

"OpenQuake: Calculate, share, explore"

Risk Modeller's Toolkit - User Guide

Copyright © 2014 GEM Foundation

PUBLISHED BY GEM FOUNDATION

GLOBALQUAKEMODEL.ORG/OPENQUAKE

Citation

Please cite this document as:

Casotto, C. (2014) OpenQuake Risk Modeller's Toolkit - User Guide. *Global Earthquake Model (GEM). Technical Report*

Disclaimer

The "Risk Modeller's Toolkit - User Guide" is distributed in the hope that it will be useful, but without any warranty: without even the implied warranty of merchantability or fitness for a particular purpose. While every precaution has been taken in the preparation of this document, in no event shall the authors of the manual and the GEM Foundation be liable to any party for direct, indirect, special, incidental, or consequential damages, including lost profits, arising out of the use of information contained in this document or from the use of programs and source code that may accompany it, even if the authors and GEM Foundation have been advised of the possibility of such damage. The Book provided hereunder is on as "as is" basis, and the authors and GEM Foundation have no obligations to provide maintenance, support, updates, enhancements, or modifications.

The current version of the book has been revised only by members of the GEM model facility and it must be considered a draft copy.

License

This Book is distributed under the Creative Common License Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) (see link below). You can download this Book and share it with others as long as you provide proper credit, but you cannot change it in any way or use it commercially.

First printing, June 2015

Contents

1	Introduction	7
1.1	Introduction	7
1.2	Current features	7
1.3	Organization	7
2	Plotting	9
2.1	Plotting damage distribution	9
2.1.1	Plotting damage distribution	9
2.1.2	Plotting collapse maps	10
2.2	Plotting hazard and loss curves	11
2.2.1	Plotting hazard curves and uniform hazard spectra	11
2.2.2	Plotting loss curves	12
2.3	Plotting hazard and loss maps	12
2.3.1	Plotting hazard maps	13
2.3.2	Plotting loss maps	13
3	Risk	15
3.1	Deriving Probable Maximum Losses (PML)	15
3.2	Selecting a logic tree branch	16
4	Vulnerability	17
4.1	Introduction	17
4.2	Definition input models	17
4.2.1	Definition of capacity curves	17
4.2.2	Definition of ground motion records	17

4.2.3	Definition of damage model criterium	17
4.3	Model generator	17
4.3.1	Generation of capacity curves using DBELA	17
4.3.2	Generation of capacity curves using SP-BELA	17
4.3.3	Generation of capacity curves using point dispersion	17
4.4	Conversion from MDOF to SDOF	18
4.4.1	Conversion based on one mode of vibration	18
4.4.2	Conversion using an adaptive approach	18
4.5	Assessment of nonlinear structural response	18
4.5.1	SPO2IDA (Vamvatsikos and Cornell 2006)	18
4.5.2	Dolsek and Fajfar 2004	18
4.5.3	Ruiz Garcia and Miranda 2007	18
4.5.4	Vidic and Fajfar 1994	18
4.5.5	Lin and Miranda 2008	19
4.5.6	Miranda (2000) for firm soils	20
4.5.7	N2 (EC8, CEN 2005)	22
4.5.8	Capacity Spectrum Method (FEMA, 2005)	22
4.5.9	DBELA (Silva et al. 2013)	23
4.5.10	Nonlinear time-history analysis in Single Degree of Freedom (SDOF) Oscillators	23
4.6	Derivation of fragility and vulnerability functions	23
4.6.1	Derivation of fragility functions	23
4.6.2	Derivation of vulnerability functions	23
5	Bibliography	25
5.1	Books	25
5.2	Articles	25
5.3	Other Sources	26
	Index	27
I	Appendices	29
A	The 10 Minute Guide to Python!	31
A.1	Basic Data Types	31
A.1.1	Scalar Parameters	31
A.1.2	Iterables	33
A.1.3	Dictionaries	34
A.1.4	Loops and Logicals	34
A.2	Functions	36
A.3	Classes and Inheritance	36
A.3.1	Simple Classes	36
A.3.2	Inheritance	37
A.3.3	Abstraction	38
A.4	Numpy/Scipy	39

Preface

To be completed by Vitor.

The goal of this book is to provide a comprehensive and transparent description of the methodologies adopted and implemented in the Risk Modeller's Toolkit (RMTK).

It is freely distributed under an Affero GPL license (more information available at this link <http://www.gnu.org/licenses/agpl-3.0.html>)

1. Introduction

1.1 Introduction

To be completed by Anirudh.

1.2 Current features

To be completed by Anirudh.

The Risk Modeller's Toolkit is currently divided into two sections:

1. These functions are intended to address the modeller's needs for defining vulnerability curves, implementing methodologies differing for level of complexity and for the input data available for the buildings under study. GEM analytical vulnerability guidelines have been integrated in this tool and some of the methodologies indicated have been already implemented in the library.
2. These functions are intended to address the needs of visualising the results of the calculations performed with the OpenQuake-engine.

1.3 Organization

This manual is designed to explain the various functions in the toolkit, to provide the theoretical background behind them, and to guide the modeller in the use of the rmtk within the "IPython Notebook" environment. This novel tool implements Python inside a web-browser environment, permitting the user to execute real Python workflows, whilst allowing for images and text to be embedded. Its use is encouraged especially for beginner python users for a more visual application of the rmtk.

The IPython Notebook comes installed from version 1.0 of IPython, that can be installed from the python package repository by entering:

```
~$ sudo pip install ipython
```

A notebook session can be started via the command:

```
~$ ipython notebook --pylab inline
```

The tutorial itself does not specifically require a working knowledge of Python. However, an understanding of the basic python data types is highly desirable. Users who are new to Python are recommended to familiarise themselves with Appendix A of this tutorial. To be completed by Anirudh.

The *rmtk* is currently subdivided into two classes of tools, the Vulnerability and Plotting tools, presented in Chapter 2 and Chapter 3 of this tutorial respectively. In the Vulnerability chapter the vulnerability methodologies implemented are classified in Non-linear Static (NLS) and Non-linear Dynamic (NLD) according to the structural analysis type performed to assess the response of the building. These two main sections (NLS and NLD) are organised as follows:

- General Introduction.
- Getting Started, where it is explained what files need to be executed to start the vulnerability analysis, and what options are available to call the preferred methodology and to input the preferred data type.
- Description of the methodologies.

Within the description of each methodology the user can find the following subsections:

- Theoretical description of the method.
- Description and examples of the inputs.
- Description of the workflow.

A summary of the algorithms available in the present version is given in Table 1.1.

Feature	Algorithm
Non-linear Static	Cr-based (Ruiz-Garcia and Miranda, 2007) Spo2ida (Vamvatsikos and Cornell, 2006) R-/mu-T-based (Dolsek and Fajfar, 2004)
Non-linear Dynamic	DPM-based (Silva et al. 2013) Ida-postprocessing (Vamvatsikos and Cornell, 2002)

Table 1.1 – *Current algorithms in the HMTK*

Plotting damage distribution

Plotting damage distribution

Plotting collapse maps

Plotting hazard and loss curves

Plotting hazard curves and uniform hazard spectra

Plotting loss curves

Plotting hazard and loss maps

Plotting hazard maps

Plotting loss maps

2. Plotting

The OpenQuake-engine is capable of generating several seismic hazard and risk outputs, such as loss exceedance curves, seismic hazard curves, loss and hazard maps, damage statistics, amongst others. Most of these outputs are stored using the Natural Risk Markup Language (nrml), or simple comma separated value (csv) files. The Plotting module of the Risk Modellers Toolkit allows users to visualize the majority of the OpenQuake-engine results, as well as to convert them into other formats compatible with GIS software (e.g. QGIS). Despite the default styling of the maps and curves defined within the Risk Modellers Toolkit, it is important to state that any user can adjust the features of each output by modifying the original scripts.

2.1 Plotting damage distribution

Using the Scenario Damage Calculator (Silva et al., 2014) of the OpenQuake-engine, it is possible to assess the distribution of damage for a collection of assets considering a single seismic event. These results are comprised of damage per building typology, total damage distribution, and distribution of collapses in the region of interest.

2.1.1 Plotting damage distribution

This feature of the Plotting module allows users to plot the distribution of damage across the various vulnerability classes, as well as to the total damage distribution. For what concerns the former result, it is necessary to set the path to the output file using the parameter `tax_dmg_dist_file`. It is also possible to specify which vulnerability classes should be considered, using the parameter `taxonomy_list`. However, if a user wishes to consider all of the vulnerability classes, then this parameter should be left empty. It is also possible to specify if a 3D plot containing all of the vulnerability classes should be generated, or instead a 2D plot per vulnerability class. For follow the former option, the parameter `plot_3d` should be set to `True`. It is important to understand that this option leads to a plot of damage fractions for each vulnerability class, instead of the number of assets in each damage state. An example of this output is illustrated in Figure 2.1.

In order to plot the total damage distribution (considering the entire collection of assets), it is necessary to use the parameter `total_dmg_dist_file` to define the path to the respective output file. Figure 2.2 presents an example of this type of output.

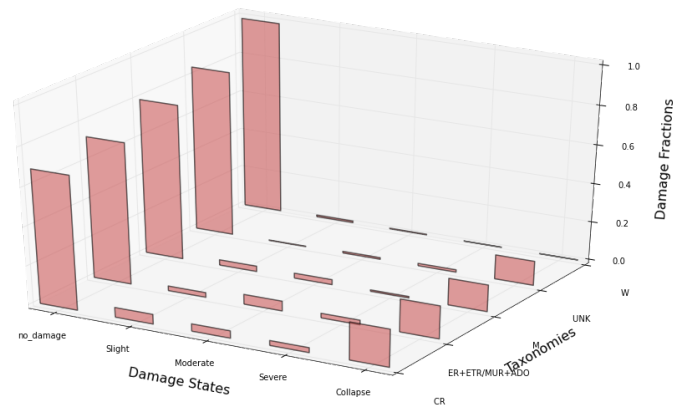


Figure 2.1 – Damage disaggregation per vulnerability class.

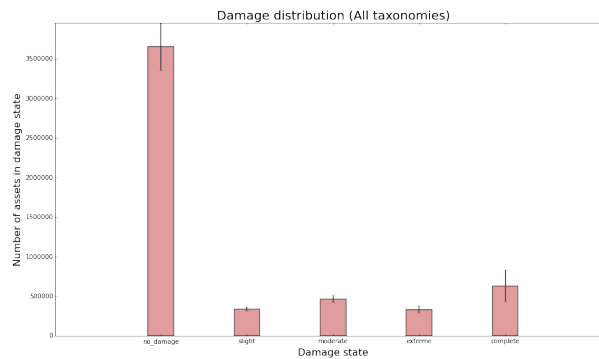


Figure 2.2 – Total damage distribution.

2.1.2 Plotting collapse maps

The OpenQuake-engine also generates an output defining the spatial distribution of the mean (and associated standard deviation) of assets in the last damage state (usually representing collapse or complete damage). The location of this output needs to be specified using the parameter `collapse_map`. Then, it is necessary to specify whether the user desires a map with the aggregated number of collapses (i.e. at each location, the mean number of collapses across all of the vulnerability classes are summed) or a map for each vulnerability class. Thus, the following options are permitted:

1. Aggregated collapse map only.
2. Collapse maps per taxonomy only.
3. Both aggregated and taxonomy-based.

The plotting option should be specified using the parameter `plotting_type`, and the location of the exposure model used to perform the calculations must be defined using the variable `exposure_model`. A number of other parameters can also be adjusted to modify the style of the resulting collapse map as follows:

- `bounding_box`: If set to 0, the Plotting module will calculate the geographical distribution of the assets, and adjust the limits of the map accordingly. Alternatively, a user can also specify the minimum/maximum latitude and longitude that should be used in the creation of the map.

- `marker_size`: This attribute can be used to adjust the size of the markers in the map.
- `log_scale`: If set to `True`, it will apply a logarithmic scale on the color scheme of the map, potentially allowing a better visualization of the variation of the numbers of collapses in the region of interest.

An example of a collapse map is presented in Figure 2.3.

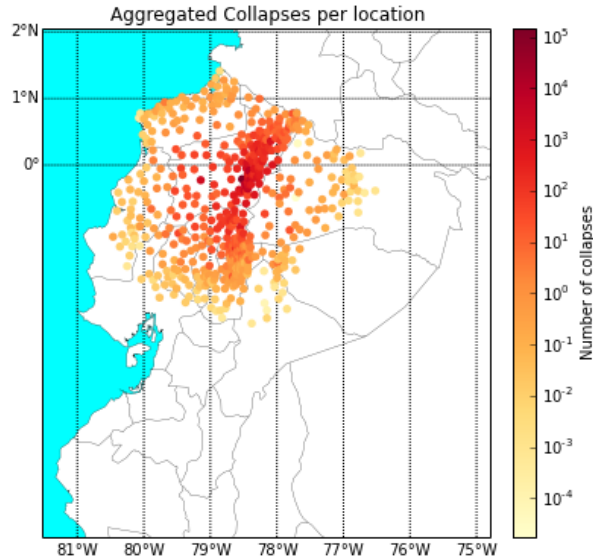


Figure 2.3 – *Spatial distribution of the mean number of collapses.*

2.2 Plotting hazard and loss curves

Using the Classical PSHA-based or Probabilistic Event-based Calculators (Silva et al., 2014, Pagani et al., 2014) of the OpenQuake-engine, it is possible to calculate seismic hazard curves for a number of locations, or loss exceedance curves considering a collection of spatially distributed assets.

2.2.1 Plotting hazard curves and uniform hazard spectra

A seismic hazard curve defines the probability of exceeding a number of intensity measure levels (e.g. peak ground acceleration or spectral acceleration) for a given interval of time (e.g. 50 years). In order to plot these curves, it is necessary to define the path to the output file in the parameter `hazard_curve_file`. Then, since each output file might contain a great number of hazard curves, it is necessary to establish the location for each hazard curve will be extracted. To visualize the list of locations comprised in the output file, the function `hazard_curves.loc_list` can be employed. Then, the chosen location must be provided to the plotting function (e.g. `hazard_curves.plot("81.213823|29.761172")`). An example of a seismic hazard curve is provided in Figure 2.4.

To plot uniform hazard spectra (UHS), a similar approach should be followed. The output file containing the uniform hazard spectra should be defined using the parameter `uhs_file`, and then a location must be provided to the plotting function (e.g. `uhs.plot("81.213823|29.761172")`). An example of uniform hazard spectra is illustrated in Figure 2.5.

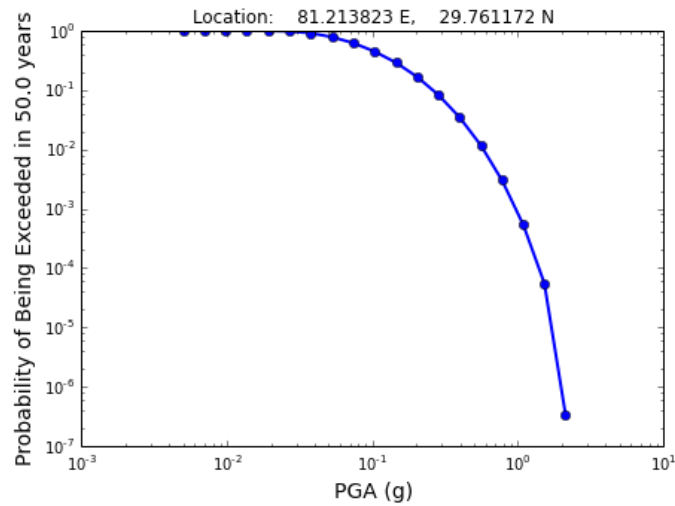


Figure 2.4 – Seismic hazard curve for peak ground acceleration (PGA).

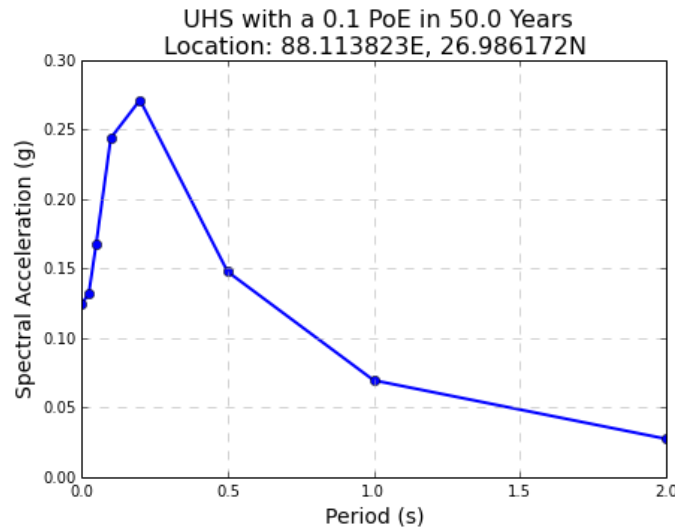


Figure 2.5 – Uniform Hazard Spectra for a probability of exceedance of 10% in 50 years.

2.2.2 Plotting loss curves

A loss exceedance curve defines the relation between a set of loss levels and the corresponding probability of exceedance within a given time span (e.g. a year). In order to plot these curves, it is necessary to define the location of the output file using the parameter `loss_curves_file`. Since each output file may contains a large number of loss exceedance curves, it is necessary to define for which assets will the loss curves be extracted. The parameter `assets_list` should be employed to define all of the chosen asset ids. These ids can be visualize directly on the loss curve output file, or on the exposure model used for the risk calculations. It is also possible to define a logarithmic scale for the x and y axis using the parameters `log_scale_x` and `log_scale_y`. A loss exceedance curve for a single asset is depicted in Figure 2.6.

2.3 Plotting hazard and loss maps

The OpenQuake-engine offers the possibility of calculating seismic hazard and loss (or risk) maps. To do so, it utilizes the seismic hazard or loss exceedances curves, to estimate the

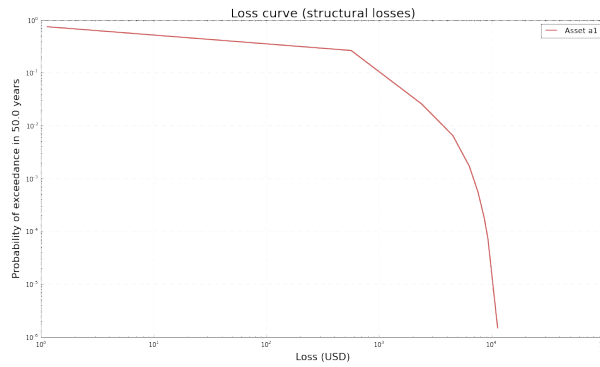


Figure 2.6 – *Loss exceedance curve.*

corresponding hazard or loss for the pre-defined return period (or probability of exceedance within a given interval of time).

2.3.1 Plotting hazard maps

A seismic hazard map provides the expected ground motion (e.g. peak ground acceleration or spectral acceleration) at each location, for a certain return period (or probability of exceedance within a given interval of time). To plot this type of maps, it is necessary to specify the location of the output file using the parameter `hazard_map_file`. An example hazard map is displayed in Figure 2.4.

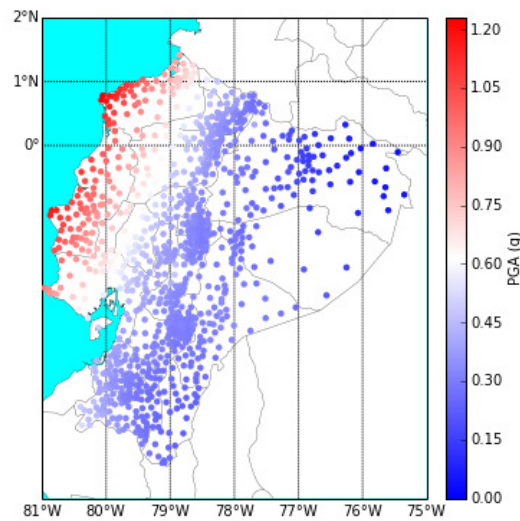


Figure 2.7 – *Seismic hazard map for a probability of exceedance of 10% in 50 years.*

2.3.2 Plotting loss maps

A loss map provides the estimated losses for a collection of assets, for a certain return period (or probability of exceedance within a given interval of time). It is important to understand that these maps are not providing the distribution of losses for a seismic event for the chosen return period, nor the losses whose sum would correspond to the aggregated loss for the same return period. This type of maps is simply providing the expected loss for a specified level of frequency for each asset. To use this feature, it is necessary to define the path of the output file using the parameter `loss_map_file`, as well as the exposure model used to perform the risk

calculations through the parameter `exposure_model`. Then, similarly to what was explained in section 2.1.2 for collapse maps, it is possible to follow three approaches to generate the loss maps:

1. Aggregated loss map only.
2. Loss maps per taxonomy only.
3. Both aggregated and taxonomy-based.

Then, there are a number of options that can be used to modify the style of the maps. These include the size of the marker of the map (`marker_size`), the geographical limits of the map (`bounding_box`), and the employment of a logarithmic spacing for the color scheme (`log_scale`). An example loss map for a single vulnerability class is presented in Figure 2.8.

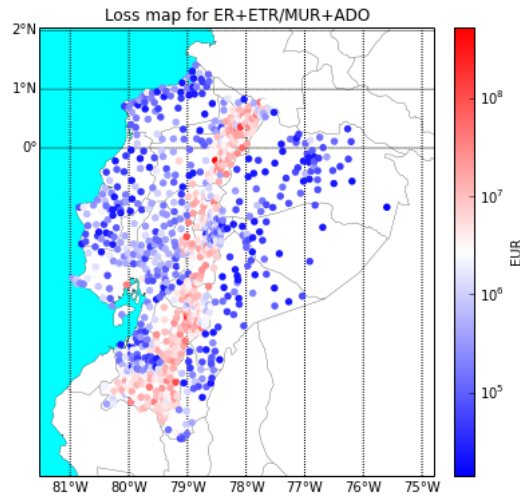


Figure 2.8 – Loss (economic) map for a probability of exceedance of 10% in 50 years.

As mentioned on the introductory section, it is also possible to convert any of the maps into a format (csv) easily readable by GIS software. To do so, it is necessary to set the parameter `export_map_to_csv` to True. As an example, a map containing the average annual losses for Ecuador has been converted to the csv format, and introduced into the QGIS software to produce the map presented in Figure 2.9.

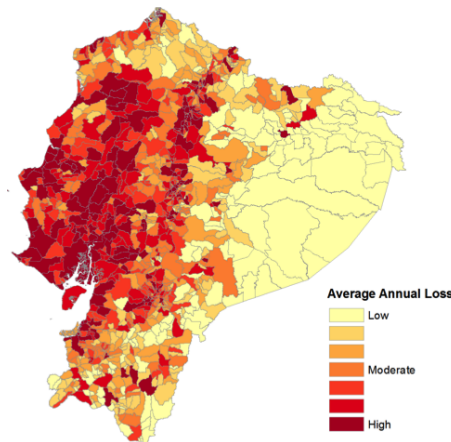


Figure 2.9 – Average annual (economic) losses for Ecuador.

3. Risk

The OpenQuake-engine currently generates the most common seismic hazard and risk results (e.g. hazard maps, loss curves, average annual losses). However, it is recognized that there are a number of other metrics that might not be of interest of the general GEM community, but fundamental for specific users. This module of the Risk Modellers Toolkit aims to provide users with additional risk results and functionalities.

3.1 Deriving Probable Maximum Losses (PML)

The Probabilistic Event-based Risk calculator (Silva et al., 2014) of the OpenQuake-engine is capable of calculating event loss tables, which contain a list of earthquake ruptures and associated losses. These losses may refer to specific assets, or the sum of the losses from the entire building portfolio (aggregated loss curves). Using this module, it is possible to derive a probable maximum loss (PML) curves (i.e. relation between a set of loss levels and corresponding return periods of exceedance), as illustrated in Figure 3.1.

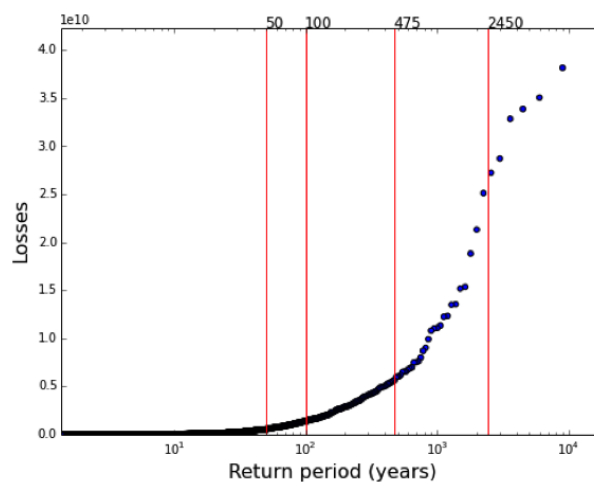


Figure 3.1 – Probable Maximum Loss (PML) curve.

To use this feature, it is necessary to use the parameter `event_loss_table_folder` to

specify the location of the folder that contains the set of event loss tables and stochastic event sets. Then, it is also necessary to provide the total economic value of the building portfolio (using the variable `total_cost`) and the list of return periods of interest (using the variable `return_periods`). This module also offers the possibility of saving all of the information in csv files, which can be used in other software (e.g. Microsoft Excel) for other purposes. To do so, the parameters `save_elt_csv` and `save_ses_csv` should be set to `True`.

3.2 Selecting a logic tree branch

When a non-trivial logic-tree is used to capture the model uncertainty in the source model or in the choice of an appropriate ground motion prediction equation (GMPE) for each of the tectonic region types in the region considered, OpenQuake can calculate the hazard curves for each end-branch of the logic-tree individually. Now, if a risk modeller wishes to estimate damage or losses using one or a few of these branches only, it is useful to compare the hazard curves for the chosen branch with the mean hazard curve. Depending upon the distance of the hazard curve for a particular branch from the mean hazard curve, the risk modeller may wish to choose the branches for which the hazard curves are closest to the mean hazard curve. This Python script and corresponding IPython notebook allow the risk modeller to list the end-branches for the hazard calculation, sorted in increasing order of the distance of the branch hazard curve from the mean hazard curve. Currently, the distance metric used for performing the sorting is the root mean square distance.

Introduction

Definition input models

- Definition of capacity curves
- Definition of ground motion records
- Definition of damage model criterium

Model generator

- Generation of capacity curves using DBELA
- Generation of capacity curves using SP-BELA
- Generation of capacity curves using point dispersion

Conversion from MDOF to SDOF

- Conversion based on one mode of vibration
- Conversion using an adaptive approach

Assessment of nonlinear structural response

- SPO2IDA (Vamvatsikos and Cornell 2006)
- Dolsek and Fajfar 2004
- Ruiz Garcia and Miranda 2007
- Vidic and Fajfar 1994
- Lin and Miranda 2008
- Miranda (2000) for firm
- N2 (EC8, CEN 2005)
- Capacity Spectrum Method (FEMA, 2005)
- DBELA (Silva et al. 2013)
- Nonlinear time-history analysis in Single Degree of Freedom (SDOF) Oscillators

Derivation of regularity and vulnerability functions

- To be completed by Anirudh.
- Derivation of regularity functions
- Derivation of vulnerability functions

4. Vulnerability

4.2 Definition input models

To be completed by Anirudh.

4.2.1 Definition of capacity curves

To be completed by Anirudh.

4.2.2 Definition of ground motion records

To be completed by Anirudh.

4.2.3 Definition of damage model criterium

To be completed by Anirudh.

4.3 Model generator

To be completed by Vitor.

4.3.1 Generation of capacity curves using DBELA

To be completed by Vitor.

4.3.2 Generation of capacity curves using SP-BELA

To be completed by Vitor.

4.3.3 Generation of capacity curves using point dispersion

To be completed by Vitor.

4.4 Conversion from MDOF to SDOF

To be completed by Vitor.

4.4.1 Conversion based on one mode of vibration

To be completed by Anirudh.

4.4.2 Conversion using an adaptive approach

To be completed by Anirudh.

4.5 Assessment of nonlinear structural response

4.5.1 SPO2IDA (Vamvatsikos and Cornell 2006)

To be completed by Chiara.

4.5.2 Dolsek and Fajfar 2004

To be completed by Chiara.

4.5.3 Ruiz Garcia and Miranda 2007

To be completed by Chiara.

4.5.4 Vidic and Fajfar 1994

This procedure aims to determine the displacements from an inelastic spectra for systems with a given ductility factor. The inelastic displacement spectra is determined by means of applying a ductility-based reduction factor (C), which depends on the natural period of the system, the given ductility factor, the hysteretic behaviour, the damping model, and the frequency content of the ground motion.

The procedure proposed by (Vidic et al., 1994) was validated by a comparison of the approximate spectra with the “exact” spectra obtained from non-linear dynamic time history analyses. Records from California and Montenegro were used as representative of “standard” ground motion, while the influence of input motion was analysed using other five groups of records (coming from different parts of the world) that represented different types of ground motions. The influence of the hysteretic models was taken into account by considering the bilinear model and the stiffness degrading Q-model. Finally, in order to analyse the effect of damping, two models were considered: “mass-proportional” damping, which assumes a time-independent damping coefficient based on elastic properties, and “instantaneous stiffness-proportional” damping, which assumes a time-dependent damping coefficient based on tangent stiffness. For most cases, a damping ratio of 5% was assumed, although for some systems a value of 2% was adopted.

It is possible to derive approximate strength and displacement inelastic spectra from an elastic pseudo-acceleration spectrum using the proposed modified spectra. In the medium and long-period region, it was observed that the reduction factor is slightly dependant on the period T and is roughly equal to the prescribed ductility (μ). However, in the short-period region, the factor C strongly depends on both T and μ . The influence of hysteretic behaviour and damping can be observed for the whole range of periods. Based on this, a bilinear curve was proposed. Starting in $C = 1$, the value of C increases linearly along the short-period region up to a value

Table 4.1 – Parameters for the estimation of the reduction factor C proposed by (Vidic et al., 1994)

Hysteresis model	Damping model	c_1	c_2	c_R	c_T
Q	Mass	1.00	0.65	1.00	0.30
Q	Stiffness	0.75	0.65	1.00	0.30
Bilinear	Mass	1.35	0.75	0.95	0.20
Bilinear	Stiffness	1.10	0.75	0.95	0.20

approximately equal to the ductility factor. In the medium- and long-period range, the C-factor remains constant. This is mathematically expressed by the following relationships:

$$C_\mu = \begin{cases} c_1 (\mu - 1)^{c_R} \frac{T}{T_0} + 1, & T \leq T_0 \\ c_1 (\mu - 1)^{c_R} + 1 & T > T_0 \end{cases} \quad (4.1)$$

where:

$$T_0 = c_2 \mu^{c_T} T_c \quad (4.2)$$

And T_c stands for the characteristic spectral period and c_1, c_2, c_R, c_T are constants dependant on the hysteretic behaviour and damping model, as defined in Table 4.1.

In order to use this methodology, it is necessary to load one or multiple capacity curves as described in Section 4.2.1, as well as a set of ground motion records as explained in Section 4.2.2. Then, it is necessary to specify a damage model using the parameter `damage_model` (see Section 4.2.3), and a damping ratio using the parameter `damping`. It is also necessary to specify the type of hysteresis (Q or bilinear) and damping (mass or stiffness) models as defined in Table 4.1, using the parameters `hysteresis_model` and `damping_model`, respectively. After importing the module `vidic_etal_1994`, it is possible to calculate the distribution of structures across the set of damage state for each ground motion record using the following command:

```
PDM, Sds = vidic_etal_1994.calculate_fragility(capacity_curves,gmrs,...
damage_model,damping)
```

Where PDM (i.e. probability damage matrix) represents a matrix with the number of structures in each damage state per ground motion record, and Sds (i.e. spectral displacements) represents a matrix with the maximum displacement (of the equivalent SDOF) of each structure per ground motion record. the variable PDM can then be used to calculate the mean fragility model as described in Section 4.6.1.

4.5.5 Lin and Miranda 2008

This methodology estimates the maximum inelastic displacement of an existing structure based on the maximum elastic displacement response of its equivalent linear system without the need for iterations, based on the strength ratio R (instead of the most commonly used ductility ratio).

In order to evaluate an existing structure, a pushover analysis should be conducted in order to obtain the capacity curve. This curve should be bilinearised in order to obtain the yield strength, f_y , the postyield stiffness ratio, α , and the strength ratio, R . With these parameters, along with the initial period of the system, it is possible to estimate the optimal period shift (i.e. the ratio

Table 4.2 – Parameters for the estimation of the reduction factor C proposed by (Lin and Miranda, 2008)

α	m_1	m_2	n_1	n_2
0%	0.026	0.87	0.016	0.84
5%	0.026	0.65	0.027	0.55
10%	0.027	0.51	0.031	0.39
20%	0.027	0.36	0.030	0.24

between the period of the equivalent linear system and the initial period) and the equivalent viscous damping, ξ_{eq} , of the equivalent linear system, using the following relationships derived by (Lin and Miranda, 2008).

$$\frac{T_{eq}}{T_0} = 1 + \frac{m_1}{T_0^{m_2}} (R^{1.8} - 1) \quad (4.3)$$

$$\xi_{eq} = \xi_0 + \frac{n_1}{T_0^{n_2}} (R - 1) \quad (4.4)$$

Where the coefficients m_1 , m_2 , n_1 , and n_2 depend on the postyield stiffness ratio, as shown in the following Table 4.2.

Using ξ_{eq} and the damping modification factor, B (as defined in Table 15.6-1 of NEHRP-2003), it is possible to construct the reduced displacement spectrum, $S_d(T, \xi_{eq})$ from which the maximum displacement demand (i.e. the displacement corresponding to the equivalent system period) can be obtained, using the following equation:

In order to use this methodology, it is necessary to load one or multiple capacity curves as described in Section 4.2.1, as well as a set of ground motion records as explained in Section 4.2.2. Then, it is necessary to specify a damage model using the parameter `damage_model` (see Section 4.2.3). After importing the module `lin_miranda_2008`, it is possible to calculate the distribution of structures across the set of damage state for each ground motion record using the following command:

```
PDM, Sds = lin_miranda_2008.calculate_fragility(capacity_curves,gmrs,...
damage_model,damping)
```

Where PDM (i.e. probability damage matrix) represents a matrix with the number of structures in each damage state per ground motion record, and Sds (i.e. spectral displacements) represents a matrix with the maximum displacement (of the equivalent SDOF) of each structure per ground motion record. the variable PDM can then be used to calculate the mean fragility model as described in Section 4.6.1.

4.5.6 Miranda (2000) for firm soils

This study by Miranda, 2000 aims to quantify the influence of soil conditions, earthquake magnitude, and epicentral distance on the inelastic displacement ratios, C_μ . For two systems with the same mass and period of vibration that have been subjected to the same earthquake ground motion. C_μ can be defined as the ratio of the maximum lateral inelastic displacement

demand of one to the maximum lateral elastic displacement demand on the other, as shown in the following equation:

$$C_\mu = \frac{\Delta_{inelastic}}{\Delta_{elastic}} \quad (4.5)$$

In this study, 264 earthquake acceleration time histories recorded in California (USA) for 12 different events were used. In order to investigate the effect of the soil conditions, the records were classified into three groups: the first one consisted of ground motions recorded on stations located on rock (i.e. average shear-wave velocities >760 m/s). The second group included the records registered on stations on very dense soil or soft rock (i.e. average shear-wave velocities between 360 m/s and 760 m/s). Finally, the third group consisted of ground motion records from stations located on stiff soil (i.e. average shear-wave velocities between 180 m/s and 360 m/s).

It was observed that for periods longer than about 1.0 s, the mean inelastic displacement ratios are approximately equal to 1, meaning that, on average, the maximum inelastic displacements are equal to the maximum elastic displacements. On the other hand, for periods smaller than 1.0 s, the mean inelastic displacement ratios are larger than 1 and strongly depend on the period of vibration and on the level of inelastic deformation. The results of the investigation yielded that for the sites under consideration (i.e. average shear-wave velocities higher than 180 m/s) neither the soil conditions, nor the earthquake magnitude, nor the distance to rupture cause significant differences on the value of C_μ . However, if directivity effects are taken into consideration, the inelastic displacement ratios for periods between 0.1 s and 1.3 s can be larger than those estimated for systems not affected by directivity.

Based on the results of the mean inelastic displacement ratios, nonlinear regression analyses were conducted to estimate the following simplified expression, which allows estimating the inelastic displacement ratio of a system:

$$C_\mu = \left[1 + \left(\frac{1}{\mu} - 1 \right) \exp(-12T\mu^{-0.8}) \right]^{-1} \quad (4.6)$$

Where μ = displacement ductility ratio and T = period of vibration.

In order to use this methodology, it is necessary to load one or multiple capacity curves and a set of ground motion records, as explained in Section 4.2.1 and 4.2.2, respectively. Then, it is necessary to specify a damage model using the parameter `damage_model` (see Section 4.2.3), and a damping ratio using the parameter `damping`. After importing the module `miranda_2000_firm_soils`, it is possible to calculate the distribution of structures across the set of damage states for each ground motion record using the following command:

```
PDM, Sds = miranda_2000_firm_soils.calculate_fragility(capacity_curves,...
gmrs,damage_model,damping)
```

Where PDM (i.e. probability damage matrix) represents a matrix with the number of structures in each damage state per ground motion record, and Sds (i.e. spectral displacements) represents a matrix with the maximum displacement (of the equivalent SDOF) of each structure per ground motion record. the variable PDM can then be used to calculate the mean fragility model as described in Section 4.6.1.

4.5.7 N2 (EC8, CEN 2005)

To be completed by Vitor.

4.5.8 Capacity Spectrum Method (FEMA, 2005)

The capacity spectrum method (CSM) was initially proposed by (Freeman et al., 1975), and it represents a simplified methodology for many purposes such as the evaluation of a large inventory of buildings, structural assessment of new or existing buildings or to identify the correlation between damage states and levels of ground motion. ATC-40, 1996 proposes three different procedures (A, B and C) for the application of the Capacity Spectrum Method. However, procedure B adopts some simplifications that might not always be valid and procedure C has a very strong graphical component, making it difficult for systematic applications. Hence, procedure A, which is characterized by its intuitiveness and simplicity, has been implemented in the RMTK.

This procedure iteratively compares the capacity and the demand of a structure, using a capacity curve (for the equivalent SDOF) and a damped response spectrum, respectively. The ground motion spectrum is computed for a level of equivalent viscous damping that is estimated as a function of the displacement at which the response spectrum crosses the capacity curve, in order to take into account the inelastic behaviour of the structure. Iterations are needed until there is a match between the equivalent viscous damping of the structure and the damping applied to the spectrum. The final intersection of these two curves approximates the target displacement response of the structure. This result is presented in Figure 4.1 for a "weak" and a "strong" ground motion record.

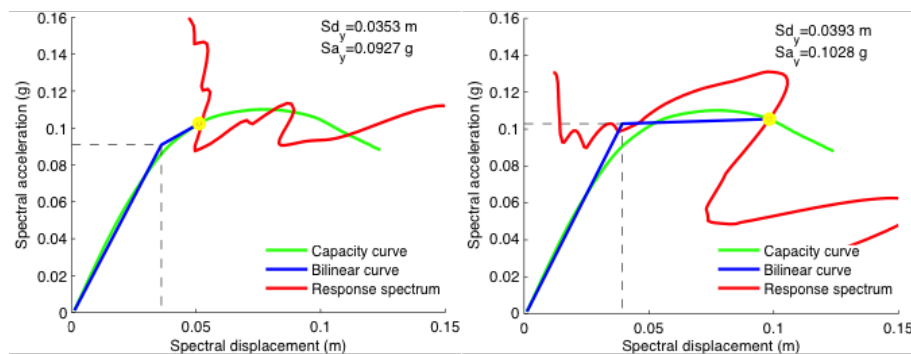


Figure 4.1 – Assessment of the target displacement for "weak" (left) and "strong" (strong) ground motion record.

The initial proposal of this method was heavily criticized due to its tendency to underestimate the deformation of the structures, which was mostly related with the model employed to calculate the equivalent viscous damping (e.g. Fajfar, 1999; Chopra and Goel, 2000). Thus, in FEMA-440, 2005, some modifications were proposed regarding the calculation of this component. Furthermore, several other models relating an equivalent viscous damping with a ductility level have been proposed in the last decades, and implemented in the RMTK. The following list describes these models, and specifies the code that should be defined in the variable `damping_model` in order to follow the associated model in the vulnerability calculations.

- (FEMA-440, 2005):.
- (Kowalsky, 1994):.
- (Iwan, 1980):.

- (Gulkan and Sozen, 1974):.
- (PriesleyEtAl2007):.
- Masonry structures:.

This performance point is equivalent to what would be obtained by subjecting the equivalent single degree of freedom to a nonlinear time history analysis. This response (i.e. performance point) can be used to allocate the structure in a damage state, based on a pre-established set of displacement thresholds. This process can be repeated several times considering other ground motion records, as well as structures (i.e. building class).

4.5.9 DBELA (Silva et al. 2013)

To be completed by Vitor.

4.5.10 Nonlinear time-history analysis in Single Degree of Freedom (SDOF) Oscillators

To be completed by Vitor.

4.6 Derivation of fragility and vulnerability functions

4.6.1 Derivation of fragility functions

To be completed by Anirudh.

4.6.2 Derivation of vulnerability functions

To be completed by Anirudh.

5. Bibliography

5.1 Books

5.2 Articles

- Chopra, Anil and Rakesh Goel (2000). "Evaluation of NSP to estimate seismic deformation: SDF systems". In: *Journal of Structural Engineering* 126.4, 482–490 (cited on page 22).
- Fajfar, Peter (1999). "Capacity spectrum method based on inelastic demand spectra." In: *Earthquake Engineering and Structural Dynamics* 28.9, 979–993 (cited on page 22).
- Gulkan, P. and M. A. Sozen (1974). "Inelastic Responses of Reinforced Concrete Structure to Earthquake Motions". In: *ACI Journal Proceedings* 71.12, pages 604–610 (cited on page 23).
- Iwan, W (1980). "Estimating inelastic response spectra from elastic spectra". In: *Earthquake Engineering and Structural Dynamics* (cited on page 22).
- Lin, Yu-Yuan and Eduardo Miranda (2008). "Noniterative equivalent linear method for evaluation of existing structures". In: *Journal of structural engineering* 134.11, pages 1685–1695. URL: [http://ascelibrary.org/doi/abs/10.1061/\(ASCE\)0733-9445\(2008\)134:11\(1685\)](http://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9445(2008)134:11(1685)) (visited on 06/18/2015) (cited on page 20).
- Miranda, Eduardo (2000). "Inelastic Displacement Ratios for Structures on Firm Sites". In: *Journal of AAPOS Structural Engineering* 126.10, pages 1150–1159 (cited on page 20).
- Pagani, Marco, Damiano Monelli, Graeme Weatherill, Laurentiu Danciu, Helen Crowley, Vitor Silva and Paul Henshaw, Lars Butler, Matteo Nastasi, Luigi Panzeri, Michele Simionato, and Daniele Vigano (2014). "OpenQuake Engine: An open hazard (and risk) software for the Global Earthquake Model". In: *Seismological Research Letters* (cited on page 11).
- Silva, Vitor, Helen Crowley, Marco Pagani, Damiano Monelli, and Rui Pinho (2014). "Development of the OpenQuake engine, the Global Earthquake Model open-source software for seismic risk assessment". In: *Natural Hazards* 72.3, pages 1409–1427. ISSN: 0921-030X, 1573-0840. (Visited on 06/04/2015) (cited on pages 9, 11, 15).
- Vidic, T., P. Fajfar, and Fischinger M. (1994). "Consistent inelastic design spectra: Strength and displacement." In: *Earthquake Engineering and Structural Dynamics* 23.5, pages 507–521 (cited on pages 18, 19).

5.3 Other Sources

- ATC-40 (1996). *Seismic Evaluation and Retrofit of Concrete Buildings*. Technical report. Applied Technology Council, Redwood City, CA (cited on page 22).
- FEMA-440 (2005). *Improvement of Nonlinear Static Seismic Analysis Procedures*. Technical report. California, USA (cited on page 22).
- Freeman, S., J. Nicoletti, and J. Tyrell (1975). "Evaluation of Existing Buildings for seismic risk - A case study of Puget Sound Naval Shipyard, Bremerton". In: *Proceedings of the 1st U.S. National Conference on Earthquake Engineering, Berkley, USA* (cited on page 22).
- Kowalsky, M. J. (1994). "Displacement-based design - a methodology for seismic design applied to RC bridge columns." Master's thesis. University of California, San Diego (cited on page 22).

Part I

Appendices

Basic Data Types

- Scalar Parameters
- Iterables
- Dictionaries
- Loops and Logicals

Functions

Classes and Inheritance

- Simple Classes
- Inheritance
- Abstraction

Numpy/Scipy

A. The 10 Minute Guide to Python!

The HMTK is intended to be used by scientists and engineers without the necessity of having an existing knowledge of Python. It is hoped that the examples contained in this manual should provide enough context to allow the user to understand how to use the tools for their own needs. In spite of this, however, an understanding of the fundamentals of the Python programming language can greatly enhance the user experience and permit the user to join together the tools in a workflow that best matches their needs.

The aim of this appendix is therefore to introduce some fundamentals of the Python programming language in order to help understand how, and why, the HMTK can be used in a specific manner. If the reader wishes to develop their knowledge of the Python programming language beyond the examples shown here, there is a considerable body of literature on the topic from both a scientific and developer perspective.

A.1 Basic Data Types

Fundamental to the use of the HMTK is an understanding of the basic data types Python recognises:

A.1.1 Scalar Parameters

- **float** A floating point (decimal) number. If the user wishes to enter in a floating point value then a decimal point must be included, even if the number is rounded to an integer.

```
1 | >> a = 3.5
2 | >> print a, type(a)
3 | 3.5 <type 'float'>
```

- **integer** An integer number. If the decimal point is omitted for a floating point number the number will be considered an integer

```
1 | >> b = 3
2 | >> print b, type(b)
3 | 3 <type 'int'>
```

The functions `float()` and `int()` can convert an integer to a float and vice-versa. Note that taking `int()` of a fraction will round the fraction down to the nearest integer

```

1 | >> float(b)
2 | 3
3 | >> int(a)
4 | 3

```

- **string** A text string (technically a “list” of text characters). The string is indicated by the quotation marks ”something” or ’something else’

```

1 | >> c = "apples"
2 | >> print c, type(c)
3 | apples <type 'str'>

```

- **bool** For logical operations python can recognise a variable with a boolean data type (True / False).

```

1 | >> d = True
2 | >> if d:
3 |     print "y"
4 | else:
5 |     print "n"
6 | y
7 | >> d = False
8 | >> if d:
9 |     print "y"
10 | else:
11 |     print "n"
12 | n

```

Care should be taken in Python as the value 0 and 0.0 are both recognised as False if applied to a logical operation. Similarly, booleans can be used in arithmetic where True and False take the values 1 and 0 respectively

```

1 | >> d = 1.0
2 | >> if d:
3 |     print "y"
4 | else:
5 |     print "n"
6 | y
7 | >> d = 0.0
8 | >> if d:
9 |     print "y"
10 | else:
11 |     print "n"
12 | n

```

Scalar Arithmetic

Scalars support basic mathematical operations (# indicates a comment):

```

1 | >> a = 3.0
2 | >> b = 4.0
3 | >> a + b # Addition
4 | 7.0
5 | >> a * b # Multiplication
6 | 12.0
7 | >> a - b # Subtraction
8 | -1.0
9 | >> a / b # Division
10 | 0.75
11 | >> a ** b # Exponentiation
12 | 81.0
13 | # But integer behaviour can be different!
14 | >> a = 3; b = 4

```



```

15 | >> a / b
16 | 0
17 | >> b / a
18 | 1

```

A.1.2 Iterables

Python can also define variables as lists, tuples and sets. These data types can form the basis for iterable operations. It should be noted that unlike other languages, such as Matlab or Fortran, Python iterable locations are zero-ordered (i.e. the first location in a list has an index value of 0, rather than 1).

- **List** A simple list of objects, which have the same or different data types. Data in lists can be re-assigned or replaced

```

1 | >> a_list = [3.0, 4.0, 5.0]
2 | >> print a_list
3 | [3.0, 4.0, 5.0]
4 | >> another_list = [3.0, "apples", False]
5 | >> print another_list
6 | [3.0, 'apples', False]
7 | >> a_list[2] = -1.0
8 | a_list = [3.0, 4.0, -1.0]

```

- **Tuples** Collections of objects that can be iterated upon. As with lists, they can support mixed data types. However, objects in a tuple cannot be re-assigned or replaced.

```

1 | >> a_tuple = (3.0, "apples", False)
2 | >> print a_tuple
3 | (3.0, 'apples', False)
4 | # Try re-assigning a value in a tuple
5 | >> a_tuple[2] = -1.0
6 | TypeError                                Traceback (most recent call last)
7 | <ipython-input-43-644687cfd23c> in <module>()
8 | ----> 1 a_tuple[2] = -1.0
9 |
10 | TypeError: 'tuple' object does not support item assignment

```

- **Range** A range is a convenient function to generate arithmetic progressions. They are called with a start, a stop and (optionally) a step (which defaults to 1 if not specified)

```

1 | >> a = range(0, 5)
2 | >> print a
3 | [0, 1, 2, 3, 4] # Note that the stop number is not
4 |                # included in the set!
5 | >> b = range(0, 6, 2)
6 | >> print b
7 | [0, 2, 4]

```

- **Sets** A set is a special case of an iterable in which the elements are unordered, but contains more enhanced mathematical set operations (such as intersection, union, difference, etc.)

```

1 | >> from sets import Set
2 | >> x = Set([3.0, 4.0, 5.0, 8.0])
3 | >> y = Set([4.0, 7.0])
4 | >> x.union(y)
5 | Set([3.0, 4.0, 5.0, 7.0, 8.0])
6 | >> x.intersection(y)
7 | Set([4.0])
8 | >> x.difference(y)
9 | Set([8.0, 3.0, 5.0]) # Notice the results are not ordered!

```

Indexing

For some iterables (including lists, sets and strings) Python allows for subsets of the iterable to be selected and returned as a new iterable. The selection of elements within the set is done according to the index of the set.

```

1 >> x = range(0, 10) # Create an iterable
2 >> print x
3 [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
4 >> print x[0] # Select the first element in the set
5 0 # recall that iterables are zero-ordered!
6 >> print x[-1] # Select the last element in the set
7 9
8 >> y = x[:] # Select all the elements in the set
9 >> print y
10 [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
11 >> y = x[:4] # Select the first four element of the set
12 >> print y
13 [0, 1, 2, 3]
14 >> y = x[-3:] # Select the last three elements of the set
15 >> print y
16 [7, 8, 9]
17 >> y = x[4:7] # Select the 4th, 5th and 6th elements
18 >> print y
19 [4, 5, 6]

```

A.1.3 Dictionaries

Python is capable of storing multiple data types associated with a map of variable names inside a single object. This is called a “Dictionary”, and works in a similar manner to a “data structure” in languages such as Matlab. Dictionaries are used frequently in the HMTK as ways of structuring inputs to functions that share a common behaviour but may take different numbers and types of parameters on input.

```

1 >> earthquake = {"Name": "Parkfield",
2                  "Year": 2004,
3                  "Magnitude": 6.1,
4                  "Recording Agencies" = ["USGS", "ISC"]}
5 # To call or view a particular element in a dictionary
6 >> print earthquake["Name"], earthquake["Magnitude"]
7 Parkfield 6.1

```

A.1.4 Loops and Logicals

Python’s syntax for undertaking logical operations and iterable operations is relatively straightforward.

Logical

A simple logical branching structure can be defined as follows:

```

1 >> a = 3.5
2 >> if a <= 1.0:
3     b = a + 2.0
4     elif a > 2.0:
5         b = a - 1.0
6     else:
7         b = a ** 2.0
8 >> print b
9 2.5

```

Boolean operations can are simply rendered as and, or and not.

```
1 >> a = 3.5
2 >> if (a <= 1.0) or (a > 3.0):
3     b = a - 1.0
4     else:
5         b = a ** 2.0
6 >> print b
7 2.5
```

Looping

There are several ways to apply looping in python. For simple mathematical operations, the simplest way is to make use of the **range** function:

```
1 >> for i in range(0, 5):
2     print i, i ** 2
3 0 0
4 1 1
5 2 4
6 3 9
7 4 16
```

The same could be achieved using the while function (though possibly this approach is far less desirable depending on the circumstance):

```
1 >> i = 0
2 >> while i < 5:
3     print i, i ** 2
4     i += 1
5 0 0
6 1 1
7 2 4
8 3 9
9 4 16
```

A for loop can be applied to any iterable:

```
1 >> fruit_data = ["apples", "oranges", "bananas", "lemons",
2                 "cherries"]
3 >> i = 0
4 >> for fruit in fruit_data:
5     print i, fruit
6     i += 1
7 0 apples
8 1 oranges
9 2 bananas
10 3 lemons
11 4 cherries
```

The same results can be generated, arguably more cleanly, by making use of the **enumerate** function:

```
1 >> fruit_data = ["apples", "oranges", "bananas", "lemons",
2                 "cherries"]
3 >> for i, fruit in enumerate(fruit_data):
4     print i, fruit
5 0 apples
6 1 oranges
7 2 bananas
8 3 lemons
9 4 cherries
```

As with many other programming languages, Python contains the statements **break** to break out of a loop, and **continue** to pass to the next iteration.

```
1 >> i = 0
2 >> while i < 10:
3     if i == 3:
4         i += 1
5         continue
6     elif i == 5:
7         break
8     else:
9         print i, i ** 2
10    i += 1
11 0 0
12 1 1
13 2 4
14 4 16
```

A.2 Functions

Python easily supports the definition of functions. A simple example is shown below. *Pay careful attention to indentation and syntax!*

```
1 >> def a_simple_multiplier(a, b):
2     """
3     Documentation string - tells the reader the function
4     will multiply two numbers, and return the result and
5     the square of the result
6     """
7     c = a * b
8     return c, c ** 2.0
9
10 >> x = a_simple_multiplier(3.0, 4.0)
11 >> print x
12 (12.0, 144.0)
```

In the above example the function returns two outputs. If only one output is assigned then that output will take the form of a tuple, where the elements correspond to each of the two outputs. To assign directly, simply do the following:

```
1 >> x, y = a_simple_multiplier(3.0, 4.0)
2 >> print x
3 12.0
4 >> print y
5 144.0
```

A.3 Classes and Inheritance

Python is one of many languages that is fully object-oriented, and the use (and terminology) of objects is prevalent throughout the HMTK and this manual. A full treatise on the topic of object oriented programming in Python is beyond the scope of this manual and the reader is referred to one of the many textbooks on Python for more examples

A.3.1 Simple Classes

A class is an object that can hold both attributes and methods. For example, imagine we wish to convert an earthquake magnitude from one scale to another; however, if the earthquake occurred after a user-defined year we wish to use a different formula. This could be done by a method, but we can also use a class:

```

1 >> class MagnitudeConverter(object):
2     """
3     Class to convert magnitudes from one scale to another
4     """
5     def __init__(self, converter_year):
6         """
7         """
8         self.converter_year = converter_year
9
10    def convert(self, magnitude, year):
11        """
12        Converts the magnitude from one scale to another
13        """
14        if year < self.converter_year:
15            converted_magnitude = -0.3 + 1.2 * magnitude
16        else:
17            converted_magnitude = 0.1 + 0.94 * magnitude
18        return converted_magnitude
19
20 >> converter1 = MagnitudeConverter(1990)
21 >> mag_1 = converter1.convert(5.0, 1987)
22 >> print mag_1
23 5.7
24 >> mag_2 = converter1.convert(5.0, 1994)
25 >> print mag_2
26 4.8
27 # Now change the conversion year
28 >> converter2 = MagnitudeConverter(1995)
29 >> mag_1 = converter2.convert(5.0, 1987)
30 >> print mag_1
31 5.7
32 >> mag_2 = converter2.convert(5.0, 1994)
33 >> print mag_2
34 5.7

```

In this example the class holds both the attribute `converter_year` and the method to convert the magnitude. The class is created (or “instantiated”) with only the information regarding the cut-off year to use the different conversion formulae. Then the class has a method to convert a specific magnitude depending on its year.

A.3.2 Inheritance

Classes can be useful in many ways in programming. One such way is due to the property of inheritance. This allows for classes to be created that can inherit the attributes and methods of another class, but permit the user to add on new attributes and/or modify methods.

In the following example we create a new magnitude converter, which may work in the same way as the `MagnitudeConverter` class, but with different conversion methods.

```

1 >> class NewMagnitudeConverter(MagnitudeConverter):
2     """
3     A magnitude converter using different conversion
4     formulae
5     """
6     def convert(self, magnitude, year):
7         """
8         Converts the magnitude from one scale to another
9         - differently!!!
10        """
11        if year < self.converter_year:
12            converted_magnitude = -0.1 + 1.05 * magnitude

```

```

13         else:
14             converted_magnitude = 0.4 + 0.8 * magnitude
15             return converted_magnitude
16 # Now compare converters
17 >> converter1 = MagnitudeConverter(1990)
18 >> converter2 = NewMagnitudeConverter(1990)
19 >> mag1 = converter1.convert(5.0, 1987)
20 >> print mag1
21 5.7
22 >> mag2 = converter2.convert(5.0, 1987)
23 >> print mag2
24 5.15
25 >> mag3 = converter1.convert(5.0, 1994)
26 >> print mag3
27 4.8
28 >> mag4 = converter2.convert(5.0, 1994)
29 >> print mag4
30 4.4

```

A.3.3 Abstraction

Inspection of the HMTK code (<https://github.com/GEMScienceTools/hmtk>) shows frequent usage of classes and inheritance. This is useful in our case if we wish to make available different methods for the same problem. In many cases the methods may have similar logic, or may provide the same types of outputs, but the specifics of the implementation may differ. Functions or attributes that are common to all methods can be placed in a “Base Class”, permitting each implementation of a new method to inherit the “Base Class” and its functions/attributes/behaviour. The new method will simply modify those aspects of the base class that are required for the specific method in question. This allows functions to be used interchangeably, thus allowing for a “mapping” of data to specific methods.

An example of abstraction is shown using our two magnitude converters shown previously. Imagine that a seismic recording network (named “XXX”) has a model for converting from their locally recorded magnitude to a reference global scale (for the purposes of narrative, imagine that a change in recording procedures in 1990 results in a change of conversion model). A different recording network (named “YYY”) has a different model for converting their local magnitude to a reference global scale (and we imagine they also changed their recording procedures, but they did so in 1994). We can create a mapping that would apply the correct conversion for each locally recorded magnitude in a short catalogue, provided we know the local magnitude, the year and the recording network.

```

1 >> CONVERSION_MAP = {"XXX": MagnitudeConverter(1990),
2                       "YYY": NewMagnitudeConverter(1994)}
3 >> earthquake_catalogue = [(5.0, "XXX", 1985),
4                             (5.6, "YYY", 1992),
5                             (4.8, "XXX", 1993),
6                             (4.4, "YYY", 1997)]
7 >> for earthquake in earthquake_catalogue:
8     converted_magnitude = \ # Line break for long lines!
9         CONVERSION_MAP[earthquake[1]].convert(earthquake[0],
10                                                earthquake[2])
11     print earthquake, converted_magnitude
12 (5.0, "XXX", 1985) 5.7
13 (5.6, "YYY", 1992) 5.78
14 (4.8, "XXX", 1993) 4.612
15 (4.4, "YYY", 1997) 3.92

```

So we have a simple magnitude homogenisor that applies the correct function depending on

the network and year. It then becomes a very simple matter to add on new converters for new agencies; hence we have a “toolkit” of conversion functions!

A.4 Numpy/Scipy

Python has two powerful libraries for undertaking mathematical and scientific calculation, which are essential for the vast majority of scientific applications of Python: Numpy (for multi-dimensional array calculations) and Scipy (an extensive library of applications for maths, science and engineering). Both libraries are critical to both OpenQuake and the HMTK. Each package is so extensive that a comprehensive description requires a book in itself. Fortunately there is abundant documentation via the online help for Numpy www.numpy.org and Scipy www.scipy.org, so we do not need to go into detail here.

The particular facet we focus upon is the way in which Numpy operates with respect to vector arithmetic. Users familiar with Matlab will recognise many similarities in the way the Numpy package undertakes array-based calculations. Likewise, as with Matlab, code that is well vectorised is significantly faster and more efficient than the pure Python equivalent.

The following shows how to undertake basic array arithmetic operations using the Numpy library

```

1 >> import numpy as np
2 # Create two vectors of data, of equal length
3 >> x = np.array([3.0, 6.0, 12.0, 20.0])
4 >> y = np.array([1.0, 2.0, 3.0, 4.0])
5 # Basic arithmetic
6 >> x + y # Addition (element-wise)
7 np.array([4.0, 8.0, 15.0, 24.0])
8 >> x + 2 # Addition of scalar
9 np.array([5.0, 8.0, 14.0, 22.0])
10 >> x * y # Multiplication (element-wise)
11 np.array([3.0, 12.0, 36.0, 80.0])
12 >> x * 3.0 # Multiplication by scalar
13 np.array([9.0, 18.0, 36.0, 60.0])
14 >> x - y # Subtraction (element-wise)
15 np.array([2.0, 4.0, 9.0, 16.0])
16 >> x - 1.0 # Subtraction of scalar
17 np.array([2.0, 5.0, 11.0, 19.0])
18 >> x / y # Division (element-wise)
19 np.array([3.0, 3.0, 4.0, 5.0])
20 >> x / 2.0 # Division over scalar
21 np.array([1.5, 3.0, 6.0, 10.0])
22 >> x ** y # Exponentiation (element-wise)
23 np.array([3.0, 36.0, 1728.0, 160000.0])
24 >> x ** 2.0 # Exponentiation (by scalar)
25 np.array([9.0, 36.0, 144.0, 400.0])

```

Numpy contains a vast set of mathematical functions that can be operated on a vector (e.g.):

```

1 >> x = np.array([3.0, 6.0, 12.0, 20.0])
2 >> np.exp(x)
3 np.array([2.00855369e+01, 4.03428793e+02, 1.62754791e+05,
4         4.85165195e+08])
5 # Trigonometry
6 >> theta = np.array([0., np.pi / 2.0, np.pi, 1.5 * np.pi])
7 >> np.sin(theta)
8 np.array([0.0000, 1.0000, 0.0000, -1.0000])
9 >> np.cos(theta)
10 np.array([1.0000, 0.0000, -1.0000, 0.0000])

```

Some of the most powerful functions of Numpy, however, come from its logical indexing:

```
1 >> x = np.array([3.0, 5.0, 12.0, 21.0, 43.0])
2 >> idx = x >= 10.0    # Perform a logical operation
3 >> print idx
4 np.array([False, False, True, True, True])
5 >> x[idx]             # Return an array consisting of elements
6                       # for which the logical operation returned True
7 np.array([12.0, 21.0, 43.0])
```

Create, index and slice n-dimensional arrays:

```
1 >> x = np.array([[3.0, 5.0, 12.0, 21.0, 43.0],
2                  [2.0, 1.0, 4.0, 12.0, 30.0],
3                  [1.0, -4.0, -2.1, 0.0, 92.0]])
4 >> np.shape(x)
5 (3, 5)
6 >> x[:, 0]
7 np.array([3.0, 2.0, 1.0])
8 >> x[1, :]
9 np.array([2.0, 1.0, 4.0, 12.0, 30.0])
10 >> x[:, [1, 4]]
11 np.array([[ 5.0, 43.0],
12           [ 1.0, 30.0],
13           [-4.0, 92.0]])
```

The reader is referred to the online documentation for the full set of functions!