REPORT
doi:10.26443/seismica.v5i1.1677

The OpenQuake Model Building Toolkit: A suite of tools for building components of a seismic hazard model

Marco Pagani [1], **Kirsty Bayliss** * [1], **Christopher Brooks** [1], **Kendra Johnson** [1], **Richard Styron** [1], **Manuela Villani** [1], **Yufang Rong** [2]

[1]Global Earthquake Model (GEM) Foundation, Pavia, Italy, [2]FM, Research Division, Norwood, MA 02062, USA

Abstract Building a probabilistic seismic hazard model is a complex task, requiring the integration of disparate datasets into one cohesive and comprehensive model. To facilitate this process, we have developed the OpenQuake Model Building Toolkit (OQ-MBTK), a collection of functions for constructing probabilistic seismic hazard models, enabling users to start from catalogue and fault data and sequentially step through the model building process to produce hazard inputs compatible with the OpenQuake (OQ) Engine. These tools allow users to build seismic source models that capture epistemic uncertainty using a logic tree, select suitable ground motions for different tectonic regions, and carry out thorough sensitivity analyses by easily investigating the consequences of model changes. Crucially, the toolkit ensures that data are treated consistently at all stages of the process. Using the MBTK to streamline the model building workflow can ensure it is reproducible and more robust.

Production Editor:
Andrea Llenos
Handling Editor:
Carmine Galasso
Copy & Layout Editor:
Aisling O'Kane

Signed reviewer(s):
Pablo Iturrieta, Peter Powers

Received:
March 31, 2025
Accepted:
February 14, 2026
Published:
April 20, 2026

1 Introduction

The essence of a Probabilistic Seismic Hazard Analysis (PSHA Cornell, 1968; McGuire, 2004; Baker et al., 2021) is the construction of the two key components: the Seismic Source Characterization (SSC) and the Ground-Motion Characterization (GMC). The former includes a comprehensive description of the position, geometry and seismogenic properties of all the sources and the associated epistemic uncertainties. The latter describes the models adopted to compute the ground motion at sites of interest and the related epistemic uncertainties. The SSC and GMC collectively form the PSHA input model.

Preparation of the SSC and GMC components entails collecting, pre-processing and homogenising earthquake and ground motion information into hazard model inputs. For the SSC, the primary data sources are earthquake catalogues, as well as geodetic (e.g., from space geodesy), geological, and paleoseismological information that shed light on active faults and deformation patterns (e.g., fault types, geometries, and slip rates). In the model building process, these data are assembled into a set of earthquake sources that describe the expected rupture geometries, as well as statistical representation of expected frequencies for the range of possible earthquake magnitudes. For the GMC, strong motion data are the most commonly used data, alongside site data (e.g., site class) at the recording station. These data are used to select and/or calibrate ground motion models that best reflect the shaking expected from earthquakes in the SSC.

The approaches and methodologies used in the model building process vary depending on the seismotectonic conditions and the source typologies in question. The traditional source typologies considered for the SSC include (1) distributed seismicity in the stable and actively deforming shallow crust; (2) shallow crustal faults and fault systems; and (3) subduction interface and in-slab sources, and there are complications and variations in each of these broad classes. For example, modelling distributed seismicity in stable areas often entails a baseline level of seismicity defined over broad areas that historically did not show significant levels of seismic activity (e.g., Johnston et al., 1994), whereas in active shallow crust areas, one challenge is how to define the spatial pattern of seismic activity based on the available information (i.e., past seismicity; geological and geodetic information) which may be incomplete. Data availability also affects the approaches used for the GMC. For example, when possible, ground motion models are developed or tailored based on data solely from the geographic and tectonic region under consideration, but when the seismic recordings are too scarce, globally applicable ones are used instead.

The case-by-case considerations required to build PSHA models call for a model building workflow that is modular, reproducible, and - ideally - linked with the code used for the calculation of hazard. Modularity ensures that alternative approaches and hypotheses can be tested or included in different branches of a logic tree structure without altering the overall workflow; a classic example is the use of different approaches for the declustering of seismicity (e.g., van Stiphout et al.,

*Corresponding author: kirsty.bayliss@globalquakemodel.org

2010). Reproducibility lends credibility and is a desirable characteristic because automated processes can be more easily used to explore various hypotheses and reconsider them in subsequent phases, such as when additional information becomes available. Moreover, reproducible models have more chance of being accepted by the broader scientific community and receiving contributions from scientists not directly involved in the original hazard model, and ultimately support the work of the experts requested to review the model. Finally, reproducibility helps to add tests to the model building process, thus reducing the possibility of introducing mistakes. The use of identical tools or functions (e.g., code) in both the model construction process and PSHA calculations helps to ensure that input models will be treated with the same assumptions in both phases. For example, small details concerning the binning of magnitude-frequency distributions (MFDs) can be overlooked when the input model is created separately from the PSHA codes. With a direct link between tools used for building models and the calculation of hazard, it is possible to share components and use the hazard calculation code to perform various tests while building the model.

Despite the considerable progress made in the last couple of decades in the community development of open-source software for the modelling of seismic hazard (e.g., Field et al., 2003; Pagani et al., 2014), to our knowledge, a comprehensive tool (or library) with these characteristics does not yet exist. There are numerous reasons that can explain this, though the most probable is a general preference in the hazard modelling community for developing in-house tools, to address more specific needs emerging in each project. This is certainly a valid strategy, though we think that new improvements should be brought back to a more general framework, making the use of new methods straightforward within other projects and by other modellers. With this objective in mind, we present the OpenQuake Model Building Toolkit (OQ-MBTK), hosted at https://github.com/GEMScienceTools/oq-mbtk. The OQ-MBTK is a suite of tools that collectively offers capabilities for building various components of a hazard model. In the following sections, we describe the packages currently included in the OQ-MBTK.

2 The OQ-MBTK pilot study and applications

The idea of a publicly accessible set of standard tools that could be used for constructing various input components of a hazard model was initially framed in the GEM pilot project called GEM1 in the early 2010s. However, the initial version of the OQ-MBTK was not completed until 2017 within a collaboration between researchers from GEM and FM (https://www.fm.com, one of the long-lasting private sponsors of the Global Earthquake Model initiative). It consisted of a set of Jupyter notebooks (see https://jupyter.org/) and Python scripts that together constituted a single workflow for the construction of a hazard input model, as described in Rong et al. (2017). Jupyter notebooks were used in the pilot study to make the tools accessible to hazard modellers who are not necessarily proficient at programming with the Python language (https://www.python.org/). Some of these older notebooks are still included in the OQ-MBTK for legacy reasons. The disadvantages of this approach included the need to maintain complicated software for running various Jupyter notebooks in a single run and an excessively large data structure which contained all the original information as well as the intermediate and final results of the workflow. The new OQ-MBTK tools instead provide functions that can be used more flexibly by importing directly in Python, in Jupyter notebooks, or directly from the command line, as described in Section 5. This allows modellers to more easily incorporate useful components into their workflows.

The OQ-MBTK was first used successfully by Rong et al. (2020) for developing a PSHA model for mainland China. Since then, and through several iterations, the tools have been used by the GEM hazard team in building hazard models for the Phillipines (Peñarubia et al., 2020), the Pacific Islands (Johnson et al., 2021), and the Dominican Republic (Johnson et al., 2024), as well as for constructing several other hazard models in the GEM Global Seismic Hazard Mosaic (Johnson et al., 2023, see https://hazard.openquake.org/gem/).

3 The current modules of the OQ-MBTK

The OQ-MBTK is the combination of several modules, each one dedicated to a specific task. A workflow showing the different modules and how they are used in the model building process is shown in Figure 1. The OQ-MBTK does not represent a finite product, but is instead a container for tools and functions that are constantly evolving. For this reason, the maturity and completeness of the different modules vary. The aim of this paper is to illustrate the main modules and the primary motivation behind this repository, to promote the use of the available tools, and to incentivize community contributions similar to the development style of the OpenQuake Engine (Pagani et al., 2014; Silva et al., 2014). When describing the OQ-MBTK modules, we focus on a select few that we believe will be of particular interest to readers, and we explain the criteria used for their development.

3.1 The Catalogue (CAT) module

The catalogue (CAT) module primarily provides functionalities for compiling a homogenised catalogue from a collection of catalogues that cover various time periods, geographic areas, and magnitude ranges. The most commonly used workflow involves four main steps and mostly resembles the procedure proposed in Weatherill et al. (2016). The user controls each step of the process through a .toml formatted configuration file (see https://toml.io/en/), an input format used in several OQ-MBTK modules.

The first step is to merge the original input catalogues

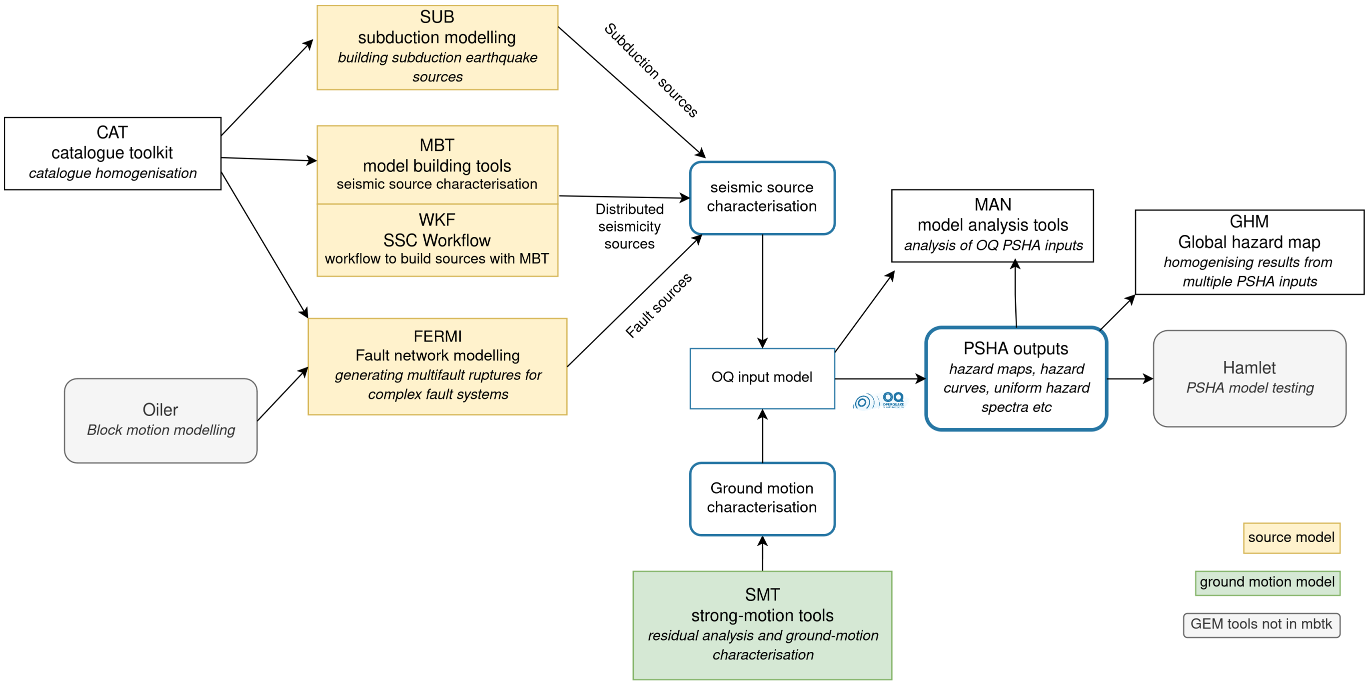

Figure 1 A workflow used by the GEM hazard team for building, testing, and applying PSHA input models, highlighting the modules of the OQ-MBTK

into a single organised collection. The formats accepted include the IASPEI Seismic Format (ISF, see http://www.isc.ac.uk/standards/isf/), the .ndk format used by the Global Centroid-Moment-Tensor (CMT) project (https://www.globalcmt.org/), and a generic .csv that is compatible with most common catalogue formats. The CAT creates an aggregated collection of seismological data by associating each new event with the ones already incorporated. This is done by defining tolerances in time and space for each new event and searching for previously added events that occurred in its surroundings. When a previous event falls into the defined space-time window, the old and the new event are merged into a group and a preferred solution based on the user's hierarchical settings is assigned. In the case of large catalogues, this procedure can be computationally demanding and time-consuming. To improve the efficiency of this process, the OQ-MBTK performs the search in space with the help of a spatial index.

The second step is selecting or deriving empirical equations for converting between different typologies of magnitude, as shown in Figure 2, and using them to homogenise the catalogue magnitudes (to use a consistent magnitude type throughout). This phase is required because the standard magnitude typology used by ground-motion models (and consequently in PSHA) is moment magnitude, while the typologies listed in catalogues vary depending on time and agency; only recent catalogues regularly report moment magnitude (e.g., ISC-GEM; see Storchak et al., 2013; Di Giacomo et al., 2018).

The third phase constructs the homogenised catalogue by selecting a single origin (time and location) and magnitude per earthquake. This is performed sequentially event by event, based on two lists – one for magnitude and one for location – ordered in decreasing order of preference. The CAT module also includes tools to check the final catalogue for remaining duplicate events by producing .geojson files for manual inspection and validation (for example inside a Geographic Information System), which help remove potential duplicates matched within wider time and space windows than the initial tolerances used to merge catalogues.

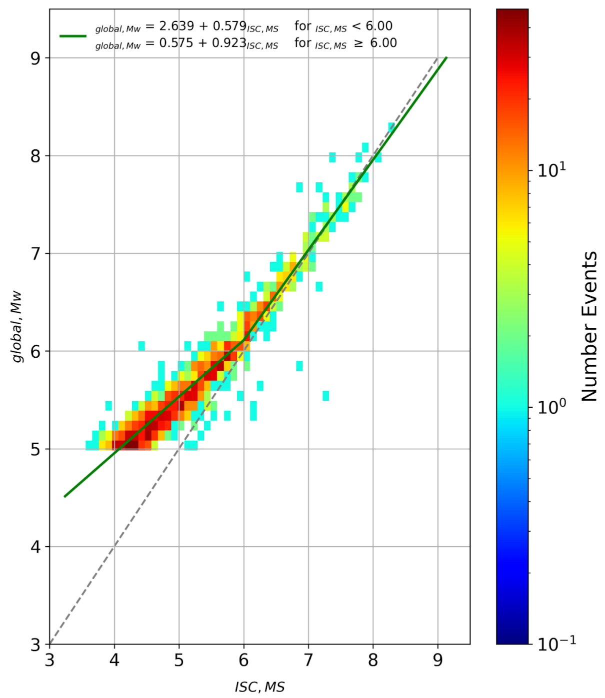

Figure 2 Example of a density plot showing the relationship between two types of earthquake magnitudes reported by two different agencies, and the relationship used to convert between them. The relationship here was derived by Weatherill et al. (2016).

3.2 The Model Building (MBT) module

The MBT module contains the implementations of many standard methods used in seismic source characterisation, as well as ad-hoc approaches developed by GEM, for constructing both fault and distributed source models. These tools operate on datasets formatted for the OQ-MBTK, including catalogues, fault databases, and strain rates, parsing these data and passing them to the fundamental steps of seismic source characterisation. For example, the MBT includes tools for analysing the completeness of a catalogue, deriving MFDs from catalogues (e.g., Weichert, 1980) or slip rates, and smoothing of seismicity based on past earthquake locations (e.g., Frankel, 1995; Helmstetter et al., 2007), as well as plotting functions that allow the user to easily visualise the characteristics of their data. Many of the MBT functions (among those for other modules) directly use the HMTK module of the OQ Engine, such as those for declustering and evaluating seismicity characteristics, ensuring consistency with GEM's older toolkits. Figure 3 shows examples of three such plots: histograms of the hypocentral depths for earthquakes within a catalogue (or sub-catalogue), the magnitude-frequency distribution (MFD) of that catalogue (modelled and observed), and the time-magnitude density plot of the catalogue used to derive the MFD. The information plotted in the left and centre panels is directly used in the seismic source produced by the MBT.

The MBT also includes tools for tectonic regionalisation following the approach described by Pagani et al. (2021). These tools isolate portions of catalogues, thereby 'classifying' the events, according to tectonic units delimited by user-specified 3D surfaces, e.g., existing Earth structure models, such as Slab2.0 (Hayes et al., 2018) and Lithos1.0 (Pasyanos et al., 2014), or ones developed by the user. This is a critical step when developing PSHA input models in complex tectonic regions such as subduction zones, where the tectonic units must be characterised separately.

3.3 The Model Building Workflow (WKF) module

The Model Building Workflow (WKF) module contains various methods for processing catalogues and building simple components of an SSC such as shallow crust distributed seismicity sources and unsegmented shallow crustal faults. The WKF allows the user to apply many of the functions in the MBT consecutively and to construct models from input using a single .toml file. This makes the creation of sources easily reproducible, simplifies the process of creating many logic tree branches, and allows for robust sensitivity analysis.

3.4 The Subduction (SUB) module

The SUB module offers tools for constructing subduction earthquake sources. The tools are iteratively improved alongside subduction zone science (e.g., allows to impose or exclude segmentation of the subduction interface) and to ensure the rupture geometries produced are compatible with PSHA applications (e.g., risk analysis; scenarios). The main functionalities unique to the SUB module (versus the MBT) are workflows for interactively defining the surface that represents the top of the slab from cross section visualisations of the slabs or directly from existing models (e.g., Slab2.0; Hayes et al., 2018), and for constructing in-slab sources following the methodology of Pagani et al. (2021). Naturally, the SUB module depends a lot on the MBT, and is in particular linked to the tectonic regionalisation. After regionalisation, the events classified as belonging to the slab and interface are used to define suitable MFDs for these sources, with the option of considering tectonic limits for the upper magnitudes. The SUB tools can then be used to generate the subduction sources, including 3D fault surfaces with floating ruptures for the interface, with appropriate MFDs derived from seismicity, tectonics or a hybrid model (i.e., a la Youngs and Coppersmith, 1985), and non-parametric sources (e.g., predefined rupture geometries and their rates) that model seismicity in the slab.

3.5 Fault Network Modelling (Fermi)

The Fault nEtwoRks ModellIng (Fermi) module, called FNM in the MBTK (for consistency with module 3-letter naming conventions), contains GEM's latest tools for modelling multi-fault ruptures in complex fault systems. These tools are inspired by the OpenSHA tools (e.g., Milner and Field, 2024) and SHERIFS (Chartier et al., 2019) approaches to earthquake rupture modelling given an input fault network, but with some added flexibility. A fault network is broken into sections of uniform length, where each section becomes a node in a graph. Then, using graph theory, all possible ruptures in the network are defined with an adjacency matrix, where further rupture plausibility constraints can be applied. Fermi can also determine rupture rates through inversion, offering options to fit these rates using slip rates and MFDs, either for individual faults or the entire fault system, with a variety of iterative solver options available. Further details on Fermi will be described in a separate paper.

3.6 The Strong-Motion Tools (SMT) module

The Strong-Motion Tools (SMT) module provides tools for selecting and comparing ground-motion models (GMMs), and supersedes the original (and now deprecated) GMPE-SMTK. The SMT comprises two submodules: the residuals module and the comparison module. The residuals module evaluates how well a given set of GMMs predict observed values of ground-shaking using a classical residual analysis. This submodule includes capabilities for plotting the distributions of residuals for a given GMM and intensity measure, and for computing GMM ranking metrics such as the stochastic area (Sunny et al., 2021) and the Euclidean distance-based metrics (Kale and Akkar, 2013). The comparison module compares GMMs using highly customisable ground-shaking scenarios (e.g., for which the user can specify all aspects of the rupture and sites that may be required

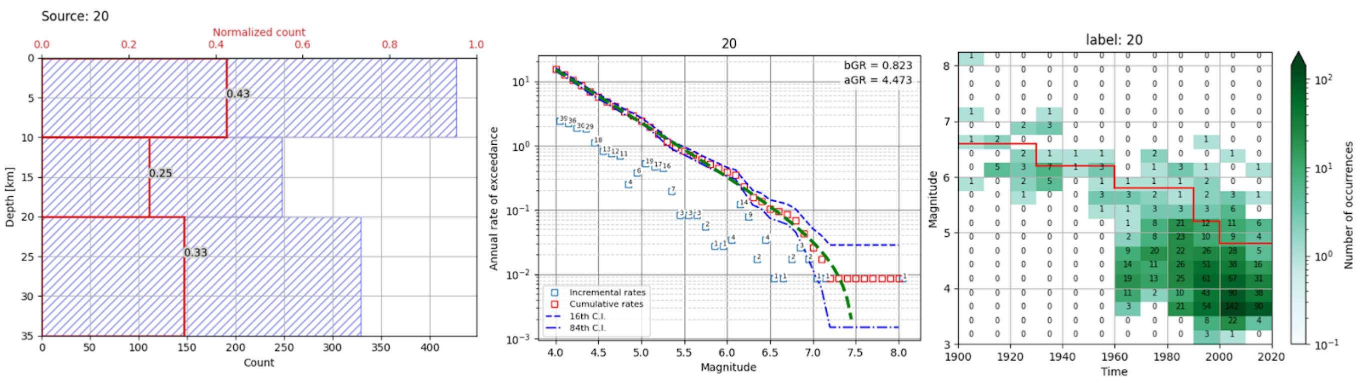

Figure 3 Examples of figures produced by the MBT module. Left: a histogram of depths for a subcatalogue that corresponds to one source zone in a model. Centre: the MFD for the catalogue, showing both incremental and cumulative observations and their confidence intervals (C.I.), and the derived MFD. Right: the time-magnitude density plot corresponding to the subcatalogue.

by a given GMM) by generating ground-motion scaling curves, response spectra (including plotting spectra from processed records against corresponding GMM predictions), Sammon's maps (Scherbaum et al., 2010), and dendrograms - a novel visual representation of the results of agglomerative clustering performed on median GMM predictions to help evaluate the degree of similarity between GMMs. Examples of the dendrograms and Sammon's maps are shown in Figure 4.

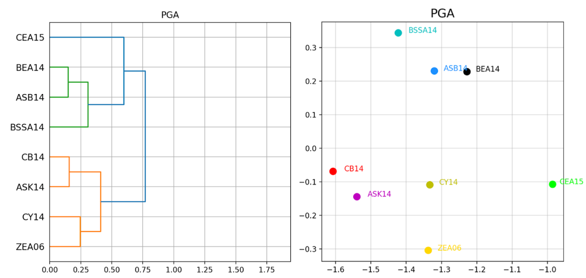

Figure 4 Example of (left) a dendrogram and (right) a Sammon's map that emphasize similarities and differences among the evaluated GMMs.

3.7 The Model Analysis (MAN) module

The MAN module contains tools for analysing the characteristics of an OQ Engine formatted input model. This module is primarily used internally by the GEM team and includes some more experimental functions that summarise model components, create representative plots of the main characteristics, or plot results; these are considered experimental because they may not work for all model instances (e.g., for the newest features supported by the OpenQuake Engine). It also requires the Generic Mapping Tools (Wessel et al., 2019, GMT), which is not a dependency of the OQ-MBTK and must be installed separately.

3.8 The Global Hazard Model (GHM) module

The GHM module collects tools for constructing global hazard maps using the hazard results computed as part of the GEM Mosaic standard calculations (see Pagani et al., 2020; Johnson et al., 2023). These tools are probably of low interest for the general audience, but they are distributed publicly nonetheless to ensure reproducibility. The main capabilities of these tools include:

- Identifying co-located hazard curves among a set of output files from different models
- Homogenising pairs (or more) of hazard curves across defined model boundaries
- Collecting hazard results (intensity measure levels) among many models with common criteria, including spectral period, reference site condition, and return period, into global .csv files
- Producing global seismic hazard maps from the above

The methodologies implemented are described in Pagani et al. (2020), with some more recent changes and improvements.

4 Coding principles and third party libraries

The OQ-MBTK is primarily coded in Python with some components written in Julia (https://julialang.org/). The plotting functionalities mostly rely on Python and Julia, while some use the Generic Mapping Tools (GMT; Wessel et al., 2019) and its Python and Julia wrappers. Note that since GMT is not an official dependency of the OQ-MBTK, some plotting functions may not work for all users. In addition, there is a command-line interface that provides access to most of the methods implemented in the various modules. The OQ-MBTK repository is under a Continuous Integration (CI) Environment, as is necessary for developing scientific code collaboratively (Silver, 2017). Overall, the package is extensively unit-tested and all code is subject to review among GEM members. The OQ-MBTK is released under a GNU Affero general public license 3.0.

5 Using the functions available in the OQ-MBTK

All the functions available in the various modules (with the current exception of Fermi) can be accessed through a command-line interface like the one available for the OQ Engine. By executing the command in Inset 1 the OQ-MBTK returns a list of available options, each one identifying a module of the list just described.

```
> oqm -h

usage: oqm [-h] {rep,wkf,cat,ccl,sub,unc,mbi}
     ...

positional arguments:
    {rep,wkf,cat,ccl,sub,unc,mbi}
        available subcommands; use oqm <
            subcmd> --help
options:
    -h, --help
```

Additional information about the functionalities within a module can be found by typing a command as exemplified in Inset 2.

```
> oqm cat -h

usage: oqm cat [-h]
    {MFDs_sample_mag_sigma,homogenise,
    create_figures,purge_earthquakes,
    check_duplicates,merge,
    completeness_analysis,
completeness_generate,create_csv}
               ...

positional arguments:
  {MFDs_sample_mag_sigma,homogenise,
  create_figures,purge_earthquakes,
  check_duplicates,merge,
  completeness_analysis,
  completeness_generate,
  create_csv}

    available subcommands;
        use oqm cat <subcmd> --help

options:
  -h, --help
```

Finally, information about a specific command can be obtained as in Inset 3

```
> oqm cat merge -h

usage: oqm cat merge [-h] settings

Merges the information contained in a
number of catalogues. The output is a
couple of .h5 files (you can read them
using pandas.read_hdf) which contain
the origins and the magnitudes of
the earthquakes in the catalogues
specified in the settings.

positional arguments:
  settings  .toml file with the settings

options:
  -h, --help
```

For documentation and end-to-end examples, see the OQ-MBTK github page (https://github.com/GEMScienceTools/oq-mbtk/), which contains simple applications of many functions as well as detailed notebooks that demonstrate the full workflow used to build subduction sources with the SUB module and perform strong-motion analysis with the SMT.

6 Conclusions

The OQ-MBTK is a comprehensive suite of tools designed to facilitate the construction and analysis of PSHA input models. This toolkit has already been instrumental in building several seismic hazard models, demonstrating its robustness and utility. However, the OQ-MBTK is not a static product; it is a dynamic and evolving collection of tools and functions. This continuous evolution ensures that the toolkit remains relevant and up-to-date with the latest advancements in seismic hazard modelling. By providing a modular and reproducible workflow, the OQ-MBTK allows for the exploration of various hypotheses and the inclusion of new methodologies without altering the overall process. The modularity and reproducibility are critical for maintaining the scientific rigour and acceptance of the models within the broader community. We hope that this contribution will inspire further engagement from the PSHA community. Through the development and sharing of the OQ-MBTK, we aim to foster a collaborative environment where improvements and innovations can be shared and integrated into a general framework. Such collaboration will not only enhance the toolkit but also contribute to the advancement of seismic hazard modelling as a whole. We invite researchers and practitioners to explore its capabilities and contribute to its development.

Acknowledgements

We thank Franco Tamanini, Harold Magistrale, Hosam Ali and Elliot Klein for internal reviews of this paper. We thank reviewers Peter Powers and Pablo Iturrieta for substantially improving the clarity of this paper, and Pablo Iturrieta for suggestions that also helped us improve the code. We thank all private and public sponsors of the GEM Foundation (www.globalquakemodel.org) for their support.

Data and code availability

The OQ-MBTK repository is available on GitHub at the following link https://github.com/GEMScienceTools/oq-mbtk. At the time of submission, Version 0.9.0 is archived at https://doi.org/10.5281/zenodo.15111998. The OQ-MBTK can be installed on most common operating systems (OS) including macOS, Linux, and Microsoft-related OS. The OpenQuake Engine (see https://github.com/gem/oq-engine) is a dependency of the OQ-MBTK.

Competing interests

The authors acknowledge that there are no conflicts of interest recorded.

References

Baker, J., Bradley, B., and Stafford, P. *Seismic Hazard and Risk Analysis*. Cambridge University Press, 1st edition, 2021. doi: 10.1017/9781108425056.

Chartier, T., Scotti, O., and Lyon-Caen, H. SHERIFS: Open-Source Code for Computing Earthquake Rates in Fault Systems and Constructing Hazard Models. *Seismological Research Letters*, 2019. doi: 10.1785/0220180332.

Cornell, C. A. Engineering seismic risk analysis. *Bulletin of the seismological society of America*, 58(5):1583–1606, 1968. doi: 10.1785/BSSA0580051583.

Di Giacomo, D., Engdahl, E. R., and Storchak, D. A. The ISC-GEM Earthquake Catalogue (1904–2014): Status after The Extension Project. *Earth System Science Data*, 10:1877–1899, 2018. doi: 10.5194/essd-10-1877-2018.

Field, E. H., Jordan, T. H., and Cornell, C. A. OpenSHA: A developing community-modeling environment for seismic hazard analysis. *Seismological Research Letters*, 74(4):406–419, 2003. doi: 10.1785/gssrl.74.4.406.

Frankel, A. Mapping seismic hazard in the central and eastern United States. *Seismological Research Letters*, 66(4):8–21, 1995. doi: 10.1785/gssrl.66.4.8.

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

Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. High-resolution time-independent grid-based forecast for M>= 5 earthquakes in California. *Seismological Research Letters*, 78(1):78–86, 2007. doi: 10.1785/gssrl.78.1.78.

Johnson, K., Villani, M., Bayliss, K., Brooks, C., Chandrasekhar, S., Chartier, T., Chen, Y., Garcia-Pelaez, J., Gee, R., Styron, R., Rood, A., Simionato, M., and Pagani, M. Global Earthquake Model (GEM) Seismic Hazard Map (version 2023.1 - June 2023). Unpublished, 2023. doi: 10.5281/zenodo.8409647.

Johnson, K., Chartier, T., Pagani, M., et al. Probabilistic seismic hazard analysis for the Dominican Republic. *Earthquake Spectra. 0(0)*, 2024. doi: 10.1177/87552930241263618.

Johnson, K. L., Pagani, M., and Styron, R. H. PSHA of the southern Pacific Islands. *Geophysical Journal International*, 224(3): 2149–2172, March 2021. doi: 10.1093/gji/ggaa530.

Johnston, A. C., Coppersmith, K. J., Kanter, L. R., and Cornell, C. A. The Earthquakes of Stable Continental Regions (No. *TR-1022261-V1). EPRI*, 1994.

Kale, O. and Akkar, S. A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): The Euclidean distance-based ranking (EDR) method. *Bulletin of the Seismological Society of America*, 103(2A):1069–1084, 2013. doi: 10.1785/0120120134.

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

Milner, K. R. and Field, E. H. A Comprehensive Fault-System Inversion Approach: Methods and Application to NSHM23. *Bulletin of the Seismological Society of America*, 114(1):486–522, 2024. doi: 10.1785/0120230122.

Pagani, M., Monelli, D., Weatherill, G., Danciu, L., Crowley, H., Silva, V., Henshaw, P., Butler, L., Nastasi, M., Panzeri, L., Simionato, M., and Vigano, D. OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*, 85:692–702, 2014. doi: 10.1785/0220130087.

Pagani, M., Garcia-Pelaez, J., Gee, R., Johnson, K., Poggi, V., Silva, V., Simionato, M., Styron, R., Viganò, D., Danciu, L., Monelli, D., and Weatherill, G. The 2018 version of the Global Earthquake Model: Hazard component. *Earthquake Spectra*, 36:226–251, 2020. doi: 10.1177/8755293020931866.

Pagani, M., Johnson, K., and Garcia Pelaez, J. Modelling subduction sources for probabilistic seismic hazard analysis. *Geological Society, London, Special Publications*, 501:225–244, 2021. doi: 10.1144/SP501-2019-120.

Pasyanos, M. E., Masters, T. G., Laske, G., and Ma, Z. LITHO1.0: An updated crust and lithospheric model of the Earth. *Journal of Geophysical Research: Solid Earth*, 119(3):2153–2173, 2014. doi: 10.1002/2013JB010626.

Peñarubia, H. C., Johnson, K. L., Styron, R. H., et al. Probabilistic seismic hazard analysis model for the Philippines. *Earthquake Spectra.*, 36(1):44–68, 2020. doi: 10.1177/8755293019900521.

Rong, Y., Pagani, M., Magistrale, H., and Weatherill, G. A. Modeling Seismic Hazard by Integrating Historical Earthquake, Fault, and Strain Rate Data. In *Proceedings of the 16th World Conference on Earthquake Engineering*. Presented at the 16th WCEE, Santiago, Chile, p. Paper n. 448, 2017.

Rong, Y., Xu, X., Cheng, J., Chen, G., Magistrale, H., and Shen, Z.-K. A probabilistic seismic hazard model for Mainland China. *Earthquake Spectra*, 36:181–209, 2020. doi: 10.1177/8755293020910754.

Scherbaum, F., Kuehn, N. M., Ohrnberger, M., and Koelher, K. Exploring the proximity of ground-motion models using high-dimensional visualization techniques. *Earthquake Spectra*, 24: 1117–1138, 2010. doi: 10.1193/1.3478697.

Silva, V., Crowley, H., Pagani, M., Monelli, D., and Pinho, R. Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seismic risk assessment. *Natural Hazards*, 72:1409–1427, 2014. doi: 10.1007/s11069-013-0618-x.

Silver, A. Collaborative software development made easy. *Nature*, 550:143–144, 2017. doi: 10.1038/550143a.

Storchak, D. A., Di Giacomo, D., Bondár, I., Engdahl, E. R., Harris, J., Lee, W. H. K., Villaseñor, A., and Bormann, P. Public Release of the ISC–GEM Global Instrumental Earthquake Catalogue (1900–2009). *Seismological Research Letters*, 84:810–815, 2013. doi: 10.1785/0220130034.

Sunny, J., DeAngelis, M., and Edwards, B. Ranking and Selection of Earthquake Ground Motion Models Using the Stochastic Area Metric. *Seismological Research Letters*, 93:787–797, 2021. doi: 10.1785/0220210216.

van Stiphout, T., Zhuang, J., and Marsan, D. Theme V - Basic Techniques/Models and Techniques for Analysing Seismicity: Seismicity Declustering. 2010. doi: 10.5078/corssa-52382934.

Weatherill, G. A., Pagani, M., and Garcia, J. Exploring earthquake databases for the creation of magnitude-homogeneous catalogues: tools for application on a regional and global scale. *Geophysical Journal International*, 206:1652–1676, 2016. doi: 10.1093/gji/ggw232.

Weichert, D. H. Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bulletin of the Seismological Society of America*, 70(4): 1337–1346, 1980. doi: 10.1785/BSSA0700041337.

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., and Tian, D. The Generic Mapping Tools Version 6. *Geochem. Geophys. Geosyst.*, 20:5556–5564, 2019. doi: 10.1029/2019GC008515.

Youngs, R. R. and Coppersmith, K. J. Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological Society of America*, 75(4):939–964, 1985. doi: 10.1785/BSSA0750040939.