# The 2020 update of the European Seismic Hazard Model - ESHM20: Model Overview

Laurentiu Danciu, Shyam Nandan, Celso Reyes, Roberto Basili, Graeme Weatherill, Celine Beauval, Andrea Rovida, Susana Vilanova, Karin Sesetyan, Pierre-Yves. Bard, Fabrice Cotton, Stefan Wiemer, Domenico Giardini

EFEHR Technical Report 001 v1.0.0
December 16th, 2021

# The 2020 update of the European Seismic Hazard Model - ESHM20: Model Overview

Laurentiu Danciu[1], Shyam Nandan[1], Celso Reyes[1], Roberto Basili[2], Graeme Weatherill[3], Céline Beauval[4], Andrea Rovida[5], Susana Vilanova[6], Karin Sesetyan[7], Pierre-Yves Bard[4], Fabrice Cotton[3], Stefan Wiemer[1], Domenico Giardini[8]

1. Swiss Seismological Service, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland,
2. Istituto Nazionale di Geofisica e Vulcanologia (INGV), 00143 Rome, Italy,
3. German Research Centre for Geociences (GFZ), Section 2.6 Seismic Hazard & Risk Dynamics, 14473 Potsdam, Germany,
4. Institut des Sciences de la Terre (ISTerre), IRD, 38058 Grenoble, France,
5. Istituto Nazionale di Geofisica e Vulcanologia (INGV), 20133 Milan, Italy,
6. Instituto Superior Tecnico (IST), 1049-001 Lisboa, Portugal,
7. Bogazici University, Kandilli Observatory and Earthquake Research Institute, Department of Earthquake Engineering, 34684, Cengelkoy, Istanbul, Turkey,
8. Institute of Geophysics , ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland

## Rights and permissions



This report aims at documenting the 2020 update of the European Seismic Hazard Model. The findings, comments, statements or recommendations expressed in this report are exclusively of the author(s) and not necessarily reflect the views and policies of the institutions listed here: i.e. i.e. Swiss Seismological Service, ETH Zurich, Istituto Nazionale di Geofisica e Vulcanologia (INGV), German Research Centre for Geociences (GFZ), Institut des Sciences de la Terre (ISTerre), Instituto Superior Tecnico (IST), Bogazici University, Kandilli Observatory and Earthquake Research Institute, Department of Earthquake Engineering, the EFEHR Consortium or the European Union. The core team of ESHM20 has tried to make the information in this product as accurate as possible. However, it does not guarantee that the information is totally accurate or complete. Therefore, you should not solely rely on this information when making a commercial decision. Users of information expressed in this report assume all liability arising from such use. While undertaking to provide practical and accurate information, the authors assume no liability for, nor express or imply any warranty with regard to the information contained hereafter.

EFEHR Consortium is committed to providing web accessible content and results of the ESHM20 via the www.efehr.org platform. If you are having difficulties with accessing this document please email efehr@sed.ethz.ch

## Citation Advice

Danciu L., Nandan S., Reyes C., Basili R., Weatherill G., Beauval C., Rovida A., Vilanova S., Sesetyan K., Bard P-Y., Cotton F., Wiemer S., Giardini D. (2021) - The 2020 update of the European Seismic Hazard Model: Model Overview. EFEHR Technical Report 001, v1.0.0, https://doi.org/10.12686/a15

# Table of contents

# Acknowledgments

We are grateful all researchers, scientist, experts and engineers contributing to the development of the 2020 European Seismic Hazard Model (ESHM20), in various way from data compilation and curation to knowledge transfer during numerous meetings and webinars:

Andrea Antonucci, António Araújo Correia, Sinan Akkar, Kuvvet Atakan, Jure Atanackov, Hamid Sadegh Azar, Stephane Baize, Julien Barriere, Paolo Bazzuro, Maria Belen Benito, Myriam Belvaux, Bjarni Bessason, Dino Bindi, Christian Bosse, Thierry Cameelbeck, Eduardo Cansado Carvalho, Michele Carafa, Fernando Carrilho, Alexandra Carvalho, Carlo Virgilio Cauzzi, Gilles Celli, Eugenio Chioccarelli, Carmen Cioflan, Pasquale Cito, Cécile Cornou, Edward Cushing, Susana Custódio, Snjezana Cvijic-Amulic, Nicola D'Agostino, Vera D'Amico, Nicolas D'Oreye, Atefe Darzi, Mathias Dolsek, John Douglas, Branko Dragicevic, Dejan Dragojević, Stephane Drouet, Blaise Duvernay, João Estevão, Donat Fäh, Ekkehard Fehling, Joao Fonseca, Mariano Garcia-Fernandez, Eulalia Gracia, Gottfried Grünthal, Benedikt Halldorsson, Florian Haslinger, Marian Herak, Jomard Herve, Stefan Hiemer, Fabrice Hollander, Shible Hussein, Iunio Iervolinio, Diethelm Kaiser, Vanja Kastelic, Amir M. Kaynia, Rexhep Koci, Annakaisa Korja, Sreeram Kotha, Svetlana Kovacevic, Milad Kowsari, Olga-Joan Ktenidou, Daniela Kuehn, Kresimir Kuk, Neki Kuka, Pierre Labbe, Steffi Lammers, Giovani Lanzano, Björn Lund, Lucia Luzi, Francesco Maesano, Elena Manea, Päivi Mäntyniemi, Basil Margaris, Christophe Martin, Warner Marzocchi, Frederic Masson, Luís Matias, Carlo Meletti, Nikolaos Melis, Albero Michelini, Jadranka Mihaljevic, Zoran Milutinovic, José Antonio Peláez Montilla, Ilaria Mosca, Roger Musson, Shemsi Mustafa, Adrien Oth, Bruno Pace, Roberto Paolucci, Christos Papaioannou, Florin Pavel, Laura Peruzza, Rui Pinho, Kyriazis Pitilakis, Valerio Poggi, Mircea Radulian, Evi Riga, Philippe Roth, Agathe Roulle, Zafeiria Roumelioti, Radmila Salic, Abdullah Sandikkaya, Antoine Schlupp, Jochen Schwarz, Anselm Smolka, Oona Scotti, Barbara Sket-Motnikar, Vitor Silva, Anne Socquet, Efthimios Sokos, Mathilde Sorensen, Carlos Sousa Oliveira, Thomas Spies, Max Stuchi, Nikolaos Theodoulidis, Mara Tiberti, Paola Traversa, Nino Tsereteli, Radu Vacareanu, Dina Vales, Roberto Vallone, Dimitrios Vamvatsikos, Kris Vanneste, Emmanuel Viallet, Daniele Vigano, Francesco Visini, Stefan Weginger, Polona Zupancic

Special thanks to the SERA project management and logistics: Kauzar Saleh Contell and Romano Meier.

We would like to thank IT team at Swiss Seismological Service, ETH Zurich for their excellent support on maintaining the servers, update the databases and upgrade the EFEHR web-page and the *hazard.EFEHR* web-platform to access the 2020 models: Philipp Kästli, Leandra Eberle, Emil Zylis, Emilia Petronio and Cyrill Bonjour.

Furthermore, we would like to also thank Helen Crowley and Jamal al Dabeek for their efforts and feedback regarding the use of the seismic hazard for risk calculation.

We would like to extend our gratitude to Michèle Marti, Nadja Valenzuela, Irina Dallo and Simone Zaugg for their creativity and coordination of the outreach activities.

In addition, we would like to thank Marco Pagani for scientific discussions and feedback, and Michele Simionato for timely OpenQuake improvements.

Finally, the authors would wish to express their appreciation to the SC8 Ad-hoc Working Groups for the excellent collaboration during the ESHM20 evaluation: Pierre Labbe, Philippe Bisch, António Araújo Correia, Hamid Sadegh Azar, Alexandra Carvalho, Matjaž Dolšek, Blaise Duvernay, João Estevão, Ekkehard Fehling, Max Gündel, Iunio Iervolino, Christophe Martin, Roberto Paolucci, Evangelia Peli, Kyriasis Pitilakis, Jochen Schwarz, Mathilde Sorensen, Evi Riga, Kris Vanneste, Dimitrios Vamvatsikos, Emanuel Viallet, Radu Vacareanu

# Executive Summary

The 2020 European Seismic Hazard Model (ESHM20) provides an update of the earthquake hazard assessment of the Euro-Mediterranean region. ESHM20 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.s 730900 of the SERA-Project (www.sera-eu.org)

The ESHM20 follows the same principles as the ESHM13, with state-of-the art procedures homogeneously applied for the entire pan-European region, without the country-borders issues. The model was built upon recently compiled datasets (i.e. earthquake catalogues, active faults, ground shaking recordings), information (tectonic and geological) and models (seismogenic sources, ground shaking). A fully probabilistic framework was adopted in the hazard model implementation and all datasets and inputs are fully cross-border harmonized.

The newly developed seismogenic source model encompass fully harmonized and cross borders seismogenic sources following the recent national earthquake hazard models. The inherent uncertainties in characterizing the earthquake rupture forecast are handled by a complex logic tree, consisting of two main models (branches): an area source-based model and a hybrid fault-smoothed seismicity model. The ground motion characteristic model is built upon the most complete ground shaking recordings in Europe and aims at capturing the effects of source and attenuation path of the expected ground shaking at a site.

The regional variability of the ground shaking as constrain by observations is captured by a novel approach of spatial clustering and the overall uncertainties are handled in a backbone logic tree. The master logic tree combines the earthquake rate forecast with the backbone ground motion models into a computational model for assessing the earthquake ground shaking at across the entire Euro-Mediterranean region. Full sets of hazard results (i.e. hazard curves and maps, uniform hazard spectra) for various combinations of return periods and descriptive statistics (mean, median and quantiles) are available.

The ESHM20 development process involved several regional workshops where the scientists and experts were consulted and their feedback was acknowledged and considered on the model outcome. Finally, ESHM20 provides the basis to derive informative hazard maps for two key engineering parameters defining the anchoring points of the seismic design spectra for the next version of the European Seismic Design Code (CEN-EC8).

The European Facilities for Earthquake Hazard and Risk (EFEHR) will maintain and further develop this model in collaboration with the GEM Foundation and the European Plate Observing System (EPOS).

# 1 Introduction

The 2020 European Seismic Hazard Model (ESHM20) was developed within the EU founded project "Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe" (SERA), under the call INFRAIA-01-2016-2017 Research Infrastructure for Earthquake Hazard in Europe. ESHM20 was initiated within the SERA-Joint Research Activities (JRA3) with the following objectives:

- Revisit the 2013 European Seismic Hazard Model (ESHM13, Wössner et al 2015) to identify the input datasets (earthquake catalogues, active faults, ground motion recordings) and the components to be updated.
- Develop and updated seismic hazard model based on technically sound methods and ensure the quantification of inherent uncertainties.
- Build upon the principles of the ESHM13: consensus and acceptance of the input models, transparency and open access of input datasets, components and output.
- Update the seismogenic source part (catalogues, faults, sources) of ESHM13 in time for the EC 8 revision, so that the 2nd generation codes are based on the most recent and most state-of-the art information.
- Revise the ESHM13 Ground Motion Prediction Equation (GMPE) logic tree in light of the newly available ground motion data.
- Interact with CEN/TC250/SC8 committee as well as with key experts at the national level to ensure the correct information and timely implementation of the ESHM20, and extend the output to serve additional engineering requirements as part of design standard EC 8 update and revision.
- Embed the ESHM20 into the Global Earthquake Mosaic (Pagani et al 2018), ensuring consistency with the overall global model.
- Enable earthquake risk modelling in Europe at regional (ESHRM20, Crowley et al 2021) scale

## 1.1 Framework and Model Building Philosophy

An open and collaborative framework was adopted for the development of the 2020 update of the European seismic hazard assessment. The work was coordinated and conducted by a core team, mainly the authors of this report, with contributions from a large community of earth scientists, seismologists, geologists, earthquake engineers, statisticians, software developers and outreach specialists (see a full list in the Acknowledgements).

The role of the collaborators is critical as they contribute in various ways: provide access to the raw data, provide access to the national hazard models, assess the quality of the datasets and models, review the science and methods, provide technical advancements, etc. Their contribution was integrated in numerous meetings including a) bilateral meetings between the core-team members and the local experts b) regional review meetings and c) community webinars.

The bilateral meetings are an excellent opportunity to interact with the local experts, discuss the scientific methods and the assumption underlying the model development, evaluate the quality and curate the main input datasets, seek consensus on the cross-border harmonization, review the main components of the results, file exchange and provide feedback. This interaction between the various actors involved at regional and national level is essential to the acceptance of the ESHM20 at the national level.

**Regional Meetings**

The objectives of the regional meetings were to enhance community understanding and stimulate open discussion on the model development. These meetings were attended in person by about 150 European experts and researchers. Furthermore, the regional meetings provided the opportunity for the ESHM20

core team to present the status update on data compilation, curation and harmonization, development of the seismogenic sources, ground motion models, model implementation, outputs and results spanning across Europe without limitations of the country boundaries. Four regional review workshops were organized in Lisbon (June, 6th to 7th ,2019), Potsdam, June (12th to 13th2019), Athens, July (2nd to 3rd, 2019) and Pavia, (October 14th, 2019). An additional workshop was organized in Luxembourg, March 27th, 2019 with focus on the ground motion characterisation model.

**Public Webinars**

A public webinar was organized on September 9, 2020 – as a scientific session of the EFEHR Consortium General Assembly. The webinar was attended online by an audience of about 250 registered participants. The feedback received during the online webinar was summarized in a general response and shared with the community. Meeting presentations, supporting materials and results made available during the review and consultation process until October 2021.

Another online webinar was organized as a special session on the 37th General Assembly of the European Seismological Commission, where all the components of ESHM20 and results were presented.

**Interaction with the CEN/TC 250/SC 8 ad-hoc working groups**

An important aim of the ESHM20 is to provide the basis for deriving the engineering products for CEN/TC 250/SC 8, i.e., the subcommittee of national experts that is responsible for the development of Eurocode 8 within the European Committee for Standardization (CEN). Two *ad-hoc* working groups i.e. AHG1 with focus on site conditions and AHG2 with focus on seismic hazard evaluation.

The scope of these two ad-*hoc g*roups was to analyse the definition of seismic action in EN 1998, interact with the ESHM20 core team, analyse the comments and feedback, with a view to proposing to SC8 a consensual version of the respective clauses in order to finalise the corresponding informative annex in EN 1998-1-1. Furthermore, the ESHM20 was evaluated with different comparison and testing activities for several countries in Europe (in alphabetic order): France, Greece, Italy, Norway, Portugal, Romania, Slovenia and Switzerland.

To support this evaluation the core team provided the support and access to all materials, input files, results, documentation and presentations. Several bilateral meetings and four plenary meetings took place in the last phase of the model development cycle (2020 to 2021). At the moment when this report is drafted, although a consensus has not yet been reached on some points, on the basis of positions expressed by the (vast) majority of participants, the recommendation of the convener of AHG2 to SC 8 was that $S_\alpha$ and $S_\beta$ maps derived from the median ESHM20 map at 475 years return period are incorporated in EN 1998-1-1 *(Doc N1111 of CEN/TC 250/SC 8).*

After discussion during the 41st SC 8 meeting, SC 8 approved the inclusion of the ESHM20 maps in Annex A of prEN 1998-1-1:2021, which will now undergo the enquiry and formal voting procedure by CEN member countries.

**Interaction with the CEN SC8 ad-hoc working groups**

An important aim of the ESHM20 is to provide the basis for deriving the engineering products for CEN/TC 250/SC 8, i.e., the subcommittee of national experts that is responsible for the development of Eurocode 8 within the European Committee for Standardization (CEN). Two *ad-hoc* working groups i.e. AHG1 with focus on site amplification and AHG2 with focus on seismic hazard evaluation.

The scope of these two ad-*hoc g*roups was to analyse the definition of seismic action in EN 1998, interact with the ESHM20 core team, analyse the comments and feedback, with a view to proposing to SC8 a consensual version of the respective clauses in order to finalise the corresponding informative annex in EN 1998-1-1. Furthermore, the ESHM20 was evaluated with different comparison and testing activities for several countries in Europe (in alphabetic order): France, Greece, Italy, Norway, Portugal, Romania, Slovenia and Switzerland.

To support this evaluation the core team provided the support and access to all materials, input files, results, documentation and presentations. Several bilateral meetings and four plenary meetings took place in the last phase of the model development cycle (2020 to 2021). At the moment when this report is drafted, although a consensus has not yet been reached on some points, on the basis of positions expressed by the (vast) majority of participants, recommendation of the convener to SC8 is that $S_\alpha$ and $S_\beta$ maps derived from the median ESHM20 map at 475 years return period are incorporated in EN1998-1-1 Annex G *(Pierre Labbe, lead of Sc8-AHG2, personal communication)*

**Interaction with the Global Earthquake Model**
During the development phase of the ESHM20, there was an excellent collaboration with the GEM's hazard and software developing team. Model calculation was done with the OpenQuake open-source software (Pagani et al 2014) which was constantly improved during the model development cycle. ESHM20 will replace the ESHM13 with the Global Earthquake Mosaic (Pagani et al 2021).

**ESHM20 release cycle**:

The model was developed between September 2017 and November 2021. In the first 2 years, we focused on compiling and curating the main datasets, designing the blueprint of the models and approaches, software development, and defining the computational workflow. During the following year, the model was implemented iteratively, with calculations undertaken on these preliminary models. In the final year the model was extensively reviewed and quality assurance procedures were established. The model release cycle was:

- ESHM20 *v08* – community preview version, released to partners, collaborators and to the hazard and risk community in September 2020
- ESHM20 *v12d* – beta version, distributed to partners, collaborators in February 2021
- ESHM20 v12e – final version, to be released in December 2021

## 1.2 Outline of Report

In summary, this report provides an overview of the ESHM20 with focus on the main input datasets, main components, calculation framework and results.
Following the background information and scope of the study, Chapter 2 of this report describes the main input datasets from earthquake catalogues, active faults, subduction sources and regional tectonic regionalization. We discuss the methods applied to each dataset, the harmonization strategies and procedures to curate these datasets.

Chapter 3 provides an overview of the unified earthquake catalogue, which is the foundation of any seismic hazard assessment. The newly developed completeness analysis of the unified earthquake catalogue is given also together with insights of the declustering analysis.

Chapter 4 discusses the overall seismogenic sources, reviews the model assumptions, describes the main sources i.e. the area sources, the active faults, background smoothed seismicity and the logic tree.

Chapter 5 addresses the development of the ground motion characteristic model. We provide the insights of the back-bone approach to characterize this uncertainty for seismicity in Europe, incorporating region-to-region source and attenuation variability based on European strong motion data.

Chapter 6 gives guidance on model implementation and hazard calculation settings for the OpenQuake software (Pagani et al 2014).

The main results are summarized in Chapter 7 and Chapter 8 indicates the structure of the inputs with focus on file formats and description of the main attributes. Links to the main repositories and data resources are provided also in this chapter.

The present report does not provide or repeat information on the overall probabilistic seismic hazard assessment (PSHA); rather our purpose is to offer additional, practical insights into the development and use of the ESHM20. A special issue on a peer-reviewed journal is in preparation at the time of writing this report.

# 1 Main Input Datasets

## 1.1 Historical Earthquake Catalogue

### 1.1.1 Background and compilation strategy

The 1000-1899 seismic catalogue compiled for the ESHM20 is the European PreInstrumental earthquake CAtalogue EPICA (https://doi.org/10.13127/epica.1.1;Rovida and Antonucci, 2021a), an update of SHEEC 1000-1899 (https://doi.org/10.6092/ingv.it-share-sheec.1000-1899; Stucchi et al., 2013) based on the same main principles - mostly transparency, repeatability, and continent-wide harmonization of data - and compilation strategies and methods.

As detailed in Stucchi et al. (2013), SHEEC 1000-1899 relied upon the data gathered in the European Archive of Historical Earthquake Data AHEAD (https://doi.org/10.6092/ingv.it-ahead; Albini et al. 2013; Locati et al. 2014; Rovida and Locati, 2015), including both macroseismic intensity data (hereafter MDP, Macroseismic Data Point) supplied by descriptive historical seismological studies and macroseismic databases, and parameters contained in regional catalogues.

For compiling the catalogue, such data were thoroughly analyzed in order to select the most representative of the knowledge of each earthquake, independently from national constraints. From the selected MDPs distributions, earthquake location and magnitude were determined with three methods based on regional attenuation models of macroseismic intensity with magnitude and source distance. In addition, parameters from regional catalogues, selected from AHEAD according to their reliability and public availability, were also considered. The earthquake parameters consisted of locations and magnitude selected among those calculated from MDPs and derived from regional catalogues according to a priority scheme.

The same procedures were applied in the compilation of EPICA, and they are summarized in this report and detailed in dedicated report currently in preparation (Rovida and Antonucci, 2021b). In particular, for the compilation of EPICA version 1.1 as the new version of SHEEC 1000-1899, all the updates concerning both i) input macroseismic datasets and catalogues, and ii) regional calibrations of the methods for parameterizing MDPs, or new methods, were surveyed.

However, since the publication of SHEEC 1000-1899 in 2013, only historical earthquake data have been subject to updates and innovations, whereas data that can significantly improve the calibration have not been published, as detailed in Basili et al., (2018). A new Boxer's calibration has been realized for Italy, only (Rovida et al., 2020a), and no new robust method for deriving location and magnitude from intensity data has been recently proposed.

For these reasons, the compilation of EPICA version 1.1 focused on new input datasets, whereas the parametrization procedures were maintained the same of SHEEC 1000-1899. EPICA version 1.1 contains 5703 earthquakes with either maximum intensity ≥5 or Mw ≥4.0 occurred between 1000 and 1899.

### 1.1.2 Input data

The input data of EPICA were provided by the European Archive of Historical Earthquake Data AHEAD, which collects, systematically organizes and makes available data on European earthquakes in the period 1000-1899 from i) regional and national macroseismic databases, ii) seismological descriptive studies on specific earthquakes, periods or areas, and iii) major modern parametric catalogues.

AHEAD connects multiple datasets referred to the same earthquake, enabling their analysis and comparison for the identification of the most updated and informative one. For both SHEEC 1000-1899

and EPICA, the selection considered the availability of MDPs, the features of the supporting historical seismological research and the update of each dataset, and ignored the national provenance of it. In the end, AHEAD provided EPICA version 1.1 with 160 sources of intensity data and 39 parametric catalogues (see Rovida and Antonucci, 2021b).

The 160 sources of intensity data supply 49852 MDPs related to 3622 earthquakes. These data sources include the nine regional nodes contributing to AHEAD, complemented with the results of several investigations on single or groups of earthquakes, areas and/or periods. As a result, 46 studies providing MDPs to 2014 earthquakes are not among those used for SHEEC 1000-1899, mostly because they are more recent or because they were not considered although already available, e.g. those from the Italian Archive of Historical Earthquake Data ASMI (https://doi.org/10.13127/asmi; Rovida et al., 2017), which did not yet exist at that time.

The 39 regional catalogues EPICA relies upon relate to 5511 earthquakes. For guaranteeing transparency, EPICA exclusively considers published catalogues. Following this principle and for the period of interest (1000-1899), only the following four catalogues were suited for EPICA in addition or substitution of those contributing to SHEEC 1000-1899: CPTI15 (Rovida et al., 2016; 2020a) for Italy, FCAT-17 (Manchuel et al., 2018) for France, Hammerl and Lenhardt (2013) for Lower Austria, and Leydecker (2011) for Germany (although almost identical to the previous version).

### 1.1.3 Earthquake parameters

Parameters in EPICA, as in SHEEC 1000-1899, are assessed from both i) MDPs, with homogeneous and repeatable procedures and ii) regional catalogues, coherently with those from MDPs (Stucchi et al., 2013). The two sets of parameters are then combined in the final parameters, and are both listed in the catalogue file together with the final parameters.

#### *1.1.3.1 Parameters from MDPs*

Location with uncertainty, epicentral intensity, and magnitude from MDPs are jointly determined for 3297 earthquakes out of the 3622 with MDPs, because the intensity distribution of the remaining 325 events are not reliable enough to obtain robust parameters.

Macroseismic parameters are assessed with the same three methods as in SHEEC 1000-1899, i.e Boxer (Gasperini et al., 1999; 2010), Meep (Musson and Jimenéz, 2008) and BW (Bakun and Wentworth, 1997). The intensity attenuation models of each of the three methods are calibrated for five attenuation regions and are the same as in SHEEC 1000-1899 (see Gomez Capera et al., 2015) except for the Apennine-Dinarides (APD) region, substituted with the Boxer's calibration updated for CPTI15 (Rovida et al., 2016; 2020a). In addition, the new I0 to Mw conversion strategy of CPTI15 is adopted in the newly defined Italian Volcanic Areas (IVA) region.

The choice of the method is mostly in favour of Boxer, with Meep and BW used as exceptions respectively for all the events in the UK and for a few offshore events. In each calibration region, a linear conversion relation from epicentral intensity I0 to Mw is derived from the same datasets used for calibrating the three MDPs methods (Gomez Capera et al., 2015; Rovida et al., 2020a).

#### *1.1.3.2 Parameters from regional catalogues*

Out of the 39 selected catalogues, 38 provide 5251 earthquakes in EPICA with location and epicentral intensity or magnitude, or both. Locations from the catalogues are adopted without any modifications, whereas magnitude is determined as in SHEEC 1000-1899, i.e.:

- Mw originally provided by catalogues is adopted without any modifications for 3316 earthquakes;
- otherwise, Mw is re-assessed from epicentral intensity I0 according to the regional conversion relations derived from the datasets used for calibrating the MDP methods (1831 events);
- for 56 cases neither Mw nor I0 is available, and Mw is converted from Ms or ML according to Bungum et al. (2003) and Grünthal et al. (2009); for additional 58 earthquakes the unspecified magnitude type from 4 catalogues is assumed as equivalent to Mw.

#### *1.1.3.3 Final parameters*

Considering the sets of parameters described above, the final parameters in EPICA are determined as follows.
The epicentral location is selected according to the following criteria:

- When only one location from either MDPs or the selected regional catalogue is available, it is adopted
- When both locations from MDPs and from a regional catalogue, priority is given to the epicentre from macroseismic data.

In EPICA, 3297 (57.8%) epicentres are from MDPs, and 2257 (39.6%) from the selected regional catalogue. In addition, 149 epicentres (2.6%), marked as "preliminary", relate to earthquakes for which the available data do not allow a robust determination of the location. Among the locations from MDPs, 3187 are from Boxer, 82 from Meep, and 28 from BW (see **Fig. 1.1**). Moment magnitude is determined as follows:

- When Mw determinations from MDPS and from a regional catalogue are both available, they are combined through the weighted mean of the two values, with arbitrary weights of 0.75 and 0.25 attributed to the Mw from MPDs and from the regional catalogue, respectively. In continuity with SHEEC 1000-1899, reverse weights were given to ECOS-09 and CPTI15.
- When only the Mw from MDPs is available, it is adopted
- When only the Mw from a catalogue is available, it is adopted.

Mw is determined as the weighted mean for 3127 earthquakes (55% of the total), it derives from a regional catalogue for 2124 (37%) earthquakes, and from MDPs for 170 (3%) of them (**Fig. 1.1**). In addition, Mw is not determined for 282 earthquakes because the data they rely upon are not robust enough.

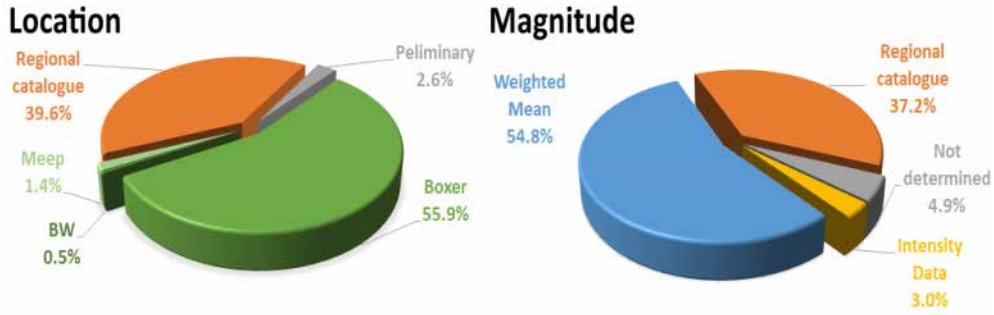



**Fig. 1.1** Origin of the epicentral coordinates (left) and of the magnitude (right) in EPICA.

### 1.1.4 EPICA, the European PreInstrumental earthquake CAtalogue

EPICA, the European PreInstrumental earthquake CAtalogue version 1.1 exploits the best knowledge of the European historical seismicity today available, accounting for the most informative datasets among those in the European Archive of Historical Earthquake Data AHEAD, and maximizes the harmonization of earthquake data and parameters at country borders by means of uniform procedures.

EPICA contains 5703 earthquakes with either maximum observed intensity ≥5 or Mw ≥4.0, for the period 1000-1899. As a whole, the 64% of the earthquakes in EPICA is supported by MDPs, and the 97% is reported in a regional catalogue. In addition, for the 60% of the earthquakes both MDPs and a parametric catalogue are available, while the 36% is known from parametric catalogues only (**Fig. 1.2**). Conversely, the 3% of the earthquakes derive from an historical seismological study and is not included in any published catalogue.

Compared to SHEEC 1000-1899 EPICA contains 1035 new earthquakes, and the number of earthquakes supported by MDPs increased from 2447 to 3622 in EPICA, with 49852 considered MDPs instead of 42581. More than half of the added earthquakes relies on newly published sources of data. In addition, 339 earthquakes from studies or catalogues already considered in SHEEC 1000-1899 are added to EPICA because of the lowered intensity/magnitude threshold or of a revision of the dataset.

Half of the earthquakes in common between EPICA and SHEEC 1000-1899 (i.e. 2332 records) derive from the same sources of data, and thus present the same parameters, whereas 2336 earthquakes listed in SHEEC 1000-1899 are included in EPICA with a new or different set of data. In addition, 49 earthquakes in SHEEC 1000-1899 are not included in EPICA because they have been recognized as fakes, or resulted to be duplications.

The increased number of earthquakes with MDPs indicates an overall improvement of the knowledge of European historical seismicity, being the MDPs distributions the results of thorough historical investigation (see Rovida et al., 2020b).

The wealth of new input data results in an improved reliability and harmonization of earthquake parameters. However, the availability and nature of historical earthquake data across Europe is still inhomogeneous, with a large variability from one area to the other and substantial gaps mostly in Northern and Central-Eastern Europe (Rovida et al. 2020b), which of course cannot but reflect on the uniformity of any Europe-wide catalogue.

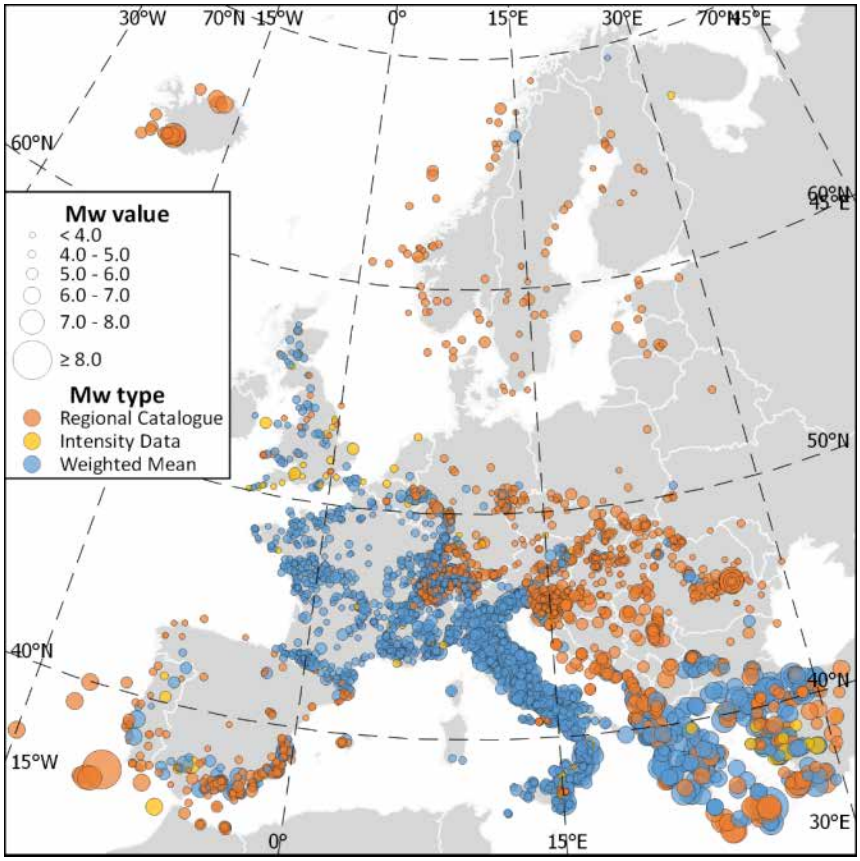



**Fig. 1.2** Earthquakes in EPICA by magnitude value and type, which reflects the typology and reliability of supporting data.

## 1.2 Instrumental Earthquake Catalogue

The "instrumental" earthquake catalogue covers the period 1900 to the end of 2014 and builds on the previous European-Mediterranean Earthquake Catalogue (EMEC) from Grünthal & Wahlström (2012). Although it is referred to as an "instrumental catalogue", the EMEC catalogue and its current update are composite catalogues constructed from local seismicity bulletins, harmonized earthquake catalogues and special studies across various regions of Europe.

Though the majority of events are instrumentally recorded, for much of the early 20$^{th}$ century the catalogues may contain earthquakes that are parameterized from macroseismic data or may contain a mix of macroseismically and/or instrumentally calibrated events depending on the region or data source in question. For many of the data sources, the earthquake magnitudes are calibrated in a mix of scales, including local magnitude ($M_L$) calibrated to different regional conditions, body-wave and surface-wave magnitudes ($m_b$ and $M_S$), moment magnitude ($M_W$) and in some cases maximum intensity ($I_{MAX}$) or epicentral intensity ($I_0$). To use the catalogue in PSHA, in addition to a range of other applications, these must be harmonized into a common scale directly taken from, or equivalent to, moment magnitude $M_W$.

The compilation procedure to update the EMEC catalogue follows closely the approach Grünthal & Wahlström (2012), which was modified slightly for application to the ESHM13 by Grünthal et al. (2013). Both catalogues begin in 1000 CE, but for the ESHM20 only the period from 1900 CE to 2014 CE is considered.

The update incorporates recently published data sets and special studies where available within certain regions and aims to extend the time-period covered from 1000 CE to the end of 2014 CE. To ensure a degree of continuity with respect to the earthquake catalogues adopted by the ESHM13, the same

hierarchy of data sources and the magnitude conversion formulae adopted are retained from the original work of Grünthal & Wahlström (2012), except where new data sets necessitate a revision to this process.

### 1.2.1 Data collection and regionalisation

The construction of the EMEC catalogue is based around a geographical/political regionalization of Europe, reflecting the extent of either political territories of each country or, in some cases, regions of influence that may be treated as a single region in the catalogue compilation.

The regionalization is shown in **Fig. 1.3**. The data collection process identifies new bulletins, compiled catalogue and special studies. The latter may consist of investigations into seismic sequences for which moment tensors were determined and corresponding moment magnitudes defined.

Events from each bulletin (in the form of a date, time, location, magnitude and data source) are added to a master database. Where new representations of existing events in the database are found, these are associated to the existing events based on simple spatio-temporal proximity criteria.

Among the new data sources integrated into the updated EMEC database are recently updated harmonized compilations in Italy (CPTI15, Rovida et al. 2020), France (F-CAT, Manchuel et al. 2018), Turkey (Kadirioglu et al. 2018), Slovenia (Zivcic, 2018) and Romania (INFP, 2018). Bulletin data comprises recent events reported by various seismic networks across Europe. A complete region-by-region description of the data sources and hierarchies is prepared in a separate report.

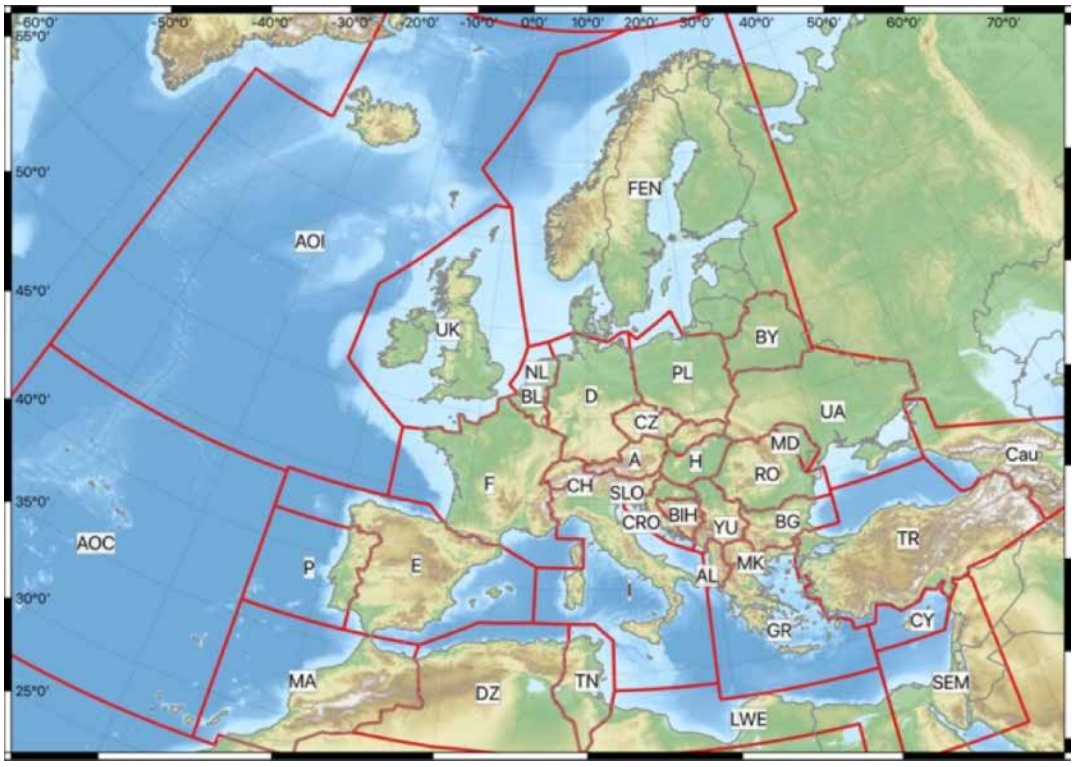



**Fig. 1.3** EMEC regions for data compilation

### 1.2.2 Harmonisation

Data are harmonized using a hierarchical strategy, prioritizing existing $M_W$ harmonizied catalogues and special studies, then considering moment tensor databases and subsequently local bulletins (without $M_W$). Finally, where no local bulletin can be identified, earthquake information is retrieved from the International Seismological Centre (ISC).

For local bulletins and ISC events, the magnitude conversions adopted by Grünthal & Wahlström (2012) are retained, with the exception of a small number of cases where updated conversion formulae are identified in the literature. It is important to emphasise that in this approach the local sources of information are treated as the most authoritative sources, particularly when considering existing harmonized bulletins.

This aims to ensure some level of consistency with local scale catalogues that have been used in national seismic hazard modelling projects within Europe, though we note that differences nonetheless inevitably emerge as magnitudes and locations may change for events on the borders of respective countries.

The updated European-Mediterranean earthquake catalogue is shown in **Fig. 1.4**, and contains 55,732 events with $M_W$ (or equivalent proxy) ≥ 3.5 in the period 1900 to the end of 2014. The density of events per magnitude bin and year is shown in **Fig. 1.5**, alongside the equivalent plot from the Grünthal & Wahlström (2012) catalogue and the difference, i.e. the increase (or decrease) in number of events per cell. The majority of new data comes from the post-2006 period and from dropping the minimum magnitude in southern Europe from $M_W$ 4.0 to 3.5.

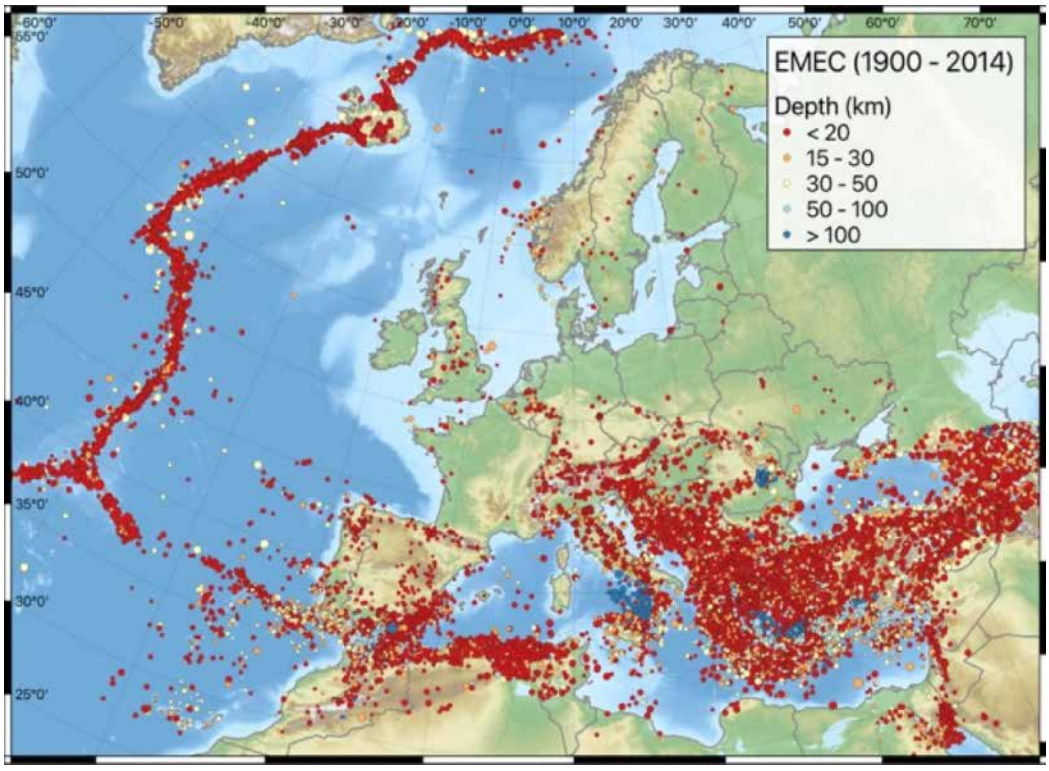



**Fig. 1.4** The complete EMEC catalogue for the period 1900 – 2012.

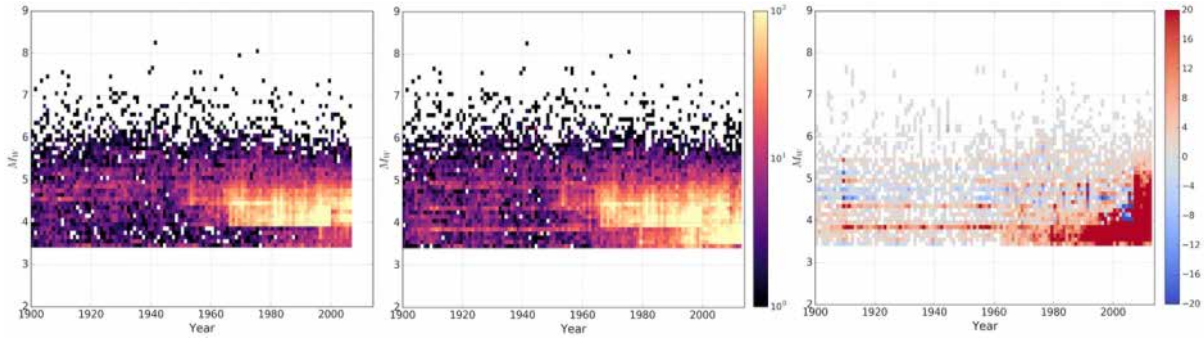



**Fig. 1.5** Density of events per year and 0.1 $M_W$ bin width for the original EMEC catalogue of Grünthal & Wahlström (2012) (left), the current version (centre) and the change in number of events per cell (right).

## 1.3 Active Crustal Faults and Subduction Zones

This section summarizes the main elements of the European Fault-Source Model 2020 (EFSM20), developed as an update of the European Database of Seismogenic Faults 2013 (EDSF13; Basili et al., 2013). For this update, we considered primarily the compilations that covered with a consistent approach significantly large regions, relying on the work of the authors of each compilation for what regards the accuracy and recentness of the information. We resorted to working on individual studies or original work only in case of undefined situations, e.g., area of overlap between two regional compilations or cases where a significant update was available or where the fault information was not covered by EDSF13 but deemed necessary.

In EFSM20, two main categories of seismogenic faults are considered: 1) crustal faults; and 2) subduction zones. The collated datasets and the main curation aspects are summarized in the next two Sections. The map of collated datasets is shown in **Fig. 1.6**. The model covers an area that encompasses a buffer of 300 km around all target European countries (except for Overseas Countries and Territories, OTCs), and a maximum of 300 km depth for slabs.

### 1.3.1 Crustal faults

The minimum set of basic fault parameters required for constructing a seismogenic source model refer to Geometry (Location: Lat, Lon, Depth; Size: Length, Width; Orientation: Strike, Dip) and Behavior (Rake and Slip Rate). These are indispensable elements for devising and applying a fault recurrence model to be expressed by a Magnitude-Frequency Distribution (MFD).

Not all fault compilations provide this characterization in full, and strategies needed to be devised to fill in the missing information. Concerning crustal faults, we identified several regional datasets that vary in date of latest release, geographical extent, level of fault characterization, and more importantly, data formats. In addition to those listed below, several others datasets were considered, such as for example those covering Romania and the northern Black Sea (Diaconescu, Craiu, Moldovan, et al., 2019; Diaconescu, Craiu, Toma-Danila, et al., 2019; Diaconescu et al., 2021). Although these datasets represented a significant advancement with respect to EDSF13, they could not be used because did not fully compliant with the requirements above recalled.

The list below summarizes the main adopted datasets in the various regions (**Fig. 1.6**).

- Dataset #01

This is the original database EDSF13, compiled in the framework of the SHARE project (Basili et al., 2013; Giardini et al., 2013) which covered Europe and the Mediterranean region. This was the starting dataset

to build the new fault model. The largest regions that remained unmodified are in the Balkans and northern Africa. Most regions were either entirely replaced by new datasets or partly revisited according to individual studies. Additions in regions that were not covered by EDSF13 are in Iceland, France, and the northern Mid-Atlantic plate boundary. The major regional updates are summarized below.

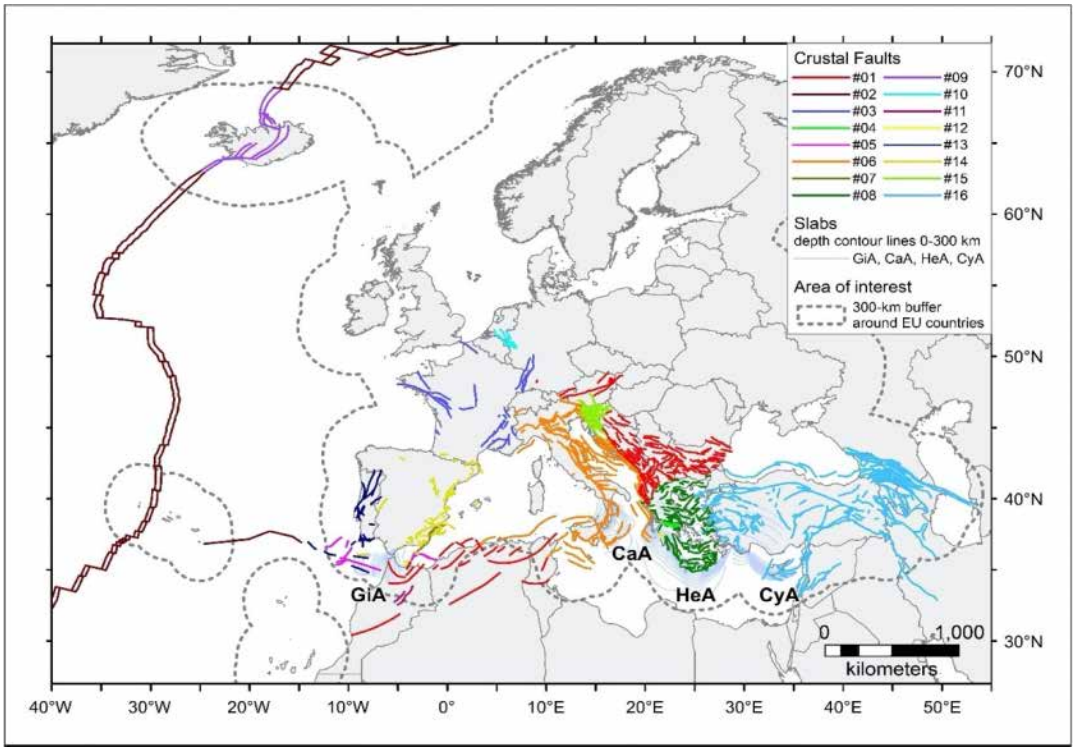



**Fig. 1.6** Map of collated fault datasets for the development of the European Fault-Source Model 2020 (EFSM20). See text for the descriptions of the various datasets. From west to east, the subduction systems are: Gibraltar Arc (GiA); Calabrian Arc (CaA); Hellenic Arc (HeA); and Cyprus Arc (CyA).

- Dataset #02

This dataset covers the Mid-Atlantic ridge and transforms. The initial geometry was derived from a global plate-boundary model (Bird, 2003) and the rest of the characterization was based on the oceanic crust age and spreading rate (Müller et al., 2008). For the transform faults the slip rate is directly derived from the spreading rate, aided by more local data for the Gloria fault (Fadil et al., 2006; Koulali et al., 2011). For the normal faults the slip rate is derived by combining the spreading rate with local information about fault spacing and heave (Escartín et al., 1999; MacDonald & Luyendyk, 1977).

- Dataset #03

This dataset, which covers the French region, is derived from BDFA (Jomard et al., 2017). Due to the different strategy of fault mapping used in BDFA, we redrew the fault traces by interpolation and reassigned several parameters based on recent works in the region.

- Dataset #04

This dataset includes a few faults in the Gulf of Corinth. The initial geometry of the faults is based on GREDASS (Caputo & Pavlides, 2013) and the slip rate were updated based on more recent works (Bell et al., 2009; Fernández-Blanco et al., 2019).

- Dataset #05

This dataset covers the offshore regions of the Gulf of Cadiz and the Alboran Sea. In this area we updated the EDSF based on several recent works, either providing updated geometries and slip rates (Koulali et al., 2011; Martínez-Loriente et al., 2018; Martínez-Loriente et al., 2013; Neres et al., 2016).

- Dataset #06

This dataset covers the Italian territory and some surrounding regions. It is mainly based on the most recent version of the DISS (DISS Working Group, 2021).

- Dataset #07

This dataset includes a few faults in the Eastern Betics region. Such faults were modified with respect to the QAFI according to recent works with substantial updates of fault geometries and slip rates (Borque et al., 2019; Gómez-Novell, Chartier, et al., 2020; Gómez-Novell, García-Mayordomo, et al., 2020; Herrero-Barbero et al., 2020).

- Dataset #08

This dataset covers the Aegean region. It is mainly based on the most recent version of GreDaSS (Caputo & Pavlides, 2013).

- Dataset #09

In Iceland we started from the same approach as the Mid-Atlantic ridge and transforms (Dataset #02) and added data and considerations based on local studies (Árnadóttir et al., 2008; Bergerat et al., 1990; Forslund & Gudmundsson, 1991; Garcia et al., 2002; LaFemina et al., 2005; Rögnvaldsson et al., 1998; Rust & Whitworth, 2019).

- Dataset #10

This dataset deals with the Lower Rhine Graben. In this area we started from EDSF13, already based on a local fault model (Vanneste et al., 2013), and updated the slip rate of several faults based on more recent data (Gold et al., 2017)

- Dataset #11

This dataset covers the northwestern African region (Morocco, Algeria, and Tunisia). In this region, we mainly relied on EDSF13 with updates of a few faults in the Moroccan region based on the GEM Global Active Faults Database (Styron & Pagani, 2020) and various other works (Akoglu et al., 2006; Gomez et al., 1996; Pastor et al., 2015; Rigby, 2008; van der Woerd et al., 2014) for refining several fault parameters.

- Dataset #12

The NOAFAULTS database (Ganas, 2021) was used to integrate the dataset in the Aegean region for faults not already included in GreDaSS.

- Dataset #13

This dataset covers the Lower Tagus Valley Fault Zone (LTVFZ), southern Portugal, based on recent works (Canora et al., 2015) and Algarve (Sanz de Galdeano et al., 2020).

- Dataset #14

This dataset covers most of the Iberian region including the Pyrenees. In this region we relied on the Quaternary Faults Database of Iberia (QAFI) database (García-Mayordomo et al., 2012; IGME, 2015). Due to the different strategy of fault mapping used in QAFI, we redrew the fault traces by interpolation.

- Dataset #15

This dataset deals with Slovenia and its surroundings. In this area we relied on the recently published Database of active faults in Slovenia (Atanackov et al., 2021).

- Dataset #16

This dataset covers Anatolia and parts of the Middle East. In this region, we relied on recent data from the project EMME and data from the national update of the Turkish hazard model (Danciu et al., 2018; Demircioğlu et al., 2018; Emre et al., 2018).

To maximize the use of available fault compilations, crustal faults are represented with a down-dip planar geometry. In several compilations, however, crustal faults are represented only by the trace of the fault upper edge. To homogenize the collated dataset, the fault upper edge was resampled at a regular spacing of ~5 km. Whenever a fault was represented by multiple strands, the upper edge was reduced to a single trace by interpolation and smoothing. The fault plane was then extruded downward based on the dip and depth values (**Fig. 1.7**).

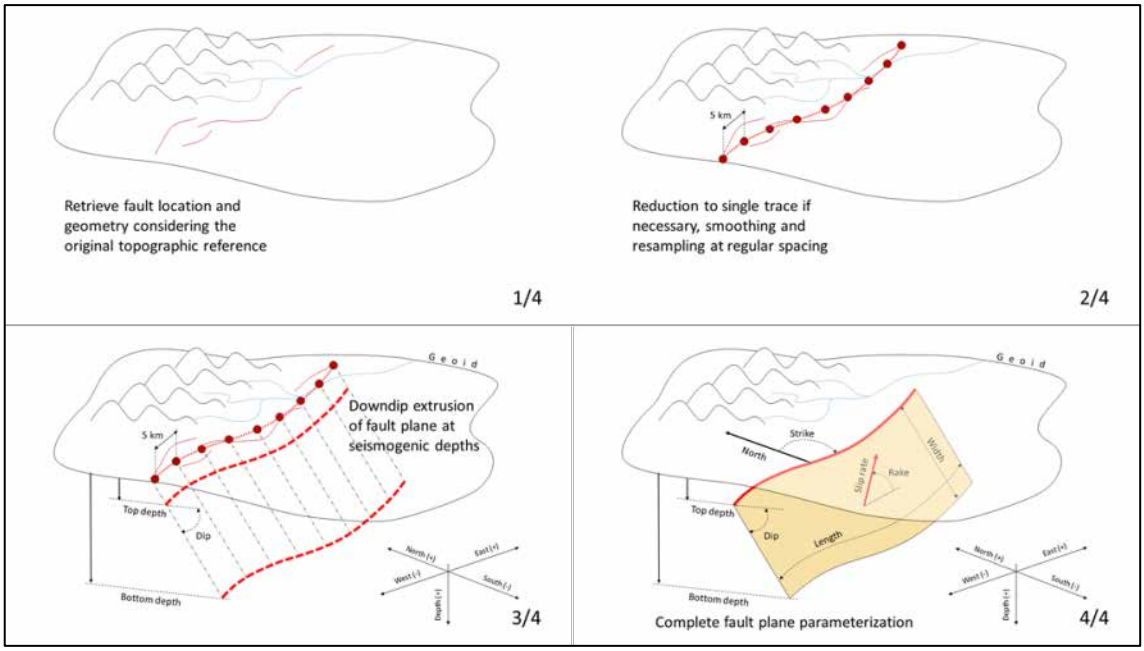



**Fig. 1.7** Schematic showing the geometric construction of the fault-source model.

For all datasets, we verified the depth datum (local ground surface or mean geoid/spheroid), then checked the possible intersections of the fault plane with the topo-bathymetry using the ETOPO1 Global Relief Model (Amante, 2009; NOAA National Geophysical Data Center, 2009) and with the base of the crust using the European Moho (Grad et al., 2009). Also, the possible intersections at depth between extruded fault planes were verified and, occasionally, this circumstance led to removal of a fault partly or even entirely. The overlaps between different datasets were harmonized based on a prioritization scheme which included considerations on newer data and national data, accuracy and justification of the fault information, coherence with the fault characteristics in the surrounding regions.

Once the fault reconstruction is completed, several geometric parameters, such as total length, end-to-end length, down-dip width, and area are calculated and stored in the GIS file attribute table.
For what concern slip rates, we paid attention to harmonize the associated uncertainty. Occasionally, the original datasets we relied upon provided single slip rate values without uncertainty, or the difference between the minimum and maximum value was very small compared with what is usually found in the literature, or did not specified minimum values larger than zero. We thus performed a statistical analysis of the relative uncertainty in all records, considering the different categories of fault type (dip-slip and strike-slip) and tectonic domain (interplate and stable continental regions).

We then adopted the average relative uncertainty (percentage of slip rate), weighted for the number of occurrences in each category, to correct the slip rate variability where it was missing or not reliable. This procedure prevents having apparent small uncertainties for specific faults which could have depended solely on the different strategy of data collection in different datasets or regions.
The maximum earthquake magnitude of the crustal faults is estimated as the magnitude value, in the moment magnitude scale, that corresponds to the largest possible rupture that a fault can host based on its dimensions and magnitude scaling relations (Leonard, 2010, 2014).

The adopted scaling relations are in the form Mw=a+b*Log(S) where S can be the size of any of the following rupture dimensions: end-to-end length, width, area, or displacement, and the parameters “a” and “b” take different values depending on the S, the sense of slip (dip-slip or strike-slip), and the seismotectonic context (interplate or stable continent). The characterization of the faults for adopting the different scaling relations was established by using the provided rake values and following the definition and geographic distribution of the stable continents (Johnston, 1994).

A characterization of the level of activity of each fault is provided by the tectonic moment rate, that can be obtained by the product of fault area multiplied by slip rate multiplied by an assumed rigidity (or shear modulus) of the upper crust.

The parameterization of the crustal faults in EFSM20 is visualized in the maps shown in **Fig. 1.8**. The distribution of occurrences of the different parameters are displayed in the histograms of **Fig. 1.9**.

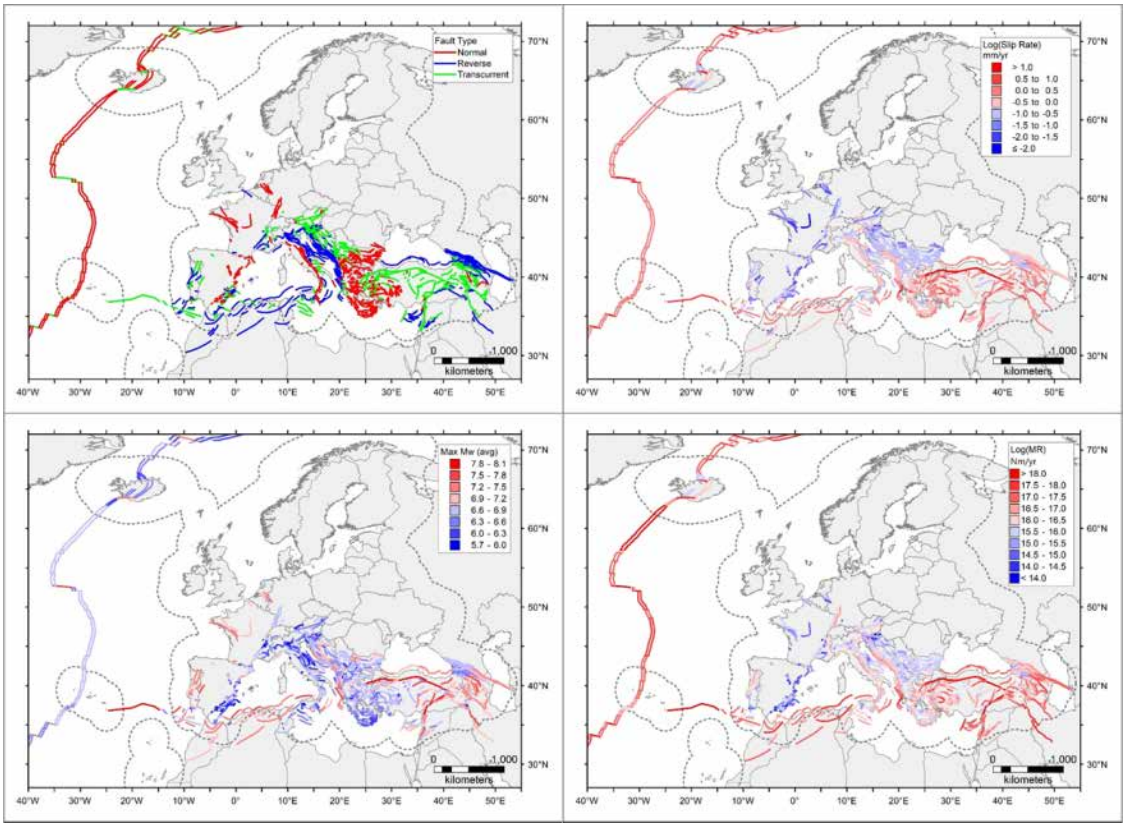



**Fig. 1.8** Maps showing different characteristics of the harmonized fault-source model. From top to bottom and left to right, the four panels show the faulting types (normal, reverse, transcurrent), the maximum magnitude, the slip rates and the moment rates. Slip rates and moment rates are color-coded based on a classification in the logarithmic scale.

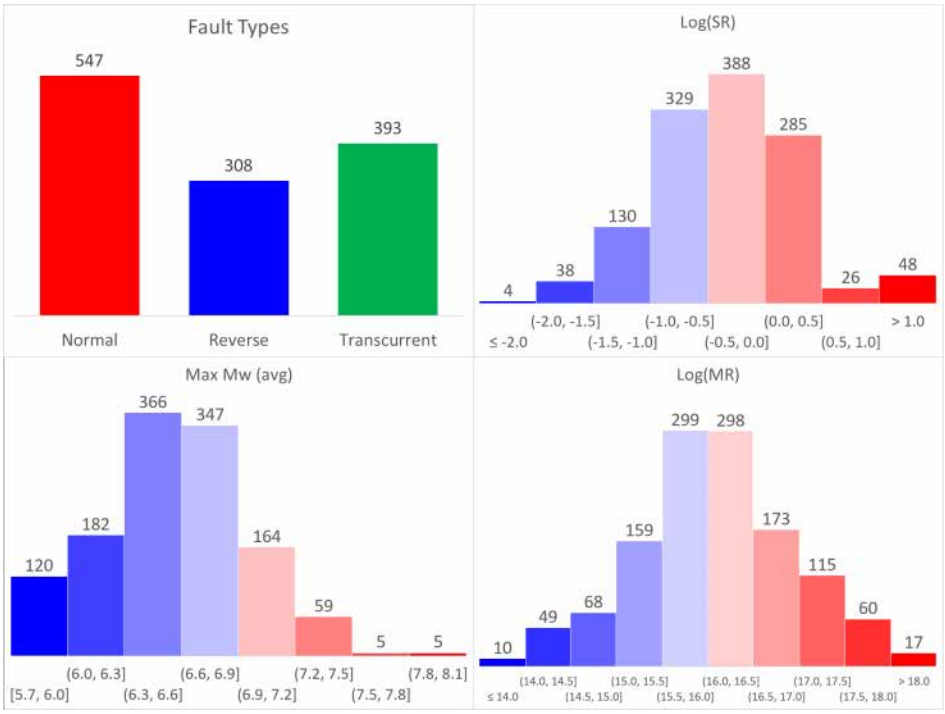



**Fig. 1.9** Histograms showing the frequency of occurrence of the different characteristics of the harmonized fault-source model. From top to bottom and left to right, the four panels show the faulting types (normal, reverse, transcurrent), the maximum magnitude, the slip rates and the moment rates. Slip rates and moment rates are color-coded based on a classification in the logarithmic scale. See correspondence with the geographic distributions fin Figure 3.

### 1.3.2 Subduction Zones

Subduction zones are complex systems where different types of potential earthquake sources exist. We rely on the assumption that the earthquakes occurring in the upper plate of a subduction zones are dealt with by the crustal fault model. We thus designed the subduction model to prospectively assess the slab interface seismicity and the intraslab seismicity. To this end the main necessary element is the tri-dimensional geometric reconstruction of the lower plate top surface. In addition, the minimum set of basic parameters required for constructing a seismogenic source model are the crustal thickness of the subducting plate, the upper and lower depths of the seismic interface, and the net convergence direction and rate.

Concerning subduction sources, we identified several datasets that vary in date of the latest release, level of characterization, and data formats. We started with EDSF13, the original database compiled in the framework of the SHARE project (Basili et al., 2013; Giardini et al., 2013) which covered the Mediterranean subduction systems, which include the Calabrian Arc, the Hellenic Arc, and the Cyprus Arc (Figure 1). This was the starting dataset to build the new subduction zone model. In addition to searching the literature for more updated data, we conducted an original work to reconstruct the subduction system of the Gibraltar Arc (**Fig. 1.6**) which was not included in EDSF13.
The geometry of the Calabrian Arc slab was entirely updated using a large dataset of seismic reflection profiles and the seismicity distribution (Maesano et al., 2017). The geometry of the Hellenic Arc and Cyprus Arc slabs were recently revisited in the framework of a tsunami hazard project (Basili et al., 2021). These datasets have also been re-examined in light of the SLAB 2 model (Hayes et al., 2018) and several others (Ganas & Parsons, 2009; Halpaap et al., 2018, 2019; Sachpazi et al., 2016).

The geometry of the Gibraltar Arc slab was reconstructed using different datasets. For the shallowest depths, we used data of bedrock markers based on the interpretation of multichannel seismic reflection profiles and wide-angle seismic surveys (Gutscher et al., 2009) assuming that the top of the slab coincides with the top of the basement. For the intermediate depths (12-40 km), we used a model of the Moho obtained from a set of diverse datasets using a probabilistic surface reconstruction algorithm (Arroucau et al., 2021) and considering typical values for the old Tethys oceanic crust in the range of 7-9 km (Sallarès et al., 2011). Then, between 40-70 km depth, we obtained the position of the slab by interpolation of seismicity clusters. Within the 140-200 km depth range, the slab was assumed to be vertical based on a tomographic model (Civiero et al., 2018), which shows a nearly vertical high-velocity p-wave anomaly down to 600 km deep.

All these 3D geometries were then resampled at regular spacing and smoothed to ensure the same spatial resolution in the different models. Considering the crustal thickness of the lower plate, the geometry reconstruction is then completed with a lattice of 10x10 km laying on the slab intermediate surface. The various steps of the slab geometric construction are shown in **Fig. 1.10**, while the resulting geometries of the four slabs are shown by oblique views in **Fig. 1.11**.

The upper and lower depths of the seismic interface was estimated from data and modeling of the 150°C and 350-450°C isotherm, the seismicity distribution, and the slab intersection with the Moho of the upper plate (Davies, 2013; Grad et al., 2009; Gutscher et al., 2006; Heuret et al., 2011; Syracuse et al., 2010; Thiebot & Gutscher, 2006).

Tectonic rates are derived from geodetic observations on the plate convergence across the subduction zones (Carafa et al., 2018; Devoti et al., 2008; Hollenstein et al., 2008; Howell et al., 2017; Nocquet, 2012; Palano et al., 2015; Reilinger et al., 2006; Stich et al., 2006; Wdowinski et al., 2006).
The maximum earthquake magnitude of the slab interface is estimated as the magnitude value, in the moment magnitude scale, that corresponds to the largest possible rupture that the interface can host

based on its dimensions and magnitude scaling relations specifically developed for this type of earthquake ruptures (Allen & Hayes, 2017).

For what concern the tectonic moment rate, in the case of the slab interface an important aspect to consider is the rigidity (shear modulus) variation with depth that has been observed in subduction zones from around the world (Bilek & Lay, 1999; Sallarès & Ranero, 2019) and already used to simulate earthquake ruptures for tsunami simulations (Geist & Bilek, 2001; Scala et al., 2020).

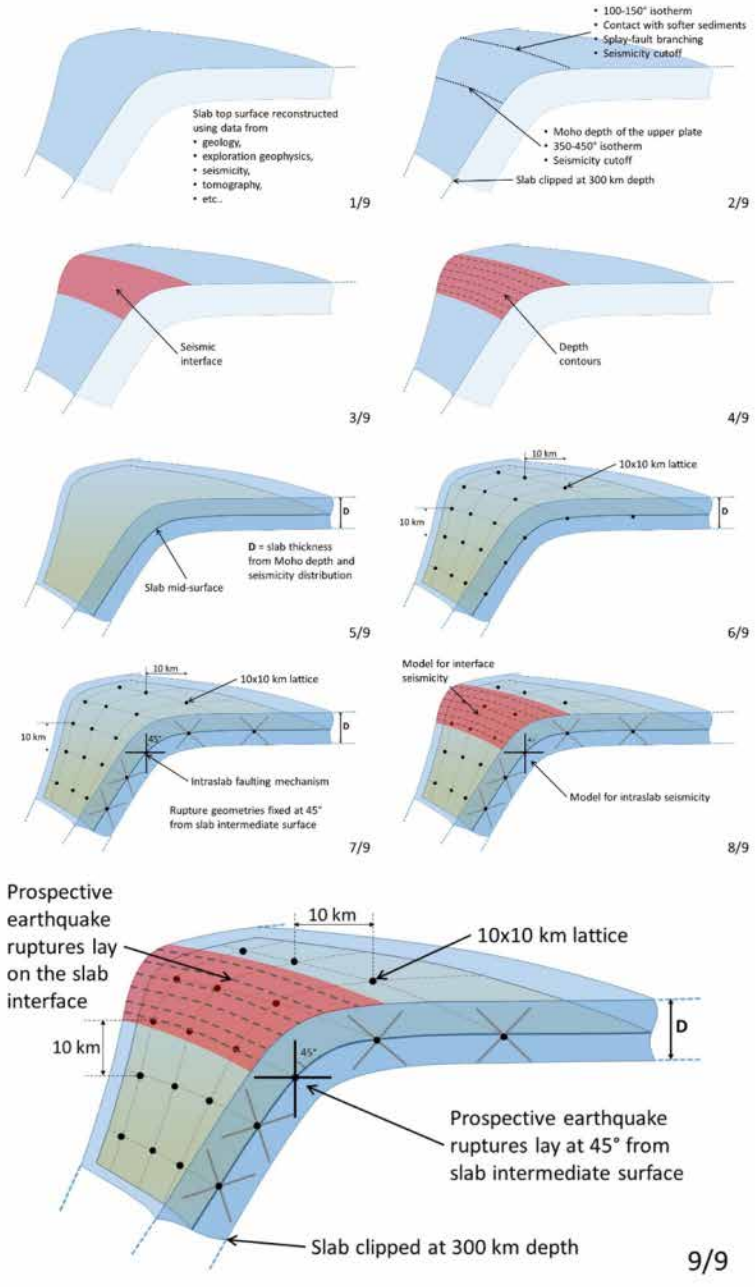



**Fig. 1.10** Schematic showing the geometric construction of the slab model.

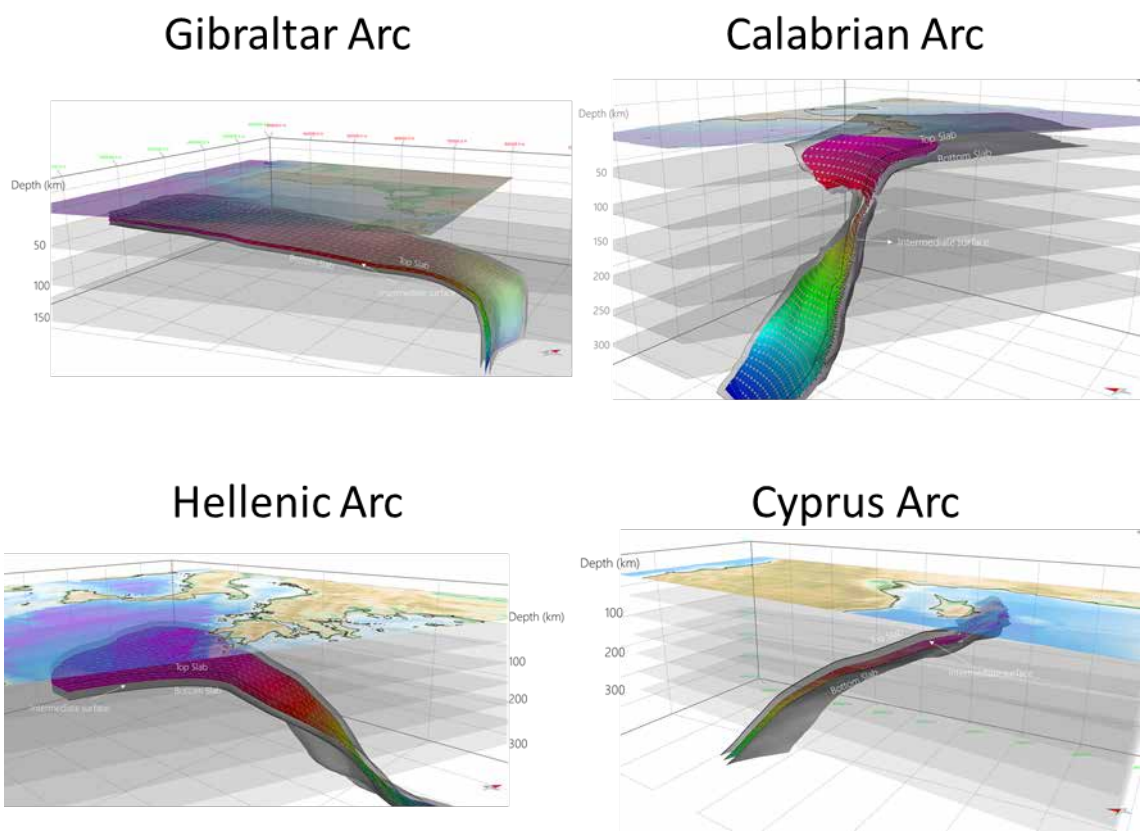



**Fig. 1.11** Oblique views of the tri-dimensional geometry of the four slab models.

## 1.4 Tectonic Regionalization

A tectonic regionalization is a subdivision of the entire source space relevant for the Euro-Mediterranean area into discrete regions that are internally as homogeneous as possible from the standpoint of the dominant tectonic processes.

This regionalization was initially built following basic principles of plate tectonics and building on previous experience of the ESHM13 (Delavaud et al., 2012) and the TSUMAPS-NEAM project (Basili et al., 2021). The following eleven tectonic settings are defined for the crustal level:

1. Active volcano;
2. Back-arc and orogenic collapse;
3. Continental rift;
4. Oceanic rift;
5. Contractional wedge;
6. Accretionary wedge;
7. Conservative plate margin;
8. Transform zones proper;
9. Shield;
10. Stable continental region;
11. Stable oceanic region.

The geographic distribution of the above-defined tectonic settings forms the regionalization shown in **Fig. 1.12**. This construct is a 2D subdivision of the crustal volume. Most of the boundaries between adjacent zones can be assumed to be nearly vertical. However, for some of them this assumption does not hold and further considerations are needed. This is especially the case of the subduction zones,

where the presence of slabs implies a 3D geometry of the regions classified as accretionary wedges. Inclined boundaries exist also elsewhere. In addition, in some cases it should be considered that the transition between one region and another could be gradual and that some regions could have an intermediate classification between those envisaged beforehand.

An important element used to constrain the first draft of the regionalization is represented by the geometry of the main plate boundaries (Bird, 2003). Subdivisions of the plate interiors was mainly based on large-scale geologic maps and tectonic classifications (Asch, 2005; Johnston, 1994; Müller et al., 2008). For the volcanoes, we relied on the global database without modifications (Global Volcanism Program, 2013).

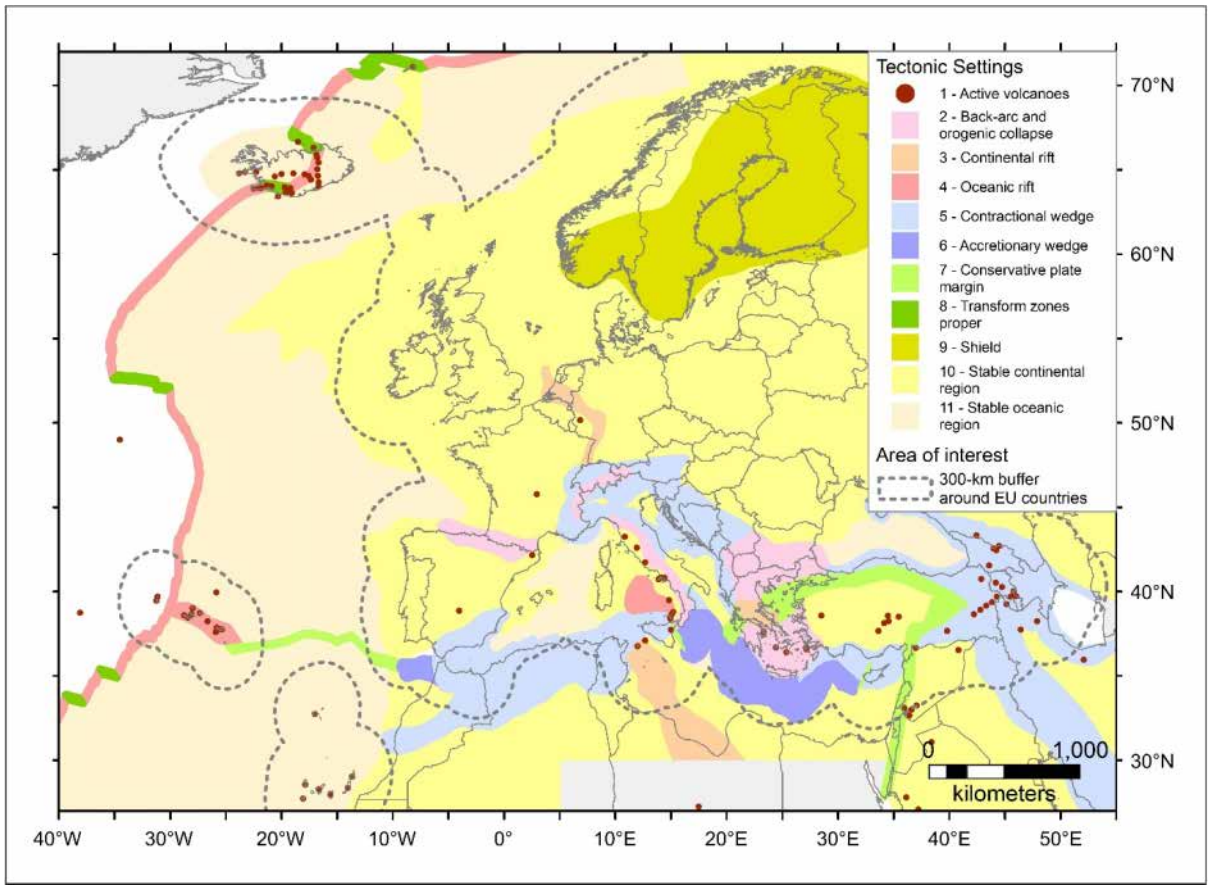



**Fig. 1.12** Map of the tectonic regionalization. Each region is color-coded depending on the tectonic setting.

# 2 Unified Earthquake Catalogue

The unified earthquake catalogue consists of two parts: the so-called "instrumental" catalogue (after 1900) based on the updated EMEC catalogue (Weatherill et al. 2021), and the European PreInstrumental Earthquake Catalogue EPICA v. 1.1 (Rovida and Antonucci, 2021; https://doi.org/10.13127/epica.1.1) earthquake catalogue (between the years 1000 CE and 1899). Both catalogues are documented in Chapter 1.

Table 1 summarizes the unified ESHM20 catalogue and its two main components. For comparison, we also include the summary of the unified ESHM13 catalogue. ESHM20 unified earthquake catalogue includes an additional period of 8 years since the end of 2006 and a larger spatial coverage. Thus, it naturally contains more events. Even for the same period (1000-2007), the ESHM20 catalog is overall richer in the number of events at different magnitude thresholds (**Fig. 2.1**).

However, the ESHM20 catalogue is not homogeneously richer everywhere in space for a given magnitude threshold. For a given magnitude threshold, the difference in the number of events, $\Delta$, between the two catalogs varies widely in space, with some regions featuring more events and others less. We also find that $\Delta$ varies substantially with the magnitude threshold.

Especially striking is the Italian region, where the ESHM20 catalogue features substantially more events at a magnitude threshold of 3.5, while at the magnitude threshold of 4.5, the ESHM13 catalog is comparatively richer as a consequence of the updated Italian catalogue CPTI15 (Rovida et al., 2020).

| | (1) Historical: EPICA v1.1 | (2) Instrumental: EMEC v20190218 | (1)+(2): Unified ESHM20 | Compare Unified ESHM13 |
|---|---|---|---|---|
| Time span | 1000/1899 | 1900/31-12-2014 | 1000/31-12-2014 | 1000/31-12-2006 |
| Magnitude span | 1.9/8.5 | 3.49/8.3 | 1.7/8.5 | 1.7/8.5 |
| Longitude span | -23.5°/32.4° | -37.0°/51.9° | -37.0°/51.9° | -31.65°/45.0° |
| Latitude span | 35.0°/69.4° | 26.9°/73.0° | 26.9°/73.0° | 33.2°/73.32° |
| # Total Number | 5703 | 55411 | 61127 | 30012 |
| # magnitude ≥ 4.5 | 2337 | 20388 | 22728 | 13284 |
| # magnitude ≥ 5.0 | 1550 | 6013 | 7565 | 5585 |
| # magnitude ≥ 5.5 | 885 | 1920 | 2805 | 2066 |
| # NaN depth | 4965 | 9451 | 14422 | 10616 |
| # NaN magnitudes | 282 | 0 | 282 | 303 |
| # depth < 40km | 5355 | 49436 | 54791 | 25666 |

**Table 1** Summary of the updated and unified earthquake catalogue used in ESHM20 and the earthquake catalogue for ESHM13 (SHEEC v3.3).

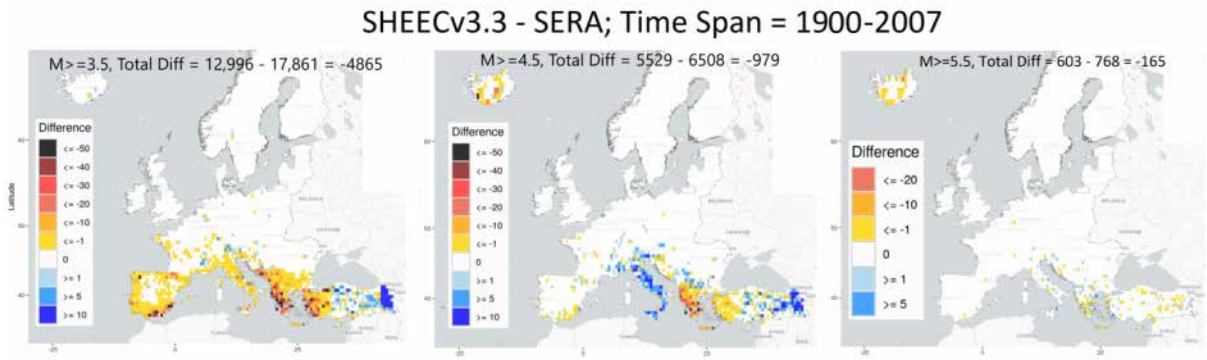



**Fig. 2.1** Spatial variation of the difference in number of events, $\Delta$, between the ESHM20 and ESHM13 catalogs for different magnitude thresholds. Warmer and cooler colors indicate regions where ESHM20 and ESHM13 catalog feature more events, respectively.

Unified Earthquake Catalogue: 2020 update for Iceland

Jónasson et al. (2021) provide an update of the instrumental earthquake catalogue for Iceland. In this updated version the authors have revised the Mw conversion acknowledging that the equations used in the 2013 earthquake catalogue (Grünthal and Wahlström, 2012; Grünthal et al., 2013) resulted in conservative Mw estimates for early events. The differences of Mw between the two catalogues is time-dependent as reported in Jónasson et al (2021), i.e. the average difference is 0.41 before 1970, 0.37 between 1970 and 1980, and 0.27 after 1980; third quartiles are 0.59, 0.47, and 0.36, respectively and the locations are reasonably accurate. Following the recommendation of the Icelandic colleagues, the core team used the earthquake catalogue by Jónasson et al. (2021) for ESHM20.

## 2.1 Catalogue Declustering

One of the assumptions of the time independent probabilistic seismic hazard analysis is that earthquake occurrence times can be represented by a stationary Poisson process. However, catalogues contain dependent events, such as foreshocks aftershocks or triggered earthquakes, and exhibit strong spatial and temporal clustering of seismicity.

Thus, the catalogues of recorded earthquakes usually reveal noticeable departures from the Poisson distribution. To obtain catalogs that follow a stationary Poisson process, the reported catalogs need to be declustered.

Declustering algorithms categorize the earthquakes in a catalogue into groups of independent and triggered occurrences and fulfil two main purposes:

- Creating a catalogue of expected mainshocks using a stationary Poisson process.
- Removing the spatial bias in the seismicity rate across the entire catalogue produced by seismicity's intensive spatio-temporal clustering.

There are numerous declustering techniques available, which reflect the complexity of assigning such labels. Often, the methods are used without considering their effect on the final results.

ESHM13 employed a straightforward windowing scheme based on Grünthal (1985) window sizes. To account for foreshock activity, the original method was slightly modified: if a greater magnitude event is discovered inside the time- and space-window of a detected mainshock, the algorithm retains the foreshock cluster, extends it to the window size of the larger event, and merges all events into one cluster, leaving just the largest event in the declustered catalog. This adjustment is reflected in the implementation of the hazard modeler toolkit (https://github.com/gem/oq-engine; Weatherill et al 2012) by the "fs time prop" input parameter (see table below).

As a first step, we used the SHEECv3.3 catalog from ESHM13 to derive this input parameter and thereby define a reference method. We discovered that setting "fs time prop" to 0.17 results in a declustered catalog that matches the published mainshock catalog for ESHM13. This method will be referred to as our reference approach (winGT fs0.17). In the framework of ESHM20 we investigated the following three main declustering methods:

1. Windowing techniques with space-time windows as defined by Gruenthal (personal communication), (Gardner & Knopoff, 1974) and (Uhrhammer, 1986). The corresponding declustering methods are referred to as winGT, winGK and winUH, respectively. All those windowing techniques are based on the idea that for each earthquake in the catalogue, the following events are labelled as aftershocks if they occur within a specified time-span, and within a distance range. Foreshocks are similarly treated and if the largest earthquake occurs later in the sequence, the foreshock is labelled as the aftershocks. Thus, the temporal-spatial windows are reset according to the magnitude of the largest earthquake in a sequence.

2. Cluster method introduced by (Reasenberg, 1985). The method aims at identifying dependent earthquakes by linking them to clusters according to spatial and temporal interaction zones. Such clusters thus typically grow in size when processing more and more events. The spatial interaction zone is chosen according to the stress distribution near the mainshock, and the temporal extension of the interaction zone is based on Omori's law. A cluster is then composed of all linked events, and the largest earthquake is labelled as the mainshock whereas smaller earthquakes are classified as fore- and aftershocks. The procedure applied here does not account for the uncertainties on the earthquake location.

3. Declustering based on a correlation metric (Zaliapin et al., 2008). The technique is based on a space-time metric to correlate earthquakes with each other. (Zaliapin et al., 2008) showed that by comparing rescaled times to rescaled distances it is possible to identify two distinct populations. These two categories consist of the time-stationary Poissonian mainshock events and the dependent events characterized by much smaller time and space inter-event distances. The method does not provide binary labels of mainshock and aftershock, but gives, for each earthquake, the associated probabilities that it is an independent event.

### 2.1.1 Sensitivity Analysis

We explore the sensitivity of the number of declustered events to different choices of declustering methods and corresponding parameters. We first considered the crustal part of the unified earthquake catalogue for ESHM20 as input for seismicity declustering. We define crustal seismicity as the part of the catalog that features (1) depth values < 40km or (2) *NaN* depth values. This definition leads to a total of 54791 crustal events. We report a large variability in terms of number of mainshocks when comparing the different techniques. This variability is more dominant than the variability given by parameter changes for a single selected method (see **Fig. 2.2**). Thus, we selected three methods for further investigation with the default input parameters given in bold font in Table 2. The results of the sensitivity analysis are given in Table 3. The method that generates the largest number of events after declustering is the Reasenberg method with 43563 events identified as main events, followed by Zaliapin method with 30982 main events and default time windows technique used in ESHM with 22708 main events.

The results of the sensitivity analysis are given in Table 3. The method that generates the largest number of events after declustering is the Reasenberg method with 43563 events identified as main events, followed by Zaliapin method with 30982 main events and default time windows technique used in ESHM with 22708 main events. However, these results do not provide a straightforward answer in which

method to be chosen. Thus, at a glance one could use the default time-window declustering technique, which was used in ESHM13. This method is was calibrated for the seismicity in central Europe, and it might provide an aggressive declustering as the number of events remaining is the lower among the considered declustering techniques. A spatial comparison of the total number of events identified as mains by the three declustering methods is given in **Fig. 2.3**. The comparison is done in grid cells of 50km and only events of magnitude greater than 4.5 are illustrated. It can be seen that when compared with the WT-Grünthal the Reasenberg method identifies more main events in stable continental regions (Central Europe, Iberia, whereas in the southern Europe this trend is reversed. On the same figure (bottom plot), the spatial comparison between the WT-Grünthal the Zaliapin method, suggests that the Zaliapin method might be suitable for very active regions (i.e. Southern Europe), as in the central Europe the number of main events identified are below the WT-Grünthal method. However, these are just some assumptions not supported by any other evidences, hence we progress with a data-driven procedure to aid the selection of the suitable declustering technique. The results of this data driven statictial analysis indicate that the Reasenber technique is an alternative method to WT-Grünthal method. Given the computational demand of the model, it was decided not to include an alternative branch for the declustering algorithm.

**Table 2** Input parameters of the main declustering techniques under investigation for use in declustering the unified earthquake catalogue. The bolded parameters are considered as the reference input parameters.

| Parameter | Declustering Method | Description | Range/(ESHM20 default) |
|---|---|---|---|
| fs_time_prop | Window | fraction of the time window used for foreshocks | [1.0, 0.5, 0.1, 0.17, 0.01, 0.001] |
| taumin | Reasenberg | Minimum look ahead time for clustered events | [1.0] day |
| taumax | | maximum look ahead time for clustered events | [10.0, 50.0 100 300] days |
| xmeff | | effective magnitude to define magnitude cutoff (with xk=0.5) | [3.5, 4.5, 5.5,6.5, 7.5] |
| rfact | | factor for interaction radius for dependent events | [10.0, 20.0 30.0] |
| fractal_dim | Zaliapin | spatial weighting factor | [1.4, 1.6, 1.8] |
| b_value | | magnitude weighting factor | [0.9, 1.0, 1.1] |
| theta | | temporal weighting factor | [0.5, 1.0, 2.0] |

**Table 3** Results of the sensitivity analysis on the declustering techniques.

| Unified catalogue | ESHM20-full | ESHM20-winGT | ESHM20-rbOrig | ESHM20-Zaliapin |
|---|---|---|---|---|
| # Total Number | 54791 | 22708 | 43563 | 30982 |
| # magnitude ≥ 5.0 | 6726 (100%) | 3972 (59%) | 5583 (82%) | 3641 (54%) |
| Instrumental part | ESHM20-full | ESHM20-winGT | ESHM20-rbOrig | ESHM20-Zaliapin |
| # Total Number | 49436 | 18784 | 38470 | 27526 |
| # magnitude ≥ 5.0 | 5245 | 2795 | 4106 | 2337 |

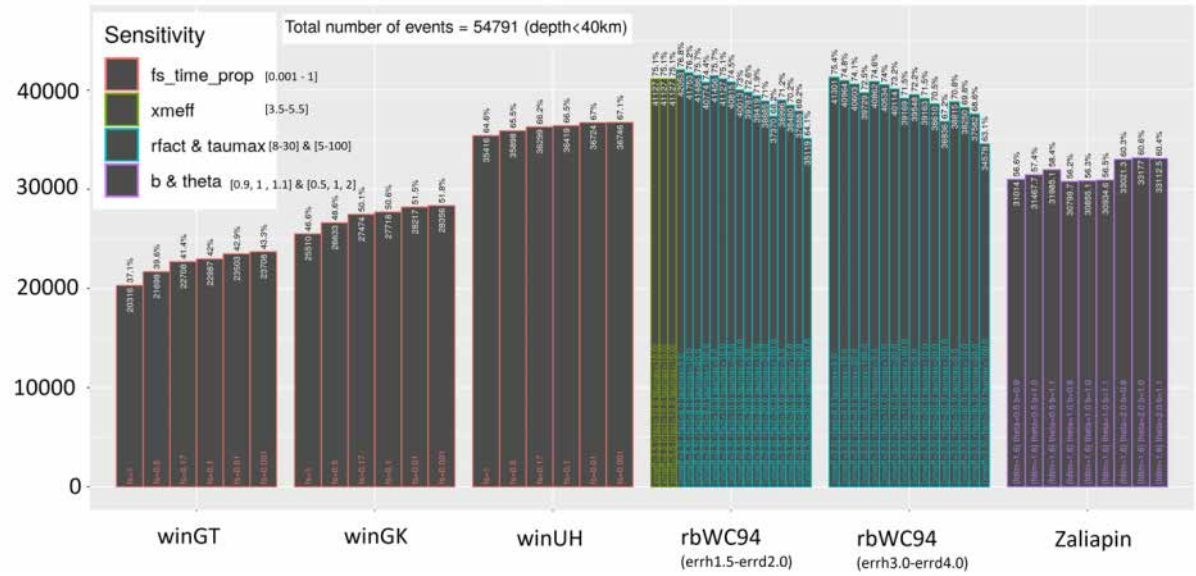



**Fig. 2.2** Sensitivity of number of declustered events to the choice of declustering methods and corresponding parameters.

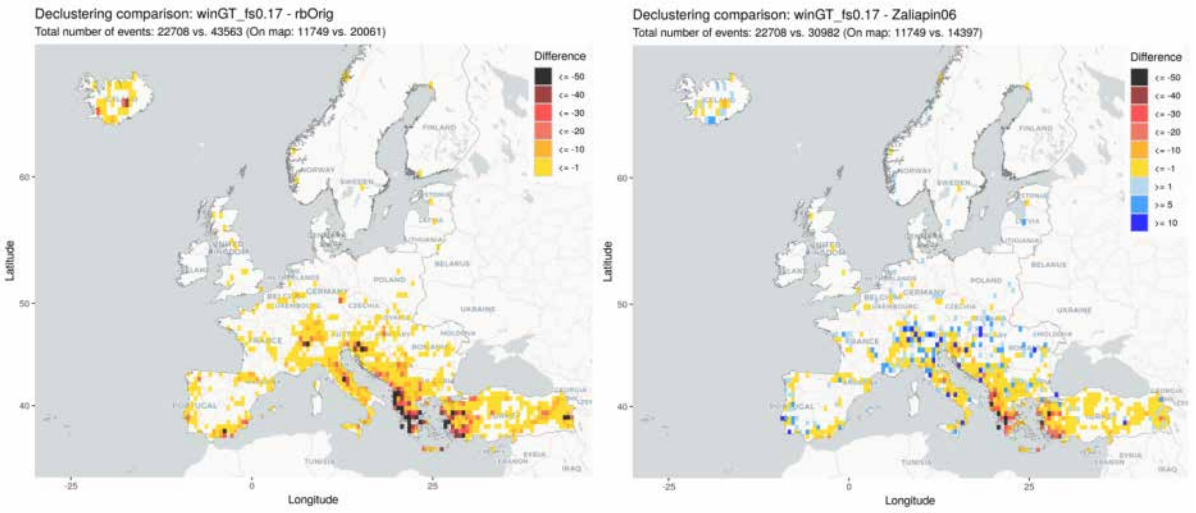



**Fig. 2.3** Spatial difference of the number of events with magnitude M >4.5 in a 50km grid cell, comparing the time-window based method used in ESHM13 (winGT fs0.17) with rb=Reasenberg (1985) (left), and with Zaliapin=Zaliapin et al. (2008) (right).

## 2.2 Catalogue Completeness

The magnitude of completeness (Mc) of earthquake catalogs is a significant metric that has critical consequences for probabilistic seismic hazard assessment, principally because the correct calculation of the parameters of the Gutenberg Richter (GR) equation is contingent on the reliability of Mc. The importance of correct estimation of Mc is illustrated in **Fig. 2.4**.

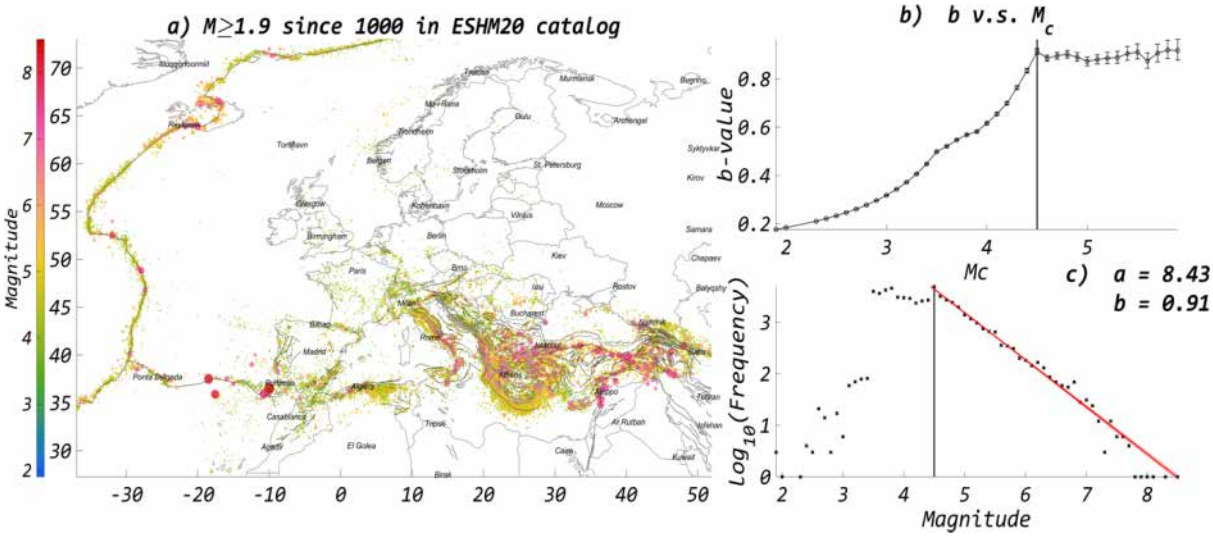



**Fig. 2.4** a) Spatial distribution and magnitude of all earthquakes reported in the ESHM20 earthquake catalog; (b) b-value as function of assumed magnitude of completeness ($M_c$); error bars indicate the 95% confidence interval of the estimate of b-value; (c) overall empirical magnitude distribution for the entire catalog; bestfit GR law obtained assuming $M_c = 4.5$ for the entire catalog; Maximum likelihood estimates of the a and b-value are indicated.

Assuming a too small value of $M_c$ leads to an underestimation in the b-value. As the value of $M_c$ is increased, the bGR-value first increases and then saturates at a constant level. Increasing the value of $M_c$ beyond the saturation point, then only increases the error in the estimate of the b-value, as the number of earthquakes with magnitudes $\geq M_c$ decreases exponentially.

Using the stability in the estimates of b-values as a proxy for the completeness of the catalog (Cao & Gao, 2002), we can consider the shallow ESHM20 catalog to have a bulk? completeness for $M \geq 4.5$. Using $M_c = 4.5$, we find that the global estimates of $(aGR, bGR)$ values for the ESHM20 catalog are (8.43, 0.91), respectively.

In PSHA, one frequently deals with earthquake catalogs very heterogeneous in space and time, with varying degrees of completeness and non-uniform histories. As a result, an average Mc for the entire region, especially large ones, and period of the catalog may not be the most appropriate method, resulting in a skewed estimation of region-specific GR parameters.

This issue is frequently addressed in PSHA by first identifying completeness zones (SZs), which are assumed to have homogeneous completeness history. Then methods that essentially employ changes in earthquake reporting rates in different magnitude bands to identify time periods and corresponding magnitudes of completeness are applied.

To the best of our knowledge, the applications of these methods are semi-quantitative, in that an expert determines both the time-intervals and the corresponding magnitude of completeness by visually analyzing the slope changes in the graph of either the cumulative number of earthquakes with time or from "Stepp" plots (Nasir et al., 2013; Stepp, 1972).

We address the primary deficiency of these methods by making them completely quantitative, replacing expert judgment with appropriate statistical tests. We primarily focus on quantifying the temporal course of earthquake frequency “TCEF” method and combine it further with the Maximum curvature method to achieve the best of both methods.

### 2.2.1 Completeness zones (CSZs)

To simplify the problem of estimation of the spatio-temporal variation of $M_c$, as a first step, the study region is divided into completeness zones (CSZs).

These CSZs are delineated starting from the same regions as those of the EMEC catalogue harmonization (See Fig. 1.3), with additional criteria based on the homogeneity of earthquake reporting through time history, i.e., the earthquake reporting rates are thought to be spatially homogeneous within each of the CSZs.

**Fig. 2.5**(a) shows the polygons outlining the boundary of the 48 defined CSZs. The unique Id assigned to each of these zones is shown within the corresponding polygon in **Fig. 2.5**(a). The color of the polygons shown in Figure 5a scales with the log of the number of earthquakes reported within each polygon. The number of earthquakes reported within the polygons varies approximately by a factor of 3, according to the underlying seismicity of the region and the duration and characteristics of the catalog.

As shown in **Fig. 2.5**(b), the duration of the earthquake catalog also varies by a factor of 2, with minimum and maximum duration being 100 and 1000 years. The variation in the duration of the catalog already points to the necessity of this spatial zonation, so as to ensure that regions with different completeness histories are analyzed independently of each other.

**Fig. 2.5**(b) also indicates the range of magnitudes reported in each CSZs. While the upper magnitude range is both a factor of maximum possible magnitude within a region as well as the completeness history, the variation in the reported lower magnitude range points to the heterogeneity of the completeness history from one CSZ to another and to the lower threshold of the regional catalogue.

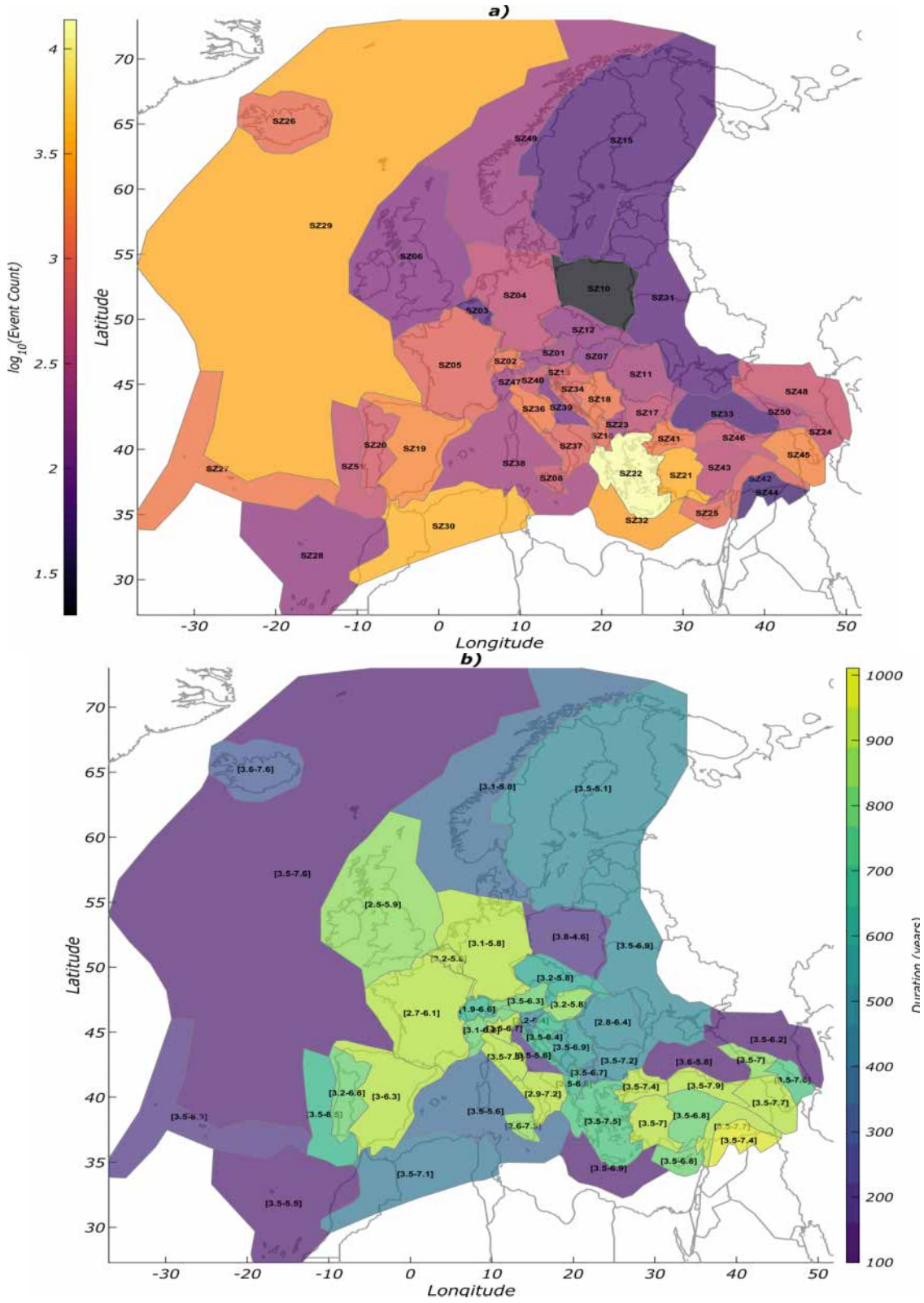



**Fig. 2.5** a) Boundaries of the completeness super-zones and their IDs; the color of each completeness zone scales with the log of the number of events reported within it in the catalog; (b) duration and magnitude range of the catalog reported within each completeness zone.

### 2.2.2 Improved temporal course of earthquake frequency (TCEF) method

Because the temporal course of earthquake frequency, herein after TCEF approach, is primarily reliant on the premise that earthquake rates are Poissonian in time, its first step consists of declustering the data.

To do this, a declustering technique is employed to first preprocess the earthquake catalogs by eliminating clustered seismicity. In this work, we decluster the catalog before finding Mc using the window-based declustering approach with (Gruenthal, 1985) scaling relations.

Following this procedure, the remaining declustered catalog is subdivided into magnitude intervals. A time series of the cumulative number of earthquakes is presented for each magnitude interval. If the catalog is adequately declustered, it should generally follow a Poissonian distribution; consequently, the time series of the total number of earthquakes should seem straight. However, changes in the reporting rate of earthquakes may cause changes in the slopes of these cumulative time series, which can be used as a proxy for changes in the catalog's completeness.

These change points are graphically identifiable in the time series of all magnitude intervals. The change points for distinct magnitude intervals are ordered using the assumption that $M_c$ diminishes with time as more and smaller earthquakes are recorded. This is done in a methodical manner, beginning with the greatest magnitude bin and progressing to the lowest magnitude bin. The constraint that the change point in a particular magnitude bin may only be greater than or equal to the change point in the previous magnitude bin is followed each time. The majority of these transition points are detected visually.

To avoid the qualitative aspect of this approach, we enhance it by including statistical checks to detect the change points. First, we begin by defining a multiple magnitude threshold that may be used to evenly split the reported magnitudes.

Secondly, all earthquakes with magnitudes greater than the highest magnitude criteria are selected. The timings of these earthquakes are then used to undertake a statistical test of uniformity. In this test, we calculate the difference between the observed empirical temporal distribution's cumulative density function (CDF) and the theoretical CDF predicted from a perfectly uniformly distributed sample.

Next, this difference is compared to the distribution of difference calculated for random data simulated from a uniform distribution. Finally, the Poissonian hypothesis may be rejected if the observed time series distance is much greater than the distances for the simulated time series.

This procedure is repeated by gradually extending the lengths of the periods chosen for the uniformity test. The moment when the test can no longer be rejected is regarded as the change point for that magnitude threshold.

The magnitude of completeness for this change-point and magnitude-threshold combination is assessed by choosing all earthquakes since this change point and over this magnitude threshold and then using the maximum curvature technique to detect $M_c$ for the chosen data.

This method is then repeated for each magnitude threshold, with the condition that the change point for a lower magnitude threshold may only be greater than or equal to the change point for a higher magnitude threshold. Furthermore, the $M_c$ corresponding to a lower magnitude threshold must be less than or equal to the $M_c$ corresponding to a higher magnitude threshold. A detailed step-by-step algorithm of the method id provided in the Nandan et al 2022 (in preparation).

Furthermore, we apply the Maximum-Curvature approach (Wiemer & Wyss, 2000) in the proposed algorithm to estimate the magnitude of completeness for each given completeness period. We chose Maximum-Curvature since it is non-parametric and rather easy to apply. However, more complex approaches for measuring the amount of completeness are available (Mignan & Woessner, 2012).

### 2.2.3 Time Series of Mc

Fig. 2.6 shows a time series of the magnitude of completeness obtained by applying the proposed approach to CSZ02 - Switzerland.

Panels (a) and (d) of each of these images depict the completeness zone and the data to which the suggested technique is applied. Panel (d) shows the location of the earthquakes, with the size of the dots according to the magnitude of the earthquakes. Panel (a) displays the magnitude-time plot of the data for the chosen SZs. Because the approach is only applied to the declustered catalog, a distinction is made between clustered (black) and background seismicity markers (orange). The completeness time series generated by applying the suggested technique to the zone is shown using a solid blue line. A dashed blue line is used to represent the completeness times for each zone proposed in previous studies or estimated for the ESHM13.

Note that these previous completeness estimates are based on different earthquake catalogs. The time intervals in Figure 6a, for example, were produced using the national catalogs utilized in the different national seismic hazard analyses for Switzerland (Wiemer et al 2015)

We discover a significant difference between the completeness time series produced in this investigation and those reported in previous studies. One probable explanation for this disparity is the difference in both the analyzed catalogs and the different method used to define the complete periods.

The time intervals of completion for all CSZs are presented in the Appendix A of this Report.

### 2.2.4 Objective quality assessment of time series of Mc

To determine whether the completeness from the various studies or the one determined using the algorithm proposed in this study is a better choice for the ESHM20 earthquake catalog, we measure the consistency between the theoretical and empirical annual frequency magnitude distribution, which can be derived by evaluating completeness time periods.

Using the duration of the complete period $T(m)$ of each magnitude bin, one can obtain the number of earthquakes in such a period in each magnitude bin $N(m)$, which can then be annualized by dividing by $T(m)$ to obtain $N_{annual}(m)$.

Finally, $N_{annual}(m)$ can be used to obtain the empirical annualized frequency magnitude distribution by converting $N_{annual}(m)$ to $N_{annual}(\geq m)$. The empirical annualized frequency magnitude distributions resulting from the two completeness assessments are shown in panel (c) and (d) of Figure 6 using orange circles. Using the data above completeness and the length of the complete periods, we can also obtain the parameters of the GR law, aGR and bGR values, which best describes the corresponding complete data. The maximum likelihood estimate of these two parameters is obtained for the two completeness assessments and are indicated at the top of panel (b) and (c) in **Fig. 2.6**. By comparing the panels (b) and (c) in **Fig. 2.6**, we find that the theoretical GR distribution obtained using the completeness assessment of this study seems (visually) more consistent with corresponding

empirical annualized frequency magnitude distributions than for the case of the regional completeness. We then quantify this goodness of fit in terms of the KS distance between the empirical and theoretical cumulative density functions.

Furthermore, **Fig. 2.7** (a) shows the plot of KS distance obtained from the completeness time steps in this study versus those obtained from the regional completeness time periods for all the SZs. Out of 48 zones, for which both types of completeness estimates are available, in 38 cases the completeness times obtained in this study result in lower KS distance, indicating a higher degree of consistency between the empirical and theoretical cumulative density functions of magnitudes. The average KS distance corresponding to the regional completeness times is 0.24, while for the periods obtained in this study is only 0.14. Using the paired Student's t-test, we verify that the mean KS distances for the regional completeness are significantly larger than those obtained from the method proposed in this study. The null hypothesis that the two means are equal can be rejected at a significance level of 0.01 with a p-value of ~$2.2\times 10^{-6}$.

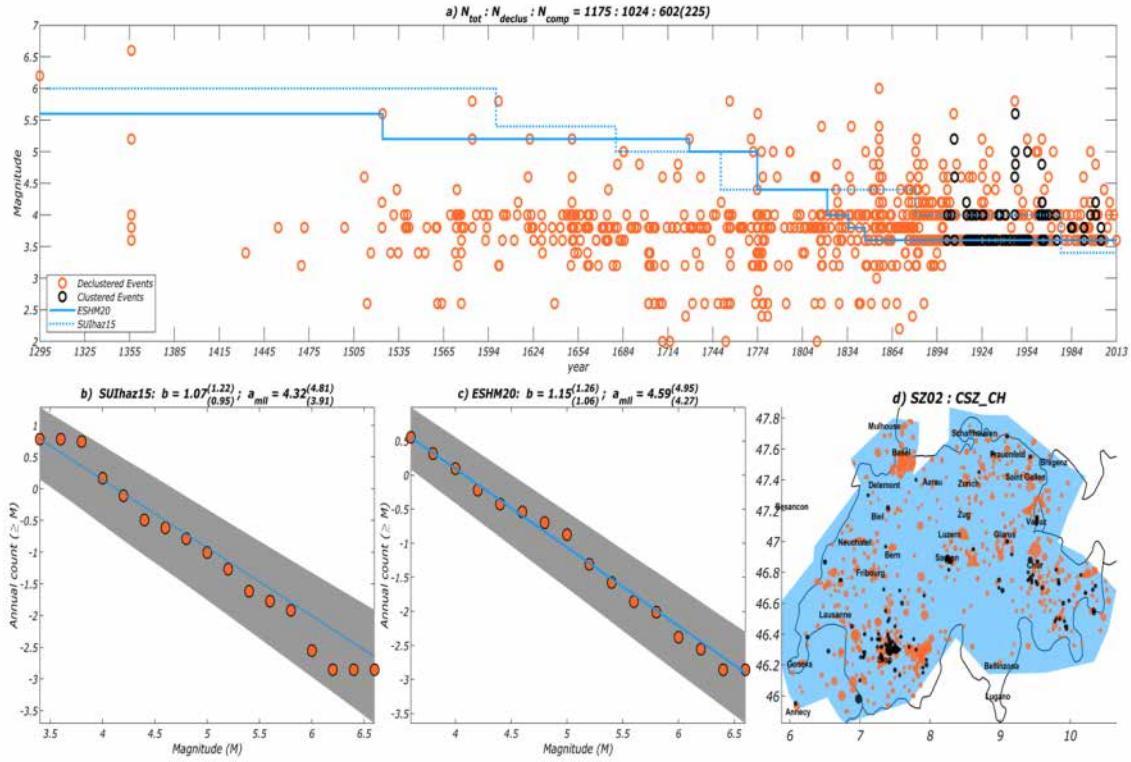



**Fig. 2.6** a) Comparison of the magnitude of completeness time steps obtained using the improved TCEF method (solid blue line) and expert-driven magnitude of completeness time steps (dashed blue line) for four SZs; Orange and black circles indicate the time and magnitude of the declustered and clustered earthquake events; Only declustered data is used to obtain the completeness time steps; (b) Consistency between the empirical annualized frequency magnitude distribution obtained using the expert-driven magnitude of completeness (orange circles) and best fit theoretical GR distribution; Maximum likelihood estimate of the a and b-value and their 95% confidence interval are indicated at the top of the panel; (c) Same as panel (b) but for the completeness time steps obtained using the improved TCEF method; (d) area of completeness super zone highlighted in blue; location of the declustered and clustered seismicity.

In **Fig. 2.7** (b), we also plot the $log\ N_{comp}$ or log of the number of complete events identified by the two types of completeness for each of the SZs. We find that the proposed method identifies fewer complete events (~1400) than the regional completeness. This tendency of being over-optimistic about completeness is also reflected in the estimates of b-value obtained (Figure 3b) being generally smaller for the case of regional completeness. Of the 48 SZs, the b-values estimated using the regional completeness are smaller for the 34 SZs. Using the Student's t-test, we again verify if the mean b-value for the regional completeness is significantly smaller than the mean bGR-value for the completeness obtained in this study. We can reject the null hypothesis that the two means are equal at a significance level of 0.01 with a p-value of 0.0021. The tendency of lower resulting b-value indicates that the regional completeness is likely more optimistic in its definition of complete events.

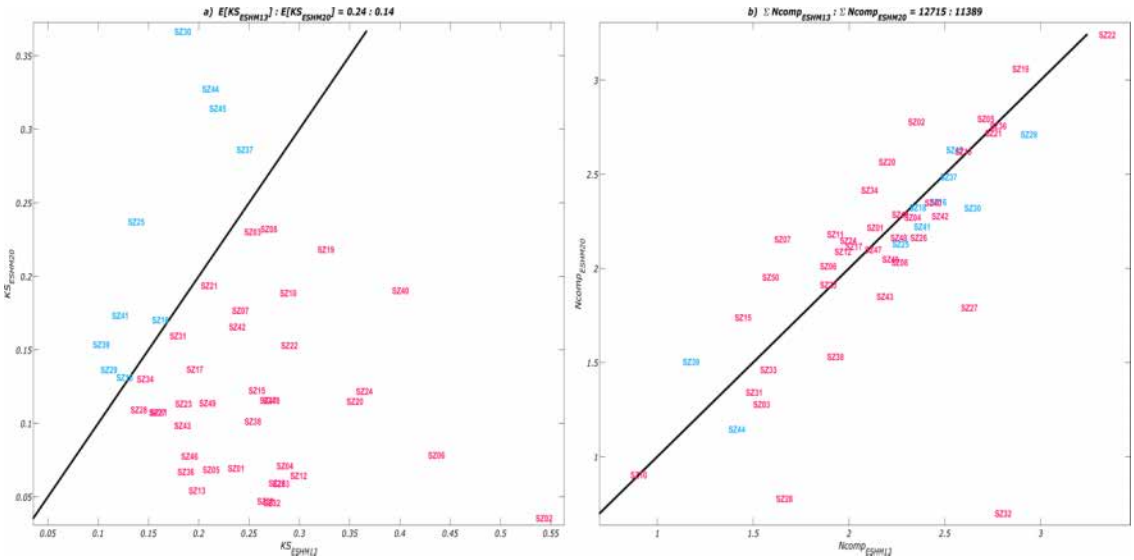



**Fig. 2.7** (a) KS distance between the empirical annualized frequency magnitude distribution and the best fit theoretical GR law obtained using the expert driven completeness and completeness time steps inferred in this study. The red and blue markers indicate that KS distance for the expert driven completeness is larger (worse) than the completeness time steps inferred in this study and vice versa; (b) Number of events above completeness times steps for the two types of completeness time steps.

# 3 Seismogenic Source Model

The seismogenic source model of ESHM20 is built upon the foundation of the ESHM13 (Wössner et al 2015), enhanced with newly developed local and regional seismogenic sources. The main seismogenic source model consists of four distinct source models aimed at capturing the spatial and temporal uncertainty of the earthquake rate forecast across the entire pan-European region:

- area sources model, which is assumed to be the pan-European consensus model, incorporating the national area sources provided by local experts and fully cross-border harmonization.
- active faults and background smoothed seismicity, a hybrid seismicity model that combines the updated active faults datasets (chapter 3, section 3.3.) with the background seismicity in regions where faults are identified. In regions without active faults – the smoothed seismicity model represents an alternative to the area sources model.
- subduction sources depicting both the subduction interface and in-slab of the Hellenic, Cyprian, Calabrian and Gibraltar Arcs.
- non-subducting deep seismicity sources describe the nested seismicity with depth in Vrancea, Romania as well as the cluster of deep seismicity in southern Iberia Peninsula.

In this Chapter an overview of these seismogenic source models, their components and properties, as well as their characterization in terms of earthquake rate forecast is provided.

## 3.1 Unified Area Sources Model

Area sources are the most common seismogenic source representation (e.g., Grünthal et al., 1999a, 1999b, 2018, Jimenez et al., 2003, Meletti et al. 2008; Vilanova et al. 2014, Danciu and Giardini 2015, Danciu et al 2018, Sesetyan et al 2018, Pagani et al 2018).

The development of the area source model is achieved by updating the area sources of ESHM13 with the contribution from the national hazard models. Since 2013, the release of the ESHM13, updates of the seismic hazard assessment at national level have been conducted across Europe.

The following countries have been providing updates of the seismogenic sources: Belgium, Bulgaria, Switzerland, Germany, Macedonia, Romania, Slovenia, Spain, Switzerland, Turkey; contribution for Northern Africa are taken from global mosaic of hazard models of Global Earthquake Model (Pagani et al 2018), the models for the Balkans are the same as in the BSHAP2 project (Salic et al 2016).

From every national model provided by the regional and national experts we have retained the area source model, or the branch model, that was recommended when more than one model was available (i.e. Model C for Germany, Grünthal et al 2018).

The area sources model describes regions of homogenous seismicity classified as: shallow crustal, volcanic, subduction in-slab and deep seismicity (i.e. Vrancea region, Romania). The delineation of the area sources was revisited and then modified to ensure cross-border harmonization. Mainly, the overlapping area sources at national borders were resolved while preserving the original information (i.e. depth distribution, style-of-faulting, maximum magnitude). As guidance for cross-border harmonization, we prioritize the national source models within the country boundaries.

The cross-border sources are guided by seismotectonic evidence, active faults, and major geologic/tectonic features if available, if not, then the seismicity patterns are used: historical earthquake locations or recent clusters of seismicity. If none is available, then area sources from the most recent seismogenic model are subjectively preferred. Furthermore, the curation of the area sources includes removing duplicates, simplifying too complex area source boundaries (i.e. area boundaries with inner spaces), and eliminating multiple segments or points.

After the necessary curation of the input files, for the final harmonization of the area, several questions had to be addressed by the core team (similarly with ESHM13 and other regional hazard models, Danciu et al 2018, Pagani et al 2018:

- In area sources with few events, how can we constrain the magnitude-frequency distributions (MFDs)? We observed that a significant number of seismogenic sources have no events of moment magnitude Mw > 3.5. In fact, about 35% of the area sources have less than 10 events, 7% of the area sources contain 10 to 15 events, 18% of the area sources have 30 to 60 events, and about 21% have greater than 60 events of Mw>3.5. This statistic considers the un-declustered earthquake catalogue without completeness constraints. Hence, to have robust earthquake statistics the use of super zones, or of intermediate layers, is required. Such super zones are constituted of a number of adjacent sources, assembled according to tectonic criteria. The well-defined magnitude-frequency distributions evaluated at the super zone level can help constrain the distributions in source zones with little data.
- Is there additional tectonic or geologic data that justify small sources with only few seismicity, or can a small source be merged with adjacent sources? A minimal number of earthquakes has to be captured within each area source, for a robust statistical estimate of the activity parameters. Area sources without known significant seismicity or known tectonic/geologic features should be considered as background sources i.e. sources with a minimum activity rate assigned as earthquake occurrence cannot be excluded.

Noteworthy, the area sources were further discussed with the local experts in bilateral meetings or workshops and their feedback and changes were iteratively integrated. Each area source model is characterized by various parameters including tectonic settings, style-of-faulting parameters, depth distribution, range of maximum magnitudes, lower and upper seismogenic depth, some of which are inherited from the ESHM13. The procedure to estimate the seismic activity parameters are given in the next section. **Fig. 3.1** illustrates the geographical distribution of the shallow crust area, as well as the deep seismicity and subduction inslab sources of Europe and Turkey.

### 1.1.1 The use of Super Zones

Super zones provide a spatial proxy to describe tectonic features, geological fault systems and seismicity patterns. Some parameters evaluated at the super zone level are then applied at the level of seismogenic sources. Super-zones cover large geographical regions, and we distinguish three types (see **Fig. 3.2**): completeness super zones (CSZ), tectonic super zones (TECTO) and the maximum magnitude zones (MAXMAG).

The delineation of the CSZ closely follows the regions used for the compilation and magnitude harmonization of the instrumental catalogue (EMEC, chapter 3 section 1.2). The primary use of the CSZ is to provide the geo-delineation of the regions for estimating the complete magnitude and time windows. Assessing the catalogue completeness is more feasible for super-zones than for individual seismogenic sources due to the increased number of earthquakes that improves the statistical analysis.

TECTO model consists of super zones extracted from the national hazard models and then updated for consistency with the unified earthquake catalogue, active faults and the tectonic regionalization given in Chapter 3 section 3.5. The main use of TECTO is to constrain regional seismicity by facilitating a robust estimation of the activity parameters given the larger number of earthquakes within each super-zone. The TECTO model also provides the basis for a smoothed seismicity model and delivers the regional b-values necessary for estimating the seismic productivity of active faults.

MAXMAG zonation is an update of the maximum magnitude super zones of ESHM13 (Meletti et al SHARE Deliverable D3.3). The main use of this geo-regionalization is the estimation of the maximum

magnitude ($M_{max}$) based on the earthquake catalogue. It shall be noted that none of the super zone models presented herein are considered as a direct input for seismic hazard calculation.

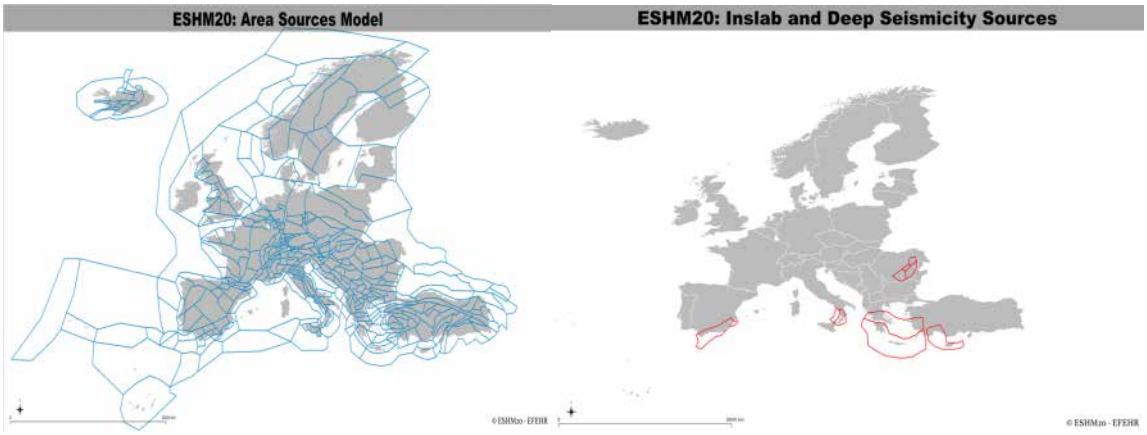



**Fig. 3.1** ESHM20: Area Sources Model to describe the shallow crust seismicity (left) and the deep seismicity and subduction inslab sources (right) across the pan European region

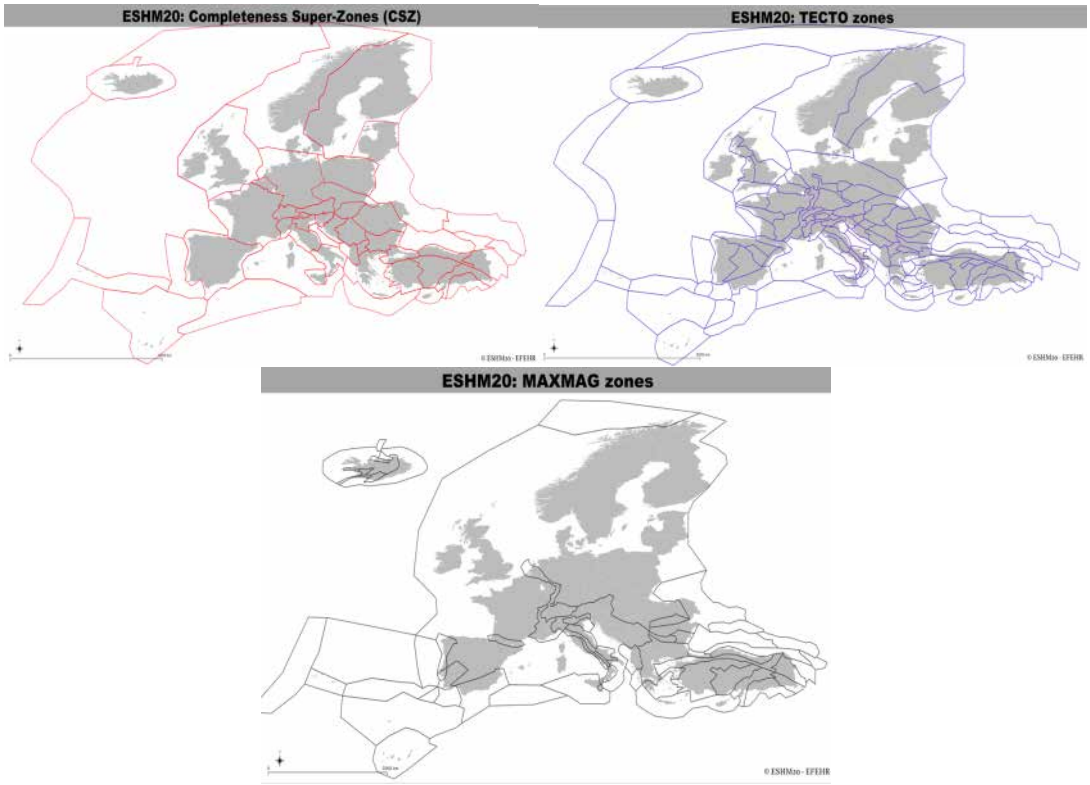



**Fig. 3.2** ESHM20 Completeness super zones (top left), tectonically delineated TECTO super zones (top right) and MAXMAG maximum magnitude super zones (bottom).

### 1.1.2 Activity Rates Parameters and Magnitude Frequency Models

The activity rates of the individual area sources are estimated based on the assumptions that the regional seismicity follows a memoryless Poisson process characterized by a stationary mean rate of occurrence described by an exponential distribution i.e. Gutenberg- Richter model: $Log_{10}\ N = a_{GR} - b_{GR}*M$. The N is the cumulative number of earthquakes per year equal to or greater than a magnitude M (in this model we use the moment magnitude Mw), and $a_{GR}$ and $b_{GR}$ are constants. The $a_{GR}$-value

represents the total seismic productivity of a given source (the log of number of events with M≥0); the $b_{GR}$-value is the negative slope of the recurrence curve expressing the average ratio of exponentially distributed small and large magnitude earthquakes.

Earthquake recurrence parameters of each area source are derived from the declustered earthquake catalogue (see chapter 2, section 2.1) using the completeness intervals as defined for each CSZ-zonation (see chapter 2, Section 2.2). The activity parameters are estimated, accounting for the uncertainties of the aGR, bGR parameters, and their confidence intervals, for both data and prediction. Methodology is outlined in Bollinger et al (1989, 1993), Weichert (1980), Berril and Davis (1980) and Herrmann (1977).

The uncertainties of the $a_{GR}$,$b_{GR}$ parameters are defined by their covariance and the associated standard error. To estimate the range of $a_{GR}$, $b_{GR}$ values we used a random sampling technique. Firstly, a multivariate normal distribution is used for generating samples of correlated $a_{GR}$ and $b_{GR}$ values, based on their values, their covariance and the individual standard error.

Next, we bootstrap 1 million samples and the discrete approximation method of probability distributions (Miller and Rice 1983) is used to obtain the range of the $a_{GR}$, $b_{GR}$ values. The $16^{th}$ , $50^{th}$ , $84^{th}$ percentiles are used to represent the lower, median and the upper values of the $a_{GR}$ and $b_{GR}$, respectively. The activity parameters were estimated for the completeness regions (CSZ), for the large-scale tectonic zones (TECTO) and as well as for the area source model (ASM).

Examples of the uncertainty range of the magnitude frequency distribution for three sources in Spain (ESAS134), France (FRAS176) and Italy (ITAS317) are given in Figure 3. Note the different number of events used in each source (*n_used* is the number of complete events used for estimating the aGR, bGR values). There is a relevant correlation between the number of events used for statistical fitting to obtain the GRs parameters and the uncertainty range (green dashed curved in the same plot); the uncertainties decrease as the number of events increases and vice-versa.

The mfds of the upper and lower bGR-value are well constrained by observations for the source with fewer events (EAS134, i.e. 33 events) and decreases as the data increases (FRAS176, i.e. 78 events, ITAS317, i.e. 220 events).

Another critical factor affecting the statistical fitting of the seismicity parameters of the source models is the magnitude bin sizes (Stromeyer and Grunthal, 2015). The maximum likelihood method (Weichert 1980) is sensitive to the number of events in the most populated bins, usually the first magnitude bins, which often describe the frequency of modern instrumental seismicity.

Herein, the size of the magnitude bin is set up to 0.2 units of magnitude, and the statistical fitting will apply for a minimum of two consecutive bins with at least three events in each bin when available, otherwise a proxy on bGR from the overlaying tectonic super zones (i.e. TECTO) is used. Furthermore, for area sources with more than 30 complete events, the GR parameters are computed with the automatic maximum likelihood method using the earthquakes falling in the area source, whereas for the area sources with less than 30 events, $b_{GR}$ from TECTO is used (Danciu et al 2018, Grünthal et al 2018, Sesetyan et al 2018).

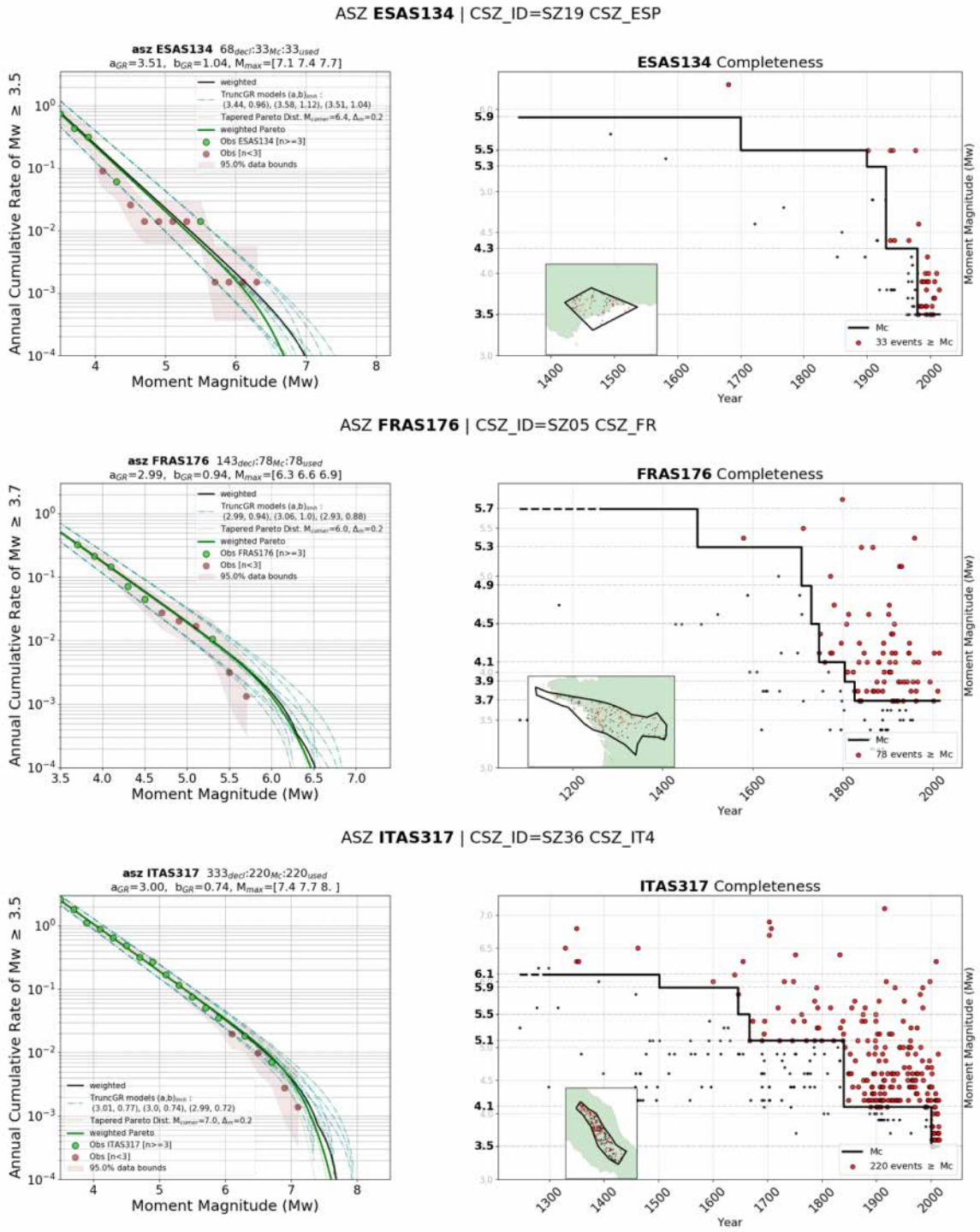



**Fig. 3.3** Examples of MFDs (left panels) for individual area sources in Spain (top), France (mid) and Italy (bottom). The right panels illustrate the magnitude-time windows intervals for which the earthquakes are assumed to be complete.

In this automatic procedure, the estimation of the $b_{GR}$-value is implemented so that first the statical fitting is done with all the magnitude bins; if the resulting $b_{GR}$ is outside a predefined range of values (i.e. 0.70 to 1.15), then the first magnitude bin is removed and the $b_{GR}$ is recomputed and checked again against the predefined $b_{GR}$ values.

The procedure allows for iteratively removing a maximum of five magnitude bins. If $b_{GR}$ does not converge to values within the validity range (0.7 to 1.15), then the $b_{GR}$ from the TECTO model is considered. The exceptions are the volcanic sources for which the $b_{GR}$ values greater than 1.15 are expected.

When a regional bGR from TECTO is used as a proxy for the area sources with few data, the seismic activity is estimated either from a direct estimation of the aGR or by re-scaling the occurrence rates as a function of the number of complete earthquakes.

Specifically, the scaling factor is the ratio between the number of complete events in the area source and the number of complete events in the corresponding TECTO super zone, i.e. the $a_{GR}$ value of an area with limited events is equal to $\log_{10}(N_{TECTO}(M>0)$ * (number of complete events in the super zone / number of complete events in the area source).

In area sources with zero complete events, the aGR is computed as a function of the regional (i.e. aGR of the super zone) seismic productivity normalized per units of area, $a_{GR}$ = ($aGR_{TECTO}$ – log10(tecto superzone-area in km2))*area source area in km2.

**Fig. 3.4** displays the MFDs of individual area sources in an example super zone in Greece. The sum of the area source MFDs is compared with the regional seismic productivity as given by the TECTO super zone MFD.

It shall be noted that there are cases when the area source contains enough events, yet the resulting bGR lies outside the validity range, and hence the $b_{GR}$ is set to the regional bGR from the TECTO model. Other exceptions worth mentioning imply long sequence of historical seismicity (i.e. Southern Italy and off-shore Portugal), and in this case a two-range GR model (Utsu 1999) is used to define such a distribution (see **Fig. 3.5**).

A double truncation Gutenberg-Richter magnitude-frequency distribution is used to characterize the decay of the annual occurrence rates with different magnitude bins (i.e. from lower to upper magnitude bounds).

Additionally, a tapered Pareto distribution (Kagan 1993, Utsu 1999) is considered as an alternative to the double truncated MFD. The main reason is to provide an alternative recurrence model with a faster decay of the rates towards the maximum magnitude. In many area sources without events above Mw5, the rates obtained in the magnitude interval 5 to 6 from the GR model could be inflated. In these cases, the tapered Pareto distribution provides alternative estimates for the rates in the moderate to large magnitude range.

A comparison between the two distributions, i.e. double truncated GR distribution and the Pareto tapered distribution model, are given in **Fig. 3.5** for an area source in Southern Spain. The differences of the two MFDs are obvious for the recurrence rates of Mw > 5.2, with the Pareto tapered distribution depicting a faster decay of the rates.

The tapered Pareto distribution requires the activity parameters aGR and bGR, and an additional parameter, the corner magnitude Mc, estimated here as a function of the observed maximum

magnitude. A default $M_{corner}$=6.0 is used for area sources with an observed maximum magnitude lower than 5.5; otherwise the $M_{corner}$ is automatically estimated as a function of the maximum observed magnitude.

The evaluation of the maximum magnitude of area sources is based on the strategy defined in ESHM13 (Meletti et al 2010) using the updated tectonic super zones (i.e. see **Fig. 3.2** for MAXMAG zonation) and the unified earthquake catalogue.

Three $M_{max}$ values are used to represent the upper bounds of the expected magnitude: the lower bound of the maximum magnitude (Mmax01) is equal to the highest magnitude observed for the zone, increased of its standard deviation; the second value (Mmax02) is assumed equal to Mmax01 plus a magnitude increment of 0.3, and finally the third value is assumed to be conservatively equal to Mmax01 plus a magnitude increment of 0.6.

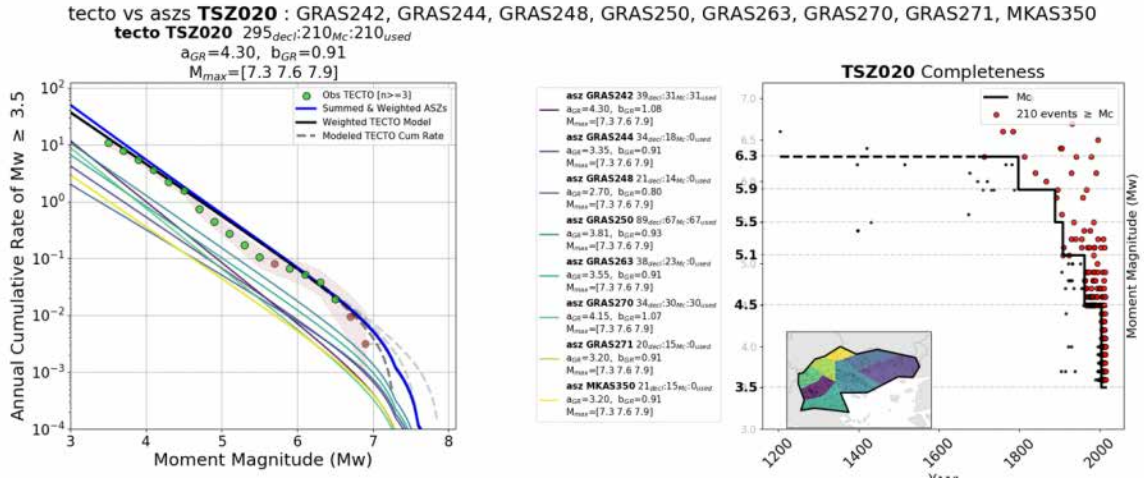



**Fig. 3.4** Comparison of the magnitude-frequency distribution of the TECTO super zone (TSZ020, in Northern Greece) with the sum of individual MFDs of area sources. The activity parameters of the TECTO zone are given in the header of the left panel; the completeness plot and the geometry of the area sources together with the activity parameters of the individual area sources are given in the right panel.

The magnitude increment was statistically estimated from the entire unified catalogue; the value 0.3 presents the statistical uncertainty (i.e. standard deviation) of the moment magnitude in the unified earthquake catalogue. Wheeler (2009) states that two added magnitude increments would imply an extension of the historical record by a factor of 1.35 to 1.7, assuming a regional b-value of about 0.9 to 1.00. With the two factors, the maximum magnitude obtained would correspond to a historical record of about ~ 4000 and ~ 5000 years (Danciu et al 2018).

For the tapered Pareto distribution, the $M_{max}$ is controlled by the Mcorner, and its value might be different from those estimated in the super-zones, as it can be observed in **Fig. 3.6**. For a fixed annual cumulative rate (i.e. $10^{-4}$ ) the Mmax values of all distributions can be compared: the lower $M_{max}$ values for Pareto's MFD are 6.1, 6.3 and 6.5 whereas the range of the remaining $M_{max}$ values span from 6.5 to 7.1.

In summary, shallow crust area sources are characterized by specific earthquake recurrence (aGR- and bGR-values) parameters with associated uncertainties, and a distribution of $M_{max}$ values. Details of the model implementation are given in chapter 7.

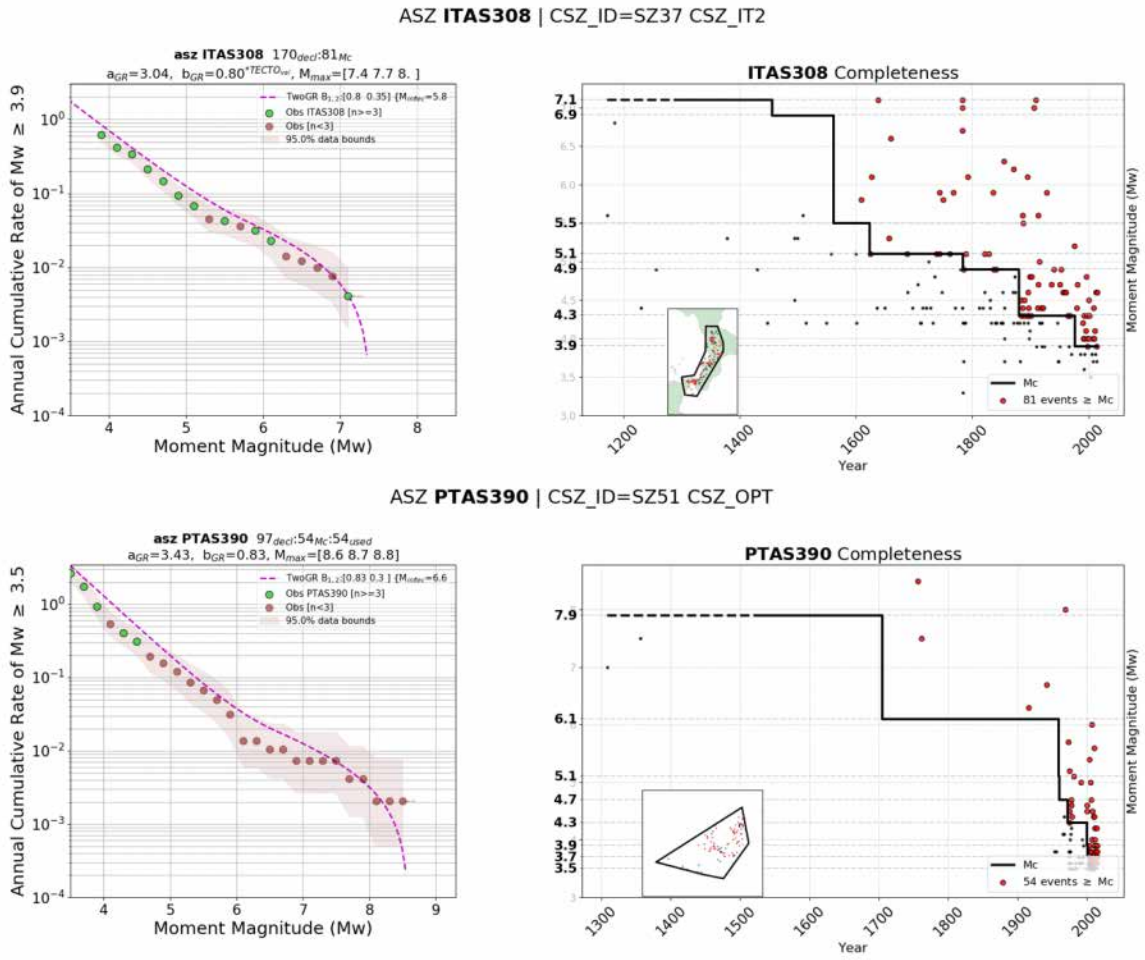



**Fig. 3.5** Illustration of the two-range magnitude frequency distributions (Utsu 1999) used for two area sources in southern Italy (upper plot) and off-Shore Portugal (lower plot).

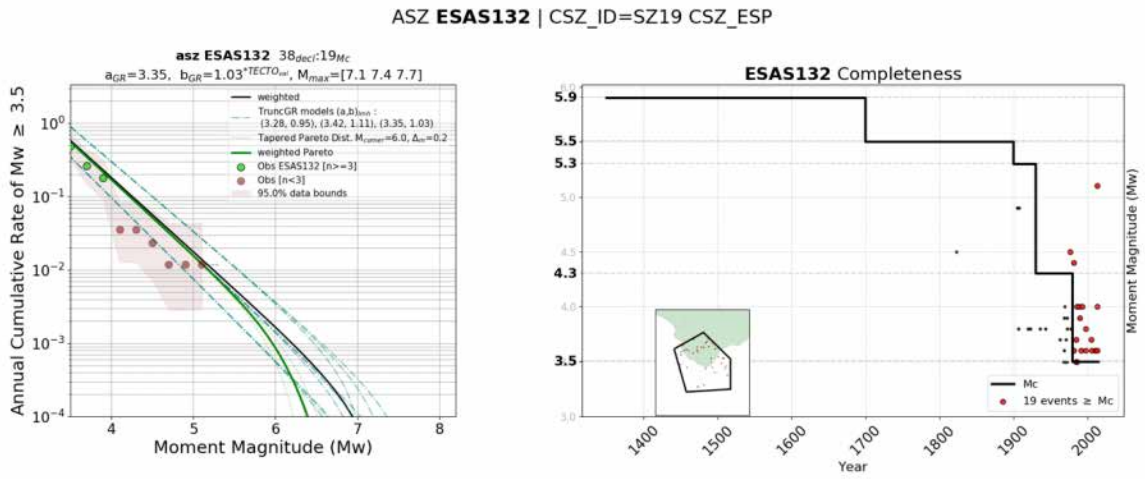



**Fig. 3.6** Example of an area source in Southern Spain, modeling of the recurrence based on two alternative magnitude-frequency distributions. The double truncated GR distribution (black line) and its uncertainties (dashed-blue lines) and the Pareto tapered distribution (green line) are given in the left panel.

## 3.2 Active faults and background seismicity model

### 3.2.1 Active Faults

The concept of active faults and background seismicity has become the preferred solution for regional seismogenic source models (see Danciu et al 2018 for EMME14, Sesetyan et al 2018 for Turkey, and global mosaic of hazard models of GEM, Pagani et al 2018).

The ESHM13 model for active faults and background area sources is abandoned in ESHM20. The main reasons are twofold: 1) background area sources were delineated to match the area source model, preventing the independency assumption of the alternative seismogenic source models, 2) uneven spatial completeness of the faults within the background area sources. In the updated version, we combine the active faults and background smoothed seismicity.

The fault-based source model is built upon the European Fault-Source Model 2020, an update of the European Database of Seismogenic Faults (EDSF) and several other recent active fault compilations in the Euro-Mediterranean regions (Basili et al 2020). The crustal faults are represented with a down-dip planar geometry that defines its three-dimensional plane in the Earth's crust. In this representation, the trace of the fault upper edge is extruded downward basing on the dip and depth values.

The minimum set of basic fault parameters required for constructing a seismogenic source model refer to Geometry (Location: Lat, Lon, Depth; Size: Length, Width; Orientation: Strike, Dip) and Behavior (Rake and Slip Rate). This set of attributes is associated with each geographic feature, and then a magnitude-frequency distribution (MFD) can be derived for each fault source.

To this end, the geologic information contained in the fault datasets can be used to estimate two key parameters: 1) the seismic productivity by converting the long-term geological or geodetic slip rates into activity rates, assuming moment conservation, and 2) a proxy for the moment magnitude upper bound of the MFD. It is assumed that the geological slip rates encompass several seismic cycles of large-magnitude earthquakes on a fault, offering advantages over historical seismicity when used to estimate earthquake frequencies. Furthermore, another important assumption is that the fault location provides a spatial proxy for occurrence of future large magnitude earthquakes (M>6.0).

In the ESHM20, we use the Model 2 (i.e. arbitrary area) of Anderson and Luco (1983) that is a truncated exponential magnitude-frequency distribution to characterize the seismic activity of each entry of the active faults. The functional form of the recurrence relationships is provided in the original paper of Anderson and Luco (1983), and not repeated here.

For all faults, the following values were considered as default: crust shear modulus (μ) value of 3.0 x 1e11 dyne/cm2; typical values (c=16.05 and d=1.50) for the magnitude-moment scaling coefficients as originally proposed by Kanamori and Anderson (1975).

A key assumption when using the active faults is that the seismic productivity of these faults will satisfy the GR relation. In this view, we use the bGR-values of the TECTO-zones that host the faults. The TECTO super zones are representative of regional seismicity. The active faults, although simplified herein, are not isolated, but form branching structures and networks.

The maximum magnitude of each fault source is obtained by use of magnitude – fault geometry scaling relationships, hereinafter referred to as fault scaling laws (FSL).

The generalized functional form between rupture dimensions (L, W, A, D) and moment magnitude (Mw) is $(X) = a + bM_w$, where X is the rupture dimension under consideration (e.g., L or W in km, A in km2, D in m) and the coefficients a and b are empirically determined. We here propose to adopt the FSL developed by Leonard (2014) for crustal ruptures and by Allen and Hayes (2017) for subduction interface ruptures.

A sanity-check revealed discrepancies between the fault capabilities to accommodate the assigned $M_{max}$. Specifically, there are small faults whose average area may be smaller than the rupture area associated with the higher Mmax values when considering all the variability of the fault parameters and the uncertainty in the FSL; as such, the rupture corresponding to this $M_{max}$ might extend beyond the fault geometry limits.

An analysis of the effects of various key parameters of seismic productivity indicates that two parameters are responsible for a large variability of the magnitude frequency distribution: the slip-rates used and the $M_{max}$ (see **Fig. 3.7**).

For a constant slip rate, increasing the $M_{max}$ will result in decrease in the recurrence rates of low-to-moderate magnitude events, this is because the large magnitude events constitute the majority of the total seismic moment rate; while increasing the $M_{max}$ will require a subtraction of many smaller earthquakes to preserve the same seismic moment budget (Youngs and Coppersmith 1985).

The fault productivity is also sensitive to the fault area, the larger the area the lower the seismic productivity and vice-versa. The effect of the fault area is correlated with the effect of the $M_{max}$, thus only the latter is retained.

To conclude, the main input parameters are the bGR-values, fault-slip rate and the maximum magnitude ($M_{max}$) and their uncertainty on the seismic productivity of each fault is captured as such: three values of the slip-rates, three values for the maximum magnitude, a single b-value (estimated from robust statistics in the TECTO zones) and an average area of each fault.

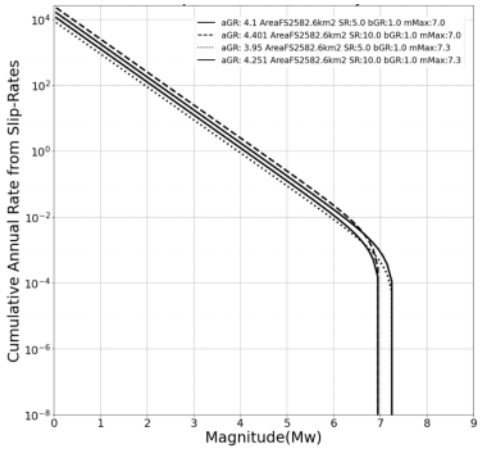



**Fig. 3.7** Effects of various key input parameters in seismic productivity of a single fault: maximum magnitude (i.e. two values of Mw=7.0 and 7.3) and variable slip-rates values (5mm/year and 10mm/year)

### 3.2.2 Smoothed Seismicity

The smoothed seismicity model is based on the TECTO zonation, for which the activity parameters (aGR- and bGR-values) were estimated on the declustered and complete earthquake catalogue using the smoothing algorithm of Nandan et al (2021, *EFEHR Technical Report, in preparation*).

Several choices are required to make earthquake forecasts based on the smoothed seismicity approach. Critical decisions include choosing the (1) adaptiveness of the bandwidth of the kernel, (2) smoothing parameters, (3) declustering algorithm, and (4) declustering parameters.

Training and validation sets are typically required to determine the best possible combination of options. The training set is used to generate the smoothed seismicity model for a given combination of the options mentioned above, while the validation set is used to rank the options in terms of mutual information gain.

We first propose a penalized-maximum-likelihood system that allows us to find optimal smoothing parameters using only the training set without invoking the need for a validation set. We then set up a pseudo-prospective test bench to validate the different combinations of choices used to obtain the smoothed seismicity forecasts.

In this test bench, we use all the data through 2006 to train a series of smoothed seismicity models (SSMs) and then compare their forecasting ability using the earthquakes between the entirety of 2007 through 2015. Based on this validation, we recommend the hyper-parameter choices for obtaining an alternative SSM for ESHM20. Finally, we perform a long-term retrospective validation of SSM based on the reference declustering algorithm, which is used in ESHM20, to demonstrate its consistency with the observed number of earthquakes in different TECTO polygons.

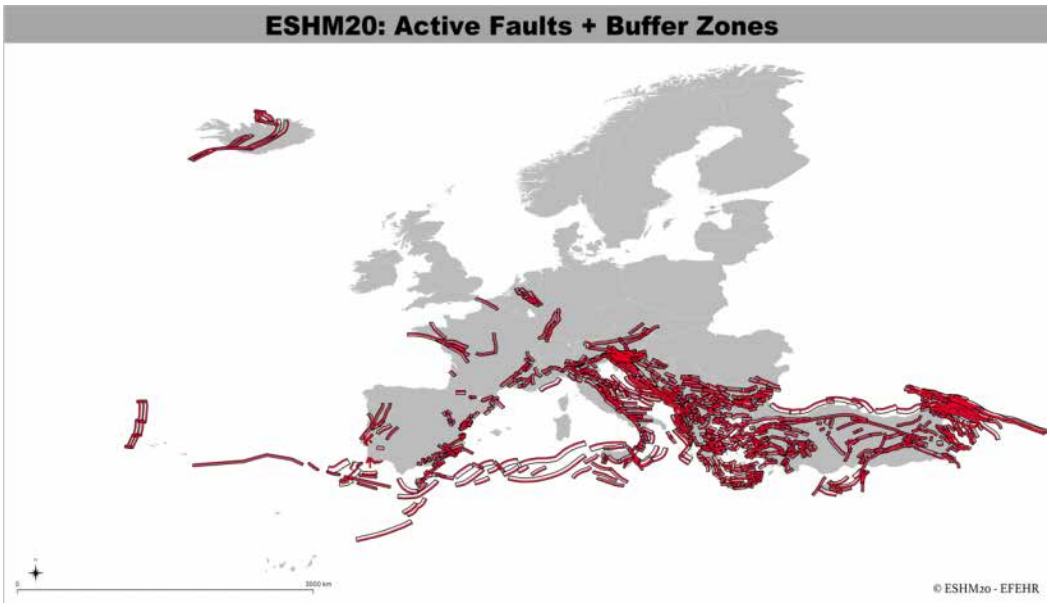



**Fig. 3.8** Conceptual sketch of active faults and background buffers for removing the double counting of seismicity.

### 3.2.3 Combining the Smoothed Seismicity with the Active Faults

In the vicinity of active faults, the background-smoothed seismicity is complementary to the seismic productivity together with the faults, whereas in regions with no faults, the smoothed seismicity is an alternative to the area source model (Northern Europe).

The main challenge when modeling the background seismicity with the active faults is the avoidance of double counting the seismicity in the fault's proximity. To avoid the double counting of the earthquake rate forecast due to active faults and background seismicity, a variable threshold magnitude was introduced. This concept requires a definition of fault proximity, which in this case is defined as a spatially symmetric buffer following the top-trace of the fault. The buffer zone results from combining the fault-width projection to the Earth surface and an arbitrary polygon that extends up to 5km along strike on the opposite side of the fault trace (see **Fig. 3.8**).

The magnitude threshold (*fMthr*) is fault dependent, and the main assumption is that events with Mw > *fMthr* are associated with the faults and a minimum of two magnitude bins is imposed for each fault. The *fMthr* for each fault is based on the Mmax (fMaxMag01) value and is estimated as following:

- faults with *fMaxMax01* $\leq$ 6.3, *fMthr=5.9*
- faults with 6.3 $\leq$fMaxMag01 $\leq$7.3, *fMthr*= fMaxMag01-dM, where the dM=0.6(the overall uncertainty of the $M_{max}$)
- faults with fMaxMag01 > 7.3, *fMthr=6.7*

Thus, within the fault buffer the occurrence rates of lower magnitude bins (<= *fMthr*) are represented by point sources, whereas the occurrence rates of the M>*fMthr* are associated with the fault. Outside the buffer, the Mmax of the smoothed seismicity is the Mmax value attributed to the corresponding TECTO zone.

Examples of comparisons between the composite cumulative annual rates of the active faults, background sources and the rates observed are done at the TECTO level and given in **Fig. 3.9**. As expected, there are three cases: 1) total productivity of all active faults matches the observed earthquake rates (TECTO zone TSZ041, Italy); 2) total productivity of all active faults overestimates the observed earthquake rates (TECTO zone TSZ029, Southern Spain) and 3) total productivity of all active faults underestimates the observed earthquake rates (TECTO zone TSZ022, South West of Bulgaria).

There are various possible reasons for such discrepancies: over or under estimation of the fault slip-rates, completeness of the active faults within each region, subjectivity in the delineation of the TECTO zone, incompleteness of earthquake catalogue at large magnitudes, short observation period for the earthquake catalogue, etc.

This visual comparison aids the sanity check of the seismic productivity of a specific seismotectonic domain. We consider that the uncertainties of the future earthquake rates are depicted by both the area source model and the smoothed seismicity plus fault model.

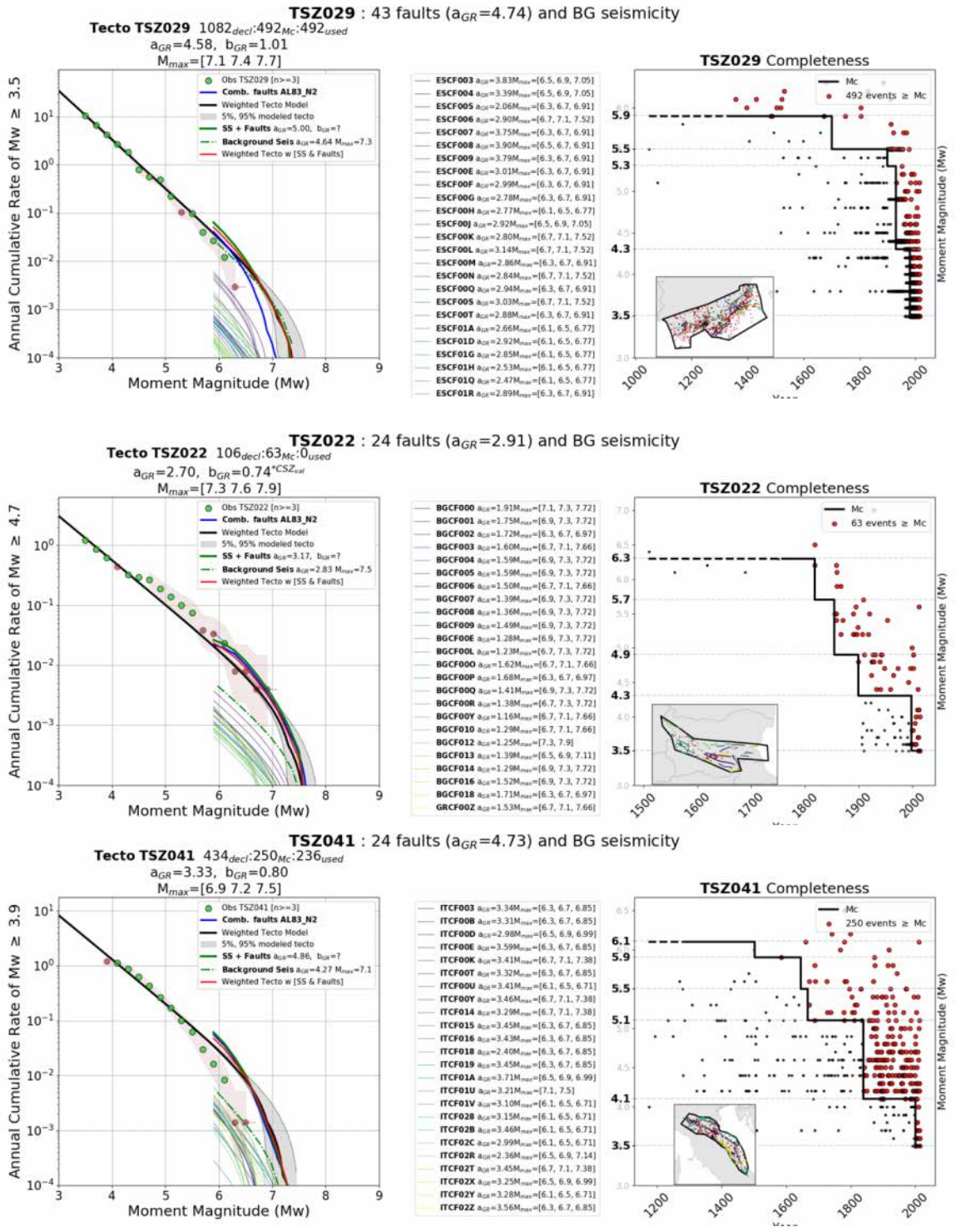



**Fig. 3.9** Comparison of total seismic productivity of all faults within a TECTO super zone, examples in Spain, Bulgaria and Italy. The left panel shows the summed MFDs (in blue) of all faults (in green) within the TECTO zone; the right panel indicates the complete events within each TECTO zone. The colored curves are the mfds for individual faults as described by their a,bGR parameters in the middle panel, the green solid line are the summed mfd of faults and background seismicity, the background seismicity mfd is depicted in a dashed green line, the black line indicates the TECTO mfd, whereas the weighted mfd between the TECTO and faults plus smoothed seismicity is described by the red curve.

## 3.3 Subduction and Deep Seismicity Sources

### 3.3.1 Subduction Interface

Subduction sources cover the parts of the Europe-Africa plate boundary where there is still the presence of oceanic crust in the lower plate. They are the Gibraltar Arc, Calabrian Arc, Hellenic Arc, and Cyprus Arc (Figure 8). The zones of active continental collision are included in the dataset of crustal faults.

After the EDSF compilation from project SHARE (Basili et al., 2013), the most recent slab geometrical reconstructions are those for the Calabrian Arc by Maesano et al. (2017) and the global compilation by Hayes et al. (2018), as well as a number of studies that treated specific aspects or parts of these subduction zones. The included subduction sources have largely been inspired by these studies.

The subduction interfaces are modelled as complex faults whose geometry is sampled from the 3D models shown in **Fig. 1.7**) considering estimates of the upper and lower boundary of the seismic interface. These same models, along with estimates of the slab thickness, are also used to derive the intraslab sources which are cut at a maximum of 300 km depth.

The recurrence parameters of the subduction zones (interface and intra-slab) are estimated based on the unified earthquake catalogue. Moreover, for the seismic interface sources we also used an alternative recurrence model based on tectonic convergence rates across the subduction zones and the double-truncated frequency-magnitude distribution model.

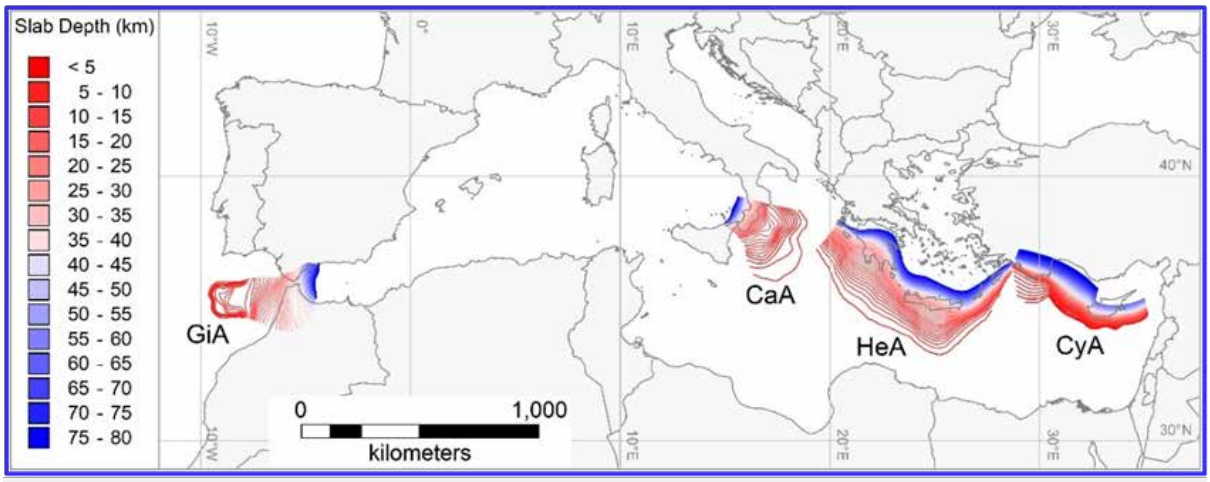
An estimate of the activity rates is obtained through an exploratory logic tree considering variable upper and lower seismogenic depths, magnitude upper bounds, depth-dependent rigidity, tectonic rates based on the geodetically observed convergence across subduction zones, seismic efficiency, and b values in an approach similar to Davies et al. (2019).

The average values of convergence rates are 1 mm/year for the Gibraltar Arc, 2.9 mm/year for the Calabrian Arc, 24 mm/year for the Hellenic Arc, and 12 mm/year for the Cyprus Arc (see Section 3.3.2). The magnitude upper bound of the magnitude-frequency distribution is constrained by the scaling relations of Allen and Hayes (2017), assuming that the entire seismic interface area may rupture in one single event.

The key parameters can be classified in two components: geometry related (the first three branching levels) and seismogenic parameters: the tectonic rates, seismic efficiency and regional bGR-values. The aim is to estimate the total variability for the number of events generated by the interface when cascading all key input parameters.

The conversion of the tectonic rates into seismic productivity was done with the model of Kagan (2002). The resulting distribution of activity rates consists of 2187 possible realizations of activity rates. However, such distribution is extremely computationally demanding for the hazard computation, thus the logic tree was trimmed down to 27 branches by combining the first two branching levels and a median crust rigidity model to describe the geometry and crust features; a single b-value is used, as well as three values of the tectonic rate and seismic coupling.

The values of the key input parameters are given in the table of **Fig. 3.10**. The maximum magnitude is coupled with the effective area of the subduction interface. The resulting magnitude frequency distributions are illustrated in **Fig. 3.11**.




| *Slab* | Upper Seism. Depth (km) | Lower Seism. Depth (km) | Conv. Rate (mm/yr) | Seism. Eff. | b-value |
|---|---|---|---|---|---|
| *GiA* | 8-10-11 | 21-27-33 | 0.8-1-1.2 | 0.1-0.2-0.3 | 1-1-1 |
| *CaA* | 9-10-12 | 19-34-39 | 2.7-2.9-3 | 0.1-0.2-0.3 | 1-1-1 |
| *HeA* | 9-10-11 | 30-34-38 | 19-23.8-28.6 | 0.1-0.2-0.3 | 1.05-1.07-1.15 |
| *CyA* | 6-8-10 | 25-31-38 | 9.8-12.2-14.6 | 0.1-0.2-0.3 | 1.05-1.07-1.15 |

**Fig. 3.10** (Top) Map view of the slab models, GiA - Gibraltar Arc, CaA - Calabrian Arc, HeA - Hellenic Arc, and CyA - Cyprus Arc. (Bottom) Key input parameters for the calculation of the magnitude-frequency distribution.

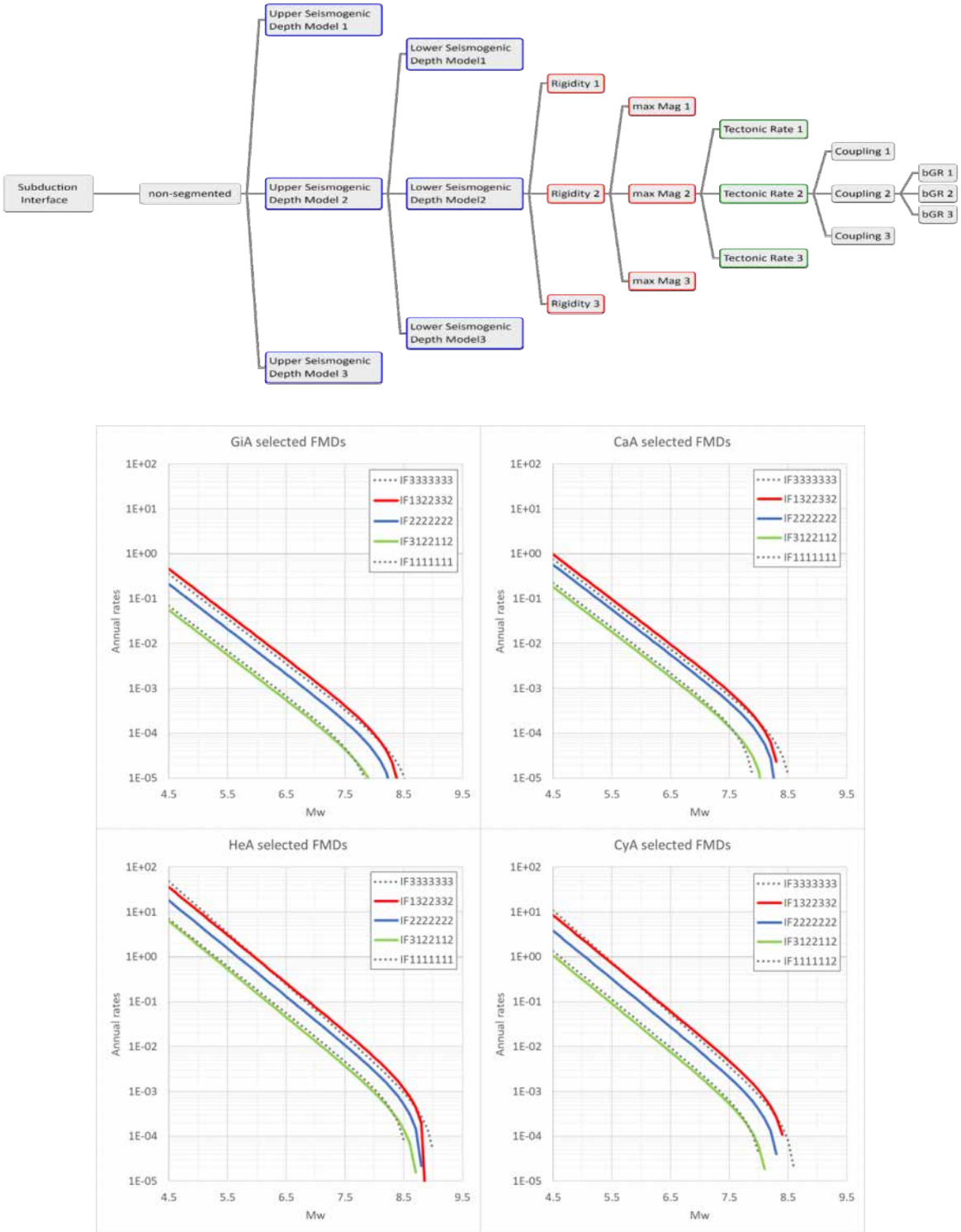



**Fig. 3.11** (Top) Alternative components used to derive a distribution for the activity rates of the subduction interface sources. (Bottom) Frequency-magnitude distributions of the four slab interfaces. The series of numbers in the distribution labels of each panel represents the selected component at each branch of the exploratory logic tree.

### 3.3.2 Deep Seismicity: Vrancea, Romania

A particular case of subduction slabs is the deep seismicity observed beneath the Vrancea Region, Romania. We keep the same seismogenic model (Danciu and Wössner 2014) proposed in the ESHM13 as illustrated in **Fig. 3.12**. The most hazardous seismicity in the Vrancea region, situated beneath the southern Carpathian Arc, is located at a depth between 70-150km. As there is currently no well-defined complex fault structure (no well-defined 3D fault plane), we model the pattern with four zones located at depth and a different zonation for the upper crust.

The seismicity parameters (GR parameters) are derived from the declustered catalogue and the data-driven derived completeness intervals applied to earthquakes occurring at a hypocentral depth greater than 70km.

Maximum magnitude for the seismicity in the Vrancea region was assessed by the historical record, i.e. we added an uncertainty to the maximum observed event and then added further increments with decreasing weights in the computation. The $M_{max}$-values for the seismicity at depths 70-90km depth are 7.5 and 7.8, the seismicity between 90-150km has $M_{max}$ values of 7.8 to 8.1.

The magnitude frequency distributions are the same as those used for the shallow crust regions, an exponential double truncation GR and tapered Pareto model, including the uncertainties.

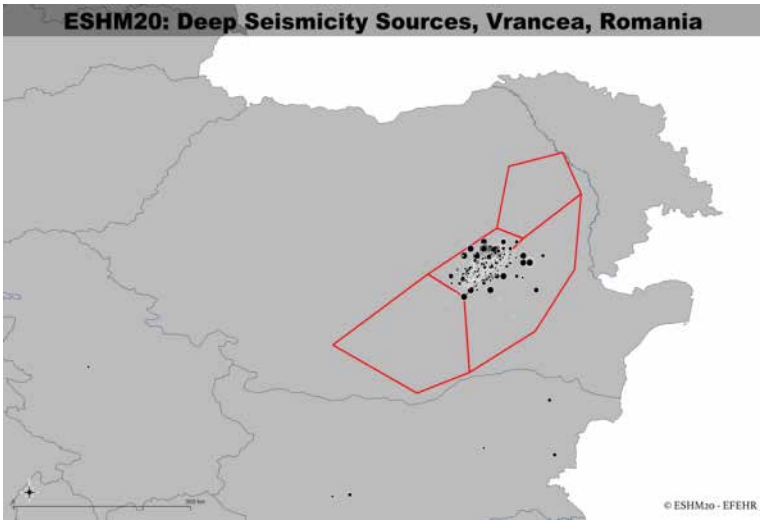



**Fig. 3.12** Deep Seismicity Sources – Azimuth dependent in Vrancea Region, Romania

## 3.4 Seismogenic Source Model - Logic Tree

Development of a seismic hazard model requires decisions and wide interpretations at every step to overcome limitations on data, randomness associated with input data, incomplete knowledge as well as challenges due to large scale modeling and calculations. One of the key constraints on the development of the logic tree was to design a logic tree that is feasible for the hazard calculation at the entire scale of the pan European region including Turkey.

In the design of the logic tree, we acknowledge the circular role of various inputs and components which might be reflected in a correlated logic tree branch. The logic tree is made of two main branching levels, area source model and smoothed seismicity plus fault model (**Fig. 3.13**). Both branching levels are based on the unified earthquake catalogue, the completeness zones, and regional bGR proxies from the TECTO model. The active fault model and the area source model are treated as independent, albeit in the area source delineation fault locations were used in some parts of the model. The weight assigned to a logic tree branch should reflect the independence between datasets, the assumptions taken on the model building and ultimately it shall represent the likelihood of that branch being the true model. Herein, weights are assigned by the experts deriving the input models, based on their judgement of the likelihood of the alternative models.

The logic tree structure of the seismogenic source model is asymmetric. The first branching level depicts the two main regional source models: the area sources model and the active fault with background smoothed seismicity model. The second branching level addresses the individual area sources for which two alternative magnitude-frequency distributions are handled as individual branches: one for the exponential double truncation GR and one for the tapered Pareto distribution. At the level of individual area sources, the main branching levels are the a and b-values and maximum magnitudes for area sources, whereas for active faults the uncertainties on the slip-rates and on the maximum magnitudes are considered.

We assume a constant seismic rate (above a given magnitude) when applying the maximum magnitude uncertainties of the area sources, and a constant seismic moment rate when considering maximum magnitude uncertainties of active faults (see Figure 2 in Youngs and Coppersmith 1985).

The uncertainty of the activity parameters for individual area sources is obtained from a simulated distribution of the activity parameters using the standard deviation of the aGR, bGR and the covariance matrix of these two parameters.

The median, and the two quantiles, i.e. $5^{th}$ and $95^{th}$ percentiles, populate the three independent branches of individual area sources for the exponential double truncated GR distribution. The weights assigned to the lower, median, upper activity parameters are 0.2, 0.6 and 0.2 respectively, representing the 5-$95^{th}$ confidence interval. For the tapered Pareto distribution only the median aGR, bGR values are used.

Three branching levels for the Mmax are symmetrically applied to all models, however with different weights. The Mmax weights for the area sources are the same as in the ESHM13 in the case of exponential double truncated GR distribution and equal weights for the tapered Pareto distribution.

For the active faults, only one value is used for the bGR-value (from TECTO) because it was found that the impact of b-value uncertainties on the magnitude-frequency-distributions is lower in comparison with the impacts of the fault slip-rates and Mmax uncertainties.

Note that, each seismic source is characterized by alternative values of various source parameters (i.e. average depth and the predominant style-of-faulting) that were not treated as epistemic but rather aleatory.

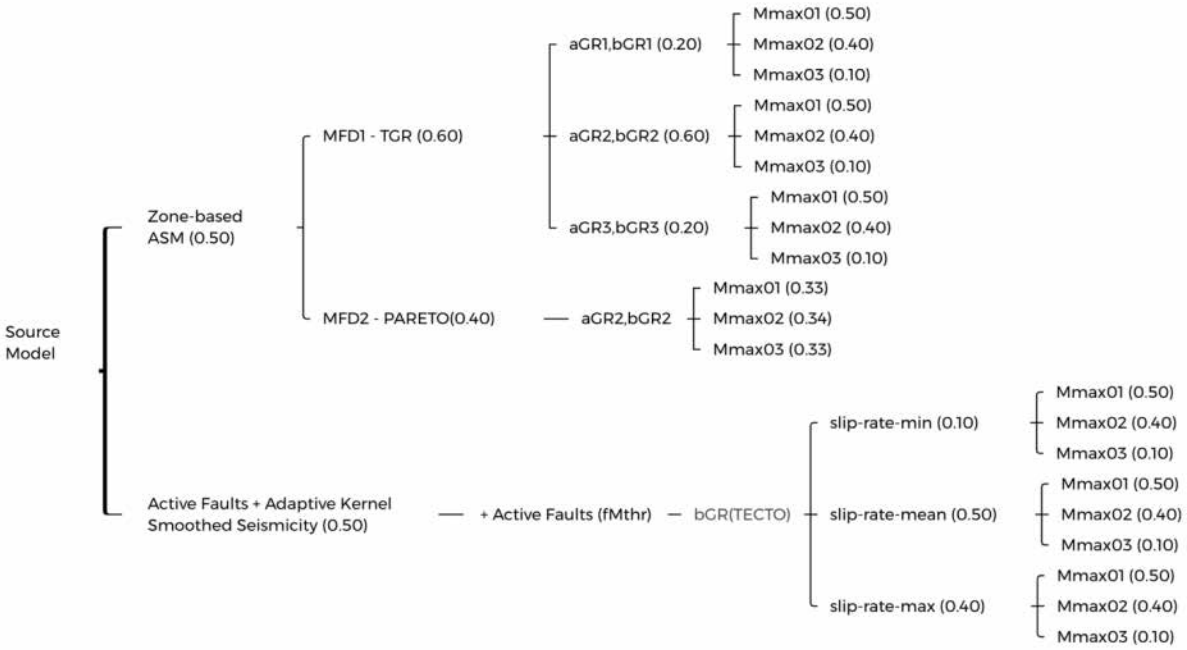



**Fig. 3.13** ESHM20: Schematic representation of the logic tree of the seismogenic sources for shallow crust earthquakes.

For deep seismicity of Vrancea, Romania and the subduction intraslab, the upper part of the logic tree is used with the same weights as indicated in Figure 13. To conclude, for the final calculation, the main seismogenic source model logic tree is supplemented with the subduction sources (both interface and in-slab) and Vrancea region in Romania for a complete model, and also combines ground motion logic tree model described in the next Chapter.

# 4 Ground Motion Model Characterization

## 4.1 Ground Motion Modelling Approach

The development of the ground motion model (GMM) logic tree for a region should aim to leverage as far as possible on the available strong motion data in order to characterize the expected ground motion from earthquakes, its aleatory variability, its epistemic uncertainty and its variability from region to region.

Previous models, including the 2013 European Seismic Hazard Model (ESHM13), have mostly attempted to do this by adopting a multi-model approach, in which a set of *best suited* GMMs are identified for each tectonic region of relevance (e.g., active shallow crustal seismicity, subduction interface, subduction in-slab, volcanic etc.) and weighted in a manner that reflects either their respective fits to the observed data (where available) or to expert views on suitability. Though practical, this approach has many theoretical limitations when it comes to their objectives of characterizing the epistemic uncertainty in the ground motions (Atkinson et al., 2014)

In the ESHM20 we adopt a different approach for characterising epistemic uncertainty, which is built around the concept of a *scaled backbone* ground motion model logic tree. In this approach a single ground motion model is calibrated (or selected from the literature) and to this model adjustment factors are applied that quantify the uncertainty in the expected ground motion as a result of the limited knowledge on the seismological properties in a region.

We not only quantify epistemic uncertainty in terms of a parameterised statistical distribution, we also set in place a framework through which new information and/or information at regional scale can adapt the backbone GMM logic tree and reduce the epistemic uncertainty over time. In the following we describe the scaled backbone ground motion model logic tree for the three main seismotectonic region types in Europe: shallow crustal seismicity (non-craton), seismicity in the stable craton region of north-eastern Europe, and subduction and deep seismicity (including the Hellenic, Calabrian, Cypriot and Gibralter arcs, as well as the Vrancea deep seismic zone). In addition, a number of *special cases* are considered in which the main approach is modified, or else decisions are made on the basis of insights and information from other data sets or studies.

A more comprehensive discussion of the theoretical and practical aspects of the ESHM20 GMM logic tree can be found in an accompanying EFEHR Report (Weatherill et al. 2022a) and in multiple journal publications that have arisen from this work (Kotha et al., 2020, 2022; Weatherill et al, 2020; Weatherill & Cotton, 2020; Weatherill et al.; 2022b).

## 4.2 Ground Motion Model Logic Tree for Shallow Seismicity

### 4.2.1 A regionally adjustable backbone Ground Motion Model (GMM) and logic tree

The backbone GMM logic tree for shallow crustal seismicity leverages heavily on the large volume of information in the Engineering Strong Motion (ESM) flatfile (Lanzano et al., 2019). With the available information a backbone ground motion model is constructed in which regional differences are quantified explicitly as random effects within the mixed effects regression process. Here, two prior regionalization models based on geology and tectonics are used to assign the earthquakes and the recording stations to their respective zones.
The two regionalisations (shown in **Fig. 4.1**) are i) a tectonic zonation based on an early version of the TECTO source model, and ii) a geological and tectonic zonation provided by Basili et al. (2019) and with small adaptations in the zone boundaries to improve the distribution of recording stations. Earthquake

sources are assigned to regions according to the TECTO model, while stations are assigned to the regions from the Basili et al. (2019) zonation.

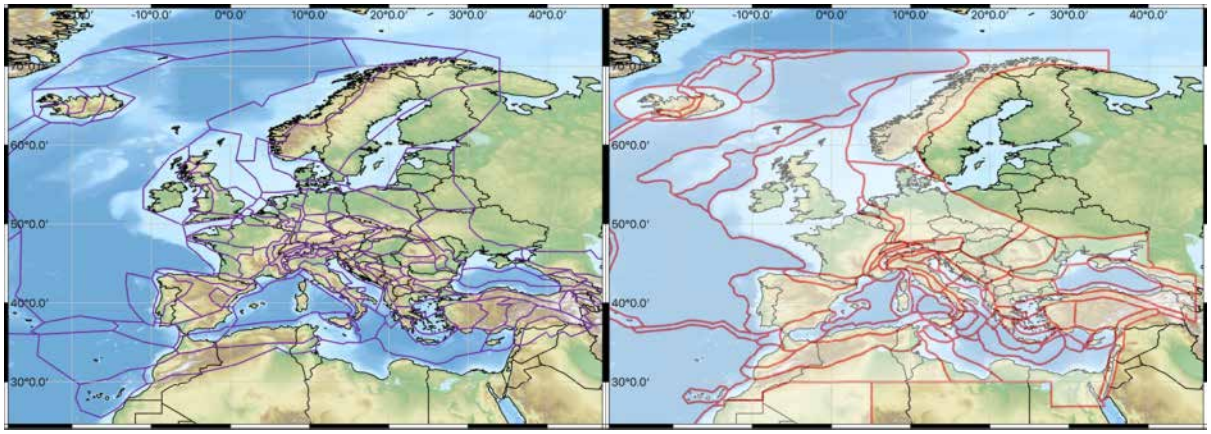


**Fig. 4.1** Regionalisation models used to calibrate the region-to-region random effects: i) TECTO model (left) and ii) adapted geological and seismotectonic model from Basili et al (2019) (right)

For the development of the model, records are selected from the ESM database according to the following criteria: i) events must be classified as non-subduction (according to the classification scheme indicated in the subsequent section), ii) reported hypocentral depth must be shallower than 40 km, iii) events must have ≥ 3 records, iv) only records with high-pass filter frequency $f_{hp} \leq 0.8\,T$ are retained for each period in the regression. The final database yields 18,222 records from 927 events ($3.1 \leq M_W \leq 7.4$) recorded at 1829 stations ($0 \leq R_{JB}(km) \leq 545$). The predictor variables of the model are moment magnitude ($M_W$), taken from the updated EMEC catalogue (section 3), hypocentral depth ($z_h$) in km, Joyner-Boore distance (km) and 30 m averaged shearwave velocity ($V_{S30}$, m/s). The Joyner-Boore distance is taken as the distance to the surface projection of the rupture plane where available (for most events with $M_W \geq 5.5$) or treated equivalently to epicentral distance otherwise. More details of the selection and data processing are given in Kotha et al. (2020).

The functional form of the backbone model is (Kotha et al. 2020):

$$\ln(\mu_{ijkl}) = e_1 + f_{R,g}(M_{W,i}, R_{JB,ij}) + f_{R,a}(R_{JB,ij}, k) + f_M(M_{W,i}) + \delta L2L_l + \delta B^0_{e,i} + \delta S2S_j + \delta W_{e,ij} \quad \text{(Eq. 6.1)}$$

$$f_M(M_{W,i}) = \begin{cases} b_1 \cdot (M_W - M_h) + b_2 \cdot (M_W - M_h)^2 & M_W \leq M_h \\ b_3 \cdot (M_W - M_h) & M_W > M_h \end{cases} \quad \text{(Eq. 6.2)}$$

$$f_{R,g}(M_{W,i}, R_{JB,ij}) = \left(c_1 + c_2 \cdot (M_W - M_{ref})\right) \cdot \ln \sqrt{\frac{(R^2_{JB} + h^2_D)}{(R^2_{ref} + h^2_D)}} \quad \text{(Eq. 6.3)}$$

$$f_{R,a}(R_{JB,ij}, k) = \frac{c_3 + \delta c_{3,k}}{100} \cdot \left(\sqrt{(R^2_{JB} + h^2_D)} - \sqrt{R^2_{ref} + h^2_D}\right) \quad \text{(Eq. 6.4)}$$

Where $\mu_{ijkl}$ is the PGA or spectral acceleration at period T (Sa(T)) from event $i$ located in event region $l$, recorded at station $j$ located in station region $k$. The definition of the horizontal component of motion is RotD50 (Boore, 2010). $M_h$, $M_{ref}$ and $R_{ref}$ are period-independent constants, which take the values $M_W$ 6.2, $M_W$ 4.5 and $R_{JB}$ 30 km respectively in Kotha et al. (2020); however, the hinge-magnitude ($M_h$) was revised in the recent paper of Kotha et al. (2022) to $M_W$ 5.7 based on updated analysis of ground motion records from larger magnitude events in Europe and globally. $h_D$ is the near-source saturation plateau, a term that is dependent on hypocentral depth ($z_h$). This is fixed to a reference value depending whether the depth falls in one of three bins: $z_h \leq 10\ km$, $10 < z_h(km) \leq 20$ or $z_h(km) > 20$, for $h_D$

takes the value 4 km, 8 km and 12 km respectively. Robust linear mixed effects regression is used, which down-weights outliers when fitting the fixed effects ($e_1, b_1, b_2, b_3, c_1, c_2$ and $c_3$).

In addition to the fixed effects, the model contains five random effects: $\delta L2L_l$ a source-region specific scaling parameter (depending on the location of the event), $\delta B_e^0$ the between-event residual *filtered for source region-to-source region variability*, $\delta S2S_j$ the site-to-site variability and $\delta c_{3,k}$ the between-region variability of the residual attenuation term $f_{R,a}$. Finally, $\delta W_{e,ij}$ describes the remaining region-, event- and site-corrected variability, representing therefore the *apparent* aleatory variability of the model.

These five random effects are described by Gaussian distributions such that: $\delta L2L_l = \mathcal{N}(0, \tau_{L2L})$, $\delta B_e^0 = \mathcal{N}(0, \tau_0)$, $\delta S2S_j = \mathcal{N}(0, \phi_{S2S})$, $\delta c_{3,k} = \mathcal{N}\left(0, \tau_{c_3}\right)$ and $\delta W_{e,ij} = \mathcal{N}(0, \phi_0)$. We exploit the separation of region-specific random effects ($\delta L2L_l$ and $\delta c_{3,k}$) from the more traditional between-event, site-to-site and site-corrected within event random effects ($\delta B_e^0$, $\delta S2S_j$ and $\delta W_{e,ij}$) in order to be able to identify and apply regional calibrations of the GMM, and at the same time removing some of the ergodicity that would otherwise result in higher aleatory variability, and thus higher hazard.

The spatial distributions of the source-region to source-region random effects ($\delta L2L_l$) and the residual attenuation region random effects ($\delta c_{3,k}$) are shown in **Fig. 4.2** and their total variability in **Fig. 4.3**. Where $\delta L2L$ is higher we expect earthquakes in these regions to produce characteristically larger ground motions than the "average" across Europe, and where it is lower, we would therefore expect weaker ground motions. Similarly, $\delta c_{3,k}$ controls the rate of decay of ground motion over longer distances ($R_{JB} \geq 80$ km), but as $c_3$ is negative then a higher $\delta c_{3,k}$ results in a slower modelled attenuation, while a lower $\delta c_{3,k}$ indicates a faster attenuation.

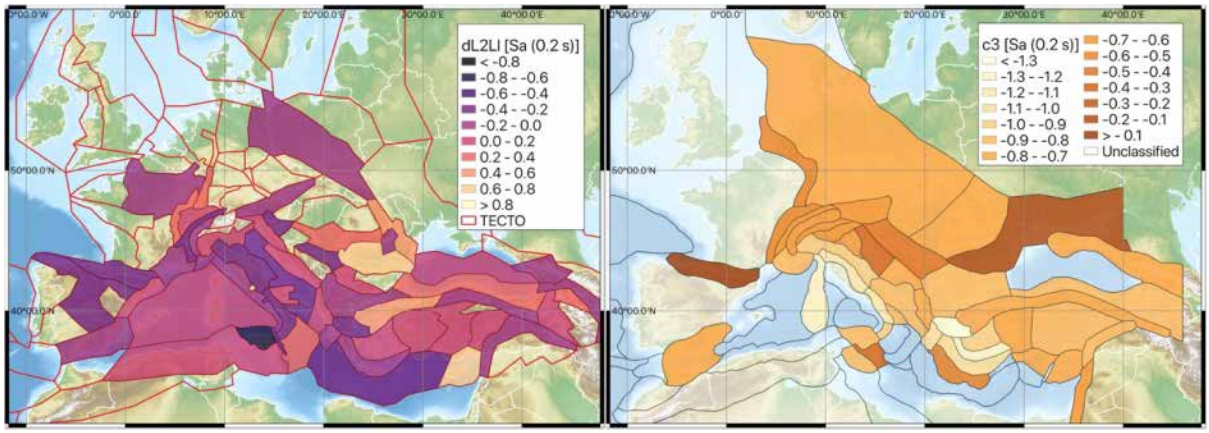



**Fig. 4.2** Distribution of source-region specific residual $\delta L2L_l$ (left) and net residual anelastic attenuation term ($c_3 + \delta c_{3,k}$) (right) for Sa (0.2 s).

The described by the terms $\tau_{L2L}$ and $\tau_{c_3}$ effectively describe the full region-to-region variability in source and residual attenuation implied by the available strong motion data, conditional upon the assumed regionalization. As can be seen in **Fig. 4.2**, however, there are many regions for which $\delta L2L_l$ and $\delta c_{3,k}$ have not been determined owing to insufficient data. In this case we do not know the specific source-region or residual attenuation region term, but rather we assume that the probability distribution of these adjustments is described by the full region-to-region variability quantified by $\mathcal{N}(0, \tau_{L2L})$ and $\mathcal{N}\left(0, \tau_{c_3}\right)$. We use this information to create the *default* scaled backbone ground motion logic tree to be applied in regions were data are absent, but we believe the probability of a given $\delta L2L_l(T)$ and $\delta c_{3,k}(T)$ values follow the aforementioned distributions. The distributions themselves are mapped into branches of the logic tree using the N-branch discrete approximation to a Gaussian distribution proposed by Miller & Rice (1983). This approximation will be used throughout the logic tree, so the corresponding weights and $\varepsilon$ values are for the N = 3, 5 and 7 branch cases are provided in Table 1.

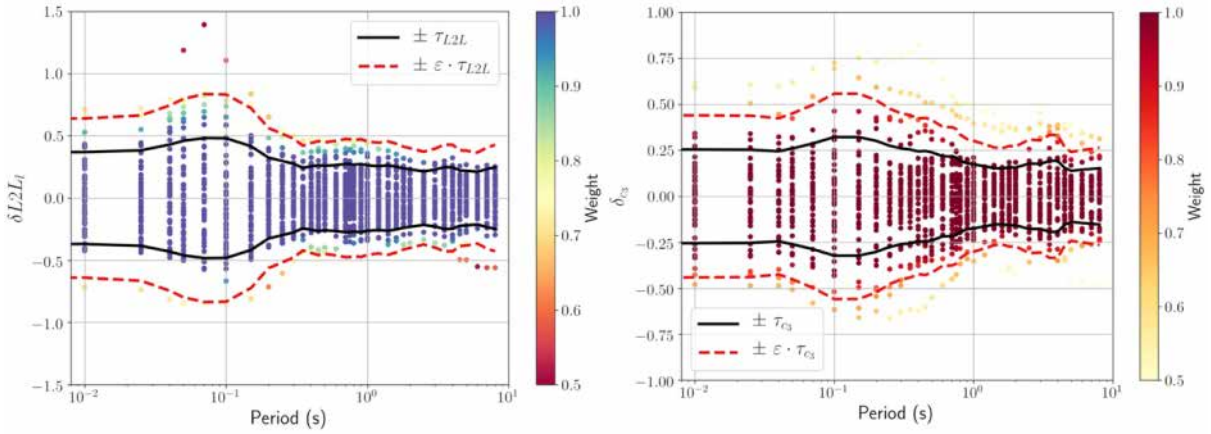



**Fig. 4.3** Distribution of region-specific random effects for each spectral period: $\delta L2L_l$ (left) and $\delta c_{3,k}$ (right). Markers are colour-scaled according to their resulting weight in the robust mixed effects regression, and the respective standard deviations shown in black line.

The *default* backbone GMM logic tree has two branch sets, which correspond to the residual attenuation region and the source region variability respectively. This is shown in the top half of **Fig. 4.4**, in which the source-region distribution is mapped into 5 discrete branches, and the residual attenuation region distribution into 3 branches.

In the case of the source-region variability we take the larger of two values: $\tau_{L2L}$ and $\sigma_\mu$, the latter of which corresponds to the statistical uncertainty in the regression determined according to the approach of Al Atik & Youngs (2014).

By taking the larger of $\tau_{L2L}$ and $\sigma_\mu$ we ensure that over the range of $M_W$ and $R_{JB}$ scenarios that are well constrained in the regression (for which $\sigma_\mu$ is small) the epistemic uncertainty is not less than the source-region to source-region variability.

At the same time, for scenarios that are poorly constrained in the regression, typically large $M_W$ and small $R_{JB}$ the statistical uncertainty is then dominant. This is intended to minimize the degree of double counting of epistemic uncertainty where region-to-region variability may be present within the statistical uncertainty.

**Table 6.1:** Weightings and an $\varepsilon$ values for the 3-, 5- and 7-branch discrete approximations to a Gaussian distribution based on Miller & Rice (1983).

| N = 3 | | N = 5 | | N = 7 | |
|---|---|---|---|---|---|
| **Weight** | $\boldsymbol{\varepsilon}$ | **Weight** | $\boldsymbol{\varepsilon}$ | **Weight** | $\boldsymbol{\varepsilon}$ |
| 0.167 | -1.732051 | 0.011257 | -2.85697 | 0.000548 | -3.750440 |
| 0.666 | 0.000000 | 0.222076 | -1.355626 | 0.030757 | -2.366759 |
| 0.167 | 1.732051 | 0.533334 | 0.000000 | 0.240123 | -1.154405 |
| | | 0.222076 | 1.355626 | 0.457144 | 0.000000 |
| | | 0.011257 | 2.85697 | 0.240123 | 1.154405 |
| | | | | 0.030757 | 2.366759 |
| | | | | 0.000548 | 3.750440 |

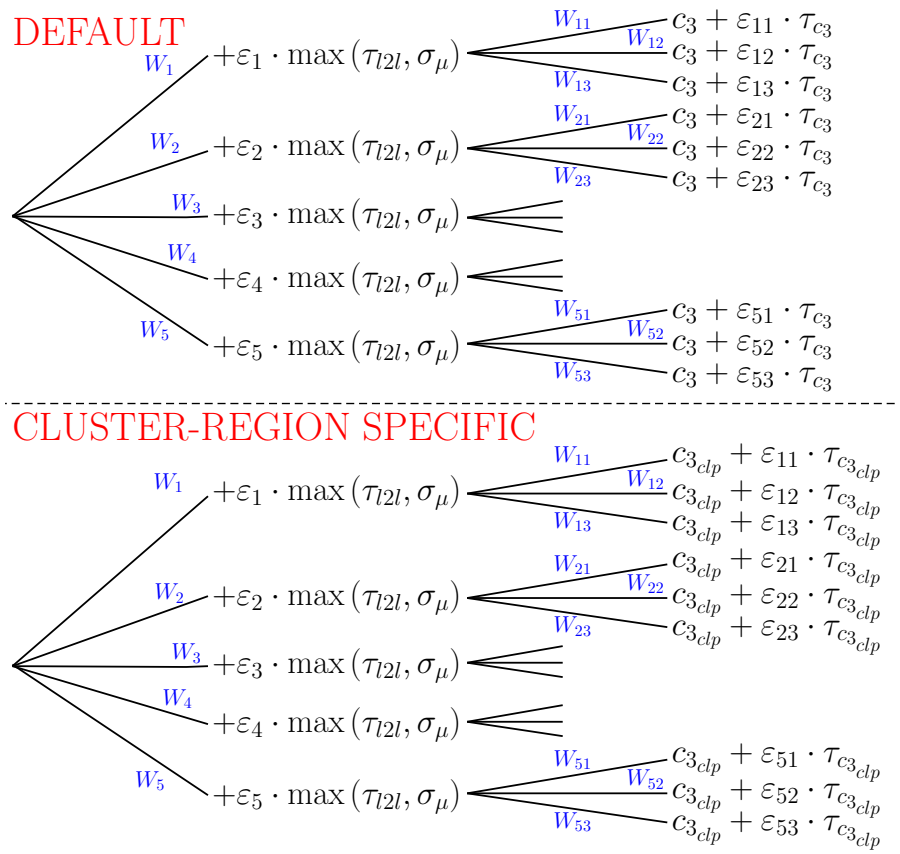



**Fig. 4.4** Full scaled backbone ground motion model logic tree for shallow crustal seismicity in Europe, excluding the stable shield region.

Explicit within the formulation of the *default* backbone GMM logic tree is the assumption that the epistemic uncertainty in the source-region and residual attenuation region is described by their respecive full region-to-region variabilities.

For regions where data are available and region-specific values have been determined, however, the local adjustments to the backbone GMM and their respective uncertainties should be used in order to regionalize the GMM and reduce the epistemic uncertainty. In the case of $\delta L2L_l$ the regional variability has been explored and interpreted, and we find that not only are the specific spatial trends difficult to reconcile with previous observations of regional variation in ground motion in Europe, for many regions the values are constrained by very few events, and in most of these cases the events may come from a single seismic sequence.

The decision is taken *not* to adopt the local $\delta L2L_l$ distributions but rather maintain the full $\tau_{L2L}$ distribution. This should of course be revisited in future when more data are available and could still be re-considered at a local level should specific countries or users wish to adopt the local values.

In contrast with $\delta L2L_l$, our assessment of $\delta_{c_3,k}$ led us to conclude that not only are these values better constrained by data, with more stations per calibrated region, they are also consistent with previous investigations into regional variability of strong motion attenuation within Europe (e.g., Kotha *et al.*, 2016; Scasserra et al., 2009).

In this case a region-specific $\delta_{c_3,k}$ term has been calculated for 42 zones (from 111 considered within the original regionalization), two of which are subsequently removed as they were found to be

constrained by only one station and their statistical uncertainty was almost equal to the total region-to-region variability.

These zones are Southern Betics (Spain) and Central Europe (covering Germany, Poland, Czechia, Slovenia and Denmark). From the 40 remaining zones, however, there are many regions showing similar period-dependent trends in $\delta_{c3,k}$, many of which are associated with moderate statistical uncertainty. The data show clear evidence of broader scale regional trends in attenuation, so rather than define a zone-specific $\delta_{c3,k}$ distribution we apply hierarchical cluster analysis to group these into five *clusters (or cluster regions)* that capture the main regional variations in attenuation.

Within each cluster region the period-dependent $\delta_{c3,k}$ values are grouped and using a Bayesian regression analysis (with the original distribution of $\tau_{c_3}$as a prior) we define a cluster region specific distribution of $\delta_{c_3}$, which we define as $\mathcal{N}(\delta c_{3,cl_p}, \tau_{c3,cl_p})$ where $clp$ refers to cluster region $p$ for $p = 1, 2, \ldots, 5$. The spatial distribution of the clusters and the distributions of $c_3$ are shown in **Fig. 4.5**. Further details on this process can be found in Weatherill *et al.* (2020a).

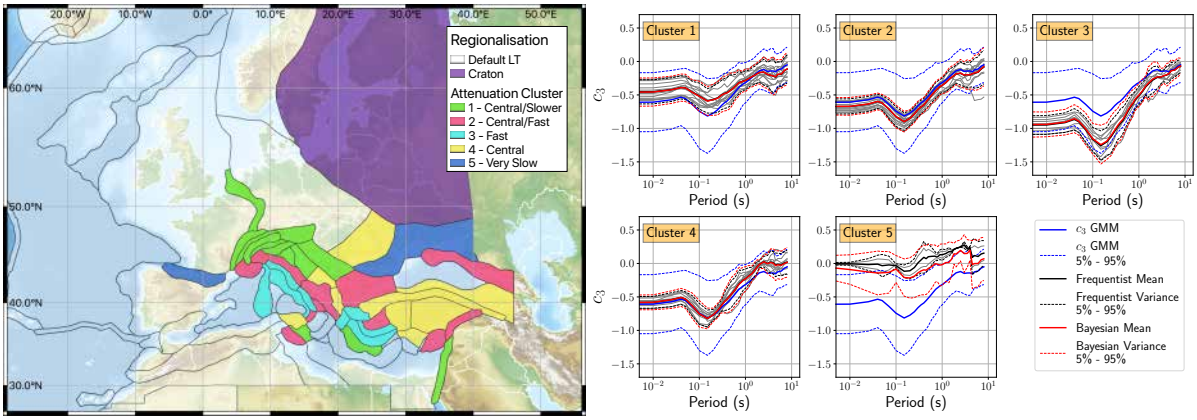



**Fig. 4.5** Spatial distributions of the five residual attenuation cluster regions (left) and their corresponding distributions of $c_3$ (right). Figure from Weatherill et al. (2020)

With the five cluster-specific distributions established and the regional trends in residual attenuation identified from the available data, the set of adjustments made to the scaled backbone ground motion model are then cluster specific.

This means that the N-branch logic tree illustrated in the top half of **Fig. 4.4** is then repeated for each of the five cluster regions, with $c_3$ and $\varepsilon \cdot \tau_{c3}$ replaced by $c_{3_{cl,p}}$ and $\varepsilon \cdot \tau_{c3,cl_p}$ for $p = 1, 2, \ldots, 5$ respectively, while the *default* logic tree with full region-to-region variability is applied to those regions indicated by the semi-transparent white shaded regions in **Fig. 4.5**. The complete logic tree for shallow crustal seismicity *excluding the stable shield region of north-eastern Europe*, is shown in Figure 6.6.

### 4.2.2 Aleatory Uncertainty

As the region-to-region variability in source and attenuation characteristics retrieved from the mixed effects regression are used to construct the GMM logic tree, these terms are removed from the aleatory variability of the scaled backbone GMM. The resulting total aleatory uncertainty ($\sigma_T$) is therefore described by $\sigma_T = \sqrt{\tau_0^2 + \phi_0^2 + \phi_{S2S}^2}$. The corresponding between-event, site-to-site and site-corrected within event standard deviations provided by Kotha et al. (2020) are homoscedastic, i.e., independent of the scenario.

Many studies of ground motion variability in recent years, however, have identified heteroskedasticity in some (or all) of these residual terms (e.g., Rodriguez-Marek et al. 2013; Al Atik, 2015). Re-analysis of the uncertainty terms by Kotha et al. (2022) also showed that the assumption of homoskedasticity with respect to earthquake magnitude was not consistent with the data at large magnitudes.

The change in variability of the residual terms $\delta B_e^0(T)$ and $\delta W_{e,ij}(T)$ with respect to $M_W$ and $R_{JB}$ are explored by taking the observed standard deviations of the residuals within fixed bins of 0.5 $M_W$ units and a set of logarithmically spaced distance bins respectively (Weatherill et al. 2020). This analysis agrees with that of Kotha et al. (2022) in identifying reductions in $\tau_0$ and $\phi_0$ at higher magnitudes, while no compelling evidence was found for distance-dependent heteroskedasticity.

For the case of $\tau_0$ the magnitude-binned standard deviations are compared with a heteroskedastic model of $\tau$ proposed by Al Atik (2015), which is itself fit to a global database of ground motions from shallow crustal events and is therefore richer in observations from larger magnitudes than the ESM database. We find that for the small-to-moderate magnitudes the observed values of $\tau_0$ in the Kotha et al. (2020) homoscedastic model agrees well with the model of Al Atik (2015).

At larger magnitudes the observed $\tau_0$ is smaller than the heteroskedastic model of Al Atik (2015); however, with so few observations from large magnitude (M ≥ 6.0) events in ESM we would consider the model of Al Atik (2015) better constrained. Therefore, we adopt the global $\tau$ model of Al Atik (2015) in place of the homoscedastic $\tau_0$ model of Kotha et al. (2020) without modification.

When considering the magnitude-dependence of observed $\phi_0$ we again compare this against a corresponding global model of $\phi_0$ proposed by Al Atik (2015). With the largest magnitude events in the ESM database being particularly well-recorded, we have greater confidence in the constraint of $\phi_0$ from the observed ESM data and thus attempt to calibrate a magnitude-dependent $\phi_0$ model.

Comparing the calibrated $\phi(M_W, T)$ values against the model of Al Atik (2015) we find good agreement at larger $M_W$ values, while for small-to-moderate magnitudes the observed $\phi_0$ are similar but slightly larger. In this case we replace the homoscedastic $\phi_0$ model of Kotha et al. (2020) with a new magnitude-dependent model calibrated upon the ESM data set. We do not adopt the Al Atik (2015) $\phi_0(M_W, T)$ model in the present case.

### 4.2.3 Site Amplification and Characterisation

The functional form of the Kotha et al. (2020) in equation 6.1 does not contain an explicit site amplification term, $f_{SITE}$. In their original paper, the authors present alternative models for $f_{SITE}$ dependent on either $V_{S30}$ or topographic slope. These are determined by regression of the $\delta S2S_S$ terms against various predictor variables, assuming a quadratic form such that $f_{SITE}(X) = g_0 + g_1 \cdot \ln\left(\frac{X}{X_{ref}}\right) + g_2 \cdot \ln\left(\frac{X}{X_{ref}}\right)^2$, where $X$ refers to either topographic slope or measured $V_{S30}$. From their robust mixed effects regression $\delta S2S_S$ is available for more than 1,800 sites with 3 or more records, 644 of which would be considered "well constrained" by 5 or more records. From the full set of 1,800 site only 419 stations report a measured $V_{S30}$, the rest requiring a $V_{S30}$ that is inferred from slope using the method of Wald & Allen (2007).

In their estimation of $f_{SITE}(V_{S30})$ only the subset of measured observations was used, meaning that the resulting variability ($\phi_{S2S}$) represents only the site-to-site variability for measured $V_{S30}$ cases, not for

Europe as a whole. The spatial distributions of the measured and inferred $V_{S30}$ stations are shown in **Fig. 4.6**.

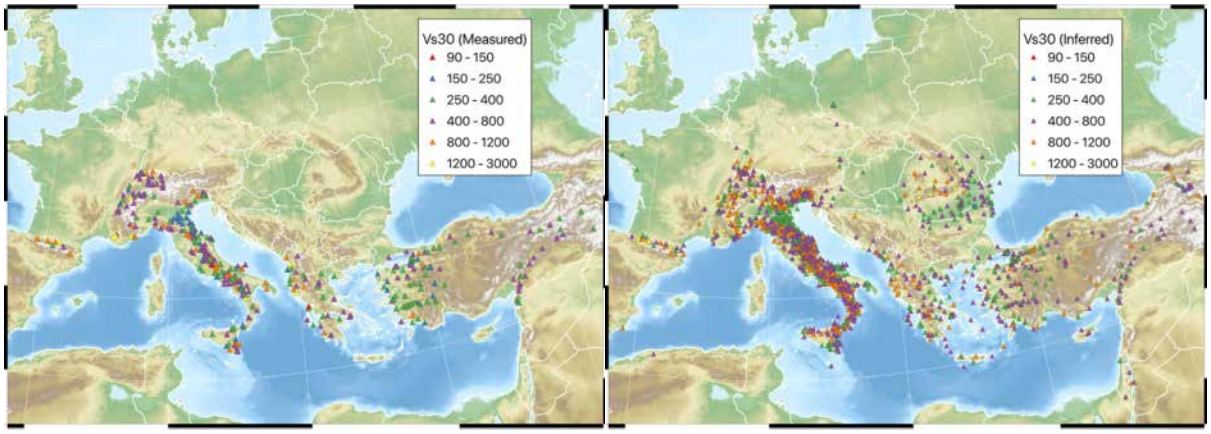



**Fig. 4.6** Distribution of sites used in the construction of the Kotha et al. (2020) model for the case when a measured $V_{S30}$ available (left) and when $V_{S30}$ is inferred from topography (right)

To capitalise on the model complete data set of $\delta S2S_S(T)$ including stations without a measured $V_{S30}$ case, a new $f_{SITE}$ regression is undertaken in which different coefficients and resulting $\phi_{S2S}(T)$ are defined depending on whether the predictor $V_{S30}$ refers to a measured or an inferred site condition.

The trend in $\delta S2S_S(T)$ against $V_{S30}$ is shown in **Fig. 4.7** for Sa (0.2 s) and Sa (1.0 s) considering both the measured and inferred cases. To define $f_{SITE}$ we use a two-segment piecewise linear function:

$$f_{site}(\theta, T) = \begin{cases} g_1(T) \ln\left(\frac{\theta}{\theta_{ref}(T)}\right) & \text{for } \theta \leq \theta_c \\ g_1(T) \ln\left(\frac{\theta_c}{\theta_{ref}(T)}\right) & \text{for } \theta > \theta_C \end{cases} + \varepsilon_s \cdot \phi_{S2S}(T) \qquad \text{(Eq. 6.5)}$$

where $\theta$ is equal to $V_{S30}$ and $g_1(T), \theta_{ref}(T)$ and $\phi_{S2S}(T)$ depend on whether the $V_{S30}$ is measured or inferred. $\theta_c$ refers to the $V_{S30}$ above which $f_{SITE}$ is constant, which in the current implementation is fixed to 1100 m/s. The fitted model is also shown in Figure 6.9. The resulting function results in a smaller increase in amplification with decreasing $V_{S30}$ for the inferred $V_{S30}$ case than for the measured $V_{S30}$ case, but this is partly compensated for the in the hazard calculations by the larger $\phi_{S2S}$ term. For the measured $V_{S30}$ case, $\phi_{S2S}(T)$ was found to be almost identical to that of the Kotha et al. (2020) $V_{S30}$-based amplification model, and the degree of amplification comparable over the range of $V_{S30}$ constrained by data. This functional form does not consider nonlinearity in amplification for strong shaking on soft soil sites as the available database contains too few records of strong earthquakes on such sites to constrain a nonlinear amplification model empirically.

The dependence of the amplification model on whether the $V_{S30}$ corresponds to a measured or inferred site has implications for its application. The ESHM20 serves multiple purposes: i) calculation of seismic hazard on reference rock (Eurocode 8 class A, $V_{S30}$ 800 m/s) that may form an input into seismic resistant design codes, ii) calculation of ground shaking on the surface soil as input into seismic risk analysis including the 2020 European Seismic Risk Model (ESRM20).

In the case of the former, we define the reference rock as corresponding to a *measured* $V_{S30}$ case, and adopt the lower $\phi_{S2S}$ value accordingly. This decision is made according to the context in which site amplification is applied in design codes.

As code amplification factors are generally calibrated with conservatism to account for uncertainty in site response, and in the current Eurocode 8 case higher amplification is required in the case in which the site has unknown $V_S$ properties, adoption of the inferred $V_{S30}$ $\phi_{S2S}$ would likely double count uncertainty already implicit in the code amplification factors.

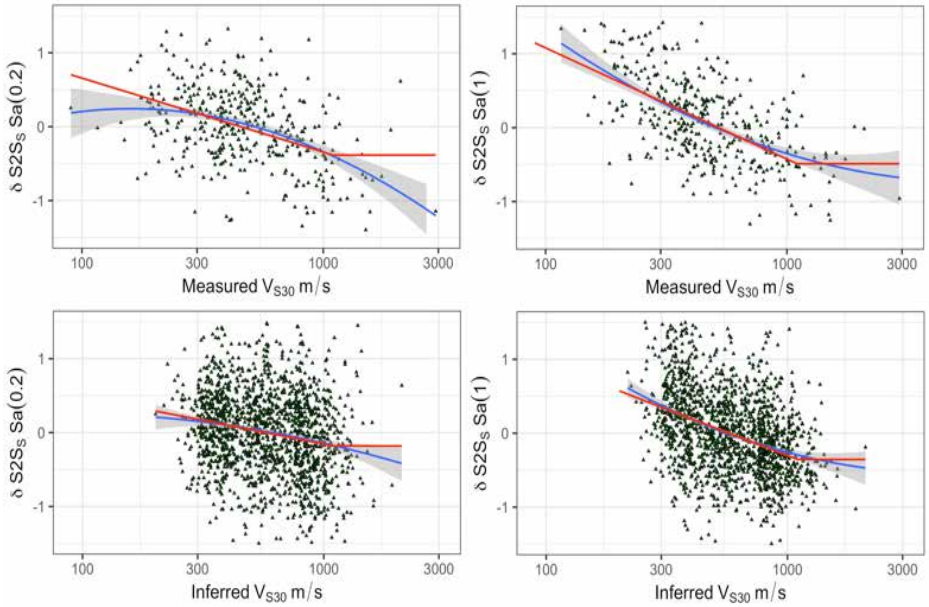



**Fig. 4.7** Relation between $\delta S2S_S(T)$ and $V_{S30}$ for Sa (0.2 s) (left column) and Sa (1.0 s) (right column) for the measured (top row) and inferred $V_{S30}$ case (bottom row). Blue lines and grey shaded region indicate a non-parametric LOWESS regression fit, while red lines indicate the fit of the two-segment linear $f_{SITE}$ model.

For the seismic risk case, including ESRM20, the corresponding measured or inferred $V_{S30}$ can be chosen depending on whether the sites in question take their properties from direct measurement, or detailed local scale microzonation, or whether they are inferred from regional scale proxies, such as topographic slope.

As the ESRM20 requires characterisation of the soil properties across all of Europe, the inferred $V_{S30}$ case is dominant. Though not described here, the ESRM20 actually adopts a form of equation 6.5 that is dependent on both slope *and* local geology, thus refining the degree of amplification depending on the geological environment. $\phi_{S2S}(T)$ values are found to be very similar regardless of whether topographically inferred $V_{S30}$ is used or slope directly (with or without local geology).

Further details on the construction, application and seismic hazard and risk implications of the site amplification model can be found in Weatherill et al. (2022, *in preparation*).

## 4.3 Scaled Backbone GMM Logic Tree for the Stable Craton

### 4.3.1 Defining and delineating the cratonic region of Europe

In the previous section it has been indicated that the *default* shallow crustal seismicity GMM logic tree is applied in regions for which we have insufficient data to calibrate $\delta L2L_l$ and/or $\delta c_{3,k}$. A fundamental assumption that underpins this, however, is that the seismological characteristics of the regions for which we lack data are within the range implied by the full region-to-region distributions determined using the ESM data (Lanzano et al 2019).

As data is lacking for much of northern Europe, we cannot necessarily say for certain to what extent this is true. We therefore consulted other geophysical data sets whose properties *may* be relevant for ground motion characterization, and from these sought to identify where regional differences exist.

In **Fig. 4.8** we show the spatial patterns of 1-second Rayleigh wave attenuation quality factor ($Q_{LG}$) (Mitchell *et al*., 2008), upper mantle shearwave velocity anomaly at 120 km depth ($\delta V_S$) (Mooney et al., 2012), heat flow (Lucazeau, 2019) and Moho depth (Grad et al. 2009; Szwillus et al. 2019).

From these data sets we observe a clear difference in geophysical properties when we contrast southern and western Europe against northeastern Europe. Specifically, in northeastern Europe we observe higher $Q_{LG}$ (implying slower attenuation), higher mantle shearwave anomaly, lower heatflow and deeper Moho. The transition between the two environments is generally clear and corresponds to the Trans-European Suture zone (TESZ).

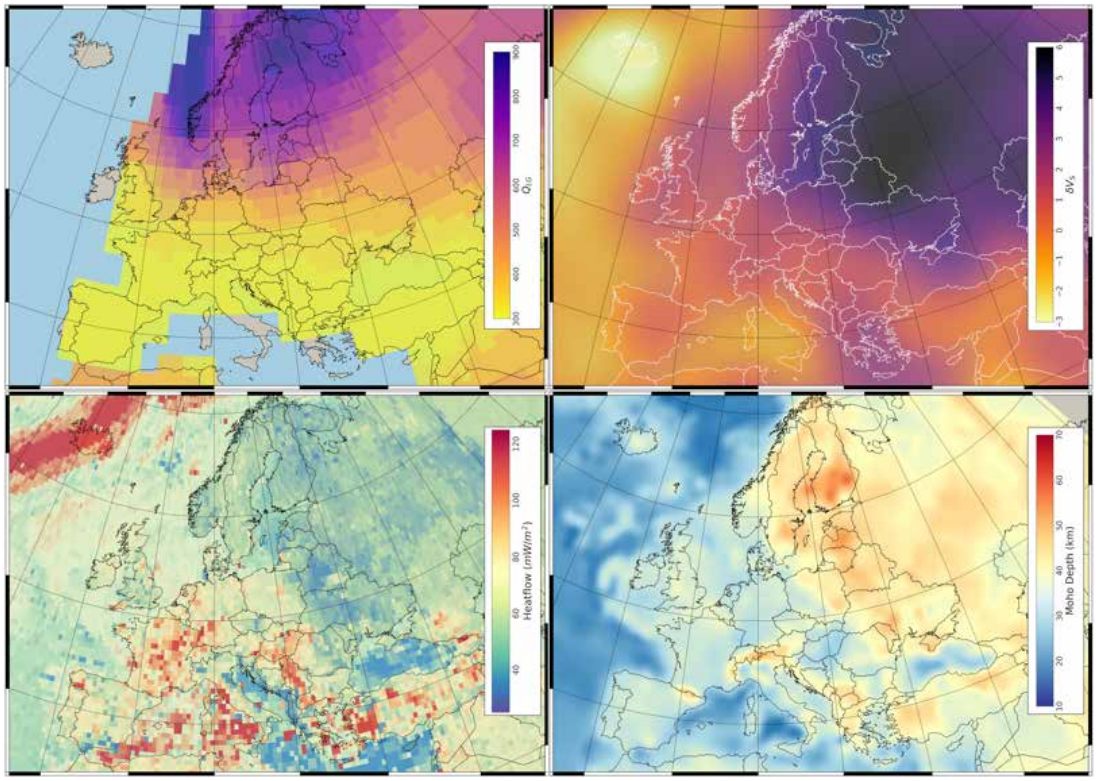



**Fig. 4.8** Geophysical data sets across Europe: $Q_{LG}$ (top left), $\delta V_S$ (top right), Heatflow (bottom left) and Moho depth (bottom right)

As the observed strong motion data in the ESM database fall entirely on the southern and western size of the TESZ, it is not possible to assume with certainty that the seismological characteristics of northeastern Europe are represented within the region-to-region variability. Instead, we find that a better analogue can be found in eastern North America.

This is not necessarily a novel insight, as previous seismic hazard models for northern Europe, including the ESHM13, have adopted GMMs from the central and eastern United States (CEUS). We make use of this analogue in order to leverage upon recent developments in ground motion modelling within the CEUS as a means of potentially regionalizing our shallow crustal scaled backbone logic tree approach further. We summarise this in the following, but full details describing the development of the GMM model and scaled backbone GMM logic tree for northeastern Europe can be found in Weatherill & Cotton (2020)

### 4.3.2 Calibrating the scaled backbone GMM for median ground motions

The Next Generation Attenuation East (NGA East) project represent a significant step forward in strong ground motion modelling for the central and eastern United States (Goulet et al. 2018). Among the wide range of outputs and publications from the project we find a suite of 20 new individual median GMMs for very hard rock ($V_S$ 3000 m/s) conditions, the characterization of a full GMM logic tree to represent epistemic uncertainty, a site amplification model (Stewart et al. 2020; Hashash et al. 2020) and an aleatory uncertainty model (Al Atik, 2015). In the absence of observed ground motion data for northeastern Europe, we utilize these products to undertake a pseudo-regionalisation of the scaled backbone ground motion model.

We begin with the 20 individual ground motion models for CEUS, which are developed using a mixture of simulated, empirical or hybrid ground motion modelling methods with different assumptions about the seismological properties of the crust.

We also add the model of Pezeshk et al. (2011), which precedes the NGA East project but adopts many of the same characteristics. Each model is assumed equally weighted. From the 21 GMMs we generate expected median ground motions (on very hard rock) for a broad range of scenario magnitudes ($4.0 \leq M_W \leq 8.0$) and distances ($1.0 \leq R_{RUP}$ (km) $\leq 500$ km). The epistemic uncertainty on the median ground motions is then described by a log normal distribution:

$$P(\ln Y(M_W, R_{RUP}, T)) = \mathcal{N}\left(\boldsymbol{\mu}(M_W, R_{RUP}, T), \boldsymbol{\sigma_\mu}(M_W, R_{RUP}, T)\right) \quad \text{(Eq. 6.6)}$$

The distribution is scenario-dependent, where $\boldsymbol{\mu}$ is the median expected ground motion and $\boldsymbol{\sigma_\mu}$ represents the epistemic uncertainty in the median.

Though this distribution could be used in this non-parametric form to capture the ground motion, we instead use this distribution of scenario ground motions to fit a parametric GMM adopting the same functional form as that of Kotha et al. (2020) (equations 6.1 – 6.4). Several modifications are made, however, the preferred distance metric of the NGA East GMMs is closest distance to rupture ($R_{RUP}$). As such $R_{ref}$ is now fixed to 1 km and $h_d$ to 5 km (no longer hypocentral depth dependent):

$$\ln Y_{median}(M_W, R_{RUP}, T) = e_1 + f_M(M_W, T) + f_{R,g}(M_W, R_{RUP}, T) + f_{R,a}(R_{RUP}, T) + \varepsilon_\mu \cdot \sigma_\mu(T) \quad \text{(Eq. 6.7)}$$

This model is fit using robust least squares regression, with no mixed effects. The standard deviation ($\sigma_\mu$) therefore represents the model-to-model variability. As with the shallow crustal GMM logic tree, we use the N-branch discrete approximation to a Gaussian distribution by Miller & Rice (1983) to then map the epistemic uncertainty distribution ($\mathcal{N}(\mu, \sigma_\mu)$) into a scaled backbone GMM logic tree.

The response spectra from the suite of NGA East models, the default (or general) shallow crustal seismicity logic tree, the proposed parametric ground motion model logic tree and a recent GMM developed for nuclear power plants in Finnland (Fülöp et al. 2020) are compared for four scenarios in **Fig. 4.9**.

The parametric GMM and its epistemic uncertainty generally capture the range of median ground motions from the NGA East model well and are in good agreement with the GMM of Fülöp et al. (2020). In contrast with the default GMM logic tree, the ground motions are significantly higher for short periods ($T < 0.3$ s) though the centre and range of the two logic trees converge for $T \geq 1.0$ s.

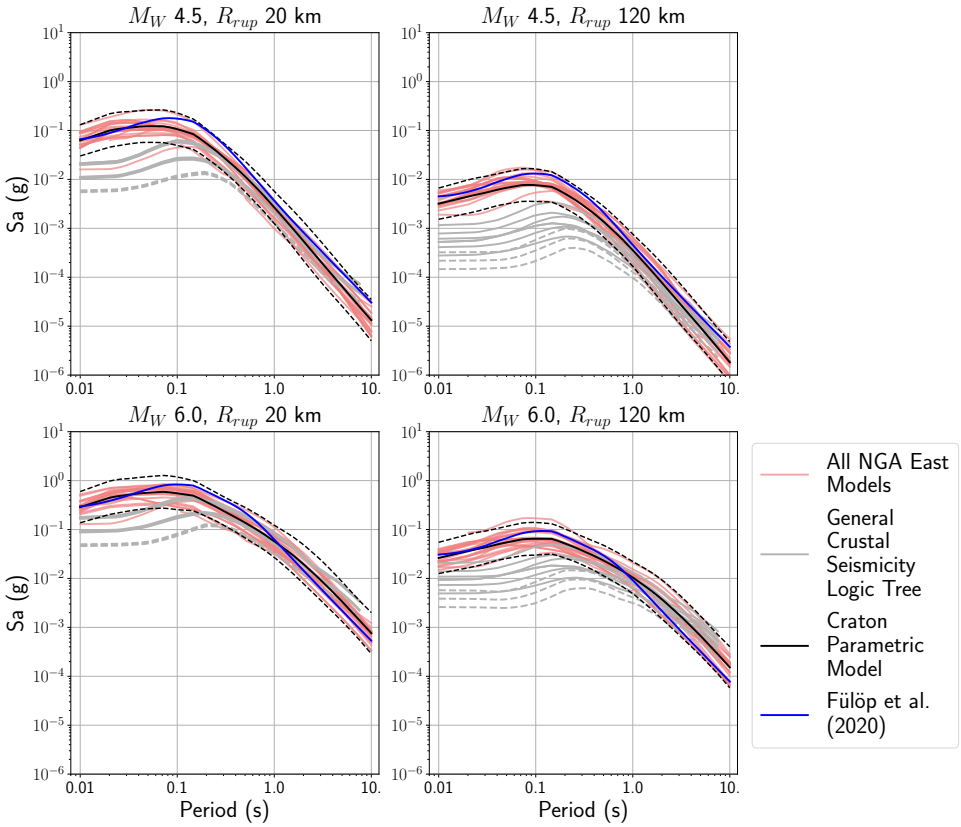



**Fig. 4.9** Comparison of the GMM logic tree of the craton GMM model (N = 3 branches) (black) with the NGA East GMM suite (coral), the *default* crustal seismicity logic tree (grey) and the Fülöp et al. (2020) GMM (blue) for a reference rock site of $V_{S30}$ 800 m/s.

### 4.3.3 Site Amplification

As the parametric form of the model is fit to the NGA East models, the ground motion values are initially defined for a very hard rock ($V_S$ 3000 m/s) condition. To determine the ground motion on the reference $V_{S30}$ 800 m/s rock condition, and to other site conditions, the CEUS site amplification models of Stewart et al. (2020) and Hashash et al. (2020) are integrated into the model. These amplification models have three components: a very hard rock to reference rock amplification factor ($F_{760}(T)$), a linear amplification factor ($f_{lin}(V_{S30}, T)$) and a nonlinear amplification factor ($f_{nl}(V_{S30}, PGA_{760}, T)$, each of which has an epistemic uncertainty that when combined forms an additional distribution that we define here as $\sigma_{\mu,S}$. One critical difference between the present model and similar GMMs that have been applied previously in Europe is that for sites on very stiff soil or rock the $F_{760}$ and $f_{lin}$ factors assume a seismic velocity profile of the site with a strong impedance contrast as depth, typical of what might be expected in regions of low tectonic deformation and glacially formed geomorphology. This impedance contrast results in a stronger amplification for high frequency motion than that of previous models that have assumed a more gradational velocity profile. The new GMM therefore predicts higher short-period motion than the previous ESHM13 selection; however, we believe this is appropriate to the region in question.

### 4.3.4 Aleatory variability

In keeping with the aleatory variability models assumed by both the NGA East ground motion project *and* the default shallow crustal seismicity logic tree, we adopt the same heteroskedastic aleatory uncertainty model from Al Atik (2015). In this case we use the Al Atik (2015) "global" models (fit to the NGA West 2 database) with magnitude-dependent $\tau$ and magnitude-dependent $\phi_0$. Site-to-site

variability ($\phi_{S2S}$) is taken from the study of Stewart et al. (2019), which is used for the ergodic seismic hazard assessment in the 2018 United Stated National Seismic Hazard Model. No epistemic uncertainty on the aleatory variability term is considered in the present case.

### 4.3.5 Complete Logic Tree for GMM in the Stable Cratonic Region of NE Europe

To define the complete GMM logic tree for this region (shown in Figure 6.12) we return to our initial premise that the scaling and attenuation of ground motions in NE Europe are fundamentally different from those of southern and western Europe. Without strong motion data within the ESM to verify this assumption, we retain in the logic tree both the possibility that the ground motion characteristics are fundamentally different, and the possibility they are not. The logic tree if first split into two branches to accommodate both hypothesises. Initially we consider the possibility that some branches of the default crustal seismicity logic tree may be appropriate in NE Europe. This is assigned a weight of 0.2. We do not, however, consider the low stress region branches or the fast attenuation branches to be feasible and thus eliminate them from the tree. The branches are re-weighted to assign $W_{11} = W_{111} = W_{121} = 0.667$ weight to the high stress ($\varepsilon_{11} = 1.732$) and to the slow attenuation branches with ($\varepsilon_{111} = \varepsilon_{121} = 1.732$), and just $W_{12} = 0.333$ weight to each of the "average" stress region and attenuation region branches ($W_{112} = W_{122} = 0.333$) respectively ($\varepsilon_{12} = \varepsilon_{112} = \varepsilon_{122} = 0$).

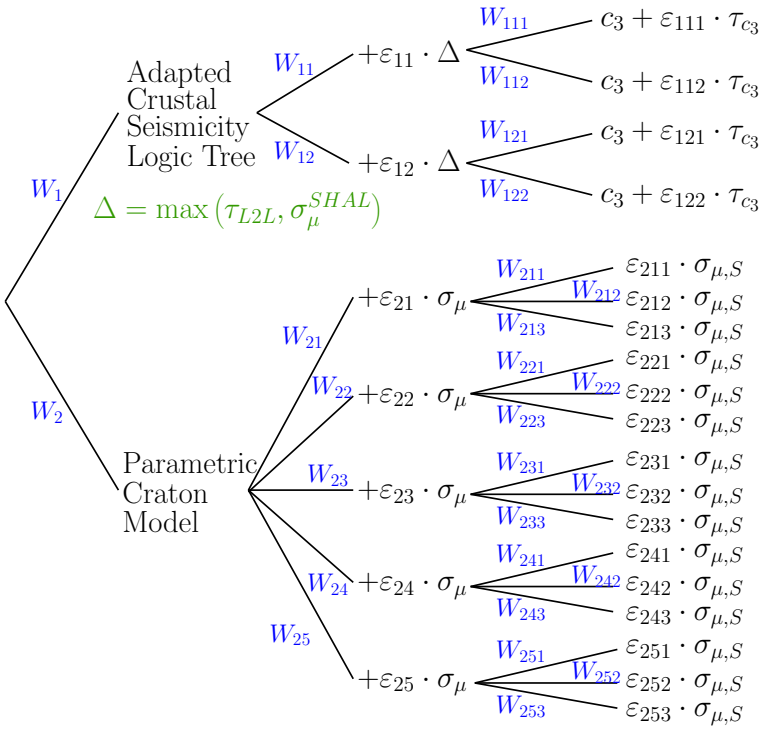



**Fig. 4.10** Complete GMM logic tree for the stable cratonic region of northeastern Europe

On the other side of the logic tree, we have the assumption that the ground motion characteristics of NE Europe *are* fundamentally different from those represented by any branch of the default crustal seismicity GMM logic tree, and therefore apply the parameteric craton GMM and its epistemic uncertainty. Two branching levels are considered to reflect both the epistemic uncertainty in the median ground motion on very hard rock ($\sigma_\mu$) and the epistemic uncertainty in the site amplification factor ($\sigma_{\mu,S}$). As the former is the dominant epistemic uncertainty, we apply a 5-branch discrete approximation to the Gaussian distribution, while for the latter we apply a 3-branch discrete approximation (see Table 6.1 for the $\varepsilon$ values and weights).

## 4.4 Subduction and Deep Seismicity

The ESM database is rich in records from shallow crustal seismicity, and this has motivated the development of the regionalized scaled backbone ground motion model approach thus far. Some of most active seismic regions in Europe, however, and indeed some of the areas of highest seismic hazard come from seismicity that does not take place in the shallow crust but rather from Europe's subduction zones (Hellenic, Calabrian, Cypriot and Gibraltar) and from the Vrancea deep seismic zone in Romania. The ground motion scaling and attenuation characteristics from earthquakes in these regions can differ significantly from those of shallow crustal seismicity, prompting the use of ground motion models specifically calibrated for subduction earthquakes. Fortunately, the ESM database contains hundreds of records from earthquakes originating in the Hellenic and Calabrian arcs, and from the Vrancea deep seismic zone (DSZ).

### 4.4.1 Identifying "non-shallow crustal earthquakes"

In its initial form the ESM database does not specifically indicate the tectonic region of the earthquake, and therefore events are not classed as shallow, subduction, etc. The first step is to sort the records from subduction regions and the Vrancea DSZ out from the other shallow crustal events. In the case of the Vrancea DSZ this is done manually, first by identifying the subset of earthquakes in the Vrancea region whose hypocentral depths are between 60 and 250 km, then verifying these by inspection to check for misclassifications or geographical outliers from the observed deep seismicity. This process yields 938 records from 64 deep source events in the Vrancea region.

For the subduction regions the identification process is more complex as the earthquakes cannot be so easily discriminated by the simple depth criteria. Instead, the southern Hellenic region and Calabria experience earthquakes in the upper 50 km that may originate from shallow seismicity in the over-riding crust, subduction interface seismicity or even seismicity within the subducting slab or the outer rise.

Given the observed uncertainty in both the depth estimates of the earthquakes and the location of the crust and subducting slab, a fuzzy classification system was designed to classify earthquakes in terms of degree of membership of each tectonic region type given the event's depth, proximity to the subduction surface and, if available, the focal mechanism.

A comprehensive description of the fuzzy classification algorithm is given in the accompanying report of Weatherill et al. (2021) and has also been applied to similar problems of earthquake classification in subduction regions by Ghasemi et al. (2020). Setting a *degree of membership* threshold of 0.6 (where 0 is unambiguously not a classified according to a given tectonic region type and 1.0 is unambiguously classified as so) we apply the fuzzy classification system to categorise events as one of: active shallow crust, subduction interface, subduction in-slab, upper mantle, outer rise and non-subduction (i.e., deep events not seemingly associated with the subduction zone).

From this classification we identify 681 records from 164 subduction interface earthquakes and 704 records 121 from subduction in-slab events. The distribution of records for the three "non-shallow" seismicity types (subduction interface, subduction in-slab and Vrancea DSZ) are shown in **Fig. 4.11**.

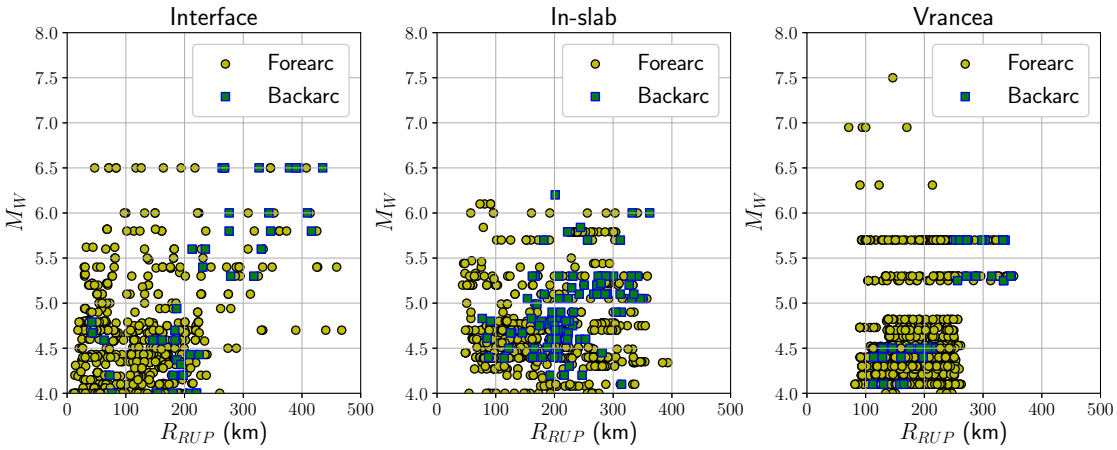



**Fig. 4.11** Magnitude and distance composition of the ground motion databases for the subduction interface, subduction in-slab and Vrancea DSZ regions of Europe.

### 4.4.2 Selecting and modifying a backbone deep/subduction seismicity GMM

In contrast to both the case of shallow seismicity (for which a substantial amount of data was available in ESM) and the cratonic region (for which no data were available), the subduction/deep records provide us with enough data to make some inferences about the ground motion characteristics of the region, but not enough to fit a full GMM. In particular, Figure 6.13 shows that while the number of records is quite high overall, the magnitude range is limited to M ≤ 6.0 to 6.5 for the subduction events while even the Vrancea database has only a limited number of records for M ≥ 6.0 earthquakes.

As larger magnitudes in the range 6.5 ≤ $M_W$ ≤ 8.5 are the most relevant for seismic hazard we instead aim to construct a backbone model by first using the available data to identify the most appropriate candidate GMM from the suite of recent global subduction models and then attempting to calibrate this model to the attenuation characteristics of the southern Mediterranean and Vrancea regions.

An initial pre-selection process is applied considering the required range of spectral periods and earthquake scenarios for the ESHM20. St the time of undertaking this analysis the Next Generation Attenuation Subduction (NGA-Sub) ground motion models (Parker et al. 2020; Kuehn et al. 2020; Gülerçe & Abrahamson, 2020) were not available. Though the NGA-Sub models do introduce some calibrations for different subduction zones around the globe, no data from the eastern Mediterranean were included and thus no calibration is currently available.

The pre-selected models are compared against the observed data in each of the three region types quantitatively using multivariate log-likelihood scoring (Mak et al., 2017). From this pre-selection process we find that the Abrahamson et al. (2016) “BC Hydro” model is the most suitable candidate for the subduction regions, meeting the project requirements in terms of scenario and period range while also giving low log-likelihood scores across the full range of spectral periods in each of the three regions. For the Vrancea DSZ the in-slab version of the Abrahamson et al. (2016) model is also found to be the best suited.

With the candidate model identified, we then use the available strong motion data to explore the trends in the between-event and within-event residuals with respect to this model. A lack of data from large magnitude events, combined with limited a limited number of sites with measured $V_{S30}$ terms, means that we do not attempt to re-calibrate either magnitude or linear site scaling terms of the GMM.

However, with the large volume of data covering the 60 to 200 km distance range we use the observed trends in within-event residuals to calibrate the anelastic attenuation term of the model to better fit the local data.

We do this by first using the site amplification component of the GMM to scale the ground motions to reference rock. Then within-event residuals are calculated with respect to an adaptation of the model in which the anelastic term ($f_{R,A}(R_{RUP}) = \theta_6 \cdot R_{RUP}$) is set to 0. The locally calibrated $\theta_6$ is retrieved for each period from the slope of a linear fit to the within-event residuals.

The observed period-dependent trend in $\theta_6$ for both the interface and in-slab models is shown in **Fig. 4.12**, where it is compared against similar anelastic attenuation scaling coefficients from other relevant subduction GMMs as well as from regional estimates of this same parameter from Abrahamson et al. (2018), a preliminary update of the 2016 model using NGA Sub data.

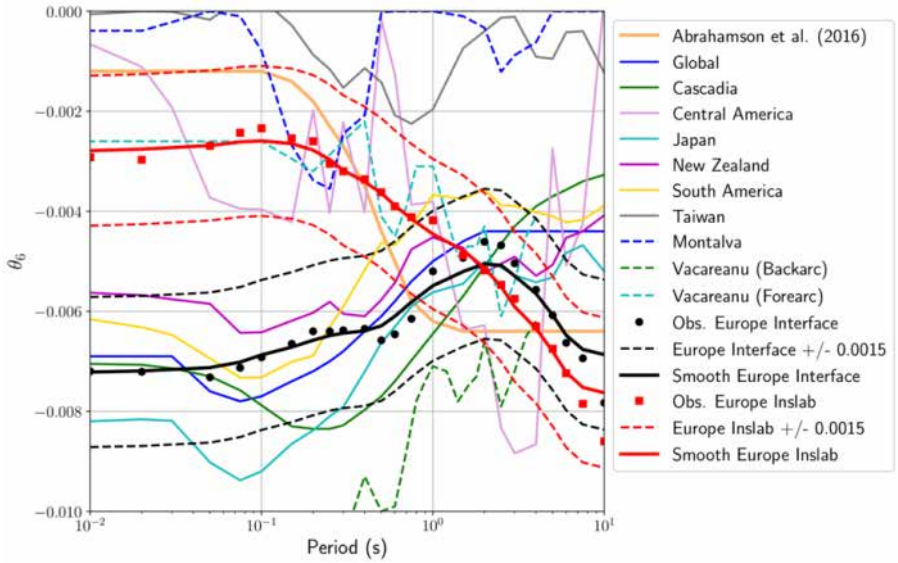



**Fig. 4.12** Variation in anelastic attenuation term ($\theta_6$) with period for the fitted cases (red and black dots for in-slab and interface events respectively) and selected subduction GMMs and local regionalisations, where "Global", "Cascadia", "Central America", "Japan", "New Zealand", "South America" and "Taiwan" correspond to the regions for which $\theta_6$ terms have been calibrated by Abrahamson et al. (2018)

### 4.4.3 Quantifying the epistemic uncertainty and constructing the logic tree

**Fig. 4.12**, demonstrates that there is a significant diverge in anelastic attenuation coefficients from different regions of the globe, particularly at short periods where they diverge more significantly between two groups. Recognising this variability, and considering the potential uncertainties in our present estimates of $\theta_6$ for the regions of interest, we define an average model (solid red and black lines in Figure 6.14) and an upper and lower bound (dashed red and black lines) based around the global variability in this term, which we define as a period-independent $\pm 0.0015$.

For consistency with the scaled backbone approaches adopted elsewhere, we consider these upper and lower bounds as notionally equivalent to the 95$^{th}$ and 5$^{th}$ percentile of a Gaussian distribution, and hence we define a three-branch distribution in the logic tree of $\theta_6' - 0.0015$, $\theta_6'$ and $\theta_6' + 0.0015$ with weights of 0.167, 0.666 and 0.167 accordingly for the subduction zones, and weights of 0.2, 0.6 and 0.2 for Vrancea where we believe slightly higher weights on the upper and lower branches may be appropriate.

The calibration of $\theta_6$ may account for epistemic uncertainty in the anelastic attenuation, but its influence is mostly apparent only at longer distances. This does not address epistemic uncertainties due to large magnitude scaling and/or regional differences in the stress parameter term of the model.

For this we take advantage of an epistemic uncertainty distribution for median ground motions ($\sigma_\mu$) proposed by Abrahamson et al. (2018), who looked at both the region-to-region variability in stress parameter ($\theta_1$). Their model presents a $\sigma_\mu$ term that is constant with period for interface models and increases slightly at shorter period for in-slab events.

As their analysis is based on the stress-parameter, this term does not account for region-to-region variability in attenuation, meaning that is can be applied as an additional set of branches on top of the anelastic attenuation uncertainty model with a minimal risk of double counting.

An example of the resulting ground motions for a three-branch approximation of the distribution of $\sigma_\mu$ is shown in Figure 6.15 for a M 7.5 event at 25 km depth (for the interface) and 60 km depth (in-slab) assuming a site condition of 800 m/s and a forearc path. The distribution of $\sigma_\mu$ is shown in **Fig. 4.13**.

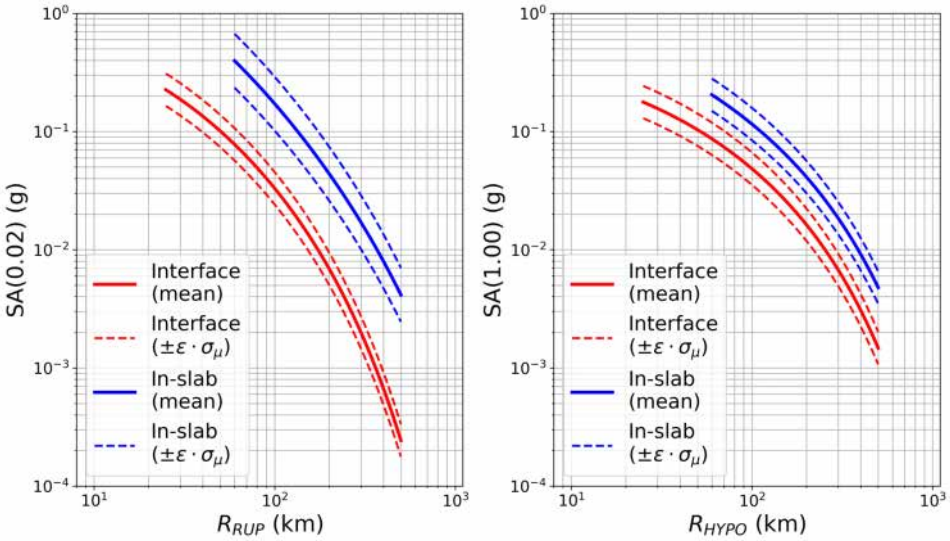



**Fig. 4.13** Ground motion attenuation for the ESHM20 subduction GMMs for a three-branch distribution of $\sigma_\mu$, where $\varepsilon$ corresponds to ± 1.732 (according to the 3-point distribution shown in Table 6.1). Mw = 7.5, $V_{S30}$ = 800 m/s and hypocentral depth is 60 km for the in-slabe model

Combining the epistemic uncertainties on the stress-parameter (characterized here by $\sigma_\mu$) and the attenuation, we define a complete ground motion model logic tree for application to the European subduction interface and in-slab regions and Vrancea DSZ. The Abrahamson et al. (2016) in-slab model is used as the backbone for both the in-slab and Vrancea DSZ regions.

As the epistemic uncertainty with the largest influence on hazard, we map $\sigma_\mu$ into a 5-branch discrete approximation of the Gaussian distribution according to Miller & Rice (1983). The scaling factors and corresponding weights for the anelastic attenuation adjustments are stated previously. The complete logic trees for each of the three environments (subduction interface, subduction in-slab and Vrancea DSZ) are shown in **Fig. 4.14**.

### 4.4.4 Modifying the forearc/backarc scaling term

The Abrahamson et al. (2016) GMM incorporates a forearc/backarc scaling term, one in which attenuation with distance is faster for sites in the backarc of the subduction zone.

Trends in within-event residuals were explored to identify if this forearc/backarc scaling is present in the observed data, and although data in the backarc are sparse there was sufficient evidence to suggest such a distinction is present. However, with so few data in the backarc we could not re-calibrate the coefficients of the scaling term within the GMM.

To define the boundary between the forearc and backarc for the Hellenic and Calabrian subduction zones we first consider the Slab 2.0 representation of the zone geometry. Then using the locations of predominant volcanic activity as the defining boundary we draw the forearc/backarc front based on the relative positions of the volcanic front with respect to the interface.

For Vrancea we follow the example of Vacareanu et al. (2015) in defining the forearc/backarc margin along the watershed of the Carpathian mountain belt. For the Gibraltar arc we cannot identify a clear front, and in the present case all sites are assumed to represent forearc sites. The lines of the fronts can be seen in **Fig. 4.15**.

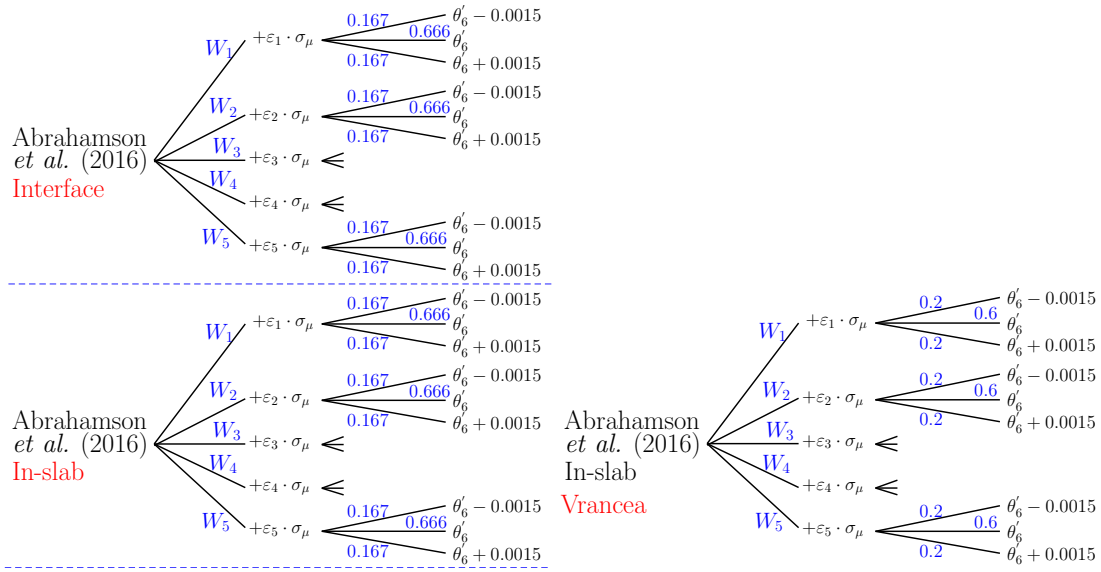



**Fig. 4.14** Complete scaled backbone logic tree formulations for the subduction interface, in-slab and Vrancea DSZ regions.

Though the analysis of the ground motion residuals does support the use of a forearc/backarc scaling term, in the Abrahamson et al. (2016) GMM this is treated as a binary term, $F_{FABA}$, taking 0 or 1 depending on whether the site is located in the forearc or backarc respectively. As the impact on the attenuation is significant, with lower ground motions in the backarc, this can result in a "cliff-edge" drop in seismic hazard across the front.

This was particularly pronounced in Romania where the Vrancea DSZ is the dominant seismic source over several hundreds of kilometres. To modulate this effect, we modify the forearc/backarc definition to describe $F_{FABA}$ as a continuous variable in the range $0 \leq F_{FABA} \leq 1$, dependent on the distance to the forearc/backarc front, thus effecting a smoother transition from one attenuation regime to the other across the front. To accommodate this a new site term is introduced, which defines the distance of the

site with respect to the forearc/backarc front (or volcanic front), $x_{vf}$, and this parameter is calculated for all three major subduction and deep seismicity regions. $F_{FABA}$ is related to $x_{vf}$ through a simple function that describes the smooth transition from "pure" forearc to "pure" backarc attenuation over a set distance range. Several functional forms were tested but the preferred relation is an S-function such that:

$$F_{FABA}(x_{vf}, a, b) = -\begin{cases} 0 & \text{for } x_{vf} \leq a \\ 2 \cdot \left(\frac{x_{vf}-a}{b-a}\right)^2 & \text{for } a < x_{vf} \leq 0.5 \cdot (a+b) \\ 1 - 2 \cdot \left(\frac{x_{vf}-a}{b-a}\right)^2 & \text{for } 0.5 \cdot (a+b) < x_{vf} < b \\ 1 & \text{for } x_{vf} \geq b \end{cases} \quad \text{(Eq. 6.8)}$$

where $a$ and $b$ are -150 km and 150 km respectively. The spatial patten of $x_{vf}$ is shown in Figure 6.17, along with an illustration of the influence of this term on the attenuation of ground motion from a scenario $M_W$ 7.7 earthquake (with the volcanic front placed at 250 km from the source).

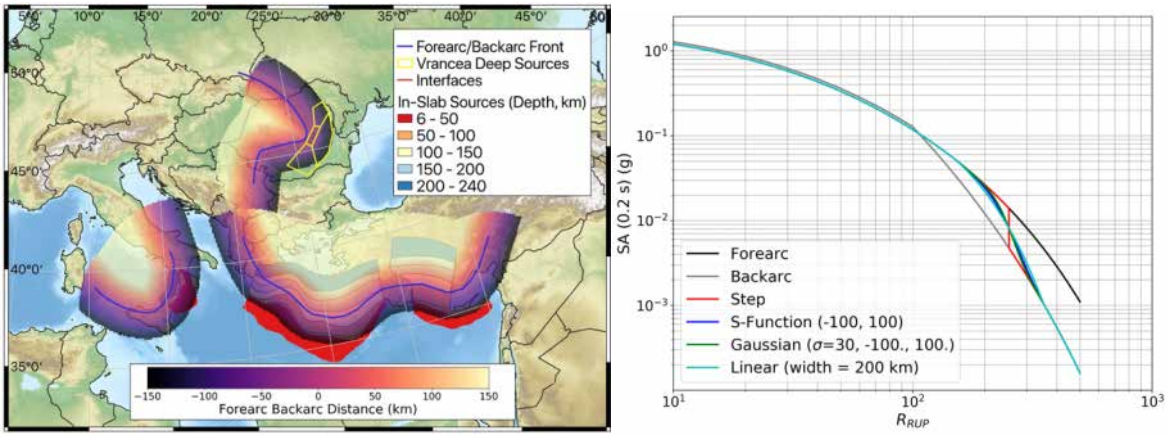



**Fig. 4.15** Location of forearc/backarc front and $x_{vf}$ distance in eastern and southern Europe (left), and impact of smoothed functions for $F_{FABA}$ on attenuation for a scenario event of $M_W$ 7.7 on rock with the volcanic front 250 km from the source

A small region of conflicting overlap can be seen at the North Macedonia/Bulgaria border region, which falls into the backarc region with respect to the Hellenic arc but also the forearc region with respect to the Vrancea DSZ. In this narrow region the sites are classified as forearc in order to ensure that the contribution to the seismic hazard from Vrancea is dominant over the very distant Hellenic sources.

No epistemic uncertainty on the functional form and distance range of the $F_{FABA}$ is modelled here, though this is accommodated in the model implementation and thus may be considered in future projects.

## 4.5 Special cases and adjustments

The three main tectonic region types that have been considered so far, i.e., shallow crustal seismicity, stable craton and subduction/deep seismicity, cover most of the seismic sources in Europe.

With the adoption of the regionalization features in the shallow crustal seismicity model GMM we are also able to capture a further level of detail than has been assumed in the previous classifications adopted by ESHM13 and some existing national models.

As this approach has been largely data driven, however, the calibrations themselves can be readily quantified and adopted in regions with sufficient data *within the available ESM dataset*. There are areas, however, where ground motion data are available in a limited capacity, but were not included within the ESM data set and therefore were not part of the main regression process.

As the ground motion models and its epistemic uncertainties *should* represent the best state-of-knowledge at the time of construction we consider two *special cases* of shallow seismicity where we adapt the logic tree on the not on the basis of the data used for the regression but rather from other data, information and interpretations as well as feedback from third parties.

### 4.5.1 North-western Europe

The general strategy adopted within the shallow default crustal logic tree is to assign regions the full region-to-region variability in stress region scaling and residual attenuation where there are insufficient data to constrain a local calibration. This would be the case much of France (excluding the Pyrenees, Alpine and Rhine regions), Norway, the United Kingdom, Spain and Portugal. Subsequent analysis by Kotha (2020) explored in detail a database of strong and weak motion records from the French seismological and Geodetic Network (RESIF), which contains a large number of stations and events that were not available within the ESM database. In the majority of cases these events from RESIF were smaller magnitude ($2.5 \leq M \leq 4.5$) than those found in ESM. For these reasons, and accounting for differences in processing that may introduce potential biases between the two data sets, we do not include the RESIF data into the initial regressions.

The analysis by Kotha (2020) was able to define $\delta c_{3,k}$ terms for additional regions in France and parts of Spain, while also providing a general comparison point for overlapping regions in southeastern France and Italy. The regional variation in $\delta c_{3,k}$ for a period close to Sa (0.1 s) is shown in **Fig. 4.16**.

From this analysis the general spatial trend of slower residual attenuation to the north and west of the Alps contrasted against progressively faster attenuation in Italy is persistent. What the analysis on the RESIF data does suggest, however, is that the slow residual attenuation seen in the Pyrenees is potentially persistent throughout much of France west of the Alps.

However, it is important to note that the distribution of stations for in the RESIF data set for this region is skewed toward southwestern France and the Pyrenees, and to the northwest in Brittany and Pays de la Loire.

For central and northern France few stations exist in the RESIF database, which leads us to conclude that we would retain the default general crustal logic tree here, conceding that this region would most likely be a transition zone between the faster residual attenuation in the southeast and the particularly slow residual attenuation in the north and southwest.

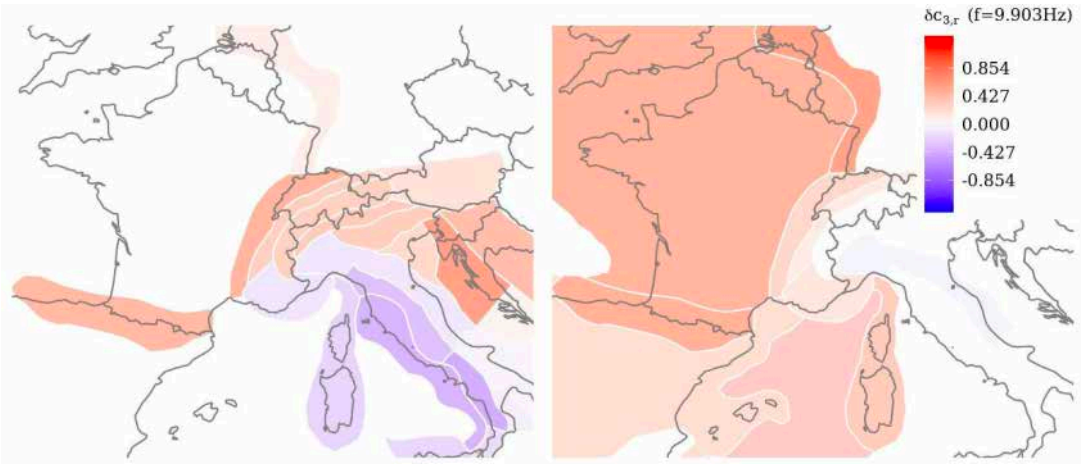



**Fig. 4.16** Variation in $\delta c_{3,k}$ at Sa (≈ 0.1 s) for France according to the ESM dataset (left) and the RESIF data set (right). Figure taken from Kotha (2020)

The regionalization adopted by Kotha (2020) for France groups most of France and the UK into a single zone, likewise for Iberia. Recognising that a more detailed insight is needed, we explore possible data sets that may explain the slower attenuation in these regions. Note that $\delta c_{3,k}$ is a *residual* attenuation, meaning that it is a statistical property of the data and cannot disaggregate between different factors that may contribute to faster or slower attenuation. In that sense, both crustal attenuation (Q) and depth to the Moho are properties that would be expected to have an effect.

Taking this into account we consider the distribution of the bedrock geology in Europe, shown in **Fig. 4.17**. From this we can identify blocks of older Phanerozoic bedrock that span southern Portugal, northern Spain, the Pyrenees, northwestern France, western UK and Norway. In these environments we have older, stable rock commonly associated with slower attenuation, which would be consistent with the observations of $\delta c_{3,k}$ that we have for the Pyrenees and western France.
We therefore delineate a set of regions following the most extensive blocks of older bedrock, within which we do not use the general default backbone model, but rather the $\delta c_{3,cl5}$ distribution that characterizes regions of slow attenuation. These are shown in the blue shaded regions in Figure 6.19.

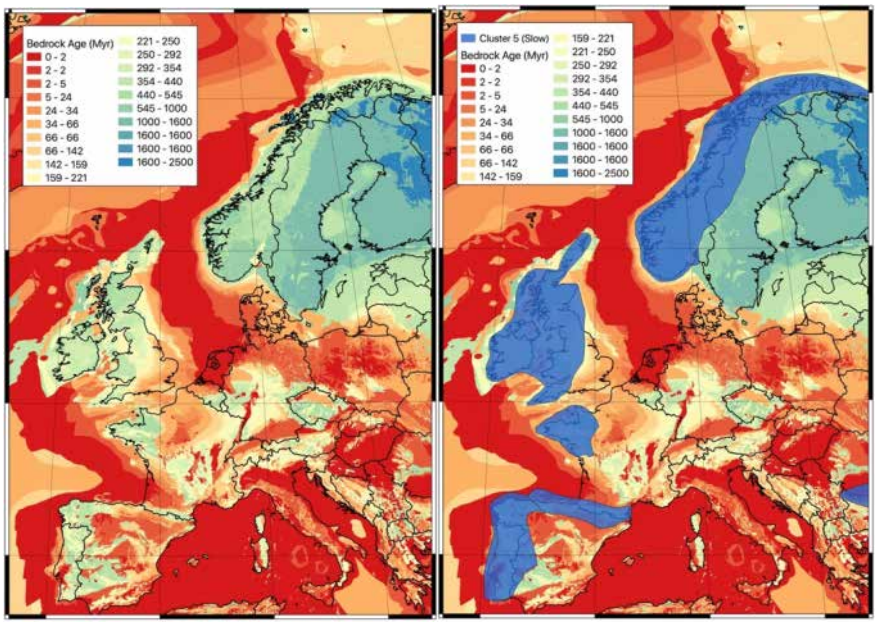



**Fig. 4.17** Map of bedrock geological age in western Europe (left) with manually assigned "cluster 5" (slow attenuation) regions highlighted in blue (right)

### 4.5.2 Iceland

The second special case area for which we have data that allow us to refine the calibration of the logic tree is Iceland. No data from Iceland was available within the ESM dataset; hence it was not included within the regionalized GMM of Kotha et al. (2020).

However, observations from Iceland were available in the previous RESORCE strong motion database resulting in 116 records from 18 events with a reported $M_W$ value. These records and events are mostly from the southwest seismic zone of Iceland and span a magnitude range from 4.5 ≤ $M_W$ ≤6.5 and a distance range 3 ≤ $R_{JB}$ ≤ 100 km (**Fig. 4.18**)

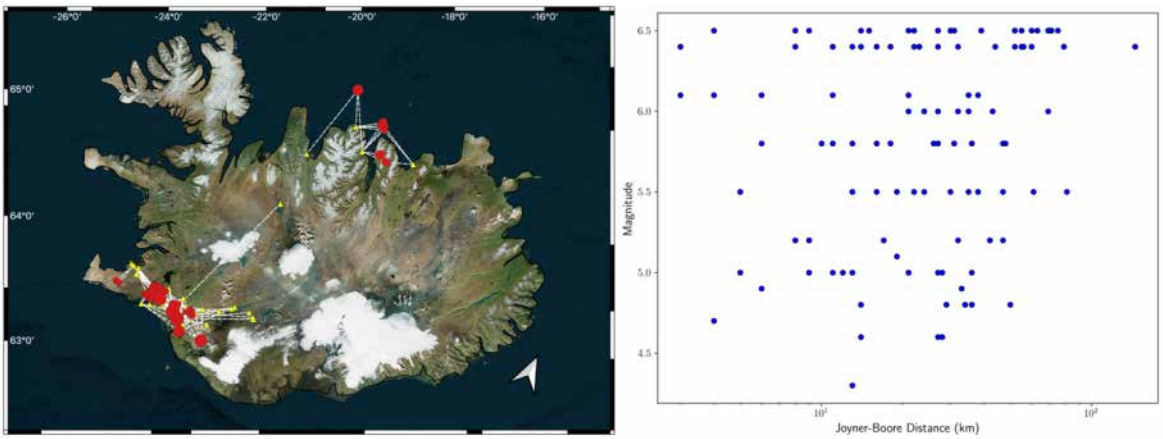



**Fig. 4.18** (left) Location of earthquakes (red dots) and stations (yellow triangles) for the set of Icelandic strong motion records from RESORCE. (right) Magnitude and distance distribution of the Icelandic records.

From this data set of Icelandic records, we retrieve the between ($\delta B_{e,i}$) and within-event residuals ($\delta W_{ij}$) with respect to the unregionalised Kotha et al. (2020) GMM, and then to the adjustments for each of the five cluster regions.

As most of the ground motion data are taken from distances shorter than 80 – 100 km the influence of the residual attenuation cluster is minimal; however, the distribution of $\delta W_{ij}$ did marginally favour cluster 3 (fast attenuation), which is expected given that this particular cluster is defined by records from western Italy and central Greece, where volcanism is having an influence on the attenuation. The attenuation region cluster applied to the active regions of central and southern Iceland is therefore cluster 3.

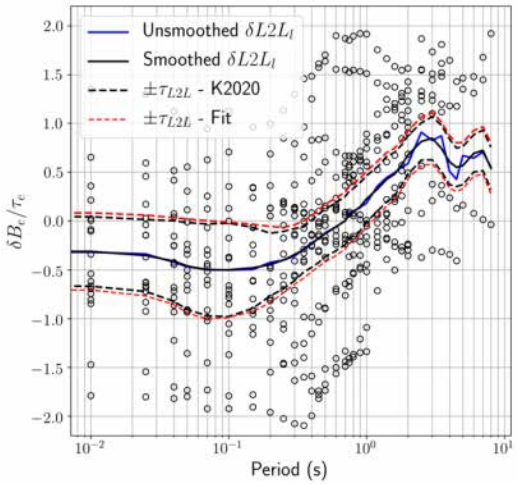



**Fig. 4.19** Distribution of between-event residuals and fit model of $\delta L2L_l$ and $\tau_{L2L}$ for Iceland

Greater differences between the observations and the model can be found in the between-event residuals. Here we find predominantly negative residuals at short periods, transitioning to moderately positive residuals at long periods.

The residuals alone do not unambiguously identify the cause of this trend, though we speculate that lower stress drops from shallower hypocentral depths in a predominantly volcanic environment may be a significant factor. Taking the mean of the normalized residuals for each period we can define a term equivalent to $\delta L2L_l(T)$, which we then smooth across the period range.

The observed variance of this parameter is found to be very close to $\tau_{L2L}^2$ from the original Kotha et al. (2020) GMM. Given these observations we define a modified logic tree for Iceland in which the median ground motions from the default shallow seismicity backbone model are adjusted according to the $\delta L2L_l$ model shown in **Fig. 4.19**, but the original $\pm\varepsilon \cdot \max(\tau_{L2L}, \sigma_\mu)$ factors are retained.

The resulting logic tree for Iceland is also shown in Figure 6.18 and this is applied to the active sources following the mid-Atlantic rift in southern and central Iceland. For the lower seismicity sources outside of this region the default shallow crustal GMM logic tree is retained.

# 5 Model Implementation and Hazard Calculation

Computation of the seismic hazard for a large-scale region, covering the entire Euro-Mediterranean region as illustrated in **Fig. 5.3** is not trivial. The main challenges are linked to the complexity of the seismogenic source model, the logic tree, the size of the region, as well as the output types: seismic hazard maps, hazard curves, uniform hazard spectra and disaggregation.

OpenQuake-engine v13.3 is used for computing the ground shaking estimates hazard of the ESHM20. The computational engine is open-source, community driven and it is developed under a test-driven framework. For computation of the ESHM20, we adopted the classical Probabilistic Seismic Hazard Analysis (PSHA) calculator that estimates the probabilities that various earthquake-induced ground motion levels will be exceeded at a given location in a given time period by solving the PSHA integration procedure (Pagani et al 2014). Note, that for the risk calculation the stochastic event set calculator is used (Crowley et al 2021).

## 5.1 Seismogenic Source Model – implementation

First and foremost, the unified earthquake catalogue's format is standardized according to the template of the OpenQuake hazard modeler toolkit (https://github.com/gem/oq-engine, Weatherill et al 2014) with following fields: *eventID, longitude, latitude, year, month, day, hour, minute, second, magnitude, depth, source_catalog, source_eventID.*

The earthquake catalogue is declustered with the Grünthal time window algorithm, and further separated in two sub-catalogues for shallow crust seismicity and deep seismicity. An additional filter is applied to the shallow earthquakes to remove the earthquakes with Mw > 6.5 on top of the subduction interface, to avoid the double counting. Similarly, the volcanic seismicity is removed from the sub-catalogue of shallow earthquakes. The activity parameters are estimated for the area sources: shallow, deep and volcanic using the corresponding earthquake sub-catalogues and the completeness time-magnitude intervals described in Chapter 3. The procedure of estimating the activity rates of the area sources is implemented in Python, as an extension of the hazard modeler toolkit (https://github.com/gem/oq-engine, Weatherill et al 2012). The methodology is outlined in Bollinger et al (1989, 1993), Weichert (1980), Berril and Davis (1980) and Herrmann (1977) and the main assumptions are described in Chapter 4. Similarly, an extension to the hazard modeler toolkit was developed to compute the activity rates for the active faults. For developing the smoothed seismicity, a MATLAB toolbox was used.

Next, we used the seismic source definitions of OQ-hazard engine User's Manual v3.13 (Pagani et al 2014, Pagani et al 2021) as the blueprints to define and parametrize the individual seismogenic sources.The minimum magnitude ($M_{min}$) is 4.5 for all models but active faults. $M_{min}$ for active faults represents the threshold magnitude ($M_{thr}$) that separates the contribution of each fault from the background seismicity. The upper and lower seismogenic depths control the extension of the ruptures within the Earth's crust. In this implementation, the mean values (i.e. MINDEPTH and MAXDEPTH attributes) are used for all source typologies, without considering their uncertainty due to its impact on the seismic hazard estimates. Sensitivity analyses conducted during the development and implementation phase of the model indicate a minor effect of the upper and lower seismogenic depth to the return periods of interest (i.e. 475yrs and 2500yrs). A summary of the seismogenic source models' attributes are provided in Appendix B: ESHM20 Input Datasets – Summary and Files Format.

### 5.1.1 Area Source Zones

The template of each area source includes its geometry attributes and the magnitude scaling relationship. This relationship controls the size of floating ruptures as a function of magnitude. Sensitivity analysis conducted on the impact of extensive ruptures vs the point ruptures to the hazard results, found that with the current ground motion model the impact of the extensive rupture was less than 5% for the return periods of interest (RP475yrs to 2500yrs). Hence, to speed up the calculation, we used the point ruptures.

The geometry includes the definition of polygon location, style-of-faulting, and depth. Style-of-faulting is defined by the predominant strike, rake and dip angle values. We used a probability mass function to characterize the aleatory uncertainty, which comprises a set of strike, dip and rake angles, each with their associated weights. This implementation is valid for each source of the area source model and the point sources of the smoothed seismicity model. The predominant style of faulting is defined for each source as it was found that it has a minor effect on the hazard estimates. Furthermore, since extensive ruptures are modelled as point ruptures, we collapsed the depth distribution to a weighted value to optimize the computational time.

The magnitude frequency distribution (MFD) for individual area sources is described by two recurrence models as given in the Figure 4.3 section 5: a) double truncated GR exponential distribution and b) Pareto tapered distribution (Kagan 2002). The double truncated GR exponential distribution (the default recurrence distribution of OpenQuake Manual, section 2.2, Pagani et al 2021) and defined by a lower and upper magnitude bounds and the two GR parameters: aGR and bGR.

Similarly, the Pareto truncated distribution is expressed by the aGR and bGR parameters, the minimum and maximum magnitudes. This distribution is not supported by the current version of OpenQuake, and it is implemented as a discrete incremental magnitude-frequency distribution.

The special case of the two GR double truncated distribution applies to two area sources in Italy i.e. ITAS308, ITAS331 and two area sources off-shore Portugal i.e PTAS390 and ZZAS506. This special MFD is defined by the $M_{min}$, aGR, two bGR values and $M_{max}$ and implemented also as a discrete incremental magnitude-frequency distribution. The incremental distribution is defined by incremental annual rates reported in the center of the magnitude bin. The magnitude bin size is 0.1 units of magnitude.

The area source of deep seismicity and subduction in-slab are implemented in the same way as the shallow crust sources.

### 5.1.2 Active Faults and Subduction Interface

Following OpenQuake User's Manual (v3.12, Pagani et al. 2021), crustal faults are modelled as simple faults, and the subduction interface zones are represented as complex faults. A simple fault describes a fault surface projected along strike and dip. A complex fault does not require a dip angle because the geometry can be described by combinations of fault edges to describe top, mid or bottom of a fault surface. This model used the default attribute values as provided by the compilers as best estimates of the multiple parameters and assumptions. Thus, in this implementation no additional uncertainties are applied to the parameters describing the fault's geometry, style of faulting, upper and lower seismogenic depth and similarly, for the subduction interface implementation.

The uncertainties of the slip-rates and maximum magnitudes are characterized as epistemic uncertainties and are used to characterize the seismic productivity of each fault. We use a doubly-truncated GR exponential distribution to describe the MFD of each fault, as well as of the subduction

interface. This MFD is defined with lower and upper magnitude bounds, and the two GR parameters: aGR and bGR.

The lower magnitude corresponds to $M_{thr}$ (a magnitude that separates the background seismicity from the fault seismic productivity), the upper bounds are described by three $M_{max}$ values. The bGR- value is retained from the TECTO zones and the aGR values are directly linked to the slip-rates values (lower, average and high). The resulting 9 end-branches describing the epistemic uncertainties of the correlated aGR and $M_{max}$ are illustrated in **Fig. 5.2**.

The MFD of the subduction interface is a doubly-truncated GR exponential distribution defined by a lower and upper magnitude, and aGR, bGR parameters estimated out of a complex logic tree with more than 2800 logic tree end-branches. To facilitate the hazard calculation, only the central model of this distribution was used. Note that the subduction interface sources are used as a proxy for generating earthquakes of magnitude Mw $\geq$ 6.5, thus the $M_{min}$= 6.5.

The smoothed seismicity is represented as point sources of a grid spanning across the entire region, which are grouped according to the corresponding area source in which they are situated. All point sources within an area source share source attributes such as style of faulting, depth, upper and lower seismogenic source. At each grid location, the magnitude frequency distribution is described as double truncated GR exponential distribution, and implemented as an evenly discretized incremental distribution. The incremental annual rates are reported in the center of each magnitude bin and magnitude bin size is 0.2 units of magnitude to speed up the calculation. In the proximity of the faults, the length of the incremental distribution is controlled by the lower magnitude bound $M_{min}$ (i.e. 4.5) and the upper magnitude bound $M_{thr}$ inherited from the overlaying fault's buffer. This avoids double counting the seismicity rate.

The smoothed seismicity model compliments the active faults logic tree branch and is provided without additional epistemic uncertainty.

### 5.1.3 Logic Tree Seismogenic Sources and OpenQuake Input Files

An important aspect of the seismogenic source model implementation for hazard calculation is the consideration of correlated uncertainties of the activity parameters i.e. aGR,bGR and Mmax. Correlation of the uncertainties would mean applying alternative branches of aGR, bGR and Mmax (e.g. low, middle, high) across all sources in the same end branch. For a single source model with three aGR,bGR and three Mmax values, this would result in 9 logic tree branches; however, the upper and lower branches would correspond to extreme cases (i.e. all sources producing the upper/lower Mmax and/or high/low activity rates simultaneously) yielding exceptionally high or low activity rates with non-trivial weights. Uncorrelated uncertainties, in contrast, require permutation of all of the combinations of aGR,bGR and Mmax for all of the sources.

Here the distribution of the total activity rates across all of the sources on each branch will reflect the sum of underlying distributions more accurately; however, the total number of end branches in this case would correspond to 9NSRCS where NSRCS is the number of sources in the source model. For a regional model on the scale of Europe, which may contain hundreds of sources, uncorrelated uncertainties yield intractably large logic trees. In the current version of OpenQuake, it is not computationally feasible to compute a logic tree with uncorrelated branches, thus we opt for correlated branches. Thus, each source model is described by a default source model for which the activity parameters are assigned to the upper, mid and lower values.

Along the first branch of the logic tree (i.e. area sources model) the correlated aGR, bGR and $M_{max}$ values are used to generate the OpenQuake input files. The combination of the range of values for the activity parameters for area source activity parameters are given in **Fig. 5.1**. Noteworthy is the combination of the parameters for the two MFD models, resulting in nine end-branches for the exponential double truncated GR and the three end-branches for the Pareto distribution.

Along the second branch (i.e. active faults) of the logic tree, the correlation between aGR and Mmax is illustrated in **Fig. 5.1**. The xml input files are generated following the combination of the aGR from slip-rates values (lower, average and high) with the $M_{max}$ values.

In summary, the logic tree of the seismogenic source model is described by source model end-branches with correlated values of the activity parameters. Each end-branch is a complete model consisting of different source typologies: area sources for shallow crust, deep seismicity, subduction inslab and volcanic regimes; point sources for the smoothed seismicity model, simple faults for shallow crust faults and complex faults for subduction interface.

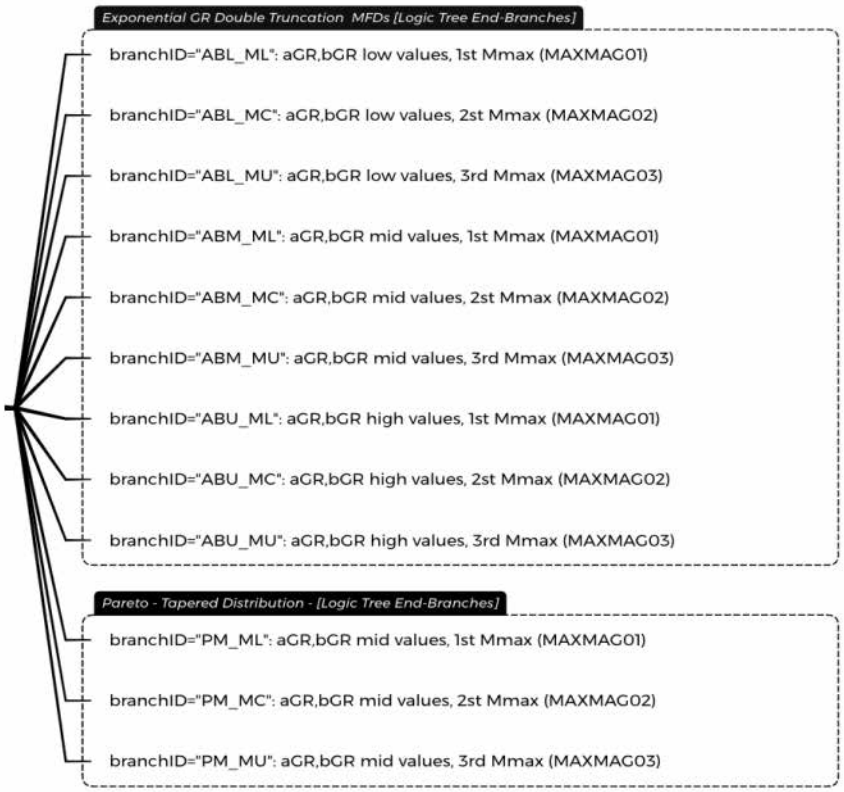



**Fig. 5.1** Illustration of the source model end-branches for the area source model. The logic tree considers correlated values for the activity parameters aGR, bGR and $M_{max}$.

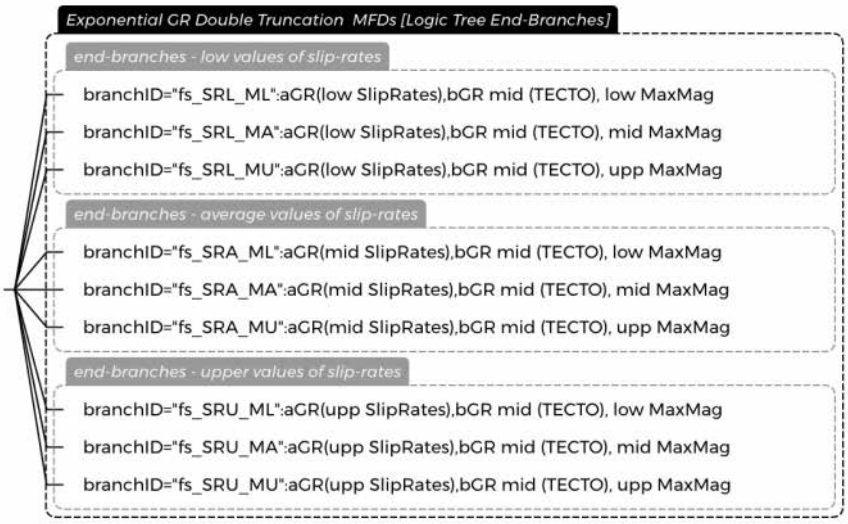



**Fig. 5.2** Illustration of the source model end-branches for the active fault sources model. The logic tree considers correlated values for the activity parameters aGR and $M_{max}$.

## 5.2 Ground Motion Model – implementation

### 5.2.1 Implementation of the Ground Motion Models

The OpenQuake implementation of the complete suite of ground motion model logic trees for the different tectonic regions (described fully in chapter 6) requires a change of approach in comparison to the ESHM13 and other models that adopt a "weights-on-multiple-models" approach to developing the GMM logic tree. The fundamental difference here is that in the present version we are considering a small number of core backbone GMMs to which we then apply a set of adjustments. The concept of a user adjustable GMM, one in which arguments can be passed by the user from the GMM logic tree configuration file that modify the operation of the GMM, has been supported in OpenQuake for several years. The new GMMs required for the ESHM20 have been implemented around this framework, with the adjustments needed to construct the full scaled backbone logic tree supported as configurable parameters. These will be illustrated in the following sections.

#### *5.2.1.1 GMM Logic Tree for Shallow Crustal Seismicity Regions*

The core class supporting the Kotha et al. (2020) GMM that forms the regionalisable scaled backbone ground motion model for crustal seismicity in ESHM20 is implemented as *KothaEtAl2020ESHM20.* This model requires as input the earthquake magnitude and hypocentral depth, the Joyner-Boore distance, $V_{S30}$, a Boolean term indicating whether the $V_{S30}$ refers to a "measured" or an "inferred" site condition, and finally the cluster region to which the site belongs. This last term indicates to the code which of the $\delta c_{3,clp}$ distributions for each of the 5 clusters should be applied to the site. This is expressed as the site property "region", for which a single integer should be input in the range 0 – 5, where values greater than 0 define the cluster region to which the site belongs, while 0 indicates that the default (unassigned) distribution should be used.

For the maximum possible flexibility, adjustments to the model based on $\delta L2L_l$ and/or $\delta c_3$ can be input in one of two ways: i) the specific values of $\delta L2L_l(T)$ and/or $c_3(T)$ can be input in the form of a dictionary that associates the specific $\delta L2L_l$ and/or $c_3$ value to the intensity measure type (e.g. PGA, Sa (T)) in question, or ii) as a value of $\varepsilon$ that indicates to the code the number of standard deviations to multiply either $\tau_{L2L}$ or $\tau_{c_3}$ (where $\tau_{c_3}$ will correspond to $\tau_{c_3,clp}$ for the cluster to which the site belongs).

The four controlling arguments are therefore:

*sigma_mu_epsilon*: The number of standard deviations ($\varepsilon$) by which to multiply the regional stress parameter scaling, such that $ln\ ln\ Y = ln\ ln\ \underline{Y} + \varepsilon \cdot (\tau_{L2L}, \sigma_{statistical})$

*c3_epsilon*: The number of standard deviations ($\varepsilon$) by which to multiple the residual attenuation scaling such that $c_3 = \underline{c_3} + \varepsilon \cdot \tau_{c_3}$, where $\underline{c_3}$ is the either original $c_3$ value of the Kotha et al. (2020) model or $c_3 + \delta_{c_3,clp}$ for the respective cluster $p$.

*dl2l:* The explicit $\delta L2L_l$ value for each period and/or intensity measure type such that $ln\ ln\ Y(T) = \underline{ln\ ln\ Y\ (T)} + \delta L2L_l(T)$

*c3*: The explicit $c_3(T)$ value to be used for each intensity measure type

Note that *sigma_mu_epislon* and *dl2l* are mutually exclusive operators, as too are *c3_epsilon* and *c3*, meaning that only one or the other should be input. If both *sigma_mu_epsilon* and *dl2l* are input, then the *dl2l* value will override the *sigma_mu_epislon.* Likewise, *c3* will override *c3_epsilon* in the case that both are input.

Finally, a non-ergodic version of the GMM can be run in which $\sigma = \sqrt{\tau_0^2 + \phi_0^2}$ and $\phi_{S2S}$ is set to 0, in the case that the model is to be used for site-specific seismic hazard analysis. This can be selected by setting the keyword *ergodic* from True (its default value) to False.

#### *5.2.1.2 GMM Logic Tree for Cratonic Regions*

The scaled backbone ground motion model for application to the stable cratonic part of Europe is implemented in the OpenQuake class *ESHM20Craton*. It is parameterized with $M_W$, rupture distance ($R_{RUP}$) and $V_{S30}$, but in contrast to the model for shallow crustal seismicity there is no need to specify whether the site in question refers to a *measured* or *inferred* case. The configurable parameters that can be passed to the GMM in order to apply the different scaling factors are:

*epsilon:* The number of standard deviations ($\varepsilon$) by which to multiply the epistemic uncertainty on median ground motion ($\sigma_\mu$) term such that $ln\ ln\ Y = ln\ ln\ \underline{Y} + \varepsilon \cdot \sigma_\mu$

*site_epsilon:* The number of standard deviations ($\varepsilon_S$) by which to multiply the epistemic uncertainty on site amplification ($\sigma_{\mu,S}$).

As with the shallow crustal seismicity scaled backbone GMM, a non-ergodic version of the GMM can be run by seeing the keyword argument *ergodic* to False.

#### *5.2.1.3 GMM Logic Tree for Subduction and Deep Seismicity Regions*

The backbone ground motion model for subduction and deep seismicity is an adaptation of the original subduction GMM of Abrahamson et al. (2016). The new GMM is implemented under the name *BCHydroESHM20Sinter* and *BCHydroESHM20SSlab*. The required parameters are $M_W$, distance to the rupture ($R_{RUP}$) for interface events or distance to the hypocentre ($R_{HYPO}$) for in-slab and/or Vrancea DSZ sources, $V_{S30}$ and distance of the site to the forearc/backarc front (or volcanic front), $x_{vf}$ (km). Configurable parameters include:

*sigma_mu_epsilon*: The number of standard deviations ($\varepsilon$) by which to multiply the epistmic uncertainty on median ground motion ($\sigma_\mu$) term such that $ln\ ln\ Y = ln\ ln\ \underline{Y} + \varepsilon \cdot \sigma_\mu$

*theta6_adjustment*: The adjustment factor ($\delta\theta_6$) to be applied to the anelastic attenuation term ($\theta_6$) such that $\theta_6' = \theta_6 + \ \delta\theta_6$

*faba_taper_model*: The choice of functional form to use for the forearc/backarc scaling region, from "step" (default, behaves as the original GMM with an immediate transition from forearc to backarc across the front), "Linear" (a linear transition model), "SFunc" (an S-function, see equation 6.8),

"Sigmoid" (a sigmoid function) or "Gaussian" (a Gaussian cumulative distribution function). Depending on the choice of taper model additional parameters configure the shape of the function (and vary depending on the function). These include *a, b* (the end limits over which the function is applied), "sigma" (the bandwidth for the Gaussian function), "width" (the length of the distance transition for the linear model"), and "c" the shape parameter of the sigmoid model.

As with the other models, a non-ergodic version of this GMM can be run by setting *ergodic* to False.

## 5.3 OpenQuake Calculation Settings

The hazard calculation is optimized for the computational grid (see **Fig. 5.3**), logic tree structure (both the seismogenic source model and the ground motion backbone model), output and the hardware capacity.
The calculation is performed for three different regions:1) the main inland Europe plus Turkey, 2) Southwest Islands and 3) Iceland. It is worth mentioning that the logic tree of the seismogenic source model is identical for all regions, but a different ground motion logic tree was used for Iceland (see chapter 6, section 6.5.2).

From the calculation point of view, we made pragmatic decisions to optimize the computational time and the hardware availability:

- single branch of the smoothed seismicity
- single branch of the subduction interface
- 12 correlated end-branches for the area source model describing the shallow crust, volcanic, deep and inslab sources
- 9 correlated end-branches for the active shallow crust faults
- reduce the number of the computational grid-points, mainly off-shore (computational geo-grid (~ 130 00 grids points)
- use point ruptures as they are not critical for the return period of interest (RP=475yrs)
- use the weighted depth for area and point sources
- use the predominant style-of-faulting for area and point sources
- reduce the size of the incremental distribution for smoothed seismicity by clipping the annual incremental rates lower than $10^{-6}$ (1/1million occurrence not contributing to the return period of interest)
- fault source mesh sampling intervals to 2km
- complex fault mesh spacing of 50km
- use of 0.2 magnitude bin for the smoothed seismicity
- 16 IMTs to cover peak ground acceleration (PGA), and response spectra acceleration with 5% damping at predominant periods in the range of 0.05s to 5s.
- ground motion discretization levels, 25 intervals from 0.0005g to 5g (i.e. x-axis of the hazard curve)
- use the point source gridding approximation(with the following parameters: *point_source_distance* set to 75, and *ps_grid_spacing* set to 50, based on https://docs.openquake.org/oq-engine/advanced/point-source-gridding.html)
- random sampling technique implemented in OpenQuake is used to sample 10000 logic tree end-branches.

With the above settings, the calculation runtime is about 20hours in a server with 128CPU and 0.5T SDRAM. The config file of the ESHM20 calculation for OpenQuake is provided together with the main input files, see chapter 7 Data Resources. The results of the hazard calculation are given in the next chapter.

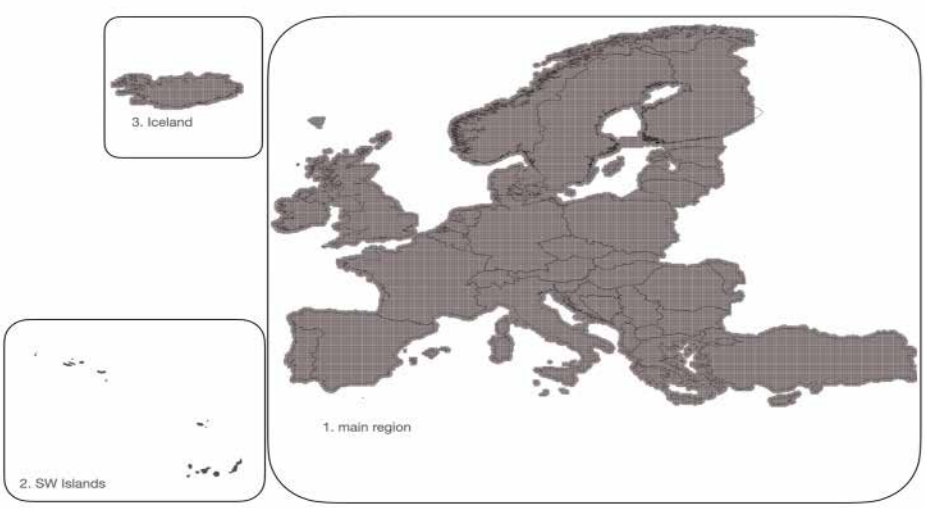



**Fig. 5.3** Computational grid for seismic hazard calculation of ESHM20 – the grid points are equally spaced at 0.1 to 0.1 degrees – total number of grid points 97920.

# 6 Main Results

The ESHM20 results depict time-independent earthquake ground-shaking exceedance levels for a uniform rock site condition of Vs,30 ≥ 800 m/s and valid for GRot50 horizontal ground motion as specified on the ground motion models (chapter 6). For additional and supplementary materials, we recommend the reader to consult *chapter 7: Data, Products and Resources*.

The main results of the ESHM20 are:

- Hazard maps for specified intensity measure types (PGA, spectra acceleration with 5% damping at predominant periods in the range of 0.05s to 5s) and five mean return periods (i.e. 50, 475, 975, 2500 and 5000 years).
- Hazard curves at every computational site, depicting the mean, median (50th) and four quantiles (5th, 16th, 84th and 95th) for all intensity measure types.
- Uniform Hazard Spectra depicting the mean, median (50th) and four quantiles (5th, 16th, 84th and 95th) and five mean return periods (i.e. 50, 475, 975, 2500 and 5000 years) estimated at every location of the computational grid.
- Disaggregation of the hazard results. Seismic hazard disaggregation is a technique that allows identifying the earthquake scenarios that significantly contribute to a specified exceedance probability of ground motion levels. Ground motion is represented by the PGA or acceleration response spectra at different periods (0.05 to 5s). It is foreseen that in the next year, the disaggregation of the hazard results will be provided as an online tool.

**Important notes**:

The ESHM20 results constitute an updated reference for the European region and Turkey, but these results do not replace nor modify the existing national design regulations and seismic provisions. Furthermore, in Europe the ground motion hazard estimates do not directly translate into design values; these must be conformed for seismic design and construction of buildings at the national level provided by the national design codes or by Nationally Determined Parameters (NDPs) in EN Eurocodes.

Hazard curves are computed up to very low annual probabilities (10-4, corresponding to 10000 years return periods) – however, caution is to be used when interpreting the curves at very low probability levels because a regional hazard model has limitations and uncertainties that cannot be resolved at a regional scale We therefore limit result representation for hazard maps to 5000 years mean return period.

## 6.1 Seismic Hazard Maps

### 6.1.1 475yrs ground shaking hazard maps

Figure 8 illustrates the spatial distribution of the PGA mean values for a 10 % probability of exceedance in 50 years, corresponding to a mean annual frequency of exceedance of about 0.002107, and to a mean return period of 475 years (1/0.002107 ~ 475 years). The mean seismic hazard map incorporates the range of uncertainty defined at the level of seismogenic sources combined with the spatially variable backbone ground motion model. The range around the mean estimates are given by the percentile estimates ($16^{th}$, $50^{th}$, $84^{th}$) and illustrated in **Fig. 6.1**(a to d).

As it might be expected, the contours of equal probabilistic ground motion follow the patterns of the regions where earthquakes were observed in the past or in the proximity of the active faults.

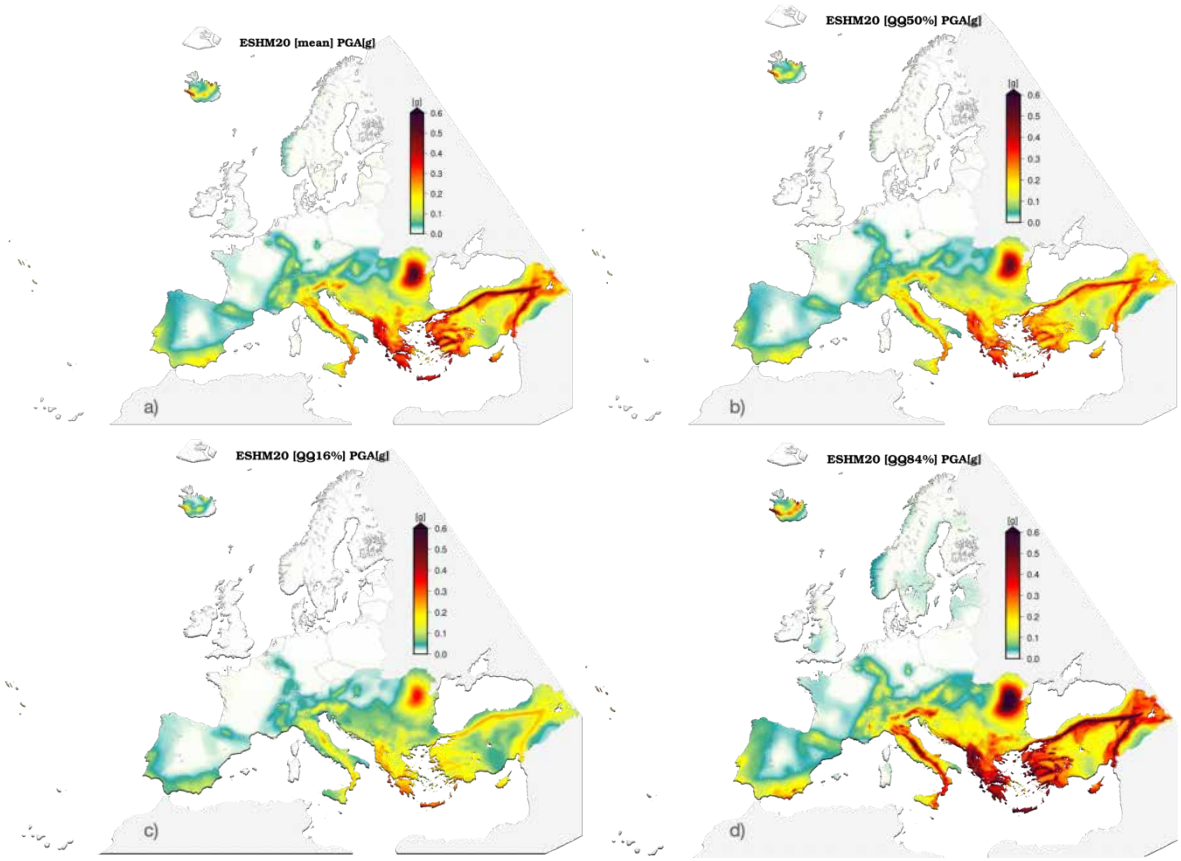


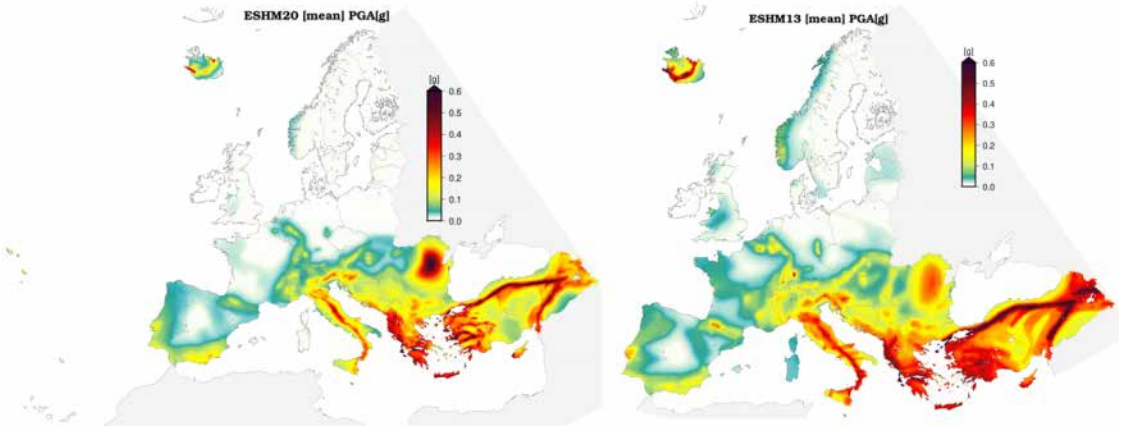

**Fig. 6.1** Ground shaking hazard maps for PGA [g] mean (a), median(b), 16% quantiles (c) and 84% quantiles (d) values for a return period of 475yrs.

## 6.2 Comparison with ESHM13

The seismic hazard maps for this update and the 2013 maps are shown for PGA for 10% probability of exceedance in 50 years on a uniform rock site condition (Vs,30>800m/s) in **Fig. 6.2**.

Noteworthy, ESHM20 as ESHM13 also, provides more than 500 hazard maps, for various intensity measure types (PGA, spectra acceleration with 5% damping at predominant periods in the range of 0.05s to 5s) and five mean return periods (i.e. 50, 475, 975, 2500 and 5000 years). In the interest of space, we provide only the comparison for the PGA mean hazard values. Additional comparison maps are provided online at the ESHM20 GitLab repository (Danciu et al 2021, https://doi.org/10.12686/a16).

Generally, the ESHM13 and ESHM20 maps have a similar spatial pattern with the new hazard maps being overall lower in most of the areas, as shown in **Fig. 6.2** together with a map of the ESHM20-ESHM13 values differ at each site. However, increased values are also observed in some areas in Romania, Albania, Greece, Western Turkey, Southern Spain and Southern Portugal. The largest reduction between the two models is observed in Iceland, which is due to the changes on the earthquake catalogues, active faults and calibrated ground motion models. These differences are likely caused by the updated seismogenic sources and new backbone ground motion models.



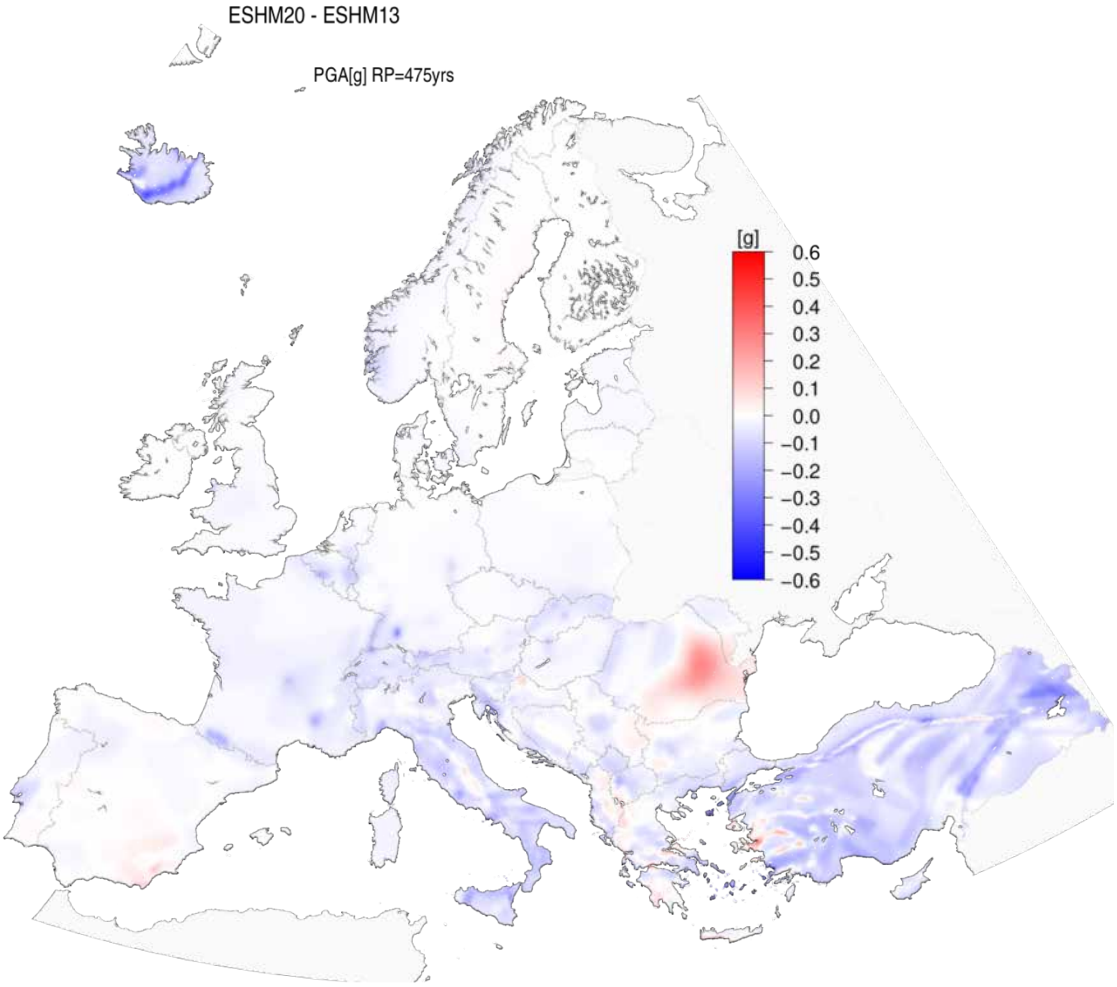




**Fig. 6.2** Spatial variability of the PGA [g] difference on mean values of the ESHM20 versus ESHM13 for a return period of 475yrs. Red color indicates an increase of PGA values when compared with the ESHM13 estimates, and the blue color indicates the opposite. The 475yrs – PGA[g] hazard maps for ESHM20 (top left) and for ESHM13 (top right) are also illustrated.

Changes in the seismogenic sources cause many of the local differences across the entire region. Regional discrepancies in the earthquake rates are likely caused by the new earthquake catalogue, new completeness time-magnitude intervals, new magnitude frequency distributions (Pareto Tapered Distribution and exponential GR distribution), updated slip-rates and maximum magnitude of the faults,

new adaptive-smoothing technique, new subduction sources, new logic tree and its implementation. To investigate these changes the *ensemble* earthquake rate forecast of the ESHM20 minus that of ESHM13 are compared at each grid site for two magnitudes, Mw 5.5 and Mw 6.5, in **Fig. 6.3**.

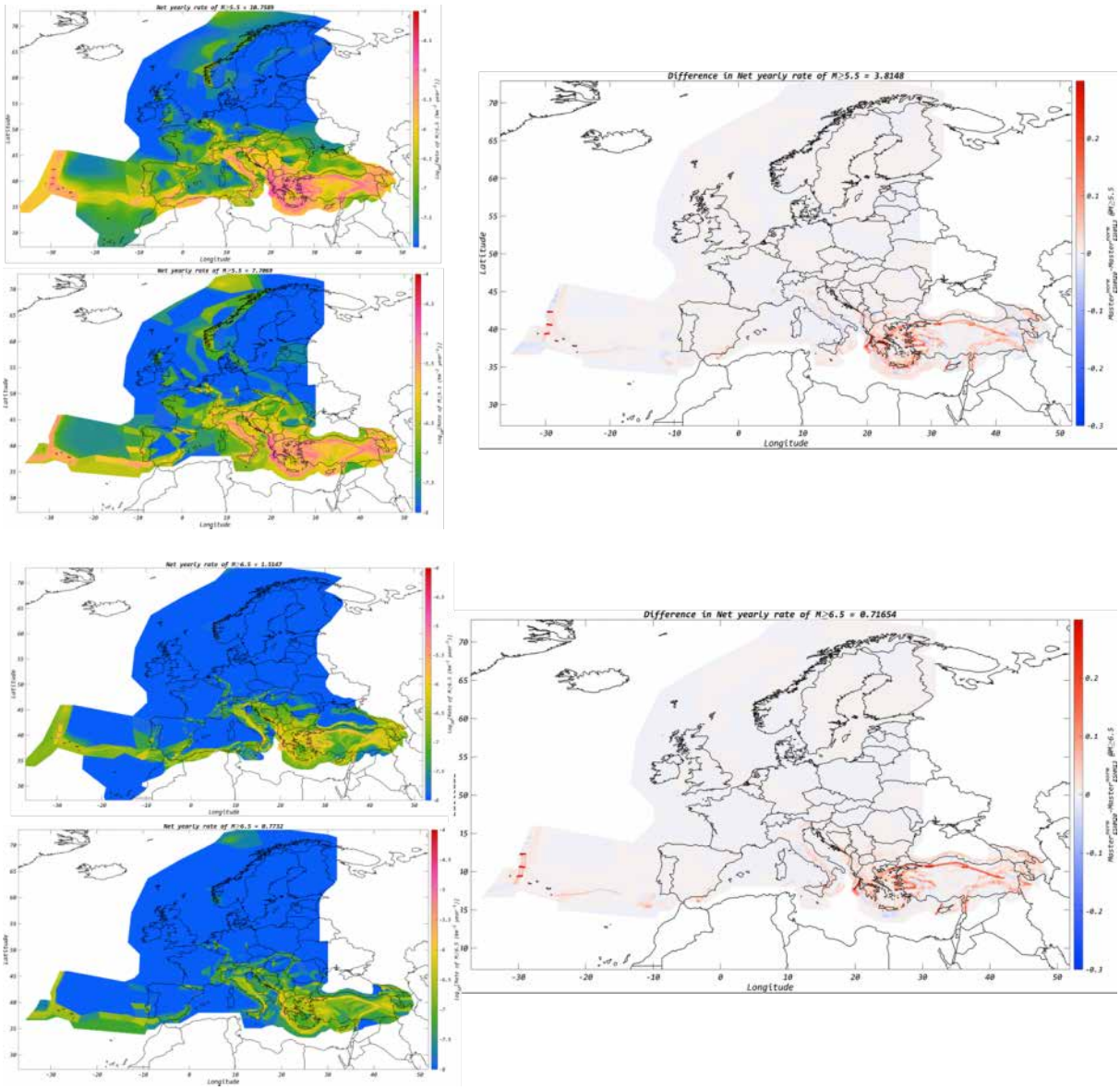



**Fig. 6.3** Maps of annual earthquake rate forecast of ESHM20 and ESHM13 ensemble models for magnitude M>5.5 and M>6.5 (left column) and the difference maps i.e. ESHM20 – ESHM13(right column).

Weatherill et al. (2020) discuss the comparison between the two ground motion models as illustrated by the trellis plot comparisons (see **Fig. 6.4**) of the new model versus those GMMs used for shallow crustal regions in the ESHM13. The latter include Akkar and Bommer (2010), Cauzzi and Faccioli (2008), Chiou and Youngs (2008), and Zhao et al. (2006) for "active shallow crustal regions", plus Toro (2002) and Campbell (2003), which were added for "stable shallow crustal regions" (excluding the Shield region).

For those magnitudes and distances well-constrained by the data (typically MW < 6.0 and 10 ≤ RJB (km) ≤ 100) the centre and range of this backbone agree relatively well with the range predicted by the ESHM13 selection for active shallow crust for periods greater than approximately 0.2 - 0.3s.

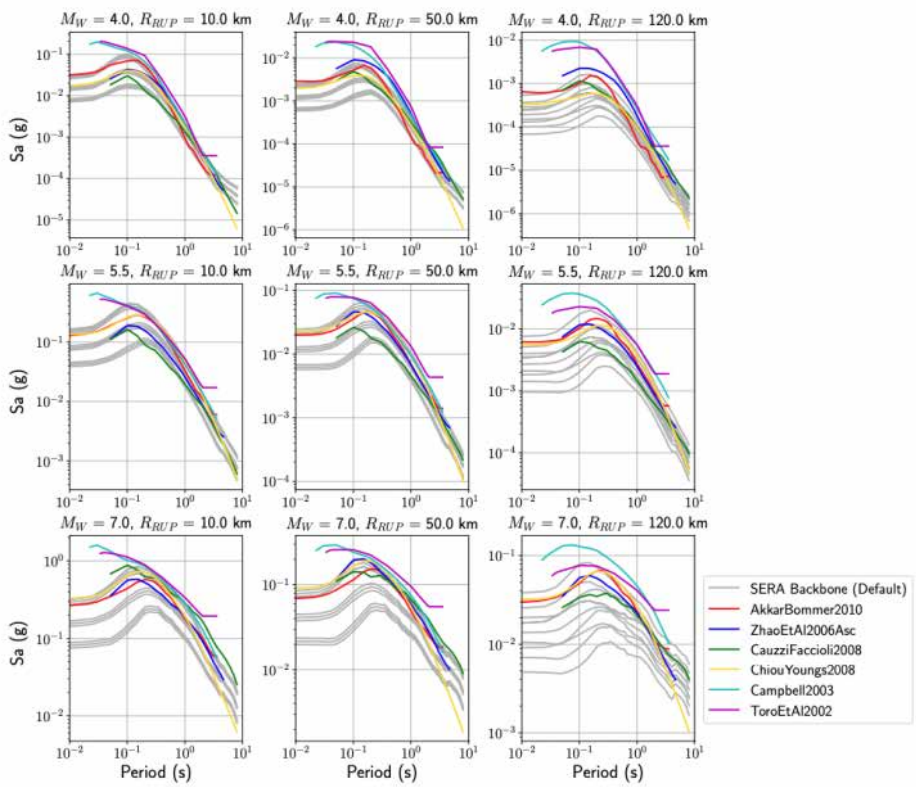



**Fig. 6.4** Trellis plots comparing the response spectra of the nine-branch default shallow logic tree against the ESHM13 shallow GMPE selection Delavaud et al. (2012) for a strike-slip earthquake with hypocentral depth between 10 and 20 km, assuming a measured site condition of VS30 800 m/s *(from Weatheril and Cotton 2020)*

The greatest differences can be seen at very short periods (< 0.1 s) where the new models seem to predict lower motions in general. Greater divergences emerge at longer distances, where the general trend of the new models is toward faster attenuation and thus the centre and body of the new GMM logic tree predict lower motions in general.

There are several that may be influential in accounting for this discrepancy. The first is that the parameterisation of site properties Engineering Strong Motion (ESM) Database (Lanzano et al., 2019) is improved with respect to previous models, resulting in improved characterisation of the site scaling at short periods. This trend toward lower short period motion on rock has been seen in ground motion models from many parts of the world, including the preceding generation of European GMMs based on the RESORCE strong motion database.

Additionally, in comparison to the previous strong motion databases for Europe, particularly that used by Akkar and Bommer (2010), the proportion of records from the central and northern Apennine region in Italy, is significantly increased and therefore skews the centre of the data set toward these conditions, which are generally faster attenuating than the rest of Europe on the whole. This effect is mitigated for the seismically active regions of southern Europe, however, by use of the regionally calibrated attenuation coefficients.

Compounding the first two factors is the fact, that the ESHM13 model selection contains GMMs constrained by data from other regions outside Europe and the Middle East, including those that may be more slowly attenuating or display particular source or site characteristics that are distinct from those of the records found in the ESM Database.

## 6.3 Hazard Contribution per Source Model

The ground shaking hazard maps estimated for each of the two input seismogenic source models, the area sources model (ASM) and the hybrid active faults plus background smoothed seismicity model (FS+BG) are illustrated in the following for a mean return period of 475 years. Spatial distribution of PGA[g] mean values for ASM is given in **Fig. 6.5**(a), for FS+GB in **Fig. 6.5**(b.) Additionally, the two complementary PGA maps for subduction inslab plus deep seismicity and subduction interface are given in in **Fig. 6.5**(c) and in **Fig. 6.5**(d), respectively.

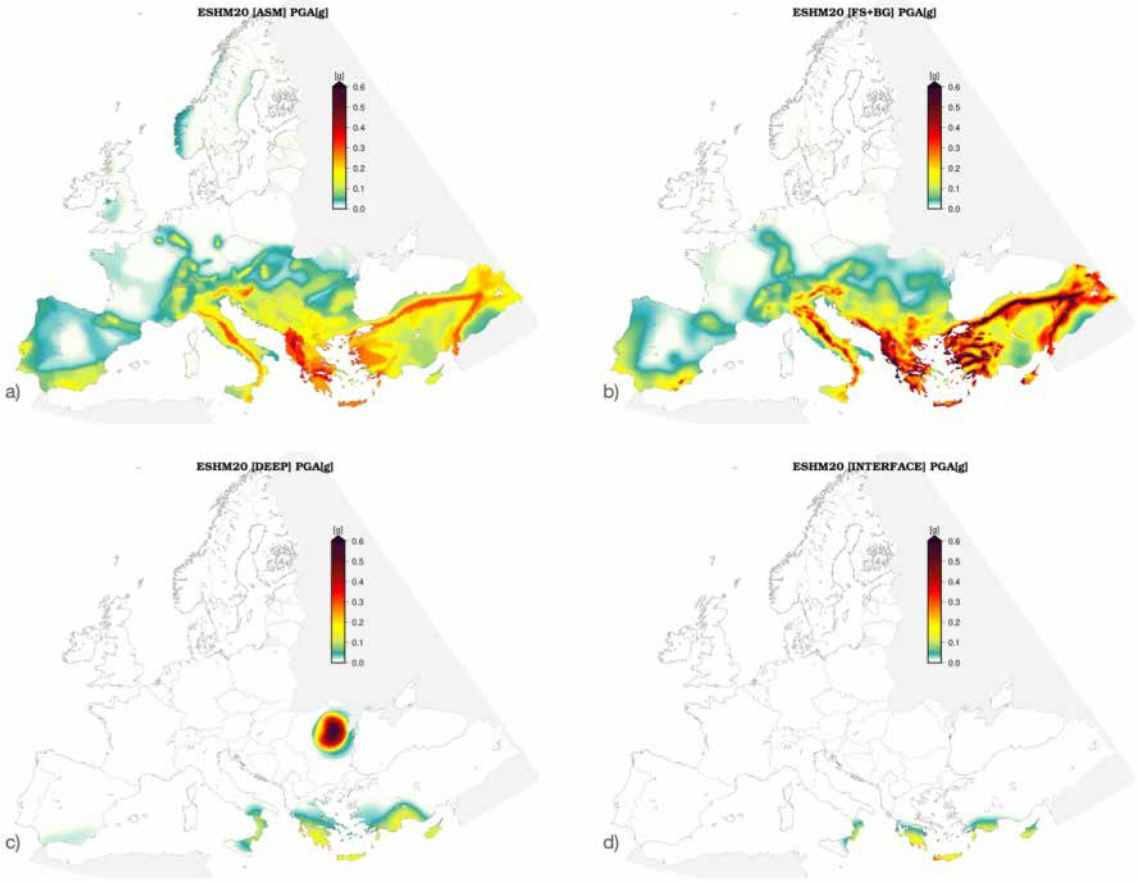



**Fig. 6.5** Ground shaking hazard maps for PGA [g] mean as an outcome of ASM(a), FS+BG(b), deep seismicity (c) and subduction interface (d) models for a return period of 475yrs.

## 6.4 Site Specific Hazard Estimates

### 6.4.1 Uniform Hazard Spectra

In this section, examples of the uniform hazard spectra for the four cities of Rhodes, L'Aquila, Lisbon and Cologne (same as in Wössner et al 2015) for a mean return period of 475 years are illustrated in **Fig. 6.6** Mean, median and four quantiles (5%, 16%, 84%,95%) as obtained from the ESHM20 full logic tree are also compared with those of ESHM13 (**Fig. 6.6**). Noteworthy is the reduction of the ESHM20 when compared with ESHM13 for Cologne and Lisbon, and the consistency of the two models for Rhodes.

For L'Aquila the median values are consistent for short spectral ordinates (<0.1s) and ESHM20 values increase above this spectral period; the mean and quantile values of ESHM20 are lower than those of

ESHM13. Overall the uncertainty range (16% to 84%) of the ESHM13 is higher than the corresponding ESHM20 range for Rhodes, Cologne and Lisbon; in fact, the observed uncertainty 16% to 84% range of ESHM13 is higher than the 9% to 95% of ESHM20. For L'Aquila, the uncertainty range of ESHM13 is significantly lower than the one of ESHM20.

As mentioned above, these differences can be described by the changes in the earthquake rates, ground motion models, but also on the way the logic tree of the two models have been implemented. For instance, in the ESHM13 only the variability due to ground motion models was explored, as the seismogenic source model was collapsed to a mean branch per model, whereas in ESHM20, the logic tree of both seismogenic source and ground motion models are randomly sampled, resulting in more robust distribution around the central values (mean and median).

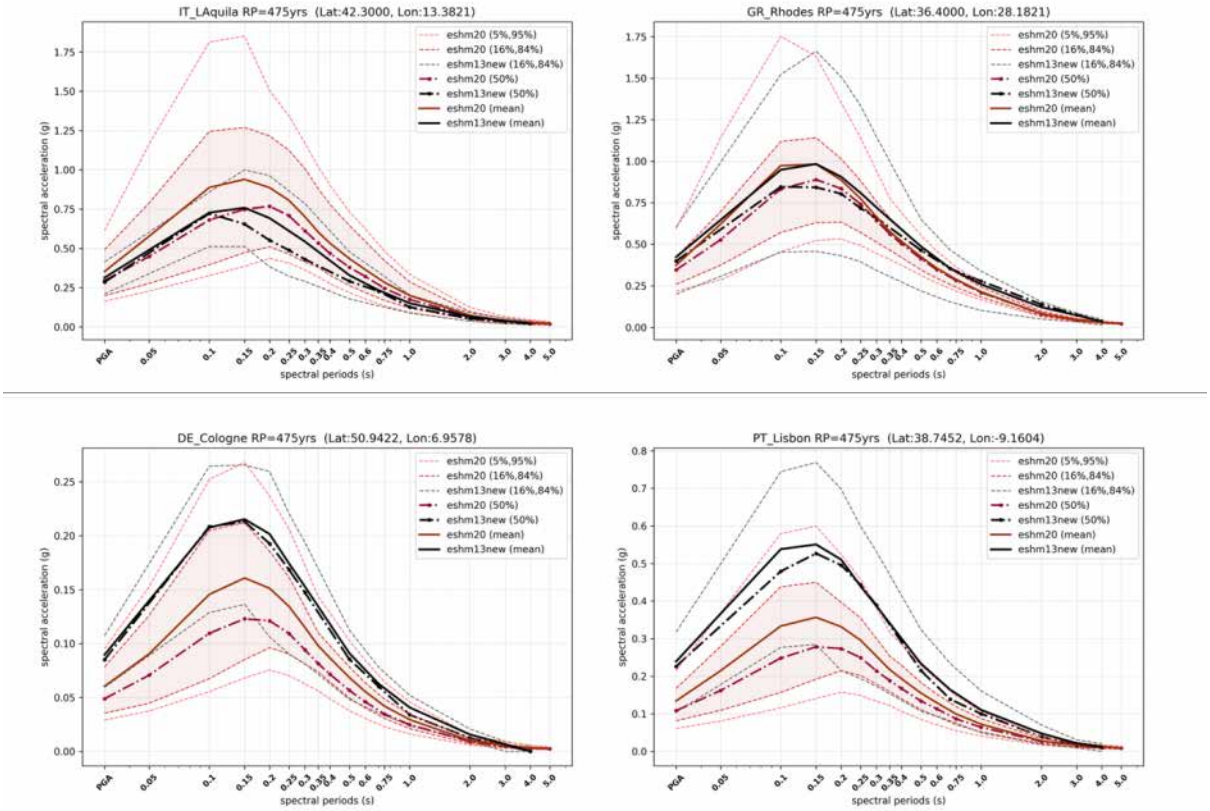



**Fig. 6.6** Uniform Hazard Spectra for a mean return period of 475yrs at the selected locations: L'Aquila (Italy), Rhodes (Greece), Cologne (Germany) and Lisbon (Portugal). Mean, median and 16% and 84% quantiles for ESHM20 and ESHM13 (Wössner et al 2015)

### 6.4.2 Hazard Curves

The hazard curves at the same cities of Wössner et al 2015, i.e. Rhodes, L'Aquila, Lisbon and Cologne, are shown in **Fig. 6.7** for mean, median and two quantiles (16% and 85%). Similarly, with the trend observed when new estimates of ESHM20 are compared with the ESHM13, there are changes that are due to differences of the main components of the two models, from the earthquake rates forecast to ground motion models, and their implementation for hazard calculation.

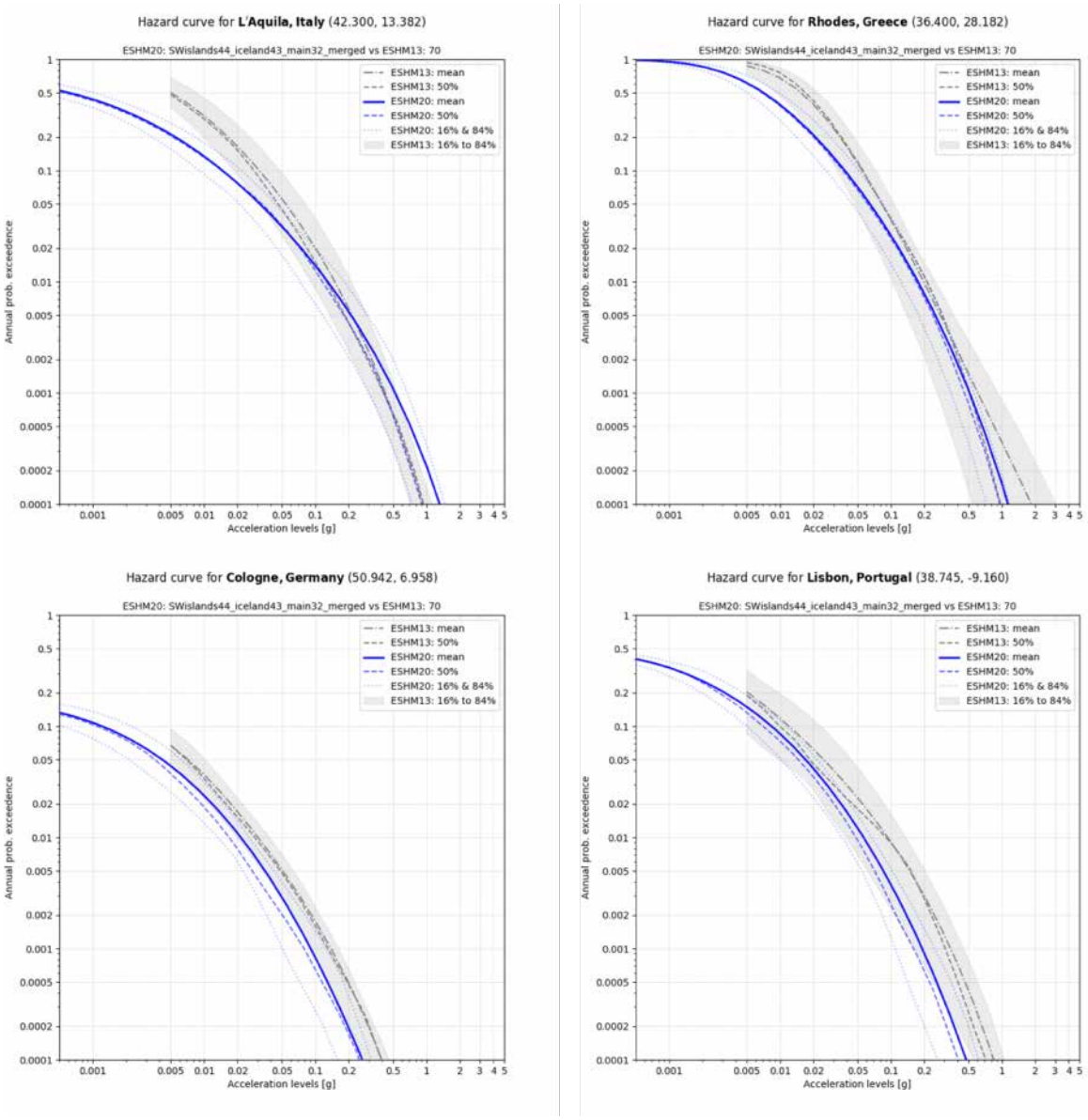



**Fig. 6.7** Hazard curves for PGA at the selected locations: L'Aquila (Italy), Rhodes (Greece), Cologne (Germany) and Lisbon (Portugal). Mean, median and 16% and 84% quantiles for ESHM20 and ESHM13 (Wössner et al 2015)

# 7 Data, Products and Resources

## 7.1 Open Data and Licensing

The EFEHR Consortium decided in its first General Assembly in September 2020 to openly release all products associated with the European hazard and risk models. The source data, input models, software and outputs of ESRM20 are thus being openly released with a Creative Commons CC-BY International 4.0 license (https://creativecommons.org/licenses/by/4.0/). This license allows re-users to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use. Each product is released with a clear notice on how it should be cited in order to abide by the license.

## 7.2 OQ input Files

The OpenQuake input files follow two distinct paths of the logic tree source model listed in **Fig.5.1** and Fig. **7.1** Each path computational path consists of a full seismogenic source model describing:

1. Area Source Branch:
    a. shallow area sources
    b. deep and subduction in-slab area sources
    c. subduction interface sources
    d. volcanic area sources
2. Fault Source + Smoothed Seismicity Branch:
    a. Active Faults for shallow crust seismicity
    b. Smoothed seismicity
    c. deep and subduction in-slab area sources
    d. subduction interface sources
    e. volcanic area sources

The hazard calculation is divided into three regions: *region_main*, *region_SWislands*, *region_iceland*. These regions are defined in the region filter shapefile and are the name of the highest-level folder. The main shape files are available at ESHM20 main input file repository https://doi.org/10.12686/A16.

#### *7.2.1.1 Region Main:*

The OpenQuake input files for Region main are available at:
https://gitlab.seismo.ethz.ch/efehr/eshm20/-/tree/master/oq_computational/oq_configuration_eshm20_v12e_region_main

- configuration files to run the hazard calculation: **config_eshm20_v12e_main_region.ini**
- list of computational grid points: **eshm20_site_model_v06d.csv**
- ground motion logic tree describing the backbone logic tree: **gmpe_complete_logic_tree_5br.xml**
- source model logic tree: **source_model_logic_tree_eshm20_model_v12e.xml**
- directory of the **source_models**, available at

#### *7.2.1.2 Region SWislands:*

The OpenQuake input files for Region SW Islands are available at:
https://gitlab.seismo.ethz.ch/efehr/eshm20/-/tree/master/oq_computational/oq_configuration_eshm20_v12e_region_swIslands

- configuration files to run the hazard calculation: **config_eshm20_v12e_oceanic.ini**
- list of computational grid points: **oceanic_site_model_5km.xml**
- ground motion logic tree describing the backbone logic tree: **gmpe_complete_logic_tree_5br.xml**
- source model logic tree: **source_model_logic_tree_eshm20_oceanic_v12e.xml**
- directory of the **source_models**

#### *7.2.1.3 Region Iceland:*

The OpenQuake input files for Region Iceland are available at:
https://gitlab.seismo.ethz.ch/efehr/eshm20/-/tree/master/oq_computational/oq_configuration_eshm20_v12e_region_iceland

- configuration files to run the hazard calculation: **config_eshm20_v12e_iceland.ini**
- list of computational grid points: **eshm20_iceland_site_model.xml**
- ground motion logic tree describing the backbone logic tree: **gmpe_complete_logic_tree_iceland_5branch.xml**
- source model logic tree: **source_model_logic_tree_eshm20_model_iceland_v12e.xml**
- directory of the **source_models**

## 7.3 Online Resources:

An overview of the ESHM2020 project, complete with links to the main elements is available from http://hazard.efehr.org/en/Documentation/specific-hazard-models/europe/eshm2020-overview/

The data is available from the EFEHR web service, available at http://hazard.efehr.org/en/web-services/

Specific web service links follow:

- Hazard Curve data is available from
- http://hazard.efehr.org/en/web-services/hazard-curve-data/
- Hazard Map data is available from
- http://hazard.efehr.org/en/web-services/hazard-map-data/
- Uniform hazard spectra data is available from
- http://hazard.efehr.org/en/web-services/uniform-hazard-spectra/

Tools
QGis available from https://www.qgis.org
GeoPandas https://geopandas.org
OpenQuake https://www.globalquakemodel.org

# References


**Instrumental Earthquake Catalogue**

Grünthal G and Wahlström R (2012) The European-Mediterranean Earthquake Catalogue (EMEC) for the last millennium. Journal of Seismology. 16: 535 – 570

Grünthal G, Wahlström R, Stromeyer D (2013) The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900 – 2006 and its comparison to the European-Mediterranean Earthquake Catalogue (EMEC). Journal of Seismology. 17: 1339 – 1344.

Kadiriogliu FT, Kartal RF, Kiliç T, et al. (2018) An improved earthquake catalogue (M ≥ 4.0) for Turkey and its near vicinity (1900 – 2012). Bulletin of Earthquake Engineering. 16: 3317 - 3338

INFP (2018) Data file of Romanian earthquakes. National Institute for Earth Physics, Bucharest, Romania (http://www.infp.ro/)

Manchuel K, Traversa P, Baumont D, et al. (2018) The French seismic CATalogue (FCAT-17). Bulletin of Earthquake Engineering. 16: 2227 – 2251.

Rovida A, Locati M, Camassi R, et al. (2020) The Italian earthquake catalogue CPTI15. Bulletin of Earthquake Engineering. 18: 2953 – 2984.

Živčić M (2018) Data file "Earthquake Catalogue of Slovenia" of the Seismology and Geology Office. Environmental Agency of the Republic of Slovenia, Ljubljana, Slovenia

**Historical Earthquake Catalogue**

Albini P., Locati M., Rovida A., Stucchi M. (2013). European Archive of Historical EArthquake Data (AHEAD). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.6092/ingv.it-ahead

Bakun W.H., Wentworth C.M. (1997). Estimating earthquake location and magnitude from seismic intensity data. Bull Seismol Soc Am, 87, 1502-1521.

Basili R., Carafa M.M.C., Kastelic V., Maesano F.E., Tiberti M.M., Rovida A., Antonucci A., Weatherill G., Lammers S., Danciu L. (2018). D25.2 - Updated databases of seismicity, faults, and strain rates for ESHM20. Deliverable 25.2, WP25 of the SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe) Project, Horizon 2020 Research and Innovation Programme, grant agreement No 730900, 70 pp.

Bungum H., Lindholm C.D., Dahle A. (2003). Long-period ground-motions for large European earthquakes, 1905–1992, and comparisons with stochastic predictions. J Seismol 7, 377–396.

Gasperini P., Bernardini F., Valensise G., Boschi E. (1999). Defining seismogenic sources from historical earthquake felt reports. Bull Seismol Soc Am, 89, 94-110.

Gasperini P., Vannucci G., Tripone D., Boschi E. (2010). The Location and Sizing of Historical Earthquakes Using the Attenuation of Macroseismic Intensity with Distance. Bull Seismol Soc Am, 100, 2035–2066.

Gomez Capera A.A., Rovida A., Gasperini P., Stucchi M., Viganò D. (2015). The determination of earthquake location and magnitude from macroseismic data in Europe. Bull Earthq Eng, 13, 5, 1249-1280, https://doi.org/10.1007/s10518-014-9672-3

Grünthal G., Wahlström R., Stromeyer, D. (2009). The unified catalogue of earthquakes in central, northern and northwestern Europe (CENEC) – updated and expanded to the last millennium, J Seismol 13, 4, 514-541, https://doi.org/10.1007/s10950-008-9144-9

Hammerl Ch., Lenhardt W.A. (2013). Erdbeben in Niederösterreich von 1000 bis 2009 n. Chr. Abhandlungen Der Geologischen Bundesanstalt, Band 67, 297 pp, Wien.Hammerl Ch., Lenhardt W.A. (2013). Erdbeben in Niederösterreich von 1000 bis 2009 n. Chr. Abhandlungen Der Geologischen Bundesanstalt, Band 67, 297 pp, Wien.

Leydecker G. (2011). Earthquake Catalogue for Germany and Adjacent Areas for the Years 800-2008. Federal Institute for Geosciences and Natural Resources, Hannover.

Locati M., Rovida A., Albini P., Stucchi M. (2014). The AHEAD portal: a gateway to European historical earthquake data. Seismol Res Lett 85, 3, 727-734. https://doi.org/10.1785/0220130113

Manchuel K., Traversa P., Baumont D., Cara M., Nayman E., Durouchoux C. (2018). The French seismic CATalogue (FCAT-17). Bull Earthquake Eng, 16, 6, 2227-2251 https://doi.org/10.1007/s10518-017-0236-1.

Musson R.M.W., Jimenéz M.J. (2008). Macroseismic estimation of earthquake parameters. NA4 deliverable D3, NERIES Project.

Rovida A., Antonucci A. (2021a). EPICA - European PreInstrumental Earthquake CAtalogue, version 1.1. Istituto Nazionale di Geofisica e Vulcanologia (INGV).Dataset. https://doi.org/10.13127/epica.1.1

Rovida A., Antonucci A. (2021b). EPICA - European PreInstrumental Earthquake CAtalogue, version 1.1. Technical report Istituto Nazionale di Geofisica e Vulcanologia (INGV), in preparation.

Rovida A., Locati M. (2015). Archive of Historical Earthquake Data for the European-Mediterranean Area. In: A. Ansal (ed.), Perspectives on European Earthquake Engineering and Seismology, Geotechnical, Geological and Earthquake Engineering 39, https://doi.org/10.1007/978-3-319-16964-4_14

Rovida A., Locati M., Antonucci A., Camassi R. (eds.) (2017). Italian Archive of Historical Earthquake Data (ASMI). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/asmi

Rovida A., Locati M., Camassi R., Lolli B., Gasperini P. (eds), (2016). CPTI15, the 2015 version of the Parametric Catalogue of Italian Earthquakes. Istituto Nazionale di Geofisica e Vulcanologia. https://doi.org/10.6092/INGV.IT-CPTI15

Rovida A., Locati M., Camassi R., Lolli B., Gasperini P. (2020a). The Italian earthquake catalogue CPTI15. Bull Earthquake Eng 18, 2953-2984. https://doi.org/10.1007/s10518-020-00818-y

Rovida A.N., Albini P., Locati M., Antonucci A. (2020b). Insights into Preinstrumental Earthquake Data and Catalogs in Europe. Seismol Res Lett 91, 5, 2546–2553. https://doi.org/10.1785/0220200058

Stucchi M., Rovida A., Gomez Capera A.A., Alexandre P., Camelbeeck T., Demircioglu M.B., Gasperini P., Kouskouna V., Musson R.M.W., Radulian M., Sesetyan K., Vilanova S., Baumont D., Bungum H., Fäh D., Lenhardt W., Makropoulos K., Martinez Solares J.M., Scotti O., Živčić M., Albini P., Batllo J., Papaioannou C., Tatevossian R., Locati M., Meletti C., Viganò D., Giardini D. (2013). The SHARE European Earthquake Catalogue (SHEEC) 1000-1899. J Seismol 17, 523-544, https://doi.org/10.1007/s10950-012-9335-2

**Active Faults and Subduction**

Akoglu, A. M., Cakir, Z., Meghraoui, M., Belabbes, S., El Alami, S. O., Ergintav, S., & Akyüz, H. S. (2006). The 1994–2004 Al Hoceima (Morocco) earthquake sequence: Conjugate fault ruptures deduced from InSAR. Earth and Planetary Science Letters, 252(3–4), 467–480. https://doi.org/10.1016/j.epsl.2006.10.010

Allen, T. I., & Hayes, G. P. (2017). Alternative Rupture-Scaling Relationships for Subduction Interface and Other Offshore Environments. Bulletin of the Seismological Society of America, 107(3), 1240–1253. https://doi.org/10.1785/0120160255

Amante, C. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis [Data set]. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5C8276M

Árnadóttir, T., Geirsson, H., & Jiang, W. (2008). Crustal deformation in Iceland: Plate spreading and earthquake deformation. Jökull, 58, 59–74.

Arroucau, P., Custódio, S., Civiero, C., Silveira, G., Dias, N., Díaz, J., et al. (2021). PRISM3D: a 3-D reference seismic model for Iberia and adjacent areas. Geophysical Journal International, 225(2), 789–810. https://doi.org/10.1093/gji/ggab005

Asch, K. (2005). The 1:5 Million International Geologial Map of Europe and Adjacent Areas - final version for the internet.- BGR, Hannover. Retrieved from https://www.bgr.bund.de/EN/Themen/Sammlungen-Grundlagen/GG_geol_Info/Karten/Europa/IGME5000/IGME_Project/IGME_Downloads.html

Atanackov, J., Jamšek Rupnik, P., Jež, J., Celarc, B., Novak, M., Milanič, B., et al. (2021). Database of Active Faults in Slovenia: Compiling a New Active Fault Database at the Junction Between the Alps, the Dinarides and the Pannonian Basin Tectonic Domains. Frontiers in Earth Science, 9, 604388. https://doi.org/10.3389/feart.2021.604388

Basili, R., Kastelic, V., Demircioglu, M. B., Garcia Moreno, D., Nemser, E. S., Petricca, P., et al. (2013, March 1). The European Database of Seismogenic Faults (EDSF) compiled in the framework of the Project

SHARE. Istituto Nazionale di Geofisica e Vulcanologia (INGV), doi: 10.6092/INGV.IT-SHARE-EDSF. Retrieved July 14, 2020, from http://diss.rm.ingv.it/share-edsf/

Basili, R., Carafa, M. M. C., Kastelic, V., Maesano, F. E., Tiberti, M. M., Rovida, A., et al. (2018). D25.2 - Updated databases of seismicity, faults, and strain rates for ESHM20. Deliverable 25.2, WP25 of the SERA (Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe) Project, Horizon 2020 Research and Innovation Programme, grant agreement No 730900 (p. 70).

Basili, R., Brizuela, B., Herrero, A., Iqbal, S., Lorito, S., Maesano, F. E., et al. (2021). The Making of the NEAM Tsunami Hazard Model 2018 (NEAMTHM18). Frontiers in Earth Science, 8, 616594. https://doi.org/10.3389/feart.2020.616594

Bell, R. E., McNeill, L. C., Bull, J. M., Henstock, T. J., Collier, R. E. L., & Leeder, M. R. (2009). Fault architecture, basin structure and evolution of the Gulf of Corinth Rift, central Greece. Basin Research, 21(6), 824–855. https://doi.org/10.1111/j.1365-2117.2009.00401.x

Bergerat, F., Angelier, J., & Villemin, T. (1990). Fault systems and stress patterns on emerged oceanic ridges: a case study in Iceland. Tectonophysics, 179(3–4), 183–197. https://doi.org/10.1016/0040-1951(90)90290-O

Bilek, S. L., & Lay, T. (1999). Rigidity variations with depth along interplate megathrust faults in subduction zones. Nature, 400(6743), 443–446. https://doi.org/10.1038/22739

Bird, P. (2003). An updated digital model of plate boundaries. Geochemistry, Geophysics, Geosystems, 4(3). https://doi.org/10.1029/2001GC000252

Borque, M. J., Sánchez-Alzola, A., Martin-Rojas, I., Alfaro, P., Molina, S., Rosa-Cintas, S., et al. (2019). How Much Nubia-Eurasia Convergence Is Accommodated by the NE End of the Eastern Betic Shear Zone (SE Spain)? Constraints From GPS Velocities. Tectonics, 38(5), 1824–1839. https://doi.org/10.1029/2018TC004970

Canora, C., Vilanova, S. P., Besana-Ostman, G. M., Carvalho, J., Heleno, S., & Fonseca, J. (2015). The Eastern Lower Tagus Valley Fault Zone in central Portugal: Active faulting in a low-deformation region within a major river environment. Tectonophysics, 660, 117–131. https://doi.org/10.1016/j.tecto.2015.08.026

Caputo, R., & Pavlides, S. (2013). Greek Database of Seismogenic Sources (GreDaSS): A compilation of potential seismogenic sources (Mw > 5.5) in the Aegean Region [Text/html,application/vnd.google-earth.kml+xml,image/jpg]. University of Ferrara, Italy. https://doi.org/10.15160/UNIFE/GREDASS/0200

Carafa, M. M. C., Kastelic, V., Bird, P., Maesano, F. E., & Valensise, G. (2018). A “Geodetic Gap” in the Calabrian Arc: Evidence for a Locked Subduction Megathrust? Geophysical Research Letters, 45(4), 1794–1804. https://doi.org/10.1002/2017gl076554

Civiero, C., Strak, V., Custódio, S., Silveira, G., Rawlinson, N., Arroucau, P., & Corela, C. (2018). A common deep source for upper-mantle upwellings below the Ibero-western Maghreb region from teleseismic P-wave travel-time tomography. Earth and Planetary Science Letters, 499, 157–172. https://doi.org/10.1016/j.epsl.2018.07.024

Danciu, L., Şeşetyan, K., Demircioglu, M., Gülen, L., Zare, M., Basili, R., et al. (2018). The 2014 Earthquake Model of the Middle East: seismogenic sources. Bulletin of Earthquake Engineering, 16(8), 3465–3496. https://doi.org/10.1007/s10518-017-0096-8

Davies, J. H. (2013). Global map of solid Earth surface heat flow: Global Surface Heat Flow Map. Geochemistry, Geophysics, Geosystems, 14(10), 4608–4622. https://doi.org/10.1002/ggge.20271

Delavaud, E., Cotton, F., Akkar, S., Scherbaum, F., Danciu, L., Beauval, C., et al. (2012). Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe. Journal of Seismology, 16(1), 1–23. https://doi.org/10.1007/s10950-012-9281-z

Demircioğlu, M. B., Şeşetyan, K., Duman, T. Y., Çan, T., Tekin, S., & Ergintav, S. (2018). A probabilistic seismic hazard assessment for the Turkish territory: part II—fault source and background seismicity model. Bulletin of Earthquake Engineering, 16(8), 3399–3438. https://doi.org/10.1007/s10518-017-0130-x

Devoti, R., Riguzzi, F., Cuffaro, M., & Doglioni, C. (2008). New GPS constraints on the kinematics of the Apennines subduction. Earth and Planetary Science Letters, 273(1–2), 163–174. https://doi.org/10.1016/j.epsl.2008.06.031

Diaconescu, M., Craiu, A., Moldovan, I. A., Constantinescu, E. G., & Ghita, C. (2019). Main active faults from eastern part of Romania (Dobrogea and Black Sea). Part II: Transversal and oblique faults system. Romanian Reports in Physics, 71(2), 708.

Diaconescu, M., Craiu, A., Toma-Danila, D., Craiu, G. M., & Ghita, C. (2019). Main active faults from Romania, Part I: Longitudinal and faults system. Romanian Reports in Physics, 71(1), 702.

Diaconescu, M., Craiu, A., Ghita, C., Moldovan, I. A., Oros, E., Constantinescu, E. G., & Marius, M. (2021). Main active faults from Romania. Part III: Fault systems from Dacia tectonic unit. Romanian Reports in Physics, 73, 710.

DISS Working Group. (2021). Database of Individual Seismogenic Sources (DISS), version 3.3.0. Istituto Nazionale di Geofisica e Vulcanologia (INGV), DOI: 10.6092/INGV.IT-DISS3.3.0. Retrieved November 30, 2021, from http://diss.rm.ingv.it/diss/

Emre, Ö., Duman, T. Y., Özalp, S., Şaroğlu, F., Olgun, Ş., Elmacı, H., & Çan, T. (2018). Active fault database of Turkey. Bulletin of Earthquake Engineering, 16(8), 3229–3275. https://doi.org/10.1007/s10518-016-0041-2

Escartín, J., Cowie, P. A., Searle, R. C., Allerton, S., Mitchell, N. C., MacLeod, C. J., & Slootweg, A. P. (1999). Quantifying tectonic strain and magmatic accretion at a slow spreading ridge segment, Mid-Atlantic Ridge, 29°N. Journal of Geophysical Research: Solid Earth, 104(B5), 10421–10437. https://doi.org/10.1029/1998JB900097

Fadil, A., Vernant, P., McClusky, S., Reilinger, R., Gomez, F., Ben Sari, D., et al. (2006). Active tectonics of the western Mediterranean: Geodetic evidence for rollback of a delaminated subcontinental lithospheric slab beneath the Rif Mountains, Morocco. Geology, 34(7), 529. https://doi.org/10.1130/G22291.1

Fernández-Blanco, D., de Gelder, G., Lacassin, R., & Armijo, R. (2019). A new crustal fault formed the modern Corinth Rift. Earth-Science Reviews, 199, 102919. https://doi.org/10.1016/j.earscirev.2019.102919

Forslund, T., & Gudmundsson, A. (1991). Crustal spreading due to dikes and faults in southwest Iceland. Journal of Structural Geology, 13(4), 443–457. https://doi.org/10.1016/0191-8141(91)90017-D

Ganas, A. (2021). NOAFAULTS KMZ layer Version 3.0.1 (2021 update) (Version V3.0.1) [Data set]. Zenodo. https://doi.org/10.5281/ZENODO.4897894

Ganas, A., & Parsons, T. (2009). Three-dimensional model of Hellenic Arc deformation and origin of the Cretan uplift. Journal of Geophysical Research, 114(B6), B06404. https://doi.org/10.1029/2008JB005599

Garcia, S., Angelier, J., Bergerat, F., & Homberg, C. (2002). Tectonic analysis of an oceanic transform fault zone based on fault-slip data and earthquake focal mechanisms: the Húsavík–Flatey Fault zone, Iceland. Tectonophysics, 344(3–4), 157–174. https://doi.org/10.1016/S0040-1951(01)00282-7

García-Mayordomo, J., Insua-Arévalo, J. M., Martínez-Díaz, J. J., Jiménez-Díaz, A., Martín-Banda, R., Martín-Alfageme, S., et al. (2012). The Quaternary Active Faults Database of Iberia (QAFI v.2.0). Journal of Iberian Geology, 38(1), 285–302. https://doi.org/10.5209/rev_JIGE.2012.v38.n1.39219

Geist, E. L., & Bilek, S. L. (2001). Effect of depth-dependent shear modulus on tsunami generation along subduction zones. Geophysical Research Letters, 28(7), 1315–1318. https://doi.org/10.1029/2000GL012385

Giardini, D., Woessner, J., Danciu, L., Valensise, G., Arvidsson, R., Sesetyan, K., et al. (2013). Seismic Hazard Harmonization in Europe (SHARE): Online Data Resource. Retrieved from http://portal.share-eu.org:8080/jetspeed/portal/, doi: 10.12686/SED-00000001-SHARE

Global Volcanism Program. (2013). Volcanoes of the World, v. 4.3.4. Venzke, E (ed.). Smithsonian Institution. https://doi.org/10.5479/si.GVP.VOTW4-2013 [Data set]. Retrieved from https://volcano.si.edu/database/webservices.cfm

Gold, R. D., Friedrich, A., Kübler, S., & Salamon, M. (2017). Apparent Late Quaternary Fault-Slip Rate Increase in the Southern Lower Rhine Graben, Central Europe. Bulletin of the Seismological Society of America, 107(2), 563–580. https://doi.org/10.1785/0120160197

Gomez, F., Barazangi, M., & Bensaid, M. (1996). Active tectonism in the intracontinental Middle Atlas Mountains of Morocco: synchronous crustal shortening and extension. Journal of the Geological Society, 153(3), 389–402. https://doi.org/10.1144/gsjgs.153.3.0389

Gómez-Novell, O., García-Mayordomo, J., Ortuño, M., Masana, E., & Chartier, T. (2020). Fault System-Based Probabilistic Seismic Hazard Assessment of a Moderate Seismicity Region: The Eastern Betics Shear Zone (SE Spain). Frontiers in Earth Science, 8, 579398. https://doi.org/10.3389/feart.2020.579398

Gómez-Novell, O., Chartier, T., García-Mayordomo, J., Ortuño, M., Masana, E., Insua-Arévalo, J. M., & Scotti, O. (2020). Modelling earthquake rupture rates in fault systems for seismic hazard assessment: The Eastern Betics Shear Zone. Engineering Geology, 265, 105452. https://doi.org/10.1016/j.enggeo.2019.105452

Grad, M., Tiira, T., & ESC Working Group. (2009). The Moho depth map of the European Plate. Geophysical Journal International, 176(1), 279–292. https://doi.org/10.1111/j.1365-246X.2008.03919.x

Gutscher, M.-A., Roger, J., Baptista, M.-A., Miranda, J. M., & Tinti, S. (2006). Source of the 1693 Catania earthquake and tsunami (southern Italy): New evidence from tsunami modeling of a locked subduction fault plane. Geophysical Research Letters, 33(8), L08309. https://doi.org/10.1029/2005GL025442

Gutscher, M.-A., Dominguez, S., Westbrook, G. K., & Leroy, P. (2009). Deep structure, recent deformation and analog modeling of the Gulf of Cadiz accretionary wedge: Implications for the 1755 Lisbon earthquake. Tectonophysics, 475(1), 85–97. https://doi.org/10.1016/j.tecto.2008.11.031

Halpaap, F., Rondenay, S., & Ottemöller, L. (2018). Seismicity, Deformation, and Metamorphism in the Western Hellenic Subduction Zone: New Constraints From Tomography. Journal of Geophysical Research: Solid Earth, 123(4), 3000–3026. https://doi.org/10.1002/2017JB015154

Halpaap, F., Rondenay, S., Perrin, A., Goes, S., Ottemöller, L., Austrheim, H., et al. (2019). Earthquakes track subduction fluids from slab source to mantle wedge sink. Science Advances, 5(4), eaav7369. https://doi.org/10.1126/sciadv.aav7369

Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. Science, 362(6410), 58–61. https://doi.org/10.1126/science.aat4723

Herrero-Barbero, P., Álvarez-Gómez, J. A., Martínez-Díaz, J. J., & Klimowitz, J. (2020). Neogene Basin Inversion and Recent Slip Rate Distribution of the Northern Termination of the Alhama de Murcia Fault (Eastern Betic Shear Zone, SE Spain). Tectonics, 39(7). https://doi.org/10.1029/2019TC005750

Heuret, A., Lallemand, S., Funiciello, F., Piromallo, C., & Faccenna, C. (2011). Physical characteristics of subduction interface type seismogenic zones revisited: SUBDUCTION-TYPE SEISMOGENIC ZONES. Geochemistry, Geophysics, Geosystems, 12(1), n/a-n/a. https://doi.org/10.1029/2010GC003230

Hollenstein, Ch., Müller, M. D., Geiger, A., & Kahle, H.-G. (2008). Crustal motion and deformation in Greece from a decade of GPS measurements, 1993–2003. Tectonophysics, 449(1–4), 17–40. https://doi.org/10.1016/j.tecto.2007.12.006

Howell, A., Jackson, J., Copley, A., McKenzie, D., & Nissen, E. (2017). Subduction and vertical coastal motions in the eastern Mediterranean. Geophysical Journal International, 211(1), 593–620. https://doi.org/10.1093/gji/ggx307

IGME. (2015). QAFI v.3: Quaternary Faults Database of Iberia. Retrieved October 9, 2018, from http://info.igme.es/QAFI

Johnston, A. C. (1994). Seismotectonic interpretations and conclusions from the stable continental region seismicity database (The Earthquakes of Stable Continental Regions—v. 1 Assessment of Large Earthquake Potential, A. C. Johnston, K. J. Coppersmith, L. R. Kanter, and C. A. Cornell (Editors) No. Report TR102261V1). Palo Alto, California: Electric Power Research Institute.

Jomard, H., Cushing, E. M., Palumbo, L., Baize, S., David, C., & Chartier, T. (2017). Transposing an active fault database into a seismic hazard fault model for nuclear facilities – Part 1: Building a database of potentially active faults (BDFA) for metropolitan France. Natural Hazards and Earth System Sciences, 17(9), 1573–1584. https://doi.org/10.5194/nhess-17-1573-2017

Koulali, A., Ouazar, D., Tahayt, A., King, R. W., Vernant, P., Reilinger, R. E., et al. (2011). New GPS constraints on active deformation along the Africa–Iberia plate boundary. Earth and Planetary Science Letters, 308(1–2), 211–217. https://doi.org/10.1016/j.epsl.2011.05.048

LaFemina, P. C., Dixon, T. H., Malservisi, R., Árnadóttir, T., Sturkell, E., Sigmundsson, F., & Einarsson, P. (2005). Geodetic GPS measurements in south Iceland: Strain accumulation and partitioning in a propagating ridge system: GPS IN SOUTH ICELAND. Journal of Geophysical Research: Solid Earth, 110(B11). https://doi.org/10.1029/2005JB003675

Leonard, M. (2010). Earthquake Fault Scaling: Self-Consistent Relating of Rupture Length, Width, Average Displacement, and Moment Release. Bulletin of the Seismological Society of America, 100(5A), 1971–1988. https://doi.org/10.1785/0120090189

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. https://doi.org/10.1785/0120140087

MacDonald, K. C., & Luyendyk, B. P. (1977). Deep-tow studies of the structure of the Mid-Atlantic Ridge crest near lat 37°N. GSA Bulletin, 88(5), 621–636. https://doi.org/10.1130/0016-7606(1977)88<621:DSOTSO>2.0.CO;2

Maesano, F. E., Tiberti, M. M., & Basili, R. (2017). The Calabrian Arc: three-dimensional modelling of the subduction interface. Sci Rep, 7(1), 8887. https://doi.org/10.1038/s41598-017-09074-8

Martínez-Loriente, S., Gràcia, E., Bartolome, R., Sallarès, V., Connors, C., Perea, H., et al. (2013). Active deformation in old oceanic lithosphere and significance for earthquake hazard: Seismic imaging of the Coral Patch Ridge area and neighboring abyssal plains (SW Iberian Margin). Geochemistry, Geophysics, Geosystems, 14(7), 2206–2231. https://doi.org/10.1002/ggge.20173

Martínez-Loriente, S., Gràcia, E., Bartolome, R., Perea, H., Klaeschen, D., Dañobeitia, J. J., et al. (2018). Morphostructure, tectono-sedimentary evolution and seismic potential of the Horseshoe Fault, SW Iberian Margin. Basin Research, 30, 382–400. https://doi.org/10.1111/bre.12225

Müller, R. D., Sdrolias, M., Gaina, C., & Roest, W. R. (2008). Age, spreading rates, and spreading asymmetry of the world's ocean crust: DIGITAL MODELS OF THE WORLD'S OCEAN CRUST. Geochemistry, Geophysics, Geosystems, 9(4), n/a-n/a. https://doi.org/10.1029/2007GC001743

Neres, M., Carafa, M. M. C., Fernandes, R. M. S., Matias, L., Duarte, J. C., Barba, S., & Terrinha, P. (2016). Lithospheric deformation in the Africa-Iberia plate boundary: Improved neotectonic modeling testing a basal-driven Alboran plate: NEOTECTONIC MODELING IN AFRICA-IBERIA. Journal of Geophysical Research: Solid Earth, 121(9), 6566–6596. https://doi.org/10.1002/2016JB013012

NOAA National Geophysical Data Center. (2009). ETOPO1 1 arc-minute global relief model. NOAA National Centers for Environmental Information. Retrieved from https://www.ngdc.noaa.gov/mgg/global/

Nocquet, J.-M. (2012). Present-day kinematics of the Mediterranean: A comprehensive overview of GPS results. Tectonophysics, 579, 220–242. https://doi.org/10.1016/j.tecto.2012.03.037

Palano, M., González, P. J., & Fernández, J. (2015). The Diffuse Plate boundary of Nubia and Iberia in the Western Mediterranean: Crustal deformation evidence for viscous coupling and fragmented lithosphere. Earth and Planetary Science Letters, 430, 439–447. https://doi.org/10.1016/j.epsl.2015.08.040

Pastor, A., Babault, J., Owen, L. A., Teixell, A., & Arboleya, M.-L. (2015). Extracting dynamic topography from river profiles and cosmogenic nuclide geochronology in the Middle Atlas and the High Plateaus of Morocco. Tectonophysics, 663, 95–109. https://doi.org/10.1016/j.tecto.2015.06.007

Reilinger, R., McClusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., et al. (2006). GPS constraints on continental deformation in the Africa-Arabia-Eurasia continental collision zone and implications for the dynamics of plate interactions. Journal of Geophysical Research: Solid Earth, 111(B5), n/a-n/a. https://doi.org/10.1029/2005JB004051

Rigby, M. (2008). Recent faulting and active shortening of the Middle Atlas Mountains, Morocco, within the diffuse African-Eurasian plate boundary (Master of Science Degree Thesis). Faculty of the Graduate School, University of Missouri-Columbia.

Rögnvaldsson, S. T., Gudmundsson, A., & Slunga, R. (1998). Seismotectonic analysis of the Tjörnes Fracture Zone, an active transform fault in north Iceland. Journal of Geophysical Research: Solid Earth, 103(B12), 30117–30129. https://doi.org/10.1029/98JB02789

Rust, D., & Whitworth, M. (2019). A unique ~12 ka subaerial record of rift-transform triple-junction tectonics, NE Iceland. Scientific Reports, 9(1), 9669. https://doi.org/10.1038/s41598-019-45903-8

Sachpazi, M., Laigle, M., Charalampakis, M., Diaz, J., Kissling, E., Gesret, A., et al. (2016). Segmented Hellenic slab rollback driving Aegean deformation and seismicity. Geophysical Research Letters, 43(2), 651–658. https://doi.org/10.1002/2015GL066818

Sallarès, V., & Ranero, C. R. (2019). Upper-plate rigidity determines depth-varying rupture behaviour of megathrust earthquakes. Nature, 576(7785), 96–101. https://doi.org/10.1038/s41586-019-1784-0

Sallarès, V., Gailler, A., Gutscher, M.-A., Graindorge, D., Bartolomé, R., Gràcia, E., et al. (2011). Seismic evidence for the presence of Jurassic oceanic crust in the central Gulf of Cadiz (SW Iberian margin). Earth and Planetary Science Letters, 311(1–2), 112–123. https://doi.org/10.1016/j.epsl.2011.09.003

Sanz de Galdeano, C., Azañón, J. M., Cabral, J., Ruano, P., Alfaro, P., Canora, C., et al. (2020). Active Faults in Iberia. In C. Quesada & J. T. Oliveira (Eds.), The Geology of Iberia: A Geodynamic Approach (pp. 33–75). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-10931-8_4

Scala, A., Lorito, S., Romano, F., Murphy, S., Selva, J., Basili, R., et al. (2020). Effect of Shallow Slip Amplification Uncertainty on Probabilistic Tsunami Hazard Analysis in Subduction Zones: Use of Long-Term Balanced Stochastic Slip Models. Pure and Applied Geophysics, 177(3), 1497–1520. https://doi.org/10.1007/s00024-019-02260-x

Stich, D., Serpelloni, E., de Lis Mancilla, F., & Morales, J. (2006). Kinematics of the Iberia–Maghreb plate contact from seismic moment tensors and GPS observations. Tectonophysics, 426(3–4), 295–317. https://doi.org/10.1016/j.tecto.2006.08.004

Styron, R., & Pagani, M. (2020). The GEM Global Active Faults Database. Earthquake Spectra, 36(1_suppl), 160–180. https://doi.org/10.1177/8755293020944182

Syracuse, E. M., van Keken, P. E., & Abers, G. A. (2010). The global range of subduction zone thermal models. Physics of the Earth and Planetary Interiors, 183(1–2), 73–90. https://doi.org/10.1016/j.pepi.2010.02.004

Thiebot, E., & Gutscher, M.-A. (2006). The Gibraltar Arc seismogenic zone (part 1): Constraints on a shallow east dipping fault plane source for the 1755 Lisbon earthquake provided by seismic data, gravity and thermal modeling. Tectonophysics, 426(1–2), 135–152. https://doi.org/10.1016/j.tecto.2006.02.024

Vanneste, K., Camelbeeck, T., & Verbeeck, K. (2013). A Model of Composite Seismic Sources for the Lower Rhine Graben, Northwest Europe. Bulletin of the Seismological Society of America, 103(2A), 984–1007. https://doi.org/10.1785/0120120037

Wdowinski, S., Ben-Avraham, Z., Arvidsson, R., & Ekström, G. (2006). Seismotectonics of the Cyprian Arc. Geophysical Journal International, 164(1), 176–181. https://doi.org/10.1111/j.1365-246X.2005.02737.x

van der Woerd, J., Dorbath, C., Ousadou, F., Dorbath, L., Delouis, B., Jacques, E., et al. (2014). The Al Hoceima Mw 6.4 earthquake of 24 February 2004 and its aftershocks sequence. Journal of Geodynamics, 77, 89–109. https://doi.org/10.1016/j.jog.2013.12.004

**Ground Motion Models**

Abrahamson NA, Gregor N, Addo K (2016) BC Hydro Ground Motion Prediction Equations for Subduction Earthquakes. Earthquake Spectra. 32(1): 23 – 44

Abrahamson NA, Keuhn N, Gülerce Z, et al. (2018) Update of the BC Hydro Subduction Ground-Motion Model using the NGA-Subduction Dataset. PEER Technical Report 2018/02.

Abrahamson N and Gülerce Z (2020) Regionalized Ground-Motion Models for Subduction Earthquakes Based on the NGA-SUB Database. PEER Technical Report No. 2020/25

Al Atik L and Youngs RR (2014) Epistemic uncertainty for NGA-West2 models. Earthquake Spectra. 30: 1301 – 1318.

Al Atik L (2015) NGA-East: ground-motion standard deviation models for Central and Eastern North America. PEER Technical Report, Report No. 2015/07

Basili R, Brizuela B., Herrero A, et al. (2019) NEAMTHM18 documentation: the making of the TSUMAPS-NEAM. In Tech. rep., InIstituto Nazionale di Geofisica e Vulcanologia (INGV). DOI: https ://doi.org/10.5281/zenodo.34066 25N

Boore DM (2010) Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. Bulletin of the Seismological Society of America. 100: 1830 – 1835

Fülöp L, Jussila V, Aapasuo R, et al. (2020) A ground-motion prediction equation for Fennoscandian Nuclear Installations. Bulletin of the Seismological Society of America. 110(3): 1211 – 1230.

Ghasemi H, Cummins P, Weatherill G, et al. (2020) Seismotectonic model and probabilistic seismic hazard assessment for Papua New Guinea. Bulletin of Earthquake Engineering. 18: 6571 – 6605.

Goulet C, Bozorgnia Y, Abrahamson NA, et al. (2018) Central and Eastern North America ground-motion characterization: NGA-East Final Report. PEER Technical Report No. 2018/08.

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

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

Kuehn N, Bozorgnia Y, Campbell KW, Gregor N (2020) Partially Non-Ergodic Ground-Motion Model for Subduction Regions using the NGA-Subduction Database. PEER Technical Report 2020/04.

Kotha S -R, Bindi D, Cotton F. (2016) Partially non-ergodic region specific GMPE for Europe and the Middle East. Bulletin of Earthquake Engineering. 14: 1245 - 1263

Kotha S -R, Weatherill G, Cotton F (2020a) A regionally adaptable ground-motion model for shallow crustal earthquakes in Europe. Bulletin of Earthquake Engineering. 18:4091 – 4125

Kotha S -R (2020) From a Regionalised Ground-Motion Model for Europe and the Middle East to Site-specific Seismic Hazard Assessments in Low-to-Moderate Seismicity Regions. Sigma 2 Deliverable SIGMA2-2019-D3-029/1.

Kotha S -R, Weatherill G, Bindi D, Cotton F (2022) Near-source Magnitude Scaling of Spectral Accelerations: Analysis and Update of the Kotha et al. (2020) Model. Bulletin of Earthquake Engineering, submitted.

Lanzano G, Sgobba S, Luzi L, et al. (2019) The pan-European Engineering Strong Motion (ESM) flatfile: compilation criteria and data statistics. Bulletin of Earthquake Engineering. 17: 561 – 582

Lucazeau F (2019) Analysis and mapping of an updated terrestrial heat flow data set. Geochemistry, Geophysics and Geosystems. 20: 4001 – 4024

Mak S, Clements RA, Schorlemmer D (2017) Empirical Evaluation of Hierarchical Ground-Motion Models: Score Uncertainty and Model Weighting. Bulletin of the Seismological Society of America. 107(2): 949 – 965.

Miller AC and Rice TR (1983) Discrete approximations of probability distributions. Management Science. 29(3): 352 – 362

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

Mooney WD, Ritsema J, Hwang YK (2012) Crustal seismicity and the earthquake catalog maximum moment magnitude mcmax in stable continental regions (SCRs): correlation with the seismic velocity of the lithosphere. Earth and Planetary Science Letters. 357 – 358: 78 – 83.

Parker GA, Stewart JP, Boore DM, et al. (2020) NGA-Subduction Global Ground-Motion Models with Regional Adjustment Factors. PEER Technical Report, Report No. 2020/03.

Pezeshk S, Zandieh A, Tavakoli B (2011) Hybrid empirical ground-motion prediction equations for eastern North America using NGA models and updated seismological parameters. Bulletin of the Seismological Society of America. 101(4): 1859 – 1870.

Rodriguez-Marek A, Cotton F, Abrahamson NA, et al. (2013) A model for single-station standard deviation using data from various tectonic regimes. Bulletin of the Seismological Society of America. 103(6): 3149 - 3163
Scasserra G, Stewart J, Bazzurro P, et al. (2009) A comparison of NGA ground-motion prediction equations to Italian data. Bulletin of the Seismological Society of America. 99(5): 2961 – 2978
Stewart JP, Parker GA, Atkinson GM, et al. (2020) Ergodic site amplification model for Central and Eastern North America. Earthquake Spectra. 36(1): 42 – 68.
Szwillus W, Alfonso JB, Ebbing J, Mooney WD (2019) Global crustal thickness and velocity structure from geostatistical analysis of seismic data. Journal of Geophysical Research Solid Earth. 124: 1626 – 1652.
Vacareanu R, Radulian M, Iancovici M, et al. (2015) Fore-Arc and Back-Arc Ground Motion Prediction Model for Vrancea Intermediate Depth Seismic Source. Journal of Earthquake Engineering. 19: 535 - 562
Weatherill G, Kotha S -R, Cotton F. (2020a) A regionally adaptable "scaled backbone" ground motion logic tree for shallow seismicity in Europe: application to the 2020 European seismic hazard model. Bulletin of Earthquake Engineering. 18: 5087 – 5117
Weatherill GA and Cotton F (2020) A ground motion logic tree for seismic hazard analysis in the stable cratonic region of Europe: regionalization, model selection and development of a scaled backbone approach. Bulletin of Earthquake Engineering. 18: 6119 – 6148

**Unified Earthquake Catalogue**

Gardner, J. K. and Knopoff, L. (1974). Is the sequence of aftershocks in Southern California, with aftershocks removed, poissonian?. Bull. Seism. Soc. Am., 64(5): 1363-1367.
Grünthal, G., Wahlström, R. and Stromeyer D. (2013). The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900-2006 and its comparison to the European-Mediterranean Earthquake Catalogue (EMEC). Journal of Seismology, 17, 1339-1344, doi 10.1007/s10950-013-9379-y
Grünthal G., Wahlström R., Stromeyer, D. (2009a), The unified catalogue of earthquakes in central, northern and northwestern Europe (CENEC) – updated and expanded to the last millennium, Journal of Seismology, 13(4), 514-541.
Mignan, A., M. J. Werner, S. Wiemer, C.-C. Chen, and Y.-M. Wu (2011), Bayesian estimation of the spatially varying completeness magnitude of earthquake catalogs, Bull. Seismol. Soc. Am., 101, doi:10.1785/0120100223.
Ogata, Y., and K. Katsura (1993), Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, Geophys. J. Int., 113, 727–738.
Zaliapin, I., A. Gabrielov, V. Keilis-Borok, and H. Wong (2008), Clustering Analysis of Seismicity and Aftershock Identification, Phys. Rev. Lett., 101(1), 1–4. 6, 22, 23

**Seismogenic Source Model**

Anderson JG, Luco JE (1983) Consequences of slip rate constraints on earthquake occurrence relations. Bull Seismol Soc Am 73:471–496.
Basili R, Valensise G, Vannoli P et al (2008) The Database of Individual Seismogenic Sources (DISS), version 3: summarizing 20 years of research on Italy's earthquake geology. Tectonophysics 453:20–43. doi:10.1016/j.tecto.2007.04.014
Hiemer, S., J. Woessner,R. Basili, L. Danciu, D. Giardini and S. Wiemer, (2014), A smoothed stochastic earthquake rate model consid- ering seismicity and fault moment release for Europe, Geophys. J. Int., 198, 1159–1172, doi: 10.1093/gji/ggu186.
Hiemer, S., Jackson, D.D.,Wang, Q., Kagan, Y.Y.,Woessner, J., Zechar, J.D. &Wiemer, S., (2013). A stochastic forecast of California earthquakes based on fault slip and smoothed seismicity, Bull. seism. Soc. Am., 103(2A), 799–810.
Kagan, Y. Y. (2002), Seismic moment distribution revisited: I. statistical results, Geophys. J. Int., 148, 520–541.
Maesano F.E., Tiberti M.M. and Basili R. (2017), The Calabrian Arc: three-dimensional modelling of the subduction interface, Scientific Reportsvolume 7, Article number: 8887 (2017), http://hdl.handle.net/2122/11221
Musson, R M W. 2011. The Effect Of Magnitude Uncertainty On Activity Rates. Bulletin of The Seismological Society Of America, Submitted.
Pagani M., J. Garcia-Pelaez, R. Gee, K. Johnson, V. Poggi, R. Styron, G. Weatherill, M. Simionato, D. Viganò, L. Danciu, D. Monelli, 10.13117/GEM-GLOBAL-SEISMIC-HAZARD-MAP-2018.1, The Global Earthquake Model (GEM) Global Seismic Hazard Map (version 2018.1)
Reasenberg, P., 1985. Second-order moment of central California seismicity, 1969–1982. Journal of Geophysical Research: Solid Earth, 90(B7), pp.5479-5495.
Rhoades, D. A., Schorlemmer, D., Gerstenberger, M. C., Christophersen, A., Zechar, J. D., and Imoto, M. (2011). Efficient testing of earthquake forecasting models. ActaGeophysica, 59(4):728–747.
Rydelek, P. A., and I. S. Sacks (1989), Testing the completeness of earthquake catalogs and the hypothesis of self-similarity, Nature, 337, 251–253.
Schorlemmer, D., and J. Woessner (2008), Probability of detecting an earthquake, Bull. Seismol. Soc. Am., 98, doi:10.1785/0120070105.
Shi, Y., and B. A. Bolt (1982), The standard error of the magnitude-frequency b-value, Bull. Seismol. Soc. Am., 72, 1677–1687.
Stucchi, M. et al., (2012). The SHARE European Earthquake Catalogue (SHEEC) 1000–1899, J. Seismol., 17(2), 523–544.
Uhrhammer R (1986) "Characteristics of northern and southern California seismicity: Earthquake Notes", v. 57, p. 21.
van Stiphout, T., J. Zhuang, and D. Marsan (2012), Seismicity declustering, Community Online Resource for Statistical Seismicity Analysis, doi:10.5078/corssa-52382934. Available at http://www.corssa.org
Wells DL, Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. Bull Seismol Soc Am 84:974–1002
Weichert, D H. 1980. Estimation of The Earthquake Recurrence Parameters for Unequal Observation Periods For Different Magnitudes. Bulletin of The Seismological Society of America, Vol. 70, 1337- 1346.
Wössner, J., and S. Wiemer (2005), Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty, Bull. Seismol. Soc. Am., 95, doi:10.1785/012040007.
Wössner, J., D. Laurentiu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili and M. B. Demircioglu (2015). The 2013 European Seismic Hazard Model: Key Components and Results, B Earthq Eng 13, 3553-3596
Youngs RR, Coppersmith KJ, (1985) Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. Bull Seismol Soc Am 75: 939–964.

**Model Implementation**

Crowley H., Monelli D., Pagani M., Silva V. , Weatherill G., (2015), OpenQuake Engine User Instruction Manual
Monelli D, Pagani M, Weatherill G, Danciu L, Garcia J (2014) Modeling Distributed Seismicity for Probabilistic Seismic-Hazard Analysis: Implementation and Insights with the OpenQuake Engine. Bull Seismol Soc Am 104(4): 1636–1649. doi: 10.1785/0120130309
Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, L. Butler, M. Nastasi, L. Panzeri, M. Simionato, and D. Vigano, OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model, Seismological Research Letters, May/June 2014, v. 85, p. 692-702, doi:10.1785/0220130087
Pagani, M., Monelli, D., Weatherill, G. A. and Garcia, J. (2014). The OpenQuake-engine Book: Hazard. Global Earthquake Model (GEM) Technical Report 2014-08, doi: 10.13117/-GEM.OPENQUAKE.TR2014.08, 67 pages.
Weatherill GA (2014) OpenQuake Hazard Modeller's Toolkit - User Guide. Global Earthquake Model (GEM). Technical Report.

# Appendix A: Earthquake Catalogue Completeness – Summary

In this appendix the time intervals of completion for all completeness super-zones (CSZs) are provided. These magnitude -time completeness intervals were used together with the declustered catalogue to estimate the activity parameters of the seismogenic source model.

Table A.1: Magnitude of completeness table for all CSZ

| CSZs ID | Completeness time steps (Year and Mc) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| SZ01 | 1152 | 1590 | 1670 | 1691 | 1794 | 1814 | 1875 | | |
| | 6.1 | 5.7 | 5.1 | 4.9 | 4.3 | 4.1 | 3.7 | | |
| SZ02 | 1295 | 1524 | 1729 | 1774 | 1821 | 1835 | 1846 | | |
| | 5.5 | 5.1 | 4.9 | 4.3 | 3.9 | 3.7 | 3.5 | | |
| SZ03 | 1000 | 1828 | 1949 | | | | | | |
| | 5.5 | 4.5 | 4.1 | | | | | | |
| SZ04 | 1479 | 1822 | 1857 | | | | | | |
| | 5.5 | 4.5 | 3.7 | | | | | | |
| SZ05 | 1014 | 1373 | 1579 | 1708 | 1730 | 1747 | 1772 | 1812 | 1826 |
| | 6.1 | 5.7 | 5.3 | 5.1 | 4.3 | 4.1 | 3.9 | 3.7 | 3.5 |
| SZ06 | 1133 | 1795 | 1816 | 1839 | | | | | |
| | 4.7 | 4.1 | 3.7 | 3.5 | | | | | |
| SZ07 | 1092 | 1753 | | | | | | | |
| | 5.5 | 3.5 | | | | | | | |
| SZ08 | 1125 | 1693 | 1828 | 1972 | | | | | |
| | 6.3 | 5.1 | 4.9 | 4.1 | | | | | |
| SZ10 | 1852 | 1926 | | | | | | | |
| | 4.3 | 4.1 | | | | | | | |
| SZ11 | 1517 | 1829 | 1847 | 1871 | 1887 | | | | |
| | 5.5 | 5.1 | 4.9 | 4.1 | 3.7 | | | | |
| SZ12 | 1329 | 1676 | 1700 | 1781 | 1844 | | | | |
| | 5.5 | 5.1 | 4.9 | 4.1 | 3.7 | | | | |
| SZ13 | 1508 | 1684 | 1813 | 1825 | 1838 | 1852 | | | |
| | 5.3 | 5.1 | 4.7 | 4.1 | 3.9 | 3.7 | | | |
| SZ15 | 1497 | 1819 | 1857 | | | | | | |
| | 4.9 | 4.3 | 3.5 | | | | | | |
| SZ16 | 1671 | 1851 | 1906 | 1951 | | | | | |
| | 6.5 | 4.9 | 4.7 | 4.3 | | | | | |
| SZ17 | 1509 | 1818 | 1858 | 1890 | | | | | |
| | 7.1 | 5.5 | 4.9 | 4.1 | | | | | |
| SZ18 | 1444 | 1837 | 1853 | 1893 | | | | | |
| | 6.9 | 5.1 | 4.9 | 4.3 | | | | | |
| SZ19 | 1308 | 1862 | 1882 | 1916 | 1963 | | | | |
| | 5.7 | 5.1 | 4.1 | 3.7 | 3.5 | | | | |
| SZ21 | 1073 | 1828 | 1839 | 1884 | 1947 | 1965 | | | |
| | 6.9 | 5.7 | 5.3 | 5.1 | 4.9 | 4.3 | | | |
| SZ22 | 1204 | 1791 | 1839 | 1859 | 1886 | 1911 | 1952 | 1963 | |
| | 7.3 | 6.5 | 5.7 | 5.5 | 5.1 | 4.9 | 4.7 | 4.3 | |
| SZ23 | 1555 | 1870 | 1904 | | | | | | |
| | 6.3 | 5.3 | 4.1 | | | | | | |
| SZ24 | 1236 | 1668 | 1868 | 1895 | 1947 | 1966 | | | |
| | 7.7 | 6.5 | 6.1 | 5.3 | 5.1 | 4.3 | | | |

| CSZs ID | Completeness time steps (Year and Mc) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SZ25 | 1182 | 1903 | 1924 | 1973 | | | | |
| | 6.7 | 5.1 | 4.9 | 4.1 | | | | |
| SZ26 | 1829 | 1872 | 1896 | 1932 | | | | |
| | 7.3 | 6.1 | 5.9 | 5.3 | | | | |
| SZ27 | 1816 | 1926 | 1939 | 1958 | 1979 | | | |
| | 8.3 | 6.7 | 6.3 | 5.7 | 5.3 | | | |
| SZ28 | 1915 | 1972 | | | | | | |
| | 5.1 | 4.9 | | | | | | |
| SZ29 | 1899 | 1928 | 1967 | 1981 | | | | |
| | 7.3 | 5.9 | 5.3 | 4.5 | | | | |
| SZ30 | 1826 | 1907 | 1983 | | | | | |
| | 6.9 | 5.1 | 4.9 | | | | | |
| SZ31 | 1517 | 1857 | 1875 | | | | | |
| | 6.9 | 5.1 | 4.9 | | | | | |
| SZ32 | 1864 | 1926 | | | | | | |
| | 6.7 | 6.3 | | | | | | |
| SZ33 | 1904 | 1927 | | | | | | |
| | 4.9 | 4.5 | | | | | | |
| SZ34 | 1281 | 1667 | 1843 | 1861 | 1895 | | | |
| | 6.3 | 6.1 | 4.9 | 4.1 | 3.9 | | | |
| SZ36 | 1005 | 1542 | 1619 | 1667 | 1829 | 1853 | 1866 | |
| | 6.9 | 5.7 | 5.3 | 5.1 | 4.5 | 4.3 | 4.1 | |
| SZ37 | 1290 | 1456 | 1561 | 1624 | 1784 | 1880 | 1976 | |
| | 7.1 | 6.9 | 5.5 | 5.1 | 4.9 | 4.3 | 3.9 | |
| SZ38 | 1809 | 1935 | 1952 | 1979 | | | | |
| | 5.3 | 5.1 | 4.9 | 4.7 | | | | |
| SZ39 | 1821 | 1892 | 1938 | 1968 | | | | |
| | 5.1 | 4.5 | 4.3 | 4.1 | | | | |
| SZ40 | 1065 | 1501 | 1688 | 1738 | 1779 | 1857 | | |
| | 6.3 | 5.9 | 5.5 | 5.1 | 4.9 | 4.1 | | |
| SZ41 | 1530 | 1850 | 1900 | | | | | |
| | 7.3 | 4.9 | 4.7 | | | | | |
| SZ42 | 1052 | 1362 | 1789 | 1857 | 1870 | 1900 | 1970 | |
| | 7.1 | 6.9 | 6.7 | 5.9 | 5.3 | 4.9 | 4.5 | |
| SZ43 | 1190 | 1717 | 1900 | | | | | |
| | 6.5 | 5.9 | 4.9 | | | | | |
| SZ44 | 1003 | 1911 | | | | | | |
| | 6.3 | 5.1 | | | | | | |
| SZ45 | 1023 | 1549 | 1641 | 1828 | 1877 | 1892 | 1903 | 1962 |
| | 7.7 | 6.9 | 6.5 | 5.9 | 5.3 | 5.1 | 4.5 | 4.3 |
| SZ46 | 1045 | 1255 | 1866 | 1900 | | | | |
| | 7.9 | 7.3 | 6.9 | 4.9 | | | | |
| SZ47 | 1182 | 1753 | 1767 | 1806 | 1866 | | | |
| | 6.1 | 4.9 | 4.7 | 4.3 | 4.1 | | | |
| SZ48 | 1923 | 1938 | 1950 | | | | | |
| | 5.9 | 4.5 | 4.3 | | | | | |
| SZ49 | 1632 | 1759 | 1834 | 1865 | 1879 | | | |
| | 5.5 | 5.3 | 4.9 | 4.7 | 3.7 | | | |
| SZ51 | 1309 | 1722 | 1962 | | | | | |
| | 8.3 | 5.9 | 4.1 | | | | | |

# Appendix B: ESHM20 Input Datasets: Summary and File Formats

This appendix presents the data summary tables describing the main input datasets used in the development of the seismogenic source model. The aim of these table is to assist the documentation of the seismogenic source model and its conversion to OpenQuake input files. The main databases that we consider are:

1. The earthquake catalogue used as basis for the computation.
2. Superzones of completeness regions
3. Superzones of the maximum magnitudes
4. Superzones of tectonic regimes
5. The parameters used for the computation of the area source model branches
6. The parameters of the fault source and background model
7. The parameters of the smoothed seismicity model

We provide the parameters of these datasets in tables when possible. The data used to create the ESHM20 input files for OpenQuake are provided as Environmental System Research Institute (ESRI) shapefiles (ESRI,1998), which is readable by programs such as ArcGIS, QGIS, and the GeoPandas Python module. Shapefile fieldnames must be 10 characters or less in length, and are limited in to text or a single numerical value per field.

These shape files of the ESHM20 input datasets reside on the gitlab repository of the hazard.EFEHR.org https://gitlab.seismo.ethz.ch/efehr/eshm20_source_model/-/tree/master/input_shapefiles.

The interested reader is encouraged to use this online repository to examine the input shape files. For the purpose of documentation, we describe the attributes of each table following the structure of the main directory:

1. eshm20_input_a_unified_eq_catalogue (see B.1)
2. eshm20_input_b_completeness_superzones (see B.2)
3. eshm20_input_c_tecto_zones (see B.3)
4. eshm20_input_d_asm_shallow_crust (see B.4)
5. eshm20_input_e_fs_model (see B.5)
6. eshm20_input_f_smoothed_seismicity (see B.6)
7. eshm20_input_g_calc_regions_nrml_filter (see B.7)
8. eshm20_input_h_simple_individual_buffer (see B.8)
9. eshm20_input_i_tecto_iceland (see B.3)
10. eshm20_input_j_asz_iceland (see B.4)
11. eshm20_input_k_tecto_deep (see B.3)
12. eshm20_input_m_asm_deep (see B.4)
13. eshm20_input_m_asz_volcanic(see B.4)

## B.1 Catalogue file:

| Attribute | Definition |
|---|---|
| **longitude** | Epicentral longitude of the event |
| **latitude** | Epicentral latitude of the event |
| **year** | year of the event |
| **month** | month of the event |
| **day** | day of the event |
| **hour** | hour of the event |
| **minute** | minute of the event |
| **second** | second of the event |
| **magnitude** | Magnitude, |
| **depth** | Hypocentral depth, in km. If unknown, randomly assigned according to tectonic region |
| **source_cat** | Catalogue from which this event was imported: {'EMECv20191114', 'EMME14_Pre1900', 'SERA_histv1.1'} |
| **eventID** | Generated ID for each event |
| **depth_orig** | Depth, in km, from original catalogue. |
| **winGT_fs01** | is Background seismicity after Gruenthal declustering |
| **winGK_fs01** | is Background seismicity after Gardiner-Knopoff declustering |
| **rb_rfact10** | is Background seismicity after Reasenberg declustering, rfact 10, tmax100, fs017, |
| **rb_rfact20** | is Background seismicity after Reesenberg declustering, rfact 20, tmax100, fs017 |
| **rb_rfact30** | is Background seismicity after Reasenberg declustering, rfact 30, tmax100, fs017 |
| **Date** | string date representation mostly "YYYY-MM-DD", but 'd-m-y' for EMME14_Pre1900 events |
| **geometry** | Point representation of the earthquake epicenter |

*Table B..1 Homogenized catalogue shapefile attributes*

## B2 Completeness Superzones (CSZ) shape file

| ZONE_ID | ID of source area |
|---|---|
| **CSZ_ID** | ID of Completeness Zonation (CSZ) source |
| **CSZ_NAME** | NAME of Completeness Zonation (CSZ) source |
| **CSZ_ID13** | ID of Completeness Zonation (CSZ) source of the ESHM13 |
| **CSZ_NAME13** | ID of Completeness Zonation (CSZ) source of the ESHM13 |
| **a_val** | Gutenberg-Richter a-value - best estimate |
| **a_hi** | Gutenberg-Richter a-value - upper curve |
| **a_mid** | Gutenberg-Richter a-value - median curve |
| **a_lo** | Gutenberg-Ricther a-value - lower curve |
| **b_val** | Gutenberg-Richter b-value - best estimate |
| **b_hi** | Gutenberg-Richter b-value - upper curve |
| **b_mid** | Gutenberg-Richter b-value - median curve |
| **b_lo** | Gutenberg-Ricther b-value - lower curve |
| **sigma_a** | Std Dev. in a-values |
| **sigma_b** | Std Dev. in b-values |
| **sigma_b** | Std Dev. in b-values |
| **cum_cov_ab** | Covariance between a-values and b-values |
| **areakm2** | Area of source zone in km^2 |

*Table B.2 Completeness superzone (CSZ) shapefile attributes*

## B.3 Tectonic Zone (TECTO) shape files

| ZONE_ID | ID of source area |
|---|---|
| **areakm2** | Area of source zone in km^2 |
| **CSZ_ID** | ID of Completeness Zonation (CSZ) source |
| **TECTO_ID** | ID of Tectonic Region (TECTO) source, same as ZONE_ID |
| **region_lev** | "TECTO" reflecting that these zones belong to the TECTO regime |
| **MAXMAG01** | Lower maximum magnitude for Gutenberg-Richter models |
| **MAXMAG02** | Middle maximum magnitude for Gutenberg-Richter models |
| **MAXMAG03** | Upper maximum magnitude for Gutenberg-Richter models |
| **a_val** | Gutenberg-Richter a-value - best estimate |
| **a_hi** | Gutenberg-Richter a-value - upper curve |
| **a_mid** | Gutenberg-Richter a-value - median curve |
| **a_lo** | Gutenberg-Richter a-value - lower curve |
| **sigma_a** | Std Dev. in a-values |
| **b_val** | Gutenberg-Richter b-value - best estimate |
| **b_hi** | Gutenberg-Richter b-value - upper curve |
| **b_mid** | Gutenberg-Richter b-value - median curve |
| **b_lo** | Gutenberg-Richter b-value - lower curve |
| **sigma_b** | Std Dev. in b-values |
| **ref_mag** | Magnitude at which the occurrence observation table starts |
| **cum_cov_ab** | Covariance between a-values and b-values |
| **mag_corner** | Corner magnitude for Pareto model |
| **dx_corner** | difference between each corner bin, for pareto models |
| **MAXMAGP1** | Effective Max Magnitude for Pareto model with lowest corner magnitude |
| **MAXMAGP2** | Effective Max Magnitude for Pareto model with corner magnitude of (mag_corner + dx_corner) |
| **MAXMAGP3** | Effective Max Magnitude for Pareto model with corner magnitude of (mag_corner + 2 * dx_corner) |
| **utsu_bval1** | b-value for magnitude bins below the inflection point magnitude utsu_Minfl, for use with UTSU (Twin-GR) model |
| **utsu_bval2** | b-value for magnitude bins greater than the inflection point magnitude utsu_Minfl, for use with UTSU (Twin-GR) model |
| **utsu_aval** | b-value for use with UTSU (Twin-GR) model |
| **utsu_Minfl** | Magnitude of inflection point. Magnitude at which the B-value changes from utsu_bval1 to utsu_bval2, for use with UTSU (Twin-GR) model |
| **geometry** | Polygon that defines this zone |

*Table B.3 Tectonic zone (TECTO) shapefile attributes*

## B.4 Area Source Zone (ASZ) shape files

| Attribute | Definition |
|---|---|
| **ZONE_ID** | ID of source area |
| **areakm2** | Area of source zone in km^2 |
| **tctCluster** | Tectonic Region Type |
| **CSZ_ID** | ID of Completeness Zonation (CSZ) source |
| **TECTO_ID** | ID of Tectonic Region (TECTO) source |
| **region_lev** | "ASZ" reflecting that these zones belong to the ASZ regime |
| **magScaleRe** | Magnitude-area rupture scaling relationship |
| **MC_ID** | ID of Completeness Zonation (CSZ) source, used for determining Magnitude of Completeness |
| **SS** | Percentage of Strike-Slip component |
| **NF** | Percentage of Normal Fault component |
| **TF** | Percentage of Thrust Fault component |
| **MINDEPTH** | Upper depth of earthquakes in the source zone in km |
| **MAXDEPTH** | Lower depth of earthquakes in the source zone in km |
| **HYPODEPTH1** | Shallow Depth hypocenter in km |
| **HYPODEPTH2** | Mid Depth hypocenter in km |
| **HYPODEPTH3** | Lower depth hypocenter in km |
| **WHDEPTH1** | Weight applied to HYPODEPTH1 |
| **WHDEPTH2** | Weight applied to HYPODEPTH2 |
| **WHDEPTH3** | Weight applied to HYPODEPTH3 |
| **WMAXMAG01** | Weighting for MAXMAG01 |
| **WMAXMAG02** | Weighting for MAXMAG02 |
| **WMAXMAG03** | Weighting for MAXMAG03 |
| **MAXMAG01** | Lower maximum magnitude for Gutenberg-Richter models |
| **MAXMAG02** | Middle maximum magnitude for Gutenberg-Richter models |
| **MAXMAG03** | Upper maximum magnitude for Gutenberg-Richter models |
| **STRIKE1** | Strike in degrees from north, used with DIP1 |
| **STRIKE2** | Strike in degrees from north, used with DIP2 |
| **STRIKE3** | Strike in degrees from north, used with DIP3 |
| **STRIKE4** | Strike in degrees from north, used with DIP4 |
| **DIP1** | Rupture dip in degrees, used with STRIKE1 |
| **DIP2** | Rupture dip in degrees, used with STRIKE2 |
| **DIP3** | Rupture dip in degrees, used with STRIKE3 |
| **DIP4** | Rupture dip in degrees, used with STRIKE4 |
| **aspRatio** | Aspect ratio for pseudo-rupture length-width scaling |
| **a_val** | Gutenberg-Richter a-value - best estimate |
| **a_hi** | Gutenberg-Richter a-value - upper curve |
| **a_mid** | Gutenberg-Richter a-value - median curve |
| **a_lo** | Gutenberg-Ricther a-value - lower curve |
| **sigma_a** | Std Dev. in a-values |
| **b_val** | Gutenberg-Richter b-value - best estimate |
| **b_hi** | Gutenberg-Richter b-value - upper curve |
| **b_mid** | Gutenberg-Richter b-value - median curve |
| **b_lo** | Gutenberg-Ricther b-value - lower curve |
| **sigma_b** | Std Dev. in b-values |
| **ref_mag** | Magnitude at which the occurrence observation table starts |
| **cum_cov_ab** | Covariance between a-values and b-values |
| **mag_corner** | Corner magnitude for Pareto model |

| | |
|---|---|
| **dx_corner** | difference between each corner bin, for pareto models |
| **MAXMAGP1** | Effective Max Magnitude for Pareto model with lowest corner magnitude |
| **MAXMAGP2** | Effective Max Magnitude for Pareto model with corner magnitude of (mag_corner + dx_corner) |
| **MAXMAGP3** | Effective Max Magnitude for Pareto model with corner magnitude of (mag_corner + 2 * dx_corner) |
| **utsu_bval1** | b-value for magnitude bins below the inflection point magnitude utsu_Minfl, for use with UTSU (Twin-GR) model |
| **utsu_bval2** | b-value for magnitude bins greater than the inflection point magnitude utsu_Minfl, for use with UTSU (Twin-GR) model |
| **utsu_aval** | b-value for use with UTSU (Twin-GR) model |
| **utsu_Minfl** | Magnitude of inflection point. Magnitude at which the B-value changes from utsu_bval1 to utsu_bval2, for use with UTSU (Twin-GR) model |
| **geometry** | Polygon that defines this zone |

*Table B.4 Area source zone (ASZ) shapefile attributes*

## B.5 Active fault source model shapefile

| Attribute | Definition |
|---|---|
| **index** | Id of this fault |
| **AR** | Rupture aspect ratio |
| **rake** | Fault Rake |
| **dip** | Fault Dip |
| **upper_seis** | Upper depth of earthquakes in the fault in km |
| **lower_seis** | Lower depth of earthquakes in the fault in km |
| **a_value** | a-value for the truncated Gutenberg-Richter Model |
| **b_value** | b-value for the truncated Gutenberg-Richter Model |
| **min_mag** | Min Mag for the truncated Gutenberg-Richter Model |
| **max_mag** | Max Mag for the truncated Gutenberg-Richter Model |
| **geometry** | MultiLineSources that specifies the fault trace geometry |

*Table B.5 Fault source model shapefile attributes*

## B.6 Background seismicity SEIFA shapefile

| Attribute | Definition |
|---|---|
| **TECTO_ID** | ID of Tectonic Region (TECTO) source, same as ZONE_ID |
| **FMDtyp** | Model type used at this grid point: tGR for Gutenberg-Richter, twinGR for Utsu |
| **CSZ_ID** | ID of Completeness Zonation (CSZ) source |
| **CSZ_NAME** | Name of the CSZ |
| **areakm2** | Area of source zone in km^2 |
| **X** | Longitude |
| **Y** | Latitude |
| **areagrid** | Area, in km^2 represented by this grid point |
| **tct_a_val** | a-value for the TECTO zone |
| **tct_b_val** | b-value for the TECTO zone |
| **tct_b_val2** | 2nd (upper) b-value for the TECTO zone, if a zone ihas FMDtyp "twinGR" |
| **max_mag** | Maximum magnitude for Gutenberg-Richter models |
| **mag_inflec** | Magnitude of inflection point. Magnitude at which the B-value changes from utsu_bval1 to utsu_bval2, for use with UTSU (Twin-GR) model |
| **bin_width** | magnitude bin width |
| **first_bin** | Center of the first magnitude bin |
| **last_bin** | Center of the last magnitude bin |
| **a_grid** | a-value for this grid point |
| **b_grid** | b-value for this grid point |
| **Mmax_ssmBa** | -- ? -- |
| | |
| **ssm_1 through ssm_30** | All fields "ssm_X" contain the incremental annual occurrence rate for the bin centered at first_bin + (x-1)*bin_width |
| **geometry** | Point representation of this location |

*Table B.6 Background seismicity shapefile attributes*

## B.7 Region Filter shapefile

| Attribute | Definition |
|---|---|
| **calcRegIds** | ID of calculation region |
| **geometry** | polygons representation each region |

*Table B.7 Region filter shapefile attributes*

## B.8 Fault buffer shapefile

| Attribute | Definition |
|---|---|
| **IDFS** | ID of fault |
| **geometry** | Polygon representing fault's region of influence upon the smoothed seismicity model |

*Table B.8 Fault buffer shapefile attributes*