ORIGINAL RESEARCH

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

Graeme Weatherill[1] · Fabrice Cotton[1,2]

Abstract

Regions of low seismicity present a particular challenge for probabilistic seismic hazard analysis when identifying suitable ground motion models (GMMs) and quantifying their epistemic uncertainty. The 2020 European Seismic Hazard Model adopts a scaled backbone approach to characterise this uncertainty for shallow seismicity in Europe, incorporating region-to-region source and attenuation variability based on European strong motion data. This approach, however, may not be suited to stable cratonic region of northeastern Europe (encompassing Finland, Sweden and the Baltic countries), where exploration of various global geophysical datasets reveals that its crustal properties are distinctly different from the rest of Europe, and are instead more closely represented by those of the Central and Eastern United States. Building upon the suite of models developed by the recent NGA East project, we construct a new scaled backbone ground motion model and calibrate its corresponding epistemic uncertainties. The resulting logic tree is shown to provide comparable hazard outcomes to the epistemic uncertainty modelling strategy adopted for the Eastern United States, despite the different approaches taken. Comparison with previous GMM selections for northeastern Europe, however, highlights key differences in short period accelerations resulting from new assumptions regarding the characteristics of the reference rock and its influence on site amplification.

Keywords Ground motion models · Stable craton · Regionalisation · Epistemic uncertainty · Europe

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

1   GFZ German Research Centre for Geosciences, Potsdam, Germany

2   University of Potsdam, Potsdam, Germany

1 Introduction

Seismic hazard and risk analysis on a regional scale requires not only the characterisation of ground motion in regions of high seismic activity, but also tectonically stable regions with little seismicity. Here the infrequency of even moderate sized earthquakes limits the number of observations of strong shaking, meaning that purely empirical ground motion models (GMMs) may be poorly constrained over the ranges of magnitudes and distances considered for seismic hazard analysis. To overcome these limitations, it has become more common in engineering practice to take GMMs that are originally constructed for active regions and adjust them to the seismological conditions of the target region. This is usually done using simulations of motion that are calibrated to source, path and site properties inferred from recordings of weak motion in the target region or via other means (e.g. Campbell 2003; Pezeshk et al. 2011; PEER 2015).

Given the paucity of strong motion observations, and the corresponding uncertainties on the seismological properties, it should follow that epistemic uncertainties in GMMs should be larger in regions of low seismicity than those of high seismicity. However, with a smaller pool of available GMMs for stable tectonic environments and few data from which better performing models can be identified, the conventional multi-model approach to characterise epistemic uncertainty may often result in the opposite trend. Despite greater uncertainty in the source, path and site properties, selection of only a small number of models (especially those adopting similar seismological assumptions) will inevitably result in a smaller spread in the epistemic uncertainty. This phenomenon was exemplified in the 2013 European Seismic Hazard Model (ESHM13) (Woessner et al. 2015), wherein only two ground motion models (Toro 2002; Campbell 2003) were selected for application to the cratonic region of northeastern Europe (the Baltic Sea and surrounding counties), in stark contrast to the more active regions of Europe for which the logic tree contained four to six GMMs that were selected on the basis of their fit to observed data (Delavaud et al. 2012).

Within the framework of the Horizon 2020 Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe (SERA) initiative, a next-generation probabilistic seismic hazard and risk model for Europe has been constructed, which builds on new ground motion data and models that have emerged since the publication of the 2013 European Seismic Hazard Model (ESHM13). Adopting the well-established probabilistic seismic hazard analysis (PSHA) procedure and the OpenQuake software for its calculation (Pagani et al. 2014), the 2020 European Seismic Hazard model (ESHM20) aims to integrate not only new insights from the exponential growth of European strong motion databases in the last decade, specifically the creation of the European Strong Motion flatfile (ESM) (Lanzano et al. 2019), but also from new ideas for the representation of epistemic uncertainty that have emerged in the seismic hazard modelling community. Recognising some of the limitations of the multi-model approach, the ESHM20 has chosen to adopt instead a scaled backbone ground motion logic tree (e.g Bommer 2012; Atkinson and Adams 2013; Atkinson et al. 2014), building on top of a coherent and reproducible framework for its construction initially proposed by Douglas (2018a). This concept aims to represent epistemic uncertainty in the ground motion model not by means of selecting multiple independent models, but rather by taking a core model or models (the backbone) and applying to this model(s) scaling factors to account for the uncertainty in the seismological properties of the target region. A more comprehensive

exposition of this approach and its advantages with respect to the multi-model approach can be found in Atkinson et al. (2014), Douglas (2018a) and Weatherill et al. (2020).

The general framework for the scaled backbone GMM logic tree for use in regions of shallow crustal seismic activity within the ESHM20 is presented by Weatherill et al. (2020), which describes the construction and regionalisation of the scaling factors built around the recent European backbone GMM of Kotha et al. (2020). As this model illustrates the philosophy and approach behind the ESHM20 strategy for ground motion modelling that will be adapted to application in northeastern Europe, the key aspects will be summarised in Sect. 2 of this paper. A regionalisation driven by observed strong motion data is a critical element of the ESHM20 approach, yet with few strong motion observations from earthquakes of engineering significance in northern Europe we need to seek other means to identify relevant differences in crustal properties that may impact upon ground motion scaling and attenuation. In Sect. 3 the differences in geophysical properties of the crust are explored in order to help delineate the cratonic region of northeastern Europe and explain why the *general crustal seismicity* logic tree adopted elsewhere on the continent may not be applicable here. From this same exploration we identify the Central and Eastern United States (CEUS) as a relevant analogue in terms of the crustal structure.

Following the regionalisation, in Sect. 4 we construct a new GMM that builds upon the recent NGA East ground motion project (Goulet et al. 2018), calibrating the coefficients of the Kotha et al. (2020) functional form for application to stable cratonic environments. We will then address the factors influencing epistemic uncertainty and how to quantify this within the scaled backbone logic tree framework. The resulting epistemic uncertainty model is compared in Sect. 5 against the ESHM13 model selection and the recent ground motion logic tree for the CEUS developed as part of the 2018 United States National Seismic Hazard Model (Goulet et al. 2017; Petersen et al. 2020). The complete ground motion model logic tree is finally proposed for application to the stable cratonic region of Europe within the ESHM20 and the potential consequences of this new model for seismic hazard demonstrated. Though the primary target of this investigation is the prediction of ground motion for seismic hazard assessment in northeastern Europe and its uncertainties, we believe that the approach illustrated here be considered as a possible strategy for application in similar tectonic environments elsewhere on the globe.

2 Scaled backbone logic tree for general crustal seismicity

The conceptual framework of the scaled backbone logic tree that is being adopted within the ESHM20 and its data-driven regionalisation is outlined by Weatherill et al. (2020), who adapt the strategy initially proposed by Douglas (2018a). In the construction of the scaled backbone GMM logic tree for general crustal seismicity Europe, Kotha et al. (2020) derive a new GMM that capitalises on the wealth of strong motion data available within the European Strong Motion (ESM) flatfile (Lanzano et al. 2019). In doing so, they quantify regional variation in strong motion scaling and attenuation via the introduction of region-specific random effects into the mixed-effects regression. For a detailed description of the model and its regionalisation, the reader is referred to the papers of Kotha et al. (2020). The general form of the ground motion model for PGA, PGV and $Sa(T)$ is as follows:

$$\ln Y = e_1 + f_m(M_W) + f_{r,g}(M_W, R) + f_{r,a}(R) + \delta L2L_l + \delta B_{el}^0 + \delta S2S_s + \delta W_{es} \qquad (1)$$

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

$$f_{r,g}(R, M_W) = (c_1 + c_2 \cdot (M_W - M_{ref})) \cdot \ln\left(\sqrt{\frac{R^2 + h_D^2}{R_{ref}^2 + h_D^2}}\right) \tag{3}$$

$$f_{r,a}(R) = \frac{c_3 + \delta c_{3,r}}{100} \cdot \left(\sqrt{R^2 + h_D^2} - \sqrt{R_{ref}^2 + h_D^2}\right) \tag{4}$$

where Y is the intensity measure of interest (e.g. PGA (in cm s^{-2}), PGV (in cm/s), $S_a(T)$ (in cm s^{-2})), R is the source to site distance in km (Joyner–Boore distance in the general crustal seismicity case) and h_D the effective depth, which is itself dependent on the hypocentral depth of the source. M_h, M_{ref} and R_{ref} are the period independent hinge magnitude, reference magnitude and reference distance, which take values of M_W 6.2, M_W 4.5 and R_{JB} 30 km respectively. Aleatory uncertainty is characterised by δB_{el}^0, and $\delta S2S_s$ and δW_{es}, which describe the event-to-event, site-to-site and site-corrected within-event variability, and are Gaussian distributed random variates with zero means and standard deviations of τ_0, ϕ_{S2S} and ϕ_{SS} respectively.

Two additional random effects are introduced: $\delta L2L_l$, which describes the source-region to source-region variability (a measure of how more or less systematically energetic earthquakes in the region are with respect to the centre of the ESM dataset); and $\delta_{C_{3,R}}$, which describes the regional variability in the residual attenuation (a measure of how much faster or slower the ground motion attenuates in comparison to the average of the ESM dataset). These two random effects are both normally distributed with zero means and standard deviations of τ_{L2L} and τ_{C_3} respectively. In a statistical sense, the distributions of these two terms describe the maximum variability in source region properties and attenuation region properties *implied by the ESM dataset*. For this purpose, two prior regionalisations based on geology and tectonics are adopted, one for the earthquake source (taken from a branch of the ESHM20 area source model) and one for the attenuation region [based on a regionalisation proposed by Basili et al. (2019)].

For those regions in the respective regionalisations that contain a sufficient number of events and/or records to constrain the random effects, it is possible to adopt the local distribution of $\delta L2L_l$ and $\delta_{C_{3,R}}$ for application in the seismic hazard analysis. Where the data are absent or insufficient to constrain the region-specific random effects, however, the full distributions of τ_{L2L} and τ_{C_3} are used. As explained by Weatherill et al. (2020), the spatial variation in $\delta L2L_l$ is complex and potentially poorly constrained, thus we retain the full uncertainty (τ_{L2L}) across Europe. For $\delta_{C_{3,R}}$, however, regions of similar residual attenuation characteristics are grouped into clusters that capture the broader scale regional attenuation variation and permit characterisation of a region-specific distribution of attenuation properties. The complete residual attenuation regionalisation is shown in Fig. 1.

A 9-branch scaled backbone logic tree is constructed for each region, with three branches describing the region-specific attenuation uncertainty ($-\varepsilon \cdot \tau_{C_3,R}$, $0 \cdot \tau_{C_3,R}$, and $+\varepsilon \cdot \tau_{C_3,R}$), and three describing the source region specific variability ($-\varepsilon \cdot \max\left(\tau_{L2L,R}, \sigma_\mu^{SHAL}\right)$, $0 \cdot \max\left(\tau_{L2L,R}, \sigma_\mu^{SHAL}\right)$, and $+\varepsilon \cdot \max\left(\tau_{L2L,R}, \sigma_\mu^{SHAL}\right)$), where σ_μ^{SHAL} is the within-model statistical uncertainty in the backbone model determined via the approach of Al Atik and Youngs (2014). As these distributions are now Gaussian,
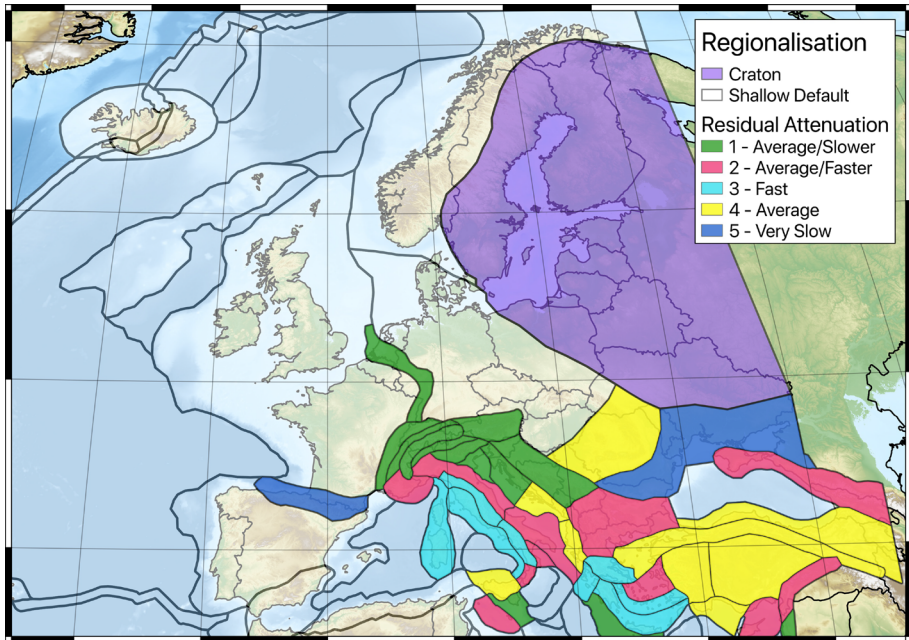
Fig. 1 Ground motion model regionalisation of the ESHM20, with the craton zones highlighted in purple, the default shallow crustal logic tree region in white, and the different cluster-analysis defined regionalisation of shallow crustal seismicity in the coloured zones labelled from 1 to 5 (Weatherill et al. 2020)

we adopt a three-branch discrete approximation using the factors proposed by Miller and Rice (1983), which results in weights of 0.167, 0.666 and 0.167 respectively and $\varepsilon \approx 1.732$. For regions where data are insufficient to constrain a random effect, $\tau_{L2L,R} = \tau_{L2L}$ and $\tau_{C_3,R} = \tau_{C_3}$, which we refer to as the 'general crustal seismicity' logic tree and apply throughout much of northern and western Europe.

The proposed formulation implies that the range of ground motion values predicted by the general crustal seismicity backbone logic is representative of the epistemic uncertainty in ground motion models. This assumes, however, that the seismogenic properties of the shallow crust are sufficiently well sampled by the ground motion records in the ESM dataset. As the ground motions themselves are geographically biased toward the more active regions of Southern Europe and the Mediterranean, there is limited data to test this assumption in northern Europe. The default backbone is intended to consider the possibility that median ground motions may be encompass higher stress drops and slower attenuation than the 'average' values in the ESM dataset. Nevertheless, it is possible that even this may be insufficient to characterise the seismological properties of earthquakes in the stable continental crust if the geological and geophysical composition of the crust in certain stable areas are fundamentally different from those found in more active regions. The critical question then is whether it can be determined that a region of shallow seismicity in Europe is of a fundamentally different nature in terms of its ground motion characteristics, compared to the range of properties represented in ESM dataset, such that the centre, body and range of median ground motion scaling described by the 'general crustal seismicity' logic tree is not sufficiently representative.

3 Regionalisation of the stable crust of Europe

With few strong motion recordings in Northern Europe from which to attempt a data-driven regionalisation of the general crustal seismicity logic tree, attention is turned instead to other data sources that could provide insights into regional differences affecting source, path and site effects in stable regions. Developing on previous efforts toward a data-driven regionalisation by Chen et al. (2018), large-scale regional and global datasets of crustal properties have been assembled, whose impacts on ground motion have a physical basis. Some datasets are defined using weak motion data, others by various different means of geophysical investigation. The purpose of such comparison is to identify where, if anywhere, within the stable seismic region of northern Europe can continental crust be said to be of a similar nature to that of southern Europe (from which we have strong motion data), and where is this not the case. The implication of the latter is that the region-to-region variability described by the full distributions of $\delta_{C_{3,R}}$ and $\delta L2L_l$ is still insufficient to capture the possible characteristics of strong motion in these localities. Furthermore, if such differences are apparent, can analogous regions be identified elsewhere in the globe?

The exploration will focus on five key geological and geophysical datasets:

1. A Eurasian model of 1-s attenuation quality factor (Q_{LG}) (Mitchell et al. 2008)
2. The mantle shearwave velocity anomaly (δV_S) with respect to the Preliminary Reference Earth Model at 175 km depth. As described by Mooney et al. (2012), lithospheric thickness generally shows a correlation with the age of the continental crust in tectonically stable regions, with the oldest Archean cratonic domains forming lithospheric roots whose thickness is in excess of 200–240 km, while the roots of younger Phanerozoic margins are notably thinner (100–150 km). A significant positive shearwave velocity anomaly at depths between 150 and 200 km is therefore indicative of older (and colder) Archean cratons that are often associated with slower attenuation and higher stress drop.
3. A global grid of estimated heatflow (Lucazeau 2019); a property directly related to the attenuation of seismic waves in crystalline material.
4. A high resolution European map of Moho depth (Grad et al. 2009) and a corresponding 1° global model from Szwillus et al. (2019).
5. A European map of bedrock geology from the 1:5,000,000 International Geological Map of Europe and Adjacent Areas (Asch 2005).

The first four datasets (Q_{LG}, δV_S, heatflow and Moho depth) are shown for Europe in Fig. 2, and the geology map (colour-scaled in terms of youngest geological age) in Fig. 3.

Across all five datasets a consistent contrast is visible between the crustal properties predominant in western and southern Europe and those of northeastern Europe, more specifically the Baltic Sea and surrounding countries. The Moho depth in particular changes rapidly across the Trans-European Suture Zone (TESZ), which extends from Denmark to the north coast of the Black Sea and delineates the boundary between the Precambrian craton to the northeast and the Phanerozoic crust in southern and western Europe. To the west of this boundary, Moho depths are typically in the range of 30–40 km, while to the east they increase sharply to 45–55 km. Similar trends can be seen in both the Q_{LG}, δV_S and heatflow data, where southern and western Europe is characterised by lower Q_{LG}, higher heatflow and weakly negative δV_S than for northeastern Europe.

From the inspection of the geophysical and geological datasets shown in Figs. 2 and 3, there is a reasonable basis upon which to assume that crustal properties in western Europe
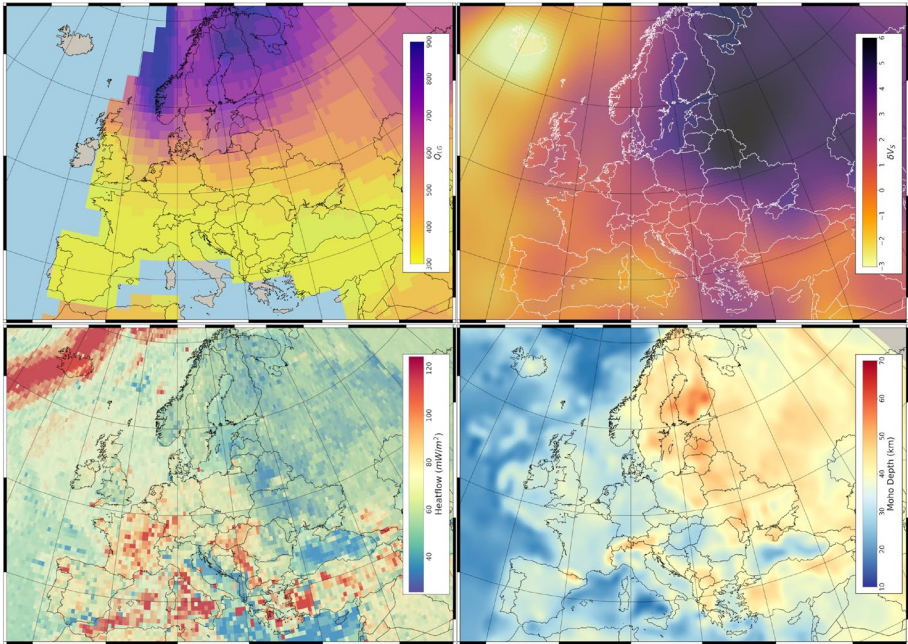
Fig. 2 Distribution of data for discriminating the *craton* region in Europe: 1-s attenuation quality factor Q_{LG} (Mitchell et al. 2008) (top left), shearwave velocity anomaly at 160 km depth (Mooney et al. 2012) (top right), heatflow (Lucazeau 2019) (bottom left) and Moho depth (Grad et al. 2009) (bottom right)
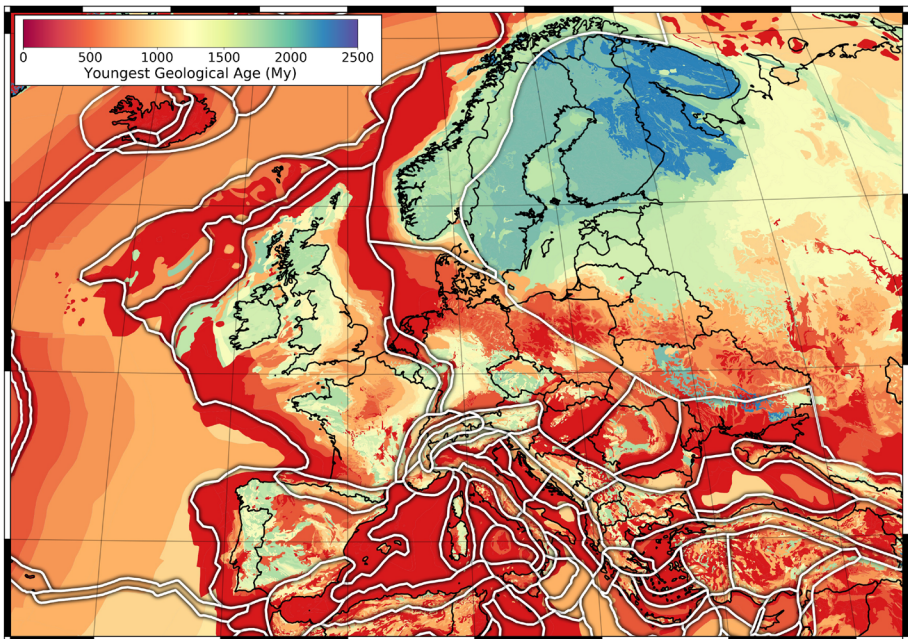
Fig. 3 Youngest age of the bedrock geology across Europe (Asch 2005), overlain by the regionalisation of Basili et al. (2019) (white lines)

are sufficiently similar to those of the Mediterranean region that they would influence the characterisation of ground motion in the manner encapsulated within the general crustal seismicity backbone logic tree. In northeastern Europe, however, this is not the case. Instead, this region is characterised by higher Q_{LG}, deeper Moho and older Paleozoic to pre-Cambrian geology; all typical features of a stable cratonic environment and whose strong motion characteristics are *potentially* unlikely to fall within the complete source and attenuation region-to-region variability implied by the ESM dataset. In terms of geographical extent, the general consensus of the datasets is that the stable cratonic region is bounded to the west and north by the Caledonian front, running the Barents Sea to the North Sea, and bounded to the south by the TESZ. Both boundaries are already implied in the eastern limits of the regionalisation of Basili et al. (2019). The exceptionally slow residual attenuation, $\delta_{C_{3,R}}$, identified by Weatherill et al. (2020) in the zone running from eastern Romania to southern Ukraine also supports the notion that this boundary marks a significant change in ground motion properties.

Geophysical datasets that are global in coverage can help to identify of tectonically analogous regions from which ground motion models and their respective uncertainties could be calibrated. Figure 4 shows the same maps of Q_{LG}, δV_S, heatflow and Moho depth [this time from Szwillus et al. (2019)] for much of the western hemisphere. With its high attenuation quality factor Q_{LG}, high mantle shearwave velocity anomaly, lower heatflow and deeper Moho, close geophysical similarity can be identified between northeast Europe and the Central and Eastern United States (CEUS). Though analogy is not necessarily novel for this region, what we aim to illustrate here is that even in regions of low seismicity and sparse strong motion data there are means by which differences in the seismological properties from one region to another can be inferred from geophysical data and potentially analogous regions found.

The analogous tectonic and geophysical characteristics of northeastern Europe and the CEUS provide a rationale for constructing an alternative backbone GMM that builds upon the recent developments in ground motion modelling in Eastern North America, namely the Next Generation Attenuation (NGA) East project Goulet et al. (2017). Before delving into the specific details, it is still important to keep in mind a relevant caveat. While several key datasets point one toward the use of CEUS GMMs in northeastern Europe, the paucity of strong motion data from this region available to the authors may yet prevent us from stating with complete certainty that the characteristics of ground motion in the CEUS are appropriate for northeastern Europe. As shall be seen in the final stable craton logic tree (Sect. 6), the possibility that the CEUS models do not adequately represent the epistemic uncertainty in ground motion in northeastern Europe should still be accommodated even if it is not favoured.

4 Constructing a GMM and scaled backbone logic tree for cratonic regions: building on the NGA east models

The NGA East project aimed to develop a new ground motion characterisation for the Central and Eastern United States for use in PSHA, in the context of an SSHAC Level 3 study (Budnitz et al. 1997; U S Nuclear Regulatory Commission 2012). This requires definition of the median and standard deviation of the distribution horizontal PGA, PGV and spectral acceleration, as well as an estimate of the epistemic uncertainty (Goulet et al. 2017, 2018). A complete overview of the various components and novel
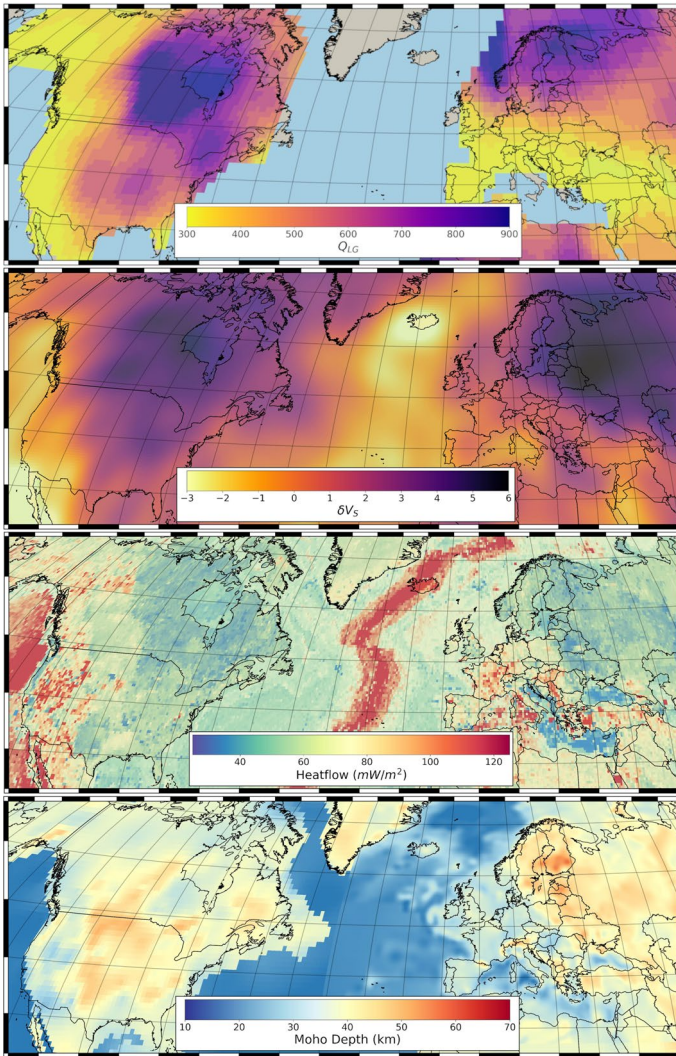
Fig. 4 Distribution of data for discrimination a *craton* region in the Western Hemisphere: 1-s attenuation quality factor Q_{LG} (Mitchell et al. 2008) (top left), shearwave velocity anomaly at 160 km depth (Mooney et al. 2012) (top right), heatflow (Lucazeau 2019) (bottom left) and European Moho depth (Grad et al. 2009) superimposed over 1° global Moho depth (Szwillus et al. 2019) (bottom right)

developments of the NGA East project is omitted here for necessary brevity, but they can be found in detail in Goulet et al. (2018) and the wealth of literature cited therein. Of relevance, however, are the suite of 20 candidate (or "seed") ground motion models developed by ten teams of authors, each team following a different approach to the characterisation of motion, with some defining multiple models to account for epistemic uncertainty in modelling approach and/or source and path properties (PEER 2015). For application to the 2018 US National Seismic Hazard Model (US NSHM), a more complete and practical quantification of the epistemic uncertainty on both the median

ground motion and the aleatory variability has been undertaken (Goulet et al. 2017), with the resulting logic tree described by Petersen et al. (2020).

The characterisation of a ground motion model for application to the craton region of Europe aims to capitalise on the outcomes of the NGA East process, without necessarily adopting the logic tree for the CEUS from Petersen et al. (2020) in full. There are several reasons for this. The first is a matter of practicality and computational efficiency, as we aim to reduce the branches to a simpler representative distribution that would aim to capture the first few moments of the epistemic uncertainty distribution. The second reason is that we seek a degree of consistency between the ground motion characterisation of the craton and that of the general crustal seismicity logic tree. Comparisons between the new GMM shown in this section and the suite of GMMs adopted by Petersen et al. (2020) for the 2018 US NSHM, and the resulting PSHA logic trees, are intended to identify the potential differences between the two approaches.

4.1 Deriving a GMM from the NGA East suite

The development of a backbone GMM for the cratonic region of Europe and its epistemic adjustment factors from the NGA East model suite is predicated on several key assumptions:

– The range of median ground motion values described by the set of NGA East models is an *unbiased* sample of the epistemic uncertainty in median ground motions *within the CEUS*. Their application in Europe is considered only for sources located within the cratonic region of northeast Europe, as identified in Sect. 3.
– Each of the 21 candidate models, which includes all present in the NGA East suite (PEER 2015) plus Pezeshk et al. (2011), is weighted in equal measure (i.e. no specific model(s) is preferred)
– For the purposes of scaling the median ground motions from the seismological reference rock of the NGA East Suite ($V_S = 3000$ m/s) to any other site condition, the amplification models described by Stewart et al. (2020) and Hashash et al. (2020) are used.
– All NGA East GMMs, and therefore all subsequently derived models herein, adopt a common aleatory uncertainty model described by Al Atik (2015). Both assumptions follow closely the approach adopted for the 2018 US NSHM and will be discussed in more detail in due course.

The range of median ground motion response spectra from the suite of 21 candidate models are shown for nine magnitude-distance scenarios (adjusted to a reference rock site condition of V_{S30} 800 m/s) in the pink lines of Fig. 5, and compared against the nine-branch general crustal seismicity backbone model in grey. It can be seen clearly that the NGA East models predict higher accelerations at short periods than even the upper branches of the general crustal seismicity backbone. At periods longer than 0.3 s (for small magnitude and short distances) to 1.0 s (for larger magnitudes and longer distances) the centre and range of the distributions are similar, however. Also shown for comparison are the spectra from the recent GMM of Fülöp et al. (2020), which adjusts the NGA East GMM of Graizer (2016) for application to the very hard rock sites (V_S 2800 m/s) in Finland using a dataset of strong motion records from small magnitude earthquakes in the region combined with response spectra from the original NGA East database. As the local calibration of model is made on the basis of seismicity mostly below the minimum magnitude considered for
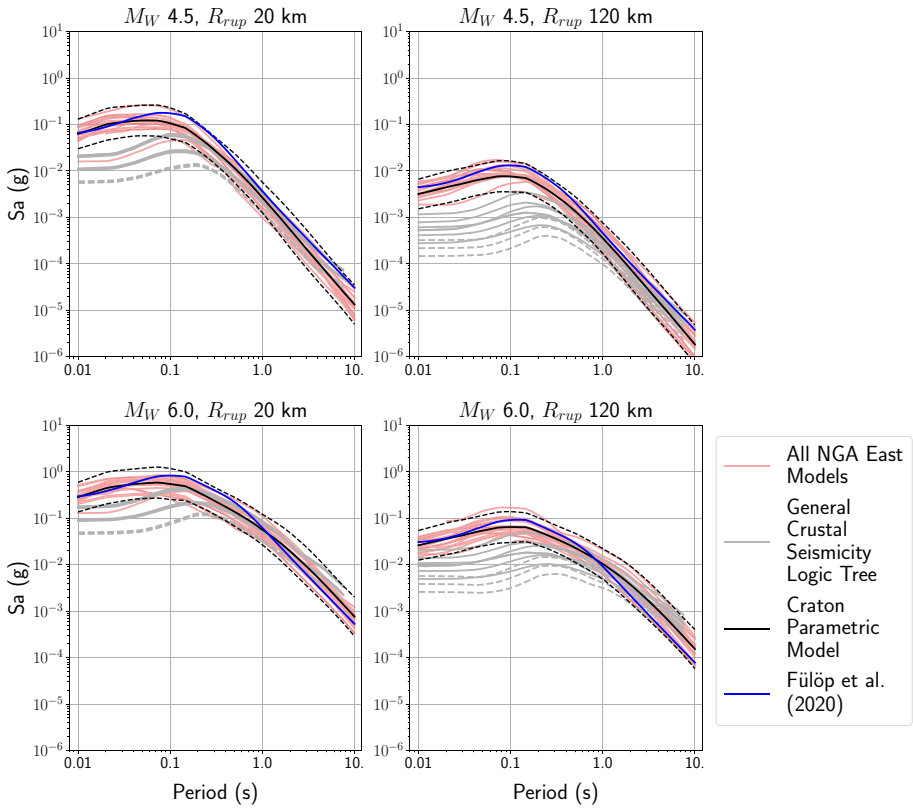
Fig. 5 Comparison of the spectra of the NGA East Model suite (pink), the default shallow crustal backbone GMM (grey), the proposed parametric ground motion model (black) and the Fennoscandian GMM of Fülöp et al. (2020) (blue). All models assuming a reference rock condition of V_{S30} 800 m/s. Low stress region and/or fast attenuation branches of the general crustal logic tree are shown by dashed grey lines

the ESHM20 ($M_W \geq 3.5$) we do not use this as a "seed model" in the following analysis; however, it does form a useful point of comparison as it would seem to predict short period motions that are toward the upper end of the range indicated by the full NGA East model suite. In Fig. 5, the ground motions from the Fülöp et al. (2020) model are adjusted to the V_{S30} 800 m/s reference rock using the same model of Stewart et al. (2020).

We note that at the time of writing very few of outcomes of the NGA East project themselves have been published in peer-reviewed scientific literature. The exceptions to this are the GMMs of Yenier and Atkinson (2015) and Graizer (2016) and the site amplification models of Stewart et al. (2020) and Hashash et al. (2020). This situation might be considered as basis for exclusion of such models, were we to apply strictly the pre-selection criteria of Cotton et al. (2006). As the main outcomes of the NGA East project have been adopted by the USGS for inclusion in the 2018 US NSHMP, a process that typically subjects candidate models to scrutiny by a larger team of scientists, and that this in itself is published in the scientific literature (Petersen et al. 2020), we believe adoption of the NGA East GMMs is appropriate in the present case.

The starting point in the construction of a scaled backbone GMM logic tree for cratonic regions is the assumption that for any given magnitude, M_W, rupture distance, R_{RUP}, and

period, T, the epistemic uncertainty on the median ground motions are described by a log-normal distribution such that:

$$P\big(\ln Y\big(M_W, R_{RUP}, T\big)\big) = \mathcal{N}\big(\boldsymbol{\mu}\big(M_W, R_{RUP}, T\big), \boldsymbol{\sigma_\mu}\big(M_W, R_{RUP}, T\big)\big) \tag{5}$$

In theory, this scenario-dependent distribution could by itself form a non-parametric GMM, which can be readily implemented in PSHA calculation software in the form of a set of tables from which the ground motion for the desired M_W and R_{RUP} is interpolated. In this case the epistemic uncertainty would be represented by the term $\varepsilon_\mu \cdot \sigma_\mu\big(M_W, R_{RUP}, T\big)$, where ε_μ is the number of standard deviations above or below the median. This would then be characterised by a discrete set of branches representing the median and pre-defined percentiles of the distribution, with their corresponding values and weights described by a discrete approximation to a standard Gaussian density function (e.g. Miller and Rice 1983). The non-parametric model has some advantages over the conventional parametric approaches in that it is independent of any prior functional form, and σ_μ is fully heteroskedastic (i.e. dependent on the magnitude and distance) so that it may be lower for scenarios where the candidate models converge and greater where they diverge. It cannot, however, always guarantee a physical behaviour, nor can it be extrapolated safely to consider scenarios outside of the range used for derivation.

Given the possible limitations of a *non-parametric* ground motion model, we prefer instead to construct a *parametric* form. Parametric GMMs are the more common form of GMM, and ensure that the ground motion with magnitude and distance should reflect physically-based assumptions of the source scaling and attenuation properties in a region. Provided that the functional forms adopted are suitable and the coefficients well-calibrated, they can be extrapolated more readily than non-parametric forms. To construct the parametric models the distributions of median ground motions described by Eq. 5 are used to generate a synthetic dataset of median ground motions for PGA and spectral acceleration. A ground motion model of the same form as the backbone shallow GMM (Eqs. 1–4) is fit to the synthetic data:

$$\ln Y_{median}(T) = e_1 + f_M\big(M_W\big) + f_{R,g}\big(R_{RUP}, M_W\big) + f_{R,a}\big(R_{RUP}\big) + \varepsilon_\mu \cdot \sigma_\mu \tag{6}$$

where f_M, $f_{R,g}$ and $f_{R,a}$ are as described by Eqs. 2, 3 and 4 respectively. As the distance metric, R, is now changed from Joyner–Boore distance (R_{JB}) to rupture distance R_{RUP}, then R_{ref} is fixed at 1 km and h_D at 5 km (independent of hypocentre depth). No random effects are considered in this regression as the aleatory uncertainty model will be treated separately. Similar to Kotha et al. (2020), robust least squares regression is used in order to minimise the weight of the outlier models for any given scenario. The median, 5th and 95th percentiles of the epistemic uncertainty distribution (σ_μ) resulting from *parametric* form of the craton model are shown in Fig. 5 as black lines, and coefficients of the model are provided in Table 1. For the current purpose a homoskedastic σ_μ is assumed.

4.2 USGS approach to characterising epistemic uncertainty in the NGA east GMMs

The construction full GMM logic tree for the CEUS is described in detail by Goulet et al. (2017), with the final modelling decisions summarised by Petersen et al. (2020). The USGS CEUS logic tree contains a total of 34 branches, combining a selection of 17 of the 20 adapted "seed" models (i.e. the original NGA East models adapted for application over the source, path and site conditions present in the Eastern US), and 17 models intended to

Table 1 Coefficients of the fitted model (referring to the natural logarithm scale)

IMT	e_1	b_1	b_2	b_3	c_1	c_2	c_3	σ_μ
PGA	0.129434	0.516399	− 0.120322	0.209373	− 1.498201	0.220432	− 0.219311	0.467518
0.010	0.441910	0.507166	− 0.101880	0.184282	− 1.567538	0.222961	− 0.217385	0.424145
0.020	0.979124	0.464490	− 0.113773	0.167234	− 1.628256	0.226151	− 0.244152	0.453414
0.025	1.043341	0.469671	− 0.113451	0.174066	− 1.609088	0.224104	− 0.257668	0.456276
0.030	1.046568	0.476295	− 0.114530	0.188789	− 1.578345	0.220698	− 0.270013	0.442618
0.040	1.007663	0.493810	− 0.115011	0.208536	− 1.522322	0.215223	− 0.287477	0.432693
0.050	0.951569	0.507031	− 0.117000	0.227663	− 1.476123	0.210021	− 0.298269	0.436895
0.075	0.766899	0.537818	− 0.125793	0.255898	− 1.390136	0.198935	− 0.306253	0.445049
0.100	0.566921	0.563265	− 0.139089	0.285966	− 1.329051	0.189119	− 0.296371	0.445057
0.150	0.316925	0.627618	− 0.168968	0.338415	− 1.252120	0.167802	− 0.266500	0.408938
0.200	0.116889	0.691137	− 0.191139	0.377390	− 1.205866	0.154400	− 0.236540	0.396718
0.250	− 0.043842	0.744830	− 0.208516	0.406489	− 1.183521	0.146981	− 0.208303	0.385803
0.300	− 0.198477	0.799805	− 0.223155	0.433866	− 1.165570	0.140633	− 0.179797	0.386776
0.400	− 0.441747	0.897281	− 0.242205	0.483912	− 1.151567	0.133979	− 0.136251	0.395065
0.500	− 0.637445	0.992673	− 0.253909	0.526939	− 1.144198	0.129944	− 0.112135	0.416677
0.750	− 1.032362	1.237960	− 0.248353	0.613138	− 1.127283	0.121478	− 0.073566	0.424884
1.000	− 1.372803	1.445804	− 0.229116	0.691619	− 1.109474	0.116811	− 0.058351	0.435249
1.500	− 1.888467	1.730211	− 0.193720	0.805619	− 1.102390	0.114304	− 0.039000	0.494395
2.000	− 2.334523	1.920451	− 0.161746	0.908051	− 1.094766	0.113859	− 0.029689	0.529657
3.000	− 3.034920	2.146848	− 0.114822	1.085141	− 1.090842	0.115717	− 0.019806	0.550852
4.000	− 3.576616	2.262688	− 0.088526	1.227766	− 1.090290	0.117770	− 0.013579	0.547912
5.000	− 4.022629	2.318744	− 0.077704	1.346637	− 1.090249	0.118983	− 0.008330	0.536941
7.500	− 4.876431	2.373219	− 0.064599	1.529693	− 1.107500	0.131643	− 0.000049	0.531853
10.00	− 5.489149	2.381481	− 0.063354	1.620020	− 1.127404	0.141292	0.005956	0.560199

sample the complete ground motion model space in a mutually exclusive and collectively exhaustive (MECE) manner using the Sammon's mapping methodology (e.g Scherbaum et al. 2010), which will be referred to here as the "Sammons" models. The weights are apportioned to the "seed" models based on the different assumptions regarding the geometric spreading as well as different simulation and/or hybrid-empirical modelling approaches (Petersen et al. 2020). The branch weights for each model are therefore period independent. A total weight of 0.333 is distributed across the 17 seed models and 0.667 across the 17 "Sammons" models. A common aleatory uncertainty and the site amplification model is assumed for all branches. Hence the epistemic uncertainty here refers to the uncertainty in the median ground motions for which we adopt the symbol σ_μ, as its meaning is the same as that described in Eqs. 5 and 6.

The "Sammons" models are built from the same suite of NGA East models describing the median ground motion for the CEUS, but applying a more complex process in order to better fulfil the MECE objective, This is explained in full mathematical detail in Goulet et al. (2017) but summarised qualitatively here. The median ground motions from the candidate GMMs are used to define the distribution of median ground motions for each of the N_D magnitude M_W and distance R_{RUP} scenarios, i. The complete distribution across all N_D scenarios is modelled by a multivariate Gaussian function with mean μ_i and covariance Σ,

such that each entry in the covariance matrix is given by $\Sigma_{ij} = \rho_{ij} \cdot \sigma_{\mu_i} \cdot \sigma_{\mu_j}$, where μ_i and σ_{μ_i} are the mean and standard deviation of the uncertainty distribution for the median ground motion at each scenario, i, and ρ_{ij} the correlation between the ith and jth scenarios. Comparison can be made at this point with the non-parametric form of GMM described in Eq. 5, albeit Eq. 5 assumes the off-diagonal terms of Σ to be zero. The corresponding mean and covariance of the multivariate normal distribution are then sampled to generate thousands of median ground motion models, which describe collectively the complete the model space. For a given spectral period, each sample model contains hundreds of scenario-dependent median ground motions, which can be considered as a single observation in a high (N_D) dimensional model space. From the high dimensional model space, the sample models then rendered into a low dimensional model space using the Sammon's mapping technique (Scherbaum et al. 2010). The resulting 2D model space is then discretised and averaged to produce a total of 17 median ground motion models that sufficiently represent the epistemic uncertainty. The weights assigned to each of the 17 models in the logic tree are determined based on a combination of the number of samples in each of the 17 subsets of the model space and the likelihood fit of the models to observed strong motion data in the CEUS. As these differ from period-to-period, the weights assigned to each of the 17 models are period-dependent in the PSHA calculation.

4.3 Comparing the parametric craton model against the USGS and ESHM13 approaches

The parametric form of the craton model as described by Eq. 6 is compared against the USGS "Seed" models (pink) and "Sammons" models (green) with respect to distance (Fig. 6) and response spectra (Fig. 7). The two magnitudes (M_W 4.5 and 6) and two distances (R_{RUP} 20 km and 120 km) represent the likely scenarios contributing most significantly to short- and long-period hazard within the craton region respectively. In general, the median and percentiles of the parametric form of the craton model match well those of the USGS models across the complete spectrum for distances most relevant to seismic hazard ($10 \leq R_{RUP}(km) \leq 80$). In the smaller magnitude case, the parametric model tends toward the lower end of the NGA East range very short distances ($R_{RUP} < 5$ km), due to the limitations of the current functional form. This virtually inconsequential for PSHA application where hypocentral depths greater than 5 km are assumed. At distances greater than 80 km we see effect of the break in distance scaling within some of the NGA East models that is not present in the form of the parametric craton model. This results in the parametric model slightly over-predicting ground motion in the 70–100 km range and slightly under-predicting it for longer distances, though all NGA East GMMs fall within the 5th and 95th percentiles. This is a constraint of the functional form of the parametric model; however, as we see in the larger magnitude case the epistemic uncertainty in the scaling with distance and the exact point of the break in scaling varies considerably from model to model. We therefore find in this case that the mean and percentiles of the parametric model envelop the range of attenuation assumptions well, and as will be seen in the next section the impacts on the mean hazard are minimal.

Figure 8 focuses on the comparison of the new parametric model against the ESHM13 selection for the craton region (i.e. Toro 2002; Campbell 2003), and the general crustal seismicity logic tree for the reference rock condition of V_{S30} 800 m/s. This comparison suggests that for much of the magnitude range and at short-to-intermediate rupture distances the median and percentiles of the parametric craton model envelope match well the spectra
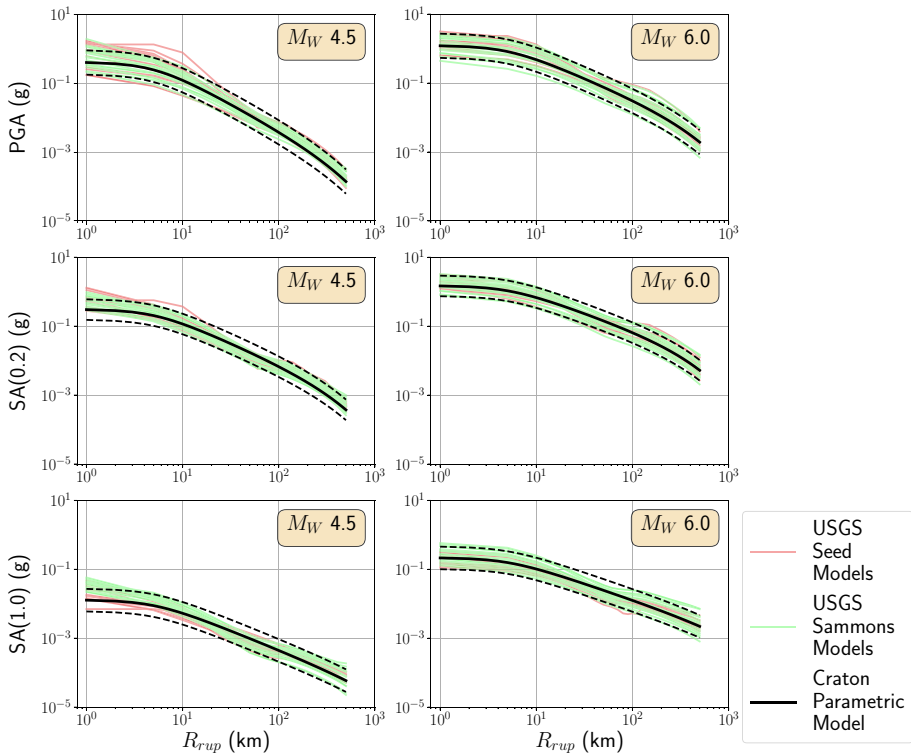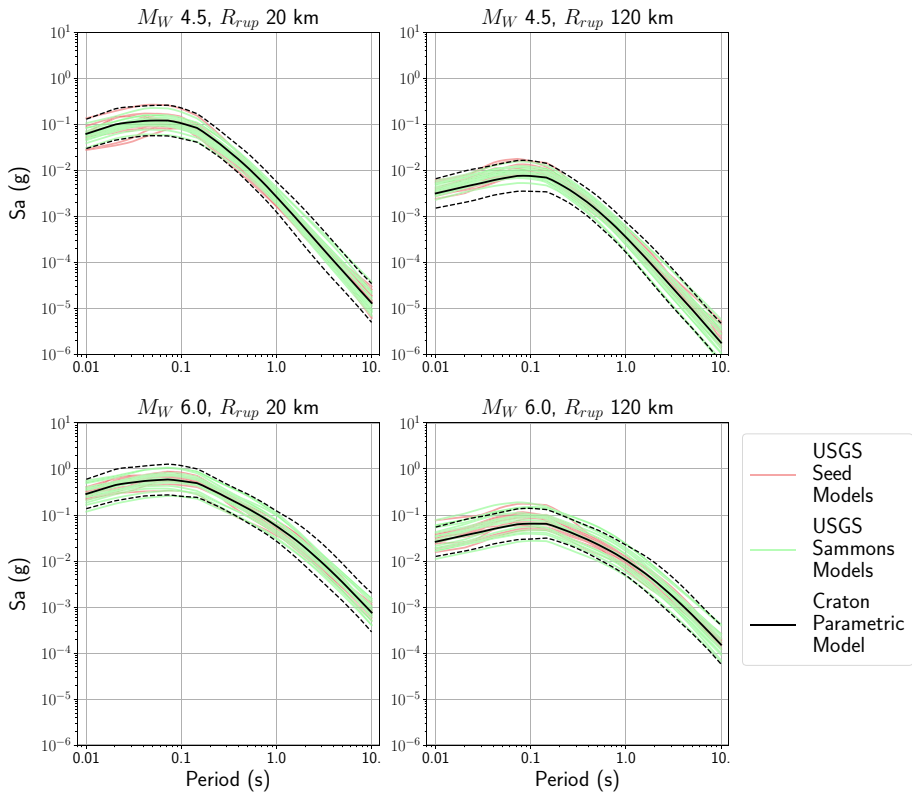
**Fig. 6** Comparison of the attenuation with distance of the median, 5th and 95th percentiles of the parametric craton model (black solid and dashed lines respectively) against the two sets of CEUS GMMs: "Seed" (pink lines) and "Sammons" (green lines), and the 9 branches of the default shallow crustal backbone logic tree, assuming a reference rock condition of V_{S30} 800 m/s

predicted by both Toro (2002) and Campbell (2003) at periods greater than 0.1 s. For very short periods the parametric model seems to indicate higher accelerations, while for longer periods at longer distances this trend is reversed. This latter observation would imply still a faster attenuation in the craton model than suggested by the ESHM13 GMM selection.

4.4 Very hard rock to reference rock amplification

The comparison of the response spectra between the parametric ground motion model developed from the NGA East GMMs here and those of the ESHM13 selection (Toro 2002; Campbell 2003) shown in Fig. 8, plus the resulting hazard outputs shown in Sect. 5, requires careful consideration of the site amplification models and the assumptions contained therein. In both the NGA East and ESHM13 cases, the original stable craton GMMs are defined for a very hard rock condition: V_S 2800 m/s for the ESHM13 case and V_S 3000 m/s for the NGA East. For the ESHM13, Van Houtte et al. (2011) adjusted the Toro (2002) and Campbell (2003) models from very hard rock to the 800 m/s reference rock using the host-to-target approach of Campbell (2003). In doing so, they assume a generic "gradient" velocity profile, with significant differences in the site-specific high frequency decay parameter κ from the very hard rock site to the rock site. Under such conditions, κ will

Fig. 7 Comparison of the spectra of the median, 5th and 95th percentiles of the parametric craton model (black solid and dashed lines respectively) against the two sets of CEUS GMMs: "Seed" (pink lines) and "Sammons" (green lines), and the 9 branches of the general crustal seismicity logic tree, assuming a reference rock condition of V_{S30} 800 m/s

have a damping effect on the high frequency ground motions at the reference rock site with respect to the hard rock site, thus resulting in lower short-period motion on the reference rock.

By contrast, Stewart et al. (2020) consider two very-hard rock to reference rock amplification factors (F_{760}), one for the case of a gradational velocity profile (F_{760}^{gr}) and one for a high-impedance case (F_{760}^{imp}). In the high-impedance case, F_{760}^{imp} produces an amplification peak in the 0.05 to 0.2 s period range. The recommended gradient model is the median of nine models, each assuming a gradual increase in shearwave velocity with depth. In the present case, the gradient model does not show a deamplification at high frequencies as only a couple of the constituent models assume kappa values as large as the $0.02 \leq \kappa \leq 0.05$ range considered by Van Houtte et al. (2011). Recognising that a V_S profile extending through the weathered zone of bedrock material is seldom available, Stewart et al. (2020) explored the amplification peaks for observed response spectra under a range of magnitude, distance and V_{S30} conditions. They found that for sites with V_{S30} greater than 600 m/s pronounced amplification peaks in the 0.05–0.2 s case were more common, while for lower V_{S30} a smaller and broader peak in the 0.1–0.3 s range was predominant. For the purpose of developing a general CEUS amplification model, a V_{S30}-dependent weighted
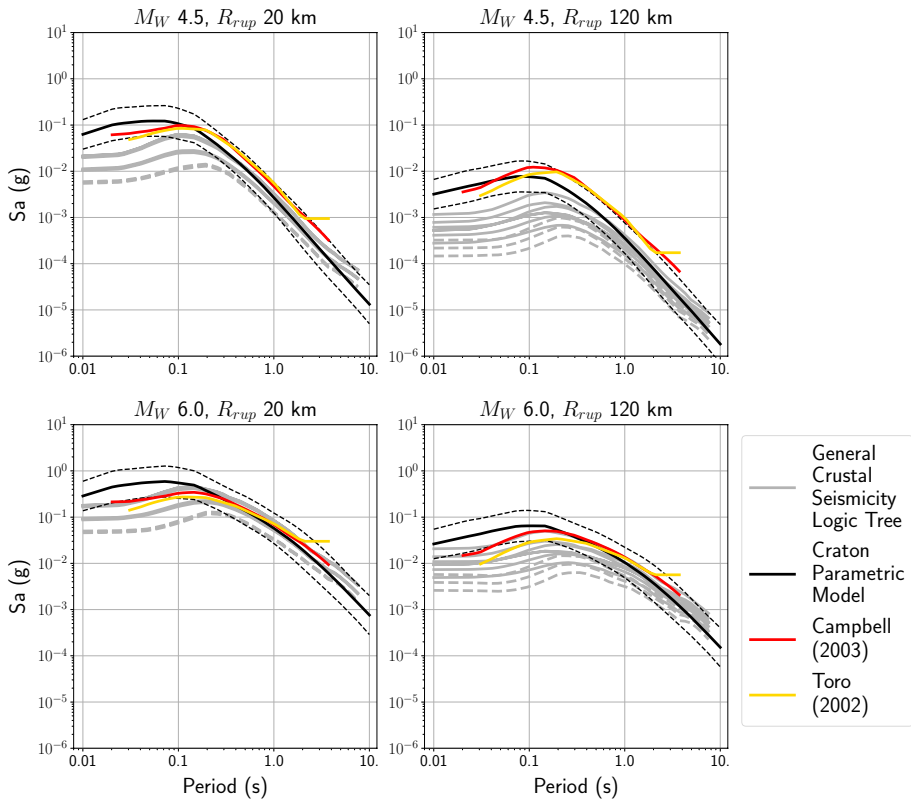
Fig. 8 Comparison of the response spectra for different magnitude and distance scenarios for the parametric craton model, the default shallow crustal logic tree [both adjusted to reference rock using the model of Stewart et al. (2020)] and the two CEUS models adopted within the ESHM13: Campbell (2003) and Toro (2002) adjusted to reference rock condition of V_{S30} 800 m/s using the factors of Van Houtte et al. (2011)

average of F_{760}^{gr} and F_{760}^{impt} was proposed, which gives preference to the high-impedance case for sites with $V_{S30} > 600$ m/s and reverses this weight to give preference to the gradient model for softer soils. For the 800 m/s reference rock case considered here, the NGA East models are therefore amplified with respect to the very hard rock, in contrast with the ESHM13 GMMs where they are de-amplified, as shown in Fig. 9.

This comparison of the implicit very-hard rock to reference rock amplification factors between the ESHM13 models (Van Houtte et al. 2011) and the new NGA East models (Stewart et al. 2020), raises a critical question regarding the nature of the reference rock considered for the Eurocode 8 application on a national or regional scale. In particular, for application of the parametric craton model in the ESHM2020, can the reference site condition be assumed to represent a high-impedance or a gradational site profile? Taking into consideration the predominantly older bedrock geology of the region (Fig. 3) and the recent glacial history, there is once again good reason to believe that the CEUS provides a suitable analogue for northeastern Europe. Stewart et al. (2020) indicate that their high-impedance profiles are representative of glacial till and/or thin layers of chemically weathered saprolite overlaying high velocity bedrock; conditions typically associated with tectonic stability, extensive glaciation and older geological domains. Such conditions can also
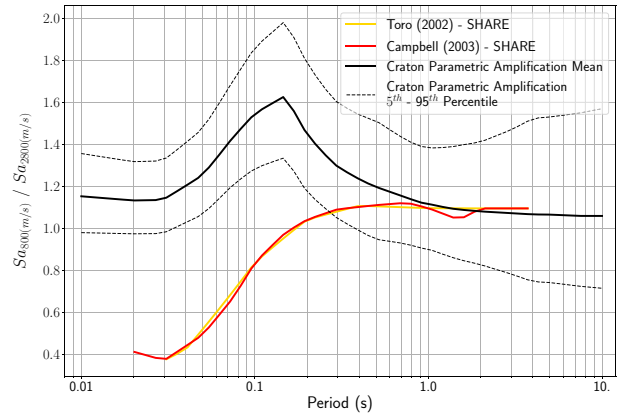
be found to be predominant in northeastern Europe, especially within the Baltic shield, and we would therefore suggest that the higher weighting to the F_{760}^{imp} assumed by the NGA East amplification model is better suited to this region for the determination of the amplification at the 800 m/s reference condition than the gradational profile assumed by Van Houtte et al. (2011). The consequence of this decision, however, is that the adoption of the parametric craton model will result in higher hazard on the reference rock in the short period motion, as will be seen in Sect. 5.

The amplification model adopted for the CEUS defines not only a median amplification, but also a V_{S30}-dependent epistemic uncertainty, represented by a log-normal distribution with standard deviation $\sigma_{\mu,S}$, which is shown in the dashed black lines in Fig. 9. This uncertainty is defined largely from expert judgement, to ensure it varies smoothly across the entire velocity range even where observations are limited and tapers to zero for V_S 3000 m/s. The possibility of de-amplification in the ground motion from very hard rock to reference rock is considered within the 5–95 percentile range of $\sigma_{\mu,S}$, except in the period range with the high impedance amplification peak. As we have no local site data from which to constrain the preferred amplification model, we implement this epistemic uncertainty in the logic tree described in Sect. 6.

4.5 Aleatory uncertainty

One of the key assumptions in the development of the craton GMM is the adoption of the Al Atik (2015) aleatory uncertainty model, which is common to all candidate NGA East models. Once again, the details of the derivation of the model are omitted here for brevity; however, the general structure of the *ergodic* aleatory uncertainty is given as: $\sigma_T = \sqrt{\tau^2 + \phi_{SS}^2 + \phi_{S2S}^2}$, where τ, ϕ_{SS}^2 and ϕ_{S2S}^2 are the event-to-event, single station (non-ergodic) and station-to-station variabilities respectively. In exploring the epistemic uncertainty on σ_T, Al Atik (2015) presents three different models for the event-to-event and single station components of variability: (1) a "global" heteroskedastic model (calibrated primarily on the larger NGA West strong motion database and dependent on magnitude), (2) a "CEUS homoskedastic" model (calibrated upon Eastern US data and independent of magnitude) and (3) a "CEUS heteroskedastic" model (calibrated upon Eastern US data and dependent upon magnitude). In addition the uncertainty in the coefficients of these models are modelled according to a χ^2 distribution.

We opt for a single aleatory uncertainty model, following similar rationale laid out in the proposed NGA East GMM logic tree for the CEUS by Goulet et al. (2017), neglecting its epistemic uncertainty. From the models presented by Al Atik (2015), we define aleatory variability using both the mean "global heteroskedastic" event-to-event and single-station forms presented by Al Atik (2015). The ϕ_{S2S} model of Stewart et al. (2019) is adopted in order to calculate the regional ergodic hazard, in line with the approach of the 2018 US NSHM (Petersen et al. 2020).

The components of aleatory variability are compared against their counterparts within the general crustal seismicity backbone model for four magnitude scenarios in Fig. 10. In this case the aleatory uncertainty model of the general crustal seismicity backbone refers to the magnitude-dependent heteroskedastic form described in Weatherill et al. (2020), rather than the original homoskedastic model initially assumed by Kotha et al. (2020). The aleatory uncertainty in the general crustal seismicity backbone model is conditional upon whether the site is assumed to have a *measured* V_{S30} or whether this is *inferred* from a proxy such as topography, the latter resulting in a higher site-to-site variability ϕ_{S2S}. As explained in Weatherill et al. (2020), GMMs used for the general crustal seismicity logic tree adopt the same heteroskedastic within-event variability (τ) model taken directly from Al Atik (2015); hence, they are identical in Fig. 10. Heteroskedastic ϕ_{SS} for the general backbone model is calibrated directly on the residuals of the Kotha et al. (2020) GMM and is found to be slightly lower
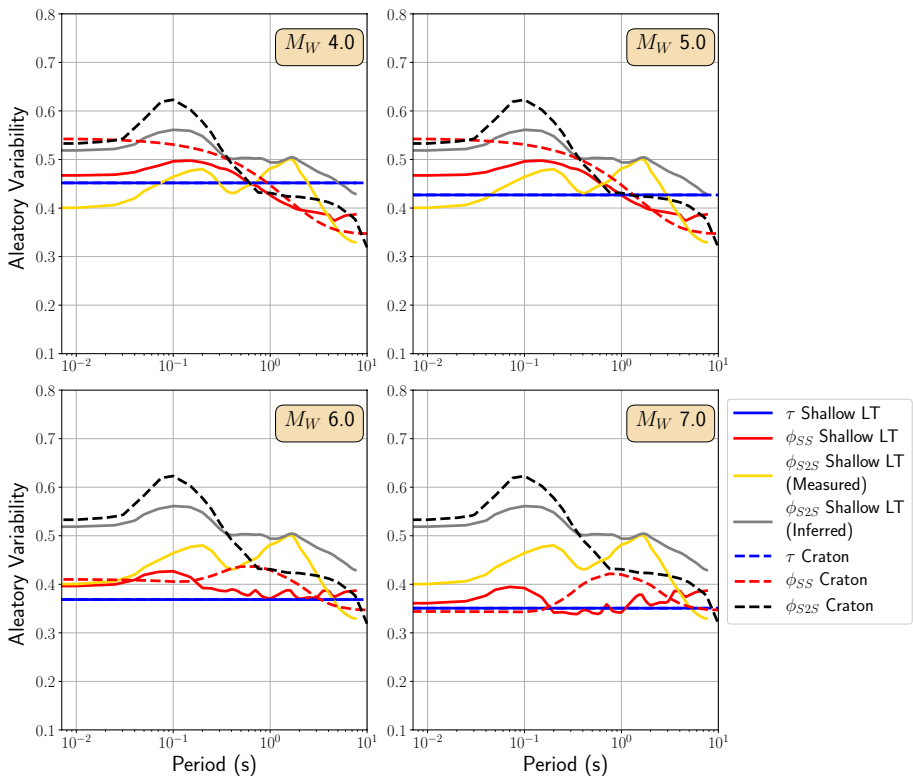
Fig. 10 Comparison of the components of the aleatory variability between the proposed craton GMM and the default shallow crustal ground motion model

for European records than that of Al Atik (2015) at smaller magnitudes. The greatest difference is seen in the site-to-site variability ϕ_{S2S}, where the values adopted for the craton model exceed both the *measured* and *inferred* V_{S30} cases at short periods, then reducing to values smaller than even the *measured* V_{S30} case for long periods.

5 Seismic hazard comparisons

The trellis plots shown in Figs. 5, 6, 7 and 8 illustrate how the parametric craton model compares with the suite adopted by the USGS and the previous ESHM13 selection. To assess the impact on the resulting PSHA calculations, however, both the median ground motion and the uncertainties (aleatory and epistemic) need to be taken into consideration. For this purpose we compare the resulting uniform hazard spectra (UHS) for the 475 and 2475 year return periods against those of US NSHMP models. The seismic hazard comparisons are made for a site located on the southern coast of Finland, using the area source branch of the ESHM13. This site is selected as it sufficiently deep within the cratonic region to be largely unaffected by ground motions from earthquakes in different tectonic environments further afield. The rate of occurrence of the area source model suggests that an earthquake of magnitude greater than or equal to M_W 4.5 occurs within 300 km of the site at an annual rate of 0.0232 (approximately once every 45 years). The minimum magnitude considered is 4.5 and the maximum magnitude M_W 7.0. To compare the parametric craton model against the US models, the epistemic uncertainty is described by a 9-branch discrete approximation to a Gaussian distribution $\mathcal{N}(0, \sigma_\mu)$ using the same Gaussian quadrature approach of Miller and Rice (1983).
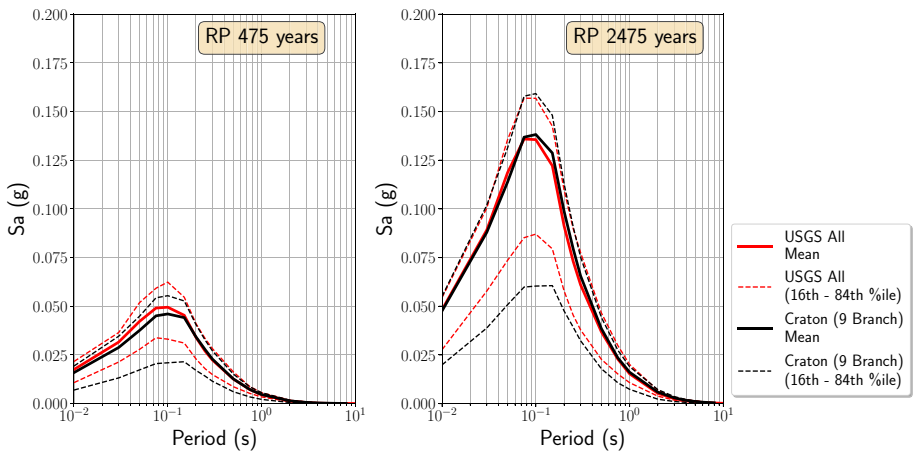
5.1 Comparison against the USGS model set

From a theoretical perspective, one would expect the closest agreement between the USGS approach and the parametric craton model for the "Sammons" models. The two UHS in Fig. 11 indicate that this is indeed the case. Close agreement in the mean UHS can be seen at both the 475 and 2475 year return periods, with the USGS models exceeding the parametric craton only in the period range 0.05–0.2 s at the longer return period. The 16–84 percentile range is slightly broader in the case of the parametric craton model, reflecting the larger σ_μ values in Eq. 6 with respect to the range of σ_μ for different M and R_{RUP} scenarios within the original NGA East suite.

The 9-branch parametric craton model is then compared against the complete US NSHMP logic tree for the CEUS, incorporating both the "Seed" and "Sammons" branches (Fig. 12). When combining the "Seed" models, which predict slightly higher mean hazard than the 9-branch parametric craton model in the short period range, and the "Sammons" models, which predict lower mean hazard, there exists a much closer agreement between full USGS CEUS logic tree approach the proposed craton model for the mean and upper percentiles of the UHS across the complete period range than is seen for either its constituent "Seed" or "Sammons" models. We acknowledge that the approach to characterise the centre, body and range of the epistemic uncertainty in the stable cratonic ground motions models is notably simpler than that proposed within Goulet et al. (2017), yet the comparisons here suggest the results to be in agreement for the mean UHS. Only the 16th percentile of the parametric craton model is notably lower, though to what extent this is significant would ultimately depend on how the results are intended for practical application.
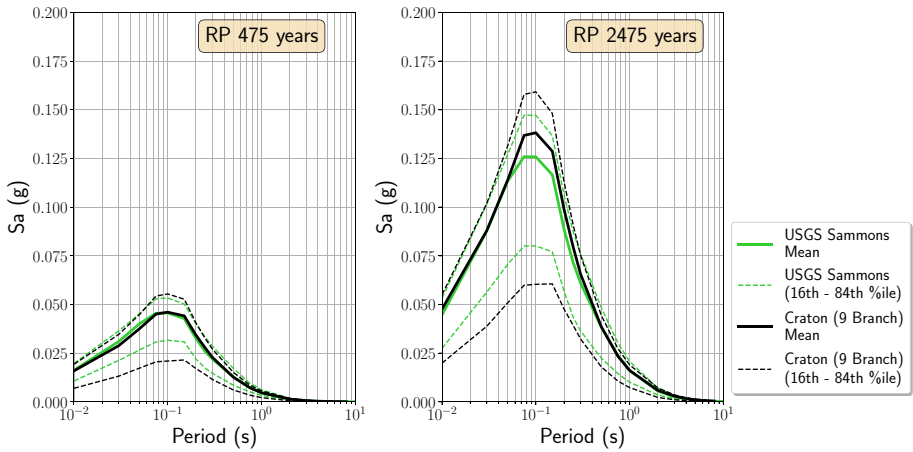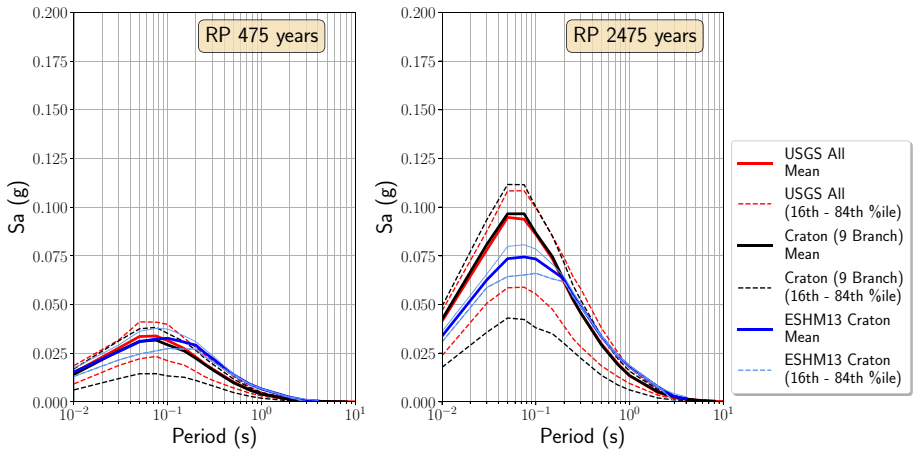
Fig. 11 Comparison of the UHS for the 9-branch parametric craton model and the USGS "Sammons" models for the 10% in 50 year return period (left) and 2% in 50 year return period (right), adjusted to a V_{S30} 800 m/s reference rock
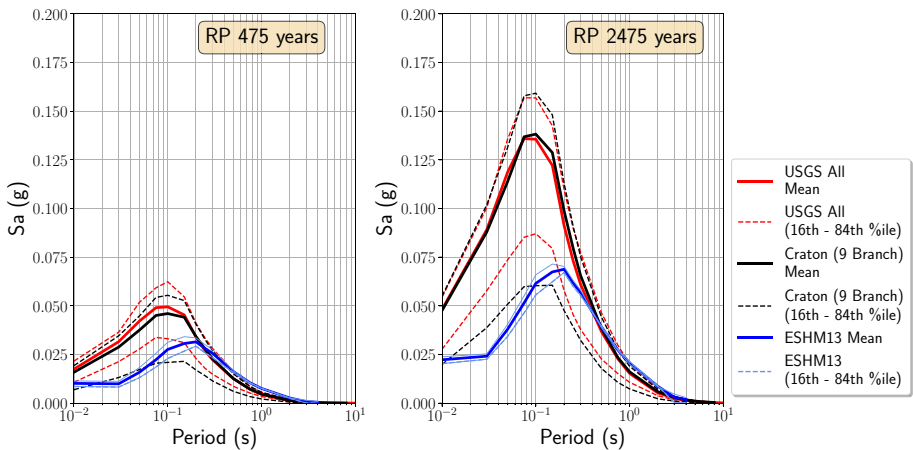

Fig. 12 Comparison of the UHS for the 9-branch parametric craton model and the complete USGS logic tree for the 10% in 50 year return period (left) and 2% in 50 year return period (right), adjusted to a V_{S30} 800 m/s reference rock

5.2 Comparison against the ESHM13 model set

The good agreement between the 9-branch logic tree for the parametric craton model and that of the CEUS ground motion logic tree demonstrates that the model proposed here suggest a consistency in the hazard results for the two approaches, the comparison against the models within the ESHM13 are perhaps more insightful to understand the impact of the new approach for the ESHM20. Both sets of models differ in how they adjust the ground motion from very hard rock to Eurocode 8 reference rock. Therefore it is critical to understand how the contrasting assumptions in this process impact on the

Fig. 13 Comparison of the UHS for the 9-branch parametric craton model, the complete USGS CEUS logic tree and the ESHM13 logic tree for $Sa(0.2s)$ (left) and $Sa(1.0s)$ assuming a very hard rock site condition of V_S 2800 m/s

Fig. 14 Comparison of the UHS for the 9-branch parametric craton model, the complete USGS logic tree and the ESHM13 logic tree for the 10% in 50 year return period (left) and 2% in 50 year return period (right), adjusted to a Eurocode 8 reference rock condition of V_S 800 m/s

hazard, and ultimately on seismic design. Figure 13 compares the resulting UHS from all three logic trees (USGS CEUS, craton 9-branch parametric and ESHM13) for a site on V_S 2800 m/s bedrock, while Fig. 14 compares the same logic trees for Eurocode 8 class A reference rock with V_{S30} 800 m/s. The UHS on very hard bedrock for the 475 year return period shows a remarkably close agreement between all three logic tree sets, albeit with the spectral peak in the USGS and parametric craton model sets shifting toward shorter periods (T < 0.1 s). At the 2475 year return period the USGS and parametric craton logic tree diverge from that of the ESHM13 selection at periods shorter

than 0.2 s, which reflects the increased aleatory variability at short periods in the more recent models compared to that of Campbell (2003) and Toro (2002).

The impact of the very hard rock to reference rock adjustment factors between the NGA East approach and that of Van Houtte et al. (2011) becomes clear when comparing the UHS on the Eurocode 8 reference rock (Fig. 14). Both the parametric craton model and the USGS models now predict notably higher hazard for short period motion at 475 years return period, with the greatest differences in the 0.05–0.1 s range. At 2475 years the divergence is even more pronounced, with the new amplification model resulting in a factor of two difference at short periods. For periods longer than 0.3 s, there is good agreement the UHS between the three approaches. This comparison serves to illustrate the extent to which the assumption of the nature of the hard-rock to reference rock adjustment can have a controlling impact on seismic hazard at short periods in regions of low seismicity.

6 The stable craton logic tree

In representing the complete epistemic uncertainty in the ground motion logic tree for stable cratonic environments, it is still necessary to return to the original hypothesis that strong motions in this region of Europe display fundamentally different source, path and site characteristics to those found within the ESM database. The parametric craton model is intended to capture the epistemic uncertainty based upon the assumption that the craton regions are significantly different from the rest of Europe, yet there is insufficient data to accept or reject this hypothesis with absolute certainty, and therefore both possibilities must be considered. The proposed logic tree for craton environments recognises this, opting to place the greater weight (0.8) on the set of branches describing the centre, body and range of the parametric craton model, and 0.2 on the possibility they the mid- and upper-branches of the general crustal seismicity backbone logic tree are appropriate. The complete logic tree is shown in Fig. 15.

Fig. 15 Proposed logic tree for application to the stable cratonic region of Europe with weights indicated in blue font. The term Δ refers to the envelop of τ_{L2L} and the statistical uncertainty (σ_μ^{SHAL}) of the Kotha et al. (2020) GMM. All other symbols are defined in the text
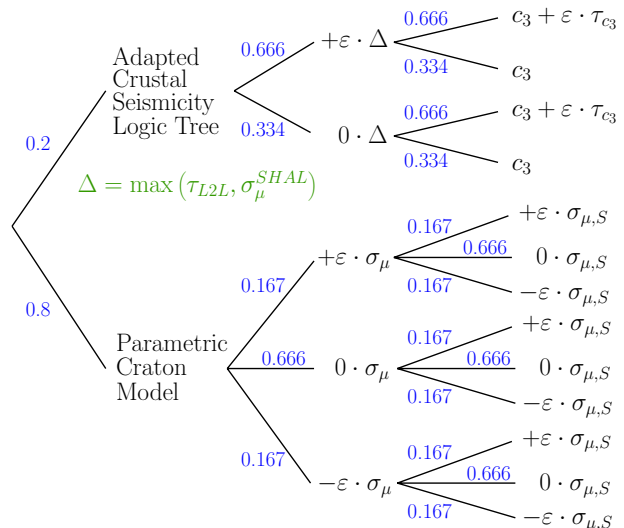
For the parametric craton model case, the resulting epistemic uncertainty is then represented by three branches ($-\varepsilon \cdot \sigma_\mu$, $0 \cdot \sigma_\mu$ and $+\varepsilon \cdot \sigma_\mu$) each weighted according to the three-point discrete approximation to the Gaussian distribution. In addition, as described in Sect. 4 we include a further set of branches describing the epistemic uncertainty in the amplification model ($\sigma_{\mu,S}$). As this term is also intended to be applied as a log-normal adjustment factor, $\sigma_{\mu,S}$ is modelled by the same three-point discrete approximation of Miller and Rice (1983).

For the case that the ground motion characteristics of the stable craton of northeastern Europe may be captured within the general crustal seismicity logic tree, we retain the high and central τ_{L2L} branches and the slow and central attenuation (τ_{C_3}) branches. We find it implausible that ground motions in northeastern Europe would systematically display lower stress drop and faster attenuation than the "average" of the shallow crustal dataset, and thus these branches of the shallow default logic tree are removed. Instead, the weights are reassigned to give high τ_{L2L} and high δ_{C_3} branches two thirds of the possible weighting, and the "central" branch only one third. To distinguish this subset of branches from the general crustal seismicity logic tree itself, we refer to these set of selected branches and their respective weights as the *adapted crustal seismicity* logic tree.

The assignment of weights between the two sets of branches (parametric craton model and adapted crustal seismicity logic tree) is not necessarily a trivial decision and the balance between them can have a significant influence on the resulting seismic hazard assessment. In Fig. 16 we illustrate the impact on the UHS for different weightings on the new parametric craton model branch set, ranging from 0.5 to 0.95. Over this weighting range we observe a change in seismic hazard at $Sa(0.1s)$ on the order of approximately 1.5 for both the 475 year and 2475 year return periods, though this change is smaller at both longer and shorter periods.

The final decision to assign the 0.8 weighting to the parametric model branches is based on an analysis of a small set of ground motion records from hard-rock (V_S 2800 m/s) sites
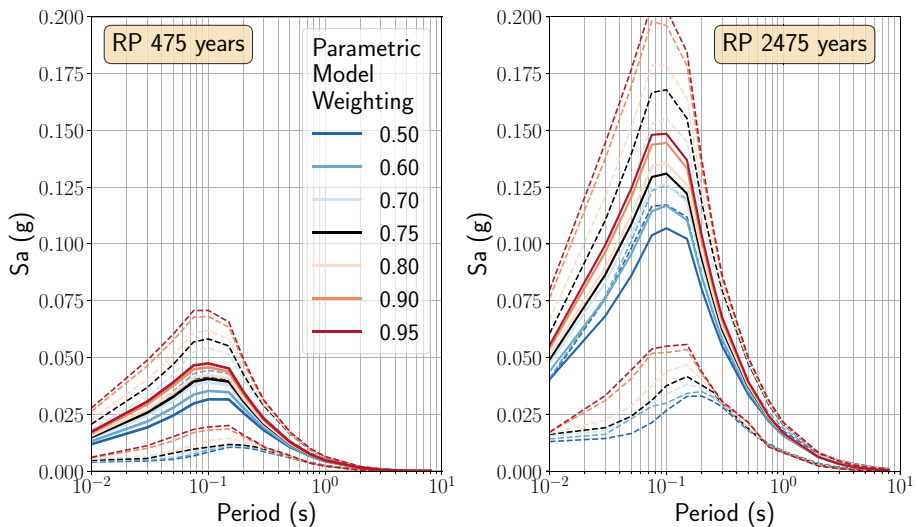

Fig. 16 Change in mean seismic hazard (solid lines) and 16th to 84th percentiles for the 475 year (left) and 2475 year (right) return period when changing the weight assigned to the new parametric craton model branch set in the logic tree

from low-magnitude earthquakes ($3.0 \leq M_L \leq 4.1$) in Fennoscandia. This dataset is published as an electronic supplement to Fülöp et al. (2020) and though we believe it is insufficient to assist in the calibration of the GMM itself, it may be a deciding factor in final weighting decision. Comparing the fit of the 13 different GMM branches to this small dataset using log-likelihood approach and corresponding weighting scheme of Scherbaum et al. (2009), we find that the total sum of the parametric logic tree branches is close to 0.8 at short-to-moderate periods (T < 0.5 s). As the models converge of their own accord for longer periods we see no reason to adjust the weights for longer periods.

Figures 17 and 18 show the comparisons of the mean and percentiles of the proposed craton logic tree with respect to that of the complete USGS CEUS model and the previous ESHM13 model. A closer agreement can be found between the proposed logic tree and that of the USGS, albeit that at shorter periods the proposed logic tree predicts smaller ground motions. This clearly reflects the influence of assimilating the branches of the adapted crustal seismicity logic tree into the craton model.

In terms of the comparisons with the ESHM13, these results suggest an overall increase in hazard at short periods with respect to the ESHM13 as well as a wider epistemic uncertainty range. Integrating some branches of the default shallow crustal logic tree modulates this slightly, but it is evident that the use of the NGA East site amplification model is producing a substantial change in the shape of the spectrum at high frequencies. Despite the change in the shape of the UHS at high frequencies, the overall difference in the amplitude of the peak of the spectra is similar at 475 years return period and the differences are not so substantial when averaged the 0.1–0.3 s range. In terms of impacts on engineering design for ordinary structures built according to Eurocode 8, despite the drastic change in shape the amplitude the resulting design spectrum might be comparable between the ESHM13 and ESHM20 models, albeit that the corner period delineating the boundary of the constant acceleration and constant velocity portions of the design spectrum should be moved toward shorter periods.
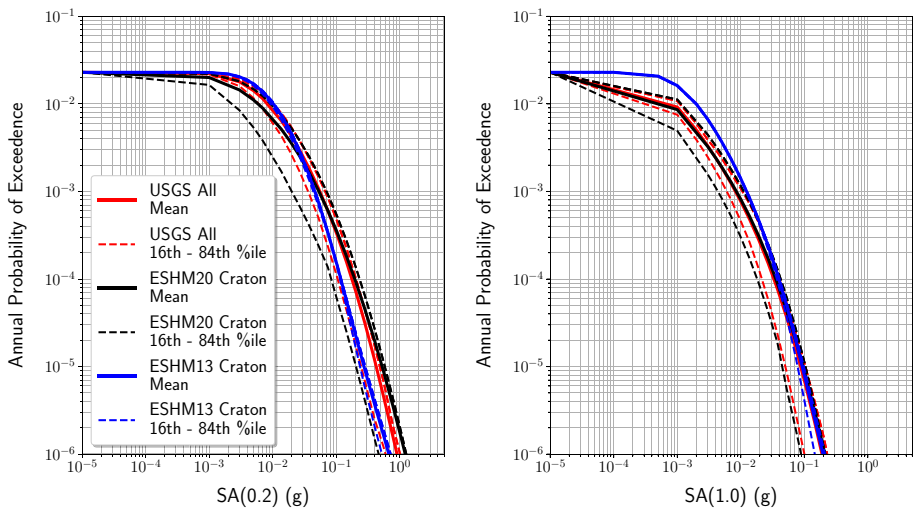
Fig. 17 Comparison of the seismic hazard curves for the proposed ESHM20 craton logic tree, the ESHM13 logic tree, and the complete USGS CEUS logic tree for $Sa(0.2s)$ (left) and $Sa(1.0s)$, both adjusted to a reference site condition of V_{S30} 800 m/s
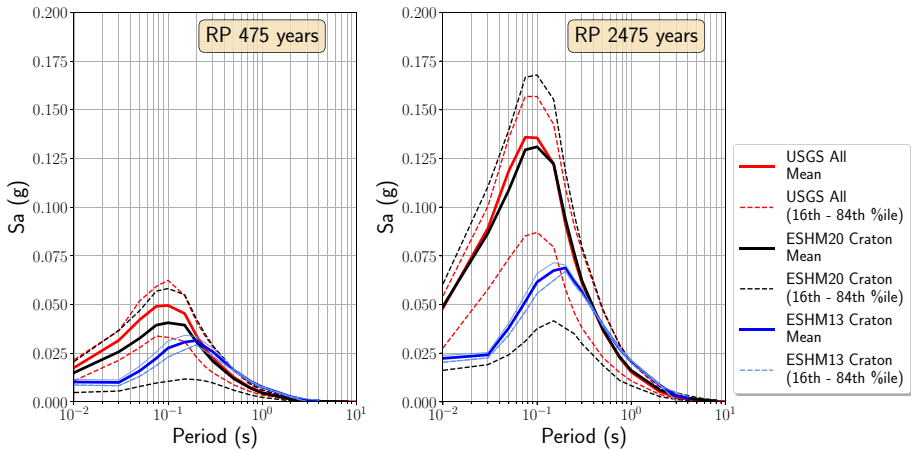
Fig. 18 Comparison of the UHS for the proposed ESHM20 craton logic tree, the ESHM13 logic tree, and the complete USGS logic tree for the 10% in 50 year return period (left) and 2% in 50 year return period (right), both for a reference site condition of V_{S30} 800 m/s

7 Conclusions

The characterisation of ground motion and its epistemic uncertainty for application to stable cratonic region of Europe requires special consideration in comparison to the broader approach adopted for shallow seismicity followed elsewhere. For much of Europe the region-to-region variability in stress parameter and attenuation implied by the ESM strong motion data constrains the epistemic uncertainty even in places of little to no data. But as discussed in Sect. 3, there are several lines of geophysical evidence that indicate an important difference in crustal properties between southern and western Europe (from where the ESM ground motion data originate) and those of northeastern Europe, and that a closer analogue may be found in eastern North America. Leveraging upon the developments of the NGA East project (Goulet et al. 2018) and the strategy for modelling ground motion in the CEUS in the recent US national seismic hazard model (Petersen et al. 2020), it has been possible to use the range of median ground motions implied by the NGA East ground motion models to re-calibrate the GMM for shallow seismicity used elsewhere in Europe to better capture the ground motion characteristics of stable cratonic environments, and correspondingly define a more appropriate distribution of epistemic uncertainty than that found within the ESHM13. All of these steps lead to a ground motion logic tree developed around a scaled backbone approach that attempts to apply the philosophy of a common ground motion model that is regionalised by available data, even to a part of Europe where we believe significant differences may exist but we lack the strong motion observations to quantify them. The complete ground motion model and logic tree are available for use in the OpenQuake-engine (Pagani et al. 2014).

The comparison of the parametric ground motion model against the recently proposed logic tree for the CEUS seems to suggest a generally good agreement in mean and upper percentiles of the hazard, despite the arguably simpler approach adopted here. The integration of some of the adapted crustal seismicity logic tree branches reduces the hazard here with respect to that of the parametric craton model alone, but we believe

this to be an appropriate compromise to reflect the epistemic uncertainty in the underlying premise that the ground motion differs substantially from that implied by the ESM dataset, which we cannot conclusively prove. The case study hazard example here suggests an increase in the hazard at short periods, and a slight decrease at longer periods, with respect to ESHM13. It must be emphasised, however, that the spatial extent over which this observation can be assumed to apply is limited and further sensitivity studies are being undertaken as part of the ESHM20.

The NGA East models and the models proposed for the Cratonic part of Europe in this paper consider that V_{S30} 800 m/s rock amplifications are associated to S-wave velocity profiles showing an impedance contrast and κ values between 0.01 and 0.02s. Are such profiles and attenuation properties consistent with similar V_{S30} rock site properties elsewhere in Europe? Recent analysis suggests that rock sites located in the shallow crustal part of Europe show similar κ values (0.01–0.02s) (Pilz et al. 2019), lower than the κ estimations considered by Van Houtte et al. (2011). Furthermore, using a database of seismic stations predominantly from central and southern Europe, Pilz et al. (2020) find that more than 40 % of sites with measured V_{S30} larger than 800 m/s show peaked horizontal-to-vertical spectral H/V ratios at frequencies greater than 1 Hz, implying strong near surface impedance contrasts that may arise due to weathering or zones of variable fracture density. Though their dataset did not include stations from Scandinavia and northeastern Europe, the presence of short period amplification at rock sites elsewhere in Europe suggests a stronger influence of high-impedance in rock ground motion than previously assumed. This may have implications for the definition of the reference rock used for seismic hazard analysis.

The acquisition of strong motion data could assist in revising the craton GMM presented here, but the time-scale over which a sufficient volume of seismic data is accumulated to have any meaningful impact on the epistemic uncertainty in northeastern Europe is likely to be on the order of decades to centuries, rather than years. As such, the need to pool insights from seismically analogous regions in order to constrain the ground motions in cratonic environments remains inevitable. This does not necessarily mean that there is no scope for refinement of the model, especially given the role that the Stewart et al. (2020) site amplification model may be playing in controlling the shape of the UHS. Inferences on the seismic properties of the rock within the shallow crust from local active and passive source microzonation studies would be highly influential in refining the site amplification model, while still retaining the current logic tree for the very hard rock (V_{S30} 3000 m/s) input. Though difficult to apply at regional scale, with well-directed microzonation studies targeting regions of highest population or economic exposure in northeastern Europe, better constraint of the seismic hazard inputs both for engineering design and for the analysis of risk could be achieved.

References

Al Atik L (2015) NGA-east: ground-motion standard deviation models for Central and Eastern North America. Tech. Rep. PEER Report No. 2015/07, Pacific Earthquake Engineering Research Center

Al Atik L, Youngs R (2014) Epistemic uncertainty for NGA-West2 models. Earthq Spectra 30(3):1301–1318

Asch K (2005) The 1:5 million international geological map of Europe and adjacent areas (digital dataset). Schweizerbart Science Publishers, Stuttgart

Atkinson GM, Adams J (2013) Ground motion prediction equations for application to the 2015 Canadian national seismic hazard maps. Can J Civ Eng 40(10):988–998

Atkinson GA, Bommer JJ, Abrahamson NA (2014) Alternative approaches to modeling epistemic uncertainty in ground motions in probabilistic seismic-hazard analysis. Seismol Res Lett 85(6):1141–1144

Basili R, Brizuela B, Herrero A, Iqbal S, Lorito S, Maesano F, Murphy S, Perfetti P, Romano F, Scala A, Selva J, Taroni M, Thio H, Tiberti M, Tonini R, Volpe M, Glimsdal S, Harbitz C, Løvholt F, Baptista M, Carrilho F, Matias L, Omira R, Babeyko A, Hoechner A, Gurbuz M, Pekcan O, Yalçıner A, Canals M, Lastras G, Agalos A, Papadopoulos G, Triantafyllou I, Benchekroun S, Agrebi Jaouadi H, Attafi K, Ben Abdallah S, Bouallegue A, Hamdi H, Oueslati F (2019) NEAMTHM18 documentation: the making of the TSUMAPS-NEAM tsunami hazard model 2018. Tech. rep., InIstituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.5281/zenodo.3406625

Bommer JJ (2012) Challenges of building logic-trees for probabilistic seismic hazard analysis. Earthq Spectra 28(4):1723–1735

Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, Morris PA (1997) Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts: main report (nureg/cr-6372, volume 1). Tech. rep., U.S. Nuclear Regulatory Commission

Campbell KW (2003) Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in Eastern North America. Bull Seismol Soc Am 93(3):1012–1033

Chen YS, Weatherill G, Pagani M, Cotton F (2018) A transparent and data-driven global tectonic regionalization model for seismic hazard assessment. Geophys J Int 213:1263–1280

Cotton F, Scherbaum F, Bommer J, Bungum H (2006) Criteria for selecting and adjusting ground-motion models for specific target regions: application to central Europe and rock sites. J Seismol 10:137–156

Delavaud E, Cotton F, Akkar S, Scherbaum F, Danciu L, Beauval C, Drouet S, Douglas J, Basili R, Sandikkaya MA, Segou M, Faccioli E, Theodoulidis N (2012) Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. J Seismol 16(3):451–473

Douglas J (2018a) Capturing geographically-varying uncertainty in earthquake ground motion models or what we think we know may change. In: Pitilakis K (ed) Recent advances in earthquake engineering in Europe, vol 46, pp 153 – 181

Fülöp L, Jussila V, Aapasuo R, Vourinen T, Mäntyniemi P (2020) A ground-motion prediction equation for fennoscandian nuclear installations. Bull Seismol Soc Am 110(3):1211–1230

Goulet C, Bozorgnia Y, Kuehn N, Al Atik L, Youngs RR, Graves RW, Atkinson GM (2017) NGA-east ground-motion models for the U. S. geological survey national seismic hazard maps. Tech. Rep. PEER Report No. 2017/03, Pacific Earthquake Engineering Research Center

Goulet C, Bozorgnia Y, Abrahamson N, Kuehn N, Al Atik L, Youngs R, Graves R, Atkinson G (2018) Central and Eastern North America ground-motion characterisation: NGA-east final report. Tech. Rep. PEER Report No. 2018/08, Pacific Earthquake Engineering Research Center

Grad M, Tiira T, the ESC Working Group (2009) The Moho depth map of the European plate. Geophys J Int 176:279–292

Graizer V (2016) Ground-motion prediction equations for Central and Eastern North America. Bull Seismol Soc Am 106(4):1600–1612

Hashash YMA, Ilhan O, Harmon JA, Parker GA, Stewart JP, Rathje EM, Campbell KW, Silva WJ (2020) Nonlinear site amplification model for ergodic seismic hazard analysis in Central and Eastern North America. Earthq Spectra 36(1):69–86

Kotha SR, Weatherill GA, Bindi D, Cotton F (2020) A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. Bull Earthq Eng 18(9):4091–4125

Lanzano G, Sgobba S, Luzi L, Puglia R, Pacor F, Felicetta C, D'Amico M, Cotton F, Bindi D (2019) The pan-European Engneering Strong Motion (ESM) flatfile: compilation criteria and data statistics. Bull Earthq Eng 17:561–582

Lucazeau F (2019) Analysis and mapping of an updated terrestrial heat flow data set. Geochem Geophys Geosyst 20:4001–4024

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

Mitchell BJ, Cong L, Ekström G (2008) A continent-wide map of 1-Hz L_g coda Q variation across Eurasia and its relation to lithospheric evolution. J Geophys Res 113:B04303

Mooney WD, Ritsema J, Hwang YK (2012) Crustal seismicity and the earthquake catalog maximum moment magnitude m_{cmax} in stable continental regions (SCRs): correlation with the seismic velocity of the lithosphere. Earth Planet Sci Lett 357–358:78–83

Pagani M, Monelli D, Weatherill G, Danciu L, Crowley H, Silva V, Henshaw P, Butler L, Nastasi M, Panzeri L, Simionato M, Vigano D (2014) Openquake engine: an open hazard (and risk) software for the global earthquake model. Seismol Res Lett 85(3):692–702

PEER (2015) NGA-east: median ground-motion models for the Central and Eastern North America region. Technical report PEER report no. 2015/04, Pacific Earthquake Engineering Research Center

Petersen MD, Shumway AM, Powers PM, Mueller CS, Moschetti MP, Frankel AD, Rezaeian S, McNamara DE, Luco N, Boyd OS, Rukstales KS, Jaiswal KS, Thompson EM, Hoover SM, Clayton BS, Field EH, Zeng Y (2020) The 2018 update of the US National Seismic Hazard Model: overview of model and implications. Earthq Spectra 36(1):5–41

Pezeshk S, Zandieh A, Tavakoli B (2011) Hybrid empirical ground-motion prediction equations for eastern North America using NGA models and updated seismological parameters. Bull Seismol Soc Am 101(4):1859–1870

Pilz M, Cotton F, Zaccarelli R, Bindi D (2019) Capturing regional variations of hard-rock attenuation in Europe. Bull Seismol Soc Am 109(4):1401–1418

Pilz M, Cotton F, Kotha SR (2020) Data-driven and machine learning of seismic reference stations in Europe. Geophys J Int 222:861–873

Scherbaum F, Delavaud E, Riggelsen C (2009) Model selection in seismic hazard analysis: an information-theoretic perspective. Bull Seismol Soc Am 99(6):3234–3247

Scherbaum F, Kuehn N, Ohrnberger M, Koehler A (2010) Exploring the proximity of ground-motion models using high-dimensional visualization techniques. Earthq Spectra 26(4):1117–1138

Stewart JP, Parker G, Al Atik L, Atkinson GM, Goulet C (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, Boore DM, Hashash YMA, Silva WJ (2020) Ergodic site amplification model for Central and Eastern North America. Earthq Spectra 36(1):42–68

Szwillus W, Afonso JC, Ebbing J, Mooney WD (2019) Global crustal thickness and velocity structure from geostatistical analysis of seismic data. J Geophys Res Solid Earth 124:1626–1652

Toro GR (2002) Modification of the Toro et al (1997) attenuation equations for large magnitude and short distances. Technical report, Risk Engineering, Inc

U S Nuclear Regulatory Commission (2012) Practical implementation guidelines for SSHAC Level 3 and 4 Hazard Studies, NUREG-2117 rev. 1. Technical report, U. S. Nuclear Regulatory Commission, Washington DC

Van Houtte C, Drouet S, Cotton F (2011) Analysis of the origins of κ (kappa) to compute hard rock to rock adjustment factors for GMPEs. Bulletin of the Seismologicsal Society of America 101(6):2926–2941

Weatherill GA, Kotha SR, Cotton F (2020) 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

Woessner J, Laurentiu D, Giardini D, Crowley H, Cotton F, Grünthal G, Valensise G, Arvidsson R, Basili R, Demircioglu M, Hiemer S, Meletti C, Musson R, Rovida A, Sesetyan K, Stucchi M, The SHARE Consortium (2015) The 2013 European seismic hazard model: key components and results. Bull Earthq Eng 13(12):3553–3596

Yenier E, Atkinson GM (2015) Regionally adjustable generic ground-motion prediction equation based on equivalent point-source simulations: Application to central and eastern north america. Bull Seismol Soc Am 105(4):1989–2009

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.