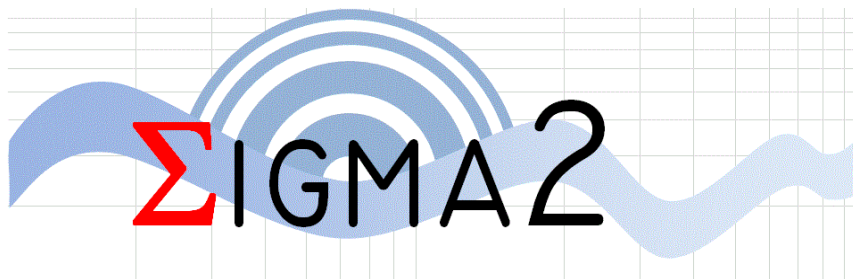



Implementation of the Drouet et al. (2020) PSHA for France into OpenQuake: Comparisons and Modelling Issues

Work Package 5 PSHA



AUTHORS		REVIEW		APPROVAL	
Name	Date	Name	Date	Name	Date
Graeme Weatherill Fabrice Cotton Guillaume Daniel Irmela Zentner	2022/03/18	Pierre Labbé Christophe Martin	2022/08/04 <i>all</i> 2022/.../.....	 Emmanuel Viallet Public access <input checked="" type="radio"/> SIGMA-2 restricted <input type="radio"/>	2023-02-17

Document history

DATE	VERSION	COMMENTS
2021/10/04	0	<i>Creation and submission to reviewers</i>
2022/03/19	1	<i>Revision in response to reviewer feedback</i>

Executive summary

As part of an ongoing initiative to develop a comprehensive and homogeneously implemented set of probabilistic seismic hazard models in Europe, we have undertaken a translation of the probabilistic seismic hazard model for Metropolitan France, published by Drouet et al. (2020), into the OpenQuake software. The process of translating a probabilistic seismic hazard model (PSHM) between software brings to light specific issues in the characterisation of the seismogenic source models and ground motion models, and in the execution of the logic tree for capturing epistemic uncertainty. These issues relate to both the explicit assumptions of the modellers as well as the implicit processes built into the software themselves, and these can yield differences in the seismic hazard curves that may be relevant for the intended users of the outputs.

This deliverable serves as a complete documentation of the translation process, focusing on the specific details needed for comprehensive characterisation of all of the seismogenic sources and ground motion models implemented within the logic tree. We provide an overview of the components of the Drouet et al. (2020) model before moving on to describe the means by which OpenQuake can characterise their behaviour and the modelling assumptions needed to complete the implementation. Though the Drouet et al. (2020) model is generally well documented, with additional information provided directly from the authors we highlight specific modelling details and their implications for seismic hazard analysis that are not otherwise found in the supporting documentation.

To establish whether the resulting seismic hazard curves from the OpenQuake implementation agree well with those of the original model, we undertook detailed branch-by-branch comparisons for five target cities (Brest, Lourdes, Marseille, Grenoble and Nice) before then comparing the mean and quantiles for these same sites. We set a target threshold for “agreement” as being no more than $\pm 10\%$ discrepancy between the curves for annual probabilities of exceedance (APoEs) greater than or equal to 10^{-4} . The branch-by-branch comparisons show that this level of agreement is achieved for the vast majority of considered branches, with reasons for discrepancies explored in further detail. At the time of preparing the deliverable, however, agreement was not achieved for the mean and quantiles, the reasons for which relate to incomplete information for the characterisation of activity uncertainty on the smoothed seismicity source model, which are elaborated upon in the body of the report.

The dual implementations of the PSHM for France in both the original software used by the authors and that of OpenQuake show that agreement between the two can only be achieved when a substantial amount of detail about the model and the processes in the underlying software are provided. We end this deliverable with a short set of recommendations for managing and standardising the process of PSHM implementation and documentation in a manner that provides greater clarity for other modellers and enhances the reproducibility of the model.

The final OpenQuake implementation of the Drouet et al. (2020) model, including input files, outputs and tools for construction, are available for download from <https://git.gfz-potsdam.de/gweather/francepsa>

1. Objectives of the Work

The seismic hazard model prepared by GEOTER (GEOTER, 2017), and published in Drouet et al. (2020), provides a comprehensive and up-to-date assessment of earthquake hazard within metropolitan France (France, hereafter). The model itself adopts recent harmonized earthquake catalogues (Cara et al., 2015; Manchuel et al., 2017), latest generation ground motion models (GMMs) for application to France, and state-of-the-art methods for characterizing epistemic uncertainty in earthquake recurrence. Given the relevance of the model for seismic hazard assessment not only in France but in Europe in general, further efforts have been made to compare and contrast the outcomes of this model with that of other probabilistic seismic hazard models that overlap geographically with the region covered by Drouet et al. (2020). Of particular interest are the German national seismic hazard model of Grünthal et al. (2018) and the forthcoming 2020 European Seismic Hazard Model (Danciu et al., 2022 *in prep.*). Additionally, comparisons with other models from neighbouring territories, e.g., Switzerland (Wiemer et al. 2016), Italy (Meletti et al. 2021) and the United Kingdom (Mosca et al. 2020), may also be undertaken by other parties in the future. Previous analysis by Mak & Cotton (2020) presented quantitative comparisons of the epistemic uncertainty distributions of the seismic hazard results provided by the GEOTER model (GEOTER, 2017) and the Germany PSHA model, for a set of sites in the region close to the France-Germany border where the two models overlap. They highlight important differences between the hazard results, which reflect the differences (and similarities) in the underlying modelling assumptions.

One factor that is often overlooked in comparisons of the hazard results, however, are differences that emerge due to different probabilistic seismic hazard assessment (PSHA) software adopted by different organization's or scientists. Though the theoretical basis for probabilistic seismic hazard assessment is well established from Cornell (1968) and McGuire (1976), and more recently re-interpreted by Field et al. (2003), the exact manner in which PSHA is implemented still varies from software to software. The differences emerge as each PSHA software developer may introduce specific ideas or algorithms to capture certain elements of both the aleatory variability and epistemic uncertainty in the earthquake process. Different algorithms or approaches can arise from different interpretations of models that can be found in the scientific literature, different ideas as to how to characterize various elements of the earthquake process, as well as potential optimisations designed to improve the computational efficiency and allow larger scale calculation. The diversity of hazard results from different PSHA software, even in "simple" or "controlled" seismic source and ground motion model configurations, have been shown in the PEER PSHA Benchmarking studies (Thomas et al. 2010; Hale et al. 2018). Recognition of software-to-software differences in seismic hazard results prompted the inclusion of Quality Assurance studies in PSHA for Nuclear Power Plants (e.g., Bommer et al., 2013; Tromans et al. 2019) in which the proposed models are implemented simultaneously in different PSHA software in order to understand and resolve potential differences in interpretations. Similar benchmarking and/or quality assurance processes have also been undertaken in national level seismic hazard projects, often as a precursory step to migrating from one particular PSHA calculation software to another (e.g., Allen et al., 2020; Abbot et al., 2020).

PSHA software comparison studies such have these have not only helped to identify potential errors in the calculation software, they also highlight those areas of the calculation in which software are likely to diverge, bringing to the fore modelling choices or theoretical discussions that have previously been overlooked or simply considered implicit. Though these discrepancies are seldom the dominant cause of differences in seismic hazard curves or maps from one model to another, they may be a contributing factor and one that introduces a certain level of unavoidable disagreement that would not be resolved even if the exact same modelling choices were made from one hazard study to another. It can therefore be a sensible course of action, where resources allow, to attempt migration of a seismic hazard model from one software to another, or to render all models into a common PSHA software implementation, as an initial step when comparing different models in a region.

For many of the recent generation of PSHA models in Europe, the OpenQuake-engine (Pagani et al., 2014a), an open-source software for probabilistic seismic hazard and risk analysis, has been the preferred calculation software. Among the models that have been implemented in this software are the 2013 European Seismic Hazard Model [ESHM13] (Wössner et al. 2015); the 2015 Swiss Seismic Hazard Model [SUIhaz2015] (Wiemer et al. 2015), the 2019 Italian Seismic Hazard Model [MPS19] and the 2020 European Seismic Hazard Model [ESHM20]. The Drouet et al. (2020) PSHA for France and the Grünthal et al. (2018) PSHA for Germany are exceptions to this. Both are implemented proprietary software belonging to their organizations, Fugro and the German Research Centre for Geosciences (GFZ) respectively. As part of a series of longer running activities to understand and compare seismic hazard models in France and Germany, and as a precursory step toward comparisons with the forthcoming ESHM20 model, we have attempted to migrate the PSHA models of Drouet et al. (2020) and Grünthal et al. (2018) into the OpenQuake framework. The present report describes the process of implementation and verification for the Drouet et al. (2020) model for France, while that of the Germany PSHA model will be addressed in a future publication.

The implementation of the Drouet et al. (2020) model into OpenQuake serves multiple objectives. The first is to incorporate it into a set of seismic hazard models for Europe that are implemented into a common and calculation software. This will help elucidate potential differences in interpretations or methodological implementations that may contribute to discrepancies in the seismic hazard calculation owing to the either the software alone or the software-enforced modelling decisions. Though the original model implementations will remain proprietary to their respective organizations, these migration efforts are accompanied by interactions with the model authors to help explain differences where they emerge. With their authorization, we expect to make public the OpenQuake implementations of the models for use by the wider seismological community, whereupon we hope that this set of models will form a crucial data set for future activities in the areas of seismic hazard model evaluation and quantitative testing.

This report will begin with a short review of the Drouet et al. (2020) PSHA model, including the seismic source characterization and ground motion characterization models as well as the resulting hazard outputs that are used as targets for comparing implementations. In the subsequent chapter we outline the relevant characteristics of the OpenQuake input models, highlighting how the information from the input files of Drouet et al. (2020) are translated into the OpenQuake format and where potential decisions or discrepancies can emerge. The two sections that follow on from that will present the comparisons and sensitivities of the seismic hazard calculations, firstly at a branch-by-branch level and then looking at the whole distribution of mean and fractiles of the logic tree. Finally, a short set of recommendations will be made to as a basis to guide future comparison studies.

The seismic hazard input data have been made available to this project from both the electronic supplement to the Drouet et al. (2020) paper and from data supplied to EDF by the original authors. We sincerely thank Stéphane Drouet (Fugro), Gabriele Ameri (Seister) and David Baumont (Seister) for providing the relevant data sets and insights into the operation of the original PSHA code. The main seismic hazard calculation engine is OpenQuake, which can be downloaded from <https://github.com/gem/oq-engine/>. Additional open-source code to construct the OpenQuake input files and post-process the results, as well as the files themselves, are available from <https://git.gfz-potsdam.de/gweather/francepsha>

2. Overview of the Drouet *et al.* (2020) PSHA Model and Target Data Set

2.1 Seismic Source Characterization (SSC) Model

The seismic hazard model of Drouet *et al.* (2020) is a model that builds upon what will be referred to here as *distributed seismicity* sources (e.g., Pagani et al., 2014). As a region of low-to-moderate seismicity and slow tectonic deformation, for much of metropolitan France the specific fault structures that produce the observed seismicity cannot be fully defined in a manner compatible for PSHA. The primary guide to the spatial extent and seismic productivity of the source is therefore the earthquake catalogue. For this purpose, the F-CAT earthquake catalogue has been adopted (Manchuel et al. 2018). The SSC considers two types of *distributed seismicity* source:

- i) **uniform area sources:** a topologically continuous set of polygons within which the seismicity is considered to be distributed uniformly (every location within the zone has an equal probability of being the epicentre of a future earthquake). The spatial extent of each polygon is defined according to an expert driven process, interpreting multiple strands of information including the observed seismicity distribution as well as local/regional-scale tectonics and geology. Earthquake recurrence models (or *magnitude frequency distributions*) are defined specifically to each polygon, though in cases of low seismicity some attributes of the recurrence model may be shared among several polygons within a larger *domaine* (or else the properties of the magnitude frequency distribution for the larger scale *domaine* may be adopted as priors in a Bayesian updating process).
- ii) **smoothed/gridded seismicity sources:** a regularly spaced grid of points covering the entire territory for which the magnitude frequency distribution is specific to that point. Such a model is built by applying a 2-D smoothing kernel to the [usually] temporally normalized distribution of seismicity. The earthquake recurrence model differs from point to point as the activity rate is determined specific to each point. Some properties of the magnitude frequency distribution (e.g., maximum magnitude (M_{MAX}) and/or b-value) may be common to groups of points.

To account for epistemic uncertainty in the spatial distribution of seismicity, including different interpretations of the regional tectonics, Drouet *et al.* (2020) consider three uniform area source models:

- 1) GEOTER seismotectonic model (GTR), based on their (GEOTER's) own interpretation of data collected from regional seismotectonic strain mechanisms. The complete model is presented in the publication of Le Dortz et al. (2019). According to Drouet et al. (2020) (references and cross-references omitted for brevity), "*The current model was first developed by defining large seismotectonic units relatively homogeneous at the national scale relying on the analysis of the geodynamic context, on maps of the Moho depth discontinuity (...), and Bouguer anomaly (...) to identify the crustal transitional zones (...). Then, within each large domain a more subtle zoning is performed based on coherent interpretation of available seismological and geological data (...).*". GTR therefore describes two models, one containing large-scale seismic domains (*DOMAINES*, or *DOM*, hereafter) and another that subdivides these domains into smaller sources.
- 2) Électricité de France seismotectonic model (EDF), an update of previous models for past seismic hazard assessment in France compiled by EDF. Adjustments to this model have been made for the purpose of the current project by way of aggregating some very small-scale zones that existed in the model for the purposes of deterministic seismic hazard assessment, which would not be suitable for PSHA.

- 3) Institut de Radioprotection et de Sûreté Nucléaire (IRSN) seismotectonic model, originally published by Baize et al. (2013).

The spatial extents of the three uniform area source models (GTR, EDF and IRSN), plus the DOMAINES, are shown in Figure 1. In the final logic tree, each of the three uniform models is assigned an equal weight.

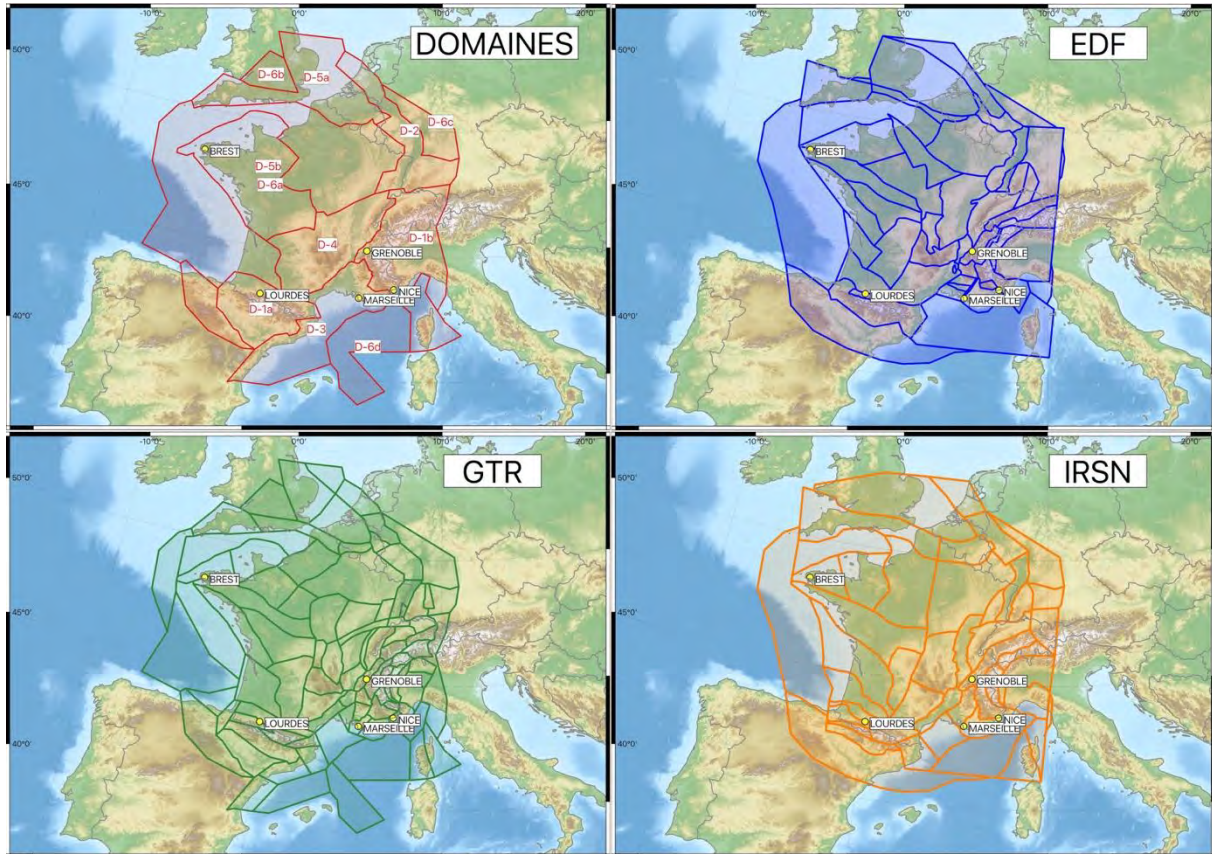


Figure 1: Seismic source zonations considered by Drouet et al. (2020), DOMAINES (top left), EDF (top right), GTR (bottom left) and IRSN (bottom right)

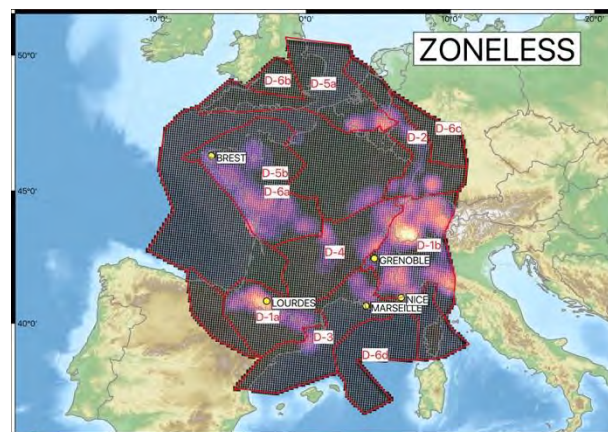


Figure 2: Spatial distribution of the “zoneless” model shown with the respective DOMAINES. Colours are scaled according to $\log_{10} a$, where a is the a -value of the Gutenberg-Richter distribution.

An alternative means of source characterization is also included by Drouet et al. (2020), which is a “ZONELESS” (ZL) or smoothed seismicity model. Here the seismic catalogue is used in conjunction with

the DOMAINES zonation, and a 2-D adaptive smoothing kernel is applied to the observed seismicity. The adaptive kernel uses the distance to the 10th closest neighbour to define the bandwidth, and though it is not specified in either the Drouet et al. (2020) paper or the GEOTER (2017) report we presume that the kernel itself is a 2-D Gaussian kernel. Important to note here is the condition that, “*Strict boundary conditions are imposed meaning that the kernels cannot cross the domains boundaries*” (Drouet et al., 2020). We interpret this to mean that the smoothing is applied only on a domain-by-domain basis, and that distributions of b-value and M_{MAX} are applied with respect to the DOMAIN in which the grid cell is found. Figure 2 shows the a-values of the gridded seismicity cells.

2.2 Magnitude Frequency Distribution

All sources in each of the source models considered (GTR, EDF, IRSN, ZONELESS) adopt a truncated Gutenberg-Richter distribution to describe the activity rates associated with each magnitude:

$$N(m) = 10^a \frac{e^{-b \ln(10)M} - e^{-b \ln(10)M_{max}}}{1 - e^{-b \ln(10)M_{max}}} \quad (1)$$

The coefficients a and b are determined using a Bayesian maximum likelihood procedure (EPRI, 2012) that accounts for regional variation and uncertainties in completeness as well as uncertainties on magnitude and location. b values are estimated firstly for the larger scales DOMAINES zones, which are then used as prior distributions for the b value estimates for each area source. For the uniform area sources in the GTR, EDF and IRSN models, the full distributions of a and b , including their respective standard deviations σ_a and σ_b and correlation $\rho_{\sigma_{ab}}$, are defined. For the zoneless models the b values are fixed according to those of the respective DOMAINE for each grid point, while a obviously varies in accordance with the smoothed seismicity rate.

The maximum magnitude, M_{MAX} , for each zone is based on the Bayesian “EPRI” approach (Johnston et al., 1994; EPRI, 2012). Distributions of M_{MAX} are defined for each of the primary domaines (D-1, D-2, ... D-6), with absolute bounds of $\min(M_{MAX}) = 5.5$ and $\max(M_{MAX}) = 7.1$ (for D-2 – D6) and $\max(M_{MAX}) = 7.5$ for D-1. The resulting posterior distributions are used to describe the uncertainty in M_{MAX} for each zone. It should be noted that in the implementation of the source models, correlation between maximum magnitude and the a and b distribution are not explicitly accounted for, as the two distributions are sampled independently. The spatial distribution of the a and b values, alongside the observed M_{MAX} for input into the “EPRI” approach and the maximum hypocentral depths of the zones are shown in Figure 3.

2.3 Source Properties

In addition to the magnitude frequency distribution, further information is needed to describe the seismogenic source properties within each zone: the hypocentral depth distribution and the rupture mechanism distribution. The critical consideration here is that in the original software the PSHA calculation itself is using *virtual faults* to define the finite rupture properties needed by the ground motion models. *Virtual faults* are synthetically generated finite fault surfaces whose depths and rupture mechanism (in terms of strike, dip and rake) are defined according to distributions of these properties for each source. The dimensions of the finite rupture, in terms of length and width, scale in proportion to the magnitude of the earthquake being considered. However, the exact manner in which the ruptures are generated, with respect to both the scaling of aspect ratio (i.e., length to width ratio) with magnitude and to the position of the hypocentre within the rupture plane as a function of both rupture size and depth, can be one of the fundamental differences between different PSHA software that adopt the *virtual fault* approaches in distributed seismicity zones¹.

¹ Not all PSHA software adopt *virtual faults* for the purpose of characterizing rupture finiteness in the calculation. Other codes may instead use empirical formulae or other means to approximate the Author(s) – Implementation of the Drouet et al. (2020) PSHA for France into OpenQuake: Comparisons and Modelling Issues- SIGMA2-2021-D5-085

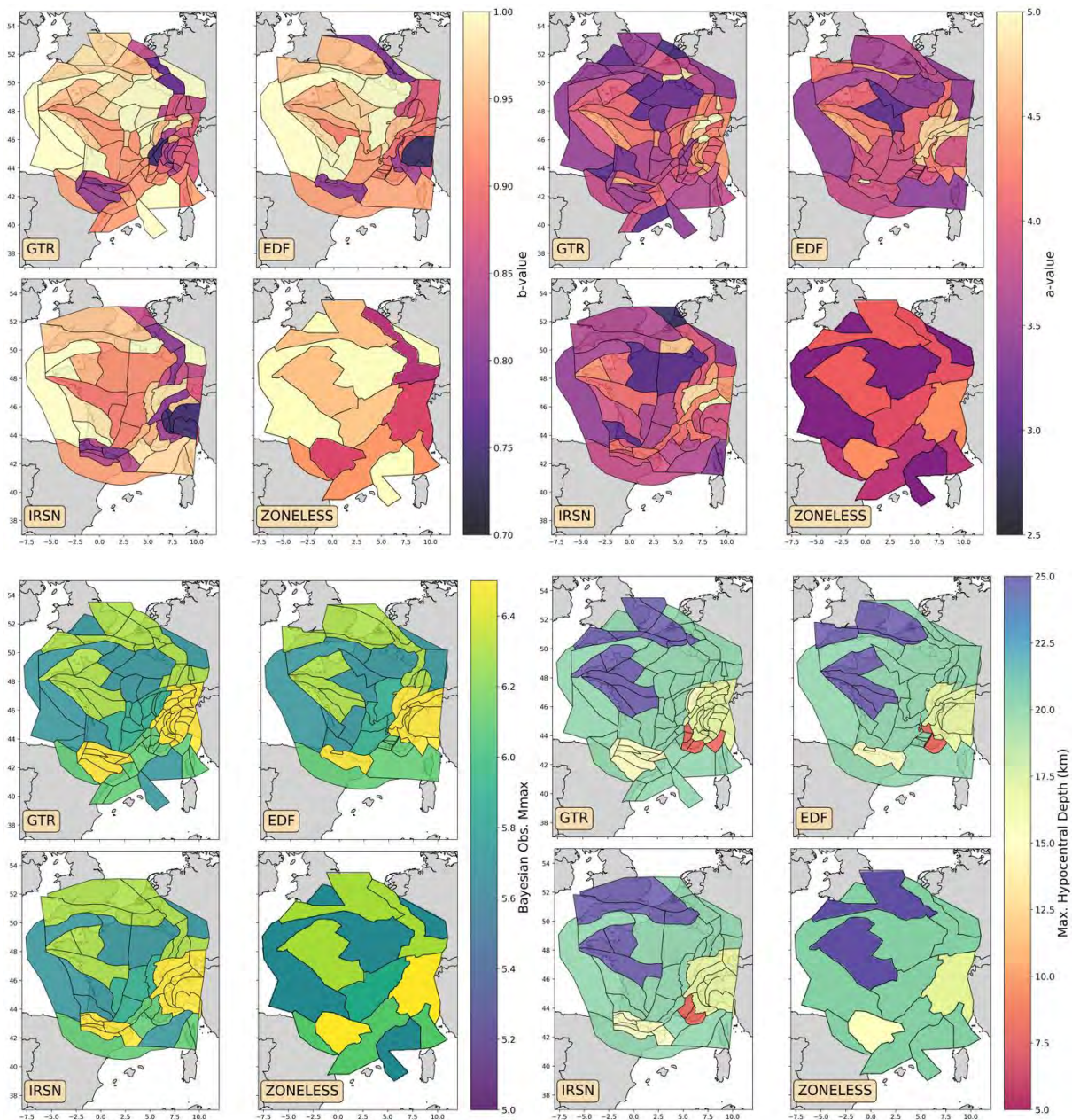


Figure 3: Comparison of the spatial variation of source model properties for the GTR, EDF, IRSN and ZONELESS (DOMAINES) models in term of b value (top left), a (top right), Bayesian observed M_{MAX} (bottom left) and maximum hypocentral depth (bottom right).

Depth Distribution

The modelled hypocentral depth distributions for each zone are determined from the statistics of observed hypocentral depths within each of the primary domaines for earthquakes with magnitudes $M \geq 4.5$. No specific trend in hypocentral depth with magnitude is found; hence distributions are held fixed for each area source. To characterise the probability distribution (in this case a discrete probability mass function) of hypocentral depths, a triangular distribution is assumed with lower and upper limits *guided* by the 5th and 95th percentiles of the observed depths in each domaine.

distributions of the finite-fault properties, which can influence the results in difference ways depending on the GMM used.

Author(s) – Implementation of the Drouet et al. (2020) PSHA for France into OpenQuake: Comparisons and Modelling Issues- SIGMA2-2021-D5-085

The limits of the hypocentral distribution are not necessarily the same as the seismogenic depth limits, which describe the range of depths in the crust to which a rupture may be allowed to propagate (rather than the range in which the ruptures can nucleate). The seismogenic thickness is constrained to the range 0.1 km depth to 30 km depth for all sources. The significance of the seismogenic thickness will become clearer as we consider OpenQuake's modelling of virtual faults

Although the GEOTER (2017) report does provide detail on the hypocentral depth distributions used for each source in the GTR, IRSN and EDF models, the assignment of depths for the ZONELESS sources is not well described. The authors were consulted for clarity and have indicated that, "... for depth, each zone may have specific parameters which can differ from the large scale domain values. Another thing to point out is that in the zoneless case depth is fixed to a single value" (Stéphane Drouet, *personal communication*). The files for the ZONELESS model indicated a fixed hypocentral depth of 10.05 km, while original files define a depth distribution specific to the large-scale zones.

Deformation Model (Mechanism) Distribution

The distribution of mechanisms is defined by Drouet et al. (2020) as a probability of a given style of faulting and, for each style-of-faulting, a corresponding range of strike and dip values. For the GTR model the probabilities for each style-of-faulting are determined based on databases of focal mechanisms and in-situ stress measurements, while for the EDF model these were defined from the interpretation of geological and seismological data, and for IRSN the deformation mechanisms were based on the seismotectonic map of France and the World stress map.

For each area source a probability of strike-slip, reverse and normal faulting is therefore given in the GEOTER (2017) report. Within each style of faulting the maximum and minimum dip ranges are specified, which are fixed to $80 \leq \delta(^{\circ}) \leq 90$ for strike-slip faulting, $50 \leq \delta(^{\circ}) \leq 70$ for normal faulting and $30 \leq \delta(^{\circ}) \leq 60$ for reverse faulting. According to Drouet et al. (2020), the fault strike can vary between $0 \leq \phi(^{\circ}) < 360$ with uniform probability for all sources; however, in digital files provided by EDF a narrower strike range is defined for certain sources in the IRSN model.

Again, for the ZONELESS model the information regarding the distribution of the mechanisms is not provided in the literature accompanying the model. The authors have instead provided us with tables indicating the corresponding style-of-faulting distributions for the DOMAINES shown in Figure 2. These distributions are then applied to all grid cells falling within each specific domain.

2.4 Ground Motion Models

The ground motion models for application to the metropolitan territory of France are identified using a pre-selection process based on the target seismotectonic context, followed by an expert-driven critical appraisal combined with likelihood-based testing of the models against observed ground motion data (e.g., Delavauld et al., 2012). According to Drouet et al. (2020), the pre-selection, review and testing process identifies four GMMs for application:

1. Ameri et al (2017) [A17] – adopting the version without a magnitude-dependent stress parameter and using Joyner-Boore distance (R_{JB}) as the preferred distance metric.
2. Abrahamson et al. (2014) [ASK14]
3. Cauzzi et al. (2015) [C15] – adopting the version with variable reference V_{S30}
4. Drouet & Cotton (2015) [DC15] – adopting the version with nearest distance to rupture (R_{RUP}) as the distance metric, and 10 MPa stress drop for large events.

The range of median ground motion values ground motion models are explored both in absolute space (in the form of trellis plots) and in low-dimensional model space (in the form of Sammons maps). The authors conclude that these models cover the centre body and range of the full distribution of ground motion values from the pre-selected models sufficiently well. All of the above listed models are available

in the OpenQuake software, and we show the OpenQuake-generated trellis plots in terms of attenuation with distance and scaling with period in Figure 4

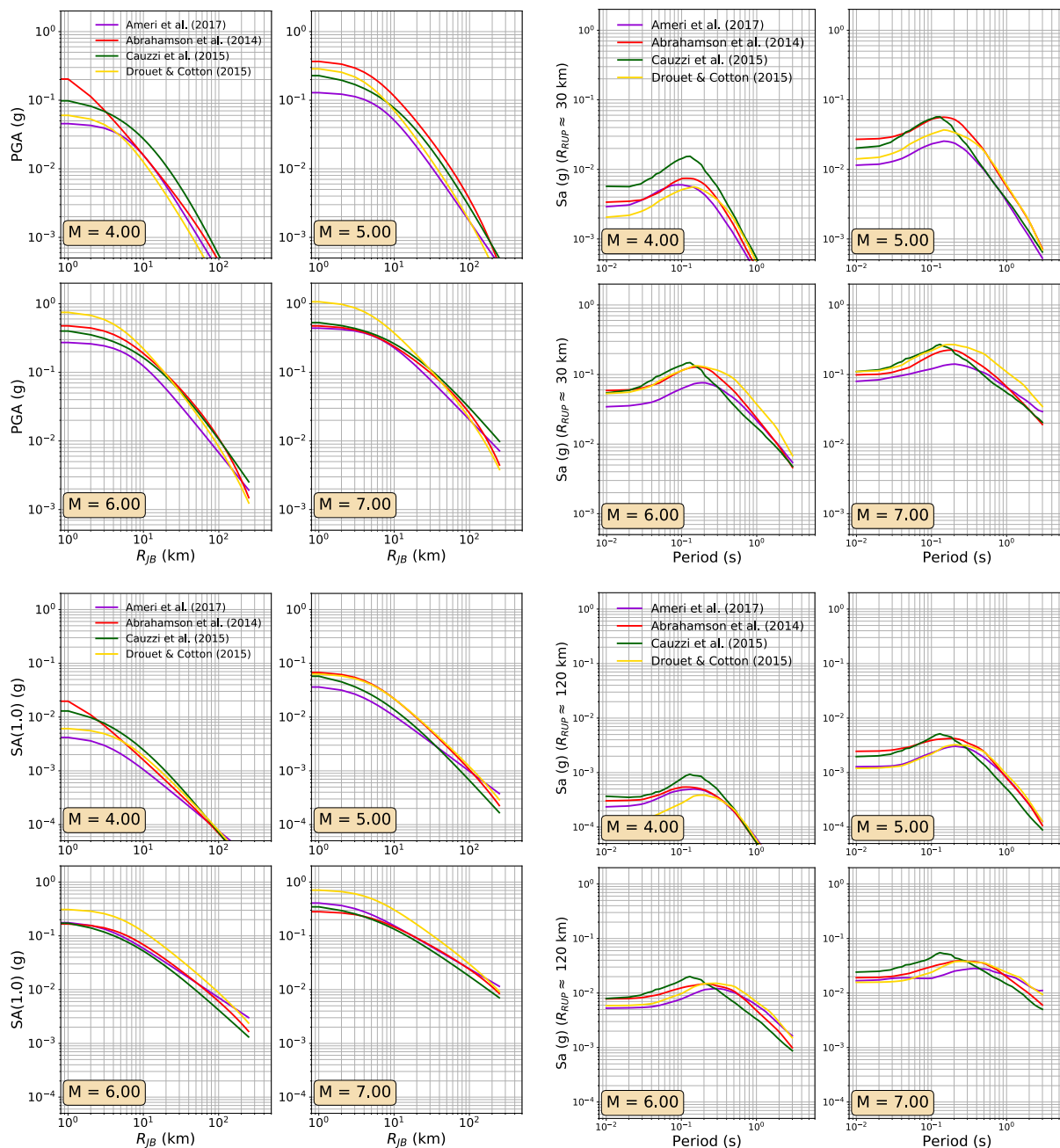


Figure 4: Trellis plots of selected ground motion models for attenuation with distance for PGA and Sa (1.0 s) (left column) and response spectral acceleration with period for a distance of R_{RUP} 30 km and 120 km (right column)

Careful note must be taken, however, that Drouet et al (2020) explain the following, “two GMPEs have been developed focusing on the use of French data: the empirical model by Ameri (2014) and the stochastic model by Drouet and Cotton (2015) (erratum published in Drouet 2017). Note that the model by Ameri (2014), available as a SIGMA project report, has been later improved and published by Ameri et al. (2017). Nevertheless, the two models are similar and the main differences concern the stress parameter model which is not used in the present study, and the sigma model for which the impact in the final results can be considered minor.”. The paper subsequently lists the Ameri et al. (2017) model within the final set of GMPEs, rather than the earlier model of Ameri (2014). We compare both models

in Figure 5 and find that although they are similar, differences between the two emerge at short source-to-site distances and at larger magnitudes. Furthermore, Ameri et al. (2017) adopt a heteroskedastic aleatory uncertainty model, which contrast with the homoscedastic model of Ameri (2014). As a result of these differences, we found that the resulting seismic hazard curves from the two models were not comparable. Upon seeking clarification from the authors, we find that it is in fact the Ameri (2014) version that is used rather than the Ameri et al. (2017) model. The Ameri (2014) GMM was added to the OpenQuake software library as part of this implementation effort.

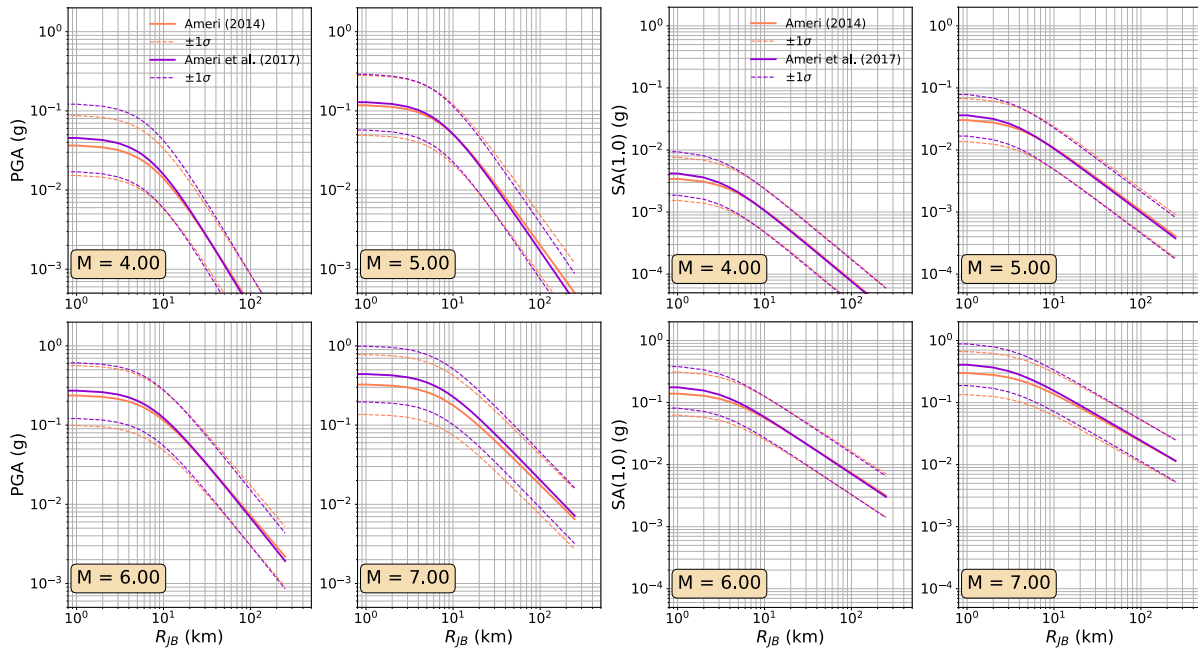


Figure 5: Comparison of the expected attenuation with distance of the Ameri (2014) GMM and Ameri et al. (2017) ground motion models.

2.5 Logic Tree

The complete logic tree for the PSHA model of Drouet et al. (2020) is shown in Figure 6. Conceptually, this logic tree is straightforward. At the first level the weights are split between a uniform area source approach (2/3 weight) and the ZONELESS smoothed seismicity approach (1/3 weight) to modelling the distributed seismicity. Within the area source approach each of the three models (GTR, EDF and IRSN) are assigned equal weights of 1 / 3 each. For the ground motion models, weights are distributed evenly across all four of the models, taking 0.25 weight each.

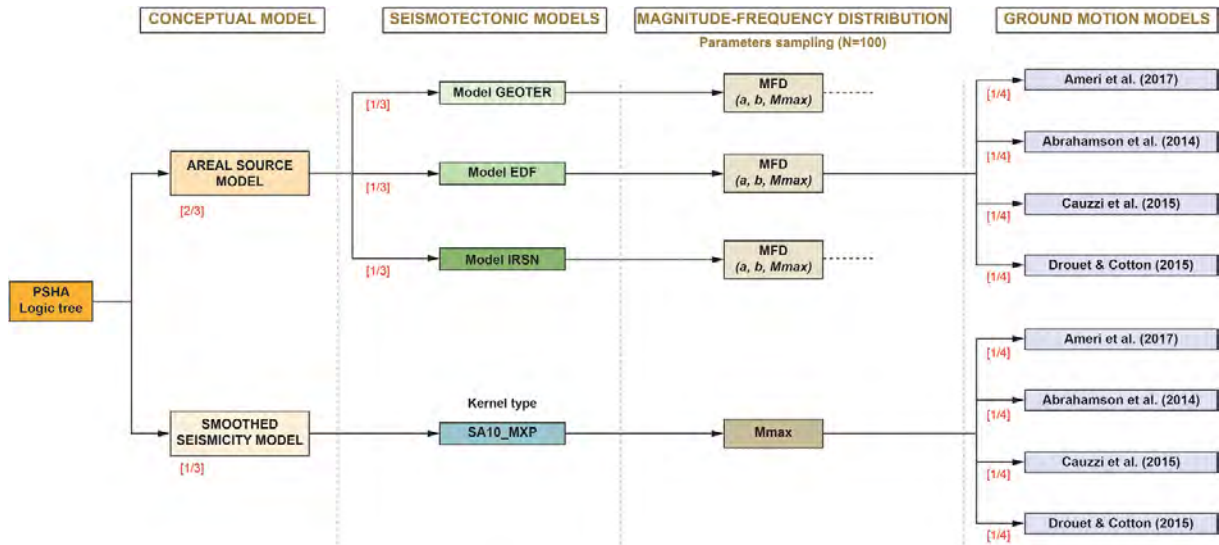


Figure 6: Seismic hazard model logic tree for France (image courtesy of Drouet et al. 2020)

The vast majority of branches in the ground motion logic tree relate to the epistemic uncertainty on the magnitude frequency distribution, which are defined from a sampling process. For the a and b values, $N = 100$ samples are taken from a multivariate Gaussian distribution:

$$a, b = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}\left(\begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \rho_{\sigma_{ab}} \sigma_a \sigma_b \\ \rho_{\sigma_{ab}} \sigma_a \sigma_b & \sigma_b^2 \end{bmatrix}\right) \quad (2)$$

with the a and b values sampled independently for each uniform area source. Similarly, $N = 100$ samples of M_{max} are taken from the posterior distribution of the Bayesian M_{max} approach. The Gutenberg-Richter recurrence parameters and the M_{max} values are sampled independently from one another. For the zoneless models the publications and data files provided indicate that the a and b values are fixed for each cell and only the M_{max} values are sampled for each of the DOMAINS (and assigned to all gridded seismicity sources within the domain). This description is was subsequently found to be incorrect and is discussed in futher detail in section 5 of this report. Altogether with the $N = 100$ samples the resulting logic tree contains 1600 branches, which in the resulting calculation are fully enumerated and not themselves subsequently sampled.

Understanding the MFD Parameter Sampling

The description of the sampling process for the parameters a , b and M_{max} is given only a cursory treatment in the paper of Drouet et al. (2020): “epistemic uncertainty on the seismic activity parameters and maximum magnitude is also propagated using Monte-Carlo sampling (100 samples). Gaussian distributions truncated at 3 sigmas are used for the seismic activity parameters, taking into account the correlation between a and b , and the M_{max} distributions computed with the Bayesian method are used”. As the two distributions (a, b) and M_{max} are sampled independently we treat the sampling process for each of these in turn.

Gutenberg-Richter a and b values

For the parameters of the truncated Gutenberg-Richter distribution we have a fully defined multivariate Gaussian distribution described in equation 2. Sampling from such a distribution is normally done via $X = \boldsymbol{\mu} + \mathbf{A}\mathbf{z}$, where X is the array of N samples in D dimensions, $\boldsymbol{\mu}$ the vector of D mean values, \mathbf{z} a set of N independent samples from a standard normal distribution, and \mathbf{A} the real matrix defined by $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$ (obtained from Cholesky factorisation of the positive-definite covariance matrix $\boldsymbol{\Sigma}$). Introduction of constraints, such as the $\pm 3\sigma$ truncation applied here, add complexity to this problem that may require

use of more demanding Markov-Chain Monte Carlo procedures – though when the number of samples required is small, a simpler (though slightly less accurate) method may be to simply reject solutions outside of the truncation range and continue re-sampling until the desired N samples are obtained.

Fortunately, authors of the model provided their sets of samples for (a, b) that are used in the original calculation. This gives us the opportunity to understand and verify that the samples themselves represent the underlying distribution *to a sufficient degree*. Figure 7 compares the sampled distribution of (a, b) for a given source (GTR model, zone “ACC”) as provided by the authors against those generated from 100,000 samples of a multivariate normal distribution (without truncation). The author-provided samples generally fall well within or on the $\pm 3\sigma$ error ellipse in this particular case. What can be seen clearly here, however, is that a stratified sampling technique has been applied in which b samples are fixed to a discrete set of evenly spaced values while the a value is sampled conditionally on the fixed b . Consultation with the authors provided the following explanation: “For each area source, $N \varepsilon_b$ values are sampled from the normal distribution and, for each value, the conditional value of ε_a is calculated based on the correlation coefficient and its uncertainty. ... ε_b values were selected to discretize the normal distribution (according to the number of standard deviations), then ε_a is sampled based on the correlation and the error on the correlation. This is why the ε_a seems more freely sampled. If the error on the correlation is small, then this effect would be largely reduced.” (Gabriele Ameri, personal communication).

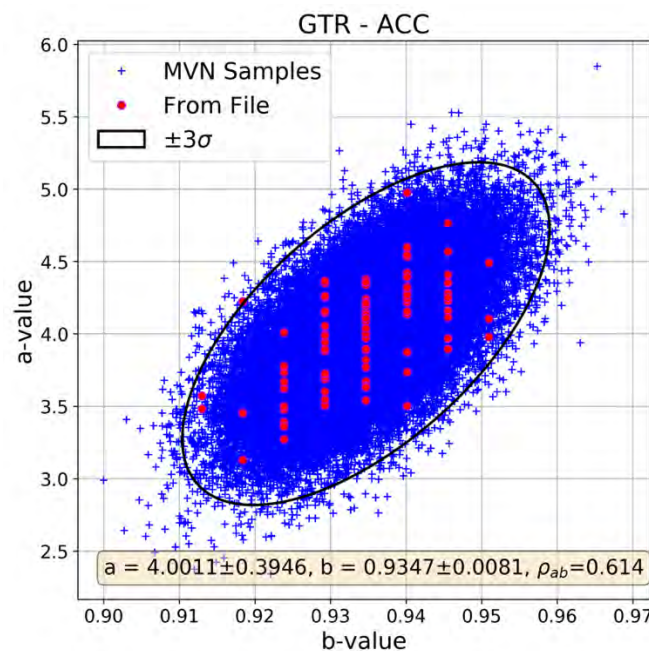


Figure 7: Comparison of samples of the Gutenberg & Richter distribution parameters for the Zone ACC from the model GTR. Red dots indicate the original samples of the authors, blue crosses the samples from an untruncated multivariate Gaussian distribution, and the black ellipse indicated the $\pm 3\sigma$ error ellipse

The details on the sampling strategy provided by the authors do give some insight into the process but are not sufficient to reproduce the steps of the sampling algorithm (not without an independent verification process). What we can do, however, is to verify that for each zone the sampled values of (a, b) are statistically consistent with the underlying distributions. For this purpose, for each source we apply a k-sample Multivariate Analysis of Variance (MANOVA) to determine whether the author-provided (a, b) sample values reject the null hypothesis that they are not different from the assumed multivariate Gaussian distribution at the $p = 0.05$ significance level. Running MANOVA for all of the sources in each of the three area source models, four particular sources were found to reject the null hypothesis (GTR SMC, GTR PRO, GTR ZNA, IRSN 5005). Visual inspection of the resulting distributions found that of

the four sources GTR SMC and IRSN 5005 were found to be broadly consistent with the underlying model. In these two cases the rejection might be due to a slight under sampling of the tails of the distribution; however, it is also statistically likely that from the couple of hundred sources considered, at least one or two may reject the null hypothesis at the requested significance level by chance. Greater inconsistencies can be seen for the GTR PRO and GTR ZNA sources, and while we do not investigate the cause of these further, we do note that these differences would likely yield significant differences in seismic hazard if one were to attempt to generate their own set of samples from the underlying distributions.

M_{max}

For the M_{max} distribution a similar stratified sampling strategy can be seen in the samples provided by the original authors. Figure 8 compares the posterior density functions of M_{MAX} for domaine D3 (South East) with the samples for this same zone provided by the authors. As a reference, we also show a density distribution taken from 100,000 samples of the posterior density function generated using inverse transform sampling. In the set of 100 M_{MAX} samples provided by the authors we observed only three values (M 6.267, M 6.6 and M 6.933) each sampled multiple times. As with the a and b value sampling, the strategy for drawing the M_{MAX} samples is not described in the documentation of the model.

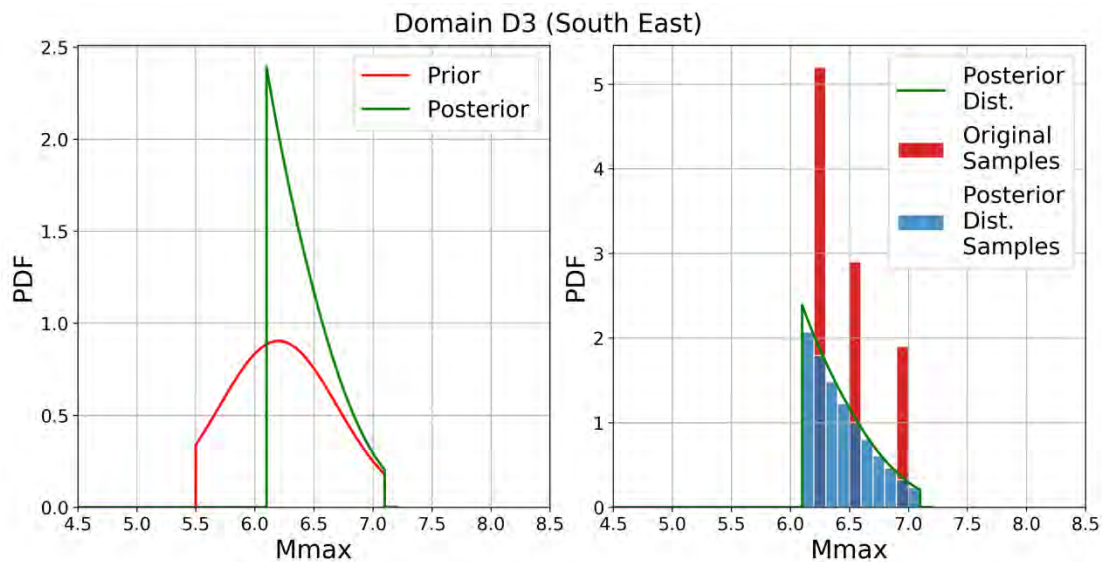


Figure 8: Truncated prior and resulting posterior M_{MAX} distribution for Domaine D3 (left), and comparison (right) of the sample M_{MAX} values for this same region provided by the authors (red bars) against the same posterior distribution and a density distribution from 100,000 samples of the posterior M_{MAX} generated using inverse transform sampling.

2.6 Data files

As the purpose of the work undertaken here is a complete implementation of the Drouet et al. (2020) PSHA model into a new software, comparisons must be made between the outputs of the new software and those of the original files. Though the model is a national model, deeper exploration of the comparisons is better undertaken at a site-specific level, where the entire suite of seismic hazard curves can be considered in more detail rather than the seismic hazard maps. Comprehensive suites of site-specific (in location only, not local soil properties) hazard results are available for 5 locations: 1) Brest (-4.5°E, 48.4°N), 2) Lourdes (0.0°E, 43.1°N), 3) Marseille (5.4°E, 43.3°N), 4) Nice (7.3°E, 43.7°N) and 5) Grenoble (5.7°E, 45.2°N). These results contain the seismic hazard curves for all branches of the model for a range of spectral periods between S_a (0.01 s) and S_a (3.0 s). In addition to these site-specific

results, mean and quantile seismic hazard curves have been provided for the complete set of 6,836 target sites used in the development of the seismic hazard maps. For these target sites the mean and quantile curves are defined for only three intensity measure types ($S_a(0.01\text{ s})$, $S_a(0.20\text{ s})$ and $S_a(1.00\text{ s})$) and for the 5th, 16th, 50th, 84th and 95th percentiles.

For the development of the seismic source models, we are provided with tables of the distributions of the earthquake recurrence models, the hypocentral depths and style-of-faulting/rupture orientation values, all of which were available in digital format and are available via the GEOTER (2017) report or the electronic supplement to Drouet et al. (2020). In addition, digital files of the source geometries have been provided directly from the authors, in addition to the output tables of the smoothed seismicity algorithm and the values of the random samples of a , b and M_{MAX} shown in this chapter.

3. OpenQuake Inputs: Source Model, Ground Motion Model and Calculation Configuration

In this section we will give an overview of OpenQuake and its inputs specifically as they relate to the PSHA model of Drouet et al. (2020) for France. As a sophisticated and adaptable software, OpenQuake can support the characterisation of many types of active faulting, including those with complex time-dependent recurrence models, adaptable scaled backbone ground motion model logic trees, local site amplification etc. These features, and many others, are not considered within the current PSHA model being explored here, so for readers interested to know and understand how they work we refer them to the OpenQuake publications including the paper of Pagani et al. (2014a) and the OpenQuake book, user manual and online help (available from <https://github.com/gem/oq-engine/>). The specific focus of the current section will be to understand how OpenQuake characterises the two types of distributed seismicity sources used by Drouet et al. (2020) (i.e., uniform area sources and gridded seismicity), which parameters can be configured, how the PSHA calculation is performed and the outputs it produces. In each of the steps we explain how the information found in the model publications and data files has been interpreted and implemented within the OpenQuake framework.

Though the primary objective of the work is the implementation of the model as close to that of the original authors, we also wish to undertake further exploration and sensitivity analysis in future studies, to understand better the resulting seismicity distributions and to use the model as a basis for PSHA testing. We have therefore developed a small Python module to construct the source models in such a manner that the user can easily adjust configurable properties. This is complemented with an additional Python module for post-processing the results of the complete logic tree analysis. Further details on the implementation and usage of these modules can be found in material supporting their public release.

3.1 General Overview of OpenQuake

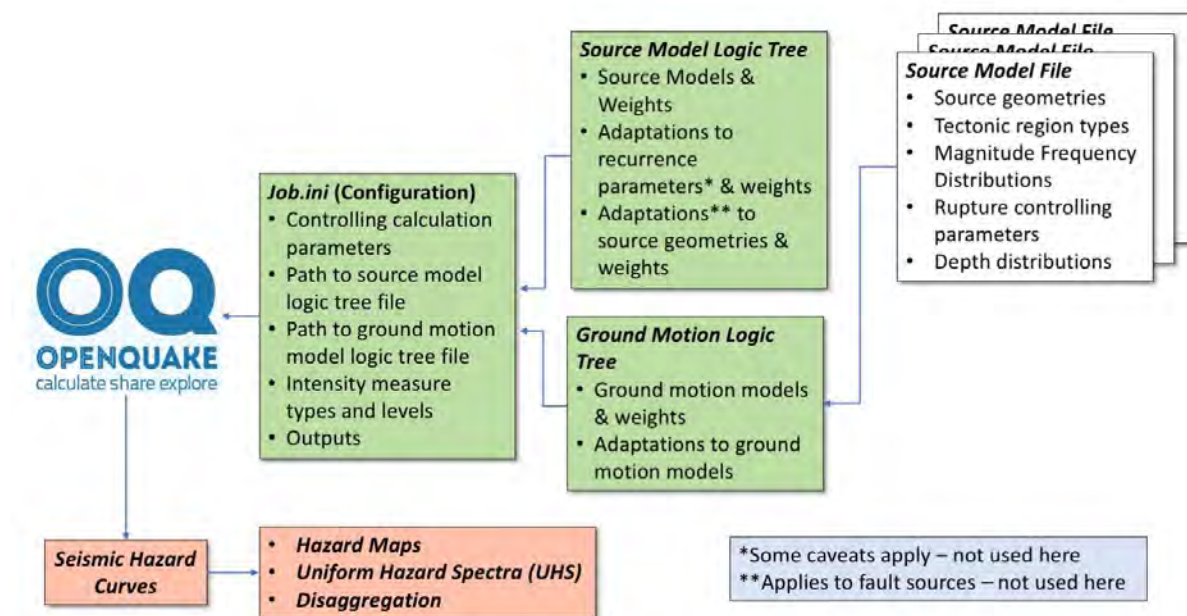


Figure 9: Schematic of data inputs, model outputs and information flows in the OpenQuake-engine PSHA calculation software.

The OpenQuake calculation engine requires several key inputs in order to define the model, and the general flow of the inputs and outputs is illustrated in Figure 9. The calculation is controlled by the

configuration (or *job.ini*) file, which indicates the type of calculation to be undertaken (e.g., “classical” PSHA, “event-based” PSHA, scenario etc.) and, depending on the calculation it may require the user to set configuration properties under the following groups:

geometry of the output target sites, which can be a single site (longitude, latitude), a list sites, a polygon and accompanying discretisation properties (rendering a set of evenly spaced sites within a 2D polygon on the Earth’s surface) or arbitrary sets of points in an external file.

logic tree execution, which can be full path enumeration (i.e., calculations are run for all branches) or sampling, i.e., logic tree branches are sampled according to their weight and only these sampled branches are run (recent options can configure the sampling strategy for greater calculation efficiency)

erf (earthquake rupture forecast) properties that can control how ruptures are built from the source model. These depend heavily on the contents of the source model, but relevant here are the *area source discretisation* spacing (the spacing between points used to characterise sources within a uniform area – see section 3.3 for more detail) and the *width of the mfd bin*, which controls the magnitude bin width for discretising the magnitude frequency distribution.

site parameters control the properties of the local site conditions for the target calculations. For complex site-specific analysis, a full site amplification model can be characterised, but for the present implementation a fixed set of site properties is adopted of all sites. These properties include the V_{S30} of the site (in m/s), a Boolean parameter indicating whether this V_{S30} refers to a “measured” property or whether it is inferred from a proxy, and parameters indicating the depth to the basin: $Z_{1.0}$ (depth *in m* to the 1 km/s shearwave velocity layer) and $Z_{2.5}$ (depth *in km* to the 2.5 km/s shearwave velocity layer).

calculation inputs, which include the path to the *source model logic tree file*, the path to the *ground motion logic tree file*, the investigation time (i.e., the time period to which the resulting probabilities of exceedance will refer) and truncations to be applied (e.g., no. of standard deviations for the sigma of the GMPE, maximum source-to-site distance). Also included here are the *intensity measure types and level* in which the user defines the set of *intensity measure types* for which they wish to calculate hazard (e.g., PGA, SA (0.1 s), SA (0.2 s), ... etc.) and for each intensity measure type the set of levels for which probabilities of exceedance will be calculated (e.g., 0.001 g, 0.002 g, 0.01 g, ... 1.0 g, 2.0 g). *For accelerations OpenQuake uses fractions of g as the unit, for peak ground velocity (PGV) and Peak Ground Displacement (PGD) it uses instead cm/s and cm respectively.*

output defines the types of results desired from the calculation, e.g., hazard maps, uniform hazard spectra etc., as well as whether to define hazard results for the mean and quantiles of the logic tree. If exporting results to file, the path to this file must also be defined here.

Though the configuration file controls the calculation, the actual hazard model inputs themselves are defined via the two logic tree files: the *source model logic tree file* and the *ground motion logic tree file*. Within these two files the user defines the information that helps to construct the complete logic tree and controls the choice of models or model parameters in the calculation, along with their respective weights. A logic tree therefore defines multiple *branch sets*, each branch set consisting of one or more *branches* representing alternative models or parameters for a particular element of the model. A branch refers to a particular model (or model parameter) choice and must be accompanied by its corresponding weight, such that all weights in a particular *branch set* sum to 1.0.

The *ground motion logic tree file* defines the set of ground motion models to be used for each particular *tectonic region type* in the model. Traditionally *tectonic region type* is usually taken to mean one of the few “standard” tectonic environments for which ground motion models are defined (e.g., active shallow crust, stable craton, subduction interface, subduction in-slab, etc.). In practice, however, the *tectonic region type* is simply a mapping that defines which GMMs should be applied to which sources in the source model; thus, every source in the source model must have the *tectonic region type* as an attribute and every tectonic region type must have a corresponding ground motion model or, more commonly,

set of ground motion models and weights defined in the ground motion logic tree file. The actual tectonic region types can be freely defined in a model, but they must be consistent between the sources and the ground motion model logic tree file.

The *source model logic tree file* will indicate to the OpenQuake-engine which source models (in their respective *source model files*) are to be used and their corresponding weights. For more complex logic tree configurations, the files can also define adjustments to specific sources or sets of sources within a particular model. These source or source group specific adjustments can include increases or decreases on the recurrence parameters, or alternative source geometries. However, as we shall explore more in due course, these can rapidly increase complexity in terms of the number of branches being considered, particularly when uncertainties on each source are treated independently.

The source model file should describe a comprehensive set of seismogenic sources, each associated with a unique ID, *tectonic region type*, geometry, magnitude frequency distribution and rake (or focal mechanism: strike, dip and rake). Depending on the nature of the source being used other information controlling the generation of the ruptures may be needed, and we will explore these in the following subsection. Sources from different tectonic regions can be included in the same file (or joined together from multiple files in the *source model logic tree file*).

The main output of OpenQuake, when running a classical PSHA calculation (e.g., Cornell-McGuire), are the seismic hazard curves. For each site and intensity measure type, OpenQuake returns the *probability of exceedance* (not the rate of exceedance) in the stated investigation time (T) for of each intensity measure level. From these curves, seismic hazard maps, uniform hazard spectra and disaggregation can be determined for specific probabilities of exceedance if required by the user (they are not done so by default). When a comprehensive and non-trivial logic tree (i.e., one containing more than one end-branch) is defined in the input files, OpenQuake also offers the user the possibility to return the mean seismic hazard curves (and associated products) from the logic tree, as well as quantiles (e.g. 16th, 50th and 84th percentile) if requested.

3.2 Point Sources

The most elemental form of seismogenic source supported by OpenQuake is the *PointSource*. A single point source defines the earthquake rupture with respect to a single location on the Earth's surface, specified by the longitude (in decimal degrees, centred on the Greenwich meridian) and latitude (in decimal degrees). Though it is referred to as a point, it is critical to recall that for the vast majority of ground motion models commonly in use, including those selected here, the earthquake rupture must be described in terms of the three-dimensional finite fault whose dimensions should scale with magnitude. To control the attributes of the ruptures generated by a point source, including the dimensions, depth distributions and orientations, further information is needed. The overall schematic of the ruptures that point sources define is shown in Figure 10.

Each point source is associated with a magnitude frequency distribution that describes the rate of occurrence of earthquakes of each magnitude within the range $M_{min} \leq M \leq M_{max}$. As shown in the previous section, Drouet et al. (2020) apply a truncated Gutenberg-Richter model for all sources (both gridded seismicity and uniform area sources), which is parameterised via the a , b and M_{max} value. The minimum magnitude (M_{min}) is fixed to M 4.5 for all sources. To generate the set of ruptures from the source that form the *earthquake rupture forecast*, the truncated Gutenberg-Richter model is discretised into bins of fixed width δM between M_{min} and M_{max} , the annual rate of occurrence (λ_{m_j}) is then determined for the bin and the magnitude *in the centre of the bin* taken as rupture magnitude. Each magnitude bin is considered in turn. If a Poisson model is assumed (as it is in all the cases we are considering here) then the probability of occurrence of the magnitude in the centre of the bin, $P_{m_j}(T)$, is determined by $P_{m_j}(T) = 1.0 - e^{-\lambda_{m_j} \cdot T}$.

To generate the finite rupture for the given magnitude the rupture needs the hypocentre to be defined. The epicentre is taken from the position of the point source, but uncertainty in hypocentral depth is accounted for in the source model input by means of a probability mass function, i.e., a set of depths (in km) and their corresponding probabilities (which must sum to 1.0), e.g., depths $z_h = [5, 10, 15]$, probabilities $P(z_h) = [0.2, 0.6, 0.2]$. The probability of occurrence of the rupture occurring at each specific hypocentre location then becomes $P_{rup_j}(m_j, Z_h, T) = P_{m_j}(T) \cdot P(Z_h)$

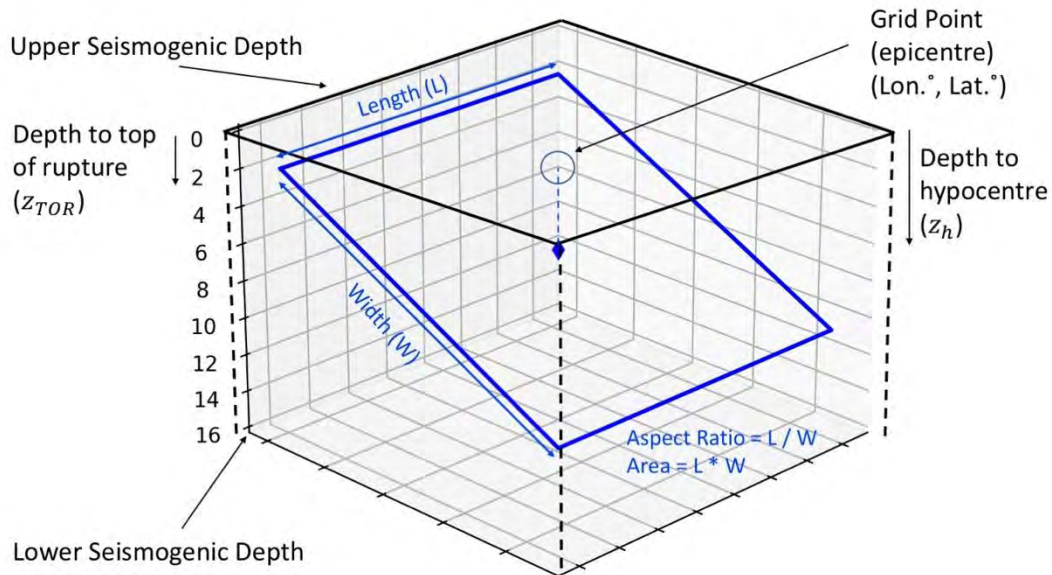


Figure 10: Definition of a “point source” in a point source model and how these correspond to attributes of the rupture required for ground motion models.

Each hypocentre is then used as an anchoring point for constructing the rupture, and *initially* it is assumed that the hypocentre corresponds to the centroid (or mid-point) of the rupture. To build the finite fault rupture the modeller must specify the choice of magnitude-to-area scaling relation and aspect ratio. The magnitude-to-area scaling relations are defined from those support by OpenQuake, and mostly take the form $A (km^2) = 10^{c+dM}$, where c and d are coefficients depending on the model and (sometimes) the style of faulting. Many of these *magnitude scaling relations* are empirical and have an associated variability (σ_A). This uncertainty is not considered in the OpenQuake calculations for distributed seismicity sources such as these. Many magnitude scaling relations in the literature provide models for other rupture parameters such as length, width or displacement. For the rupture generation, however, OpenQuake considers only magnitude-to-area scaling relations. With the rupture area defined, the length and width are controlled by the *rupture aspect ratio*, which must be specified by the user in the source model. No uncertainty on the *rupture aspect ratio* is modelled explicitly; however, the behaviour of the aspect ratio with respect to magnitude is more nuanced.

Finally, with the anchoring location and dimensions of the rupture configured, the final step is to configure the orientation. To do this the user must input into the source model the *nodal plane distribution*. This is input as a probability mass function, with a set of nodal planes (each defined by a fixed strike ϕ , dip θ and rake λ) and their corresponding probabilities $P(\phi, \theta, \lambda)$. Here, aleatory uncertainty is characterised by considering each different rupture orientation, the probability of which becomes $P_{rup_j}(m_j, Z_h, T) = P_{m_j}(T) \cdot P(Z_h) \cdot P(\phi, \theta, \lambda)$.

The rupture shown in Figure 10 contains two additional sources of information: the upper seismogenic depth and the lower seismogenic depth. These two properties are properties of the source rather than the rupture, and they define constraining surfaces that prevent the rupture from propagating to depths

that would be considered by the modeller too shallow or too deep for brittle failure of the crust. These two surfaces must be defined in the point source configuration, and they influence how the rupture dimensions scale at large magnitudes. When the top and bottom depths of the ruptures (Z_{TOR} and Z_{BOR} respectively) are able to be fully contained within the seismogenic thickness, then the initial rupture aspect ratio is retained, and the hypocentre is located in the centroid of the rupture plane. For larger magnitudes, if vertical extent of the rupture width is less than the seismogenic thickness ($Z_{BOR} - Z_{TOR}$) but the rupture itself would exceed the upper or lower seismogenic depths, then the rupture is translated down- or up-dip until its top or bottom edges match the upper or lower seismogenic depth respectively. At this point the absolute location of the hypocentre (longitude, latitude and depth) remains fixed, but its relative position within the rupture plane is translated up-dip or down-dip. If the rupture is so large that its vertical extent is greater than the seismogenic thickness then the *aspect ratio assumption is broken*, and the rupture is re-scaled so that its width is constrained by the seismogenic thickness, but its length is increased in order to conserve the rupture area; hence the actual aspect ratio in this case will be greater than the initially set aspect ratio.

The two fundamental assumptions controlling large magnitude rupture behaviour, i.e., ruptures are constrained by the seismogenic thickness and for larger magnitudes the rupture area is preserved at the expense of aspect ratio, are specific characteristics of OpenQuake and can differ from one PSHA software to another. This behaviour is consistent with observations of rupture properties, as the breaks in rupture length, width and aspect ratio scaling for larger magnitude ruptures are manifestations of a rupture propagation process constrained by the seismogenic thickness of the crust (Stafford, 2014). This behaviour may, in some locations, account for differences between different PSHA software when calculation seismic hazard at a given site with the same input models. The differences emerge as the distributions of source-to-site distance (and therefore median ground motion) may be slightly shifted toward either longer or shorter distances, or may simply be more variable, depending on the software. As these differences tend to affect larger magnitudes to a greater extent than smaller magnitudes, the differences may manifest in the curves more at the lower annual probabilities of exceedance.

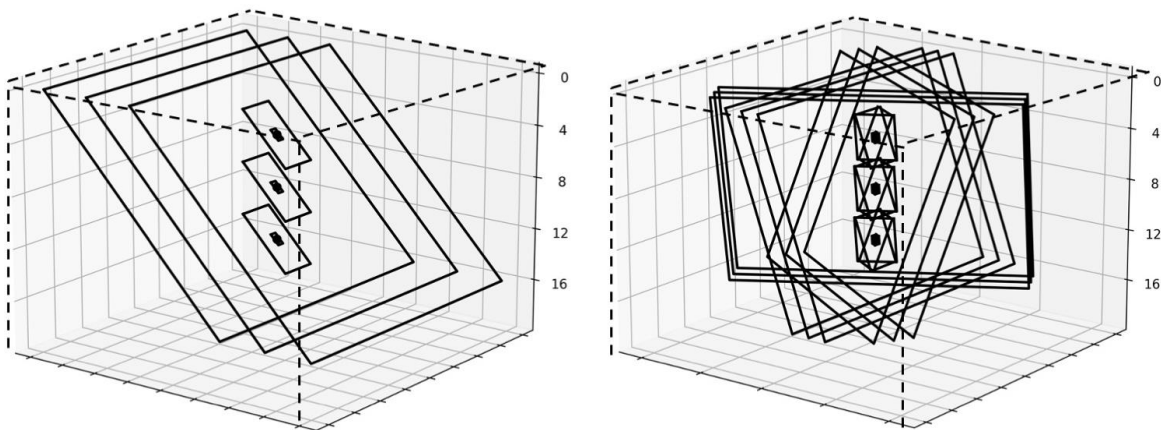


Figure 11: Illustration of the construction and position of virtual faults within a point or area source for i) multiple hypocentral depths and a single rupture plane orientation (left) and ii) multiple hypocentral depths and multiple rupture plane orientations (right)

Figure 11 illustrates how the ruptures for a point source model may be generated in the case of 1) three depth levels and a single rupture orientation, and 2) three depth levels and three rupture orientations. In that idealised example the rupture dimensions are scaling with magnitude for three different magnitudes: 4.0, 5.5, 7.0. It should become obvious that with each additional hypocentral depth layer and nodal plane orientation considered, the total number of ruptures generated increases quadratically. The simple point source with three magnitudes and three depths illustrated in Figure 11 has a total of 9 ruptures in the first case and 27 ruptures in the second case. In the seismic hazard calculation, the

probability of exceedance of ground motion requires the consideration of every rupture j from every source, i , in order to determine the total probability of ground motion at a site exceeding the specified value. To do so, for each rupture we need to calculate the rupture distances to the site(s) and the expected ground motion and its aleatory variability. The two properties *hypocentral depth distribution* and *nodal plane distribution* are referred to as *quadratic* parameters, as increasing the number of samples (or values) in the distribution will increase quadratically the number of computations required in the calculation. A careful balance must therefore be struck between the aim to represent as much of the uncertainty as is feasible by sampling well the distribution of orientations and depths where few constraints exist, while at the same time trying to avoid making the calculations unfeasibly large to run by considering massive numbers of ruptures.

3.3 Uniform Area Sources

The general definition of a uniform area source was outlined in section 2; the key characteristic being that within the area source every location is considered equally probable as an epicentre of a future earthquake. For this area a single magnitude frequency distribution is defined, as too are distributions of hypocentral depth and rupture orientation. As it is a distributed seismicity source, no specific fault surface is identified (though some models may choose to use area sources that align closely with the spatial extents of known faults or fault systems). Though the source describes a spatial extent, in practice the same considerations apply that we need to generate finite ruptures in order to characterise the seismic source for the ground motion model and its scaling with magnitude. The geometric shapes of area sources can assume a significant complexity, including convexity and (in some cases) potentially even topological discontinuity. It is therefore extremely challenging to define a closed-form solution to describe the distribution of source-to-site distances. In many PSHA software, including OpenQuake, the decision is made to discretise the polygons of the area sources into a mesh of evenly spaced and equally probable points. The spacing of the mesh of points into which the area sources are discretised is set by the user in the main configuration file, though this can be over-ridden on a source-by-source basis if necessary. Within the calculation, the area source with a given magnitude frequency distribution is split into a mesh of N_{PTS} point sources, whose corresponding rate of occurrence for any given magnitude is determined via λ_{m_j}/N_{PTS} . From this point on, the generation of the ruptures within the calculation proceeds in the manner described for the point source. The general illustration of how the point sources are distributed across the area source, and how the corresponding ruptures are generated, can be seen in Figure 12

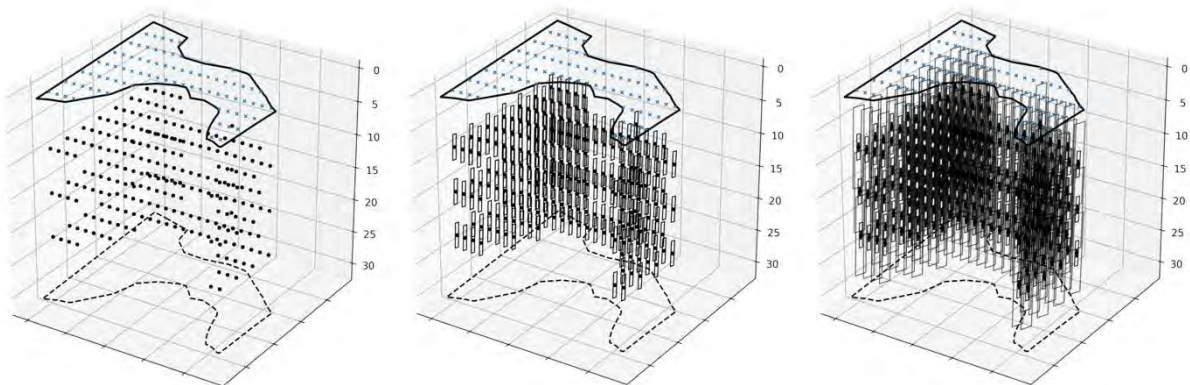


Figure 12: Schematic illustration of rupture generation in a uniform area sources from hypocentre locations (left), initial ruptures for small magnitude events (middle) and the full rupture set (right). The upper and lower seismogenic depths are indicated by the solid and dashed black polygons respectively, epicentres of the grid points are shown in blue crosses and the hypocentres from a three-depth hypocentre distribution shown as black dots.

As the uniform area source can be considered then effectively as a collection of equally probable point sources, the attributes that control the scaling of the rupture dimensions, the rupture orientation and the depth distribution are identical to those of the point sources. The only differences are in the definition of the geometry and in the specification of the spacing of the grid used to discretise the source into points. For an OpenQuake user, the area source geometry is input into the files as a set of (longitude, latitude) nodes. The nodes can be entered in clockwise or semi-clockwise order; however, the polygon must be topologically regular (i.e., vertices cannot cross one-another) and it should not be closed (OpenQuake closes it). Care should be taken with the area source discretisation parameter, as this too is a *quadratic parameter*, meaning that reducing the mesh spacing by a factor of two will result in a four-fold increase in the number of ruptures being considered.

While the generation of the ruptures from the point sources within a uniform area zone follows that described for the point source, it also is the case that for larger ruptures constrained within the seismogenic thickness the aspect ratio assumption is broken, and the ruptures will increase in length. The actual extent of the rupture will therefore be very likely to exceed the boundaries of the polygon, potentially by some distance for the very large ruptures. We describe this as the area source being “leaky”. Not all PSHA software assume or tolerate “leaky” boundaries, and some make this configurable. In OpenQuake, however, all area sources are assumed “leaky” and this is not configurable. There are several reasons as to why these “leaky” boundaries are preferable to impermeable ones, and these relate to the underlying assumptions about rupture scaling that OpenQuake makes. With impermeable bounds, the constraints of the boundary impose magnitude-dependent change in the spatial distribution of ruptures, as sites closer to the edges of the zones would not be able to generate larger ruptures without breaking the zone boundary. This would be addressed either explicitly by translating such ruptures so that they fall into the zone, or implicitly by simply re-weighting the rates on the grid points that can accommodate larger ruptures. Either way, “density” of the sources for larger magnitudes would migrate inward toward the centre of the zone meaning that the area is no longer uniform. Furthermore, for the very large ruptures it is possible that no source location or orientation would allow the rupture to fit within the bounds of the zone. In this case either the probability of the rupture would be set to zero, meaning that your geometry would impose an artificial M_{MAX} rather than the seismicity, or the rupture would be constrained within the source but then the scaling between magnitude and area would be broken. Given OpenQuake’s fundamental assumption of preserving magnitude-area scaling, “leaky” boundaries are the most consistent option.

3.4 Multi-Point Sources

Point sources and uniform area sources are some of the most elemental distributed seismicity source typologies supported by OpenQuake, but more recent improvements to the software have allowed a little more flexibility in source characterisation than earlier versions. Of most relevance here is the multi-point source. Traditionally, the point source typology required specification of the seismogenic depths, magnitude frequency distributions, nodal plane distributions and hypocentral depth distributions on a point-by-point basis. This is certainly flexible, but in practice it is rare that *all* of this information differs from one point to the next, and it is more common that groups of points will share many of the same attributes (such as depth distribution, nodal plane distribution, thickness, scaling relation etc.), even if their specific rate or magnitude frequency distribution changes point-by-point. This is the case for the ZONELESS model of Drouet et al. (2020), where much of the information regarding the nodal plane distribution, hypocentral depth distribution and even parts of the magnitude frequency distribution (namely b and M_{max}) are defined according to the DOMAINE. The multi-point source operates as a hybrid between the point source and uniform area source models, allowing the magnitude frequency distribution to vary from point-to-point but assigning common properties for a region that would otherwise be typical of the uniform area source.

3.5 Implementing the Drouet et al. (2020) Source Models in OpenQuake

The source models are constructed based on a mixture of data sets provided by the original authors, information and interpretation of the source models found within the Drouet et al. (2020) and GEOTER (2017) publications. After an initial implementation based on the data files provided and the information publicly available in the literature, the authors were consulted directly to clarify specific points relating to the behaviour of the sources in the original PSHA software. A summary of each of the source attributes required for the respective models, the information to constrain these attributes provided by the authors, and specific points relating to the OpenQuake implementation are shown in Table 1. Clarifications provided by direct communication with the authors are highlighted in blue text.

Table 1: Attributes of the Drouet et al. (2020) source models, their representation and construction in the original PSHA software (based on information in the publications and/or from the authors), and their representation and construction in OpenQuake

Source Attribute	Drouet et al. (2020)	OpenQuake
Uniform Area Sources (GTR, EDF and IRSN)		
<i>Basic Identifiers (ID, Name)</i>	Specified in supplement to paper and GEOTER (2017)	As per Drouet et al. (2020) and GEOTER (2017)
<i>Tectonic Region Type</i>	<i>Unspecified</i>	« Stable Shallow Crust »
<i>Geometry</i>	Polygon (from shapefile)	Polygon (converted from shapefile)
<i>Upper and Lower Seismogenic Depth</i>	Fixed to 0.1 km and 30 km respectively	Fixed to 0.1 km and 30 km respectively
<i>Magnitude Frequency Distribution</i>	a and b values sampled from multivariate Gaussian distribution (see section 2). M_{MAX} sampled from Bayesian posterior distribution. M_{MIN} fixed to 4.5	Adopts <i>exact</i> set of a , b and M_{MAX} values sampled by authors. M_{MIN} fixed to 4.5
<i>Magnitude Scaling Relation</i>	Not defined in publications. <i>Wells & Coppersmith (1994) confirmed by authors</i>	Wells & Coppersmith (1994) – dependent on style-of-faulting
<i>Rupture Aspect Ratio</i>	Not defined in publications. <i>Authors indicate that it is based on length to width ratio implied by Wells & Coppersmith (1994)</i>	Set to 1.0 initially
<i>Nodal Plane Distribution</i>	Style of faulting probabilities and corresponding strike range and dip ranges defined in Tables 2.6 to 2.8 of GEOTER. Strike and dips vary between limits of the range. <i>Authors confirm 4 random strike values and one random dip</i>	For each style-of-faulting with a non-zero probability : 1) define rake as 0° for strike-slip, -90° for normal and 90° for reverse ruptures. 2) Enumerate 4 evenly-spaced strikes between minimum and maximum 3) Single dip taken as mid-point of range
<i>Hypocentral depth distribution</i>	Triangular distribution assumed with $\min(Z_h)$ and $\max(Z_h)$	Three-point triangular distribution with $\min(Z_h)$ and $\max(Z_h)$

	defined as per Tables 2.6 – 2.8 of GEOTER (2017). <i>Number of depths not given in publications, but authors confirm a three-point distribution is used with probabilities 0.2, 0.6 and 0.2 respectively</i>	defined as per Tables 2.6 – 2.8 of GEOTER (2017) (in digital form). Probabilities assigned as 0.2, 0.6 and 0.2 respectively
ZONELESS (Multi-Point Source)		
<i>Basic Identifiers (ID, Name)</i>	Unspecified	Multi-point ID and name from the DOMAINE names
<i>Tectonic Region Type</i>	Unspecified	« Stable Shallow Crust »
<i>Geometry</i>	Grid of 10 km spaced points	Multi-point grouping according to DOMAINE
<i>Upper and Lower Seismogenic Depth</i>	Fixed at 0.1 km and 30 km respectively	Fixed at 0.1 km and 30 km respectively
<i>Magnitude Frequency Distribution</i>	<i>a and b values specified for each 10 km spaced point (b taken from the DOMAINE)</i> M_{MAX} for each point sampled from the posterior distribution of the DOMAINE M_{MIN} fixed at 4.5	<i>a and b values specified for each 10 km spaced point, from input files provided by the authors</i> M_{MAX} values taken from samples provided by the authors. M_{MIN} fixed at 4.5
<i>Magnitude Scaling Relation</i>	Not defined in the publications.	Wells & Coppersmith (1994) – dependent on style-of-faulting
<i>Rupture Aspect Ratio</i>	Not defined in the publications. <i>Authors indicate that it is based on length to width ratio implied by Wells & Coppersmith (1994)</i>	Set to 1.0
<i>Nodal Plane Distribution</i>	Not defined in the publications. <i>An additional file containing the style-of-faulting distribution for the DOMAINES is supplied.</i>	Based on the style-of-faulting for each DOMAINE, for each style-of-faulting with a non-zero probability : 1) define rake as 0° for strike-slip, -90° for normal and 90° for reverse ruptures. 2) Enumerate 4 evenly-spaced strikes between minimum and maximum 3) Single dip as mid-point of range
<i>Hypocentral depth distribution</i>	Not described in the publications.	Depth defined as 3-point triangular distribution according to the DOMAINES.

From Table 1 it can be seen that author input was critical in two aspects, neither of which were discussed in the publications themselves: i) sampling of the parameters of the magnitude frequency distribution,

Author(s) – Implementation of the Drouet et al. (2020) PSHA for France into OpenQuake: Comparisons and Modelling Issues- SIGMA2-2021-D5-085

and ii) properties of the virtual faults and their scaling with magnitude. In the case of the former, we were unable to replicate the sampling strategy from the information provided by the authors. To achieve the best possible agreement between the two implementations, we have taken the exact values of a , b and M_{MAX} resulting from the sampling procedure implemented by the authors. In this sense the 100 samples of the MFD should be thought of a deterministic set of values, and not necessarily one that can be replicated or approximated if using other sampling strategies. In the case of the properties of the virtual faults and their scaling with magnitude, it is arguably more understandable that this information is omitted, as specific details relate to the exact manner in which the virtual faults are computed within the original software.

3.6 Ground Motion Model (and Site) Implementation

In comparison to the source model implementation, the ground motion models themselves are arguably more straightforward. In the ideal case, a starting point for comparison would be to use the GMMs from the two software to generate tables of expected ground motion values and their aleatory variability terms. Time and resources prevented us from undertaking this exercise at the present. We therefore assume that in both PSHA software the GMMs have been implemented correctly *as the GMM authors intended*. For OpenQuake, the GMM implementations are routinely accompanied by a set of verification tables, which are tables of expected ground motion and random effects standard deviations that have been generated from an independent implementation of the software. In the ideal case, this implementation should be provided directly from the authors, though in some cases this is not always possible and other open-source implementations are sought from third parties. More detailed explanation of the OpenQuake GMM implementation and testing process is found in Pagani et al. (2014b).

Assuming the GMM implementations are correct, where there is potential scope for divergence between PSHA codes is in the specific details of the models. In Table 2 we compare components of the functional forms of the four selected models: Ameri (2014) [A14] / Ameri et al. (2017) [A17], Abrahamson et al. (2014) [ASK14], Cauzzi et al. (2015) [C15] and Drouet & Cotton (2015) [DC15]. Between the models there are several different distance metrics being adopted: Joyner-Boore distance (R_{JB}) for A14/A17, and minimum distance to rupture (R_{RUP}) for C15 and DC15, while ASK14 requires multiple distance metrics (R_{RUP} , R_X and R_{y0}) on account of the complex hanging wall term (which we shall return to shortly). The mixture of distance metrics between the models highlights the importance of using virtual faults for distributed seismicity sources rather than empirical conversion models from point-source to finite-fault distance metrics. Empirical relations between the two can introduce inconsistencies between the distance metrics for large ruptures (such as $R_{JB} = 0$ and $R_X < 0$), which may produce unexpected amplifications of the motion. However, this can mean that the control of the virtual faults and the proximity of the anchoring points with respect to the target site for hazard can have a larger influence on the seismic hazard.

In addition to the distance metrics, there are evident differences between the models owing to different site parameterisations (e.g., Eurocode 8 site classes versus V_{S30}), the inclusion of style-of-faulting factors, geometric and anelastic attenuation scaling and more. For most of the models, however, reasonable comparisons can be made given that we consider only a reference V_{S30} of 800 m/s (Eurocode 8 class A) rock. The ASK14 GMM is far more complex in terms of both functional form and parameterisation, however, and this can not only introduce some inconsistencies in the implementations between different PSHA software, it also means more specification of the site properties are needed. For ASK14 we need to define the depth to the 1.0 km/s shearwave velocity layer ($Z_{1.0}$) that corresponds to the V_{S30} 800 m/s site property we are assuming, and we need to specify whether the V_{S30} refers to a *measured* or an *inferred* property (the variability being higher in the case of the latter). In our implementation we determine $Z_{1.0}$ from the empirical relation between V_{S30} and $Z_{1.0}$ assuming a California site model (Equation 18 of Abrahamson et al., 2014):

$$Z_{1.0} \approx \exp\left(-\frac{7.67}{4} \cdot \ln\left(\frac{V_{S30}^4 + 610^4}{1360^4 + 610^4}\right)\right) \quad (3)$$

For a V_{S30} of 800 m/s this gives a $Z_{1.0}$ of 31.07 m. Consultation with Stéphane Drouet confirmed that this same approach was used to define an input $Z_{1.0}$ and that the V_{S30} is assumed to correspond to a *measured* condition.

In addition to the site parameterisation, the hanging wall amplification model of ASK14 GMM also allows room for discrepancies between implementations in PSHA software to arise, albeit in this case such discrepancies are anticipated by the authors. The differences relate to the strike-parallel taper of the hanging wall term $T_5(R_x, R_{y0})$:

$$T_5(R_x, R_{y0}) = \begin{cases} 1 & \text{for } R_{y0} - R_{y1} \leq 0 \\ 1 - \frac{R_{y0} - R_{y1}}{5} & \text{for } 0 < R_{y0} - R_{y1} < 5 \\ 0 & \text{for } R_{y0} - R_{y1} \geq 5 \end{cases} \quad \text{where } R_{y1} = R_x \tan 20$$

Here R_{y0} refers to the distance from the site to a line perpendicular to the ends of the rupture (taking 0 if within the length of the rupture plane. Recognising that this introduces a new distance metric for PSHA software to calculate, Abrahamson et al. (2020) suggest that, if necessary, this term can be replaced with:

$$T_5(R_{JB},) = \begin{cases} 1 & \text{for } R_{JB} = 0 \\ 1 - \frac{R_{JB}}{30} & \text{for } R_{JB} < 30 \\ 0 & \text{for } R_{JB} \geq 30 \end{cases}$$

In OpenQuake it is the $T_5(R_x, R_{y0})$ formulation that is implemented and the authors of Drouet et al. (2020) confirm this too is the case in the original software. We should therefore not expect discrepancies arising from this term.

Table 2 : Comparison of the components of the functional forms of the four selected GMMs

Model Component	Ameri (2014)/Ameri et al. (2017)	Abrahamson et al. (2014)	Cauzzi et al. (2015)	Drouet & Cotton (2015)
Magnitude Scaling $f_M(M)$	Two-segment polynomial with a single hinge magnitude M_h : $\begin{cases} b_1(M - M_h) + b_2(M - M_h)^2 \\ b_3(M - M_h) \end{cases}$	Three-segment polynomial with two hinge-magnitude (see Equation 2 of Abrahamson et al. (2014) for complete description)	Quadratic: $m_1M + m_2M^2$	Quadratic: $b_2(M - 8) + b_3(M - 8)^2$
Geometric Attenuation $f_{R,G}(R, M)$	$[c_1 + c_2M_{ref}] \cdot \log_{10} \left(\frac{\sqrt{R_{JB}^2 + h^2}}{R_{ref}} \right)$	$[a_2 + a_3(M - M_f)] \cdot \ln(R)$ where $R = \sqrt{R_{RUP}^2 + c_{4M}^2}$ and c_{4M} is a magnitude-dependent fictitious depth term. $a_3(M - M_f)$ is dependent on each of the sections in the magnitude scaling polynomial	$(r_1 + r_2M) \log_{10}(R_{RUP} + r_3)$	$(b_4 + b_5M) \cdot \ln \left(\sqrt{R^2 + b_6^2} \right)$ Where $R = R_{RUP}$ in the selected version of the model (different coefficients are given for different distance metrics)
Anelastic Attenuation $f_{R,A}(R)$	None	$a_{17} \cdot R_{RUP}$	None	$b_7 \cdot R$ Where $R = R_{RUP}$
Source Depth Scaling $f_Z(Z_h, Z_{TOR})$	None	$\begin{cases} a_{15} \cdot \frac{Z_{TOR}}{20} & \text{for } Z_{TOR} < 20 \text{ km} \\ a_{15} & \text{for } Z_{TOR} \geq 20 \text{ km} \end{cases}$	None	None
Style-of-faulting f_{SOF}	$f_1 \cdot E_N + f_2 \cdot E_R + f_3 \cdot E_S$ Where E_N , E_R and E_S take the values 1 for normal faulting, reverse faulting or strike-slip faulting respectively, or 0 otherwise	$\begin{cases} a_{xx} & \text{for } M > 5 \\ a_{xx}(M - 4) & \text{for } 4.5 \leq M \leq 5 \\ 0 & \text{for } M < 4.0 \end{cases}$ Where a_{xx} are specific coefficients for the reverse and normal faulting case	$f_N \cdot F_N + f_R \cdot F_R + f_{SS} \cdot F_{SS}$ Where F_N , F_R and F_{SS} take the values 1 for normal faulting, reverse faulting or strike-slip faulting respectively, or 0 otherwise	None
Hanging Wall/Near-Fault Scaling $f_{HW}(\dots)$	None	$a_{13} \cdot T_1(\delta) \cdot T_2(M) \cdot T_3(R_X, W, \delta) \cdot T_4(Z_{TOR}) \cdot T_5(R_X, R_{y0}, OR R_{JB})$ Where $T_1(\dots), T_2(\dots)$ etc. are functions dependent on rupture geometry terms	None	None

		dip (δ), width (W), Z_{TOR} and distance terms R_x and R_{y0}		
Shallow Site Response (Linear) $f_{S,lin}(V_{S30})$	$s_1 \cdot F_A + s_2 \cdot F_B + s_3 \cdot F_C + s_4 \cdot F_D$ Where F_A to F_D take the value 1 for Eurocode 8 site class A, B, C or D, and 0 otherwise	$(a_{10} + bn) \cdot \ln\left(\frac{V_{S30}^*}{V_{lin}}\right)$ Where bn are constants, $V_{S30}^* = \min(V_{S30}, V_1)$ and V_1 and V_{lin} are period-dependent and	$b_V \cdot \log_{10}\left(\frac{V_{S30}}{V_A}\right)$ Where V_A is a period-dependent reference condition	Coefficients fixed for a reference rock of $V_{S30} = 800$ m/s
Shallow Site Response (Nonlinear) $f_{S,nl}(V_{S30}, a_{ref})$	None	$a_{10} \ln\left(\frac{V_{S30}^*}{V_{lin}}\right) - b \cdot \ln(Sa_{1180} + c) + b \cdot \ln\left(Sa_{1180} + c \left(\frac{V_{S30}^*}{V_{lin}}\right)^n\right)$ Where Sa_{1180} is the median spectral acceleration on V_{S30} 1180 m/s hard rock	None	None
Basin Depth Term $f_{Z1.0}(Z_{1.0})$	None	$a_{XX} \ln\left(\frac{Z_{1.0} + 0.01}{Z_{1ref}(V_{S30}) + 0.01}\right)$ Where a_{XX} depends on the V_{S30} and Z_{1ref} is the average Z_1 for a given V_{S30} , which related via an empirical correlation model for California or Japan	None	None
Inter-event Variability τ	Constant τ in Ameri (2014), but magnitude-dependent $\tau(M)$ in Ameri et al. (2017)	Magnitude-, V_{S30} and Sa_{1180} dependent τ (owing to propagation of between event variability in the nonlinear soil term)	Constant τ	Constant τ
Intra-event variability ϕ	Constant ϕ	Magnitude-, V_{S30} and Sa_{1180} dependent ϕ (owing to propagation of within-event variability in the nonlinear soil term). Additionally, the coefficients differ depending on whether the V_{S30} refers to a <i>measured</i> or <i>inferred</i> term	Constant ϕ	Constant ϕ

3.7 Implementing the full Logic Tree

The particular configuration of the logic tree in the Drouet et al. (2020) model presents a challenge for OpenQuake as it is not possible to implement it cleanly using the available options for executing logic trees. The reason for this is that the generation of the a , b and M_{MAX} values is done via a sampling each of the sources independently rather than permuting pre-defined values or adjustments to values for all of the possible sources. To implement the logic tree then, we model the epistemic uncertainty in the MFD parameters by generating 100 source model files per source model (400 altogether), and in each source model the a , b and M_{MAX} values on each source will correspond to $i = 1, 2, \dots, N_{BR}^{MFD}$ of the $N_{BR}^{MFD} = 100$ MFD branches, each realisation receiving a weighting of 0.01. The total set of source model inputs will therefore be a list of 400 individual files, the weights of which in the source model logic tree will be the $0.01 \cdot SRC_{WEIGHT}$, where SRC_{WEIGHT} is the weight of each of the three area source or zoneless source models.

The means of implementing the logic tree using multiple source models that repeat the same geometries and other characteristics but vary the the a , b and M_{MAX} on each model can achieve the objective of ensuring compatibility with the approach adopted by the authors. Unfortunately, however, this incurs a substantial computational cost. This is because by treating each permutation as a separate model OpenQuake will need to rebuild the *earthquake rupture forecast* for each model individually, even if the only property that changes from one model to the next is the rates in the given magnitude bins. This means that a large amount of the computational time will be spent repeatedly generating the same rupture surfaces and then calculating the source-to-site distances and expected ground motions. In Section 5 we will talk more about the computational challenges related to implementing the full logic tree for the entire country and what we have done to try to overcome them.

To retrieve the mean and quantile seismic hazard curves from the logic tree, OpenQuake calculates statistics in terms of the probabilities of exceedance (PoEs) rather than the intensity measure levels. For the mean, this corresponds to the weighted arithmetic mean PoE for each intensity measure level. To determine the quantiles, the PoE for each intensity measure level are ordered from low to high and their corresponding weights cumulatively summed to 1.0. The PoE for quantile Q (in the range $0 \leq Q \leq 1$) is determined by identifying the corresponding PoEs whose cumulative weights are immediately below and above the quantile Q and then interpolating between these values to return the PoE corresponding to Q . Once again, Drouet and colleagues were contacted directly for information as to how the mean and quantiles the seismic hazard are calculated in their in-house software, and they confirm that they understand the process to be the same.

3.6 Calculation Configuration

With the source and ground motion models implemented, the final pieces of information relate to the calculation configuration. Here, only a few parameters are left to set and mostly these control the truncations and discretisations to be applied to the calculation. The first truncation is the maximum rupture-to-site distance (km) to consider in the calculation, R_{MAX} , such that $P(A > a_g | R > R_{MAX}) = 0$. After consulting with the authors, we set this term to 200 km. In OpenQuake R_{MAX} refers to the maximum rupture to site distance (R_{RUP}). Depending on whether this filter is applied to R_{RUP} or R_{JB} there may be some sensitivity in the total rate of the exceedance of ground motion at a site, as deeper ruptures from the point sources close to this distance limit may be filtered out. The impact of this on the hazard at return periods of interest should be minimal, however.

The second truncation parameter that can be controlled is the truncation of the ground motion standard deviation (σ). In some PSHA applications it has been common practice to truncate the Gaussian distributions describing a GMM's aleatory uncertainty (e.g., Strasser et al. 2008). In OpenQuake this term can be set by the user, and if this is done then the truncations are applied symmetrically, i.e., to

both the standard deviations above and below the mean. Initial calculation runs with truncation set $\pm 3\sigma$ showed significant divergence at longer return periods between the OpenQuake hazard curves and those from the original authors. After consulting with the authors again, we find that no truncation was applied, and we therefore do the same.

As the source model comprises mostly uniform area sources, we need to indicate the spacing of the discrete mesh used to convert these into the equivalent set of point sources (see section 3.4). This is set to 10 km for the current application. Given that the sources use a truncated Gutenberg-Richter magnitude frequency distribution, we also need to set the width of the magnitude bins into which this distribution should be discretised in order to define the occurrence rate of ruptures of each magnitude bin. For this we choose a bin width of 0.1 magnitude units. Again, both of these properties are *quadratic* parameters, so a balance is struck between aiming for the highest calculation precision and computational feasibility. The above values are set based on our experience of running calculations at a spatial scale such as this.

The standard configuration of the input file (job.ini) is given as follows :

```
[general]
description = DROUET ET AL 2020, VS30=800 m/s: 5 Cities
calculation_mode = classical
random_seed = 23 # Unused in the calculations

[geometry]
# Brest, Grenoble, Lourdes, Marseille, Nice
sites = -4.49 48.39, 5.7224 45.1715, -0.05 43.1, 5.37 43.2964, 7.2663 43.7034

[logic_tree]
number_of_logic_tree_samples = 0

[erf]
rupture_mesh_spacing = 5 # Unused in the calculations
area_source_discretization = 10
width_of_mfd_bin = 0.1

[site_params]
reference_vs30_type = measured
reference_vs30_value = 800
reference_depth_to_2pt5km_per_sec = 1.0 # Unused in the calculations
reference_depth_to_1pt0km_per_sec = 31.07

[calculation]
source_model_logic_tree_file = path/to/logic_tree_source_model.xml
gsim_logic_tree_file = path/to/logicTree_GMPE.xml
investigation_time = 1.0 # To return annual probabilities of exceedance
intensity_measure_types_and_levels = {
    "PGA": [1.0E-4, 1.57999995E-04, 2.50999990E-04, 3.98000004E-04,
            6.30999974E-04, 1.00000005E-03, 1.58499996E-03, 2.51200004E-03,
            3.98099981E-03, 6.31000008E-03, 9.99999978E-03, 1.58489998E-02,
            2.51189992E-02, 3.98110002E-02, 6.30960017E-02, 0.100000001,
            0.158489004, 0.251188993, 0.398106992, 0.630957007, 1.00000000,
            1.58489299, 2.51188588, 3.98107195],
    "SA(0.01)": [... as PGA],
    "SA(0.02941)": [... as PGA],
    "SA(0.05)": [... as PGA],
    "SA(0.1)": [... as PGA],
    "SA(0.2)": [... as PGA],
    "SA(0.3)": [... as PGA],
```

```
"SA(0.4)": [... as PGA],  
"SA(0.5)": [... as PGA],  
"SA(1.0)": [... as PGA],  
"SA(2.0)": [... as PGA]}  
truncation_level = # Blank indicates that no truncation is applied  
maximum_distance = 200
```

```
[output]  
export_dir = ./path/to/results_folder  
mean_hazard_curves = true  
hazard_maps = false  
individual_curves = true  
quantile_hazard_curves = 0.05, 0.16, 0.5, 0.84, 0.95  
uniform_hazard_spectra = true  
poes = 0.002105, 0.000404. # Annual PoE corresponding to 10 % and 2 % PoE in 50 years
```

Specific calculation settings will change depending on the calculations being performed, e.g., site location, paths to the logic tree files, intensity measure types and levels, required quantiles etc. Digital versions of these files will be made available as part of the general model release.

4. Comparisons and Sensitivities – Branch-by-Branch

To attempt to verify that the OpenQuake implementation of Drouet et al (2020) can provide results that are sufficiently comparable to those of the original authors, we break the process down into two stages. In the first stage, we undertake comparisons branch-by-branch at those target sites for which we have results from all logic tree branches available (Brest, Lourdes, Grenoble, Marseille, Nice). The branch-by-branch comparisons help to identify where systematic discrepancies can be seen, which would later propagate through to the comparisons of the mean and quantiles.

As described in section 2 and 3, the majority of branches in the ground motion logic tree relate to uncertainty on the parameters of the magnitude frequency distribution and that in the current implementation we adopt the exact values that were sampled by the original authors. A complete suite of comparisons for all 1,600 branches would be redundant in this case, as we are most interested in understanding if there are specific combinations of source and/or ground motion models that require specific investigation. We therefore show the comparisons for a single branch of the magnitude frequency distribution (sample #5 from the set) and focus on the differences arising for each of the four ground motion models (A14, ASK14, C15, DC15) and each of the source models (GTR, EDF, IRSN, ZONELESS). The results shown in the comparisons focus on two intensity measure types: S_a (0.01 s) and S_a (1.0s).

To undertake the comparisons, we first aim to define the criteria that would constitute a good level of agreement between the two software. As the complete seismic hazard curve is a composite of the probabilities of exceedance from all of the events within the target distance of the site, one would not necessarily expect to obtain agreement for the entire range of probabilities for which the hazard curve is defined, which can be as low as 10^{-6} or 10^{-7} . The PEER Benchmark tests illustrate that even for relatively simple source model configurations, different PSHA software can diverge at lower annual probabilities of exceedance, even if agreement is otherwise good at the higher annual probabilities. For the comparisons that follow we would consider good agreement to refer to agreement within $\pm 10\%$ of the original curves from Drouet et al. (2020) for annual probabilities greater than or equal to 10^{-4} (approximately 10,000 years return period).

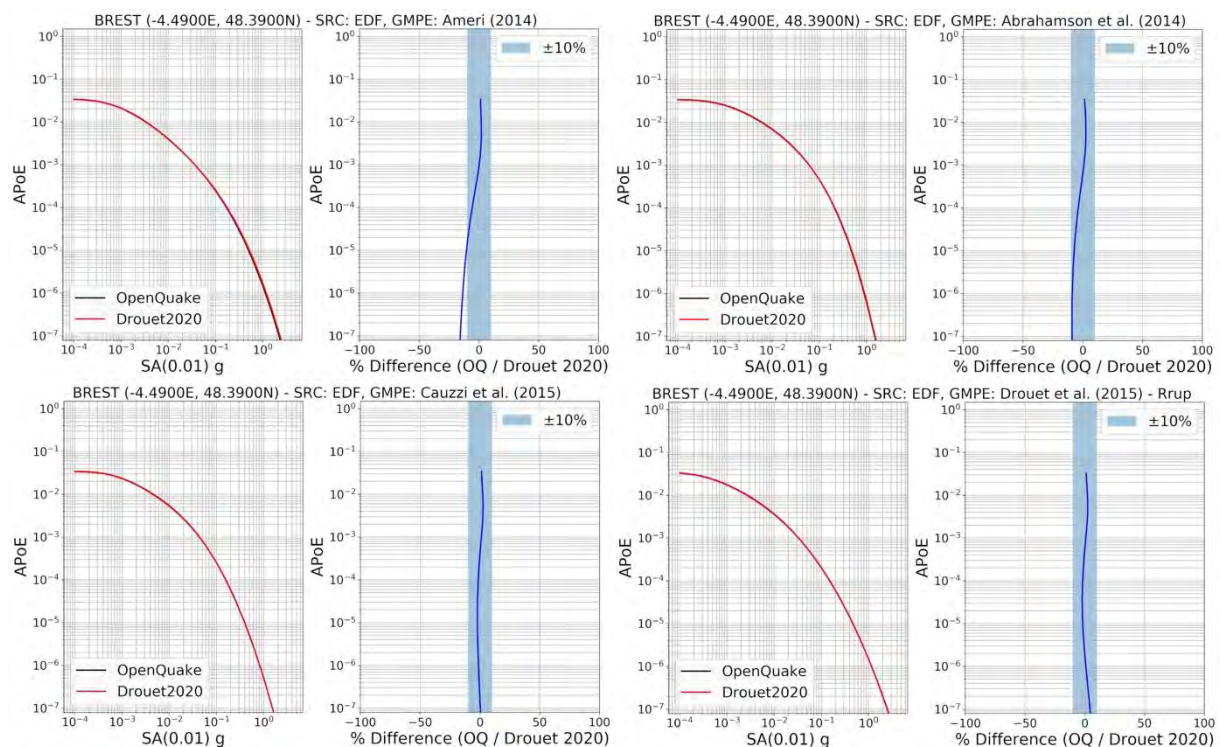
In the comparisons that follow the full probability range of the curves are shown, reaching to APoE as low as 10^{-7} in some cases. We emphasise that a model of the type presented in Drouet et al. (2020) is generally intended for application to return periods relevant for current ordinary constructions according to the national or local design code. This would usually correspond to return periods ranging from 95 years (APoE $\approx 10^{-1}$) to 10,000 years (10^{-4}). Seismic hazard models applied to structures of critical importance or high consequence of failure would normally be adapted to the specific characteristics of the site in question, and with a more comprehensive characterisation of epistemic uncertainty. Only in those cases would agreement to APoE as low as 10^{-6} or 10^{-7} be relevant. Nevertheless, we believe it of interest to the reader to show the entirety of the APoE range of the curves, which illustrate cases where the target agreement can be achieved for the such low APoEs, as well as cases where agreement of $\leq \pm 10\%$ is achieved for the target level here but soon deteriorates for longer return periods, and even some cases where the opposite holds and poorer agreement may be seen in the 10^{-2} to 10^{-4} APoE range only for agreement to improve at lower APoEs.

4.1 Site-by-Site Comparisons

Brest

The comparisons for all four source models are shown in Figures 13a to 13d, with each figure illustrating the hazard curves and the difference curves. All hazard curves are shown for annual probabilities of exceedance (APoEs) between 10^{-7} and 1.0. The key outcome here is that the target criteria ($\pm 10\%$ agreement for APoEs $> 10^{-4}$) is achieved for all of the curves. For most of the model combinations the agreement of $\pm 10\%$ is actually achieved for lower APoEs too, down to 10^{-6} . The most robust agreements can be found for the C15 and DC15 GMMs, while those of A14 and ASK14 have a greater tendency to fall outside of the 10 % range at lower APoEs and for the longer periods. In the cases where the disagreement is greater than $\pm 10\%$, OpenQuake is predicting lower values.

The disagreement of the A14 model at low APoEs and longer spectral periods is perhaps harder to explain than that of ASK14, given that the model has a relatively simple parameterisation and functional form, and the coefficients for the model were taken directly from digital files provided by the authors. Minor differences in the GMM implementation could be an explanatory factor, such as coefficient rounding or the definition of standard gravity g when converting between cm/s/s to fractions of g (OpenQuake uses $g = 9.80665 \text{ m s}^{-2}$ rather than $g = 9.81 \text{ m s}^{-2}$).



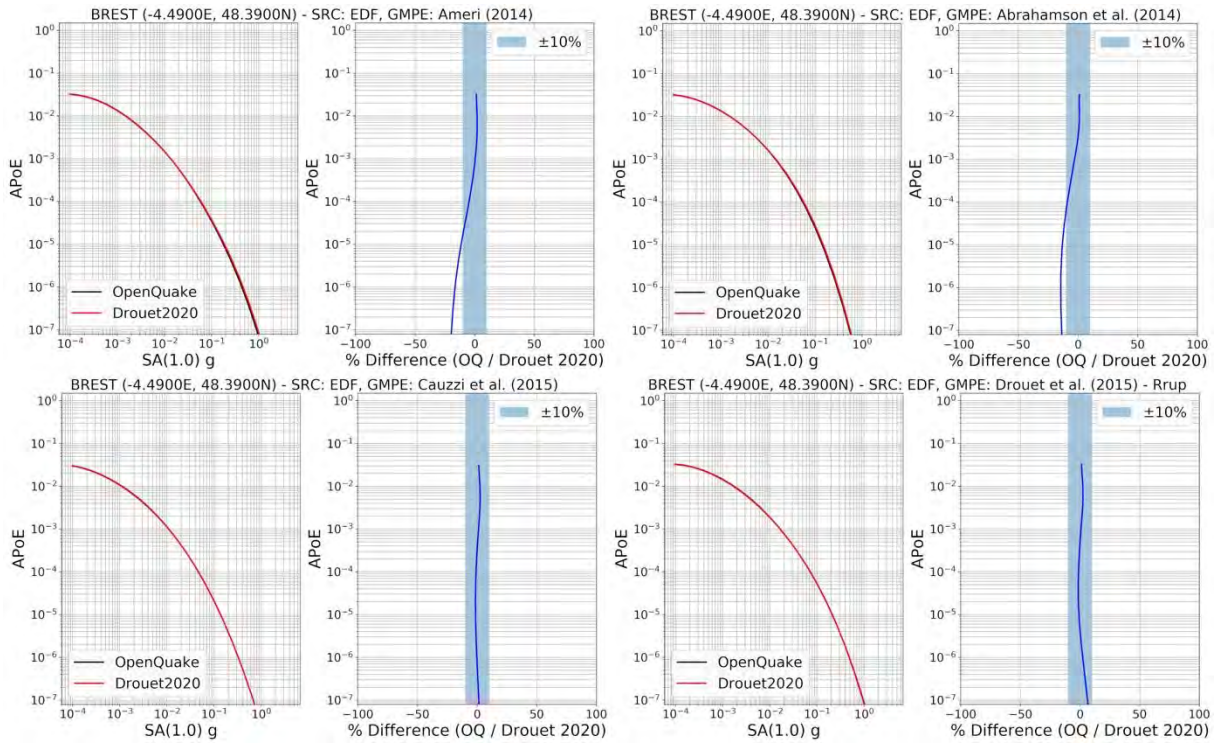
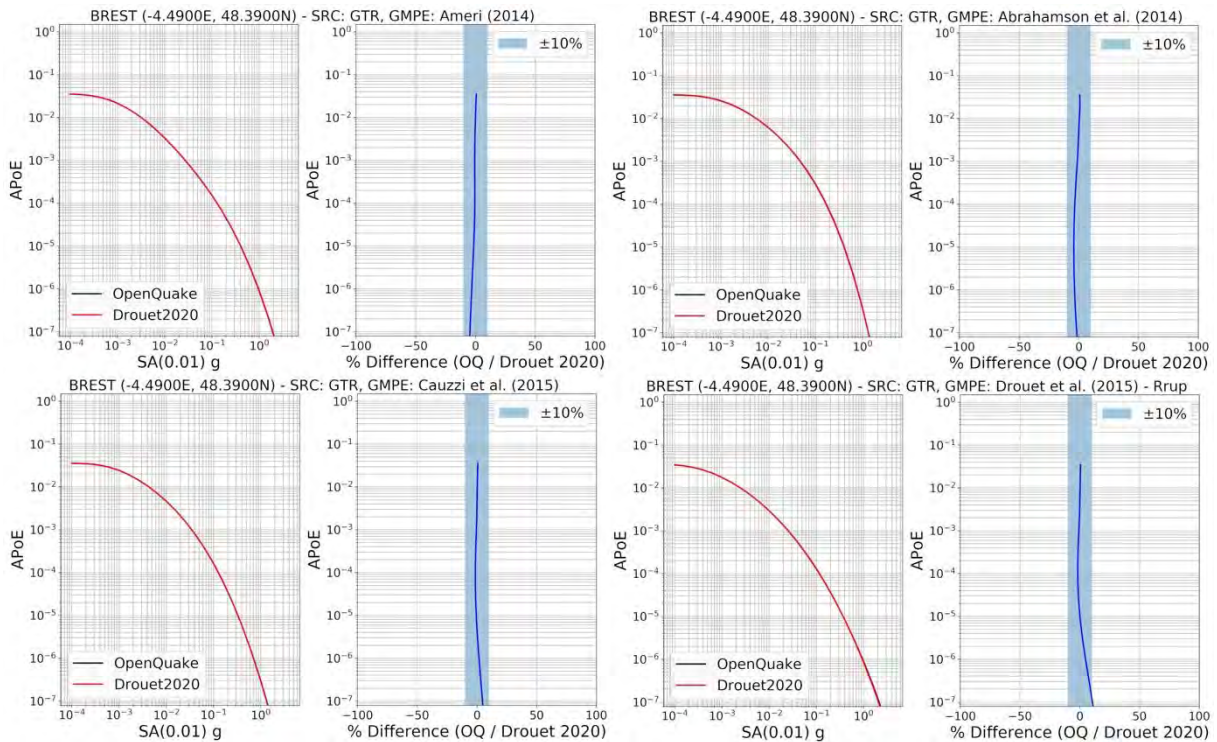


Figure 13a – Comparison of the hazard curves for Brest using the EDF source model and four GMMs. Upper four figures for Sa (0.01 s) and lower four figures for Sa (1.0 s). GMMs are A14 (top left), ASK14 (top right), C15 (bottom left) and DC15 (bottom right). Each pair of figures shows the original hazard curves on the left and the percentage difference between OpenQuake and the original curves from Drouet et al. (2020) on the right. For the difference curves the shaded area indicates the $\pm 10\%$ range.



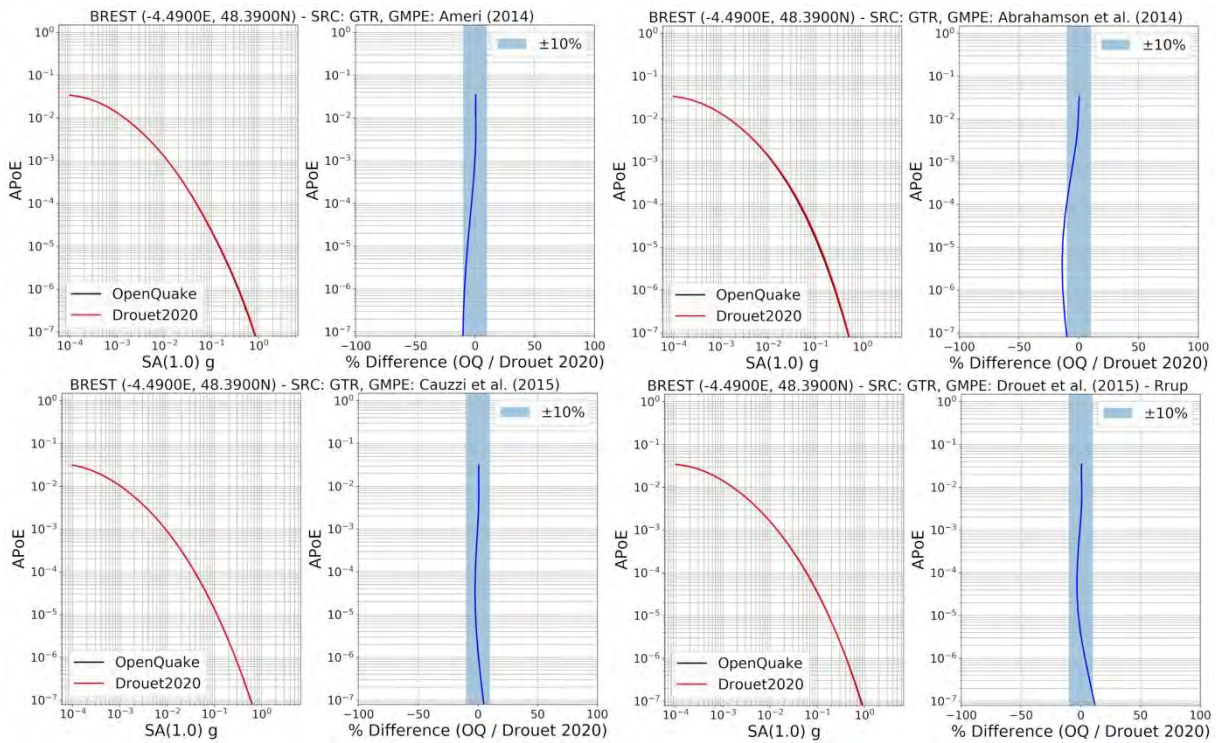
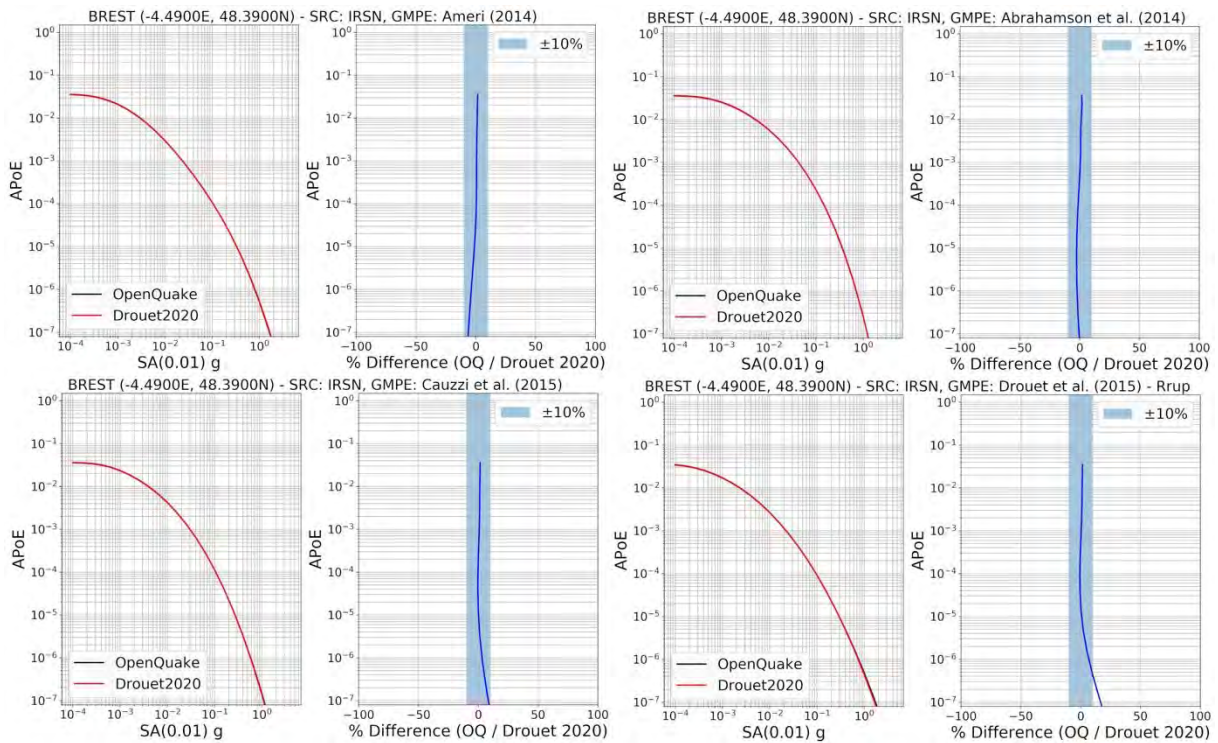


Figure 13b: As Figure 13a, for the GTR source model



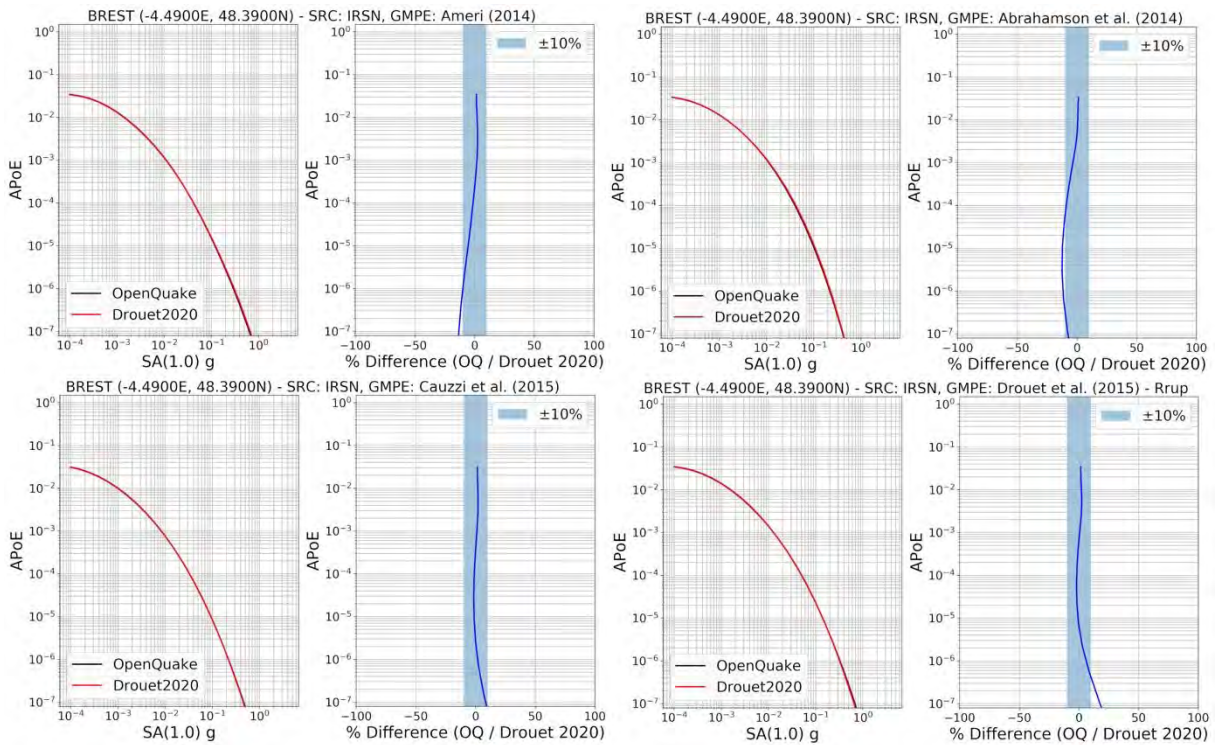
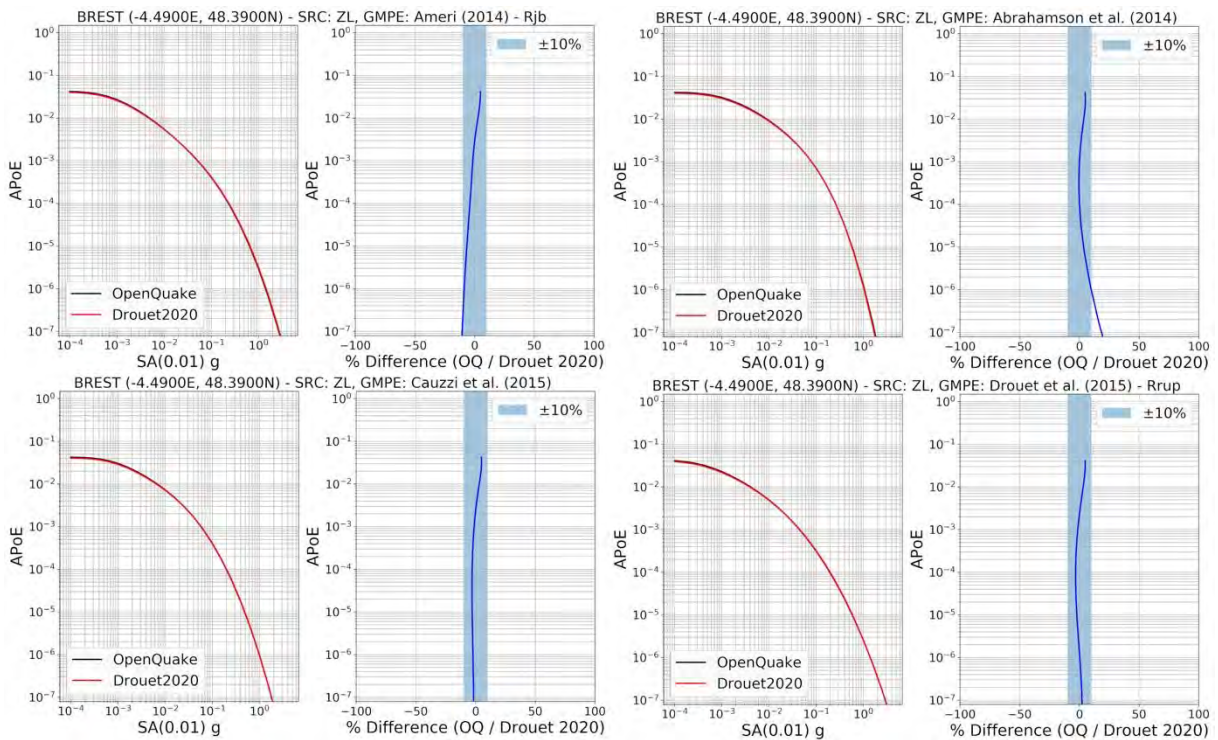


Figure 13c: As Figure 13a, for the IRSN source model



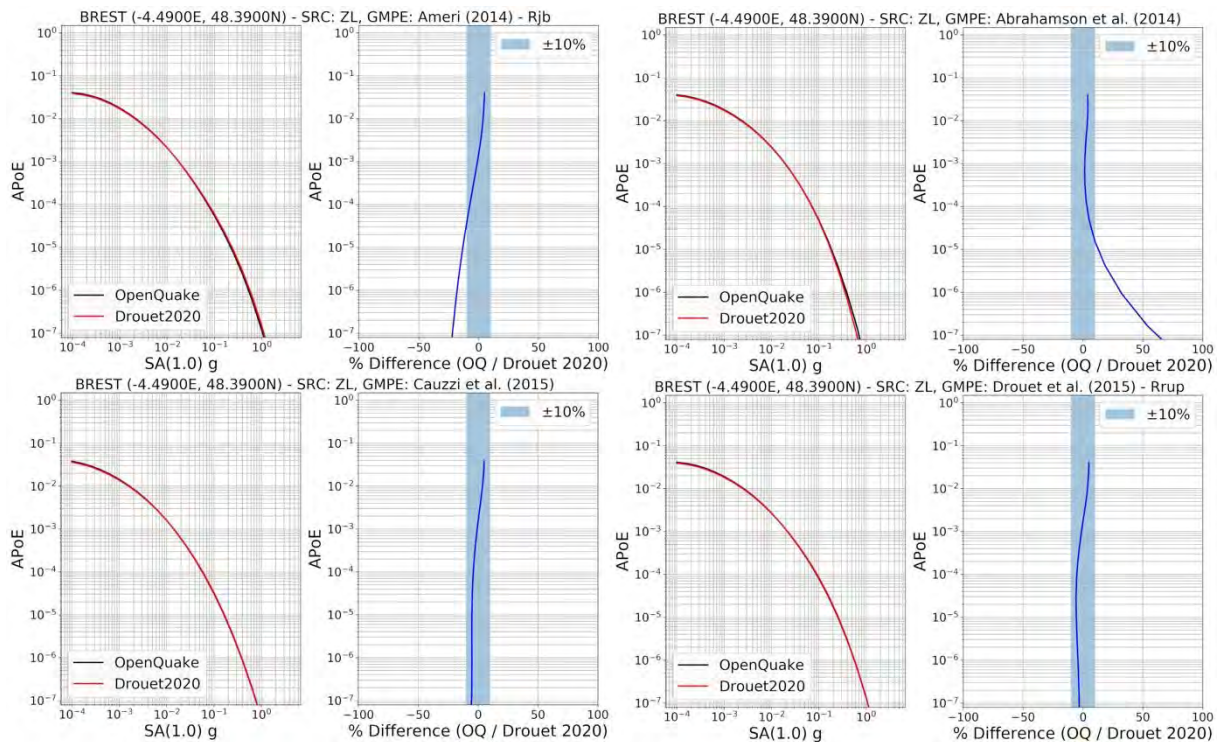


Figure 13d: As Figure 13a, for the ZONELESS source model

A minor observation, but one that will become more relevant later on, is that for the three area source models the disagreement between the original and OpenQuake curves converges to 0 at the highest annual probabilities of exceedance. Assuming that all relevant truncations have been applied correctly, this indicates that the total rate of occurrence of all ruptures within the 200 km matches correctly between the two models. This would imply that the characterisation of the sources is in good agreement and that the differences would likely come from either the manner in which the source-to-site distance is characterised, the evaluation of the density function of the magnitude frequency distribution, or the actual ground motion calculations themselves. For the ZONELESS model, however, the OpenQuake hazard curves are slightly higher (albeit within the threshold) for the highest APoE. This indicates that the total rates of occurrence of all ruptures affecting the site are marginally higher in the OpenQuake implementation.

Lourdes

Lying in the middle of region of elevated seismicity in France, the seismic hazard curves for Lourdes, shown in Figures 14a to 14d, are higher compared to those for Brest. Here we see two cases in which the acceptance threshold exceeded, as differences in hazard for both the Ameri (2014) and Abrahamson et al. (2014) GMMs exceed the $\pm 10\%$ threshold at APoEs as high as 10^{-3} for the EDF model. The Cauzzi et al. (2015) and Drouet & Cotton (2015) models manage to stay within the threshold for this model, although they too come close to it in the $10^{-3} \geq \text{APoE} \geq 10^{-4}$ range. However, for the GTR, IRSN and ZONELESS models the hazard curves from all of the GMMs seem to be in better agreement, with the exception of the Abrahamson et al. (2014) GMM in the ZONELESS case, which seems to diverge more at longer periods and lower APoEs. Once again, the total rate of ruptures affecting the site seems higher in the ZONELESS model than in the other models.

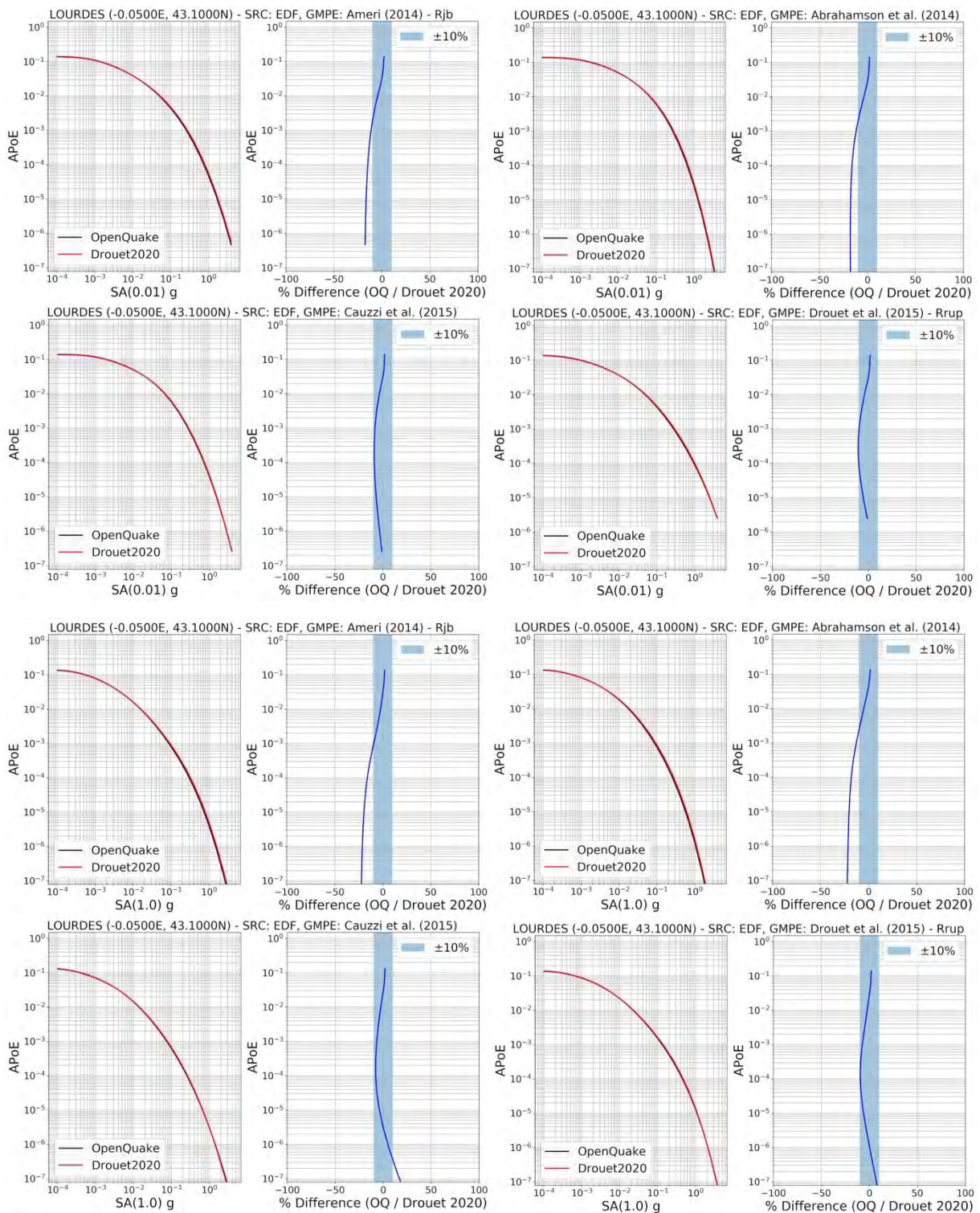


Figure 14a: Comparison of the hazard curves for Lourdes using the EDF source model and four GMMs. Upper four figures for $S_a(0.01\text{ s})$ and lower four figures for $S_a(1.0\text{ s})$. GMMs are A14 (top left), ASK14 (top right), C15 (bottom left) and DC15 (bottom right).

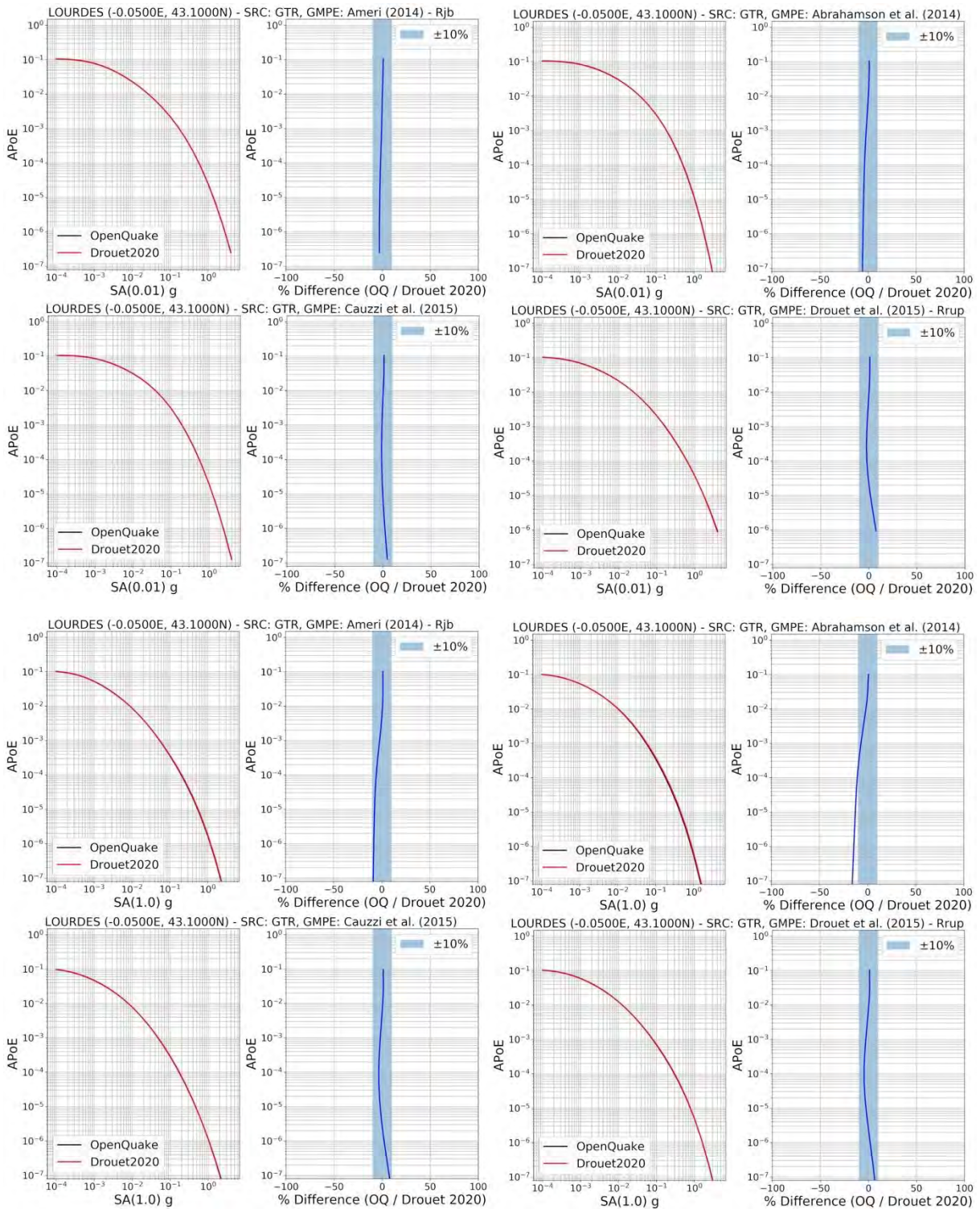


Figure 14b: As Figure 14a, for the GTR source model

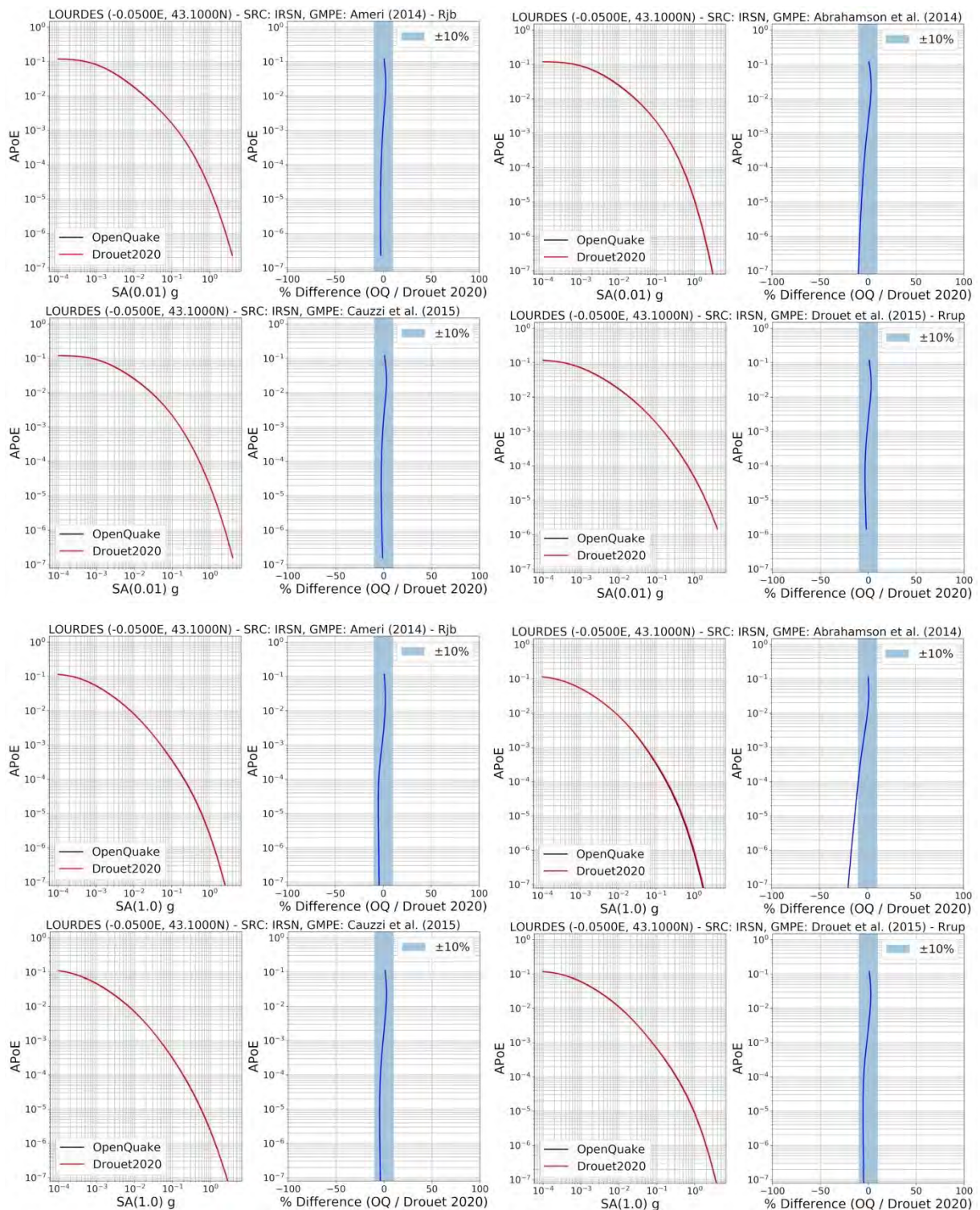


Figure 14c: As Figure 14a, for the IRSN source model

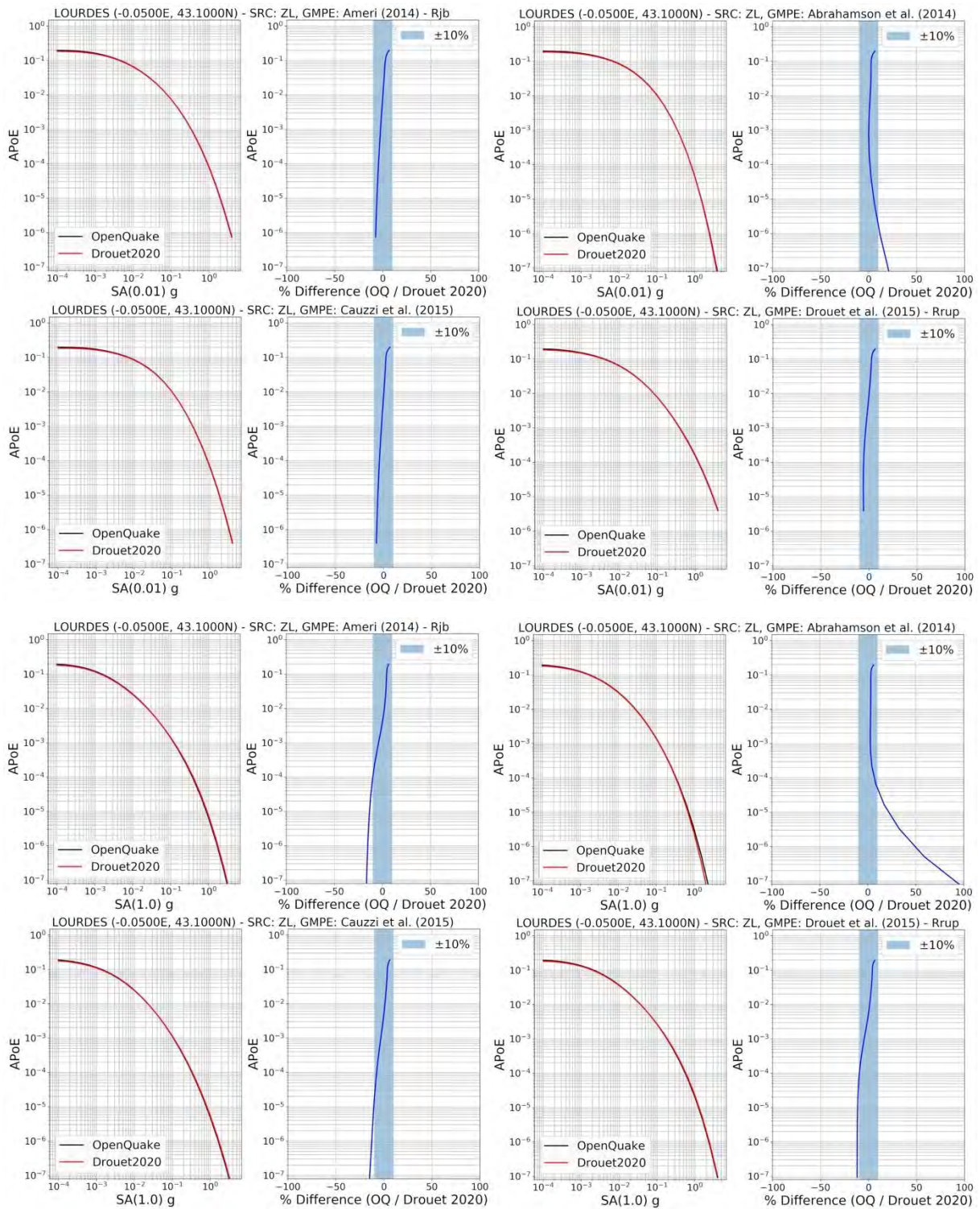


Figure 14d: As Figure 14a, for the ZONELESS source model

Marseille

The hazard curves for Marseille are shown in Figures 15a and 1d, and while the agreement is adequate (all curves fall within the $\pm 10\%$ threshold for $APoE \geq 10^{-4}$), some differences emerge here that contrast with the Lourdes and Brest case. In particular, we see here, for the first time, cases where the disagreement between the original and OpenQuake implementations seems to be more pronounced for

the Cauzzi et al. (2015) and Drouet & Cotton (2015) models than they are for Ameri (2014) and Abrahamson et al. (2014). This can be seen for the GTR and IRSN source models. There is also some suggestion that the total rates of occurrence might be higher in the EDF model than in the other area models, though this too is not as obvious as it is in the ZONELESS model. In many of the cases that do fall within the threshold, OpenQuake seems to be higher at high APoEs and lower than the original curves at longer APoEs, implying a steepening of the hazard curve.

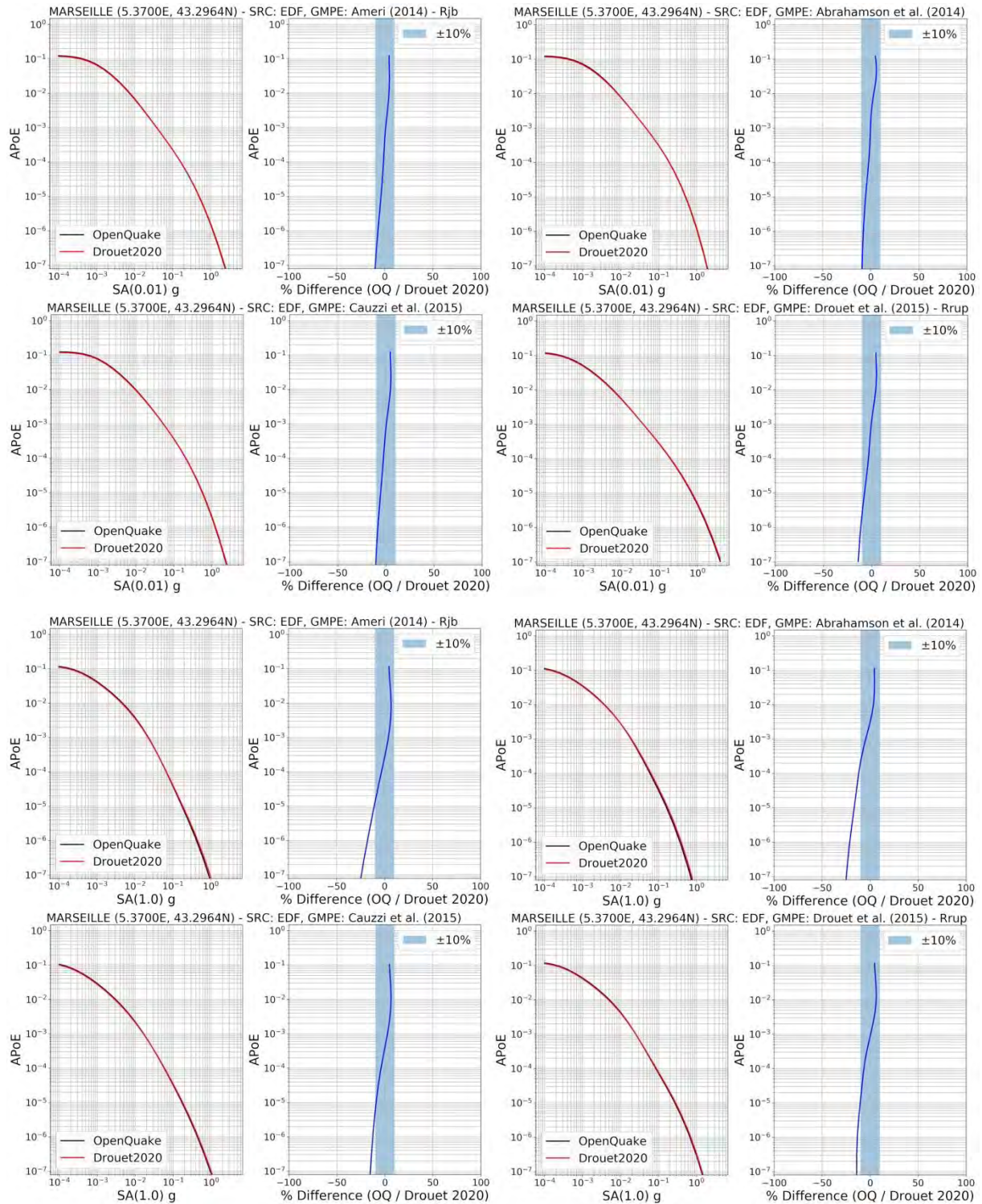


Figure 15a: Comparison of the hazard curves for Marseille using the EDF source model and four GMMs. Upper four figures for S_a (0.01 s) and lower four figures for S_a (1.0 s). GMMs are A14 (top left), ASK14 (top right), C15 (bottom left) and DC15 (bottom right).

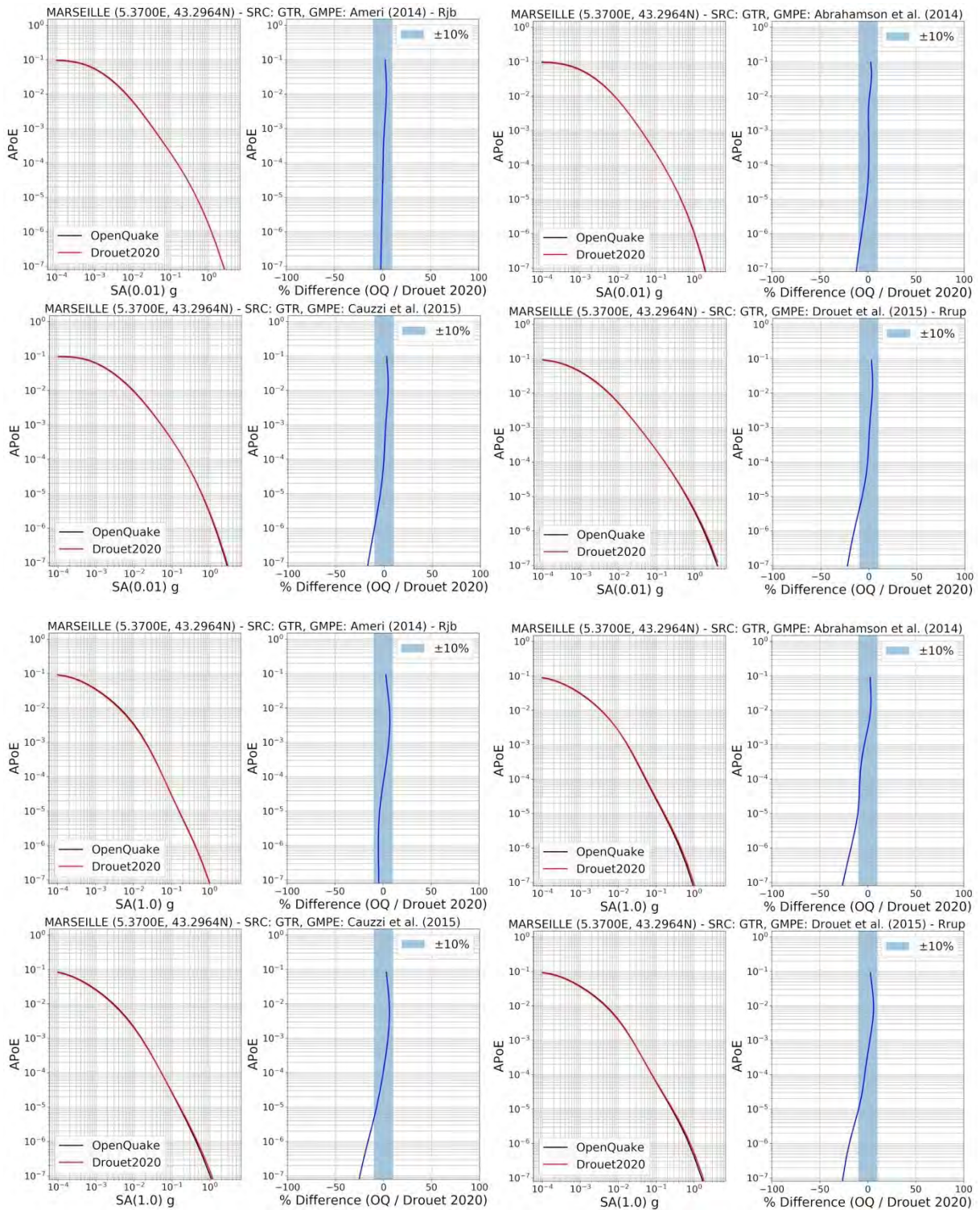


Figure 15b: As Figure 15a, for the GTR source model

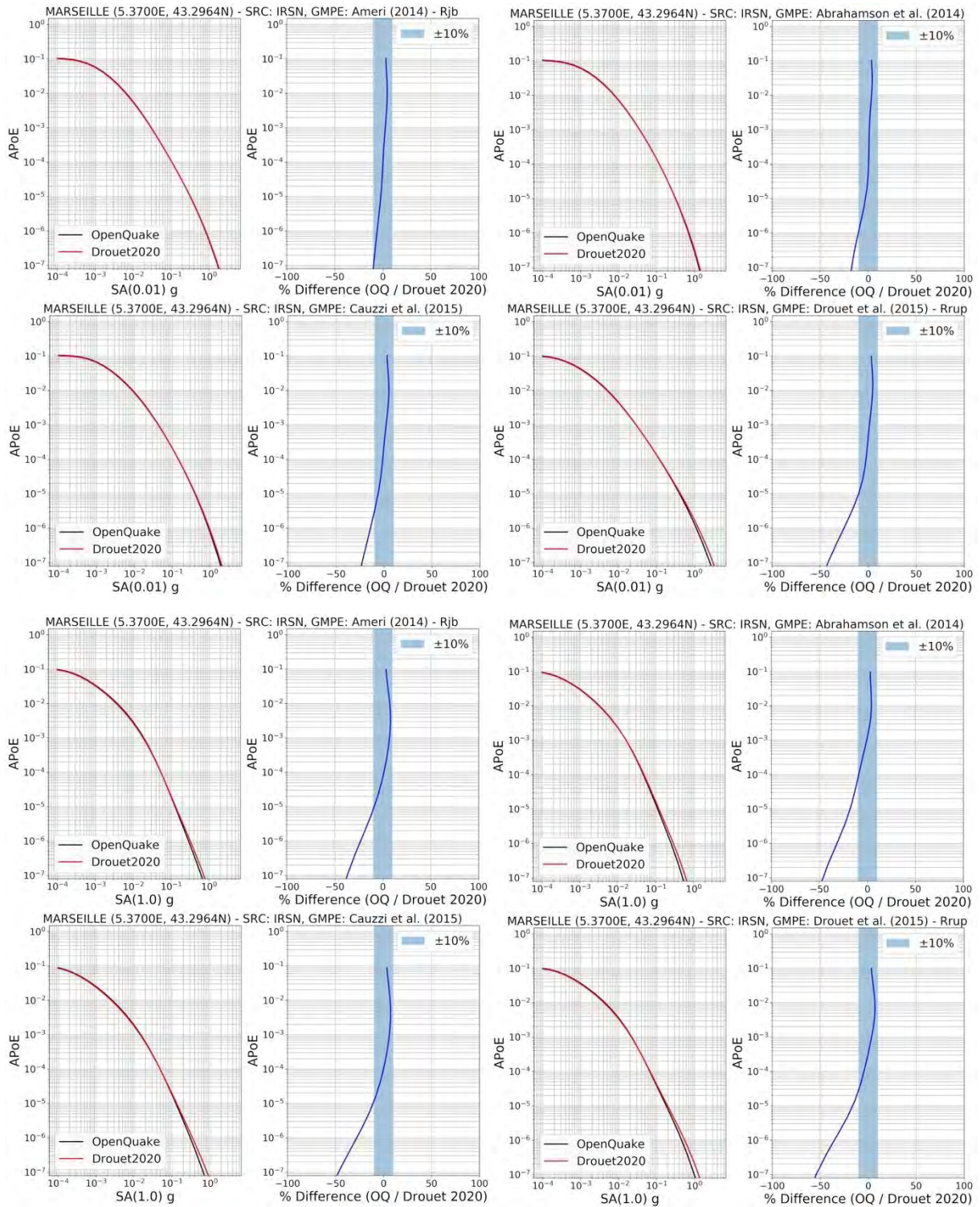


Figure 15c: As Figure 15a, for the IRSN source model

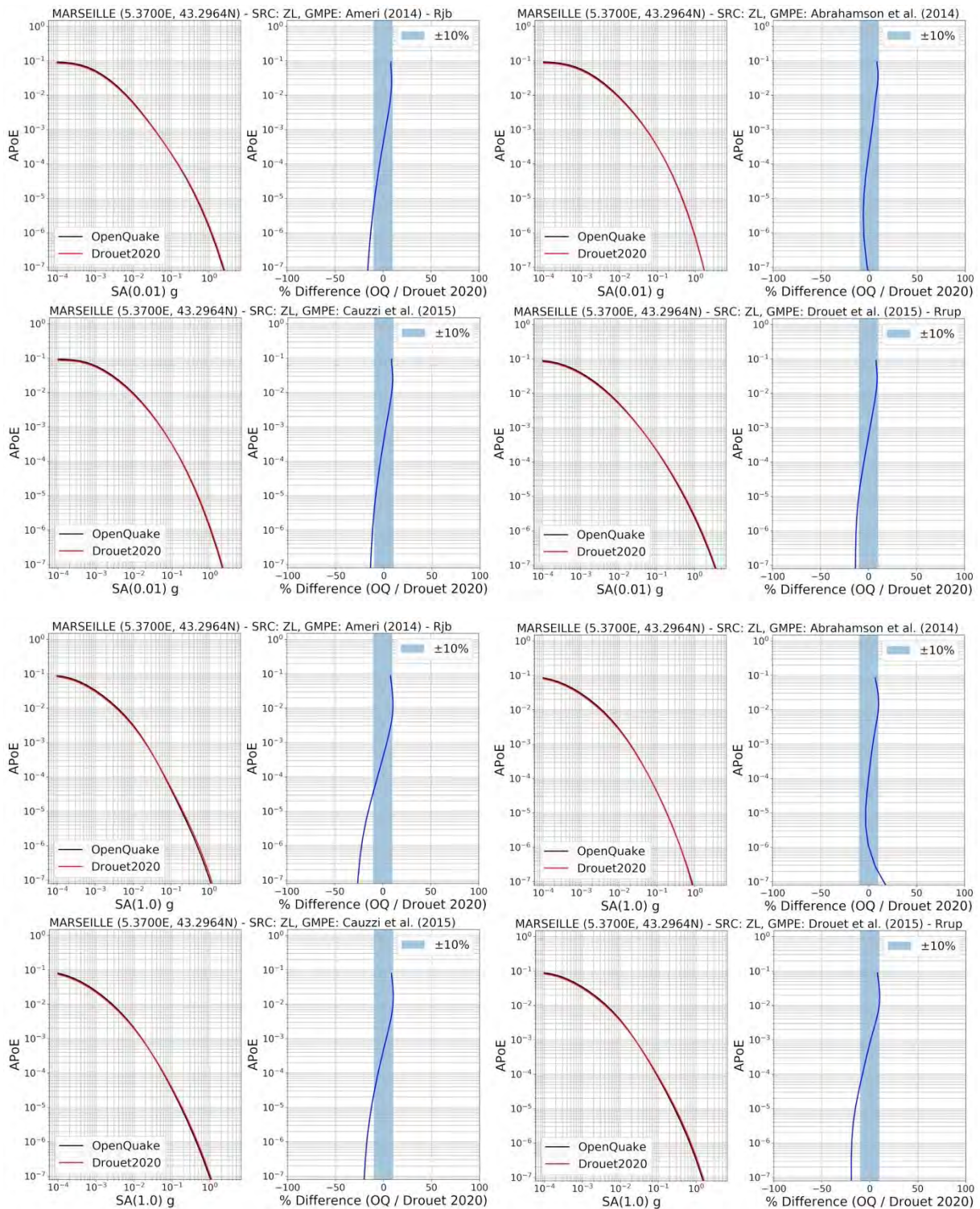


Figure 15d: As Figure 15a, for the ZONELESS source model

Grenoble

The hazard curve comparisons for Grenoble (Figures 16a to 16d) are quite mixed overall, with some cases achieving excellent agreement but others exceeding the $\pm 10\%$ threshold at $APoE \geq 10^{-4}$. The “misfitting” cases are the Ameri (2014) and Abrahamson et al. (2014) model at longer spectral

periods in the EDF and IRSN models. Otherwise, the trends are consistent with those seen in Lourdes and Brest.

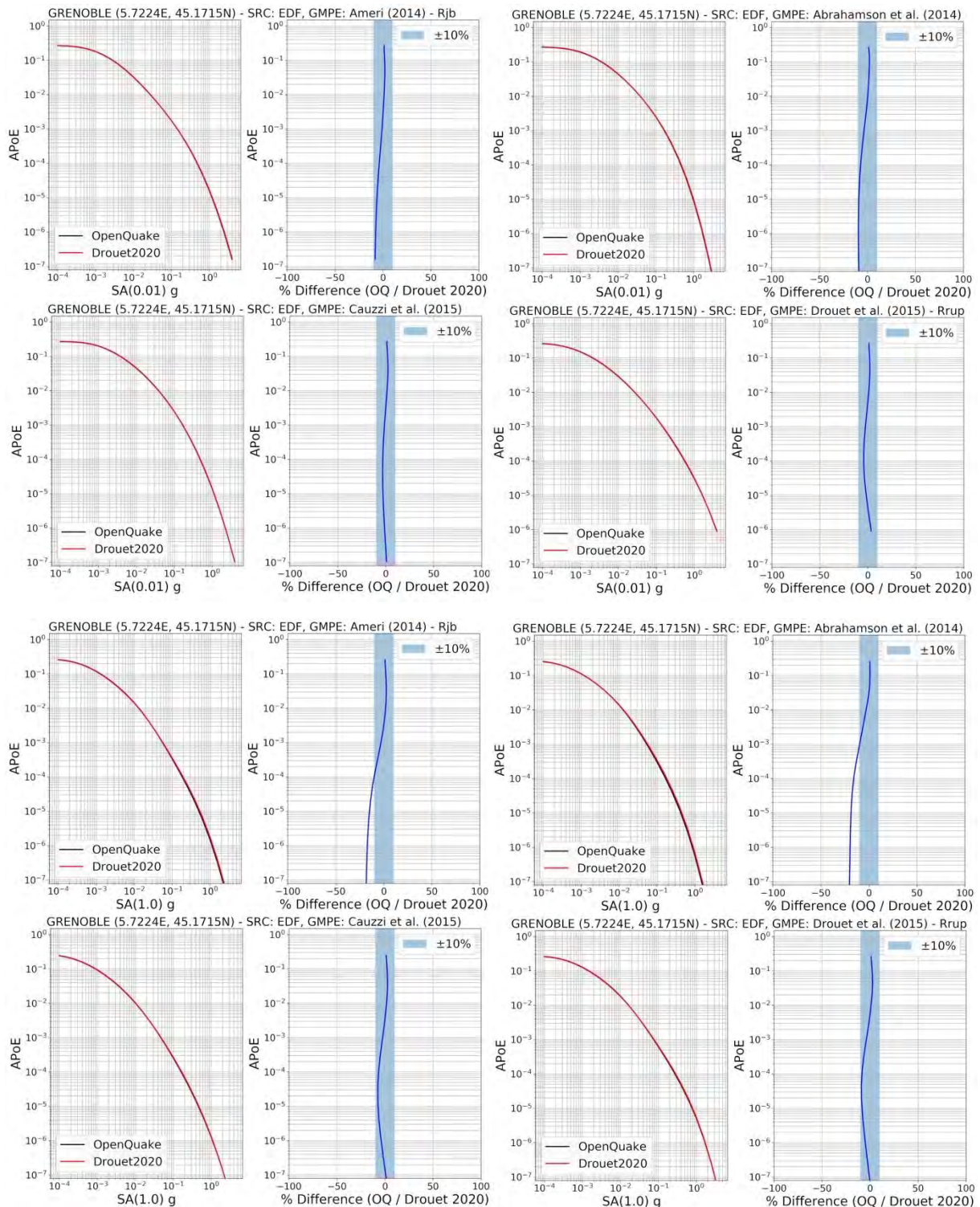


Figure 16a – Comparison of the hazard curves for Grenoble using the EDF source model and four GMMs. Upper four figures for $S_a(0.01\text{ s})$ and lower four figures for $S_a(1.0\text{ s})$. GMMs are A14 (top left), ASK14 (top right), C15 (bottom left) and DC15 (bottom right).

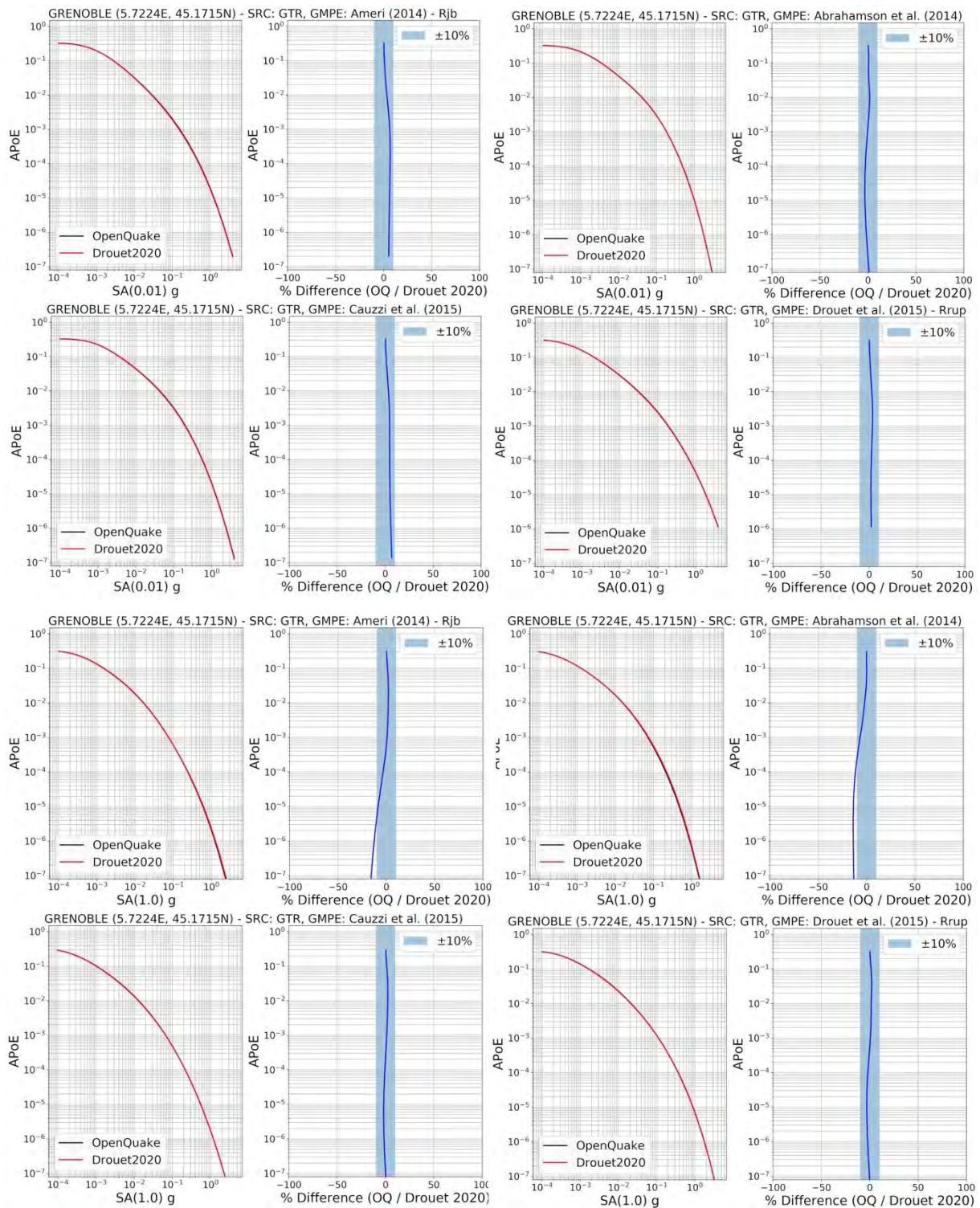


Figure 16b: As Figure 16a, for the GTR source model

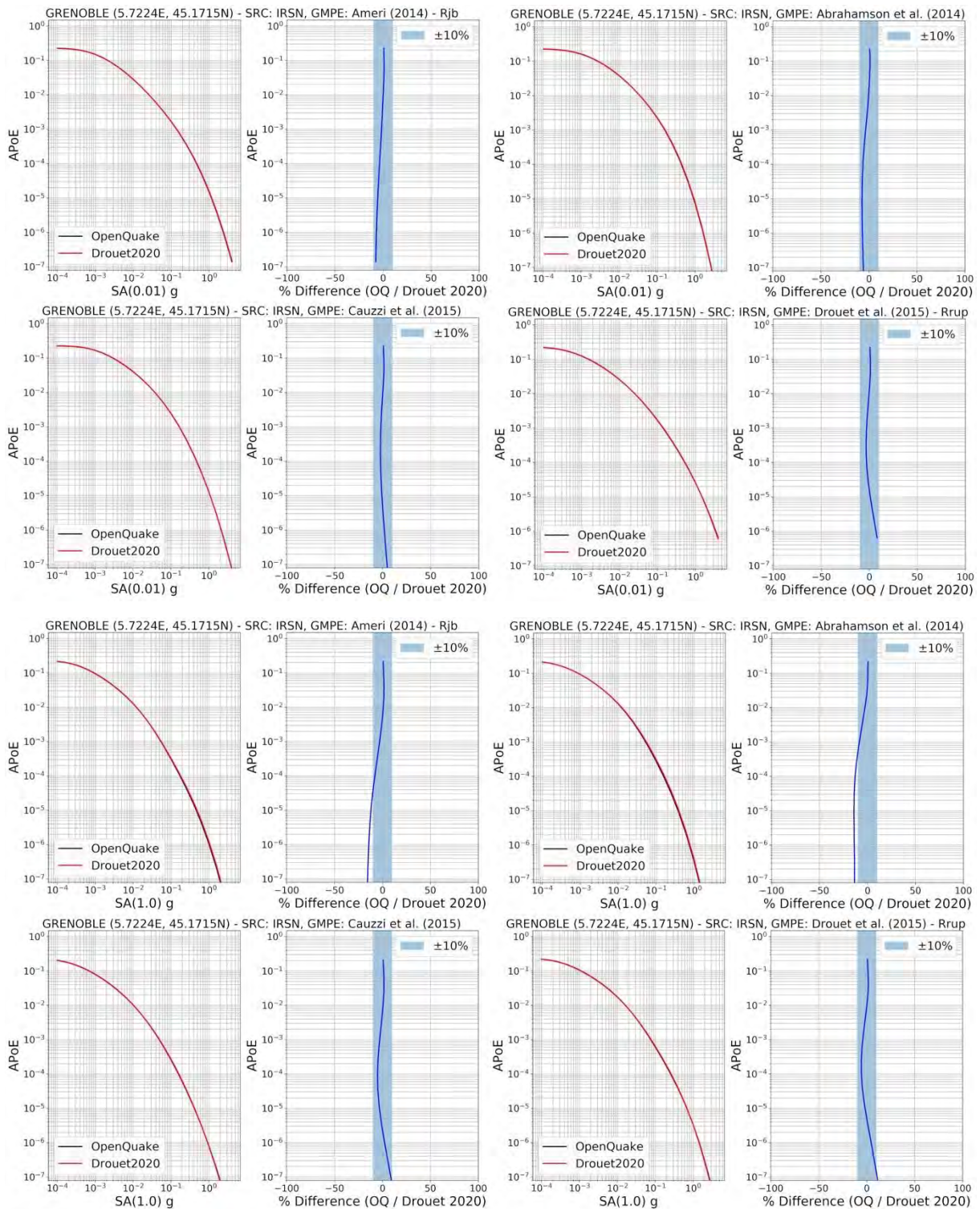


Figure 16c: As Figure 16a, for the IRSN source model

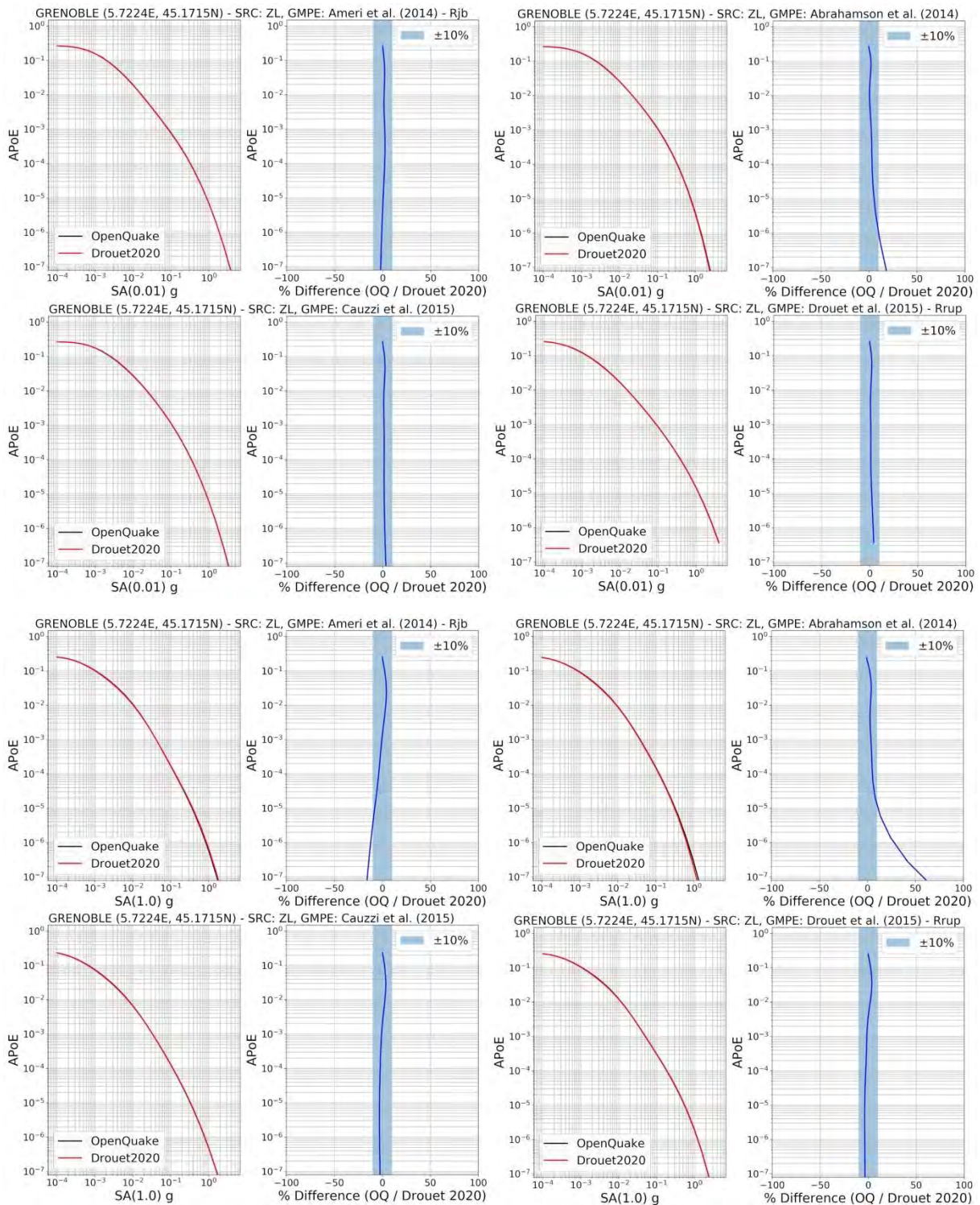


Figure 16d: As Figure 16a, for the ZONELESS source model

Nice

As the final of the five case study sites, the comparisons shown in Figures 17a to 17d for Nice share some similarities with those for Grenoble. Here, particularly good agreement is seen for most GMMs for the GTR and IRSN source models, while for the EDF model the target agreement is not reached in the Abrahamson et al. (2014) case (and only just achieved for the other GMMs).

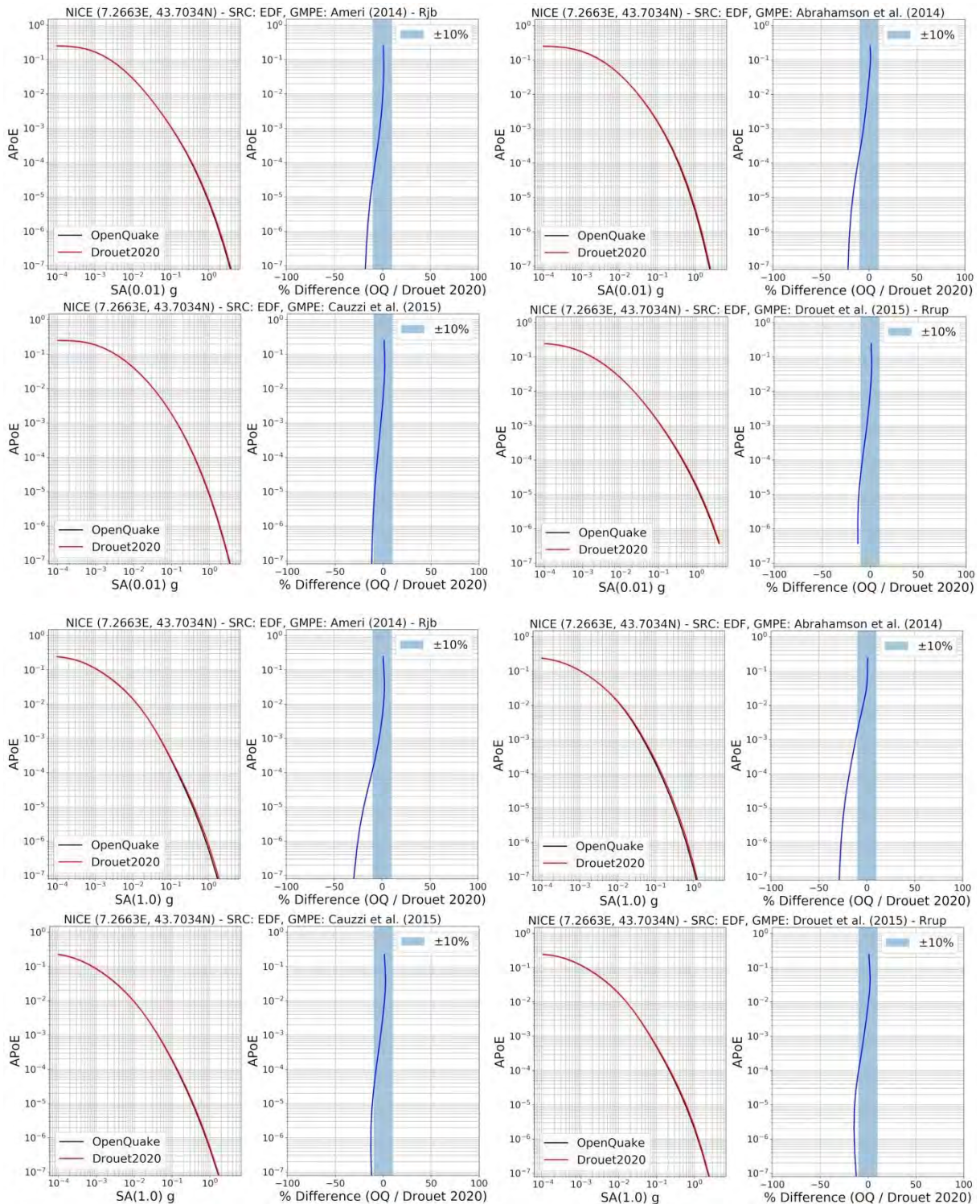


Figure 17a – Comparison of the hazard curves for Nice using the EDF source model and four GMPEs. Upper four figures for Sa (0.01 s) and lower four figures for Sa (1.0 s). GMPEs are A14 (top left), ASK14 (top right), C15 (bottom left) and DC15 (bottom right)

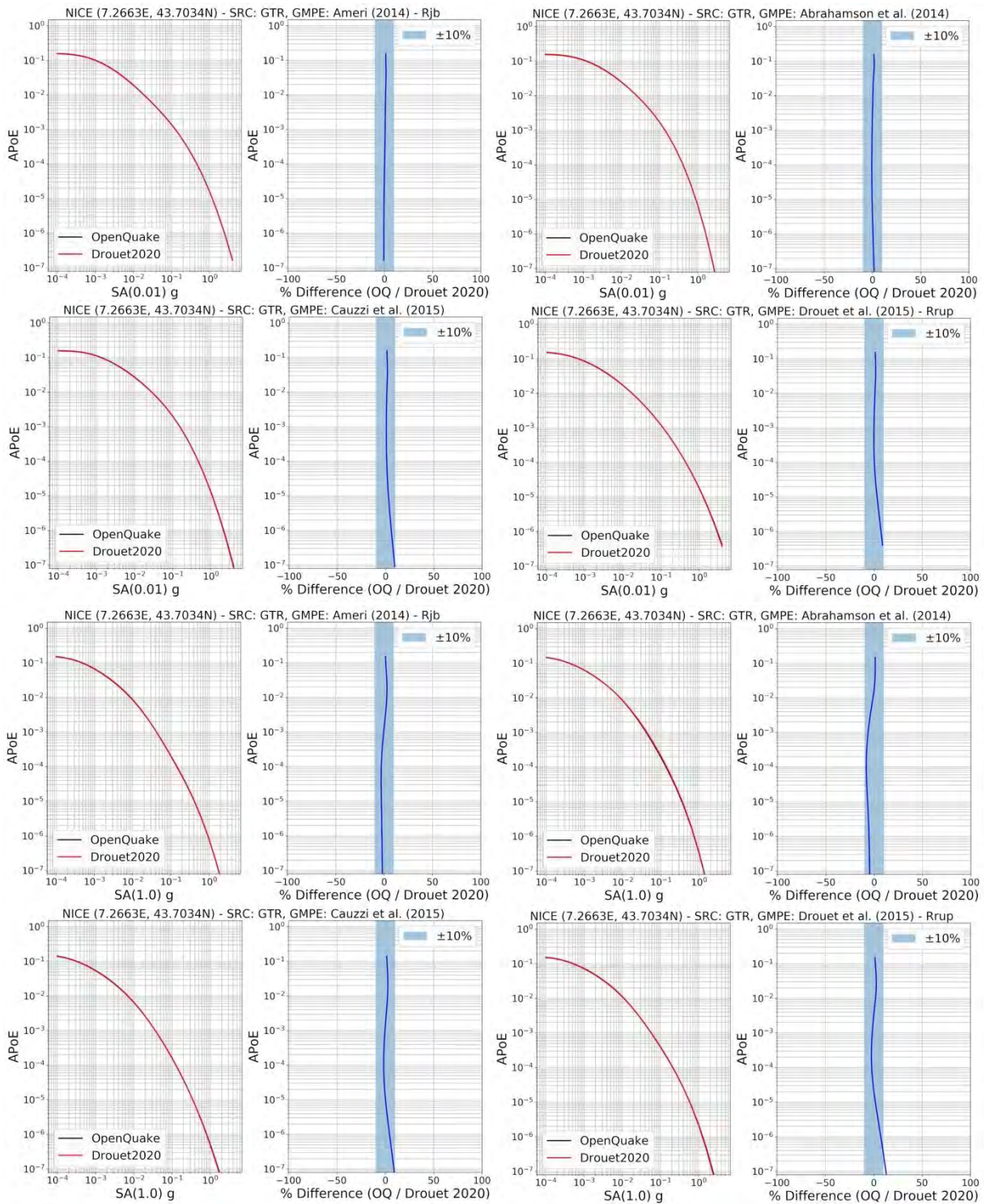


Figure 17b: As Figure 17a, for the GTR source model

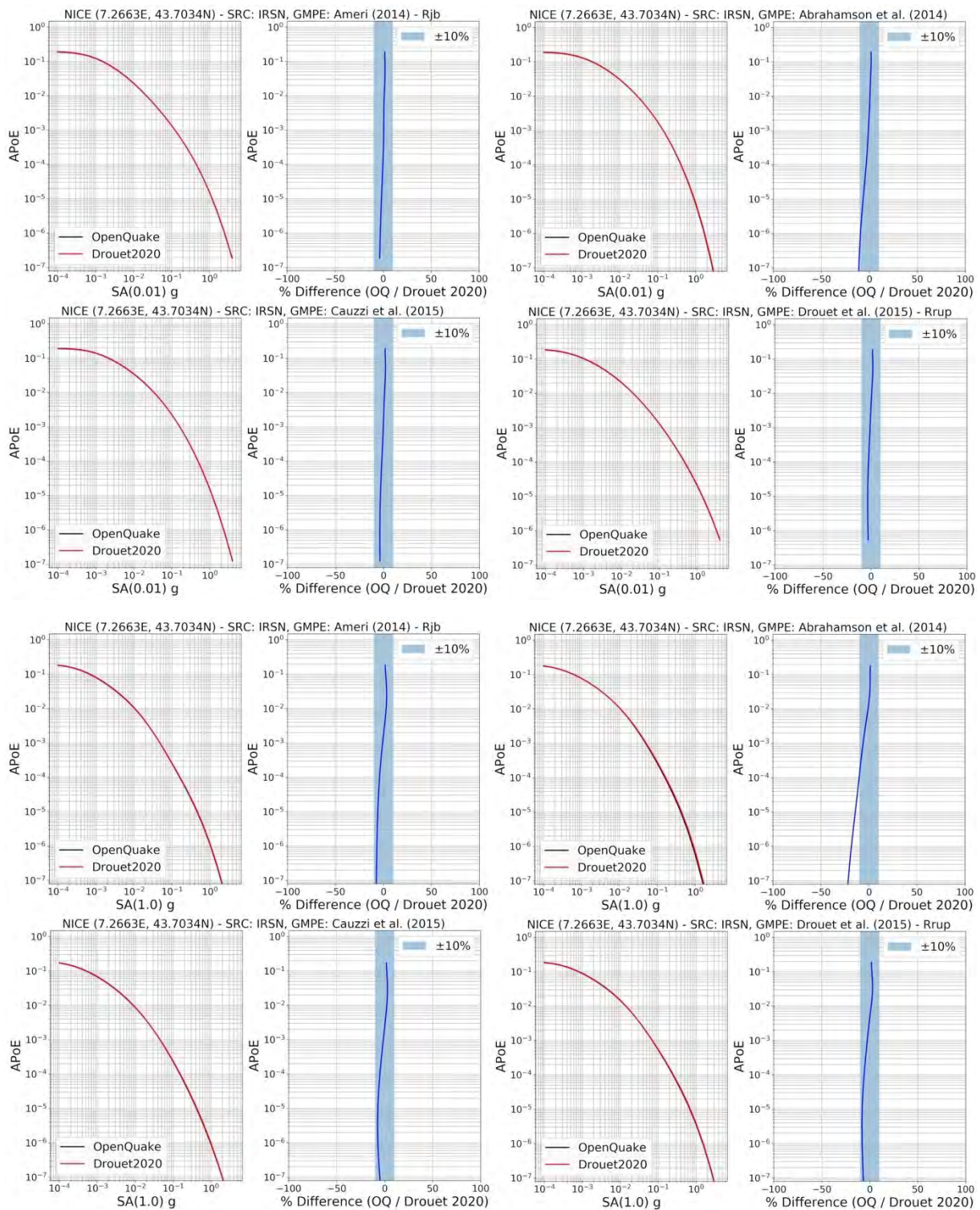


Figure 17c: As Figure 17a, for the IRSN source model

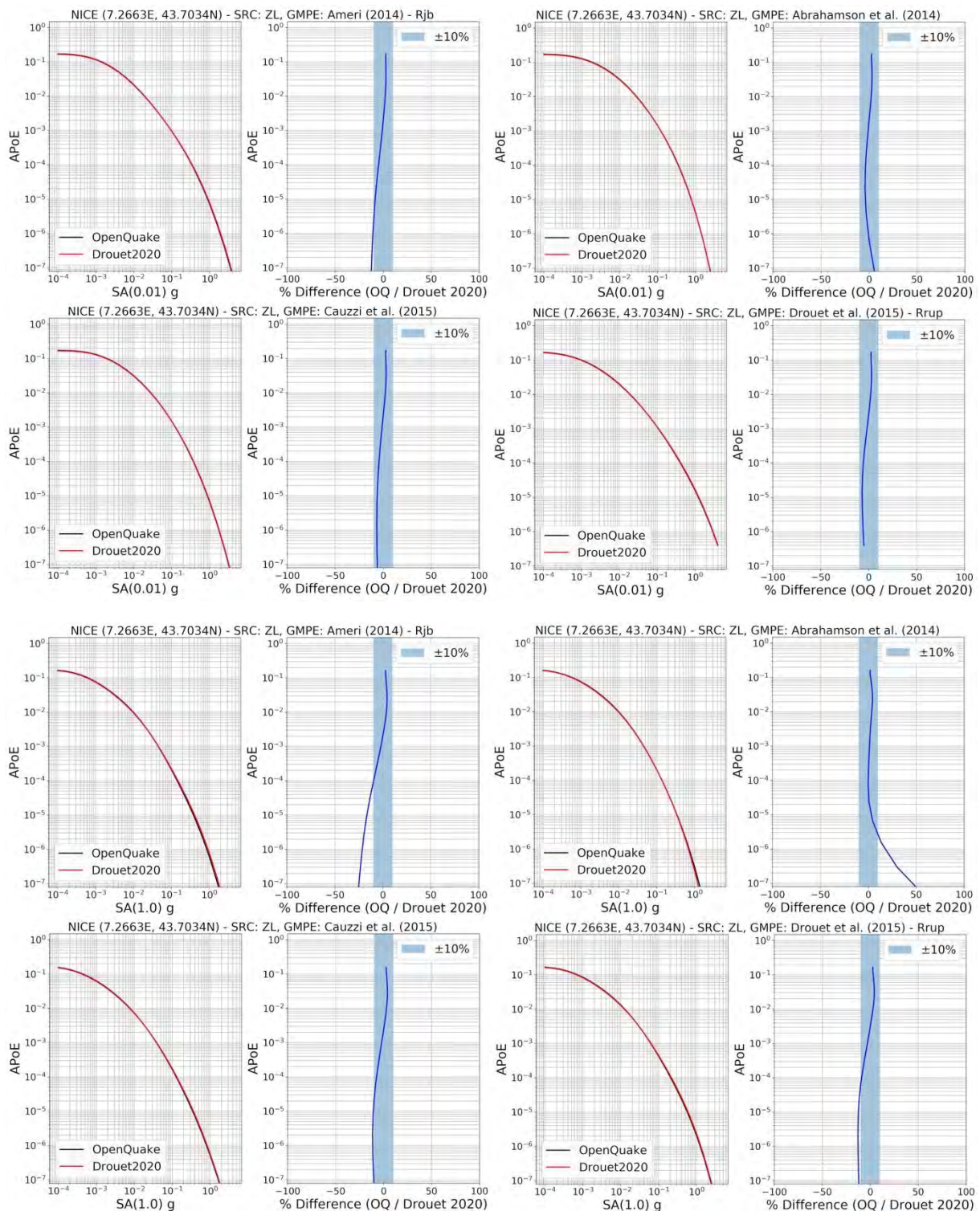


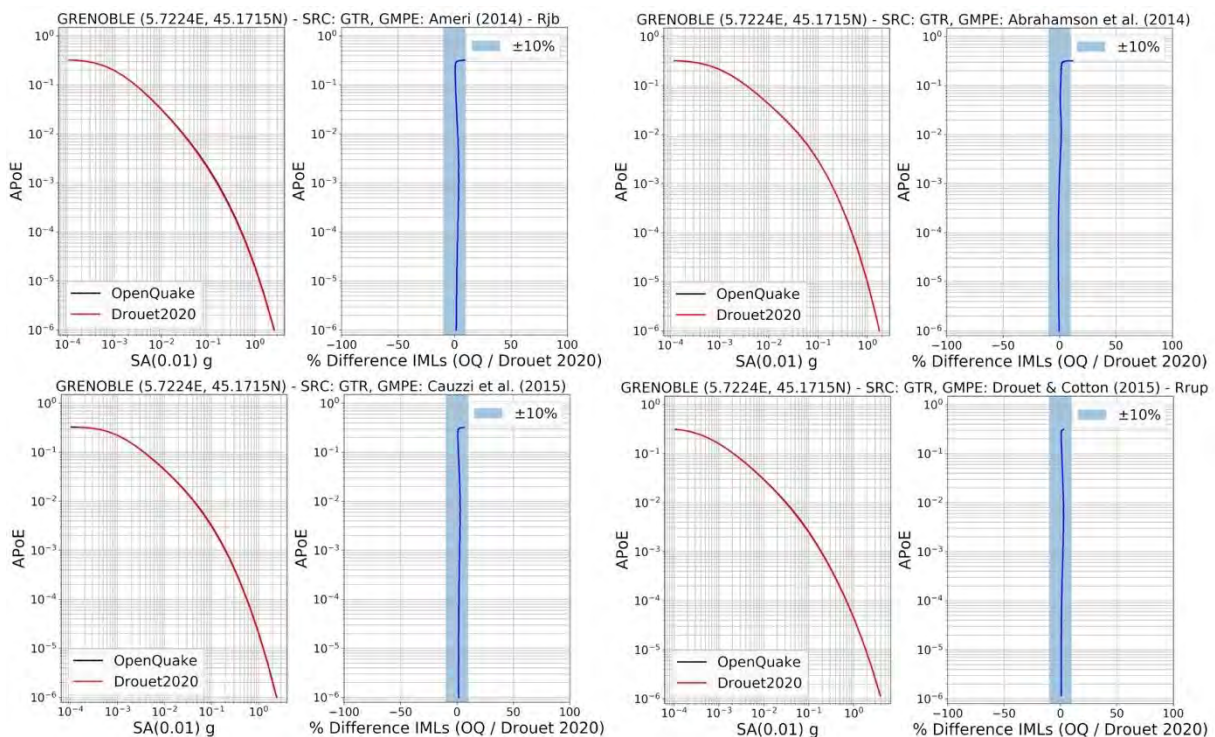
Figure 17d: As Figure 17a, for the ZONELESS source model

Comparing Annual Probabilities of Exceedance (APoEs) or Intensity Measure Levels (IMLs)

In the branch-by-branch comparisons shown for the five target cities the differences in the curves, and the criterion by which we define “agreement”, are quantified in terms of differences in the annual probabilities of exceedance for a fixed set of intensity measure levels (IMLs). This particular framing is chosen as it is confirmed that both OpenQuake and the original software calculate and return the seismic

hazard curves in this form. This also allows us to select *exactly* the same set of intensity measure levels for each intensity measure type as inputs into the calculation and reduce the possibility of discrepancies emerging from issues such as the interpolation of PoE or IML. For engineering applications, however, the criterion for agreement may be better framed in terms of the differences in intensity measure level for the same probabilities of exceedance. This does, of course, require interpolating both sets of curves to a reference set of target probabilities of exceedance from which the intensity measure levels can be calculated.

In Figure 18 we illustrate the comparison in terms of difference in expected intensity measure levels for the Grenoble site and the GTR source model, which can be compared against Figure 16b. Here we can see that just as in the case in which we compared APoE, for which the OpenQuake implementation generally passed the acceptance criteria, the same criteria of no more than $\pm 10\%$ difference in IML for APoE $\geq 10^{-4}$ is fulfilled. A minor divergence is seen at the extreme high probability end of the curve as a result of the extrapolation of the curves to the highest probabilities considered in the target range. This is not relevant for return periods of engineering interest however. Comparing the difference curves even more closely with Figure 16, it could even be argued that the agreement in this case is improved, with agreement kept almost entirely within the $\pm 5\%$ difference range for the APoE $\geq 10^{-4}$. This analysis has been repeated for other sites and branches, which we omit here for brevity, and while small differences in agreement in terms of IMLs do emerge for particular sites and models in the overwhelming majority of branches we find that when the curves pass the agreement criterion in terms of APoE then they do so in terms of IML. We therefore conclude that at a branch-by-branch level the agreement is satisfactory whether given in terms of APoE or IML.



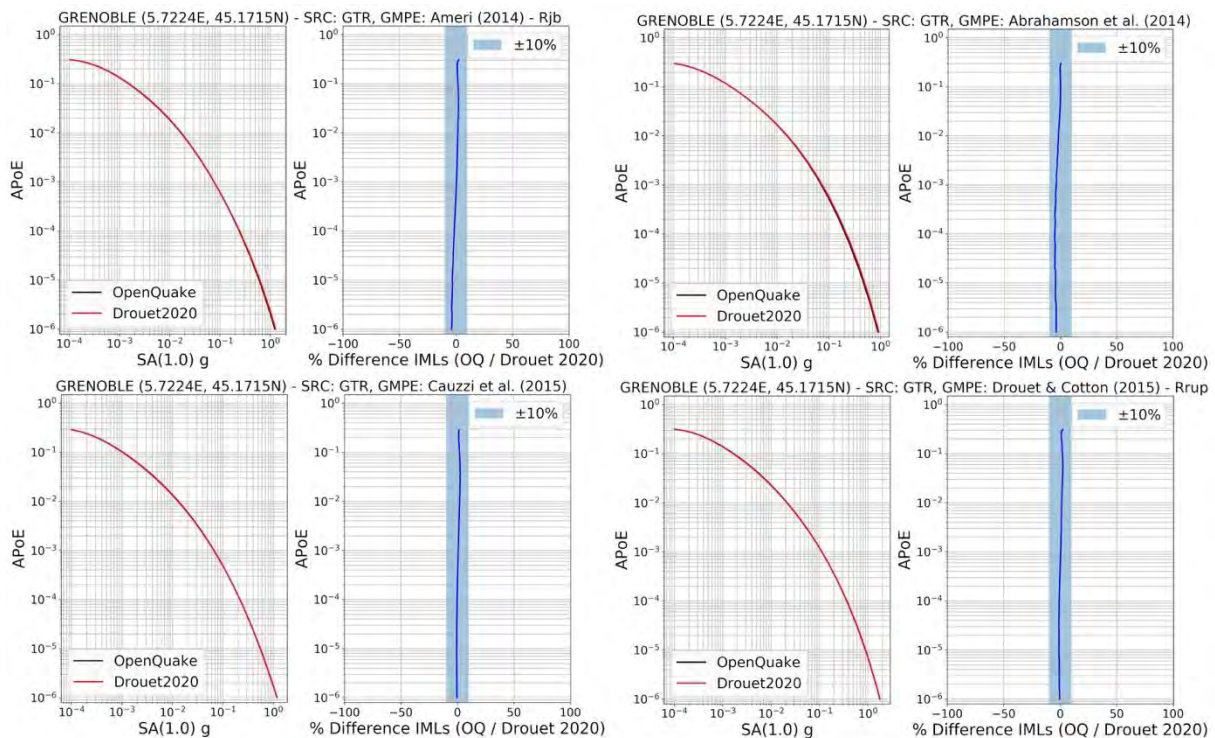


Figure 18: Comparison of seismic hazard curves and differences in terms of IMLs rather than APoE for the Grenoble test case and GTR source model according to the same configuration and criteria of Figure 16.

4.2 Identifying and Interpreting the Possible Causes of the Differences

The total set of branch-by-branch comparisons for the case study sites shown in Figures 13 to 17 illustrate that for APoEs $\geq 10^{-4}$ the target agreement of $\pm 10\%$ is achieved in the significant majority of cases. This would suggest a good agreement can be achieved overall for the target return periods of interest. In many cases where the accepted agreement level is reached for APoE $\geq 10^{-4}$, the agreement deteriorates at lower APoEs. This is not unexpected as we get to smaller and smaller probabilities of exceedance the minor differences in operation of the software come to be more visible. This has been seen in the PEER Tests (Thomas et al. 2021; Hale et al. 2020) and was a key reason as to why we set APoE $\geq 10^{-4}$ as the target. It is also interesting to observe cases in which the agreement is excellent at all APoEs, even if for the same site there may be significant differences for other source models or GMMs. The fact that not all trends are persistent across the sites can make their interpretation difficult, as one cannot necessarily ascribe the changes to single factor.

Without being able to conduct a set of more controlled seismic hazard calculations, or to access the source code of the in-house software used by Drouet et al. (2020) in their PSHA calculations, it is impossible to elucidate which of the potential causes of discrepancies between the two software account for each of the differences seen in the branch-by-branch comparisons. We can therefore only speculate by outlining those differences of which we are aware and those that can commonly occur between different PSHA software. Though errors and misunderstandings can certainly occur, we will assume for the time being that i) the GMMs have been implemented correctly in both software, ii) the clarification provided to us by the authors are accurate.

Area source discretisation and distance calculation

The calculation of finite rupture to site distances for uniform area sources is probably one of the most fundamental differences between PSHA software, yet it is also arguably the most opaque. As discussed in section 3, OpenQuake discretises each of the areas into a set of points, which then become the

anchoring locations that form the centres of the finite ruptures (see previous description for more detail). We do not know exactly how the in-house software used by Drouet et al. (2020) determines the distance distribution, but we do have some insights from the authors: “*The spatial discretization is done through swinging arcs with increasing size away from the site. ... there are 15 ‘layers’ from 1 to 200 km*” (Stéphane Drouet, *personal communication*, 24/02/2021). *The size of the circles is proportional to the grid mesh. The grid meshing is very dense at short distance.* (Gabriele Ameri, *personal communication*, 26/02/2021) (see Figure 19a).

The use of “swinging arcs” for the distance calculations originates from the FRISK88 software², and it works by defining a set of N circles of increasing radius around the site, partitioning the rate of seismicity from a given source based on the area of overlap between the arc and the source in question. In the original FRISK88 software manual it is indicated that the finite rupture size is not considered for area sources, though it is possible subsequent generations of the software have made corrections for this. For example, in the Grünthal et al. (2018) model the finite rupture distances are determined using magnitude-dependent empirical formulae to convert R_{EPI} and R_{HYPO} into R_{JB} and R_{RUP} respectively (Christian Bosse, *personal communication*). Given the circle sizes indicated in Figure 19a, we have some reason to believe that the distances are accounting for this in some manner. The distances are then binned and normalized to describe the distance density function $f_r(R)$. This description is not entirely comprehensive, but it illustrates that a fundamental difference in defining $f_r(R)$ exists between the original software and OpenQuake.

The swinging arcs description does not explain how R_x and R_{y0} distance would be calculated for use with the Abrahamson et al. (2014) GMM, and whether or not consistency between definitions for a given earthquake can be retained (such as ensuring R_x is positive when $R_{JB} = 0$), as it can for the virtual faults in OpenQuake. Exacerbating the differences is the fact that OpenQuake does not discretise the distances into a normalised distribution $f_r(R)$, but rather it calculates $P(a \geq A|rup_{ij})$ explicitly for each rupture. Depending on the spacing of the distance bin, the binning can result in a slight bias toward longer distances at small magnitudes and shorter distances at larger magnitudes, when compared with direct calculation on a rupture-by-rupture basis. This *may* account for why the PSHA curves for the Abrahamson et al. (2014) models are often steeper in OpenQuake, as the smaller earthquakes tend to control the hazard at higher APoEs.

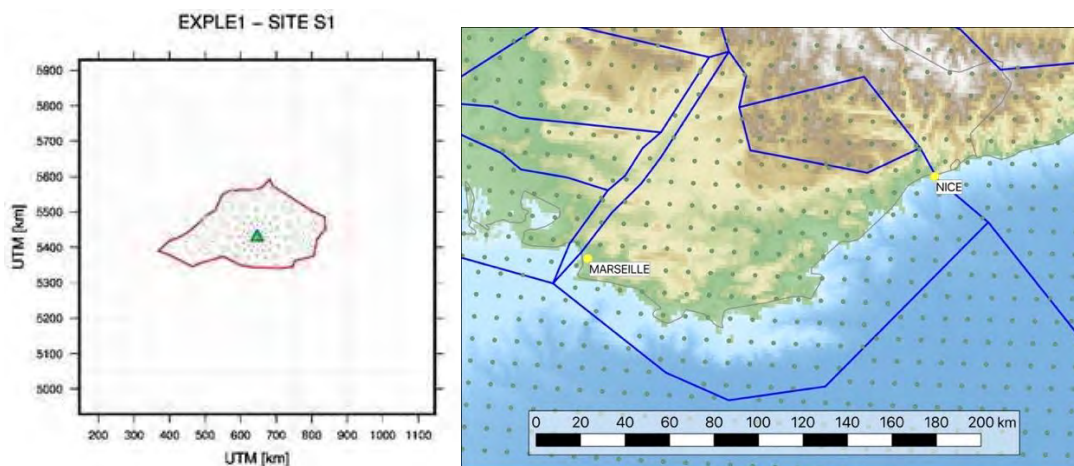


Figure 19: Illustration of swinging arc discretisation of a uniform area source (left) [Gabriele Ameri, *personal communication*] and OpenQuake fixed point grid spacing for the Marseille – Nice region using the EDF source model and a 10 km grid spacing.

² We assume that an FRISK88 derivative software has been used here as Fugro indicate that they conduct PSHA using a range of software (EZ-FRISK, FRISK88 and HAZ). EZ-FRISK originates from FRISK88, while the HAZ family of PSHA software do not, to our knowledge, use swinging arcs.

The comparison between the swinging arcs approach and OpenQuake's approach should not necessarily be interpreted as implying a greater degree of accuracy in OpenQuake's distance calculations in the present case. Indeed, by using a fixed discretisation spacing and by rendering the grid points independently from one uniform area source to the next, OpenQuake can sometimes create higher or lower densities of grid points close to the edges of source zones, particularly if the geometry is quite irregular or concave. This is illustrated in Figure 19b, where some uneven spacings can be seen around the borders of zones. The extent to which this can influence the seismic hazard at a site can depend on many factors and may, in some cases, be minimal and only come into play for very low APoEs. The proximity to a source boundary, the complexity of the source geometry, orientations of the virtual ruptures and the area discretisation spacing can all result in differences in the distance distributions when compared with the swinging arcs approach, and the influence of these differences on the resulting seismic hazard curves can vary significantly from site to site.

Initial Aspect Ratio and Rupture Dimension

Though the authors have provided some clarity regarding the rupture orientation, in light of their information and the operation of the swinging arcs approach for distance computation, is it not entirely clear how the ruptures dimensions are determined. Stéphane Drouet indicated the Wells & Coppersmith (1994) scaling relation is used and that “*maybe length and width are computed from scaling relationships*” (Stéphane Drouet, *personal communication*, 24/02/21). If this is the case, then we would interpret that aspect ratio is magnitude-dependent and determined as the ratio of the rupture length from the expected magnitude-length scaling relation to the rupture width from the rupture-width scaling relation. They confirm that rupture widths are constrained to the seismogenic thickness and that when the thickness is reached then the ruptures are allowed to increase in length, and therefore increase in aspect ratio. In contrast, in the OpenQuake implementations we use the Wells & Coppersmith (1994) magnitude to rupture area relation and an initial aspect ratio of 1.0.

In Figure 20 we compare the scaling with magnitude of the aspect ratios predicted by ratio of the Wells & Coppersmith (1994) rupture-length and rupture-area relations against those aspect ratios predicted by OpenQuake using an initial ratio of 1.0 but limited to the seismic thickness of 29.9 km indicated by Drouet et al. (2020). For this purpose, the Wells & Coppersmith (1994) subsurface rupture length (RLD) scaling relation is used, while for OpenQuake a characteristic dip angle of 45°, 60° and 85° are used for the reverse, normal and strike-slip cases respectively (i.e. the mid-points of the dip ranges).

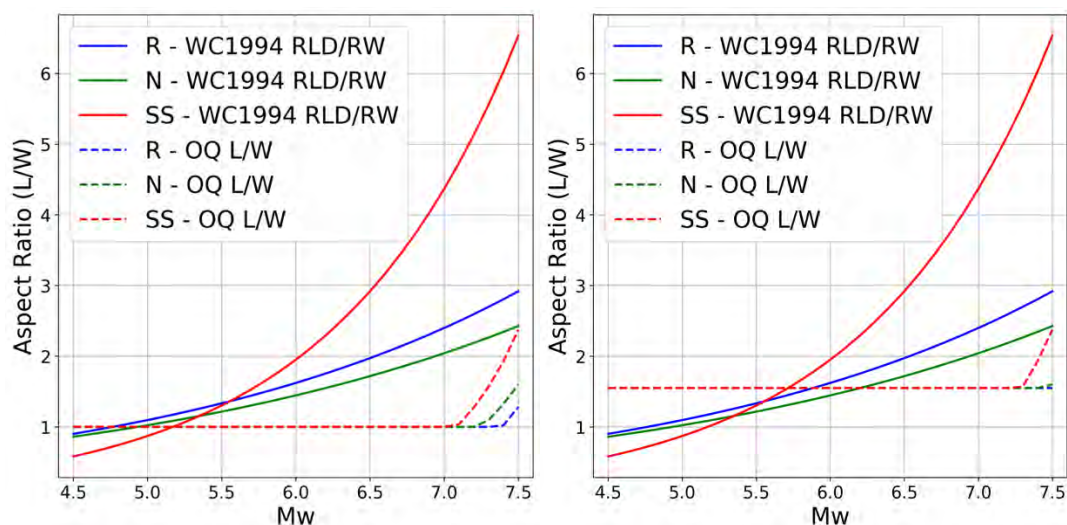


Figure 20: Implied aspect ratios from the rupture length and rupture width magnitude scaling relations of Wells & Coppersmith (1994) [solid lines] and those implied by OpenQuake from the Wells & Coppersmith (1994) magnitude-area scaling relation for a seismogenic thickness of 29.9 km and an initial aspect ratio of 1.0 (left) and 1.5 (right).

Figure 20 (left) reveals two important differences about the implied rupture aspect ratio of the two different PSHA calculations. Firstly, by taking the ratio of the magnitude-length and magnitude-width scaling relations, the implied aspect ratio in the original software is significantly higher than that of the OpenQuake ruptures for magnitudes $M \geq 5.5$. The second point to note is that given the relatively thick seismogenic layer, the initial aspect ratio of 1.0 is only “broken” for events of $M_w \geq 7.1$, which occur only in a small number of sources in the more active domains in western Switzerland and the Pyrenees (admittedly these are where the target sites of Lourdes, Nice and Grenoble are found!). Also, given that the aspect ratios suggested by the approach described by Stéphane Drouet are quite large for large magnitudes, it is unlikely that any ruptures would reach the maximum width implied by the seismogenic thickness and would therefore not actually break aspect ratio for larger magnitudes at all.

We will explore the sensitivity of hazard curve to the choice of initial aspect ratio in the next sub-section, but what is important to note here is that if the author’s description is correct then a substantially different scaling of aspect ratio is to be expected between the two codes. In Figure 20b we show the comparisons this time assuming an initial aspect ratio of 1.0. Here we can see that although this produces higher aspect ratios for events with $M \leq 5.5$ the aspect ratios in the range $5.5 \leq M \leq 6.25$ are comparable. As the influence of the initial aspect ratio on the expected ground motion distances is greater for the larger magnitude range, then there is a case for setting an initial aspect ratio of 1.5 in order to get closer agreement between the two software.

4.3 Sensitivity of Curves to Modelling Decisions

The comparisons of the curves for the individual cities at a branch-by-branch level show that for the vast majority of the branches, agreement within $\pm 10\%$ can be achieved for annual probabilities of exceedance as low to 10^{-4} or 10^{-5} . Nevertheless, “perfect” agreement does not seem to be achievable for all of the branches. Broadly speaking, the causes of the disagreements can be split into three categories:

- i) *irreconcilable* discrepancies owing to fundamental differences in the operation of the software. Here we believe that the most influential factors are the use of *swinging arc* distance calculations in the original PSHA software, compared with the fixed grids of points for each zone adopted by OpenQuake discussed previously. We would call this irreconcilable because OpenQuake would not be able to reproduce this behaviour without making fundamental changes to the operation of the code. Other factors that might fall in this category could be the methods for evaluating the density functions of the aleatory variability of the GMM (some software may use simpler approximations to the error function in the Gaussian cumulative density function in order to gain speed), the manner in which the truncated Gutenberg-Richter distributions are discretised, or the way in which distances across the surface of the Earth are calculated (including the underlying assumption of the shape of the Earth’s spheroid).
- ii) *Implementation* discrepancies, which can include actual errors or bugs in the code itself, such as incorrectly implemented GMMs, or incidental discrepancies such as rounding of coefficients or use of 32-bit floating point values.
- iii) *Free modelling parameters/choices* that allow the user to configure the operation of the PSHA calculation in ways that are either not specified in the documentation of a PSHA model. It may be the case that the software may not even allow user control of a particular parameter.

Irreconcilable discrepancies and *implementation* discrepancies may be difficult to identify and will usually require the use of tests such as the PEER Tests (Hale et al., 2018), in the case of the former, or component-by-component quality assurance tests in the case of the latter. Such comparisons are not made in this study. Instead, we focus on the potential *free modelling parameters/choices* that we have either made ourselves in order to best capture the intended behaviour of the model, or for which we have had guidance from the authors but may choose to implement differently in order to gain, for example, computational speed. In this subsection we focus on four parameters that control the size and

orientation of the virtual faults used in the PSHA calculation and show the sensitivity of the hazard curves to these parameters when adopting “reasonable” ranges of values.

The four sensitivity studies we undertake look at the choice of magnitude scaling relation, the initial aspect ratio, the number of depths used to define the triangular hypocentral depth distribution, and the number of strike and/or dip orientations that need to be used to capture the uncertainty in rupture plane orientation. In the following sensitivity studies, we adopt one single branch of the model (source model GTR, magnitude frequency sample branch #5), but shows the impact on the hazard curves for each of the four GMMs.

4.3.1 Sensitivity to Magnitude Scaling Relations

As described in the explanation of the virtual faults used for distributed seismicity sources (section 3), in OpenQuake the area of the virtual fault is related to magnitude via the *magnitude-area* scaling relation. Note that in the present case the standard deviations of the magnitude-area scaling relations are not considered within the calculations, only the median area. There are several available for the user to choose from, and though we don't list all of them here we consider the following for the sensitivity study:

1. **Wells & Coppersmith (1984) – style of faulting dependent (WC1994)**. This model is one of the oldest empirical magnitude-area scaling relations in the literature, but still in widespread use today. This model was selected for use in the OpenQuake implementation (see Table 3.1) and is hence used as the “reference” in this exploration.

$$\log_{10} A (km) = \begin{cases} -3.42 + 0.9 M_W & \text{for strike slip events} \\ -3.99 + 0.98 M_W & \text{for reverse faulting events} \\ -2.87 + 0.82 M_W & \text{for normal faulting events} \end{cases} \quad (4)$$

2. **PEER** (Hale et al, 2018). This model is a simple, idealized scaling relation used for the PEER Tests

$$\log_{10} A (km) = -4.0 + M_W \quad (5)$$

3. **“Point”** – is a null form of magnitude area scaling relation intended to collapse the area to a “true point” for all magnitudes, i.e. $A (km) \cong 0$. This model could be used in cases where one wishes to nullify the influence of the virtual fault properties and take all distances with respect to the epicentre/hypocentre.
4. **CEUS** is a model calibrate for use in the 2011 Central and Eastern United States Seismic Source Characterisation (CEUS-SSC) model (EPRI, 2012):

$$\log_{10} A(km) = -4.366 + M_W \quad (6)$$

5. **Leonard (2014) – Interplate** is one of two self-consistent magnitude scaling relations defined by Leonard (2014). This version is to be applied to inter-plate settings:

$$\log_{10} A (km) = \begin{cases} -3.99 + M_W & \text{for strike-slip events} \\ -4.00 + M_W & \text{for dip-slip events.} \end{cases} \quad (7)$$

6. **Leonard (2014) – SCR** is the other of the magnitude-area scaling relations defined by Leonard (2014). This version is to be applied to stable continental regions:

$$\log_{10} A (km) = \begin{cases} -4.18 + M_W & \text{for strike-slip events} \\ -4.19 + M_W & \text{for dip-slip events.} \end{cases} \quad (8)$$

Figures 21a - 21e show for each target city the sensitivity of the seismic hazard curves from these five models with respect to the those from the selected magnitude area scaling relation of Wells &

Coppersmith (1994). As style-of-faulting is defined for the sources we assume that the style-of-faulting dependent versions of some scaling relations are used, rather than the style-of-faulting independent versions.

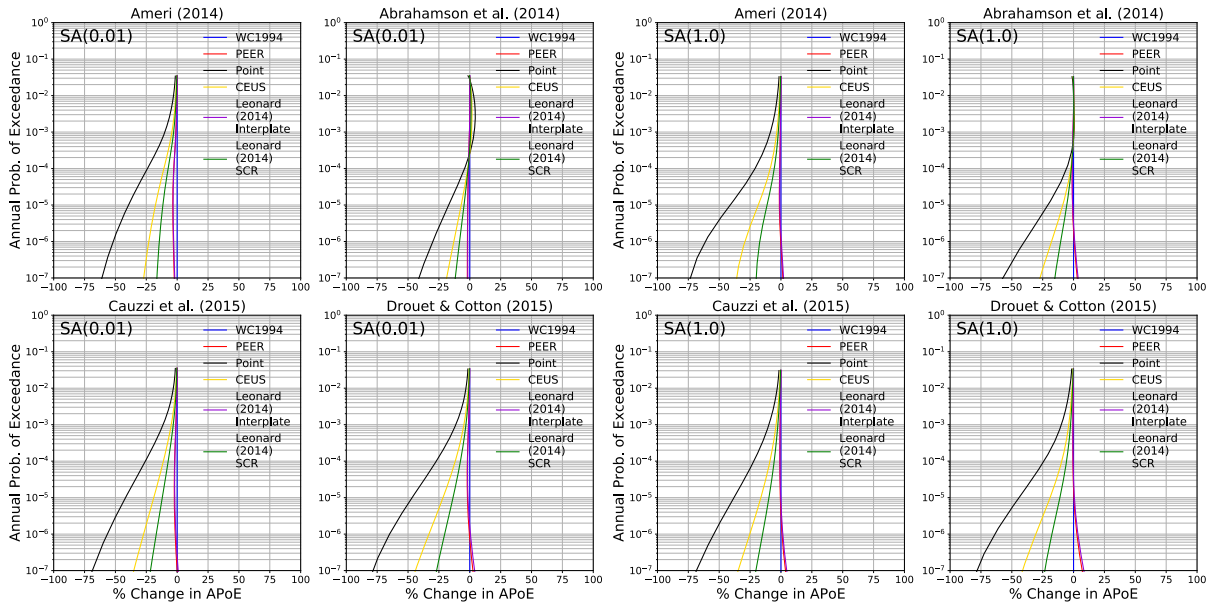


Figure 21a: Sensitivity of the seismic hazard curves for Brest to the choice of magnitude-area scaling relation, with respect to the adopted magnitude-area scaling relation of Wells & Coppersmith (1994). The IMT is $S_a(0.01)$ for the two left-most columns and $S_a(1.0)$ for the right most columns.

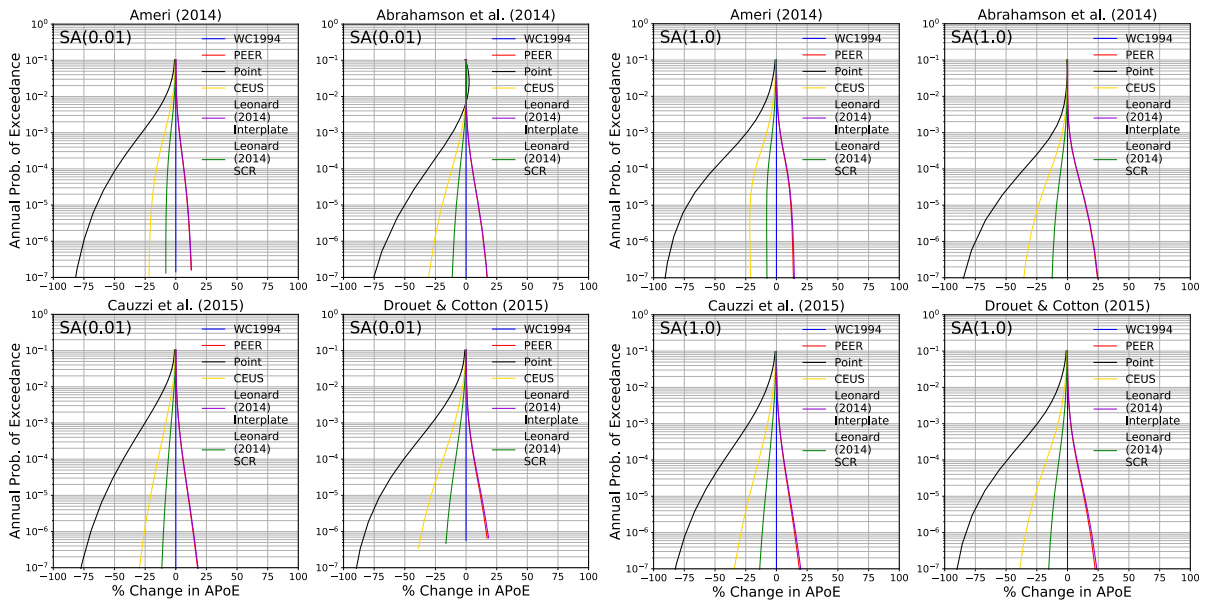


Figure 21b: As Figure 21a for the Lourdes site

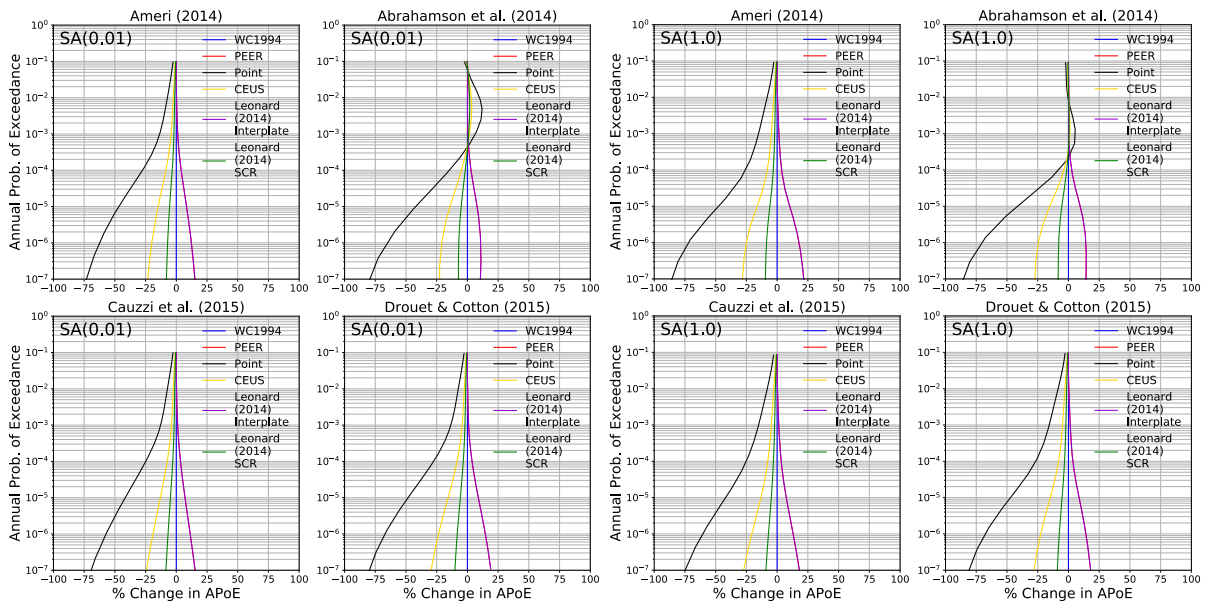


Figure 21c: As Figure 21a for the Marseille site

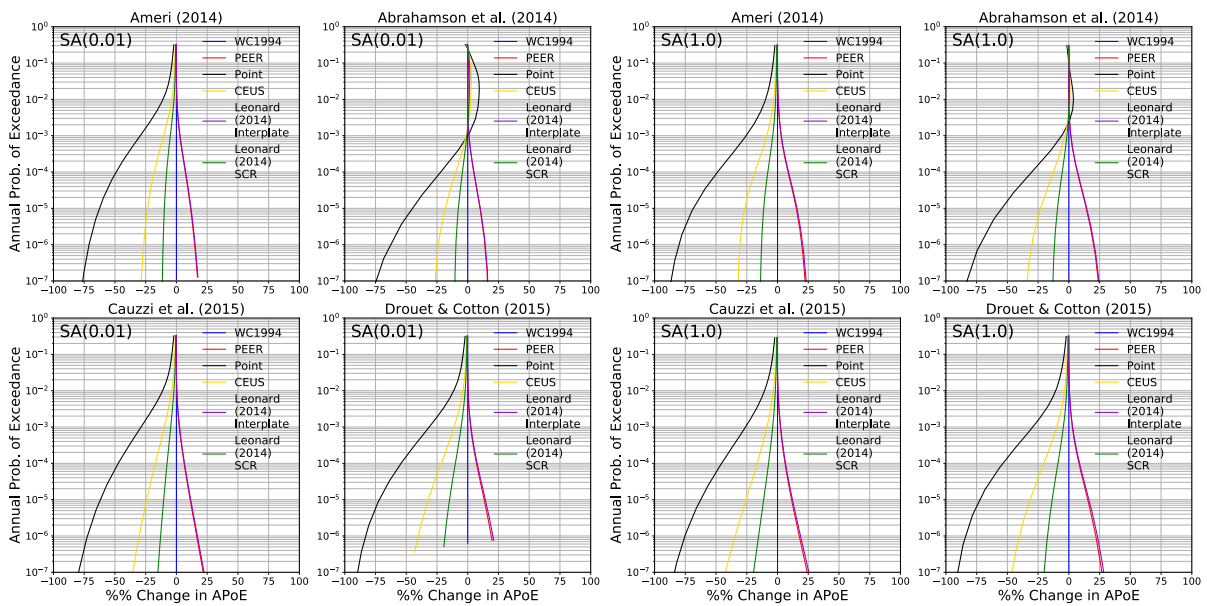


Figure 21d: As Figure 21a for the Grenoble site

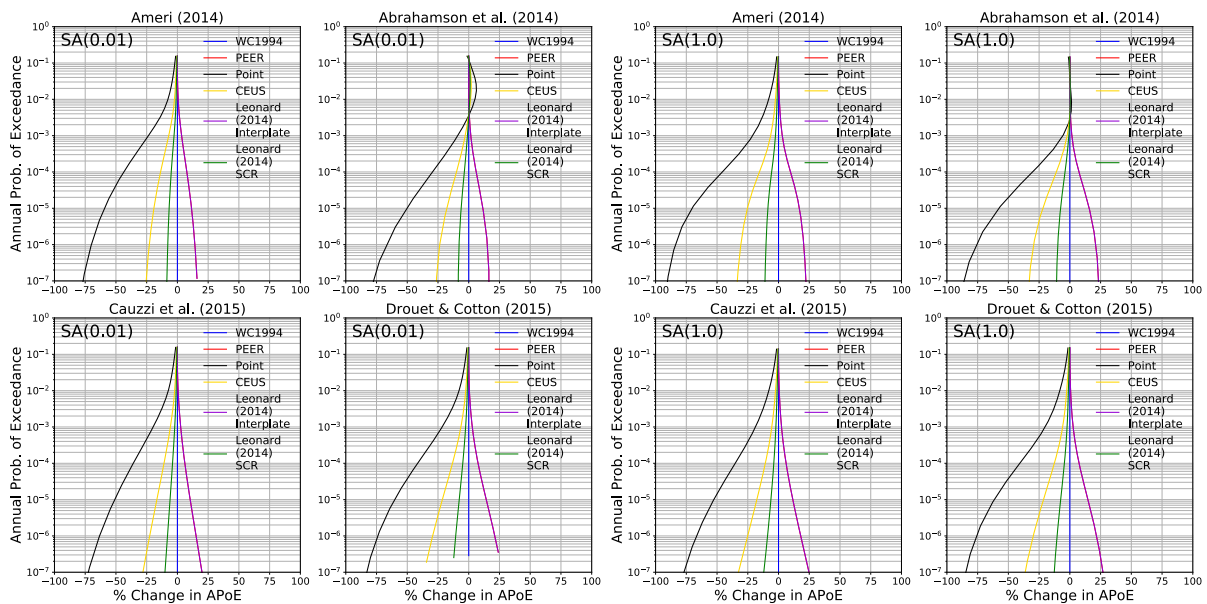


Figure 21e: As Figure 21a for the Nice site

The comparisons of the MSR sensitivity for the five sites reveal some insightful trends and demonstrate that for certain magnitude-area scaling relations differences in seismic hazard on the order of 25 % or more can be seen at longer return periods. The first trend to highlight is that for most of the sites the Wells & Coppersmith (1994) MSR lies in the middle of the range. Neglecting the “Point” MSR case, which will always result in lower seismic hazard curves as source-to-site distances will be systematically longer, the WC1994 model yields higher hazard curves than for those MSRs intended for stable continental regions (CEUS and Leonard (2014) SCR), and lower hazard curves than the models intended for the inter-plate regions (PEER and Leonard (2014) Inter-plate). Overall, the differences remain within $\pm 10\%$ for APoEs greater than 10^{-4} for most sites and GMMs, though this can vary from site to site. There is little obvious difference in sensitivity between short and long spectral periods.

4.3.2 Initial Aspect Ratio

Recalling again from Section 3, in OpenQuake the user can define the *initial* aspect ratio of the virtual faults. We indicate *initial* here as this ratio applies when the vertical extent of the virtual fault is within the defined crustal thickness, and this aspect ratio will be broken at larger magnitude in order to accommodate longer ruptures for larger magnitudes. In the current implementation this has been set to 1, meaning that the ruptures are square until they reach the extent of the seismogenic thickness. For the sensitivity study we compare four values: 0.5, 1.0, 1.5 and 2.0, the results of which are shown for each of the four GMMs and each of the five target cities and GMMs in Figure 22a - 22e.

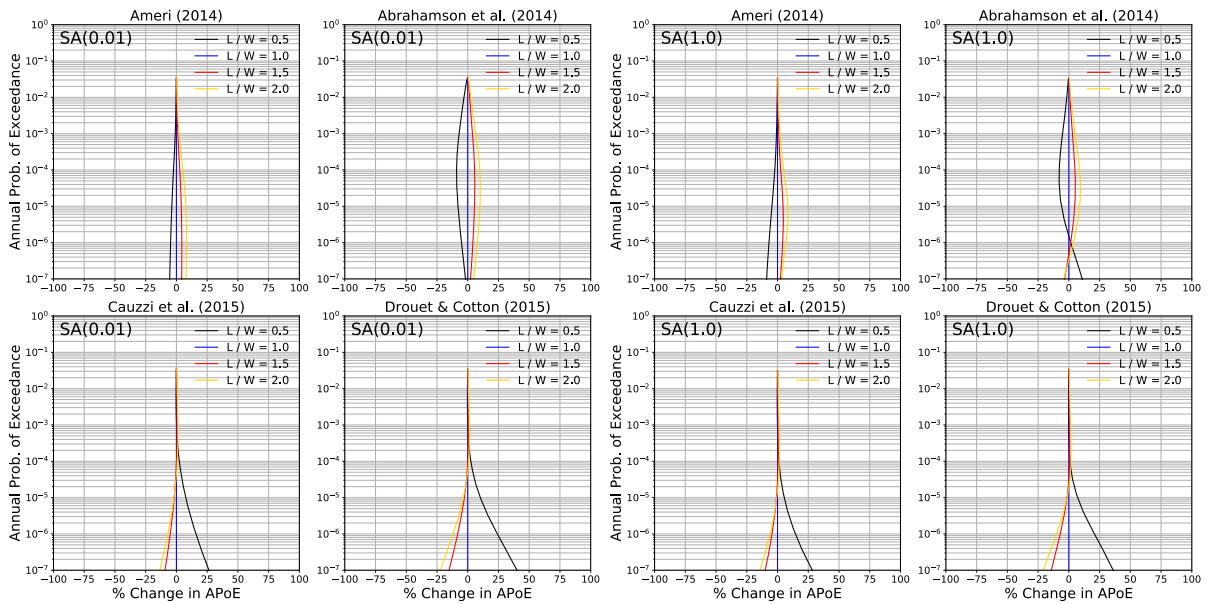


Figure 22a : Sensitivity of the seismic hazard curves for Brest with respect to the choice of initial aspect ratio, assuming an aspect ratio of 1.0 as the reference curve. The IMT is $S_a(0.01)$ for the two left-most columns and $S_a(1.0)$ for the right most columns.

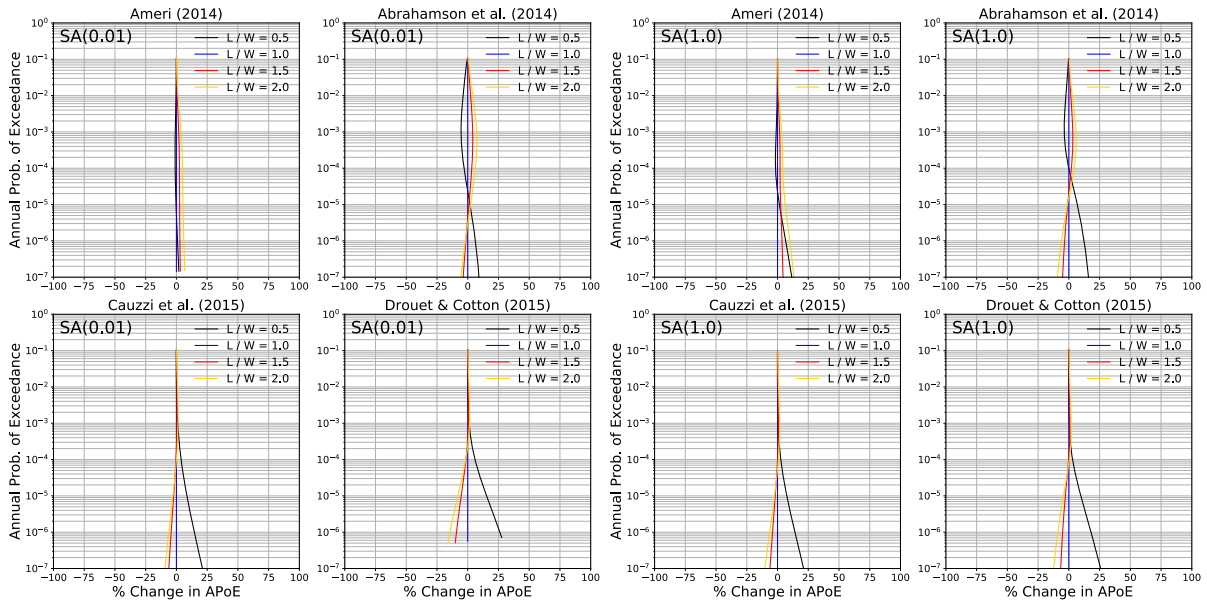


Figure 22b: As Figure 22a for the Lourdes site

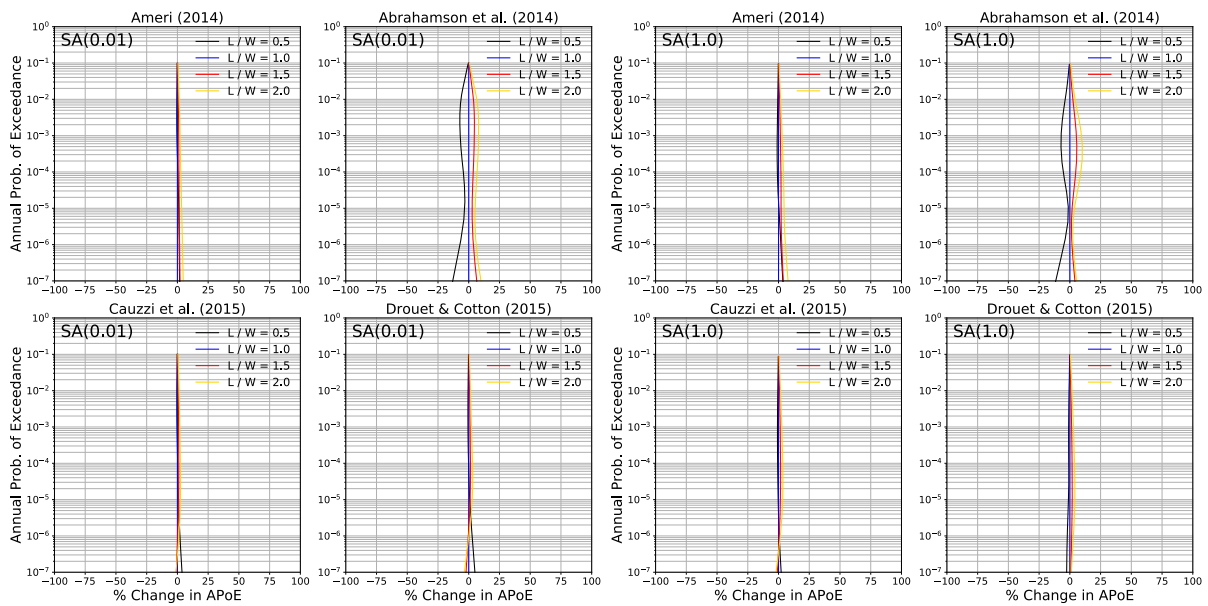


Figure 22c: As Figure 22a for the Marseille site

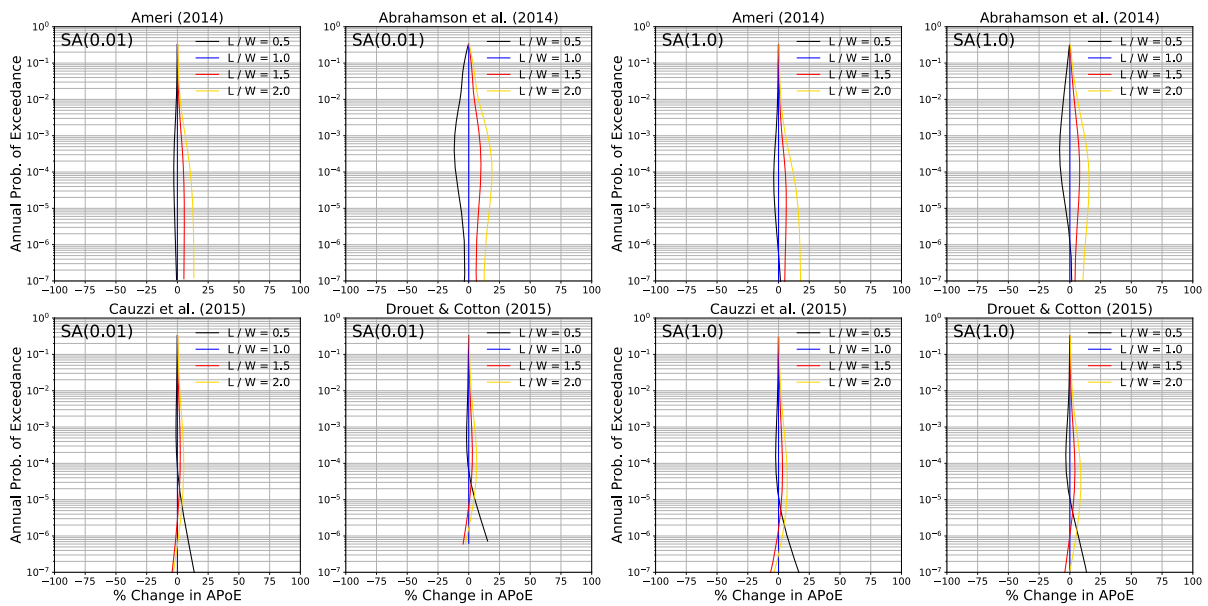


Figure 22d: As Figure 22a for the Grenoble site

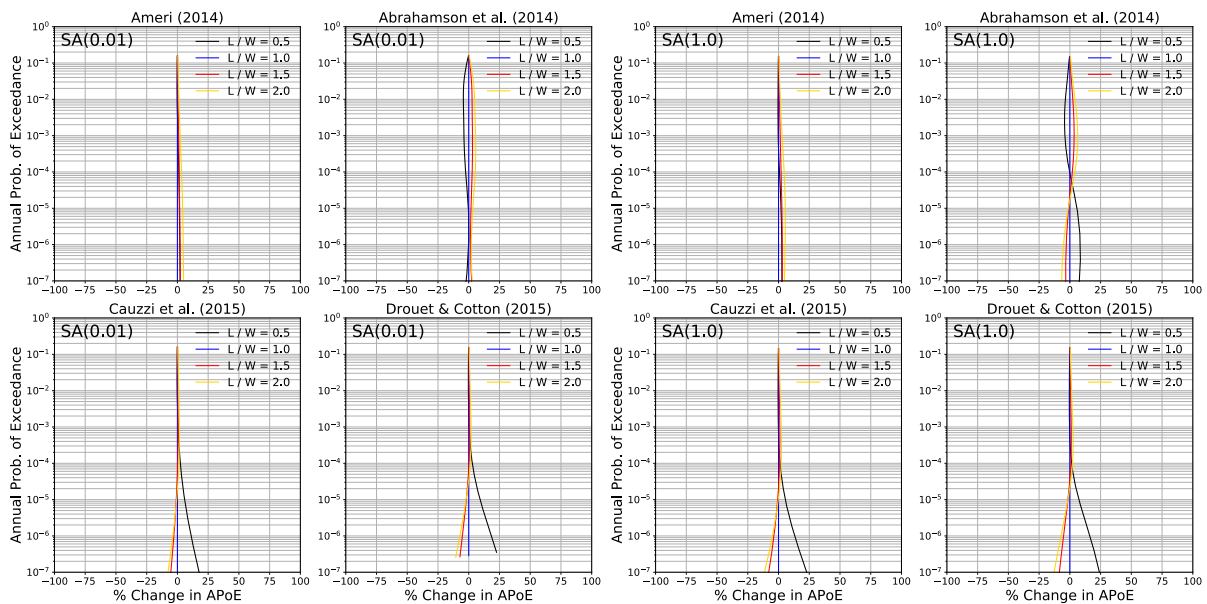


Figure 22e: As Figure 22a for the Nice site

Figures 22a to 22e show that for almost all of the sites and GMMs, the hazard curves for APoEs greater than 10⁻⁴ are generally not so sensitive to the choice of initial aspect ratio. Indeed, for the A14, C15 and DC15 models the differences for are generally only on the order of a few percent, diverging only at the very longest return periods. The hazard curves using the ASK14 GMM show the reverse trend, with sensitivities up to our $\pm 10\%$ “target” tolerance at APoEs in the range 10⁻³ to 10⁻⁵. This is likely to be a manifestation of the hanging wall effects in this model, the near source ground motion may start to become more sensitive to the rupture dimensions at magnitudes $M \geq 5.5$. The curves do, however, confirm that for initial aspect ratios in the middle of the range, e.g., 1.0 to 1.5, the differences in the curves are minimal.

4.3.3 Number of Depth Layers

To define the hypocentral depths and their associated probabilities for the uniform area sources, a triangular distribution with N layers is used (Drouet et al., 2020). The specific number of layers used by Drouet et al. (2020) is not indicated in the paper, although the authors have confirmed that $N = 3$ are used (with probabilities of 0.2, 0.6 and 0.2) respectively. This value has been adopted in the hazard results shown earlier in the section. As this parameter can have a substantial impact on the computation time required, a sensitivity study is useful to determine where are reasonable balance can be struck between the number of layers needed to capture the distribution sufficiently and the associated computational cost.

In the current sensitivity analysis, we consider five different values: $N = 1$ (i.e., a single layer in the centre of the depth distribution), $N = 2$, $N = 3$, $N = 5$ and $N = 7$ layers. The results are shown for the four GMM at each of the five cities in Figures 23a - 23e.

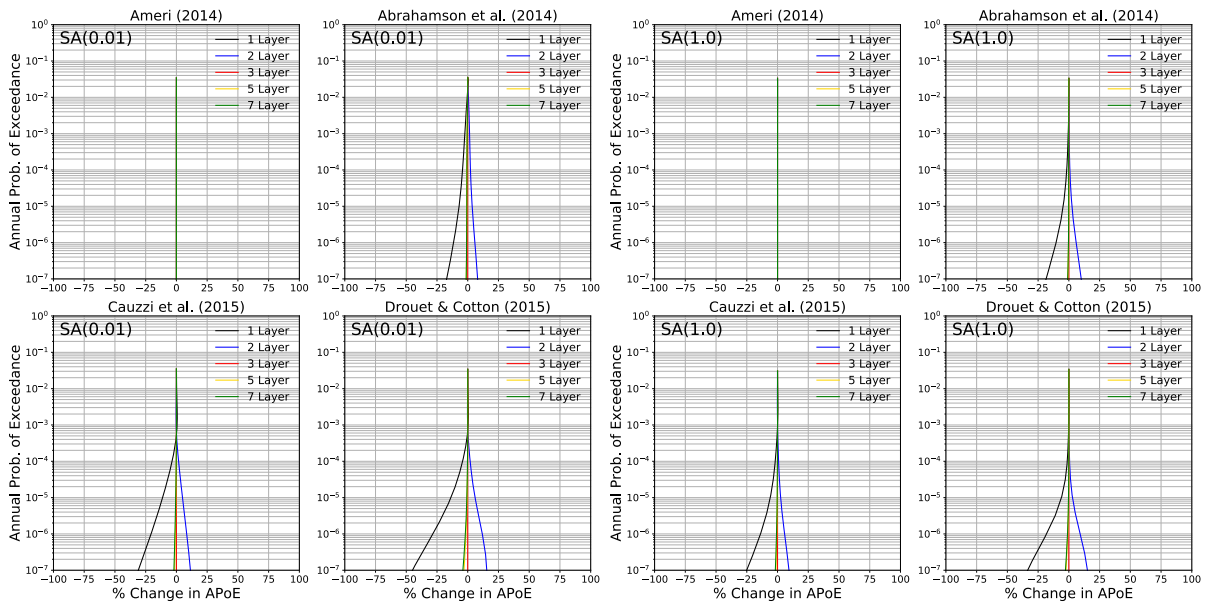


Figure 23a: Sensitivity of the seismic hazard curves for Brest with respect to the choice of number of hypocentral depth layers, assuming $N = 3$ layers as the reference case. The IMT is $S_a(0.01)$ for the two left-most columns and $S_a(1.0)$ for the right most columns.

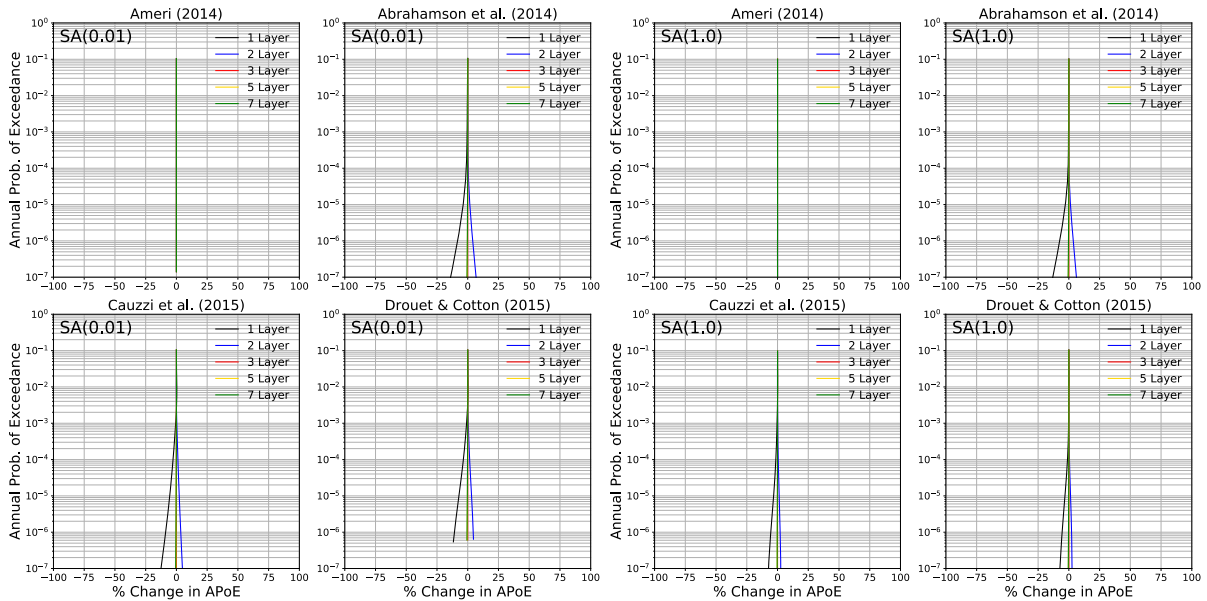


Figure 23b: As Figure 23a for the Lourdes case

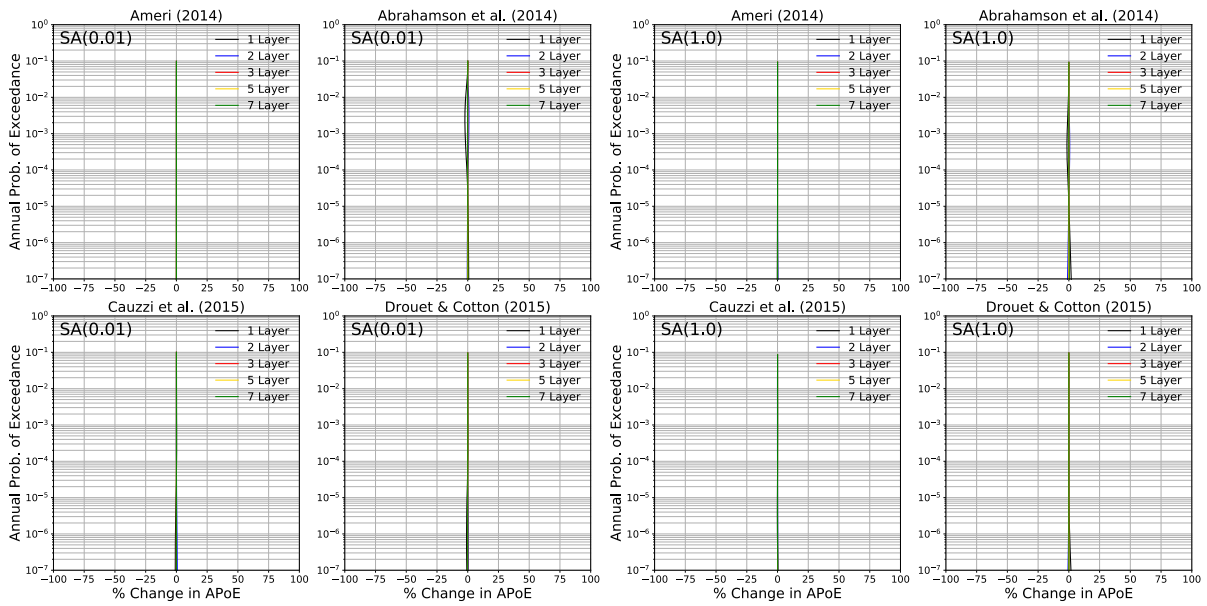


Figure 23c: As Figure 23a for the Marseille case

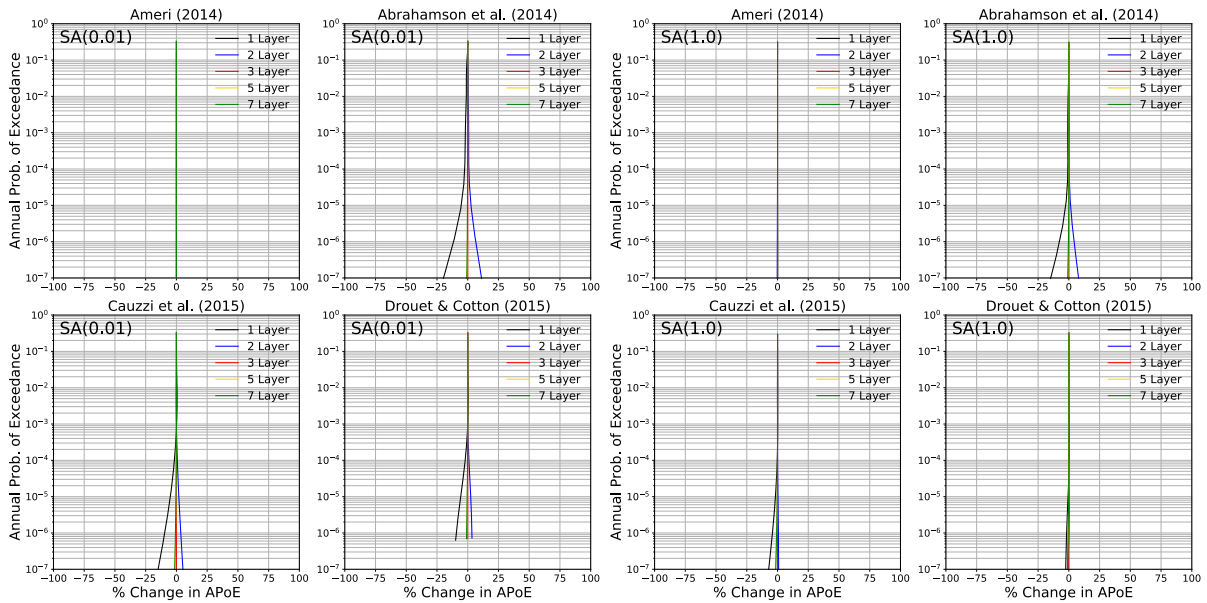


Figure 23d: As Figure 23a for the Grenoble case

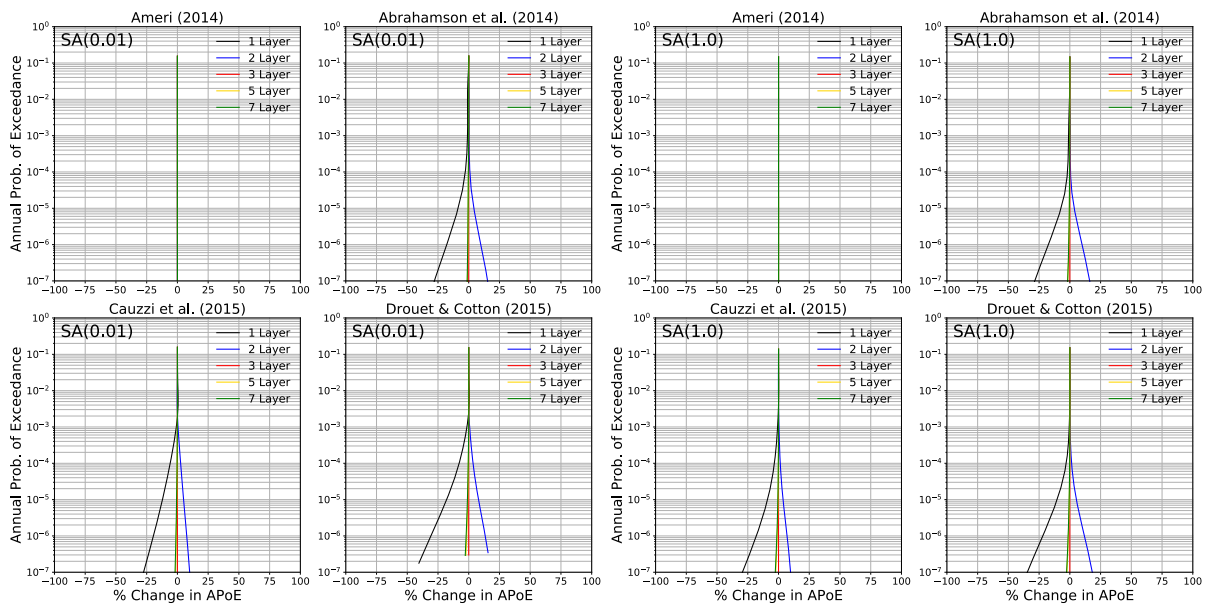


Figure 23e: As Figure 23a for the Nice case

For the number of layers in the hypocentral depth distribution, all sites and all GMMs show the same trend of being mostly insensitive to the number of depth layers for APoEs greater than 10^{-4} , with differences on the order of only a couple of percent. When $N = 1$ and $N = 2$ layers are used, significant differences between the curves begin to manifest at lower APoEs, with $N = 1$ layer tending to under-predict the hazard and $N = 2$ slightly over-predicting it. The degree of divergence even at these low APoEs is site dependent, with Grenoble and Marseille showing less sensitivity to N at low APoEs, while Brest, Lourdes and Nice are a little more sensitivity. What can be seen clearly is that for $N \geq 3$ the curves are virtually identical even at APoEs as low as 10^{-6} . This would suggest that using 3 hypocentral depth layers is sufficient for the calculation, and the results would be largely insensitive to an increasing number of depth layers, meaning that there would be little improvement in hazard for the increased computational cost.

4.3.4 Number of nodal plane orientations

The final parameter being considered in the sensitivity studies is the number of orientations of the virtual faults to be considered in the calculation. Recall from Section 3 that Drouet et al. (2020) indicate in their paper that fault strike varies between 0° and 360° , while dip varies between ranges that depend on the mechanism (80° to 90° for strike-slip faulting, 30° to 60° for reverse faulting and 50° to 70° for normal faulting). Clarification from the authors revealed that they define four random values of strike within the range and one value of dip. In order to capture fully the uncertainty in rupture orientation, one might be tempted to define a large number of strikes and dips. The number of nodal plane orientations, however, is a *quadratic parameter*, meaning that for a doubling of the number of rupture orientation the computational cost quadruples! Therefore, running an exhaustive set of orientations may make the calculation computationally unfeasible.

In this sensitivity study we consider five different configurations of the number of strikes and number of dips: 1) one strike orientation and one dip angle, 2) three strike orientations and one dip (this is our reference case as it is closest to the four strikes and one dip configuration adopted in the previous calculations), 3) three strike orientations and three dips, 4) six strike orientations and three dips, and 5) nine strike orientations and three dips. Where more than one strike or dip orientation is considered, these are evenly spaced across the range, e.g., for strike between 0° and 360° we consider three orientations of 0° , 120° and 240° . The results of the sensitivity study for the four GMMs and five target sites are shown in Figures 24a to 24e.

Author(s) – Implementation of the Drouet et al. (2020) PSHA for France into OpenQuake: Comparisons and Modelling Issues- SIGMA2-2021-D5-085

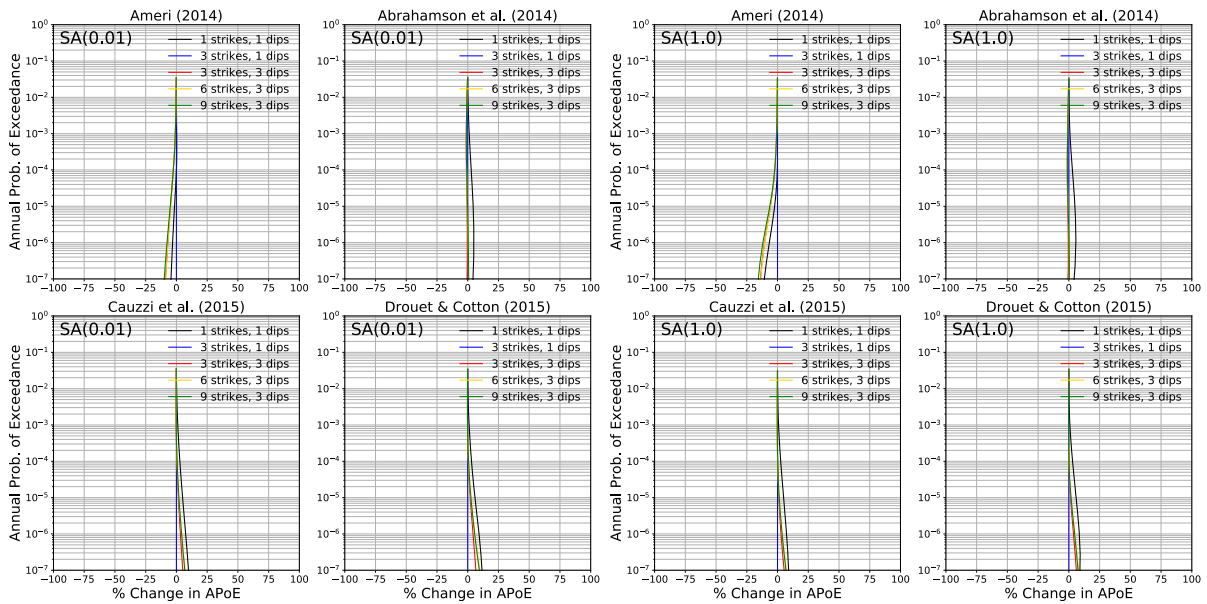


Figure 24a: Sensitivity of the seismic hazard curves for Brest with respect to the choice of number of nodal plan orientations, assuming three strikes orientations and one dip angle as the reference case. The IMT is $S_a(0.01)$ for the two left-most columns and $S_a(1.0)$ for the right most columns.

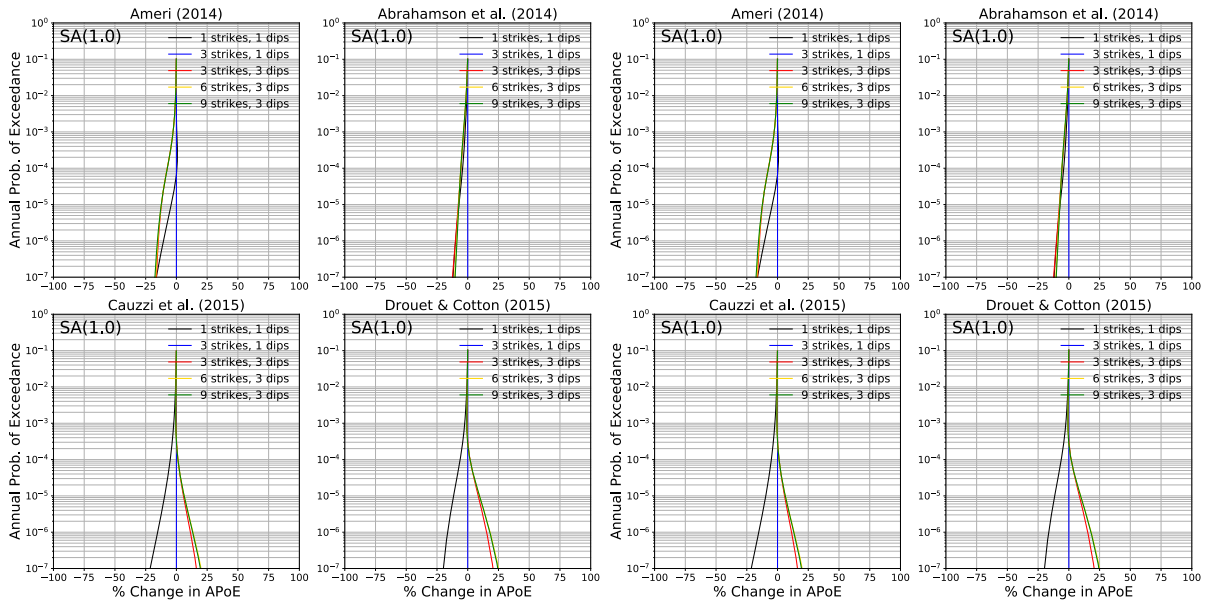


Figure 24b, As Figure 24a for the Lourdes case

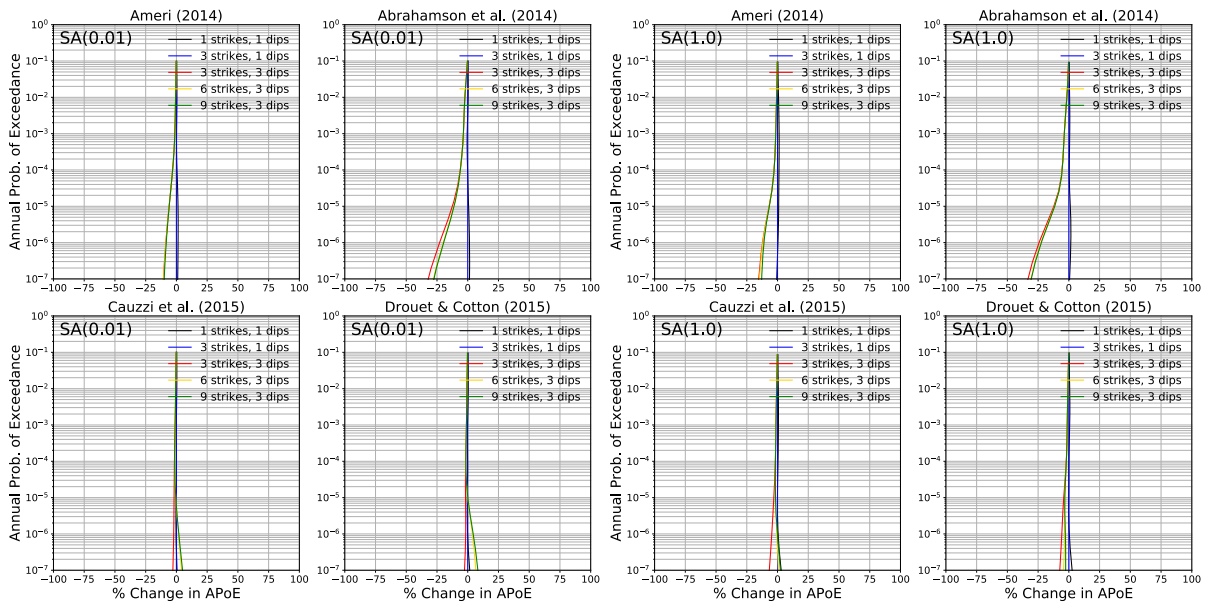


Figure 24c: As Figure 24a for the Marseille case

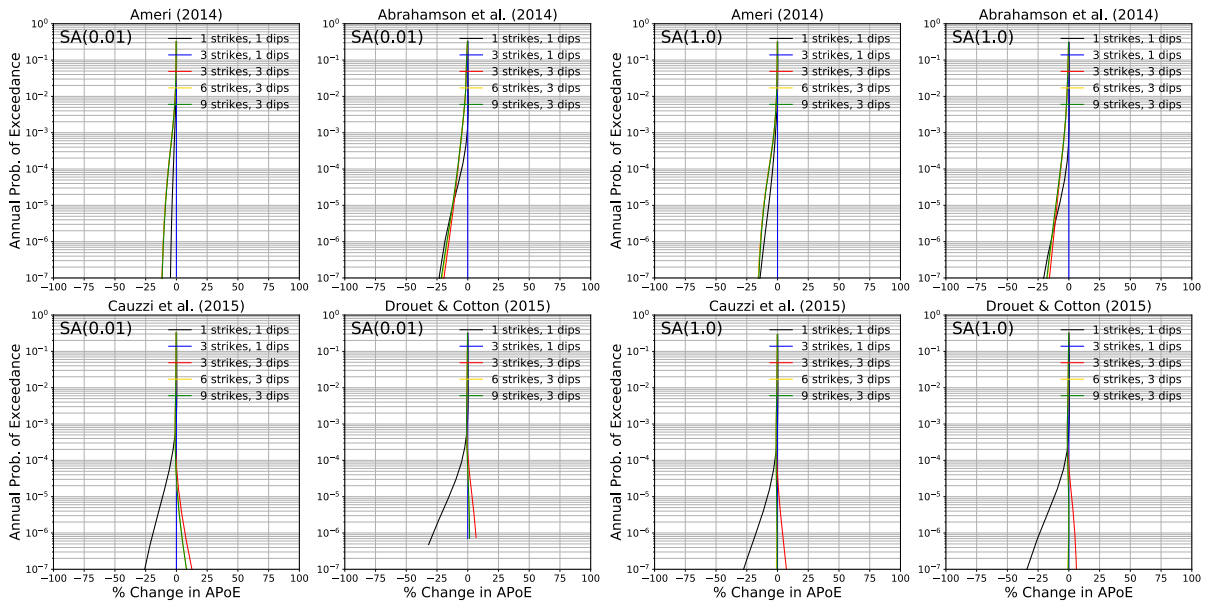


Figure 24d: As Figure 24a for the Grenoble case

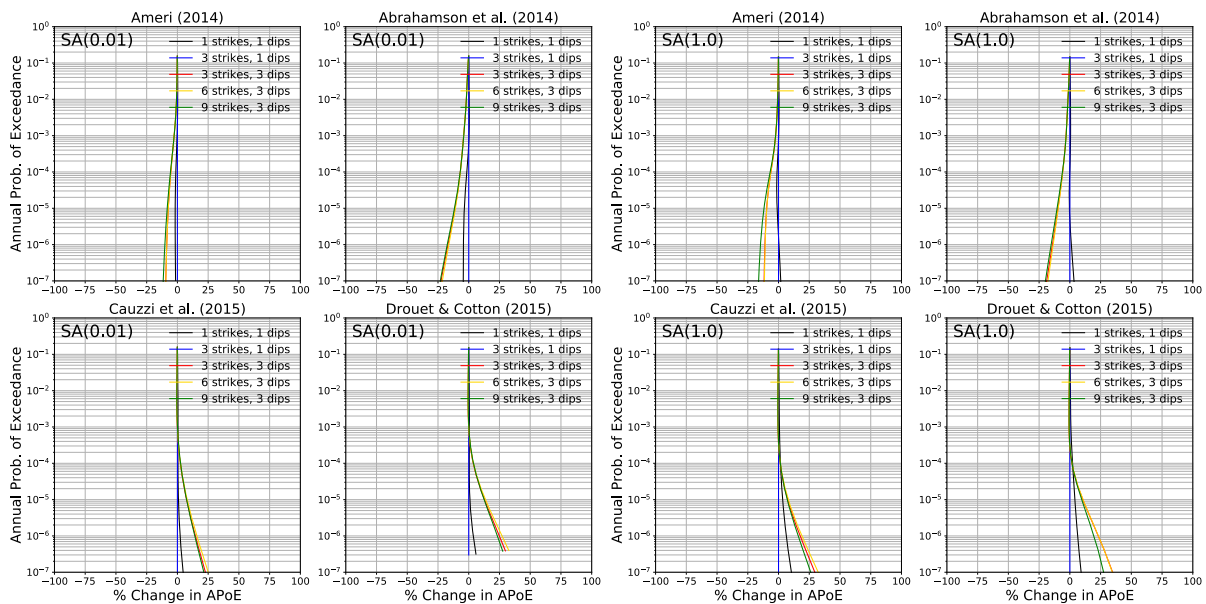


Figure 24e: As Figure 24a for the Nice case

The sensitivity of the seismic hazard curves to the number of nodal plane orientations is notably greater than that of initial aspect ratio and number of depth layers. In addition, unlike the sensitivity to magnitude-area scaling relation, depending on the site and the GMM the curves can be more systematically biased toward higher or lower APoEs if too few orientations are considered. For APoEs greater than 10^{-4} it would generally seem to be the case that sensitivity to the number of orientations is within the $\pm 10\%$ range, albeit close to the limits of this range in some cases. For lower APoEs the curves can diverge significantly, with differences as high as $\pm 25\%$ at APoE 10^{-6} in some cases. Generally speaking, the curves become less sensitive to the number of strike values when at least three or more strike orientations are considered. The greatest change, however, seems to occur when increasing the number of dip orientations from one to three. This would seem to suggest that the combination of four strike orientations but only one dip orientation that have been adopted by the original authors (and considered here for the previous results) is not necessarily a “stable” configuration. Ideally, at least three strike and three dip orientations would be needed, though perhaps the latter may vary depending on the predominant style-of-faulting in a region.

4.4 Implications of the comparisons of the sensitivity studies

The set of sensitivity studies for the four model “free” parameters has been helpful to identify where these often-under-reported modelling choices can have a significant influence on the results. In the majority of cases, the impact of varying the parameters within the reasonable ranges of values did not result in exceeding the agreement criteria of $\pm 10\%$ for APoE $\geq 10^{-4}$. However, for some individual parameters the variability came close to this threshold, and it is not unreasonable to suggest that when taken together a set of different *but seemingly reasonable* modelling choices would prevent the modeller from reaching this specific level of agreement on a branch-by-branch basis. These illustrate the need for comprehensive documentation of seismic hazard models, including seismic hazard code configuration settings. We will return to this point in the concluding chapter

5. Comparisons and Sensitivities – Mean and Quantiles

The branch-by-branch comparisons shown in the previous chapter suggest that though some hazard curves for some individual branches may not agree within the desired criteria, these are a minority of cases. Furthermore, when looking at the complete suite of curves it is difficult to discern too much systematic bias that could lead one to believe that OpenQuake tends to produce higher or lower hazard than the original software. For this we need to compare the mean and quantiles of the seismic hazard curves.

As noted in section 3.7, and will be discussed in due course later in this section, when evaluating the logic tree OpenQuake retrieves the weighted arithmetic mean of the probabilities of exceedance (annual probabilities in the present case). We discussed with the original model developers whether the original PSHA software adopts this same definition of the mean and quantiles, and they confirm that the original code also returns the weighted arithmetic mean hazard curve.

5.1 Running Calculations with the Full Logic Tree

In Section 3 we outlined one of the challenges in running the calculation, which is that by representing the 100 MFD branches as alternative source models we incur an excessive amount of computation. To run the model, we have been using a server with 120 cores and 1 Tb of RAM. OpenQuake's parallelisation strategies are extremely efficient as it splits the calculations into subsets of ruptures, whose resulting distance and ground motion calculations are then distributed across the number of available (or assigned) cores. For site-specific seismic hazard calculations, or calculations for a small number of sites located close to one another, the maximum distance filter helps to limit the computations to consider only ruptures within the maximum source to site distance. As this is set to 200 km in the present case, a significant majority of all of the possible ruptures generated by the model are filtered out and the computation can become feasible within a reasonable period of time. At a national scale, however, OpenQuake cannot filter these ruptures out as almost all with fall within the distance range of some sites. Each realisation of the source model will generate tens of millions of ruptures within its earthquake rupture forecast, all of which will need to be processed. In the first iteration of the full model, we found that it was not possible to run the calculation with the available resources of our server. We have therefore made two adjustments to the way in which the calculations are run, one using an available feature of the OpenQuake calculation engine, the other requiring an unconventional approach to running calculations. In the site-specific examples that will be shown later in this section, we do not apply these adjustments, except where shown for comparison purposes. For running the full calculation at a regional scale, we do apply the adjustments, and these will be discussed more in due course.

Point Source Distances

The “point source distance” is an optional optimisation strategy that OpenQuake can apply to distributed seismicity sources, such as the gridded seismicity or uniform area sources being used here. As explained in section 3, for such sources the earthquake itself is defined not as an actual point but as a finite rupture from which the relevant source-to-site distances are computed. When many hypocentral depths and/or nodal planes are considered for each source, this results in a massive number of finite ruptures. However, for sources located far away from the target site, the impact of the rupture finiteness on the resulting hazard curve diminishes significantly, particularly for low to moderate magnitudes whose rupture planes may only be a kilometre or less in terms of length or width. In these cases, the impact of the hypocentral depths or the nodal plane orientations may be absolutely minimal and could be approximated by an “average” plane, one whose orientation and depth would correspond to the weighted mean of those specified by the probability mass function, without significant impact on the seismic hazard. The “point source distance” parameter allows the user to apply this approximation to point sources that are further away from the sites than the distance the user indicates. The “point source distance” parameter can be a single fixed value, or it can be magnitude dependent. The magnitude-

dependent option means that users can apply the correction at shorter distances for small magnitudes and increase the distance for larger magnitudes. This is favourable in many cases as it ensures that the same source can use the average rupture plane for small magnitudes where the impact would be trivial, yet still ensure that the full distribution of rupture depths and orientations are maintained for larger events, where this could impact more on the seismic hazard.

The use of the point source distance can offer a significant speed up to the calculations, especially when considering a single site or small number of closely located sites. As this adaptation of the calculations is not only specific to OpenQuake but also a relatively recent addition, little guidance exists as to how best to configure the parameters, particularly if choosing to adopt a magnitude-dependent version. After several iterations comparing the sensitivity of seismic hazard curves from single source models, we adopt the magnitude-dependent “point source distance” scheme in Table 3 and Figure 25.

Table 3: Point Source Distances

Magnitude	Point Source Distance (km)
4.5	30
4.75	40
5.0	50
5.25	62
5.5	75
5.75	88
6.0	100
6.5	150
7.0	200
7.5	200

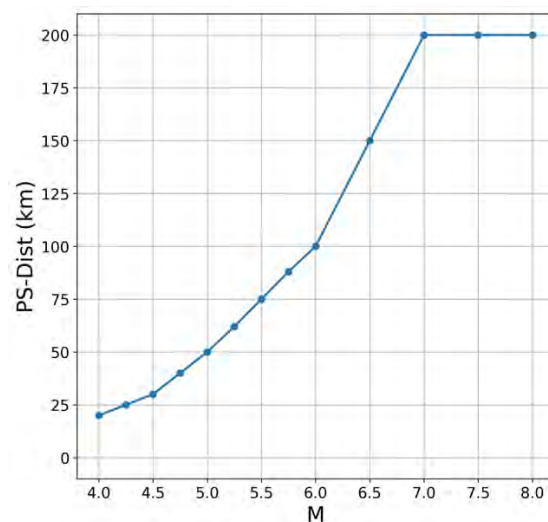


Figure 25: Scaling of Point Source distances with magnitude

Using the configuration shown, initial experiments running the entire logic tree for single sites yielded almost a 75 % reduction in computational time, with a tolerable impact on the curves. Figure 26 compares a set of hazard curves for the Grenoble test site and IRSN source model on one side using the full distribution of finite faults and the other using the “point source distance”. Though differences in the curves are visible, the degree of mismatch between the original and the OpenQuake curves is still within the acceptable threshold.

Running the calculations

To run the complete logic tree for all of the 6,836 sites, we found it necessary to split the calculations in order to avoid exceeding the RAM capacity of our server. We did this by splitting the logic tree into 10 smaller subsets of 160 logic tree branches, each of which contained the four source models, four GMMs but only 10 MFD branches. After completion of the 10 calculations, the resulting hazard curves for all sites, intensity measure types and logic tree branches were retrieved from OpenQuake’s datastore file (an hdf5 binary file that OpenQuake uses for storing calculation information in order to allow later extraction of results) and then combined and re-weighted into a single hdf5 binary file that contains all of the seismic hazard curves. From this combined binary file, we post-process the results using a customised Python script that uses the same functions taken directly from OpenQuake that the software would use to extract the mean, quantiles and UHS from the calculations. A small Python module that

allows users to extract their own desired results from this large output data file was prepared and this will be made public upon release of this deliverable along with the data itself.

The complete calculation provided seismic hazard curves for 16 intensity measure types: PGA and spectral acceleration at periods: 0.01 s, 0.03 s, 0.05 s, 0.075 s, 0.1 s, 0.15 s, 0.2 s, 0.25 s, 0.3 s, 0.4 s, 0.5 s, 0.75 s, 1.0 s, 2.0 s, 3.0 s. Results are shown only for the Sa 0.01 s, 0.2 s (not for the city scale curves) and 1.0 s case.

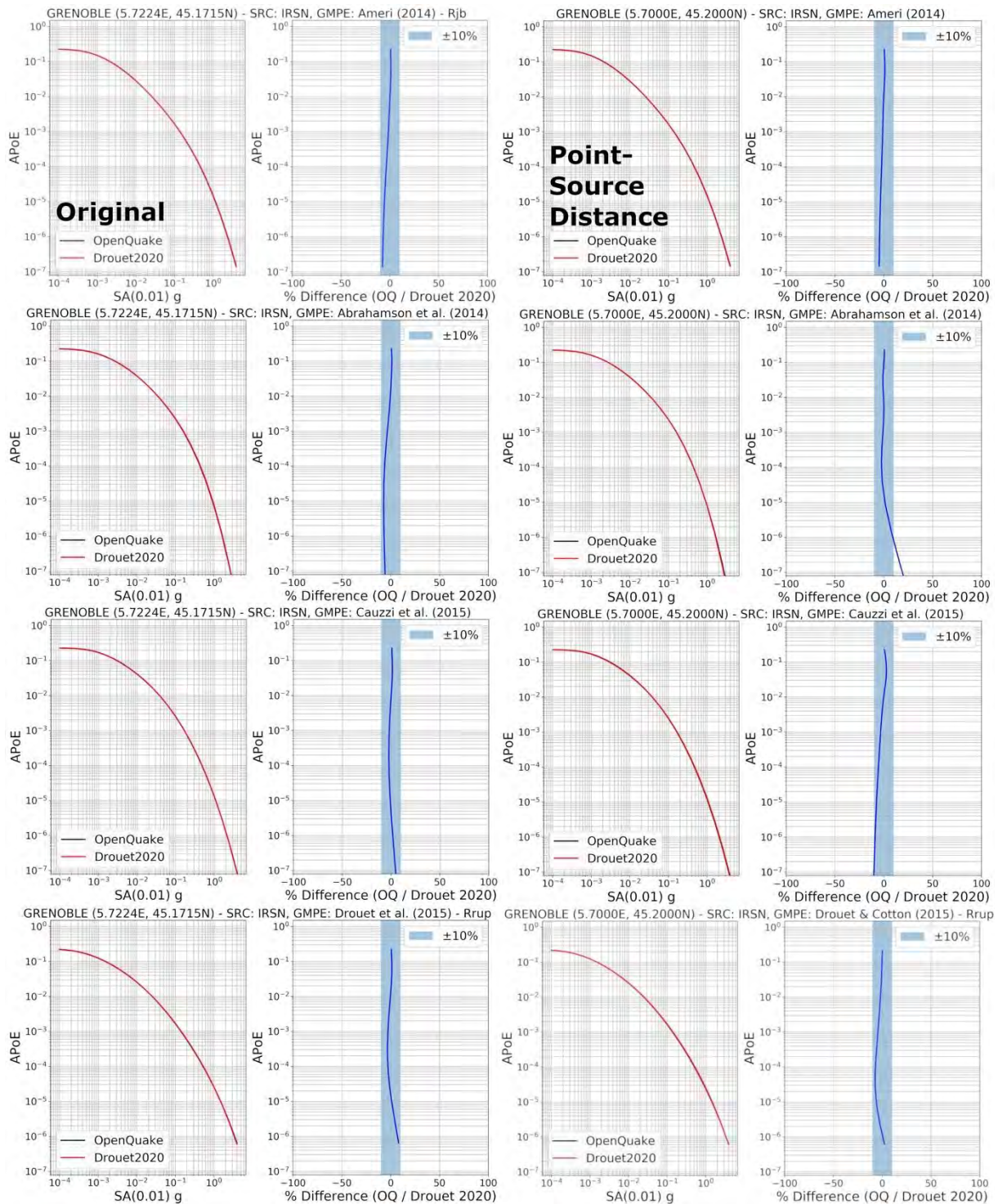
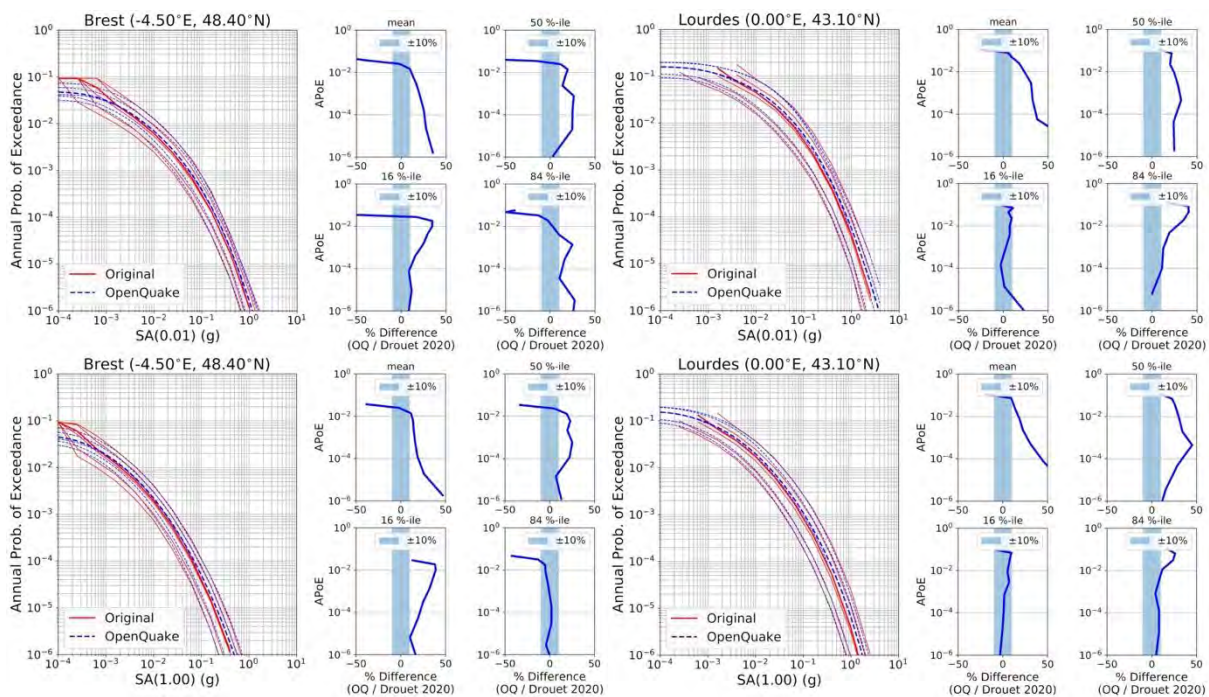


Figure 26: Impact of the « point source distance » approximation on seismic hazard curves for the Grenoble sites using the IRSN source model. Left columns show the comparisons between the original and OpenQuake

implementation when the full finite rupture distribution is used, and the right columns show the same comparison when the "point source distance" is used.

5.2 Mean and Quantile Comparisons for the Selected Cities

Figures 27 shows the comparisons of the mean and quantile seismic hazard curves for each of the five target cities (Brest, Lourdes, Marseille, Grenoble and Nice). The 5th, 16th, 50th (median), 84th and 95th percentiles are calculated as quantile set, though difference curves are only shown for the mean, 16th, 50th (median), 84th percentile cases. As with the branch-by-branch results, our target degree of mismatch is $\pm 10\%$ for APoE greater than 10^{-4} ; however, in the figures it is evident that some processing artefacts have affected the calculation of the original curves at the very highest APoEs. We would therefore revise our target to a mismatch of $\pm 10\%$ for APoEs in the range 10^{-2} to 10^{-4} (100 to 10,000 year return period).



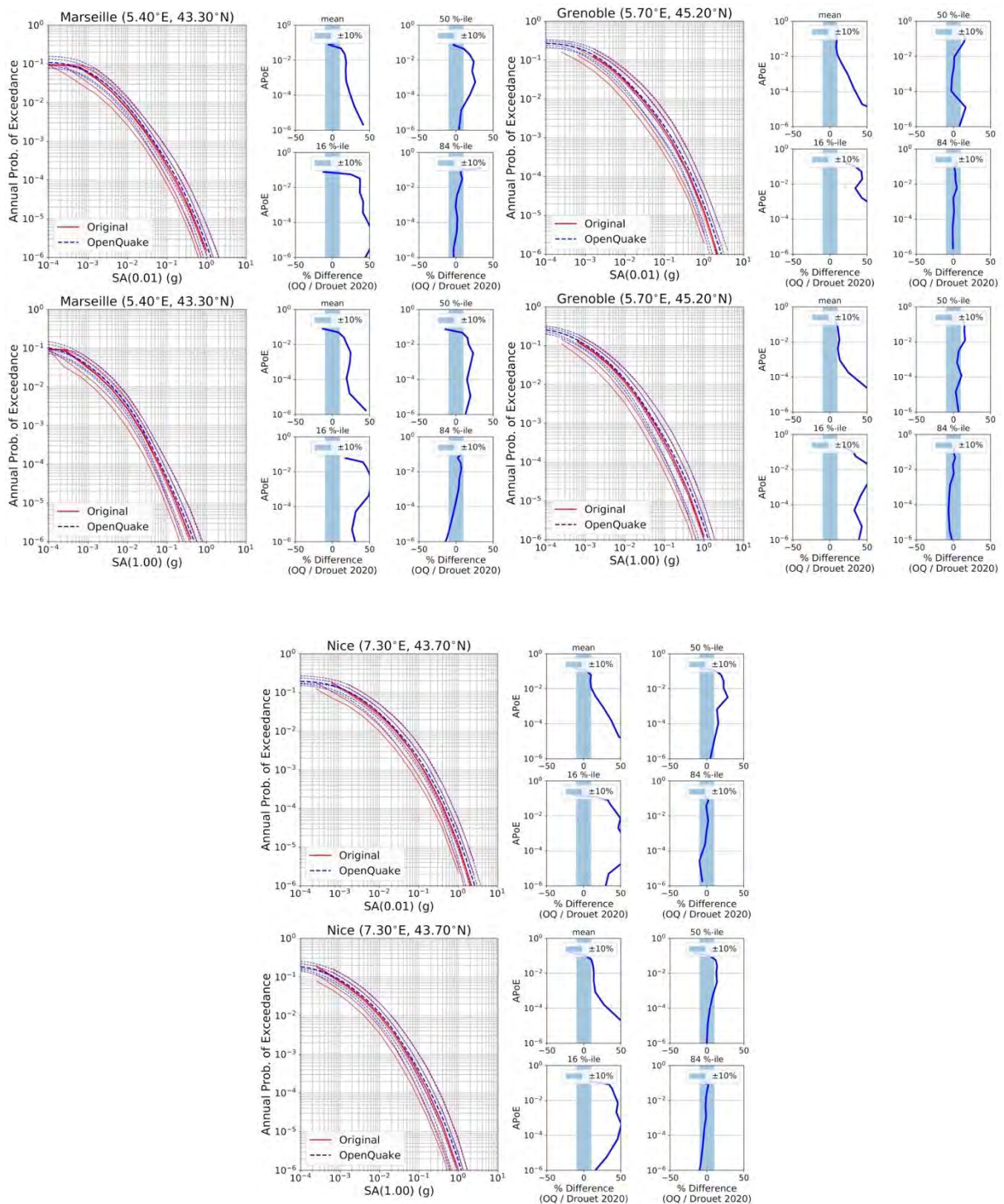


Figure 27: Seismic hazard curve comparisons and percentage differences between the mean and 16th, 50th and 84th percentiles from the OpenQuake implementation and original implementation – for Brest (top row, left), Lourdes (top row, right), Marseille (middle row, left), Grenoble (middle row, right) and Nice (bottom).

The most striking result of these comparisons is that for these five target sites we seem to be unable to achieve the target level of agreement for the mean and *some* (but not all) of the quantiles. Though the specific degree of mismatch varies from site-to-site we find that the mean curves from OpenQuake are generally greater than those coming from the original software. This was a surprising result as the actual branch-by-branch comparisons had not shown systematic trends on this scale. Indeed, if any trend had been visible then it was for a slight under-prediction of the APoEs for certain source and ground motion

models than that of the original software. However, these results also provide a slightly more perplexing picture. Though the agreement is poor between the curves from the mean and lower (5th and 16th) percentiles, the curves are generally in better agreement for the upper (84th and 95th) percentiles. The degree of agreement in the median varies from poor to acceptable, depending on the site in question.

The degree of disagreement in the mean and quantile curves, when the actual branch-by-branch results had shown reasonable agreement, warranted further investigation. Initially we had suspected differences in the computation of the quantiles; however, further discussions with the authors indicated that the same approach for computation of these terms is adopted for both software. The insights for this discrepancy come when we compare the distributions of the full set seismic hazard curves for each of the four source models. This is shown for the Grenoble site in Figure 28, and here a critical source of discrepancy becomes apparent. The centre and range of the curves for the EDF, GTR and IRSN sites are in good agreement between the original software and the OpenQuake implementation. For the ZONELESS model, however, the OpenQuake curves show a much narrower range, diverging only at the very lowest APoEs.

For the Grenoble case shown in Figure 27 the ZONELESS model produces generally lower seismic hazard results than the area source models. As the current OpenQuake results are not capturing the full range of values of seismic hazard for these sets of branches, it is missing the branches that would effectively fill out the lower end of the quantile range. As the upper end of the quantile range is coming from the area source models that are in good agreement, we are able to achieve the target discrepancy threshold for the 84th (and 95th) percentiles. Without the branches on the lower end, however, the mean is over-estimated and the lower quantiles more so. That the variable degree of agreement in the median curves from site-to-site is likely a reflection of the degree to which the suites of curves from the ZONELESS model are lower than for the uniform area sources. Indeed, the degree of mismatch in the 16th percentiles is significantly lower for the Brest and Lourdes sites than for Marseille, Grenoble and Nice. In the Brest and Lourdes case the cities are located close to regions of higher activity rate in the zoneless model, while the other cities lie in places where the local grid points show lower activity yet the overall regional seismic activity (as implied by the area sources) is higher.

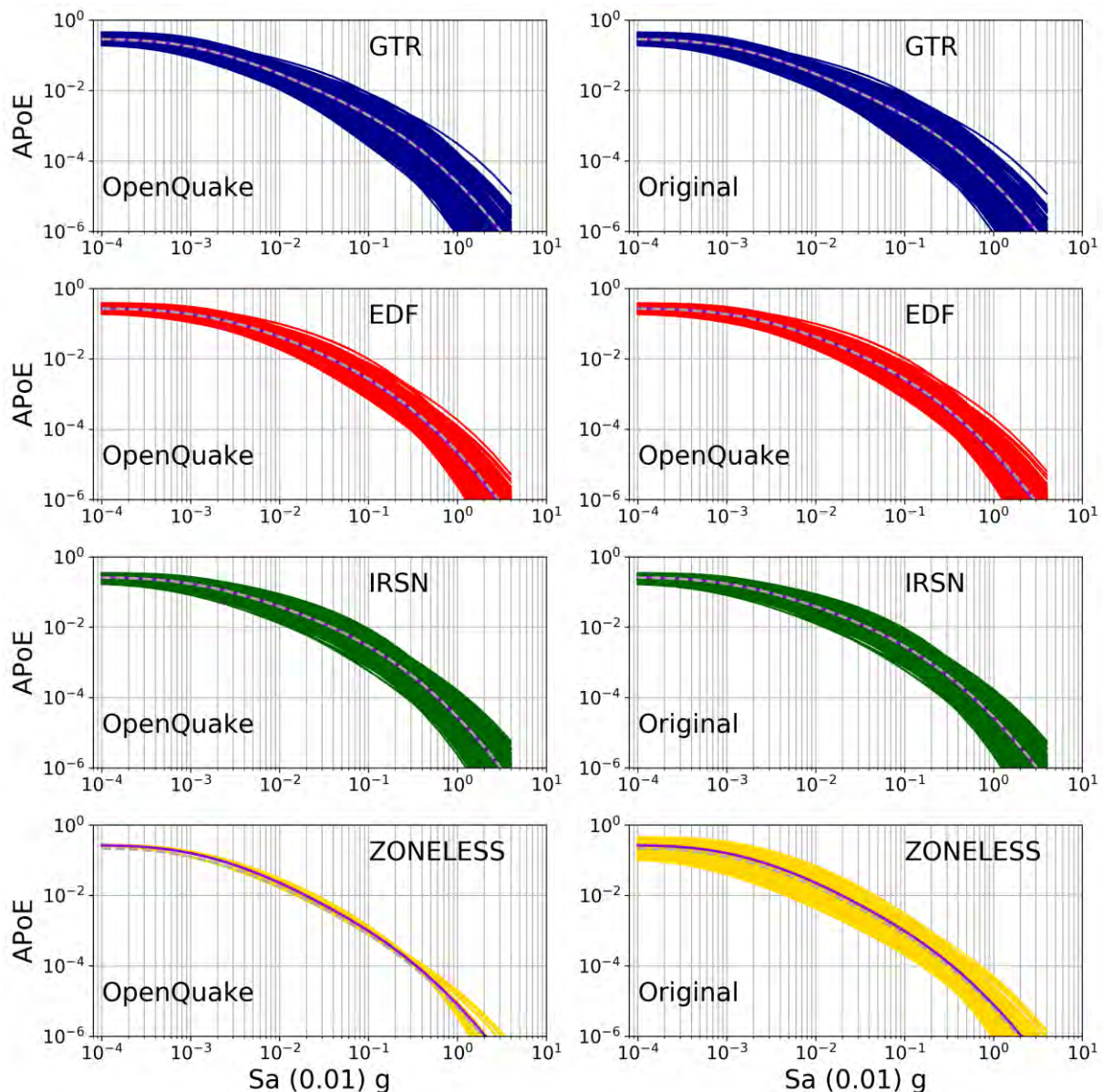


Figure 28: Comparison of the range seismic hazard curves for each of the seismic sources (EDF, GTR, IRSN and ZONELESS) for the Grenoble site. Grey dashed curves indicates the mean curve from the original software, while the purple curve indicates the mean from the OpenQuake implementation

Given that a clear discrepancy has been identified, and that this discrepancy *can* explain the disagreement we are seeing in the mean and quantile curves, what is the cause of this? Here we believe there is missing or omitted information in the literature provided by the authors. The logic tree presented in Figure 20 of Drouet et al. (2020) (reproduced as Figure 8 in this report) indicates that though uncertainties on a , b and M_{MAX} are considered for the area source models (randomly sampled in the manner shown in section 2 here) only the uncertainty on the M_{MAX} is sampled for the zoneless model. The data files we received from the authors also seemed to confirm this case, as sample values were provided for M_{MAX} for the zoneless model but not the a and b values. The OpenQuake curves for the zoneless model shown in Figure 28 reflect this with their divergence at the lower APoEs, but converging to a fixed point at the higher APoEs (i.e., a constant activity rate for each branch). The curves from the original software, by contrast, show a clear variability in activity rate that is illustrated by the range of APoEs for the lowest intensity measure levels.

Further consultation with the authors revealed that, in contrast to the description of the logic tree in the Drouet et al. (2020) paper and as presented in Figure 6, uncertainty on the a and b values is also considered in the zoneless model. The exact manner in which this is done is described as follows. Each grid cell i within the zoneless model is associated with an a value (a_i) and b value specific to the region (from the DOMAINE model) in which the cell is located $b_{i|R}$. In the smoothed seismicity files provided by the authors, both are associated with their respective uncertainties $\sigma_{a,i}$ and $\sigma_{b,i|R}$. For each of the $j = 1, 2, \dots, 100$ branches that are used to account for the uncertainty in the magnitude frequency distribution, the uncertainty in a and b is sampled at every site such that $a_{i,j} = a_i + \varepsilon_{i,j} \cdot \sigma_{a,i}$ and $b_{i,j} = b_{i|R} + \varepsilon_{i,j} \cdot \sigma_{b,i|R}$, where $\varepsilon_{i,j}$ is a random variate sampled from a standard normal distribution $\mathcal{N}(0,1)$ truncated between ± 3 standard deviations. A critical point to note here is that $\varepsilon_{i,j}$ is sampled independently for each grid cell; however, the same $\varepsilon_{i,j}$ multiplies both the $\sigma_{a,i}$ and $\sigma_{b,i|R}$ values, thus implying a multivariate distribution with perfect correlation in $\sigma_{a,i}$ and $\sigma_{b,i|R}$.

Unfortunately, whereas we had been able to retrieve from the authors the sampled values of M_{MAX} used in the original calculations, and the sampled a and b values for the uniform area sources, it was not possible to retrieve the exact samples of $a_{i,j}$ and $b_{i,j}$ for the ZONELESS model as this process was inbuilt into the code and not exported. Nevertheless, based on the descriptions given by the authors we repeated the construction of the source models, this time replicating the sampling process described above. Figure 29 shows the same comparisons of the curve distribution for the ZONELESS branches at the Grenoble site, as presented in Figure 28, this time accounting for sampling of $a_{i,j}$ and $b_{i,j}$. Here we can now see that the distribution of the hazard curves is comparable to those of the original model, which would suggest that the sampling process can be adequately reproduced even if the exact values are not retained. However, in spite of the fact that the spread of the distribution is now in better agreement, we still observe that the centre of the distribution from the OpenQuake curves is still higher than that of the original curves. Though not shown here, analysis of similar distributions from elsewhere in France suggest that the “average” of the OpenQuake implementation of the ZONELESS branches still exceeds that of the original implementation by between 20 - 40 %, depending on the location and the annual probability of exceedance.

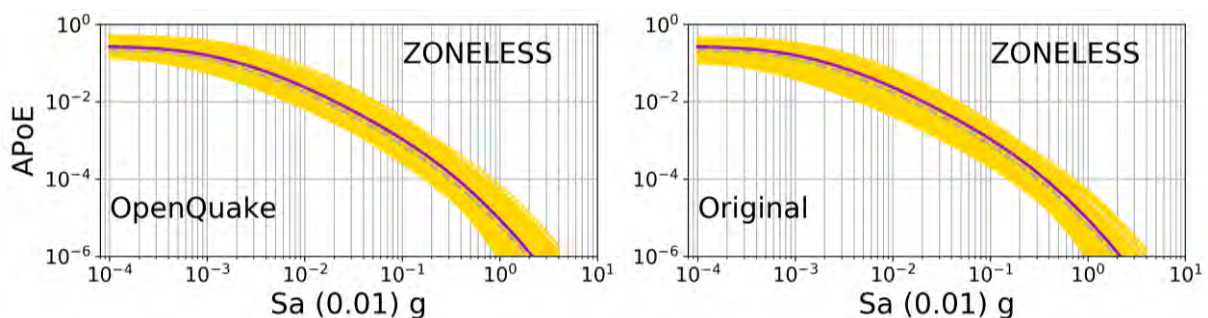


Figure 29: Comparison seismic hazard curves of the ZONELESS branches for the Grenoble site (as shown in Figure 28), this time accounting for epistemic uncertainty in a and b according to the process described in the text.

The results shown here indicate that though we have identified and potentially resolved a critical discrepancy between the original implementation and the OpenQuake implementation of the model, the desired level of agreement has still not yet been reached. Further iterations also explored for these specific branches several of the possible factors shown in the sensitivity analysis within the previous section, e.g., depth distribution, nodal plane distribution, scaling relation etc. No substantial improvement in has been observed, and further interactions with the model developers have not managed to shed light on the potential cause of the discrepancies. It is our conclusion that at the present time we cannot reduce the mismatch further without detailed element by element tests between the original software and OpenQuake in order to resolve more systematically each potential factor in the process. This cannot be done without access to the proprietary code used by the original model developers, however. We

have therefore taken the decision to disseminate the OpenQuake input files for the calculation in their present form within a version control system, alongside code used to construct the models from the available data. Limitations and caveats are made clear in the accompanying documentation to the code. It is our hope that future efforts will be made to identify the cause of the differences between the models, and if the reason or reasons are identified and resolved then the files will be updated at a future point in time. The additional results that follow in this chapter must be interpreted in light of the acknowledged discrepancy.

5.3 Seismic Hazard Maps and Uniform Hazard Spectra

Though the city comparisons have shown that in spite of the good agreement between the implementations at a branch-by-branch level the mean and quantiles were found to disagree, we nevertheless proceed to compare the results for the entire country. In Figures 29 to 33 we show both the OpenQuake-generated seismic hazard maps and the difference maps with respect to the original implementation. These maps are shown for the mean, 16th, 50th and 84th percentile for the 475-year return period (Figures 30 - 34) and the mean for the 2475-year return period.

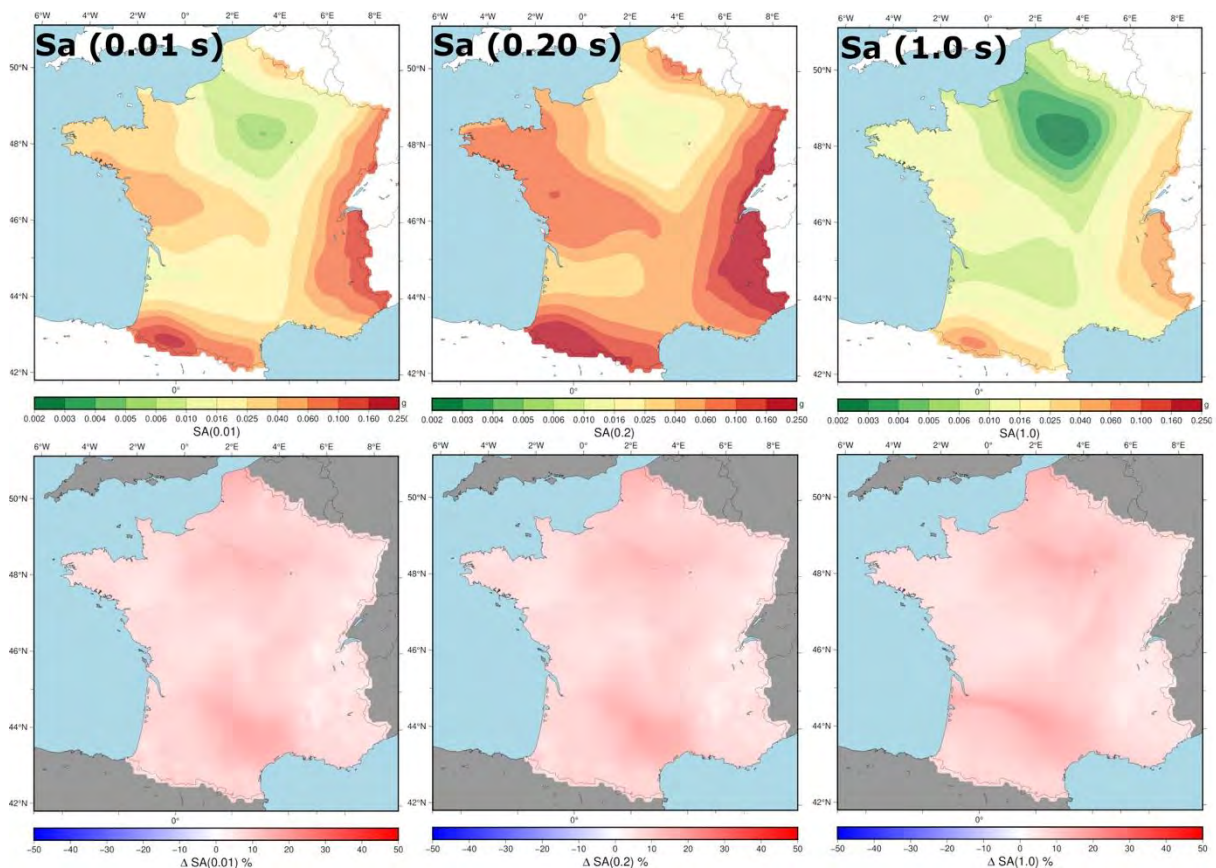


Figure 30: 475 year return period seismic mean hazard maps for $S_a(0.01\text{ s})$, $S_a(0.2\text{ s})$ and $S_a(1.0\text{ s})$ (top row) from the OpenQuake implementation of Drouet et al. (2020), and corresponding percentage difference maps with respect to the original Drouet et al. (2020) results (bottom row).

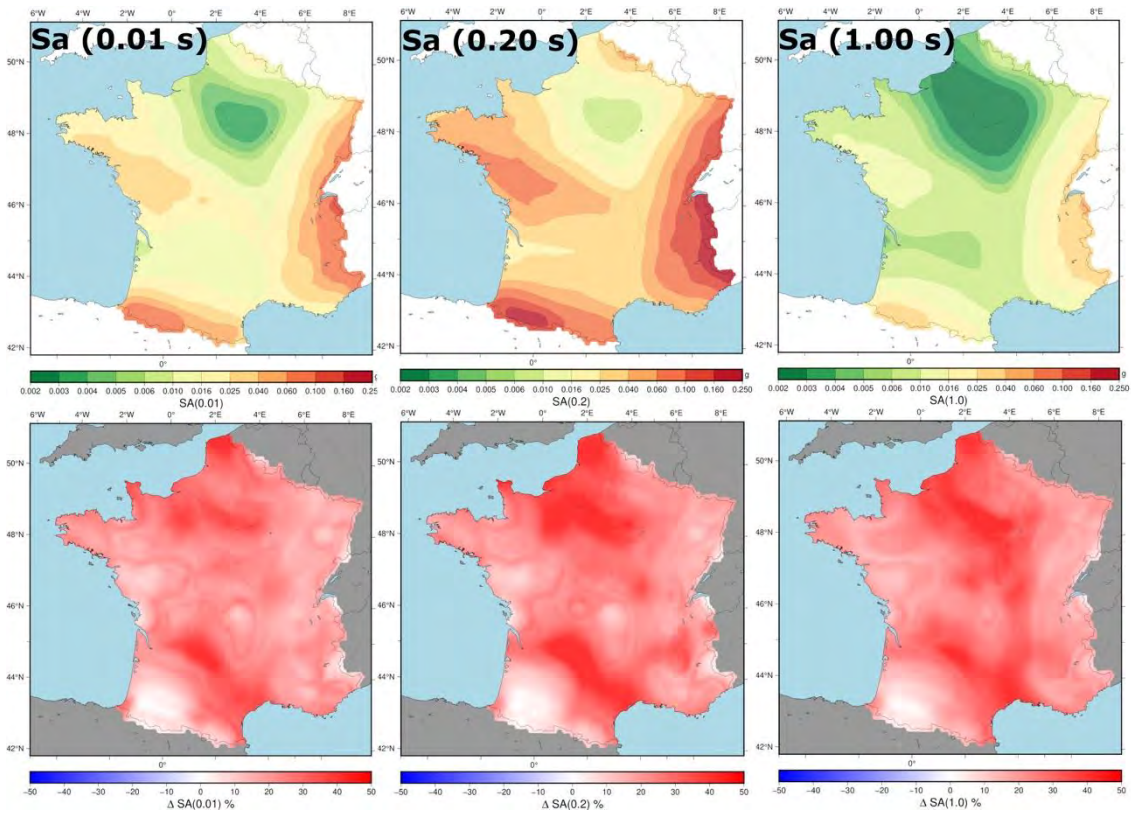


Figure 31: As Figure 30 for the 16th percentile

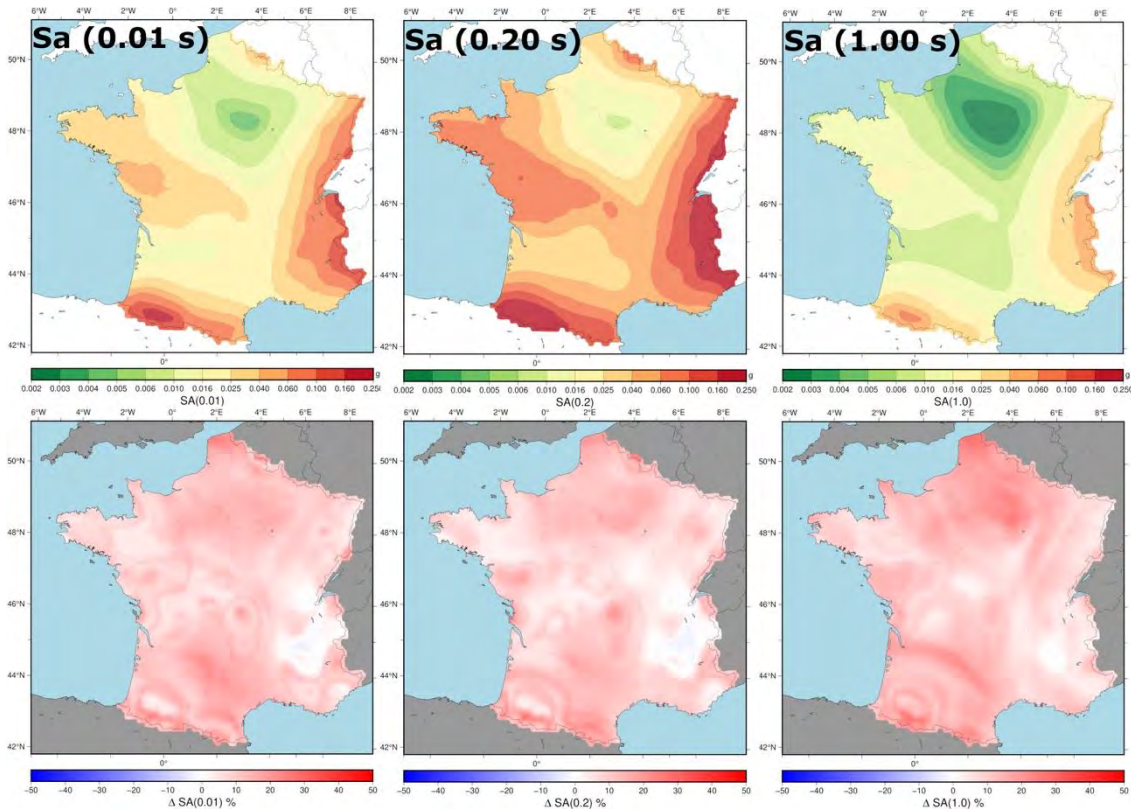


Figure 32: As Figure 30 for the 50th percentile

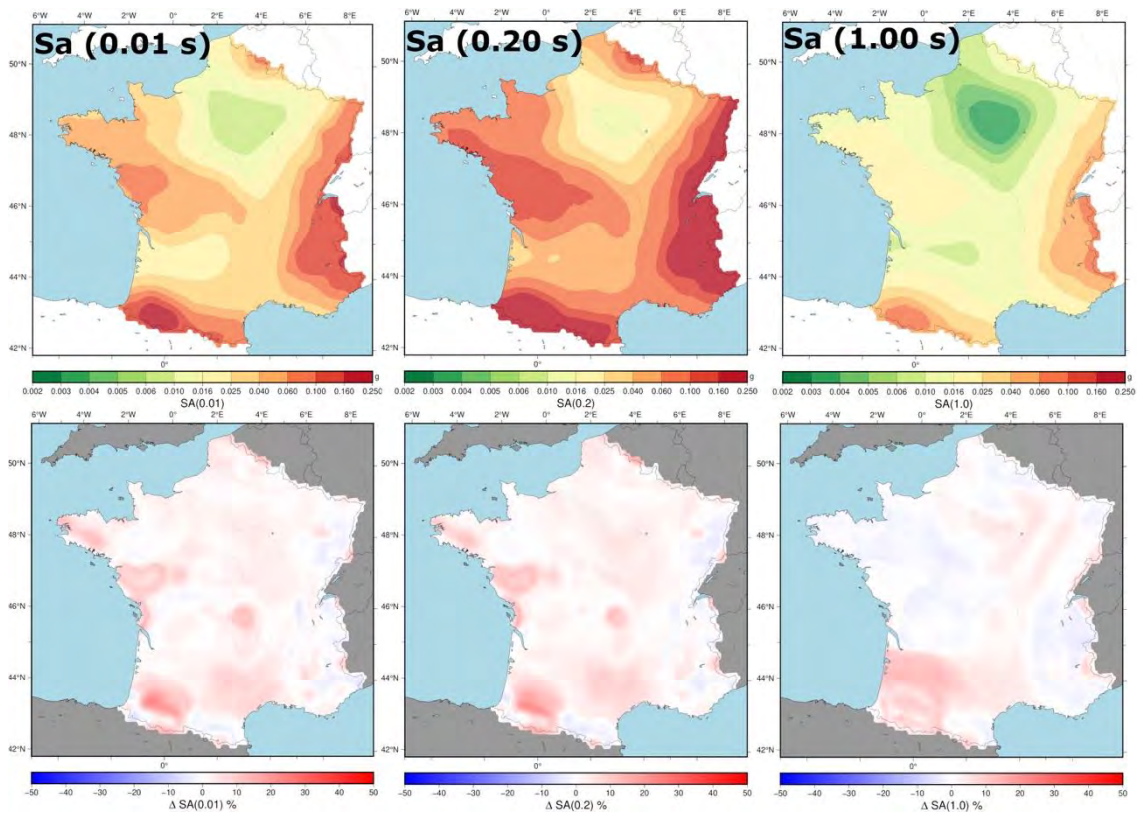


Figure 33: As Figure 30 for the 84th percentile

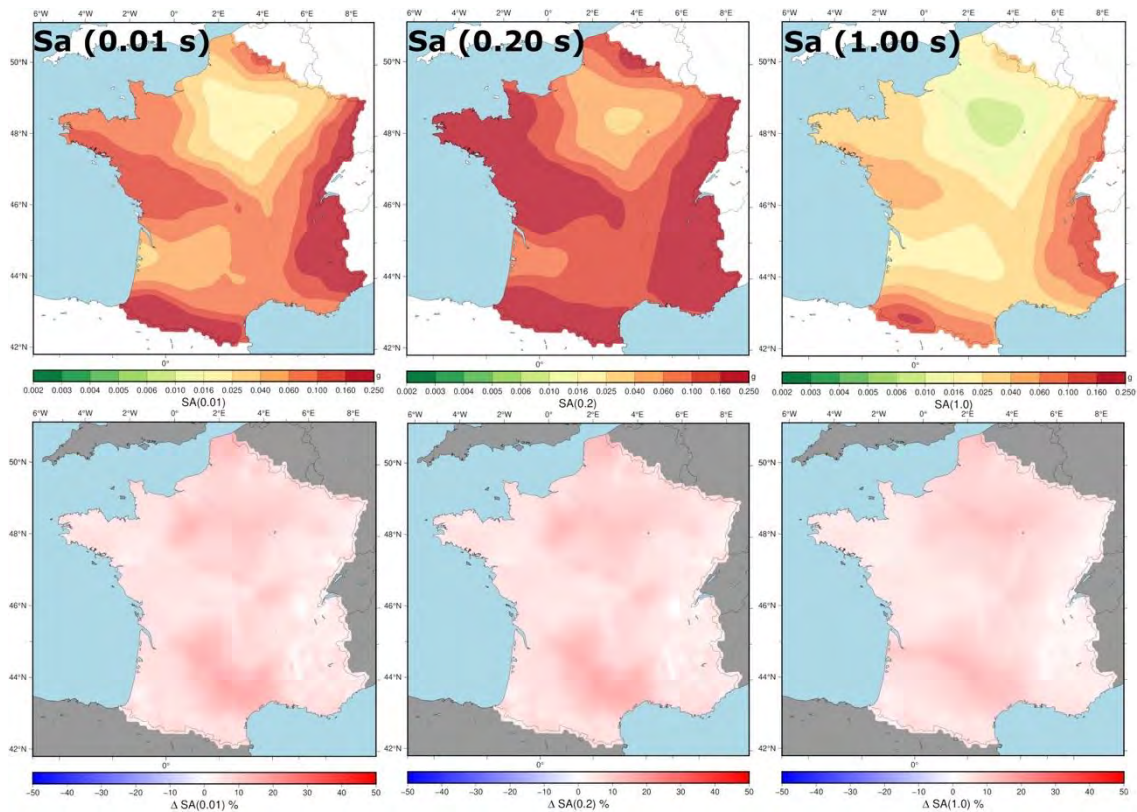


Figure 34: As Figure 30 for the 2,475 year return period

The seismic hazard maps, and in particular the difference maps, bear out the same trends that were inferred from the city comparisons. Even with the discrepancies that have been discussed, at the 475-year return period OpenQuake returns a mean seismic hazard that is mostly within 10 to 20 % higher than the original values from Drouet et al. (2020). Better agreement is seen in much of France for the 84th percentile and substantially poorer agreement at the 16th percentile. The spatial patterns reflect the relative influence of the smoothed seismicity model with respect to the uniform area sources. For example, the largest disagreements between the mean hazard maps can be seen in an arc that runs through Occitanie and Nouvelle-Aquitaine, from the southern Mediterranean to the Atlantic coast. In this region the smoothed seismicity activity rates are particularly low, while in the area source models there is more variability but the implied activity rates are higher. The absence of the lower activity rate branches in the ZONELESS model result in a higher mean hazard from the OpenQuake version. This situation appears to be reversed in the higher activity region of the Pyrenees. There the 16th percentile is in good agreement and it the 84th percentile where OpenQuake is producing higher seismic hazard. This would suggest that the smoothed seismicity branches are not among the lower quantiles but lie closer to the centre of the distribution. In southeastern France, where the smoothed seismicity rates are generally higher and in closer agreement with the rates from the area sources, the disagreement in the mean values is less as the branches of the ZONELESS model may be more closely aligned with the range of values from the uniform area sources.

5.4 Calculation and Interpretation of the Mean Ground Motion

So far in the report we have shown distributions of seismic hazard for the five target cities, including the full range of values from all branches as well as the mean, median and other quantiles. In showing these results, however, this opens up a question of what is meant by the mean hazard. As has been shown earlier in this report we can interpret the distribution of seismic hazard curves in terms of the distribution of probabilities of exceedance for a given intensity measure level, or as a distribution of intensity measure level given a probability of exceedance. The difference is illustrated in Figure 35, in which we highlight how for a given intensity measure level with a determined APoE the exact definition of the uncertainty in the hazard can differ when considering the distribution along the axis of the intensity measure level given a fixed APoE (in this case $S_a(0.1\text{ s})$) or along the axis of the APoE assuming a fixed intensity measure level. When describing the “mean” hazard we may decide if this mean is determined from the intensity measure levels or from the probabilities of exceedance. By default, OpenQuake’s mean hazard curves refer to the mean probabilities of exceedance for each of the input intensity measure levels. To retrieve the mean from the intensity measure levels we interpolate the curves (in log-log space) to a fixed set of APoEs (including the target APoE) and retrieve the mean intensity measure level of each APoE.

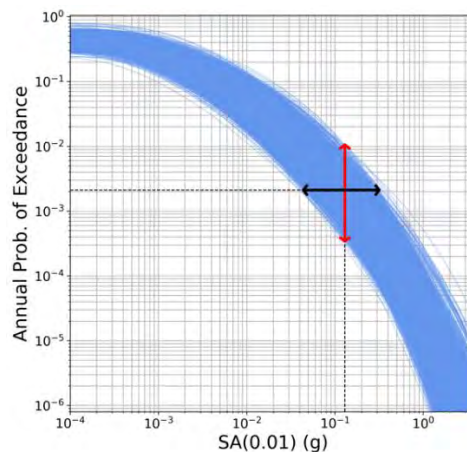


Figure 35: Example distribution of seismic hazard curves from a logic tree, with red arrows indicating distribution of APoEs for a fixed IML, and the black arrow the distribution of IMLs for a fixed APoE

The choice of whether to retrieve the mean from the APoEs or the IMLs is not the only factor that may be of influence in the resulting hazard, as shown in Bommer & Scherbaum (2008). As seen in Figure 35, the distributions of the curves are found to be close to a normal distribution in log-log space. As a measure of central tendency of the distribution, the geometric mean could also be considered a viable alternative to the weighted arithmetic mean.

$$\mu = \exp\left(\frac{1}{N_{BR}} \sum_{i=1}^{N_{BR}} w_i \cdot \ln y_i\right) \quad \text{where } \sum_{i=1}^{N_{BR}} w_i = 1 \quad (9)$$

Once again, however, by OpenQuake will calculate the mean curve as the weighted arithmetic mean of the APoEs rather than the weighted geometric mean.

From the choices listed above, we can have four different ways of retrieving the mean seismic hazard curves from the distribution of curves illustrated in Figure 35.

- 1) Arithmetic mean of the APoEs (with respect to fixed IMLs) [as done by OpenQuake]
- 2) Geometric mean of the APoEs (with respect to fixed IMLs)
- 3) Arithmetic mean of the IMLs (with respect to fixed APoEs)
- 4) Geometric mean of the IMLs (with respect to fixed APoEs).

The “mean” hazard curves using each of these four different methods are compared against the reference (option 1) for the Grenoble test site in Figure 36 alongside their relative increase or decrease in hazard with respect to APoE.

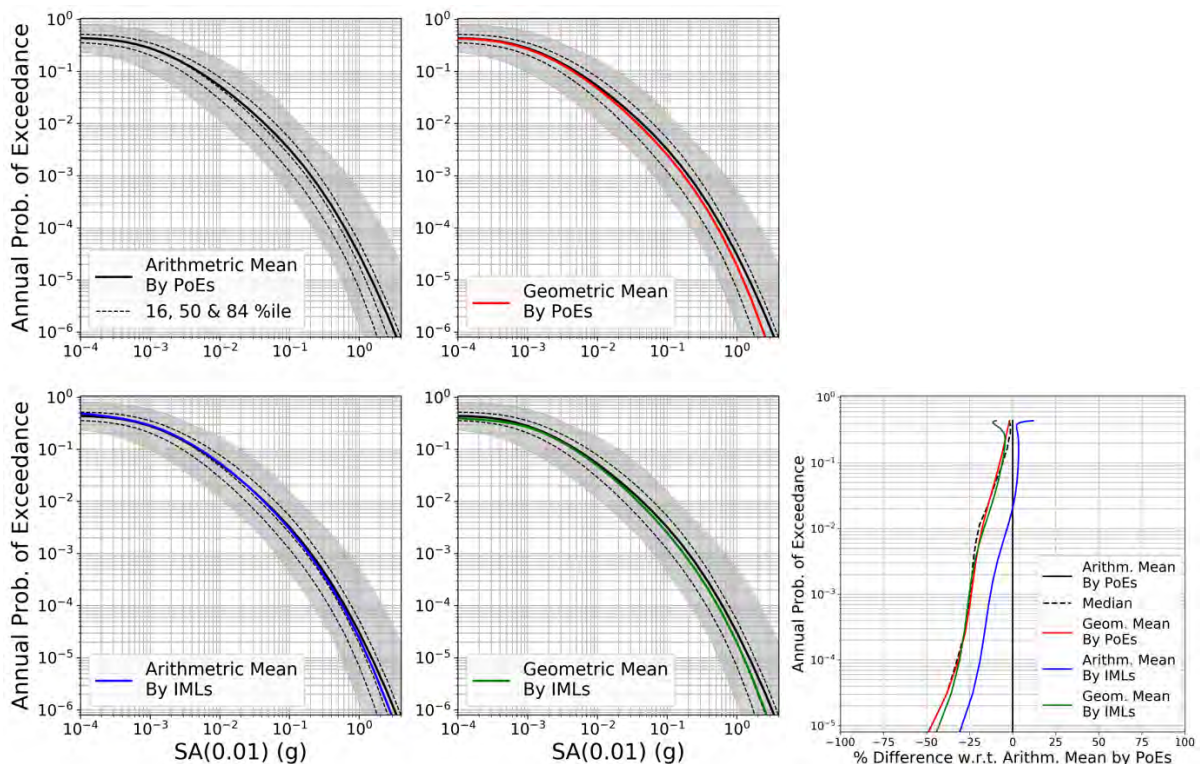


Figure 36: Comparison of mean hazard curves for Grenoble by each of the four different approaches shown in the text, with the current (reference) OpenQuake and original software approach shown in black across all plots along with the current 16th, 50th (median) and 84th percentile. Relative differences in the mean curves for the different approaches with respect to the reference approach are shown in the figure on the bottom right.

The comparison of the four different approaches show that the currently adopted approach of calculating the arithmetic mean of the probabilities of exceedance tends to yield the highest mean hazard curves. It can also be seen that when using the arithmetic mean there exists a more significant difference between the resulting hazard curves when choosing to determine the mean based on the probabilities

of exceedance or on the intensity measure levels. Adopting instead the geometric mean appears to minimise the influence of this decision, at least in the present case.

Another important insight that the comparison of the different estimators of the mean curve reveals is that the curves of the geometric mean hazard (either by APoE or IML) agree well with the median hazard curves. The calculations of the quantiles are unaffected by the choice of geometric or arithmetic mean, and though some differences may materialise when using APoE or IML as the axis for the distribution, these differences are smaller and depend more on the exact logic tree, so not as clear as generalisable. The observation that the weighted geometric mean curve from the logic tree agrees well with the median is something that can be expected from the distributions of the epistemic uncertainties being captured in the model. For this some background is necessary.

In earlier uses of logic trees for seismic hazard in low to moderate seismicity regions the most common epistemic uncertainties represented took the form of alternative area source models (similar to the GTR, EDF and IRSN options here) and alternative ground motion models. The assignment of weights may have been undertaken based on judgement, or in the cases of GMMs potentially on quantitative metrics describing the relative fit of the models to what few data may be observed. In these cases, the expected APoEs for the branches and their corresponding end weight either assume no specific probability distribution or end up being more closely approximated by a uniform distribution ($\mathcal{U}[0,1]$). Where the hazard values and weights tend more toward a uniform distribution then the mean and median hazard should converge regardless of whether it is the logarithm of the APoE/IML or the values themselves that are taken. Indeed, depending on the relative weights given to some models (e.g., strong preference for model A over models B or C) one could even encounter cases where the median exceeded the mean hazard. With a large number of models (and therefore end branches) then the central limit theorem begins to dominate, and we eventually see the distributions taking on Gaussian properties and we can become more secure that the means, median and quantiles are adequate to represent the main moments of the resulting distribution of the hazard values.

This still does not address the question of whether the geometric or arithmetic mean should be used. The picture changes, however, when epistemic uncertainties are introduced in order to capture specific distributions of the underlying parameters of influence, such as the case with the a , b and M_{MAX} branches. This becomes relevant to the question of interpretation of the meaning of the weights, i.e. they can no longer be thought of as subjective indications of relative merit but rather as discrete approximations to a specific underlying probability distribution (e.g. Miller & Rice, 1983). Now the epistemic uncertainty in the rate of occurrence of earthquakes with magnitude M is described by a marginal distribution with variance $\sigma^2(M)$ emerging from the multivariate a and b distributions (Stromeyer & Grünthal, 2015). The distributions of activity rates therefore become normal in logarithmic space, i.e., lognormal. In the case of the Drouet et al. (2020) logic tree, the distribution of the seismic hazard values (for the sake of argument, we'll assume in terms of APoEs) therefore represent the superposition of 16 evenly weighted distributions of seismic hazard in which the activity rates for each magnitude are explicitly modelled as lognormal. The result is that the hazard curves themselves tend toward a more explicit lognormal distribution in which the difference between the mean and median curves should reflect the difference in the mean of a lognormal distribution ($\exp(\mu + \frac{\sigma^2}{2})$) and the median of the lognormal distribution ($\exp(\mu)$). In this case, the argument for taking the geometric mean has a stronger theoretical basis than for taking the arithmetic mean.

As a final point of note on the topic, the assumption of lognormality in the distributions of the variables that underpin the hazard model become even more explicit when moving to a scaled backbone approach for characterising the median ground motion. A full treatise on the approach and relative merits of a scaled backbone approach is beyond the scope of this deliverable. Briefly, however, the scaled backbone approach does not represent epistemic uncertainty by selecting and weighting alternative GMMs, but instead by adopting an appropriate reference (or backbone) GMM and applying adjustments

to the model to reflect the epistemic uncertainty in the seismological properties that control the scaling of ground motion in region (as well as some statistical uncertainty from the limitations of the data set used to regress the model). The epistemic uncertainty in expected ground motion ($\ln Y = \mu$) of a backbone GMM is therefore represented by a Gaussian distribution (again in logarithmic space) with variance σ_μ^2 . To map these into alternative branches of the logic tree, the adjustment factors (ε_μ) and their respective weights to explicitly approximate the standard normal distribution into a discrete set of weighted branches that capture the first few moments of the distribution (e.g., Miller & Rice, 1983). In Europe, this has been recently adopted in the 2020 European Seismic Hazard Model, where the GMM logic tree is built around a *regionalised* scaled backbone approach (Weatherill et al., 2020). Here, the epistemic uncertainty in the median ground motion reflects the region-to-region variability in stress parameter, $\mathcal{N}(0, \tau_{L2L})$, and in the residual (anelastic) attenuation parameter, $\mathcal{N}(0, \tau_{c3})$. Here then, the distribution of expected ground motions are explicitly lognormally distributed.

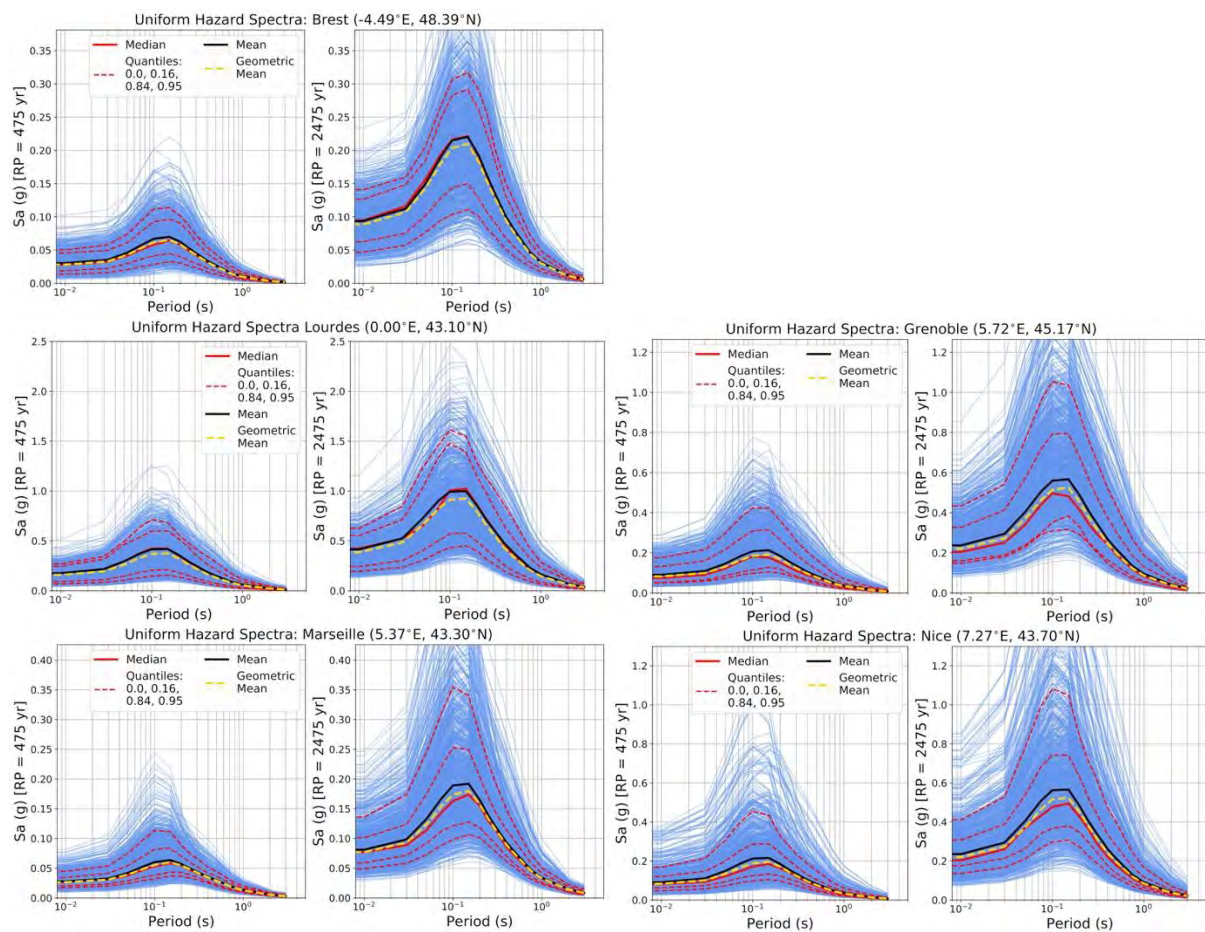


Figure 37: Uniform hazard spectra for the 475 year and 2475 year return period for Brest (top row), Lourdes (middle row, left), Grenoble (middle row, right), Marseille (lower row, left) and Nice (lower row, right). Mean UHS indicated in black, median as a solid red line, other quantiles as dashed red lines and the geometric mean as a dashed yellow line.

In light of the discussion above, the uniform hazard spectra (UHS) for five target sites are shown in Figure 37, in which we highlight both the arithmetic mean and median UHS (determined from probability of exceedance) and the geometric mean UHS. Though the differences between the two mean definitions are small compared to the distances between the quantiles, these may still be relevant for engineering design when only a single definition of seismic hazard is needed. Further efforts to explore the implications of all of the above issues in the framework of seismic risk analysis, and the conditions under which these differences assume great importance, is the subject of ongoing and future projects.

6. Conclusions and Recommendations

6.1 The Complete Implementation

The full migration of the Drouet et al. (2020) from the original PSHA software to OpenQuake serves multiple objectives. The first is to have a seismic hazard model for France that is implemented into a common framework that has been adopted by other national seismic hazard models in neighbouring countries (e.g., Italy, Germany and Switzerland) as well as the 2020 European Seismic Hazard Model (ESHM20). Included with the deliverable are the complete set of OpenQuake input files, which can be used by others to explore the model in depth and retrieve detailed results according to their needs. The second objective, however, was to explain in detail the migration process, to illustrate the differences that can emerge and the information that is needed, and to outline the sorts of comparisons and investigations that should form part of not only a migration process but a more comprehensive process of comparing PSHA models for a region.

We had set target criteria by which we would determine whether or not the models were in agreement, namely $\pm 10\%$ difference for $APoE > 10^{-4}$, with respect to the original curves provided by the authors. The comparison plots, both for the individual curves and for the mean and quantiles, show that this has been achieved. For some sites this target has been exceeded, meaning that either a smaller difference was present and/or that the agreement was maintained to lower $APoE$, though this was not achievable uniformly throughout the country. The target was not necessarily achieved for all branches of the model, though the resulting agreement of the mean and quantiles would suggest that over- or underestimations of seismic hazard on certain branches largely cancel one another out. More important in this case, at present, is the discrepancy in the sampling of activity rates for the ZONELESS branches, which we aim to resolve before release of the model.

To a certain extent, the Drouet et al. (2020) represents a simple model for undertaking an independent implementation in a different PSHA software. This is not a reflection of the rigour of process by which the recurrence the model itself was created, nor to detract from the novelties contained therein. Where it is “simple” is in the adoption of only two source types (uniform area sources and zoneless gridded seismicity), both of which are types of distributed seismicity sources and, as such, share common characteristics that are relatively straightforward to apply once they are defined. Similarly, the GMM logic tree adopts only four ground motion models, meaning that the number of end-branches of the logic tree is constrained. Other PSHA models, including those for Germany, Switzerland, Italy and Europe, may contain features such as explicit characterisation of active fault sources or more complex GMM logic trees that apply different adjustments or scaling factors to existing models. These features, particularly the characterisation of active faults, can add further complexities that produce differences from one software to another.

The initially apparent simplicity of the Drouet et al. (2020) model is, however, arguably deceptive, and certain aspects presented significant computational challenges for the OpenQuake implementation. The greatest of these challenges is in the execution of the logic tree with 100 branches per model to account for the epistemic uncertainty in a , b and M_{MAX} . In the current state of the OpenQuake software, logic trees such as this present a particular challenge. At the time of undertaking these comparisons, OpenQuake does not have a means by which the rates of the ruptures in the earthquake rupture forecast can be adjusted *without* re-computation of the earthquake rupture forecast (and correspondingly the distances and expected ground motions) on each branch. This resulted in an especially demanding computation for the whole logic tree, the majority of which was spent re-building the earthquake rupture forecast, re-computing the sources to site distances and resulting ground motions.

For site-specific calculations, or calculations over limited areas, OpenQuake’s efficiencies in terms of pre-filtering sources and ruptures away from the sites, and a well-designed rupture-by-rupture

parallelisation strategy, make execution of the complete logic tree feasible in a short period of time. To run the entire country, however, these advantages diminished and ultimately the calculations had to be run in an unconventional manner, which still took a long time to run (more than 14 days on a 120-core server with 1 Tb of RAM) and required additional work to pre- and post-process the results. Epistemic uncertainties on the recurrence parameters of the MFD are not novelties in PSHA, though they can present certain challenges in their own right. By sampling the distributions prior to the calculation, Drouet et al. (2020) have avoided the problem of zone-to-zone correlation in epistemic uncertainty and/or logic trees that rapidly increase in size. These problems have plagued logic trees such as this in other PSHA models. Nevertheless, as epistemic uncertainty on recurrence parameters for uniform area sources is commonplace and can no longer be neglected or collapsed, as was the case for the ESHM13 (Wössner et al. 2015), one recommendation from this work would be for OpenQuake's developers and contributors to seek a way to execute logic trees such as this without the extensive re-computation of the earthquake ruptures. We do not underestimate the scale of effort that may be required to adapt the code in such a manner; however, given the emergence of many models in low-to-moderate seismicity regions that are implementing more complex epistemic uncertainty models for magnitude frequency distributions, we strongly encourage exploration of the possibility of adapting the code, or seeking for alternative strategies that could more efficiently yield comparable outputs.

For the Drouet et al. (2020) model itself, we highlight several aspects relating to the PSHA implementation details, for which further information was needed that could not be found in either the journal paper or the GEOTER (2017) report. These are:

- 1) The strategy for sampling the a , b and M_{MAX} values for the 100 MFD branches of the logic tree. Here the authors are transparent in providing the coefficients of the full multivariate normal distribution, from which samples could readily be drawn. From the comparisons of the distributions of the parameters provided by the author and those retrieved from conventional Monte Carlo sampling (explained in section 2), it is evident that a stratified sampling strategy was applied. Explanation of this sampling strategy is not provided in the supporting documentation for this model, and a user unaware of this fact would not easily be able to reproduce the results without adopting a similar strategy. Ultimately, we have adopted here the samples provided by the authors, and we would recommend that these values are considered as the fixed values for the model. Though we demonstrated that the authors samples are consistent with the underlying multivariate normal distributions for uncertainty on a and b , we have reason to believe that a model whose a and b values are drawn from the same distributions but using a different and more conventional sampling strategy would likely produce different results.
- 2) Parameters to describe the construction of the virtual faults for both the uniform area and the zoneless sources. These included the number of depth layers considered, the magnitude-area scaling relation, the aspect ratios, the behaviour of the aspect ratio at large earthquakes as a result of the constraints imposed by the seismogenic thickness, the number of rupture orientations and dip angles considered and the strategy for sampling these. As we saw from the sensitivity analysis, though these factors may not be dominant (each introducing discrepancies within the $\pm 10\%$ threshold) they do influence the degree of agreement that can be achieved without receiving further information and clarification from the authors.
- 3) For the zoneless model additional information did need to be obtained from the authors to clarify how the recurrence models were defined and how the virtual faults are controlled. Documentation on the full implementation of the zoneless model in the PSHA software was lacking overall.
- 4) Certain details of the calculation configuration such as the site properties needed for the Abrahamson et al. (2014) GMM, and the application (or not) of truncations such as maximum distance or sigma truncation.

Though we have drawn attention to specific information that was not present in the available documentation, this should not be taken necessarily as a criticism of the author's approach or the available documentation itself. Indeed, we strongly commend Stéphane Drouet, Gabriele Ameri and colleagues not only for the amount of information they did make publicly available in the first place, but also for their willingness to provide further clarifications and data sets upon request. The degree of transparency and the availability of the digital data sets is particularly good in this case and sets an example for clarity and transparency that other national seismic hazard models should follow. The issues raised in this migration of the model are seldom discussed in detail, or at least in the detail necessary for reproduction of a PSHA model, in scientific papers. Even project reports or similar detailed deliverables can omit all of the information needed.

6.2 A Template for PSHA Model and Software Comparisons?

In section 1 we described the precedents for undertaking multiple implementations of a seismic hazard model in different PSHA software. These can be import and insightful exercises that can form part of a quality assurance process for PSHA at critical facilities, yet even for models for national seismic hazard assessments for design codes for *ordinary* structures they can play an important role. This is especially true if an organization wishes to migrate their seismic hazard models from one software to another and needs to understand what changes in seismic hazard can emerge through this process alone. With this being said, however, comparison studies such as this can be time-consuming, and the level of expertise needed, combined with the number of iterations required to reach an acceptable level of agreement, can make endeavours such as these costly. These costs can be compounded by factors such as an organisation's unwillingness or inability to divulge details of their proprietary tools, time pressure, and the impediments to transparency and reproducibility that scientific publications can impose in order to keep manuscripts to a manageable length, or details that end up being omitted in order to emphasise scientific novelty.

Before making recommendations as to how PSHA models can be improved for the purpose of broader implementation and verification, it is important to re-clarify what the purpose of this specific endeavour was, and with that to recognise the limitations. The efforts made to translate the Drouet et al. (2020) model from its original format into OpenQuake take the form of an advanced reproducibility and replicability study. Advanced here means that we are choosing to attempt to replicate results with a different set of tools rather than simply ensuring that the same results can be obtained with a simple repetition of the experiment. The context is important here, as we do not have access to the original codes through which the calculations are run, yet the results of these codes are the benchmark against which the new implementation must be compared. We have been fortunate that the original model developers have been forthcoming in providing input files, answering specific questions that we have raised and, to a certain extent feasible, providing insights into the operation of the PSHA software.

As we have undertaken here what is fundamentally a replicability study, however, we have not contracted or engaged the original model developers directly in order to request the unit-by-unit level comparisons and verifications that would be expected had the project been undertaken for, for example, a nuclear power plant quality assurance verification study (e.g., Bommer et al., 2014). Had that been the case, then a suite of intermediate results of the calculations would have been produced through which discrepancies would be identified and explored until their cause found. If it were found that the discrepancy would be attributed to an error or misinterpretation in one particular software, be it the proprietary or open code, then this would be rectified, the issue resolved, and the comparison continued. Where irreconcilable differences emerge then the contribution of these differences to the total discrepancy between the curve could be reported and, in the best case, the question as to whether a preferred approach should be adopted referred to the stakeholder or whether this should reflect an *operational epistemic uncertainty*. But this description of a verification or quality assurance study is an

ideal case that is desirable when the consequences to the stakeholder (or to society itself) may be severe if errors emerge.

If the purpose of the study is not necessarily a complete quality assurance validation study, what purpose does this serve and what lessons can be learned to help improve PSHA developments in the future. The purpose here was a reproducibility study, one that takes a reference model that has been developed for application in a region for *ordinary* building design, to translate it into a format separable from the propriety tools used in its construction and to understand the implications for the resulting seismic hazard curves when undertaking this translation. There are several key conclusions that we can draw for this case, which would very likely transfer to other models around the globe too: i) published documentation for the complete seismic hazard model (the journal paper *and* the technical report) did not, and often does not, provide all of the information necessary to define fully the input parameters needed to implement the model nor the specific calculation strategies, ii) the seismic hazard model is not entirely separable from the software in which it is implemented and translation will always yield differences that may not always be possible to resolve, iii) consultation and engagement with the authors is usually necessary for a systematic approach to resolving discrepancies. More generally, a model cannot be adequately migrated to a different software if its input parameters are not thoroughly and explicitly documented (ideally through the availability of the input files), if the computational strategies of the codes are not understood, and if systematic component-by-component level comparisons of codes are not undertaken. The first and last of these three points are particularly critical to address for two reasons: firstly they are impediments to reproducibility and, secondly, they imply that development and release of a model may incur a long-term cost to the authors in terms of providing support to other scientists and engineers in order for their model to be used effectively by the end-users if they wish for the model to be recognised as reproducible. This cost is, in most cases, unsustainable in the long term and so it is in addressing the issue of documentation and transparency that we find there is the greatest scope for improving reproducibility in PSHA practice.

In light of the context provided by this project, we suggest that authors of PSHA models, both for critical facilities as well as national models for seismic design codes, adopt a more standardized and comprehensive documentation strategy. This strategy can apply regardless of whether or not the models are intended for public release, and it comprises a mix of written documentation and digital information. We propose that the documentation of a model contain the following information:

1. A written report explaining the general model and its development – as is usually done at present. Some precedents for something approaching a standard template for such a report have been seen in implementation and verification reports for some PSHA studies at critical facilities. This process could be broadened, however, and comprehensive templates for model documentation developed to wider use.
2. Digital geospatial files in an Open Geoscience Consortium recognized format, containing all relevant spatial information about the seismic sources *and the seismic hazard target sites*.
3. For the selected ground motion models (including applied adjustments or scaling factors) a set of verification tables should be produced *directly from the implementations in the PSHA software*. These verification tables should provide the expected ground motion and the standard deviation, including between event, within-event and, where relevant, site-to-site variability, for a well sampled set of scenarios that span sufficiently the range of magnitudes, distances, site conditions and intensity measure types considered in the model.
4. For the source types relevant to the model, idealized test case calculations should be undertaken using the PSHA software in question. These test case calculations should be *atomic*, meaning that they should be testing only on single element of the PSHA process, for example considering only a single area source, MFD, ground motion model and target site. Fortunately, clear and well-documented examples of these atomic level tests that each PSHA software *should* be able to implement can be found in the PEER PSHA Software Verification

Tests (Thomas et al., 2010; Hale et al., 2018) and we recommend specifically adopting those described by Hale et al. (2018). For each test the modeler should provide the resulting seismic hazard curve(s) and a standard template document should be available for them to indicate what additional information needed to be input, or which parameters omitted, in order to run the test.

5. For a set of representative target sites, which might correspond to all target sites considered in the model, expected seismic hazard curves should be provided in digital format for all individual branches of the logic tree. This information was provided to us by the authors in the present case and was critical in identifying specific issues that would otherwise have been difficult to discern if considering only the mean and/or quantiles.
6. For a set of representative target sites, which might correspond to all target sites considered in the model, a set of mean and quantile seismic hazard curves should be provided in digital format. These are important to understand how the mean and quantiles are being calculated by the PSHA software, and to identify any potential differences in approach or interpretation.

Where the PSHA software allow, one could also consider adding in probabilistic seismic risk computations for given idealized structures (with fixed fragility, vulnerability and replacement cost) as a means of identifying whether differences in the seismic hazard translate to significant differences in the corresponding probabilities of exceedance of loss.

Specific details on the templates, such as the information that should be contained and the written and digital formats, will be proposed in future studies.

6.3 Future Directions

The activities presented in this deliverable form part of a larger ongoing process of harmonising different PSHA models for Europe into a common software framework. Work is ongoing to undertake the same process to migrate the Grünthal et al. (2018) model into OpenQuake, at it is hoped that other models may follow on from this. The aim behind this work is to be able to compare and explore the models and a manner that ensures consistency in the definitions of each component of the model and to identify to what extent differences between models are attributable to of the implicit and explicit assumptions of the software. As a fully open-source software, OpenQuake has been the obvious target software for this, as it is possible to understand clearly how certain processes are working by exploring the code in detail. With an extensive database of software harmonised models, we will move in future work toward the domain of testing PSHA against observations of ground motions and/or earthquake effects. We aim to make this process as transparent and reproducible as possible, providing tools for this purpose in addition to a common set of models to which testing strategies can be applied and compared.

Acknowledgements

The authors are sincerely grateful to Stéphane Drouet, Gabriele Ameri and David Baumont for the information and assistance that they have provided for the OpenQuake implementation of the model. The OpenQuake software is developed and maintained by the Global Earthquake Model and we wish to thank Michele Simionato for providing additional support. Finally, we wish to thank Christoph Martin and Pierre Labbé for their thorough and insightful reviews of this work, and Emmanuel Viallet for sharing with us his experience and guidance throughout the process.

6. Reference:

1. Abbott E, Horspool N, Gerstenberger M, et al. (2020). Challenges and opportunities in New Zealand seismic hazard and risk modeling using OpenQuake. *Earthquake Spectra*, 36(1_suppl). 210-225. doi:10.1177/8755293020966338
2. Abrahamson, N. A., Silva, W. J., and Kamai, R. (2014). Summary of the ASK14 Ground Motion Relation for Active Crustal Regions. *Earthquake Spectra* 30(3), 1025–1055. doi: 10.1193/070913EQS198M.
3. Akkar, S., Sandikkaya, M. A., and Bommer, J. J. (2014). Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering* 12(1), 359–387. doi: 10.1007/s10518-013-9461-4.
4. Allen, T. I., Halchuck, S., Adams, J., and Weatherill, G. A. (2020). Forensic PSHA: Benchmarking Canada's Fifth Generation seismic hazard model using the OpenQuake-engine. *Earthquake Spectra*,= 36(1_suppl), 91-111. doi:10.1177/8755293019900779
5. Ameri, G. (2014). Empirical Ground Motion Model adapted to the French context. Deliverable 524. SIGMA: SIGMA-2014-D2-131.
6. Ameri, G., Drouet, S., Traversa, P., Bindi, D., and Cotton, F. (2017). Toward an empirical ground motion prediction equation for France: accounting for regional differences in the source stress parameter. *Bulletin of Earthquake Engineering*, 15, 4681–4717 <https://doi.org/10.1007/s10518-017-0171-1>
7. Baize, S., Cushing, E. M., Lemeille, G., Jomard, H. (2013). Updated seismotectonic zoning scheme of Metropolitan France, with reference to geologic and seismotectonic data. *Bull. Soc. Géol. France*. 184(3). 225 – 259.
8. Bindi, D., Massa, M., Luzi, L., Ameri, G., Pacor, F., Puglia, R., and 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 dataset. *Bulletin of Earthquake Engineering* 12(1), 391–430. doi: 10.1007/s10518-013-9525-5.
9. Bindi, D., Cotton, F., Kotha, S. R., Bosse, C., Stromeyer, D., and Grünthal, G. (2017). Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. *Journal of Seismology* 21(5), 1201-1218. doi: 10.1007/s10950-017-9661-5.
10. Bommer, J., and Scherbaum F. (2008) The Use and Misuse of Logic Tree in Probabilistic Seismic Hazard Analysis. *Earthquake Spectra* 24(4). 997 - 1009
11. Bommer, J., Strasser, F. O., Pagani, M., and Monelli, D. (2013) Quality assurance for logic-tree implementation in probabilistic seismic hazard analysis for nuclear applications: a practical example. *Seismological Research Letters*. 84(6), 938 - 945
12. Cauzzi, C., Facciolo, E., Vanini, M., and Bianchini, A. (2015). Updated predictive equations for broadband (0.01 to 10 s) horizontal response spectra and peak ground motions, based on a global dataset of digital acceleration records. *Bulletin of Earthquake Engineering* 13(6), 1587-1612. doi: 10.1007/s10518-014-9685-y.
13. Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulleting of the Seismological Society of America*. 58. 1,583 – 1,606
14. Derras, B., Bard, P. Y., and Cotton, F. (2014). Towards fully data driven ground-motion prediction models for Europe. *Bulletin of Earthquake Engineering* 12(1), 495–516. doi: 10.1007/s10518-013-9481-0.
15. Douglas, J., Ulrich, T., Bertil, D., and Rey, J. (2014). Comparison of the Ranges of Uncertainty Captured in Different Seismic-Hazard Studies, *Seismological Research Letters* 85(5), 977-985. doi: 10.1785/0220140084.

16. Drouet, S. and Cotton, F. (2015). Regional Stochastic GMPEs in Low-Seismicity Areas: Scaling and Aleatory Variability Analysis—Application to the French Alps, *Bulletin of the Seismological Society of America* 105(4), 1883-1902. doi: 10.1785/0120140240.
17. Drouet, S., Ameri, G., Le Dortz, K., Secanell, R., and Senfaute, G. (2020). A Probabilistic seismic hazard map for the metropolitan France, *Bulletin of Earthquake Engineering*, 18, 1865 – 1898
18. EPRI (2012) Technical report: Central and Eastern United States Seismic Source Characterization for Nuclear Facilities. EPRI, Palo Alto, CA, U.S. DOE, and U.S. NRC: 2012
19. Field, E. H., Jordan, T. H., and Cornell, C. A. (2003). OpenSHA: A Developing Community-modeling Environment for Seismic Hazard Analysis. *Seismological Research Letters*. 74(4). 406 – 419.
20. GEOTER (2017). Probabilistic seismic hazard maps for the French metropolitan territory France. Fugro Document No.: GTR/EDF/0517-1603.
21. Grünthal, G., Stromeyer, D., Bosse, C., Cotton, F., and Bindi, D. (2018). The probabilistic seismic hazard assessment of Germany—version 2016, considering the range of epistemic uncertainties and aleatory variability. *Bulletin of Earthquake Engineering* 16, 4339-4395. doi: 10.1007/s10518-018-0315-y.
22. Hale, C., Abrahamson, N., and Bozorgnia, Y. (2018). Probabilistic Seismic Hazard Analysis Code Verification. PEER Report No. 2018/03. Pacific Earthquake Engineering Research Center, University of California, Berkeley.
23. Johnston, A.C., Kanter, L.R., Coppersmith, K.J., Cornell, C.A. (1994) The earthquakes of stable continental regions. Tech. Rep. Electric Power Research Institute (EPRI), Palo Alto, California
24. Le Dortz, K., Combes, P., and Carbon D. (2019) Alternative seismotectonic zonation for probabilistic and deterministic seismic hazard assessment for Metropolitan France. In *10th Colloque National AFPS, 24 – 27 septembre, Strasbourg*, p 15.
25. Leonard, M. (2014). Self-consistent earthquake fault-scaling relations: Update and extension to stable continental strike-slip faults. *Bulletin of the Seismological Society of America*. 104(6). 2953 - 2965
26. Mak, S. and Cotton, F. (2020). PSHA Updating and Comparison between French and German PSHA. Sigma 2 Deliverable
27. Manchuel, K., Traversa, P., Baumont, D., Cara, M., Nayman, E., Durouchoux, C. (2018) The French seismic CATatalogue (FCAT-17). *Bulletin of Earthquake Engineering*. 16. 2227 - 2251
28. McGuire, R., (1976) Fortran program for seismic risk analysis. United States Geological Survey Open File Report 76-67.
29. Meletti, C., Marzocchi, W., D’Amico, V., Lanzano, G., Luzi, L., Martinelli, F., Pace, B., Rovida, A., Taroni, M., Visini, F. and the MPS19 Working Group (2021). The new Italian Seismic Hazard Model (MPS19). *Annals of Geophysics*, 64(1), doi:10.4401/ag-8579
30. Mosca, I., Sargeant, S., Baptie, B., Musson, R. M. W., Pharaoh, T., (2020). National seismic hazard maps for the UK: 2020 update. *British Geological Survey Open Report*, OR/20/053, 1 - 141
31. Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L., Simionato, M., and Viganò, D. (2014a). OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*. 85(3), 692 – 702
32. Pagani, M., Monelli, D., Weatherill, G., and Garcia, J. (2014b) Testing procedures adopted in the development of the hazard component of the OpenQuake-engine. *Global Earthquake Model (GEM) Technical Report 2014-09*. doi: 10.13117/GEM.OPENQUAKE.TR2014.09, 1 - 73
33. Rey, J., Beauval, C., and Douglas, J. (2018). Do French macroseismic intensity observations agree with expectations from the European Seismic Hazard Model 2013? *Journal of Seismology* 22(3), 589-604. doi: 10.1007/s10950-017-9724-7.

34. Stafford, P. J. (2014). Source-scaling Relationships for the Simulation of Rupture Geometry within Probabilistic Seismic-Hazard Analysis. *Bulletin of the Seismological Society of America*. 104(4). 1620 – 1635.
35. Strasser, F. O., Bommer, J. J., and Abrahamson, N. A. (2008). Truncation of the distribution of ground-motion residuals. *Journal of Seismology*. 12. 79 – 105.
36. Tasan, H., Beauval, C., Helmstetter, A., Sandikkaya, A., and Guéguen, P. (2014). Testing probabilistic seismic hazard estimates against accelerometric data in two countries: France and Turkey. *Geophysical Journal International* 198, 1554-1571. doi: 10.1093/gji/ggu191.
37. Thomas, P., I. Wong, and N. A. Abrahamson (2010). Verification of Probabilistic Seismic Hazard Analysis Computer Programs. PEER Report 2010/106. College of Engineering, University of California, Berkeley: Pacific Earthquake Engineering Centre
38. Tromans, I. J., Aldama-Bustos, G., Douglas, J., Lessi-Cheimariou, A., Hunt, S., Daví, M., Musson, R. M. W., Garrard, G., Strasser, F. O. and Robertson, C., (2019). Probabilistic seismic hazard assessment for a new-build nuclear power plant site in the UK. *Bulletin of Earthquake Engineering*, 17, 1 – 36. <https://doi.org/10.1007/s10518-018-0441-6>
39. Wells, D. L., and Coppersmith, K. J. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*. 84(4). 974 – 1002.
40. Wiemer, S., Danicu, L., Edwards, B., Marti, M. *et al.* (2016). Seismic Hazard Model 2015 for Switzerland (SUIHaz2015). Swiss Seismological Service (SED), ESH Zürich, doi: 10.12686/a2
41. Woessner, J., Laurentiu, D., Giardini, D., Crowley, H., Cotton, F., Grünthal, G., Valensise, G., Arvidsson, R., Basili, R., Demircioglu, M. B., Hiemer, S., Meletti, C., Musson, R. W., Rovida, A. N., Sesetyan, K., Stucchi, M., and the SHARE Consortium (2015). The 2013 European Seismic Hazard Model: key components and results. *Bulletin of Earthquake Engineering* 13(12), 3553–3596. doi: 10.1007/s10518-015-9795-1.