EQRM: Geoscience Australia's Earthquake Risk Model

Technical Manual

Version 3.0

David Robinson, Glenn Fulford and Trevor Dhu

Geoscience Australia

February 24, 2011

this is a blank page

this is a blank page

Acknowledgements

The authors wish to thank John Schneider, leader of the Risk Research Group at Geoscience Australia. John's expertise in the areas of earthquake hazard and risk analysis have proven invaluable throughout the project. Without his ongoing support and advice the EQRM would not be half the product that it is today.

Andres Mendez from Aon Re in Chicago is acknowledged for providing software that formed the backbone of the earthquake catalogue generation. This generous contribution, as well as his high level of collaborative support, was fundamental to the project's success.

Bradley Horton is appreciated for the programming support that he provided through a number of contracts with Ceanet. The professionalism with which Bradley conducted his work extended beyond the call of duty on many occasions. Bradley's efforts have lead to great improvements in the useability and robustness of the EQRM. Duncan Gray, Kurt Pudniks and Ken Dale, all of Geoscience Australia, are also acknowledged for their coding efforts in particular modules of the EQRM software.

The authors have sought advice on specific topics from a range of people involved in the earthquake hazard and risk fields. The following are acknowledged for their expertise and assistance; George Walker of Aon Re, Professor John McAneney of Risk Frontiers, Walt Silva of Pacific Engineering and Analysis, Don Windeler and his team at Risk Management Solutions (RMS), Mike Griffith at The University of Adelaide, Gary Gibson of the Seismology Research Centre and Brian Gaull from Guria Consulting.

A number of people at Geoscience Australia have contributed to the development of the EQRM. Members of the Newcastle and Lake Macquarie project team lead by Trevor Jones are acknowledged for their assistance in shaping the early design of the EQRM. Mark Edwards and Ken Dale are thanked for the countless engineering questions that they have fielded over the four year project. Matt Hayne, project leader of the Risk Assessment Methods Project is thanked for creating the productive and supportive environment that allowed this project to flourish. EQRM users, Cvetan Sinadinovski, Annette Patchett, Ken Dale and Augusto Sanabria are thanked for their feedback and for continually inspiring the incorporation of new functionality. Ole Nielsen and Duncan Gray are acknowledged for the provision of software engineering advice and for generally improving the coding abilities of the authors. Angie Jaensch, Greg Michalowski and Fiona Watford are acknowledged for drafting some of the figures. Finally, the authors wish to thank Adrian Hitchman, Ken Dale and Augusto Sanabria for reviewing the drafts of this report. Their reviews lead to substantial improvements to this manual.

Contents

Chapter 1

Introduction

1.1 Overview

The EQRM application is a computer model for estimating earthquake hazard and earthquake risk. Modelling earthquake hazard involves assessing the probability that certain levels of ground motion will be exceeded. Modelling of earthquake risk involves estimating the probability of a building portfolio experiencing a range of earthquake induced losses. For any number of synthetic earthquakes, the EQRM application can be used to estimate:

1. the ground motion and its likelihood of occurrence (earthquake hazard),
2. the direct financial loss and its likelihood of occurrence (earthquake risk), and
3. less reliably the number of casualties and injuries and their likelihood of occurrence (earthquake risk).

The EQRM application is Geoscience Australia's centerpiece for modelling earthquake hazard and risk. Its use formed the basis for Geoscience Australia's recent reports on Earthquake risk in the Newcastle and Lake Macquarie (**?**) and Perth (**?**) regions.

The process for computing earthquake hazard can be described by the following steps:

1. the generation of a catalogue of synthetic earthquakes (or events) (see Chapter 3),

2. the propagation or attenuation (see Chapter 5) of the 'seismic wave' from each of the events in 1 to locations of interest (see Chapter 4),

3. accounting for the interactions between the propagating 'seismic wave' and the local geology or regolith (see Chapter 6), and

4. accounting for the probability of each event and the estimation of hazard (see Chapter 9).

The process for computing earthquake risk shares the same first three steps as the earthquake hazard. The fourth step and onwards can be described as follows:

4. estimating the probability that the portfolio buildings (see Chapter 4) will experience different levels of damage (see Chapter 7),

5. the computation of direct financial loss as a result of the probabilities computed in 4 (see Chapter 8), and

6. assembling the results to compute the risk (see Chapter 9).

1.2 Using this manual

This manual describes the EQRM application; it has been designed to serve the following three purposes:

1. describe the theory and methodology behind the EQRM application;

2. explain how to use the EQRM application to model earthquake hazard and risk; and

3. provide enough information to assist those who may wish to access individual modules and modify them.

A number of features have been included to assist readers of the manual. These features include:

- **Text highlighting -** Important parameters, file names and code are identified by a typeset font.
- **Index -** An index has been introduced to allow speedy of navigation of the manual.

- **Chapter summary -** Each chapter is concluded with a 'Key functions, flags and parameters' section that summarises the key code related items that were discussed in the chapter.

1.3 About this manual

Chapter 2: The EQRM application introduces the EQRM software. The chapter describes the directory structure and introduces important input parameter files. Setup and operational processes are discussed and the conversion of the EQRM to a stand-alone executable described. Readers who do not wish to familiarise themselves with details of the EQRM application may wish to skip this chapter.

Chapter 3: Earthquake source generation discusses the creation of an earthquake catalogue. At the core of any hazard or risk assessment is the simulation of synthetic earthquakes and the creation of a catalogue of synthetic events. The chapter describes how the simulation of earthquakes can focus around a specific earthquake of interest (a scenario based simulation) or how it can model the effect of 'all foreseeable' events (a probabilistic simulation).

Chapter 4: Grids and building databases describes how calculations of risk and hazard are conducted on a point by point basis. In the case of earthquake hazard, a grid of points is created and hazard computed at each of the grid nodes. Typically a regular grid is used to improve the production of hazard maps, however this is not a requirement. Earthquake risk calculations are also performed on a grid. Each node of the grid represents the centroid of a building within a portfolio of interest. The chapter demonstrates how the grid points are associated with other attributes such as site–class or building construction type.

Chapter 5: Attenuation discusses the propagation (or attenuation) of the motion from each of the synthetic earthquakes in the event catalogue to locations of interest. Different measures of distance between an event and a site of interest are introduced. The chapter also describes the attenuation models that can be used with the EQRM application.

Chapter 6: Regolith amplification discusses how local geology, or regolith, can be incorporated in an earthquake hazard or risk assessment. The chapter illustrates how the presence of regolith can lead to amplification (or de-amplification) of both ground and building motion, and consequently result in higher (or lower) levels of hazard and risk.

Chapter 7: Building damage describes how the EQRM application estimates the probability that each building will experience different levels of damage. The

chapter gives a brief discussion of the theory behind making such estimates as well as outlining how to extrapolate the estimates to an entire portfolio of buildings in a region of interest.

Chapter 8: Losses illustrates how to model the direct financial loss as a result of the damage estimates described in *Chapter 7*. The chapter also provides a brief description of how the EQRM application can estimate injuries and casualties.

Chapter 9: Displaying hazard and risk results describes how the results from the EQRM application can be summarised and displayed. There are many ways to display estimates of earthquake hazard and risk. Earthquake hazard is commonly modelled in terms of a probability (often 10%) of a particular ground motion (usually acceleration) being exceeded in some time frame (often 50 years). In the case of a scenario based simulation, earthquake risk can be modelled in terms of dollar estimates of building loss. Such losses can be categorised by building type, geography or any other category of interest. In the case of a probabilistic simulation, earthquake risk is often modelled in terms of one or both of the following:

1. the loss per year averaged over a long period of time, typically 10000 years. This is known as the annualised loss.

2. the probability that different levels of loss will be exceeded. The plot of loss against such exceedance probabilities is referred to as a probable maximum loss curve.

Chapter 2

The EQRM application

The Earthquake Risk Model (EQRM) is capable of:

1. earthquake scenario ground motion modeling;
2. scenario loss forecasts;
3. probabilistic seismic hazard analysis (PSHA); and
4. probabilistic seismic risk analysis (PSRA).

It is a product of Geoscience Australia, an Australian Government Agency.

This chapter describes the EQRM application. Input files and parameters are discussed and directions on how to run the EQRM provided. Readers who are interested in only the EQRM methodology and not the EQRM software package may wish to skip this chapter.

The input files required by the EQRM depend on the nature of the simulation conducted. For example, the inputs for a scenario loss simulation are different to those required for a probabilistic seismic hazard analysis. Table 2.1 provides a summary of the inputs required by the EQRM. The EQRM Demos in `*/eqrm_core/demo` provide examples of each input file and demonstrate how to run the EQRM for each of the four main simulation types. The following section provide an overview of each of the main input files.

Except for the control file all input files are assumed to be in the input directory, `input_dir`, specified in the EQRM control file. If a file in not there it is looked for in `*/eqrm_core/resources/data`.

Table 2.1: Input files required for different types of simulation with the EQRM. The asterisks indicate optional input files, the requirement for which depends on settings in the EQRM control file

	hazard	**risk**
scenario	EQRM control file amplification factors* hazard grid	EQRM control file amplification factors* building database
probabilistic	EQRM control file source file(s) event type control file amplification factors* hazard grid	EQRM controlfile source file(s) event type control file amplification factors* building database

2.1 The EQRM Control File

The EQRM control file is the primary input file for an EQRM simulation. It:

1. contains a series of input variables (or parameters) that define the manner in which the EQRM is operated; and
2. initialises the simulation.

For example, there is a parameter to control whether the EQRM models hazard or risk. Other parameters can be used to identify return periods or indicate whether site amplification is considered. A description of all of the parameters is given below.

Note that it not essential to supply all parameters for each simulation. For example, if amplification is not being used (i.e. `use_amplification = False`) it is not necessary to supply the remaining amplification parameters. Furthermore, default values are set by the EQRM for several parameters. These are indicated below when applicable. Omission of these input parameters in the EQRM control file will lead to use of the default values. For example, the default value for `atten_threshold_distance` is 400 km.

The following also provides suggested values for several parameters. Users are free to change these values as desired. The developers are merely suggesting the value they would use in most circumstances. For example, the suggested value for `loss_min_pga` is 0.05.

The term preferred is used to indicate those parameters that the developers believe to be most appropriate. For example, the preferred value for `csm_hysteretic_damping` is `curve`. In this case the alternative choices of `None` and `trapezoidal` would typically only be used for experimental purposes.

A simulation can be started by executing an EQRM control file or by executing `analysis.py` and passing a control file in. The control file is a Python script, so Python code can be used to manipulate parameter values. Note though that all variables other than the parameter values must be deleted to avoid passing unknown variables into EQRM.

Acronyms:
PSHA is probabilistic seismic hazard analysis
PSRA is probabilistic seismic risk analysis
GMPE is ground motion prediction equation
PGA is peak ground acceleration (usually in units of g)
RSA is response spectral acceleration (usually in units of g)
CSM is capacity spectrum method

General Input:

`run_type`:
Defines the operation mode of the EQRM:
'hazard' ⇒ Scenario RSA and PSHA (probabilistic hazard);
'risk' ⇒ Scenario Loss and PSRA (probabilistic risk);

`is_scenario`:
Event simulation type:
True ⇒ a specific scenario event (Use Scenario input);
False ⇒ probabilistic simulation, PSHA or PSRA (Use Probabilistic input)

`max_width`:
Maximum width along virtual faults i.e. synthetic rupture width can not exceed `max_width` (km).

`site_tag`:
String used in input and output file names. Typically used to define the city or study of interest (e.g. `newc` is used in the demos).

`site_db_tag`:
DEFAULT = ''
String used to specify the exposure or building data base. The file name is `sitedb_<site_tag><site_db_tag>.csv`

`return_periods`:
List whose elements represent the return periods to be considered for PSHA.

`input_dir`:
Directory containing any local input files.

`output_dir`:
Directory for output files. This directory will be created if not present

`use_site_indexes`:
DEFAULT = `False`
True ⇒ sample sites with indices in `site_indexes` (for testing simulations);
`False` ⇒ No sub-sampling.

`site_indexes`:
List whose elements represent the site indices to be used (if `use_site_indexes` = `True`). The index of the first row of data (i.e. first data row in site file) is 1.

zone_source_tag:
Extra tag for specifying a unique source zone file. If zone_source_tag is defined the filename for the zone source file is <site_tag>_zone_source_<fault_source_tag>.xml. Otherwise it is <site_tag>_zone_source.xml.

fault_source_tag:
Extra tag for specifying a unique source fault file. If fault_source_tag is defined the filename for the fault source file is <site_tag>_fault_source_<fault_source_tag>.xml. Otherwise it is <site_tag>_fault_source.xml.

event_control_tag:
Extra tag for specifying a unique event control file. If fault_source_tag is defined the filename for the event control is <site_tag>_event_control_<event_control_tag>.xml Otherwise it is <site_tag>_event_control.xml.

Scenario Input:

scenario_azimuth:
Azimuth of the scenario event (degrees from true North).

scenario_latitude:
Latitude of rupture centroid.

scenario_longitude:
Longitude of rupture centroid.

scenario_magnitude:
Moment magnitude of event.

scenario_depth:
Depth to event centroid (km).

scenario_dip:
Dip of rupture plane (degrees from horizontal).

scenario_number_of_events:
The desired number of copies of the event to be generated. Typically, copies are taken if random sampling is used to incorporate aleatory uncertainty in GMPE (i.e. atten_variability_method= 2), amplification (i.e. amp_variability_method= 2) or the CSM (csm_variability_method= 3).

Ground Motion Input:

`atten_collapse_Sa_of_atten_models`:
DEFAULT = `False`
Set to `True` to collapse the surface acceleration's when multiple GMPEs are used.

`atten_variability_method`:
DEFAULT = 2
Technique used to incorporate GMPE aleatory uncertainty:
None ⇒ No sampling;
1 ⇒ spawning;
2 ⇒ random sampling;
3 ⇒ $+2\sigma$;
4 ⇒ $+\sigma$;
5 ⇒ $-\sigma$;
6 ⇒ -2σ.

`atten_periods`:
Periods for RSA. Values must ascend.

`atten_threshold_distance`:
DEFAULT = 400
Threshold distance (km) beyond which motion is assigned to zero.

`atten_spawn_bins`:
DEFAULT = None
Number of bins created when spawning.

`atten_override_RSA_shape`:
DEFAULT = `None`
Use GMPE for PGA only and change shape of RSA. If 'None' use RSA as defined by GMPE, otherwise if
`'Aust_standard_Sa'` ⇒ use RSA shape from Australian earthquake loading standard;
`'HAZUS_Sa'` ⇒ use RSA shape defined by HAZUS;

`atten_cutoff_max_spectral_displacement`: DEFAULT = `False` `True` ⇒ cutoff maximum spectral displacement. `False` ⇒ no cutoff applied to spectral displacement.
`atten_pga_scaling_cutoff`: DEFAULT = 2 The maximum acceptable PGA in units g. RSA at all periods re-scaled accordingly.
`atten_smooth_spectral_acceleration`: DEFAULT = `False` `True` ⇒ Smooth RSA; `False` ⇒ No smoothing applied to RSA.

Amplification Input:

`use_amplification`:
If set to `True` use amplification associated with the local regolith. Nature of amplification varies depending on the GMPE. If GMPE has a V_{S30} term then this will be used to compute RSA on regolith. Otherwise, RSA is computed on bedrock and amplification factors used to transfer this to regolith surface.

`amp_variability_method`:
DEFAULT = 2
Technique used to incorporate amplification aleatory uncertainty:
None $\Rightarrow$ No sampling;
2 $\Rightarrow$ random sampling;
3 $\Rightarrow$ $+2\sigma$;
4 $\Rightarrow$ $+\sigma$;
5 $\Rightarrow$ $-\sigma$;
6 $\Rightarrow$ -2σ.
7 $\Rightarrow$ -2σ.

`amp_min_factor`:
SUGGESTED = 0.6
Minimum accepted value for amplification factor. This minimum will not be used for Vs30 models.

`amp_max_factor`:
SUGGESTED = 10000
Maximum accepted value for amplification factor. This maximum will not be used for Vs30 models.

Building Classes Input:

`buildings_usage_classification`:
Building usage classification system - `'HAZUS'` or `'FCB'`

`buildings_set_damping_Be_to_5_percent`:
SUGGESTED = `False`
If `True` use a damping B_e of 5% for all building structures.

Capacity Spectrum Method Input:

`csm_use_variability`:
SUGGESTED = `True`
`True` ⇒ use the variability method described by `csm_variability_method`;
`False` ⇒ no aleatory variability applied.

`csm_variability_method`:
SUGGESTED = 3
Method used to incorporate variability in capacity curve:
None ⇒ No sampling;
3 ⇒ Random sampling applied to ultimate point only and yield point re-calculated to satisfy capacity curve 'shape' constraint.

`csm_standard_deviation`:
SUGGESTED = 0.3
Standard deviation for capacity curve log–normal PDF.

`csm_damping_regimes`:
SUGGESTED = 0
Damping multiplicative formula to be used:
0 ⇒ PREFERRED: use R_a, R_v, and R_d;
1 ⇒ use R_a, R_v and assign $R_d = R_v$;
2 ⇒ use R_v only and assign $R_a = R_d = R_v$.

`csm_damping_modify_Tav`:
SUGGESTED = `True`
Modify transition building period i.e. corner period T_{av}:
`True` ⇒ PREFERRED: modify as in HAZUS;
`False` ⇒ do NOT modify.

`csm_damping_use_smoothing`:
SUGGESTED = `True`
Smoothing of damped curve:
`True` ⇒ PREFERRED: apply smoothing;
`False` ⇒ NO smoothing.

`csm_hysteretic_damping`:
SUGGESTED = `'curve'`
Technique for Hysteretic damping:
`None` ⇒ no hysteretic damping
`'trapezoidal'` ⇒ Hysteretic damping via trapezoidal approximation;
`'curve'` ⇒ PREFERRED: Hysteretic damping via curve fitting.

`csm_SDcr_tolerance_percentage`:
SUGGESTED = 1.0
Convergence tolerance as a percentage for critical spectral displacement in non-linear damping calculations.

`csm_damping_max_iterations`:
SUGGESTED = 7
Maximum iterations for nonlinear damping calculations.

Loss Input:

`loss_min_pga`:
SUGGESTED = 0.05
Minimum PGA(g) below which financial loss is assigned to zero.

`loss_regional_cost_index_multiplier`:
SUGGESTED = 1
Regional cost index multiplier to convert dollar values in building database to desired regional and temporal (i.e. inflation) values.

`loss_aus_contents`:
SUGGESTED = 0
Contents value for residential buildings and salvageability after complete building damage:
0 ⇒ contents value as defined in building database and salvageability of 50%;
1 ⇒ 60% of contents value as defined in building database and salvageability of zero.

Save Input:

`save_hazard_map`:
DEFAULT = `False`
`True` ⇒ Save data for hazard maps (Use for saving PSHA results). Specifically spectral acceleration with respect to location, return period and period.

`save_total_financial_loss`:
DEFAULT = `False`
`True` ⇒ Save total financial loss.

`save_building_loss`:
DEFAULT = `False`
`True` ⇒ Save building loss.

`save_contents_loss`:
DEFAULT = `False`
`True` ⇒ Save contents loss.

`save_motion`:
DEFAULT = `False`
`True` ⇒ Save RSA motion (use for saving scenario ground motion results).

`save_prob_strucutural_damage`:
DEFAULT = `False`
`True` ⇒ Save structural non-cumulative probability of being in each damage state. Note this is only supported for a single event.

The following grey shaded box provides an example of an EQRM controlfile to undertake a PSHA.

```
"""
  EQRM parameter file
   All input files are first searched for in the input_dir, then in the
   resources/data directory, which is part of EQRM.

 All distances are in kilometers.
 Acceleration values are in g.
 Angles, latitude and longitude are in decimal degrees.

 If a field is not used, set the value to None.


"""

from eqrm_code.parse_in_parameters import eqrm_data_home, get_time_user
from os.path import join


# Operation Mode
run_type = "hazard"
is_scenario = False
max_width = 15
site_tag = "newc"
site_db_tag = ""
return_periods = [10, 50, 100, 10000]
input_dir = r".\input/"
output_dir = r".\output\prob_haz/"
use_site_indexes = True
site_indexes = [2255, 11511, 10963, 686]
fault_source_tag = "no_fault"

# Scenario input

# Probabilistic input

# Attenuation
atten_models = ['Sadigh_97']
atten_model_weights = [1]
atten_collapse_Sa_of_atten_models = True
atten_variability_method = 2
atten_periods = [0.0, 0.30303000000000002, 1.0]
atten_threshold_distance = 400
atten_override_RSA_shape = None
atten_cutoff_max_spectral_displacement = False
atten_pga_scaling_cutoff = 2
atten_smooth_spectral_acceleration = None
atten_log_sigma_eq_weight = 0

# Amplification
use_amplification = True
amp_variability_method = 2
amp_min_factor = 0.6
amp_max_factor = 10000

# Buildings

# Capacity Spectrum Method

# Loss

# Save
```

```
save_hazard_map = True
save_total_financial_loss = False
save_building_loss = False
save_contents_loss = False
save_motion = False
save_prob_structural_damage = None

# If this file is executed the simulation will start.
# Delete all variables that are not EQRM parameters variables.
if __name__ == '__main__':
    from eqrm_code.analysis import main
    main(locals())
```

2.2 The Source Files

The EQRM source files for probabilistic modeling (PSHA and PSRA) come in two forms. These are:

- source zones, and
- faults.

The EQRM can be run with either of these inputs separately or both together

2.2.1 Source Zone File

If `<zone_source_tag>` is defined then the file name is:
`<site_tag>_zone_source_<zone_source_tag>.xml`

Otherwise the file name is:
`<site_tag>_zone_source.xml`

The source zone file is used to describe one or more areal source zones. Earthquakes are assumed to be equally likely to occur anywhere within a source zone. The magnitude recurrence relationship for each source zone is defined by a bounded Gutenberg-Richter relationship. The following grey shaded box provides an example of a source zone file. A description of the parameters follows.

```
<source_model_zone magnitude_type="Mw">
    <zone area = "5054.035"
  event_type = "crustal fault">
    <geometry
    dip = "35"
    delta_dip = "0"
    azimuth = "180"
    delta_azimuth = "180"
    depth_top_seismogenic = "7"
       depth_bottom_seismogenic = "15.60364655" >
       <boundary>
  -32.4000   151.1500
  -32.7500   152.1700
  -33.4500   151.4300
  -32.4000   151.1500
         </boundary>
    </geometry  >
        <recurrence_model
         distribution = "bounded_gutenberg_richter"
```

```
         recurrence_min_mag = "3.3"
         recurrence_max_mag = "5.4"
         A_min = "0.568"
          b = "1">
          <event_generation
          generation_min_mag = "4.5"
          number_of_mag_sample_bins = "15"
          number_of_events = "0"
           /></recurrence_model>
    </zone>


    <zone area = "57731.425"
  event_type = "crustal fault">
    <geometry
    dip = "35"
    delta_dip = "0"
    azimuth = "180"
    delta_azimuth = "180"
    depth_top_seismogenic = "7"
       depth_bottom_seismogenic = "15.60364655">
        <boundary>
-31.0000  149.5000
  -32.4000  149.5000
  -32.4000  151.1500
  -32.7500  152.1700
  -32.7500  152.7600
  -32.7000  152.8000
  -32.0000  153.1100
  -31.0000  153.2900
  -31.0000  149.5000
</boundary>
    </geometry>
        <recurrence_model
        distribution = "bounded_gutenberg_richter"
         recurrence_min_mag = "3.3"
         recurrence_max_mag = "5.4"
        A_min= "2.53"
         b = "1.14" >
          <event_generation
          generation_min_mag = "4.5"
          number_of_mag_sample_bins = "15"
          number_of_events = "0" /></recurrence_model>
    </zone>


    <zone area = "56703.105"
  event_type = "crustal fault">
    <geometry
    dip = "35"
    delta_dip = "0"
    azimuth = "180"
    delta_azimuth = "180"
    depth_top_seismogenic = "7"
       depth_bottom_seismogenic = "15.60364655">
        <boundary>
  -35.0000  149.5000
  -32.4000  149.5000
  -32.4000  151.1500
  -33.4500  151.4300
  -32.7500  152.1700
  -32.7500  152.7600
  -34.4000  151.3500
  -34.7400  151.1500
```

```
-35.0000   151.1000
-35.0000   149.5000
      </boundary>
  </geometry>
      <recurrence_model
       distribution = "bounded_gutenberg_richter"
        recurrence_min_mag = "3.3"
        recurrence_max_mag = "5.4"
       A_min= "2.48"
        b = "1.14" >
         <event_generation
         generation_min_mag = "4.5"
         number_of_mag_sample_bins = "15"
         number_of_events = "0" /></recurrence_model>
  </zone>


  <zone area = "3046.615"
event_type = "crustal fault">
  <geometry
  dip = "35"
  delta_dip = "0"
  azimuth = "180"
  delta_azimuth = "180"
  depth_top_seismogenic = "7"
     depth_bottom_seismogenic = "15.60364655">
      <boundary>
-32.9250   151.4000
-32.7500   151.7500
-33.2500   152.2500
-33.5000   151.9000
-32.9250   151.4000
      </boundary>
  </geometry>
      <recurrence_model
       distribution = "bounded_gutenberg_richter"
        recurrence_min_mag = "5.41"
        recurrence_max_mag = "6.5"
       A_min= "0.0016"
        b = "1." >
         <event_generation
         generation_min_mag = "4.5"
         number_of_mag_sample_bins = "15"
         number_of_events = "0" /></recurrence_model>
  </zone>


  <zone area = "72204.957"
event_type = "crustal fault">
  <geometry
  dip = "35"
  delta_dip = "0"
  azimuth = "180"
  delta_azimuth = "180"
  depth_top_seismogenic = "7"
     depth_bottom_seismogenic = "15.60364655">
      <boundary>
-31.0000   149.5000
-32.9250   149.5000
-32.9250   151.4000
-32.7500   151.7500
-33.2500   152.2500
-33.2500   152.3300
-32.7000   152.8000
```

```
  -32.0000   153.1100
  -31.0000   153.2900
  -31.0000   149.5000
        </boundary>
    </geometry>
        <recurrence_model
        distribution = "bounded_gutenberg_richter"
         recurrence_min_mag = "5.41"
         recurrence_max_mag = "6.5"
        A_min= "0.014"
         b = "1.118" >
          <event_generation
          generation_min_mag = "4.5"
          number_of_mag_sample_bins = "15"
          number_of_events = "0" /></recurrence_model>
    </zone>


    <zone area = "44149.044"
  event_type = "crustal fault">
    <geometry
    dip = "35"
    delta_dip = "0"
    azimuth = "180"
    delta_azimuth = "180"
    depth_top_seismogenic = "7"
       depth_bottom_seismogenic = "15.60364655">
        <boundary>
  -35.0000   149.5000
  -32.9250   149.5000
  -32.9250   151.4000
  -33.5000   151.9000
  -33.2500   152.2500
  -33.2500   152.3300
  -34.4000   151.3500
  -34.7400   151.1500
  -35.0000   151.1000
  -35.0000   149.5000
        </boundary>
    </geometry>
        <recurrence_model
        distribution = "bounded_gutenberg_richter"
         recurrence_min_mag = "5.41"
         recurrence_max_mag = "6.5"
        A_min= "0.0086"
         b = "1.118" >
          <event_generation
          generation_min_mag = "4.5"
          number_of_mag_sample_bins = "15"
          number_of_events = "0"/></recurrence_model>
    </zone>
</source_model_zone>
```

General inputs (`source_model_zone`)

- `magnitude_type`: Earthquake magnitude used to derive the recurrence parameters. NOTE - the EQRM only supports moment magnitude `Mw`.

General zone inputs (`zone`)

- `area`: Area of the source zone in km^2
- `name`: Name for the source zone
- `event_type`: Pointer to the collection of inputs described in the event type controlfile.

Geometry inputs (`geometry`)

- `azimuth`: Center azimuth for randomly generated synthetic ruptures (degrees).
- `delta_azimuth`: Range over which randomly generated azimuths will be sampled. That is, the azimuth of all synthetic earthquake will be randomly drawn from a uniform distribution between `azimuth`±`delta_azimuth` (degrees).
- `dip`: Center dip for randomly generated synthetic ruptures (degrees).
- `delta_dip`: Range over which randomly generated dips will be sampled. That is, the dip of all synthetic earthquake will be randomly drawn from a uniform distribution between `dip`±`delta_dip` (degrees).
- `depth_top_seismogenic`: Depth (km) to the top of the seismogenic zone in km. No component of a synthetic rupture will be located above this value.
- `depth_bottom_seismogenic`: Depth (km) to the bottom of the seismogenic zone in km. No component of a synthetic rupture will be located below this value.
- `boundary`: Boundary of the areal source zone as defined on the surface of the Earth in longitude (column 1) and latitude (column 2). The first and last points must be the same to close the polygon.
- `excludes`: Boundary of any regions in the source in which events are not required. Boundary defined on the surface of the Earth in longitude (column 1) and latitude (column 2). The first and last points must be the same to close the polygon.This parameter is optional. the source zone file may have no `exclude` zones, a single entry or multiple entries.

Recurrence inputs (`recurrence_model`)

- `distribution`: Distribution used to define the magnitude recurrence relations. Note that the EQRM current only supports a Bounded Gutenberg-Richter recurrence relationship for source zones (i.e. `bounded_gutenberg_richter`)
- `recurrence_min_mag`: Minimum magnitude used to define the recurrence relationship
- `recurrence_max_mag`: Maximum magnitude used to define the recurrence relationship. Typically, this is the magnitude of the largest earthquake expected in the zone.
- `A_min`: Expected number of earthquakes with magnitude `recurrence_min_mag` or higher in the source zone per year.
- `b`: Gutenberg-Richter b value for bounded Gutenberg-Richter recurrence relationship
- `generation_min_mag`: Minimum magnitude for synthetic earthquake generation. The EQRM will only generate synthetic earthquakes with magnitudes equal to or greater than `generation_min_mag`.
- `number_of_mag_sample_bins`: Number of magnitude bins used to discretise the recurrence relationship in the magnitude range `generation_min_mag` to `recurrence_max_mag`
- `number_of_events`: Number of syntectic ruptures to be generated in the source zone.

2.2.2 Source Fault File

If `<fault_source_tag>` is defined then the file name is:
`<site_tag>_fault_source_<fault_source_tag>.xml`

Otherwise the file name is:
`<site_tag>_fault_source.xml`

The source faults file is used to describe one or more faults (including crustal faults and subduction interfaces) and/or one or more dipping slabs for intraslab earthquakes. Earthquakes are assumed to be equally likely to occur anywhere

within the fault (or slab). The magnitude recurrence for faults can be defined by a bounded Gutenberg-Richter relationship or a combination of bounded Gutenberg-Richter and Characteristic. The magnitude recurrence for the intraslab earthquakes must be defined by a bounded Gutenberg-Richter relationship. The following grey box provides an example of a source fault file with the following source types:

1. *crustal fault* with recurrence defined by a bounded Gutenberg-Richter relationship (`fault 1`),
2. *crustal fault* with recurrence defined by a combined bounded Gutenberg-Richter (for small earthquakes) and a characteristic recurrence for larger earthquakes (`fault 2`),
3. a *3D dipping volume* to represent intraslab earthquakes in the subducting slab (`intraslab 1`).

Many of the parameters in the source fault file are identical to those described in Section 2.2.1 and are not described separately here. A description of the new parameters is provided below.

```
<source_model_fault magnitude_type="Mw">
  <fault
  name = "Newie South Fault"
  event_type = "crustal fault">

    <geometry
        dip= "45"
        out_of_dip_theta = "0"
        delta_theta = "0"
        depth_top_seismogenic = "2"
        depth_bottom_seismogenic = "10"
        slab_width = "0">
        <trace>
            <start lat="-33.1" lon="151.5" />
            <end lat="-33.1" lon="151.8" />
        </trace>
    </geometry>

    <recurrence_model
       distribution = "bounded_gutenberg_richter"
       recurrence_min_mag = "3.5"
       recurrence_max_mag = "6.5"
       slip_rate = "10.0"
       b = "1">
      <event_generation
        generation_min_mag = "3.5"
        number_of_mag_sample_bins = "15"
        number_of_events = "1500" />
    </recurrence_model>

  </fault>

  <fault
```

```
  name = "Newie North Fault"
  event_type = "crustal fault">

    <geometry
        dip= "45"
        out_of_dip_theta = "0"
        delta_theta = "0"
        depth_top_seismogenic = "2"
        depth_bottom_seismogenic = "10"
        slab_width = "0">
        <trace>
            <start lat="-32.9" lon="151.5" />
            <end lat="-32.9" lon="151.8" />
        </trace>
    </geometry>

    <recurrence_model
       distribution = "characteristic"
       recurrence_min_mag = "3.5"
       recurrence_max_mag = "6.5"
       slip_rate = "10.0"
       b = "1">
      <event_generation
        generation_min_mag = "3.5"
        number_of_mag_sample_bins = "15"
        number_of_events = "1500" />
    </recurrence_model>

  </fault>

  <fault
  name = "Newie Subduction Zone"
  event_type = "subduction interface">

    <geometry
        dip= "30"
        out_of_dip_theta = "90"
        delta_theta = "10"
        depth_top_seismogenic = "5"
        depth_bottom_seismogenic = "100"
        slab_width = "0">
        <trace>
            <start lat="-28.0" lon="152.5" />
            <end lat="-38.0" lon="152.5" />
        </trace>
    </geometry>

    <recurrence_model
       distribution = "bounded_gutenberg_richter"
       recurrence_min_mag = "4.0"
       recurrence_max_mag = "8.5"
       A_min = "10"
       b = "1">
      <event_generation
        generation_min_mag = "4.0"
        number_of_mag_sample_bins = "25"
        number_of_events = "2000" />
    </recurrence_model>

  </fault>

</source_model_fault>
```

Parameters unique to the source fault file

- `dip`: Dip of fault, defined as angle in degrees from horizontal.
- `out_of_dip_theta`: Out of plane dip, used for intraslab events. Angle between fault plane and out of dip rupture plane (degrees).
- `delta_theta`: Bounds the range of dips for intraslab events. That is, all synthetic ruptures will have uniformly random sampled dips in the range `dip + out_of_dip_theta` $\pm$ `delta_theta` (degrees).
- `slab_width`: Width of slab (km) when using a fault source to represent intraslab earthquakes in the subducting slab.
- `trace`: Surface trace of the fault along the surface of the Earth. Note that it is the projection of the fault along the direction of dip. It is defined by the latitude (`lat`) and longitude (`lon`) of the start and end of the trace.
- `slip_rate`: Slip rate of fault in mm per year.
- `distribution`: Distribution used to define the magnitude recurrence relations. For faults the EQRM supports (i) a Bounded Gutenberg-Richter recurrence relationship (`bounded_gutenberg_richter`) or (ii) a combined Bounded Gutenberg-Richter and Characteristic model (`characteristic`). For intraslab earthquakes the EQRM supports only `bounded_gutenberg_richter`.
- `recurrence_max_mag`: Maximum magnitude used to define the recurrence relationship. Typically, this is the magnitude of the largest earthquake expected on the fault (or in the subducting slab).

2.3 Event Type Control File

Filename: `<site_tag>_event_control_<event_control_tag>.xml`

The event type control file is a second level control file facilitating the variation of selected EQRM parameters with event types. The mechanism for this is an `event_type` parameter which links the `event_type_control` file with individual sources (i.e. specific zones, faults or dipping slabs) in the `fault_source` and/or `zone_source` files.

```
<event_type_controlfile>
  <event_group
     event_type = "background">
    <GMPE
       fault_type = "normal">
      <branch model = "Toro_1997_midcontinent"  weight = "0.3"/>
      <branch model = "Atkinson_Boore_97" weight = "0.4"/>
      <branch model = "Sadigh_97" weight = "0.3"/>
    </GMPE>
    <scaling scaling_rule = "Wells_and_Coppersmith_94"
    scaling_fault_type = "unspecified" />
  </event_group>

  <event_group
     event_type = "crustal fault">
    <GMPE
       fault_type = "reverse">
      <branch model = "Campbell08" weight = "1"/>
    </GMPE>
    <scaling scaling_rule = "Wells_and_Coppersmith_94"
    scaling_fault_type = "reverse" />
  </event_group>


  <event_group
     event_type = "interface">
    <GMPE
       fault_type = "reverse">
      <branch model = "Zhao06_crustalinterface" weight = "0.5"/>
      <branch model = "Atkinson03_interface" weight = "0.5"/>
    </GMPE>
    <scaling scaling_rule = "Wells_and_Coppersmith_94"
    scaling_fault_type = "reverse" />
  </event_group>

  <event_group
     event_type = "intraslab">
    <GMPE
       fault_type = "reverse">
      <branch model = "Zhao06_slab" weight = "0.5"/>
      <branch model = "Atkinson03_inslab" weight = "0.5"/>
    </GMPE>
    <scaling scaling_rule = "Wells_and_Coppersmith_94"
    scaling_fault_type = "unspecified" />
  </event_group>

</event_type_controlfile>
```

Parameters in the event type control file are separated into event groups. These are blocks of input parameters defined by `<event_group ... </event_group>`. Each of these blocks is linked to a specific source in the source zone or source fault files using `event_type`.

The parameters enclosed within `<GMPE ... </GMPE>` define the use of ground motion prediction equations. These parameters include:

- `fault_type`: fault mechanism used with the GMPE. Allowable options are `normal`, `reverse` and `strike_slip`.

- `branch`: specifies a branch for the GMPE logic tree. There may be a single `branch` in which case a single GMPE is used or multiple `branch`es in which case multiple GMPEs are used in a logic tree. Inside each branch the user must specify the chosen GMPEs (`model`: see below for a list of options) and the weights (`weight`) for each branch. The weights for all branches in a given `GMPE` block must sum to 1.

The parameters enclosed within `<scaling` and `/>` control the magnitude to size scaling during the generation of synthetic ruptures. These parameters include:

- `scaling_rule`: Defines the set of scaling rules which link M_w to area, length and/or width. Currently the only set of scaling rules supported by the EQRM are those defined by Wells and Coppersmith (1994). Therefore, the only allowable options are `Wells_and_Coppersmith_94` and `modified_Wells_and_Coppersmi`
- `scaling_fault_type`: Fault mechanism used with the scaling rule. Allowable options are `normal`, `reverse`, `strike_slip` and `unspecified`. Typically, `scaling_fault_type` will be the same as the GMPE `fault_type`, however the EQRM does not enforce this.

Current options for the GMPE are:

"Gaull_1990_WA" ⇒ Gaull *et al.* (1990);
"Toro_1997_midcontinent" ⇒ Toro *et al.* (1997) model for mid-continent USA;
"Atkinson_Boore_97" ⇒ Atkinson *et al.* (1997);
"Sadigh_97" ⇒ Sadigh *et al.* (1997);
"Youngs_97_interface" ⇒ Youngs *et al.* (1997) interface (Z_T=0);
"Youngs_97_intraslab" ⇒ Youngs *et al.* (1997) intraslab (Z_T=1);
"Combo_Sadigh_Youngs_M8" ⇒ combined Youngs *et al.* (1997) and Sadigh *et al.* (1997);
"Boore_08" ⇒ Boore *et al.* (2008);
"Sommerville09_Yilgarn" ⇒ Sommerville (2009) Yilgarn Craton;
"Sommerville09_Non_Cratonic" ⇒ Sommerville (2009) Average Non Cratonic model.
"AllenSEA06" ⇒ Allen *et al.* (2006) model for South Eastern Australia
"Liang_2008" ⇒ Liang *et al.* (2006)
"Atkinson06_hard_bedrock" ⇒ Atkinson and Boore (2006) model for hard bedrock (V_{s30}=760 ms^{-1})
"Atkinson06_bc_boundary_bedrock" ⇒ Atkinson and Boore (2006) model for V_{s30} at the NEHRP BC boundary
"Campbell03" ⇒ Campbell (2003) hybrid empirical model
"Abrahamson08" ⇒ Abrahamson *et al.* (2008) NGA model
"Chiou08" ⇒ Chiou and Youngs (2008) NGA model
"Campbell08" ⇒ Campbell and Borzorgnia (2008) NGA model
"Akkar_2010_crustal" ⇒ Akkar and Bommer (2010) model for Mediterranean and Middle East
"Zhao_2006_interface" ⇒ Zhao *et al.* (2006) model for earthquakes on the subduction interface near Japan
"Atkinson_2003_intraslab" ⇒ Zhao *et al.* (2006) model for earthquakes in the subducting slab near Japan
"Atkinson_2003_interface" ⇒ Atkinson and Boore (2003) model for earthquakes on the subduction interface
"Zhao_2006_intraslab" ⇒ Atkinson and Boore (2003) model for earthquakes in the subducting slab

2.4 Site Files

The EQRM requires a site file at which either hazard or loss will be modeled.

2.4.1 Hazard Site File

Filename: `<site_tag>_par_site.csv`

The site file for hazard is a csv file containing a list of points at which the hazard (PSHA simulation) or ground motion (scenario simulation) will be computed. An example is given below in the grey shaded box:

```
LATITUDE,LONGITUDE,SITE_CLASS,VS30
-6.4125,110.879166,D,346
-6.4125,110.887497,D,350
-6.4125,110.895836,D,356
-6.4125,110.904167,C,431
-6.4125,110.912498,C,532
-6.4125,110.92083,C,514
-6.4125,110.929169,C,483
-6.4125,110.962502,D,282
-6.4125,110.970833,D,216
-6.4375,110.904167,B,760
-6.4375,110.912498,B,760
-6.4375,110.92083,B,760
-6.4375,110.929169,B,760
```

Parameters in the hazard site file:

- `Latitude`: Latitude of the points of interest.
- `Longitude`: Longitude of the points of interest.
- `SITE_CLASS`: Regolith site class. Typically, this is defined by a letter. Note that the value of this parameter must match with an amplification factor defined in the amplification file (see Section 2.5)
- `VS30`: Average velocity in the top 30 m (i.e. V_{s30}). This is used to incorporate regolith for GMPEs with a V_{s30} term.

2.4.2 Risk Site File (Building Database)

Filename: `sitedb_<site_tag><site_db_tag>.csv`

The site file for risk is a csv file representing a building portfolio. It contains a list of points at which the risk (PSHA simulation) or loss (scenario simulation) will be computed. An example is given below in the grey shaded box:

```
BID,LATITUDE,LONGITUDE,STRUCTURE_CLASSIFICATION,STRUCTURE_CATEGORY, ...
    ...HAZUS_USAGE,SUBURB,POSTCODE,PRE1989,HAZUS_STRUCTURE_CLASSIFICATION, ...
    ...CONTENTS_COST_DENSITY,BUILDING_COST_DENSITY,FLOOR_AREA,SURVEY_FACTOR, ...
    ...FCB_USAGE,SITE_CLASS,
1,-32.945,151.7513, W1BVTILE, BUILDING, RES1, MEREWETHER,2291,0, W1, ...
    ...344.4451,688.8903,150,9.8,111, C,
2,-32.9442,151.7512, S3, BUILDING, RES3, MEREWETHER,2291,0, S3, ...
    ...430.5564,861.1128,480,1,131, C,
3,-32.9419,151.7495, W1TIMBERMETAL, BUILDING, RES1, MEREWETHER,2291,0, W1, ...
    ...292.7784,585.5567,120,9.8,111, D,
4,-32.9414,151.7492, URMLTILE, BUILDING, RES1, MEREWETHER,2291,0, URML, ...
    ...378.8897,757.7793,80,9.8,111, D,
5,-32.9412,151.7486, W1TIMBERTILE, BUILDING, RES1, MEREWETHER,2291,0, W1, ...
    ...292.7784,585.5567,120,9.8,111, C,
6,-32.9409,151.7498, URMLMETAL, BUILDING, REL1, MEREWETHER,2291,0, URML, ...
    ...925.6963,925.6963,150,1,421, D,
7,-32.9431,151.7558, S3, BUILDING, RES3, MEREWETHER,2291,0, S3,...
    ...430.5564,861.1128,288,1,131, D,
8,-32.9431,151.7549, W1TIMBERMETAL, BUILDING, COM8, MEREWETHER,2291,0, W1,...
    ...1087.155,1087.155,600,1,451, D,
9,-32.9416,151.7545, C3L, BUILDING, RES3, MEREWETHER,2291,0, C3L,...
    ...430.5564,861.1128,720,1,131, E,
10,-32.9386,151.7609, C1LMEAN, BUILDING, COM1, THE JUNCTION,2291,1, C1L,...
    ...548.9594,548.9594,4500,1,211, G,
```

Parameters in the building database:

- `BID`: Integer site identifier for EQRM (typically the same as column 10)
- `LATITUDE`: Latitude of building
- `LONGITUDE`: Longitude of building
- `STRUCTURE_CLASSIFICATION`: Index to building construction type... **expanded** HAZUS list (Section 4.3.1)
- `STRUCTURE_CATEGORY`: Type of structure (e.g. building or bridge)
- `HAZUS_USAGE`: Index to HAZUS usage classification (Section 4.3.2)
- `SUBURB`: within which building is located
- `POSTCODE`: Postcode within which building is located
- `PRE1989`: Logical index stating whether the building is pre- (0) or post- (1) the 1989 Newcastle earthquake
- `HAZUS_STRUCTURE_CLASSIFICATION`: Index to building construction type... HAZUS list (Section 4.3.1)
- `CONTENTS_COST_DENSITY`: Replacement cost of contents in dollars per square meter (Section 4.3.3)

- `BUILDING_COST_DENSITY`: Replacement cost of building in dollars per square meter (Section 4.3.3)
- `FLOOR_AREA`: Total floor area in square meters (summed over all stories)
- `SURVEY_FACTOR`: Survey factor indicating how many 'real' buildings the database entry represents
- `FCB_USAGE`: Index to FCB usage classification (Section 4.3.2)
- `SITE_CLASS`:Regolith site class. Typically, this is defined by a letter. Note that the value of this parameter must match with an amplification factor defined in the amplification file (see Section 2.5)
- `VS30`: **WARNING**: Not yet operational for risk!

Typically the building database used with the EQRM represents a subset of the true portfolio of interest. When creating a database that sub-samples a larger portfolio, individual database entries are used to represent more than one 'real' building. Such sub-sampling is undertaken to reduce run times and memory requirements. Results from an EQRM loss simulation are scaled to the full portfolio using the `survey factor`.

Building construction types

Buildings have been subdivided into a number of building types each with their own set of engineering parameters uniquely defining the median capacity curve and the random variability around the median. The building construction types are based upon the HAZUS definitions (FEMA, 1999),with some further subdivisions recommended by Australian engineers for Australian building construction types (Stehle *et al.*, 2001).

In essence, the seven basic HAZUS types are

- Timber frame (W)
- Steel frame (S)
- Concrete frame (C)
- Pre-cast concrete (PC)
- Reinforced masonry (R)

- Unreinforced masonry (URM)
- Mobile homes (MH)

There are further subdivisions of the HAZUS types into subtypes according to numbers of stories in the building. These are given in Table 4.2.

The new Australian sub-types, developed by Australian engineers, create further subdivisions of the HAZUS types (Stehle *et al.*, 2001). In particular, the timber frame category (W1) is subdivided into wall types (timber or brick veneer walls) and roof types (metal or tiled); the unreinforced masonry types (URML and URMM) into roof type (metal, tile or otherwise), and the concrete frame types are subdivided into soft-story or non-soft story types. Soft-story refers to buildings that may have a concrete basement or parking area but wood frame stories.

In total, we currently have 56 possible construction types although some are rarely used. For example; the original HAZUS W1 is still there, however this is rarely used in favor of the more detailed classification into W1TIMBMETAL, W1BVTILE, etc. A complete list of all the building construction types is given in Table 4.3.

Building usage types

The cost models used by the EQRM require knowledge of the building's use in society. For example the value of a factory's contents will vary from the value of a residents house. Similarly, the cost associated with building a hospital and the cost of building a local shop may differ even if the same materials are used because the buildings may be built to different standards. To transfer this information to the EQRM the building database stores information about each building's usage. There are two different schemes that can be used; the functional classification of building (FCB) usage (ABS, 2001) and the HAZUS usage classification (FEMA, 1999).

The FCB usage is summarised in Table 4.4 and the HAZUS usage classification is summarised in Table 4.5. The EQRM control file parameter `buildings_usage_classification` can be used to switch between the two usage classifications.

Replacement costs

The replacement cost in dollars per square meter for each building and the replacement cost of the contents of each building are contained within the building

Table 2.2: Definitions of the basic HAZUS building construction types.

code	description	Stories
W1	timber frame $< 5\,000$ square feet	(1–2)
W2	timber frame $> 5\,000$ square feet	(All)
S1L	steel moment frame	Low-Rise (1–3)
S1M		Mid-Rise (4–7)
S1H		High-Rise (8+)
S2L	steel light frame	Low-Rise (1–3)
S2M		Mid-Rise (4–7)
S2M		High-Rise (8+)
S3	steel frame + cast concrete shear walls	(All)
S4L	steel frame + unreinforced masonry in-fill walls	Low-Rise (1–3)
S4M		Mid-Rise (4–7)
S4H		High-Rise (8+)
S5L	steel frame + concrete shear walls	Low-Rise (1–3)
S5M		Mid-Rise (4–7)
S5H		High-Rise (8+)
C1L	concrete moment frame	Low-Rise (1–3)
C1M		Mid-Rise (4–7)
C1H		High-Rise (8+)
C2L	concrete shear walls	Low-Rise (1–3)
C2M		Mid-Rise (4–7)
C2H		High-Rise (8+)
C3L	concrete frame + unreinforced masonry in-fill walls	Low-Rise (1–3)
C3M		Mid-Rise (4–7)
C3H		High-Rise (8+)
PC1	pre-cast concrete tilt-up walls	(All)
PC2L	pre-cast concrete frames with concrete shear walls	Low-Rise (1–3)
PC2M		Mid-Rise (4–7)
PC2H		High-Rise (8+)
RM1L	reinforced masonry walls + wood or metal diaphragms	Low-Rise (1–3)
RM1M		Mid-Rise (4+)
RM2L	reinforced masonry walls + pre-cast concrete diaphragms	Low-Rise (1–3)
RM2M		Mid-Rise (4–7)
RM2H		High-Rise (8+)
URML	unreinforced masonry	Low-Rise (1–2)
URMM		Mid-Rise (3+)
MH	Mobile homes	(All)

Table 2.3: Complete list of all building construction types (with those that are rarely used in italics).The integers corresponding to each building construction type represent the integer index used in the building database Column 4 for expanded HAZUS types (column 12 for HAZUS only types).

1: W1	15: S5H	29: RM1L	43: C1LSOFT
2: W2	16: *C1L*	30: RM1M	44: C1LNOSOFT
3: S1L	17: *C1M*	31: RM2L	45: C1MMEAN
4: S1M	18: *C1H*	32: RM2M	46: C1MSOFT
5: S1H	19: C2L	33: RM2H	47: C1MNOSOFT
6: S2L	20: C2M	34: *URML*	48: C1HMEAN
7: S2M	21: C2H	35: *URMM*	49: C1HSOFT
8: S2H	22: C3L	36: MH	50: C1HNOSOFT
9: S3	23: C3M	37: W1MEAN	51: URMLMEAN
10: S4L	24: C3H	38: W1BVTILE	52: URMLTILE
11: S4M	25: PC1	39: W1BVMETAL	53: URMLMETAL
12: S4H	26: PC2L	40: W1TIMBERTILE	54: URMMMEAN
13: S5L	27: PC2M	41: W1TIMBERMETAL	55: URMMTILE
14: S5M	28: PC2H	42: C1MMEAN	56: URMMMETAL

database (see Section 4.3). Typically these costs are a function of the usage classification of the building and are hence also dependent on whether the HAZUS or FCB classification system is used. The EQRM does not cross check how the costings were created. In some instances it may be appropriate to use costings created from one usage classification with the EQRM using the other usage mode (effects cost splits - see below) and in some instance it may not be appropriate to do so. Users are encouraged to familiarise themselves with database metadata to ensure that they are using the EQRM appropriately for their own application.

Table 2.4: Functional classification of building (FCB) (**?**) and integer index used in the building database column 15.

	Residential: Separate, kit and transportable homes
1	111: Separate Houses
2	112: Kit Houses
3	113: Transportable/relocatable homes
	Residential: Semi-detached, row or terrace houses, townhouses
4	121: One storey
5	122: Two or more storeys
	Residential: Flats, units or apartments
6	131: In a one or two storey block
7	132: In a three storey block
8	133: In a four or more storey block
9	134: Attached to a house
	Residential: Other residential buildings
10	191: Residential: not otherwise classified
	Commercial: Retail and wholesale trade building
11	211: Retail and wholesale trade buildings
	Commercial: Transport buildings
12	221: Passenger transport buildings
13	222: Non-passenger transport buildings
14	223: Commercial carparks
15	224: Transport: not otherwise classified
	Commercial: Offices
16	231: Offices
	Commercial: Other commercial buildings
17	291: Commercial: not otherwise classified
	Industrial: Factories and other secondary production buildings
18	311: Factories and other secondary production buildings
	Industrial: Warehouses
19	321: Warehouses (excluding produce storage)
	Industrial: Agricultural and aquacultural buildings
20	331: Agricultural and aquacultural buildings
	Industrial: Other industrial buildings
21	391: Industrial: not otherwise classified
	Other Non-Residential: Education buildings
22	411: Education buildings
	Other Non-Residential: Religion buildings
23	421: Religion buildings
	Other Non-Residential: Aged care buildings
24	431: Aged care buildings
	Other Non-Residential: Health facilities (not in 431)
25	441: Hospitals
26	442: Health: not otherwise classified
	Other Non-Residential: Entertainment and recreation buildings
27	451: Entertainment and recreation buildings
	Other Non-Residential: Short term accommodation buildings
28	461: Self contained, short term apartments
29	462: Hotels (predominately accommodation), motels, boarding houses, hostels or lodges
30	463: Short Term: not otherwise classified
	Other Non-Residential: Other non-residential buildings
31	491: Non-residential:not otherwise classified

Table 2.5: HAZUS building usage classification (**?**) and integer index used in the building database column 5.

	Residential
1	RES1: Single family dwelling (house)
2	RES2: Mobile home
3	RES3: Multi family dwelling (apartment/condominium)
4	RES4: Temporary lodging (hotel/motel)
5	RES5: Institutional dormitory (jails, group housing - military, colleges)
6	RES6: Nursing home
	Commercial
7	COM1: Retail trade (store)
8	COM2: Wholesale trade (warehouse)
9	COM3: Personal and repair services (service station, shop)
10	COM4: Professional and technical services (offices)
11	COM5: Banks
12	COM6: Hospital
13	COM7: Medical office and clinic
14	COM8: Entertainment and recreation (restaurants, bars)
15	COM9: Theaters
16	COM10: Parking (garages)
	Industrial
17	IND1: Heavy (factory)
18	IND2: Light (factory)
19	IND3: Food, drugs and chemicals (factory)
20	IND4: Metals and mineral processing (factory)
21	IND5: High technology (factory)
22	IND6: Construction (office)
	Agriculture
23	AGR1: Agriculture
	Religion/Non/Profit
24	REL1: Church and non-profit
	Government
25	GOV1: General services (office)
26	GOV2: Emergency response (police, fire station, EOC)
	Education
27	EDU1: Grade schools
28	EDU2: Colleges and Universities (not group housing)

2.5 Amplification File

Filename: `<site_tag>_par_ampfactors.xml`

Local soil conditions (or regolith) are capable of amplifying bedrock (or hard rock) ground motion. Consequently, it can be important to incorporate regolith in hazard and/or risk studies. The choice to use regolith is controlled by the EQRM control file parameter `use_amplification`. The manner in which regolith (or amplification) is considered depends on the GMPEs used. If a GMPE explicitly incorporates regolith with a V_{S30} term, then the EQRM will use this. Otherwise the RSA is computed on bedrock and then amplified to the regolith surface using a transfer function (or amplification factor). An example of an input file containing amplification factors is provided in the grey box below.

```
<amplification_model name = "example_par_ampfactors">
  <moment_magnitude_bins>
        4.50000000 5.50000000
  </moment_magnitude_bins>
  <pga_bins>
        0.00000000 0.15290000 0.2548000 0.35680000
  </pga_bins>
  <site_classes>
        CDE
  </site_classes>
  <periods>
        0.00000000 0.40000000 0.50000000 2.00000000
  </periods>

  <site_class class="C">
        <moment_magnitude mag_bin="4.50000000">
          <pga pga_bin="0.00000000">
                <log_amplification
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.00000000">
                                0.13976194 0.13976194 0.25464222 0.25464222
                </log_amplification>
                <log_std
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.00000000">
                                0.01000000 0.01000000 0.01000000 0.01000000
                </log_std>
          </pga>

          <pga pga_bin="0.15290000">
                <log_amplification
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.15290000">
                                0.09531018 0.09531018 0.23111172 0.23111172
                </log_amplification>
                <log_std
```

```
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.15290000">
                                0.01000000 0.01000000 0.01000000 0.01000000
                </log_std>
        </pga>

        <pga pga_bin="0.2548000">
                <log_amplification
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.2548000">
                                0.03922071 0.03922071 0.20701417 0.20701417
                </log_amplification>
                <log_std site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.2548000">
                                0.01000000 0.01000000 0.01000000 0.01000000
                </log_std>
        </pga>
        <pga pga_bin="0.35680000">
                <log_amplification
                        site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.35680000">
                                -0.02020270 -0.02020270 0.17395331 0.17395331
                </log_amplification>
                <log_std site_class = "C"
                        moment_magnitude = "4.50000000"
                        pga_bin = "0.35680000">
                                0.01000000 0.01000000 0.01000000 0.01000000
                </log_std>
        </pga>
      </moment_magnitude>

      <moment_magnitude mag_bin="5.50000000">
              ...
      </moment_magnitude>
  </site_class>

  <site_class class="D">
      <moment_magnitude mag_bin="4.50000000">
              ...
      </moment_magnitude>

      <moment_magnitude mag_bin="5.50000000">

      </moment_magnitude>
  </site_class>

  <site_class class="E">
      <moment_magnitude mag_bin="4.50000000">
              ...
      </moment_magnitude>

      <moment_magnitude mag_bin="5.50000000">
              ...
      </moment_magnitude>
  </site_class>

</amplification_model>
```

The amplification of seismic ground motion depends on the composition of the regolith. The EQRM accounts for variation in regolith material by assigning amplification factors to different site classes. The EQRM also recognises that amplification of seismic waves is a non-linear process. That is, the degree of amplification is a function of the level of ground motion. To account for this non-linearity, the EQRM allows users to specify a number of amplification factors which are grouped according to the level of bedrock ground motion (as measured by PGA) and the size of the event (as measure by M_w).

The amplification factor file must specify the following parameters at the beginning:

- `moment_magnitude_bins`: centroids of the moment magnitude M_w bins for which amplification factors are defined.
- `pga_bins`: centroids of the PGA bins for which amplification factors are defined.
- `site_classes`: List of site classes for which amplification factors are defined. The EQRM assumes that each site class is defined by a single letter (e.g. `site_class` = B).
- `periods`: RSA periods at which the amplification factors are defined. Note that these periods need not be the same as those in `atten_periods` from the EQRM control file. The EQRM will interpolate as required.

The xml amplification factor file is then composed of a sequence of blocks, each of which defines:

- the `site_class` using the parameter `class`.
- the `moment_magnitude` bin centroid using the parameter `mag_bin`.
- the `pga` bin centroid using the parameter `pga_bin`.

Finally the inside of each block specifies:

- `log_amplification`: the logarithm of the median amplification factor defined at each of the RSA periods in `periods`.

- `log_std`: the standard deviation of the amplification factor. The EQRM assumes that the amplification factor is log-normally distributed when using this standard deviation. The standard deviation can be set to an arbitrarily small number such as 0.01 (as shown in grey shaded box above) when not known. Use of this standard deviation is controlled by the EQRM control file parameter `amp_variability_method` which can also be set to `None`.

Chapter 3

Earthquake source generation

3.1 Overview

The EQRM conducts probabilistic seismic hazard analysis (PSHA) and probabilistic seismic risk analysis (PSRA) using an event based approach. This means that the ground motions (hazard) and loss (risk) are computed for each event individually and the results separately aggregated to form probabilistic estimates. The event based approach differs from the traditional approach to PSHA which integrate over all magnitude and distance combinations to attain probability levels for exceeding a particular level of ground motion in a pre-defined period of time. The traditional approach is introduced by **?** and summarised by **?**. A core component of any event based analysis with the EQRM is the generation of a simulated event (or earthquake) catalogue. The generation of the earthquake catalogue relies upon an existing model for the seismicity in the region. Typically the model of seismicity comes from an interpretation of historical earthquakes, geology and neotectonics. The current version of the EQRM application requires a source model that consists of a set of areal source zones (defined by a polygon) with associated Gutenberg-Richter recurrence relationships and the option of including fault sources (defined by a plane) with associated Gutenberg-Richter or characteristic earthquake recurrence relationships **?**.

Areal source zones typically capture the background seismicity in an area and are decribed by the acitivty rate within each source zone through its Gutenberg-Richter a and b values. In event-based PSHA calculations, such as `EQRM`, synthetic ruptures will be generated stochastically throughout the areal source zone. However, a user may have sufficient information on the geometry and earthquake recurrence rates of a known fault and may want to place synthetic ruptures on

the fault by defining a *fault source*. In this instance, instead of placing the synthetic ruptures within the areal source zone, they will be placed on the known fault plane.

In addition, a novel technique has been developed to generate synthetic ruptures within a 3D volume that may represent seismotectonic features such as the intraplate environment of a subducting slab. Such events are often referred to as intraslab earthquakes and have been known to generate larger and devastating earthquakes (e.g. 30 September 2009 Padang earthquake in Indonesia).

This chapter describes the process of creating an earthquake catalogue for probabilistic seismic hazard analysis and for defining an individual scenario event. The chapter is organised as follows. First the method of generating realistic ruptures in source zones and fault zones is discussed in Section . Second, the approach for determining recurrence rates is presented in Section . Lastly, Section describes scenario event generation.

3.2 Creating an earthquake catalogue for probabilistic seismic hazard analysis

This section describes the method to generate a synthetic earthquake catalog of plausible events. The event catalog can be created for events that are defined within an areal source or located on a known fault source. The recurrence relationship for the events can be defined as a Gutenberg-Richter or Characteristic earthquake relationships.

A simulated event is represented by a plane (or rupture) in 3D space that signifies the region where slip has occurred. The geometry of a rupture plane is shown in Figure 3.1. The important parameters of a simulated event are its location, geometry, magnitude and activity (or likelihood of occurrence). The rupture trace is the surface projection of the simulated event along the direction of dip. The position and geometry of the rupture trace is defined by its start (r_s^{lat}, r_s^{lon}) and end (r_e^{lat}, r_e^{lon}) points, its azimuth r_{azi} and its length r_l. The position and geometry of the rupture plane are defined by its width r_w, dip r_{dip} and the position of its centre (or rupture centroid). The rupture centroid is defined in cartesian coordinates (r_x, r_y, r_z) in km using a local coordinate system with origin at the rupture start point (see Figure 3.1). The orientation of the local coordinate system is such that the $x - axis$ is oriented along the rupture trace with positive

direction pointing towards the rupture end point, the $z-axis$ is pointing vertically downwards and the $y-axis$ is oriented along the surface of the 'earth' such that the axes form a right handed triad. The vertical projection of the rupture centroid is also described by its latitude r_c^{lat} and longitude r_c^{lon}. The event magnitude is represented as a moment magnitude and is generated by the process described in Section 3.2.8. The activity (or likelihood of occurrence) of the simulated event is described by the event activity which represents the number of times a given simulated event (conditional on magnitude and position) occurs in one year (see Section 3.2.8).

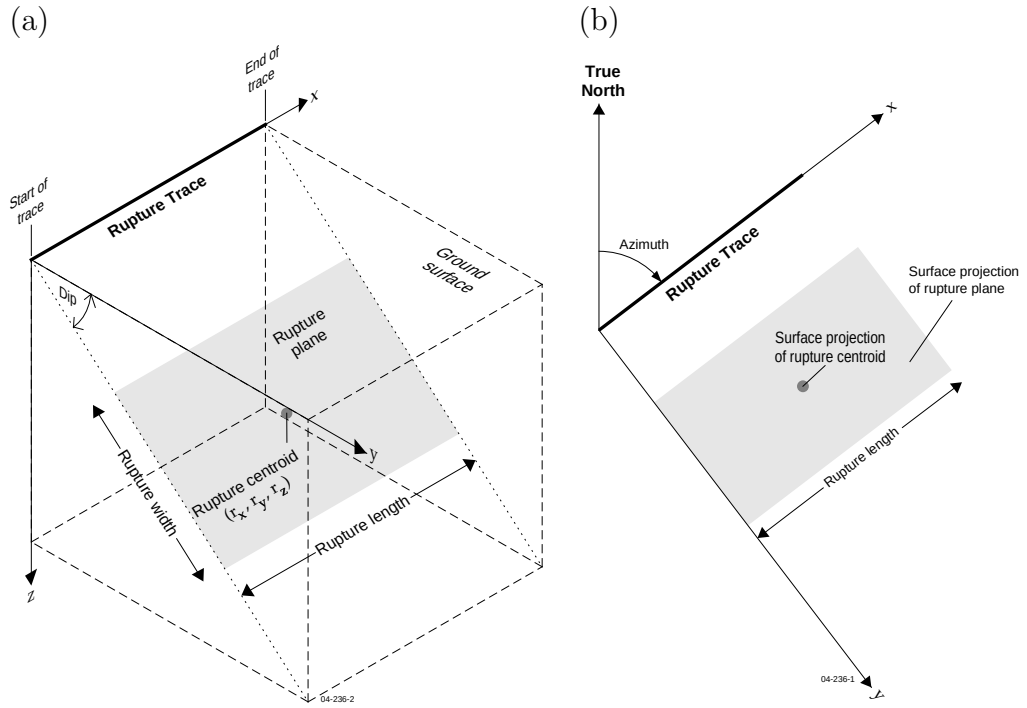


Figure 3.1: Orientation and dimension of the rupture plane in (a) 3D and (b) 2D vertical projection to ground surface.

3.2.1 Simulated events and virtual faults

The terms simulated event, simulated earthquake and simulated rupture are congruent and will be used interchangeably throughout this document. Sometimes the adjective 'simulated' will be omitted for brevity. An adjective 'actual' will be used in place of 'simulated' to refer to a historic earthquake (i.e. one that has actually occurred rather than one that is simulated).

Table 3.1: Columns in `evntdb`. Note that some of the columns are not defined until after `fuse4_hzd` is run (see Section 3.4).

column	description
1	an integer corresponding to the generation source zone from which this event was generated
2	r_s^{lat} - latitude of rupture trace start point (decimal degrees)
3	r_s^{lon} - longitude of rupture trace start point (decimal degrees)
4	r_e^{lat} - latitude of rupture trace end point (decimal degrees)
5	r_e^{lon} - longitude of rupture trace end point (decimal degrees)
6	r_ϕ - azimuth of rupture trace (decimal degrees from True North)
7	r_{dip} - dip of rupture plane ($0 <$ decimal degrees < 90 from the surface)
8	A_{min} - number of earthquakes of magnitude m_{min} or greater per year
9	integer pointer to attenuation model for use with the synthetic earthquake.
10 to 18	place holder (currently not used)
19	w_e - weight derived from event spawning
20	r_ν - event activity (the number of magnitude r_m earthquakes expected per year scaled to the number being simulated)
21	r_m - event magnitude (moment magnitude)
22	r_ε - source epsilon from event spawning
23	r_c^{lat} - latitude of event centroid (decimal degrees)
24	r_c^{lon} - longitude of event centroid (decimal degrees)
25	r_z - event centroid depth z (km from ground surface)
26	r_x - event centroid x (km along rupture trace from the start of trace)
27	r_y - event centroid y (km perpendicular from trace in direction of dip)
28	r_l - length of rupture (km)
29	r_w - width of rupture (km measured along dip)
30	a unique integer for each synthetic earthquake in the generation zone (polygon). Note that the integer count begins again for each generation zone.
31	place holder (currently not used)

A virtual fault refers to a plane in 3D space upon which an event can occur. The virtual fault can be located in within an areal source or defined by a known fault. The EQRM works by first creating a virtual fault, either within an areal source zone, or by a user defined know fault. The rupture is then placed on the virtual fault (the rupture is not allowed to exceed the bounds of the virtual fault). The introduction of a virtual fault is the mechanism by which the EQRM application constrains the location and extent of each rupture. The depth to the top of a virtual fault is defined as the depth to the seismogenic region f_z (see Table **??**). Other geometrical parameters of virtual faults include the width f_w and length f_l.

3.2.2 Location of synthetic earthquakes in areal sources

The location of each event is assigned randomly within an areal source zone and on a fault source. This assignment is done in such a way that the event has an equal probability of being located anywhere in the 'generation polygon' or on the fault.

The first phase of locating the events involves positioning the centre point of each rupture trace (r_c^{lat},r_c^{lon}). A rectangular box is generated over the top of the source zone polygon. The box is bounded by the minimum and maximum latitude and the minimum and maximum longitude of the source zone itself.

The rupture centroid of each event is then randomly placed within the box using a discrete uniform probability density function (Figure 3.2b). The event is the checked to identify whether it falls within the source polygon or not. Events that fall within the source polygon are kept and those that lie outside it are discarded.

The second phase of locating events involves positioning the start and end point of the rupture trace. Assume for the moment that both the length and azimuth of the rupture trace are already known (see Section 3.2.9 and Section 3.2.10 respectively for a description of how these are computed). It is then a trivial process to compute the start and end of the rupture trace ($r_s^{lat}, r_s^{lon} and r_e^{lat}, r_e^{lon}$). It is important to note that the EQRM will allow the start or end points of the rupture trace to lie outside of the source polygon providing the rupture centroid is located within the polygon. The source zone polygon is only used to constrain the rupture centroid (r_c^{lat}, r_c^{lon}).

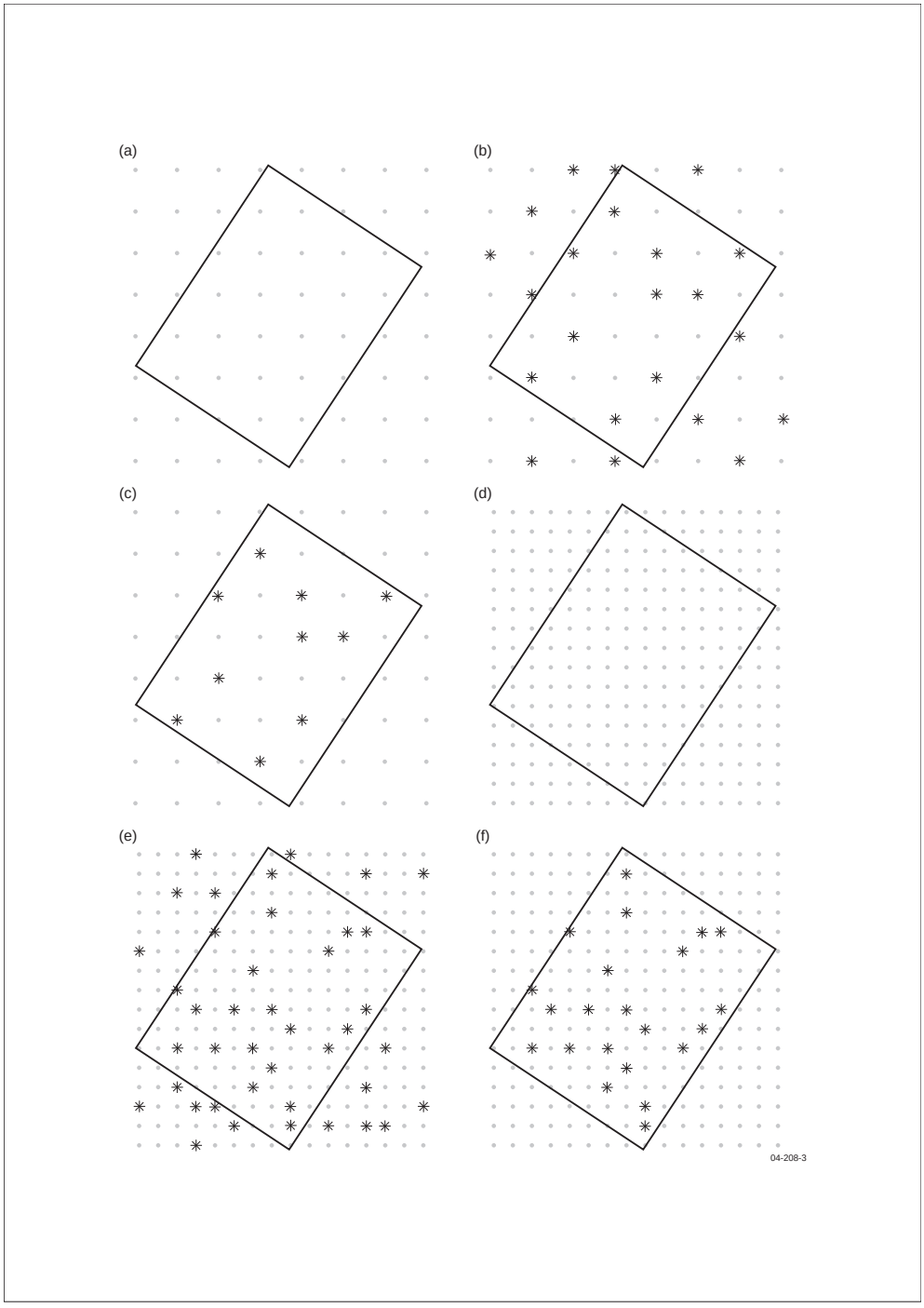


Figure 3.2: Random selection of rupture centroids for a rectangular 'generation polygon'. (a) to (c) illustrate the first iteration and (d) to (f) the second iteration after contraction of the grid spacing.

3.2.3 Location of events on fault sources

As discussed earlier in this chapter, fault sources in the `EQRM` can represent three different styles of faulting:

- Crustal faults
- Subduction interface faults
- Intraslab faults in the subducting slab

The method for placing synthetic ruptures on crustal faults and subduction interface faults is the same, but intraslab faults required a different approach.

Crustal Faults and Subduction Interface Faults

Assuming again that we have already defined the length and width of the rupture plane, we then solve to find the location of the rupture trace along the predefined fault trace by:

$$D_s = (f_l - r_l) \times \mathtt{X} \tag{3.1}$$

where `X` is a random variable between 0→1. Following this step the new start and end of the rupture trace can be obtained using the rupture length, azimuth of the fault.

We now use D_s from equation 3.1, and the precalculated rupture length and fault azimuth to calculate the new longitude and latitude of the start and end of the rupture trace (**??**).

We now follow a similar process (to equation **??**) to determine the depth of the rupture centroid. Since we already have the fault width we can randomly assign the depth of the rupture centre while making sure the rupture plane does not extend beyond the top and bottom of the fault plane defined in the source file (**??**).

We first find the depth range which the rupture centroid can lie within $[r_z^{min}, r_z^{max}]$ such that the rupture plane does not extend above or below the bounds of the seismogenic zone $[f_z^{top}, f_z^{bot}]$:

$$r_z^{min} = \text{minimum depth of rupture centroid} \tag{3.2a}$$
$$r_z^{max} = \text{maximum depth of rupture centroid} \tag{3.2b}$$
$$r_z^{min} = r_z^{top} + 0.5 r_w \times sin(r_\theta \times \text{rad}) \tag{3.2c}$$
$$r_z^{max} = r_z^{bot} - 0.5 r_w \times sin(r_\theta \times \text{rad}) \tag{3.2d}$$

where $f_\theta = r_\theta$. Note that for intrapate ruptures which will be discussed later $f_\theta \neq r_\theta$. We can now randomly assign the rupture centroid depth using a uniform distribution between r_z^{min} and r_z^{max} using the `numpy` package in `python`:

$$r_z = (r_z^{max} - r_z^{min}) \times \texttt{X} + r_z^{min} \tag{3.3}$$

where `X` is a random varaible from 0→1. The local x and y coordinates (r_x and r_y) of the rupture centroid can be calculated by the employing the same equations as for the source zones:

$$r_y = r_z \frac{cos(f_\theta \times rad)}{sin(f_\theta \times rad)} \tag{3.4}$$

$$r_x = \frac{r_l}{2} \tag{3.5}$$

Where the rupture centroid location in local coordinates (r_x, r_y) can be converted to longitude and latitude (r_c^{lon}, r_c^{lat}).

Intraslab Faults

The concept of creating realistic ruptures that represent events in a subducting slab is achieved by considering these events as out of plane ruptures. The subducting slab is defined as a 3D plane from which rupture planes extend out from at angles defined by the user (FIGURE PLOT). For this functionality and out of dip value must be defined (δ_θ). This parameter defines the angle between the dipping fault and the rupture plane (figure 3.3). If the rupture plane is parallel to the fault plane this value will be zero. If the rupture plane is it an angle to the dipping plane then it will be non-zero. When handling a non-zero δ_θ we also need to consider another parameter, Δ_θ, which defines the range of dips we will sample across to obtain the rupture dip (r_θ, figure 3.3). We first take a uniform

sample from the range $[\delta_\theta - \Delta_\theta, \delta_\theta + \Delta_\theta]$ to determine the out of plane angle (ω) and the dip of the rupture plane (r_θ) :

$$\omega = 2\Delta_\theta \times \texttt{numpy.random.random\_sample}() + \delta_\theta - \Delta_\theta \tag{3.6}$$

$$r_\theta = \omega + f_\theta \tag{3.7}$$

When dealing with out of plane ruptures we need to consider two cases:

- $0 \leq r_\theta \leq 90$:
  In this instance the rupture plane will dip in the same direction as the fault plane and we need to project the new rupture plane to the surface to define the surface trace of the rupture.
- $90 < r_\theta \leq 180$:
  In this case the rupture plane dips in the opposite direction to the fault plane and the local coordinate system must be transformed because the plane is no longer dipping to the right handside of the trace when looking from the start to the end of the trace. We must project the new rupture plane to the surface and then redefine the start and end locations of the rupture trace. This is because we always work in a right-handed local coordinate system with the rupture plane dipping to the right of the trace when looking along the trace from the start location to the end locatio (figure 3.3).
- $180 < r_\theta \leq f_\theta + \omega$:
  In this case the rupture plane dips in the same direction of the fault plane, but the rupture trace is projected in the negative y direction of the local coordinate system.

3.2.4 Case if $0 \leq r_\theta \leq 90$

We follow a similar approach to that in section **??** where we first define the rupture plane dimensions (area r_A, width r_w and length r_l). We force the rupture dimensions to fall within the length of the fault, but this time we must also not allow it to extend outside of the width of the dipping slab (S_w) defined in the source file. As shown in figure 3.4 we can find the maximum width of the rupture plane (r_w^{max}) from:

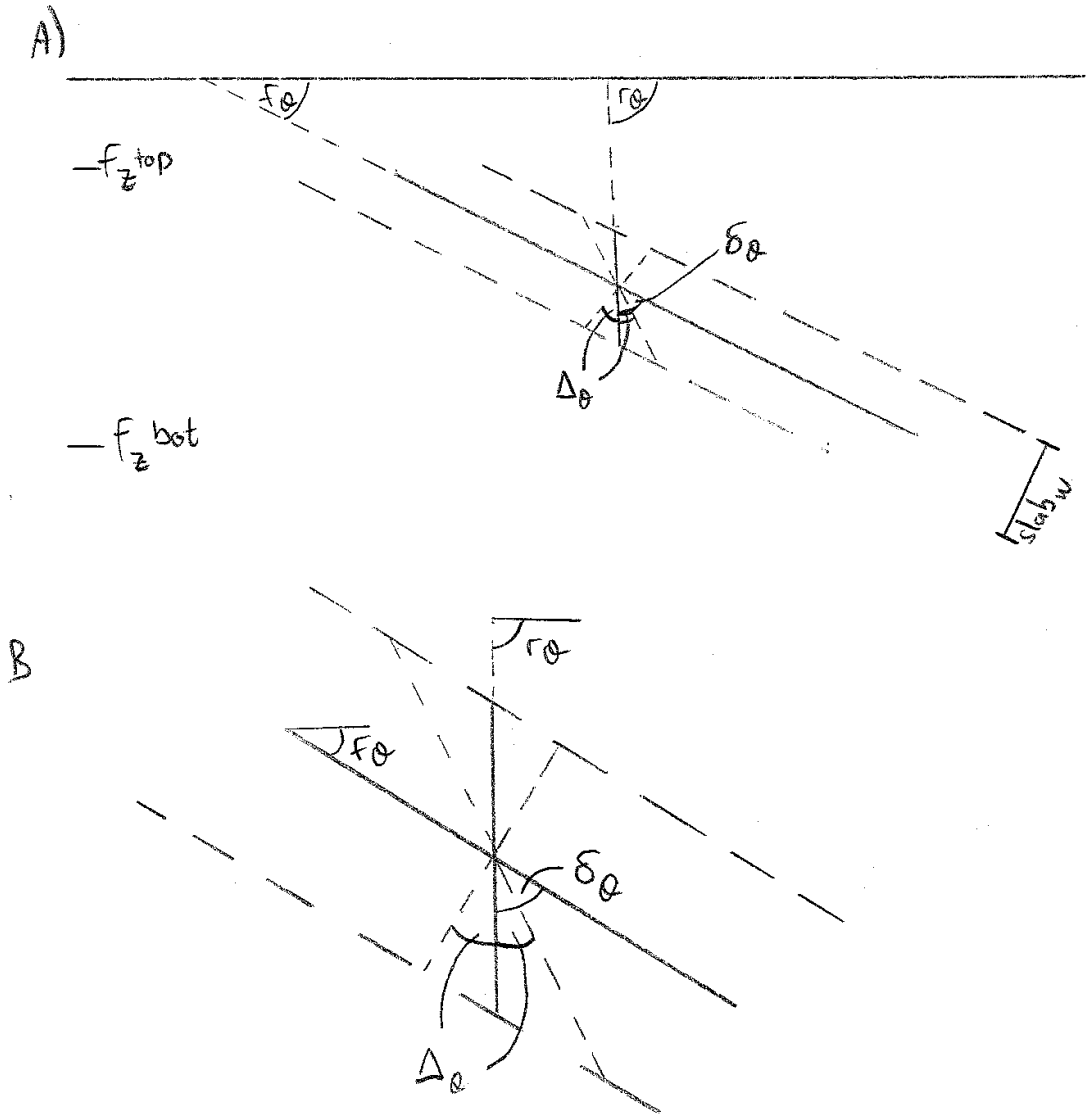


Figure 3.3: A schematic diagram of the geometry of out of plane ruptures. Panel A shows the rupture plane dipping out of the fault plane that defines the dipping slab. Panel B is a zoom in of the angles that define the rupture plane geometry. The user defines the fault dip (f_θ), which in this case represents the direction of the downgoing slab, and the out of plane dip (δ_θ) and a sampling range (Δ_θ) over which the EQRM will uniformly distribute the rupture plane.

$$r_w^{max} = \begin{cases} \frac{S_w}{sin(\omega \times rad)} & \text{if } \omega < 90 \\ S_w & \text{if } \omega = 90 \\ \frac{S_w}{sin((180-\omega) \times rad)} & \text{if } \omega > 90 \end{cases} \tag{3.8}$$

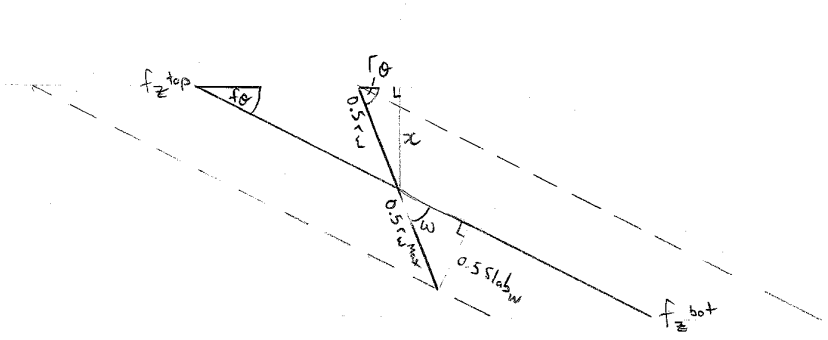


Figure 3.4: The geometry of determining the maximum rupture width (r_w^{max}) shown in the lower trigonometric geometry and the r_z^{min} and r_z^{max} shown in the upper part and equation 3.13.

We use equations **??** and 5.9 to calculate the rupture area (r_A) and rupture width (r_w^{WC}) from the Wells and Coppersmith, 1994 scaling laws. But for out of plane ruptures we must limit the rupture width (r_w) not to the fault width as is the case of on plane ruptures, but to not extend beyond the slab boundary. Therefore we use a slightly modified version of equation **??** to solve for the rupture width:

$$r_w = \mathbf{min}\{r_w^{WC}, r_w^{max}\} \tag{3.9}$$

now we can calculate the rupture length (r_l), making sure it does not extend beyond the fault length (f_l) from:

$$r_l = \mathbf{min}\left\{\frac{r_A}{r_w}, f_l\right\} \tag{3.10}$$

We use equation **??** to calculate D_s, the distance (km) along the fault trace for the start of rupture trace and equations **??** and **??** to find the position of the start and end of the rupture trace. Note that we use new notation to describe the location of the start and end of the rupture trace along the fault trace ($\hat{r}_e^{lon}$, $\hat{r}_e^{lat}$, $\hat{r}_e^{lon}$, $\hat{r}_e^{lat}$) this is because soon we will need to define the new location of the rupture trace (r_s^{lon}, r_s^{lat} and r_e^{lon}, r_e^{lat}):

$$\hat{r}_s^{lon}, \hat{r}_s^{lat} = \texttt{get\_new\_ll}(f_s^{lon}, f_s^{lat}, f_\phi, D_s) \tag{3.11}$$

$$\hat{r}_e^{lon}, \hat{r}_e^{lat} = \texttt{get\_new\_ll}(\hat{r}_s^{lon}, \hat{r}_s^{lat}, f_\phi, r_l) \tag{3.12}$$

Next we need to redefine the depth range of the random depth allocation similar to that of equation 3.2 to make sure the rupture trace does not go above or below the seismogenic zone of the fault. This is slightly different to equation 3.2 because the rupture plane is now dipping out of the fault plane (figure 3.4).

$$r_z^{min} = \text{minimum depth of rupture centroid} \tag{3.13a}$$
$$r_z^{max} = \text{maximum depth of rupture centroid} \tag{3.13b}$$
$$r_z^{min} = f_z^{top} + 0.5 r_w \times sin(r_\theta \times \text{rad}) \tag{3.13c}$$
$$r_z^{max} = f_z^{bot} - 0.5 r_w \times sin(r_\theta \times \text{rad}) \tag{3.13d}$$

Using equation 3.3 we can now randomly assign the rupture centroid depth using a uniform distribution between r_z^{min} and r_z^{max} using the `numpy` package in `python`:

$$r_z = (r_z^{max} - r_z^{min}) \times \texttt{numpy.random.random\_sample}() + r_z^{min} \tag{3.14}$$

We now need to project the rupture trace to the surface to obtain the surface trace, and then the rupture centroid. We begin by finding the location of the rupture centroid referenced to the start position of the original fault plane:

$$\hat{r}_y = r_z \frac{cos(f_\theta \times rad)}{sin(f_\theta \times rad)} \tag{3.15}$$

and then:

$$\hat{r}_x = \frac{r_l}{2} \tag{3.16}$$

We now need to determine the start and end position of the surface trace for the rupture plane. We do this by first finding the location referenced to the start location of the rupture (r_s^{lon}, r_s^{lat}) along the fault plane (equations **??** and **??**) then convert to latitude and longitude.

$$\hat{r}_s^x = 0 \tag{3.17}$$

and now we substitute the dip of the rupture plane (r_θ) into equation 3.15

$$\hat{r}_s^y = \hat{r}_y - r_z \frac{cos(r_\theta \times rad)}{sin(r_\theta \times rad)} \tag{3.18}$$

$$\hat{r}_e^x = r_l \tag{3.19}$$

$$\hat{r}_e^y = \hat{r}_s^y \tag{3.20}$$

Now we can calculate the longitude and latitude for the start and end of the rupture trace.

Now all that is left is to redefine the location of the rupture centroid in local x and y coordinates relative to the origin (r_s^{lon},r_s^{lat}) of the new local coordinate system:

$$r_x = 0.5 r_l \tag{3.21}$$

$$r_y = r_z \frac{cos(r_\theta \times rad)}{sin(r_\theta \times rad)} \tag{3.22}$$

and then transform this into longitude and latitude:

$$r_c^{lon}, r_c^{lat} = \texttt{azimuthal\_orthographic\_xy\_to\_ll}(r_x, r_y, r_s^{lon}, r_s^{lat}, f_\phi, \mathrm{R} = 6367.0). \tag{3.23}$$

Note that $r_z = \hat{r}_z$.

3.2.5 Case if $90 < r_\theta \leq 180$

In this instance the rupture plane is dipping in the opposite direction to the fault plane. Recall that if we look along the rupture trace the plane always dips to the right hand side, therefore in the case here of rupture plane the dips opposite to the project of the surface trace we need to re-project the rupture plane to the surface then swap the start and end coordinates.

After calculating r_w^{max} from equation 3.8 we then must obtain $r_\theta$such that it lies between 5 and 90:

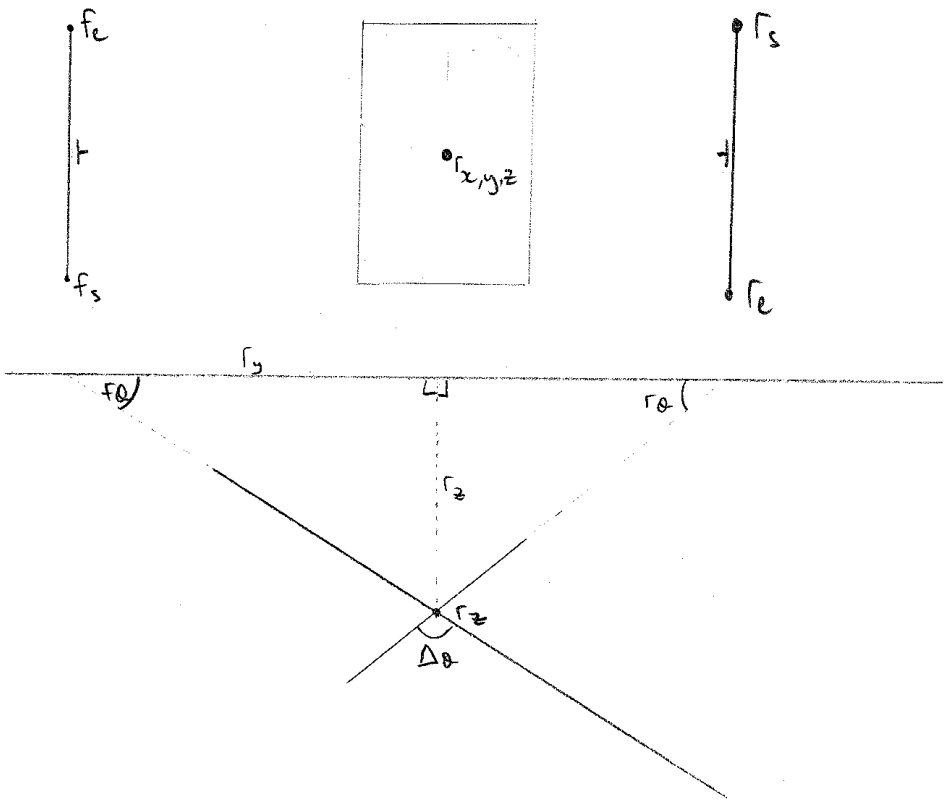


Figure 3.5: Fault and rupture plane geometry for. Lower panel shows cross section view and upper panel shows a plan view of the surface projection. Note that in this case we need to re-project the rupture plane to the surface and redefine the start and end of the rupture trace such that the plane dips to the right hand side of the fault trace when looking from the start to end of the trace.

$$r_\theta = \mathbf{max}\{180 - f_\theta + \omega, 5\}. \tag{3.24}$$

We can follow the same approach as section 3.2.4 from equations 3.9 to 3.17. We need to slightly modify the next steps as we need to add instead of subtract r_y to the right handside of the equation:

$$\hat{r}_s^y = \hat{r}_y + r_z \frac{cos(r_\theta * rad)}{sin(r_\theta * rad)} \tag{3.25}$$

And now we can again follow the method in section 3.2.4 from equations 3.19 to

3.20. But for equations **??** and **??** we transform the coordinate system so the rupture start and end locations are swapped:

$$r_e^{lon}, r_e^{lat} = \texttt{azimuthal\_orthographic\_xy\_to\_ll}(\hat{r}_s^x, \hat{r}_s^y, \hat{r}_s^{lon}, \hat{r}_s^{lat}, f_\phi, \mathrm{R} = 6367.0) \tag{3.26}$$

$$r_s^{lon}, r_s^{lat} = \texttt{azimuthal\_orthographic\_xy\_to\_ll}(\hat{r}_e^x, \hat{r}_e^y, \hat{r}_s^{lon}, \hat{r}_s^{lat}, f_\phi, \mathrm{R} = 6367.0) \tag{3.27}$$

Now we redefine the location of the rupture centroid in local x and y coordinates relative to the new origin (r_s^{lon},r_s^{lat}) of the new local coordinate system and then get the longitude and latitude of this point using equations 3.21 to 3.23.

3.2.6 Case if $180 < r_\theta \leq f_\theta + \omega$

If the initial r_θ value is greater than 180 but less than $f_\theta + \omega$ then the rupture plane dips in the same direction of the fault plane but the rupture trace lies behind the fault trace (figure 3.6).

Here we first obtain the dip of the rupture plane and force the dip to be greater than 5° so that the rupture trace projects to the surface:

$$r_\theta = \mathbf{max}\{180 - f_\theta + \omega, 5\}. \tag{3.28}$$

We follow the same methods as above up to equation 3.17 but now we change the following equation to solve for ($\hat{r}_y$) which in this case will be negative:

$$\hat{r}_s^y = \hat{r}_y - r_z \frac{cos(r_\theta \times rad)}{sin(r_\theta \times rad)}. \tag{3.29}$$

The end of the rupture trace in local coordinates ($\hat{r}_e^x$ and $\hat{r}_e^y$) referenced to the fault trace can be found from:

$$\hat{r}_e^x = r_l \tag{3.30}$$

$$\hat{r}_e^y = \hat{r}_s^y \tag{3.31}$$

Now we can follow equations **??** to 3.23 to get the remaining parameters.

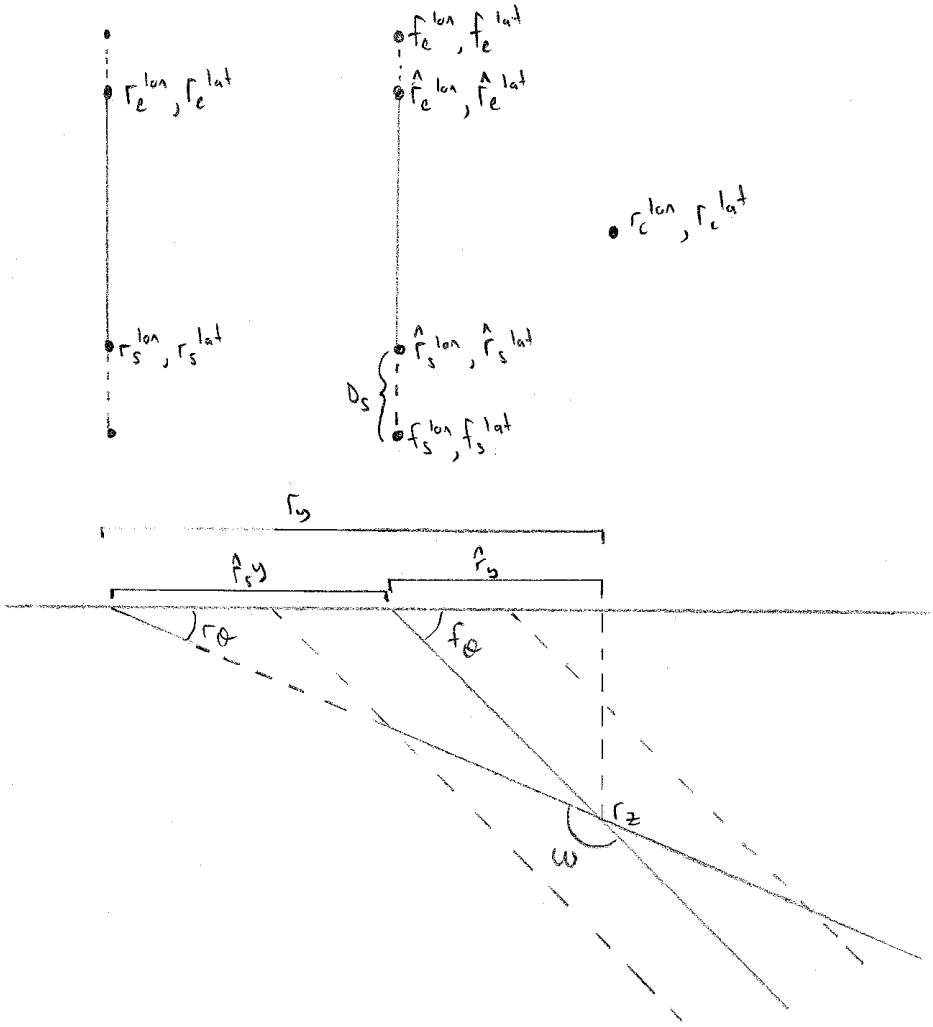


Figure 3.6: Fault and rupture plane geometry for. Lower panel shows cross section view and upper panel shows a plan view of the surface projection. Note that in this case we need to re-project the rupture plane to the surface and redefine the start and end of the rupture trace.

3.2.7 Overlapping source zones

The concept of overlapping zones in a source model is useful for accommodating background seismicity. In principle, the code can accommodate overlapping seismic zones, i.e. where two seismic zones share a common geographical region. However, it should be noted that earthquakes will be distributed randomly in the common region for each zone that exists. That is, a simulation using overlapping source zones (such as that shown in Figure 3.7(a)) may lead to a greater number of simulated earthquakes in the common region than may be warranted by the source model. This only poses a problem if the recurrence relationship (i.e. the a, b parameters) being used has not been modified to account for the 'double counting' in the simulated catalogue (generally this will not have been done). To overcome this problem, it is advisable to define the source zones in such a way that there are no overlapping regions. Figure 3.7 demonstrates two different techniques that can be used to incorporate overlapping source zones in the EQRM. Both of the illustrated techniques are based on creating a doughnut.

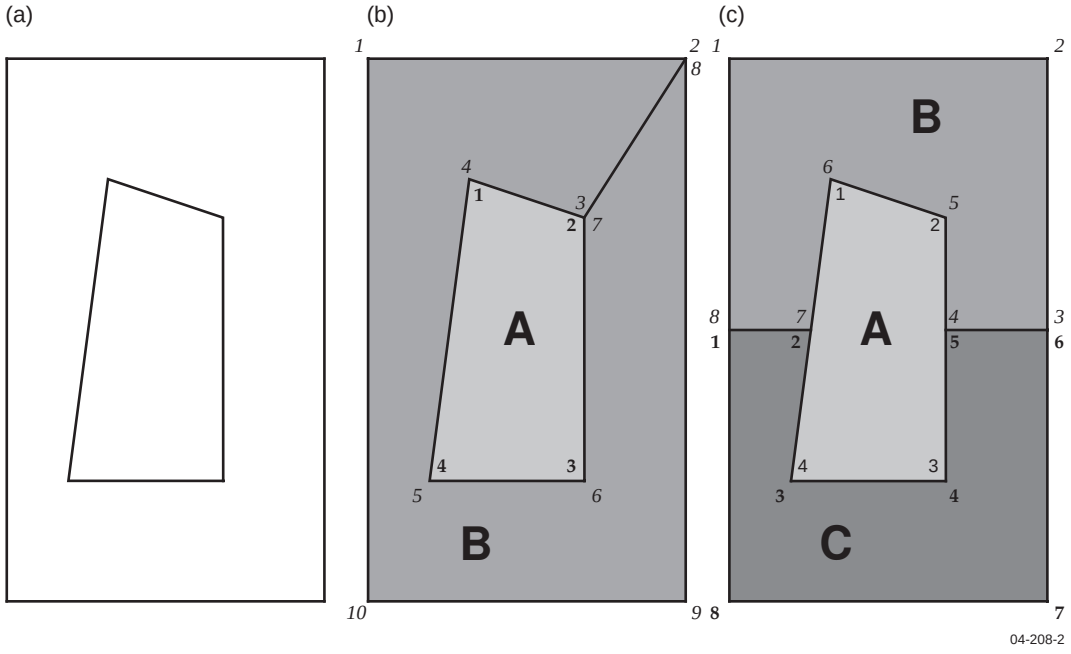


Figure 3.7: Overlapping source zones shown in (a) can be incorporated into the EQRM by cutting out a doughnut (b) or splitting the outer polygon (c). The numbers in different fonts illustrate how the polygon vertices could be listed in the `<site_loc>_par_sourcezones.txt` file.

3.2.8 Magnitude selection and event activity

A 'stratified' Monte-Carlo technique is used to assign the event magnitudes. The stratified nature of the technique ensures that the full range of magnitudes is adequately sampled. The stratified Monte-Carlo technique is illustrated in Figure 3.8 for the general case. This approach is distinctly different to a brute force Monte-Carlo technique that would preferentially sample the more probable lower magnitude events. Such an approach would require the sampling of a large number of small events to ensure that a handful of large events are sampled.

The algorithm described in this section is the algorithm applied when a single source model is used and the 'generation polygons' are the source zones of the source model. Some minor modifications are required when multiple source models are used or the 'generation polygons' differ from the source zones. These modifications are described in Section 3.3.

The Probability Density Function (PDF) for the event magnitudes is based on the bounded Gutenberg-Richter law (**?**). The EQRM application simulates a number N_s of earthquake events. Considering the i^{th} source zone, the algorithm for choosing the magnitude of each event can be summarised as follows:

1. Bound the domain of the PDF with m_{min} and m_{max} (Figure 3.8a).
2. Separate the interval $[m_{min}, m_{max}]$ into N_b bins, and return the bin centroids (Figure 3.8b).
3. For each of the $N_{s,i}$ events randomly select a bin from a discrete uniform distribution, where $N_{s,i}$ refers to the number of events in the i^{th} source zone. This effectively leads to $N_{s,i}/N_b$ earthquakes in each magnitude bin and ensures that the full range of magnitudes is adequately sampled (Figure 3.8c).
4. For each of the $N_{s,i}$ events randomly select a magnitude denoted r_m, from a continuous uniform distribution that spans the complete range of magnitudes in the bin. Note that Step 3 ensures that the entire range of magnitudes is adequately sampled whereas this step ensures that all magnitudes can be attained (Figure 3.8d).
5. For each of the $N_{s,i}$ events calculate the event activity $r_\nu(r_m^*)$, where r_m^* refers to the bin centroid for the event in question and

$$r_\nu(r_m^*) = \frac{N_b}{N_{s,i}} \times \lambda(\texttt{min\_mag\_cutoff}) \times P_{GR}(r_m^* - \Delta < R_m < r_m^* + \Delta). \quad (3.32)$$

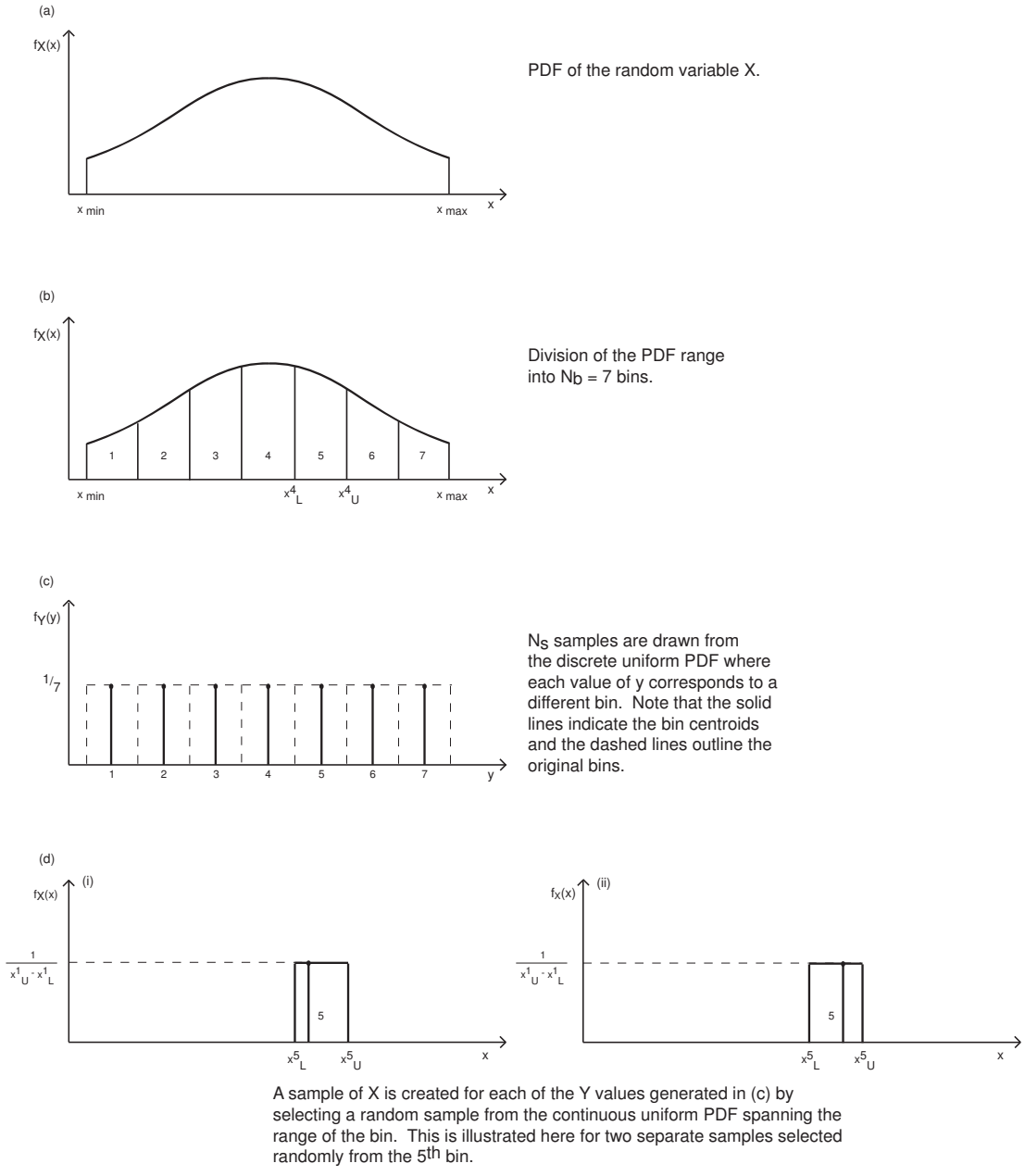


Figure 3.8: Stratified Monte Carlo technique for sampling a PDF. In this case a normal distribution is shown but any PDF can be used.

The first term in Equation 3.32, $\frac{N_b}{N_{s,i}}$, approximates the reciprocal of the number of synthetic events in the bin. This term ensures that the event activity for each synthetic event scales as the number of generated events is modified and hence 'brings the synthetic results back to the real world'. The second term in Equation 3.32, λ(`min_mag_cutoff`), represents the number of earthquakes with magnitude greater than or equal to `min_mag_cutoff` and is computed by evaluating the bounded Gutenberg-Richter recurrence relation

$$\lambda(m) = A_{min} \frac{e^{-\beta(m-m_{min})} - e^{-\beta(m_{max}-m_{min})}}{1 - e^{-\beta(m_{max}-m_{min})}}, \tag{3.33}$$

(**?**). The final term in Equation 3.32 represents the actual probability that a real event will fall into the r_m^* bin in a given year and is computed by evaluating

$$P_{GR}(r_m^* - \Delta < R_m < r_m^* + \Delta) = \frac{f_M(r_m^*)}{\sum_{j=1}^{N_b} f_M(r_{m,j}^*)} \tag{3.34}$$

(Appendix **??**), where

$$f_M(r_m^*) = \frac{\beta e^{-\beta(r_m^* - m_{min})}}{1 - e^{-\beta(m_{max}-m_{min})}}, \tag{3.35}$$

$$\beta = b \ln(10) \tag{3.36}$$

(**?**) and Δ is the half width of the bins. Note that r_ν represents the frequency of occurence in terms of a number per year and can be thought of as the synthetic version of λ for the simulated event catalogue.

Figures 3.9 to 3.11(a) illustrate histograms of the event magnitudes simulated for three different source zone combinations. Note that the fifteen bars in Figures 3.9 and 3.10(a) correspond to the fifteen different magnitude bins. The event activity r_v for the same source zone combinations are shown in Figures 3.9 to 3.11(c).

3.2.9 Dimensions and position of the rupture plane

The width and length of the rupture and position of the rupture centroid are computed by `mag2rup_v` using empirical rules based on the moment magnitude, r_m of the event (Mendez, 2002 [*pers. comm.*]) . The location of the rupture

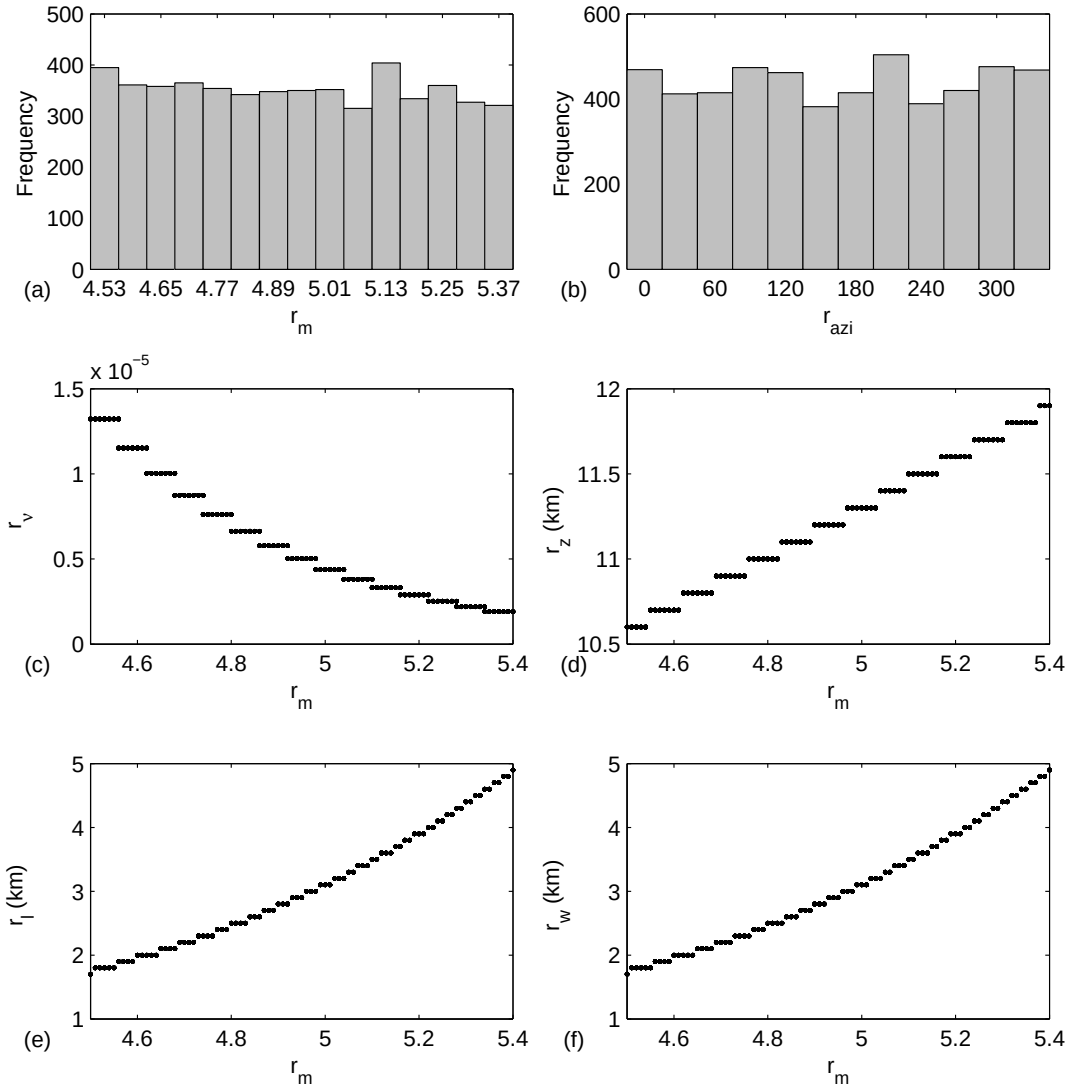


Figure 3.9: Simulated events for the Newcastle Triangle Zone (see **?**). Plots (a) and (b) show histograms of the rupture magnitude r_m and rupture azimuth r_{azi} respectively. Plots (c), (d), (e) and (f) illustrate the relationship between the event magnitude r_m and the event activity r_ν, the depth to the rupture centroid r_z, the length of the rupture r_l and the width of the rupture r_w respectively. The desired number of events $ntrgvector(1)$ was 5000.

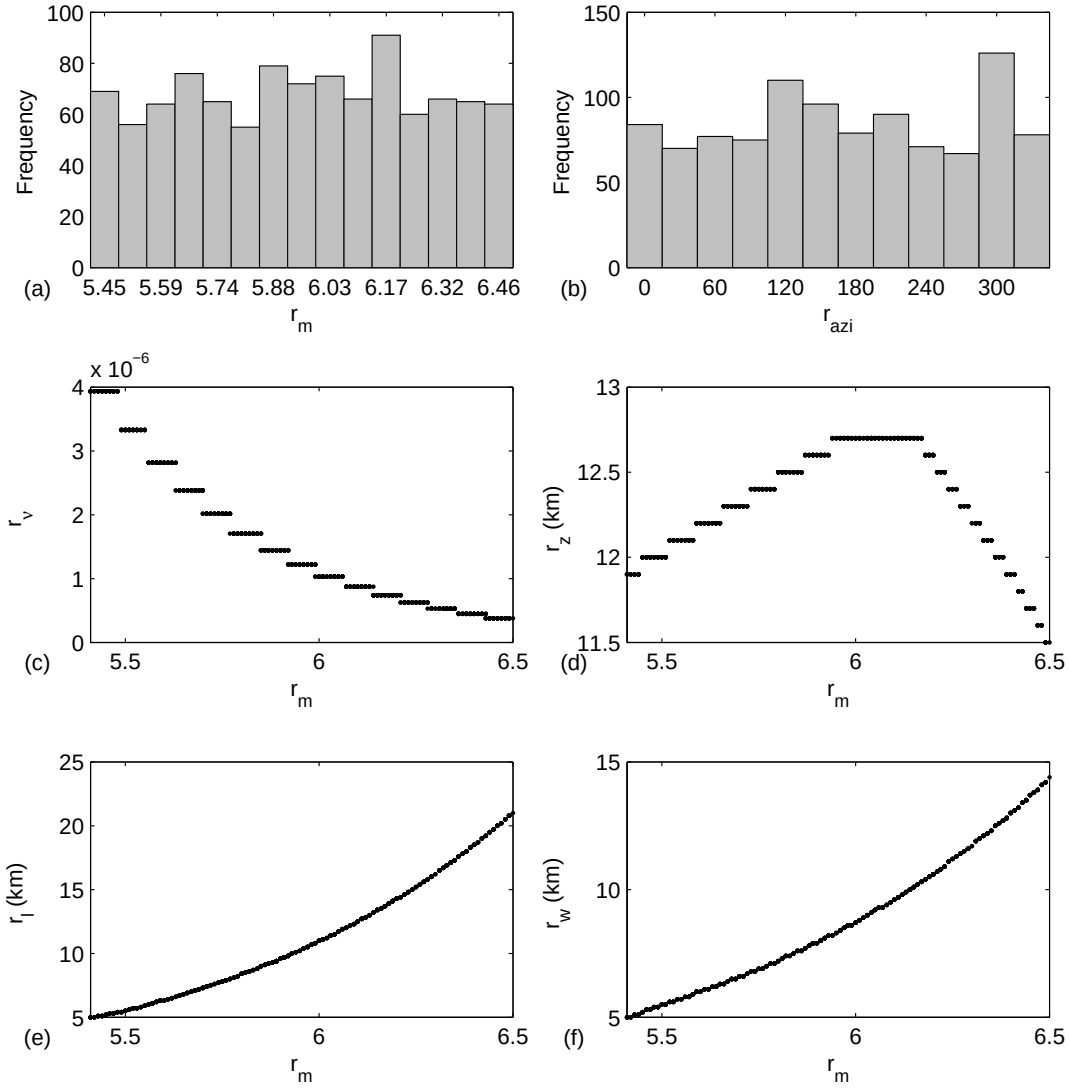


Figure 3.10: Simulated events for the Newcastle Rectangle Zone (see **?**). Parts (a) to (f) are the same as those described in Figure 3.9. The desired number of events $ntrgvector(4)$ was 3000.

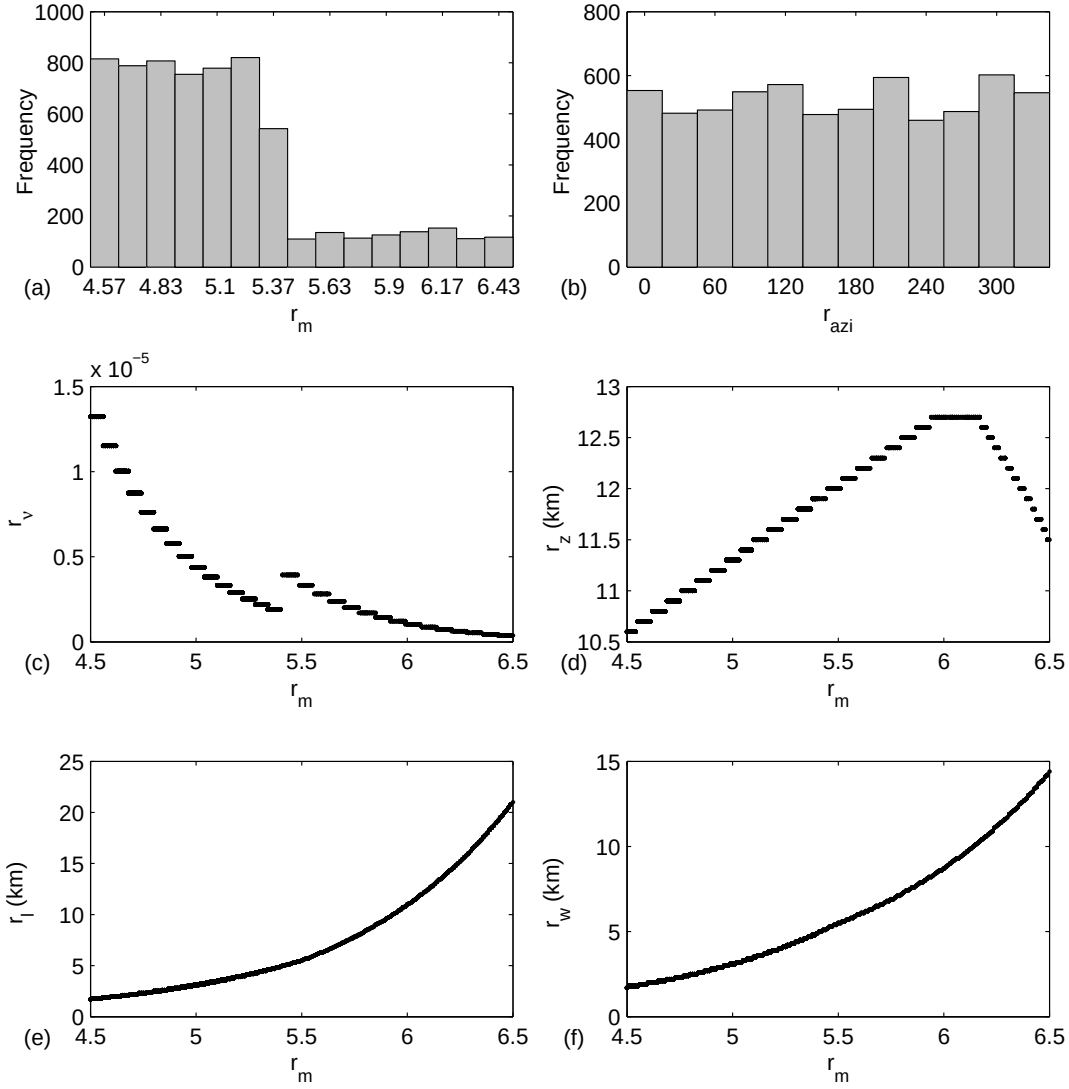


Figure 3.11: Simulated events when the Newcastle Triangle Zone and Newcastle Fault Zone are considered together (see **?**). Parts (a) to (f) are the same as those described in Figure 3.9. The desired number of events was defined by $ntrgvector = [5000, 3000]$ where the source zones 1 and 2 refer to the the Newcastle Triangle Zone (NTZ) and the Newcastle Fault Zone respectively.

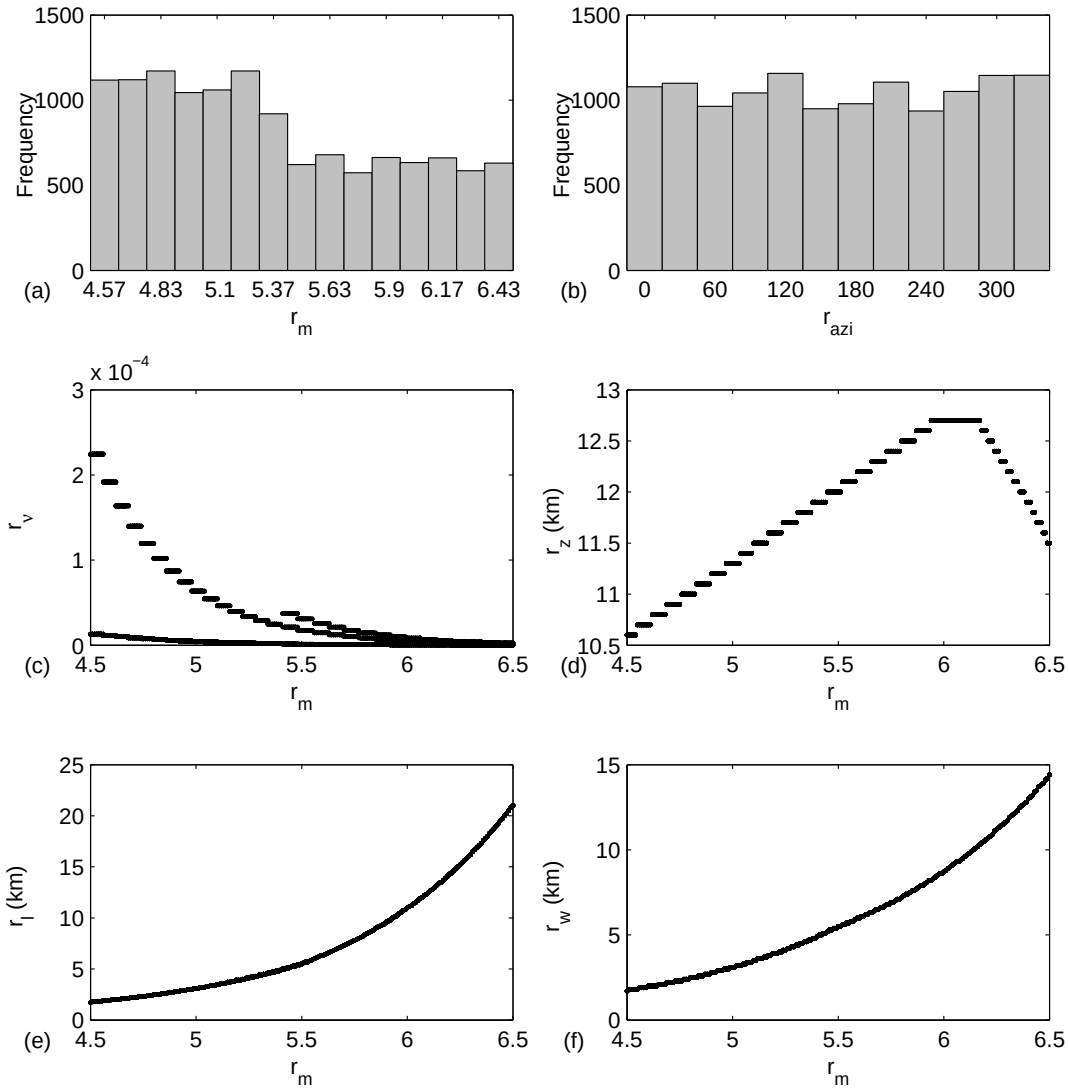


Figure 3.12: Simulated events for all six of the source zones used in the Newcastle and Lake Macquarie study (see **?**). Parts (a) to (f) are the same as those described in Figure 3.9. The desired number of events was defined by $ntrgvector = [5000, 1000, 1000, 3000, 1000, 1000]$ where the source zones 1 to 6 refer to the the Newcastle Triangle Zone (NTZ) the Tasman Sea Margin Zone Subset 1 (TSMZ1), the TSMZ2, the Newcastle Fault Zone, the TSMZ3 and the TSMZ4 respectively.

plane is constrained to lie within a virtual fault which is described by the depth to its top (equivalent to the depth to the seismogenic region f_z, see Section **??**), its length f_l, its width f_w and its dip r_{dip}. Note that by default the dip of the rupture plane is defined to be the dip of the virtual fault. Firstly, the area of the rupture plane A_r is defined as

$$A_r = 10^{r_m - 4.02} \tag{3.37}$$

which represents a subtle change from that defined by **?**. Empirical relationships for the width r_w and length l_r of the rupture plane were defined by Mendez (2002 [*pers. comm.*]). The width is defined in a two step process as follows

$$f_1 = \begin{cases} 1 & \text{if } r_m \leq 5.5 \\ \frac{1}{\sqrt[4]{1+2(r_m - 5.5)\sin(r_{dip})}} & \text{if } r_m > 5.5 \end{cases} \tag{3.38}$$

and

$$r_w = \min\{f_1\sqrt{A_r}, f_w\}. \tag{3.39}$$

Recall from Section **??** that the rupture centroid (r_x, r_y, r_z) is defined in terms of a local coordinate system. The depth of the rupture centroid is determined in a two step process as follows:

$$f_2 = \begin{cases} 1 & \text{if } r_m \leq 4 \\ 1 + \frac{r_m - 4}{2} & \text{if } 4 < r_m \leq 6 \\ 2 & \text{if } r_m \geq 6 \end{cases} \tag{3.40}$$

and

$$r_z = \min\{f_z + \frac{f_2}{3} f_w \sin(r_{dip}), f_z + f_w \sin(r_{dip}) - \frac{1}{2} r_w \sin(r_{dip})\}. \tag{3.41}$$

The other two coordinates of the rupture centroid are given by

$$r_y = r_z \cot(r_{dip}) \tag{3.42}$$

and

$$r_x = \frac{r_l}{2}. \tag{3.43}$$

The depth to the rupture centroid r_z and the length and width of the rupture plane are illustrated for four separate simulations in Figures 3.9 to 3.12(d), (e) and (f) respectively.

3.2.10 Azimuth and dip of rupture

The rupture azimuth of each event may be forced to lie within a user defined range

$$\phi - d_\phi \leq r_\phi \leq \phi + d_\phi$$

where ϕ and d_ϕ are `azi` and `d_azi` in `setdata` respectively. The default values for ϕ and d_ϕ are 180^o and 180^o respectively. Recall that the azimuth is further adjusted through the process described in Section 3.2.2, unless d_{azi} is assigned to a negative number. Note that in the current version of the EQRM application this adjustment process ignores the above azimuth bound. That is, the azimuth of one or more events may be adjusted outside the above range in an attempt to force the event's end point to lie within the source zone. The distribution of simulated azimuth for four different simulations is illustrated in Figures 3.9 to 3.12(b).

The dip of the rupture, r_{dip}, is assigned in `setdata` and may take only a single value for all source zones. The dip of the rupture trace is measured from the ground surface and the direction of dip is such that the plane is located in the region of $y > 0$ in the local coordinate system. Note that it is possible to modify the software to select the dip randomly in a similar fashion to that described for the azimuth above. However, it is worth noting that the effect of a change in dip varies based on the attenuation model (see Section 5). For example, when using an attenuation model that depends on the Joyner-Boore distance, the practical effect of a change in dip can be compensated by a horizontal translation of the rupture trace. In such cases the random location of the rupture trace negates the need to select the dip randomly.

3.3 Using multiple source zones - Incorporating epistemic uncertainty

Recall that Section 3.2 introduced two techniques for passing source models to the EQRM:

1. the `<site_loc>_par_sourcezones.txt` and `<site_loc>_par_sourcepolys-.txt` file pair, and
2. the `<site_loc>_par_source.mat` file.

Recall also that the notion of a 'generation polygon' was introduced where the generation polygons are simply the individual source zones when Technique 1 is used, and 'generation polygon' refers to the polygons in the `generation` variable when Technique 2 is used. This definition ensured that the discussion of earthquake location (Section 3.2.2), rupture geometry (Section 3.2.9), azimuth and dip (Section 3.2.10) described above is applicable to both of the two techniques for providing source information to the EQRM. When Technique 2 is used, however, a minor modification must be made to the algorithm described for magnitude and event activity determination (see Section 3.2.8). To explain the differences in the algorithm and to explain the incorporation of multiple source zone models (epistemic uncertainty with source models) the reader will need to understand the contents of the `<site_loc>_par_source.mat` file.

The `<site_loc>_par_source.mat` file contains the following two variables:

1. The `generation` variable is a MATLAB structure containing the fields `zones` and `polys`. The field `zones` is a matrix containing the same format as that described in Table **??**. The number of rows refers to the number of 'generation polygons' to be used in the generation process. It is common to have 2 'generation polygons' arranged in a doughnut shape with an inner and outer polygon. The inner polygon is typically centered on the region of interest and can be used to simulate a larger number of earthquakes (hence higher spatial concentration) than the outer polygon. The field `polys` is a matrix using the same format as Table **??** to describe the vertices of the polygons in the field `zones`.

2. The `source` variable is a MATLAB structure array containing one element for each source zone model to be used. The number of elements in the `source` structure array corresponds to the number of source zone models to be used in the simulation. Each element has fields `zones`, `polys` and `weight`. The fields `zones` and `polys` are similar to those described in 1 above, however this time they represent the individual source source zones for the source model. The field `weight` represents the weight to be applied to the source zone model.

When using Technique 2 the `generation` is used to define the 'generation polygons' and hence the location, geometry, azimuth and dip of the rupture. The variable `generation` is also used to assign magnitudes to each synthetic event using Steps 1 to 4 from Section 3.2.8. Note that a synthetic earthquake generated in this fashion may lie in none, one or multiple source zones from the different source models. The calculation of an event activity involves identifying which zones the earthquake lies within and then computing the weighted sum of

individual event activities for the zones within which it lies. The algorithm for computing the event activity for a single source zone is described in Step 5 of Section 3.2.8. Note that an earthquake must fall within the polygon that defines the source zone and its magnitude must be within the magnitude bounds for the source zone before it is considered to lie within the source zone. More details of this process can be found by looking at the functions `calc_event_activity` for the calculation of event activity and `mke_aseq4mat` for all of processes.

3.4 Spawning events

There is an option within the EQRM application to spawn (or copy) events. Such copies are required by some techniques for incorporating aleatory uncertainty in later stages of the PSHA and PSRA. For example, Section 5.6.1 describes an incorporation of aleatory attenuation uncertainty that does not require spawning of the catalogue whereas Section 5.6.2 describes a process that does require spawning. The option is performed by `fuse_4hzd`. Actually `fuse_4hzd` is an essential component of any earthquake hazard or risk run with the `EQRM` application i.e. it must be run regardless of whether the user wishes to spawn events or not. The spawning of events is controlled by `src_eps_switch` in `setdata`. If `src_eps_switch`$= 1$ or -1 then all events with magnitude greater than m_{bnd} (`mbnd` in `setdata`) are copied $n_{samples}$ (`nsamples` in `setdata`) times with their event activity adjusted according to a weight w_e. If `src_eps_switch`$= 0$ no spawning is undertaken. When spawning the weight w_e is derived by truncating and re-normalising a standard normal distribution to $\pm n_\sigma$ (`nsigma` in `setdata`). The process is summarised below.

We know that the standard normal distribution $N \sim (0, 1)$ has a PDF given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad -\infty < x < \infty. \tag{3.44}$$

To truncate and re-normalise $N \sim (0, 1)$ to $\pm n_\sigma$ we must evaluate $P(-n_\sigma\sigma \leq X \leq n_\sigma\sigma)$. The error function

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt. \tag{3.45}$$

is related to the cumulative area under the standard normal distribution via

$$erf(x) = 2P(X \leq x) - 1. \tag{3.46}$$

The function `mke_ctpdf` computes $P(-n_\sigma\sigma \leq X \leq n_\sigma\sigma)$ using the error function as follows

$$\begin{aligned} P(-n_\sigma\sigma \leq X \leq n_\sigma\ \sigma) &= P(X \leq n_\sigma) - P(X \leq -n_\sigma) \\ &= 2P(X \leq n_\sigma) - 1 \\ &= erf(n_\sigma). \end{aligned} \tag{3.47}$$

Then, `mke_ctpdf` computes the truncated and re-normalised PDF by evaluating

$$\tilde{f}_X(x) = \frac{1}{P(-n_\sigma\sigma \leq X \leq n_\sigma\sigma)} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad -n_\sigma\sigma < x < n_\sigma\sigma. \tag{3.48}$$

The observant reader will notice that because the EQRM application works in the discrete world, simply evaluating $\tilde{f}_X(x)$ will not suffice. The approach used by `mke_ctpdf` to discretise $\tilde{f}_X$ and compute the weights $\{w_{e,i}\}_{i=1}^{n_{samples}}$ is identical to that used to compute the event activity r_ν (see Section 3.2.8) and is described in Appendix **??**. That is

$$w_{e,i} = \frac{\tilde{f}_X(x_i)}{\sum_{j=1}^{n_{samples}} \tilde{f}_X(x_j)}. \tag{3.49}$$

Once the weights $\{w_{e,i}\}_{i=1}^{n_{samples}}$ are computed the event activity r_ν for each of the event copies is re-defined as follows:

$$r_{\nu,i} = r_{\nu,original} \times w_{e,i} \quad \text{for } i = 1 \text{ to } n_{samples}, \tag{3.50}$$

for each copy of the original event and its associated weighting. A source epsilon term r_ε is also defined for each copy of the original event i.e. $\{r_{\varepsilon,i}\}_{i=1}^{n_{samples}}$. Each of the $r_{\varepsilon,i}$ are defined such that they correspond to the i^{th} bin centroid from the $n_{samples}$ bin spanning $\pm n_\sigma$ (see Section 5.6.2 for the mathematical definition).

3.5 Analysing a scenario event

The EQRM application (through `mke_evnts`) incorporates an option for considering a particular (or scenario) event. This option is instigated when the user defines `determ_flag`= 1 in `setdata`. The earthquake catalogue takes the same form as that described in Table 3.1. The scenario event is constrained by user defined values for magnitude (`determ_mag`), position (`determ_lat` and `determ_lon`), depth (`determ_r_z`) and azimuth (`determ_azi`).

3.6 Key functions, flags and parameters

Name	Type	Description
mke_evnts	Function	Master file for creating an event catalogue.
mke_aseq4mat	Function	Handles a significant proportion of the synthetic event generation on behalf of mke_evnts.
m2grpdfb	Function	Computes the PDF, f_M for the bounded Gutenberg-Richter distribution.
mag2rup_v	Function	Computes rupture dimension and 'relative' location based on magnitude.
fuse_4hzd	Function	Master file for spawning events.
mke_ctpdf	Function	Truncates and discretises $N \sim (0, 1)$.
azi	Par	Centre of azimuth range for events.
d_azi	Par	Half width of azimuth range.
dip	Par	Dip of virtual fault.
ntrgvector	Par	Vector whose elements represent the desired number of events to be simulated in each source zone.
wdth	Par	Width of virtual fault.
nbins	Par	Number of bins for event magnitude sampling.
src_eps_switch	Par	Events are (not) spawned if src_eps_switch $= 1$ $(= 0)$.
mbnd	Par	Only events with magnitude greater than mbnd are spawned.
nsamples	Par	Creates nsamples copies of each spawned event.
nsigma	Par	Range of distribution to be considered when spawning.
determ_flag	Par	EQRM application used in scenario mode if determ_flag $= 1$.
determ_mag	Par	Magnitude of scenario event.

`determ_lat`	Par	Latitude of scenario epicenter (corresponds to r_c^{lat}).
`determ_lon`	Par	Longitude of scenario epicenter (corresponds to r_c^{lon}).
`determ_r_z`	Par	Depth to scenario epicenter in km (corresponds to r_z).
`determ_azi`	Par	Azimuth of scenario event (corresponds to r_ϕ).

Chapter 4

Grids and building databases

4.1 Overview

The calculation of earthquake hazard and risk are spatial problems. This chapter describes how grids are required to compute earthquake hazard and how building databases are represented to model earthquake risk. Computationally the EQRM operates by looping over the grid (building database) points one at a time i.e. the hazard or risk is computed iteratively at each point. Tools to assist the user create building databases can be found in `*/eqrm/datacvt`.

4.2 Hazard grids

The term 'hazard grid' merely refers to a set of spatially located points at each of which earthquake hazard can be computed (see Chapter 9 for a definition of earthquake hazard). The set can contain one or more points. Typically however, the grids are uniform and regularly spaced in the horizontal and vertical direction. The EQRM can work with two different hazard grid types, the choice of which is controlled by the `setdata` parameter `grid_flag`.

When `grid_flag` = 1 the EQRM loads the file `<site_loc>_par_site.mat` which contains the following variables:

1. `SiteLocations`: a double array containing a row for each grid point and two columns. The first and second columns contain the latitude and longitude of the grid point respectively.

2. site_classes: a character column array containing one row for each grid point. The content of each row is a single letter identifying the regolith site class of the grid point (see Chapter 6).

Typically such a grid originates in the GIS world where a grid of points is created and assigned to site classes. This information is imported into MATLAB where it is manipulated into the above format and saved ready for use. Note that hazard values computed on a <site_loc>_par_site.mat must be exported to the GIS environment for plotting (i.e. they can not be plotted by the EQRM).

The <site_loc>_par_site.mat is an older format that has been kept for back compatibility. The preferred technique for creating hazard grids is through the Input Manager. The Input Manageris a GUI, specifically designed to create earthquake hazard grids. It can be accessed through the eqrm_param_gui by pressing the Input Managerbutton. The Input ManagerGUI allows the user to create a grid over a region of interest. To do this the Input Managerneeds to access a polygon input file that describes the region of interest as a single polygon (e.g. perth_par_study_region_box.txt) or as a set of site classes covering the study region (e.g. perth_par_site_class_polys.mat). Note that in the first case the EQRM will not access any information about site classes and hence amplification can not be included. The user must indicate how many points are desired in the latitude and longitude directions. The name of the created grid file is <site_loc>_par_site_uniform.mat and it can be used with the EQRM by setting grid_flag = 2. Its format is more complicated than the above mentioned <site_loc>_par_site.mat, however the user does not need knowledge of the format since this is handled by the Input Manager. Note that earthquake hazard estimates produced on a <site_loc>_par_site_uniform.mat file can be plotted using the EQRM hazard plotting tools (see Chapter 9).

4.3 Building databases

The building database contains attributes for the portfolio of buildings under consideration in the risk assessment. The building database is a MATLAB data file with filename sitedb_<site_loc>.mat containing the following variables:

1. all_postcodes: a list of all postcodes currently included in the EQRM for use in linking database postcode id numbers.
2. all_suburbs: a list of all suburbs currently included in the EQRM for use in linking database suburb id numbers.

Table 4.1: Columns of the building database file `sitedb_<site_loc>.mat`.

Column	Description
1	Integer site identifier for EQRM (typically the same as column 10)
2	Latitude of building
3	Longitude of building
4	Index to building construction type... **expanded** HAZUS list (Section 4.3.1)
5	Index to HAZUS usage classification (Section 4.3.2)
6	Total floor area in square meters (summed over all stories)
7	Survey factor indicating how many 'real' buildings the database entry represents
8	Index to suburb name for building (variable: `all_suburbs`)
9	Index to postcode name for building (variable: `all_suburbs`)
10	Integer site identifier for comparison against original database
11	Logical index stating whether the building is pre- (0) or post- (1) the 1989 Newcastle earthquake
12	Index to building construction type... HAZUS list (Section 4.3.1)
13	Replacement cost of building in dollars per square meter (Section 4.3.3)
14	Replacement cost of contents in dollars per square meter (Section 4.3.3)
15	Index to FCB usage classification (Section 4.3.2)

3. `b_sitemat` a matrix containing one row for each building and 15 columns describing attributes of the building. Details of the columns are given in Table 4.1.

4. `b_soil`: a character column array containing one row for each grid point. The contents of each row is a single letter identifying the regolith site class of the grid point (see Chapter 6).

Typically the building database used with the EQRM represents a subset of the true portfolio of interest. When creating a database that sub-samples a larger portfolio, individual database entries are used to represent more than one 'real' building. Such sub-sampling is undertaken to reduce run times and memory requirements. Results from an EQRM simulation can be re-converted to the full portfolio using the `survey factor`defined in `sitedb_<site_loc>.mat` (see Section 8.2.2).

Tools to assist the creation of building databases for use with the EQRM can be found in `*/eqrm/datacvt/buildingdb`.

4.3.1 Building construction types

Buildings have been subdivided into a number of building types each with their own set of building parameters uniquely defining the median capacity curve and the random variability around the median (see Chapter 7). The building construction types are based upon the HAZUS definitions (**?**), with some further subdivisions recommended by Australian engineers for Australian building construction types (**?**).

In essence, the seven basic HAZUS types are

- Timber frame (W)
- Steel frame (S)
- Concrete frame (C)
- Pre-cast concrete (PC)
- Reinforced masonry (R)
- Unreinforced masonry (URM)
- Mobile homes (MH)

There are further subdivisions of the HAZUS types into subtypes according to numbers of stories in the building. These are given in Table 4.2.

The new Australian sub-types, developed in the Australian engineers workshop, further subdivide some of the HAZUS types (**?**). In particular, the timber frame category (W1) is subdivided into wall types (timber or brick veneer walls) and roof types (metal or tiled); the unreinforced masonry types (URML and URMM) into roof type (metal, tile or otherwise), and the concrete frame types are subdivided into soft-story or non-soft story types. Soft-story refers to buildings that may have a concrete basement or parking area but wood frame stories.

In total, we currently have 56 possible construction types although some are rarely used. For example; the original Hazus W1 is still there, however this is rarely used in favor of the more detailed classification into W1TIMBMETAL,

Table 4.2: Definitions of the basic HAZUS building construction types.

code	description	Stories
W1	timber frame $< 5\,000$ square feet	(1–2)
W2	timber frame $> 5\,000$ square feet	(All)
S1L	steel moment frame	Low-Rise (1–3)
S1M		Mid-Rise (4–7)
S1H		High-Rise (8+)
S2L	steel light frame	Low-Rise (1–3)
S2M		Mid-Rise (4–7)
S2M		High-Rise (8+)
S3	steel frame + cast concrete shear walls	(All)
S4L	steel frame + unreinforced masonry in-fill walls	Low-Rise (1–3)
S4M		Mid-Rise (4–7)
S4H		High-Rise (8+)
S5L	steel frame + concrete shear walls	Low-Rise (1–3)
S5M		Mid-Rise (4–7)
S5H		High-Rise (8+)
C1L	concrete moment frame	Low-Rise (1–3)
C1M		Mid-Rise (4–7)
C1H		High-Rise (8+)
C2L	concrete shear walls	Low-Rise (1–3)
C2M		Mid-Rise (4–7)
C2H		High-Rise (8+)
C3L	concrete frame + unreinforced masonry in-fill walls	Low-Rise (1–3)
C3M		Mid-Rise (4–7)
C3H		High-Rise (8+)
PC1	pre-cast concrete tilt-up walls	(All)
PC2L	pre-cast concrete frames with concrete shear walls	Low-Rise (1–3)
PC2M		Mid-Rise (4–7)
PC2H		High-Rise (8+)
RM1L	reinforced masonry walls + wood or metal diaphragms	Low-Rise (1–3)
RM1M		Mid-Rise (4+)
RM2L	reinforced masonry walls + pre-cast concrete diaphragms	Low-Rise (1–3)
RM2M		Mid-Rise (4–7)
RM2H		High-Rise (8+)
URML	unreinforced masonry	Low-Rise (1–2)
URMM		Mid-Rise (3+)
MH	Mobile homes	(All)

W1BVTILE, etc. A complete list of all the building construction types is given in Table 4.3. The `setdata` flag `hazus_btypes_flag` can be used to select between the use of the HAZUS building type classification and the Australian engineers extended HAZUS building type classification.

Table 4.3: Complete list of all building construction types (with those that are rarely used in italics).The integers corresponding to each building construction type represent the integer index used in the building database Column 4 for expanded HAZUS types (column 12 for HAZUS only types).

1: W1	15: S5H	29: RM1L	43: C1LSOFT
2: W2	16: *C1L*	30: RM1M	44: C1LNOSOFT
3: S1L	17: *C1M*	31: RM2L	45: C1MMEAN
4: S1M	18: *C1H*	32: RM2M	46: C1MSOFT
5: S1H	19: C2L	33: RM2H	47: C1MNOSOFT
6: S2L	20: C2M	34: *URML*	48: C1HMEAN
7: S2M	21: C2H	35: *URMM*	49: C1HSOFT
8: S2H	22: C3L	36: MH	50: C1HNOSOFT
9: S3	23: C3M	37: W1MEAN	51: URMLMEAN
10: S4L	24: C3H	38: W1BVTILE	52: URMLTILE
11: S4M	25: PC1	39: W1BVMETAL	53: URMLMETAL
12: S4H	26: PC2L	40: W1TIMBERTILE	54: URMMMEAN
13: S5L	27: PC2M	41: W1TIMBERMETAL	55: URMMTILE
14: S5M	28: PC2H	42: C1MMEAN	56: URMMMETAL

The building capacity curves used in the Newcastle and Lake Macquarie study (**?**) are provided in Appendix **??**, Table **??**. Other examples of the building capacity curves, as well as tools for manipulating them, can be found in the directory `*/eqrm/datacvtbuildingpars`.

4.3.2 Building usage types

The cost models used by the EQRM require knowledge of the building's use in society. For example the value of a factory's contents will vary from the value of a residents house. Similarly, the cost associated with building a hospital and the cost of building a local shop may differ even if the same materials are used because the buildings may be built to different standards. To transfer this information to the EQRM the building database must store information about each building's usage. There are two different schemes that can be used; the functional classification of building (FCB) usage (**?**) and the HAZUS usage classification (**?**).

Table 4.4: Functional classification of building (FCB) (**?**) and integer index used in the building database column 15.

	Residential: Separate, kit and transportable homes
1	111: Separate Houses
2	112: Kit Houses
3	113: Transportable/relocatable homes
	Residential: Semi-detached, row or terrace houses, townhouses
4	121: One storey
5	122: Two or more storeys
	Residential: Flats, units or apartments
6	131: In a one or two storey block
7	132: In a three storey block
8	133: In a four or more storey block
9	134: Attached to a house
	Residential: Other residential buildings
10	191: Residential: not otherwise classified
	Commercial: Retail and wholesale trade building
11	211: Retail and wholesale trade buildings
	Commercial: Transport buildings
12	221: Passenger transport buildings
13	222: Non-passenger transport buildings
14	223: Commercial carparks
15	224: Transport: not otherwise classified
	Commercial: Offices
16	231: Offices
	Commercial: Other commercial buildings
17	291: Commercial: not otherwise classified
	Industrial: Factories and other secondary production buildings
18	311: Factories and other secondary production buildings
	Industrial: Warehouses
19	321: Warehouses (excluding produce storage)
	Industrial: Agricultural and aquacultural buildings
20	331: Agricultural and aquacultural buildings
	Industrial: Other industrial buildings
21	391: Industrial: not otherwise classified
	Other Non-Residential: Education buildings
22	411: Education buildings
	Other Non-Residential: Religion buildings
23	421: Religion buildings
	Other Non-Residential: Aged care buildings
24	431: Aged care buildings
	Other Non-Residential: Health facilities (not in 431)
25	441: Hospitals
26	442: Health: not otherwise classified
	Other Non-Residential: Entertainment and recreation buildings
27	451: Entertainment and recreation buildings
	Other Non-Residential: Short term accommodation buildings
28	461: Self contained, short term apartments
29	462: Hotels (predominately accommodation), motels, boarding houses, hostels or lodges
30	463: Short Term: not otherwise classified
	Other Non-Residential: Other non-residential buildings
31	491: Non-residential:not otherwise classified

Table 4.5: HAZUS building usage classification (**?**) and integer index used in the building database column 5.

	Residential
1	RES1: Single family dwelling (house)
2	RES2: Mobile home
3	RES3: Multi family dwelling (apartment/condominium)
4	RES4: Temporary lodging (hotel/motel)
5	RES5: Institutional dormitory (jails, group housing - military, colleges)
6	RES6: Nursing home
	Commercial
7	COM1: Retail trade (store)
8	COM2: Wholesale trade (warehouse)
9	COM3: Personal and repair services (service station, shop)
10	COM4: Professional and technical services (offices)
11	COM5: Banks
12	COM6: Hospital
13	COM7: Medical office and clinic
14	COM8: Entertainment and recreation (restaurants, bars)
15	COM9: Theaters
16	COM10: Parking (garages)
	Industrial
17	IND1: Heavy (factory)
18	IND2: Light (factory)
19	IND3: Food, drugs and chemicals (factory)
20	IND4: Metals and mineral processing (factory)
21	IND5: High technology (factory)
22	IND6: Construction (office)
	Agriculture
23	AGR1: Agriculture
	Religion/Non/Profit
24	REL1: Church and non-profit
	Government
25	GOV1: General services (office)
26	GOV2: Emergency response (police, fire station, EOC)
	Education
27	EDU1: Grade schools
28	EDU2: Colleges and Universities (not group housing)

The FCB usage is summarised in Table 4.4 and the HAZUS usage classification is summarised in Table 4.5. The `setdata` flag `b_usage_type_flag` can be used to switch between the two usage classifications.

4.3.3 Replacement costs

The replacement cost in dollars per square metre for each building and the replacement cost of the contents of each building are contained within the building database (columns 13 and 14 respectively: see Section 4.3). Typically these costs are a function of the usage classification of the building and are hence also dependent on whether the HAZUS or FCB classification system is used. The EQRM does not cross check how the costings were created. Note that in some instances it may be appropriate to use costings created from one usage classification with the EQRM using the other usage mode (effects cost splits - see below) and in some instance it may not be appropriate to do so. Users are encouraged to familiarise themselves with database metadata to ensure that they are using the EQRM appropriately for their own application.

4.4 Key functions, flags and parameters

Name	Type	Description
*/datacvt /buildingpars	Dir	Tools for manipulating and preparing engineering parameter files.
*/datacvt /buildingdb	Dir	Tools for manipulating and preparing building databases.
*/datacvt /econsoclosspars	Dir	Tools for manipulating and preparing costing split ratios.
Input Manager	GUI	Tool for creating earthquake hazard grids.
eqrm_param_gui	GUI	Main EQRM GUI.
<site_loc>_par_study _region_box.txt	Par File	Describes study region with a box - can be used to generate hazard grids.
<site_loc>_par_site _class_polys.mat	Par File	Describes study region as a set of site classes - can be used to generate hazard grids.
<site_loc>_par_site _uniform.mat	Par File	Earthquake hazard grid.
sitedb_<site_loc>.mat	Par File	Building database.
b_usage_type_flag	Par	Flag for selecting the usage classification system.
hazus_btypes_flag	Par	Flag for selecting the building construction type classification system.
buildpars_flag	Par	Flag for selecting engineering parameters.
grid_flag	Par	Flag for selecting the type of earthquake hazard grid to be used.

Chapter 5

Ground-Motion Prediction Equations

5.1 Overview

Ground-Motion Prediction Equations (GMPEs) play key role in the evaluation of seismic hazard in the region of interest. In this chapter, we mainly focus on how EQRM deals with GMPEs and corresponding uncertainties. However, to help the interested user to become more familiar with the state-of-art and issues on using GMPEs in the framework of seismic hazard analysis, each section comes with a set of useful references. The remainder of this chapter is organised as follows. First, the background theory of GMPEs is addressed in Section 5.2. Then, an introduction on the selected GMPEs is brought in Section 5.3. Section 5.4 describes how the selected GMPEs are implemented in EQRM. Finally, capturing of uncertainties in EQRM, which is a major challenge in estimation of strong motions, forms the last section of this chapter.

5.2 Background theory

The Ground-Motion Prediction Equations (GMPEs), also referred to as attenuation models, are used to describe the attenuation of the ground-motion parameter of interest with respect to parameters of the earthquake source, propagation path and local site conditions, collectively referred to as seismological parameters. These equations are obtained from regression analysis on the recorded or

synthetic values of the parameter of interest. A GMPE can be described by the general expression as:

$$\ln Y = f(x, \theta) + \varepsilon, \tag{5.1}$$

where Y is the median of the ground-motion parameter of interest, x is the vector of seismological parameters, θ is the vector of the GMPEs regression coefficients, and ε is a random error term with a mean of zero and a standard deviation of $\sigma_{\ln Y}$. In this regard, for a given set of seismological parameters, a GMPE describes the probability density function (PDF) of the ground-motion parameter as a lognormal distribution with median and standard deviation equal to Y, and $\sigma_{\ln Y}$, respectively ($f(Y|x)$). This is a basic requirement by seismic hazard analysis (SHA) in order to calculate the probability of exceeding desired ground-motion level for a given set of seismological parameters.

In engineering practice, traditionally, the most desired ground-motion parameters are Peak Ground Velocity (PGV), Peak Ground Acceleration (PGA), and 5% damped Pseudo Spectral Acceleration (PSA or SA)of horizontal components. There are a number of ways to define the response variable based on two horizontal components, such as considering the perpendicular horizontal components to be independent, and geometric mean of horizontal components. Although using the geometric mean has the advantage that the standard deviation for the ground-motion models is smaller than for other measures (Beyer and Bommer 2006), this measurement nevertheless depends on the orientations of the sensors installed in the field. Recently, alternative definitions of ground motion that are independent of the sensor orientation was introduced by Boore et al. (2006). These parameters, so called as GMRotI50, and GMRotD50 are based on a set of geometric means computed from the recorded orthogonal horizontal motions after rotating them through a non-redundant angle of 90.

The most widely used seismological parameters by GMPEs are earthquake magnitude (source parameter), source-to-site distance measure (path parameter), and description of local site condition (site parameter).

The earthquake magnitude is used to measure the size of the earthquake. There are different scales to define magnitude. The most commonly magnitude scales that are used in the development of the GMPEs throughout the world are moment magnitude, M_W and surface wave magnitude, M_S. However, many GMPEs use local magnitude scale, M_L, for smaller magnitude earthquakes. The M_W scale has been used by most of the recent GMPEs. It is based on the moment of the earthquake (M_0), which is a measure of the seismic energy radiated by an earthquake:

$$M_W = \frac{2}{3} \log M_0 - 10.7. \tag{5.2}$$

The main advantages of M_W scale are that it is not saturated at the upper end and it can be determined from geological faulting measurements or from seismic waves.

Source-to-site distance measure is used to characterise the attenuation of the ground motion amplitude versus distance from the earthquake source. Different measures of distance are used by different GMPEs. These measures are illustrated in Figure 5.1 and can be summarized as follows:

Point-source distance measures:

- Hypocentral distance (R_{hypo}): Distance from the hypocenter of an earthquake to the site of interest.
- Epicentral distance (R_{epi}): Distance from the epicenter of an earthquake to the site of interest.

Finite-source distance measures:

- Fault distance (R_{rup}): Closest distance between rupture plane and site of interest.
- Joyner-Boore distance (R_{jb}): Closest distance between surface projection of the rupture plane and site of the interest.
- Seismogenic distance (R_{seis}): Closest distance between seismogenic part of the rupture plane and site of the interest.

In case of small earthquake that can be represented by a point source, the point source measures can be considered as source-to-site distance measure. However, the point source measures are not suitable to describe the attenuation of ground motion away from large earthquakes. In this case, the finite source measures are preferred.

Local site conditions profoundly affect the ground motion record characteristics. Several numerical and/or empirical techniques have been proposed to estimate local site responses. Such studies show distinct amplification levels in the sites of different geological and geotechnical characteristics; therefore, it is an ordinary practice to categorize sites into different general classes. Different sites of the same class are supposed to have local site responses with similar main characteristics. The most commonly used site categorization schemes by GMPEs are based on geomatrix, surface geology, and average shear wave velocity over 30m (V_{S30}).

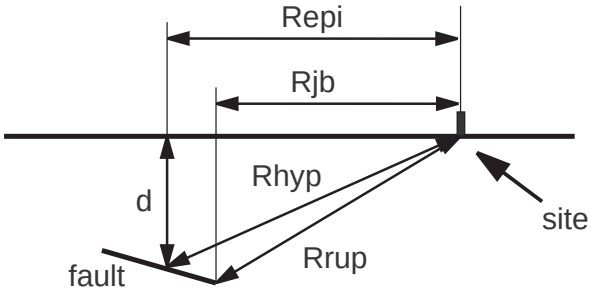


Figure 5.1: Diagram showing the distance measures used in various attenuation formulae.

Actually, many of the GMPEs follow exactly the same or similar classification criteria as suggested by NEHRP (Table 5.1). V_{S30} has also been included directly in the functional form of several attenuation models. It is worth mentioning that in spite of V_{S30} popularity as a site parameter, it has its own shortcomings. Several studies have suggested using other site parameters such as the average shear wave velocity over the depth equal to a quarter-wavelength of the period of interest, and the average shear wave velocity of 100-200m.

Table 5.1: Site categories in NEHRP provisions

Site Category	Description	V_{S30}
A	Hard rock	$> 1500m/s$
B	Firm to hard rock	$760 - 1500m/s$
C	Dense soil, soft rock	$360 - 760m/s$
D	Stiff soil	$180 - 360m/s$
E	Soft clays	$< 180m/s$

5.3 Implemented GMPEs in EQRM

In recent years the number of GMPEs is considerably increased due to development of strong motion networks and the large number of available strong motion records. Douglas (2003) provides a review on more than 120 GMPEs for the estimation of peak ground acceleration and 80 GMPEs for the estimation of response spectra. Generally, GMPEs can be grouped into two different principle tectonic regions: shallow crustal earthquakes and subduction zone earthquakes. The shallow crustal earthquakes can be further divided into earthquakes in active tectonic regions (e.g. Japan, California, Iran), and earthquakes in stable continental regions (e.g. Australia, Eastern North America). The subduction zone earthquakes can be further divided into intraslab events and earthquakes along the interface of subducting plates (e.g. Japan, Chile, Indonesia). For details of the main characteristics of each category see Campbell (2003). For each tectonic region, a handful of GMPEs are available for use within EQRM software. Their main characteristics are summarized in Table 5.2. All of these equations are widely used in engineering and engineering seismology. They reflect a broad variety of tastes in terms of their functional forms and databases.

Table 5.2: Main characteristics of the implemented GMPEs

Model Name	Magnitude Type	Magnitude Range	Distance Type	Distance Range	Site Condition	Horizontal Component Type	Period Range	Complexity
Shallow crustal events in active tectonic regions								
?	M_W	4-7.4	R_{rup}	0-100	GMX	Geometric mean	0-4	9
Zhao et al. (2006)	M_W	4.9-8.3	R_{rup}	0-300	$V_{S30}Class$	Geometric mean	0-5	16
Abrahamson and Silva (2008)	M_W	5.0-8.0	R_{jb}	0-200	V_{S30}	GMRotI50	0-10	17
Boore and Atkinson (2008)	M_W	5.0-8.0	R_{rup}	0-200	V_{S30}	GMRotI50	0-10	14
Campbell and Bozorgnia (2008)	M_W	4.0-8.5	R_{rup}	0-200	V_{S30}	GMRotI50	0-10	17
Chiou and Youngs (2008)	M_W	4.0-8.5	R_{rup}	0-200	V_{S30}	GMRotI50	0-10	25
Akkar and Bommer (2010)	M_W	5.0-7.6	R_{jb}	0-100	V_{S30} Class	Geometric mean	0-3	14
Shallow crustal events in stable continental regions								
?	M_L	?	R_{hypo}	10-500	V_{S30} Class	Random	0-2	?
?	M_W	4.0-7.25	R_{hypo}	10-500	V_{S30} Class	Random	0-2	5
?	M_W	5.0-8.0	R_{jb}	1-500	Hard Rock	Geometric mean	0-2	9
?	M_W	5.0-8.2	R_{rup}	1-1000	Hard Rock	Geometric mean	0-4	13
Atkinson and Boore (2006)	M_W	4.0-8.0	R_{rup}	1-1000	V_{S30}	Random	0-5	14
Liang et al. (2008)	M_L	4.0-7.0	R_{epi}	10-200	Hard Rock	Random	0-30	6
Somerville (2009)	M_W	5-7.5	R_{jb}	0-500	Hard Rock	?	0-10	10
Subduction zone events								
Youngs et al. (1997)	M_W	5.0-8.2	R_{rup}	10-500	GMX	Geometric mean	0-3	7
Atkinson and Boore (2003)	M_W	5.0-8.3	R_{rup}	10-500	V_{S30} Class	Random	0-3	9
Zhao et al. (2006)	M_W	5.0-8.3	R_{rup}	30-300	$V_{S30}Class$	Geometric mean	0-5	16

Considering the general tectonic setting of the region of interest, the user may either select GMPEs from those listed in Table 5.2, or implement his or her desirable model. It should be also noted that GMPEs have to be selected based on their magnitude and frequency validity ranges. In other words, GMPEs should not be used outside the ranges of their underlying dataset (Bommer et al., 2007). General guidelines on selection of GMPEs for seismic hazard analysis are provided by Cotton et al. (2006) and Bommer et al. (2010). In addition, if strong motion data is available in the target region, it can be used to verify candidate GMPEs. To compare ground motions or GMPEs between regions, the main approaches are the direct comparison of median predictions from GMPEs for different regions (Campbell and Bozorgnia, 2006; Stafford et al., 2008), analysis of variance (Douglas, 2004a, b), and evaluation of the consistency of data distributions with respect to a GMPE using likelihood concepts (Scherbaum et al., 2004; Stafford et al., 2008) or information-theory approach (Scherbaum et al., 2009).

The complexity of the implemented GMPEs varies based on the database size and developers preferences. In Table 5.2, the complexity of each GMPE is represented as the number of the regression coefficients. In addition to magnitude scaling, distance scaling and local site condition terms, some of the GMPEs have extra terms to model other phenomena affecting the ground motion amplitudes. Table 5.3 lists such terms for GMPEs developed for active crustal regions. As it can be seen, among the selected GMPEs, Abrahamson and Silve (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008) models can be considered as the most complex ones with many input parameters required. They notably require extra input parameters to model: (1) observed higher ground-motions on the hanging-wall of the fault-plane, (2) larger ground-motions generated by buried rupture earthquakes, and (3) basin effect which is believed to have significant impact on amplification of ground motion especially at longer period range. The mentioned GMPEs as well as Boore and Atkinson (2008) model account for nonlinear site amplification that is due to nonlinear soil response beyond a certain level of deformations. These models are developed as part of the Next Generation of Attenuation of Ground-Motions (NGA) project in 2008 for shallow crustal earthquakes in California. In many regions, however, it is not possible to determine all the necessary input parameters for such complicated models. General guidelines on estimating unknown input parameters when implementing NGA GMPEs in engineering practice are presented in Kaklamanos et al. (2010). The details on how EQRM feeds each of the implemented GMPEs are given in the next section.

Table 5.3: Functional forms of GMPEs for active tectonic regions

	Ground Motion Prediction Equations						
Term	Sa97	Zh06	AS08	BA08	CB08	CY08	AB10
Near-fault Saturation	•	•	•	•	•	•	•
Magnitude-dependent Distance Decay			•	•	•	•	•
Style of Faulting	•	•	•	•	•	•	•
Depth-to-Top of Rupture			•		•	•	
Hanging-Wall			•		•	•	
Basin Effect			•		•	•	
Nonlinear Site Response			•	•	•	•	

5.3.1 Further Comments on Specific GMPEs

Gaull...

Hadi to add

Somerville09

Complete this section

- decsribe the two options available (NonCratonic and Yilgarn)
- cite the report
- create table with the NonCratonic coeeficients (i.e. Table **??**)

Table 5.4: Coeeficients for `Somerville09_Non_Cratonic`. The Non-Cratonic model represents an everage model for all non cratonic regions of Australia (**update citation**. For all practical purposes it can be used in all regions except the Yilgran Craton. The complete table of coefficients is include here because it was omitted from the original publication where it should have been table 7-3g.

Period	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
PGA	1.03780	-0.03970	-0.79430	0.14450	-0.00618	-0.72540	-0.03590	-0.09730
0.010	1.05360	-0.04190	-0.79390	0.14450	-0.00619	-0.72660	-0.03940	-0.09740
0.020	1.05680	-0.03920	-0.79680	0.14550	-0.00617	-0.73230	-0.03930	-0.09600
0.030	1.13530	-0.04790	-0.80920	0.15000	-0.00610	-0.76410	-0.05710	-0.09210
0.040	1.30000	-0.07020	-0.83150	0.15920	-0.00599	-0.82850	-0.09810	-0.08530
0.050	1.47680	-0.09310	-0.83330	0.15600	-0.00606	-0.86740	-0.12740	-0.09130
0.075	1.70220	-0.05160	-0.80720	0.14560	-0.00655	-0.87690	-0.10970	-0.08690
0.100	1.65720	0.15080	-0.77590	0.13100	-0.00708	-0.77830	0.01690	-0.05980
0.150	1.94440	-0.09620	-0.75000	0.11670	-0.00698	-0.69490	-0.13320	-0.12530
0.200	1.82720	-0.06230	-0.73430	0.11940	-0.00677	-0.64380	-0.09570	-0.11920
0.250	1.74380	-0.02530	-0.72480	0.11950	-0.00646	-0.63740	-0.06250	-0.11650
0.3003	1.80560	-0.27020	-0.73190	0.13490	-0.00606	-0.66440	-0.17470	-0.14340
0.400	1.88750	-0.37820	-0.70580	0.09960	-0.00589	-0.58770	-0.24420	-0.21890
0.500	2.03760	-0.79590	-0.69730	0.11470	-0.00565	-0.59990	-0.48670	-0.29690
0.750	1.93060	-0.80280	-0.74510	0.11220	-0.00503	-0.59460	-0.50120	-0.34990
1.000	1.60380	-0.47800	-0.86950	0.07320	-0.00569	-0.41590	0.06360	-0.33730
1.4993	0.47740	0.90960	-1.02440	0.11060	-0.00652	-0.19000	1.09610	-0.10660
2.000	-0.25810	1.37770	-1.01000	0.10310	-0.00539	-0.27340	1.50330	-0.04530
3.0003	-0.96360	1.14690	-0.88530	0.10380	-0.00478	-0.40420	1.54130	-0.11020
4.000	-1.46140	1.07950	-0.80490	0.10960	-0.00395	-0.46040	1.41960	-0.14700
5.000	-1.61160	0.74860	-0.78100	0.09650	-0.00307	-0.46490	1.24090	-0.22170
7.5019	-2.35310	0.35190	-0.64340	0.09590	-0.00138	-0.68260	0.92880	-0.31230
10.000	-3.26140	0.69730	-0.62760	0.12920	-0.00155	-0.61980	1.01050	-0.24550
PGV	5.07090	0.52780	-0.85740	0.17700	-0.00501	-0.61190	0.80660	-0.03800

5.4 Implementation

In the EQRM application the attenuation calculations are conducted within the main driving function `do_analysis` which computes the attenuation for one site at a time for all events and all fundamental (or building) periods. That is, for each site there exists a large matrix of modelled $S_a(T_o, r_m, R)$ (this matrix is `SA` or `SArock` in the code) whose rows correspond to the events (i.e. an r_m-R pair) and whose columns correspond to the T_o. That is

$$SA = \begin{bmatrix} S_a(T_{o,1}, r_{m,1}, R_1) & S_a(T_{o,2}, r_{m,1}, R_1) & \dots & S_a(T_{o,N_T}, r_{m,1}, R_1) \\ S_a(T_{o,1}, r_{m,2}, R_2) & S_a(T_{o,2}, r_{m,2}, R_2) & \dots & S_a(T_{o,N_T}, r_{m,2}, R_2) \\ \vdots & \vdots & \ddots & \vdots \\ S_a(T_{o,1}, r_{m,N_s}, R_{N_s}) & S_a(T_{o,2}, r_{m,N_s}, R_{N_s}) & \dots & S_a(T_{o,N_T}, r_{m,N_s}, R_{N_s}) \end{bmatrix}$$

where N_s refers to the number of events and N_T refers to the number of periods. Note that the r_m-R pair comes from the event catalogue and a calculation of distance, as described in this section. The T_o are determined by the user-defined vector parameter `periods` in `setdata`. Computationally the above calculation of attenuation is conducted in two parts as follows:

1. The preparation of the attenuation coefficients (`prep_attn`) conducted before entering a loop over sites. This preparation involves interpolating the attenuation coefficients to the fundamental periods defined by `periods`.

2. The computation of the $S_a(T_o, r_m, R)$ at each site (`do_attenuation`) conducted site-by-site within a loop over sites.

Each attenuation model returns an estimate of:

1. the mean of $log(S_a(T_o, r_m, R))$ (equivalent to the logarithm of the median of $S_a(T_o, r_m, R)$) and hereafter denoted by $\mu_{log(S_a)}(T_o, r_m, R)$ or $\mu_{log(S_a)}$ for brevity, and

2. the standard deviation of $log(S_a(T_o, r_m, R))$, hereafter denoted by $\sigma_{log(S_a)}(T_o, r_m, R)$ or $\sigma_{log(S_a)}$ for brevity.

The required input parameters by implemented GMPEs in EQRM, to model source, path and site effects are listed in Table 5.5. The abbreviations used in this Table are defined in Table 5.6. As it can be seen GMPEs use different definitions for explanatory and response variables. Regarding response parameter,i. e. ground-motion parameter of interest, currently EQRM estimates 5% S_a using GMPEs listed in Table 5.2. The conversion equations between different definitions of response variables are provided by Beyer and Bommer (2006). For each GMPE, EQRM should provide inputs respecting the original definition used within each model. In EQRM, the recurrence relationship and hence the generated earthquake catalogue are based on moment magnitude scale. However, among the implemented GMPES Gaul et al. (1990), and Liang et al. (2008) models use M_L scale. In this case, adjustment is done using the correlation relationship between and as suggested by Johnston (*pers. comm.*, 2001):

$$M_L = \frac{0.473 + \sqrt{0.473^2 - 4 \times 0.145(3.45 - M_W)}}{2 \times 0.145}; \tag{5.3}$$

Equation 2 is developed for stable continental regions and is based on earlier work published by the same author (**?**). It is important to note that uncertainties in the explanatory variables are map to the response variable. In other words, in

equation 5.1 if an explanatory parameter, x, has been assigned using a correlation conversion with associated measure of aleatory variability, σ_x, then this should be map into the aleatory variability of the GMPEs, σ_y, by using the expression:

$$\sigma_{Total} = \sqrt{\sigma_y^2 + (\frac{\partial log(Y)}{\partial x})^2 \sigma_x^2} \tag{5.4}$$

Table 5.5: Input parameters of the implemented GMPEs in EQRM

	Ground Motion Prediction Equations																
Parameter	Sa97	Zh06	AS08	BA08	CB08	CY08	AB10	Ga90	AB97	To97	Ca03	AB06	Li08	So09	Yo97	AB03	Zh06
Source parameters																	
M_L								•					•				
M_W	•	•	•	•	•	•	•		•	•	•	•		•	•	•	•
h															•	•	•
Z_{TOR}			•		•	•											
W			•														
δ			•		•	•											
$SOFFlag$	•	•	•	•	•	•	•										
Path parameters																	
R_{epi}													•				
R_{hypo}								•	•								
R_{jb}			•	•	•	•	•			•				•			
R_{rup}	•	•	•		•	•					•	•			•	•	•
R_x			•			•											
$HWFlag$			•		•	•											
Site parameters																	
$SiteFlag$	•	•					•	•	•						•	•	•
V_{S30}			•	•	•	•						•					
$Z_{1.0}$			•			•											
$Z_{2.5}$					•												
PGA_{rock}			•	•	•	•						•					

Table 5.6: Definition of Input parameters

Parameter	Description
M_W	Moment magnitude
M_L	Local magnitude
h	Focal depth (km)
W	Down-dip rupture width (km)
δ	Fault dip (degree)
SOF	Style Of Faulting (function of rake angle, λ (degree))
R_{epi}	Distance from the epicenter (km)
R_{hypo}	Distance from the hypocenter (km)
R_{jb}	Closest distance to the surface projection of the rupture plane (km)
R_{rup}	Closest distance to the rupture plane (km)
R_x	Horizontal distance to top edge of rupture measured perpendicular to the strike (km)
$HWFlag$	Hanging-Wall
$SiteFlag$	?
V_{S30}	Average shear-wave velocity over the top 30 m (m/s)
$Z_{1.0}$	Depth to shear-wave velocity equal to 1.0 km/s (km)
$Z_{2.5}$	Depth to shear-wave velocity equal to 2.5 km/s (km)
PGA_{rock}	Peak Ground Acceleration on rock (cm/s^2)

Traditionally in seismic hazard studies the most difficult adjustments are conversions between different source-to-site distance measures. Scherbaum et al. (2004) and later Kaklamanos et al. (2010) proposed correlation relations among popular distance metrics. It is important to note that, as indicated by Scherbaum et al. (2005), large penalty can be paid in terms of increased sigma values due to conversion of source-to-site distance measures. However, in EQRM distance measures conversions are not necessary, since virtual faults are generated through simulation process. This enables EQRM to not only directly calculate the proper source-to-site distance measure for each GMPE, but also to estimate all the source geometry related parameters, i. e. h, W, δ, R_x. The only remaining parameters to be determined are site parameters. The V_{S30} parameter is the required INPUT parameter that should be provided by the user. In the absence of shear-wave velocity measurements in the target region, one may use topographic gradient approach for estimation of V_{S30} as suggested by Allen and Wald (2009). In addition

fine scale geological maps and recorded strong motions may also be used to estimate V_{S30}. The Z1.0 and Z2.5 parameters are used by some of the NGA GMPEs to model the basin effects. Currently these parameters are not included in the INPUT file, and are estimated as follows: Z1.0 is estimated from Vs30 following the correlation relation suggested by Chiou and Youngs (2008):

$$ln(Z_{1.0}) = 28.5 - 0.4775ln(V_{S30}^{8} + 38.7^{8}) \tag{5.5}$$

Z2.5 is approximated from Z1.0 following the correlation relation suggested by Campbell and Bozorgnia (2007):

$$ln(Z_{2.5}) = 0.519 + 3.595Z_{1.0} \tag{5.6}$$

It should be noted that by using these relations, we assume that the basin depths at target region and California are similar. However if it would not be the case this may introduce bias in predictions at longer periods, where the basin effects are most pronounced.

5.5 accounting for uncertainties

Regarding GMPEs, seismic hazard studies incorporate two types of uncertainty: aleatory variability and epistemic uncertainty. By definition, aleatory variability is the natural random variation of observed ground-motions with respect to explanatory variables that cannot be reduced with increasing knowledge about the earthquake process. It is worth to mention that the aleatory variability can not be reduced with respect to a particular model, however in practice it is desirable to reduce the aleatory variability. The aleatory variability is represented by the standard deviation of the PDF for the GMPEs. Epistemic uncertainty is the uncertainty raised from our lack of knowledge about earthquake process. With increased data and knowledge, ideally, the epistemic uncertainty can be reduced to zero. The epistemic uncertainty is represented by means of logic-tree. In logic-tree approach, several GMPEs would be considered, and each equation is assigned a weighting factor that is interpreted as the relative likelihood of that equation being correct. It should be noted that selection of proper ground motion models has stronger influence on the final results than assigning weights to each model. The incorporation of aleatory variability and epistemic uncertainty in EQRM are discussed in sections 5.6 and 5.7 respectively.

Table 5.7: Different techniques for incorporating attenuation variability in the EQRM.

`var_attn_method`	Description
None	No sampling
1	PDF sampling or spawning (see Section 5.6.2)
2	random sampling (see Section 5.6.1)
3	$+2\sigma$ from median
4	$+\sigma$ from median
5	$-\sigma$ from median
6	-2σ from median

5.6 Incorporating aleatory uncertainty

Incorporating uncertainty is a critical component of any PSHA. The inclusion of aleatory uncertainty in the EQRM for attenuation is facilitated by the EQRM controlfile parameter `atten_variability_method`. The aleatory uncertainty is based on estimations of $\sigma_{log(S_a)}$ (see Sections **??** and **??**). Sections 5.6.1 and 5.6.2 describe two of the most commonly used techniques for incorporating attenuation aleatory uncertainty when using the EQRM. Table 5.7 illustrates all of the aleatory uncertainty options that are available with the EQRM.

5.6.1 Random sampling of a response spectral acceleration

This technique involves selecting a single 'random' response spectral acceleration (S_a) from its PDF by:

1. Selecting a random number n_{rand} from the standard normal distribution $N \sim (0, 1)$.
2. Computing the S_a as follows

$$log(A_{S_a}(T_o, r_m, R)) = \mu_{log(S_a)} + n_{rand}\sigma_{log(S_a)}, \tag{5.7}$$

   where $\sigma_{log(S_a)}$ is usually assumed to be σ_a.

Random sampling from a PDF is illustrated in Figure 5.2. It is important to note the following about random selection:

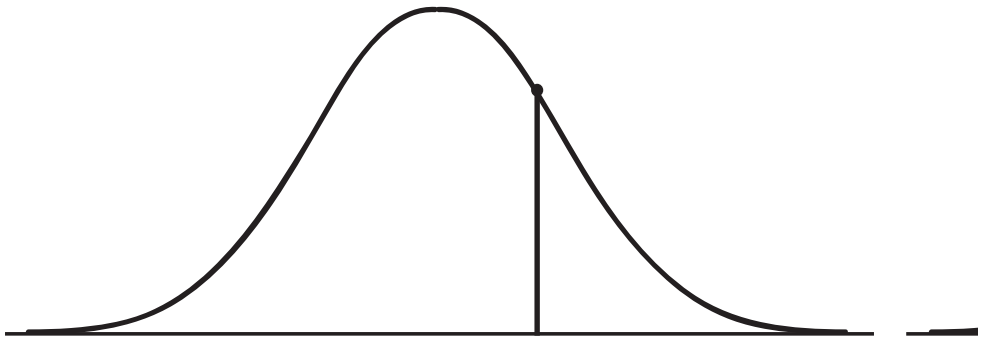


Figure 5.2: Selection of three different random samples from a PDF.

1. Two sites that are close to one another may give dramatically different estimates of ground motion for the same event. Earthquake hazard and risk values that are computed using random sampling exhibit a low level of spatial correlation.

2. The number n_{rand} can theoretically take on any value between $\pm\infty$. Therefore it is possible that estimates of S_a using Equation 5.7 can be unrealistically large. The EQRM overcomes this problem by introducing a scaling factor PGA_{cut} =`atten_pga_scaling_cutoff` which is applied to the RSA as follows:

$$log(A_{S_a,new}(T_o, r_m, R)) = R_{scale} \times log(A_{S_a,old}(T_o, r_m, R)) \tag{5.8}$$

   where

$$R_{scale} = \begin{cases} 1, PGA \leq PGA_{cut} \\ \frac{PGA_{cut}}{log(A_{S_a}(0, r_m, R))}, PGA > PGA_{cut} \end{cases} \tag{5.9}$$

3. There is no effort to account for the likelihood (or probability) of selecting a particular n_{rand} i.e. a particular value of the S_a is not weighted against its likelihood. It can be argued that if enough earthquakes are simulated a range of different n_{rand} will be taken (high and low) and the overall hazard and/or risk values will converge to the true ones.

5.6.2 Sampling the probability density function of the response spectral acceleration (spawning)

An alternative technique for incorporating uncertainty relies on sampling the PDF of the S_a. This technique is used by Aon Re (Mendez, *pers. comm.*, 2003) and is often referred to as spawning. An integral component of sampling the PDF is the spawning of events with the event catalogue (see Section 3.4). Essentially the approach involves taking a user defined number of copies of each event. Every event copy is then available for calculation of hazard (or risk) using a different sample from the attenuation PDF.

1. Firstly the user must define a lower magnitude bound m_{bnd}, a number of samples n_{smples} and a PDF range n_σ (see Section 3.4 for a complete definition). Recall that `fuse_4hzd` redefines the event activity r_ν using a weight w_e, derived by truncating and re-normalising a standard normal distribution to $\pm n_\sigma$.

2. The S_a is computed as follows

$$log(A_{S_a,i}(T_o, r_m, R)) = \mu_{log(S_a)} - \epsilon_i \sigma_{log(S_a)} \quad \text{for } i = 1 \ldots n_{smples} \tag{5.10}$$

where

$$\epsilon_i = -n_\sigma + +(i-1)\Delta \quad \text{for } i = 1 \ldots n_{smples} \tag{5.11}$$

and

$$\Delta = \frac{2n_\sigma}{n_{smples} - 1}. \tag{5.12}$$

Note that Equations 5.10 and 5.11 ensure that the S_a associated with each of the event copies are evenly spread across the domain of the attenuation PDF (Figure 5.3). Recall from Section 3.4 that the event activities r_ν giving rise to each of the S_a described in Equation 5.10 are modified by `fuse_4hzd` as follows:

$$r_{\nu,i} = r_{\nu,original} \times w_{e,i} \tag{5.13}$$

For example, assume that there is an event with $r_\nu = 0.05$ that gives rise to $\mu_{log(S_a)} = [0.3g, 0.5g, 0.2g]$ at the periods $T_o = [0s, 0.3s, 1s]$ respectively. Also assume that the estimates of $\mu_{log(S_a)}$ have the following standard deviations $\sigma_{log(S_a)} = \sigma_a = [0.03, 0.04, 0.01]$. Defining $m_{bnd} = 0$, $n_{smples} = 5$ and $n_\sigma = 2.5$ gives

$$\Delta = \frac{2 \times 2.5}{5 - 1} = 1.25, \tag{5.14}$$

and

i	$w_{e,i}$	ϵ_i	r_ν	$log(A_{S_a,i})$
1	0.0219	-2.5	0.0011	$[0.225g, 0.4g, 0.175g]$
2	0.2285	-1.25	0.0114	$[0.2625g, 0.45g, 0.1875g]$
3	0.4991	0	0.025	$[0.3g, 0.5g, 0.2g]$
4	0.2285	1.25	0.0114	$[0.3375g, 0.55g, 0.2125g]$
5	0.0219	2.5	0.0011	$[0.375g, 0.6g, 0.225g]$

It is important to note the following about spawning:

1. Sub-sampling the PDF provides a smoother (and repeatable) estimate of the S_a than random selection (Section 5.6.1).
2. The range of possible S_a values is bounded by the n_σ.
3. Estimates of all S_a are weighted against an event activity r_ν that accounts for the likelihood of a particular ground motion being observed.

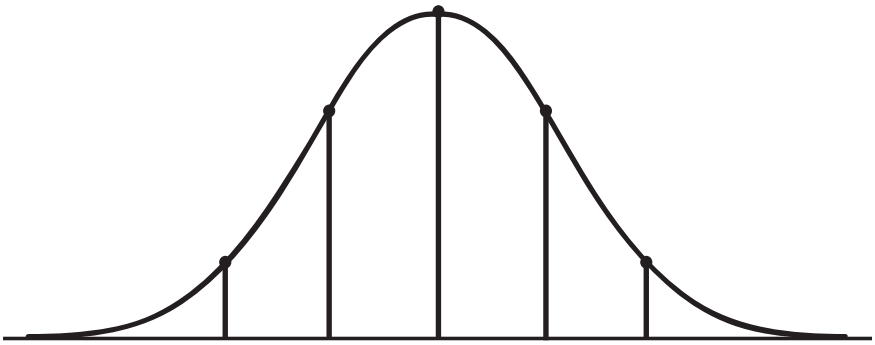


Figure 5.3: Samples drawn from a PDF using the sampling by spawning technique with $n_{smples} = 5$ and $n_\sigma = 2.5$. The width between the vertical bars is Δ. Note that there is one sample for each of the five (spawned) copies of the original event.

4. The nature of spawning means that there are more evaluations of the GMPE. This means that it is typically slower and more memory intesive since the S_a array is larger.

5.6.3 Recomendation for sampling GMPE aleatory Uncertainty

David to complete this section

- use spawning for hazard - exaplin why
- use random sampling for risk - exaplain why

5.7 Using multiple GMPEs - Incorporating epistemic uncertainty

The use of attenuation models in the EQRM is controlled by the `setdata` parameter `attenuation_flag`. Recall from Section **??** that `attenuation_flag` is a two row matrix with one column for each attenuation model to be used. The format of `attenuation_flag` is

$$\begin{bmatrix} P_1 & P_2 & \dots & P_n \\ w_{a,1} & w_{a,2} & \dots & w_{a,n} \end{bmatrix},$$

where the integers $\{P_i\}_{i=1}^n$ are pointers to the attenuation model (see Table **??**) and the real numbers $\{w_{a,i}\}_{i=1}^n$ are weights for the respective attenuation models. Note that the weights w_i must either be all positive or all negative and

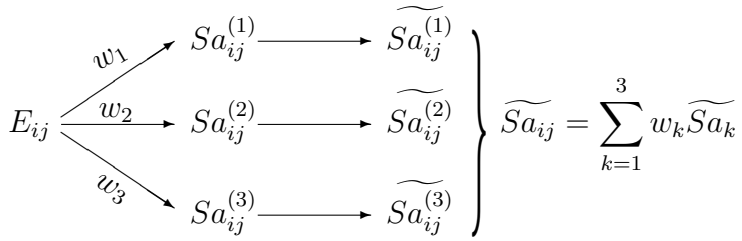
$$\left| \sum_{i=1}^{n} w_{a,i} \right| = 1. \tag{5.15}$$

The mechanism for using more than one attenuation model is similar to the mechanism used for spawning (see Section 5.6.2). That is, the event catalogue is copied n times (once for each attenuation model) and the appropriate attenuation model applied to its respective copy. The values of ground motion (loss for risk) can then be treated independently for the hazard/risk calculation (if $\sum_{i=1}^{n} w_{a,i} = -1$) or aggregated before the hazard/risk assessment is undertaken (if $\sum_{i=1}^{n} w_{a,i} = 1$). This notion of logic tree collapse is described by Figure 5.4.

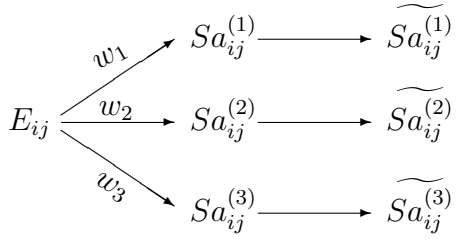
5.8 Collapse versus no-collapse

The techniques adopted for incorporating spawning (see Section 5.6.2) and using multiple ground motion models are similar in the sense that they require multiple evaluations of ground motion for each event. These are carried through the various stages of an EQRM simulation such as amplification, building damage and loss calculation (see following chapters). At the end of the simulation these values can then be collapsed back to a single 'best estimate' or kept in a un-collapsed form and used to determine the hazard and/or risk curve (see Chapter 9). These two options are illustrated in Figure 5.4 for both a hazrd and risk simulation. Note that when collapsing the samples the extreme values are removed from the hazard or risk curves. For this reason we recomend the no-collapse option.

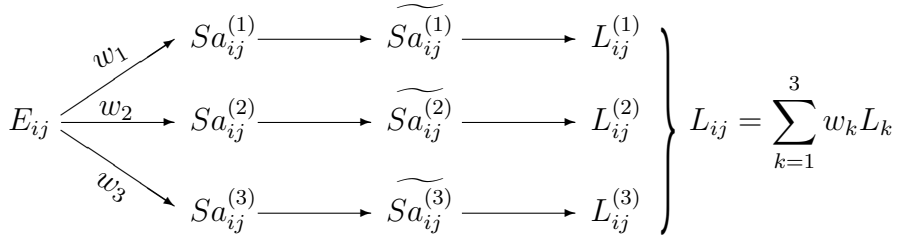
(a)



(b)



(c)



(d)

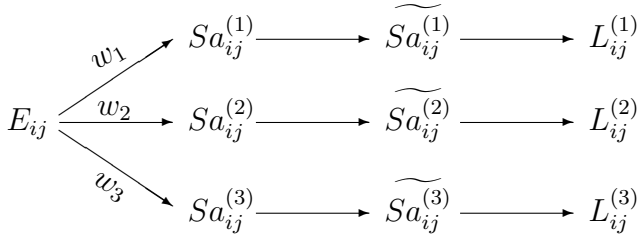


Figure 5.4: The application of spawning to a single synthetic earthquake E_{ij};(a) the collapse of spawned samples for hazard (`src_eps_switch`= 1); (b) hazard spawning without collapse (`src_eps_switch`= 2); (c) the collapse of spawned samples for risk (`src_eps_switch`= 1); and (d) risk spawning without collapse (`src_eps_switch`= 2). The illustrated procedure is repeated for N_s events in the event catalogue and the techniques of Chapter 9 used to assess the hazard or risk.

Chapter 6

Regolith amplification

6.1 Overview

Regolith is defined as the soil, geological sediments and weathered rock that overly the un-weathered bedrock. It is well documented that the presence of regolith can increase the level of ground shaking experienced during an earthquake (**?**; **?**). For example, studies in the San Fernando Valley and Los Angeles Basin, U.S.A., have demonstrated that the damage patterns observed during the 1994 Northridge, California earthquake can be strongly correlated to site-response of local regolith (**?**). Consequently, including the effect of regolith on earthquake ground shaking is an important component of any seismic hazard or risk analysis.

6.1.1 Background theory

An amplification factor can be used to transfer the earthquake motion from the bedrock to the regolith surface. Amplification factors are influenced by the regolith at the site of interest, the magnitude r_m of the event and the PGA for the event-site combination. For simplicity, the geographical region of interest is usually separated into five or six site-classes inside which the regolith is assumed to be the same. The EQRM application does not compute the amplification factors (or level of amplification). Such calculations must be conducted off-line and the results (or amplification factors) made available to the EQRM application. The interested reader is referred to **?** for a detailed description of the equivalent linear technique for computing amplification factors. **?** discuss the classification of site classes and the application of amplification factors to a probabilistic seismic hazard analysis. Figure 6.1 illustrates the geotechnical cross section for all of the

site classes in the Newcastle and Lake Macquarie region, and Figure 6.2 shows an example of the amplification factors use in the same study.

6.2 The amplification factor input file

The amplification factors are saved in a file with a name of the form `<site_loc>_par_ampfactors.mat` where `site_loc` is a variable in `setdata`. For example if `site_loc` = 'newc' then the file is called `newc_par_ampfactors.mat`. The contents of the `<site_loc>_par_ampfactors.mat` include 5 'variable types' that are explained in Table 6.1. Appendix ?? lists the contents for `newc_par_ampfactors.mat` as used in the Newcastle and Lake Macquarie study (?).

6.3 Implementation

As with the attenuation, the implementation of the amplification factors is conducted within `do_analysis` (for both hazard and risk assessments) in a two step process as follows:

1. The preparation of the amplification factors (`prep_amps`) conducted before entering a loop over sites.
2. Application of the amplification factors to compute $A_{S_a,soil}(T_o, r_m, R)$ at each site (`do_amplification`) conducted within a loop over sites.

The preparation of amplification factors involves interpolating each to the periods defined in the `setdata` parameter `periods` (see Section ??) and is conducted by the function `prep_amps`. The application of the amplification factors is carried out immediately after the evaluation of the attenuation model and is implemented within `do_amplification`. The process is described for a specific site located within a known Site Class:

1. Bin all of the events in the event catalogue into the bedrock PGA and moment magnitude bins defined by the bin centroids in `pga_bin` and `momag_bin` respectively. Note that the end points of the 'central' bins are assumed to be half way between the bin centroids. The first and last bins extend to negative and positive infinity respectively.

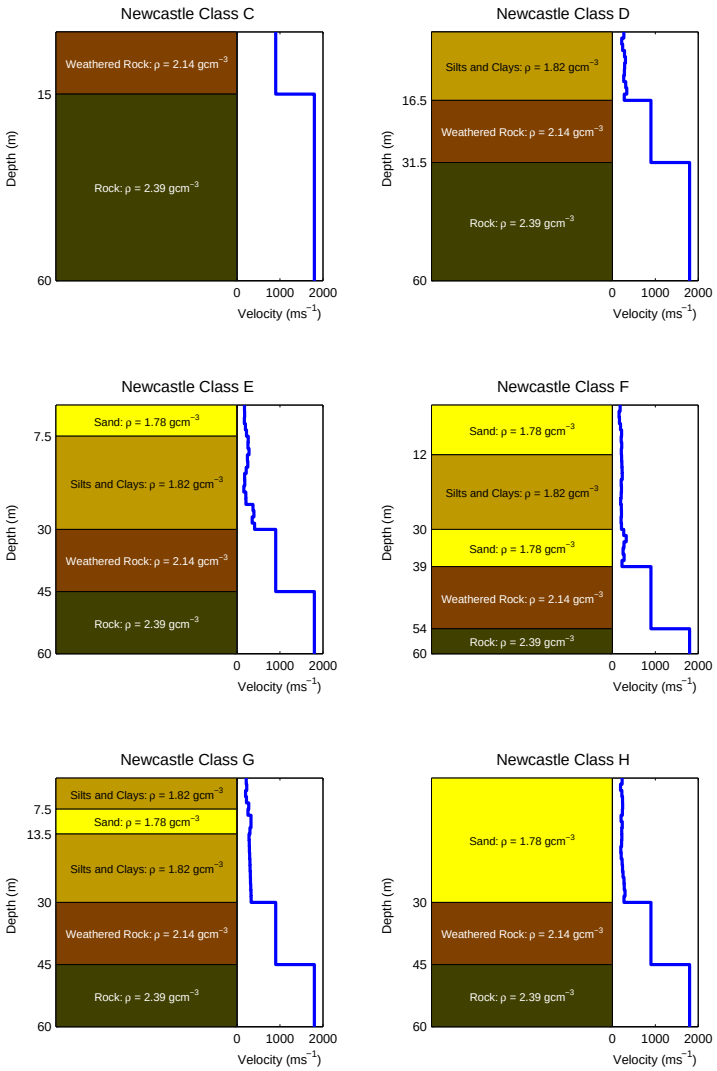


Figure 6.1: Cross sections of the 6 site classes used in the Newcastle and Lake Macquarie study (**?**).

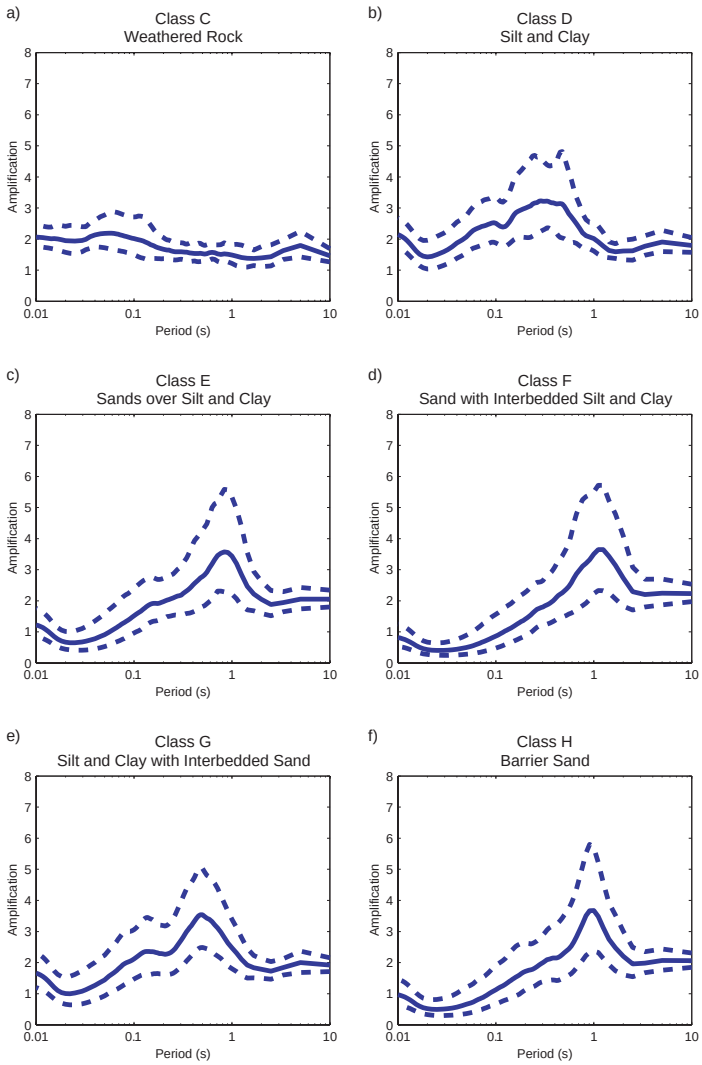


Figure 6.2: Example amplification factors for the site classes shown in Figure 6.1. Note that the amplification factors are for moment magnitude 5.5 and PGA 0.25g.

Table 6.1: Summary of the variables in `<site_loc>_par_ampfactors.mat`. The {1}, {2} and {3} represent input text strings that are pointers to the site class, moment magnitude bin and bedrock PGA bin respectively, for both the amplification factors and their standard deviation. The `<site_loc>_par_ampfactors.mat` file usually contains multiple 'copies' of `ln_site{1}_rockpga{2}_momag{3}` and `stdln_site{1}_rockpga{2}_momag{3}`, whereas it should only contain a single 'copy' of the other variables. All variables are row vectors. The superscript '$*$' shown with some variables refers to the fact that these variables have been 'binned'.

Variable	Description
`amp_period`	Fundamental periods at which the amplification factors (and their standard deviations) are defined.
`ln_site{1}_rockpga{2}_momag{3}`	The mean of the logarithm of the amplification factors $\mu_{amp}(T_o, r_m^*, PGA^*)$(at the periods defined in `amp_period`) for the site class {1}, bedrock PGA bin {2} and moment magnitude bin {3}
`momag_bin`	Bin centroids of moment magnitude bins
`pga_bin`	Bin centroids of bedrock PGA bins
`stdln_site{1}_rockpga{2}_momag{3}`	The standard deviation of the logarithm of the amplification factors $\sigma_{amp}(T_o, r_m^*, PGA^*)$(at the periods defined in `amp_period`) for the site class {1}, bedrock PGA bin {2} and moment magnitude bin {3}

2. Use a nested loop over all PGA bins (upper loop) and moment magnitude bins (lower loop) and apply the amplification factor as follows:

$$S_{a,soil}(T_o, r_m, R) = F_{amp}(T_o, r_m^*, PGA^*) \times S_{a,bedrock}(T_o, r_m, R), \qquad (6.1)$$

where r_m^* and PGA^* are the centroids of the bins containing r_m and $S_{a,bedrock}(0, r_m, R)$ respectively. The F_{amp} represents the exponential of 'some selection' of amplification factor from the distribution $N \sim (\mu_{log(F)}(T_o, r_m^*, PGA^*), \sigma_{log(F)}(T_o, r_m^*, PGA^*))$. The actual selection made will depend on the method used to incorporate uncertainty (see Section 6.4).

Notice that the use of $N \sim (\mu_{log(F)}, \sigma_{log(F)})$ equates to the assumption that the amplification factors are log-normally distributed.

6.4 Incorporating aleatory uncertainty

In theory, both the random selection and PDF sampling methods described in Section 5.6 could be used to incorporate aleatory uncertainty in amplification factors. Some care will need to be taken if the PDF sampling technique is to be used for incorporating uncertainty at several levels of computation (e.g. for attenuation and amplification) due to the need to spawn the event catalogue (see Section 3.4) at every level. Such iterative spawning may lead to a 'blow-out' in the size of the event catalogue and hence the computation time. sCurrently only the random selection technique is available for use with the amplification factors in the EQRM application.

As with the application of an attenuation model, the EQRM applies `pga_cutoff` to $S_{a,soil}(T_o, r_m, R)$ after applying the amplification factor. This accounts for any unrealistically high selection of amplification factor. A secondary upper cut-off facility exists and involves the use of the `setdata` parameter `MaxAmpFactor`. The `MaxAmpFactor` parameter is applied directly to the amplification factor as follows:

$$F_{new}(T_o, r_m^*, PGA^*) = \begin{cases} F_{old} & \forall \;\; T_o \text{ s.t. } F_{amp,old} < \texttt{MaxAmpFactor} \\ \texttt{MaxAmpFactor} & \forall \;\; T_o \text{ s.t. } F_{amp,old} \geq \texttt{MaxAmpFactor} \end{cases} \tag{6.2}$$

or brevity the functional reference to r_m^* and PGA^* in F_{new} and F_{old} in Equations 6.2 and 6.3 refers to the fact that these values represent the bin centroids of `momag_bin` and `pga_bin` respectively. In most simulations the secondary cut-off is not used since `pga_cutoff` handles scaling for an upper bound. Ignoring `MaxAmpFactor` is achieved by setting it to to a very high number (e.g. `MaxAmpFactor`=10000). In contrast, there is another parameter in `setdata` known as `MinAmpFactor` which is commonly utilised within the EQRM application. This parameter stops the selection of any unrealistically low selections of the amplification factor. The `MinAmpFactor` is commonly set to around 0.6. It is applied as follows:

$$F_{new}(T_o, r_m^*, PGA^*) = \begin{cases} F_{old} & \forall \;\; T_o \text{ s.t. } F_{amp,old} > \texttt{MinAmpFactor} \\ \texttt{MinAmpFactor} & \forall \;\; T_o \text{ s.t. } F_{amp,old} \leq \texttt{MinAmpFactor} \end{cases} \tag{6.3}$$

Note that a value of $F < 1$ leads to a de-amplification of the bedrock motion.

6.5 Key functions, flags and parameters

Name	Type	Description
`do_analysis`	Function	The main driving function for modelling earthquake hazard or risk.
`prep_amps`	Function	Loads and prepare the amplification factors.
`do_amplification`	Function	Applies the amplification factors and compute `SAsoil`.
`pga_cutoff`	Function	Re-scales the S_a such that PGA $\leq$ `pgacutoff`
`periods`	Par	Periods at which to compute S_a.
`MaxAmpFactor`	Par	Maximum cutoff for all amplification factors (seldom used).
`MinAmpFactor`	Par	Minimum cutoff for all amplification factors (usually 0.6).
`pgacutoff`	Par	The cutoff (or maximum) PGA for use with `pga_cutoff`.

Chapter 7

Building damage

7.1 Introduction

This chapter describes the modelling of building damage. The method is based on the HAZUS methodology (**?**) with a few modifications to suit the Monte-Carlo approach used in the EQRM. The modifications are discussed at the end of this chapter.

The Capacity Spectrum Method (**?**; **?**) is used to obtain the peak building displacement and acceleration corresponding to each earthquake. The peak displacement and acceleration are the dependent variables for fragility curves (**?**) which give probabilities of being in certain damage states, for different types of damage, structural (damage to main structure) and drift-sensitive and acceleration-sensitive non-structural damage (e.g. damage to non-structural internal walls).

The main MATLAB function to calculate building damage is `do_analysis` which reads the earthquake catalogue, applies the attenuation and soil amplification, models the building damage and computes direct financial loss.

The MATLAB function `prep_sites` loads the buildings database file (Chapter 4) and `prep_build_vuln` reads the building parameters file defining engineering parameters. The MATLAB function `do_build_vuln` is a wrapper, called inside the main loop (over sites). The main work is done by `make_build_dam_rand` which calculates the peak values of spectral displacement and acceleration as the intersection of the building capacity curve with the suitably damped demand curve for the vector of all spectral accelerations from each synthetic earthquake.

7.2 The capacity spectrum method

The capacity spectrum method is used to find the peak displacement and peak acceleration for each earthquake-building pair. To use this technique each building is approximated by a single degree of freedom oscillator (SDOF) (**?**).

A building's capacity (or response) to an incoming seismic wave, is defined by the capacity curve, a monotonic increasing curve giving the S_a as a function of the spectral displacement S_d. The response spectral acceleration (S_a) normally plotted against building period is plotted against spectral displacement S_d to describe the 'earthquake motion'. When plotted this way the $S_a vs S_d$ plot is known as the demand curve. The peak displacement and acceleration experienced by the building as a result of the 'induced motion' is modelled by the intersection of the capacity curve with the demand curve (Figure 7.1).

The response spectral acceleration curve is normally defined by the attenuation model with 5% damping. The actual damping generated by a building differs from 5% and is a function of the building attributes such as structural type (see Chapter 4). The capacity spectrum method allows for this variation by adjusting the damping (see Section 7.2.2).

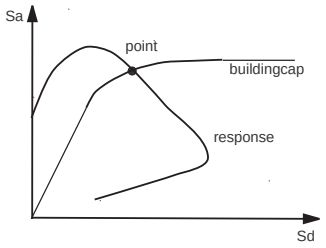


Figure 7.1: The capacity spectrum method.

Figure 7.1 illustrates the intersection of the capacity curve with the demand curve. In the elastic region (or linear component) of the capacity curve the intersection between the capacity curve and the appropriately damped demand curve occurs at a period corresponding to the natural period of the building. If the intersection occurs in the nonlinear plastic deformation region, hysteretic damping results in a larger component of the earthquakes energy being absorbed through damping

in the building. In an equivalent linear approach, this effectively reduces the natural period of the building.

7.2.1 The building capacity curve

The building capacity curve is defined by two points: the yield point (D_y, A_y) and the ultimate point (D_u, A_u). For displacements below the yield point the building response is elastic and the capacity curve represented by a straight line. For displacement amplitudes greater than the yield point nonlinear effects such as plastic deformation cause the rate of increase to reduce. The curve asymptotes towards the ultimate point (see Figure 7.2).

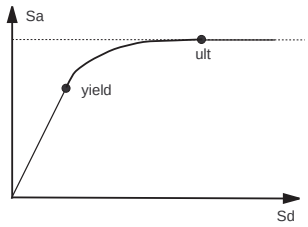


Figure 7.2: The graph of a typical building capacity curve, defined by the yield point (D_y, A_y) and the ultimate point (D_u, A_u). Note that the straight line between the yield point and the origin represents elastic behavior of the building.

The yield point and ultimate point are defined in terms of the building parameters as:

$$\begin{aligned} A_y &= \tfrac{C_s \gamma}{\alpha_1} \quad [g], \\ D_y &= \tfrac{1000}{4\pi^2} 9.8 A_y T_e^2 \quad [mm], \\ A_u &= \lambda A_y \quad [g], \\ D_u &= \lambda \mu D_y \quad [mm], \end{aligned}$$

where the building parameters

C_s = design strength coefficient (fraction of the building weight),
T_e = natural elastic building period (seconds),
α_1 = fraction of building weight participating in the first mode,
α_2 = fraction of the effective building height to building displacement,
γ = over-strength factor—yield to design strength ratio,
λ = over-strength factor—ultimate to yield strength ratio,
μ = ductility factor,

are given for several classes of building construction types (Section 4.3.1). Note that the parameter α_2 is not used here, rather it is used to define the appropriate damage ratios (see Equation 7.9).

Fitting the building capacity curve

The capacity curve is composed of three parts: a straight line to the yield point, a curved part from the yield point to the ultimate point, and a horizontal line from the ultimate point.

An exponential function is used to represent the curved part of the building capacity curve. When defining the curved section it is desirable that it has and identical slope to the elastic part at the yield point and that its slope approaches zeros at the ultimate point. We cannot satisfy all these conditions with just three constants, however, we can satisfy the condition of the curves matching at the ultimate point approximately.

The form of the exponential curve is

$$y = c + ae^{-bx}$$

where the constants a, b, and c are given by

$$c = A_u, \quad b = \frac{k}{A_u - A_y}, \quad a = (A_y - A_u)e^{bD_y} \tag{7.1}$$

where $k = A_y/D_y$.

To obtain this, the following equation follows from the curves joining at the ultimate points

$$c + ae^{-bD_u} = A_u.$$

If we neglect the e^{-bD_u} term, then $c = A_u$. From the equation for the yield point we then obtain $A_u - A_y = -ae^{-bD_y}$. From the equation that the slopes match at the yield point we obtain $k = -abe^{-bD_y}$, where $k = A_y/D_y$ is the slope of the elastic linear part of the capacity curve. Eliminating the exponential term from both these equations gives the value of b, and the expression for a follows. The condition that the slope is zero at the ultimate point is consistent with the assumption of neglecting the term e^{-bD_u}. This will be true provided

$$\frac{D_u/D_y}{A_u/A_y - 1} \gg 1.$$

which will be generally true when the ultimate point is far from the yield point.

Variability of the capacity curves

The uncertainty of the peak response of the building to a given ground-shaking is incorporated by using a log-normal distribution of building capacity curves with a log-normal standard deviation parameter $\beta = 0.3$. This value is taken from **?** for pre-code buildings with a value of 0.25 recommended for buildings constructed in alignment with building codes. HAZUS gives no explanation or reference to how these values were obtained. It was recommended in the Engineers' workshop, **?**, that a value $\beta = 0.4$ be used for Australian buildings, but there was no justification given for this, so a value of $\beta = 0.3$ has been retained.

The variability of the building capacity curve is described by a log-normal distribution for the vertical component of the ultimate point. That is, $A_u = \bar{A}_u \times e^{\beta\phi}$, where $\bar{A}_u$ is the median value of A_u and ϕ is a random number selected from a standardised normal distribution (see `bcap_rand`).This sampling is analogous to the random sampling technique used elsewhere in the EQRM (e.g. Section 5.6.1) because the median of a log-normally distributed random variable (e.g. $barA_u$) is the same as the exponential of the mean of the log (e.g. $exp(\mu_{ln(A)})$). This is illustrated in Figure 7.3. The yield point (D_y, A_y) on the random capacity curve is chosen from the equations $D_u = \lambda\mu D_y$, with $D_u = \bar{D}_u$, and $D_y = 9.8A_yT_e^2$.

7.2.2 Damping the demand curve

The Response Spectral Acceleration curve (S_a), as obtained from an attenuation formula, is specified for 5% damping. Recall that S_a describes the response of an idealised (SDOF) building. The response of the actual building is incorporated by modifying the damping. This is undertaken in the following two parts:

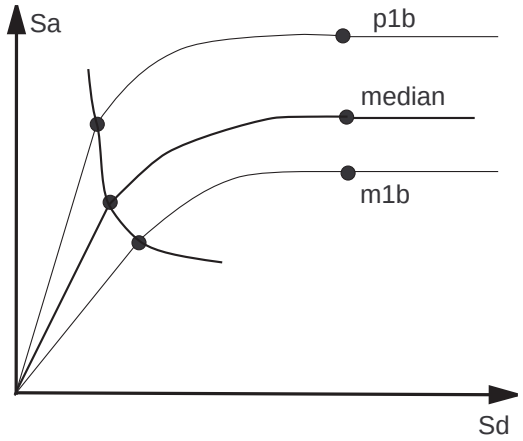


Figure 7.3: Illustration of the random selection of a building capacity curve from a log-normal distribution of building capacity curves. The median, 84th percentile ($+\beta$) and 16th percentile ($-\beta$) are shown.

1. modification of elastic damping, and
2. hysteretic damping.

Modification of elastic damping

The HAZUS approach uses 5% damping for all buildings. However, the modifications suggested at the Australian buildings workshop (**?**) suggest using elastic damping values higher than 5% with corresponding hysteretic damping coefficients made smaller. The values of the elastic damping have been determined from **?**.

Damping formulae are the same as used in **?**, which are obtained from **?**. These are empirically derived formulae, as multiplicative factors, defined for three regimes

according to building period. The three regimes correspond to the acceleration-sensitive, velocity-sensitive and displacement-sensitive areas of the response spectrum, denoted R_a, R_v and R_d respectively. The effect of elastic damping on the demand curve is illustrated in Figure 7.4.

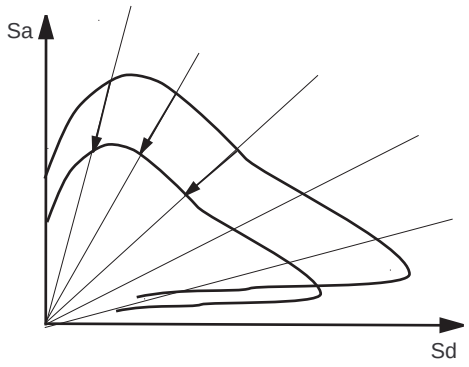


Figure 7.4: Damping of the demand curve.

The damping formula are:

$$R_a = \frac{2.12}{3.21 - 0.68\ln(100B_T)}, \quad 0 \le T < T_{av} \tag{7.2}$$

$$R_v = \frac{1.65}{3.21 - 0.68\ln(100B_T)}, \quad T_{av} \le T < T_{vd} \tag{7.3}$$

$$R_d = \frac{1.39}{1.82 - 0.27\ln(100B_T)}, \quad T \ge T_{vd} \tag{7.4}$$

where B_T is the effective damping, expressed as a decimal (not a percentage). Note that when $B_T = 0.05$, (5% damping) then $R_a = R_v = R_d = 1$.

In the code the flag `damp_flags(1)=1` can be used to set $R_d = R_v$, (HAZUS doesn't use R_d), or `damp_flags(1)=2` sets R_a and R_d to R_v. This might be used used if there are problems with convergence. The default value `damp_flags(1)=0` uses all 3 factors.

The transition building periods (corner periods), T_{av} and T_{vd}, are given (according to HAZUS) as

$$T_{av} = \frac{S_a(T_o = 1.0)}{S_a(T_o = 0.3)}, \qquad T_{vd} = 10^{(r_m - 5)/2} \tag{7.5}$$

where r_m is the moment magnitude. These are called corner periods because, in an 'ideal' tripartite plot (i.e. S_v vs building period, in log-log space), these periods correspond to corners (**?**).

HAZUS also modifies T_{av} to $T_{av\beta}$ where

$$T_{av\beta} = \frac{R_a}{R_v} T_{av}.$$

It is not clear to the authors why this is needed. According to discussions with Mark Edwards (*pers. comm.*, 2002) it is needed to account for the damping that has already occurred. However, since the damping ratios are applied to the original 5% spectra, at each stage of the iteration, it should not be needed.

In the code, the flag `damp_flags(2)=0` (default value) is used to modify T_{av} to $T_{av\beta}$, at each iteration, whereas `damp_flags(2)=1` does not do the modification. The corner periods are calculated in the file `calc_corner_periods`.

Note that the formulae for the corner periods are chosen (by HAZUS) to be consistent with the artificial standardised demand curve that HAZUS uses. Therefore they are not necessarily appropriate for response curves that are computed from attenuation formulae. It is believed that they will be a reasonable approximation. However, because our response spectra are not consistent with the corner periods there may be some problems with discontinuities leading to convergence problems with the iterative approach used to deal with inelastic behavior (see Section 7.2.2). To try and minimise this a small amount of smoothing of the demand curve is applied near the corner periods.

For future work, it may be worth investigating using damping formula which vary continuously as a function of period. Also, in **?** formula are given for a variability component of the damping. A random damping factor has not yet been implemented. It would not be difficult, but it first needs to be determined whether it is already taken into account by the random capacity curves.

Tables of the elastic damping parameter values, for different building construction types, are given in the Appendix **??**; Table **??**.

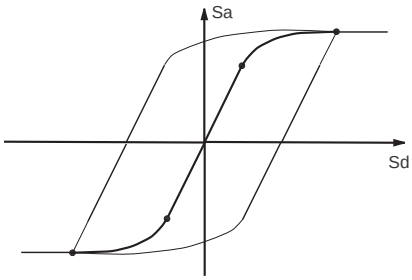


Figure 7.5: Hysteresis area.

.

Hysteretic damping

When the intersection point occurs in the inelastic region the damping applied to the response spectral acceleration also has an additional component due to hysteresis. This hysteretic damping accounts for the fact that the buildings ability to absorb energy changes after it has been pushed into the inelastic region. The effective damping term B_T has an elastic component B_E and a component due to hysteretic damping B_H . This latter component is zero when the intersection point is in the elastic region. The hysteretic damping term is determined from the area enclosed by the hysteresis curve A_H, as shown in Figure 7.5. The effective damping is given by

$$B_T = B_E + B_H, \tag{7.6}$$

where

$$B_H = \frac{A_H}{2\pi DA} \tag{7.7}$$

and B_E is defined by the process described in Section 7.2.2.

Recall that the curved part of the hysteresis boundary is approximated by

$$y = ae^{-bx} + c,$$

where

$$c = A_u, \qquad b = \frac{k}{A_u - A_y}, \qquad a = (A_y - A_u)e^{bD_y},$$

and the slope of the elastic part is,

$$k = \frac{A_y}{D_y}.$$

The area is calculated by subdividing into three separate areas (see Figure 7.6)

$$A_H = 2(A_1 - A_2 + A_3).$$

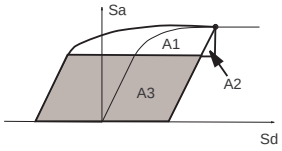


Figure 7.6: Diagram showing the sub-areas used for the hysteresis area calculation.

We also define the coordinates (see Figure 7.7)

$$x_2 = D - \frac{A}{k}, \qquad x_1 = 2x_2 + \frac{y_1}{k}, \qquad y_1 = A_u - A_y.$$

The sub-areas are calculated as

$$\begin{aligned} A_1 &= cx_1 + \frac{a}{b}(1 - e^{-bx_1}) \\ A_2 &= \frac{y_1^2}{2k} \\ A_3 &= 2A_y(D - D_y). \end{aligned}$$

The calculation of the hysteretic area is implemented in the MATLAB function `hyst_area_rand`. Note that this function has arguments to deal with a randomly sampled capacity curve. Currently the code calculates hysteretic damping for all

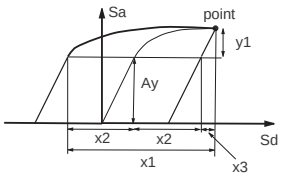


Figure 7.7: Coordinates used for hysteresis area calculation.

events including those in the elastic region (where the hysteretic damping is zero), because the calculations are vectorised. Furthermore the code iterates until the most nonlinear event has converged or the maximum iteration limit is reached. Some efficiencies might be gained by first filtering out those events in the elastic region, and then later filtering out those events which have converged.

Tables of the hysteretic damping coefficients κ_L, κ_M and κ_L are given in Appendix **??**. Table **??** provides the hysteretic damping coefficients κ_L, κ_M and κ_L corresponding to parameters used in the Newcastle and Lake Macquarie study (**?**).

7.2.3 Finding the intersection point

The algorithm used to find the intersection point of the suitably damped demand curve with the building capacity curve is quite simple. However, it is vectorised over all the synthetic earthquakes. The intersection point is generally not found exactly, rather it is approximated.

The algorithm can be summarised, as follows:

1. Locate points on demand curve and building capacity curve which bound the intersection point, as shown in Figure 7.8.
2. Use linear interpolation to find a first approximation to the intersection point.

3. Use a vertical step to force the intersection to lie on the building capacity curve.

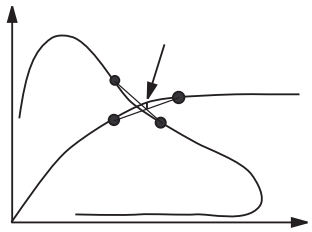


Figure 7.8: Illustration of the algorithm used to approximate the intersection point of the demand SA curve and the building capacity curve.

It is important that the approximation to the intersection point lies on the building capacity curve for the calculation of of the hysteretic area.

Further refinements could be made to improve the accuracy of the location of the intersection point. These might involve iterative procedures which use the tangent of the capacity curve intersecting the line between the two points on the demand curve, to iteratively spiral into the true intersection point. However, a more sophisticated scheme, such as this, may be difficult to code since it would have to account for all possible variations in the form of the demand curve (after soil amplification and damping) and might not always converge. Therefore further refinements are probably not warranted since the current scheme appears to be sufficiently accurate.

7.3 Fragility curves

Fragility curves give the cumulative probability of a particular building being in or exceeding a given damage state given a seismic demand parameter, such as building peak displacement or acceleration. There are separate fragility curves for each of the 4 cumulative damage states: Slight, Moderate, Extensive and Complete and the 3 types of damage: (a) Structural Damage (based on peak

displacement), (b) Non-structural damage-drift sensitive (also based on peak displacement) and (c) Non-structural damage-acceleration sensitive (based on peak acceleration).

?, page 5-12, Table 5.2 provides a table showing typical non-structural components of buildings as drift sensitive or acceleration sensitive. **?**, page 5-13 to 5-23 also provide qualitative descriptions of what the damage states for structural and non-structural damage correspond to. For example *slight structural damage* for wood framed construction corresponds to small cracks in door and window openings and *complete structural damage* corresponds to an immediate danger of structure collapse.

In **?**, pages 5-19 it is stated that non-structural (acceleration sensitive) damage to components at or near ground level may be better characterised by peak ground acceleration (PGA) rather than peak spectral acceleration S_a^*. To this end, HAZUS suggests a combined use of PGA and S_a^* to model damage for near ground components. Currently the EQRM considers only S_a^* when computing damage, however the HAZUS suggestion could easily be incorporated into future version of the EQRM.

7.3.1 Form of fragility curves

Each fragility curve is determined by two parameters: the threshold value of displacement (or acceleration) and a variability parameter. The form of the fragility curve is a cumulative log-normal distribution,

$$F(s_{dam} \| \bar{S}^*, \epsilon) = \ln \left[\Phi \left(\frac{S^*}{\bar{S}^*}; \epsilon \right) \right] \tag{7.8}$$

where s_{dam} refers to the damage state of interest, Φ is the standard cumulative normal distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} \, dt,$$

S^* denotes either peak spectral displacement, S_d^* or peak spectral acceleration, S_a^*, $\bar{S}^*$ is the median damage state threshold value of S^* and ϵ is a variability parameter.

The value ϵ represents the uncertainty of the damage state. It is the square root of the variance of the logarithm of the data (i.e. the log-standard deviation). A zero value for ϵ means the the curve approximates a step function, and so below the threshold value the building is not in the given damage state and above the

threshold it is definitely in the damage state. The larger the value of ϵ the more spread out the curve is, reflecting our less certain knowledge of what state the building is in (Figure 7.3.1).

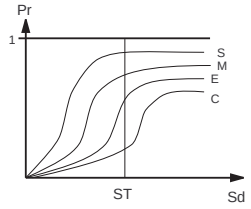


Figure 7.9: A typical fragility curve, giving the cumulative probability of being in or exceeding a certain damage state as a function of the peak displacement (or acceleration).

7.3.2 Damage state thresholds

The damage state thresholds are the median values that determine the damage states.

For structural damage and non-structural drift-sensitive damage the damage state thresholds are determined from provided drift ratios for each building construction type. For non-structural acceleration-sensitive damage the damage state thresholds are obtained from specified acceleration thresholds. These are tabulated for different versions of the engineering parameters in `buildfacts-<version>.xls`, in directory `*/datacvt/buildpars`. For example

$$S_T = \alpha_2 h \delta \tag{7.9}$$

is used to compute the median damage state threshold for structural damage, where δ is the drift ratio, h is the provided height of the building for the given building construction type and α_2 is a building construction parameter corresponding to the fraction of the building height (roof) at the location of push-over mode displacement. The parameters α_2 and h are given in Table **??**. Note that h is given in feet, however, in the code this is converted into mm (since S_d is

considered throughout the code in mm). Tables of drift ratios, and acceleration values, for different sets of building parameters are given in the Appendix **??**.

7.3.3 Variability of the damage states

The uncertainty of a building being in a given damage state for a given peak displacement, is characterised by the variability parameter ϵ in equation Equation 7.8. As this uncertainty becomes smaller, the fragility curve steepens and becomes more like a step function, as shown in Figure 7.10.

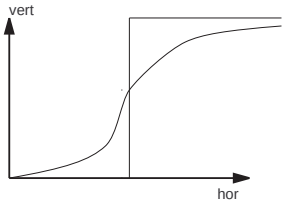


Figure 7.10: As the variability of a fragility curve is reduced the fragility curve steepens and becomes more like a step function.

We follow **?** in setting

- for structural damage $\epsilon = 0.4$,
- for non-structural, drift sensitive damage $\epsilon = 0.5$,
- for non-structural, acceleration sensitive damage, $\epsilon = 0.6$.

Note that HAZUS does not give any references to the derivation of these values.

7.3.4 Incremental probabilities

The fragility curves give the probability of being in the given damage state or lower. That is, they are cumulative probabilities. For our Monte Carlo simulation

approach we need the incremental probabilities of the building being in the given damage state. For example, suppose we have the fragility curve for the extensive damage state, $\Pr(d \leq E)$. To find $\Pr(d = E)$, we calculate

$$\Pr(d = E) = \Pr(d \leq E) - \Pr(d \leq M).$$

This is illustrated in Figure 7.11.

7.4 Differences from HAZUS methodology

A key difference in our implementation of the Capacity Spectrum Method to that used by HAZUS is using the full response spectrum for intersecting with the building capacity curve rather than a design spectra based on only a few building periods. For example, the HAZUS approach only uses periods 0.3 and 1.0 seconds. Our approach has the advantage that the full structure of the response spectrum and all the available information for the soils amplification factors, at all periods, is taken into account rather than at only two periods (Chapters 5 and 6). There are also some minor differences in the way that demand curve damping is applied.

Another difference is in how the fragility curves are used. In HAZUS, the fragility curves incorporate not only the variability of the damage state thresholds, but also the variability of the capacity curve and the ground-shaking. A 'convolution procedure', as described in **?**, is used to convolve the various log-normal probability distributions for ground-shaking and building capacity curve by calculating intersection points for a range of randomly selected demand and capacity curves, then fitting a log-normal distribution. In the EQRM the fragility curves includes only the variability for being in the given damage state and not the variability associated with ground shaking or building type. The variability associated with the fragility curve is defined by a cumulative log-normal distribution with variability parameters 0.4, 0.5 and 0.6 for structural damage, non-structural damage (drift sensitive) and non-structural damage (acceleration sensitive), respectively. The variability associated with ground shaking (Sections 5.6 and 6.4) and building type (Section 7.2.1) are incorporated elsewhere in the Monte-Carlo simulation.

7.4.1 Extra features

The risk component of the EQRM also has a number of extra features or alternative operation modes that are not offered by the HAZUS methodology. Two of the more significant extra features are:

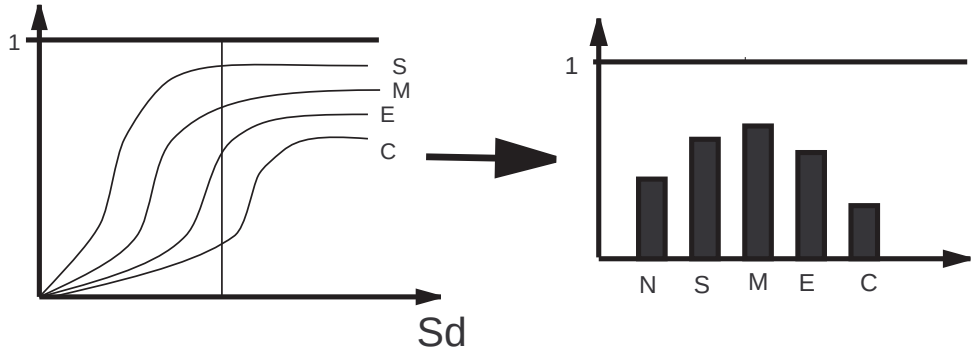


Figure 7.11: Converting the cumulative probabilities in the fragility curves to incremental probabilities describing the probability that the building component is in a given damage state.

1. The use of **uniform hazard spectra** instead of demand curves. This effectively drives the risk calculation by hazard values rather than individual earthquakes. The process is discussed further in **?**. The EQRM can be used in this fashion by setting the `setdata` parameter `run_type` to 3.

2. The use of **Modified Mercali Intensity (MMI) damage curves** rather than the capacity spectrum method. When the user selects $100 \leq$ `attenuation_flag` ≤ 199 the EQRM will use an MMI attenuation model (currently only Gaull is coded) and will compute risk using in-built damage curves that relate percentage damage to MMI. Note that the damage curves are exponential in shape and that the percentage damage estimates represent damage to the entire structure rather than components of the structure. **?** provide an example of the use of damage curves with the EQRM.

7.5 Key functions, flags and parameters

Name	Type	Description
*/datacvt/buildpars	Dir	Directory containing tools to assist the preparation of engineering parameters.
do_analysis	Function	Higher order function of EQRM responsible for computing hazard and/or risk.
prep_locations	Function	Loads and prepares earthquake hazard grids and building databases.
prep_build_vuln	Function	Loads and prepares engineering and financial loss parameters.
do_build_vuln	Function	Models building damage, financial loss and social loss.
make_build_dam_rand	Function	Models building damage.
build_cap_rand	Function	Calculates building capacity curve as a function of spectral displacement S_d.
make_fragility	Function	Calculates the fragility curve and outputs probabilities of being in various damage states.
calc_corner_periods	Function	Calculates the corner periods of the demand curve.
cvt_bfactors	Function	Converts Excel building paramter files to required *.MAT format.
hyst_area_rand	Function	Incorporates hysteretic damping (called by nonlin_damp_rand).
buildfacts _<version>.mat	Par File	Excel version of engineering parameters used by EQRM (e.g. buildfactors.xls or buildfactors-wshop.xls). Note that these Excel files are converted to the EQRM required *.MAT files using cvt_bfactors.
run_type	Par	Defines the operation mode of the EQRM (Section **??**).
damp_flags	Par	Vector to control nature of linear damping (Section **??**).
qa_switch_vun	Par	Controls diagnostics for capacity spectrum method (Section **??**).

Chapter 8

Losses

8.1 Overview

The EQRM considers two types of loss:

1. **direct financial loss** defined as the cost involved in replacing damaged building components and/or contents; and
2. **social loss** defined as the number (or probability) of casualties and injuries as a result of a simulated scenario.

This chapter describes the direct financial and social loss modules.

8.2 Direct financial loss

8.2.1 General financial loss equations: loss for a single building

Recall that the capacity spectrum method assumes that each building comprises three main components, namely a structural, non-structural drift sensitive and non-structural acceleration sensitive component. Section 7 described how the damage experienced by each building is computed separately for each of the components. It is therefore necessary to partition the replacement cost of the building

Table 8.1: Examples of costing splits for the FCB and HAZUS usage classification.

FCB Usage Classification	
111 – W1BVTILE:	
structural	23.44%
non-structural drift sensitive	50.00%
non-structural acceleration sensitive	26.56%
491 – C1MSOFT:	
structural	15.32%
non-structural drift sensitive	34.23%
non-structural acceleration sensitive	50.45%
HAZUS Usage Classification	
RES1 – W1BVTILE:	
structural	23.44%
non-structural drift sensitive	50.00%
non-structural acceleration sensitive	26.56%
COM5 – URMLMETAL:	
structural	13.79%
non-structural drift sensitive	34.48%
non-structural acceleration sensitive	51.72%

into the replacement cost for each of the three components. The proportion chosen for each building component is a function of the buildings construction and usage type as well as the usage classification system. For example Table 8.1 illustrates the proportion of the replacement value corresponding to a couple of different buildings for both the HAZUS and FCB classification system. The `setdata` flag `b_usage_type_flag` can be used to select the usage classification system and consequently the desired cost proportions.

Let $P_{i\alpha}$ denote the probability of being in a damage state $\alpha = (1, 2, 3, 4) = (S, M, E, C)$ corresponding to Slight, Moderate, Extensive and Complete damage, where the index $i = (1, 2, 3) = (s, n_d, n_a)$ corresponds to the type of damage: drift sensitive structural damage (s), drift sensitive non-structural damage (n_d) and acceleration sensitive non-structural damage (n_a). Contents damage is factored in later.

We also define L_i as the financial loss corresponding to the 3 building components

(structural damage, $i = s = 1$; drift-sensitive non-structural damage, $i = n_d = 2$; acceleration-sensitive non-structural damage $i = n_a = 3$; and L_4 as the financial loss due to damage of contents.

Let R_i denote the replacement cost component of the building per unit floor area, for $i = (1, 2, 3) = (s, n_d, n_a)$. Thus $R = R_1 + R_2 + R_3$ is the total replacement cost (per unit floor area) of the building (excluding contents). The financial loss, for a single building, excluding contents, is then given as the weighted sums of the probabilities

$$
\begin{aligned}
L_1 &= C_0 \sum_{\alpha=1}^{4} f_{\alpha,1} R_1 A P_{\alpha,1} = \sum_{\alpha=1}^{4} f_{\alpha,1} g_1 R A P_{\alpha,1}, \\
L_2 &= C_0 \sum_{\alpha=1}^{4} f_{\alpha,2} R_2 A P_{\alpha,2} = \sum_{\alpha=1}^{4} f_{\alpha,2} g_2 R A P_{\alpha,2}, \textit{and} \\
L_3 &= C_0 \sum_{\alpha=1}^{4} f_{\alpha,3} R_3 A P_{\alpha,3} = \sum_{\alpha=1}^{4} f_{\alpha,3} g_3 R A P_{\alpha,3},
\end{aligned}
$$

where A is the floor area of the building (in m^2). Note that R is the replacement cost of the building, $f_{\alpha,i}$ is a repair cost fraction of replacement cost for the given damage state, g_i is the damage component replacement value as a fraction of the replacement value, and C_0 is a regional cost factor. The total loss of the building, excluding contents, is $L = L_1 + L_2 + L_3$.

Note that for percentage loss (loss divided by the value of the building) the quantity $c_0 R$ cancels.

For example, the total repair cost for a building (excluding contents) in damage state $\alpha = 3 = E$ is $f_{3,1} R_1 + f_{3,2} R_2 + f_{3,3} R_3$. The probabilities $P_{\alpha,i}$ correspond to the probability of the building component $i = (s, n_d, n_a)$ being in damage state $\alpha = (S, M, E, C)$.

The regional cost factor, C_0, is a normalising factor to calibrate the replacement costs if necessary. For example in the Newcastle risk assessment (**?**) the HAZUS cost values were used (see Table 8.2) and converted to the Newcastle region using $C_0 = 1.4516$. In this particular case the C_0 was computed by assuming that a $100\,\mathrm{m}^2$ brick veneer residential house (RES1, W1BVTILE) had a replacement cost of AUS$1,000 per m^2 (this value of $1000 having been obtained from NRMA web site for the NSW region). In other studies, such as the Perth Cities case study, the replacement costs (see Table 4.1) were defined for the region and no further correction was required. Note that the parameter C_0 is `ci` in the `setdata` file.

The repair cost factors $f_{\alpha,i}$ are the proportions of the replacement costs (for each building component $i = (1, 2, 3) = (s, n_d, n_a)$ per floor area. For structural damage,

$$f_{\alpha,1} = (2\%, 10\%, 50\%, 100\%), \tag{8.1}$$

for non-structural damage (drift-sensitive),

$$f_{\alpha,2} = (2\%, 10\%, 50\%, 100\%), \tag{8.2}$$

and for non-structural damage (acceleration-sensitive)

$$f_{\alpha,3} = (2\%, 10\%, 30\%, 100\%), \tag{8.3}$$

These values are taken from **?**. For example, the repair cost for the acceleration-sensitive components in the extensive damage state are given by the product $f_{33}R_3$, so that if the replacement cost for complete damage is \$500 per square metre, the repair cost would be $30/100 \times \$500$.

The directory `*/eqrm/datacvt/econsoclosspars` contains tools to assist the preparation of direct financial loss parameters. The files `RcPerWrtBuildCFCB-usageEdwards.xls` and `RcPerWrtBuildCHazususageEdwards.xls` contain the replacement cost component of the building per unit floor area R_i for the FCB and HAZUS usage types respectively. The MATLAB function `per_replace-_cost_wrt_bc_usage` converts these excel files to `*.MAT` that can be used by the EQRM. The EQRM assumes the replacement cost components to be independent of the construction type. There are a few exceptions, which have not been implemented. These are based on the tables given in **?**, in particular, Table 25.2a (page 15-12), Table 13.3 (page 15-14) and Table 15.4 (page 15-15). In principle the tools in `*/eqrm/datacvt/econsoclosspars` allow the replacement cost components to be a function of both usage and construction type. Note that `per_replace_cost_wrt_bc_usage` attributes proportions of the building's total value to its different components. Recall form Table 4.1 that the replacement costs for each building is stored with the building database.

The contents damage, L_4, is based only on the probabilities for acceleration sensitive non-structural damage being in Slight, Moderate, Extensive or Complete states, and on the total contents repair costs R_4 defined by the building database (see Table 4.1). The cost L_4 is then added to the building damage cost L to get the overall loss L^* which includes contents.

As with the building components, the loss for contents damage, is expressed as a weighted sum of probabilities of the acceleration sensitive components for each damage state,

$$L_4 = C_0 \sum_{\alpha=1}^{4} f_{\alpha,4} R_4 P_{\alpha 3}, \tag{8.4}$$

where $P_{\alpha 3}$ is the probability of the acceleration sensitive component of the building being in damage state $\alpha = (1, 2, 3, 4) = (S, M, E, C)$ and $f_{\alpha,4}$ is the repair cost fraction of the replacement value for the contents in damage state $\alpha = (S, M, E, C)$. This factor is expressed as a percentage in Table 8.2 and has been taken from **?**, Table 15.6, page 15-21. Furthermore, for contents damage,

$$f_{\alpha 4} = (1\%, 5\%, 25\%, 50\%). \tag{8.5}$$

These values assume that 50% of the contents in complete damage can be salvaged with similar proportions for lesser damage (**?**). If the `setdata` parameter `aus_contents_flag` is set to 1 the following modifications are made as to suggested by George Walker (Aon Re):

1. HAZUS usage only: The contents value of usage types 1 to 10, 24 and 29 (see Table 4.5) are re-assigned to 60% of R_4.
2. $f_{\alpha 4}$ is re-assigned to $(2\%, 10\%, 50\%, 100\%)$ which assumes that no contents will be salvaged from complete damage of Australian buildings.

8.2.2 Aggregated loss and survey factors

Each building in the database represents a sample from its surrounding area. There is a building survey factor associated with each building, which represents the additional number of buildings that are represented by the modelled building (see Table 4.1). Typically the extra buildings are of the same type and are located in close proximity to the modelled building.

The aggregated loss is therefore the weighted sum of the losses of each building in the database with its corresponding survey factor.

The aggregated financial loss values are saved to a file of form `<site_loc>_db_savedecloss.mat`. If the `setdata` parameter `save_ecloss_flag` = 1 a combined loss for the building and contents is saved; whereas if `save_ecloss_flag` = 2 the losses for the building and contents are saved separately.

8.2.3 Cutoff values

The damage and financial loss models estimate small (but finite) damage for very small ground accelerations. This arises due to the asymptotic nature of the fragility curves. A cutoff value has been implemented in the code to prevent such

small values being calculated. The cutoff is in terms of PGA and is controlled by the `setdata` parameter `pga_mindamage`. For all those events having peak ground acceleration (pga) values smaller than `pga_mindamage` at the building location, the financial loss is set to zero. Typically it is assumed that no damage worth reporting occurs for ground accelerations smaller than `pga_mindamage` $= 0.05g$.

8.3 Social losses

The EQRM includes a module for computing the injuries and casualties associated with a scenario simulation (see Section 3.5). The code for considering social losses associated with probabilistic simulations has not yet been written. The social loss module is activated by setting the `setdata` parameter `save_socloss_flag` to 1. When activated, the social loss module produces a file `<site_loc>_db_savedsocloss.mat` that can be processed by the MATLAB script `newc89plothistsoc`.

For the Newcastle 1989 simulation the results are slightly greater than those observed in the actual earthquake (e.g. median of 45 casualties vs 13 casualties actually recorded). Further investigation of the model is recommended before it is widely used. Firstly, a check on the accuracy of the distribution of the population would be useful. Secondly, some thought should be given to a model which distributes the population randomly. It is expected that this will result in a greater spread of the distribution of injuries.

Chapter 13 of **?** provides a detailed description of the methodology behind social loss calculations.

8.4 Key functions, flags and parameters

Name	Type	Description
*/eqrm/datacvt/econsoclosspars	Dir	Tools to assist the preparation of direct financial loss parameters.
per_replace_cost_wrt_bc_usage	Function	Converts the RcPerWrtBuildCFCB-usageEdwards.xls and RcPerWrt-BuildCHazususageEdwards.xls to the required format for the EQRM.
socloss	Function	Function for calculating injury probabilities.
newc89plothistsoc	Script	MATLAB script to plot injury probabilities (for Newcastle scenario only).
RcPerWrtBuildCFCB usageEdwards.xls	Par File	Replacement cost component of the building per unit floor area for the HAZUS usage classification.
RcPerWrtBuildCHazus usageEdwards.xls	Par File	Replacement cost component of the building per unit floor area for the FCB usage classification.
ci	Par	Regional cost index multiplier.
aus_contents_flag	Par	Contents value for residential buildings and salvageability after complete building damage.
pga_mindamage	Par	minimum PGA(g) below which financial loss is assigned to zero.
save_ecloss_flag	Par	Options for saving direct financial loss data.
save_socloss_flag	Par	Options for saving casualty and injury data.
<site_loc>_db_savedecloss.mat	Output	Output file containing the loss estimates for an EQRM risk simulation.

Table 8.2: Calculated replacement costs (AUD m^2) of building usage types.

	Structural	NS-drift	NS-accel	Contents %	ex-contents	total
RES1	234	500	266	50	1000	1500
RES2	172	266	266	50	703	1055
RES3	172	531	547	50	1250	1875
RES4	172	547	547	50	1266	1898
RES5	234	500	516	50	1250	1875
RES6	219	484	484	50	1187	1781
COM1	234	219	344	100	797	1594
COM2	172	141	219	100	531	1062
COM3	172	359	531	100	1062	2125
COM4	219	375	547	100	1141	2281
COM5	250	625	937	100	1812	3625
COM6	266	656	969	150	1891	4727
COM7	203	484	719	150	1406	3516
COM8	156	562	859	100	1578	3156
COM9	141	406	609	100	1156	2312
COM10	219	62	78	50	359	539
IND1	125	94	578	150	797	1992
IND2	125	94	578	150	797	1992
IND3	125	94	578	150	797	1992
IND4	125	94	578	150	797	1992
IND5	125	94	578	150	797	1992
IND6	125	94	578	100	797	1594
AGR	94	16	94	100	203	406
REL	266	437	641	100	1344	2687
GOV1	187	344	516	100	1047	2094
GOV2	266	594	875	150	1734	4336
ED1	219	562	375	100	1156	2312
ED2	172	937	453	150	1562	3906
RES1-T	199	425	226	50	850	1275
RES1-DB	258	550	292	50	1100	1650

Chapter 9

Hazard and risk results

9.1 Overview

Earlier chapters have described how to generate synthetic earthquakes, propagate (or attenuate) the response spectral acceleration and compute loss. This chapter describes how to 'aggregate' the above information to estimate earthquake hazard and earthquake risk. A number of diagrams are used to demonstrate the common techniques for visualising earthquake hazard and risk.

9.2 Calculating hazard and risk

Consider a random variable Y such that;

- Y refers to the response spectral acceleration S_a at a particular site in the case of hazard, or
- Y indicates the loss, either at a particular simplicity we will assume that Y represents the aggregated loss as a percentage of building stock value.

A common way to represent Y is in terms of a probability of exceedance in one one year. To achieve this we assume that earthquakes occur as a Poisson process (**?**). This means that the earthquake process has no memory; or in other words, the probability of an earthquake today does not depend on whether or

not an earthquake occurred yesterday. Mathematically, this assumption can be represented as follows:

$$\Pr(T, Y \geq y) = 1 - e^{\lambda_Y(y)t}, \tag{9.1}$$

where t is a time interval in years (typically 1 year) and $\lambda_Y(y)$ is the annual exceedance rate. The return period is given by Equation 9.2:

$$R_Y(y) = \frac{1}{\lambda_Y(y)}. \tag{9.2}$$

9.2.1 Computing the annual exceedance rate

Chapter 3 described the Monte-Carlo approach used to generate earthquakes. Chapters 5 to 8 describe how to compute an S_a or loss value for each of the synthetic events. It follows, therefore that there exists a set of N_s numbers $\{Y_i\}_{i=1}^{N_s}$ with corresponding event activities $\{r_{\nu_i}\}_{i=1}^{N_s}$ (see Section 3.2.8), where N_s is the number of simulated events.

The annual exceedance rate $\lambda_Y(y)$ is computed from the event activity through the following process:

1. re-order the values $\{Y_i\}_{i=1}^{n}$ from largest to smallest and re-order the corresponding event activities such that the $Y_i - r_{\nu_i}$ pairs are not separated.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \begin{bmatrix} r_{\nu_1} \\ r_{\nu_2} \\ \vdots \\ r_{\nu_n} \end{bmatrix} \rightarrow \begin{bmatrix} y_{1^*} \\ y_{2^*} \\ \vdots \\ y_{n^*} \end{bmatrix}, \begin{bmatrix} r_{\nu_1^*} \\ r_{\nu_2^*} \\ \vdots \\ r_{\nu_n^*} \end{bmatrix} \tag{9.3}$$

2. evaluate $\lambda_Y(y)$ by computing the cumulative sum of event activities.

$$\begin{bmatrix} \lambda_{Y_{1^*}} \\ \lambda_{Y_{2^*}} \\ \vdots \\ \lambda_{Y_{n^*}} \end{bmatrix} = \begin{bmatrix} r_{\nu_1^*} \\ r_{\nu_1^*} + r_{\nu_2^*} \\ \vdots \\ r_{\nu_1^*} + r_{\nu_2^*} + \ldots + r_{\nu_n^*} \end{bmatrix}, \tag{9.4}$$

The asterisk in Equations 9.3 and 9.4 refer to re-ordered values.

9.3 Earthquake hazard results

The EQRM estimates the earthquake hazard $E_H(R_Y)$ at the return periods R_Y defined in the `setdata` vector `rtrn_per` by interpolating the values $\{\lambda_{Y_i^*}\}_{i=1}^{N_s}$ and $\{Y_i^*\}_{i=1}^{N_s}$ to the values $1/R_Y$ (see Equation 9.2). In the case of earthquake hazard the random variable Y refers to the response spectral acceleration and is therefore a function of building period T_o. The earthquake hazard information is saved in the output file `<site_loc>_db_hzd.mat`. It is important to emphasise that only the hazard information is saved, the information for individual synthetic earthquakes is discarded. This means that the earthquake hazard information can not be de-aggregated in terms of causative events.

The EQRM offers three tools for visualising the earthquake hazard:

1. the earthquake hazard map,
2. the hazard exceedance curve, and
3. the uniform hazard spectra.

9.3.1 Hazard maps

Earthquake hazard maps can be used to illustrate the earthquake hazard across a spatial region. The earthquake hazard $E_H(R_Y)$ is plotted for a specific building period T_o and return period R_Y. Figure 9.1 provides an example of a hazard map for the Newcastle and Lake Macquarie region.

It is necessary to draw a number of hazard maps corresponding to different T_o and R_Y in order to understand the hazard across a spatial region. Traditionally return periods considered for building design correspond to a 2% and/or 10% probability of exceedance within 50 years. An exceedance probability of 10% in 50 years equates to a return period of roughly 500 years and an exceedance probability of 2% in 50 years corresponds to roughly 2500 years.

The EQRM post processing tool `plot_hzd_map` can be used to generate hazard maps. The tool `create_gis_output` can be used to export the data in `<site_loc>_db_hzd.mat` to a convenient file for plotting in GIS.

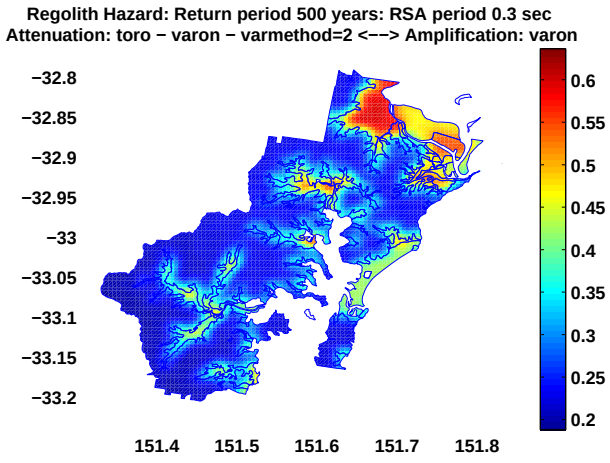


Figure 9.1: Newcastle and Lake Macquarie hazard map for return period of 500 years and S_a period of 0.3 seconds

9.3.2 Hazard exceedance curves

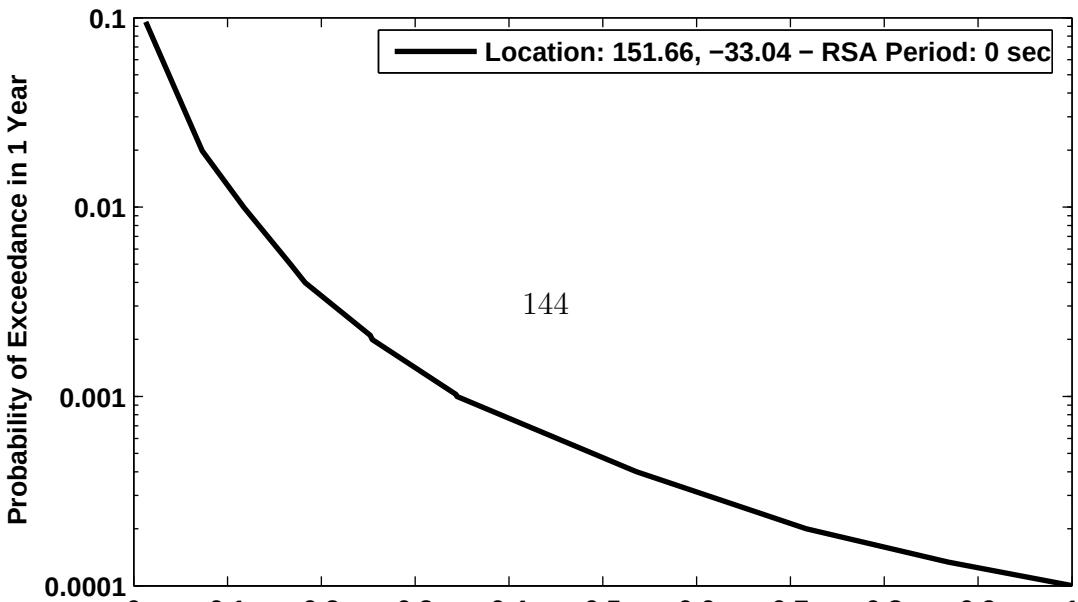
The exceedance probability curve for hazard (also known as the hazard exceedance curve or hazard curve) represents a technique for visualising the earthquake hazard at a single location. The hazard exceedance curve expresses $P(Y \geq y)$ at a single location as a function of the earthquake hazard ($Y = S_a$). Note that unlike the hazard map, the hazard exceedance curve presents the hazard corresponding to the 'full' range of probability levels (return periods). Figure 9.2 provides an example of a hazard exceedance curve for the Newcastle central business district. The EQRM post processing tool `plot_singlesite_hzd` can be used to generate hazard exceedance curves.

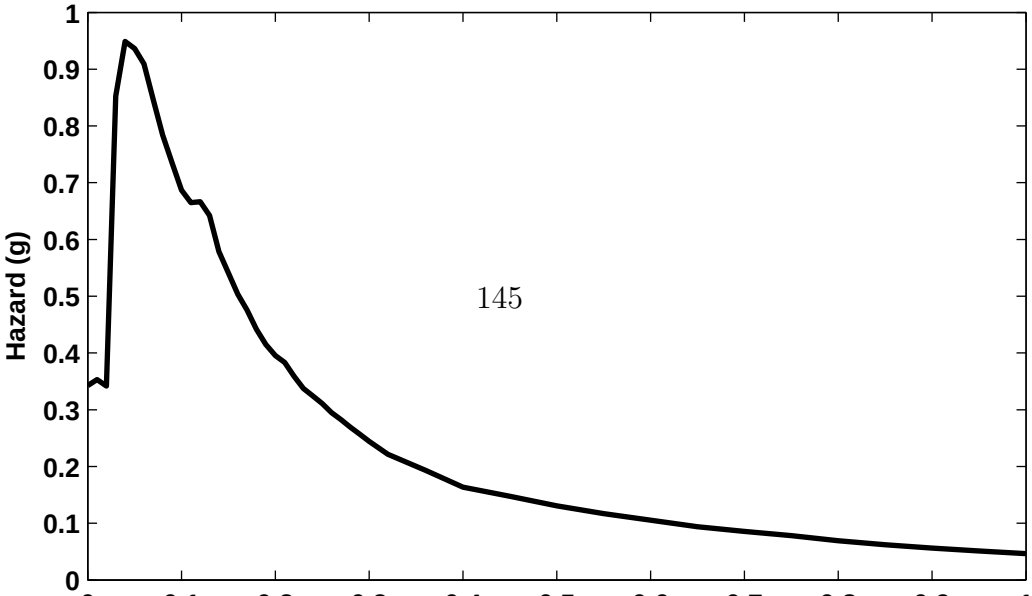
9.3.3 Uniform hazard spectra

The uniform hazard spectra represents another technique for visualising the earthquake hazard at a single location. The hazard for a specified probability level (return period) is plotted for all S_a periods. Figure 9.3 provides an example of an uniform hazard spectra for the Newcastle central business district. The EQRM post processing tool `plot_singlesite_hzd` can also be used to generate uniform hazard spectra.

9.4 Earthquake risk results

The EQRM returns the loss for each earthquake-building pair. This information is saved in the file `<site_loc>_db_savedecloss.mat` along with other information such as building values, the event activity for each synthetic earthquake and the distance between the building-earthquake pair. It is important to note that this technique for data storage is fundamentally different to the technique used for storing the earthquake hazard outputs. Recall that the earthquake hazard outputs store only the earthquake hazard values and not the S_a at each site. Saving the information for each earthquake-building pair provides greater flexibility for risk visualisation and allows the de-aggregation of risk results into a variety of forms. As a result risk output files are typically much larger than hazard outputs and computer memory can become problematic. When the EQRM runs out of memory in a risk simulation the user needs to reduce the number of synthetic earthquakes and/or the number of buildings in the building catalogue. If reducing the number of earthquake-building pairs is not possible the EQRM can be run in a cut-down mode that saves risk results rather than earthquake-building pairs by assigning the `setdata` parameters `save_ecloss_flag` to 3.


Location: 151.66, −33.04 − RSA Period: 0 sec
Probability of Exceedance in 1 Year
0.1
0.01
0.001
0.0001


Hazard (g)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

The EQRM is equipped with tools for visualising earthquake risk in the following forms:

1. the risk exceedance curve (PML),
2. annualised loss, and
3. a variety of disaggregated annualised loss values.

The manner in which the `<site_loc>_db_savedecloss.mat` is processed will vary slightly depending on the plotting tool being used. However, all of the plotting tools make use of the basic process described by Equations 9.3 and 9.4 where the random variable Y describes the loss. A plot manager (or wrapper) `wrap_risk_plots` has been developed to manage the plotting of all risk results. Essentially `wrap_risk_plots` intelligently calls the individual plotting tools by reducing the repetition of common operations.

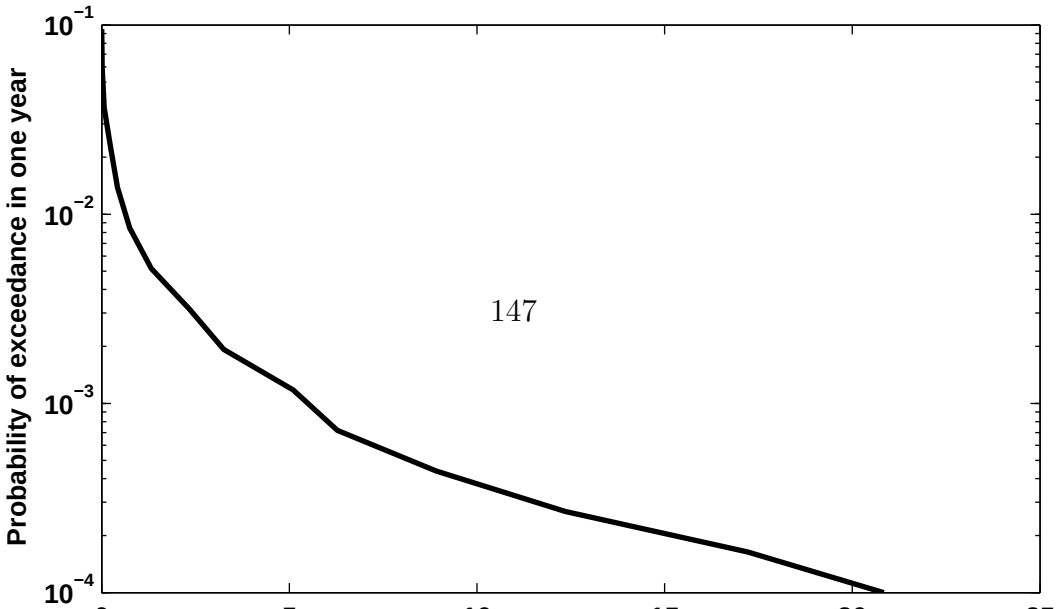
9.4.1 Risk exceedance curve

The exceedance probability curve for risk is analogous to the exceedance probability curve for hazard. The curve represents the maximum loss that is expected to be exceeded for given levels of probability. Unlike the hazard exceedance curve the risk exceedance curve need not be presented for a single site. It is more common to represent the percentage loss for an entire portfolio of buildings, however risk exceedance curves for single buildings can also be constructed. The risk exceedance curve expresses $P(Y \geq y)$ as a function of the earthquake risk $Y = E_L$, where E_L may be the percentage loss to the portfolio or the loss to a single building. Figure 9.4 provides an example of a risk exceedance curve for the Newcastle and Lake Macquarie portfolio. The EQRM post processing tool `wrap_risk_plots` can be used with the individual plotting tool `plot_pml` to plot a risk exceedance curve.

9.4.2 Annualised loss

The annualised loss represents the loss per year averaged over all of the synthetic earthquakes. It can be computed by integrating underneath the risk exceedance curve. Since this is not necessarily intuitive the proof is given below.

Consider the risk exceedance curve as the probability of a loss being exceeded in a one year time frame $P(Y \geq y)$. Let $f_{E_L}(\ell)$ denote the PDF (probability


Probability of exceedance in one year
10^{-1}
10^{-2}
10^{-3}
10^{-4}

density function) of the aggregate loss and $F_{E_L}(\ell)$ denote the CDF (cumulative distribution function). Let us define $F_{E_L}(\ell)$ as the CDF,

$$F_{E_L}(\ell) = \int_0^{\ell} f(x)\, dx.$$

and $G(\ell)$ as the exceedence probability function,

$$G_{E_L}(\ell) = 1 - F_{E_L}(\ell) = \int_{\ell}^{\infty} f_{E_L}(y)\, dy.$$

The mean of the random variable E_L or annualised loss is given by the expectation

$$E(L) = \int_0^{\infty} y f_{E_L}(y)\, dy.$$

Using integration by parts, and using $f_{E_L}(y) = -G'_{E_L}(\ell)$, we get

$$E(L) = \int_0^{\infty} -y G'_{E_L}(y)\, dy = \big[-y G_{E_L}(y) \big]_0^{\infty} + \int_0^{\infty} G_{E_L}(y)\, dy.$$

Assuming $xG_{E_L} \to 0$ as $x \to \infty$ (which is true for an exponential distribution, or distribution which behaves asymptotically as exponential) we have

$$E(L) = \int_0^{\infty} G_{E_L}(y)\, dy$$

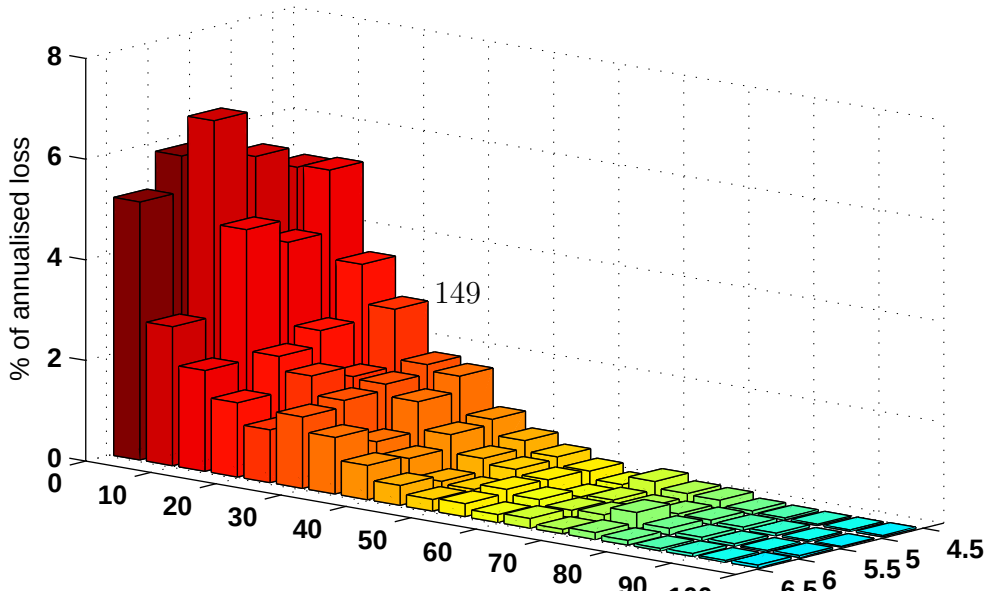
which demonstrates that the mean value of the loss is the area under the risk exceedance curve (PML).

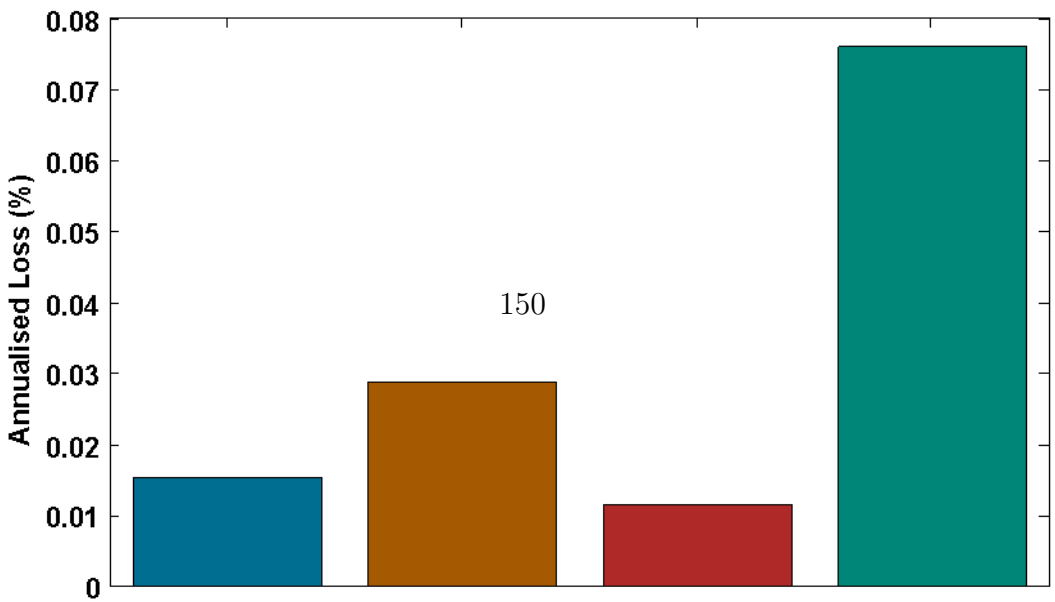
The EQRM post processing tool `wrap_risk_plots` can be used with the individual plotting tool `calc_annloss` to compute the annualised loss.The annualised loss associated with the PML in Figure 9.4 is 0.03%.

9.4.3 Disaggregated annualised loss

Storing the loss results for all of the earthquake-building pairs allows the de-aggregation of annualised loss into a range of classes. The EQRM offers several tools to assist the de-aggregation into common classes all of which are managed through the plotting tool `wrap_risk_plots`. These are described as follows:

1. **disaggregation by distance and magnitude** allows the visualisation of magnitude-distance combinations which have the greatest influence on risk


% of annualised loss
8
6
4
2
0
0
10
20
30
40
50
60
70
80
90
6.5
6
5.5
5
4.5


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Annualised Loss (%)
150

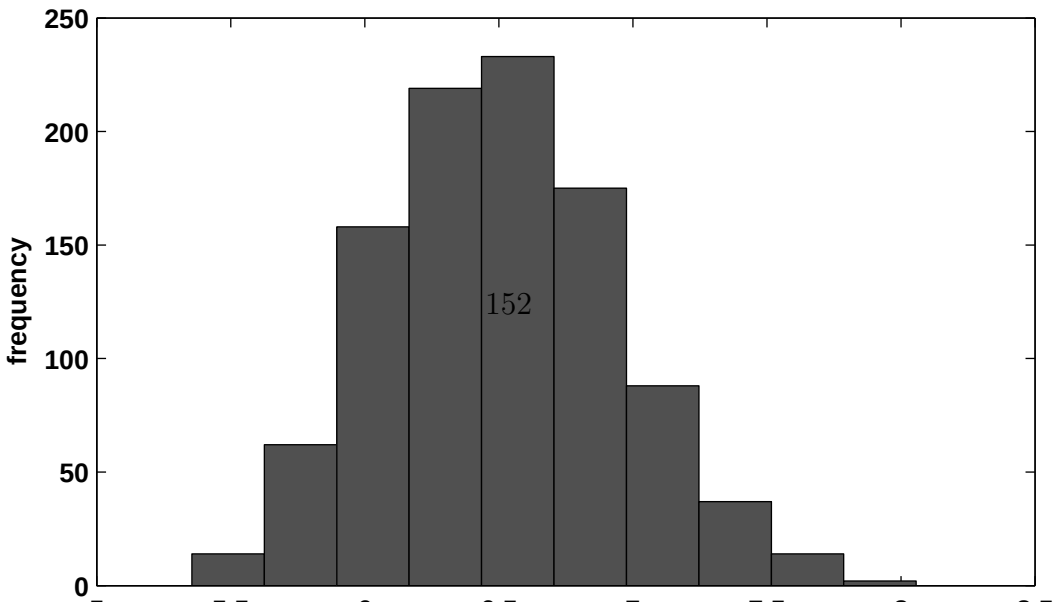
assessments. Figure 9.5 illustrates the disaggregation by distance and magnitude for risk in the Newcastle and Lake Macquarie region. The individual plotting tool `calc_annloss_deagg_distmag` can be used with `wrap_risk_plots` to de-aggregate the risk by distance and magnitude.

2. **disaggregation by building construction type** separates the annualised loss into the four major construction types; wood framed, un-reinforced masonry, concrete and steel framed. Figure 9.6 illustrates the annualised loss disaggregated by structural type for the Newcastle and Lake Macquarie region. The individual plotting tool `calc_annloss_deagg_structural` can be used with `wrap_risk_plots` to de-aggregate the risk by structural type.

3. **disaggregation by suburb** separates the annualised loss into different suburbs and allows the visualisation of how earthquake risk varies spatially. The individual processing tool `calc_annloss_deagg_sub` can be used with `wrap_risk_plots` to de-aggregate the risk by suburb. Note that `calc_annloss_deagg_sub` produces a CSV file that can be plotted in GIS or some other environment.

9.5 Earthquake scenario results

The output for an earthquake scenario loss simulation is also stored in the file `<site_loc>_db_savedecloss.mat`. The format for this file is exactly the same as the format for the probabilistic risk output. Note, however, that each of the earthquakes in `<site_loc>_db_savedecloss.mat` represent different realisations of the same scenario (see Section 3.5). The primary technique for visualising an EQRM scenario output is via a histogram such as that shown in Figure 9.7 for a simulation of the Newcastle 1989 earthquake. Each bar of the histogram in Figure 9.7 represents the frequency of realisations which predicted a loss value within the bars domain. Histograms such as Figure 9.7 can be produced using the individual plotting tool `calc_scen_loss_stats` with `wrap_risk_plots`. Note that `calc_scen_loss_stats` also produces basic statistics for the scenario realisations such as the mean and median.


250
200
150
100
50
0
frequency
152

9.6 Key functions, flags and parameters

Name	Type	Description
<site_loc>_db_hzd.mat	Output	EQRM output file for a probabilistic earthquake hazard run.
<site_loc>_db_savedecloss.mat	Output	EQRM output file for a probabilistic (or scenario) risk run.
rtrn_per	Par	Vector whose elements represent the return periods to be considered.
plot_hzd_map	Function	Post processing tool for plotting probabilistic earthquake hazard maps.
create_gis_output	Function	Post processing tool to export probabilistic earthquake hazard data.
plot_singlesite_hzd	Function	Post processing tool for plotting hazard exceedance curves or uniform hazard spectra.
wrap_risk_plots	Function	Post processing tool to handle user interaction with risk processing/plotting tools.
plot_pml	Function	Post processing tool for plotting and computing the PML.
calc_annloss	Function	Post processing tool for computing the annualised loss.
calc_annloss_deagg_distmag	Function	Post processing tool for plotting the annualised loss disaggregated by distance and magnitude.
calc_annloss_deagg_structural	Function	Post processing tool for plotting the annualised loss disaggregated by building structural type.
calc_annloss_deagg_sub	Function	Post processing tool for computing the annualised loss disaggregated by suburb.
calc_scen_loss_stats	Function	Post processing tool for analysing and plotting loss data for earthquake scenario runs.