
***METIS***

*Research and Innovation Action (RIA)*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 945121

Start date : 2020-09-01 Duration : 48 Months

# New PSHA methodologies: code developments and documentation

Authors : Dr. Marco PAGANI (GEM), Marco Pagani (GEM), Richard Styron (GEM), Kendra Johnson (GEM), Michele Simionato (GEM)

METIS - Contract Number: 945121

Project officer: Katerina PTACKOVA

| | |
|---|---|
| Document title | New PSHA methodologies: code developments and documentation |
| Author(s) | Dr. Marco PAGANI, Marco Pagani (GEM), Richard Styron (GEM), Kendra Johnson (GEM), Michele Simionato (GEM) |
| Number of pages | 74 |
| Document type | Deliverable |
| Work Package | WP4 |
| Document number | D4.4 |
| Issued by | GEM |
| Date of completion | 2024-02-07 13:19:25 |
| Dissemination level | Public |

**Summary**

This report summarises the results achieved within three tasks of the METIS project Work package 4 ?Seismic Hazard?. The first task, named ?Approaches for the Propagation of Epistemic Uncertainty? had the goal of improving the approaches currently adopted for the processing of epis- temic uncertainties in Probabilistic Seismic Hazard Analysis (PSHA). We expanded the capabilities of the OQ Engine to process logic tree by adding the option of us- ing a Latin Hypercube sampling approach in lieu of the more traditional Monte Carlo approach and we added the possibility - while processing the logic tree - of deciding where to apply the weights assigned to the various realizations admitted. We also developed a completely new approach for the propagation of epistemic uncertainties that computes seismic hazard for each individual source and combines the results us- ing discrete distributions via a convolution approach. According to the tests completed so far this new method is computationally efficient and it provides results consistent with the ones provided by more traditional approaches. In the second task titled ?Vector-valued PSHA and CS Approach? we implemented in the OQ Engine the most complete, rigorous and complex approach available for the calculation of the spectrum, the so-called method 4 from Lin et al. [2013b]. We also developed a new approach for the calculation of VPSHA that combines the logic used by the so-called ?indirect? approach with a higher precision of the results computed. The last task considered in this report is the one on the ?Modelling earthquake se- quences for considering aftershocks?. In this case we implemented tools for adjusting the rates of existing main shock based input models for the OQ Engine to account for the contribution of aftershocks. We also added to the engine new functions allowing the calculation of hazard using these models.

Approval

| Date | By |
|---|---|
| 2024-02-07 13:20:12 | Dr. Marco PAGANI (GEM) |
| 2024-02-07 13:59:55 | Dr. Irmela ZENTNER (EDF) |

METIS

Seismic Risk Assessment for Nuclear Safety

Research & Innovation Action

NFRP-2019-2020

# D4.4 - Extended PSHA methodology and tools

Version N°1

Authors: Marco Pagani (GEM), Richard Styron (GEM), Kendra Johnson (GEM), Michele Simionato (GEM)

This project has received funding from the Horizon 2020 programme under grant agreement n°945121. The content of this presentation reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

# Disclaimer

The content of this deliverable reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

# Document Information

| | |
|---|---|
| Grant Agreement | 945121 |
| Project Title | Methods And Tools Innovations For Seismic Risk Assessment |
| Project Acronym | METIS |
| Project coordinator | Dr. Irmela Zentner, EDF |
| Project duration | 1st September 2020 – 31st August 2024 (48 months) |
| Related work package | WP 4 – Seismic hazard |
| Related task(s) | |
| Lead organisation | GEM Foundation |
| Due date | Jan 23 |
| Submission date | Dec 23 |
| Dissemination level | Public |

# History

| Version | Submitted by | Reviewed by | Date | Comments |
|---|---|---|---|---|
| 1 | M. Pagani and R. Styron | M. Villani | 4/5/2023 | |
| 2 | M. Pagani and R. Styron | Z. Gülerce | 25/5/2023 | |
| 3 | M. Pagani and R. Styron | I. Zentner | 28/12/2023 | |

# Contents

# List of Figures

# List of Tables

# Glossary and Acronyms

## Acronyms

**a-value**

Gutenberg-Richter a-value.

**ApE**

Annual Probability of Exceedance.

**b-value**

Gutenberg-Richter b-value.

**CS**

Conditional Spectrum.

**GR**

Gutenberg-Richter.

**IML**

Intensity Measure Level.

**IMT**

Intensity Measure Type.

**METIS**

Methods And Tools Innovations For Seismic Risk Assessment.

**MFD**

Magnitude-Frequency Distribution.

**MRD**

Mean Rate Density.

**MRE**

Mean Rate of Exceedance.

**OQ Engine**

OpenQuake Engine.

**PGA**

Peak Ground Acceleration.

**PMF**

Probability Mass Function.

**PSHA**

Probabilistic Seismic Hazard Analysis.

**$R_X$**

horizontal distance to top edge of rupture measured perpendicular to the strike.

**$R_{y,0}$**

horizontal distance off the end of the rupture measured parallel to strike.

**SSHAC**

Senior Seismic Hazard Analysis Committee.

**VPSHA**

Vector-Valued Probabilistic Seismic Hazard Analysis.

**Ztor**

Depth to the top of rupture.

# Special Terms

**GMM**

Ground-Motion Model (also defined as Ground-Motion Prediction Equation).

**Ground-Motion Characterization**

The suite of ground-motion models required to compute seismic hazard for a particular area and the associated epistemic uncertainties.

**Hazard Curve**

A hazard curve is the primary result provided by a Probabilistic Seismic Hazard Analysis (PSHA). For a specific site, it defines the probability in an investigation time (or the annual frequency) of exceeding a set of intensity measure levels for a given intensity measure type (e.g. Peak Ground Acceleration (PGA)).

**Hazard Input Model**

It is a collection of files that altogether specify the type of analysis to be completed, the position the characteristics of the investigated sites, the earthquake sources with the associated epistemic uncertainties, the ground-motion models and their epistemic uncertainties.

**Seismic Source Characterization**

The description of the geometry, position, earthquake occurrence characteristics and epistemic uncertainties of the earthquake sources required to compute the seismic hazard for a given area.

**Uniform Hazard Spectrum**

A derived product of a Probabilistic Seismic Hazard Analysis (PSHA). It is generally computed for a set of spectral accelerations with different periods of vibration, the ordinates contain points with the same probability of exceedance in the investigation time.

# Mathematical symbols

| Variable Definition | Symbol |
|---|---|
| Index variable for the group of sources with correlated uncertainties | $c$ |
| Number of groups of sources with correlated uncertainties | $C$ |
| Index variable for intensity measure levels | $l$ |
| Number of intensity measure levels | $L$ |
| Intensity Measure Level | $\kappa$ |
| Annual frequency of exceedance | $\phi$ |
| Intensity Measure Type (e.g. PGA, SA(T)) | $\xi$ or $im$ |
| Probability Mass Function for the annual frequency of exceedance | $P(\phi)$ |
| Index variable for the patterns of correlated uncertainties | $q$ |
| Number of patterns of correlated uncertainties in a group | $Q$ |
| Index variable for logic tree realisations | $r$ |
| Number of realisations admitted by the logic tree | $R$ |
| Number of realisations for source $i$ | $R_i$ |
| Index variable for sources | $s$ |
| Number of sources | $S$ |
| Number of realizations admitted by a single source-specific logic trees in a hazard model | $\|\tau\|$ |
| Single source-specific logic trees in a hazard model | $\tau$ |
| Number of realizations admitted by the set of source-specific logic trees in a hazard model | $\|\mathbf{T}\|$ |
| Index variable for logic tree paths | $x$ |
| Number of logic tree paths | $X$ |
| Weight for one logic tree realisation | $w_r$ |

# Summary

This report summarises the results achieved within three tasks of the METIS project Work package 4 “Seismic Hazard”.

The first task, named ‘Approaches for the Propagation of Epistemic Uncertainty’ had the goal of improving the approaches currently adopted for the processing of epistemic uncertainties in Probabilistic Seismic Hazard Analysis (PSHA). We expanded the capabilities of the OQ Engine to process logic tree by adding the option of using a Latin Hypercube sampling approach in lieu of the more traditional Monte Carlo approach and we added the possibility - while processing the logic tree - of deciding where to apply the weights assigned to the various realizations admitted. We also developed a completely new approach for the propagation of epistemic uncertainties that computes seismic hazard for each individual source and combines the results using discrete distributions via a convolution approach. According to the tests completed so far this new method is computationally efficient and it provides results consistent with the ones provided by more traditional approaches.

In the second task titled ‘Vector-valued PSHA and CS Approach’ we implemented in the OQ Engine the most complete, rigorous and complex approach available for the calculation of the spectrum, the so-called method 4 from Lin et al. [2013b]. We also developed a new approach for the calculation of VPSHA that combines the logic used by the so-called ‘indirect’ approach with a higher precision of the results computed.

The last task considered in this report is the one on the ‘Modelling earthquake sequences for considering aftershocks’. In this case we implemented tools for adjusting the rates of existing main shock based input models for the OQ Engine to account for the contribution of aftershocks. We also added to the engine new functions allowing the calculation of hazard using these models.

# Part I

# Task 4.3.2.: Approaches for the Propagation of Epistemic Uncertainty

# 1. Methodologies for the Propagation of Epistemic Uncertainty

The modelling of epistemic uncertainty is nowadays routinely performed in almost every Probabilistic Seismic Hazard Analysis (PSHA). Nonetheless, the way in which this is carried out varies between different studies depending on their goals, the extent of the area covered by the analysis and the code used for the calculation. The number of sites is in general a factor controlling the complexity of the logic tree [Kulkarni et al., 1984, Bommer and Scherbaum, 2008, Bommer, 2012] because a large investigation area presupposes a larger number of sources and - as a results - a more complicated logic tree structure and a much larger number of realizations. Therefore, the size of the calculation and the storage requirements normally increase with the number of sites.

To overcome these limitations in the calculation of hazard in regional and national studies, modelers often decide to 'collapse' the seismic source characterisation component of the logic tree or to assume some level of correlation within the logic tree structure [e.g. Woessner et al., 2015]. The 'collapsing' process involves the calculation of a mean model - or a smaller set of models - that has the potential to provide hazard results close to the true mean, that is the result obtained by properly processing the entire logic tree structure. This is a practical solution when the goal of the analysis is to provide a robust mean hazard but it eliminates the possibility of assessing the overall variability of results admitted by whole set of epistemic uncertainties considered.

Accepting some correlation between uncertainties is a second solution that helps simplify the logic tree structure although in this case the exploration of epistemic uncertainty is limited and the statistics provided possibly biased. For example, if we consider the epistemic uncertainty on the maximum magnitude of a double-truncated Gutenberg-Richter distribution, the OpenQuake Engine [Pagani et al., 2014] can be instructed to create realizations of the source model where the maximum magnitude for all the sources - or a subset of them - is either incremented or decremented. Such a strategy will most probably give some indication on the tails of the distribution of hazard but - certainly - it will not provide an appropriate description of the overall shape of the distribution describing the full spectrum of variability.

With respect to these subjects, the research we performed within Methods And Tools Innovations For Seismic Risk Assessment (METIS) covered two separate directions. In the first one, we worked at expanding the capabilities of the OQ Engine to process logic trees. In the second one, we explored new approaches for a more efficient propagation of epistemic uncertainties.

## 1.1. Introduction

In this section we describe various strategies for processing the logic tree. For this purpose, we consider a simple hazard analysis involving one site, one scalar intensity measure type (e.g. PGA, herein indicated as IMT) and, one set of sources $\boldsymbol{S} = \{src_1, \ldots, src_S\}$. The hazard input model contains epistemic uncertainties in the seismic source and ground motion characterisation. The modelling of epistemic uncertainty is done using a set of logic trees, one for each source included in the model, that can be described as follows, $\boldsymbol{T} = \{\tau_1, \ldots, \tau_S\}$,.

Some of the uncertainties assigned to a source can be shared amongst other sources to form groups with correlated uncertainties. A typical example are the ground motion models for all the sources belonging to the same tectonic region. In this case, as already explained, it is necessary to consider correlation of uncertainties amongst groups of sources when combining the results.

The results of the hazard analysis for the case in question is a set of tuples each one composed by a hazard curve $\boldsymbol{\phi}$ and a weight $w$. The cardinality of this set is $R$. The weights in the various tuples obey to the - well known - constraint $w : \sum_{i=1,R} w_i = 1$. The hazard curve is a vector with the annual frequencies of exceedance (or probabilities of exceedance in a given investigation time) $\boldsymbol{\phi} = \{\phi_1, \ldots, \phi_L\}$. The cardinality of $\boldsymbol{\phi}$ is $L$, which is the same cardinality of the vector with the intensity measure levels $\kappa = \{\kappa_1, \ldots, \kappa_L\}$.

Given an intensity measure type $\xi$, the overall goal of the logic-tree processing is to compute the total hazard at the site for all realization admitted, that is the set of the annual frequencies of exceedance, $\boldsymbol{\phi}_{|\mathbf{T}|}$. For each realization, this set is simply computed by summing the annual frequencies of exceedance for the same intensity measure level produced by $S$ sources considered.

$$\phi_l = \sum_s \phi_{s,l} \tag{1.1}$$

In the most common case, where epistemic uncertainties must be accounted for, the result will be a set of values $\phi_{l,r} | l \in [1, R]$.

## 1.2. Improvements to the Logic Tree Processing in the OpenQuake Engine

The OpenQuake Engine had been traditionally providing the users with the possibility of defining in the input files a number of epistemic uncertainties affecting the Seismic Source Characterization and the Ground-Motion Characterization. Regarding the former, OQ Engine provides a number of predefined uncertainties that according to instructions provided by the users are applied to all the seismic sources in a Hazard Input Model or to a subset of them. For example, in the case of sources whose occurrence if described by a double truncated Gutenberg-Richter (GR) distribution, the epistemic uncertainty on the maximum magnitude can be specified by a set of magnitude values and the associated weights. Subsets of sources to which apply a certain uncertainty can be defined either using their IDs, by specifying a tectonic region or a

***Table 1.1*** *– Logic tree processing options offered by the OpenQuake Engine. The column 'key' contains the key to be used to define in the OQ Engine configuration file the parameter* `sampling_method`

| name | key | sampling of realizations | calculation of statistics |
|---|---|---|---|
| early weights | `early_weights` | Sampling of realizations performed using Monte Carlo sampling and considering the weight of each possible realization | Statistics computed considering equal weight for each realisation |
| late weights | `late_weights` | Sampling of realizations performed using Monte Carlo sampling and equal weight s for all the models | Statistics computed considering the weight of each realisation |
| early latin | `early_latin` | Sampling of realizations performed using Latin Hypercube sampling and considering the weight of each possible realization | Statistics computed considering equal weight for each realisation |
| late latin | `late_latin` | Sampling of realizations performed using Latin Hypercube sampling and equal weight s for all the models | Statistics computed considering the weight of each realisation |

source typology. With respect to the Ground-Motion Characterization, to account for epistemic uncertainties the modeller can specify in the input a list of ground-motion models for each tectonic region.

Traditionally, the OQ Engine offered two approaches for the processing of the user-defined logic-tree. The first approach called 'full path enumeration' computes the hazard for all the possible realisations admitted by the logic tree and calculates the relative statistics by taking into account the weight associated with each realisation. This is an approach that works well when the logic tree in the Hazard Input Model contains a limited number of realisations and the modeler' interest is mostly on getting a more robust estimate of the mean hazard rather than exploring the full range of uncertainties. A typical use case is the calculation of hazard at regional scale. The second approach called 'sampling' supports cases where the number of realizations admitted by the logic tree is too large and therefore only a subset of models is used for computing the hazard. In this case, OQ Engine randomly picks a subset of all the possible realisations and computes the associated results. The sampling strategy implemented selects the models (i.e. the logic tree realizations or end branches) based on their corresponding weight using a traditional Monte Carlo approach. In the following phase the OQ Engine computes the hazard and their statistics (e.g. the mean hazard curve for a given IMT) without accounting for the weight of each result.

Within the METIS project we added to the OQ Engine a different sampling scheme that selects models (i.e. logic tree realizations) independently from their weight and

computes the hazard statistics in the post-processing phase taking into account the weight related to each realization.

To further improve and make more efficient the exploration of the full spectrum of uncertainties defined in the logic tree, we also added to OQ Engine the possibility of performing the sampling not just via a traditional Monte Carlo strategy but also via Latin Hypercube Sampling.

Table 1.1 summarises the options now supported by the OQ Engine for processing the logic tree.

### 1.2.1. OQ Engine implementation

The module in the OQ Engine that performs the processing of the logic tree is available in the `lt` module. In particular, the function performing the sampling - called random - can be inspected here.

## 1.3. A New Methodology for Propagating Epistemic Uncertainties in PSHA

Current PSHA practice distinguishes between hazard analyses performed for a single site and analyses covering large territories such as the one of a nation. The difference between the two analyses is not just semantic since the two options entail different goals, problems, methods and calculation approaches.

In the case of site specific studies for critical facilities, particular emphasis is put on capturing the full spectrum of epistemic uncertainties. Everyone involved or reading about PSHA studies following the Senior Seismic Hazard Analysis Committee (SSHAC) guidelines, knows well the phrase stating the goal of a hazard analysis is "to represent the center, the body, and the range of technical interpretations that the larger informed technical community would have if they were to conduct the study." [Budnitz et al., 1997].

From a more practical perspective, this means that in a hazard calculation for a given site and IMT, each source will produce a set of Hazard Curves that overall describe the variability of the hazard due to uncertainties of different nature. Assuming independence between the several hazard curves produced by each source, the total number of hazard curves is a simple counting problem solved with the so called 'multiplication principle' (or rule of product). If $X_i$ is the number of hazard curves computed for source $i$ the total number of curves $X_T$, that is the number of realizations admitted by the combining the logic trees for all the sources contributing to hazard at the site, will correspond to

$$X_T = \prod_{i=1}^{S} X_i \tag{1.2}$$

where $S$ is the total number of seismic sources. As already explained in the previous section, in case of complex logic trees, the calculation of all the realisations in common PSHA software might not be feasible and, therefore, only a subset of total curves admitted is effectively computed, generally using a Monte Carlo sampling. The final annual frequency of exceedance for one of the $X_T$ hazard curves admitted is computed considering a specific combination of realizations (i.e. one for each source contributing to the final hazard). In particular, for each intensity measure level the final annual frequency of exceedance is the sum of the annual frequency of exceedance computed for each of the sources contributing to the hazard and the weight associated with each realisation corresponds to the product of the weights of the realisation sampled for each one of the sources considered.

In the following section we illustrate a new methodology for propagating the epistemic uncertainties that instead of either considering the full set of realizations admitted by the logic trees or a sample of it, considers all the realizations admitted by each source-specific logic tree and propagate them more efficiently while accepting some approximations in the description of the annual frequencies of exceedance due to the binning process adopted.

### 1.3.1. Methodology description

The approach proposed requires that the calculations of hazard at the site of interest are performed by individually processing each source and its associated epistemic uncertainties. The final hazard is computed *a-posteriori* using an *ad-hoc* post-processing procedure. For a proper execution of the latter operation, the modeller must provide as input the branch-sets (i.e. an individual component of the logic tree describing one specific uncertainty) shared amongst at least two source-specific logic trees. A simple example of correlated uncertainties occurs when the one tectonic region in a input hazard model contains more than one source or when two fault sources share the same the same variability of the seismogenic thickness. This uncertainty is described by a set of values of the thickness and the associated weights. In the hazard analysis, for the sources sharing this correlated uncertainty, we will use the same values of thickness and the same weights. If different weights are used for the same set of parameters (e.g. values of the seismogenic thickness), correlation cannot be assumed anymore.

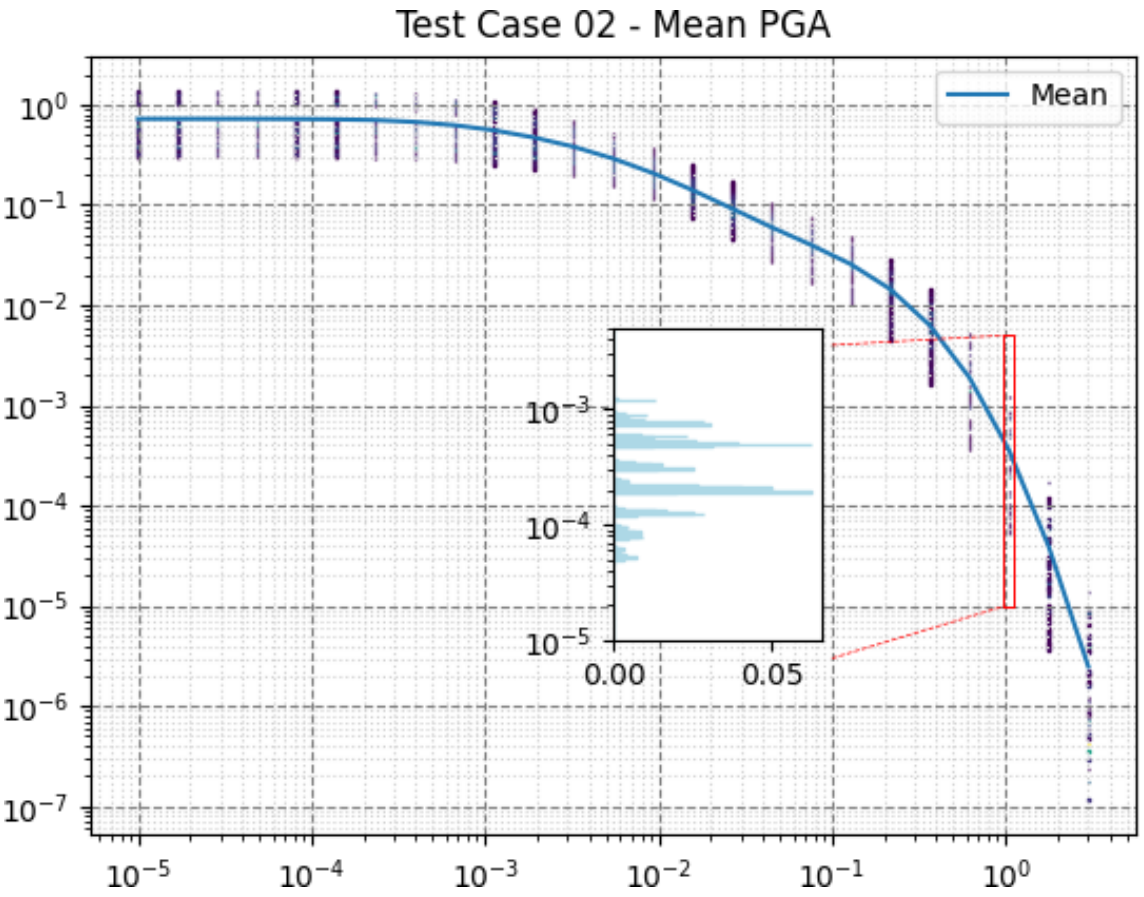


***Figure 1.1*** *– Example of PMF for one intensity measure level. The dots distributed along a vertically elongated pattern represent the values $\phi_{s,r,l}$ of the annual frequencies of exceedance produced by all the end branches of the logic tree considered for this example.*

To introduce the new methodology we start from the results of the PSHA computed for a single site using a set of sources and their associated epistemic uncertainties and one IMT (see also Section 1.1 at page 16). We consider an intensity measure

level $\kappa_l$ and the corresponding values of the annual frequency of exceedance in each of the $R$ hazard curves for source $src_s$, that is $\boldsymbol{\phi}_{s,r,l}$. In conventional approaches for processing the output of a PSHA taking into account epistemic uncertainty, $\boldsymbol{\phi}_{s,r,l}$ is used to obtain common statistics such as the mean or the median annual frequency of exceedance. In the methodology here discussed, we obtain from the weights $w_r$ associated with the various $\boldsymbol{\phi}_{s,r,l}$ a discrete distribution that is a PMF $P(\phi)$ describing the probability of having a certain value of the annual frequency of exceedance for a given intensity measure level $\phi$. That is

$$P_{s,l}(\phi) = \sum_{r=1}^{R} w_r \quad \text{if} \quad b_x < \phi_{s,l,r} < b_{x+1} \tag{1.3}$$

where $\{b_1, \ldots, b_M\}$ are the edges of the bins describing the discrete PMF and $\{w_1, \ldots, w_R\}$ are the weights assigned to the realisations of the source-specific logic tree for source $s$. Figure 1.1 shows an example of the histogram created for the annual frequency of exceedance of the PGA corresponding to 1 g. We repeat the process of building the histogram $P_{s,l}(\phi)$ for all the intensity measure levels $\kappa_l$ of the $R$ results produced by an individual source; we repeat this operation for all the $S$ sources in the hazard input model.

Since for each intensity measure level and individual realization, the final annual frequency of exceedance is the sum of the annual frequency of exceedance computed for each of the sources contributing to the hazard, we can consider this as the sum of $S$ independent random variables with a distribution described by the corresponding $P_{s,l}(\phi)$. As explained in common probability theory books, the distribution of the sum of two independent random variables $x$ and $y$, corresponds to the convolution of the distributions for $x$ and $y$. In the new approach proposed, therefore, the calculation of the final distribution of the annual frequency of exceedance for one intensity measure level is performed by convolving the $S$ discrete probability mass functions. Note that the use of convolution is possible only under the condition that the realizations produced by the various source-specific logic trees are independent. If this constraint is ignored, the final distribution of the annual frequency of exceedance will be slightly broader than the correct one. Moreover, it should be noted this possible mismatch occurs also in the case when the existing correlations are ignored while propagating the epistemic uncertainties using a sampling approach.

In many real hazard analyses the assumption of independence does not hold. The simplest example demonstrating the interdependence of components of the logic tree used for different sources is the one where two sources belonging to the same tectonic region (say for example Active Shallow Crust) contribute to the hazard for a given site. Therefore, in presence of correlation between epistemic uncertainties in the logic trees of two or more sources, we split the sources into groups $\{G_1, \ldots, G_C\}$ where $C$ is the number of groups with correlated uncertainties. Each group contains the sources sharing correlated epistemic uncertainties. For example, in absence of uncertainty in the seismic source characterisation, the sources belonging to the same tectonic region form a group because they share the same ground-motion logic tree. The number $C$ of groups in this case corresponds to the number of tectonic regions controlling the hazard at the investigated site. If a source does not share correlated uncertainties with other sources, it forms a group with a single element.

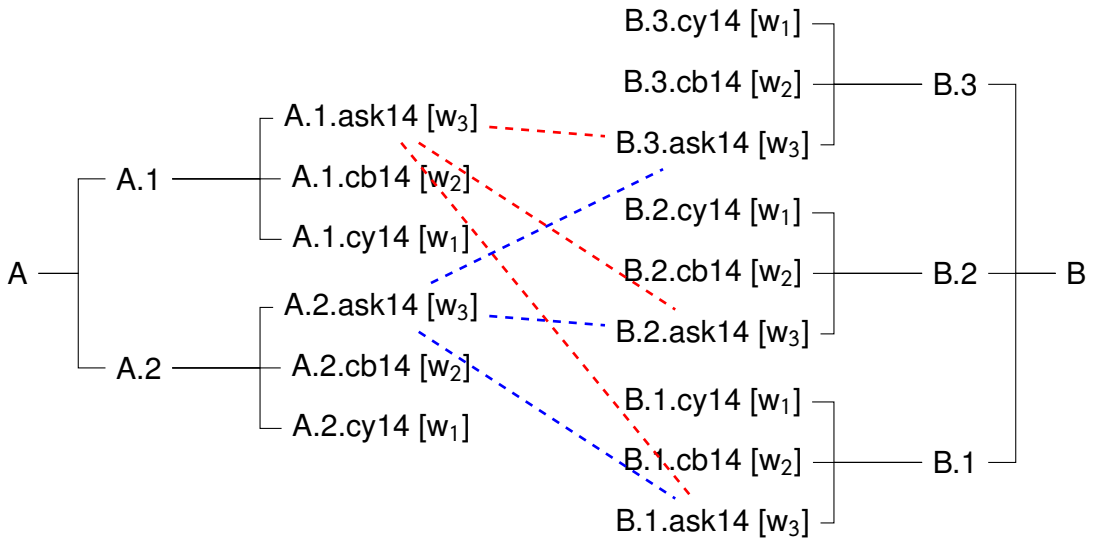


***Figure 1.2*** *– Schematized logic trees for sources A and B sharing one branch of correlated uncertainties. The branch in common is the one describing the epistemic uncertainty in the modelling of ground motion. The red dashed segments connect one realization of the logic tree for source A with all the correlated branches in the logic tree for source B.*

For each group we find the paths admitted by the correlated uncertainties $\{path_1 \text{-} \ldots path_X\}$. Each path is a unique tuple of indexes, where each index specifies a branch of the branch sets describing the correlated uncertainties. Using these paths, we can select the hazard curves computed for for each combination of the correlated epistemic uncertainties by the sources included in the group. For example, Figure 1.2 shows the logic trees for two sources sharing epistemic uncertainties on the ground motion model. If we do not account for the correlations, the combination of these two source specific logic trees admits 54 realizations. One example of path describing one of the these 54 realizations is the following: 'A.1.ask14 + B.3.cy14'. However, source A and B share the same uncertainty at the second branching level. This entails that we can combine a realization from the logic tree for source A only with the correlated realisations of source B. Using the example described in Figure 1.2 this means that we can combine the hazard curve computed for the branch 'A.1.ask14' only with the hazard curves in the logic tree for source B computed using the 'ask14' GMM. This condition can be represented with a pattern used to identify paths of correlated uncertainties. For example, the combination of 'A.*.ask14' and 'B.*.ask14' selects all the realisations for which the hazard in the tectonic region for A and B was computed using the 'ask14' GMM. The realisations belonging to the group described by this pattern will include the following six realisations (i.e. paths):

- A.1.ask14 + B.1.ask14
- A.1.ask14 + B.2.ask14
- A.1.ask14 + B.3.ask14
- A.2.ask14 + B.1.ask14
- A.2.ask14 + B.2.ask14

- A.2.ask14 + B.3.ask14

This subset of realisations for source A and B contains independent realisations for each source hence, similarly for the case described above with uncorrelated uncertainties, we can compute a histogram for of the annual frequency of exceedance using the hazard curves for one source and belonging to this group. If we consider the example above, the histogram for source A is based on the results computed for branch 'A.1.ask14' and 'A.2.ask14' while for source B the branches considered are 'B.1.ask14', 'B.2.ask14' and 'B.3.ask14'. Since the realisations used to build these two histograms are independent, we can convolve them to compute the total hazard for this group and pattern $P_{l,c,q}(\phi)$. $P_{l,c,q}(\phi)$, a probability mass function, defines the probability that, for a given intensity measure level $l$, the realisations in subgroup $c$ obtained for a specific combination $q$ of the correlated epistemic uncertainties gives a certain value of the annual frequency of exceedance.

In our example, if we repeat this operation for the remaining two groups (i.e. the one comprising results for 'cy14' and for 'cb14'), we get as many $P_{l,c,q}(\phi)$ as the number of patterns of correlated uncertainties (that is 3). Interestingly, the total number of paths (i.e. hazard curves), created by combining the paths of the individual source-specific logic tree while taking into account correlation, corresponds to 18; this number is much smaller than the one obtained while ignoring correlation.

Each $P_{l,c,q}(\phi)$ has an associated weight $w_{l,c,q}$ corresponding to the product of the weights of the correlated uncertainties used to create it. In our example, the weight of each group is the weight of the corresponding ground-motion model. Overall, the set of $P_{l,c,q}(\phi)$ describes alternative distributions of the annual frequency of exceedance provided by the various combinations of the correlated uncertainties. In accordance with this interpretation, we combine this set of histograms $P_{l,c,q}(\phi)$ for all the $Q$ groups. Specifically, we compute the $P_{l,g}(\phi)$ as a mixture distribution corresponding to the weighted sum of $Q$ probability mass functions

$$P_{l,c}(\phi) = \sum_{q=1}^{Q} P_{l,c,q}(\phi) w_{l,c,q} \tag{1.4}$$

What remains to be completed in the final processing step is the combination of the $P_{l,c}(\phi)$ each one describing the distribution of the annual frequency of exceedance in group $c$ and intensity measure level $l$. Given that these distributions are independent the last processing step can be easily accomplished via convolution to obtain $P_l(\phi)$.

### 1.3.2. Implementation

The library containing the tools implementing the methodology just described - LEAP (Library for Epistemic uncertAinty Propagation) - is currently in a repository in the Gitlab hosted at GEM (access available to the members of the consortium upon request). The library will be released publicly once the METIS project will end and the scientific paper illustrating the library submitted.

### 1.3.3. Tests

To illustrate the capabilities of the methodologies for the propagation of epistemic uncertainties proposed, we present herein a few application examples and tests.

We describe herein a number of tests we implemented in the 'unc' library to verify the ability of the proposed methodology for propagating epistemic uncertainties compatible with the more traditional approaches adopted.

#### Test 1

The first test case considers a hazard analysis for a sigle site and two fault sources that do not share uncertainties, that is, we assume they are completely independent.

Source A is a simple fault source with epistemic uncertainties in the Seismic Source Characterization on the Gutenberg-Richter *a* and *b* values (see listing 1.1) and four GMMs in the Ground-Motion Characterization (see listing 1.2). Overall, the number of realizations admitted by the logic tree structure for source A is 24.

***Listing 1.1*** *– Source A Seismic Source Characterization logic tree*

```
Branch Set [bs1]: sourceModel
└── Branches:
    └── b11: ssm_a.xml
Branch Set [bs2]: abGRAbsolute
├── applyToSources: a
└── Branches:
    ├── b21: 4.5 1.1
    ├── b22: 4.5 1.0
    ├── b23: 4.5 0.9
    ├── b24: 4.6 1.1
    ├── b25: 4.6 1.0
    └── b26: 4.6 0.9
```

***Listing 1.2*** *– Source A Ground-Motion Characterization logic tree*

```
Branch Set [bs1]: gmpeModel
├── applyToTectonicRegionType: Active Shallow Crust
└── Branches:
    ├── b1: AbrahamsonEtAl2014
    ├── b2: BooreEtAl2014
    ├── b3: CampbellBozorgnia2014
    └── b4: ChiouYoungs2014
```

Source B is also a simple fault source with various epistemic uncertainties Seismic Source Characterization (see listing 1.3) and the same epistemic uncertainties of source A in the Ground-Motion Characterization (see listing 1.2). Overall, the number of realizations admitted by the logic tree structure for source B is 48.

***Listing 1.3*** *– Source B Seismic Source Characterization logic tree*

```
Branch Set [bs1]: sourceModel
└── Branches:
    └── b11: ssm_b.xml
Branch Set [bs2]: truncatedGRFromSlipAbsolute
```

```
├── applyToSources: b
└── Branches:
    ├── b21: slipRate: 10.0 rigidity: 32.0
    ├── b22: slipRate: 13.0 rigidity: 32.0
    └── b23: slipRate: 16.0 rigidity: 32.0
Branch Set [bs3]: setMSRAbsolute
├── applyToSources: b
└── Branches:
    ├── b31: WC1994
    └── b32: Leonard2014_Interplate
Branch Set [bs4]: setLowerSeismDepthAbsolute
├── applyToSources: b
└── Branches:
    ├── b41: 40
    └── b42: 50
Branch Set [bs4b]: recomputeMmax
├── applyToSources: b
└── Branches:
    └── b4b1: 0
```

The input model that combines source A and B, overall admits 288 logic tree paths. Figure 1.3 shows a comparison between the mean and median hazard curve computed by the OQ Engine by considering the contributions provided by the two sources considered. Overall, the match between the hazard curves computed the two methodologies considered is satisfactory.

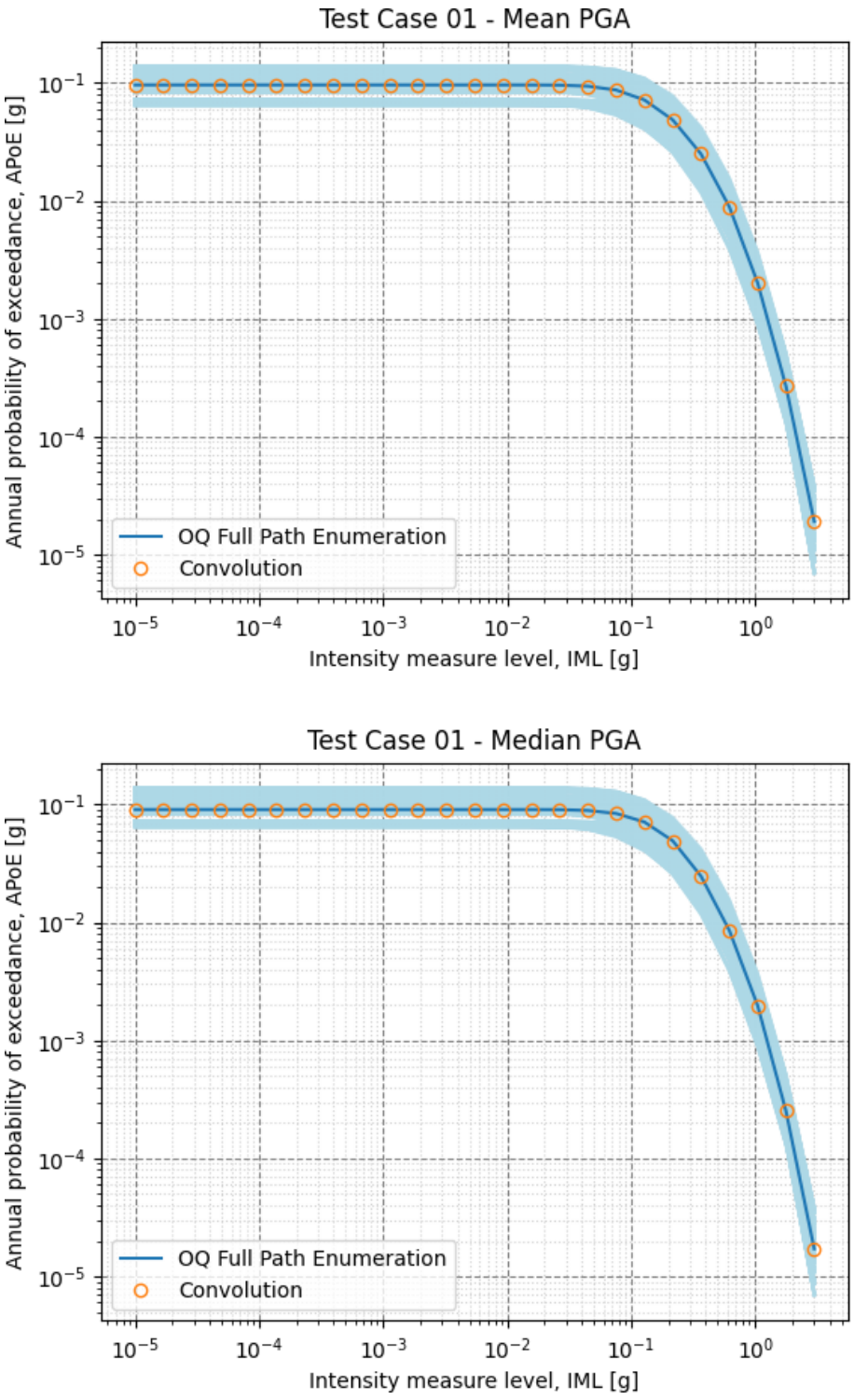


***Figure 1.3*** *– [top panel] Comparison between the mean hazard curve computed by the OQ Engine for the site considered in test 01. The curves in light blue are the hazard curves computed considering individual realizations of the logic tree. [bottom panel] As in the top panel, but in this case the comparison is between median hazard curves.*

### Test 2 - Hazard curves

The second test case considers a hazard analysis for a single site and four sources that share some uncertainties. In particular, we assume that sources B and C share the same uncertainty on the lower seismogenic depth and that sources A and B belong to the same tectonic region and, therefore, they share the same Ground-Motion Characterization.

The OQ Engine is currently unable to cope with a model with this complexity. To compute the set of expected results we use an *ad-hoc* Monte Carlo based sampling procedure.

The Seismic Source Characterization for source A is the same used for source A in the test 1 (see section 1.3.3). Listings 1.4 and 1.5 describe the uncertainties considered in the Seismic Source Characterization for sources B and C. For source D we do not consider epistemic uncertainties in the Seismic Source Characterization.

***Listing 1.4*** *– Source B Seismic Source Characterization logic tree*

```
Branch Set [bs1]: sourceModel
└── Branches:
    └── b11: ssm_b.xml
Branch Set [bs2]: truncatedGRFromSlipAbsolute
├── applyToSources: b
└── Branches:
    ├── b21: slipRate: 10.0 rigidity: 32.0
    ├── b22: slipRate: 13.0 rigidity: 32.0
    └── b23: slipRate: 16.0 rigidity: 32.0
Branch Set [bs3]: setMSRAbsolute
├── applyToSources: b
└── Branches:
    ├── b31: WC1994
    └── b32: Leonard2014_Interplate
Branch Set [bs4]: setLowerSeismDepthAbsolute
├── applyToSources: b
└── Branches:
    ├── b41: 40
    └── b42: 50
Branch Set [bs4b]: recomputeMmax
├── applyToSources: b
└── Branches:
    └── b4b1: 0
```

***Listing 1.5*** *– Source C Seismic Source Characterization logic tree*

```
Branch Set [bs1]: sourceModel
└── Branches:
    └── b11: ssm_c.xml
Branch Set [bs2]: truncatedGRFromSlipAbsolute
├── applyToSources: c
└── Branches:
    ├── b21: slipRate: 10.0 rigidity: 32.0
    ├── b22: slipRate: 13.0 rigidity: 32.0
```

```
    └── b23: slipRate: 16.0 rigidity: 32.0
Branch Set [bs3]: setMSRAbsolute
├── applyToSources: c
└── Branches:
    ├── b31: WC1994
    └── b32: Leonard2014_Interplate
Branch Set [bs4]: setLowerSeismDepthAbsolute
└── Branches:
    ├── b41: 40
    └── b42: 50
Branch Set [bs4b]: recomputeMmax
├── applyToSources: c
└── Branches:
    └── b4b1: 0
```

Sources A and B belong to the active shallow crust tectonic region, we assigned source C to subduction interface and source D to stable continental crust. The listings 1.6 and 1.7 show the GMMs for the three tectonic regions considered in this analysis. The logic tree for stable continental crust is the same used for active shallow crust (see listing 1.6).

***Listing 1.6** – Active Shallow Crust Ground-Motion Characterization logic tree*

```
Branch Set [bs1]: gmpeModel
├── applyToTectonicRegionType: Active Shallow Crust
└── Branches:
    ├── b1: AbrahamsonEtAl2014
    ├── b2: BooreEtAl2014
    ├── b3: CampbellBozorgnia2014
    └── b4: ChiouYoungs2014
```

***Listing 1.7** – Subduction Interface Ground-Motion Characterization logic tree*

```
Branch Set [bs1]: gmpeModel
├── applyToTectonicRegionType: Subduction Interface
└── Branches:
    ├── b11: ParkerEtAl2020SInter
    ├── b12: AbrahamsonGulerce2020SInter
    └── b13: KuehnEtAl2020SInter
```

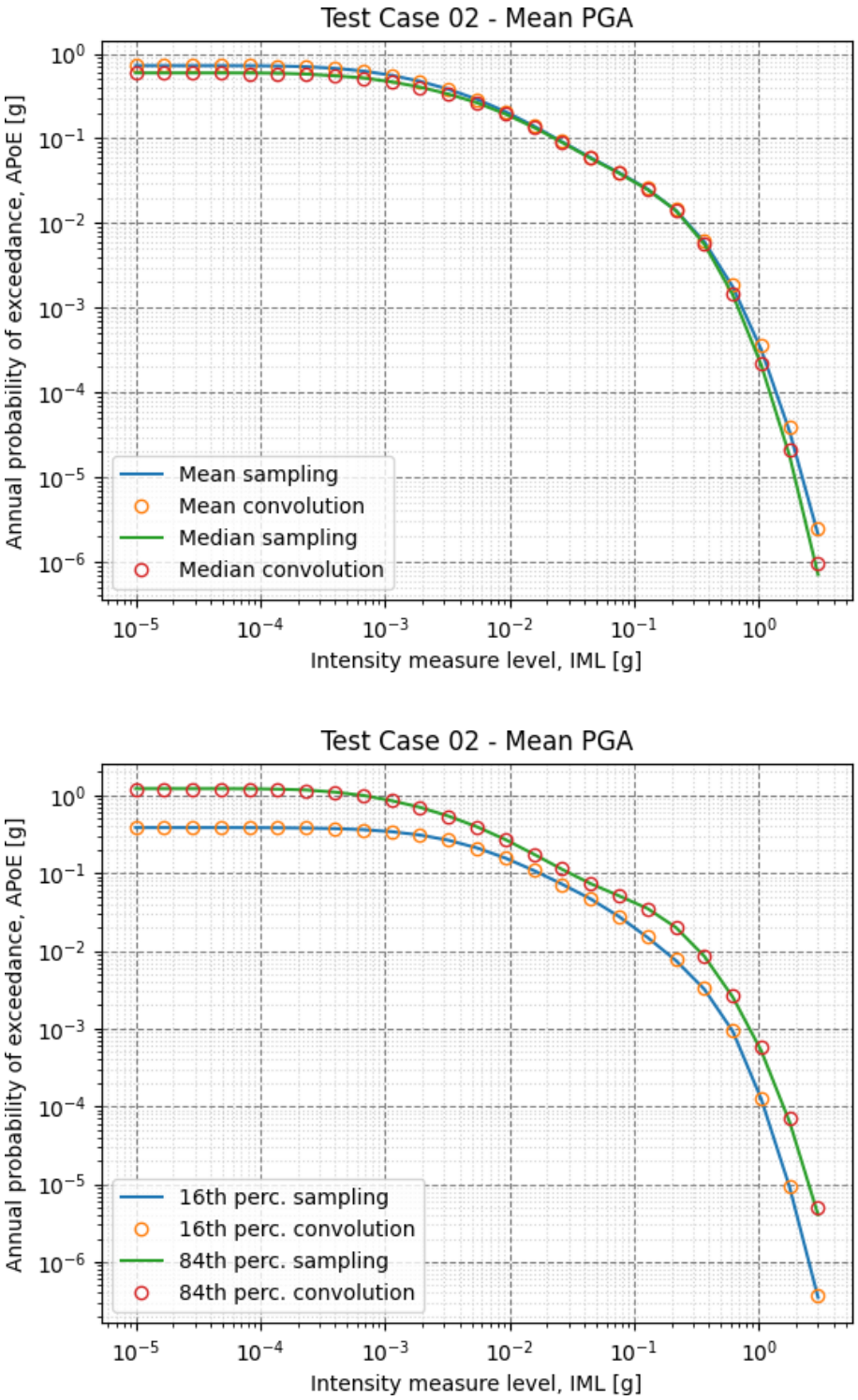


***Figure 1.4*** *– [top panel] Comparison between the mean and median hazard curves computed using Monte Carlo sampling and the corresponding results computed using the new method proposed, called 'convolution' in the plots. [bottom panel] Same as above but for the 16th and 84th percentiles.*

### Test 1 - Hazard Disaggregation

We use test 1, previously used for checking the hazard curves computed with the new methodology proposed, to also check the calculation of the mean disaggregation.

Figure 1.5 shows the comparison between the annual frequency of exceedance computed by OQ Engine for the various magnitude bins (with a width of 0.1 units) and the results obtained with the new methodology proposed. Overall, the results in this

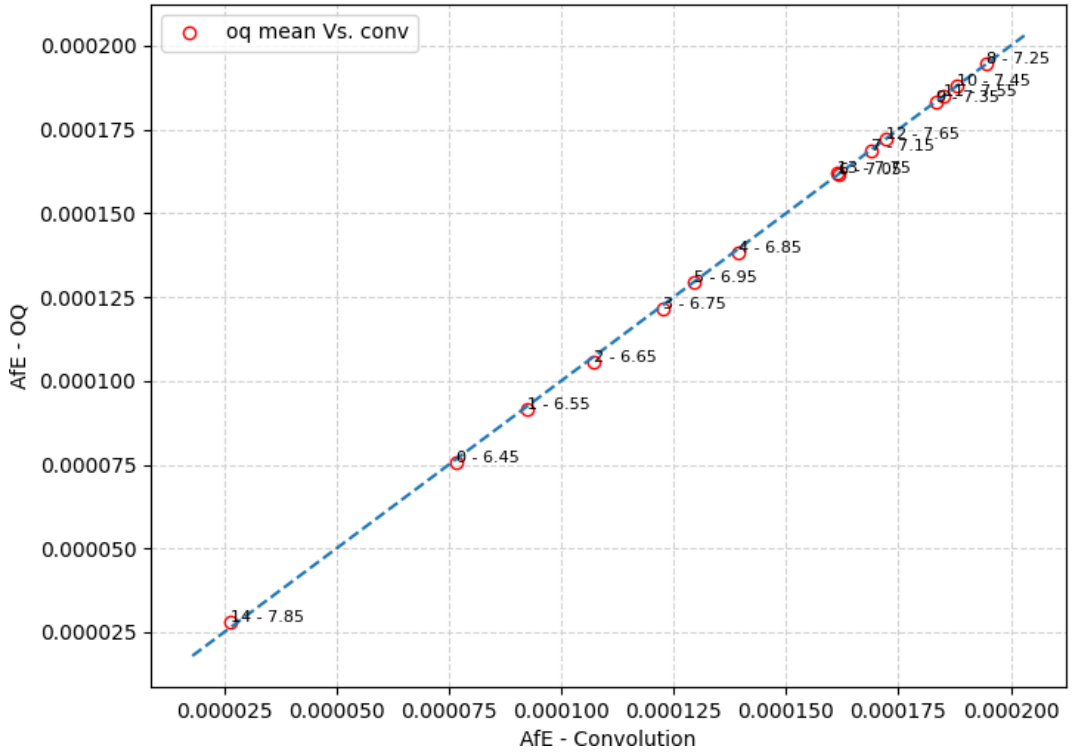


***Figure 1.5*** *– Comparison between the annual frequency of exceedance for magnitude bins with edges comprised between 6.4 and 8.1, computed using OQ Engine and the approach for the propagation of epistemic uncertainty presented here.*

Figure show a good match between the original results from OQ Engine and the ones computed using the alternative approach proposed in this document.

Figure 1.6 shows the mean magnitude-distance disaggregation for PGA with 10% probability of exceedance in 50 years computed with the new methodology proposed and with the OQ Engine. The match between the two results is satisfactory as well as the value of the annual frequency of exceedance for the modal combination of magnitude and distance (see red cross and information provided in the inset in the top right corner).

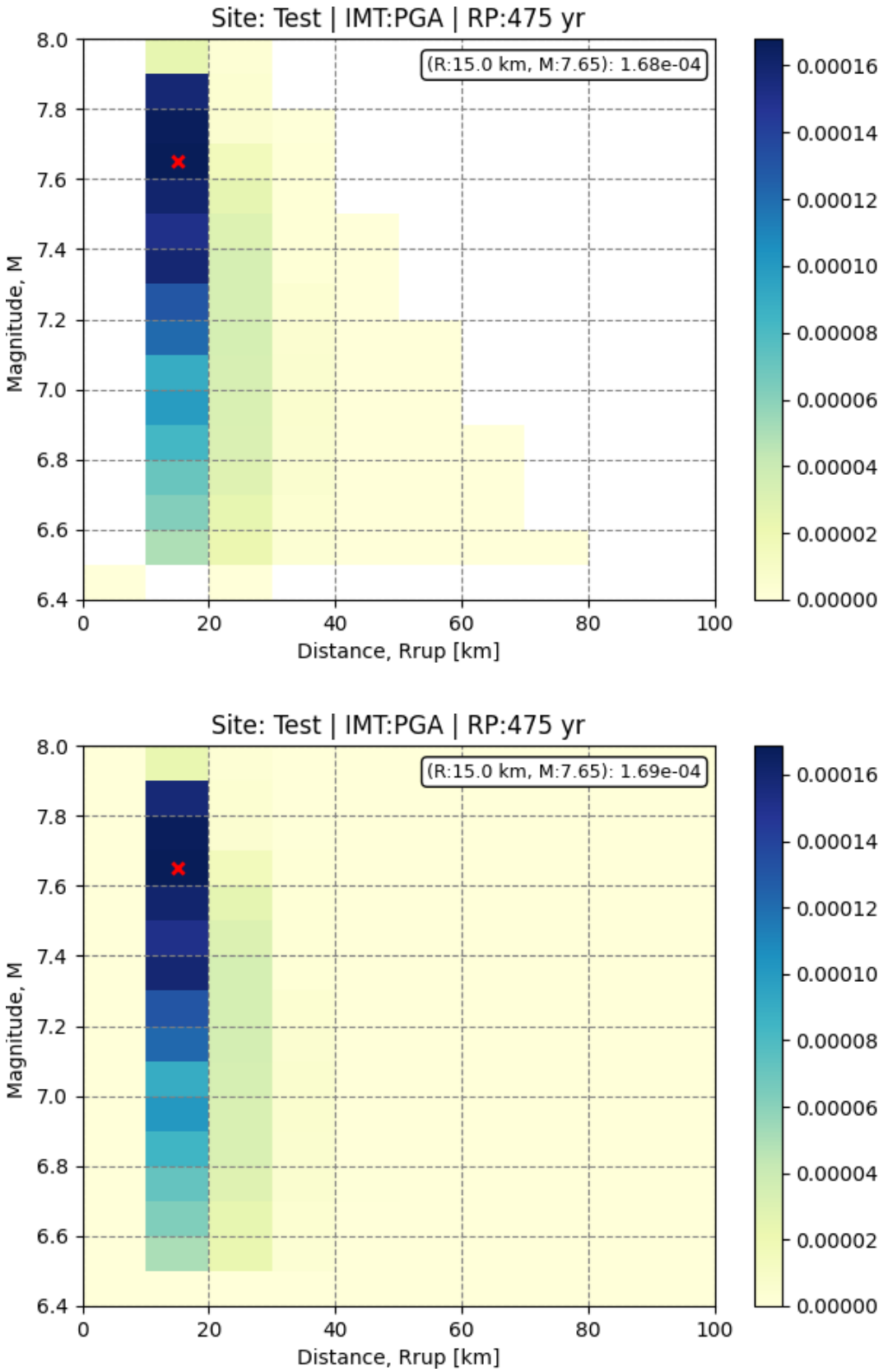


***Figure 1.6*** *– Comparison between the mean magnitude-distance disaggregation for PGA with 10% probability of exceedance in 50 years computed with the new method proposed [top] and the OQ Engine [bottom].*

**Performance analysis**

To appraise the efficiency of the new methodology proposed we provide in Figure 1.7 the summary of a performance test implemented in the LEAP library.

This performance test uses test 1 described in Section 1.3.3 to assess the calculation time and the amount of memory used by the proposed method compared to a more traditional sampling approach. The upper left panel shows the calculation time needed to compute the mean hazard curves as a function of the number of samples used (the sampling approach used is a traditional Monte Carlo sampling). The calculation time required by the method for propagating epistemic uncertainty based on convolution is lower than the one required by a sampling with 10000 samples.

The lower left panel shows the memory consumption of the sampling method as a function of the number of samples used. The method for propagating epistemic uncertainty based on convolution is extremely efficient in terms of memory consumption.

The two panels on the right show, the ratios between the hazard curves computed with the sampling and the convolution-based approached and the mean hazard curve computed with the OQ Engine and - in the lower panel - the various hazard curves computed. Overall the precision of the method proposed is good for this idealised case study.

## 1.4. Conclusions

In this task we worked on two separate fronts. On one side we worked at improving the methods already available in the OQ Engine for the propagation of epistemic uncertainties by adding different sampling strategies. On the other side, we developed a new methodology for computing hazard and propagating epistemic uncertainties that relies on the convolution of discrete probability distributions and the calculation of mixture distributions. According to the tests implemented the method provides results close to the ones computed using more traditional approaches, with a quite use of memory. This gives us confidence that the method will help coping more efficiently with models characterised by a large number of epistemic uncertainties as it is the case for hazard analyses of critical facilities.

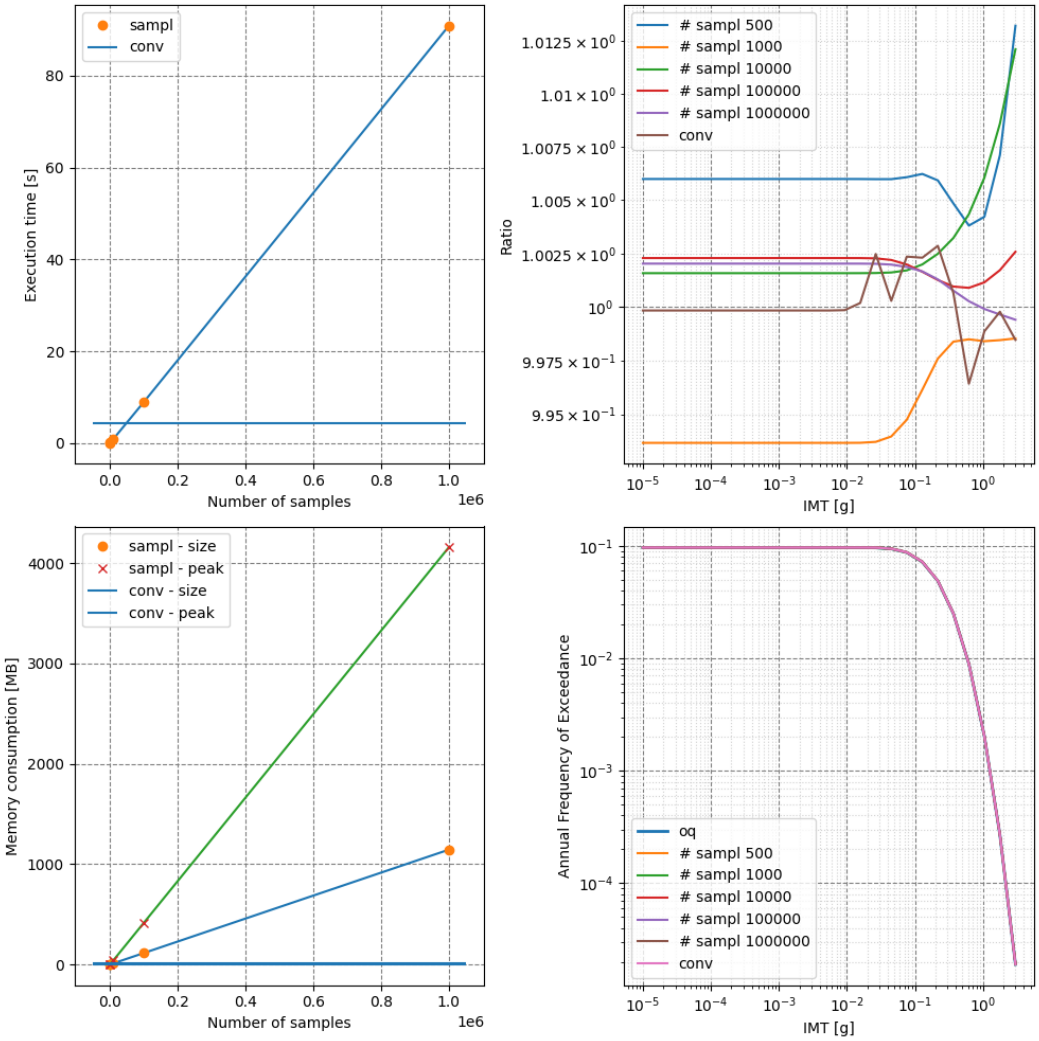


***Figure 1.7*** *– [top left panel] Execution time for the calculation of the mean hazard curves as a function of the number of samples. The time requested by the convolution based approach is represented by the blue horizontal line. [bottom left panel] As for the previous panel but in this case we show the memory consumption. [top right panel] Ratios between the hazard curves for PGA computed for different numbers of samples and the mean hazard curve computed with the OQ Engine full path enumeration. The curve in brown shows the ratio of the mean hazard curve computed using the convolution-based approach and the hazard curve computed with the OQ Engine full path enumeration. [bottom right panel] Mean hazard curves computed for the various cases. The intensitty measure type used in all these tests is PGA.*

# Part II

# Task 4.4.1: Vector-valued PSHA and CS Approach

# 2. Vector-Valued Probabilistic Seismic Hazard Analysis: Methodology and OpenQuake-Engine Implementation

VPSHA was introduced in the late 1990s to compute the joint occurrence (or exceedance) of a vector of intensity measures [Bazzurro and Cornell, 2002, and references therein]. Compared to the results computed using a more traditional PSHA approach based on scalar intensity measures, VPSHA provides a more comprehensive description of the different ground motion characteristics required for particular studies such as risk analyses of various typologies of structures or evaluations of the slope stability under seismic excitation.

Despite the power of the methodology and the advantages it provides compared to scalar PSHA, the application of VPSHA, so far, has not been as large as one would intuitively expect. The causes of this disagreement between the power of the method and its little use in common practice are a combination of various reasons. Bazzurro et al. [2010] explains that the limited use of VPSHA can be attributed to:

- The small number of codes implementing it and their restricted accessibility.
- The number of IMTs supported in the implementations available (generally only two).
- The inability of these codes to disaggregate the joint hazard as traditionally done in scalar PSHA.
- The limited number of prediction models for some Intensity Measure Levels (IMLs) and the lack of correlation models between several IMLs.
- The computational burden and the difficulty of using VPSHA in risk and record selection.

In this Chapter we describe a new approach for the calculation of VPSHA that follows the same principles of the 'indirect' method of Bazzurro et al. [2010] while avoiding its main limitations.

## 2.1. The VPSHA Methodology

As stated in Bazzurro and Cornell [2002] the goal of VPSHA is to compute the mean rate of occurrence (or exceedance) for a combination of IMLs of a set of IMTs.

The approaches for computing these results so far presented are fundamentally two. The one originally proposed, the so called 'direct' method [Bazzurro and Cornell, 2002], requires a modification of the kernel used for computing the traditional scalar PSHA. This approach is the one providing the most precise results but it is computationally intensive.

The second approach [Bazzurro et al., 2010] aims to reduce the calculation burden required by the 'direct' method by utilizing the results of the disaggregation approach to compute the Mean Rate Density (MRD). The disadvantage of this approach stands in the fact that, to provide sensible results, the code implementing it has to cope with a large multidimensional matrix to store the disaggregation results. This is needed to exhahustively consider the explanatory variables required by modern GMMs [e.g. Abrahamson et al., 2014].

In the context of the METIS project, we developed a methodology that combines the idea proposed by Bazzurro et al. [2010] and overcomes the main problem of the 'indirect' approach. In this new approach we mimic the logic of stacking the contributions from different ruptures used in the 'indirect' approach but, instead of using as independent variables in the disaggregation matrix the explanatory variables used by GMMs, we accumulate in a four-dimensional matrix the rates of occurrence of discrete combinations of the median values of the two IMTs selected and the corresponding total standard deviations. To avoid confusion with the more conventional disaggregation matrix, we call this matrix the 'kernel' matrix.

### 2.1.1. Previous Approaches

The Mean Rate of Exceedance (MRE) for a traditional scalar PSHA is computed by numerically solving the so-called hazard integral [Cornell, 1968, McGuire, 2004, Gülerce and Abrahamson, 2010, Baker et al., 2021] which is normally written as follows

$$\lambda(IM > im) = \sum_{s=1}^{n_{sources}} \sum_{r=1}^{n_{ruptures}} \nu_{s,r} P(IM > im | rupture_{r,s}) \tag{2.1}$$

where $\lambda(IM > im)$ is the annual rate of exceedance of the intensity measure level $im$, $P(IM > im|rupture_{r,s})$ is the conditional probability of exceeding the intensity measure level $im$. $\nu_{s,r}$ is the rate of occurrence of the rupture $r$ in source $s$. $P(IM > im|rupture_{r,s})$ is computed from the complementary cumulative distribution of the logarithm ground motion, which is habitually assumed to follow a Gaussian distribution with mean and standard deviation provided by the selected GMM.

If we consider a logic tree to account for the epistemic uncertainties in the Seismic Source Characterization and Ground-Motion Characterization, we compute the final annual rate of exceedance as a weighted average of the $\lambda(IM > im)$ obtained for each branch from equation 2.2 where $w_b$ is the weight assigned to branch $b$.

$$\lambda(IM > im) = \sum_{b=1}^{n_{branches}} w_b \lambda_b(IM > im) \tag{2.2}$$

By resolving Equations 2.1 and 2.2 for a set of IMLs, we obtain the alleged hazard curve.

If, without any loss of generality, we neglect epistemic uncertainties and focus on the calculation of hazard by considering only aleatory uncertainty, the MRD for a scalar IMT is computed by differentiating the hazard curve from Equation 2.1 or, more directly, by using a modified version of equation 2.1 where we replace the conditional probability of exceedance of $im$ with its discrete probability of occurrence $P(IM = im|rupture_{r,s})$.

### Direct VPSHA Approach

We can extend the scalar hazard integral described in the previous section to calculate of the joint probability of occurrence of $im_1$ and $im_2$ (i.e. the mean rate density) as follows:

$$\lambda(IM_1 = im_1, IM_2 = im_2) = \sum_{s=1}^{n_{sources}} \sum_{r=1}^{n_{ruptures}} \nu_{s,r} P(IM_1 = im_1, IM_2 = im_2|rupture_{r,s}) \tag{2.3}$$

where $P(IM_1 = im_1, IM_2 = im_2|rupture_{r,s})$ is the discrete joint probability of occurrence of $im_1$ and $im_2$ given the occurrence of $rupture_{r,s}$. $P(IM_1 = im_1, IM_2 = im_2|rupture_{r,s})$ is computed using a multinomial Gaussian distribution.

### Indirect VPSHA Approach

Bazzurro et al. in 2010 proposed the 'indirect' method to foster a wider use of the VPSHA methodology. The indirect method is founded on the use of the disaggregation methodology [e.g. Bazzurro and Cornell, 2002] and the assumption that a multinomial normal distribution describes appropriately the joint probability of the logarithms of correlated IMTs [Jayaram and Baker, 2008].

According to Bazzurro et al. [2010], the main advantages of this approach are the use - with limited modifications - of already existing PSHA codes computing scalar results, a lower computational load than the 'direct' approach, the possibility to compute the joint hazard for a larger number of intensity measure types and the possibility to disaggregate the joint hazard.

Given a set of intensity measure types $IMT$ with cardinality $|IMT|$, the information required for the application to a given site of the indirect method corresponds to [Bazzurro et al., 2010]:

- $|IMT|$ hazard curves.
- The variance-covariance matrix for the IMLs in $IMT$.
- The disaggregation of hazard for the explanatory variables used by the GMMs included in hazard input. In the most complex cases this set might comprise several parameters in addition to ground motion variability, magnitude and distance, such as the Depth to the top of rupture (Ztor), the horizontal distance to top edge of rupture measured perpendicular to the strike ($R_X$) or the horizontal distance off the end of the rupture measured parallel to strike ($R_{y,0}$).

### 2.1.2. Other approaches proposed

Li and Cai [2023] proposed a simplified methodology for computing VPSHA based on the concept of 'equivalent earthquake' that is a virtual earthquake defined using the information collected via a traditional disaggregation analysis [see also Iervolino et al., 2010]. This methodology while computationally efficient is not well suited for cases where the selected GMMs require a large number of explanatory variables.

### 2.1.3. New Proposed Approach

The approach proposed for the computation of the MRD follows the same logic presented for the development of the 'indirect' method while avoiding either to cope with highly multidimensional disaggregation sparse matrices or the approximations entailed by the use of a disaggregation matrix based on a subset of the explanatory variables required by the selected GMMs.

To illustrate the method we consider a simple case of a VPSHA for one site and two IMTs - without accounting for epistemic uncertainty - and define a four dimensional 'kernel' matrix $K$ with edges

- $\{im_{1,1}, \ldots im_{1,n_{im_1}}\}$
- $\{im_{2,1}, \ldots im_{1,n_{im_2}}\}$
- $\{\sigma_{1,1}, \ldots \sigma_{1,n_{\sigma_1}}\}$
- $\{\sigma_{2,1}, \ldots \sigma_{2,n_{\sigma_2}}\}$

where $\sigma_{1,x}$ are binned valued of the total standard deviation of ground motion for the first IML and $\sigma_{2,x}$ are the corresponding bin edges for the second IML considered.

For each rupture of each source in the hazard input model, we can calculate the median value of ground motion for the two IMTs considered and the corresponding total standard deviation of ground-motion. With these values we identify the corresponding cell in the four dimensional 'kernel' matrix $K$ and we add to this cell the rate of occurrence for the corresponding rupture as follows

$$K_{i,j,k,l} = \forall_{s=1}^{n_{sources}} \forall_{r=1}^{r_{ruptures,s}} \sum \nu_{s,r} \text{if } \left[ (im_{1,i} < im_1 < im_{1,i+1}) \wedge (im_{2,j} < im_1 < im_{1,j+1}) \wedge (\sigma_{1,k} < \sigma_1 < \sigma_{1,k+1}) \wedge (\sigma_{2,l} < \sigma_2 < \sigma_{2,l+1}) \right] \quad (2.4)$$

where $\nu_{s,r}$ is the rate of occurrence of the rupture $r$ from source $s$. The kernel matrix $K_{i,j,k,l}$, in general, is a highly sparse matrix with elements concentrated along a the main diagonal.

The calculation of the MRD is then completed by iterating over the elements of the kernel matrix, that is, over all the combinations of median ground motions and standard deviations admitted by the hazard input model as determined by $K_{i,j,k,l}$.

## 2.2. OpenQuake Engine Implementation

The VPSHA calculator is implemented in OQ Engine as a post-processing tool. This means that the vector-valued PSHA analysis can be performed by reusing the information collected in the datastore of a previous classical PSHA for the same site and set of IMTs. One current limitation of this calculator is that is does not support non-paramteric sources.

The functions computing the mean rate density are located in the ‘calc‘ module of the OQ Engine hazard library (see the ‘mrd.py‘ file here). The ‘mrd‘ module contains two functions called ‘update_mrd‘ and ‘update_mrd_indirect‘ that update the matrix ‘mrd‘ by adding the contributions of the set of ruptures GMMs provided as input.

Focusing on the implementation of the indirect method, the construction of the kernel matrix is completed within the loop at line 183 of the ‘mrd‘ module. The contributions of the various combinations of median ground-motions and standard deviations for the intensity measure types considered are aggregegared in the following loop at line 194.

Overall, this implementation is computationally efficient but the component that computes the cumulative density function of the multivariate Gaussian. This component currently relies on an implementation of the multivariate Gaussian available in the Scipy library that is notoriously slow. In the future we will explore the use of alternative and more efficient implementations.

## 2.3. Tests

To illustrate the capabilities of the VPSHA methodology proposed, we present herein a few application examples and tests. The OQ Engine implementation of VPSHA comes with one unit test implemented to ensure that the code produces the expected results.

The test uses two point sources and no epistemic uncertainty in the Seismic Source Characterization and two ground motion models that is two versions of the Kotha et al. [2020] GMM. We compute the MRD for spectral acceleration at 0.2 and 1.0 second and we compare the marginals obtained by summing the matrix along the axes against the original mean rate densities for each of the two IMTs considered. We obtain these mean rate densities from the hazard curves.

Figure 2.1 shows a comparison between the scalar mean rate densities computed with a traditional scalar PSHA and the marginal of the join mean rate densities computed with a VPSHA. The comparison in the upper panel displays the marginal computed with the VPSHA direct approach and the reference scalar mean rate densities computed from the hazard curves. The match between each marginal and the corresponding benchmark is good over the entire range of intensity measure levels. The comparison in the lower panel shows instead the marginal computed with the VPSHA kernel matrix approach and the reference scalar mean rate densities computed from the hazard curves. The match between each marginal and the corresponding benchmark is acceptable for spectral accelerations larger than 10−4. Below this value, in this example the results of the kernel matrix method are slightly lower than the reference mean rate densities.

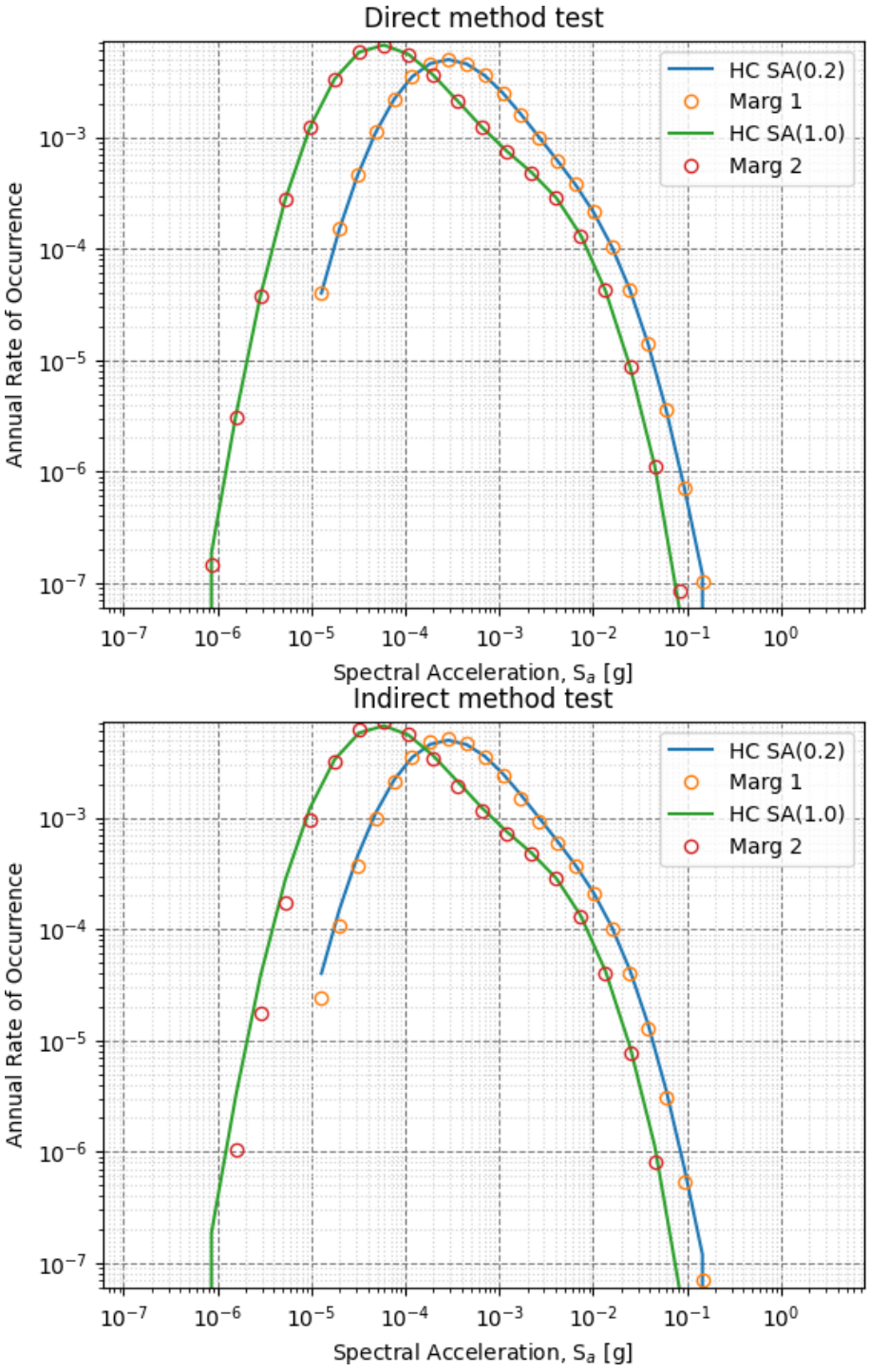


***Figure 2.1*** *– [top] Comparison between the scalar mean rate densities computed for the two selected IMTs and corresponding marginal obtained from the joint mean rate density (claculated with the direct method) [bottom] Similar comparison to the one in the upper panel. In this case the VPSHA was performed using the new kernel matrix method.*

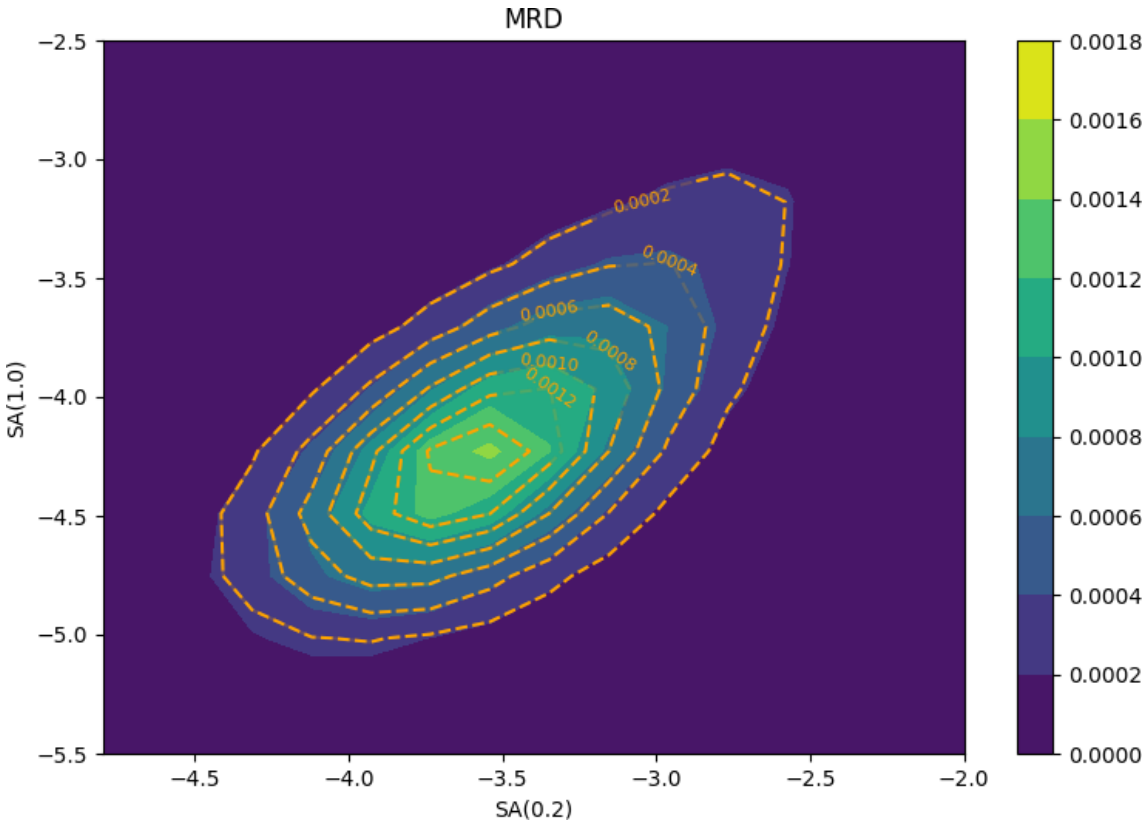


***Figure 2.2*** *– Comparison between the joint mean rate density computed with the direct (underlying mesh) and the kernel matrix method (orange dashed contours).*

Figure 2.2 displays a comparison between the joint mean rate densities computed with the direct and kernel matrix method. The one computed with the direct method is represented with contours filled while the one obtained with the kernel matrix method is represented with dashed orange contours. Overall, the two distributions show a quite similar pattern although some differences can be observed for low intensity measures levels and around the peak of the distribution approximatively located at (-3.5, -4.25) (note that the axes represent the logarithm - base 10 - of ground motion).

# 3. Conditional Spectrum: Methodology and OpenQuake-Engine Implementation

PSHA is a methodology that computes seismic hazard by combining the contributions of all earthquake scenarios and associated ground motion variability of engineering importance for the structure located at the investigated site. The relevance of a rupture from an engineering perspective, depends on its magnitude, its distance from the site (i.e. these two variables determine the intensity of the shaking at the site for a given rupture) and the frequency of occurrence compared to the expected lifetime of the structure.

One of the main disadvantages of PSHA, is that the hazard results (e.g. Uniform Hazard Spectrum, hazard curves) do not have a one-to-one relationship with an earthquake with well defined properties (e.g. rupture geometry, earthquake size) or to an earthquake recording [e.g Reiter, 1991]. This ineffectiveness poses problems to all those analyses requiring strong-ground motion time-histories (e.g. site-response analyses, non-linear dynamic analyses of structures).

## 3.1. Methodology

Lin et al. [2013a] summarises the various approaches available for the calculation of the CS. Overall, they propose four approaches with an increasing level of complexity going from the first one to the last one.

### 3.1.1. Method four

This method computes the exact value of the CS by integrating contributions from the ruptures in the realizations admitted by the Hazard Input Model logic tree.

To compute the CS mean and standard deviation we implemented in the OQ Engine equations 14 and 15 of Lin et al. [2013a]. In particular, the CS, $\mu_{\ln Sa(T_i)|\ln Sa(T^*)}$, corresponds to the weighted sum of the conditional spectra computed for all the ruptures generated by the various realisations defined by the logic tree

$$\mu_{\ln Sa(T_i)|\ln Sa(T^*)} = \sum_k \sum_j p_{j,k} \mu_{\ln Sa_{j,k}(T_i)|\ln Sa(T^*)} \tag{3.1}$$

where $i$ is a variable defining the period of the spectral acceleration, $k$ is the index of the realisation, $j$ is the index of a rupture in model $j$ and $p_{j,k}$ is the contribution of

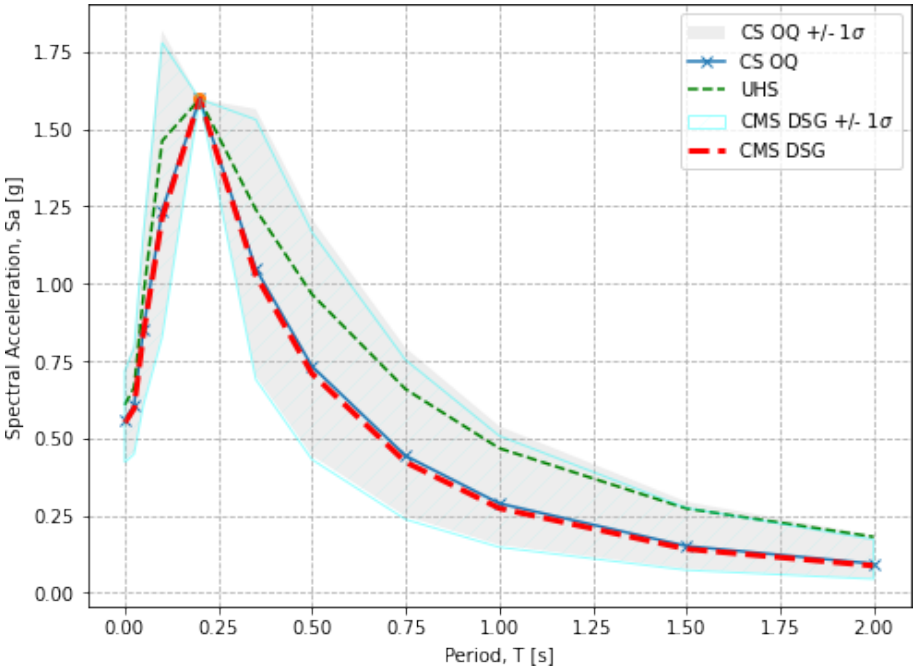


***Figure 3.1*** *– Comparison between the CS mean and standard deviation and the corresponding CS results computed with the OQ Engine. The green dashed curve shows the Uniform Hazard Spectrum with 10% probability of exceedance in 50 years.*

rupture $j$ in model $k$ to the total rate of exceedance of the target spectral acceleration at period $T^*$.

The mean and standard deviation of each rupture in each realization of the logic tree corresponds to:

$$\mu_{\ln Sa_{j,k}(T_i)|\ln Sa(T^*)} = \mu_{\ln Sa_{j,k}(Rup_j,T_i)} + \rho(T^*,T_i)\epsilon_j(T^*)\sigma_{\ln Sa_k(Rup_j,T_i)} \tag{3.2}$$

## 3.2. OpenQuake Engine Implementation

The module in the OQ Engine implementing the “Aggregation Approach to Method 4 (Exact CS)” in Lin et al. [2013a] is the ‘cond_spectra’ one, available in the ‘calc’ folder within the OQ Engine hazard library.

The main function (i.e. the one that can be called from a python script) is the cond_spectra one.

## 3.3. Examples and Tests

To illustrate the capabilities of the CS methodology implemented, we present herein a few application examples.

### 3.3.1. Example 0

The first example considers one of the simplest Hazard Input Model conceivable and is intended to test the basic implementation of the CS in the OQ Engine. This contains

one characteristic fault source with a magnitude frequency distribution comprising a single magnitude value. The Hazard Input Model does not include epistemic uncertainties. The ground-motion model used is the one of Chiou and Youngs [2008].

We compute the CS for one target IMTs – spectral acceleration for a vibration period of 0.2s – and one Annual Probability of Exceedances (ApEs) corresponding to 0.002105 i.e. a 10% probability of exceedance in 50 years. In this test, we expect the CS mean and standard deviation computed with the OQ Engine to be close to the ones obtained following a simpler a calculation approach based on the disaggregation results.

Figure 3.1 shows a comparison between the CS mean and standard deviation computed using the results of disaggregation and the CS computed with the OQ Engine. As expected, the mean and standard deviation computed with the two approaches are nearly identical. It is worth to note that in order to achieve a good match between the results, in the calculation of results based on disaggregation we had to use almost identical values of magnitude, distance and other explanatory variables usually non considered in the disaggregation of hazard.

### 3.3.2. Example 1

The second example is also based on a model without epistemic uncertainties. The Hazard Input Model contains one simple fault source generating earthquakes with magnitude between 6.5 and 7.0. The ground-motion model used is the one of Chiou and Youngs [2008].

We compute the CS for one target IMTs – spectral acceleration for a vibration period of 0.2s – and one ApEs corresponding to an annual probability of exceedance of 0.002105.

Figure 3.2 shows a comparison between the CS mean and standard deviation computed using the disaggregation results and the CS computed with the OQ Engine.

### 3.3.3. Example 2

The last example presented contains epistemic uncertainties in the Seismic Source Characterization and Ground-Motion Characterization logic tree. This is the case in which we expect the largest differences between the CS mean and standard deviation computed with the methodology implemented in the OQ Engine and the CS mean and standard deviation.

The Seismic Source Characterization contains two models for the same fault. In the first model the Magnitude-Frequency Distribution (MFD) is a double truncated Gutenberg-Richter distribution with with magnitude comprised between 6.5 and 7.0, Gutenberg-Richter b-value (b-value) of 0.9 and b-value corresponding to 4.2. In the second case the b-value is 1.0 and Gutenberg-Richter a-value (a-value) 3.9. The weights of the models are 0.6 and 0.4, respectively.

The Seismic Source Characterization comprises two GMMs, the one of Chiou and Youngs [2008] and the model developed by Campbell and Bozorgnia [2008] with a weight of 0.7 and 0.3, respectively.

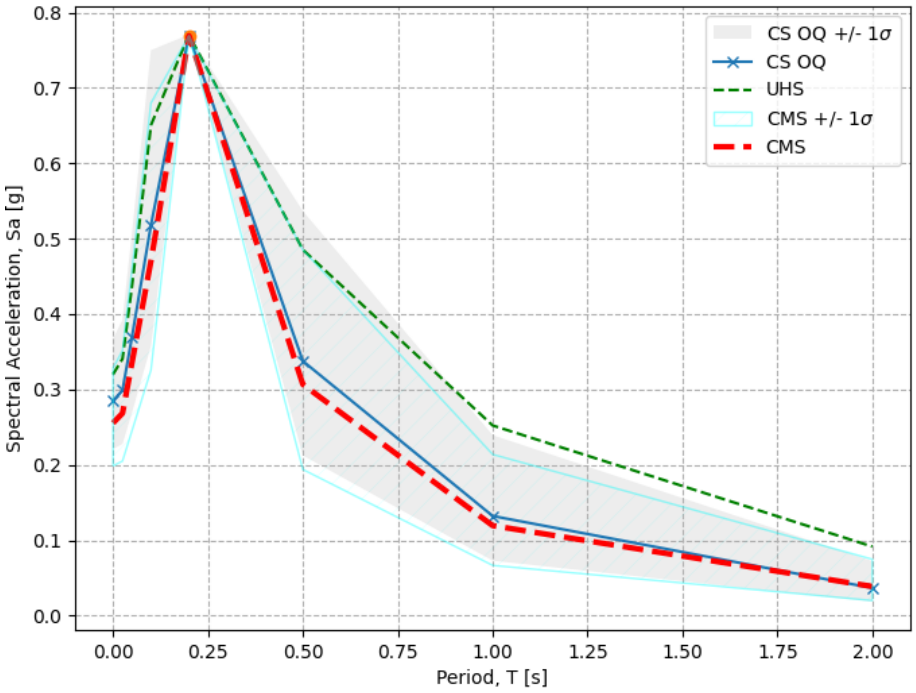


***Figure 3.2*** *– Comparison between the CS mean and standard deviation and the corresponding CS results computed with the OQ Engine. The green dashed curve shows the Uniform Hazard Spectrum with 10% probability of exceedance in 50 years.*

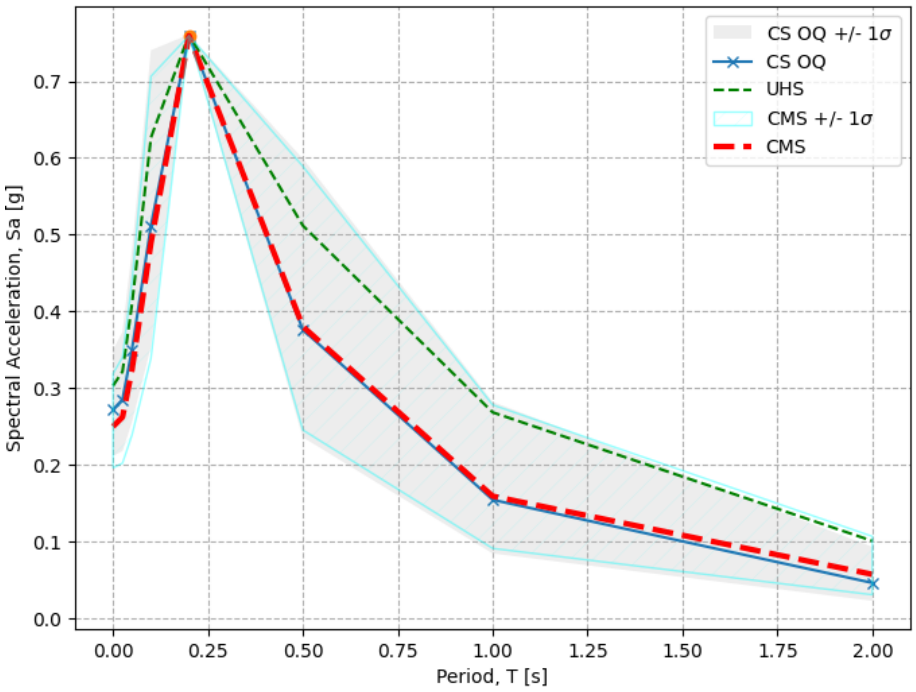


***Figure 3.3*** *– Comparison between the CS and its standard deviation and the corresponding CS computed with the OQ Engine. The green dashed curve shows the Uniform Hazard Spectrum with 10% probability of exceedance in 50 years. The mean magnitude and distance (closest distance to the rupture) computed are 6.77 and 33.55 km.*

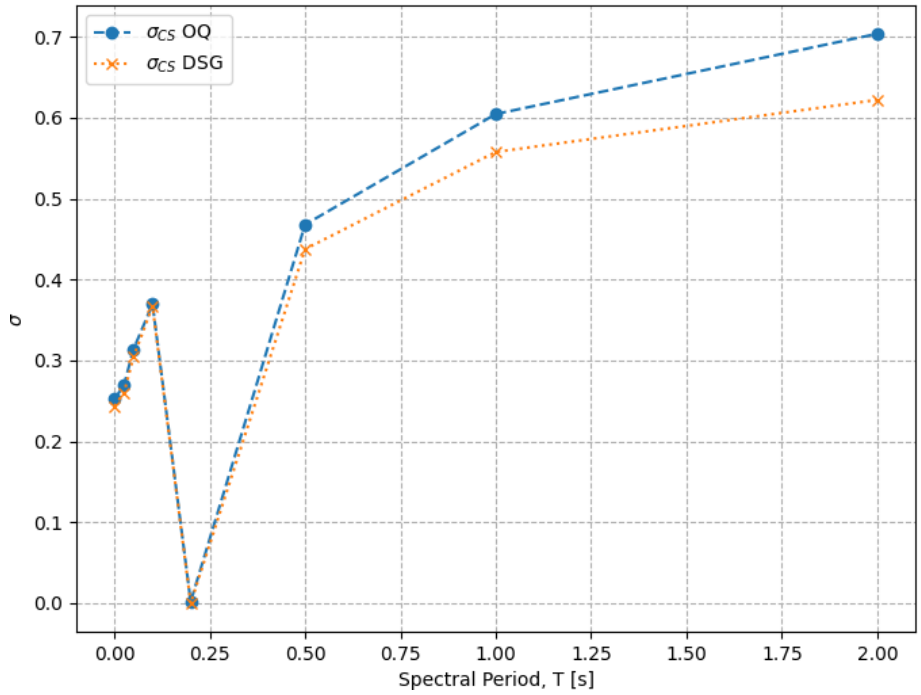


***Figure 3.4*** *– Comparison between the CS standard deviation obtained using the disaggregation results and the corresponding CS computed with the OQ Engine. The mean magnitude and distance (closest distance to the rupture) computed are 6.75 and 34.22 km.*

Figure 3.3 shows a comparison between the CS mean and standard deviation computed using the results of disaggregation and the CS computed with the OQ Engine.

Figure 3.4 shows a comparison between the CS standard deviation computed with the OQ Engine following method 4 of Lin et al. [2013b] and the standard deviation computed using the results of the disaggregation analysis. In this example, the differences between the two results are moderate.

# Part III

# Task 4.4.2: Modelling earthquake sequences for considering aftershocks

# 4. Work Package 4.4.2: Classical PSHA with Aftershocks

## 4.1. Introduction

### 4.1.1. Purpose

The broad purpose of Work Package 4.4.2 is to investigate the impacts of including aftershocks into PSHA, and to compare the results provided by different teams implementing different approaches. In the past, there were several attempts to consider aftershocks in risk assessments. These works consisted in simply adding the aftershock to the mainshock hazard. However, this does not correctly represent the physics and required safety parades. The combined hazard due to main and aftershock needs to be accounted for in site specific risk assessment for nuclear safety assessments.

### 4.1.2. Types of aftershock PSHA

Aftershocks are not typically considered in regional, long-term PSHA, either in classical PSHA studies or in stochastic event-based studies [e.g. Pagani et al., 2014, Assatourians and Atkinson, 2019]. Therefore, there is no methodological standard for including aftershocks. However, two methods have been proposed previously, using stochastic event-based approaches. A third method, the incorporation of aftershocks into classical PSHA, has been developed and implemented for this project.

Another category of aftershock PSHA exists, where the analysis is restricted to a single mainshock, chosen deterministically (and often chosen based on a recent or historical event), with stochastic aftershocks. Examples of this kind of analysis include studies by Yeo and Cornell [2009] and Gee et al. [2022]. This type of study, focusing on a single mainshock, is fundamentally different than the regional, long-term PSHA methods described in this document which integrate over tens of thousand to millions of mainshocks, which occur probabilistically.

#### Sequence-based Aftershock PSHA

The most straightforward way of including aftershocks in PSHA is to create long synthetic catalogs of mainshocks (for example from a standard seismic source model), and then add sequences of aftershocks to each mainshock. Then, the hazard can be calculated as from a typical stochastic event set analysis. This changes both the

spatial distribution and magnitude-frequency distribution of the seismic source model to reflect the additional density of earthquake production, particularly around sources capable of producing larger earthquakes with more populous aftershock sequences.

### Epidemic-Type Aftershock Sequences (ETAS)

A modification of the Sequence-based Aftershock PSHA methods is to incorporate the ETAS concept from statistical seismology [Ogata, 1988], that each earthquake creates a sequence of aftershocks, each of which may stimulate its own sequence of aftershocks. A stimulated or triggered aftershock has a small chance of being larger than its progenitor, and therefore generating a much more populous aftershock sequence than the one that birthed it. In this way, a very small background rate of 'independent' earthquakes may generate a much richer synthetic seismic catalog, comparable to the one that is observed [Sornette and Werner]. By creating many long catalogs that, in principle, fully sample or approximate all ruptures represented in a seismic source model with representative rates, a regional stochastic time-dependent PSHA model that incorporates aftershocks may be implemented [Field et al., 2017, Iacoletti et al., 2022].

### Classical PSHA with Aftershocks

A third, and somewhat different approach, is to incorporate aftershocks directly into a classical PSHA model. This idea was developed for this project, and so has not seen the level of exploration or implementation that the others have.

The fundamental idea of Classical PSHA with Aftershocks is to identify the 'time-independent' rate that each rupture would occur as an aftershock of each mainshock rupture in the source model, and from this calculate the annual occurrence rate of the aftershock rupture. There are two ways of doing this: The first is to create aftershock ruptures for each mainshock rupture *de novo*; the second is to create persistent ruptures (geometries and magnitudes), so that the final occurrence rate of any aftershock rupture is the sum of the rates that it would occur as an aftershock for all mainshock ruptures. This work takes the latter approach, and extends it by re-using each mainshock ruptures as potential aftershock ruptures, based on the concept that all possible damaging earthquakes are already modeled in the seismic source model. Therefore, no new ruptures are needed. This has both benefits and drawbacks to the implementation, as discussed below in Section 4.2.2.

### Comparison

The first two methods have theoretical differences but in practice are quite similar. The second is clearly a modification of the first.

There may be no difference in catalogs generated from a standard stochastic event set with aftershocks and one generated from ETAS methods; indeed, one hopes that if each is well calibrated to the observed seismicity that the studies are supposed to replicate, both fit quite well. The difference is in the internal consistency of the methods; ETAS methods are thought by many to be more complete or realistic in that much or most of the seismicity is self-generating.

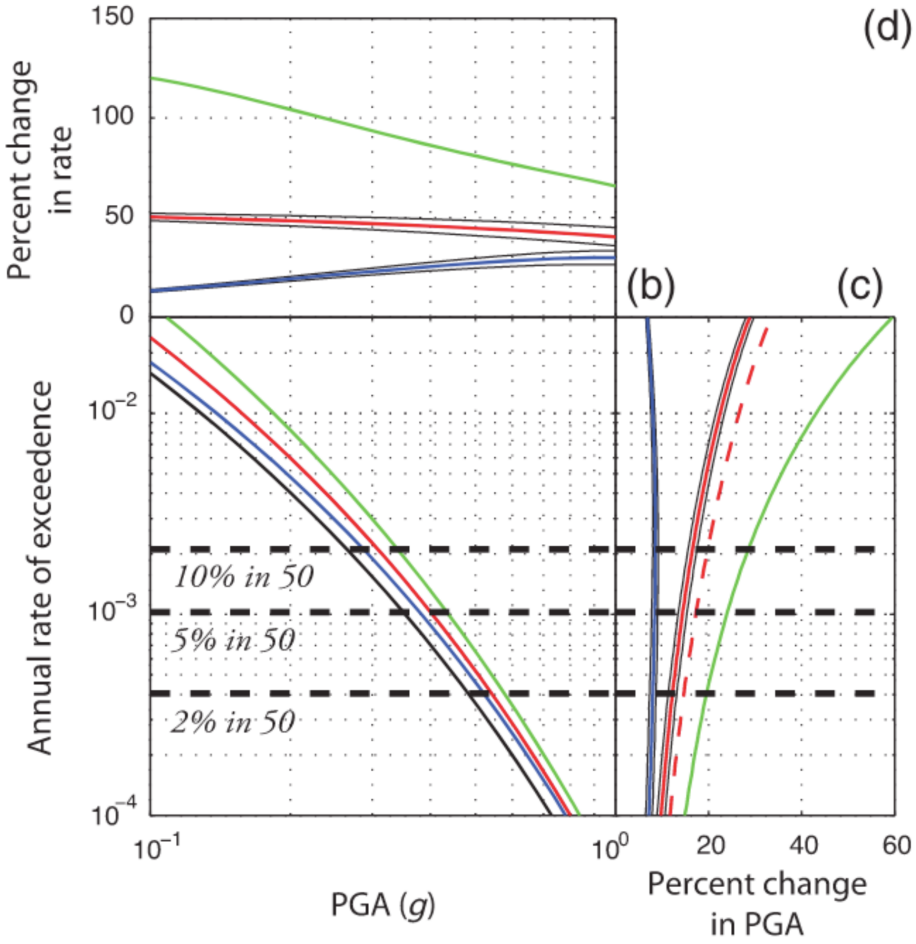


***Figure 4.1*** *– Comparison between hazard results from different methods by Boyd [2012] for San Jose, California. The black curve represents the mainshock-only hazard. The green curve represents the hazard when the source model is built from an undeclustered catalog. The blue curve represents the hazard from a clustered model, where aftershocks are temporally associated with the mainshocks. The red curve represents a stochastic event-based approach where the aftershocks are temporally independent from the mainshocks, which is in principle most similar to the methods developed here.*

However, from the perspective of computing seismic hazard, the methods are exactly the same once a stochastic event set has been generated.

The third method diverges from the other two more substantially, both in theory and in computational considerations.

The primary difference between the sequence-based methods and the classical approach is that in the former, the aftershocks are temporally associated with the mainshocks, while in the latter, they are not. Interestingly, this does not affect the broader statistics of earthquake occurrence. Instead, the difference manifests in the statistics of ground motion occurrence within a given time window. Because the time windows over which hazard is typically calculated (which may be years to decades for a long-term PSHA) are long with respect to the time difference between a mainshock and the vast majority of its aftershocks, and because most aftershocks are smaller than their mainshocks and therefore are likely to cause lower ground motions than the mainshock, it is rare that the ground motions from aftershocks will exceed those simply

from mainshocks in the time windows of interest, so the resulting hazard curve (the ground motion exceedances) are only slightly increased by considering aftershocks.

In contrast, with the classical PSHA approach, because of the temporal dissociation of mainshocks and aftershocks built into the classical time-independent methodology, aftershocks may cause an increase in ground motion exceedances because they can occur 'at any time' (in a statistical sense), so some fraction of the time they will be occurring in different time windows than the mainshocks.

Therefore, we consider *a priori* that the Classical PSHA with Aftershocks methods developed here are a maximum credible estimate of the hazard, compared to the other methods. This interpretation is consistent with the analysis by [Boyd, 2012], who compared similar methods (with some important differences) to the non-ETAS sequence–based method (Figure 4.1).

On the other hand, the classical methodology offers a major benefit compared to the sequence-based methods, that is of particular value for PSHA of long-lived critical infrastructure. This benefit is that the computation is relatively fast, once the aftershock rates have been calculated. Classical PSHA methods are quite fast even for very long return periods, because they integrate directly over the set of ruptures (regardless of their rates of occurrence) and do not require proportionally long stochastic event sets to evaluate rare ruptures.

Finally, though this was not directly implemented as part of this project, it is possible to generate stochastic event sets from the source characterization and methods implemented as part of the Classical PSHA with Aftershocks approach, with some minor code additions to deal with the time decay of aftershock production. This will be described more fully below.

## 4.2. Implementation of Classical PSHA with Aftershocks

The Classical PSHA with Aftershocks method considers all of the aftershocks to occur on existing ruptures within the seismic source model. This removes the need to do additional source characterization beyond defining basic statistical parameters for aftershock generation. Therefore any existing seismic hazard model suitable for classical PSHA in OpenQuake can be run with aftershocks considered as well.

### 4.2.1. Theory

#### Aftershock productivity

Aftershocks are assumed to follow a truncated Gutenberg-Richter distribution [e.g. Felzer et al., 2004], where total number of earthquakes $N$ above some magnitude $M$ is quantified as

$$N = 10^{a-bM}. \tag{4.1}$$

$N$ is a function of the mainshock magnitude $M_{main}$:

$$N(M_{main}) = c10^{\alpha M_{main}} \tag{4.2}$$

[Felzer et al., 2004] where $\alpha$=1.0, and $c \approx 0.02$, based on the global analysis by Felzer et al. [2004]. $c$ is a scalar aftershock productivity parameter, while $\alpha$ controls the scaling with mainshock magnitude.

Please note that Equation 4.1 is the number of aftershocks integrated over the long-term, so that it is a time-independent parameter (and therefore suitable for our time-independent methods. Our internal testing has demonstrated that the decay rate of aftershocks will not affect the results in the time-independent classical PSHA; in effect, an arbitrarty time window sampling a sequence of mainshocks and aftershocks will contain the same number of events, regardless of the aftershock decay rate: If the decay rate is short relative to the window size, the window will capture almost all of the aftershocks that occur from mainshocks within the time window, but very few aftershocks from mainshocks before the time window; however, if the decay rate is long compared to the time window, the window may not contain all of the aftershocks of earthquakes that occur within the time window, but it will contain many aftershocks that occur from earthquakes before the time window.

The $a$, $b$, and $M_{max}$ parameters for the truncated Gutenberg-Richter distribution must be defined as well. $b$ is calibrated based on observed seismicity, jointly with $c$, as described in Section 4.2.2. $M_{max}$ is more challenging to determine from seismicity data, as it is necessarily tied to the declustering process and explicit choices about whether a smaller event shortly before a larger event is a foreshock, or whether it is a mainshock with a larger aftershock. To be clear, this classification does not come from the data but is an explicit and somewhat semantic decision made by the modeler. Nonetheless, there are two obvious choices for the aftershock $M_{max}$ parameter: it may be the regional mainshock $M_{max}$ (in which case it will be as large as larger as any given mainshock), or it may be related to the mainshock magnitude, for example equal to that magnitude (or perhaps smaller by a delta, as following Båth's law). $a$ is calculated algebraically by relating Equations 4.1 and 4.2.

#### Aftershock occurrence rates for a single mainshock

After the magnitude-frequency distribution for each mainshock has been defined, all the ruptures within some distance (e.g., four times the mainshock rupture length) of the mainshock are selected for consideration as aftershocks.

The conditional occurrence probabilities for each aftershock rupture given the mainshock rupture are calculated based on the aftershock magnitude-frequency distribution and the set of ruptures within the distance threshold. The criteria involved in this determination are the distances from the mainshock rupture to each aftershock rupture, and the number of potential aftershock ruptures within each magnitude bin in the magnitude-frequency distribution.

For each aftershock rupture, an un-normalized conditional rupture probability $p'(rup_i|main)$ is defined based on the distance $d_i$ from the mainshock rupture to the $i^{th}$ aftershock rupture (accounting for the finite geometries of each), so that the probability decreases exponentially with distance:

$$p'(rup_i|main) = e^{d_i} \tag{4.3}$$

This probability is then normalized by the number of expected aftershocks in that magnitude bin $N(M_{aft})$ as derived from Equation 4.2, and the sum of the un-normalized

probabilities for all ruptures in the magnitude bin,

$$p(rup_i|main) = p'(rup_i|main) \cdot \frac{N(M_{aft})}{\sum p'(rup)} \tag{4.4}$$

such that the sum of the normalized rupture probabilities is equal to $N(M_{aft})$.

Finally, the unconditional occurrence rate $r(rup_i)$ of some aftershock rupture $i$ is calculated as the conditional occurrence probability times the occurrence rate of the mainshock:

$$r(rup_i) = p(rup_i|main) \cdot r(main)\,. \tag{4.5}$$

#### Aftershock occurrence rates for all mainshocks

The total aftershock occurrence rates for each rupture, considering all mainshocks, is simply the sum of the $r(rup_i)$ values for every mainshock in the model (or, any mainshock above a threshold value for generating aftershocks). These rates are added to the occurrence rate of each mainshock before the classical PSHA computation is performed.

### 4.2.2. Implementation details

In practice, the computation of hazard including aftershocks is split into two calculations. The first is the calculation of the aftershock occurrence rates, which requires the estimation of aftershock productivity parameters. The second is the actual hazard calculation.

Calculating aftershock occurrence rates is much more computationally-intensive than calculating the hazard (which takes the same time and computational resources as the hazard calculation without the aftershocks, as the addition of aftershock occurrence rates is instantaneous). However, the calculation of aftershock occurrence rates only needs to be done once after the mainshock seismic source characterization is complete, and it may be split into separate calculations as well. The calculation of aftershock occurrence rates can and should be thought of as a step in the seismic source characterization rather than in the hazard calculation.

The code to calculate the aftershock occurrence rates is included in the OpenQuake Model Building Toolkit, and the code to add these rates to the source model during the calculation of seismic hazard is included in the OpenQuake Engine. However, the interface for the hazard calculations in the OpenQuake Engine is not finalized and so it is still a bit complicated to perform the analysis.

#### Estimation of aftershock productivity parameters

The first step in the process of calculating the aftershock occurrence rates is to estimate the aftershock productivity parameters. This step should come after the seismic catalog has been declustered, and requires the association of aftershocks removed in the declustering process with the causative mainshock. This association is not always part of the declustering procedure, depending on the methods used, so additional post-declustering analysis of the mainshock and aftershock catalogs may be necessary.

We use a nonlinear Monte Carlo inversion for the aftershock productivity parameters $b$ and $c$, defined in Equations 4.1 and 4.2, respectively.

First ranges are defined for $b$ and $c$, and then some large number of samples of each are drawn from uniform distributions based on those ranges. Each sample value of $b$ is paired with a sample value of $c$.

For each $(b, c)$ pair, we iterate over all of the mainshock-aftershock clusters from the observed seismic catalog, and calculate the likelihood of observing the empirical magnitude-frequency distribution of the cluster given the truncated Gutenberg-Richter distribution for the values of $M_{main}$, $b$, and $c$ as described in Section 4.2.1, considering the year of the mainshock and the catalog completeness for that year. The total likelihood of each $(b, c)$ pair is the geometric mean of the $(b, c)$ pair likelihoods for each mainshock cluster.

This method produces a distribution of likelihoods, so that the single best (most likely) estimate for $(b, c)$ can be chosen, or numerous values can be sampled proportional to their likelihood (i.e., Bayesian sampling) in order to incorporate epistemic uncertainties into the aftershock production rates.

Currently, the aftershock productivity parameter estimation methods are not implemented directly into the OpenQuake Engine or Model Building Toolkit, but simply in a standalone Python script.

### Calculation of aftershock occurrence rates

The calculation of aftershock occurrence rates follows this procedure: First, pairwise distances are computed between mainshock ruptures (i.e., all ruptures in a seismic source model above some magnitude threshold) and potential aftershock ruptures. Then, the mainshock ruptures are iterated over and the unconditional aftershock occurrence rates are calculated for each mainshock rupture, following the mathematics described in Section 4.2.1.

*Rupture distance calculations*

Calculating the pairwise rupture distances is a very computationally expensive procedure for nontrivial hazard models. The number of ruptures in a typical source model is often in the millions, and may be 10 or 100 times that for the most complex national– to continental–scale models. Furthermore, the distance between the ruptures that we are interested in is the minimum difference between the finite rupture surfaces, not the hypocenters. The most straightforward way to calculate the minimum distance between the two rupture surfaces is to create a mesh of points representing each rupture surface, calculate the pairwise distance between the points in each set, and retain the minimum of this.

If a source model has 1,000,000 ruptures and 100,000 of those are above the threshold magnitude for calculating aftershocks (for example $M \geq 6$, where $M_{min} = 5$ for the whole model), and the surface meshes of each rupture has a mean of 50 points, then the total number of distance calculations is $(50 \cdot 1,000,000) \cdot (50 \cdot 100,000) = 250,000,000,000,000$ (i.e., $2.5 \times 10^{14}$ or 250 quadrillion) distance calculations. Obviously in order for this to be a tractable computation, efforts have to be made to cut orders of magnitudes of calculations off of this total, and to ensure that the remaining calculations are as efficient as possible.

The reduction in the number of calculations is to filter the source model, so that

sources that are farther than some maximum distance are not incorporated. If a seismic source model has many moderate-sized sources, such as models that are based around fault sources, this filtering greatly reduces the number of extraneous computations. If a model has very large area or multipoint sources, then this reduction is not particularly effective. Further efforts in this direction on splitting up large sources may greatly enhance the efficacy, but that sort of optimization is out of the scope of the METIS project.

The efficiency of the computations that must be done is also of prime importance. The major work towards efficiency is the coding of the algorithms for calculating distances in *Numba*, a compiled subset of the Python language that allows for low-level memory and CPU management similar to C or Fortran. This allowed us to implement very efficient multithreaded algorithms for the distance calculations, in a manner which is not possible with traditional Python and Numpy. These algorithms also include aggregation of rupture mesh point sets to maximize multicore CPU utilization, as well as multi-stage reduction of intermediate results, to reduce RAM usage.

The final major development was to use a strategy of caching rupture distances to the hard drive using an HDF5 format so that the distance calculations (which are independent of the aftershock production parameters) need only be calculated once, and an exploration of the parameter space may then be done without the need to replicate these expensive computations.

*Aftershock rate calculations*

The calculation of aftershock rates is less intensive than the calculatio of rupture distances, but is still significant for most hazard models, particularly if one wants to explore the impacts of different values of the aftershock productivity parameters.

Much like the code written for the rupture distance calculations, the most critical parts of the aftershock rate calculations are implemented using Numba, and optional caching of intermediate results to disk using HDF5 is implemented.

### Generation of stochastic event sets with aftershocks

The methods developed here for Classical PSHA with Aftershocks may be easily modified or extended to support the creation of stochastic event sets for single mainshock-aftershock sequences or for model-wide sequences. All that is needed is the calibration of the time-decay parameter of aftershock generation (which is not necessary for the time-independent classical method).

For any mainshock taken from the seismic source model, the code is in place to calculate the conditional earthquake occurrence probabilities for each potential aftershock ($p(rup_i|main)$ from Equation 4.4).

If a specific mainshock rupture is selected, for example for a single event-based seismic hazard and/or risk analysis, then the set of potential ruptures can be selected and the $p(rup_i|main)$ values for all ruptures can be calculated. Then, a temporal sequence can be constructed by modifying the $p(rup_i|main)$ values based on the time since the mainshock. Finally, stochastic sampling based on the final probabilities can be easily performed through standard methods.

If a model–wide stochastic event set including aftershocks is desired, then first a mainshock stochastic event set can be created from the seismic source model using existing tools in the OpenQuake Engine. Then, for each of those mainshocks, the

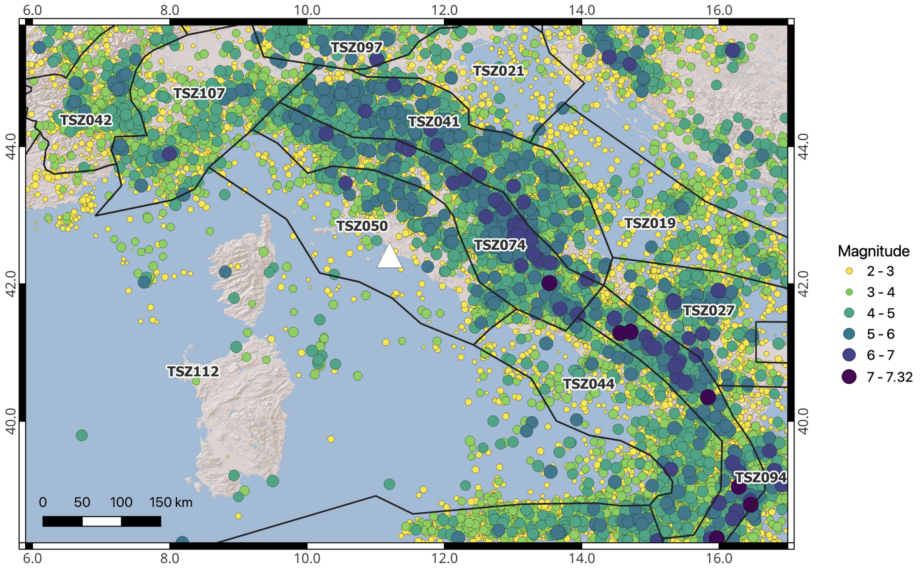


***Figure 4.2*** *– Map of the METIS test site and surrounding region. Seismicity is from the catalog compiled for the project. The polygons denote the source regions, with labels from the ESHM20 super zones. The white triangle denotes the test site.*

aftershock sequences can be created as described in the previous paragraph.

## 4.3. Results for the METIS Test Site

Though a number of PSHA models were used during the development of the code, the results presented here are for the METIS test site, which is located in the Lazio region of Italy. See Deliverable D4.7 for more details.

### 4.3.1. METIS test model

The METIS test model is a relatively simple model developed for this project, in order to test various methods rather than to implement the most robust and complete hazard model for the region (which would be unwieldy for methodological development).

The model is based on a smoothed seismicity model based around a new catalog, derived from the CPTI15 v.5 and HORUS catalogs (Figure 4.2). The seismic source zones are from ESHM20 super zones. The sources are made from smoothed seismicity from the declustered catalog. The aftershocks removed from the catalog are used for estimating aftershock productivity parameters.

### 4.3.2. Estimating aftershock production parameters

Aftershock productivity parameters were estimated through the methods described in Section 4.2.2. 10,000 samples were drawn from a uniform distribution of $b$-values from 0.0 to 1.5, and $c$ values from 0.0 to 2.0. The results are shown in Figure 4.3. It is clear that there is a swath of higher-likelihood values with low $c$-values ($c \leq 0.25$), and corresponding $b$-values above 0.75. There is a negative correlation between $b$ and $c$

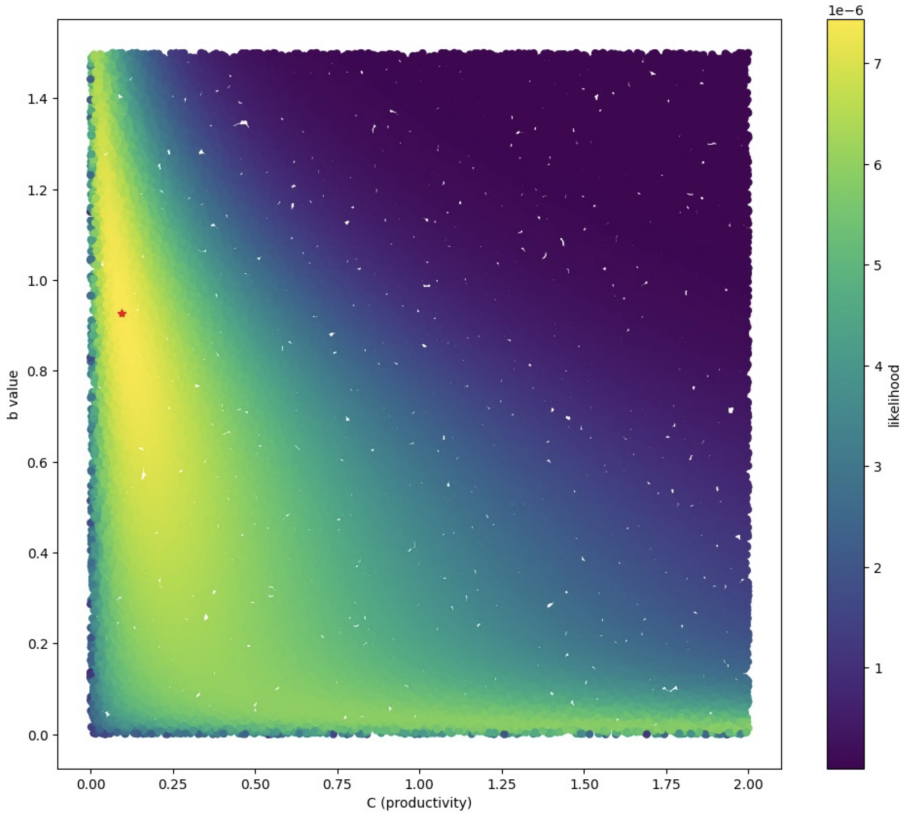


***Figure 4.3*** *– Likelihoods for the aftershock productivity parameters b and c based on the mainshock-aftershock clusters for the METIS test model. The star represents the maximum-likelihood sample.*

values, as high values for both of these lead to excessive generation of aftershocks as compared to the observations. The maximum-likelihood result is $b \approx 0.93$ and $c \approx 0.1$

#### Tradeoff between aftershock *b*-value and $M_{max}$

As mentioned previously, the upper bound for each aftershock magnitude-frequency distribution needs to be given to the model, but this is not necessarily available from the data. One result of our analysis (Figure 4.4) is that if the aftershock $M_{max}$ is set at a regional $M_{max} = 7.5$, regardless of the magnitude of the mainshock, the *b*-value is held consistent with the mainshock *b*-value (if they are specified to be the same during the modeling process). However, if the $M_{max}$ is set at $M_{main}$ for each aftershock sequence but *b*-value is still set at a constant, there is in reality a change in the regional *b*-value because of oversampling of small to moderate aftershocks from small to moderate mainshocks. This is consistent with the global analysis by Felzer et al. [2004] even though it was not part of our *b*-value estimation algorithm, which was based on clusters in which the largest earthquake was the mainshock.

### 4.3.3. Aftershock hazard at the METIS test site

Hazard curves at the METIS test site with and without aftershocks are shown in Figure 4.5. It is clear from the results that the inclusion of aftershocks substantially increases

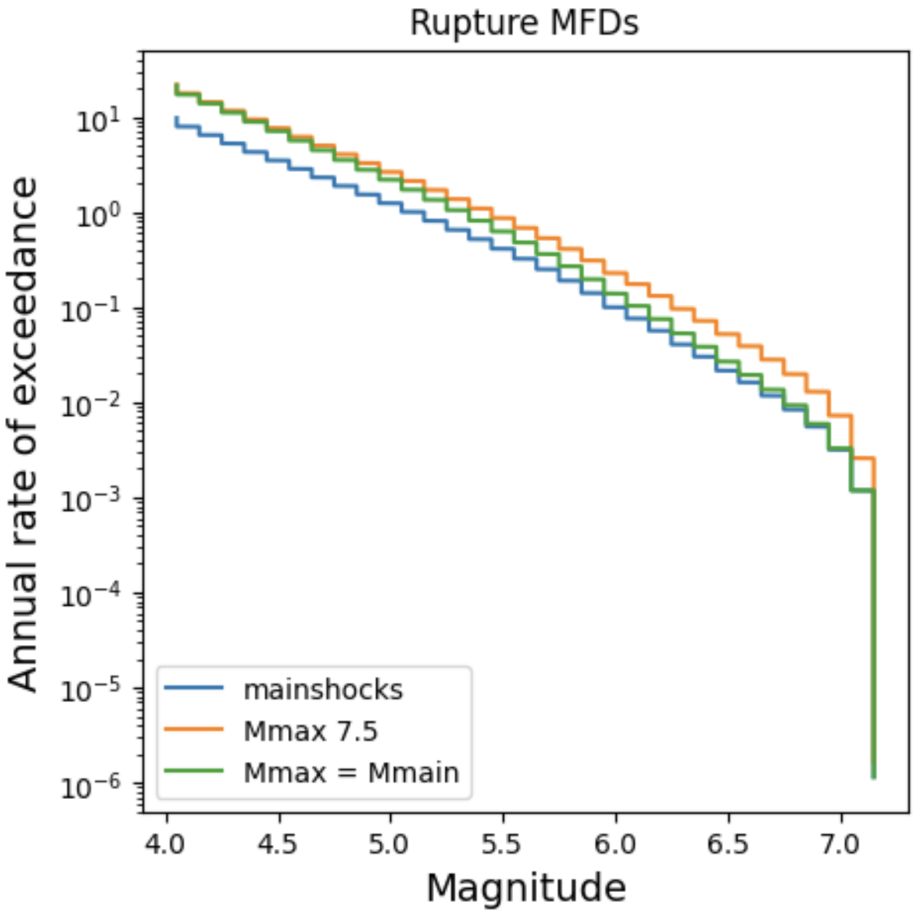


***Figure 4.4*** *– Magnitude-Frequency Distributions for the METIS test model, considering mainshocks only (blue), mainshocks and aftershocks with an aftershock $M_{max}$ = 7.5, and an aftershock $M_{max}$ = $M_{main}$ for each mainshock.*

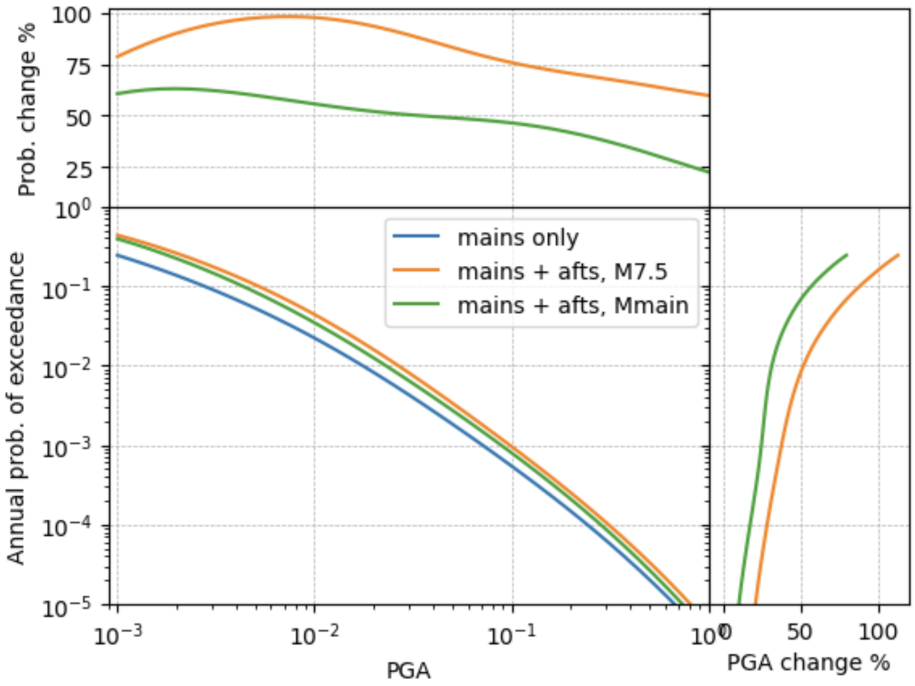


***Figure 4.5*** *– Hazard curves for the METIS test site, considering three scenarios: The mainshock-only hazard curve is shown in blue. The scenario with the aftershock $M_{max}$ = 7.5 for all sequences is shown in orange. The scenario with the aftershock $M_{max} = M_{main}$ is shown in green.*

the hazard, though the amount of the increase differs based on the return period of interest as well as the aftershock $M_{max}$. At the shortest return periods, the percentage in the increase of PGA is up to 75% if aftershock $M_{max} = M_{main}$, and over 100% if aftershock $M_{max} = 7.5$. The fractional change decreases with increasing return period and increasing base PGA values, but at the longest return periods the increase is still 25%–50%.

These changes in PGA is quite high. Further analysis is needed to determine whether the magnitude of the changes is a natural result of the increase in seismicity considered in the hazard model regardless of temporal associations, or whether having time-dependence (temporal clustering) would reduce the increase in hazard, as in the study by Boyd [2012]. The generation of stochastic event sets based on these methods, which is not yet implemented but may be (as described in Section 4.2.2), should aid in the understanding of the importance of the temporal component.

Furthermore, the seismicity produced by seismic source model is parameterized based on analysis of the mainshock and aftershock seismic catalogs for this test site, but limited efforts were put into comparing the final modeled seismicity with the observations. This will be a necessary step to perform before the methods developed here should be used for real hazard calculations.

## 4.4. Conclusions and Future Work

Methods were developed and implemented for performing Classical PSHA with Aftershocks using the OpenQuake Engine and the OpenQuake Model Building Toolkit.

These methods were then used on the METIS test site in Lazio, Italy. In contrast to some previous analysis and perhaps expectations, the results indicate a substantial increase in seismic hazard due to the incorporation of aftershocks. However, questions remain as to how much this large increase is due to embedded assumptions of the Classical PSHA approach (in particular the temporal dissociation of mainshocks and aftershocks), or whether it may be an accurate model. Additional work verifying seismicity generation and exploring the impacts of temporal clustering should be performed before this analysis becomes routinely employed.

The methodology here proposed provides an upper bound to the seismic hazard it can be expected for a given site when both the contributions of mainshocks and aftershocks are taken into account.

This appoach does not fulfil the requirements of hazard studies for nuclear facilities that account for the aftershock contributions although it provides components that can be used in more tailored approaches that are still under discussion within the METIS project. In particular, these tailored analyses must account for the change in the status of the plant after the occurrence of the mainshock (i.e. shutdown status rather than regular operation status) and the possible change in the damage/vulnerability functions of some components in cases where the mainshock is partially damaging them.

Based on the discussions entertained until the completion of this document, there are various approaches that could be tested. One of them implies computing the probability of exceeding a given value of shaking within a short time-window located just after the occurrence of the mainshock for a selected set of scearios. A different - and more general - approach might compute conditional hazard curves for a suite of mainshock cases.

# Bibliography


N. A. Abrahamson, W. J. Silva, and R. Kamai. Summary of the ASK14 Ground Motion Relation for Active Crustal Regions. *Earthquake Spectra*, 30(3):1025–1055, aug 2014. doi: 10.1193/070913eqs198m.

Karen Assatourians and Gail M Atkinson. Implementation of a smoothed-seismicity algorithm in monte carlo psha software eqhaz and implications for localization of hazard in the western canada sedimentary basin. *Seismological Research Letters*, 90(3):1407–1419, 2019.

G. M. Atkinson, J. J. Bommer, and N. A. Abrahamson. Alternative Approaches to Modeling Epistemic Uncertainty in Ground Motions in Probabilistic Seismic-Hazard Analysis. *Seismological Research Letters*, 85(6):1141–1144, November 2014. ISSN 0895-0695. doi: 10.1785/0220140120.

J. Baker, B. Bradley, and P. Stafford. *Probabilistic seismic hazard and risk analysis*. Cambridge University Press, Cambridge New York, NY, 2021. ISBN 9781108348157.

P. Bazzurro and C. A. Cornell. Vector-valued probabilistic seismic hazard analysis. In *Proceedings 7th U.S. National Conference on Earthquake Engineering*, 2002.

P. Bazzurro, J. Park, and P. Totong. Vector-Valued Probabilistic Seismic Hazard Analysis of Correlated Ground Motion Parameters. Technical report, Report submitted to the USGS, 2010. URL https://earthquake.usgs.gov/cfusion/external_grants/reports/G09AP00135.pdf.

J. J. Bommer. Challenges of Building Logic Trees for Probabilistic Seismic Hazard Analysis. *Earthquake Spectra*, 28(4):1723–1735, nov 2012. doi: 10.1193/1.4000079.

J. J. Bommer and F. Scherbaum. The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis. *Earthquake Spectra*, 24(4):997–1009, November 2008.

Oliver S Boyd. Including foreshocks and aftershocks in time-independent probabilistic seismic-hazard analyses. *Bulletin of the Seismological Society of America*, 102(3): 909–917, 2012.

R. J. Budnitz, G. Apostolakis, D. M. Boore, S. L. Cluff, K. J. Coppersmith, C. A. Cornell, and P. A. (1997) Morris. Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and the Use of Experts. Technical Re-

port NUREG/CR-6372, Nuclear Regulatory Commission, Washington D.C., United States, 1997. Two volumes.

K. W. Campbell and Y. Bozorgnia. NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5Ranging from 0.01 to 10 s. *Earthquake Spectra*, 24(1):139–171, 2008.

B. S.-J. Chiou and R. R. Youngs. An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *Earthquake Spectra*, 24: 173–215, 2008.

C. A. Cornell. Engineering Seismic Risk Analysis. *Bulletin of the Seismological Society of America*, 58:1583–1606, 1968.

Karen R Felzer, Rachel E Abercrombie, and Goran Ekstrom. A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America*, 94(1):88–98, 2004.

Edward H Field, Kevin R Milner, Jeanne L Hardebeck, Morgan T Page, Nicholas van der Elst, Thomas H Jordan, Andrew J Michael, Bruce E Shaw, and Maximilian J Werner. A spatiotemporal clustering model for the third uniform california earthquake rupture forecast (ucerf3-etas): Toward an operational earthquake forecast. *Bulletin of the Seismological Society of America*, 107(3):1049–1081, 2017.

Robin Gee, Laura Peruzza, and Marco Pagani. The power of the little ones: Computed and observed aftershock hazard in Central Italy. *Earthquake Spectra*, 38(1):702–724, 02 2022. ISSN 8755-2930. doi: 10.1177/87552930211036913. URL https://doi.org/10.1177/87552930211036913.

Z. Gülerce and N. A. Abrahamson. Vector-Valued Probabilistic Seismic Hazard Assessment for the Effects of Vertical Ground Motions on the Seismic Response of Highway Bridges. *Earthquake Spectra*, 26(4):999–1016, 2010. doi: 10.1193/1.3464548.

Salvatore Iacoletti, Gemma Cremen, and Carmine Galasso. Validation of the epidemic-type aftershock sequence (etas) models for simulation-based seismic hazard assessments. *Seismological Society of America*, 93(3):1601–1618, 2022.

I. Iervolino, M. Giorgio, C. Galasso, and G. Manfredi. Conditional hazard maps for secondary intensity measures. 100(6):3312–3319, 2010. ISSN 0037-1106. doi: 10.1785/0120090383.

N. Jayaram and J. W. Baker. Statistical tests of the joint distribution of spectral acceleration values. *Bulletin of the Seismological Society of America*, 98(5):2231–2243, 2008. doi: 10.1785/0120070208.

S. R. Kotha, G. Weatherill, D. Bindi, and F. Cotton. A regionally-adaptable ground-motion model for shallow crustal earthquakes in Europe. *Bulletin of Earthquake Engineering*, 18(9):4091–4125, 2020. doi: 10.1007/s10518-020-00869-1.

R. B. Kulkarni, R. R. Youngs, and K. J. Coppersmith. Assessment of confidence intervals for results of seismic hazard analysis. In *Proceedings of the Eighth World Conference on Earthquake Engineering*, pages 263 – 270, 1984.

B. Li and Z. Cai. A Simplified Method for Performing Vector-Valued Probabilistic Seismic Hazard Analysis. 113(1):348–360, 2023. doi: 10.1785/0120220138.

T. Lin, S. C. Harmsen, J. W. Baker, and N. Luco. Conditional Spectrum Computation Incorporating Multiple Causal Earthquakes and Ground-Motion Prediction Models. *Bulletin of the Seismological Society of America*, 103(2A):1103–1116, mar 2013a. doi: 10.1785/0120110293.

Ting Lin, Curt B. Haselton, and Jack W. Baker. Conditional spectrum-based ground motion selection. part II: Intensity-based assessments and evaluation of alternative target spectra. *Earthquake Engineering & Structural Dynamics*, 42(12):1867–1884, may 2013b. doi: 10.1002/eqe.2303.

R. K. McGuire. *Seismic Hazard and Risk Analysis*. Earthquake Engineering Research Institute, 2004.

Yosihiko Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27, 1988.

M. Pagani, 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*, 85(3):692–702, May 2014. ISSN 0895-0695, 1938-2057. doi: 10.1785/0220130087.

L. Reiter. *Earthquake Hazard Analysis*. Columbia University Press, 1991. 254 pages.

F. Scherbaum and N. M. Kuehn. Logic tree branch weights and probabilities: Summing up to one is not enough. *Earthquake Spectra*, in press, 2011.

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

D. Sornette and M. Werner. Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, båth's law, and observed aftershock sequences. 110(B8). doi: 10.1029/2004jb003535.

J. Woessner, D. Danciu, D. Giardini, H. Crowley, F. Cotton, G. Grünthal, G. Valensise, R. Arvidsson, R. Basili, M. Demircioglu, et al. The 2013 European Seismic Hazard Model: Key Components and Results. *Bulletin of Earthquake Engineering*, 13(12): 3553–3596, 2015.

Gee Liek Yeo and C Allin Cornell. A probabilistic framework for quantification of aftershock ground-motion hazard in california: Methodology and parametric study. *Earthquake Engineering & Structural Dynamics*, 38(1):45–60, 2009.

# Appendices

# A. Computing the CS with the OQ Engine

The OQ Engine computes CSs using an ad-hoc calculator which follows a computational flow similar to the one used for the calculation of hazard results with a "classical" approach. This CS calculator is intended to be used for a limited number of sites. Computing a CS with the OQ Engine requires a hazard input model. For a description of the structure and formats used to describe a hazard input model for OQ Engine the reader can refer to the section Using the Hazard Module of the OpenQuake Engine Manual.

## A.1. Configuration file

There are two main chnages in the configuration file required for the calculation of the CS with the OQ Engine.

The first change is for the `calculation_mode` parameter which in must be equal to `conditional_spectrum`. This parameter is generally included in the "general section". The A.1 listing contains an example of the `general` section in a configuration file for the OQ Engine that can be used for the calculation of a CS.

The second modification pertains an ad-hoc set of parameters essential for the calculation of the CS. These include:

- `cross_correlation` Defines the model computing the correlation coefficient between residuals for two spectral periods.
- `imt_ref` Defines target intensity measure type $S_a(T^*)$
- `imt_ref` Defines the probabilities of exceedance in the `investigation_time` to be used for the calculation of the CSs

***Listing A.1*** *– Example of the general section in the OQ Engine configuration file*

```
[general]
description = Test 00
calculation_mode = conditional_spectrum
random_seed = 23
```

***Listing A.2*** *– Example of the conditional spectrum section in the OQ Engine configuration file*

```
[conditional_spectrum]
cross_correlation = BakerJayaram2008
imt_ref = SA(0.2)
poes = 0.002105, 0.000404
```

## A.2. Calculation

The calculation of the CS can be executed from the command line (or through the web interface) using the normal syntax used for more traditional analyses (e.g. `oq engine -run job.ini`).

## A.3. Output Files

The computed CS results can exported from the OQ Engine data-store with one simple command executed in a terminal. For example, a .csv file of the CSs computed in the last OQ Engine calculation can be exported in the `out` subfolder with the following command: `oq engine -eos <calculation_id> <directory>`. Amongst various files, in the `out` you will find one called `conditional-spectrum-<target_imt_id>_<calc_id>.csv` where `<target_imt_id>` is an integer specifying the target IMT and `<calc_id>` is the ID of the calculation executed with the OQ Engine.

The listing A.3 shows an example of the .csv file with the computed CS. The first row contains auxiliary information such as the version of the OQ Engine engine used, the coordinates of the site. The second row includes the names of the columns; `sa_period` refers to the vibratory period considered, while the `val<poe_id>` and `val<std_id>` columns contain the mean and the standard deviation of the conditional spectrum compute for the probability of exceedance with index `<poe_id>`. For example, `<poe_id>` equal to 0 indicates the first probability of exceedance assigned to the `poes` parameter in the OQ Engine configuration file.

More detailed information can be accessed in the OQ Engine data-store file (i.e. an .hdf5 formatted file - for more info see the hdf5 website). Information and examples on how to access information stored in the OQ Engine data-store are available in the "OpenQuake for Advanced Users" manual available at https://docs.openquake.org/oq-engine/advanced/.

***Listing A.3*** *– Example of .csv file containing the computed conditional spectrum*

```
#,,"generated_by='OpenQuake engine 3.15.0-git91dead60ba', ...
sa_period,val0,std0
0.00000E+00,5.59671E-01,2.67185E-01
2.50000E-02,6.05945E-01,2.84829E-01
5.00000E-02,8.53945E-01,3.26447E-01
1.00000E-01,1.23737E+00,3.85023E-01
2.00000E-01,1.59719E+00,0.00000E+00
3.50000E-01,1.04855E+00,4.00405E-01
5.00000E-01,7.33086E-01,5.00310E-01
7.50000E-01,4.43261E-01,5.80321E-01
1.00000E+00,2.89336E-01,6.21800E-01
1.50000E+00,1.51410E-01,6.64289E-01
2.00000E+00,9.42085E-02,6.80138E-01
```

# B. Computing VPSHA with the OQ Engine

The OQ Engine computes VPSHAs by post-processing the results provided by a traditional classical PSHA calculation. This VPSHA calculator is intended to be used for a limited number of sites. Computing a VPSHA with the OQ Engine requires a hazard input model. For a description of the structure and formats used to describe a hazard input model for OQ Engine the reader can refer to the section Using the Hazard Module of the OpenQuake Engine Manual.

## B.1. Configuration file

The OQ Engine configuration file required for performing a VPSHA contains the same parameters needed for a classical PSHA analysis (see for example Classical Area Source Demo Example) plus one additional section specifying the parameters for the vector component of the analysis. The listing B.1 shows an example of the additional section required. It contains only two parameters: `postproc_func` and `postproc_args`. The `postproc_func` parameter defines the type of post-processing analysis to be completed; in this case the value corresponds to `compute_mrd`. The parameter `postproc_args` includes the variables needed to specify the cross-correlation model (i.e. `cross_correlation`), the two IMTs that will be used to compute the MRD (`imt1`, `imt2`) and the edges of the bins defining the kernel matrix (see page 38).

***Listing B.1*** *– Example of the post-processing section in the OQ Engine configuration file required to perform the VPSHA analysis*

```
[postproc]

postproc_func = compute_mrd
postproc_args = {
  'imt1': 'PGA',
  'imt2': 'SA(0.05)',
  'cross_correlation': 'BakerJayaram2008',
  'seed': 42,
  'meabins': [0.1, 0.2, 0.3, 0.4, 0.5, 0.6],
  'sigbins': [0.2, 0.3, 0.4, 0.5, 0.6, 0.7]}
```

***Listing B.2*** *– Example of the general section in the OQ Engine configuration file. <filename> is the name of the datastore file (the usual format is like this: calc_1.hdf5)*

```
from openquake.commonlib.datastore import datastore
dstore = datastore.read(<filename>)
mrd = dstore['mrd'][:]
```

## B.2. Calculation

The calculation of VPSHA can be executed from the command line (or through the web interface) using the normal syntax used for more traditional analyses (e.g. `oq engine -run job.ini`).

## B.3. Output Files

The result of the VPSHA is a matrix with the MRD. This is available in the `mrd` dataset of the OQ Engine datastore.

We provide an example of how to read the `mrd` from the OQ Engine datastore in the B.2 listing.

# C. Approaches for Epistemic Uncertainty Propagation: some calculation examples

## C.1. Components of a logic tree

We give here a short illustration of the terminology used in the OpenQuake Engine to describe logic trees, since this is the platform used to implement the methods later described. The OpenQuake Engine [Pagani et al., 2014, Silva et al., 2014] (herein indicated with the acronym OQ), is an open-source code for seismic hazard and risk analysis developed the by Global Earthquake Model (GEM) Foundation in collaboration with various scientists and engineers.

In the OQ Engine, a Logic Tree (LT) is a hierarchical structure describing the epistemic uncertainties affecting different components or methods used to build a PSHA model. The uncertainty on a parameter or a modelling component (e.g. the declustering algorithm) is described by a set of mutually exclusive and collectively exhaustive branches [Bommer and Scherbaum, 2008] contained in a branch-set. The weights (or probabilities, e.g. Scherbaum and Kuehn [2011]) of the branches in a branch-set must sum to one.

Figure C.1 shows one example of a logic tree structure. The logic tree containing the epistemic uncertainties affecting the seismic source characterisation has always an initial uncertainty on the seismic source model where at least one initial set of sources is defined. In the schematized example considered in Figure C.1, this branch-set is represented by the branch with the 0.1 label.

An uncertainty (i.e. a branch-set) can be applied following various rules; for example, it can be applied to a single source or to a previous branch or to all the sources belonging to a tectonic region. In the latter case, when there is more than one source belonging to a given tectonic region, we have a case of correlated uncertainty. In this instance, the same uncertainty applies to a number of sources at the same time. This means that when a branch is selected during the processing of the logic tree, the parameter (or model) assigned to that branch must be used with all the sources sharing that uncertainty. A common example of correlated uncertainty is when epistemic uncertainty on the ground-motion is modelled using a multi-model approach [e.g. Atkinson et al., 2014]. In this case, when we compute hazard using one realisation of the ground-motion logic tree (i.e. we pick one ground-motion model for one tectonic region), the same model must be used for all the sources belonging to that

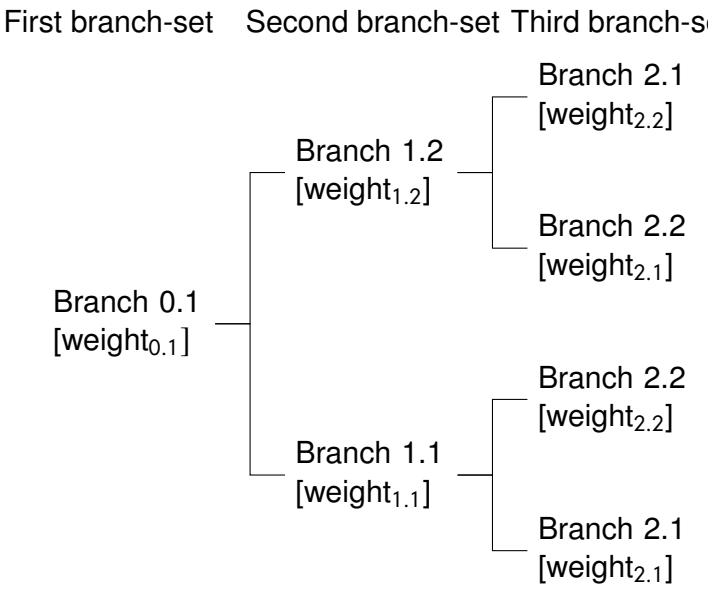


***Figure C.1*** *– Main components of a logic tree structure as defined in the OQ Engine.*

tectonic region. For example, if the branch-set modelling ground motion uncertainty for active shallow crust contains two models and applies to three sources, the number of realisations admitted will be two, one computed using the first ground motion model with all the three sources and the second one obtained with the second ground-motion model. On the contrary, if the three sources were assigned to different tectonic regions we would not have a case of correlated uncertainty. In this case, if for each tectonic region there was two ground-motion models, the total number of realisations would be eight.

## C.2. Logic Tree Sampling

The quality assurance tests included in the OQ Engine repository contain a number of examples of simple analyses where the various sampling strategies described in Chapter 1.

The case 71 shows an example of processing the logic tree that combines the Latin Hypercube sampling strategy with an early sampling strategy.

The case 7 shows the use of sampling the logic tree in an event-based hazard analysis also based on Latin Hypercube sampling strategy with an early sampling strategy.

Last processed: Wednesday 7th February, 2024 @ 12:23