of data in each category to ensure stable convergence in the mixed-effects regression. While lithology may act as a strong control on the site amplification, adoption of the lithological classification results in too few data per category from which obtain a stable estimate of the random effects coefficients. Similarly, use of all of the stratigraphic categories presents the same problem, meaning that some grouping of categories was necessary in order to ensure statistically significant random effects coefficients. Different strategies for grouping stratigraphic units and/or lithological units were tested, and the resulting random effects coefficients and standard deviations compared. Groupings based on stratigraphy rather than lithology were preferred, as clearer and more coherent trends in amplification function with geological age could be discerned, while groupings of lithological units were more arbitrary and yielded fewer clear and explicable differences between groups. For the purpose of developing the amplification model for ESRM20, the geological category is assigned to one of the seven following geological eras using the stratigraphic classification from the harmonised geological maps (Fig. 1): Pre-Cambrian (43 sites), Paleozoic (80 sites), Jurassic-Triassic (172 sites), Cretaceous (305 sites), Cenozoic (407 sites), Pleistocene (328 sites) and Holocene (338 sites). These seven categories describe more than 95% of the total data set, and the vast majority of identified geological units in the European geological map. The remainder of cases are classified as "Unknown".

The distribution of sites with respect to slope and geological era is shown in Fig. 11, and a key trend to note here is the general skewing toward the low end of the range of slopes for the younger geological units (tertiary or later). This is to be expected from the geomorphology, with Holocene environments predominantly characterized by flatter lowland river plains or, in upland areas, sediment-filled valley bottoms. For the older geological units, a wider range of slope values can be seen, and here the higher slope values tend to originate from valley sides and/or outcrops of older rock. Naturally in such a data set there are likely to be a certain proportion of misclassifications that can emerge as a result of the resolution of the geological maps or slight errors and inconsistencies in either the station location or in the georeferencing of the maps. In some cases, the assignments of the geological era to the station were corrected by eye using aerial imagery (e.g., stations on harbour walls and jetties, artificial or reclaimed land etc.).

The mixed effects regressions are undertaken comparing three different predictor variables: *measured* V_{S30}, *inferred* V_{S30} and slope. The correlations and fitted regression models
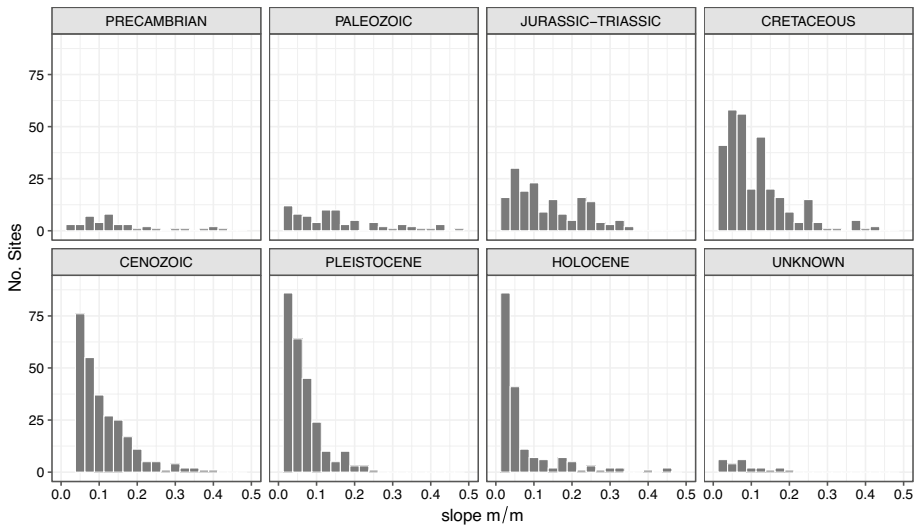
Fig. 11 Distribution of stations for which a $\delta S2S_S$ value could be determined, according to slope and geological classification

are shown in Figs. 12, 13 and 14 for the three predictor variables respectively. In the *measured* V_{S30} case, only the subset of 423 stations can be considered, resulting in substantially fewer available $\delta S2S_S$ values for each geological class: Pre-Cambrian (9 sites), Paleozoic (9 sites), Jurassic-Triassic (31 sites), Cretaceous (70 sites), Cenozoic (108 sites), Pleistocene (96 sites) and Holocene (98 sites). Though statistically significant correlations between $\delta S2S_S$ and measured V_{S30} can be seen clearly for the sites on Cretaceous rock or younger, for the Jurassic-Triassic units and earlier there are too few observations from which to establish an empirically robust model.

For the case of inferred V_{S30} and/or slope (Figs. 13 and 14) the overall correlation between $\delta S2S_S$ and the predictor variable is poorer across all geological classes, in line with what we have seen previously for the conventional regression case without geology as a random effect. For short period motion the geological class has seemingly little influence on the prediction of amplification, with most geological classes yielding similar models. In the case of the older geological environments the number of available data from which to establish any correlation is low and the resulting models may not be statistically significant. At longer periods, the differences between the geological units become more apparent, with stronger correlations between $\delta S2S_S$ and *inferred* V_{S30} or slope seen in younger Cenozoic, Pleistocene and Holocene environments than in older environments. In general, the gradients of the slope seem to increase in broad correspondence to the geological unit age, becoming progressively steeper for each unit between the Cretaceous and the Holocene eras.

The general trends in the correlations between slope or *inferred* V_{S30} and $\delta S2S_S$ are influenced by many factors, most notably the geomorphology and the geographical locations of stations within each category. Consider, for example, the sub-set of stations on Holocene sediments. These tend to correspond to sites that are located on either thick sedimentary basins such as the Po Plain or Danube river basin, upland valleys within mountainous regions, or smaller river deltas. Though mostly associated with flat local topography,
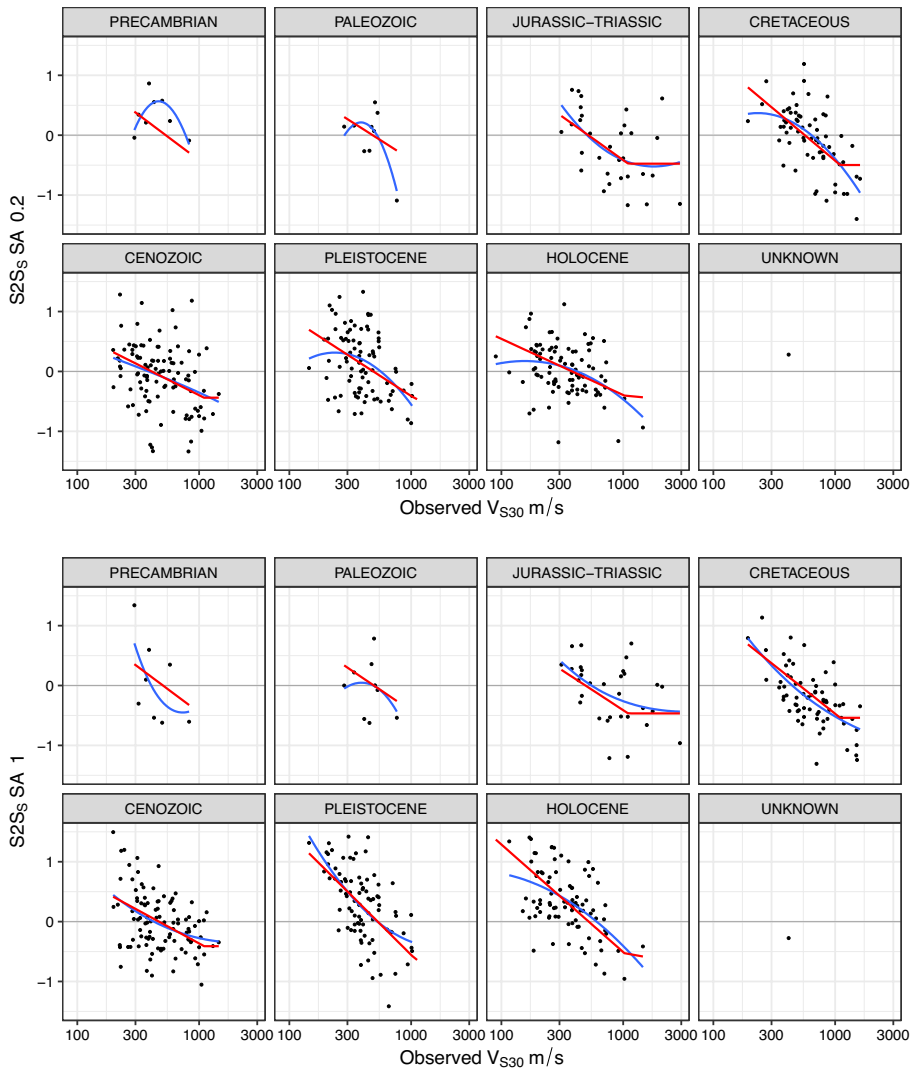
Fig. 12 Correlation between $\delta 2S_S$ and *measured* V_{S30} for each geological unit for Sa (0.2 s) (upper) and Sa (1.0 s) lower. Blue lines indicate the fit of a LOESS regression model (grey shaded area indicating the 5–95% confidence interval) and the red lines the fit of the robust mixed effects regression

stations on Holocene sediments with higher slope tend to be located at the edges of valleys (mostly in upland areas).

For older geological environments the stations tend to come from mountainous regions (e.g., Alps, Pyrenees, Dinarides and Carpathians etc.) with the largest contribution to the Cretaceous class being the dense network of seismic recording stations in the Apennine mountains of central Italy. In terms of the relative degree of amplification between classes, we generally observe a greater degree of amplification at short periods for these older geological classes (Precambrian through to Cretaceous) than for the younger classes (Cenozoic
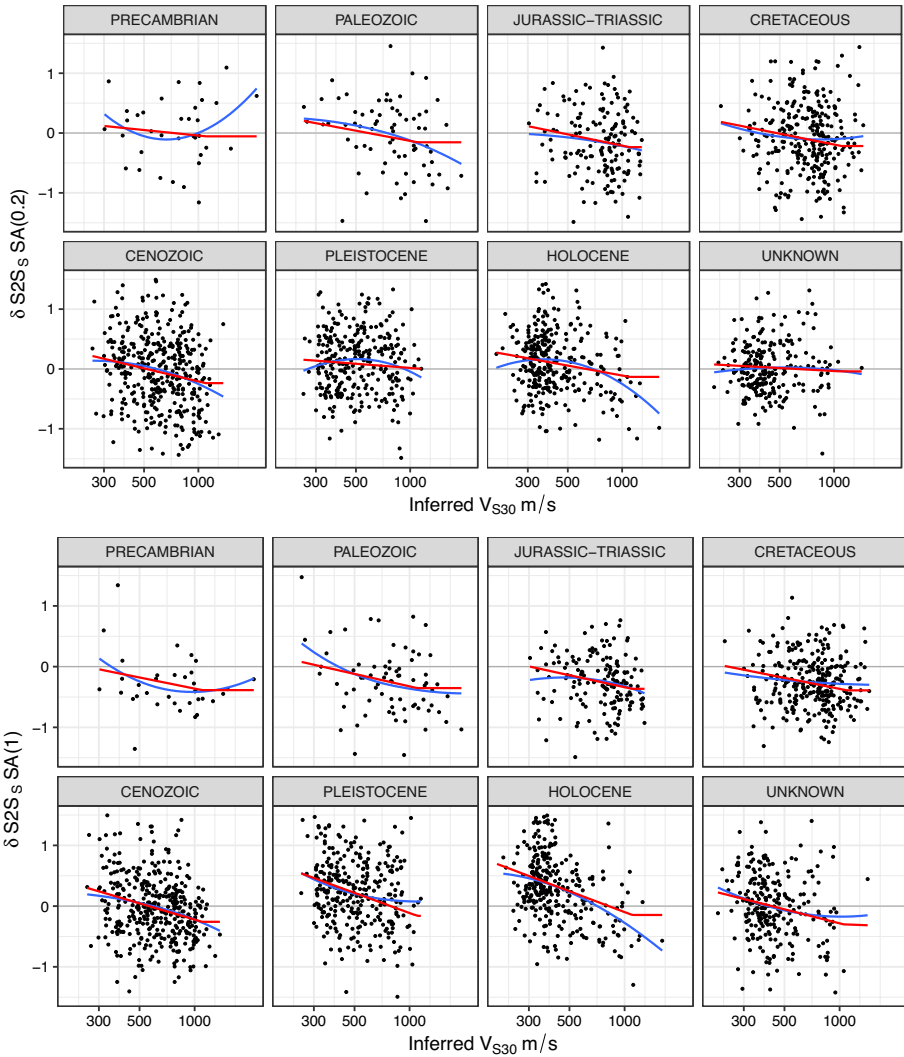
Fig. 13 As Fig. 12 for the case of topographically *inferred* V_{S30}

and later). This trend is seen particularly strongly in the case of the *measured* V_{S30} observations, but it is also visible to varying degrees in both models using the *inferred* V_{S30} and the slope as predictors. Care should obviously be taken, however, not to over-interpret these trends and acknowledge the limitations of assigning the geological conditions to the stations based on a large-scale and therefore potentially low-resolution data set. Similarly, the number of observations in these older geological categories are an order of magnitude lower than those from Tertiary and Quaternary sites and are particularly skewed toward certain regions. Generalisation of these trends to all similar geological environments in Europe, while necessary for the current application, still carries the potential for systematic error that can only be identified and corrected with the availability of more stations and more ground motion recordings.
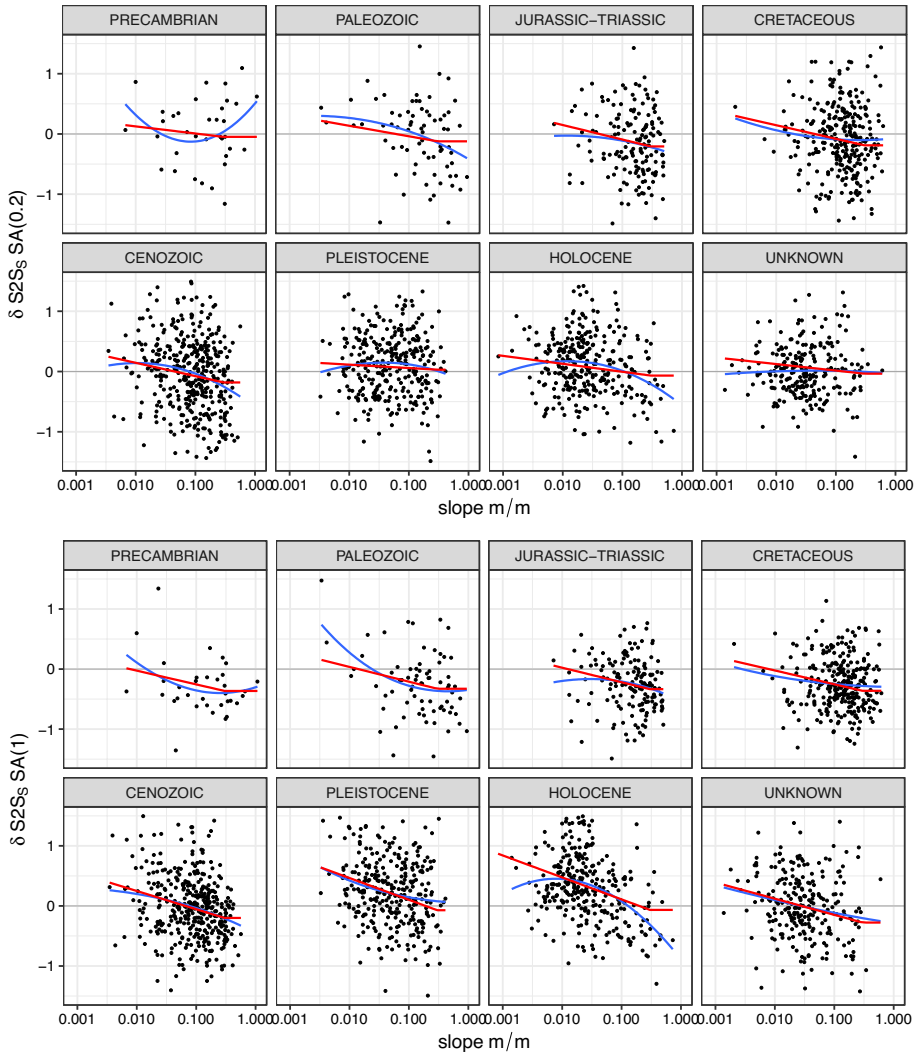
Fig. 14 As Fig. 12 for the case of slope

Figures 12, 13, 14 have illustrated the general trends the data set of $\delta S2S_S$ with respect to spectral period, geological class and predictor variable. To give a more complete picture of the resulting amplification model, however, we compare the resulting amplification levels with period for different site conditions and geological classes for the case of *measured* and *inferred* V_{S30} (Fig. 15) and slope (Fig. 16). For the V_{S30} case we adopt V_{S30} 800 m/s as a *reference* condition to which the amplifications refer, while for slope we use 0.3 m/m. This reference slope takes no specific meaning as a point of comparison; however only 0.6% of sites across Europe exceed this slope value. When comparing the *measured* and *inferred* V_{S30} cases, it is clear that overall larger amplifications are seen for the measured case and that amplification peaks shift from shorter periods in the older geological environments to longer periods in the younger ones. This highlights why increasing ϕ_{S2S} alone

for the *inferred* V_{S30} case would not necessarily capture the differences in the resulting amplification. The cause of the peak in amplification around 0.2–0.3 s period for older rock conditions and low V_{S30} values is difficult to explain and may simply represent extrapolation of the model to V_{S30} conditions too far below those input into the regressions. In reality, however, such conditions rarely coincide, so we avoid introducing constraints into the regression, which may result in greater instability in the results, or capping the amplification to a lower value.

Comparison of the absolute amplification values for the models using V_{S30} as a predictor against those using slope as a predictor is challenging, as we lack a reference slope value that can be considered directly equivalent to the reference V_{S30}. Nevertheless, the general trends with respect to both the slope and geological category are consistent with one another, albeit with some minor differences. For example, a minor peak in amplification around T = 0.5 s can be seen for the Holocene and Pleistocene classes in the *inferred* V_{S30} case, which is far less pronounced in the slope and geology model. A slightly larger degree of amplification appears to be present for the Cenozoic class when slope is used as a predictor, though the differences are moderate.

The influence that the geological unit has on the resulting amplification models illustrates the advantages of this approach with respect to the more conventional usage of topographically inferred V_{S30} as the sole proxy for site amplification. Though many studies have attempted to integrate geology and slope as a predictor of V_{S30}, they do not address the influence that the geology has on the prediction of amplification. The results also show that the degree of amplification is lower when using either slope or *inferred* V_{S30} in comparison to using *measured* V_{S30}.

5.1.3 Impact on site-to-site variability

The term $f_{SITE}(\theta, T)$ given in Eq. 4 describes the mean amplification in ground motion given the predictors in question. The lower degree of amplification that emerges for the same V_{S30} values when inferred from proxies compared to those from measurements, however, clearly demonstrates an incompatibility between the amplification models from mappable proxies and those from measured site data. We are still missing a key component that is extremely relevant for probabilistic seismic risk analysis, however, which is the resulting site-to-site variability. Figure 16 compares the site-to-site variability (ϕ_{S2S}^{θ}) from the different approaches here, including and excluding geology as a random effect, and the resulting differences in the total aleatory variability of the ground motion model (i.e., $\sigma_T^2 = \tau_E^2 + \phi_0^2 + \phi_{S2S}^2$). In the left-hand side of Fig. 17, the ϕ_{S2S} models for the case when geology is included as a random effect are shown in dashed lines, and those excluding it as solid lines. As a point of comparison, we also show the original ϕ_{S2S} terms for the different predictors adopted by Kotha et al. (2020): V_{S30} ("Orig. V_{S30}"), slope ("Orig. Slope") and no site predictor at all ("Orig."). Both the model of Kotha et al. (2020) and that proposed here use robust linear mixed effects regression, and for the V_{S30} predictor case Kotha et al. (2020) limits the sites to only those for which V_{S30} is *measured*. They also did not apply smoothing to the coefficients, which is something that is present in the new model. The resulting ϕ_{S2S} values from the slope-dependent case of Kotha et al. (2020) and V_{S30} case should be equivalent to the fixed effect model with slope as a predictor and the fixed effects model with *measured* V_{S30} as a predictor. Allowing for a small degree of mismatch due to the application of smoothing in the model proposed here, this does appear to be the case. This suggests that the differences in functional form assumed for f_{SITE} are
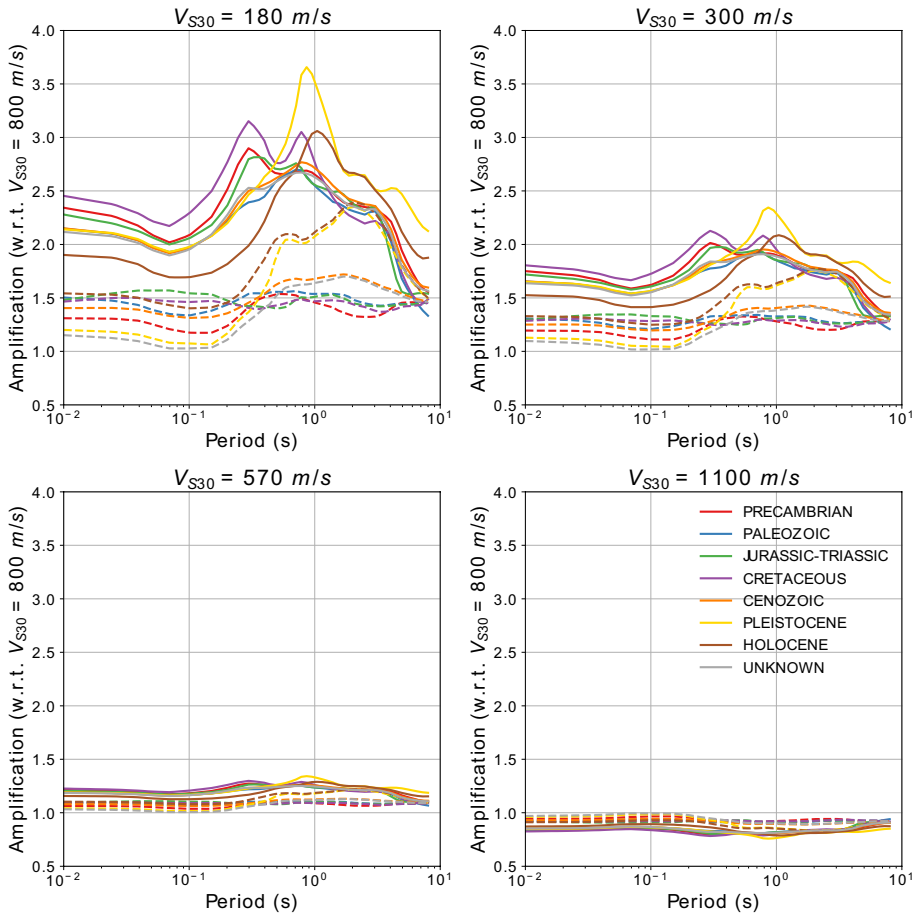
Fig. 15 Comparison of amplification factors with respect to V_{S30} 800 m/s rock for different geological classes and V_{S30} values. Solid lines indicate the *measured* V_{S30} case and dashed lined the *inferred* V_{S30} case

having no discernable impact on the resulting variability. Neither the original Kotha et al. (2020) model nor the model implemented here find compelling evidence of heteroskedasticity with respect to V_{S30} or any other predictor in the residuals of the site amplification model ($\delta S2S_S^{\theta}(T)$).

The contrast between the red lines (*measured* V_{S30}) and the blue and black lines (*inferred* V_{S30} and slope respectively) illustrate the differences between the use of directly measured parameters to predict amplification and those derived from mappable proxies. The black lines indicating the use of slope directly are obscured behind those of the *inferred* V_{S30} (shown in blue), and differences are on the order of one or two percent or less, which suggests there is no reduction in site-to-site variability for using *inferred* V_{S30} as a predictor of amplification rather than slope itself. The resulting difference in the total standard deviation when considering the *measured* V_{S30} rather than the *mappable* parameters (i.e., *inferred* V_{S30} or slope) is on the order of about 5–6% at short periods reducing to 1–2% in the 1–3 s period range.
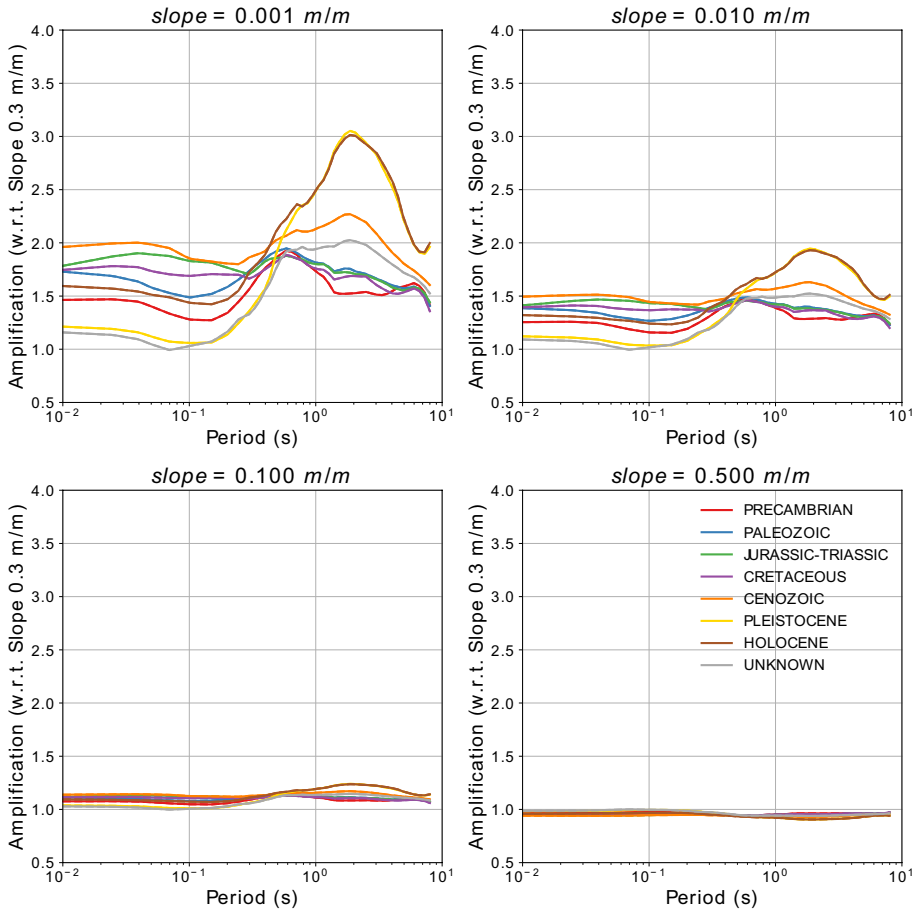
Fig. 16 Comparison of amplification factors with respect to a *reference* slope of 0.3 m/m for different geological classes and slope values

An important result regarding the site-to-site variabilities shown in Fig. 17, however, is the difference in the values when geology is included as a random effect and when it is omitted. Here we see that the introduction of geology produces a small reduction in ϕ_{S2S} at long periods (T ≥ 1.0 s) with respect to the ordinary fixed effects case, virtually no reduction at intermediate periods (0.4 ≤ T (s) ≤ 1.0) and a small increase in ϕ_{S2S} at short periods. This observation should be contrasted with the similar derivation of ϕ_{S2S} of Weatherill et al. (2020a), wherein we applied the same approach to Kik-net data. In that case, we did achieve a reduction in variability across the whole period range when introducing geology as a random effect, albeit the reduction was greater at longer periods and barely discernible at short periods. In the European case, despite the larger number of stations in each of the defined stratigraphic classes, we appear to achieve little gain in reducing uncertainty, and indeed appear to incur larger uncertainty for doing so at shorter periods.

From a theoretical perspective, an increase in residual variability with the introduction of a random effect within a mixed effects regression can be a sign of problems in the underlying data set. These often correspond to certain "perils" of mixed effects regression
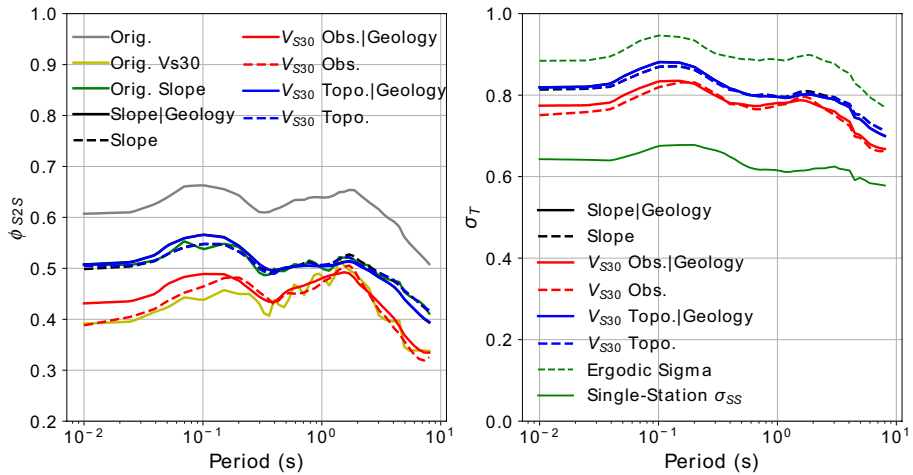
Fig. 17 Comparison of the site-to-site variability of the GMM for each of the difference predictors (left), and the resulting GMM total variability (right). "Orig." here refers to the original values provided in Kotha et al. (2020). Black lines indicating the slope cases lie behind the blue lines

described by Silk et al. (2020), namely in this case (i) a correlation between the fixed effects variables (in this case V_{S30} or slope) and the groups used as a random effect (i.e., V_{S30} and/or slope do show some correlation with geological age), (ii) imbalances in the number of samples for each level of random effect (see Fig. 10), and/or (iii) non-normality in the means of the random effects categories. An additional factor that contrasts the present analysis with that of Weatherill et al. (2020b) is the use of robust linear mixed effects regression firstly in fitting the original GMM from which $\delta S2S_S$ is defined and, secondly, in fitting the f_{SITE} term here. This approach reduces the random effects variances in the regressions by applying lower weights to data points further than $\pm 1.345\ \sigma$ from the fixed effects mean values. The a posteriori weights for $\delta S2S_S$ from the robust linear mixed effects model were used as prior weights on the data in both the robust linear regression without geology as a random effect and in the robust linear mixed effects regression. Other factors that contrast with the Kik-net data are the lower resolution geological data set, meaning that there may be a higher likelihood of misclassification of the stations, and potentially a higher degree of variability in site terms within each geological category as the stations in each category sample different regions of Europe.

From a purely statistical perspective it can be argued that if the introduction of a random effect into the regression yields an increase in the resulting residual variability then the more parsimonious course of action is to neglect the random effect and use ordinary (or in this case robust) linear regression instead. While the issue of the variance here deserves further analysis in future, we would argue in favour of retaining the mixed effects formulation in the current model for several reasons. i) The differences in the predicted amplification from one geological category to another are significant in terms of their impact on the resulting loss calculations. ii) The trend of increased variability with inclusion of geology as a mixed effect applies only to short period motion, and decreases are seen at longer periods of relevance. iii) The overall impact on the total aleatory variability is small and as the vulnerability models used in the ESRM20 converge around four intensity measure types (PGA, Sa (0.3 s), Sa (0.6 s), Sa (1.0 s)), the periods most affected by the increase (around T = 0.1 s) would not influence the loss estimates.

6 A complete regional site amplification model for Europe

6.1 The complete amplification model

The characterization of site amplification for seismic risk analysis at a European scale builds on the combination of the data sets presented in Sects. 3 and 4, with the empirically driven models of site amplification with respect to the Kotha et al (2020) GMM in Sect. 5. For the target resolution of 30 arc-seconds, every cell in Europe is assigned a value of slope, topographically inferred V_{S30} and geological unit (from the 8-category stratigraphic classification adopted previously). For earthquakes from shallow seismicity (excluding the stable craton region) the model of f_{SITE} developed using the regression between $\delta S2S_S$ and slope with geological unit as a random effect (Eq. 4) is adopted in place of the simple V_{S30} dependent predictor. The resulting mean amplification for each 30 arc-second cell, determined this time with respect to a "reference" condition of 0.3 m/m slope on a Precambrian rock unit, is shown in Figs. 18 and 19 for spectral acceleration at Sa (0.2 s) and Sa (1.0 s) respectively. The longer period amplification map highlights the influence of the soft sedimentary basins of the Po Plain, the Danube basin, the Pannonian basin (eastern Hungary) and the lower Rhine. Conversely, in many upland regions, such as the Alps, Pyrenees, Dinarides and Caledonides, we observe far less amplification, reflecting the influence not only of the higher slope values but of the predominance of older rock close to the surface.

The approach adopted for characterising the ground motion at the surface is specifically defined for application in a probabilistic seismic risk context. It ensures that for every 30 arc-second cell in Europe we are able to, in some way, account for local site effects and we do so in a manner that is consistent with the GMM adopted for shallow seismicity in the ESHM20. From an implementation perspective, the approach is practical, computationally efficient and accounts for the additional uncertainty in the ground motion that comes from adopting *inferred* proxies at a regional scale. By accounting for the differences in the relation between the proxy site property and the resulting amplification for different geological units, we can capture more of the spatial complexities that may exist from one region to another.

Though useful for understanding how the topography and geology influence the pattern of amplification across Europe, the maps shown in Figs. 18 and 19 give only a partial picture of the amplification model itself and how it will be used in the seismic risk calculations. The degree of amplification is connected not only with the ground motion model to which it should be applied, but it is also dependent on whether the site property represents a *measured* quantity at a given location or a regional mapped (or *inferred*) quantity. In the case of Figs. 18 and 19, the model is representative of an *inferred* property and though the "average" amplification is lower, the uncertainty is higher. This latter property is accounted for in the seismic hazard and risk calculation but not in the maps presented here. The maps shown cannot be considered as "general" maps of amplification that can be readily transferred to other GMMs or other contexts. This is an important caveat to be aware of if intending to integrate models such as this into, for example, a ShakeMap system.

The practicality of the predominantly data driven approach adopted here enables the method to be applied at European scale. There are, however, some limitations in this approach that could be refined with improved data in the future. By calibrating f_{SITE} exclusively on the $\delta S2S_S$ with respect to the form of the Kotha et al. (2020) GMM without a specific site term, we cannot separate easily the respective contributions of different effects. For example, short period $\delta S2S_S$ at certain sites may contain records from smaller
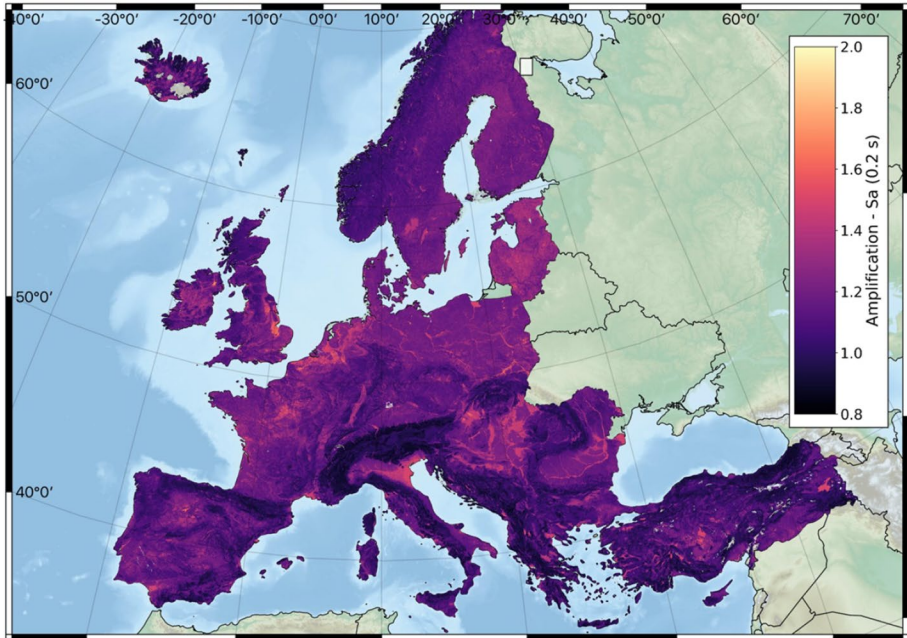
Fig. 18 Spatial pattern of mean amplification at a spectral period of T = 0.2 s with respect to a slope of 0.3 m/m on Precambrian rock. Amplification is not determined for areas classified as "Ice" or "Lake/Inland Water"
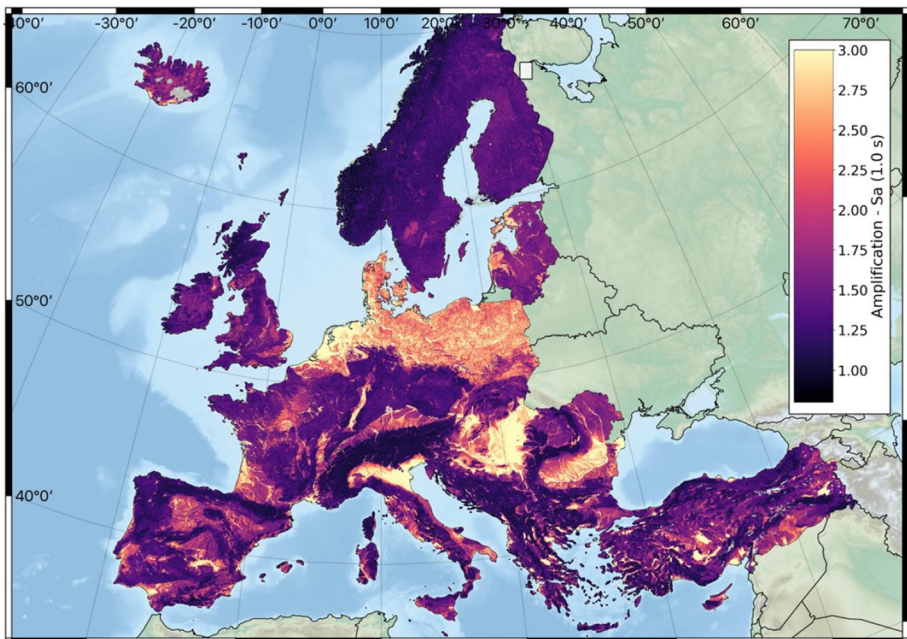

Fig. 19 As Fig. 18, for a spectral period of T = 1.0 s

earthquakes producing predominantly linear amplification, as well as from larger earthquakes for which nonlinear amplification is present. In the underlying ESM dataset, there are few observations from large earthquakes on soft soil sites; an observation confirmed by Guéguen et al. (2019) who found less evidence of soil nonlinearity in ESM with respect to the other databases. This ultimately means that evidence of nonlinearity is difficult to identify in the observed data, the amplification models defined here are likely to reflect mostly the linear component of amplification. An attempt was made to integrate the non-linear term of the Seyhan and Stewart (2014) amplification model into f_{SITE}, revising the fit of the linear model in Eq. 4 to be calibrated only on stations with records whose PGAs did not exceed 0.02 g. The resulting model, however, produced an unduly large degree of non-linearity and expected motions on soft soils were far lower than those produced by other models with nonlinear amplification terms (e.g., Akkar et al. 2014a; Boore et al. 2014; Chiou and Youngs 2014). Though we do not suggest that nonlinearity has no influence upon amplification within Europe, the data are simply insufficient to constrain it in the current model. Future research in this area towards the calibration of a nonlinear amplification model is certainly warranted.

For longer period motion the amplification model here cannot distinguish which effects are due to amplification from shallow soils and which may be due to deep basin effects. In contrast to other databases such as those of the NGA West 2 (Ancheta et al. 2014) or Kik-net (NIED 2019), the ESM contains no information about the depth to bedrock for the stations present in the database. Though it was subsequently possible to retrieve the full V_S profiles for several hundred stations in Italy, Turkey and Greece, we find that in many cases the available profiles for stations in the deep basin regions, particularly in Northern Italy, do not reach the engineering bedrock (e.g. V_S 800 m/s–1000 m/s). Though precedent exists for attempting to identify basin depth using horizontal-to-vertical spectral ratios (HVSRs) from ambient noise, such information was not available as metadata within the flatfile. We did not attempt to use the fundamental frequency, f_0, from the HVSR of the earthquake response spectra to infer basin depth, even though this information could be obtained from the flatfile, owing to the substantial uncertainty this would produce in the resulting basin depth estimates. We did explore residuals of the f_{SITE} model to assess if a resulting trend could be identified with respect to inferred sediment thickness from the global model of Pelletier et al. (2016), but as was observed by Weatherill et al. (2020b) no robust trend could be found.

Given the dependency on mappable proxy data, and the potential limitations that come from constraining f_{SITE} exclusively on empirical data, the amplification model presented here should be considered as a reference baseline application. We believe it has several advantages over the current predominant practice of using topographically inferred V_{S30} as a direct substitute for observed V_{S30} in the GMM, but it does not constitute a replacement for detailed local microzonation studies when they are available. We emphasise strongly though, that there is nothing in this approach, or its application to seismic risk calculations in general, that prevents integration of detailed local scale site amplification models where they are available. These could take the form of better refined maps of *measured* V_{S30}, which can easily replace the *inferred*-property site model, or even sites with known amplification functions. Resources, time and the often proprietary and inconsistent microzonation data prevented us from pursing a systematic collation of local site amplification models on this occasion, but we encourage such an endeavour in the future as we now have a clear means by which this can be integrated into the risk calculation.

7 Site amplification approach for other ground motion models within the ESHM20

The amplification model adopted here is calibrated on the site-to-site residuals of the Kotha et al. (2020) GMM, which forms the backbone GMM for shallow seismicity in active and low seismicity non-cratonic regions in the ESHM20 logic tree. The complete logic tree, however, also takes into account ground motions in the stable shield region of northeastern Europe and from deep active subduction and non-subduction seismicity (e.g., the Vrancea deep seismic zone). In the case of the former, no data was available from ESM to constrain the local seismogenic properties, and instead a different approach for the construction of the backbone GMM was adopted that capitalises on the outcomes of the NGA East project (Goulet et al. 2018), further details of which can be found in Weatherill and Cotton (2020). For the subduction and deep seismicity earthquakes, while the ESM database did provide a sufficient set of records to identify and calibrate a backbone GMM and its epistemic uncertainties, in this case the Abrahamson et al. (2016) "BC Hydro" model, the number of stations was too small to attempt the sort of site model calibration shown here.

In both the stable shield region of the Baltic sea and surrounding countries as well as the subduction/deep seismicity regions of the Hellenic, Calabria, Cypriot arcs and the Vrancea deep seismic zone, we adopt the site amplification terms directly from the models themselves and apply them to the topographically inferred V_{S30} values for all the cells within the regions for which these GMMs influence. A reasonable question to ask when adopting this approach, is whether or not we are being consistent in the treatment of site uncertainty by mixing the conventional and new approaches here. In the craton region, Weatherill and Cotton (2020) adopted the NGA East site amplification model of Stewart et al. (2020) and Hashash et al. (2020) without modification. This model does not make a separation between *measured* and *inferred* ϕ_{S2s} as we have done here; however, the epistemic uncertainty in the site response was taken into account by including logic tree branches to account for the epistemic uncertainty in the median site scaling term (Stewart et al. 2020). Comparisons of the *inferred* ϕ_{S2S} values in the current model with those adopted for the Central and Eastern United States by Stewart et al. (2019) showed that the two were similar in terms of absolute value and the relative change with period.

For subduction and deep seismicity regions we find that total within event-variability (ϕ) of the Abrahamson et al. (2016) subduction model is fixed for all periods at 0.6, a value lower than that of both the shallow crustal seismicity and craton model estimates found here. A corresponding ϕ_{S2S} term for this model is provided in the original BC Hydro report, which is given as 0.43. This value is comparable to the *measured* V_{S30} case used here, which is unsurprising given that the derivation of the ϕ_{S2S} term was based on the subset of stations in the BC Hydro database with more than 10 records, which is mostly limited to data from Taiwan and Japan where the vast majority V_{S30} values for the sites were taken from direct measurement. Recent analysis by Abrahamson and Gulurce (2020) of within-event residuals from the more extensive NGA Subduction database finds not only a degree of region-to-region variability in ϕ_{S2S} but also a distance dependence in ϕ_0 that results in higher within-event variability for longer distance records than for the previous Abrahamson et al. (2016) model. Their resulting estimates of ϕ yield values comparable to the present *inferred* V_{S30} or slope and geology models here for regions such as South America, Central America and Japan, but closer

to the original Abrahamson et al. (2016) models for New Zealand, Taiwan, Alaska and Cascadia.

When combined, the more recent investigations build a more complex picture of ϕ_{S2S} globally than the previous models would infer and suggest there is a case for adopting higher values than those assumed by the Abrahamson et al. (2016) GMM in some regions. In spite of this, we cannot calibrate a more appropriate ϕ_{S2S} from the data available and attempts to increase it to bring it closer to the *inferred* V_{S30} case do not have a sufficiently rigorous basis and may overestimate site-to-site variability from subduction earthquakes in Europe. Therefore, for subduction and deep seismicity events we adopt the V_{S30} values implied from topography in Sect. 3 and apply no further modification of the GMM. We recognise that this may underestimate the within-event variability from subduction earthquakes, however, and suggest this as an area that warrants further research in light of the publication of the NGA Subduction database (Bozorgnia and Stewart 2020) and associated GMMs.

8 Comparing site amplification characterisation in the context of probabilistic seismic hazard and risk analysis

The site model presented in the previous sections allows us to characterise ground motion at the surface for each location in Europe. Though there are few alternative approaches available to us for modelling site response at such a large spatial scale, it is important to understand what the impact this approach has on the resulting seismic hazard and risk calculations when compared against other methods that could be applied where better data are available, for example from a city scale microzonation. The aim of such a comparison is not necessarily to determine the *optimum* approach that could be taken at a site, but to understand where the resulting hazard and risk curves for ground motion at the surface might sit within the range of hazard and/or loss curves that could be produced when applying methods that depend on measured site data. Essentially, we wish to know if the complete site characterisation approach we are adopting for regional scale application, with the uncertainties fully propagated into the hazard and risk calculations, yields estimates of hazard or loss that would be significantly different from those determined using approaches that could be applied, were better data available.

For this purpose, we have assembled a data set of locations across Europe for which we have complete V_s profiles. At these locations we can run seismic hazard and risk computations adopting both the regional scale approaches used for the ESRM20 (i.e., amplification based on slope and geology or amplification based on topographically inferred V_{S30}) with *some* that could be achieved at a site-specific level. The complete database of V_S profiles covers Italy (212 locations), Turkey (219 locations) and Greece (19 locations) and is taken from the European Strong Motion Database (https://esm-db.eu/#/home) and the European Geotechnical Database (http://egd-epos.civil.auth.gr/). Though a measured V_{S30} can be identified from each of these profiles, not all profiles are sufficiently deep to identify 800 m/s V_S layer or the 1000 m/s V_S layer. We therefore limited analysis to just the subset of 172 profiles for which an observed V_{S30} and depth to 1000 m/s V_S could be obtained. This will obviously exclude more sites in deeper sedimentary basins, particularly in the lower basin of the River Po, Northern Italy, than on stiffer soil and rock; however, in terms of geological composition the resulting database still provides an adequate sample of profiles on different geological conditions.

We separate the different site amplification modelling approaches into those based on *regional* data and those based on the V_S profile data. For the *regional* data we consider two approaches, the first is based exclusively on *inferred* V_{S30} (without geology) using Eq. 4, and the second based directly on slope and geology using the proposed amplification model for Europe. In both cases the ϕ_{S2S} values reflect an *inferred* site case and would therefore be higher *theoretically* than the ϕ_{S2S} when the site properties are taken from the V_S profile. In the present case we do not undertake a state-of-the-art site-specific PSHA analysis in which both the ground motion model and its site-specific amplification function are calibrated to the profile (e.g., Rodriguez-Marek et al. 2014; Al Atik et al. 2014). Instead, we use methodologies that could be applied based on a reasonable microzonation study of a city in which the shearwave velocity and depth to a reference bedrock can be mapped. These methods are as follows:

1. Amplification of the ground motion based on the *proposed* Eurocode 8 amplification model for the case when both the depth to the reference V_S 800 m/s rock layer (h_{800}) and the average shearwave velocity in the soil column above the reference layer ($V_{S,h}$) are known (where $V_{S,h} = V_{S,30}$ in the case that $h_{800} \geq 30$ m), which we refer to as EC8.
2. Amplification of the ground motion based on the proposed Eurocode 8 amplification factors for the case when the $V_{S,h}$ is not known, i.e. the Eurocode 8 site class is inferred from other information (EC8 Default)
3. Characterisation and amplification of the ground motion according to the design code amplification model of Pitilakis et al. (2018)
4. The empirical nonlinear amplification model of Sandikkaya and Dinsever (2018), based on V_{S30} and depth to the 1 km/s V_S layer ($Z_{1.0}$).
5. The ESHM20 amplification model based on *measured* V_{S30} (Eq. 3)

Methods (1) to (3) are based on design code amplification factors, which require some adaptation for application into the framework needed for comparison with the empirical approaches based on the ground motion model (Method 5) directly or some a posteriori amplification of ground motion (e.g., Sandikkaya and Dinsever 2018). To apply these methods, we firstly need to assign each site to the corresponding design code class. In the case of Eurocode 8 this is done directly using the V_S profile (EC8), while the Pitilakis et al. (2018) classification requires characterisation of the fundamental period of the H/V spectral ratio (T_0), which is not available for the profiles in question. Where the T_0 is necessary to distinguish the particular site class we assign the resulting site class by judgement taking into consideration the available V_S profile, the depth to the rock layer and the corresponding impedance contrast. We acknowledge that the application of the Eurocode amplification factors in this context is different from how these factors would often be applied in practice, which is to the ground motion corresponding to the design level seismic hazard. Nevertheless, as such factors are calibrated from both observed measurements and model predictions of site amplification including nonlinearity where relevant (Paolucci et al. 2021), we believe that they represent a legitimate model of site amplification that may be adopted in local or city scale risk analysis, against which our approach can be compared.

To apply the design-code amplification factors it is necessary to define the acceleration on reference rock (V_S 800 m/s) for the short period coefficient (S_S) and the long-period

coefficient (S_1). In the case of the long-period coefficient this is equivalent to the spectral acceleration at $T = 1.0$ s. For the short period coefficient (S_S) no standard definition of this term has been agreed at the time of writing. From the definition of the design code shape, S_S should be 2.5 times greater than the very high frequency acceleration Sa ($T \leq 0.03$ s), so we adopt this definition and anchor S_S to 2.5 times the PGA. As we do not attempt a vector seismic hazard calculation (e.g., Bazzurro and Cornell 2002) the PGA is preferred as the short period anchor rather than, say, a short-period acceleration corresponding more closely to the peak of the design spectrum, in order to minimise the correlation between the short- and long-period anchoring acceleration values. The design code amplification is incorporated into the seismic hazard calculation in a similar manner to a ground motion model, meaning that for any given period of spectral acceleration, T, we first predict the PGA and Sa (1.0 s) on reference rock to determine S_S and S_1, then use these values to determine amplification at T and then apply this to the acceleration on rock at period T. Given the dependency of the amplification on the acceleration on bedrock, this approach is not directly equivalent to applying the design code amplification factors to the uniform hazard spectrum at a given return period (as might be done in practice), but rather aims to treat the design code amplification factors as nonlinear amplification models with respect ground motion on a reference rock surface.

8.1 Impact on seismic hazard analysis

From the 172 sites at which seismic hazard was calculated, in Fig. 20 we show results from selected examples representing a "rock" site, a "stiff soil" site and a deep "soft soil" site. All three sites correspond to areas of moderate to high hazard. For each of these sites a *measured* and *inferred* V_{S30} have been determined, along with Eurocode 8 and Pitilakis et al. (2018) site classes, $Z_{1.0}$ and topographic slope. These initial sites are chosen as they correspond to cases where the *inferred* and *measured* V_{S30} are in broad agreement with one another. The dashed lines in the figures correspond to the two cases based on *regional* data. As a general observation, we first note that the spread of amplification values with the range of annual probabilities of exceedances (APoEs) of interest can be quite variable depending on the site properties and the spectral period. Arguably the lowest spread can be seen for the stiff soil sites and the highest for the soft soil sites. At the rock site (EC8 A, Jurassic-Triassic) the Pitilakis et al. (2018) models appear to result in the highest hazard, though this may be due to the fact that the EC8(P18) classification at this particular site is B1, compared to the A classification assumed for Eurocode 8. For the soft soil sites, the Sanikkaya and Dinsever (2018) model produces the highest hazard, possibly reflecting the strong influence of the basin depth term. In several cases it is noteworthy that the curves for the Eurocode 8 model are barely visible in Fig. 19, meaning that they yield highly similar level of seismic hazard as other approaches.

They key result of interest for us here, is that the hazard curves based on amplification for the regional models (ESHM20 *inferred* V_{S30} and ESHM20 slope/geology) fall within the range of values suggested by other approaches that could be applied were the site properties better known. This result should be contrasted against the implied amplification for the *measured* and *inferred* V_{S30} cases shown in Figs. 14 and 15. Those figures indicate that the median amplification for the *measured* V_{S30} case is notably higher than for the inferred

cases. Seismic hazard curves, however, are controlled not only by the median amplification but also by changes to the aleatory uncertainty. By increasing the ϕ_{S2S} values for the regional models we compensate for the lower median amplification and, as a result, retrieve comparable probabilistic seismic hazard curves. This is a critical point to emphasise in interpreting the model, as the amplification itself cannot be decoupled from the definition of the aleatory uncertainty. If the *inferred* or *regional* properties are to be used then the additional uncertainty must be propagated into the ground motion model, and ultimately into the hazard and risk analysis. Failure to do so will inevitably result in an underestimation of seismic hazard.

8.2 Impact on seismic risk analysis

The comparison of the seismic hazard curves in Fig. 20 illustrate the differences in probabilities of exceedance of ground motion when adopting different site amplification models, but for the ESRM20 the target is not the hazard but ultimately the probability of exceedance of losses. The complete set of seismic risk results based on the site amplification model developed here can be found in Crowley et al. (2021) and will not be repeated here. Instead, we extend the seismic hazard analysis for the three aforementioned sites into the domain of seismic risk by comparing the differences between the site amplification models in terms of loss exceedance curves for three structures. The selected structure is reinforced concrete, with a load bearing wall lateral load resisting system with moderate ductility: CR-LWAL-DUM according to the Global Earthquake Model (GEM) Building Taxonomy (https://github.com/gem/gem_taxonomy). To understand how the different periods of spectral acceleration are affected we consider three structures whose fragility models depend on different intensity measures: 3-storey (IM = PGA), 7-storey (IM = Sa (0.6 s)) and 12-storey (IM = Sa(1.0 s)). These cover a range of low- to medium-rise reinforced concrete structures with a "moderate" level of seismic resistance. The CR-LWAL-DUM is selected as the periods of the desired IMs over the range of building heights span well the spectral period range we are considering here. The general results we show here would be generalisable to any other fragility and vulnerability models considered within the ESRM20 (available from https://doi.org/10.5281/zenodo.4062410). Loss curves in terms of annual probability of exceeding a given loss ratio of the structure are calculated using the classical seismic risk calculator of OpenQuake (Silva et al. 2014; Silva 2018). The vulnerability model for each structure models the distribution of loss ratio for each given intensity measure level as a lognormal distribution. We also include coefficient of variation, which we determine using the model of Silva (2019). The seismic loss curves are shown for the three different sites ("rock", "stiff soil" and "soft soil") and building heights (H3, H7 and H12) in Fig. 21.
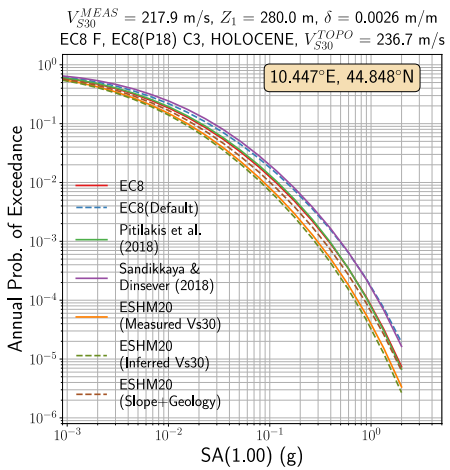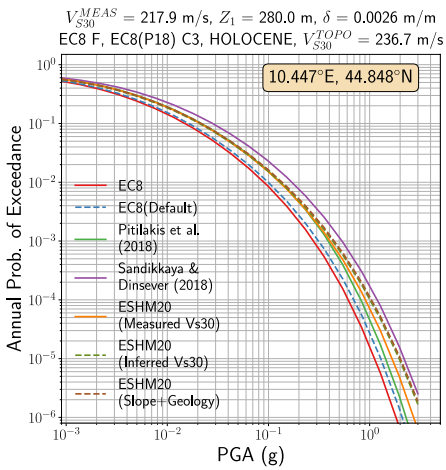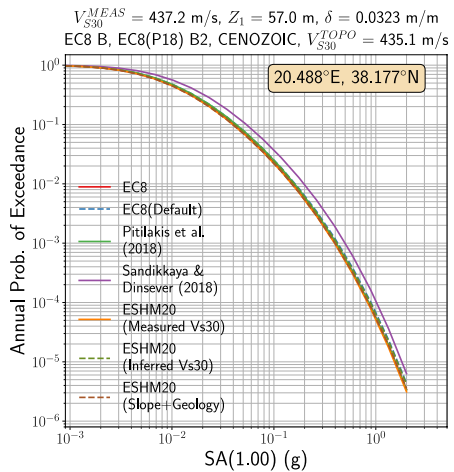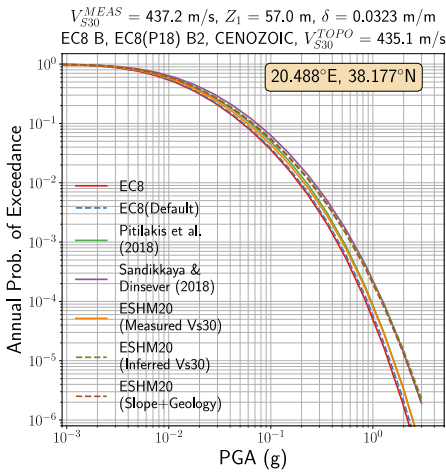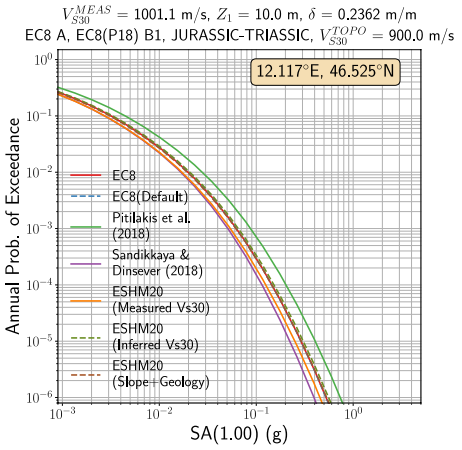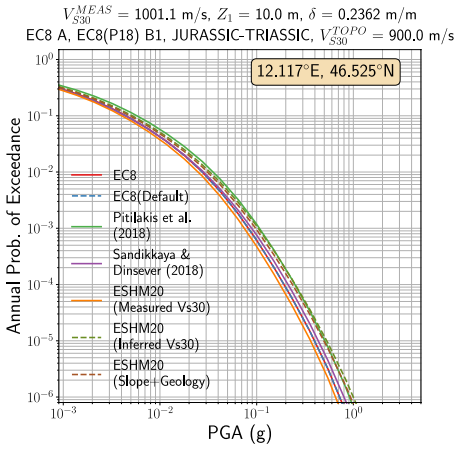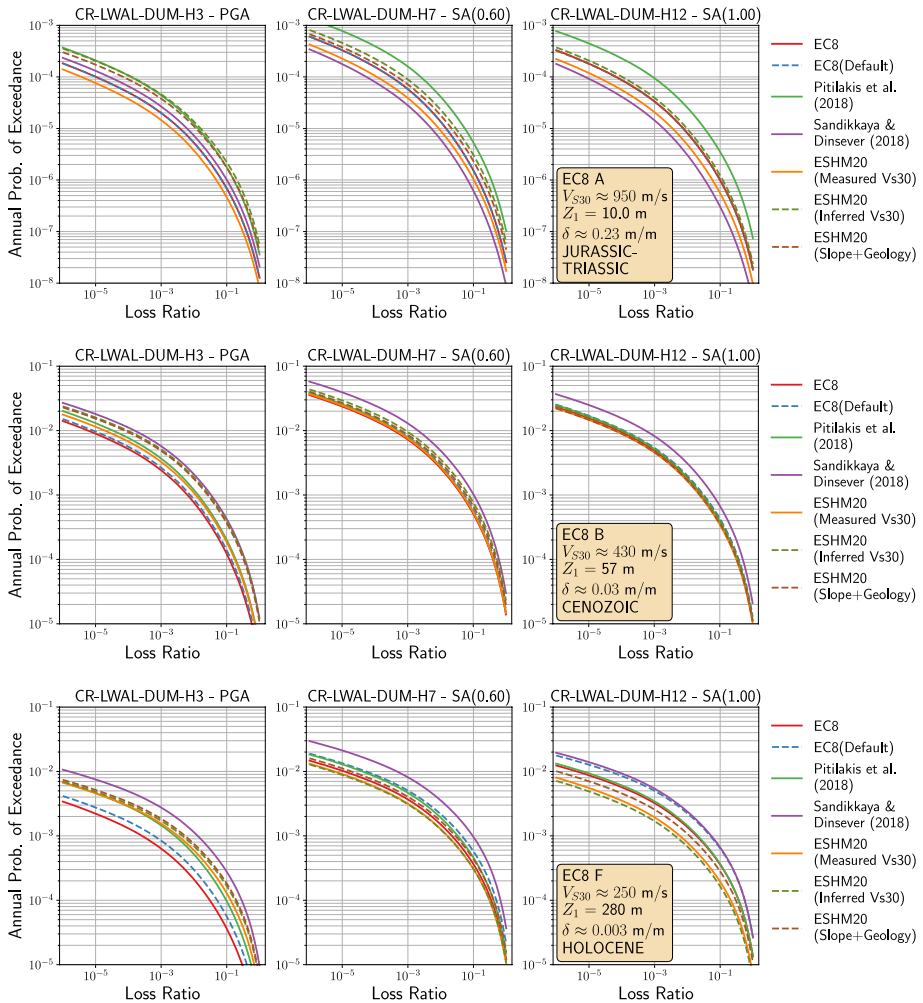
Fig. 21 Comparison of seismic loss curves in terms of annual probability of exceedance of loss ratio assuming different approaches for site amplification. The three different structure heights are considered (3-, 7- and 12 storey) for the same structure located on sites of different soil types: rock (top), stiff soil (middle) and soft soil (bottom)

By viewing the comparison amplification models through the lens of seismic risk rather than hazard, we can see that differences between methods are exacerbated. For the "rock" site, where the Pitilakis et al. (2018) amplification model tends to produce higher seismic hazard, the loss estimates are substantially larger than those resulting from other amplification methods. On the "stiff soil" and "soft soil" sites Sandikkaya and Dinsever (2018) model is the outlier, producing greater losses for medium-rise structures than other approaches. As with the seismic hazard curves, we generally find the slope and geology amplification model yielding results that are toward the middle of the range of values

produced from the different approaches. In none of the cases shown is the ESHM20 slope and geology method an outlier, though for very short period motion on rock it is among the curves indicating higher losses, along with the Pitilakis et al. (2018) amplification factors.

These comparisons of the different amplification methods are intended to help understand whether or not the method modelling site amplification for *inferred* properties yield significantly different hazard and risk results when compared against methods that could be applied when *some* detailed site data is available (e.g., measured V_{S30}, depth to bedrock, Eurocode 8 site class etc.). For the examples shown in Fig. 21, although the loss curves may appear similar in many cases the ratio of the higher losses to the lowest loss is close to a factor of 2–3 at lower loss ratios, increasing to around 5–7 at the highest loss ratios (even as large as a factor of 12–15 for the largest outliers). Both the *inferred* V_{S30} model and the slope and geology model produce hazard or risk results that are comfortably within this range, albeit skewed slightly toward the lower end with a factor of 1.5–2 above the minimum loss curves depending on the period and soil type. Though we do not propose that the simplified amplification models should be adopted in place of some of the alternative approaches when the detailed site information is available, these results suggest that at the proposed "simplified" approaches to site amplification yield hazard and loss curves that lie comfortably *within* the considerable range of values that can be obtained when using different approaches to site amplification, including cases when better data are available. However, the comparisons shown here do not necessarily represent the wide variety of site conditions that can occur in the real world. In other locations greater disparities can be found, particularly in upland regions where low V_{S30} measurements can arise in areas of potentially steeper topography and older geological rock owing to nearby fluvial features or recent (Holocene) mobilisation of sediments at the foot of hillsides. In these cases, neither a slope/geology approach nor a topographically inferred V_{S30} approach would be able to capture well the fine-scale variation in site response.

9 Conclusions

The process implemented for the development of the site response model within the ESHM20 and ESRM20 primarily occupies itself with the challenge of application at scale, rather than at a site-specific level. The key outcome is that we are able to characterise ground shaking at the surface at every 30 arc-second grid cell in the ESRM20 target countries and have implemented this in a manner that is practical for the seismic risk calculations and integrates uncertainties that have until now commonly been neglected. Several standalone products have emerged from this effort, which may be of value to the seismology and engineering communities in general: (i) a harmonised map of surficial geology, (ii) a new European model of V_{S30} constructed from slope and geology, (iii) a database of V_{S30} observations compiled from multiple sources, (iv) ground motion models (and a ground motion model logic tree) for Europe with site scaling terms calibrated on the largest data set of strong motions and site metadata available across the region, which can be adapted to the appropriate context of the calculation.

The data sets themselves are released as part of the complete suite of products to emerge from the ESRM20 and are available for download from http://risk.efehr.org/site-model/. The ground motion models and their corresponding site scaling terms have been implemented in the OpenQuake-engine (http://github.com/gem/oq-engine), a state-of-the-art open source software for seismic hazard and risk analysis that is the primary calculation tool for the ESHM20 and ESRM20 calculations. In addition, as the input data for seismic hazard and risk calculations requires a more complete suite of information, a dedicated open-source Python toolkit is available to help users access all of the site input data and construct the site model for their calculation needs. Further information on how the site data can be optimised to best suit the exposure model in a seismic risk calculation can be found in Dabbeek et al. (2021), and the toolkit ("exposure2site") is available for download from https://gitlab.seismo.ethz.ch/efehr/esrm20_sitemodel.

The various calibrations of the f_{SITE} and ϕ_{S2S} terms for the shallow earthquake ground motion model pertain to different applications, and care must be taken to ensure that the most appropriate is used. To summarise these briefly:

1. For the development of the seismic hazard maps with respect to the Eurocode 8 reference rock, or similar applications in national seismic hazard maps that may form the input for seismic design codes, the amplification model $f_{SITE}\left(V_{S30}^{OBS}, T\right)$ from Eq. 3, and its corresponding ϕ_{S2S}^{OBS}, should be used. This is the case for the seismic hazard products on reference rock for ESHM20. In this case we expect the V_{S30} to be known at the site, and when adopting soil amplification factors from the design code itself a pre-determined level of conservatism is already accounted for, so any further inflation of ϕ_{S2S} would likely overestimate uncertainty.

2. For characterisation of ground shaking at the Earth's surface at regional scale, either in the context of a probabilistic seismic hazard and/or risk analysis or for rapid assessment of shaking and losses in a Shakemap/PAGER framework, we advise using $f_{SITE}(\theta, T, geology)$ defined in Eq. 4, where θ refers either to *inferred* V_{S30} (V_{S30}^{INF}) or to topographic slope depending on the desired application. Critically, however, the ϕ_{S2S} term *must* correspond to the ϕ_{S2S}^{INF} case in order to reflect the increased uncertainty in the resulting amplification factor. This approach is adopted in the ESRM20 using the 30 arc-second topographic slope and geology data set.

3. Amplification maps in Figs. 17 and 18 are shown for illustrative purposes and should not be adopted as a basis for amplification of ground motions for other contexts without the inclusion of the appropriate uncertainties. Such maps will therefore not be distributed as an official product of ESRM20. We acknowledge, however, that interested users of the amplification model may wish to explore it in more detail and that future efforts for site characterisation at national or European scale may wish to use this model as a reference for comparison. We therefore include functionality to construct the amplification maps, with respect to a user-defined reference slope and geology condition, into the aforementioned "exposure2site" toolkit.

As we have illustrated throughout the development of the model, the methodology for mapping site conditions at large scale permits first order estimations of site conditions and should not be applied for local purpose (e.g., microzonation). These methods

lead to information that is useful where no precise site information is available, and their objective is to integrate site conditions in large scale seismic hazards and risks maps in order to inform and to harmonize such information.

The implementation of the site characterisation methodology for ESRM20 serves to highlight some of the future developments that would be needed in order to reconcile local scale and regional scale approaches for the future generations of seismic risk analysis in Europe. The most fundamental of these would be a harmonised database of microzonation studies and a greatly expanded effort to improve the site metadata for current (and future) seismic recording stations, which themselves would benefit from enhanced resolution and availability of digital geological data sets. While the benefits of improved site metadata would be seen in improved calibrations of the amplification terms in the manner implemented here, the microzonation studies would likely have a greater impact in terms of the resulting loss estimates owing to a reduction of site-to-site variability in regions of highest economic exposure. The framework developed here is entirely amenable to a modular approach whereby detailed characterisation of site properties for a city from microzonation data can supplant the regional scale models wherever they are available.

The illustrative probabilistic seismic hazard and risk applications of the proposed site response model in Sect. 7, and the comparison with other approaches that would be feasible with more detailed site information, demonstrate that with the appropriate propagation of uncertainty the resulting loss estimates are comparable. This provides some re-assurance that the proposed approach would not necessarily produce loss results that diverge substantially from those that would be achieved using other methods that require more detailed site information. Though the results shown here do not necessarily exhaust the range of possible local conditions that one might encounter at a European scale, they do highlight the importance of appraising site characterisation approaches not just in terms of the resulting amplification but also in terms of the contexts in which they are to be applied. In doing so, one is able to appraise both the practicality of implementation of any proposed site characterisation strategy as well as the impact of its uncertainties. We would encourage future developments in site response modelling for seismic risk to adopt and expand strategies such as these in the future.

Appendix A

See Tables 3, 4, 5

Table 3 Simplified lithologies and codes for the European lithological map shown in Fig. 1

Simplified lithological description	Lithological code
Alluvium	ALLU
Conglomerates	CGRT
Clay	CLAY
Clastic Rocks, pelites, shales and silts	CLST
Coal basins	COAL
Formations with dominant flints	FLINT
Impactite	IMPCT
Felsic, mafic and ultramafic igneous rocks	IGNEOUS
Lakes	LAKES
Limestone, dolomites, chalk and carbonaceous formations (peat)	LMST
Marls	MARL
Metamorphic rocks (from micaschists to migmatites)	METAM
Mud–unconsolidated sediments	MUD
Sands and Gravels	SAGR
Formations with salts	SALT
Schists and slates	SCHST
Undifferentiated sediments	SDMT
Soils and/or Loess	SFOR
Sandstones	SSTN
Tectonic breccias and mylonites	TECTO
Tills	TILL
Unconsolidated formations	UNCS
Volcanic formations	VOLC
Anthropic formations	WASTE

Table 4 Simplified stratigraphic information (epoch and systems) and corresponding codes for the pan-European stratigraphic map shown in Fig. 2

Detailed Epoch	Synthetic system	Stratigraphic code
HOLOCENE	HOLOCENE	HC
PLEISTOCENE	NEOGENE	NG
PLIOCENE	NEOGENE	NG
MIOCENE	NEOGENE	NG
NEOGENE	NEOGENE	NG
OLIGOCENE	PALEOGENE	PG
EOCENE	PALEOGENE	PG
PALEOGENE	PALEOGENE	PG
CENOZOIC	CENOZOIC	CN
CRETACEOUS	MESOZOIC	MS
JURASSIC	MESOZOIC	MS
TRIASSIC	MESOZOIC	MS
MESOZOIC	MESOZOIC	MS
PALEOZOIC	PALEOZOIC	PZ
PHANEROZOIC	PHANEROZOIC	PH
PRECAMBRIAN	PRECAMBRIAN	PK
UNKNOWN	UNKNOWN	UK

Table 5 Distributions of measured V_{S30} for each of the lithological units shown in Figs. 3 and 4, with corresponding codes defined in Table 3

Lithological codes	Statistical values of V_{S30} (m/s) for each lithological unit				
	Boxplot Q1	Median (Q2)	Boxplot Q3	Mean value	Standard deviation
MUD	362	397	450	435	177
SAND	192	271	450	345	188
SAGR	300	451	621	476	242
ALLU	237	290	403	348	183
SILT	478	576	673	576	138
CLAY	293	390	542	438	208
SDMT	311	374	580	460	211
CGRT	237	426	559	420	272
COAL	473	485	496	485	16
PLTE	421	587	753	587	235
CHLK	290	324	398	393	176
SCHST	345	510	758	604	363
TECTO	314	376	460	390	147
CLST	338	452	525	464	203
MARL	378	448	829	664	480
SALT	351	420	514	436	163
METAM	358	501	798	666	503
VOLC	315	399	485	425	148
SSTN	346	447	582	485	192
LMST	381	555	847	660	382
FELSIC	375	446	628	499	197
UMAFIC	362	433	611	482	141
MAFIC	799	853	895	845	96

Appendix B

See Fig. 22
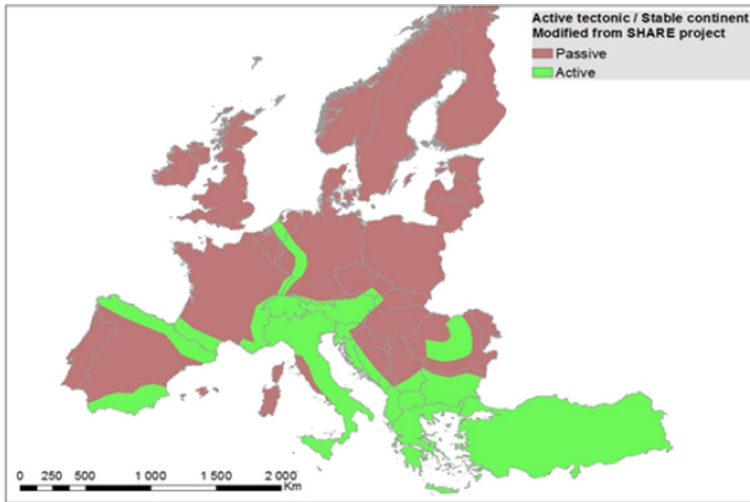
Fig. 22 Classification of active and stable regions used for development of the topographically inferred V_{S30} model based on Delavaud et al. (2012)

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to declare.

References

Abrahamson NA, Silva WJ, Kamai R (2014) Summary of the ASK14 ground motion relation for active crustal regions. Earthq Spectra 30(3):1025–1055

Abrahamson NA, Gregor N, Addo K (2016) BC hydro ground motion prediction equations for subduction earthquakes. Earthq Spectra 32(1):23–44

Abrahamson NA, Gülerce Z (2020) Regionalized ground-motion models for subduction earthquakes based on the NGA-SUB database. Tech. Rep. PEER Report No. 2020/25. Pacific Earthquake Engineering Research Center

Ahdi SK, Stewart JP, Ancheta TD, Kwak DY, Mitra D (2017) Development of VS profile database and proxy-based models for VS 30 prediction in the Pacific Northwest region of North America. Bull Seismol Soc Am 107(4):1781–1801

Akkar S, Sandikkaya A, Bommer JJ (2014a) Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. Bull Earthq Eng 12:359–387

Akkar S, Sandikkaya A, Senyurt M, Azari Sisi A et al (2014b) Reference database for seismic ground-motion in Europe (RESORCE). Bull Earthq Eng 12:311–339

Al Atik L, Kottke A, Abrahamson N, Hollenback J (2014) Kappa (κ) scaling of ground-motion prediction equations using an inverse random vibration theory approach. Bull Seismol Soc Am 104(1):336–346

Allen TI, Wald DJ (2009) On the use of high-resolution topographic data as a proxy for seismic site conditions (V_{S30}). Bull Seismol Soc Am 99(2A):935–943

Ancheta TD, Darragh RB, Stewart JP et al (2014) NGA-West2 database. Earthq Spectra 30(3):989–1005

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48

Bazzurro P, Cornell CA (2002) Vector-valued probabilistic seismic hazard analysis. In: Proceedings of the 7th U.S. National Conference on Earthquake Engineering, Boston MA, July 21 – 25, 2002

Bergamo P, Hammer C, Fäh D (2021) On the relation between empirical amplification and proxies measured at Swiss and Japanese stations: systematic regression analysis and neural network prediction of amplification. Bull Seismol Soc Am 111(1):101–120

Bindi D, Massa M, Luzi L, Ameri G, Pacor F, Puglia R, Augliera P (2014) Pan-European ground motion prediction equations for the average horizontal component of PGA, PGV and 5 %-damped PSA at spectral periods up to 3.0 s using the RESORCE data set. Bull Earthq Eng 12:391–430

Boore DM, Stewart JP, Seyhan E, Atkinson G (2014) NGA-West2 equations for predicting PGA, PGV, and 5 % damped PSA for shallow crustal earthquakes. Earthq Spectra 30(3):1057–1085

Bozorgnia Y, Stewart JP (2020) Data resources for the NGA-Subduction Project. Tech. Rep. PEER Technical Report No. 2020/02. Pacific Earthquake Engineering Research Center

Cassard D et al (2015) ProMine mineral databases: new tools to assess primary and secondary mineral resources in Europe. In: Weihed P (ed) 3D, 4D and predictive modelling of major mineral belts in Europe. Mineral resource reviews. Springer, Cham

Chen G, Magistrale H, Rong Y, Cheng J, Binselam SA, Xu X (2021) Seismic site condition of mainland China from geology. Seismol Res Lett 92(2A):998–1010

Chiou BS-J, Youngs RR (2014) Update of the Chiou and Youngs NGA model for the average horizontal component of peak ground motion and response spectra. Earthq Spectra 30(3):1117–1153

Crowley H, Dabbeek J, Despotaki V, Rodrigues D, Martins L, Silva V, Romão X, Pereira N, Weatherill G, Danciu L (2021) European Seismic Risk Model (ESRM20). In: EFEHR Technical Report 002 V1.0.0, https://doi.org/10.7414/EUC-EFEHR-TR002-ESRM20

Dabbeek J, Crowley H, Silva V, Weatherill G, Paul N, Nievas CI (2021) Impact of exposure spatial resolution on seismic loss estimates in regional portfolios. Bull Earthq Eng 19(14):5819–5841

Delavaud E, Cotton F, Akkar S et al (2012) Toward a ground-motion logic tree for probabilistic seismic hazardassessment in Europe. J Seismol 16:451–473. https://doi.org/10.1007/s10950-012-9281-z

Derras B, Bard P-Y, Cotton F (2017) 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. Earth Planets Space 69:133. https://doi.org/10.1186/s40623-017-0718-z

Di Capua G, Peppoloni S, Amanti M, Cipolloni C, Conte G (2016) Site classification map of Italy based on surface geology. Geol Soci Lond Eng Geol Spec Publ. 27(1):147–158

ETHZ, Swiss Seismological Service (SED) at ETH Zurich (2015) The site characterization database for seismic stations in Switzerland. Federal Institute of Technology, Zurich. https://doi.org/10.12686/sed-stationcharacterizationdb

Fäh D, Huggenberger P (2006) INTERREG III, Erdbebenmikrozonierung am südlichen Oberrhein. Zusammenfassung für das Projektgebiet Gebiet in der Schweiz. CD and Report (in German; available from the authors)

Fäh D, Havenith H, Roten D, Alvarez S and Giardini D (2007) Site effects in the rhone valley, switzerland: measurements, observations, modelling and verification of the building code spectra. Interreg Project SISMOVALP Seismic Hazard and Alpine Valley Response Analysis. Final Report Swiss Seismological Service ETH Zurich, 18.1.2007

Forte G, Chioccarelli E, De Falco M, Cito P, Santo A, Iervolino I (2019) Seismic soil classification of Italy based on surface geology and shear-wave velocity measurements. Soil Dyn Earthq Eng 122:79–93

Foster KM, Bradley BA, McGann CR, Wotherspoon LM (2019) A V_{S30} map for New Zealand based on geological and terrain proxy variables and field measurements. Earthq Spectra 35(4):1865–1897

Garofalo F, Foti S, Hollender F, Bard PY et al (2016) InterPACIFIC project: Part II: inter-comparison between surface-wave and borehole methods. Soil Dyn Earthq Eng 82:241–254

Gehl P, Bonilla F, Douglas J (2011) Accounting for site characterization uncertainties when developing ground-motion prediction equations. Bull Seismol Soc Am 101(3):1101–1108

Goulet C, Bozorgnia Y, Abrahamson N, et al (2018) Central and Eastern North America ground-motion characterisation: NGA-East final report. Tech. Rep. PEER Report No. 2017/03, Pacific Earthquake Engineering Research Center

Guégen P, Bonilla F, Douglas J (2019) Comparison of soil nonlinearity (In Situ Stress-Strain Relation and G/Gmax Reduction) observed in strong-motion databases and modeled in ground-motion prediction equations. Bull Seismol Soc Am 109(1):178–186

Hashash YMA, Ilhan O, Harmon JA et al (2020) Nonlinear site amplification model for ergodic seismic hazard analysis in Central and Eastern North America. Earthq Spectra 36(1):69–86

Havenith H-B, Fäh D, Polom U, Roulle A (2007) S-wave velocity measurements applied to the seismic microzonation of Basel, Upper Rhine Graben. Geophys J Int 170:346–358

Iwahashi J, Pike RJ (2007) Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. Geomorphology 86(3–4):409–440

Johannesson H (2014) Geological map of Iceland 1:600,000 – Bedrock Geology – Icelandic Institute of Natural History. https://en.ni.is/resources/publications/maps/geological-maps#&gid=null&pid=1

Koller M (2016) robustlmm: an R package for robust estimation of linear mixed-effects models. J Stat Softw 75:1–24

Kotha S-R, Bindi D, Cotton F (2016) Partially non-ergodic region specific GMPE for Europe and Middle East. Bull Earthq Eng 14:1245–1263

Kotha S-R, Cotton F, Bindi D (2018) A new approach to site classification: mixed-effects ground motion prediction equation with spectral clustering of site amplification factors. Soil Dyn Earthq Eng 110:318–329

Kotha S-R, Weatherill G, Cotton F (2020) A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. Bull Earthq Eng 18:4091–4125

Kuehn NM, Abrahamson NA (2018) The effect of uncertainty in predictor variables on the estimation of ground-motion prediction equations. Bull Seismol Soc Am 108(1):358–370

Lanzano G, Sgobba S, Luzi L et al (2019) The pan-European engineering strong motion (ESM) flatfile: compilation criteria and data statistics. Bull Earthq Eng 17:561–582

Lee CT, Tsai BR (2008) Mapping Vs30 in Taiwan. TAO Terr Atmos Ocean Sci 19(6):6

Lemoine A, Douglas J, Cotton F (2012) Testing the applicability of correlations between topographic slope and VS 30 for Europe. Bull Seismol Soc Am 102(6):2585–2599

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

McPherson A, Hall L (2013) Site classification for earthquake hazard and risk assessment in Australia. Bull Seismol Soc Am 103(2A):1085–1102

Miller AC, Rice TR (1983) Discrete approximations of probability distributions. Manage Sci 29(3):352–362

NIED K-NET (2019), KiK-net, National Research Institute for Earth Science and Disaster Resilience, https://doi.org/10.17598/NIED.0004

Noorlandt R, Kruiver PP, de Kleine MP, Karaoulis M, de Lange G, Di Matteo A, Bommer JJ (2018) Characterisation of ground motion recording stations in the Groningen gas field. J Seismol 22(3):605–623

Panzera F, Bergamo P, Fäh D (2021) Canonical correlation analysis based on site-response proxies to predict Site-specific amplification functions in Swittzerland. Bull Seismol Soc Am 111(4):1905–1920

Paolucci R, Aimar M, Ciancimino A, Dotti M et al (2021) Checking the site categorization criteria and amplification factors of the 2021 draft Eurocode 8 Part 1–1. Bull Earthq Eng 19:4199–4234

Pelletier JD, Broxton PD, Hazenberg P et al (2016) A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modelling. JAdv Model Earth Syst 8:41–65

Pilz M, Cotton F, Razafindrakoto H, Weatherill G, Spies T (2021) Regional broad-band ground-shaking modelling over extended and thick sedimentary basins: an example from the Lower Rhine Embayment (Germany). Bull Earthq Eng 19:581–603

Pitilakis K, Riga E, Anastasiadis A et al (2018) Towards the revision of EC8: proposal for an alternative site classification scheme and associated intensity dependent spectral amplification factors. Soil Dyn Earthq Eng 126:105137

Poggi V (2011) The use of surface waves for site characterization and seismic hazard analysis. PhD Thesis. Swiss Federal Institute of Technology Zürich (ETHZ). http://seismo.org/Biblio/PhD_Thesis/Poggi_PhD_Thesis_2011.pdf

Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F, Stafford PJ (2014) Application of single-station sigma and site response characterization in a probabilistic seismic hazard analysis for a new nuclear site. Bull Seismol Soc Am 104(4):1601–1619

Rodriguez-Marek A, Bommer JJ, Youngs RR, Crespo MJ, Stafford PJ, Bahrampouri M (2021) Capturing epistemic uncertainty in site response. Earthq Spectra 37(2):921–936

Roullé A, Auclair S, Dewez T, Hohmann A., Lemoine A, Rey J (2010) Cartographie automatique des classes de sol à l'échelle régionale à partir d'un modèle numérique de terrain ou de surface. Rapport final. BRGM/RP-58853-FR

Sandikkaya MA, Dinsever LD (2018) A site amplification model for crustal earthquakes. MDPI Geosci 8:264

Seyhan E, Stewart JP (2014) Semi-empirical nonlinear site amplification from NGA-West2 data and simulations. Earthq Spectra 30(3):1241–1256

Seyhan E, Stewart JP, Ancheta TD, Darragh RB, Graves RW (2014) NGA-West2 site database. Earthq Spectra 30(3):1007–1024

Silk MJ, Harrison XA, Hodgson DJ (2020) Perils and pitfalls of mixed effects regression in biology. PeerJ 8:e9522. https://doi.org/10.7717/peerj.9522

Silva V (2018) Critical issues on probabilistic earthquake loss assessment. J Earthq Eng 22(9):1683–1709

Silva V (2019) Uncertainty and correlation in seismic vulnerability functions of building classes. Earthq Spectra 35(4):1515–1539

Silva V, Crowley H, Pagani M, Monelli D, Pinho R (2014) Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment. Nat Hazards 72(3):1409–1427

Silva V, Amo-Oduro D, Calderon A, Costa C et al (2020) Development of a global seismic risk model. Earthq Spectra 36(1):372–394

Stafford PJ (2014) Crossed and nested mixed-effects approaches for enhanced model development and removal of the ergodic assumption in empirical ground-motion models. Bull Seismol Soc Am 104(2):702–719

Stewart JP, Klimis N, Savvaidis A, Theodoulidis N et al (2014) Compilation of a local VS profile database and its application for Inference of VS30 from geologic- and terrain-based proxies. Bull Seismol Soc Am 104(6):2827–2841

Stewart JP, Afshari K, Goulet CA (2017) Non-ergodic site response in seismic hazard analysis. Earthq Spectra 33(4):1385–1414

Stewart JP, Parker GA, Atkinson GM, et al (2019) Site-to-site standard deviation model for Central and Eastern North America. Technical Report, University College Los Angeles. https://escholarship.org/uc/item/2sc5g220

Stewart JP, Parker GA, Atkinson GM et al (2020) Ergodic site amplification model for Central and Eastern North America. Earthq Spectra 36(1):42–68

Thompson EM, Wald DJ, Worden CB (2014) A V_{S30} map for California with geologic and topographic constraints. Bull Seismol Soc Am 104(5):2313–2321

Vilanova SP, Narciso J, Carvalho JP, Lopes I, Quinta-Ferreira M, Pinto CC et al (2018) Developing a geologically based V_{S30} site-condition model for portugal: methodology and assessment of the performance of Proxies. Bull Seismol Soc Am 108(1):322–337

Wald DJ, Allen TI (2007) Topographic slope as a proxy for seismic site conditions and amplification. Bull Seismol Soc Am 97(5):1379–1395

Weatherall P, Marks KM, Jakobsson M, Schmitt T et al (2015) A new digital bathymetric model of the world's oceans. Earth Space Sci 2(8):331–345

Weatherill G, Cotton F (2020) A ground motion logic tree for seismic hazard analysis in the stable cratonic region of Europe: regionalisation, model selection and development of a scaled backbone approach. Bull Earthq Eng 18(14):6119–6148

Weatherill G, Kotha S-R, Cotton F (2020a) A regionally-adaptable "scaled backbone" ground motion logic tree for shallow seismicity in Europe: application to the 2020 European Seismic Hazard Model. Bull Earthq Eng 18(11):5087–5117

Weatherill G, Kotha SR, Cotton F (2020) Re-thinking site amplification in regional seismic risk assessment. Earthq Spectra 36(1_suppl):274–297

Wessel P, Smith WH (1998) New, improved version of generic mapping tools released. EOS Trans Am Geophys Union 79(47):579–579

Wills CJ, Clahan KB (2006) Developing a map of geologically defined site-condition categories for California. Bull Seismol Soc Am 96(4A):1483–1501

Yenier E, Sandikkaya MA, Akkar S (2010) Report on the fundamental features of the extended strong-motion databank prepared for the SHARE project. Seismic Hazard Harmonisation in Europe (SHARE) Deliverable 4.1

Yilmaz MT, Ansari A, Harmandar E (2014) Simple geological categories as proxies to VS30 in Turkey and Iran. In: The Proceedings of the second European conference on earthquake engineering and seismology, Istanbul August 25–29, 2014

Zhu C, Cotton F, Kawase H, Haendel A, Pilz M, Nakano K (2022) How well can we predict site response so far? Site specific approaches. Earthq Spectra 38(2):1047–1075

Authors and Affiliations

Graeme Weatherill[1] · **Helen Crowley[2]** · **Agathe Roullé[3]** · **Bruno Tourlière[4]** · **Anne Lemoine[3]** · **Cécile Gracianne[3]** · **Sreeram Reddy Kotha[5]** · **Fabrice Cotton[1]**

Helen Crowley
helen.crowley@eucentre.it

Agathe Roullé
a.roulle@brgm.fr

Bruno Tourlière
btperso@hotmail.com

Anne Lemoine
a.lemoine@brgm.fr

Cécile Gracianne
c.gracianne@brgm.fr

Sreeram Reddy Kotha
sreeram-reddy.kotha@univ-grenoble-alpes.fr

Fabrice Cotton
fabrice.cotton@gfz-potsdam.de

[1]   GFZ German Research Centre for Geoscience, Potsdam, Germany

[2]   European Centre for Training and Research in Earthquake Engineering (EUCENTRE), Pavia, Italy

[3]   BRGM-French Geological Survey, Direction Risques et Prévention, Orléans, France

[4]   BRGM-French Geological Survey, Direction des Géoressources, Orléans, France

[5]   University Grenoble Alpes, University Savoie Mont Blanc, CNRS, IRD, University Gustave Eiffel, ISTerre, 38000 Grenoble, France