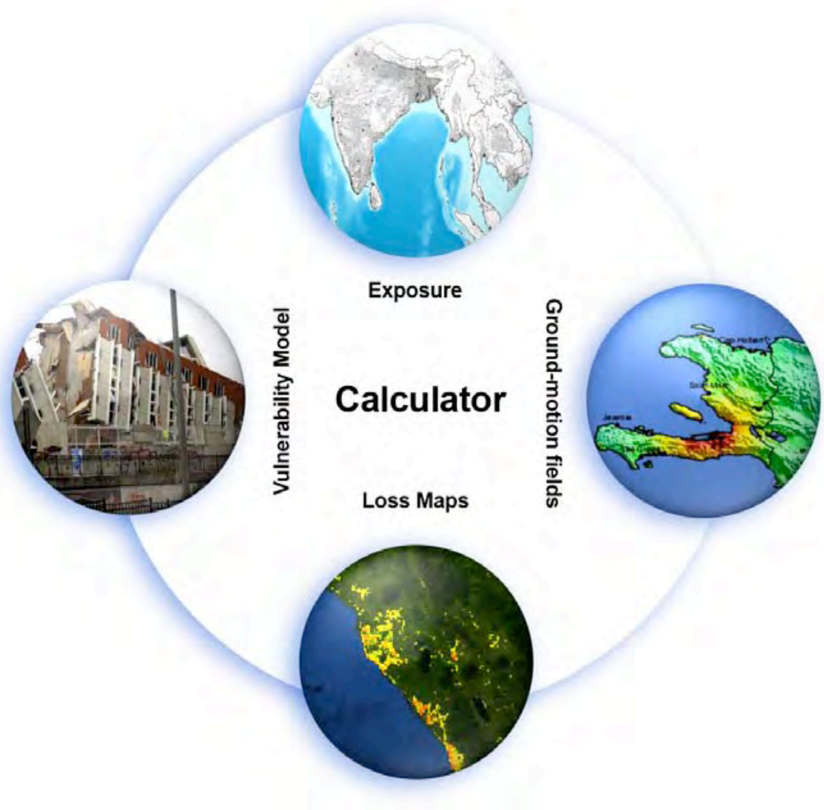
GEM Technical Report 2010-5

GEM1 Seismic Risk Report

Part. 1: H. Crowley, M. Colombi, J. Crempien, E. Erduran, M. Lopez, H. Liu, M. Mayfield, M. Milanesi
Part. 2: H. Crowley, A. Cerisara, K. Jaiswal, N. Keller, N. Luco, M. Pagani, K. Porter, V. Silva, D. Wald, B. Wyss



GEM1 Seismic Risk Report

Part 1

By H. Crowley[1], M. Colombi[1], J. Crempien[2], E. Erduran[2], M. Lopez[1], H. Liu[3], M. Mayfield[3], M. Milanesi[1]

Part 2

By H. Crowley[1], A. Cerisara[1], K. Jaiswal[4], N. Keller[5], N. Luco[4], M. Pagani[6], K. Porter[4], V. Silva[1], D. Wald[4], B. Wyss[5]

October 2010

1. *European Centre for Training and Research in Earthquake Engineering (EUCENTRE), Pavia, Italy*
2. *NORSAR, Kjeller, Norway*
3. *GFZ Helmhotz Centre, Potsdam, Germany*
4. *National Earthquake Information Center, US Geological Survey, Golden, Colorado, USA*
5. *GEM Foundation, Pavia, Italy*
6. *Swiss Seismological Service, ETH Zurich, Zurich, Switzerland*

Citation: Crowley H., Colombi M., Crempien J., Erduran E., Lopez M., Liu H., Mayfield M., Milanesi M. (2010) GEM1 Seismic Risk Report: Part 1, GEM Technical Report 2010-5, GEM Foundation, Pavia, Italy.

Crowley H., Cerisara A., Jaiswal K., Keller N., Luco N., Pagani M., Porter K., Silva V., Wald D., Wyss B. (2010) GEM1 Seismic Risk Report: Part 2, GEM Technical Report 2010-5, GEM Foundation, Pavia, Italy.

www.globalquakemodel.org

ABSTRACT

This report provides a summary of the GEM1 Risk activities; it is divided into two parts, the first relating to a study of existing seismic risk software and previous global risk initiatives, and the second concentrating on the development of a computational engine for global risk assessment, based on those lessons learned. Part 1 looks at a number of existing software for seismic risk assessment that are either open source or have been supplied to the GEM1 Risk team. The main aim and achievement of the study was to identify specific aspects of existing software (related to methodology, usability, graphical representation, efficiency and so on) which should be implemented in a global risk engine for GEM. Furthermore, the following key characteristics required for the GEM risk engine were identified:

- Open-source software development: the source code of the global risk calculator should be available to any user and the development of the code should be a product of the efforts of a community, and not just limited to a small working group.
- Platform independent: it should be possible to use the tool(s) in any operating system.
- Flexible: this code needs to be developed with the purpose of creating a platform for risk assessment, rather than another static risk application. This platform should allow users to evaluate risk due to seismic hazard, but also allow other types of hazard such as floods or hurricanes, to be considered.
- Dynamic: this calculator should allow users to update their results based on newer models, datasets or hazard inputs.
- Modular and expandable: this risk calculator should be developed in a way that any user can easily implement and combine different methodologies. To make this attribute possible, an object-oriented philosophy should be adopted.
- Scalable: the tool should allow one to perform risk assessment at different levels of resolution from an urban level to a global scale.

Part 2 describes the preliminary development of a risk engine and risk user interface for the OpenGEM platform. The development of this global risk calculator had the preliminary goal of producing an engine capable of calculating human losses on a worldwide scale, but that could also be scaled down for local risk analyses, including building damage assessment. This report also provides a detailed description of aspects such as the structure of this tool, the databases that were handled in GEM1 for exposure and vulnerability and the calculations and procedures that have so far been performed. A sample test of deterministic-event and classical PSHA-based losses for several different regions are thus also included within this document, in order to show the possible outputs that can be obtained from this calculator. Finally, a chapter is provided on the future direction that will be taken for the development of the risk engine and the required future specifications that have so far been identified.

Keywords: GEM1 risk; seismic risk; software; local and global applications.

ACRONYMS

ATC	Applied Technology Council
CDF	Cumulative distribution function
CE	Conversion equation
COV	Coefficient of variation
DCW	Digital Chart of the World
ERF	Earthquake rupture forecast
ESRI	Environmental Systems Research Institute
FAO	Food and Agriculture Organization
GAUL	Global Administrative Unit Layer
GADM	Global Administrative Areas
GDP	Gross domestic product
GLC	Global Land Cover
GMPE	Ground motion prediction equation
GRUMP	Global Rural and Urban Mapping Project
HC	Hazard curve
IML	Intensity measure level
IMT	Intensity measure type
LC	Loss curve
LM	Loss map
LR	Loss ratio
LRC	Loss ratio curve
LREM	Loss ratio exceedance matrix
LRM	Loss ratio map
LRPM	Loss ratio probability matrix
MAG	Model Advisory Group
ML	Mean loss
PAGER	Prompt Assessment of Global Earthquakes for Response
PDF	Probability density function
PE	Probability of exceedance
PGA	Peak ground acceleration
PGV	Peak ground velocity
PO	Probability of occurrence
PSHA	Probabilistic Seismic Hazard Assessment
SA	Spectral acceleration
SD	Standard deviation
USGS	United States Geological Survey
WHE	World Housing Encyclopaedia

ACKNOWLEDGEMENTS

The GEM1 Risk team would like the acknowledge and thank all those that have supported them in producing this report by providing source code, example applications, support, answering doubts, and so on, and in particular: Trevor Allen, Edward Anderson, Cyril Bonjour, Alfredo Campos Costa, Alexandra Carvalho, Mustafa Erdik, Ned Field, Stuart Gill, Duncan Gray, Ufuk Hancilar, Yaver Kamer, Andrew King, Dominik Lang, Conrad Lindholm, Iain Matcham, Sergio Molina, Mario Ordaz, Stefano Parolai, David Robinson, Mathilde Sorensen, Maria Luisa Sousa, Sergey Tyagunov, Friedemann Wenzel, Max Wyss, Jochen Zschau.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Part 1. Seismic Risk Review

By H. Crowley, M. Colombi, J. Crempien, E. Erduran, M. Lopez, H. Liu, M. Mayfield, M. Milanesi

1 Introduction

1.1 Seismic Risk

Seismic risk can be broadly defined as the probability of harmful consequences, or loss of life or injury, property damage, social and economic disruption or environmental degradation, resulting from interactions between seismic hazards and vulnerable conditions. There are a number of seismic hazards that can be considered in risk assessment, but this report focuses on strong ground shaking. Vulnerable conditions can be described using vulnerability functions or fragility and consequence functions. Vulnerability functions describe the probability of losses, given a level of ground shaking, whereas fragility functions describe the probability of exceeding different limit states (such as damage or injury levels) given a level of ground shaking, and consequence functions (or damage-loss conversion functions) provide the probability of loss given a damage state (e.g. collapse).

The calculation of seismic risk in the codes studied herein has been undertaken through a number of approaches, with the two main categories of vulnerability being those that consider the built environment through an indirect approach to loss calculation (i.e. the calculations go first via damage estimation) and those that directly calculate the loss based on the hazard (Figure 1.1). Within these approaches, empirical (based on past observation), analytical or even expert opinion methods might be used to describe the vulnerable conditions.

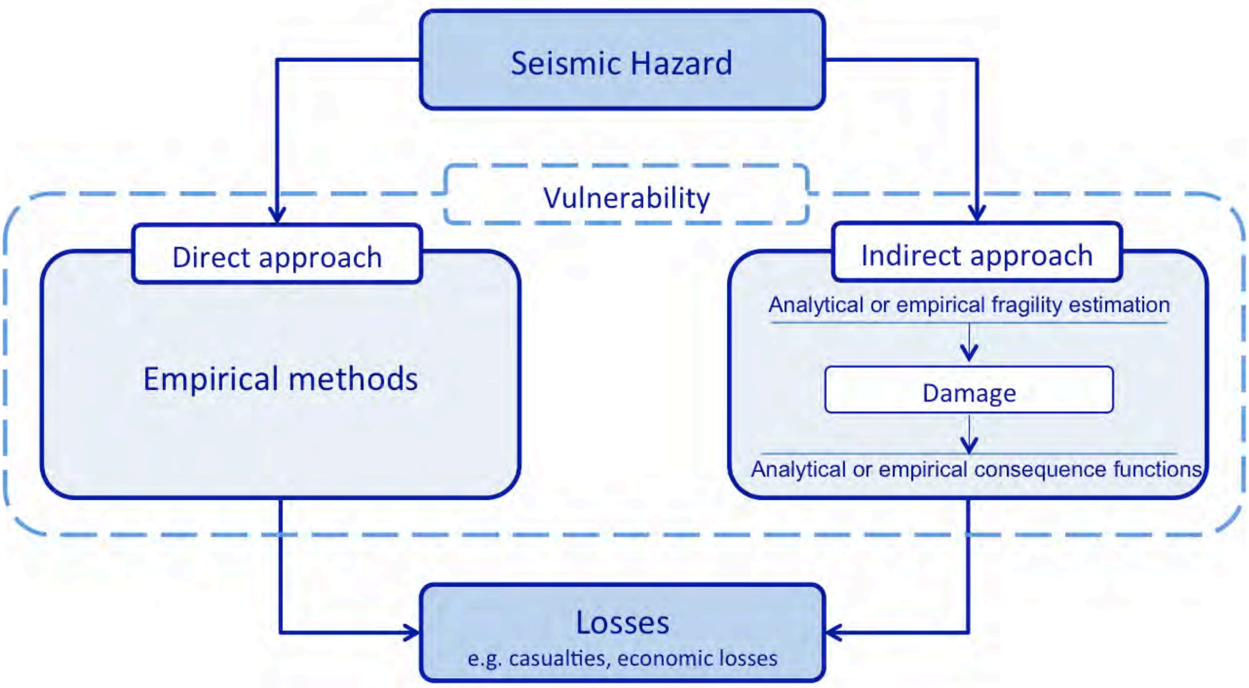


Figure 1.1 Illustration of the types of seismic vulnerability methodologies

Different outputs of “seismic risk” can be obtained depending on the way in which seismic hazard is modelled. The following hazard descriptions and the resulting risk outputs are frequently found in seismic risk codes.

Table 1.1 Different possible risk outputs as a function of the hazard description

Seismic Hazard Description	Risk Output
Deterministic events, maps of ground shaking	Damage/loss maps (i.e. spatial distribution of damage and/or loss) that are conditional on a given event (earthquake) occurring
Hazard maps	Damage/loss maps that are conditional on the hazard with a given return period
Hazard curves	Single asset "loss curves" (better known as loss exceedance curves) and damage/loss for a given probability of exceedance/return period
Probabilistic events-based ground-motion fields	Multiple asset (aggregate) loss curves and aggregate damage/loss maps with a given probability of exceedance/return period

The different risk outputs have varying uses; when a single earthquake scenario is considered, this can be of use for emergency planning or post-earthquake rapid loss assessment; when a hazard map or hazard curve is used this can be of use to provide a simple comparative estimate for different assets of the levels of conditional or unconditional risk, respectively; the use of probabilistic events-based ground-motion fields (for example, based on Monte Carlo simulation) allows one to produce aggregate loss exceedance curves of portfolios of buildings / exposed assets (noting that the aleatory variability of the ground-motion prediction equation should ideally be modelled with separation of inter- and intra-event variability and spatial correlation modelling of the intra-event variability). Appendix A summarises the different hazard approaches used by each code studied herein.

1.2 Existing Seismic Risk Software

The activities in GEM1 Risk began with the study of existing seismic risk software with the objective of demonstrating the existing capabilities for residential buildings and setting the guidelines/specifications for seismic risk assessment in GEM.

The activities of the GEM1 Risk team were initially defined at the GEM1 Kick-off meeting in Canberra in March 2009 where it was decided that a number of existing risk codes (Table 1.2) should be added to a GEM1 sandbox for testing and application.

In Table 1.2, "open-source" is used to identify those codes for which the source code is openly and publicly available, "standalone application available" means that an executable version is publicly available, "standalone application provided" means that an executable version of the code was provided to the GEM1 team, and "source code provided" means that this code is not freely downloadable to all, but that the source code has been provided to the GEM1 team. Of all the codes considered, the only code that has been uploaded to a public repository for open source community development is EQRM, though the actual development to date has mainly been restricted to a core group at GeoScience Australia.

The GEM1 Risk team also decided to add HAZUS and MAEviz to the selected software as the former was a pioneering software in seismic risk assessment and formed the basis for many subsequent codes, whilst MAEviz is an open-source software which is highly developed from an IT perspective. The HAZUS software was obtained (it is freely distributed on a CD to anyone who requests it), but the hardware requirements were rather demanding compared to the other software reviewed (e.g. need for 3GHz processor, 2GB RAM, 120 GB available storage) that the team decided not to carry out applications with HAZUS. Nevertheless, it is evident that the methodology behind this software has been the basis for many of the codes tested herein and it is thus implicitly part of the review.

Table 1.2 Overview of existing risk software selected for evaluation in GEM1

Risk Software	Version	Institution	Main Contact	Second Contact	Availability
SELENA	4.0 and 5.0	NORSAR	Conrad Lindholm	Dominik Lang	Open-source (www.norsar.no/pc-35-68-SELENA.aspx)
EQRM	1.0svn1393	GeoScience Australia	Duncan Gray	Trevor Allen	Open-source (http://sourceforge.net/projects/eqrm)
ELER	2.0	KOERI	Mustafa Erdik	Ufuk Hancilar	Standalone application available (ftp://www.orfeus-eu.org/pub/software/ELER)
QLARM	1.1.7	WAPMERR	Max Wyss	Cyril Bonjour	Source code provided upon request (http://qlarm.ethz.ch)
CEDIM (CREST)	3.1	CEDIM	Sergey Tyagunov	Friedemann Wenzel	Source code provided upon request
CAPRA	0.9.8.0	World Bank	Edward Anderson/ Stuart Gill	Mario Ordaz	Source code provided upon request (http://ecapra.org)
RiskScape	0.1.4	GNS Science	Andrew King	Iain Matcham	Standalone application provided (www.riskscape.org.nz)
LNECLoss		LNEC	Alfredo Campos Costa	Maria Luisa Sousa	Source code provided upon request
MAEviz	3.1.1	MAE Center	-	-	Open-source (www.mae.cee.uiuc.edu/software_and_tools/maeviz.html)
OpenRisk		SPA Risk LLC	Keith Porter	-	Source code provided upon request (www.risk-agora.org)

1.3 Structure of Part 1 of Report

Part 1 of this report (from Chapter 2 to 11) includes a description of each of the risk codes that have been studied, considering in particular: the methodologies that they employ; the key IT details in terms of programming language, software development, graphical user interface, documentation, etc.; the main attributes of the exposure, hazard and vulnerability modules, focusing mainly on the input requirements; and finally the output of the software. A summary of these attributes for each code is provided in Appendix A.

All of the codes were applied by the GEM1 Risk team in test-bed applications in Los Angeles, Zeytinburnu, Bucharest, Lisbon, Managua, Marmara Region, Hawke's Bay and Shelby County. However, these are not included in this report as the main goal of carrying out such applications was to learn more about the software and not to provide reliable loss results. A much larger amount of data collection and validation would be needed before publishing the results of such applications.

Chapter 12 describes some recent global risk assessment programmes, focusing in particular on seismic risk. The study of these programmes has been helpful to ensure that the GEM1 Risk team could begin to provide the necessary advancements to global seismic risk assessment, which will then need to continue within GEM.

A critical summary of the findings of this study is provided at the end of Part 1 (Chapter 13) and makes proposals regarding the attributes that the GEM seismic risk software should feature, some of which were introduced into the risk engine developed by the GEM1 Risk team and described in Part 2 of this report. The GEM1 risk engine was used in some test bed applications and for global risk assessment, though it is noted that the results presented herein are for proof-of-concept purposes only and should not be deemed as reliable estimations of seismic risk.

2 SELENA

2.1 Summary of Software

SELENA (SEismic Loss EstimatioN using a logic tree Approach) is an earthquake loss estimation tool developed by the International Centre for Geohazards ICG through NORSAR, (Norway) and the University of Alicante (Spain). A Beta version of SELENA 1.0 was released in early January 2007. Since then, many versions have been released with version 4.0 and 5.0 being the versions which are currently available. SELENA version 4 is a MATLAB™-based tool used to compute the seismic risk in urban areas and it is open source as long as the user has MATLAB. In version 5 MATLAB m-code has been translated into C-code which allows SELENA to run without using MATLAB. This latter is a standalone version. A GIS-based software (e.g., ArcGIS) can be utilized at multiple levels of resolution to display predicted losses graphically. However, RISe (Risk Illustrator for Selena) has been produced which is the associated GIS viewer that is part of the SELENA package. It allows for easy viewing of the results from the SELENA analysis.

This software is based on the HAZUS methodology [FEMA, 1999; 2003] and the versions reviewed herein use the Capacity Spectrum Method, the MADRS Method and the Displacement Coefficient Method to identify the non-linear response of buildings. The user has to supply various input data such as: built area or number of buildings of the different model building types (MBT); earthquake sources; empirical ground-motion prediction relationships; soil maps and corresponding ground-motion amplification factors; capacity curves and fragility curves corresponding to each of the model building types; and finally cost models for building repair or replacement. SELENA will compute the probability of being in each one of the five damage states for the given building types. The structural damage states are divided into five classes: none, slight, moderate, extensive and complete. This probability is subsequently used with the built area or the number of buildings to express the results in terms of damaged area or number of damaged buildings. Finally the damage is converted into economic losses and human casualties.

While the HAZUS approach is tailored to the U.S., SELENA could potentially be applied to user-specified regions provided the ground-motion distributions and building capacity functions are appropriately modified. This is due to the fact that the software package is constructed in a manner that clearly separates the regional data from the calculation procedures.

2.2 Methodology

This software is based on the HAZUS methodology [FEMA, 1999; 2003] and it uses the Capacity Spectrum Method (CSM) as proposed in ATC 40 [ATC, 1996], the MADRS Method, which is the modified capacity spectrum method, and the Displacement Coefficient Method (DCM) as proposed in FEMA 440 [ATC, 2005]. In the current versions, the user can choose to apply any of these methods. An innovation with respect to the HAZUS methodology consists in the use of a logic tree approach based on the weighting of the input parameters which allows epistemic uncertainty to be accounted for. The final results are multiplied by their corresponding weights and they are fitted to a normal distribution in order to get the median values as well as the fractiles. Currently, the logic tree represents uncertainties in the earthquake source, attenuation relationships, shaking maps, soil type, vulnerability and economic values of damaged built area (Figure 2.1). The user has to provide a *.txt table named cpfile.txt to choose which methodology has to be used. This input file defines the type of analysis method to be used in order to calculate the performance point and whether the damage results are dependent on the number of damaged buildings or the damaged building area. The 1st column of the table is the type of analysis method used (1-CSM; 2-MADRS, 3 DCM), the 2nd column is the type of damage results (1-square metres; 2-number of buildings) and the 3rd column is the human losses method (1-basic method; 2-HAZUS methodology). The specific details on these methods can be found in the manual which has been produced for SELENA [Molina *et al.*, 2009a,b].

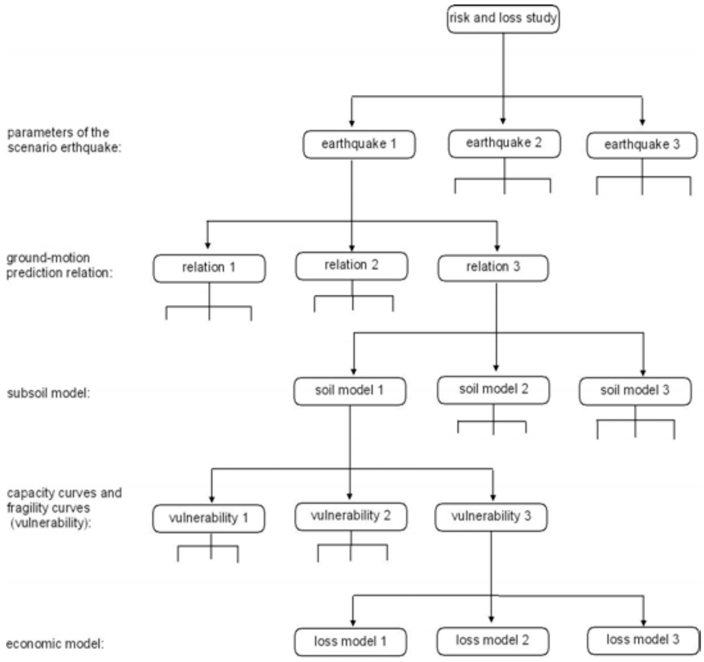


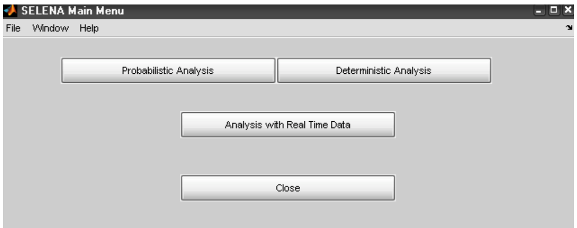
Figure 2.1 Epistemic uncertainty framework considered in the SELENA software [Molina *et al.*, 2009a]

2.3 IT Details

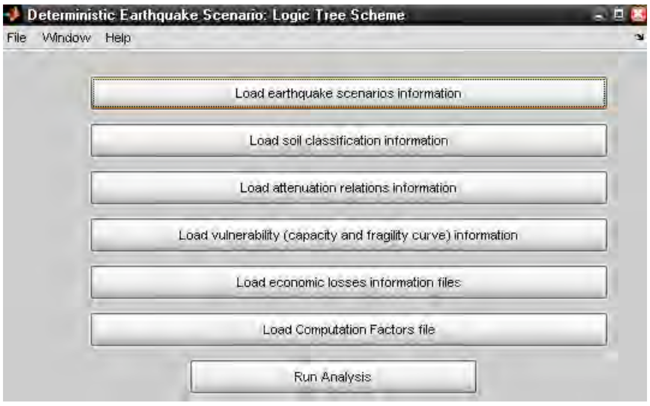
The software requirement of the SELENA version 4 package mainly consists of MATLAB which is a commercial and proprietary software. Since MATLAB is a cross-platform programming language, SELENA can run on the following operating systems without any code modification: Windows (x64), Linux (x86-64), Mac OS X and Solaris 64. Version 5 of SELENA is written in the C programming language which is a free software covered by the GNU General Public License (GPL) and to be built requires libgsl and libgslcblas of the GNU Scientific Library and a minor modification to the Makefile. Furthermore a Geographic Information System (GIS), such as Arcview, is an optional choice for the representation of results; however, as mentioned before, RISe (Risk Illustrator for Selena) is available which is a separate GIS tool that uses the Google Earth platform to illustrate all geo-referenced input and output files of SELENA. RISe is part of the SELENA package.

SELENA's input files can be generated as tables in MS-Excel or MS-Access and then they can be easily exported in ASCII-table files with all the required information given in the matrices. SELENA's output files have been created in ASCII-table files and they can easily imported into MS-Excel or MS-Access. They are geo-referenced and for this reason a GIS is recommended. To be plotted in a GIS, the output files have to be exported to *.dbf formatted files and they must contain at least information concerning longitude, latitude and geounit ID. The geounit is the minimum geographical unit, e.g. the census tract, grid cell or even individual building. The level of resolution of the predicted parameter depends on the size of the geounit which can be defined by the user. The source code is very well commented such that a user can easily go through the script and change it and in this way, apply their own modifications and adjustments. The running time of SELENA will be a function of the computer specifications and the level of detail considered in the analysis. However, Version 5 is able to run scenario cases up to 50 times faster than the SELENA-MATLAB version.

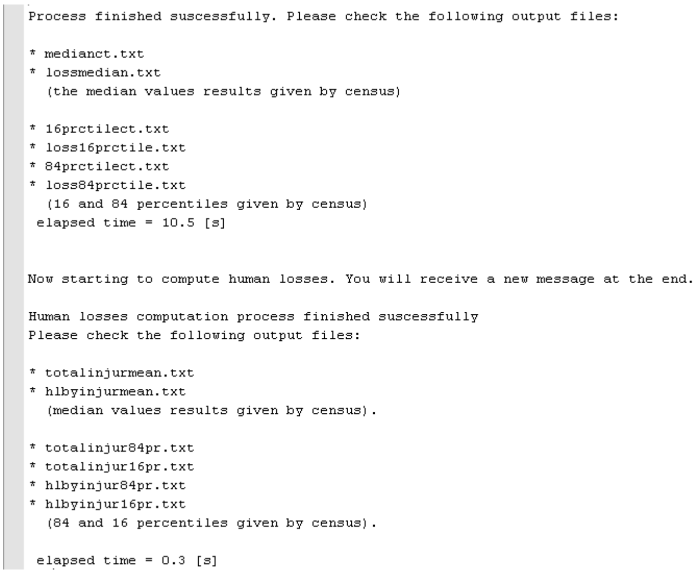
Finally, SELENA version 4 has a simple GUI and it is possible to create a Log file (see Figure 2.2c). In Figure 2.2 a couple of screenshots of the program (version 4) are shown; SELENA version 5 offers both a command line interface and a GUI version, which has been programmed in Qt Creator (Figure 2.3). For what concerns the documentation of the Application Programming Interface (API), a manual API has been produced by the SELENA team and shared with GEM1.



(a)

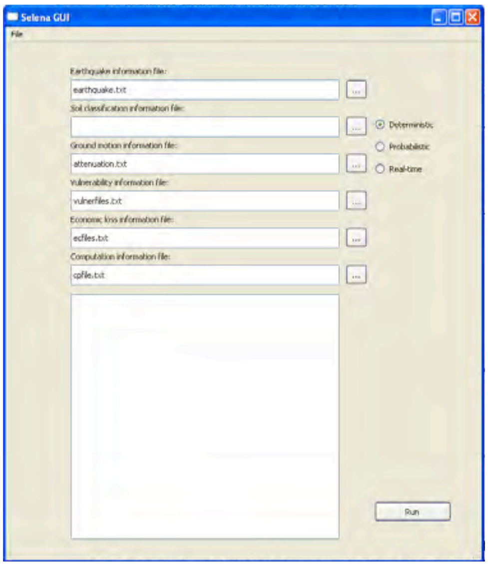


(b)

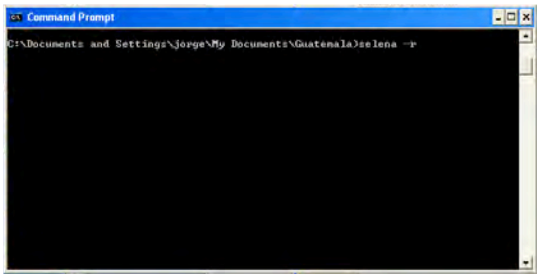

```
Process finished successfully. Please check the following output files:

* medianct.txt
* lossmedian.txt
  (the median values results given by census)

* 16prctilect.txt
* loss16prctile.txt
* 84prctilect.txt
* loss84prctile.txt
  (16 and 84 percentiles given by census)
 elapsed time = 10.5 [s]


Now starting to compute human losses. You will receive a new message at the end.

Human losses computation process finished successfully
Please check the following output files:

* totalinjurmean.txt
* hlbyinjurmean.txt
  (median values results given by census).

* totalinjur84pr.txt
* totalinjur16pr.txt
* hlbyinjur84pr.txt
* hlbyinjur16pr.txt
  (84 and 16 percentiles given by census).

 elapsed time = 0.3 [s]
```

(c)

Figure 2.2 (a) and (b) Screenshots of the GUI of SELENA version 4 (c) Log file screenshot of SELENA version 4



(a)



(b)

Figure 2.3 (a) Screenshot of the SELENA version 5 stand-alone GUI (b) Screenshot of the command prompt of SELENA version 5

2.4 Exposure Module

The user has to input data on both the building and population inventory. With regards to the building inventory, the user can choose to compute risk using the total built area (in square metres) or the total number of buildings according to the occupancy of each model building type for each geographical unit. Hence, the occupancy (residential, industrial, commercial, etc.) is considered by the program. Furthermore, it is necessary to input parameters describing the capacity curve and the fragility functions for each model building type. The model building type is user-defined as long as the user can input corresponding capacity and fragility curves. Finally, to compute economic losses it is necessary to upload economic information about cost of repair or replacement of each model building type. For what concerns the population inventory, the user has to provide the total number of people living in the geographical unit, the population percentages staying indoors or outdoors in relation to the time of day, the percentage of occupancy class for each of the different model building types and the casualty rate of severity for the different damage states. The inventory is stored in a lot of different *.txt tables (see the following examples). It is a fairly demanding job to prepare the input files and to transform a preliminary inventory into the format required by SELENA.

Each inventory table is directly linked with the hazard by the geounit; hence, the user has to upload data in the same units as the hazard. The level of resolution of the inventory is user-defined, the unit could be a regional unit, a provincial unit, a local unit and it could be an arbitrary polygon (regular or irregular). The user can configure in a totally independent way the inventory which is not hard-coded in any way.

SELENA can work with either the area or the number of buildings. The following describes the input tables required for building area.

- builtarea.txt table: this contains the total built area of each model building type in square metres for each geounit;
- ocupmbt(j).txt table: these input files contain the built area (in square metres) according to occupancy in each geounit for the different model building type *j*. For each model building type a *.txt table has to be prepared;
- headerocc.txt table: this provides the necessary header for the input file ocupmbt(j).txt;
- header.txt table: this is necessary in order to create the damage output files. The first four columns of the header remain always the same and they represent the geounit, the coordinates and the soil type. All other columns represent the damage state of each MBT. The number of these additional columns always has to be a multiple of 5 due to the fact that SELENA considers five damage states. The last label Numb stands for a column with an order number.

With regards to the population data, to compute social losses, the following tables need to be supplied:

- population.txt table: this contains the total population in the studied area for each geounit;
- ocupmbtp.txt table: this indicates the share of each model building type and its occupancy over the entire building stock;
- poptime.txt table: this contains the population percentages staying indoors or outdoors as a function of the time of the day;
- injury(i).txt table: these files contain the casualty rates of severity *i* (*i*=1,2,3,4 - light, moderate, heavy, death) in percentages for the different damage state (slight, moderate, extensive and complete).

Finally, for what concerns economic losses the following input files need to be prepared:

- ecfiles.txt table: this refers to the sub-files elosssd(i).txt (slight damage), elossmd(i).txt (moderate damage), elossed(i).txt (extensive damage), elosscd(i).txt (complete damage) and indicates their corresponding weight for the logic tree methodology. Uncertainty concerning the economic model is considered and different models can be uploaded;

- elosssd1.txt, elossmd1.txt, elossed1.txt and elosscd1.txt tables: these files contain the economic information in order to compute economic losses due to slight, moderate, extensive and complete structural damage to a specific model building type.

2.5 Hazard Module

SELENA allows the seismic ground-motion parameters to be calculated in three different ways, which are termed:

- Deterministic analysis;
- Probabilistic analysis;
- Analysis with real-time data.

In order to compute a deterministic analysis the values of the spectral ordinates in each geographical unit are calculated using a deterministic earthquake scenario and ground-motion prediction equations (GMPEs). To characterize the scenario earthquake, the following inputs are required: longitude and latitude, focal depth, surface wave magnitude M_s and moment magnitude M_w, fault orientation from North, dip angle and fault mechanism. Then, the user has to define the type of design spectrum (IBC-2006 [IBC, 2006], Eurocode 8 Type I – II [CEN, 2003], Indian Code [BIS, 2002]) and SELENA provides the predicted ground-motions parameters (spectral accelerations and spectral displacements) which are calculated by selectable GMPEs. A considerable number of GMPEs are already incorporated in the SELENA code, however any user-provided relation can be easily implemented. All provided prediction relations refer to rock site conditions. However, in a separate calculation step, the user can supply soil parameters useful to set up soil amplification factors to amplify the seismic demand.

In order to compute a probabilistic analysis the values of the spectral ordinates in each geographical unit are taken from probabilistic hazard maps. The user can also supply different ground-motion fields corresponding to different earthquake scenarios. Then, each ground-motion field is characterized by a specific weight of the logic tree. In addition to the spectral ground-motion values of PGA, spectral acceleration (Sa) at 0.3 seconds and spectral acceleration at 1.0 second for each geounit, the geographical coordinates of the centroids have to be supply for each scenario. Then, the user has to provide the type of design spectrum and the soil parameters if hazard or ground-motion fields have been developed for rock conditions.

In order to compute an analysis with real-time data the values of the spectral ordinates in each geographical unit have to be assigned. The real-time ground-motion data is unlikely to be available in the centroids of all geounits. This is the major problem of this forecasting model. SELENA has applied a special procedure to solve this problem, however any user-provided procedure can be implemented. In SELENA the points with available data that are within a 5-km radius around each centroid are checked. If at least 5 points are within this radius, the mean values and corresponding 16% and 84% fractiles of the spectral ordinates of all stations are computed and assigned to the centroid. If not, a new circle of 10-km radius is considered. These provided spectral ordinates already cover soil amplification effects and no additional input regarding the soil parameters are required.

The SELENA code considers only the ground shaking hazard and other earthquake-related hazards such as liquefaction, fault rupture, landslide, tsunami or seiche are not yet considered.

The hazard data have to be uploaded as external files. They are *.txt tables and they are different depending on the forecasting model used.

For the deterministic analyses the user has to provide:

- earthquake.txt table: this contains information about the earthquake to be used in the study and a numerical code for the spectral shape to be used (1 = United States IBC-2006, 2 = Eurocode 8 Type 1, 3 = Eurocode 8 Type 2, 4 = Indian Code IS 1893 Part 1);
- attenuation.txt table: this contains information about the empirical ground-motion prediction equations.

For what concerns probabilistic analyses, the user has to provide:

- shakefiles.txt table: this indicates the corresponding weights for the logic tree and a numerical code for the spectral shape to be used;
- shakecenter(i).txt table: this contains the spectral ground-motion values (PGA, Sa at 0.3 seconds and Sa at 1 second in g) , the coordinates of the geounit in order to visualize the results in a GIS and the soil type.

For the real-time data analyses the user has to provide:

- realtimefile.txt table: this indicates the moment magnitude of the respective earthquake considered and a numerical code for the spectral shape to be used;
- realtimegrid.txt table: this contains information on peak ground acceleration, spectral acceleration at 0.3 s and spectral acceleration at 1 s.

The fact of using external tables makes the software very flexible. The user can change and modify the hazard input without any problems. However this flexibility could lead to significant effort to run analyses using many ground-motion fields. This is true especially when the user is running probabilistic analyses. For instance, it is worth noting that if 100 ground-motion fields are necessary, 100 different input hazard files have to be prepared.

2.6 Vulnerability Module

In order to determine the seismic performance of a building in SELENA a comparison between the capacity of the structure and the seismic demand is necessary. For what concerns the capacity of the structure, the user has to create a *.txt table named capacity.txt. This table refers to one particular set of building capacity curves. Moreover, it contains the value of the elastic damping in percentage for each one of the model building type according to the recommendations of Newmark and Hall [1982], the spectral displacement corresponding to the elastic limit (in m), the κ value for short, moderate and long duration earthquakes and a comment on denomination of the respective model building type. The κ value is a degradation factor that defines the effective amount of hysteretic damping as a function of earthquake duration. In SELENA the capacity curve is completely user-defined. The user has to provide a *.txt table composed of two columns. The first column represents the spectral displacement (in m), while the second column represents the spectral accelerations (in m/s^2). This way, the user can choose to use bilinear curve, trilinear curve, smooth curve, etc. (Figure 2.4).

These files have to be uploaded as external files and the capacity curves are explicitly defined by the user. This makes the software flexible and the user can change the curves at any time. However, it is noted that the user has to upload as many tables as the number of model building types used.

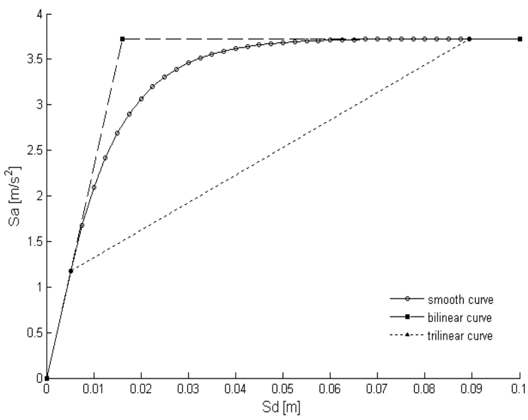


Figure 2.4 Plot of three different types of capacity curves: exponential (or smooth), bilinear, trilinear

The seismic performance of a building, which is the response of the system studied, is represented by the so-called performance point. This point is the intersection between the capacity curve of a structure and the response spectrum. An iterative process is necessary to estimate this point so that the damping of the response spectrum and the structural response are the same. As mentioned before, SELENA can use three different methodologies, Capacity Spectrum Method, the MADRS Method and the Displacement Coefficient Method to calculate the performance point. In Figure 2.5 the performance point calculated with CSM method of a low-rise concrete moment resisting frame designed with a moderate code (C1L_MC) is shown.

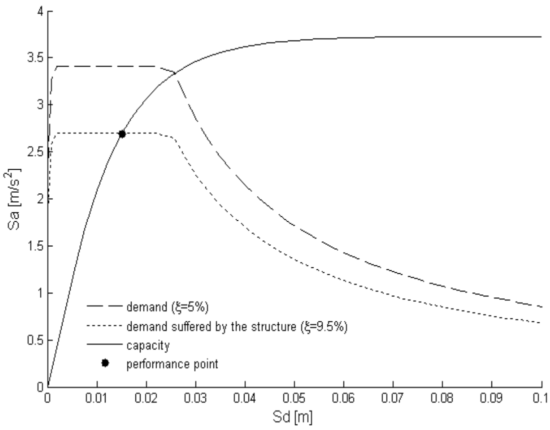
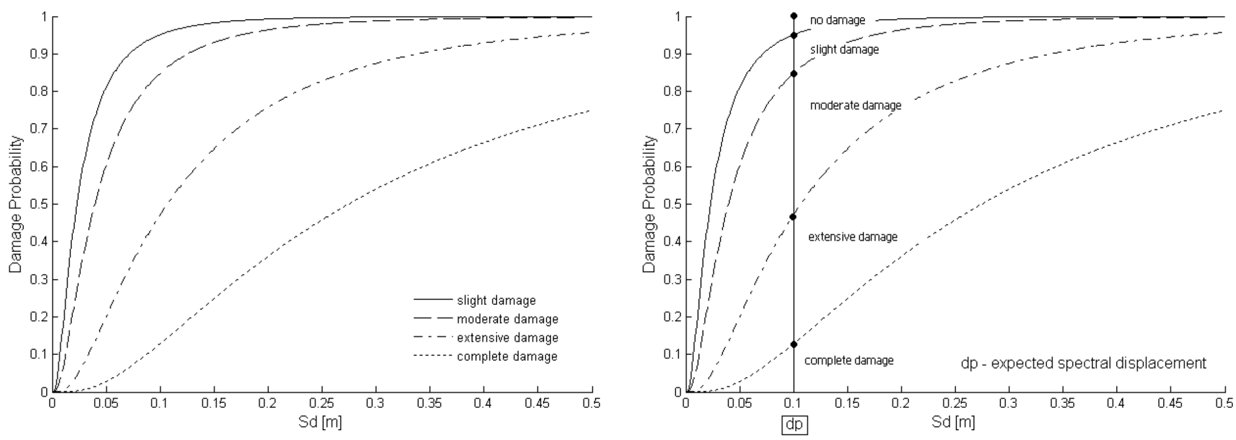


Figure 2.5 Performance point of a low-rise concrete moment resisting frame designed with a moderate code

SELENA uses fragility curves to calculate the damage probability distribution for a class of structures. These curves are represented as the conditional probability of being in, or exceeding a particular damage state given by the spectral displacement or other seismic demand parameter. The parameters defining the fragility functions for a certain building type are closely connected to its capacity curve. In SELENA fragility curves are lognormally distributed and the values related to the HAZUS building types are already calculated and summarized in tables. Nevertheless, the user can also provide their own parameters. The user has to create a *.txt table called fragility.txt. In the 1st column of the table there is the index of the model building type. Then the user has to provide xmedian (xμ) which is the median value of spectral displacement in units of metres at which the building reaches the threshold of the damage states: slight (*s*), moderate (*m*), extensive (*e*) and complete (*c*). Then he/she has to provide xbeta (xβ) which is the standard deviation of the natural logarithm of spectral displacement of the damage states: slight (*s*), moderate (*m*), extensive (*e*) and complete (*c*). Finally, the last column is a comment on the denomination of the model building type. These values are reported in the HAZUS manual or in the SELENA technical manual.

These files have to be uploaded as external files and then the fragility curves are implicitly evaluated by the code. In the code, the lognormal distribution is implemented and the user has just to upload mean and standard deviation for each damage state. In order to change the distribution, one has to modify the source code.

In Figure 2.6 the expected displacement response is overlaid with the fragility curves in order to compute the damage probability in each one of the different damage states. In this example, the expected displacement (performance point) is taken equal to 0.1 metres and the damage distribution is found accordingly.



(a) (b)

Figure 2.6 (a) Fragility curves of a low-rise concrete moment resisting frame designed with moderate code, (b) expected displacement response overlaid with the fragility curves

Finally, uncertainties in capacity and fragility curves can be handled using the logic tree approach. The user can upload different capacity curves (capacity*i*.txt) and different fragility curves (fragility*i*.txt). Then, for each pair of *.txt tables the user has to indicate the corresponding weights of the logic tree through a *.txt file called vulnerfiles.txt (see the following example).

- vulnerfile.txt table:

  0.5 capacity1.txt fragility1.txt

  0.5 capacity2.txt fragility2.txt

Evidently, the summed up weights must yield a value of 1.

The economic losses in SELENA for building repair are computed in the following way:

$$EconomicLosses = CI \cdot \sum_{i=1}^{OccupType} \sum_{j=1}^{BuildingType} \sum_{k=1}^{DamageState} FA_{ij} \cdot PMBTSTR_{j,k} \cdot RCS_{i,j,k} \tag{2.1}$$

Where:

- CI: regional cost multiplier (currently set to 1.0);
- FA_{ij}: built area of the model building type j in the occupancy type i (in [m^2]);
- $PMBTSTR_{j,k}$: damage probability of a structural damage k for the model building type j;
- RCSi,j,k: cost of repair or replacement (by [m^2]) of the structural damage k for occupancy type i and model building type j.

In the current version of SELENA only the direct economic losses caused by structural damage are computed. It also should be regarded that economic losses will only be calculated if the user chooses the analysis type dependent on damaged building area, not on the number of damaged buildings. As can be noted by Equation (2.1) the economic losses are independent of social losses and they are strictly connected to the level of damage suffered by the structure.

The human losses are given by the formulae of Coburn and Spence [2002] and they consider the level of severity i:

$$K_i = K_{si} + K_i' + K_{2i} \tag{2.2}$$

Where:

- K_s: number of casualties due to structural damage;
- K': number of casualties due to non-structural damage;

- K_2: number of casualties due to follow-on-hazards (landslides, fires, etc.).

The level of severity i ranges from light injuries (i=1), moderate injuries (i=2), heavy injuries (i=3) and death (i=4). However, the loss model applied in this version of SELENA only considers the direct losses caused by structural damage (K_s).

The number of human losses can be computed using two different methodologies depending on the available information: the basic methodology and the HAZUS methodology.

For what concerns the basic methodology, the number of casualties due to structural damage K_s can be calculated by:

$$K_{si} = \sum_{j=1}^{BuildingType} \sum_{k=1}^{DamageState} CSR_{j,k} \cdot PMBTSTR_{j,k} \cdot Population_j \tag{2.3}$$

Where:

- $CSR_{j,k}$: casualty rate of severity i for damage state j as provided by the user in the input files;
- $PMBTSTR_{j,k}$: damage probability of a structural damage k for the model building type j;
- $Population_j$: number of people in each model building type j.

The total number of people is computed as follows:

$$Population_j = \sum_{j=1}^{BuildingType} TotalPop \cdot PopbyOccup \cdot Occupbymbt \tag{2.4}$$

Where:

- TotalPop: total number of people living in the geounit as provided by the user in the input files;
- PopbyOccup: percentage of people staying indoors or outdoors dependent on the time of the day as provided by the user in the input files;
- Occupbymbt: percentage of occupancy class for each one of the different model building type as provided by the user in the input files.

For what concerns the HAZUS methodology, the total population is classified into five classes: residential, commercial, education, industrial and hotel population. It is calculated for three times of the day: 2:00 am (night time), 2:00 pm (day time) and 5:00 pm (commute time). In the following table, the relationships to estimate the distribution of the population are shown, where:

- POP: census tract population taken from HAZUS;
- DRES: daytime residential population inferred from census data;
- NRES: night-time residential population inferred from census data;
- COMM: number of people commuting inferred from census data;
- COMW: number of people employed in the commercial sector;
- INDW: number of people employed in the industrial sector;
- GRADE: number of students in grade schools (usually under 17 years old);
- COLLEGE: number of students of college and university campuses in the census tract (over 17 years old);
- HOTEL: number of people staying in hotels in the census tract;
- PRFIL: fraction representing the proportion of commuters using automobiles inferred from profile of the community (0.60 for dense urban areas, 0.80 for less dense urban areas or suburban areas and 0.85 for rural). Default value is 0.80;
- VISIT: number of regional residents who do not live in the studied area, visiting the census tract for shopping and entertainment. Default is set to zero.

Table 2.1 Relationships for estimating population distribution (taken from HAZUS)

Distribution of people in census tract			
Occupancy	2:00 am	2:00 pm	5:00 pm
		Indoors	
Residential	(0.999)0.99(NRES)	(0.70)0.75(DRES)	(0.70)0.50(NRES)
Commercial	(0.999)0.02(COMW)	(0.99)0.98(COMW)+ (0.80)0.20(DRES)+ 080(HOTEL)+0.80(VISIT)	0.98[0.50(COMW)+ 0.10(NRES)+0.70(HOTEL)]
Educational	-	(0.90)0.80(AGE_16)+ 0.80(COLLEGE)	(0.80)0.50(COLLEGE)
Industrial	(0.999)0.10(INDW)	(0.90)0.80(INDW)	(0.90)0.50(INDW)
hotels	0.999(HOTEL)	0.19(HOTEL)	0.299(HOTEL)
		Outdoors	
Residential	(0.001)0.99(NRES)	(0.30)0.75(DRES)	(0.30)0.50(NRES)
Commercial	(0.001)0.02(COMW)	(0.010.98(COMW)+ (0.20)0.20(DRES)+ 0.20(VISIT)+ 0.50(1-PRFIL)0.05(POP)	0.02[0.50(COMW)+ 0.10(NRES)+0.70(HOTEL)]+ 0.50(1-PRFIL)[0.05(POP)+ 1.0(COMM)]
Educational	-	(0.10)0.80(AGE_16)+ 0.20(COLLEGE)	(0.20)0.50(COLLEGE)
Industrial	(0.001)0.10(INDW)	(0.10)0.80(INDW)	(0.10)0.50(INDW)
hotels	0.001(HOTEL)	0.01(HOTEL)	0.001(HOTEL)

The total number of casualties for each severity level can be estimated using damage probabilities and casualty rates. For what concerns the outdoor casualties, slight damage and collapse damage do not generate any outdoors casualties.

The social losses are independent of economic losses and they are strictly connected to the level of damage suffered by the structure.

2.7 Output

Five structural damage states are considered in SELENA, the same as HAZUS: no damage, slight damage, moderate damage, extensive damage, complete damage. The final damage results are given as absolute square metres of the respective damaged building type. In this way, the user is able to process these results in any desired format, i.e., the percentage of built area normalized by the total built area. Further he/she can process these results using a spreadsheet program(e.g. MS Excel). However, the final results can also be given as the absolute numbers of damaged buildings.

The number of damage grades is hard-coded and to change the level of damage the user has to change the code. The output files are *.txt tables and they report the damage both in terms of damage probabilities and in terms of damaged building area (or number of damaged buildings). The results are given for each model building type for each geounit and the user can easily compute the total damage for each unit and for the total region of interest.

In the tables below the damage output files given by SELENA are described in detail.

- dout.txt table: this contains the damage probabilities of each model building type for the five different damage states (no damage, slight, moderate, extensive, complete). The file is structured according to the background file header.txt. Obviously, the sum of damage probabilities for one model building type must yield 1;

- sqmctdout.txt table: this contains the damage results in terms of damaged building area of each model building type for the five different damage states (no damage, slight, moderate, extensive, complete). The file is structured according to the background file header.txt.

The damage results obtained with the SELENA software can be plotted in a GIS given that the coordinates are known. However, the software RISe, which has been developed for SELENA, is recommended to plot the results (Figure 2.7).

The economic loss output file is a *.txt table and it is reported for each geounit. It contains the total economic loss (in a user-defined currency) in each geounit.

The social loss output files are *.txt tables and they are reported for each geounit and for each injury type and for each time of the day. The user can easily compute the total damage for each unit and for the total region of interest.

In the following a list of the tables provided by the software are explained:

- hlbyinjury.txt table: this contains the distinct numbers of human losses (from slightly injured to dead) for the three different daytime scenarios (2:00, 10:00, 17:00) in each geounit;
- totalinjury.txt: this contains the cumulative numbers of human losses (from slightly injured to dead) for the three different daytime scenarios (2:00, 10:00, 17:00) in each geounit.

All these output files can be plotted in Google Earth using RISe and the results can be visualized in 2D and 3D.

Finally, SELENA allows the disaggregation of risk results into a variety of forms such as:

- Disaggregation by geographical unit;
- Disaggregation by damage states;
- Disaggregation by building type;
- Disaggregation by injury type (low, median, heavy and death);
- Disaggregation by the time of the day (2:00 am, 10:00 am and 5:00 pm);
- Disaggregation by the branch of the logic tree utilized.

Figure 2.7 Damage results from SELENA presented in the RISe software

3 EQRM

3.1 Summary of Software

The EarthQuake Risk Model (EQRM) is a computer model for estimating earthquake hazard and risk. It is a product of Geoscience Australia, an Australian Government Agency, and for this reason Australia is the current geographical area of application. Initially, the software was compiled to be used in the MATLAB environment. In further developments, EQRM was transformed into the object-oriented programming language, Python. Its last version is eqrm_version1.0svn1393 and it is developed in order to run multi risk analyses. A GIS-based software can be used to display predicted damages and losses graphically due to the fact that the EQRM outputs can be geo-referenced with a simple join with the inventory table.

This software is based on the HAZUS methodology and it utilizes the Capacity Spectrum Method. It is currently also tailored for use in Australia. However, since it is an open source program, classes inside the package can be modified, replaced or improved and then it can be used for any region of the world as long as the user can estimate the required building parameters (capacity, fragility, etc.).The user has to supply a number of input databases concerning the hazard, the vulnerability and the exposure. If the user would like to execute the risk component directly, without using the hazard module, the following data will be needed: the hazard, the building engineering parameters, the building construction types, the square metres of construction, the building usage and the number of buildings. The replacement cost of buildings and the contents are necessary to compute the economic losses. The main outputs are structural damage and economic loss. However, as stated above, since it is an open source program, classes inside the package can be modified and expanded.

3.2 Methodology

The software is based on the HAZUS methodology [FEMA, 1999; 2003] and it uses the Capacity Spectrum method [ATC, 1996]. However, some modifications have been made to suit the Monte-Carlo approach used in EQRM.

EQRM uses the full response spectra for intersecting with the building capacity curve, rather than a spectral shape fit to two spectral acceleration values, as is carried out in the original HAZUS methodology. The advantage of this approach is that the full response spectrum and all the available information about the soil amplification factors, at all periods, is taken into account. Moreover, there are also minor differences in the way the damping is calculated due to a different hysteretic model being implemented in EQRM.

Another difference is in how the fragility curves are represented. In HAZUS, the fragility curves incorporate the variability of the damage state thresholds, the variability of the capacity curve and the variability of the ground shaking. In EQRM the fragility curves include only the variability of being in a given damage state and not the variability associated with ground shaking or capacity curves. The variability associated with the ground shaking is incorporated in the Monte-Carlo simulations whilst that associated with the capacity curve can be changed by the user using a flag in the input file.

3.3 IT Details

EQRM can be installed on a personal computer loaded with either a Windows or Linux operative system. The software requirement of EQRM originally consisted of MATLAB, which is a commercial and proprietary software. Then, the use of an object-oriented language was favoured and the latest releases have been coded in Python which is an interpreted language that is easy to develop/modify/adapt and especially useful in a scientific context. In addition, to be Python dependent, the package relies on other software such as Numpy, Scipy, Numeric, Shapely, minGW. The last software is only necessary for the Windows operative system. Furthermore a GIS software, such as Arcview, is an optional choice to display the results, though some external tools in MATLAB are provided with the software.

EQRM's input and output files can be generated as *.csv tables and ASCII files. These files can easily be transformed into a database and then they can be plotted on a GIS program. To be plotted in a geographical information system they have to contain information concerning latitude and longitude for each geounit (i.e. the minimum geographical unit). Also in EQRM the level of resolution of the predicted parameter depends on the size of the geounit, which can be defined by the user. Moreover the user can create a *.par file to change the default parameters which are hard-coded. The full path of the setdata file, which is the *.par file, must be passed to EQRM when the program is running. It contains a series of input variables (or parameters) that define the manner in which EQRM operates. For example, there is a flag to control whether the EQRM models hazard or risk (run_type=1 means scenario response spectrum analysis and probabilistic hazard, run_type=2 means scenario loss and probabilistic risk). Other flags can be used to set engineering parameters. In the following, an example of part of a *.par file is shown. In this case the flags about the operation mode and the Capacity Spectrum Method (CSM) are reported:

```
[Operation_Mode]
run_type=[2]
[CSM]
var_bcap_flag=[0]
stdcap=[0.3]
bcap_var_method=[3]
damp_flags=[0,0,0]
Harea_flag=[3]
max_iterations=[7]
SDRelTol=[1]
```

The algorithm is transparent and the users can modify the source code by adding their own functions, analyses or comments. Unfortunately, the manual is not updated for the Python version and in this way the user is not always easily supported (though the queries of the GEM1 Risk team were always rapidly answered by the developers of the software). EQRM uses a command line interface and it does not have a GUI. There is no documented API, but the GEM1 team were able to produce this from the source code using "EPYDOC" which demonstrated that the code is very well commented throughout. Finally, EQRM produces a progress of the calculations it is carrying out (see Figure 3.1). For this reason, it is simple to create a log file as the user just has to redirect this output to another file. Moreover, it is worth noting that this version has the possibility to run in "batch mode". It supports multiple ground-motion fields for processing in the risk module.

```
do site  1  of  44896
get_building_states SD(mm) SA(g) [[ 10.0392]] [[ 0.323]]
get_building_states SD(m) = 0.0100392 SA(m/s2) = 3.16754795
do site  2  of  44896
get_building_states SD(mm) SA(g) [[ 6.9662]] [[ 0.2291]]
get_building_states SD(m) = 0.0069662 SA(m/s2) = 2.246703515
do site  3  of  44896
get_building_states SD(mm) SA(g) [[ 6.9662]] [[ 0.2291]]
get_building_states SD(m) = 0.0069662 SA(m/s2) = 2.246703515
do site  4  of  44896
get_building_states SD(mm) SA(g) [[ 18.1205]] [[ 0.2923]]
get_building_states SD(m) = 0.0181205 SA(m/s2) = 2.866483795
do site  5  of  44896
get_building_states SD(mm) SA(g) [[ 14.2333]] [[ 0.282]]
get_building_states SD(m) = 0.0142333 SA(m/s2) = 2.7654753
do site  6  of  44896
get_building_states SD(mm) SA(g) [[ 14.2333]] [[ 0.282]]
get_building_states SD(m) = 0.0142333 SA(m/s2) = 2.7654753
do site  7  of  44896
get_building_states SD(mm) SA(g) [[ 14.2333]] [[ 0.282]]
get_building_states SD(m) = 0.0142333 SA(m/s2) = 2.7654753
do site  8  of  44896
get_building_states SD(mm) SA(g) [[ 11.6287]] [[ 0.3197]]
get_building_states SD(m) = 0.0116287 SA(m/s2) = 3.135186005
do site  9  of  44896
get_building_states SD(mm) SA(g) [[ 12.4324]] [[ 0.3305]]
get_building_states SD(m) = 0.0124324 SA(m/s2) = 3.241097825
do site  10  of  44896
get_building_states SD(mm) SA(g) [[ 18.252]] [[ 0.2639]]
get_building_states SD(m) = 0.018252 SA(m/s2) = 2.587974935
```

Figure 3.1 Log file screenshot of EQRM

For version control, during the development process and as a repository for publishing the source code for a user community, the development documents and the code are stored in a Subversion repository. Subversion is an open source version control system. It allows users to keep track of and document changes made over time to any type of electronic data, and to provide controlled access to files. The source code has been uploaded to the open source software website SourceForge.net which is geared towards getting developers involved in open source software projects.

3.4 Exposure Module

The user has to provide a building inventory which contains, for each building or building class, information about the postcode, the suburb, the coordinates of the building (or class of buildings) and the attributes of the building/building class (construction type, occupancy classification, total floor area in square meters, number of buildings, replacement cost). This software supports the HAZUS model building types with some further subdivisions recommended by Australian engineers for Australian building construction types. In total, 56 possible construction types are available. Notwithstanding that the model building types are user-defined as long as the user can supply capacity and fragility parameters. The occupancy is considered in the program because EQRM requires knowledge of the building's use in order to estimate the cost. For example, the value of the contents within an industrial building will vary from the value of the contents in a residential house.

The inventory is stored in a unique *.csv file which is very convenient. In this way it is simple to transform a general preliminary inventory into the format required by EQRM. The level of detail is user-defined, the unit could be a regional unit, a provincial unit, a local unit and it could be an arbitrary polygon (regular or irregular). The user can independently configure the inventory which is not hard-coded in any way.

3.5 Hazard Module

EQRM allows the seismic ground-motion to be estimated for a deterministic event (where the user must define the event and the ground-motion prediction equation, GMPE, to be used) or for multiple events that are generated from the input PSHA input model through Monte Carlo simulation. For what concerns the GMPE and the amplification factors the user can choose to consider or not the variability using the setdata.par file. EQRM currently considers only ground shaking and it does not consider different types of hazard such as landslides, tsunami, floods, hurricanes, etc.

EQRM uses the full response spectrum for intersecting with the building capacity curve rather than a design spectrum based only on a few building periods as is done in HAZUS. The advantage of this approach is that the full response spectrum and all the available information about the soil amplification factors, at all periods, is taken into account. Moreover, the user has also the option to input a design spectrum method using the setdata.par file.

EQRM requires an external *.csv file for the hazard input. The first two rows will be comments while the third row has to contain the value of the period for which the hazard is calculated. Then, the user can input the response spectral values. The user has to remember that it should be one row of spectral accelerations, in g, for each record contained in the inventory file. Therefore, if there are 1000 building records in the inventory, for instance, the hazard file will have 1000 rows of spectral acceleration information. There is no identification code in this input file to join the inventory table with the hazard table and so the user has to prepare the hazard in the same order as the inventory table. For example, the first row of the hazard table which contains response spectral values will join with the first row of the inventory. In some cases the user will thus have to pay a great deal of attention in preparing this input file.

3.6 Vulnerability Module

EQRM estimates the performance of buildings using the Capacity Spectrum Method. The capacity curve is not user-defined in EQRM though the user can modify the source code by adding their own functions to calculate the capacity curve. In the tested version, to compute the capacity, the user has to supply some building parameters to estimate the yield point and the ultimate point of the curve. These parameters are the design strength coefficient (C_s), the natural elastic building period (T_e), the fraction of building weight participating in the first mode (α_1), the fraction of the effective building height to building displacement (α_2), the over-strength factor-yield to design strength ratio (γ), the over-strength factor-ultimate to design strength ratio (λ) and the ductility factor (μ). These values are stored in a single table in the database of the program and are given for several classes of building construction types. In the following equations the yield point and the ultimate point are given in terms of the previous parameters.

$$A_y = \frac{C_s \gamma}{\alpha_1} \text{ [g]} \tag{3.1}$$

$$D_y = \frac{1000}{4\pi^2} 9.8 A_y T_e^{\,2} \text{ [mm]} \tag{3.2}$$

$$A_u = \lambda A_y \text{ [g]} \tag{3.3}$$

$$D_u = \lambda \mu D_y \text{ [mm]} \tag{3.4}$$

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 (See Figure 3.2). An exponential function is used to represent the curved part of the building capacity curve.

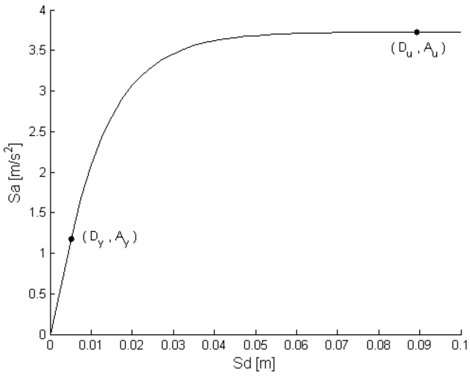


Figure 3.2 Capacity curve used by EQRM for the model building type C1L moderate code

The seismic performance of a building, which is the response of the system studied, is represented by the performance point. The way EQRM calculates the performance point is slightly different from SELENA even though it uses the capacity spectrum method. The main difference consists in the calculation of hysteresis area which is a fundamental parameter to estimate the final damping. EQRM uses the area under the smooth hysteresis curve whereas SELENA uses a parallelogram obtained from the bilinear approximation of the smooth pushover curve. For this reason the final results of the test-bed application were found to be slightly different (see Section 12, where the results of the different approaches are described in detail).

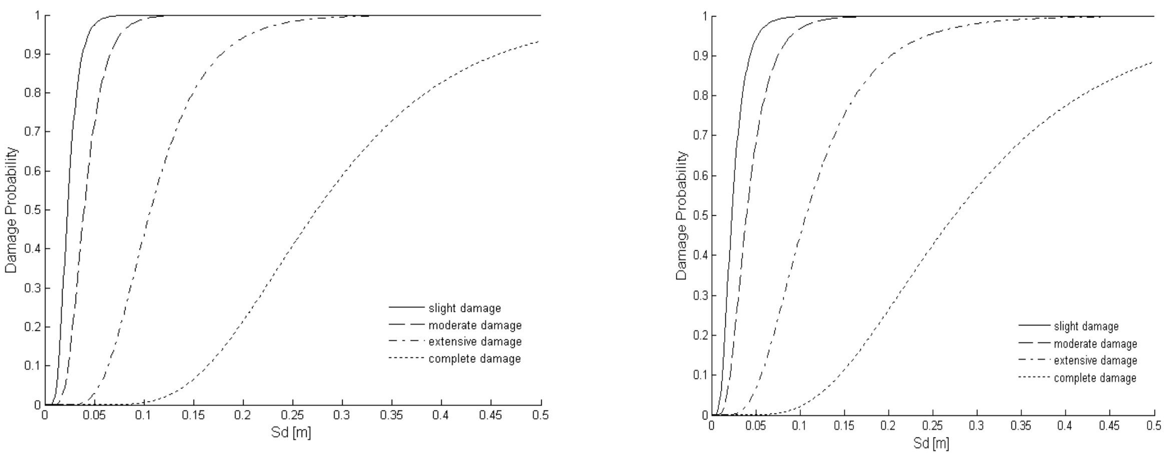
To input the capacity of the model building types the user has to supply just one *.csv table with a few building parameters. These parameters are those useful to compute the yield point and the ultimate point of the capacity curve which are described above, the degradation factors which are a function of short, moderate and long earthquake duration and the elastic (pre-yield) damping.

EQRM uses fragility curves to calculate the damage probability distribution for a class of structures. The fragility curves are lognormally distributed. The median value of spectral displacement at which the building reaches the damage states slight, moderate, extensive and complete are the values loaded into the database of the program, while the standard deviation for structural damage is taken equal to 0.4 [FEMA, 1999] for each curve. Moreover, if the user wants to include also the variability of the capacity curve, the standard deviation is taken equal to 0.5 for each curve. The user can choose the methodology in the setdata.par file where var_bcap_flag=0 means that the variability with the building capacity curve is not included and var_bcap_flag=1 means that it is included (Figure 3.3). Then, the fragility curves are implicitly evaluated by the code. To change the distribution and the variability, the user has to change the source code.

For what concerns non-structural fragility curves, the standard deviation is 0.5 for non-structural drift sensitive and 0.6 for non-structural acceleration sensitive. In Figure 3.3 the fragility curves for structural damage used for a low-rise concrete moment resisting frame building designed with a moderate code are shown.

The parameters defining the fragility functions for a given model building type for structural damage are connected to its capacity curve, therefore, the user has to add to the fragility parameters to the capacity parameters table described above. These values are the median value of spectral displacement at which the building reaches the different damage states and the median value with regards to non-structural damage (drift sensitive and acceleration sensitive).

The essential code to join the building parameters table with the inventory is the column called ‘structure classification’ which contains the name of the model building type which is also reported in the inventory table.



(a) (b)

Figure 3.3 Fragility curves used by EQRM: (a) variability in capacity is not included (b) variability in capacity is included

As carried out in SELENA, the code enters the fragility curves with the performance point so that it can compute the percentages of damage in each damage state.

The economic losses are defined as the cost involved in replacing damaged building components and/or contents. The replacement cost of the building can be studied for the structural, non-structural drift sensitive and non-structural acceleration sensitive components. The proportion chosen for each building component is a function of the building construction and usage type as well as the usage classification system. For example Table 3.1 illustrates the proportion of the replacement value corresponding to a couple of different buildings for both the HAZUS usage classification [FEMA, 1999] and the functional classification of buildings (FBC) [Australian Bureau of Statistics, 2001 (www.abs.gov.au)]. The setdata.par file can be used to select the usage classification system and consequently the desired cost proportions. To calculate the economic loss, the user has to input, as external *.csv files, two different tables: one with the percentages used for HAZUS classification and one with the percentages used for FCB classification. These files contain the information shown in Table 3.1 for each building type.

Table 3.1 Example of the damage values calculated by EQRM

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	
111 - W1BVTILE	
structural	23.44%
non-structural drift sensitive	50.00%
non-structural acceleration sensitive	26.56%
491-C1MSOFT	
structural	13.79%
non-structural drift sensitive	34.48%
non-structural acceleration sensitive	51.72%

In this way the economic losses 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 total loss of the building and its respective components, though excluding contents, is calculated through the following equations:

$$L = L_1 + L_2 + L_3 \tag{3.5}$$

$$L_1 = C_o \sum_{\alpha=1}^{4} f_{\alpha,1} R_1 AP_{\alpha,1} = \sum_{\alpha=1}^{4} f_{\alpha,1} g_1 RAP_{\alpha,1} \tag{3.6}$$

$$L_2 = C_o \sum_{\alpha=1}^{4} f_{\alpha,2} R_2 AP_{\alpha,2} = \sum_{\alpha=1}^{4} f_{\alpha,2} g_2 RAP_{\alpha,2} \tag{3.7}$$

$$L_3 = C_o \sum_{\alpha=1}^{4} f_{\alpha,3} R_3 AP_{\alpha,3} = \sum_{\alpha=1}^{4} f_{\alpha,3} g_3 RAP_{\alpha,3} \tag{3.8}$$

Where:

- L_1, L_2, L_3 are the financial loss of the structural damage, drift sensitive non-structural damage and acceleration sensitive non-structural damage, respectively;
- $P_{\alpha,I}$ is the probability of being in a damage state α = (1,2,3,4) corresponding to slight, moderate, extensive and complete damage. The index *i* corresponds to the type of damage: drift sensitive structural damage, drift sensitive non-structural damage and acceleration sensitive non-structural damage;
- R_i denotes the replacement cost component of the building per unit floor area. Thus R=R1+R2+R3 is then total replacement cost of the building (per unit floor area) excluding contents;
- A is the floor area of the building in squares metres;
- $f_{\alpha,1}$ is a repair cost as a fraction of the replacement value. They are taken from FEMA [1999] and they are different for structural damage, non-structural damage (drift-sensitive) and non-structural damage (acceleration-sensitive);
- C_o is a regional cost factor to calibrate the replacement costs if necessary.

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

The contents damage L_4 is based only on the probabilities for acceleration sensitive non-structural damage and it is calculated by the following equation:

$$L_4 = C_o \sum_{\alpha=1}^{4} f_{\alpha,4} R_4 P_{\alpha,3} \tag{3.9}$$

The cost L_4 is then added to the building damage cost L to get the overall loss L* which includes contents.

The economic losses are independent of social losses (not considered in the code) and they are linked to the level of damage suffered by the structure.

3.7 Output

Five structural damage states are considered in EQRM: no damage, slight damage, moderate damage, extensive damage and complete damage. In the output file, no damage is not reported but it can be easily deduced. The final results are given as the percentage of buildings in each damage state for each model building type. The number of damage grades is hard-coded and to change the level of damage the user has to change the code. The output files are ASCII tables and they report the results for each model building type for each geounit and with this the user can easily compute the total damage for each unit and for the total region of interest.

For what concerns the economic losses, the software provides three ASCII files: one for the economic losses of the building excluding contents (L), one for the content losses (L_4) and the last one for the overall losses (L*). The losses are reported in a user-defined currency.

EQRM offers some external tools to visualize results which are given in *.csv format. These tools are written in MATLAB language and the user can use them as long as he/she has a MATLAB license. Notwithstanding that, the users can easily use an external geographic information system to plot their results. It is worth noting that the EQRM team is currently working towards converting the MATLAB plotting tools to Python.

4 ELER

4.1 Summary of Software

Under the JRA-3 component of the EU FP-6 NERIES Project, a methodology and software (ELER - Earthquake Loss Estimation Routine) for the rapid estimation of earthquake shaking and losses in the Euro-Mediterranean region has been developed. This multi-level methodology is capable of incorporating regional variability and sources of uncertainty stemming from ground-motion predictions, fault finiteness, site modifications, inventory of physical and social elements subjected to earthquake hazard and the associated vulnerability relationships. Although primarily intended for post-earthquake rapid loss estimation, the routine is also equally capable of developing scenario earthquake-based loss assessments.

ELER v2.0 is the version of the software that was made available to the GEM1 Risk Team. It has to be considered that this software is not currently distributed as an open source software in the same way as other codes. It is a compiled version using the MATLAB code and tools. Since the program is compiled the user does not need to have a MATLAB license to execute the analysis, but it also means that the user cannot make modifications to the code. The source code of ELER was shared with the GEM1 team, however. ELER has a Graphical User Interface(GUI) to input the data and to present the output. The output consists of maps and tables presenting general (aggregated) results. With regards to the hazard tool, the user is allowed to download ground-motion fields in *.kml format or *.shp format. For what concerns losses, the software provides data as *.shp files.

The ELER software mainly has two modules which are the Earthquake Hazard Assessment module and the Earthquake Loss Assessment module. The Earthquake Hazard Assessment includes estimation of ground motion and intensity distributions using ground-motion prediction equations, correlation between intensity and ground-motion parameters and soil condition information. The Earthquake Loss Assessment module uses the ground motion and intensity information from the Earthquake Hazard Assessment module, together with a demography and building inventory. Anyway, the user can upload their own external ground-motion field or hazard following the required input format. This module includes three levels (Level 0, Level 1 and Level 2) of analysis for the estimation of building damage and/or casualties. The Level 0 analysis estimates casualties based on magnitude and macroseismic intensity information with a seismological method. The Level 1 analysis estimates casualties and building damage based on macroseismic intensity information with an empirical engineering-based method. The Level 2 analysis estimates casualties and building damage based on ground-motion and spectral parameters (with an analytical engineering-based method).

It has to be mentioned that a new version of ELER, ELER v3.0, is under development and it contains some interesting improvements compared to ELER v 2.0. For instance, a new tool that allows the user to upload their own GMPE is provided. Then, economic losses (Level 1 and Level 2) and lifeline damage based on peak ground velocity distributions can be computed by the program. Moreover, there is a new tool for Level 2 building database formation. The Xls2BDB tool converts a normal Excel document to a Level 2 building database file accepted by ELER. Finally, performance optimizations in some functions have improved the overall calculations times by 100-150% for some methods.

4.2 Methodology

The basic methodologies used in the ELER software are divided into three levels of analysis; a schematic view of the analysis procedures is given in Figure 4.1.

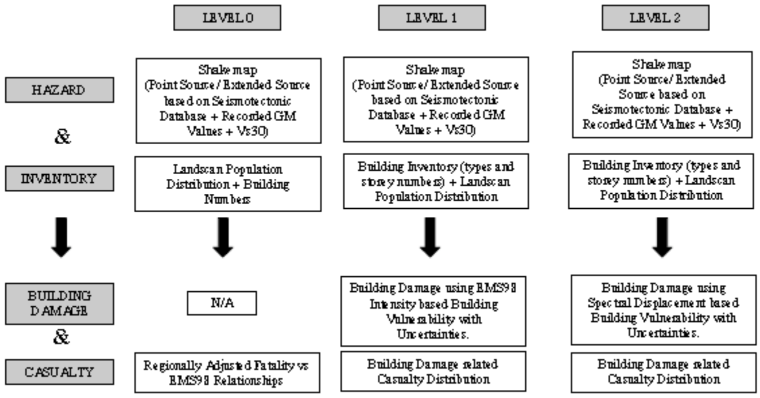


Figure 4.1 Flow chart for multi-level analysis methodology of ELER [Kamer *et al.*, 2009]

Both the Level 0 and Level 1 analysis of the ELER software are based on obtaining intensity distributions analytically and estimating the total number of casualties either using regionally adjusted intensity-casualty or magnitude-casualty correlations (Level 0) or using the damage calculated for a regional building inventory database (Level 1).

The Level 2 analysis of the ELER software is similar to HAZUS methodology and is essentially intended for earthquake loss assessment (building damage and consequential human casualties) in urban areas. It employs a variety of analytical techniques that include: Capacity Spectrum Method [ATC, 1996], Modified Acceleration-Displacement Response Spectrum Method [ATC, 2005], Reduction Factor Method [Fajfar, 2000] and Displacement Coefficient Method [ATC, 2005]. The building taxonomy is based on the European building typology developed within the EU-FP5 RISK-UE project as well as the model building types of HAZUS-MH [FEMA, 2003].

4.3 IT Details

In ELER a MATLAB programming environment is used both for the computational tool as well as for the display of the results. As mentioned before, ELER v2.0 is a compiled software in the MATLAB environment; therefore, there is no need to install the MATLAB software. Since MATLAB is a cross-platform programming language, ELER can run on the following operating systems without any code modification: Windows (x64), Linux (x86-64), Mac OS X and Solaris 64.

The compiled version of ELER has a GUI from where the input and output procedures can be performed; however, there is not a command line interface if a user would like to perform automated, scheduled or event triggered runs. There is also a companion screen which opens together with the GUI screen, and which is a log window and shows event messages related to the processes being carried out.

Input files have different formats depending on the analysis being carried out. For instance, in the Hazard section, ground-motion fields can be created by inputting the event data and the actual station recording (if available) by using *.xml files or manually. The Risk section comprises three levels of analysis: Level 0 uses as input *.mat files to define the Intensity Grid and the Source, which is a point vector of the epicentre [lat, lon], or a *.mat file containing the fault matrix, or a text file containing the fault coordinates (see Figure 4.2a). Level 1 needs *.mat files to account for the Intensity Grid, as in Level 0, the "VQ Table" which is a table containing the vulnerability and ductility values for each building type, and the Building Database which is a shapefile (*.shp) containing the number of buildings of each building typology in each grid cell (see Figure 4.2b). Level 2 makes use of *.mat files to define the shear wave velocity and spectral acceleration values together with a *.shp file (the Building Database) containing the number of buildings of each building typology in each grid cell (see Figure 4.2c).

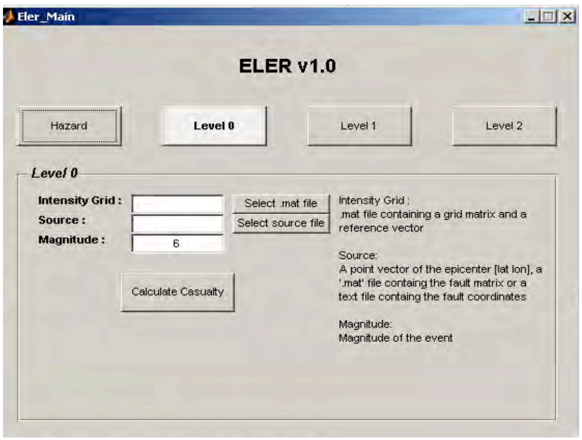
The *.mat files are composed of:

- a grid matrix of intensities, shear wave velocities or spectral acceleration in %g (Intensity Grid, V_{s30} or SA, respectively);

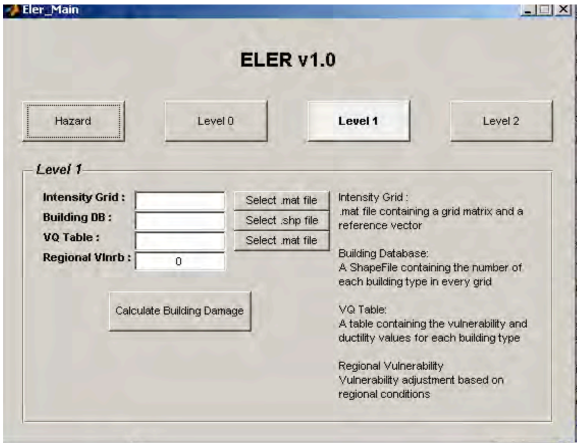
- a reference vector where the first element of the vector defines the number of cells per degree while the second and third elements specify the latitude and longitude of the upper left corner of the grid;
- the magnitude of the earthquake;
- the coordinates of the source.

In Figure 4.3 an example of the components of the *.mat file is shown.

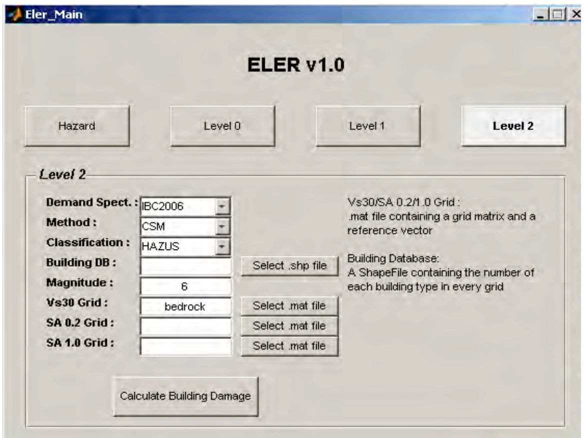
Note that grid format must follow the MATLAB convention to manage Regular Data Grids which by definition are rectangular, not sparse, matrices that contain double values. MATLAB stores them in column order, with their southern edge as the first row and their northern edge as their last row.



(a)



(b)



(c)

Figure 4.2 Input GUI screens for a) Level 0, b) Level 1 and c) Level 2 of ELER

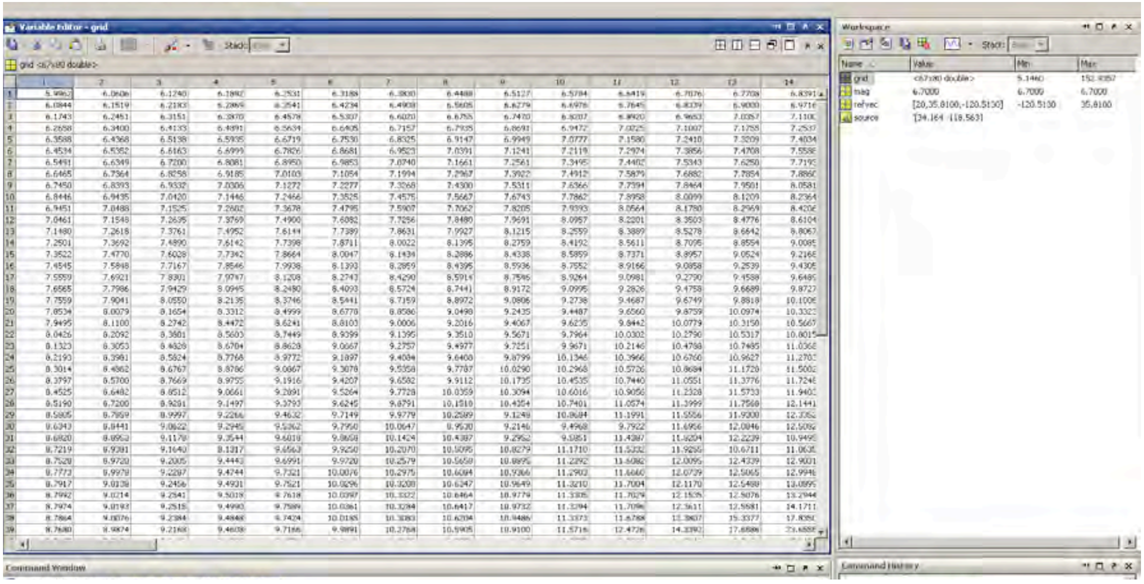


Figure 4.3 Components of the *.mat file of ELER for the SA 0.2s grid

4.4 Exposure Module

The data contained in the inventory file varies depending on the level of analysis the user decides to perform.

In Level 0, default inventory consists of population density (LandScan^{TM} Population Distribution Data at a 30 arc-second resolution), city names, locations and populations for the whole Euro-Med region. This information is already saved in the program; therefore, the user does not need to provide any inventory.

In Level 1, the building inventory data required for analysis consist of a grid-based building distribution, classified in terms of a RISK-UE or European building taxonomy. The building database file is basically a shapefile (*.shp) containing the building distribution for each cell. Additionally this file may contain the population of each cell for the computation of casualties. If there is no Population field in the building database, casualty estimations are calculated with the approximation of the regional population (obtained from the LandScan population distribution).

Finally, Level 2 analysis is essentially intended for earthquake loss assessment (building damage and consequential human casualties) in urban areas. As such, the building inventory data for the Level 2 analysis consists of grid-based urban building (RISK-UE and HAZUS taxonomy) and demographic inventories. In Table 4.1 the structure of the building database shapefile is presented. This table is the same that can be used for Level 1.

Table 4.1 Structure of building database shapefile for ELER

ID	Model building types (W1_PC; S1L_MC, etc.)	Total_Bld	Population	Pop_Day
1	The number of each building type should be given in its corresponding field	Total number of buildings in each cell	If this field exists casualty will be calculated according to each cells population, otherwise regional population approximations will be used	Number of people located in the building during the day

4.5 Hazard Module

The Hazard Module in ELER can be run independently or together with the loss assessment modules, though the required ground-motion parameters for each level of loss assessment can also be provided from an external file. The computation of ground-motion parameters in the Hazard module is based on, but not limited to, the ShakeMap methodology developed by Wald *et al.* [1999a; 2003, 2006]. The user specifies all parameters, options and modes of the ground motion computation through the graphical user interface of the Hazard module.

The input specification of the Hazard Module is done by the “Hazard” button in the main GUI window (Figure 4.2). The flowchart of hazard and a snapshot of the Hazard GUI are given in Figure 4.4a and b, respectively.

The user has two options to enter the event data. The first option is to use an *.xml file containing the event parameters and station information, if available. The second option is to enter event data manually from the Hazard GUI.

The Source Type panel defines the source mechanism associated with the event. For small magnitude events the source can be given as a point, for large magnitude events the user can specify the source type as a finite fault. In the Auto Assign option the user is required to select a fault database containing the candidate faults.

The Site Correction panel determines how the effect of the local site conditions will be incorporated into the calculations of ground-motion parameters, though the user can also decide not to apply any correction. When using the NGA at surface choice in the menu, rather than calculating bedrock values and then amplifying these with respect to site conditions, ELER software uses attenuation relationships taking V_{s30} as an input parameter to calculate the ground-motion values directly at the surface. For what concerns this V_{s30} Grid, ELER comes with a default site condition map covering the entire Euro-Mediterranean region. The default site condition map has been compiled from the USGS Global V_{s30} Map Server. Custom site condition maps should be in the form of V_{s30} grids. In MATLAB grids are defined by a matrix containing the values of each cell and a reference vector which is used to map each cell to its corresponding geographical location. The first element of the reference vector defines the number of cells per degree while the second and third elements specify the latitude and longitude of the upper left corner of the grid.

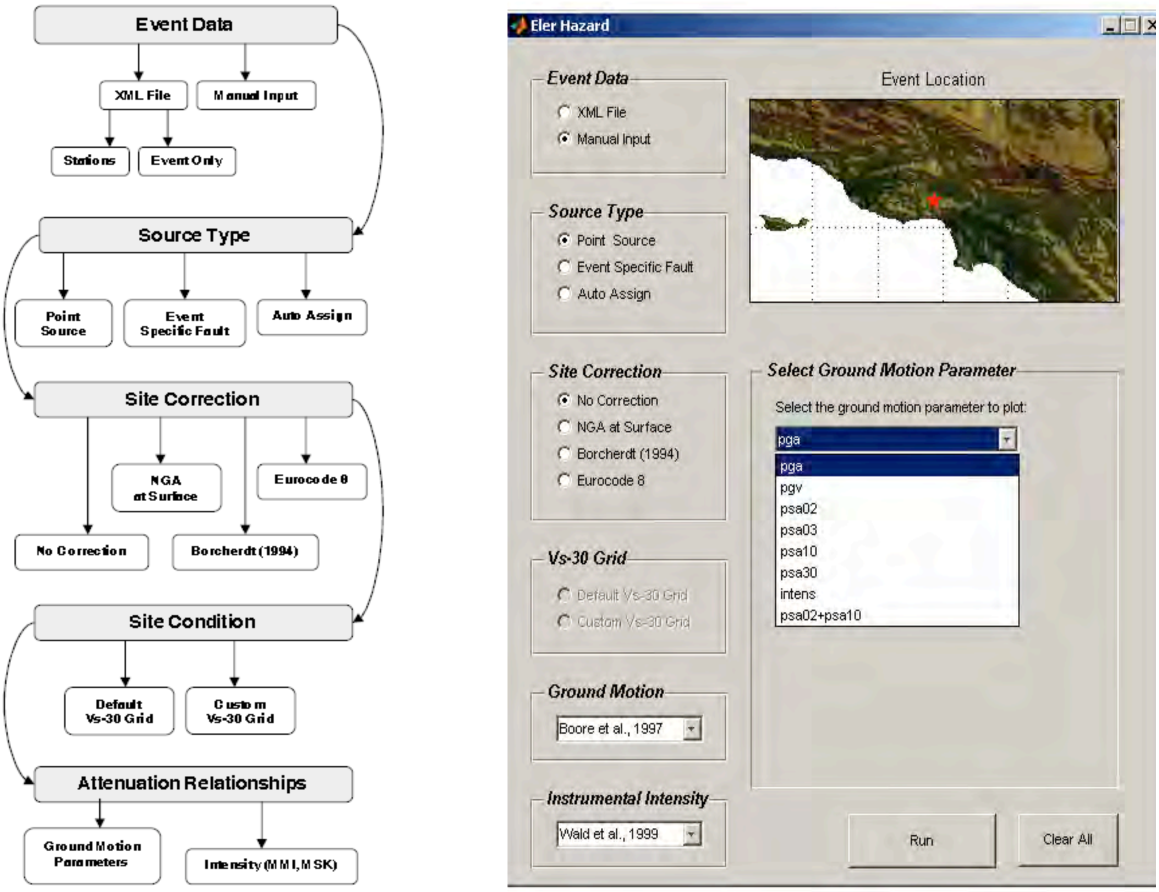
Then, the user should select a ground-motion prediction equation (GMPE) which is used to estimate instrumental ground-motion parameters such as PGA and spectral accelerations. Each GMPE has its unique set of input parameters resulting from the regression analysis; the common parameters such as event magnitude, distance to source and site condition are set automatically whilst the remaining parameters such as fault type, hanging wall effect etc. should be specified by the user.

In the Instrumental Intensity Estimation drop down menu, the user can select the method to use to convert the instrumental ground-motion parameters to macroseismic intensity.

The software produces both ground-motion fields (MSK’98 Intensity) and maps for ground-motion parameters such as PGA, PGV and SA at specific periods. They can be saved in different formats such as *.mat, *.kml or *.shp.

It is worth noting that in ELER the hazard values may or may not be located at the centroid of the geocell employed in the inventory file; therefore, there is no need to change the ground-motion field file if there is any variation in the size or location of the geocells.

A helpful tool has been developed by the authors to convert text file containing grid data in xyz format into MATLAB matrices to be used in ELER. This tool called Text2Grid can be used for creating ground-motion grids to be used in Level 0, 1 and 2 for creating custom site condition maps for use in the Hazard module or for creating ground-motion fields in the required format. Skipping the Hazard module, the user can create their own hazard in a simple way using this converter.



(a) (b)

Figure 4.4 (a) Flowchart of the Hazard GUI [Kamer *et al.*, 2009], (b) a screen shot of the Hazard GUI

4.6 Vulnerability Module

The vulnerability in ELER is computed in different ways according to the level the user is running.

Level 0 uses magnitude-casualty and intensity-casualty correlations that directly correlate social losses to the ground-shaking intensity or the magnitude of the event. The empirical approaches that are used in this level to estimate the human casualty are:

- Vacareanu *et al.* [2004] Approach. It relates the number of deaths to the magnitude of the earthquake. A correlation of human losses with the earthquake magnitude is provided in Equation (4.1). Representing human losses from several earthquakes that occurred in different countries as a function of the earthquake magnitude, a correlation has been derived:

$$D = ce^{1.5M} \tag{4.1}$$

  Where:

  - D is the number of deaths;
  - M is the magnitude of the earthquake;
  - c is a coefficient assuming different values for lower, median and upper bounds (respectively c = 0.002, c = 0.06, c = 0.4).

As can be seen in Figure 4.5 uncertainties in such correlations can reach up to two orders of magnitude.

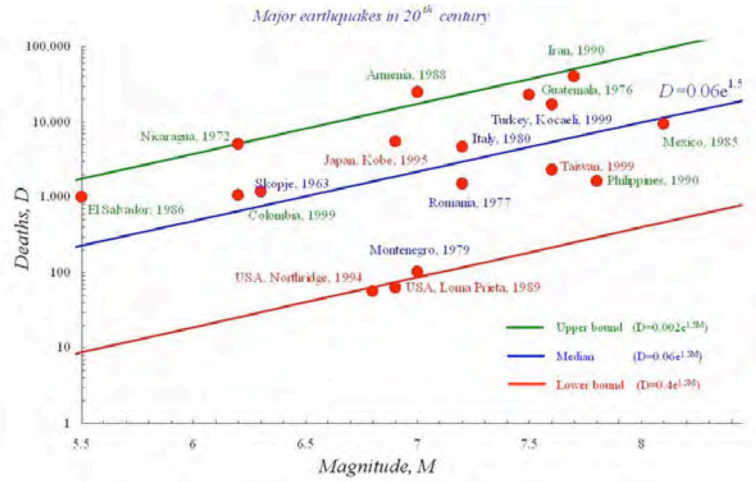


Figure 4.5 Correlation of deaths with earthquake magnitude [Vacareanu *et.al*., 2004]

- RGELFE [1992] Empirical Approach. Based on empirical data, RGELFE [1992] provides for major cities, the following fatality rates for various levels of intensity: 0.0014%, 0.031%, 0.48% and 6.8% for intensity zones VI, VII, VIII and IX respectively. As can be observed from that study, this kind of crude casualty prediction models exhibit a high level of regional dependency.
- Samardjieva and Badal [2002] Empirical Approach. Samardjieva and Badal [2002] conducted a regression analysis between the number of casualties (in log scale) and the earthquake magnitude for different ranges of population density in the affected region using events from 1945 till 2000.

It has to be kept in mind that for this level of analysis the default inventory consists of population density (LandScan Population Distribution Data, 30 Sec-arc for the Euro-Med Region), city names, locations and populations, for the Euro-Med Region.

Level 1 loss estimation engine of ELER is based on macroseismic damage estimation tools and aims at the assessment of both the building damage and casualties. The intensity-based empirical vulnerability relationships of Giovinazzi and Lagomarsino [2005] are used for damage assessment and they are described in detail in Section 5. To convert Excel files containing vulnerability and ductility values of buildings types into MATLAB arrays that are necessary to implement the empirical vulnerability relationships, the Xls2Mat tool is provided by the authors of ELER. The software will check the building database for the building types specified in this file. Then, in order to calculate casualty estimation the knowledge of building occupancy data and the probability of several levels of injury and death for different building types with given states of building damage is required. This however, is not easily attainable due to the limited quality and lack of information on earthquake casualty data. There are three approaches used in this level:

- Coburn and Spence [1992] Method. For the estimation of the fatalities due to structural damage (the Ks parameter), which is the controlling factor for most destructive earthquakes, Coburn and Spence [1992] proposed the equation given below:

$$K_{sb} = TC_b \times \left[M_1 \times M_2 \times M_3 \times (M_4 + M_5(1 - M_4)) \right] \tag{4.2}$$

Where

- TC_b is the total number of collapsed buildings of type b;
- M_1 is the factor taking into account regional variation of population per building;
- M_2 is the factor taking into account variation of occupancy depending on the time;
- M_3 is the factor taking into account percentage of trapped occupants under collapsed buildings;
- M_4 is the factor taking into account different injury levels of trapped people;

– M_5 is the factor taking into account change of injury levels of trapped people with time.

The number of buildings and population in each cell are the main parameters to estimate casualties. To obtain casualty estimations from the number of buildings in different damage states, an average number of population per building should be known. To estimate this, the user should define an average number of dwelling units per building type, which is usually a function of the number of floors. Using the user-defined average number of dwellings per building type and the grid based population data entered by the user, or the default LandScan population data of the region, the program computes an average number of population per dwelling unit, which in turn can be used to check if the estimated number of dwellings per building type was correct.

- Risk-UE Casualty Vulnerability Relationships. The casualty vulnerability relationships used in the Risk-UE project are based on the findings of Bramerini *et al.* [1995] that studied the statistics on casualties, severely injured and homeless people in Italy.
- KOERI [2002] Method. Casualty rates, especially deaths, depend largely on the probability of the building being in the "complete" damage state. Casualty data in urbanized areas from Turkish earthquakes indicate much higher fatalities in heavily damaged multi-storey RC buildings. Data from the 1992 Erzincan earthquake indicate 1 death and 3 hospitalized injures per collapsed or heavily damaged RC building [Erdik, 1993]. Similar statistics are also valid for the 1999 Kocaeli earthquake. About 20,000 RC buildings were collapsed or heavily damaged and the total dead count was around 19,000. The death to hospitalized injury ratio in this earthquake was 1:2.5. For the assessment of human casualties from damage data computed from intensity based vulnerabilities, ELER assumes that the number of deaths will be equal to the number of buildings with damages in D4 and D5 level. The number of hospitalized injuries is found by multiplying the death figure by 4 based on ATC-13 [ATC, 1985] recommendations.

Level 2 analysis is essentially intended for earthquake loss assessment (building damage, consequential human casualties and macro economic loss quantifiers) in urban areas. In ELER the seismic performance of a building is represented by the performance point. The user can choose four different methodologies to calculate this point: Capacity Spectrum Method, Modified Acceleration-Displacement Response Spectrum Method, Reduction Factor Method and Displacement Coefficient Method. The difference between these methodologies is in the calculation of the final reduction factor to modify the elastic spectrum. This value is fundamental to estimate the performance point. The user can choose the methodology to employ using the GUI for Level 2 (Figure 4.2). For what concerns the capacity spectrum, the ELER code needs the yield value and the ultimate value for each model building type and it calculates a bilinear curve which has an initial linear section where the slope depends on the typical natural frequency of vibration of the building class, and rises to a plateau level of spectral acceleration at which the maximum attainable resistance to static lateral force has been reached. As an example, a capacity spectrum is shown in Figure 4.6. As can be seen, the capacity spectrum is controlled by the points of design, yield and ultimate capacities. These points can be correlated with the damage limit states.

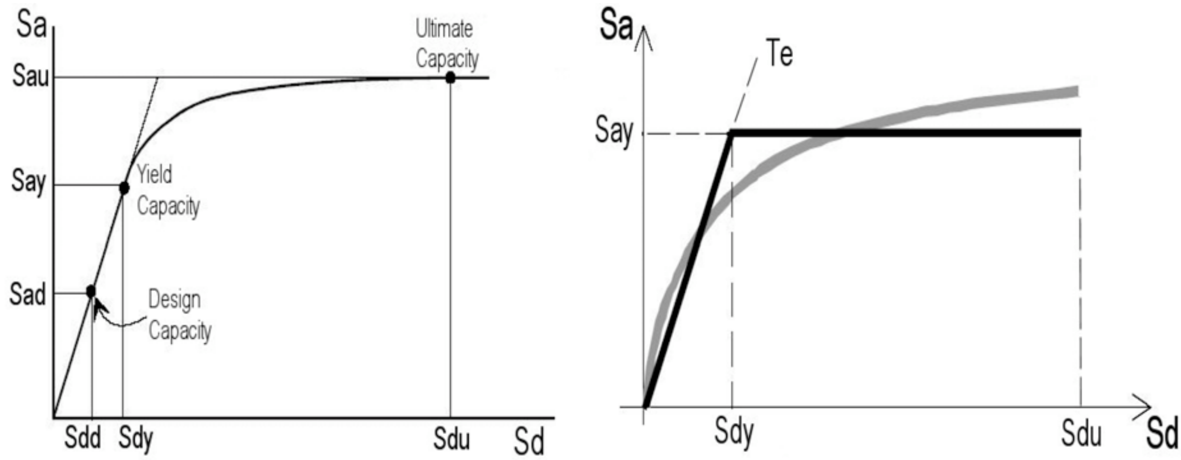


Figure 4.6 Typical structural capacity spectrum (left) and its simplified form (right) [Demircioglu *et al.*, 2009]

For the RISK-UE building typologies and model building types of HAZUS'99, the capacity curve parameters as described above are already provided in the ELER database. These parameters are stored in internal MATLAB tables. An example for three different model building types designed with pre-code and low-code is shown in Figure 4.7. Although not shown here, the values for structures designed with moderate and high codes are also stored. The user can easily change the values of all of these parameters. Furthermore, the user can easily upload an external capacity curve point by point.

```
BuilgingCapacityData = [
%    | pre-code                                          | low-code
%    +---------------------------------------------------+-----------------------------------------------------
%#   | dy     ay     du     au     de     beta0 alpha   | dy     ay     du     au     de     beta0 alpha |
%      (m)    (m/s2) (m)    (m/s2) (m)                     (m)    (m/s2) (m)    (m/s2) (m)
1      0.0061 1.9620 0.1097 5.8860 0.0043 15    0         0.0061 1.9620 0.1097 5.8860 0.0043 15    0
2      0.0041 0.9810 0.0597 2.4525 0.0028 10    0         0.0041 0.9810 0.0597 2.4525 0.0028 10    0
3      0.0038 0.6082 0.0699 1.8345 0.0027  5    0         0.0038 0.5886 0.0582 1.8639 0.0027  5    0
```

Figure 4.7 Capacity curve parameters of the model building type of HAZUS provided in the database of ELER

ELER uses fragility curves to calculate the damage probability distribution of classes of structures. In ELER, fragility curves are lognormally distributed and the type of distribution is hard-coded. The lognormal standard deviation, which describes the total variability of the fragility curve, has been modelled by the variability of the capacity curve, demand spectrum and damage state threshold. This is the same approach used by HAZUS. The values related to HAZUS and RISK-UE building type are already calculated and summarized in MATLAB tables. An example of the building fragility database for the HAZUS classification for structures designed with pre-code and low-code is shown in Figure 4.8. Although not shown here, the values for structures designed with moderate and high codes are also stored. It is worth noting that ELER v2.0 provides a tool for the users to create their own building parameters database which is different from the tables hard-coded in the software itself. This tool called Building Database Formation is shown in Figure 4.9. As it can be noticed, the user can insert the definition of building types, analytical methodology (each method has its own parameter to be defined), definition of fragility parameters for each damage state, definition of building capacity parameters and definition of additional parameters.

```
%% Building Fragility Database
BuildingFragilityData = [
%   | pre                                           | low
1     0.40 1.01 1.00 1.05 3.09 1.07  7.56 1.06     0.50 0.93 1.25 0.98 3.86 1.02  9.45 0.99
2     0.69 1.04 1.71 0.97 5.29 0.90 12.96 0.99     0.86 0.97 2.14 0.90 6.62 0.89 16.20 0.99
3     1.04 0.85 1.65 0.82 3.50 0.80  8.64 0.95     1.30 0.77 2.07 0.78 4.38 0.78 10.80 0.96
```

Figure 4.8 Fragility curve parameters of the model building type of HAZUS provided in the database of ELER

The casualty estimation in Level 2 analysis is based on the HAZUS99 [FEMA, 1999] and HAZUS-MH [FEMA, 2003] methodology. The output from the module consists of a casualty breakdown by injury severity level, defined by a four level injury severity scale [Durkin and Thiel, 1993; Coburn and Spence, 1992; Cheu, 1995].

The HAZUS casualty rates were obtained by revising those suggested in ATC-13 [1985] using limited post-earthquake fatality data. The casualty model itself in fact is based on the models suggested by Coburn and Spence [1992], Murakami [1992] and Shiono et al. [1991]. However, unlike other approaches, the methodology is in event-tree format and thus is capable of taking into account non-collapse related casualties. To estimate the casualties from structural damage, the model combines a variety of inputs from other HAZUS modules including the probability of being in the damage state and the relationship between the general occupancy classes and the model building type with specific casualty inputs provided for each damage state (D1-slight, D2 moderate, D3 Extensive, D4 Complete, D5 complete with collapse structural damage) in combination with occupancy data and time event.

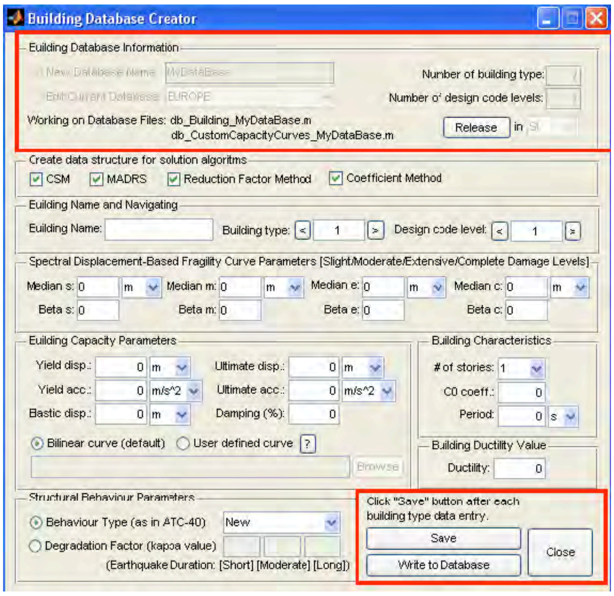


Figure 4.9 Building Database Creator tool

The probability of suffering i- severity (i=1:4) level is calculated by:

$$pSI = \sum_{k=1}^{5} w_{SI,k} p_k \tag{4.3}$$

Where:

- p_{si} is the probability for people involved in an earthquake to suffer a level *i*- severity (*i*=1:4);
- p_k, is the probability of a damage D_k,(*k*=1:5) occurrence;
- $w_{si,k}$ is the casualty rate considered for p_k probability.

The expected number of occupants in severity level *i* (EN_i) is the product of the number of occupants of the building at the time of earthquake (*Noccupants*) and the probability of an occupant suffering severity level *i* (P_{si}).

$$EN_i = Noccupants * P_{SI} \tag{4.4}$$

Casualty rates for reinforced concrete and masonry structures as given in HAZUS99 are provided in the ELER database.

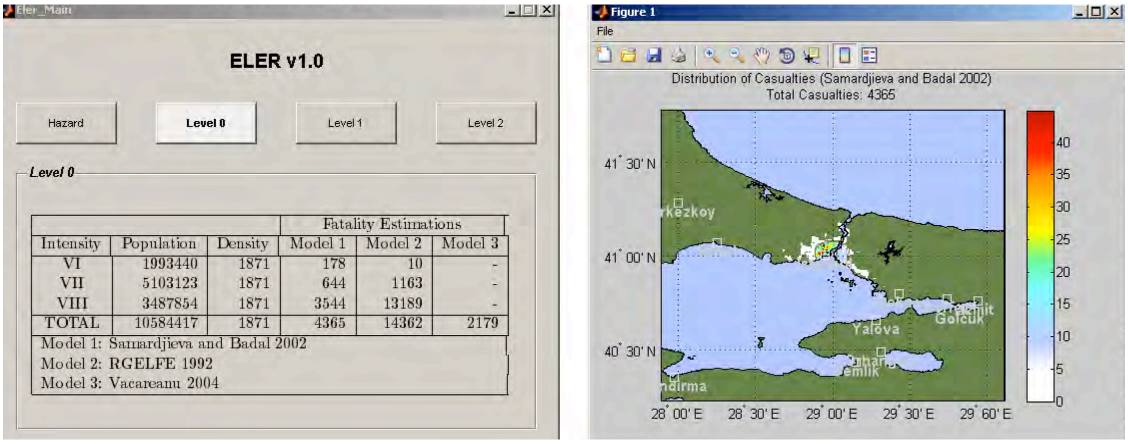
The methodology used in Level 2 for the estimation of number of casualties is the same methodology suggested by HAZUS [1999]. If, in addition to the grid based building inventory, a grid based population distribution is defined by the user, the software computes the number of dwelling units (using user defined estimated number of dwellings per building type) and an average population per dwelling unit for each cell. Then, casualties for any given building type, building damage level and injury severity level can be calculated by the following equation:

K_{ij} = Population per building*Number of Damaged Buildings in damage state j*Casualty Rate for severity level i and damage state j (4.5)

At present three casualty models are included in ELER. These are HAZUS [1999], HAZUS-MH and the KOERI casualty model for Turkey developed by Erdik and Aydınoglu [2002] using the casualty data from 1992 Erzincan and 1999 Kocaeli earthquakes. Mainly two building types are considered in all three models: RC and masonry. As in Level 1, if a user defined grid based population is not available, the program calculates an average population per dwelling unit for the whole study area using the default LandScan population and calculates casualties accordingly.

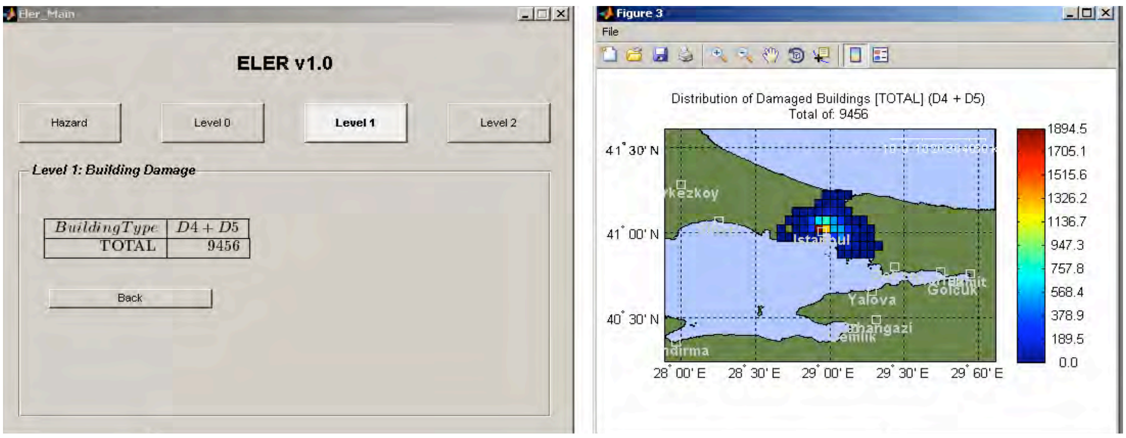
4.7 Output

The outputs are different according to the level the user is running. As mentioned before, Level 0 considers only the social losses and it is empirical based. The software gives a table with the summary of the results computed using different methodologies and a *.fig that visualizes the results (Figure 4.10). In Level 1, five damage states are considered: slight damage (D1), moderate damage (D2), substantial to heavy damage (D3), very heavy damage (D4) and Destruction (D5). The damage is directly derived by the intensity value and the levels are hard-coded. An example of building damage estimation output is given in Figure 4.11, while in Figure 4.12 an example of the social losses, which are linked with the damage suffered by the structure and which are computed using the KOERI casualty model is shown. Finally, Level 2 considers four structural damage states (slight, moderate, extensive and complete). These states are considered both for the building model types of HAZUS [1999] as well as for the European building taxonomy. Figure 4.13 shows the spatial distribution of damage for Los Angeles. Results are also presented in a table format showing the total number of buildings that have been damaged for a particular damage level and a plot showing the spatial distribution of damage. The program calculates the damage estimations only once and the results for all building types and damage states are stored in *.dbf file. The program then plots the results according to the selected building type and damage state. When the user selects a different building type ELER creates a new figure from the stored results.



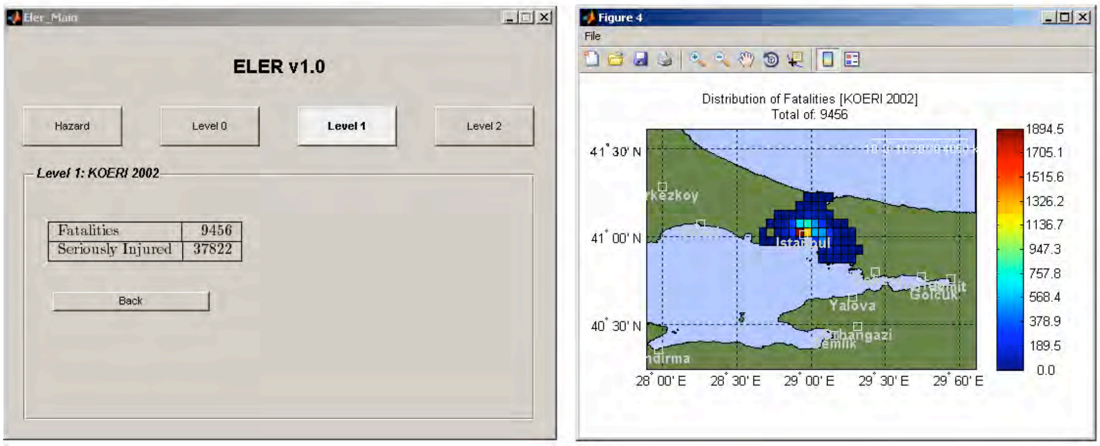
(a) (b)

Figure 4.10 Output windows for Level 0 (a) table with a summary of social losses (b) map showing the distribution of the losses



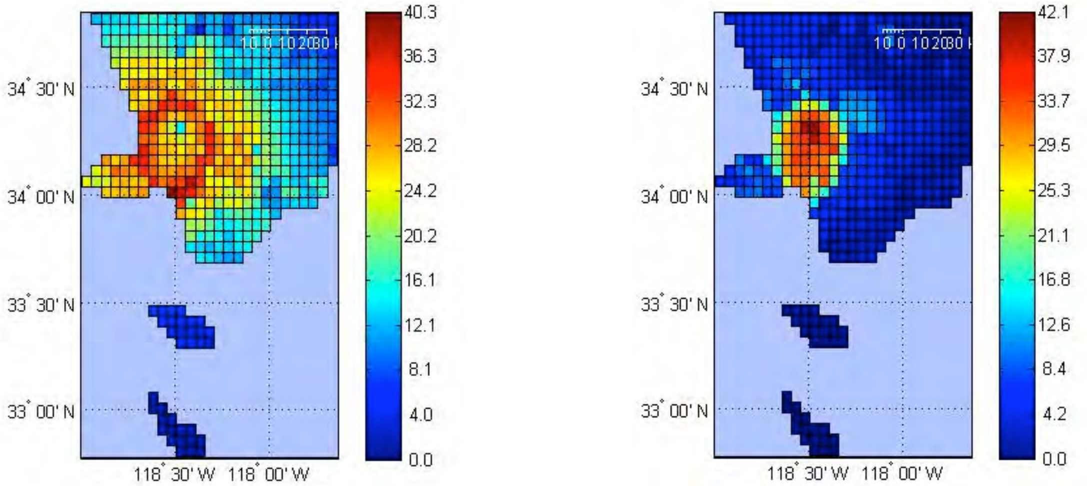
(a) (b)

Figure 4.11 Building damage output for Level 1 analysis option (a) table showing the total number of buildings which are in damage states D4 and D5 (b) map plotting the distribution of damage building belonging to the damage states D4 and D5



(a) (b)

Figure 4.12 Social loss output for Level 1 analysis option, employing the KOERI methodology (a) table showing the fatalities and seriously injured (b) map plotting the distribution of fatalities.



(a) (b)

Figure 4.13 Spatial distribution of damage for Los Angeles obtained with ELER (a) slight (b) moderate

5 QLARM

5.1 Summary of Software

The earthQuake Loss Assessment for Response and Mitigation (QLARM) software is a loss estimation tool developed by the World Agency of Planetary Monitoring and Earthquake Risk Reduction (WAPMERR) in collaboration with the Swiss Seismological Service (SED) at the Swiss Federal Institute of Technology Zurich (ETH). The World Agency of Planetary Monitoring and Earthquake Risk Reduction was created in 2001 in Geneva as a non-profit organization. It has the mission to estimate losses after earthquakes in near-real-time as rapidly and as accurately as possible. Subsequently, WAPMERR reports disasters to the Swiss Agency for Development and Cooperation (SDC) and to the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). They distribute loss estimates by e-mail in near-real-time for any earthquake with M ≥ 6 worldwide with the main scope being to trigger rapid humanitarian response at a worldwide scale, with a specific focus on developing countries. This is an open and free service. WAPMERR has used QUAKELOSS1 for estimating losses after earthquakes 2003-2009. Since 2010, QLARM is the second-generation loss tool to estimate building damage and human losses due to earthquakes anywhere in the world. A development installation of QLARM is available on the following website http://qlarm.ethz.ch. The Software is available as source code and as fully set-up virtual machine under Linux.

The version of QLARM used by the GEM1 Risk team is a prototype and the software and datasets are "open" though a user name and a password are compulsory to have access to data and tools. Anyone who is interested to the software can freely ask for an ID and a password. The IT framework of this software is developed all within the Eclipse development platform and a Subversion repository is used for version control. QLARM is coded with the object-oriented programming language Java. Its data tier is implemented using the relational database PostgresSQL and object-relational mapping based on Hibernate. Its user interface is implemented as web pages based on the XML publishing framework Apache Cocoon and a UMN-MapServer-empowered web mapping tool.

QLARM is a modular and flexible framework that can be extended and tailored by expert users to their needs. It offers:

- A data model, capable to host multiple "world data sets" (sets of exposure data, i.e. population, building stock and vulnerability, and site conditions), earthquakes, related ground motion, and loss results. Exposure-data is organized hierarchically and tagged with country, "settlement" (which may be any political subdivision carrying a name), and "settlement part" (which may be any physical subdivision of a settlement (also a single building) with known distribution of population and buildings into different building types). Within one world dataset, granularity of the information, population, and building stock properties may vary over time. In addition, population numbers may fluctuate in daily, weekly, and seasonal cycles.
- An interface for source description / source geometry. Implemented source geometries are point source, and (by summer 2010) line source in 3D.
- An interface for ground-motion attenuation. Implemented are a series of global and regional attenuation relationships for PGA and Macroseismic Intensity.
- An interface for site amplification. Currently, it is just possible to read additive and multiplicative amplifications from world datasets.
- An import/export functionality for ground motion: Instead of calculating the ground-motion field from source and attenuation, measured or 3rd-party modelled ground-motion maps may be imported in USGS ShakeMap (XML) format.

- An extensible list of building stock classes with their fragility characteristics.
- An interface for loss calculators. Currently, two types of semi-empirical loss calculators (based on fragility matrices) are implemented.

If operating based on a point source description and the implemented loss models, the input needed for a loss calculation are the earthquake origin hour, the coordinates of the epicentre, the depth and the magnitude. The program then calculates the ground shaking as a function of the distance from the hypocentre. In the database of QLARM, the population of about 2 million settlements and subdivisions is available and each settlement has a profile of building fragility. The degree of damage due to the calculated shaking is determined for each of the defined building stock classes, and from that the resulting numbers of fatalities and injured persons are estimated.

The output of QLARM consists of an estimated range of numbers of fatalities and one of injured persons, and a map showing the expected average degree of damage in the affected settlements. In addition, a list of settlements can be supplied with the expected numbers of fatalities and injured, as well as the percentage of buildings expected to fall into each of the 5 classes of damage.

5.2 Methodology

The methodology of the implemented examples is taken from the loss models of "Trendafiloski" (loss assessment approach, with collapse rates from building class). These loss models are based on vulnerability curves, fragility functions and casualty matrices pertinent to EMS-98 vulnerability classes as a function of the seismic intensity. Using macroseismic intensity and the damage and loss data from past events, this vulnerability module has been calibrated for different countries and regions worldwide. QLARM/"Trendafiloski" estimates are most reliable in areas where earthquakes occur frequently, since the calibration is based on past earthquakes for which losses are known. The output of QLARM is the total number of fatalities and injured persons by settlement, percentage of buildings in six damage grades and a map showing mean damage for each settlement. Uncertainties are handled by providing computed ranges, with maximum and minimum boundaries, for numbers of fatalities and people injured. To estimate damage and losses, the user has to input a deterministic event. Then, the program calculates the ground shaking, accounting for site amplification, using global and local attenuation relationships and information on the soil conditions. To calculate the damage and social losses, the following information is needed: distribution of buildings into vulnerability classes based on their response to ground-shaking, the distribution of people into the building vulnerability classes, the casualty matrix, the population for each census tract, and the population as a function of the time of the day and as a function of the season. In general, it is rarely the case that all the information reported above is available and complete, however the software can also run with partial information.

5.3 IT Details

The software characteristics of QLARM mainly consist of Java which is an object-orientated programming language with great flexibility and good support for modern web, communication, and xml technologies. Moreover, Java can be easily run on different operating systems. There are different open source implementations of Java compilers, and the developers of QLARM have used Sun Microsystems implementation Java 1.5, available under the terms of the GNU General Public License.

For what concerns the data repository, the relational database PostgresSQL is used. It provides functions for storing spatial representations of data records, elaborates triggering systems and allows a flexible handling for different interfaces for future integration in another system. Moreover, PostgresSQL is an open source software. Then, for mapping the data stored in the relational database to java objects, Hibernate is used. It is an object-relational mapping (ORM) solution for the Java language and it provides a framework for mapping an object-oriented domain model to a traditional relational database. Also, Hibernate is available as open source software that is distributed under the GNU Lesser General Public

License. QLARM provides a mapping interface using a PostgreSQL/PostGIS backend, a UMN MapServer middleware and a web frontend for displaying hazard and loss results. As a further step, these data may also be provided to GIS clients using WMS services. The software components of QLARM are represented in the schema of Figure 5.1.

The software offers a command line interface to calculate losses based on QuakeML-formatted single event or earthquake catalog information. For interactive usage, as well as for viewing calculation results, it has a web interface. For the implementation of the Web Interface, Apache Cocoon is used. The aim of the web interface is to provide the functions for handling user interactions and also to facilitate exchange of data to other services. Apache Cocoon is available as open source software under the Apache license.

For running the Java-based components of the QLARM application a servlet container is needed. Apache Tomcat is a web container that provides an environment for the Java code to run in cooperation with a web server. It runs on various operating systems including Linux and Windows; for QLARM, Version 6 is used.

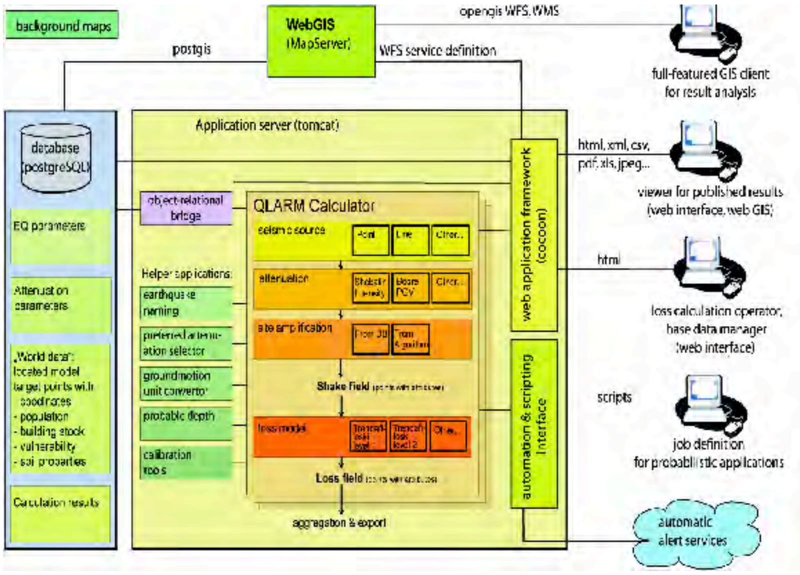


Figure 5.1 Overview of the software components of QLARM (http://qlarm.ethz.ch/SW_architecture.html)

The software does not have a manual, but some documentation is integrated in its web frontend, and the user can download some brief reports which explain the different steps used in the analyses. It uses the standard logging of the servlet container, implementing 2 loglevels. API documentation is in javadoc. It has a web user interface which is very helpful for going through the steps of the software (Figure 5.2), however the user interface does not give adequate access to all functionality. For instance, large exposure datasets are currently created and maintained easiest using direct SQL access to the database. For summer 2010, an XML import/export functionality for exposure data is planned.

For what concerns the input data, the user can choose hazard, exposure and vulnerability models through a number of scroll down menus or he/she can add and modify some features by hand using special web pages.

With regards to the output, the results are displayed on the browser or the user can choose to download them as *.xls files. Moreover a map of the results can be visualized on the web (Figure 5.9).

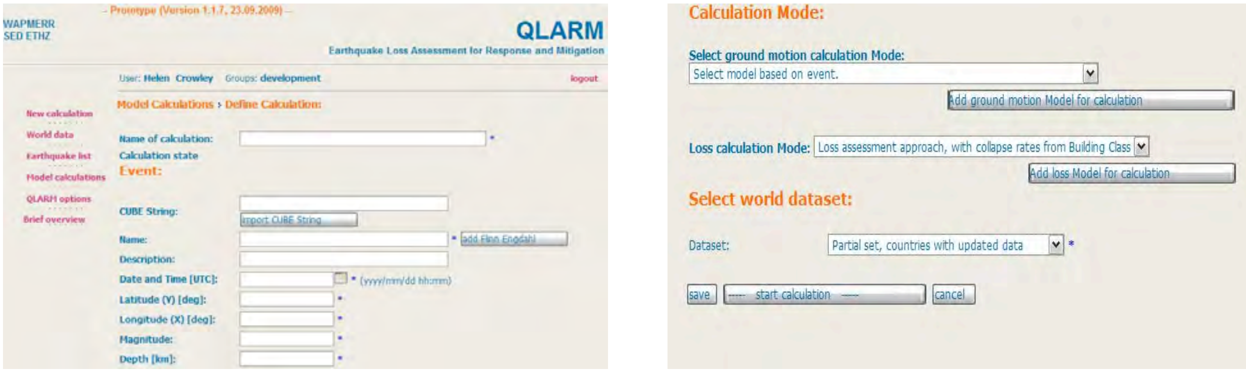
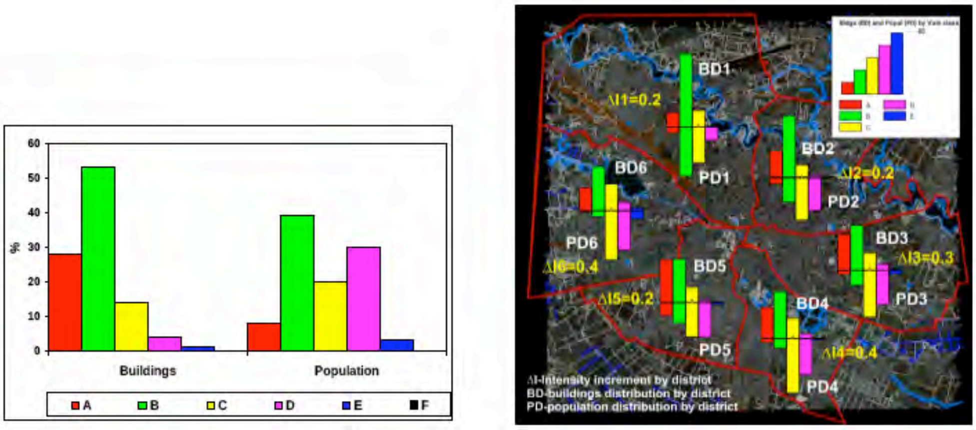


Figure 5.2 QLARM Screenshots which show how to run a new calculation

5.4 Exposure Module

At present, the user cannot easily upload their own inventory data into QLARM using the web interface, but he/she has to write them directly to the database. Thus, only users with administrative rights on a QLARM installation can introduce new world data sets. However, interactive modification of pre-existing exposure data is supported by the web interface. Built-in example datasets regarding buildings, populations, and settlements are distributed with the program. QLARM has two different approaches to define the exposure resolution depending on the available data. The first one is the so-called 'point city model' which is used for the cases where only summary data for a named settlement is available. The second one is the so-called 'discrete city model' which is used for the cases where data regarding settlement sub-divisions (districts or similar) are available. In Figure 5.3 the difference between the two approaches is shown. In Figure 5.3a the point model for the city of Bucharest is represented by a single distribution of buildings into classes A through F and a single distribution of people into these buildings. In Figure 5.3b the discrete city model is represented. In this case, the distributions into classes of vulnerability are different for each administrative district, with histograms above the horizontal black line indicating distribution of buildings and histograms below the horizontal black line indicating distribution of people in these buildings.



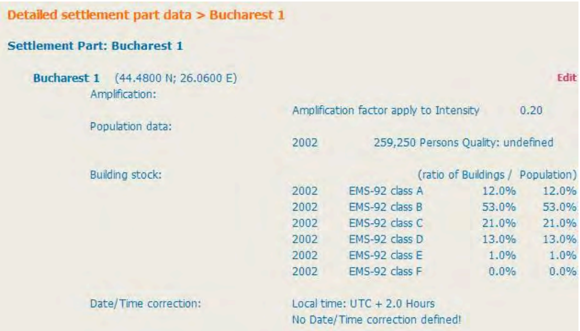
(a) (b)

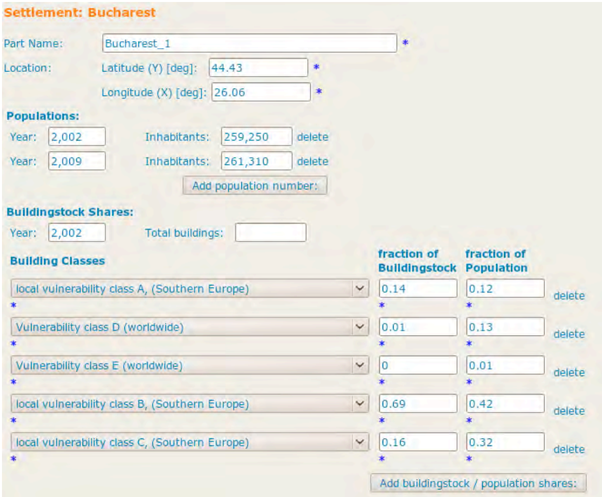
Figure 5.3 City models for Bucharest: (a) point city model (b) discrete city model (http://qlarm.ethz.ch/)

In the inventory database, the QLARM developer introduces the soil amplification factors, the distributions of building stock and population into vulnerability classes, and the most recent population numbers by settlement or district.

Regarding the different building types of the example datasets, QLARM considers the relative vulnerability classes on the European Macroseismic Scale EMS-98 [Grünthal, 1998]. The different occupancy levels (residential, commercial, industrial, etc.) are not taken into account separately from the vulnerability classes, however they are implicitly represented in loss calculation by the way of settlement-part specific daily and weekly fluctuations of the population. The software calculates the building and population distributions using percentages of the number of buildings and population belonging to a particular vulnerability class. It is worth noting that the distribution of buildings and population into vulnerability classes are in principle individual to each city. However, in the QLARM example database, they are modelled depending on the city size. These distributions are related to the population numbers that are different according to the city size, in country- and region-specific size classes.

The user can build his/her own world dataset or choose from example data in QLARM. As stated before, this dataset includes the amplification factor, the population data and building stock data (see Figure 5.4a). Moreover, the user can modify and add some features to the existing dataset by clicking on the 'edit' button. This can be done by hand using the web pages shown in Figure 5.4b. It is noted that from the web frontends available by now, the building stock composition of a settlement may be changed.





(a) (b)

Figure 5.4 (a) Web page showing the detailed settlement data and (b) web pages to modify an existing world dataset

5.5 Hazard Module

The user can just provide the epicentre, depth, and magnitude of a single earthquake and then select a ground-motion prediction equation (GMPE), or he/she can directly import ground-motion data in USGS ShakeMap format. The user can interactively set GMPE parameters, however, he/she cannot introduce new GMPEs from the web frontend (analogous to OpenSHA, new GMPEs need to be implemented as a Java class). QLARM calculates the ground-shaking for each settlement present in its database as a function of distance from the source. This is done using global and local GMPEs. The ground motion at a given distance from the source is influenced by the source parameters (epicentre, depth and magnitude), the wave propagation modelling and the local site geology. The QLARM database contains a number of well-known GMPEs that the user may choose from. Furthermore, QLARM recommends regional GMPEs that have been reviewed by its team and a standard world-average relationship.

The user can choose to express the seismic demand in terms of macroseismic intensity or instrumental parameters (for instance, PGA or PGV). However, it is worth noting that QLARM is still under development and continues to be updated; future revisions will include spectral acceleration as a ground-motion parameter.

For the time being, the user can choose between three intensity-based relationships and four global relations concerning PGA/PGV. These relationships are:

- Ambraseys [1985] used for western Europe and Maghreb (intensity-based);

- ECOS [2006] used for Switzerland (intensity-based);
- Huo and Hu [1992] with local amplification;
- Boore *et al*. [1997] for shallow crustal seismicity;
- Atkinson and Boore [2006] for stable continental shields;
- Youngs *et al*. [1997] for subduction zones.

To estimate the amplification factor used to modify the ground motion, QLARM needs information about the soil conditions. It uses two approaches to compute soil amplification:

- Local approach: this is based on data regarding soil properties, microzonation, and geological maps from which it derives amplification factors for each settlement;
- Global approach: this is based on V_{s30} values derived from topographic slopes [Allen and Wald, 2007], which is implemented when no detailed information on soils is available. In these cases an average V_{s30} value is calculated from the values on the grid of data (Global V_{s30} Map Server of the USGS) at a certain radius of each settlement and converted into an amplification factor.

Only ground-shaking hazard is considered by QLARM and other earthquake-related hazards (Tsunami, fire, etc) are not taken into account.

The user can make all choices related to the hazard through a scroll down menu, though fragility matrices are editable only in the database, and formulas necessary to compute the calculation are coded in java modules.

5.6 Vulnerability Module

The QLARM example dataset base EMS-98 vulnerability classes A through F and the vulnerability classification system is shown in Figure 5.5. In spite of the detailed distinction of each type of building, it is recognized that the seismic behaviour of buildings, in terms of apparent damage, may be subdivided by six vulnerability classes. Notwithstanding this, it is worth noting that it is possible that even if each type of structure is characterized by a prevailing vulnerability class, it is possible to find buildings with a better or worse seismic behaviour, depending on their construction characteristics.

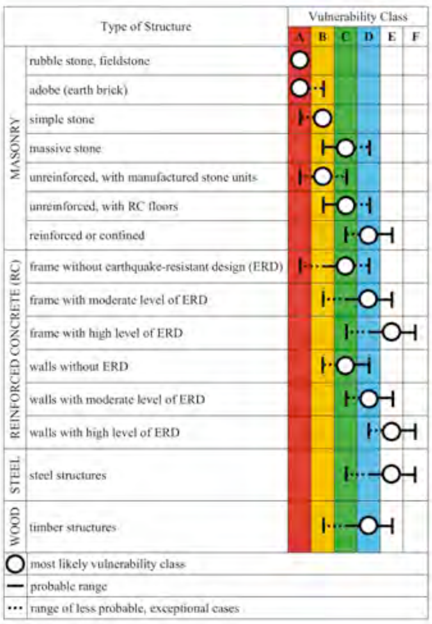


Figure 5.5 EMS-98 vulnerability classification system [Tyagunov et al., 2006b]

In the “Trendafiloski” loss models currenly implemented, the vulnerability method is intensity-based and the fragility curves are used to estimate the damage through the seismic intensity level. The program thus goes directly from the intensity to the damage without calculating any structural response such as, for instance, the performance point. The first step to estimate the vulnerability is to calculate an analytical function which correlates the mean damage grade with the seismic intensity and the vulnerability index [Giovinazzi and Lagomarsino, 2005].

In Figure 5.6 the “Trendafiloski” vulnerability models for each vulnerability class are shown. These models correlate the mean damage grade μ_D with the seismic intensity I and the vulnerability index (V_I) as shown in Equation (5.1). The membership V_I is not deterministic and defines the most probable class and its plausible and ultimate bounds.

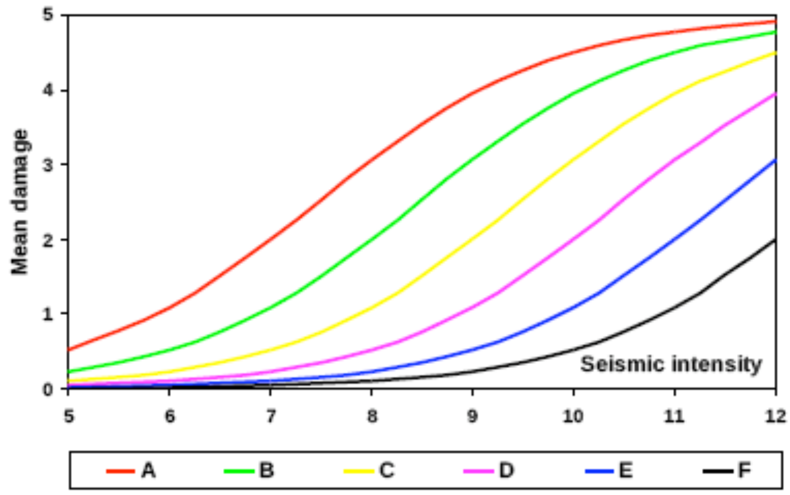


Figure 5.6 QLARM vulnerability models (http://qlarm.ethz.ch/)

$$\mu_D = 2.5\left[1+\tanh\left(\frac{I+6.25V_I-13.1}{2.3}\right)\right] \tag{5.1}$$

Once the mean damage grade is calculated, fragility functions are then defined using a cumulative beta distribution. Thus, the discrete beta density probability function is calculated and “Trendafiloski” creates a damage probability matrix to get the damage distribution. At the same time, the mean damage grade for the city model is calculated by weighting the mean damage grades for all the vulnerability classes with the distribution of the buildings.

The user may upload their own fragility curve and building class using the web interface shown in Figure 5.7. The data has to be uploaded by hand and the user can choose to modify an existing building class or add their own.

The “Trendafiloski” loss module estimates the human losses using the casualty event-tree model proposed by Stojanovsky and Dong [1994]. In the following equations the probability of occurrence of casualty state C_k (k=1,5) for seismic load I_J are shown. They are calculated as a product of the damage probabilities for seismic load I_J and the casualty probabilities for damage grade D_i.

$$P(C_K I_J) = \sum_{i=1}^{3} P(D_i I_J)P(D_i C_K) + P(D_{NC} I_J)P(D_{NC} C_K) + P(D_C I_J)P(D_C C_K)\Big| J = 1,n \tag{5.2}$$

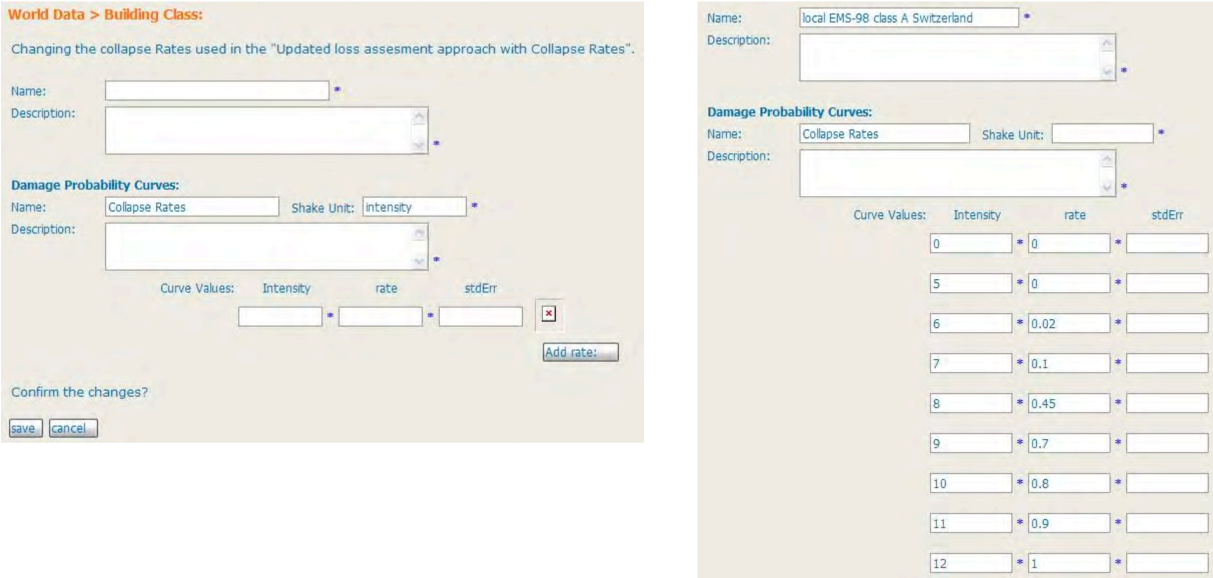
$$P(D_C I_J) = k_C(I_J)\left[P(D_4 I_J) + P(D_5 IJ)\right] \tag{5.3}$$

$$P(D_{NC} I_J) = (1 - k_C(I_J))\left[P(D_4 I_J) + P(D_5 IJ)\right] \tag{5.4}$$

Where:

- $P(D_i I_J)$ is the probability of occurrence of damage grades i=1 to 3 for the seismic intensity I_J;
- $P(D_{NC} I_J)$ is the probability of having no collapse among the buildings with damage grades 4 and 5;

- $P(D_C I_J)$ is the probability of having collapse among the buildings with damage grades 4 and 5;
- $k_C(I_J)$ is the collapse model which determines the percent of collapsed buildings as a function of seismic intensity I_J. For the time being, “Trendafiloski” uses a discrete collapse model for vulnerability classes A to E for seven regions worldwide using the collapse rates for 26 countries worldwide provided by the World Housing Encyclopedia (www.worldhousing.net);
- $P(D_i C_k)$ is the probability of having casualty state C_k due to damage grade D_i.



(a) (b)

Figure 5.7 (a) Web page to add a new building class and (b) web page to modify an existing building class

With regards to the casualty rate, C_k, “Trendafiloski” uses the HAZUS 2003 indoor casualty rates as a default, which the developers are modifying based on observed fatality and injured rates [Wyss and Trendafiloski, 2009]. The social losses are independent of economic losses, that are not considered in the code, and they are correlated to the damage suffered by the structure. It is worth noting that QLARM may handle alternative loss models, however they would need to be implemented in Java.

5.7 Output

The number of damage stages given in the result depends on the loss calculator. With “Trendafiloski”, six structural damage states are considered in QLARM, as given by the EMS-98 scale: no damage (D_0), slight damage (D_1), moderate damage (D_2), heavy damage (D_3) very heavy damage (D_4) and destruction (D_5). The number of damage grades is hard-coded in the software. The final damage results are given in terms of the percent of buildings in six damage grades and the mean damage grade for each settlement. QLARM considers 5 limit states of social losses, however, for the web frontends, they are summarized to two: fatalities and patients. Social loss results include a low and high value that encompass the estimated range in number of fatalities and patients, for each settlement.

The output is displayed in the browser and it shows a summary of the data used to run one test-bed application (Figure 5.8a) and a table with the final damage and loss results (Figure 5.8b). Furthermore, the user can choose to download the results as an *.xls file. Moreover, a map showing mean damage by settlement can be visualized on the web (Figure 5.9).

Calculation > Detailed View

Progress:	**done**, calculation finished successfully
Calculation: Lima Test Scenario	
World data set:	Partial set, countries with updated data
Groundmotion model:	PGA: Youngs et Al 1997
Loss model:	Level 1 loss assessment, with intensity based colapse rates form Builing Class
Event:	
Date/Time	2009/09/22 10:56:00 UTC 2009/09/22 05:56:00 GMT-05:00
Region	Lima Test Scenario A
Location (LAT/Y;LON/X)	(-12.18 ; -77.20)
Depth, km	33.0
Magnitude:	8.0
Note:	**DISCLAIMER: The implementation of the used ground motion model was not yet tested, and may calculate INCORRECT results!** **DISCLAIMER: For converting Intensity to PGA a not yet tested function was used, so the calculation may produces INCORRECT results!**

Result

fatalties (high)	6482	Pers.
patients (high)	78772	Pers.
fatalties (low)	2980	Pers.
patients (low)	39389	Pers.

Show_Map xls

Detail Results:

	population	distance	pga	intensity	buildings no damage	buildings slight damage	buildings moderate damage	buildings heavy damage	buildings very heavy damage non collapse	buildings very heavy damage collapse	fatalities (low)	fatalities (high)	patients (low)	patients (high) (•)	mean damage grade
	1000 Pers.	km	m/s/s	EMS-98	%	%	%	%	%	%	Pers.	Pers.	Pers.	Pers.	-
Callao	813.3	16	3.95	8	24.0	26.8	24.5	16.6	4.0	4.1	482	980	5,598	10,750	1.6
L32 San Juan de Lurigancho	919.4	30	3.21	8	32.0	29.2	22.0	12.3	2.8	1.8	198	457	3,036	6,234	1.3
L35 San Martín de Porres	594.1	23	3.56	8	27.6	28.0	24.1	14.5	3.3	2.5	201	440	2,685	5,417	1.5
L15 Lima downtown	315.4	22	3.63	8	22.2	24.6	24.8	18.2	4.9	4.4	204	422	2,342	4,542	1.7
L9 Comas	525.8	32	3.13	7	31.9	29.3	22.4	12.1	2.7	1.6	102	240	1,581	3,285	1.3
L17 Los Olivos	324.2	26	3.40	8	25.8	27.0	24.2	15.4	3.8	2.8	126	276	1,611	3,236	1.5
L7 Chorrillos	297.1	21	3.69	8	26.7	28.1	23.9	14.8	3.5	3.0	114	244	1,523	3,016	1.5
L33 San Juan de Miraflores	379.3	25	3.47	8	29.8	29.2	22.9	13.0	2.9	2.1	91	206	1,453	2,950	1.4
L2 Ate Vitarte	474.8	36	2.96	7	33.7	29.2	21.8	12.0	2.8	1.5	84	200	1,364	2,839	1.3

(a) (b)

Figure 5.8 (a) Web page which shows the summary of the data used to run the test-bed application and (b) web page which shows the table with the final results

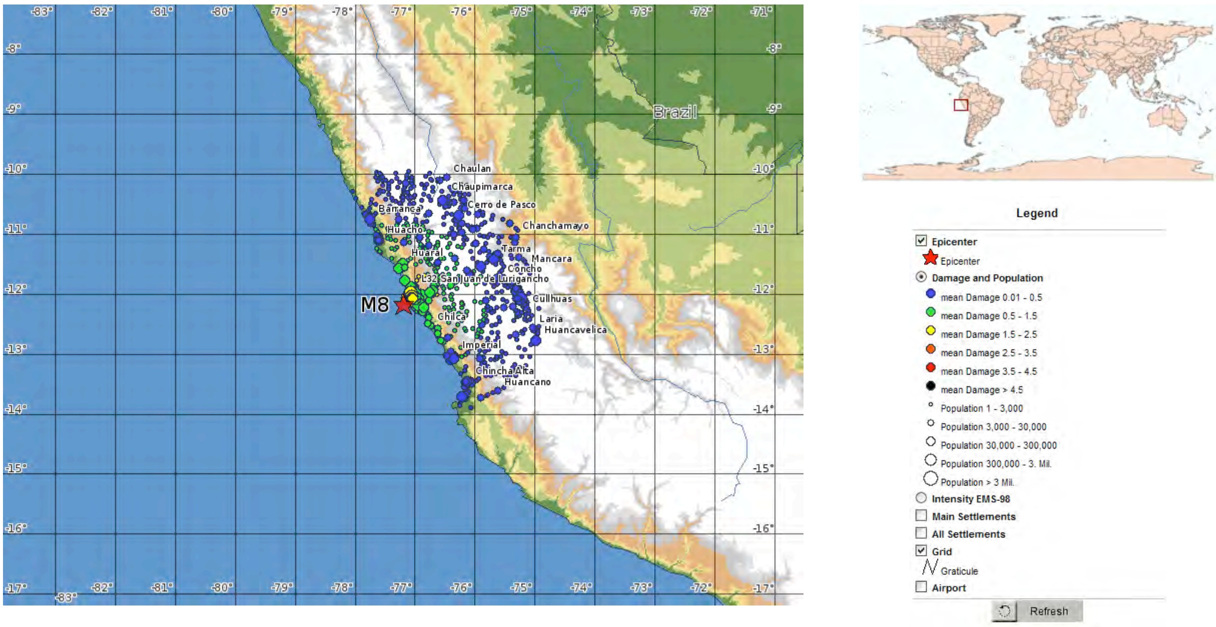


Figure 5.9 QLARM map of mean damage by settlement

QLARM provides reasonable results for cities with datasets stored in the program. Computed results tend to follow expected trends where, for a particular settlement, a larger magnitude event or buildings of higher vulnerability will result in greater damage, which will result in greater losses. Furthermore, for certain tested cities, the computed loss results have been found to be similar to the actual reported values from the past earthquake considered.

6 CEDIM

6.1 Summary of Software

CEDIM (The Center for Disaster Management and Risk Reduction Technology) is an interdisciplinary research centre in the field of disaster management, founded by Helmholtz Centre Potsdam - German Research Centre for Geoscience (GFZ) and Karlsruhe Institute of Technology (KIT). The main goal of CEDIM is to advance the scientific understanding of natural and man-made hazards assessment and to develop disaster management solutions for the early detection and reduction of risk.

The CEDIM software (CEDIM Risk Estimation Tool – CREST) was developed in the frame of the project Risk Map Germany (2003-2005) [Tyagunov *et al*., 2006a] and afterwards (with some modifications) it was also used for damage and risk analysis for other earthquake prone areas of the world.

The CEDIM software that has been provided for the GEM1 team and is reviewed herein combines exposure, vulnerability, and hazard together to evaluate risk from earthquakes. It deals with vulnerability classification using the European Macroseismic scale 1998 [EMS-98].

The user has to supply input data concerning the hazard, the vulnerability and the exposure. The input data are the hazard, the geographical location of each geographical unit or grid cell, the total number of buildings, the percentage of buildings of each vulnerability class and the population of each geographical unit or grid cell. The repair and replacement cost of buildings and the contents are necessary to compute the economic losses. The main outputs are the structural damage, social losses and economic losses.

6.2 Methodology

The provided version of CEDIM software uses an intensity-based methodology, as defined by the European Macroseismic Scale 1998, EMS-98, [Grünthal, 1998] for seismic risk and loss assessment. The software considers macroseismic intensity, along with vulnerability and exposed assets, as the main input parameters which all damage and loss computations are based upon. The macroseismic intensity scale, as defined by EMS-98, has twelve levels described by the effects of different degrees of observed earthquake shaking. The location and epicentral intensity of a scenario earthquake are the input required for an analysis using CEDIM.

The current version of the software only considers vulnerability of residential buildings. EMS-98 divides buildings into six vulnerability classes based on structural type, which can then be modified based on other factors that affect vulnerability, such as date of construction, number of floors, etc. Risk is computed based on the percentage of total number of buildings in each EMS-98 vulnerability class, within a geographical unit. The vulnerability distribution is then used to compute the damage to buildings, distributed between six damage grades that range from no damage to complete destruction. A damage probability matrix is provided in the code, based on damage grade and intensity, for each vulnerability class. The damage probability matrices are used to compute the percent of buildings in each damage grade, within a geographical unit. Direct economic losses, due to structural damage, and social losses, due to injury and death, can be computed based on damage and building and population inventory for each geographical unit.

The provided version of CEDIM software can only use the intensity-based methodology. The user cannot change the methodology, as this would require a much more complex code and inputs. The intensity-based methodology is hard-coded in the software. The number and definition of vulnerability classes and damage grades are hard-coded in the software, to describe the expected response of buildings based on the EMS-98 system.

6.3 IT Details

CEDIM code requires the use of a GIS software. The programming language of the code used to calculate shaking intensity, damage, and loss within the GIS environment is Visual Basic Application (VBA script). The software does not offer a command line interface and it does not have a personalised GUI, though the GIS software GUI can be used to carry out the calculations using the Visual Basic calculators. The algorithms used for the hazard, damage and loss calculations are transparent and fairly straightforward, though they are not particularly well commented; nevertheless, the user can go through the script and change it to apply their own modifications and adjustments. Run times are fast, but individual runs are not tracked and no log files are produced. There is no documented API and no manual of the CEDIM code.

CEDIM's input files can be generated as tables using a spreadsheet software, such as MS-Excel or MS-Access. Then, they can be easily imported into the GIS software as *.dbf files and joined to the *.shp file of the area being assessed. Output can be easily exported back into a spreadsheet software. The hazard input and the results of the analyses in terms of damage and loss are automatically plotted in the GIS software; Figure 6.2 presents the map outcomes in CEDIM. Figure 6.1 shows screenshots of the GUI and of how the calculators are used for damage and loss calculations.

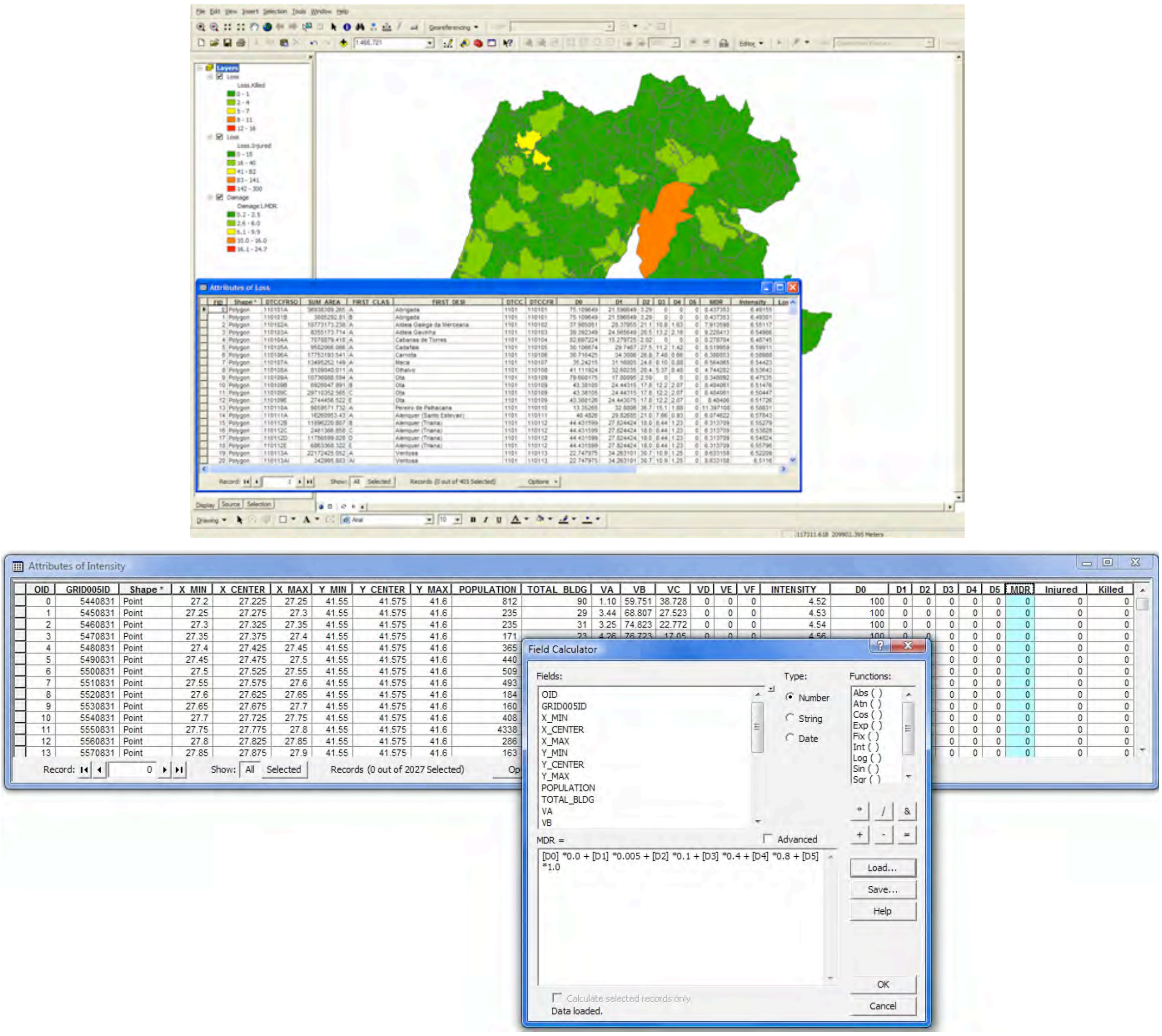


Figure 6.1 CEDIM screenshots of the GUI with map view (above) and the field calculator within the attributes table (below)

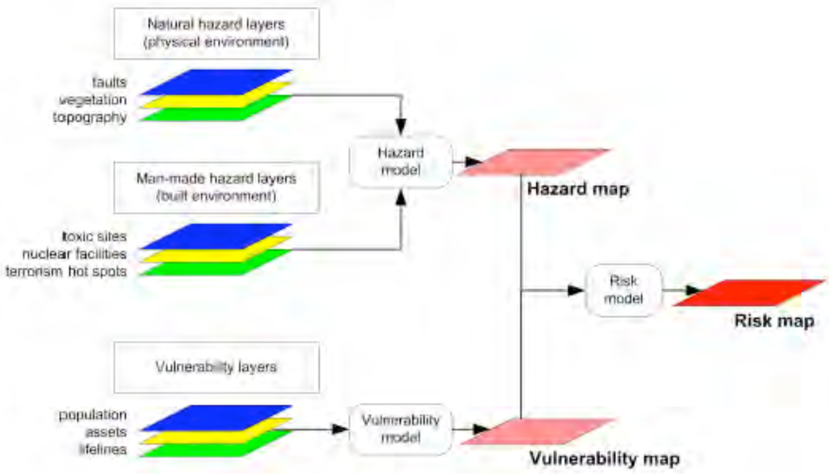


Figure 6.2 Main map outcomes in CEDIM [Müller et al., 2006]

6.4 Exposure Module

The user is required to provide building and population inventory for each geographical unit. A geographical unit consists of one census tract and is represented in the program as a single grid cell, or polygon. It is possible to carry out the assessment at any level of resolution, provided the building stock and population data are available at that level of detail. All computational parameters (e.g., hazards, vulnerability, and exposed assets) and damage/loss calculations are uniformly distributed within the area of a grid cell. The code does not account for uncertainties in the inventory data, as the data input by the user is used directly for computations of risk.

European Macroseismic Scale 1998 (EMS-98) building taxonomies are used to group buildings with similar damage/loss characteristics into six vulnerability classifications, denoted A through F. The vulnerability classifications are based on structural type taking into account other factors contributing to the vulnerability of buildings, such as date of construction, number of floors, etc. The provided version of the program only considers one occupancy level, that being residential buildings. With regards to the building inventory, the damage is computed based on the percentage of total number of buildings in each EMS-98 vulnerability class, within a geographical unit. To compute economic losses it is necessary to input economic information about cost of repair or replacement of all buildings within a geographical unit (replacement costs per person). With regards to the population inventory, the user must provide the total number of people living in each geographical unit. The population and computed damage are used to compute social losses, based on a casualty model. All components of the loss calculations are at the same resolution, which is based on the resolution of the inventory.

The inventory is stored in database (*.dbf) files, corresponding to a shape (*.shp) file that consists of a map of the area being considered, divided into geographical units. The *.shp and *.dbf files can be input into any GIS software. All inventory will then be stored in a single attributes table. The level of detail considered in the program is user-defined. Census tracts can be used, but the unit could also be regional, provincial, local, or some other unit, as long as building stock and population data are available at that level of detail. The user inputs the central coordinates of each geographical unit (or they can be automatically calculated within GIS), along with the population and total number of buildings occupying the area. The properties of building classes are not hard-coded, though as discussed below they have to be assigned to EMS-98 vulnerability classes, and the user can configure the inventory outside of the program.

6.5 Hazard Module

In order to perform a seismic risk and loss assessment, the introduction of ground-motion intensity is required. The software allows for the introduction of an earthquake scenario, by means of attenuation relationships or ground-motion fields, for one or multiple hazard sources. A deterministic analysis will predict the level and impact of ground shaking from a particular scenario event, but will not take into account the likelihood of the scenario occurring. A probabilistic analysis will allow for the consideration of multiple hazard sources and the prediction of the probability of exceeding a given intensity at a certain distance in a certain time frame.

The provided version of the CEDIM software only considers the hazard of ground shaking and, while other earthquake-induced hazards (e.g., liquefaction, landslide, or tsunami) may be present, the version of the code does not account directly for these additional hazards. The code considers macroseismic intensity as the ground-motion parameter which all damage and loss computations are based upon. The macroseismic intensity scale, as defined by EMS-98, has twelve levels described by the effects of different degrees of observed earthquake shaking.

In the case of a deterministic analysis, there are two options for introducing hazard. The first option is the use of an attenuation relationship, which is written in the code and used to calculate the median value of intensity of ground shaking at the location of the centroid of each geographical unit. For this type of analysis, the user must input the epicentral coordinates, focal depth, and earthquake magnitude (or epicentral intensity). These parameters, which are then used to calculate the attenuation of intensity, can represent a potential hypothetical future earthquake or a past historic earthquake.

It is important to note that some attenuation relationships are applicable only to certain regions or soil conditions. The code must be altered by the user in order to change the attenuation relationship to one that is applicable to the area that is being considered. The macroseismic intensity scale inherently accounts for site effects, as intensity is a measure of observed effects of an earthquake. If ground shaking is increased due to soil amplification or topographical conditions, this will be seen as a localized increase in observed intensity. However, when intensity is computed based on observed intensity at a different location, it may be necessary to use an attenuation relationship that accounts for the differing site conditions between the two locations. An attenuation relationship that accounts for local site effects can be written into the code.

A second option for introducing hazard in a deterministic analysis is for the user to use a ground-motion field and input the median value of intensity of shaking for the centroid of each geographical unit. In this case, the user could import the intensity data from the ground-motion field into the attributes table in the GIS software, in the form of an external database file. It is likely that the resolution of the ground-motion field and the building stock/population inventory will not be the same. An external program must be used to provide the ground-motion field data at the resolution of the inventory, before the intensity data can be added to the attributes table. This can be accomplished by writing a code that provides a weighted average of observed intensities at the nearest points to the location of the centroid of each geographical location.

In the case of a probabilistic analysis, seismic hazard maps need to be imported by the user; the code does not currently carry out probabilistic seismic hazard assessment.

It is possible to use CEDIM software to manually perform a near-real-time assessment of risk and loss, shortly after an earthquake has occurred. This can be done as soon as the epicentral coordinates, focal depth, and epicentral intensity can be accurately estimated. However, this type of analysis would require that information regarding building stock and population in the considered area be readily available.

6.6 Vulnerability Module

CEDIM code assumes that all computational parameters are uniformly distributed within a single geographical unit. This means that communities (or their parts) are considered as units at risk and the vulnerability composition of all buildings within a community is being considered, rather than the seismic performance of a single building [Tyagunov *et al*., 2006a].

The provided version of the CEDIM code uses a relatively simple combination of damage probability matrices and empirical equations. The vulnerability analysis is performed in terms of six vulnerability classes, denoted A through F, as defined by EMS-98. The vulnerability class designation describes the ability of a building to perform seismically, where class A signifies a building that is most susceptible to damage and class F signifies a building that is least susceptible. EMS-98 classifies vulnerability by providing typical ranges of, as well as the most probable, vulnerability class for various structure types. For a given structure type, the most probable classification can be adjusted based on various factors, such as date of construction and number of floors, etc. Figure 5.5 shows the EMS-98 vulnerability classification scheme. CEDIM considers six damage grades, as defined by EMS-98: (0) no damage, (1) negligible to slight damage, (2) moderate damage, (3) substantial to heavy damage, (4) very heavy damage, and (5) destruction. Each damage grade represents a combination of structural and non-structural damage. Table 6.1 presents qualitative and quantitative descriptions for each damage grade. A damage probability matrix is created in the code, based on damage state and intensity, for each vulnerability class. Therefore, each matrix describes the probable distribution of damage to buildings of equal vulnerability at a certain level of seismic intensity [Tyagunov *et al*., 2006a]. The damage probability matrices are then used to compute the percent of buildings in each damage grade, within a geographical unit. The mean damage ratio (MDR), defined as total cost of repair divided by total cost of reconstruction, is calculated for each geographical unit from the computed distribution of damage grades, as

$$MDR = [D_0] \times 0.0 + [D_1] \times 0.005 + [D_2] \times 0.1 + [D_3] \times 0.4 + [D_4] \times 0.8 + [D_5] \times 1.0 \tag{6.1}$$

where, D_0, D_1, D_2, D_3, D_4 and D_5 are the percentage of buildings in damage grades 0, 1, 2, 3, 4, and 5, respectively.

The final damage results are reported both as distribution of different damage grades and as mean damage ratio, in percent. According to EMS-98, an idealized damage occurrence probability distribution would be represented by damage grades normally distributed about the mean damage value. Of course, as the damage grades are limited to be between 0 and 5, the shape of the distribution will shift as one of these bounds is approached [Grünthal, 1998]. A sample damage probability matrix is provided in Table 6.2.

The damage is not disaggregated into damage for each building type; rather, it is reported for the total of all buildings within a geographical unit. The damage results are output by the code as a column within the attributes table. This allows the user to process the results and view the results geospatially on a map of the considered area. The results could also be exported into a spreadsheet program (e.g., MS Excel) for processing. The number of damage grades is hard-coded and, in order to change the damage grade divisions, the user must alter the code.

Table 6.1 EMS-98 damage grades, for CEDIM [Tyagunov *et al*., 2006a]

Damage Grade	Damage Ratio, %
Grade 0: No damage	0
Grade 1: Negligible to slight damage	0-1
Grade 2: Moderate damage	1-20
Grade 3: Substantial to heavy damage	20-60
Grade 4: Very heavy damage	60-100
Grade 5: Destruction	100

Table 6.2 Sample damage probability matrix for vulnerability class B, for CEDIM [Tyagunov *et al*., 2006b]

Damage Grade	Damage Ratio, %	Central Damage Factor, %	Seismic Intensity				
			V	VI	VII	VIII	IX
0	0	0	92.5	50	7.5	***	***
1	0-1	0.5	7.5	42.5	42.5	7.5	***
2	1-20	10	***	7.5	42.5	42.5	7.5
3	20-60	40	***	***	7.5	42.5	42.5
4	60-100	80	***	***	***	7.5	42.5
5	100	100	***	***	***	***	7.5
Mean Damage Ratio, %			0.04	0.96	7.5	27	59

Vulnerability functions for vulnerability classes A through D are presented in Figure 6.3, as a function of mean damage ratio and seismic intensity, where mean damage ratio represents the ratio of repair costs to replacement costs. For instance, at an intensity of 10, the mean damage ratio for a geographical unit that consists of only buildings in vulnerability class A would be 100%. This would indicate that all buildings can be described by damage grade 5: destruction. At a given intensity, for a geographical unit consisting of a distribution of buildings in different vulnerability classes, the mean damage ratio could be determined as a combination of the vulnerability functions shown in the figure. The resulting mean damage ratio would indicate that multiple damage grades may be experienced by the different buildings.

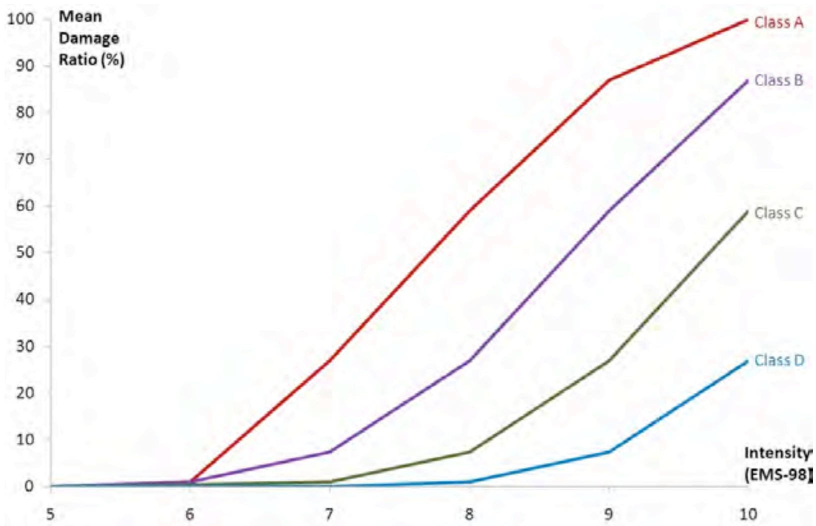


Figure 6.3 Sample vulnerability functions, for CEDIM [Tyagunov *et al*., 2006b]

CEDIM code employs empirical functions, which use weighted averages of damage grade distribution, to consider both economic and social losses. The code can calculate direct monetary losses due to structural damage to buildings and social losses due to injury and death, based on damage distribution and population and building stock information.

CEDIM computes economic losses as follows:

$$\text{Economic Losses} = [\text{Population}] \times [\text{Replacement\_Costs\_per\_Person}] \times [\text{MDR}] / 100 \quad (6.2)$$

Where:

- Population is the total number of people residing in the geographical unit;
- Replacement_Costs_per_Person can be estimated on the base of statistical data or calculated as the total cost of building replacement for the geographical unit divided by the number of people residing in the geographical unit;
- MDR is the mean damage ratio for the geographical unit.

The current version of CEDIM only considers the direct economic losses due to structural damage. Therefore, as seen in Equation (6.2), economic losses are dependent on the computed damage output and are independent of social losses.

The economic losses results are reported in a user-defined currency (e.g., Euros). The economic losses are not disaggregated into losses for each building type; rather, the results are reported for the whole geographical unit. The results are output by the code as a column within the attributes table. This allows the user to process the results and view the results geospatially on a map of the considered area. The results could also be exported into a spreadsheet program (e.g. MS Excel) for processing.

With regard to social losses, CEDIM considers both injury and death, calculated based on a casualty model as:

$$\text{Social Losses from Injury} = [\text{Population}] \times ([D_2] \times 0.01 + [D_3] \times 0.02 + [D_4] \times 0.1 + [D_5] \times 1.0)/100 \quad (6.3)$$

$$\text{Social Losses from Death} = [\text{Population}] \times ([D_3] \times 0.0025 + [D_4] \times 0.01 + [D_5] \times 0.2)/100 \quad (6.4)$$

Where:

- Population is the total number of people residing in the geographical unit;
- D_2, D_3, D_4 and D_5 are the percentage of buildings in damage grades 2, 3, 4, and 5, respectively.

The coefficients of this casualty model have been developed empirically and are comparable to similar coefficients developed by HAZUS.

The immediate social losses results are reported as number of people injured and killed. The social losses are only disaggregated into injury and death, with results reported for the whole geographical unit. The results are output by the code as columns within the attributes table. This allows the user to process the results or export the results into a spreadsheet program (e.g. MS Excel) for processing.

All damage and loss results are computed within the attributes table of the GIS software by loading calculation files, containing the necessary parameters, into the field calculator. Figure 6.4 shows a comparison of damage, injury, and life loss ratios for different damage grades.

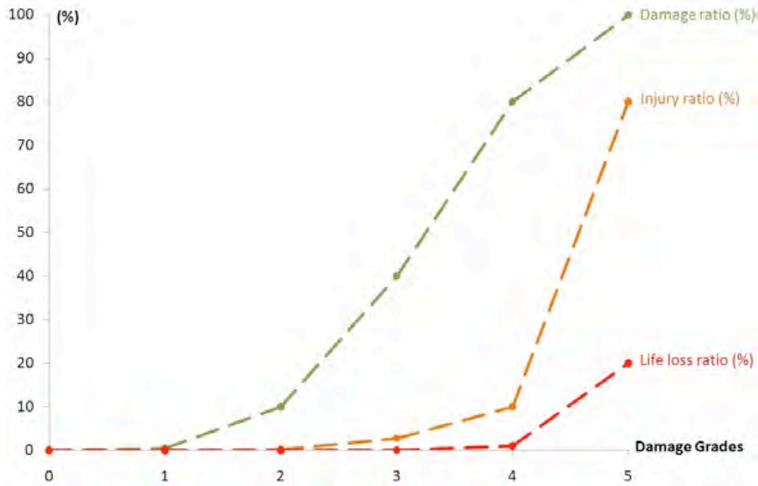


Figure 6.4 Sample damage, injury, and life loss ratios, for CEDIM [Tyagunov *et al*., 2006b]

6.7 Output

Damage, economic losses, and social losses results are output by CEDIM code. These results include percent of buildings in each of the six damage grades, mean damage ratio, direct economic losses (monetary value) due to structural damage to residential buildings, and direct social losses (number of people injured and number of people killed, in positive integer value) due to damage. The output are not disaggregated for an individual building or building type; rather, they are reported as single values that are evenly distributed within a geographical unit (grid cell). The reported values represent the mean for the geographical unit, with no other statistical information provided.

Each category of output is provided by the code as a column within a single attributes table of the GIS software. The user can process the results and view the output geospatially on a map of the considered area. Multiple maps can be created by the user, each representing a different category of output. These maps are a good visualization tool for viewing the spatial distribution of damage and loss results, because the user can easily differentiate between areas of minimal and excessive damage and/or losses, based on the colour distribution of the map. The user also has the option to export the output results into a spreadsheet program, to create a single file for processing results.

CEDIM code provides reasonable results of damage and losses. While some test-bed application results computed by CEDIM matched well with reported damage and losses for a past event, others did not; however, the results always followed an expected trend. For a particular region or geographical unit, a larger magnitude event results in greater damage. Building types of higher vulnerability also result in greater damage. Greater damage results in increased economic and social losses. The code itself and the methodology behind it are considered reasonable for the purposes of seismic risk estimation.

7 CAPRA

7.1 Summary of Software

CAPRA (Central American Probabilistic Risk Assessment) is a multi-hazards software for computing loss estimates. It has been developed with funding from The World Bank, by the ERN-LA consortium, formed by the following companies and institutions: ERN Ingenieros Consultores (México), ITEC (Colombia), INGENIAR (Colombia) and CIMNE (Spain). The main regional applications have focused on Central America, specifically Costa Rica and Nicaragua. The main program is CAPRA-GIS (v 0.9.8.0), a GIS based software and the main module in the CAPRA methodology, which performs loss estimation for a number of natural hazards including earthquakes.

CAPRA-GIS is capable of performing hazard analysis once the required input files are prepared; the open-source probabilistic seismic hazard software CRISIS can be used to create the specific hazard files, though it would also be possible to use another hazard code, as long as the output is presented in the standard format that CAPRA requires. Once the hazard files are prepared in the required format they can be easily read and imported into CAPRA-GIS.

The GUI of CAPRA-GIS is easy to understand. The exposure data is handled as shapefiles, which can be produced with any GIS software. With respect to the vulnerability of the exposed inventory, CAPRA-GIS can read simple files in ASCII format which contain the vulnerability data for given hazard levels. These vulnerability files can be obtained using another stand-alone tool called ERN-Vulnerabilidad, developed by ERN, or they can be user-defined with any methodology chosen by the user, provided they are input in the required standard format. The vulnerability functions should depend directly on the hazard output parameters (intensities), so there is no need to input additional data for the structural typologies, such as capacity curves. The main output of CAPRA-GIS is the expected annual damage ratio (the ratio of repair cost to the real monetary value of the structure) of each structure that is geo-referenced. When it is multiplied by the exposed value yields the total economic loss for that specific structure. Another result that CAPRA delivers is an aggregate loss exceedance curve for the whole building portfolio. The code has built-in visualization tools to view the exposed building stock, the hazard, the economic loss per building, and the total loss for the entire area of study.

7.2 Methodology

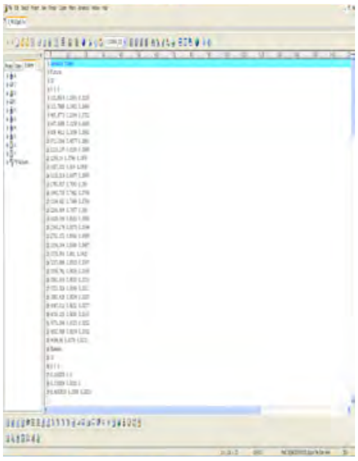
CAPRA uses vulnerability functions that provide the expected damage as the ratio of the repair cost to the total value of the geo-referenced structure for a given level of intensity, together with the mean and standard deviation which allow to fully define the probability distribution of damage given an intensity. This way the uncertainty of the relationship is taken into account. With this in mind, the total economic loss per building, for a given earthquake scenario, is computed simply by multiplying the aforementioned ratio by the total value of the building. The user needs to provide the vulnerability functions based on technical expertise and the preferred methodologies such as the case of capacity spectrum methods, equivalent SDOF methods, etc. It is worth mentioning that there is a tool available, ERN-Vulnerabilidad, which is capable of computing these vulnerability functions in the specified format for CAPRA-GIS (expected damage). This program can use different methodologies, in particular the methodology proposed by Miranda [1999]. This method estimates the maximum inter-storey drift-ratio (IDR_{max}) and uses this response combined with specific fragility functions in terms of IDR, to obtain the expected damage.

7.3 IT Details

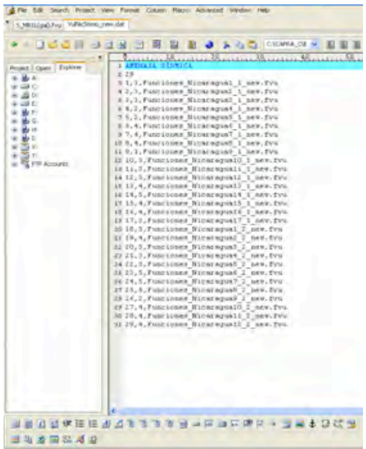
CAPRA-GIS can be installed in any PC with Windows XP or higher versions for what concerns the operating system. Regarding the programming language, it was developed using Visual Basic.Net, and the code is not open-source yet, but it was made available to the GEM1 Risk team. The program needs three main sets of files in order to run, one related to the hazard, one related to the exposure and another related to the vulnerability.

These files are summarized in the following:

- The *.ame file is used to characterize the hazard in the specific region of study;
- A *.shp file is used to characterize the whole exposed building stock in the area of study;
- The *.fvu files contains the structural vulnerability functions and the corresponding human loss functions for each building type. A *.dat file relates each fragility function with the structural typology code used to describe the exposure data in the *.shp file.



(a)



(b)

Figure 7.1 (a) Screenshot of the input file *.fvu that specifies the vulnerability function in CAPRA-GIS for a specific building typology and (b) Screenshot of the *.dat file, which specifies the code of each vulnerability function used in a seismic risk analysis in CAPRA-GIS

CAPRA does not have a documented API. The GUI of CAPRA-GIS is quite simple and easy to understand. It provides very similar tools to a standard GIS software such as pan, zoom, etc. and it allows to visualize the *.shp and *.ame files. However, once any of these functions are used, the visualization tool takes a few seconds to update itself, especially when the exposure data is extensive. It does not have a command line interface and all runs are done using the GUI.

There is no formal manual available in the public domain yet; only an on-line help dedicated to answering FAQs. An interesting contribution is a short tutorial movie, which includes instructions for future users on how to operate the program with the GUI.

The run time for a single event scenario with a building exposure of over 160.000 structures is under 60 seconds on a Intel Core Duo processor with 1.58 GHz, 3.50 GB of RAM, and once a run is made, there are log files (Figure 7.2) that track every single run, so that if there is a re-run with some minor modifications in the input data, the previous results will not be deleted.

```
*******************PARAMETROS DE CALCULO*******************

* TIPO DE ANALISIS:
POR UNICO ESCENARIO

* EXPOSICION PARA CALCULO:
----------------------------------------------------------
Archivo                  | Fecha              | Elementos
----------------------------------------------------------
EDIFICIOS_M..DENE2009.SHP | 24/03/09 05:16:16  | 160452
----------------------------------------------------------

* AMENAZA ASOCIADA :
----------------------------------------------
Archivo                  | Fecha
----------------------------------------------
NIC_EARTHQUAKE.AME       | 02/07/09 02:59:16
vulnicsismo_new.dat      | 02/07/09 12:24:48
----------------------------------------------

* ESCENARIO ASOCIADO :
13590. CAc4_SF32_M=6.93

*******************PROCESO DE CALCULO*******************

* FECHA DE CALCULO:
OCTUBRE 14 2009 03:56:28 P.M.

* CALCULO DE DAÑO TERMINADO :
Tiempo de procesamiento: 00:00:59

*******************RESULTADO DE CALCULO*******************

No. de amenazas:          1
No. de puntos:            50

Escenarios de todas las temporalidades. Las pérdidas incluyen todos los archivos de exposición
Temporalidad          Escenario                       Frecuencia    EP          VarP        a           b           EXP
   01    13590  CAc4_SF32_M=6.93                      1.17470E-04 6.38101E+09 4.62260E+19 7.60436E-01 1.11194E+01 9.96863E+10

Curva de pérdidas globales (todos los archivos de exposición
Pérdida      Total     Temporalidad 1
9.97E+04     1.17E-04   1.17E-04
1.32E+05     1.17E-04   1.17E-04
1.74E+05     1.17E-04   1.17E-04
2.29E+05     1.17E-04   1.17E-04
3.02E+05     1.17E-04   1.17E-04
3.99E+05     1.17E-04   1.17E-04
5.27E+05     1.17E-04   1.17E-04
6.95E+05     1.17E-04   1.17E-04
9.17E+05     1.17E-04   1.17E-04
1.21E+06     1.17E-04   1.17E-04
1.60E+06     1.17E-04   1.17E-04
2.11E+06     1.17E-04   1.17E-04
2.78E+06     1.17E-04   1.17E-04
3.67E+06     1.17E-04   1.17E-04
4.84E+06     1.17E-04   1.17E-04
6.39E+06     1.17E-04   1.17E-04
```

Figure 7.2 Screenshot of a log file that is related to a specific risk computation in CAPRA-GIS

7.4 Exposure Module

The inventory needs to be provided in one or more standard shape files (*.shp). This information can be formatted using standard GIS tools (such as Arc-GIS), and minor changes to these files can be done within CAPRA-GIS. This information needs to be characterized in spatial components such as polygons, polylines or points. The minimum exposure information required within the shape file is depicted in Table 7.1.

Table 7.1 Description of each input parameter for CAPRA-GIS code

Field	Description
VALFIS	Monetary value of each building
VALHUM	Human occupation on the specified building
SE_SISMO	Structural typology of each building

An example of the shape file containing the exposure data is depicted in Figure 7.3.

Id	CONSECUTI	OBJECTID	OBJECTID_1	CODCATEDIF	MANZ	ID_BARRIO	NUM_PISOS	OCUPEDIF	USO	VALFIS	VALHUM	SE_SISMO
3	4	5	24	1003U601021	1003U601021	512	1	8	670	18073.22	8	7
4	5	6	25	1003U601021	1003U601021	512	1	1	670	27672.84	1	7
5	6	7	26	1003U601022	1003U601022	512	1	3	670	10929.74	3	13
6	7	8	27	1003U601022	1003U601022	512	1	1	670	25414.13	1	7

Figure 7.3 Extract of a *.shp file which can be used as inventory in CAPRA-GIS

All of this information is fed directly into the *.dbf file, which is directly associated to the *.shp file. The *SE_SISMO* field contains the structural type for each building and with this information, CAPRA-GIS relates each building with the corresponding vulnerability function, which is then used to compute the expected damage and loss. The *VALHUM* field contains the number of occupants in each building, which is used to compute human losses using a specific vulnerability function which is in the *.fvu file, and that relates the human loss to given a specific ground-motion intensity parameter (GMIP), such as spectral acceleration. It is worth noting that CAPRA can be run with different type of asset as long as the user has the necessary vulnerability function.

Figure 7.4 Screenshot of the CAPRA *.shp file for the Managua area

7.5 Hazard Module

CAPRA-GIS is a multi-hazards software and it can compute different natural hazards including earthquakes.

CAPRA-GIS does not compute directly the hazard but instead requires the user to input a specific standardised hazard file which includes a collection of stochastic scenarios associated to a specific annual frequency of occurrence, the intensity distribution and the variability across the region of interest for each scenario (i.e. probabilistic intensity maps). It is possible to use a number of different Ground-Motion Intensity Parameters (GMIP's) with CAPRA such as PGA, spectral accelerations for different fundamental periods (as to later interpolate the spectral accelerations at the periods of the structures that belong to the corresponding building stock), etc, provided that a consistent vulnerability function that uses the same GMIP to express the probability of damage is input into the program. All of the hazard information is organized in a standardized *.ame file, so that CAPRA-GIS may read and visualize this data. The current version of the PSHA software CRISIS2007 (version 7.2) can create *.ame files that can immediately be used in CAPRA-GIS. In the following figure an intensity map and the metadata description are shown.

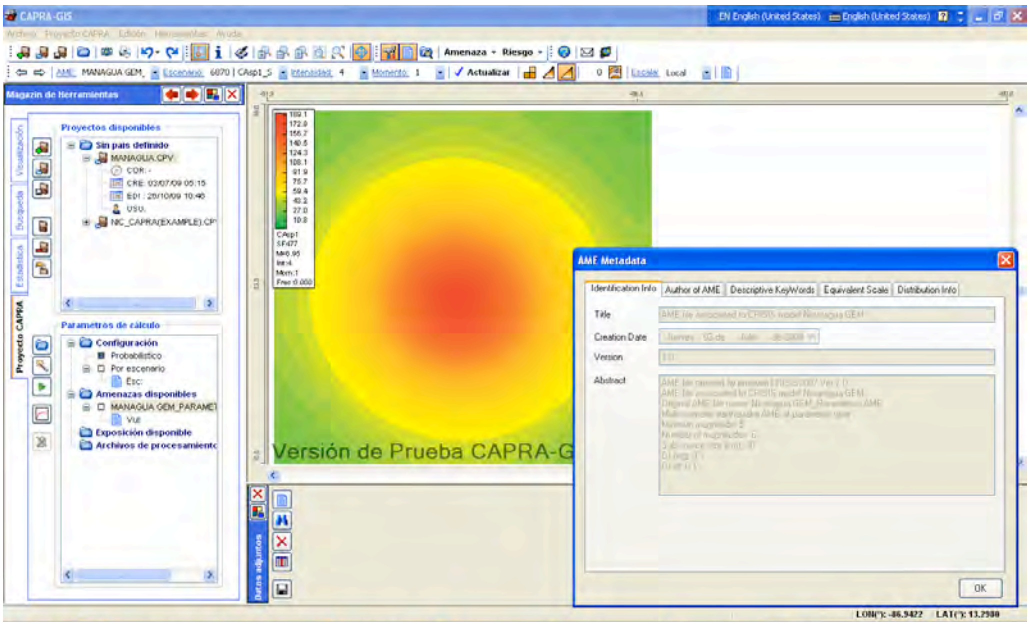


Figure 7.5 Screenshot of an intensity map from the *.ame file in CAPRA and the metadata description

7.6 Vulnerability Module

A vulnerability function for each building type needs to be provided to the CAPRA-GIS as an input. These vulnerability functions need to be consistent with the hazard curves in terms of the GMIP used to express the damage ratios. Functions can be developed with another stand-alone tool called ERN-Vulnerabilidad (Figure 7.6) developed also by ERN. This code considers different methodologies to compute these vulnerability functions, in particular the methodology proposed by Miranda [1999] (though it is also possible to use other methods to create the vulnerability functions, such as the Capacity Spectrum Method), which computes the maximum inter-storey drift ratio (IDR_{max}) of all storeys in a particular building for a given ground intensity level. This is combined with a fragility function that computes the expected damage in terms of IDR, and using a composition of functions, these are merged into a single fragility function dependent on the earthquake intensity level. The variance is adjusted to ensure that the variance is zero for no seismic demand and infinite demand, since the expected damage is surely zero and complete (100%) for these demand levels, respectively. The parameters to adjust the variance are determined by expert judgement.

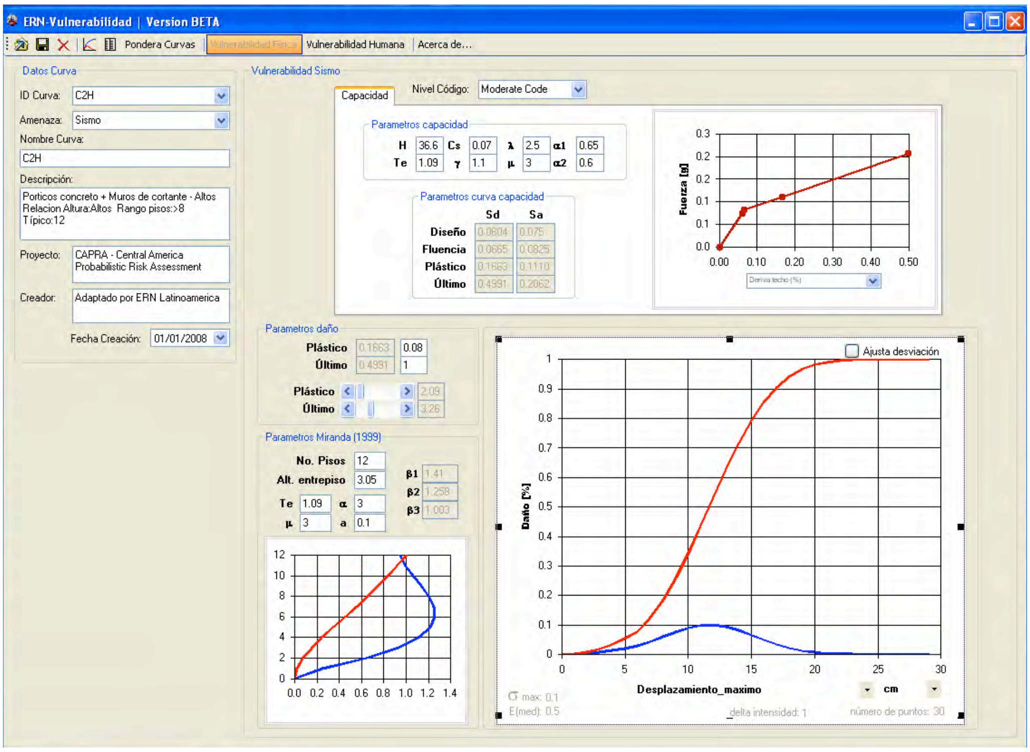


Figure 7.6 Screenshot of ERN-Vulnerabilidad code, which is used for developing fragility functions for different structural typologies

In reality, CAPRA is not tied to any specific method to compute vulnerability functions; it is just the format of the vulnerability function that has been standardised. ERN-Vulnerabilidad uses the method of Miranda [1999], but it can also use other analytical methods as long as the vulnerability functions are prepared *a priori* in the required format. If the user already has fragility functions (based on discrete damage states) they can be relatively easily converted to the vulnerability functions required by CAPRA, as long as the ratio of cost of repair to cost of replacement can be estimated for each damage state. The required vulnerability function format is presented in Figure 7.7. Once the expected damage is obtained, for a given earthquake scenario, from Figure 7.7, it is multiplied by the value of the assets (which might be in economic or human terms) to compute the average loss, and then the variance is used to calculate the variance of the loss.

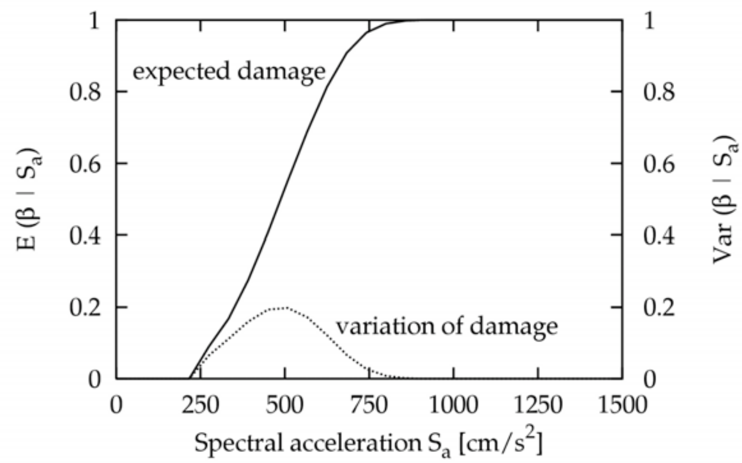


Figure 7.7 Vulnerability function of mean damage given a level of intensity, and variability of this function, as required for CAPRA

7.7 Output

Due to the use of direct vulnerability functions in CAPRA, the damage distribution is not calculated and only the mean damage ratio can be obtained. Notwithstanding that, it is worth noting that, since for a given intensity the full loss probability distribution is known, any percentile of the loss distribution could be computed. Moreover, the user can back-calculate the damage distribution if he/she knows the percentages used to compute the mean damage ratio. The economic and social losses are calculated in the same way using the appropriate vulnerability functions. The following outputs are given by the program:

- Loss Exceedance Curve (LEC): The annual frequency with which a certain monetary or human-loss threshold value is exceeded (Figure 7.8);
- Probable Maximum Loss (PML): Total loss for a particular annual exceedance frequency;
- Aggregated Average Annual Loss (AAAL): Expected loss per year, which is the sum of all the expected annual losses of the values at risk;
- Average Annual Loss (AAL): Expected annual loss per building.

CAPRA-GIS is also able to perform earthquake losses for specific scenarios, such as historical earthquakes or future events that are likely to occur.

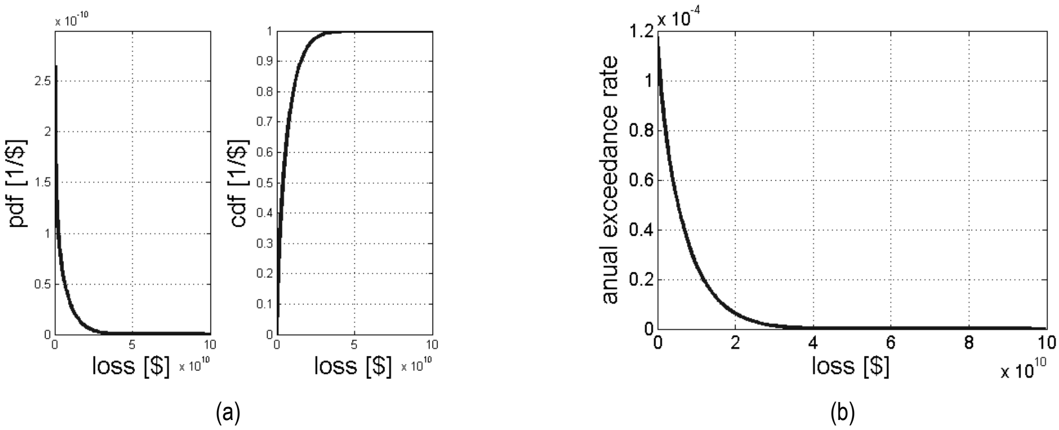
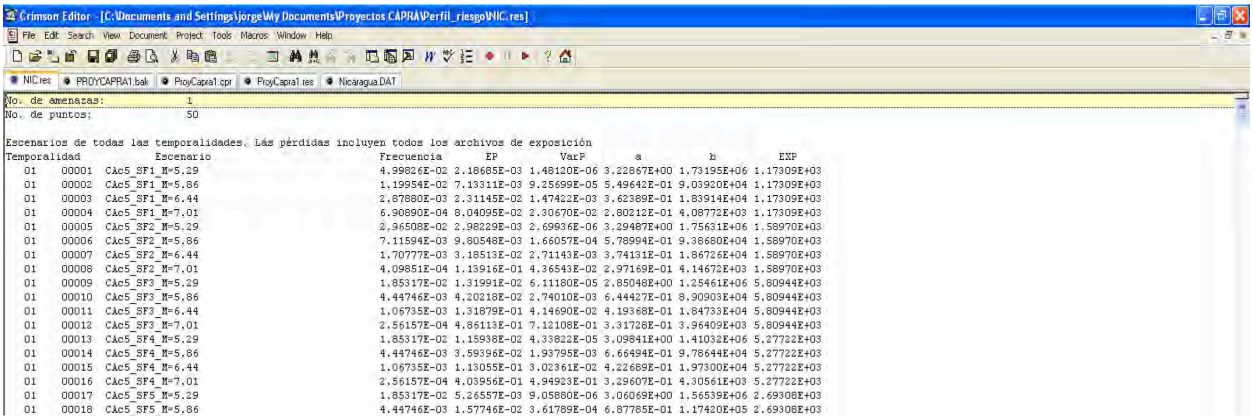


Figure 7.8 (a) PDF and CDF of the total loss characterized as a beta distribution. (b) Loss curve for a single earthquake scenario.

The results per building are appended to the *.shp file, which are the total annual loss value per building and per peril, and the corresponding normalized ratio of annual loss per building or expected annual damage, as well as the annual human loss per building and per peril. These can be visualized (see Figure 7.9) using the GIS tools in CAPRA.

Figure 7.9 Expected mean damage for each building, obtained using CAPRA-GIS



No. de amenazas: 1
No. de puntos: 50

Escenarios de todas las temporalidades. Lás pérdidas incluyen todos los archivos de exposición

Temporalidad		Escenario	Frecuencia	EP	VarP	a	b	EXP
01	00001	CAc5_SF1_M=5.29	4.99826E-02	2.18685E-03	1.48120E-06	3.22867E+00	1.73195E+06	1.17309E+03
01	00002	CAc5_SF1_M=5.86	1.19954E-02	7.13311E-03	9.25699E-05	5.49642E-01	9.03920E+04	1.17309E+03
01	00003	CAc5_SF1_M=6.44	2.87880E-03	2.31145E-02	1.47422E-03	3.62389E-01	1.83914E+04	1.17309E+03
01	00004	CAc5_SF1_M=7.01	6.90890E-04	8.04095E-02	2.30670E-02	2.80212E-01	4.08772E+03	1.17309E+03
01	00005	CAc5_SF2_M=5.29	2.96508E-02	2.98229E-03	2.69936E-06	3.29487E+00	1.75631E+06	1.58970E+03
01	00006	CAc5_SF2_M=5.86	7.11594E-03	9.80548E-03	1.66057E-04	5.78994E-01	9.38680E+04	1.58970E+03
01	00007	CAc5_SF2_M=6.44	1.70777E-03	3.18513E-02	2.71143E-03	3.74131E-01	1.86726E+04	1.58970E+03
01	00008	CAc5_SF2_M=7.01	4.09851E-04	1.13916E-01	4.36543E-02	2.97169E-01	4.14672E+03	1.58970E+03
01	00009	CAc5_SF3_M=5.29	1.85317E-02	1.31991E-02	6.11180E-05	2.85048E+00	1.25461E+06	5.80944E+03
01	00010	CAc5_SF3_M=5.86	4.44746E-03	4.20218E-02	2.74010E-03	6.44427E-01	8.90903E+04	5.80944E+03
01	00011	CAc5_SF3_M=6.44	1.06735E-03	1.31879E-01	4.14690E-02	4.19368E-01	1.84733E+04	5.80944E+03
01	00012	CAc5_SF3_M=7.01	2.56157E-04	4.86113E-01	7.12108E-01	3.31728E-01	3.96409E+03	5.80944E+03
01	00013	CAc5_SF4_M=5.29	1.85317E-02	1.15938E-02	4.33822E-05	3.09841E+00	1.41032E+06	5.27722E+03
01	00014	CAc5_SF4_M=5.86	4.44746E-03	3.59396E-02	1.93795E-03	6.66494E-01	9.78644E+04	5.27722E+03
01	00015	CAc5_SF4_M=6.44	1.06735E-03	1.13055E-01	3.02361E-02	4.22689E-01	1.97300E+04	5.27722E+03
01	00016	CAc5_SF4_M=7.01	2.56157E-04	4.03956E-01	4.94923E-01	3.29607E-01	4.30561E+03	5.27722E+03
01	00017	CAc5_SF5_M=5.29	1.85317E-02	5.26557E-03	9.05880E-06	3.06069E+00	1.56539E+06	2.69308E+03
01	00018	CAc5_SF5_M=5.86	4.44746E-03	1.57746E-02	3.61789E-04	6.87785E-01	1.17420E+05	2.69308E+03

Figure 7.10 Expected total damage for each hazard scenario obtained using CAPRA-GIS

The main results of CAPRA-GIS are provided in a *.res file (Figure 7.10); a file that can be opened using any text editor and contains, for each scenario analyzed, the expected loss (EP), the variance of the loss (VarP), and parameters *a* and *b* that fully define the beta distribution of the loss for a given scenario. Also, the exposed value of the whole portfolio of structures (EXP), which is also the upper limit of the assigned beta distribution of the total loss, where the lower limit of this distribution is 0.

8 RISKSCAPE

8.1 Summary of Software

In June 2004 the New Zealand Foundation for Research, Science and Technology (FRST), the research funding agency for central government, provided funding for the development of a Regional RiskScape Model.

The organizational structure adopted was a 50:50 joint venture comprising GNS Science and the National Institute of Water & Atmospheric Research (NIWA). Both NIWA and GNS Science are Crown Research Institutes (CRI) and have comprehensive expertise in the field of natural hazards and the resultant impacts on communities and their assets.

The Regional RiskScape is an easy-to-use multi-hazard risk analysis tool. It converts hazard information into the likely impacts for a region, for example, damage and replacement costs, casualties, economic losses, disruption, and number of people affected. Its default geographical setting is New Zealand but it can be expanded to other regions in the world.

Presently, RiskScape 0.1.4 (the version reviewed herein) is a compiled version using the Java programming language and tools. The source code was made available to the GEM1 Risk team for testing, however the source codes of the tool sets and build tools had been removed. During the drafting of this report, RiskScape was undergoing major changes and it is expected that an updated version will be released in the very near future.

8.2 Methodology

The structure of the RiskScape software is shown in Figure 8.1. The software is able to manage five hazard types: earthquake shaking, inundation related to river flooding, inundation related to tsunami, wind storm and volcanic ashfall.

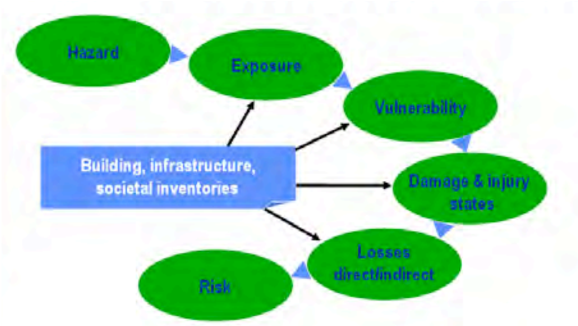


Figure 8.1 The main components and the sequence of operations of the Regional RiskScape tool run from simulating the natural hazard through to calculating losses and impacts from the vulnerable areas of exposed communities and their assets

Regarding earthquake hazard and vulnerability, RiskScape uses an empirical methodology where ground shaking intensity, measured by the Modified Mercalli Intensity scale, is calculated and then this is related to the mean damage ratio for different types of buildings. To calculate the social losses, the number of collapsed buildings of a given building typology is first estimated and then casualty factors are applied to their occupants. For instance, for timber houses, 1% of people present at the time in all collapsed houses may be killed, while 10% may be injured.

Vulnerability functions have been developed for different kinds of assets such as buildings, contents, vehicles, water-supply network, sewage network, storm water network and the population (injuries and fatalities).

8.3 IT Details

The system is written with the open-source software Java, and does not require a GIS licence to run it. The input and output files are, however, GIS-compatible. A simplified flowchart showing the software process is presented in the following figure.

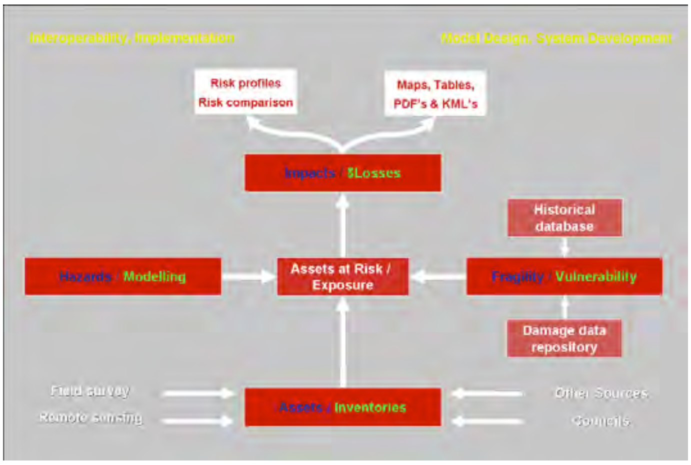


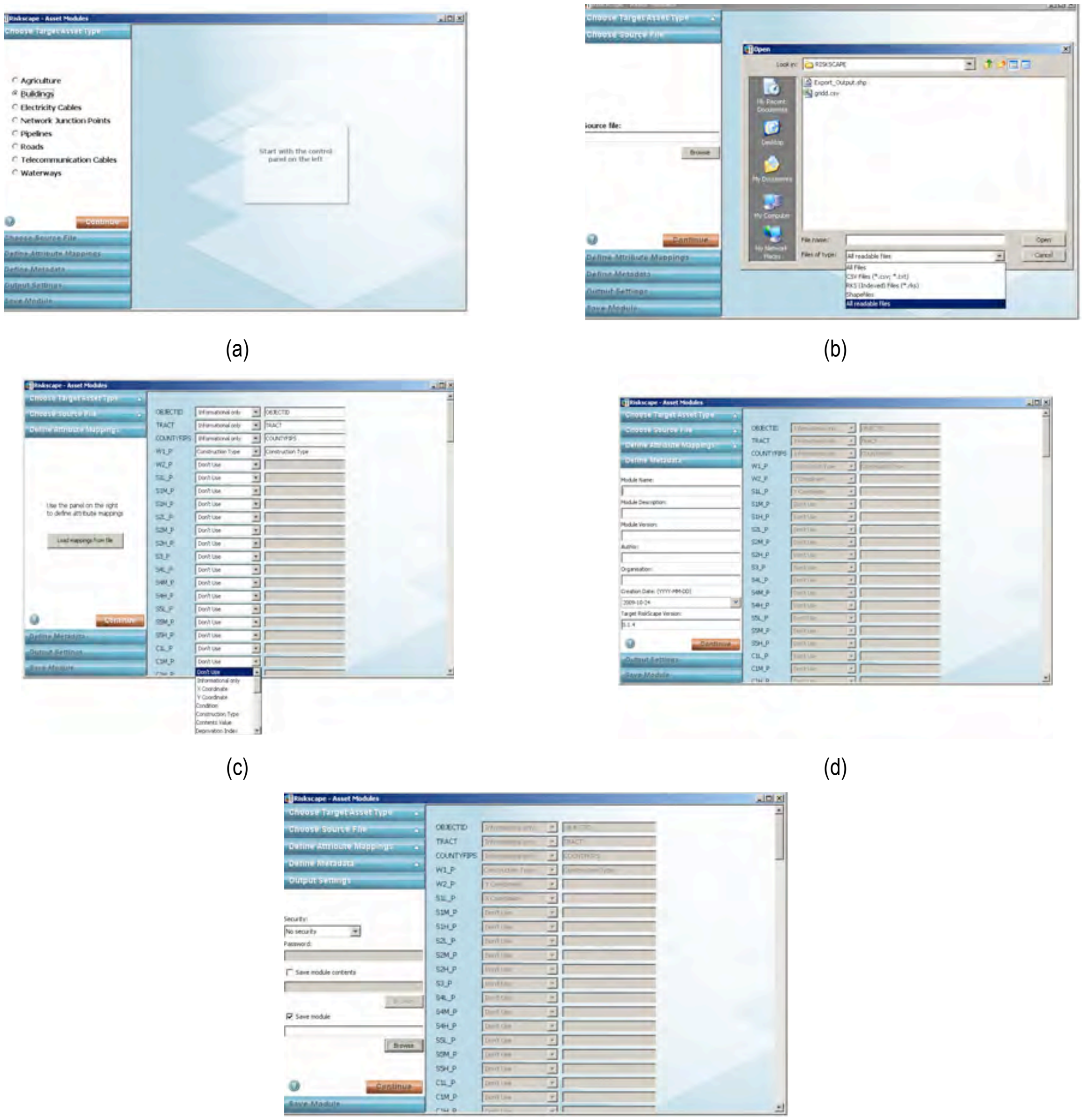
Figure 8.2 Flowchart showing the software process of RiskScape

RiskScape features two GUIs. One GUI (Figure 8.3a), which is the program itself, guides the user along the different steps to compute an analysis from the beginning to the end of the process; such steps comprise choosing the hazard and the hazard model, defining the model parameters, selecting the assets, selecting the aggregation, choosing the type of loss, selecting the loss parameters, computing the damage and producing the report. This figure shows the final screen, after the analysis has been performed, where the user can choose the results presented in four different manners *.kml, *.shp, *.xls or *.pdf files. The second GUI (Figure 8.3b), named RiskScape Tools, is used to input user-defined data which will be stored in the database of the program. The different modules are: Asset Builder, to input the inventory type, data and attributes; Hazard Builder, to input the hazard and hazard model; Aggregation Builder, to input different files that will be used to perform the aggregation of the output for different assets; Fragility Builder, to create user defined fragility curves; Query Mapper, to generate a code which will perform various user-defined analyses. It is noted that the home page is to be fully revamped with the next release.

(a) (b)

Figure 8.3 (a) Final screen of the RiskScape analysis process (b) GUI of the RiskScape tools

An example to introduce data to the program is given in what follows by using the Asset Builder. This example is given to show how the users are well guided through the software processing their analyses. The first screen, Figure 8.4a, gives the user the option to select the asset from the asset types that RiskScape recognizes (asset and attributes types will be explained in more detail in the following section). Then, the source file, which can be in either *.csv, *.shp, *.txt or *.rks format, should be input (Figure 8.4b). The third step is to select, from the different fields in the data file, the ones that RiskScape will use as attribute types (Figure 8.4c). After that, the metadata information for the data file is given (Figure 8.4d) and finally the file is saved (Figure 8.4e).



(a) (b) (c) (d) (e)

Figure 8.4 Steps to input data by using the RiskScape Tools (a) Select target asset type, (b) Select Source File, (c) Select Attribute Mappings, (d) Define Metadata, (e) Output Settings and Save file

8.4 Exposure Module

In RiskScape, a ‘type’ describes a specific kind of asset, which for a given hazard will have a certain fragility function, allowing RiskScape to remain generic. The following asset types can be modelled in RiskScape: buildings, electricity cables, telecommunication cables, network junction points, roads, agriculture, waterways and pipelines.

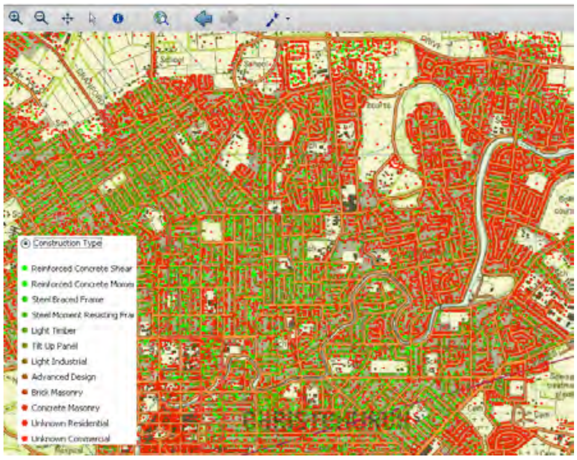
There are a lot of attribute types which apply to the asset types mentioned above. The inventory has to contain information about these specific attributes. With regards to the buildings, Riskscape needs to have information about the condition of the structure (e.g. sound or deficient), the construction type, the content values, the deprivation index, employee daily income, floor area and height, floor type, footprint area, occupancy, parapet, replacement cost, roof cladding class, number of storeys, use category, vehicle value, vehicles, wall cladding class and year of construction. These attributes are described in detail in the help provided by the software. In Table 8.1 an example of the table shown in the help is reported.

Table 8.1 Attribute types that may be affected in a seismic event (RiskScape)

Attribute type	Applies to	Category	Unit	Equivalent to	
Construction Type	Buildings	1 : Reinforced Concrete Shear Wall 2 : Reinforced Concrete Moment Resisting Frame 3 : Steel Braced Frame 4 : Steel Moment Resisting Frame 5 : Light Timber 6 : Tilt Up Panel 7 : Light Industrial 8 : Advanced Design 9 : Brick Masonry 10 : Concrete Masonry 11 : Unknown Residential 12 : Unknown Commercial			
Floor Area	Buildings	-	m²	-	-
Storeys	Buildings	-	no. of storeys	-	-
Year of Construction	Buildings Electricity Cables Telecommunication Cables Network Junction Points Pipelines	-	year	-	-

The model building types should be user-defined as long as the user can join them with a fragility curve. The level of detail is also user-defined and it can be regional, provincial or local. It always depends on the available information the user owns. Instead, as a default, RiskScape contains detailed inventory and population information for at least three locations in New Zealand (Figure 8.5a). Information on the general distribution of the numbers and broad age of structure and numbers of people living in these communities were obtained from the New Zealand census. The user does not need to create an inventory module for these three places because they are stored into the software itself. In order to get information such as construction type, building age, floor height, roof type and routes for utilities such as water supply, sewerage, road and power, a number of sources of data were used such as Statistics New Zealand, Councils and complementary surveys were carried out. In Figure 8.5b an example of the buildings classified according to the construction type is shown.

(a)



(b)

Figure 8.5 (a) Three pilot locations in New Zealand (b) Example of building inventory classified according to the construction type

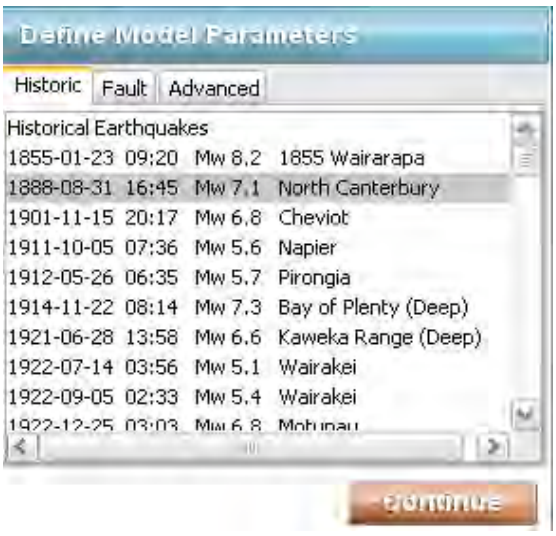
8.5 Hazard Module

The RiskScape software is able to manage five hazard types: earthquake, inundation related to river flooding, inundation related to tsunami, storm and volcanic ashfall.

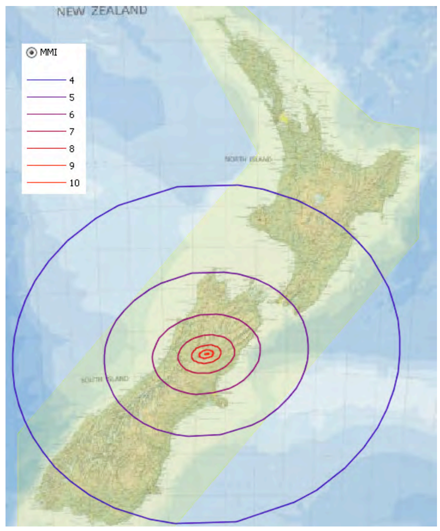
The hazard modelling to account for earthquakes follows the subsequent steps:

- Define the earthquake, strength and place. The user can choose to define a point source, a fault or he/she can choose, with a scroll down menu, a historical earthquake (Figure 8.6a);
- Estimate the ground motions at various distances using attenuation functions in terms of Modified Mercalli Intensity (MMI scale);
- Allow for the effects of different types of soil (amplification effects and indirect effects like liquefaction and landslides).

By default the Earthquake hazard module currently is based on calculated intensities in terms of the Modified Mercalli Intensity scale calibrated for New Zealand and it is hard-coded in the software (Figure 8.6b). Nevertheless, the user can input his/her own hazard data (e.g., ground-motion prediction equation) by using the hazard builder contained in the RiskScape Tools, in a similar fashion to the way the inventory was loaded into the database.



(a)



(b)

Figure 8.6 (a) Scroll down menu used to select and historical earthquake (b) Example of hazard computed by the software

RiskScape software can currently only be used for deterministic scenario earthquakes in New Zealand. It will be possible to estimate losses using select scenario earthquakes for other locations in the future once inventory data becomes available. It is noted that RiskScape cannot currently perform probabilistic analyses, though this will be included in future versions.

8.6 Vulnerability Module

The vulnerability of an asset exposed to a particular hazard is dependent on the magnitude of the hazard (e.g. macroseismic intensity) and the characteristics (attributes) of the asset (e.g. roof cladding, structural strength of a building). The vulnerability is described by the expected degree of loss, which can be expressed as a damage ratio. The mean damage ratio is calculated from the relative proportion of the repair cost to the total replacement cost, with 1.0 being totally damaged. RiskScape uses an empirical methodology that uses intensity to estimate mean damage ratio for different types of buildings (Figure 8.7). This means that the program goes directly to the damage without calculating any structural response such as, for instance, the performance point. To calculate the social losses, the number of collapsed buildings of a given building typology is first estimated and then, casualty factors are applied to their occupants. Figure 8.7b is used to directly compute the mean collapse rate, which is then used to compute social losses. Similarly to the inventory and hazard input, fragility curves can also be defined by the user through the fragility creator in the RiskScape Tools (Figure 8.8). The fragility types that can be employed by RiskScape and created with the aforementioned tool are the following: Affected Businesses; Affected People; Business Disruption; Loss of Income; Displaced People; Fatalities; Functional Downtime; Injuries – Serious; Injuries – Minor; Injuries; Vehicle Cost; Cost of Contents; Monetary Cost; Uninhabitable Buildings; Damage; Unaggregable Damage.

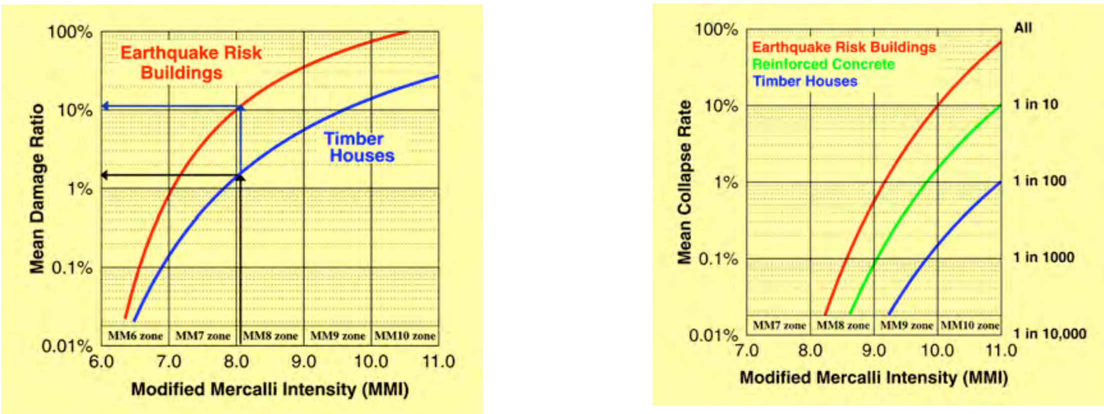


Figure 8.7 Vulnerability curves used in RiskScape to compute (a) mean building damage ratio (b) social Losses

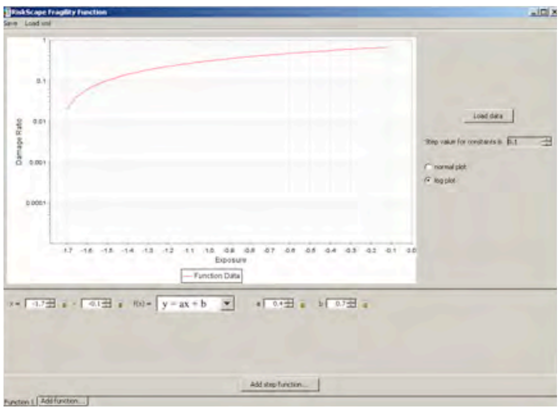


Figure 8.8 Fragility curve builder tool (RiskScape)

8.7 Output

RiskScape currently calculates the mean damage ratio and does not consider any other level of damage. The social losses are calculated only based on collapsed buildings. RiskScape output is given in four different manners *.kml, *.shp, *.xls (Figure 8.9) or *.pdf files which the user can choose in the final screen of the analysis process. The *.kml file for the economic losses is presented in Figure 8.10; this 3D plot is very professional-looking, but it does not appear to have a legend. The direct economic losses, that include both monetary cost and cost of contents, are calculated in RiskScape by multiplying the mean damage ratio by the total value of the assets. Figure 8.11 presents the *.shp of the economic losses and damage rate. Figure 8.12 presents two pages of the *.pdf provided by the code. Finally, RiskScape also calculates fatalities, minor injuries and serious injuries at different times of the day: night time, daytime and full occupancy. (Figure 8.13).

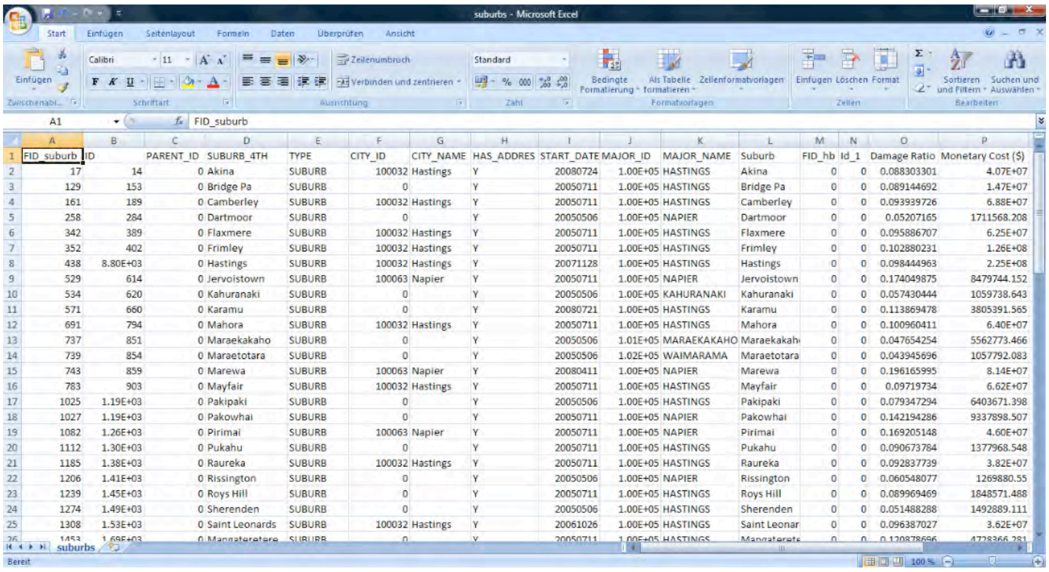


Figure 8.9 Example of *.xls output of RiskScape

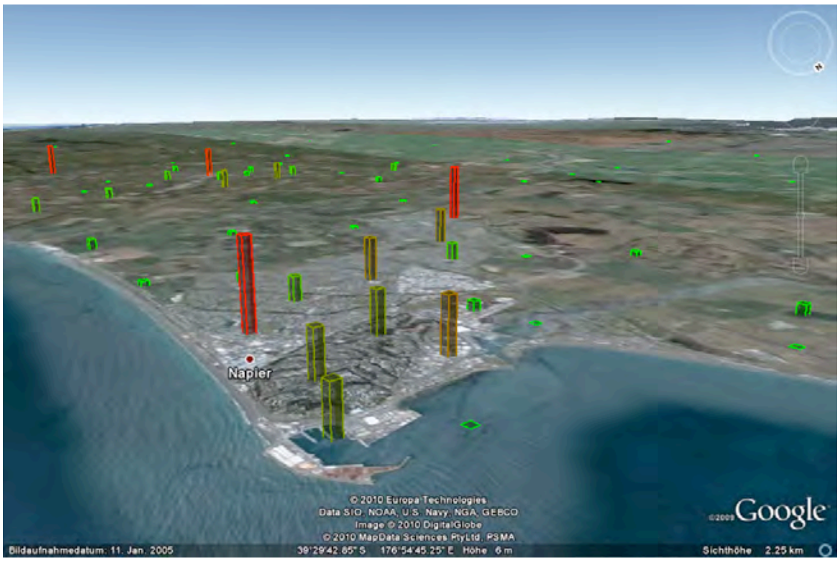


Figure 8.10 Example of *.kml output of RiskScape. Economic losses

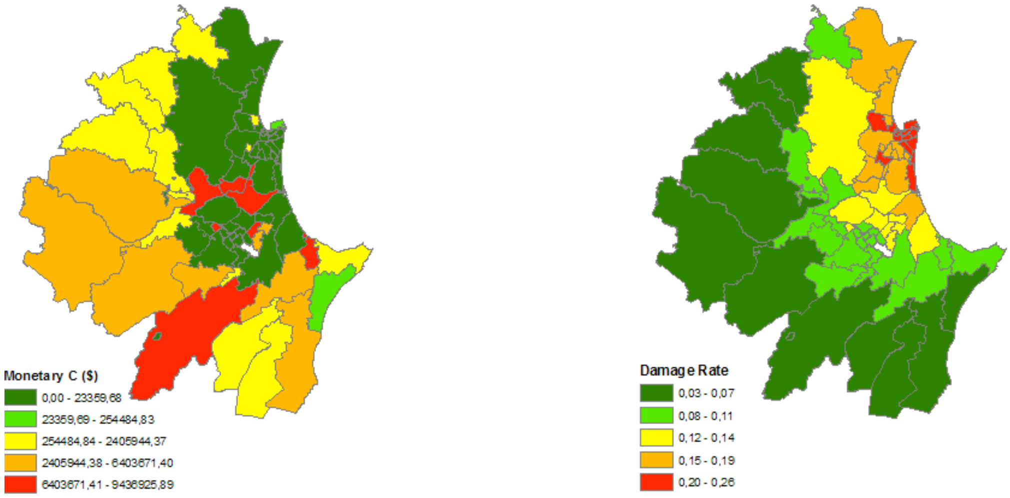


Figure 8.11 Example of *.shp output of RiskScape

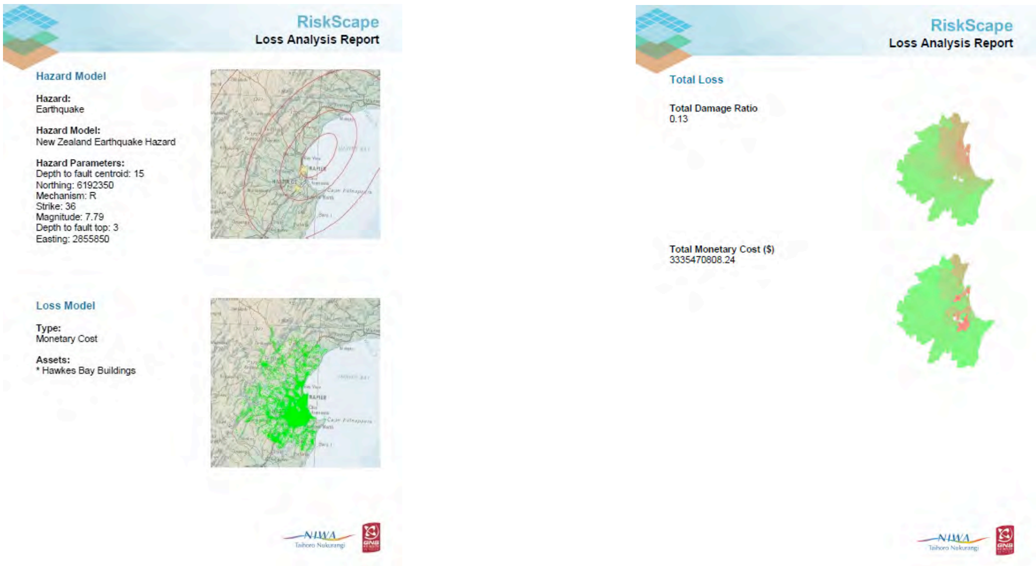


Figure 8.12 Example of *.pdf output of RiskScape

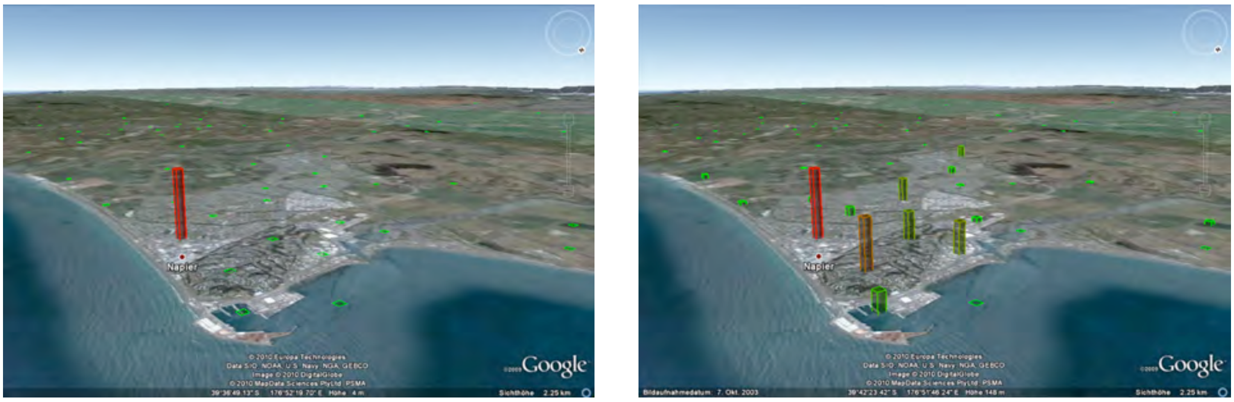


Figure 8.13 Social loss results in RiskScape (a) Number of Fatalities, (b) Number of minor injured people

9 LNECLOSS

9.1 Summary of Software

LNECLoss software is developed at LNEC (Laboratório Nacional de Engenharia Civil – National Laboratory for Civil Engineering) in Lisbon in the framework of seismic risk mitigation projects in Portugal. The Portuguese national civil protection authority and research European projects like LESSLOSS (FP6) are some of the main supporters of this software as it can be used for emergency planning, emergency management and seismic risk analysis and mitigation. It can be used to study and consequently reduce the effects of future earthquakes on population and building structures. This tool aims to be supplied to local and regional Portuguese authorities to provide a decision support system to evaluate seismic risk and to initiate mitigation programs.

It is a straightforward and an automatic methodology to estimate seismic losses that can be integrated on a Geographic Information System. The software contains databases suitable for Portugal, but it could potentially be applied to any user-specified regions. It can compute buildings damage, economic losses and social losses using different methodologies, both empirical and mechanical. All the results are given in *.txt tables and the user can modify, import data and manipulate them in any way. Moreover, it considers both seismic action at the bedrock and seismic action at the surface. In its framework there is the interaction of different specialized sub-fields of research such as engineering seismology, earthquake engineering, geotechnical engineering, probability theory and computer and GIS development tools. All these features make it very useful and complete.

9.2 Methodology

The software is divided into 5 different modules that are independent of each other: (i) seismic action at bedrock, (ii) seismic action at surface, (iii) vulnerability / building damage, (iv) human losses and (v) economic losses. Being a modular structure, LNECLoss is constantly being updated in a very simple way, both in terms of data and methodologies.

This software considers both empirical and mechanical approaches to model building damages. The first approach relies on the seismic ground motion defined as a macroseismic intensity whereas the second requires that seismic ground motion is defined as a spectral quantity. It is worth noting that it is possible to model spectrally the characteristics of seismic motion, for a given scenario, and to use empirical relationships to transform peak ground-motion quantities in macroseismic intensity and compute damages applying an empirical approach.

Considering seismic motion, its spectral characteristics can be modelled using empirical attenuation laws or a finite-fault non-stationary stochastic seismological model, based on random vibration theory. LNECLoss introduced some major improvements to take into account site effects due to soil dynamic amplification, evaluated by means of an equivalent stochastic nonlinear one-dimensional ground response analysis of stratified soil profile units designed for the region.

The flowchart of the software is shown in Figure 9.1. For what concerns the seismic action, it is feasible to calculate the values of PGA, PGV, PGD and response spectrum both at the bedrock and at the surface. Alternatively, intensity maps can be used to compute risk. The buildings damage can be computed following statistical models or mechanical models. The former considers four different empirical studies [Di Pasquale & Orsini, 1997; Zuccaro & Papa, 2002; Tiedemann, 1992; Giovinazzi & Lagomarsino, 2004] and the latter is based on the capacity spectrum method of HAZUS [FEMA, 1999]. There are also different methodologies to compute losses. The user can choose the most suitable approach for his/her purpose. With regards to the social losses, three different models are stored inside the program [Tiedemann, 1992; Coburn & Spence, 2002; FEMA, 1999], while two models are considered for economic losses [SSN [1]; FEMA, 1999].

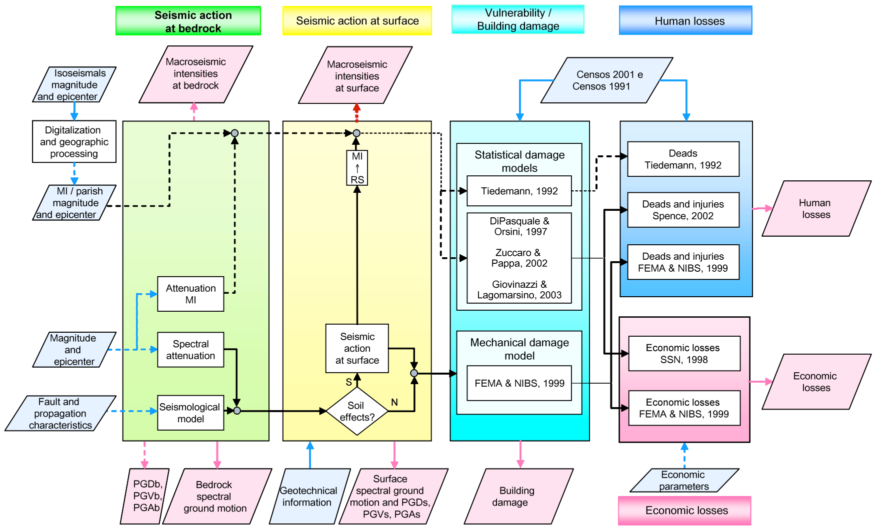


Figure 9.1 Flow chart of LNECLoss (MI = macroseismic intensity, RS = Response spectra) [adapted from Sousa, 2007]

9.3 IT Details

LNECLoss is written in a Fortran programming language and it can be integrated on a GIS. The Geographic Information System integration is used as a high level language and as a GUI for input and output data and maps (Figure 9.2). To compute analyses in GIS, the calculations are performed through Fortran code compiled as DLLs. Moreover, the algorithm is transparent and the user can easily modify the Fortran's routines and he/she can adapt the code to his/her purposes. Unfortunately, there are not any manuals and in this way the user is not always easily supported (though the queries of the GEM1 Risk team were always rapidly answered by the developers of the software). Notwithstanding this, there are some documents and papers which explain the software itself.

LNECLoss's input and output files can be generated as *.txt files. These files can easily be transformed into a database and integrated in a GIS program. The level of detail is completely user-defined and it depends on the size of the geounit. The software is really flexible because each module is independent. For instance, the user can choose to begin to model seismic ground motion and use the output of the correspondent routine as the input of the building damage module. Alternatively, the user can upload their own ground motion map to feed the building damage module. Furthermore, the user can use the data_source_regional.txt file to change the parameters and modify the choices he/she has to take before running analyses. This is very useful to have a global idea of the chosen options to perform the analysis. In Figure 9.3 an extract of this file is shown. There are different flags which control several approaches. For instance, the user can choose between different types of fault editing the number corresponding to the fault he/she needs (1 for known fault, 2 for empirical fault, 3 for Carvalho 2007, 4 for Bommer *et al*. 1998).

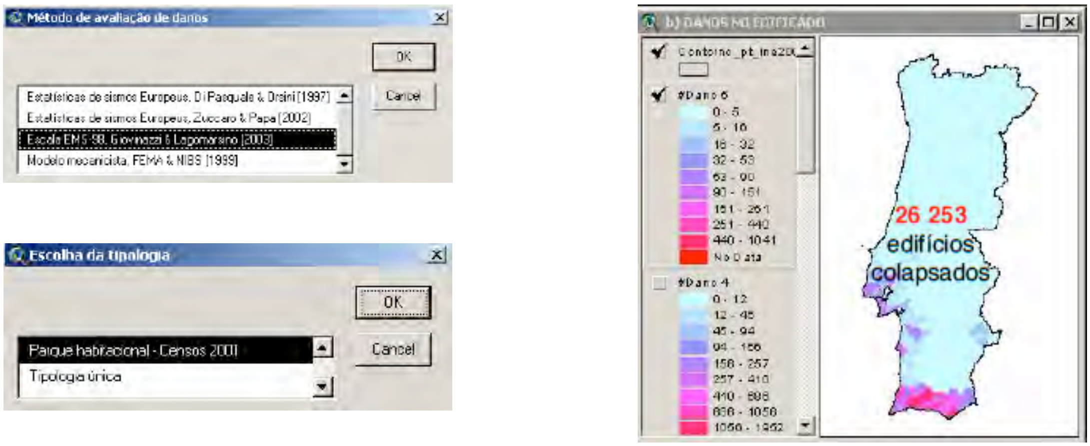


Figure 9.2 Screenshots of the GUI of the GIS LNECLoss [Sousa, 2006]

```
site coordinates:
189168.00 //  site cartographic cocordinates in meters
14349.3   //  site cartographic cocordinates in meters
10.0d0    //dist=Distance to epicenter [km]
1        //  Flag =0 Statistical methodology;  =1 Hazus
7.0d0     //  Macroseismic intensity
soil properties:
1         //  soil profile: =0 if rock, =1 if different than rock
2         //  number nestratos
0, 20     //  soil properties, if not rock
1.7, 2.2  //  rho
110, 600  //  v0min [m/s]
170, 600  //  v0max  [m/s]
14, 0     //  nlayers
40, 0     //  ip
fault parameters:
4         //  falha = 1 (known fault); Falha = 2 (empirical fault); Falha = 3 (Carvalho 2007 (attenuation));
Falha = 4 ( Boomer et al, 98)
1         //  slip = 0 (homogeneous) =1 (random) =2 (file)
```

Figure 9.3 An extract of the data_source_regional.txt data

9.4 Exposure Module

The user has to provide different building inventories according to the approaches he/she wants to run (empirical or mechanical). If empirical methods are chosen the building inventory has to be provided in terms of building typologies. Four vulnerability classes are used by the Di Pasquale & Orsini [1997] and Zuccaro & Papa [2002] relationships which are based on the MSK intensity scale. Giovinazzi & Lagomarsino [2002; 2004] method identifies 15 building typologies, described in EMS-98 vulnerability table, plus some behaviour modifier factors, to evaluate the typological vulnerability index. If the mechanical method is chosen the user should define the capacity and fragility parameters (like in HAZUS) that best describe the building typologies within the inventory. In the latter, the user is not forced to assign a vulnerability class to his/her model building types. However he/she needs more information about the structural parameters of the building. The statistical option Tiedemann [1992] follows the mechanical typological classification as the independent variable of this model is the seismic coefficient that is also used in the mechanical model. In any case, the user has to provide a residential building database that is geographically organized. This database has to include the mean total floor area of a building typology for each geounit. Moreover, the number of floors and the era of construction are taken into account. In addition to the building inventory, the user has to supply population inventory to compute social losses. This database comprises the number of inhabitants with the same level of geographic organization of the building inventory. It considers the number of inhabitants living in buildings classified according to their age, structural elements and number of floors. The last database is related with the economic losses and it includes the average floor area, repair and replacement cost for each geounit and model building type.

The inventories are stored in an external *.txt file and it is extremely convenient. In this way it is simple to transform a general preliminary inventory into the format required by LNECLoss. The level of detail is user-defined, the unit could be a regional unit, a provincial unit, a local unit and it could be an arbitrary polygon (regular or irregular). The user can configure independently the inventory, which is not hard-coded in any way. This is due to the flexibility of the software and to its modular structure.

9.5 Hazard Module

The ground motion simulation module is divided into two sub-modules, again independent of each other. The first sub-module deals with the seismic action at the bedrock, while the second one analyses the seismic action at the surface. The user can decide to use both of them to estimate the seismic action or he/she can run just the first module and then go directly to the evaluation of the building damage.

With regards to the seismic action at the bedrock, the user has different choices. Both intensity-based method and response spectral ordinates are considered in the code and the user has to modify data_source_regional.txt file to fix this decision. For what concerns the mechanical approach, the user has to decide the typology of the source he/she wants to use. There are two options. The first one considers the seismic source as a point. For this reason it is necessary to enter the coordinates of the epicentre, the depth and the magnitude. Then, different attenuation laws can be applied. There are some attenuation laws that are already hard-coded in the software. Nevertheless, the user can easily add their own relationships. The second option considers the seismic source as a fault. In this case the user has to supply the geometry of the source (length, width, strike, dip) or optionally, the magnitude and epicentre which will allow the program to design a fault according to magnitude, with epicentre as the centre of the fault. It is possible to use attenuation laws or the stochastic approach. If the user wants to utilize attenuation laws, the distance source-site will be associated to the closer distance between the fault and the site. Finite-fault stochastic simulations require the knowledge of some source parameters and crustal properties of the region [Carvalho *et al.*, 2008]. In the two cases described above, the output could be the PGA, the PGV, the PGD or the response spectrum.

For what concerns the seismic action at the surface, LNECLoss allows the response spectrum to be estimated at the surface as long as the user knows the conditions of the soil under the site of interest. Given a response spectrum at the bedrock, the program computes the Power Spectral Density Function (PSDF). Then, given stratified soil profile units it uses a transfer function to estimate the PSDF for any location at the surface level taking into account the nonlinear behaviour of the stratified geotechnical site conditions. After that, the code reconverts the PSDF into a response spectrum to be used in the following steps of the analysis.

The parameters of the seismic action (e.g., coefficients of the attenuation laws), except for the features of the source, are hard-coded in the software. To modify them the user has to modify the code. On the other hand, the source data are stored in an external *.txt file. In the following figure the path of the seismic action is shown.

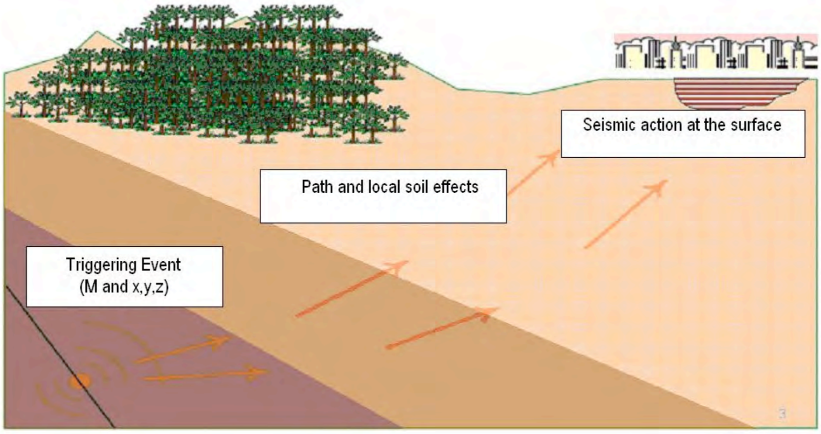
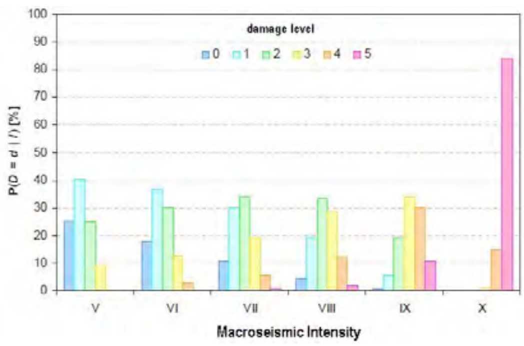


Figure 9.4 Schema of the seismic action path from the source to the surface [Source: adapted from Carvalho, 2007]

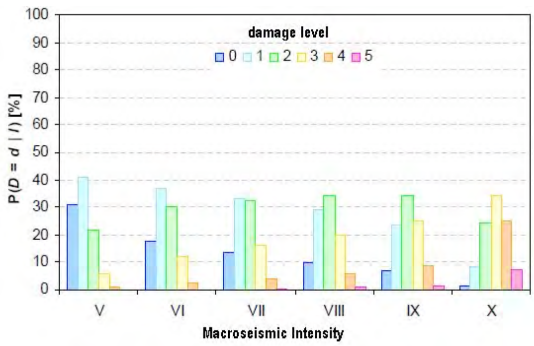
9.6 Vulnerability Module

As stated above, the user can choose to follow empirical methods or a mechanical approach to evaluate building damage.

Three statistical buildings damage evaluations are hard-coded in the software: Di Pasquale & Orsini [1997], Zuccaro & Papa [2002] and Giovinazzi & Lagomarsino [2004]. These methodologies are based on empirical matrices and they are intensity-based. The fragility curves, correspondent to 5+1 damage grades, are used to estimate damage conditioned by a seismic intensity level. The statistical damage models are less flexible than the mechanical one, because the user is forced to adopt pre-defined vulnerability classes to describe the inventory. Each approach mainly differs from the other for the values of the empirical matrices and for the kind of vulnerability classes used. Giovinazzi & Lagomarsino matrices are based on EMS-98 vulnerability classes, from A to F, (see Section 5) and they consider 15 structural types according to the EMS-98 vulnerability table. Furthermore, two behaviour modifiers proposed by those authors were also considered in LNECLoss. Those behaviour modifiers are the buildings state of preservation and number of floors. Di Pasquale & Orsini [1997] and Zuccaro & Papa [2002] matrices are based on MSK scale and they consider four vulnerability classes (A, B, C1 and C2). Given the era of construction and the building typology the user can classify the inventory into the four classes and, then, it is possible to apply the damage matrices. In Figure 9.5 an example of the damage probability matrix for the vulnerability class A is shown. Figure 9.5a represents the matrix for the Di Pasquale & Orsini [1997] approach, while Figure 9.5b shows the matrix used in Zuccaro & Papa [2002]. Moreover, Figure 9.6 shows the fragility curves for a building belonging to class A for both methods. The probability distribution is a binomial distribution whose parameters are different according to the method used. The parameters are hard-coded in the software even if the user can modify them at any time.

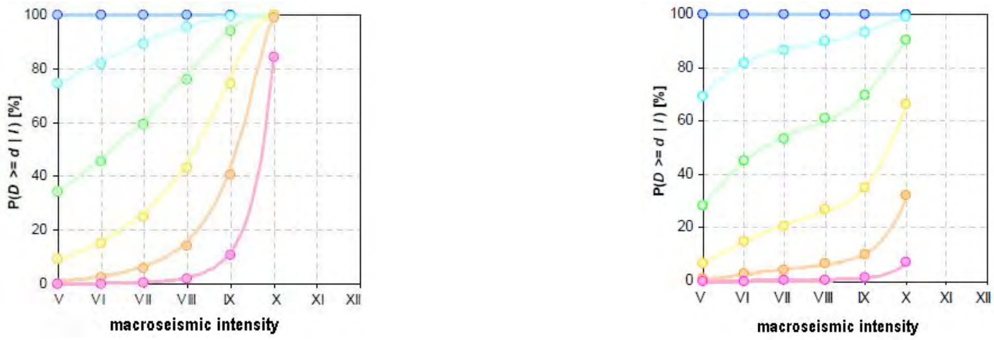


(a)



(b)

Figure 9.5 Damage probability matrix for vulnerability class A. (a) Di Pasquale & Orsini [1997] and (b) Zuccaro & Papa [2002] [Source: Sousa, 2006]



(a) (b)

Figure 9.6 Fragility functions for class A (a) Di Pasquale & Orsini [1997] and (b) Zuccaro & Papa [2002] [Source: Sousa, 2006]

The mechanical approach hard-coded in LNECLoss is based on the capacity spectrum method developed in the HAZUS methodology [FEMA, 1999]. However an innovative technique of the evaluation of the so-called performance point is introduced into the code. This technique takes into account an iterative procedure that estimates sequential demand spectra, with increasing damping, reflecting structure degradation during its cyclic response. While in HAZUS the modifications of spectral demand are represented by the reduction factors, in LNECLoss those modifications are performed through an iterative equivalent non-linear stochastic methodology. Progressive building responses are obtained over the demand spectra till the convergence with median capacity curve is achieved. The capacity curves are completely user-defined as long as the user can define the yield point and ultimate point of the curve. The software reads these values and it computes an elliptical capacity curve between these two points. The fragility curves are lognormal distributed and they are also user-defined as long as the user can provide the values of the median of the curve and its standard deviation. These data are stored in an external table called Structural_building_parameters.txt (Figure 9.7). In this case, the fragility functions are defined in terms of structural response and not in terms of seismic intensity. These curves define five different limit damage states (no damage, slight damage, moderate damage, severe damage and collapse).

```
0.107    //Ay [g]       | capacity curves parameters
0.074    //Dy [cm]      |
0.214    //Au [g]       |
0.739    //Du [cm]      |
0.400    //Kshort
0.200    //Kmod
0.000    //Klong
20.000   //LimDurShort [s]
2500.0   //LimDurMod [s]
0.750    //alpha2
3.2      //h[m]
0.0024 0.0048 0.009 0.021 //delta    | fragility curves parameters
1.180 1.180 1.180 1.180//parbeta     |
```

Figure 9.7 Structural_building_parameters.txt file where capacity and fragility parameters are stored

The social losses in LNECLoss can be computed following three different approaches: FEMA [1999], Coburn & Spence [2002] and Tiedemann [1992]. Most methods to estimate human casualties as a consequence of earthquakes are based on the correlation between building damage and the number of people killed, injured or homeless. Unfortunately, those estimations are always affected by a great level of uncertainty because the same seismic intensity causes an heterogeneous number of victim in different countries. The first approach is the one used for HAZUS methodology and it is strictly connected to the level of damage suffered by the structure. Furthermore, the social losses are classified in terms of level of severity (light injuries, hospitalization, life threatening and death). It is also worth noting that LNECLoss is only considering the direct losses caused by structural damage, because only structural damage is computed by the software. The second approach generally overestimates the ratio between number of deaths and the number of injured people as it should be applied to epicentral regions of severe earthquakes. The third approach is a simplified one, that relates the

Death Rate (DR) with the Mercalli Modified Intensity (MMI) for different typologies of buildings. If the user chooses this last option, the human losses routine estimates for each geounit and for each building typology the Death Rate for a given scenario.

The economic losses are also taken into account. LNECLoss considers the damages in each typology and damage state and it computes damaged building floor areas that can be multiplied by the repair and replacement cost to obtain economic losses. In the following figure the procedure to compute the seismic risk assessment for a deterministic scenario is shown.

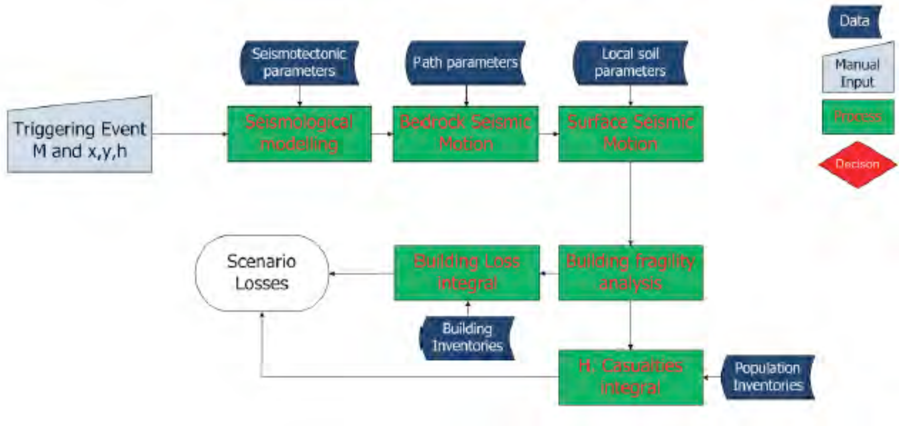


Figure 9.8 Risk assessment for a deterministic scenario

9.7 Output

Five structural damage states are considered in LNECLoss when a mechanical approach to evaluate building damage is used, whereas six damage levels are considered when an empirical approach is chosen. The results are given in terms of number of buildings and in terms of percentages belonging to each damage level. The social losses are given in terms of percentage of people suffering different injury levels according to the method which is used to run the analyses. Finally, the economic losses are provided by the software. Beyond the absolute numbers, the percentages for each result are provided. The results obtained with LNECLoss are stored in external *.txt table and they can be easily exported in a GIS to visualize results. In the following figure some examples of the kind of visualization computed by the software are shown.

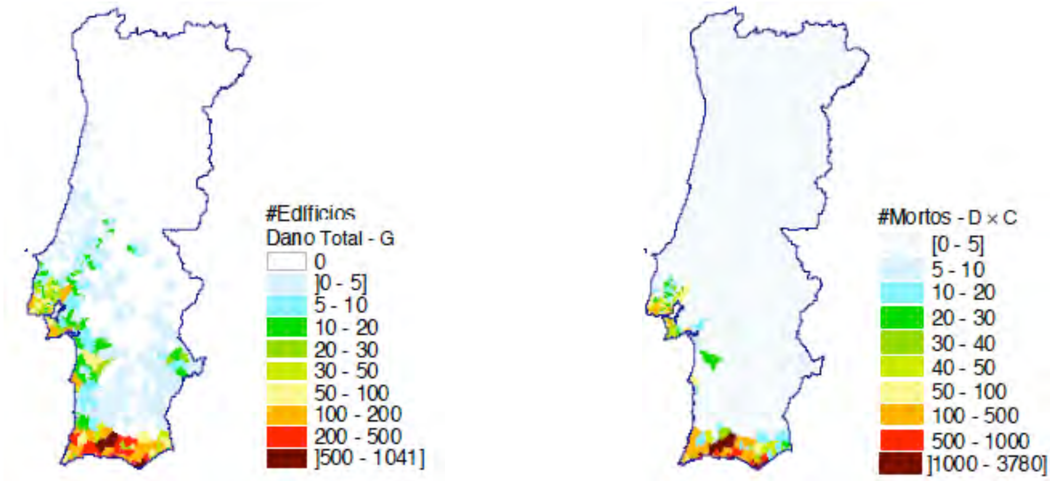


Figure 9.9 Visualization of the results given in a GIS. (a) Number of collapsed building computed with Giovinazzi and Lagomarsino damage model. (b) death computed with Di Pasquale and Orsini damage model followed by Coburn and Spence human loss model [Sousa, 2006].

10 MAEVIZ

10.1 Summary of Software

MAEviz is a tool to support seismic risk reduction. The aim of MAEviz is “to reduce the gap between researchers, practitioners and decision making”. It is a seismic risk assessment software developed by the Mid-America Earthquake (MAE) Center and the National Center for Supercomputing Applications (NCSA).

Moreover, MAEviz has an open-source framework which employs the advanced workflow tools to provide a flexible and modular path. It can run over 50 analyses ranging from direct seismic impact assessment to the modelling of socio-economic implications. It provides 2D and 3D mapped visualizations of source and result data and it provides tables, charts, graphs and printable reports for result data. It is designed to be quickly and easily extensible. When new science, source data or methodologies are discovered, these can be added to the system by developers or end users via a plug-in system. Furthermore, the extensibility allows MAEviz to be used for other future scenarios, such as other hazardous events like hurricanes, fires, landslides, etc.

There are some input files which are hard-coded in the software. However the possibility for the user to upload their own input files does exist. They are stored in *.csv tables and as shapefiles in order to visualize directly the results on maps. The software provides output files stored in tables, charts, report and maps. Moreover, the user can calculate statistics with the results using special tools provided by the program.

10.2 Methodology

MAEviz implements Consequence-Based Risk Management (CRM) using a visual, menu-driven system to generate damage estimates from scientific and engineering principles and data (Figure 10.1). It can also estimate impacts on transportation networks, social or economic systems.

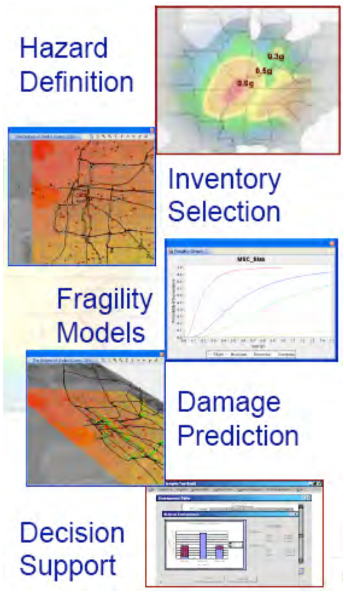


Figure 10.1 Path followed by MAEviz in computing predictions

It requires the following as inputs: hazard, inventory and fragility models. These information are useful to estimate the damage and the losses. It can consider different types of assets such as buildings, bridges and lifeline (gas, water, electric facilities, etc.). With regards to buildings, it estimates structural and non structural damage, economic losses and liquefaction damage. With regards to bridges, it computes damage, loss of functionality and repair cost analysis. For what concerns lifelines it calculates the network damage and the repair rate analysis. Finally, it computes socio-economic losses such as shelter needs, fiscal and business interruption.

10.3 IT Details

MAEviz is an open source software which can be downloaded from the web. It is written in Java Language and using Eclipse RCP. Eclipse RCP is a platform for building and deploying rich client applications. It includes Equinox, a component framework based on the OSGi standard, the ability to deploy native GUI applications to a variety of desktop operating systems, such as Windows, Linux and Mac OSX and an integrated update mechanism for deploying desktop applications from a central server. RCP made it possible to make MAEviz modular and extensible. In fact, MAEviz provides a framework to add new data and algorithms or update existing ones. It also uses NCSA GIS Baseline that is a rich client application. This latter is composed by three main functions: data management (typing, ingestion, access, provenance tracking), visualization (support for 2D and 3D views, zoom, selection, highlighting), analysis execution (support for local multithreaded execution, visual dataflow system in development). The extensions to NCSA GIS and RCP are all provided by plug-ins. The technologies used by MAEviz are all open source software like geotools, vtk (Visualization Toolkit), Jasper Reports (for generating various kind of outputs), JfreeChart and Ktable. MAEviz is also planning/developing a 'cyberinfrastructure/cyberenvironment' that wants to provide universal access to the calculated results through 'grid computing', both for the calculation part and the rendering part (Figure 10.2a). By using as main transport SOAP/HTTP, the concept of repositories, and grid computing, the project is not only a client that makes risk analysis, but an entire platform that enables the user to use data from different places (data sources).

MAEviz can easily integrate spatial information, data and visual information to perform seismic risk and analysis. It supports *.shp file that is really useful to visualize immediately input and output data. For what concerns output data it provides different formats of results such as *.dbf table, .shp file and *pdf file. The level of resolution depends on the size of the geounit that can be defined by the user. The geographical unit could be a regional unit, a provincial unit, a local unit and it could be an arbitrary polygon (regular or irregular). Finally, MAEviz has a user-friendly GUI that guides the user throught the analyses (Figure 10.2b). In the catalog box, in the lower left corner, the databases are stored and the user can utilize this box to select default data or their own data that are saved herein. In the scenario box, in the upper left corner, the analyses that have been run can be selected and can be visualized on the right box of the GUI.

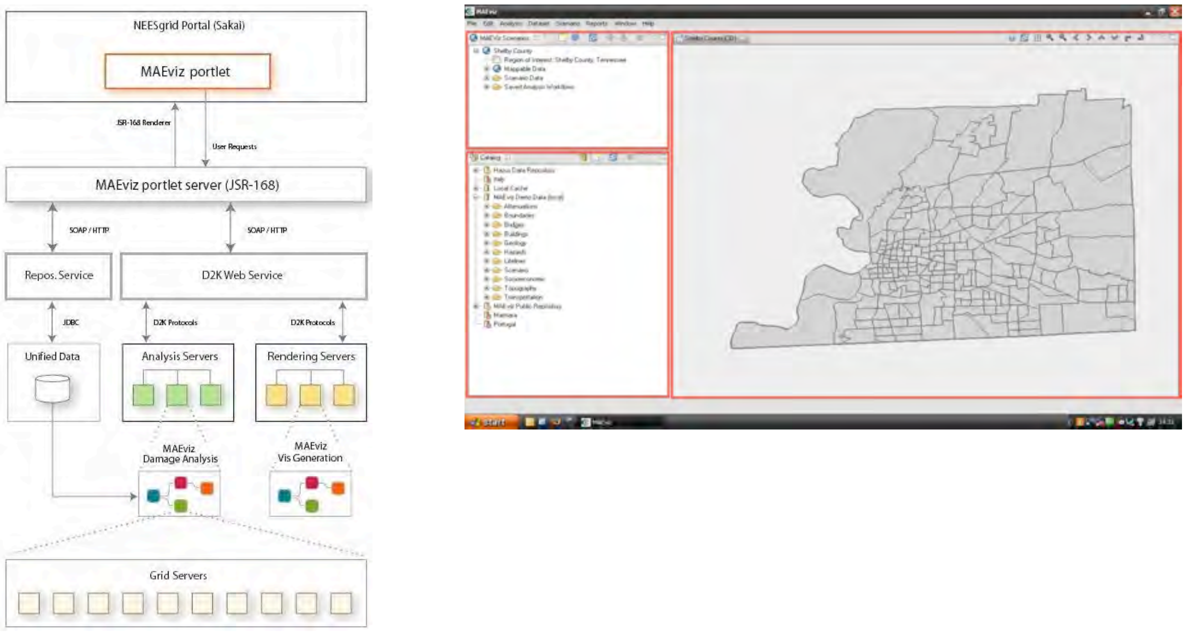


Figure 10.2 (a) MAEviz conceptual architecture (b) MAEviz GUI

10.4 Exposure Module

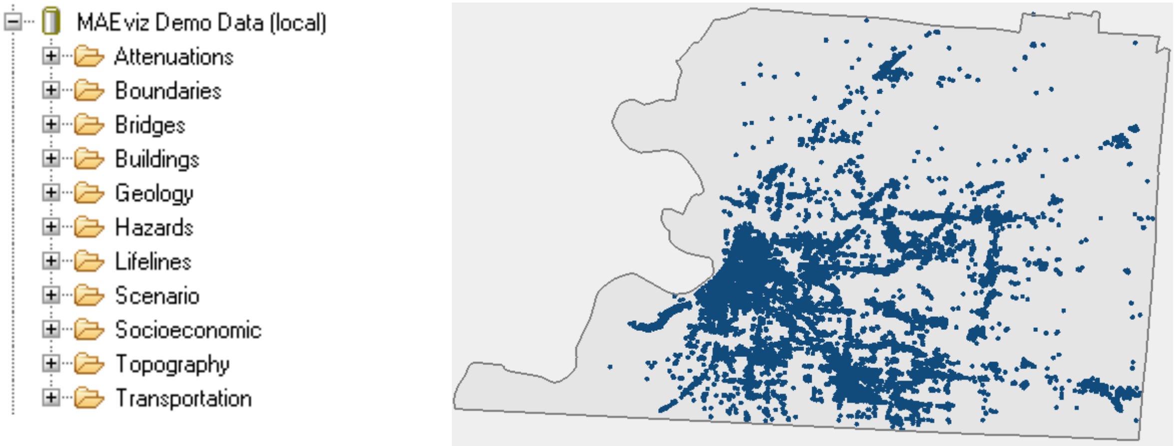
The MAEviz tool provides as default some inventories stored in tables and shapefiles. However the user can upload their own inventory in the 'catalog box' of the GUI (Figure 10.2b). The inventory has to contain, for what concerns buildings, information about the construction type, number of storeys, occupancy level, year of construction and building area. Moreover, the user can upload data about bridges and lifelines as the software consider different type of assets. The level of detail is user-defined and the unit could be a regional unit, a provincial unit, a local unit and it could be an arbitrary polygon (regular or irregular).

The default taxonomy is the one reported in Table 10.1. Buildings have to be classified into one of the 11 structure types reported in the table, however the user can upload a dataset following a different taxonomy as long as he/she has information on the necessary data required to compute the damage (capacity, fragility, etc.).

Table 10.1 Structural Type Categories

General Structure Type	Code
Concrete Moment: Resisting Frame	C1
Concrete Frame with Concrete Shear Wall	C2
Mobile Homes	MH
Concrete Tilt-up	PC1
Precast Concrete Frame	PC2
Reinforced Masonry	RM
Steel Frame	S1
Light Metal Frame	S3
Unreinforced Masonry	URM
Light Wood Frame	W1
Commercial Wood Frame	W2

The inventory has to be provided as a shapefile and it has to be uploaded in the catalog box' (Figure 10.2b) following a simple procedure. The user is always supported by a graphical interface. This way, all the inventories are available in the catalog box as shown in the Figure 10.3a. The advantage of having shapefiles stored in the software is that a direct visualization is possible (Figure 10.3b).

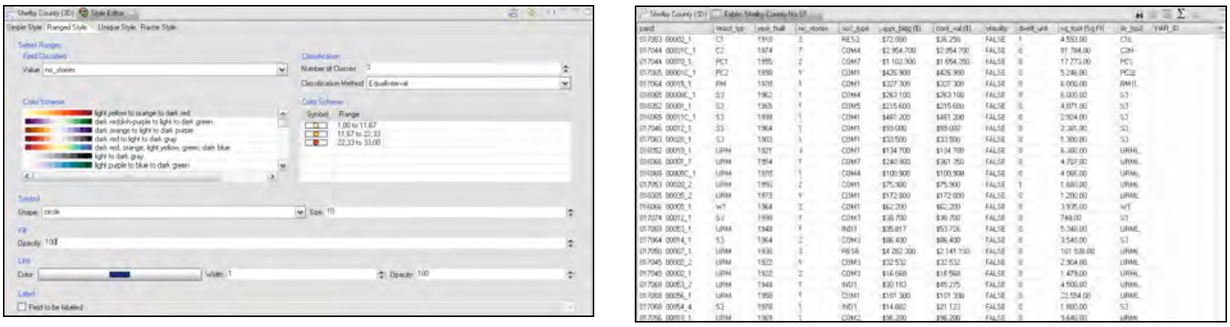


(a)

(b)

Figure 10.3 (a) Catalog box where data are uploaded. (b) An example of the shapefile of the building inventory in the Shelby County, Tennesse.

The user can change the layer style of the inventory (colour, size, symbol shape, etc.). He/she can decide to visualize the buildings using different colours and range for the number of floors, the building area, etc. (Figure 10.4a). He/she can also visualize the inventory table (Figure 10.4b) or create graphs and statistics with regards to the data.



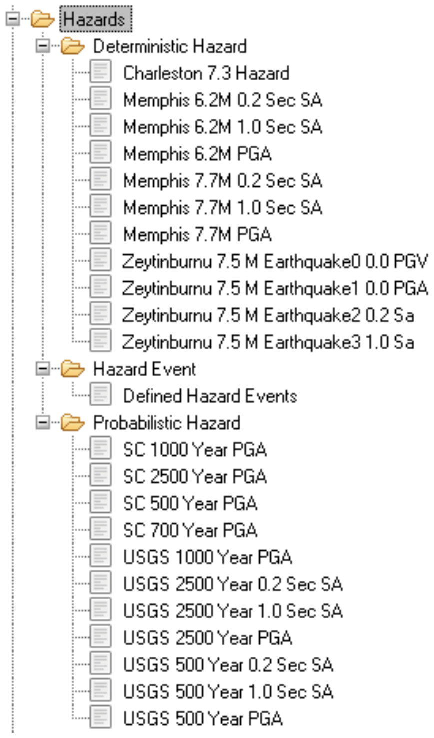
(a)

(b)

Figure 10.4 (a) Window to change Style editor (b) An example of the inventory table for the Shelby County, Tennesse.

10.5 Hazard Module

MAEviz can take into account liquefaction hazard in addition to ground shaking. The hazard is response spectral based and local site effects are taken into account. There are some default scenarios and probabilistic hazard maps in the catalog box (Figure 10.5a), however the user can upload their own hazard following a graphical interface. In Figure 10.5b an example of the ground-motion field the user can upload is shown. It is geo-referenced in order to visualize results on maps.

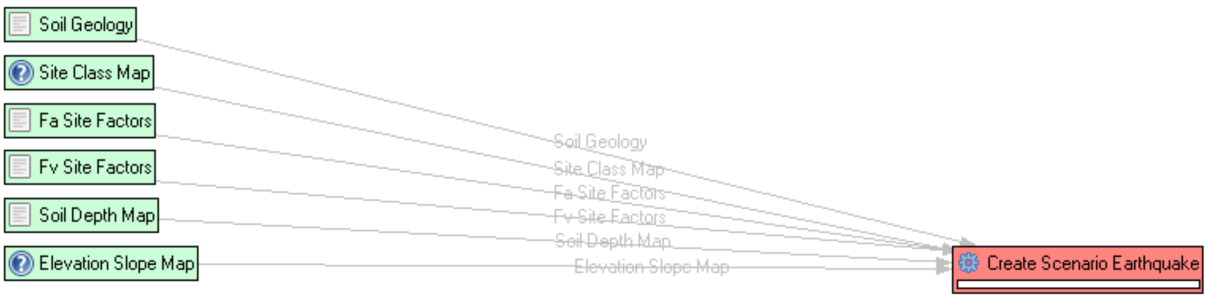



```
NCOLS          37
NROWS          25
XLLCORNER      -90.11501
YLLCORNER      35.00500000000002
CELLSIZE       0.010009765999999875
NODATA_VALUE   -9999.0
0.23199999332427979 0.2310000061988306 0.2310000061988306
0.23399999737739563 0.23499999940395355 0.2310000061988306
0.2329999953508377 0.23399999737739563 0.2310000061988306
0.2310000061988306 0.23199999332427979 0.2290000021457672
0.2290000021457672 0.23000000417232513 0.22699999809265137
0.2240000069141388 0.2290000021457672 0.2290000021457672
0.21899999678134918 0.22699999809265137 0.2310000061988306
0.2160000056028366 0.22300000488758087 0.22699999809265137
0.21400000154972076 0.2199999988079071 0.2240000069141388
0.21199999749660492 0.21699999272823334 0.2199999988079071
0.210999995470047 0.2150000035762787 0.21299999952316284
0.20900000631809235 0.21299999952316284 0.2070000022649765
0.20800000429153442 0.21299999952316284 0.20800000429153442
0.20800000429153442 0.21400000154972076 0.20800000429153442
0.20999999344348907 0.2150000035762787 0.20800000429153442
0.21299999952316284 0.21400000154972076 0.20800000429153442
0.21199999749660492 0.2070000022649765 0.2060000023841858
0.20499999821186066 0.19699999690055847 0.19900000095367432
0.19200000166893005 0.18700000643730164 0.18899999558925630
0.1809999942779541 0.1809999942779541 0.1850000023841858
0.1809999942779541 0.1809999942779541 0.1850000023841858
0.1809999942779541 0.18199999630451202 0.18199999630451202
0.17800000309944153 0.18000000715255737 0.1809999942779541
0.17399999499320984 0.17900000512599945 0.18400000035762787
0.17399999499320984 0.17800000309944153 0.18400000035762787
```

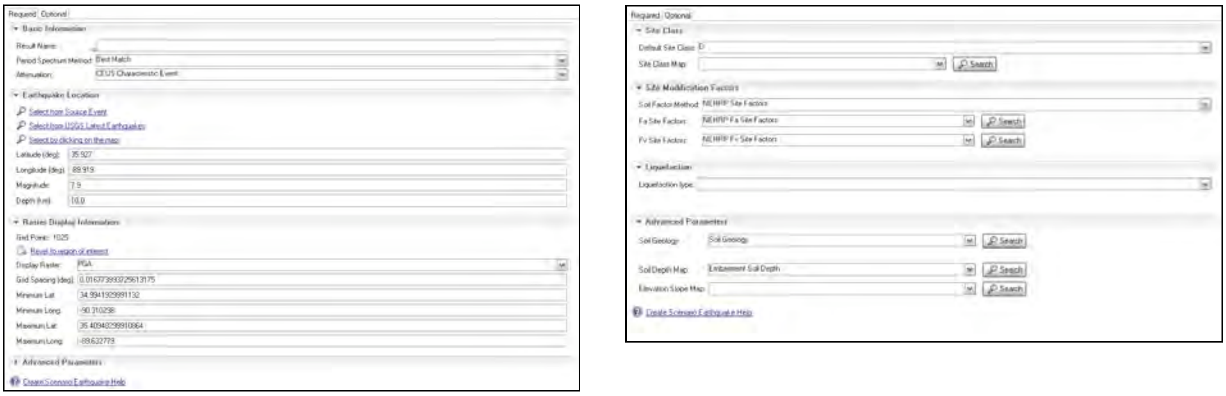
(a) (b)

Figure 10.5 (a) Default Hazard (b) An extract of the ground-motion field

The user can also choose to create the scenario using an analysis. Analyses in MAEviz consist of any sort of calculation that generates data. In Figure 10.6a the hazard process browser is shown. When the 'Create Scenario Earthquake' tool is red, some data are missing and the user has to fill the boxes shown in Figure 10.6a and Figure 10.6b. The user has to provide the attenuation relationships necessary to compute the scenario, the spectrum type, the earthquake location, the coordinates of the region of interest and some advanced parameters such as the fault type, the dip angle, etc. Most of these data are set by default, however the use can modify data adding their own parameters. For what concerns optional parameters, the user can also provide site class, site modification factors, liquefaction and geology to improve his/her hazard. When all the data are provided, the 'Create Scenario Earthquake' tool becomes green, and the user can execute the hazard. As stated above, MAEviz can also compute liquefaction hazard. Based on the hazard chosen by the user, there is a tool called 'Generate LPI Map' useful to create a liquefaction hazard map.



(a)



(b) (c)

Figure 10.6 Hazard tool analysis. (a) structure of the analysis (b) required data to compute the hazard (c) optional data to compute the hazard.

10.6 Vulnerability Module

A fragility dataset is hard-coded in MAEviz. Building fragilities have been developed by the MAE Center for construction typical of the Mid-America region. These vulnerability functions were derived by conducting structural analyses that accounted for aleatory uncertainty related to both structural features and excitation uncertainty. Fragility provides the conditional probability of being in, or exceeding a particular damage state given by the seismic demand parameter. The fragility curves are lognormal distributed and the cumulative lognormal distribution which identify the fragility is reported in the following equation:

$$P(LS|S_a) = \Phi\left(\frac{\ln(S_a) - \lambda}{\beta}\right) \tag{10.1}$$

Where:

- LS is the given limit state;
- S_a is the seismic demand parameter which is determined from hazard analysis. It is in g;
- λ is the median value of the fragility curve. It is in g;
- β is the standard deviation value of the fragility curve.

λ and β are determined for each model building type (Table 10.1).

MAEviz provides alternative fragility formulations and these formula take the form of fragility surfaces rather than fragility curves as the one given by Equation (10.1). These surfaces are representative of analyses conducted for a range of building heights.

$$P(LS|S_a) = \Phi\left(\frac{\ln(S_a) - (\alpha_{11} + \alpha_{12}T)}{(\alpha_{13} + \alpha_{14}T)}\right) \text{ when } 0.87 \le T \text{ (sec)} \tag{10.2}$$

$$P(LS|S_a) = \Phi\left(\frac{\ln(S_a) - (\alpha_{11} + \alpha_{12}0.87)}{(\alpha_{13} + \alpha_{14}0.87)} + (0.87 - T)\frac{\ln(S_a) - \alpha_{21}}{\alpha_{22}}\right) \text{ when } 0 < T < 0.87 \text{ (sec)} \tag{10.3}$$

Where α terms define the fragility surface and they are supplied in the fragility dataset used during analysis.

For each model building type, the user has to map fragility to identify the ID code and then go to the fragility database and localize the necessary parameters for each structure. Three fragility curves are provided. This is the reason why four damage states are taken into account: insignificant, moderate, heavy and complete damage. The MAEviz damage states of moderate, heavy and complete damage map approximately to HAZUS moderate, extensive and complete. The MAEviz Insignificant damage state is approximately equivalent to the combination of HAZUS None and Slight damage states.

The discrete probabilities of damage states are obtained by taking the difference between adjacent curves. In Figure 10.7, the fragility curves obtained used Equation (10.1) for a mid-rise steel braced frame are reported.

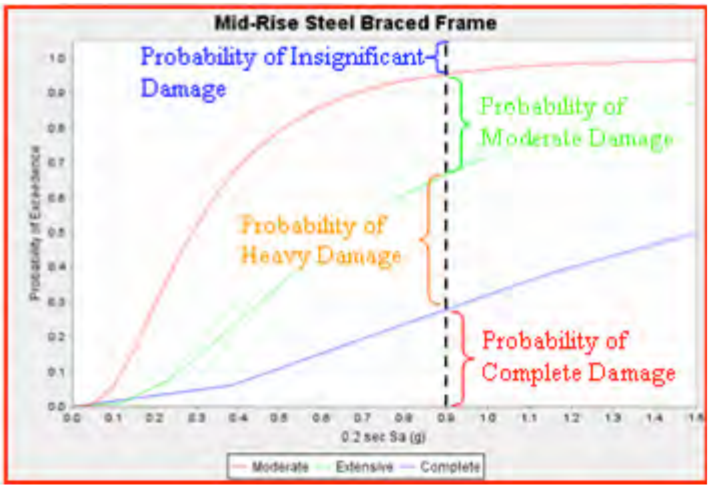


Figure 10.7 Fragility curves for mid-rise steel braced frame

HAZUS fragilities are also supplied in the Fragility dataset of MAEviz. Notwithstanding that, the user should be aware that the hazard used when evaluating these fragilities is obtained by performing a transformation from elastic spectral acceleration to elastic spectral displacement without regard for inelasticity in building response, although the HAZUS fragility parameters are calibrated for inelastic response parameters.

Even if there are some fragility curves hard-coded in the software, the user can upload their own fragility curve dataset following a graphical interface procedure. However, the user should be always aware that the hazard is as explained above. These data have to be provided as *.xml files like the one shown in the following figure.


```
- <fragility-dataset>
  - <fragility-dataset-sets>
    - <fragility-set>
        <fragility-set-properties Parameters="6" Description="Mid-Rise Concrete Frame w/ Shear Walls" DamageType="Structural" Author="Elnashai and Jeong"
          LimitStates="Moderate: Extensive: Complete" TEqnType="1" DemandType="PGA" Code="High-Code" TEqnParam0="0.56" DemandUnits="g" TEqnParam2="0"
          TEqnParam1="0" EquationType="1" GroundMotions="50th percentile: Memphis Uplands: 6.5 @ Marked Tree AR" StructureType="Concrete Frame w/ Shear
          Walls" Stories="5" ID="SF_C2_383" />
      - <fragility-set-labels>
          <fragility-set-label>Moderate</fragility-set-label>
          <fragility-set-label>Extensive</fragility-set-label>
          <fragility-set-label>Complete</fragility-set-label>
        </fragility-set-labels>
      - <fragility-set-fragilities>
        - <fragility-curve fragility-curve-median="-0.718" fragility-curve-beta="0.851" fragility-curve-type="logNormal" curve-
            type="ncsa.maeviz.fragilitycurves.PeriodStandardFragilityCurve">
            <PeriodEqnType>1</PeriodEqnType>
            <PeriodParam0>0.56</PeriodParam0>
            <PeriodParam1>0.0</PeriodParam1>
            <PeriodParam2>0.0</PeriodParam2>
          </fragility-curve>
```


Figure 10.8 Extract of the fragility database provided by MAEviz

In the fragility dataset, the required parameters to estimate the natural period T_1 of a building to which a fragility set has to be assigned are reported. They are represented by 'TEqnType' and TEqnParamX' where X can be 0, 1 or 2. In the following table, formula to estimate T_1 are shown.

Table 10.2 Equations to estimate the natural period of a building

When TEqnType = 1	T_1=a	Where a = TEqnParam0
When TEqnType = 2	T_1=a* Stories	Where a = TEqnParam0

When TEqnType = 3	T_1=b*(a* Stories)^c	Where a = TEqnParam0 b = TEqnParam1 c = TEqnParam2

The default damage method is to find the hazard at the period that is closest to the one required by the fragility for each building (T_1). If the hazard input for the fragility of a building is spectral acceleration (Sa) and the user selects NEHRP or HAZUS damage method, the spectrum method will be used to find an interpolated hazard value based on the building's period.

Once probabilities of damage states have been determined, a mean damage factor has to be calculated. The mean damage ratio is the fraction of building value expected to be lost for the selected scenario. The final structural damage results are given in terms of the mean damage ratio - Equation (10.4). To calculate the mean damage ratio and its variance, MAEviz uses the following equations:

$$\mu_D = \sum_{i=1}^{4}\left[P(DS_i)\mu_{D|Ds_i}\right] \tag{10.4}$$

$$\sigma^2{}_D = \sum_{i=1}^{4}\left[P(DS_i)(\sigma^2{}_{D|DS_i} + \mu^2{}_{D|DS_i})\right] - \mu^2{}_D \tag{10.5}$$

Where the mean of damage factors ($\mu_{D|DSi}$) and the standard deviation of damage factors ($\sigma_{D|DSi}$) are values reported in the Damage Ratio dataset which is hard-coded in the catalog box of the software. However the user can upload in the catalog box a *.csv table using their own ratios.

MAEviz can compute more than 50 analyses, ranging from direct seismic impact assessment to the socio-economic implications. These analyses concern buildings, bridges, lifeline, hazard, decision support and socioeconomic analyses.

For what concerns the building structural damage, the user has to provide the inventory, the hazard, the fragility dataset, the damage ratio dataset and the method to compute the damage. These are the required information, though the user can provide some additional information to improve the result. For instance, he/she can supply data concerning liquefaction. These data have to be provided by a scroll down menu using a graphical interface similar to the one reported in Figure 10.6 which is for the hazard.

With regards to the non-structural and contents losses, data and algorithms used in MAEviz have been adapted from the HAZUS-MH. For consistency with building damage states, only four non-structural damage states are employed. The probability of exceedance is determined also in this case by evaluating a cumulative lognormal distribution. The hazard parameter could be Sa or Sd (spectral displacement) depending on acceleration-sensitive or drift-sensitive components. The HAZUS fragilities are calibrated for hazard parameters which are expected to be obtained from the Capacity Spectrum Method (considering inelastic behavior), but CSM is not currently implemented in MAEviz. Consequently, the user has to be aware that linear elastic hazards are used when evaluating nonstructural and contents damage. Once probabilities of damage states have been determined, a mean damage factor has to be calculated in the same way it is computed for the structural damage. In this case, the Damage Factor dataset has to contain data concerning the acceleration-sensitive non structural damage factor, the drift-sensitive non structural damage factor and the contents damage factor. To summarize, the user has to provide the hazard, the inventory, the fragility and the damage factors to compute this kind of damage. As an optional element, the user can choose to consider liquefaction.

MAEviz computes the direct economic losses using the results of previously executed analyses for structural and optionally non structural damage. The algorithm implemented in MAEviz to compute economic losses uses the following equations:

$$Loss_i = M_i\left(\alpha_i^{SD}\mu_{D_i^{SD}} + \alpha_i^{NA}\mu_{D_i^{NA}} + \alpha_i^{ND}\mu_{D_i^{ND}} + \alpha_i^{CL}\mu_{D_i^{CL}}\right) \tag{10.6}$$

$$\sigma^2_{Loss} = \sum_{i=1}^{N}\left(\left(M_i^{SD}\right)^2\sigma^2_{D_i^{SD}} + \left(M_i^{NA}\right)^2\sigma^2_{D_i^{NA}} + \left(M_i^{ND}\right)^2\sigma^2_{D_i^{ND}} + \left(M_i^{CL}\right)^2\sigma^2_{D_i^{CL}}\right) \tag{10.7}$$

Where:

- M_i is the total replacement cost of a particular item *i*. This value does not include contents;
- μ are terms obtained from the damage analysis results for each components (structural damage SD, non structural damage acceleration-sensitive NA, non structural damage drift-sensitive ND and contents CL);
- α_i are obtained from Occupancy Damage Multipliers dataset which is hard-coded or could be upload by the user in the catalog box in a *.csv table. These multipliers account for the value associated with the different types of components (structural, acceleration-sensitive non structural, drift-sensitive non structural and contents), as a percentage of the replacement value. Sets of multipliers are applied to inventory items based on specific occupancies.

The user has to provide the structural damage distribution and the occupancy damage multipliers. Optionally, he/she can supply the non structural damage distribution. In the following Figure an example of the economic losses is shown.

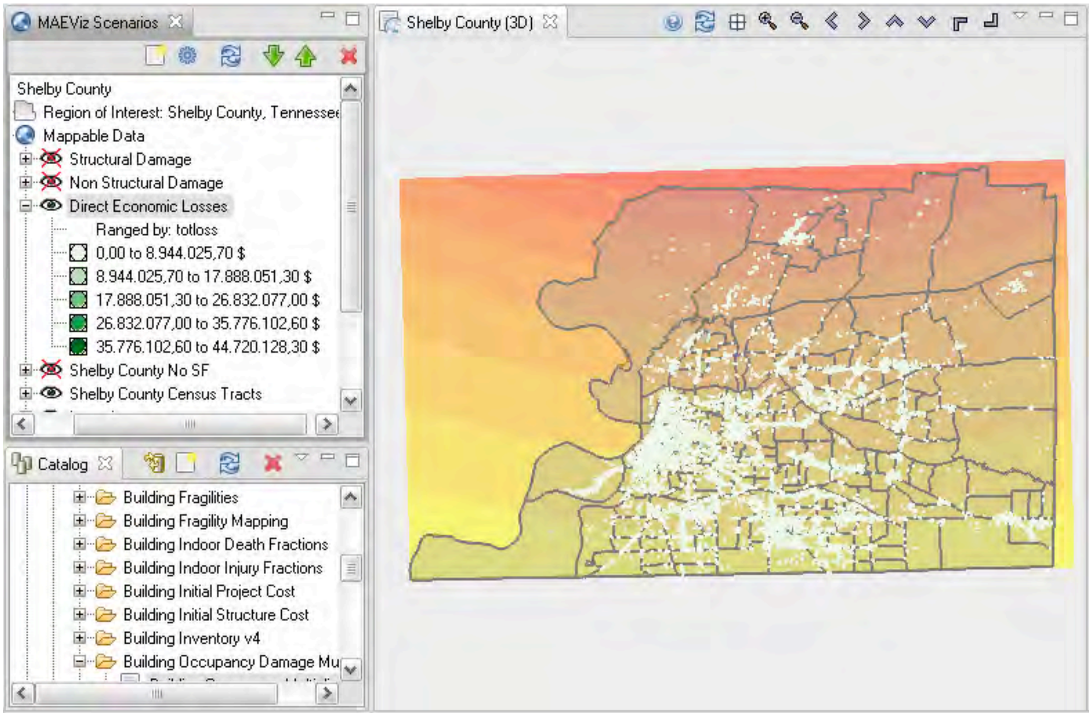


Figure 10.9 Example of economic losses in MAEviz.

There are additional economic analyses implemented in MAEviz. The building cost benefit analysis computes the cost benefit ratio considering the difference in economic losses due to the upgrade. The building repair cost calculates the cost to repair a structure based on the structural damage and building type. Finally, there is the building retrofit cost estimation which computes the cost of all available retrofits for a building.

Moreover there are a number of socioeconomic analyses ranging from casualties to fiscal impact. The social losses analysis computes the number of casualties for each severity level (from I to IV). The user has to supply the building structural damage, the indoor casualty fraction table which contains the casualty fractions for each severity level for each building type and damage state, the building collapse rate which specifies the collapse rate for each structure type and the time of the earthquake (night, day or average). These tables are hard-coded but the user can upload their own table in the catalog box. MAEviz also provides the ability to aggregate data to a geographic boundary, such as the census-tract level. The user has to select the aggregation polygon (for instance, census tract) and the aggregated feature such as structural damage, economic losses or social losses. Finally, he/she has to select the fields to aggregate such as the mean damage, the collapse damage, the moderate damage, etc. In Figure 10.10 an example of aggregate data is shown. It represents the aggregation of the structural mean damage for each census tract. Furthermore, the user can also filter data and create their own selections.

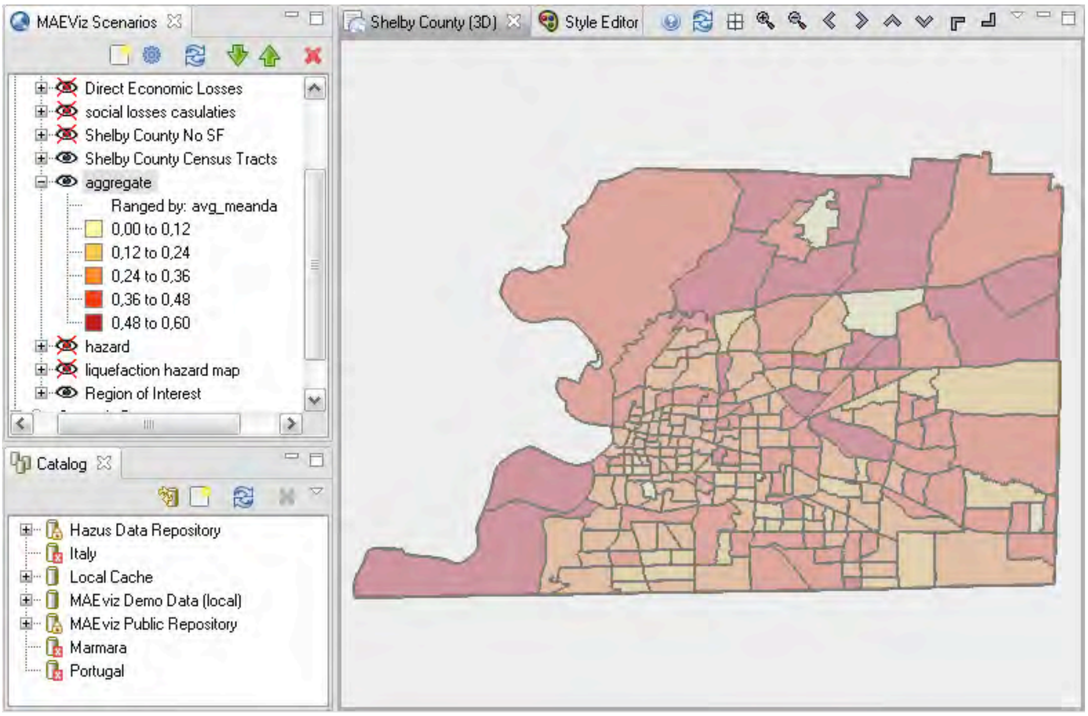


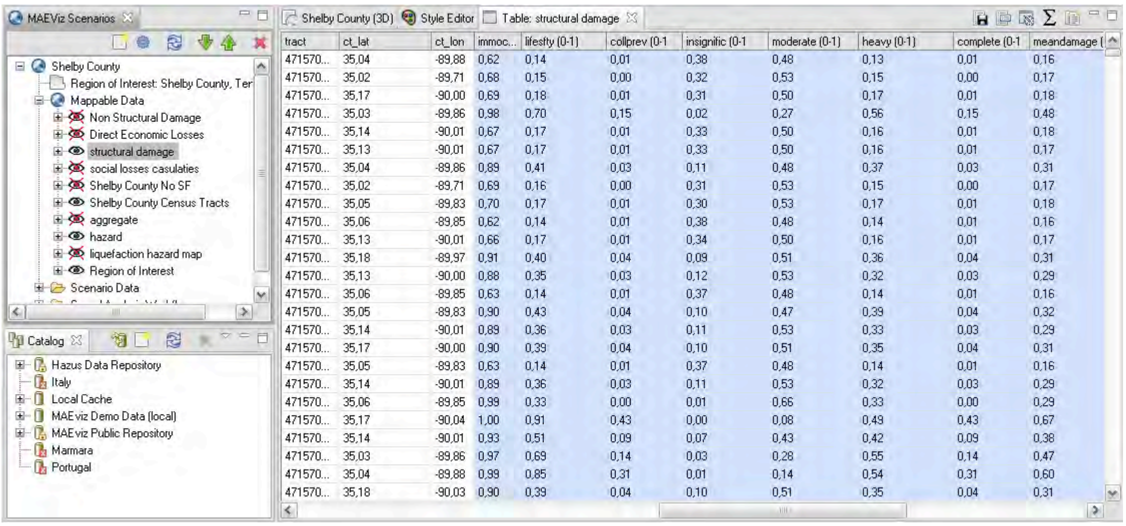
Figure 10.10 Example of aggregated data. Structural mean damage is aggregated for census tract

10.7 Output

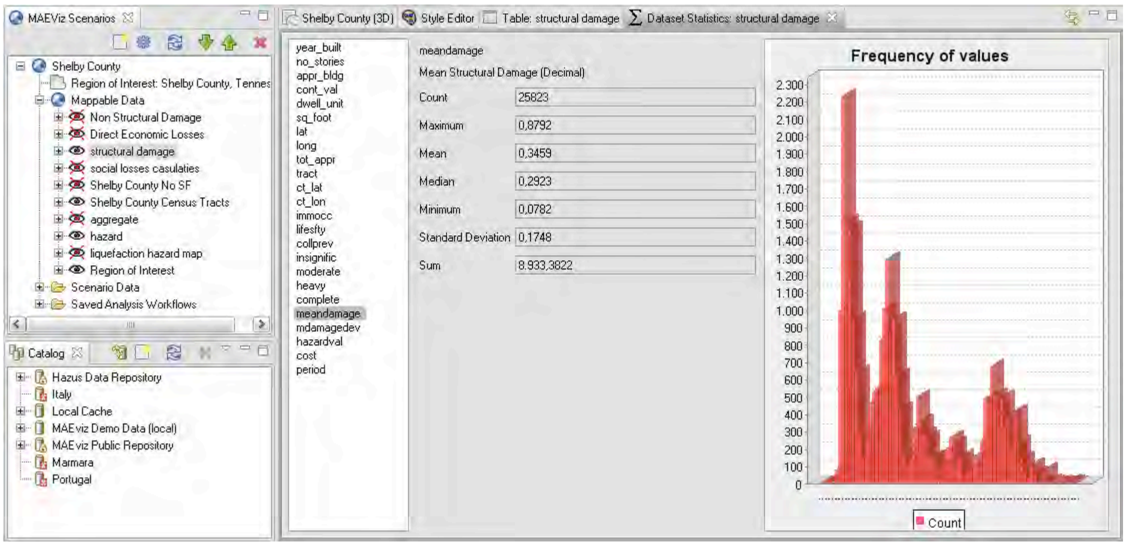
Four damage states are considered in MAEviz: insignificant, moderate, heavy and complete damage. The user can visualize immediately on the browser the structural or non structural damage for these damage states. He/she can choose to visualize the insignificant, the moderate, the heavy, the complete state or the mean damage state. The results are given in a map. The user can plot each analysis in maps such as economic losses, social losses, fiscal impact, etc.

Moreover the user can choose, for each analysis, to display the attributes tables (Figure 10.11a) and he/she can compute statistics (Figure 10.11b) and create plots. The output tables can be exported in excel tables or *.dbf tables.

MAEviz provides also the chart for each of the analyses. It is possible to choose between pie charts, bar charts or line charts (Figure 10.12). Finally, a report can be provided by the software. It is a detailed report or a simple summary report which includes a map image, a chart and a table with the results. This is provided as a *.pdf. In Figure 10.13 the summary report for structural damage is shown.

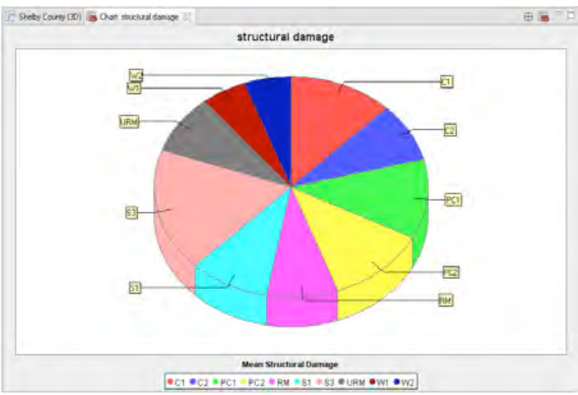


(a)

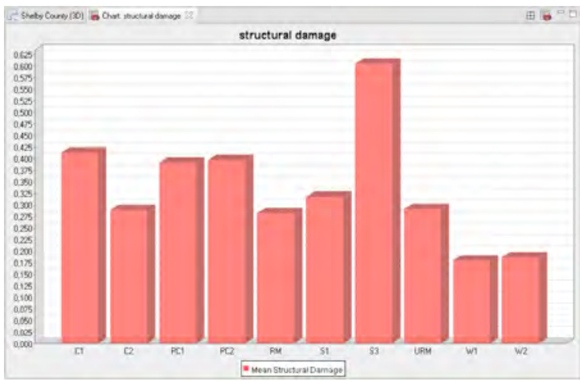


(b)

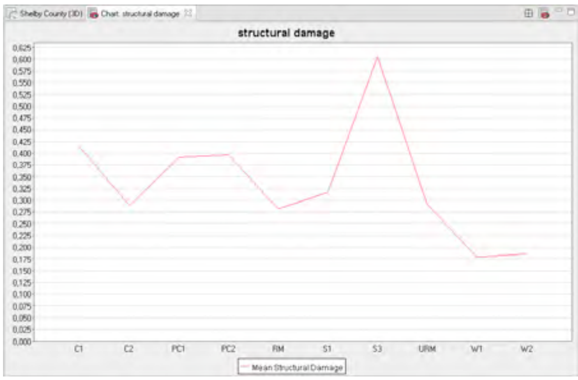
Figure 10.11 (a) Example of attributes table for the structural damage and (b) statistics based on the mean damage



(a)



(b)



(c)

Figure 10.12 Example of the charts which can be computed with MAEviz (a) pie chart, (b) bars chart and (c) line chart

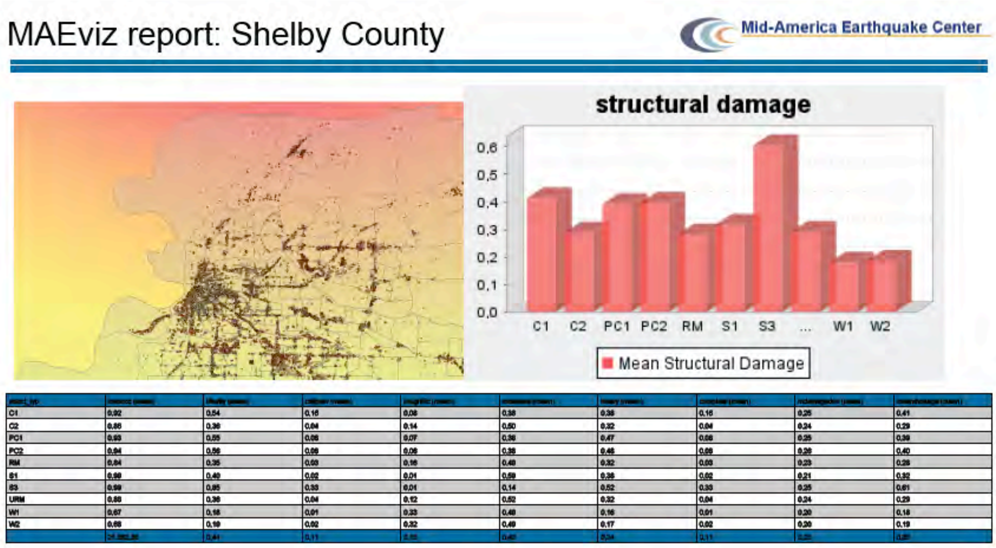


Figure 10.13 Example of a summary report provided by MAEviz for the structural damage

11 OPENRISK

11.1 Summary of Software

OpenRisk is an application for quantifying risk and losses. It is distributed through the Alliance for Global Open Risk Analysis (AGORA), a non-profit virtual organization begun in early 2007 by approximately 35 scholars and professionals in Japan, USA and Europe. AGORA currently has 298 registered members. It offers a web interface where its members can upload and exchange electronic data, tools, documents, etc. OpenRisk is not a *product* of AGORA, but it uses AGORA as a convenient distribution point for reaching potential users and collaborators.

OpenRisk aims to be an object-oriented, web and gui-enabled, open source and freely available software code for conducting multihazard risk analysis. OpenRisk largely represents risk-related extensions of the USGS's and SCEC's OpenSHA seismic hazard analysis software. Developments of the OpenRisk software to date focus on probabilistic seismic risk to individual facilities and portfolios of facilities for shaking damage only, although some initial work on wind has produced a database of US wind hazard. Future developments will expand on portfolio risk algorithms, scenario risk to single buildings and portfolios of buildings, and may deal with 2nd-generation performance-based earthquake engineering analysis of single facilities and other hazards.

For the moment, perhaps for the long term, OpenRisk's seismic components are limited to the spatial extent of OpenSHA, which historically has focused on Southern California, but will soon be extended to North America and the world. Ultimately OpenRisk may treat any arbitrary hazard (hurricane, wind, tornado, etc.) anywhere in the world, as long as the user can supply the hazard and vulnerability data.

OpenRisk currently comprises these stand-alone applications and databases:

1. a single-site BCR (benefit-cost-ratio) calculator
2. a single-site LEC (loss-exceedance-curve) calculator
3. a portfolio EAL (expected annualized loss) calculator
4. a fragility-function calculator
5. a portfolio LEC calculator
6. a US national exposure inventory that provides building count, building area, building replacement cost, and number of occupants, by census tract, structure type, and occupancy type.
7. a database of seismic vulnerability functions for US and similar construction.

GEM1 has only been provided with items 1-4, since the portfolio LEC calculator was not in development at the time the present work began. These items are explained in detail in the following sections. Item 5, which GEM1 has not examined, is briefly summarized in the next section, but no further detail is provided. Items 6 and 7 are largely outside the scope of the present work and are only briefly mentioned hereafter.

11.2 Methodology

The first application, the single-site BCR calculator, computes the expected annualized loss (EAL, with loss measured in terms of repair cost) for a given asset at a given geographic location before and after a retrofit or between as-is and what-if cases. Then it calculates benefit as the difference between the present value of the EAL under the two cases (as-is and what-if), and finally it computes the benefit cost ratio as the ratio of the calculated benefit to the cost of retrofit. The numerical integration employed for calculating EAL is specified in Porter *et al.* [2006].

The single-site LEC calculator creates a loss-exceedance-frequency curve for a single asset at a single location. It does so by numerically integrating the product of the frequency of experiencing a specified level of shaking and the probability of exceeding a specified loss level conditioned on that shaking. The integration is performed over the range of shaking levels to produce the frequency of exceeding the specified loss level. It is repeated for each of many loss levels. The calculation results in a set of pairs—loss level and exceedance frequency—which are then joined in a curve, the loss-exceedance-frequency curve. Loss is measured in terms of repair cost as a fraction of replacement cost.

The portfolio EAL calculator computes the EAL for an arbitrary number of assets and sums them, since the portfolio EAL is the simple sum of site EALs. The methodology for each asset is the same as for the single-site BCR calculator, and much of the same code is re-used for this item.

The fragility function calculator was developed for the Applied Technology Council and it is available for research use. This application calculates fragility functions using a variety of methods documented in Porter *et al.* [2007].

The fifth, not further examined here and currently in alpha testing, using a numerical integration technique called moment matching to calculate the probability distribution of portfolio loss for a specified scenario earthquake. It considers several uncertainties: inter- and intra-event shaking variability; systematic error in seismic vulnerability; asset-to-asset variability in seismic vulnerability, and systematic uncertainty in asset value. (The author acknowledges that other uncertainties exist that are not treated.) The software then calculates the probability of portfolio loss exceeding a specified level as a fraction of total mean portfolio value, multiplies by the occurrence frequency of the specified scenario event, and sums over an exhaustive list of scenario events. No Monte Carlo simulation is used. The moment-matching technique makes the simulation efficient, greatly more so than Monte Carlo simulation. The software uses existing OpenSHA code as the basis for its calculation of seismic hazard.

11.3 IT Details

The software characteristics of OpenRisk consist entirely of Java, which is an object orientated programming language. There are different open source implementations of Java compilers and they are available under the terms of the GNU General Public License. Moreover, Java can be easily run on different operating systems. The OpenRisk source code is available to anyone who contributes new source code under the same terms.

OpenRisk's input files comprise the available ground-motion prediction equations, earthquake rupture forecast, seismic vulnerability functions, and in the case of the portfolio EAL calculator, the portfolio data file. Other inputs include the selection of GMPE, ERF, and in the single-site BCR and LEC calculators, location, structure types, replacement costs, discount rates, site soil class, which are chosen in the OpenRisk's GUI through a drop-down menu. In the case of the fragility function calculator, the specimens, excitations, and damage observations are manually entered (or copied and pasted) into a spreadsheet-like table, and other metadata about the specimens likewise manually entered or copied and pasted into text boxes in the GUI. OpenRisk's output files have been created in ASCII files or they have been directly plotted in a Cartesian plane. The ASCII data can be easily imported in MS-Excel or MS-Access.

Finally, OpenRisk has a GUI system and it does not offer a command line interface (although command line versions are trivial to create given the object-oriented design). In Figure 11.1 some screenshots are shown. It is also possible to create a Log file (Figure 11.2).

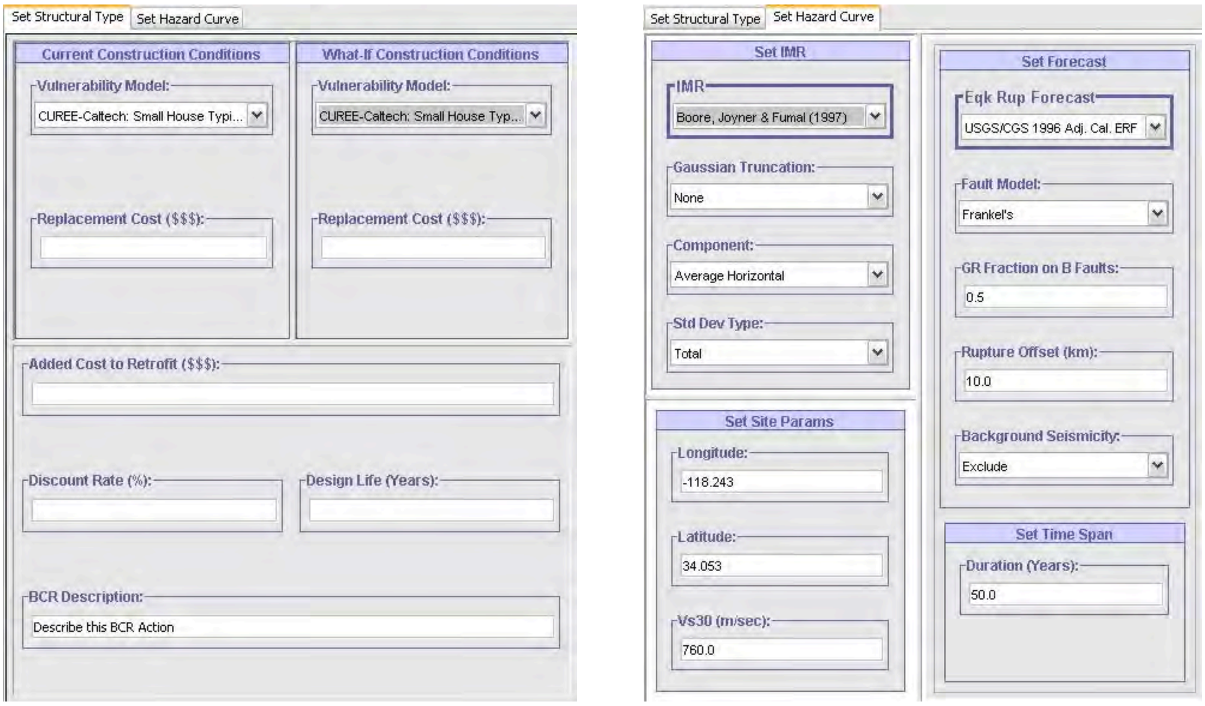


Figure 11.1 Screenshots of (a) BCR calculator - structural type and (b) BCR-calculator - hazard

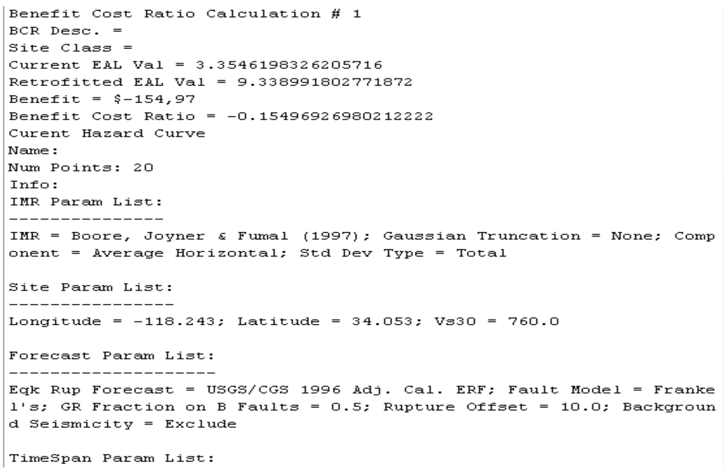

```
Benefit Cost Ratio Calculation # 1
BCR Desc. =
Site Class =
Current EAL Val = 3.3546198326205716
Retrofitted EAL Val = 9.338991802771872
Benefit = $-154,97
Benefit Cost Ratio = -0.15496926980212222
Curent Hazard Curve
Name:
Num Points: 20
Info:
IMR Param List:
---------------
IMR = Boore, Joyner & Fumal (1997); Gaussian Truncation = None; Comp
onent = Average Horizontal; Std Dev Type = Total

Site Param List:
----------------
Longitude = -118.243; Latitude = 34.053; Vs30 = 760.0

Forecast Param List:
--------------------
Eqk Rup Forecast = USGS/CGS 1996 Adj. Cal. ERF; Fault Model = Franke
l's; GR Fraction on B Faults = 0.5; Rupture Offset = 10.0; Backgroun
d Seismicity = Exclude

TimeSpan Param List:
```

Figure 11.2 Log File screenshot of OpenRisk

11.4 Exposure Module

The versions of the OpenRisk software provided to GEM1 in 2009 were limited to 1 set of classes which refers to wood frame houses. This set is sub-divided into eight different model building types with characteristics from Southern California. They are described as Small House Typical, Small House Retrofit, Large House Typical, Large House Waist Wall, Large House Immediate Occupancy, Large House Rigid Diaphram, Townhouse Typical and Townhouse Limited Drift. These are the structure types developed for the CUREE-Caltech Woodframe Project; see Porter *et al.* [2006]. The properties of these building classes are hard-coded and not externally configurable.

The user can choose between one of these buildings by their vulnerability curve using a drop-down menu (Figure 11.3).

The user cannot upload their own inventory with multiple assets and multiple locations. It is possible to analyze just one single model building type by one and one site by one. For this reason, it has not been possible to run an urban or regional test-bed application using this software.

Since 2009 the OpenRisk developers have revised the software in two ways that are relevant here: (1) its allows for the vulnerability functions to be specified in a file, and (2) have added the exhaustive set of US vulnerability functions mentioned earlier, based on the “cracking an open safe” method documented in Porter (2009a, b).

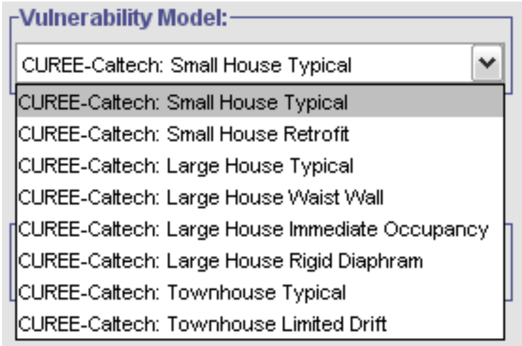
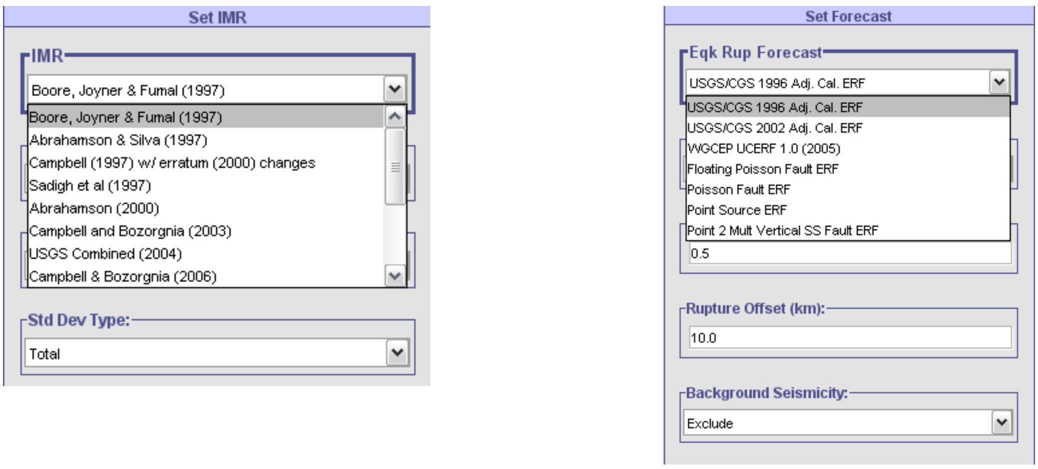


Figure 11.3 Available model building types of wood frame construction

11.5 Hazard Module

The hazard curve is calculated in the BCR calculator (Figure 11.1b). Hazard calculation must produce values of same Intensity Measure Type as in vulnerability functions. The hazard function computed by OpenRisk defines the mean annual rate at which events occur causing intensity equal or greater than an Intensity Measure Level. It has to be considered that the hazard requires an internet connection to run due to the fact that it is linked to OpenSHA. OpenRisk does not have its own hazard module, but is integrated with OpenSHA.

In this application the user has to select the intensity measure relationship (i.e. attenuation relationship). In the 2009 versions of the software, the user can choose between 8 relationships (see Figure 11.4a), the site location (longitude and latitude – V_{s30}) and the earthquake rupture forecast (a model of the seismic sources and their seismicity; there are 7 rupture forecasts, see Figure 11.4b). Using these information the BCR calculator computes the hazard which is given in terms of points. Using this information the BCR calculator computes the hazard which is given in terms of points (Figure 11.5b).



(a) (b)

Figure 11.4 Set hazard curve with BCR applications. (a) Choices of Intensity Measure Relationships (b) Choice of Earthquake Rupture Forecast.

```
Curent Hazard Curve
Name:
Num Points: 20
Info:
IMR Param List:
---------------
IMR = Boore, Joyner & Fumal (1997); G
aussian Truncation = None; Component
= Average Horizontal; Std Dev Type =
Total

Site Param List:
----------------
Longitude = -118.243; Latitude = 34.0
53; Vs30 = 760.0

Forecast Param List:
--------------------
Eqk Rup Forecast = USGS/CGS 1996 Adj.
 Cal. ERF; Fault Model = Frankel's; G
R Fraction on B Faults = 0.5; Rupture
 Offset = 10.0; Background Seismicity
 = Exclude

TimeSpan Param List:
--------------------
Duration = 50.0
```

(a)

```
X, Y Data:
0.1     0.989029
0.2     0.8331241
0.3     0.59727305
0.4     0.413964
0.5     0.28466284
0.6     0.1946346
0.7     0.1327126
0.8     0.0905982
0.9     0.062135052
1.0     0.042915117
1.1     0.029892242
1.2     0.02101212
1.3     0.014907554
1.4     0.010673021
1.5     0.0077081514
1.6     0.005612954
1.7     0.004119011
1.8     0.003044651
1.9     0.002265782
2.0     0.0016968455
```

(b)

Figure 11.5 Output of the BCR calculator with regards to the hazard. (a) The summary of the user choices (b) the hazard curve

For the time being, OpenRisk considers only ground-shaking hazard, but in the future it aims to become a multihazard software. The local site effects are taken into account by the software. The software looks up site classification if not already specified.

The hazard calculation is hard-coded though the user can make a number of choices related to the hazard through a scroll down menu. The output is a text file such as the one shown in Figure 11.5 and it includes a brief description of the choices taken by the user and the hazard curve in term of points.

11.6 Vulnerability Module and Output

OpenRisk is not a standalone seismic risk software like the others presented in this report. As stated before, it is composed of separate standalone applications, each with their own purpose. In this section these three applications will be discussed in further detail.

11.6.1 Single-Site Benefit Cost Ratio calculator (BCR)

OpenRisk provides a BCR calculator that considers repair costs alone. This application needs as input the structural type and the hazard curve (Figure 11.1).

For what concerns the structural type the user has to select a vulnerability model to represent the as-is and what-if conditions. Then, he/she has to enter an estimate of the replacement cost of the asset under each case and an incremental cost (which may or may not be the difference between the what-if and as-is replacement costs). Finally, he/she has to select a discount rate (r) and a design life for purposes of calculating the present value of the reduction in the future losses (t). With regards to the vulnerability the user can choose between 8 CUREE-Caltech wood frame construction vulnerability curves (2 for Small House, 4 for Large House and 2 for Town House). The vulnerability functions implemented inside the calculator are in terms of Mean Damage Factor versus Intensity Measure Type for every typology of buildings. Moreover, there is also a corresponding coefficient of variation for each curve. (As noted above, the software has since been updated to allow for the selection of any of the US structure types, considering lateral force resisting system (16 categories), height category (3 categories), code era (4 categories), occupancy classification (28 categories). Since not all combinations exist in the US (e.g., high code unreinforced masonry bearing wall), there are a total of 3,584 combinations, each with a repair-cost vulnerability function. Although the developer has created vulnerability functions for both repair cost (Porter 2009b) and casualty rate (Porter 2009a), only the repair-cost vulnerability functions have been implemented in OpenRisk.

For what concerns the hazard the user has to select the intensity measure relationship (i.e. attenuation relationship). He/she can choose between 8 relationships, the site location (longitude and latitude – V_{s30}) and the earthquake rupture forecast (a model of the seismic sources and their seismicity; there are 7 rupture forecasts).

This application computes the expected annualized loss (EAL) for a given asset at a given geographic location before and after retrofit or between as-is and what-if cases. EAL is a function of vulnerability, hazard, replacement cost and the remainder term R as it is shown in the following equation:

$$EAL = V\sum_{i=2}^{m}\left(y_{i-1}G_{i-1}(1-\exp(g_i\Delta s_i)) - \frac{\Delta y_i}{\Delta s_i}G_{i-1}\left(\exp(g_i\Delta s_i)\left(\Delta s_i - \frac{1}{g_i}\right)+\frac{1}{g_i}\right)\right)+R \qquad (11.1)$$

Where:

- V denotes the replacement cost of the facility. V_0 is used for V to calculate as-is EAL, and V_r is used for V to calculate what-if EAL;
- y(s) denotes a mean vulnerability function defined as the mean damage factor repair cost divided by replacement cost for a building or other facility as a function of intensity denoted by s. A value y_i represents the mean damage factor given shaking intensity s_i. It is assumed that the vulnerability function varies linearly between values of s;
- G(s) denotes the mean hazard function defined as the mean annual rate at which events occur causing intensity equal or greater than s. A value G_i represents the mean exceedance frequency of given shaking intensity s_i. It is assumed that the hazard function varies exponentially between values of s and g is the slope in the log-linear domain of the mean hazard function;
- s denotes the intensity;
- R is a remainder term for values of $s>s_m$. For large s_m, R can be neglected.

The values of y, G and s can be stored in vectors of size *m*.

Then this application calculates benefit (B) as difference between the present value of the EAL under the two cases (as-is and what-if) and the benefit-cost ratio (BCR) as the ratio of the calculated benefit to the cost of the retrofit or other what-if alternative.

$$B = (EAL_0 - EAL_r)\frac{(1-e^{-rt})}{r} \qquad (11.2)$$

$$BCR = \frac{B}{C} \qquad (11.3)$$

Where:

EAL_0 is the as-is EAL while EAL_r is the what-if EAL;

r is an after-inflation discount rate;

t is the design life of the facility;

C denotes the construction of a retrofit (for an alteration of an existing facility) or it denotes the marginal cost of an alternative design (for choosing between two designs for new facilities).

The vulnerability and hazard calculation are hard-coded and the user can make all choices through a scroll down menu. The output is a text as the one showed in the following figure and it includes the values of the losses (Figure 11.6), a brief description of the choices taken by the user (Figure 11.5a) and the hazard curve in term of points (Figure 11.5b).

```
Current EAL Val = 233.47479506929682
Retrofitted EAL Val = 238.14429097068276
Benefit = $-120,92
Benefit Cost Ratio = -0.06045984221828769
```

Figure 11.6 Economic Losses computed by the BCR

11.6.2 Single-Site Loss-exceedance-curve calculator (LEC)

OpenRisk provides a LEC calculator to compute damage. This application needs as input the structural type and the hazard curve. The user has to select a vulnerability model and he/she can choose between 8 CUREE-Caltech wood frame construction vulnerability curves (2 for Small House, 4 for Large House and 2 for Town House). Then the user has to set the intensity-measure relationship (i.e. attenuation relationship), the site location (longitude and latitude – V_{s30}) and the earthquake rupture forecast (a model of the seismic sources and their seismicity). The input are more or less the same as the single-site BCR calculator. The vulnerability functions implemented inside the calculator are in terms of Damage Factor for every typology of buildings. Moreover, there is also a corresponding coefficient of variation for each curve and a loss exceedance matrix is hard-coded on the code.

This calculator creates a loss-exceedance frequency curve for a single asset at a single location. In the following figure the loss-frequency calculator is shown.

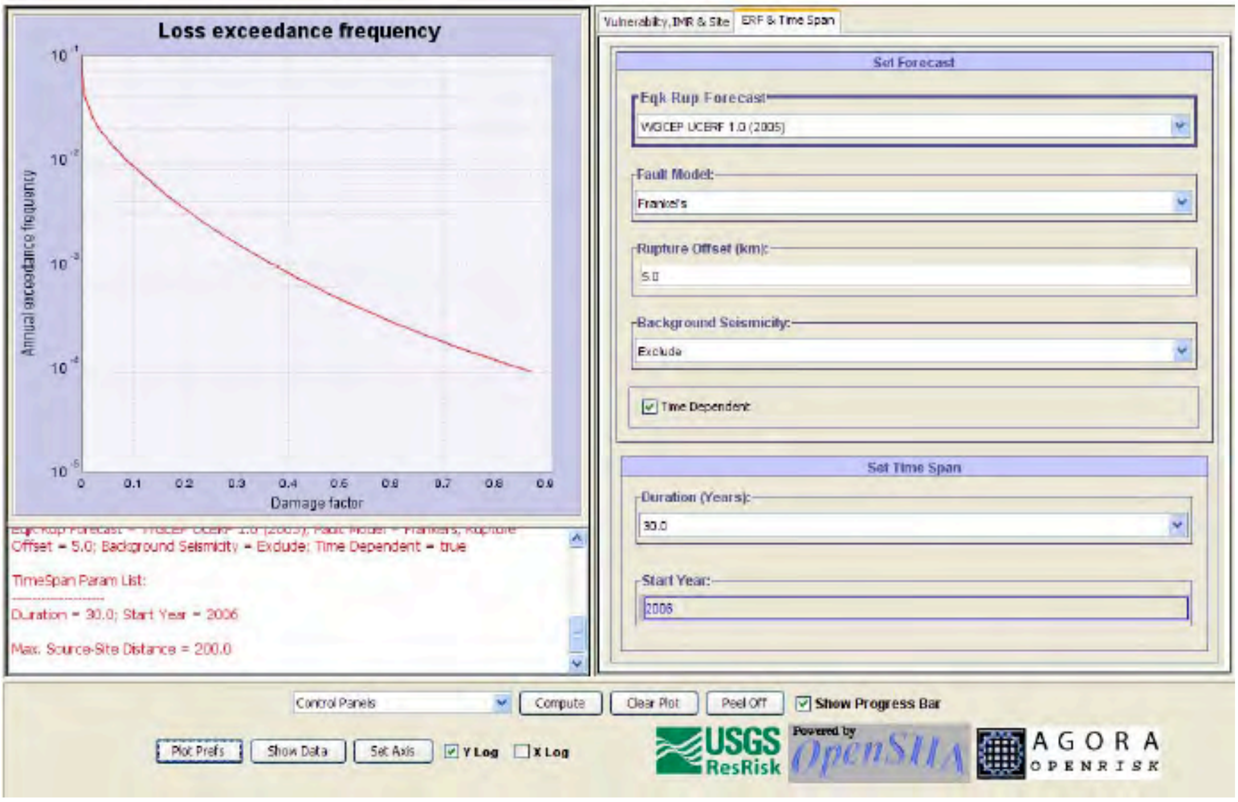
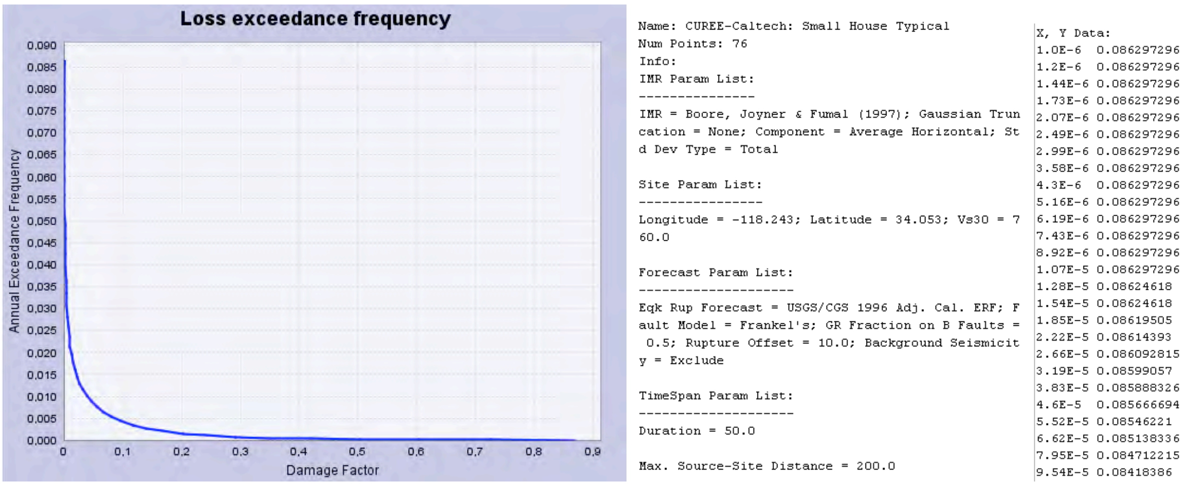


Figure 11.7 LEF calculator

The calculation of the loss exceedance curve is hard-coded and the user can make all choices through a scroll down menu. The user can visualize the plot of the curve (Figure 11.9a) and modify axis, font, colours, etc. or he/she can visualize the data and a brief description of the choices (Figure 11.9b and Figure 11.9c) in a text which can be copied and pasted into an outside documentation.

Figure 11.8 LEC calculator's panel to set the plot



(a) (b) (c)

Figure 11.9 Output of the Loss exceedance frequency calculator. (a) plot, (b) description of the choices taken by the user, (c) an extract of the data of the exceedance curve

11.6.3 Fragility Function Calculator

OpenRisk also provides a fragility-function calculator to determine component fragility. This calculator offers various methods for creating fragility functions from different types of data, depending on the data the user has available [Porter *et al*. 2006 and Porter *et al.*, 2007]. In the following figure a screenshot of this calculator is shown.

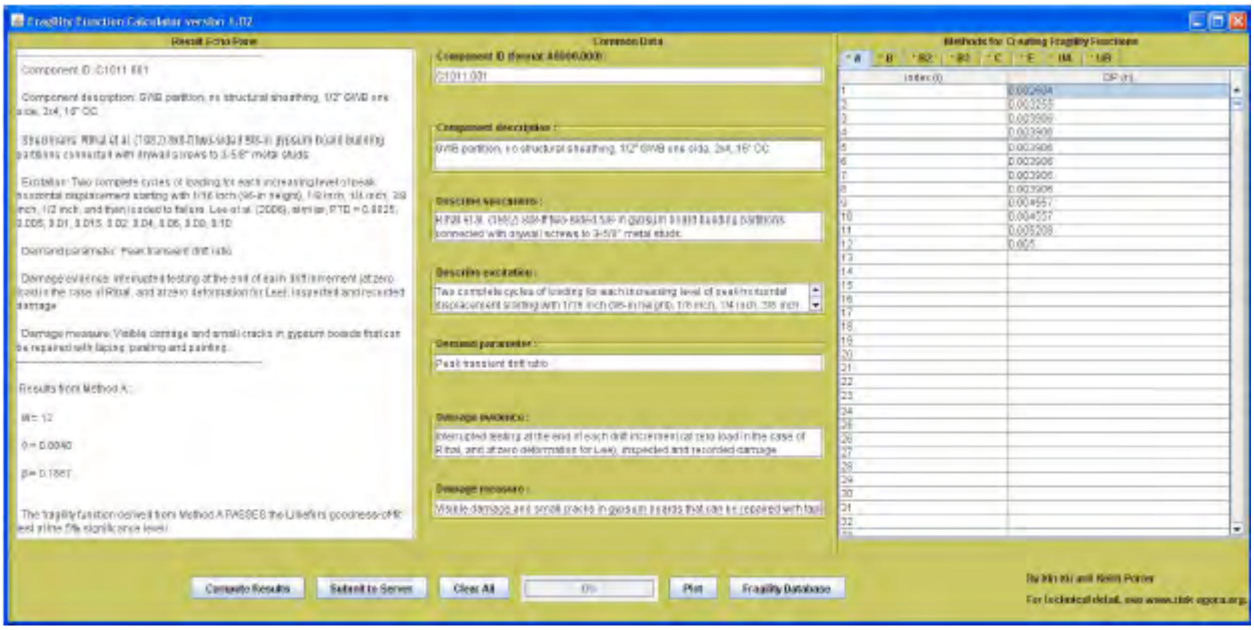


Figure 11.10 Fragility Function calculator

In the left panel there are the data-echo and output. It is a text box and can be selected and copied for pasting into external documentation. The data-echo is the summary of the data description given in the middle panel and the output are the median and the deviation standard of the fragility curve calculated by the program.

In the middle panel the user has to describe data. This text does not affect the calculation:

- Component ID: level-5 component ID. See Porter [2005] for details;
- Component description: brief description of components the tested specimens are deemed to represent;
- Describe specimens: what specimens were observed, where and when;
- Describe excitation: describe if the specimens were tested in a laboratory or in a field, name of accelerograms, directions of excitation, etc.
- Demand parameter: describe the DP type that fragility curve uses as input (peak transient drift, PGA, etc.);
- Damage evidence: describe how the damage state was observed;
- Damage measure: define the damage state observed.

In the right panel the user has to insert the specimens and the corresponding data. Depending on available data the user can choose 8 types of analysis [Porter *et al*., 2007]:

- A: the user has to insert the DP values at which each specimen failed in the column labelled DP. The output is the fragility curve in terms of median and standard deviation. A Lilliefors goodness-of-fit test is also performed. The user can also see the results graphically;
- B: the user has to insert the DP values to which each specimen was subjected in the column labelled DP. If the specimen failed the user has to put 1 in the failure indicator (fi), 0 otherwise. The methods B2 and B3 are slightly different from method B. The output is the fragility curve;
- C: the user has to enter DP to which each specimen was subjected in the column labelled DP. If the specimen showed no damage enter 0 in the failure indicator. If it is experienced damage not suggestive of imminent failure enter 0.1. If it experienced damage suggestive of imminent failure, enter 0.5.The output is the fragility curve;

- E: the user has to enter the results of expert opinion forms in the table, one row per expert. 'Median DP' is the expert's estimate of median capacity. 'Lower DP' is the expert's estimate of the 10^{th} percentile of capacity. 'Level' is the expert's self-judged level of expertise;
- UA-UB: a fragility function already exists and the user has to enter its median and logarithmic standard deviation. Then the user has to insert new specimen data (either as method A or B).

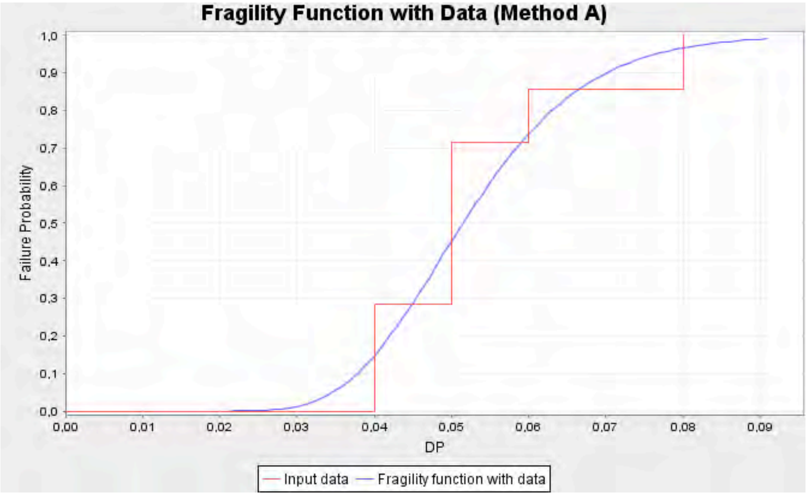


Figure 11.11 Example of fragility function computed with method A

12 Global Risk Assessment

This Chapter describes briefly the methodologies and results of some recent global risk assessment programmes, focusing in particular on seismic risk. The aim is to identify the areas/aspects of such past work where improvements are needed, and thus to ensure that the GEM1 Risk team can begin to provide the necessary advancements.

12.1 UNDP's Disaster Risk Index

The Disaster Risk Index (DRI) of the United Nations Development Programme (UNDP [1]), in partnership with UNEP[2]/DEWA[3]/GRID-Europe[4], is a statistical methodology that has global coverage and national scale resolution. The DRI is applied to earthquakes, tropical cyclones and flooding and in part to volcanoes, landslides and drought. The index is used to compare the impact of hazards in terms of fatalities, as the available databases on fatalities were more complete than those on economic loss and it was felt to be a good parameter for comparison between hazards as one person killed by a cyclone is comparable to one person killed by a flood. The Emergency Disasters Data Base (EM-DAT) from the Centre for Research on the Epidemiology of Disasters (CRED) [5] was used to obtain fatality data, whilst the population data was based on a merging of the databases from the Gridded Population of the World (GPW2) and UNEP models. The study is published in the report Reducing Disaster Risk: A Challenge for Development [UNDP, 2004].

The first stage in the DRI method is to overlay population maps on hazard maps for the different natural hazards and to thus calculate the exposure to each hazard type. For earthquakes, the affected area from a given earthquake was estimated using a radial length from the epicentre as a function of the magnitude, and after which the duration of shaking is assumed to be close to zero. The physical exposure is then estimated by multiplying the exposed population by the frequency of a given hazard event; this gives an estimation of the average population exposed to a hazard event per year. The relative vulnerability is given by the number of fatalities divided by the number of people exposed. Overall a strong correlation between the number of victims and the physical exposure was observed.

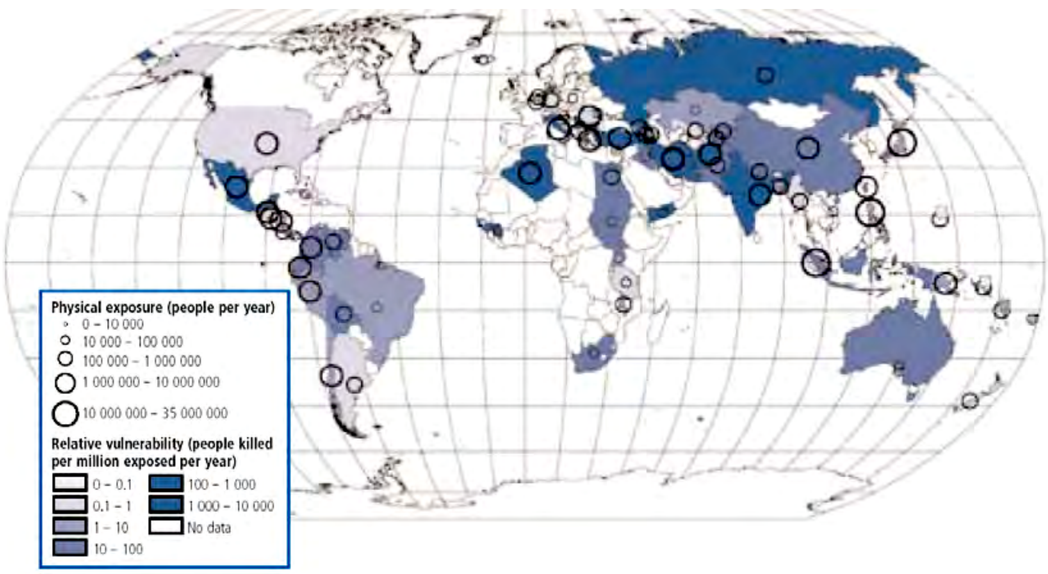


Figure 12.1 Physical Exposure and Relative Vulnerability to earthquakes 1980-2000 [UNDP, 2004]

The relative vulnerability index does not explain why different countries are unequally vulnerable or what role the level of development plays. In order to study this a number of vulnerability indictors (e.g. population growth, urban growth, GDP per inhabitant) were also studied, based on expert opinion. The regression analysis of vulnerability indicators showed that, statistically, physical exposure and the rate of urban growth acted together in being associated with the risk of death from earthquakes. That is, the risk of dying in an earthquake was greater in countries with rapid urban growth.

12.2 Natural Disaster Hotspots: A Global Risk Analysis

The Disaster Hotspots project [Dilley *et al*., 2005] was an initiative led by the World Bank and Colombia University (under the umbrella of the ProVention Consortium [6]). The aim of the project was to identify those places where risks of disaster-related mortality and economic losses are highest on the basis of the exposure of people and GDP to major hazards and on historical loss rates. Hotspots was undertaken at a global level with sub-national scale of resolution based on grid cells.

Three indices of disaster risk have been developed:

- disaster-related mortality risks, assessed for global gridded population;
- risks of total economic losses, assessed for global gridded GDP per unit area; and
- risks of economic losses expressed as a proportion of the GDP per unit area for each cell.

The data for mortality-related risks are assessed on 2.5' x 2.5' latitude-longitude grid of global population (an early version of GPW3 was used). Economic risks are assessed at the same resolution using a gridded surface of Gross Domestic Product (GDP) per unit area following the procedure proposed by Sachs *et al*. [2001].

For earthquakes, the results of the Global Seismic Hazard Program (GSHAP) [7] in terms of peak ground acceleration with a return period of 475 years was used and the Advanced National Seismic System Earthquake Catalogue from 1976 to 2002 was used to obtain estimates of the frequency of earthquakes with a magnitude greater than 4.5 on the Richter Scale.

In order to reduce the computational effort of the analyses, grid cells with population densities less than 5 persons per square kilometre and without significant agriculture were masked out of the analyses. Historical losses from EM-DAT across all events from 1981-2000 for each hazard type were used to obtain mortality and economic loss weights for each of the 7 identified regions (Africa, East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, North America, South Asia) and for the four economic wealth classes (lower, lower middle, upper middle, high) within these regions.

The risk assessment procedure used for both mortality and economic loss is reported in the steps below, using just mortality as an example.

1. The appropriate measure of total global losses from 1981-2000 is extracted from EM-DAT (in the mortality case, the number of fatalities) by hazard, h: Mh .

2. Using the GIS data on the extent of each hazard, the total population estimated to live in the area affected by that hazard in the year 2000 is estimated: Ph.

3. A simple mortality rate for the hazard is then computed: rh = Mh / Ph. Assuming the 1981-2000 period is representative, this rate is an estimate of the proportion of persons killed during a 20-year period in the area exposed to that hazard. Since the numbers are very small, they are expressed per 100,000 persons in 2000.

4. For each GIS grid cell i that falls into a hazard zone h, the location-specific expected mortality is computed by multiplying the global hazard-specific mortality rate by the population in that grid cell: Mhi = rh*Pi.

5. If the various combinations of region and country-wealth status are denoted by j, then the estimated mortality in a given grid cell is now Mhij = rhj * Pi.

6. The global hazard data compiled for the analysis measures the degree of hazard in terms of frequency, for earthquakes. The various degree of hazard measures are used to redistribute the total regional mortality from EM-DAT across the grid

cells in the area of the region exposed to each hazard. For example, if a grid cell was hit four times by a severe earthquake during the 20-year period, the regional mortality rate is multiplied by four to yield an accumulated mortality for that grid cell. More generally, if the degree of hazard measure is denoted by W, and assuming that the weighting scheme is identical across region/wealth-class combinations j, the accumulated mortality in the grid cell is: M'hij = rhj * Whi * Pi.

7. To avoid literal interpretation of the hazard disaster risk hotspot index as the number of persons expected to be killed in a 20-year period and in recognition of the many limitations of the underlying data, the resulting measure was converted into an index from one to ten using a classification of the global distribution of unmasked grid cell values into deciles (as shown in the following figures).

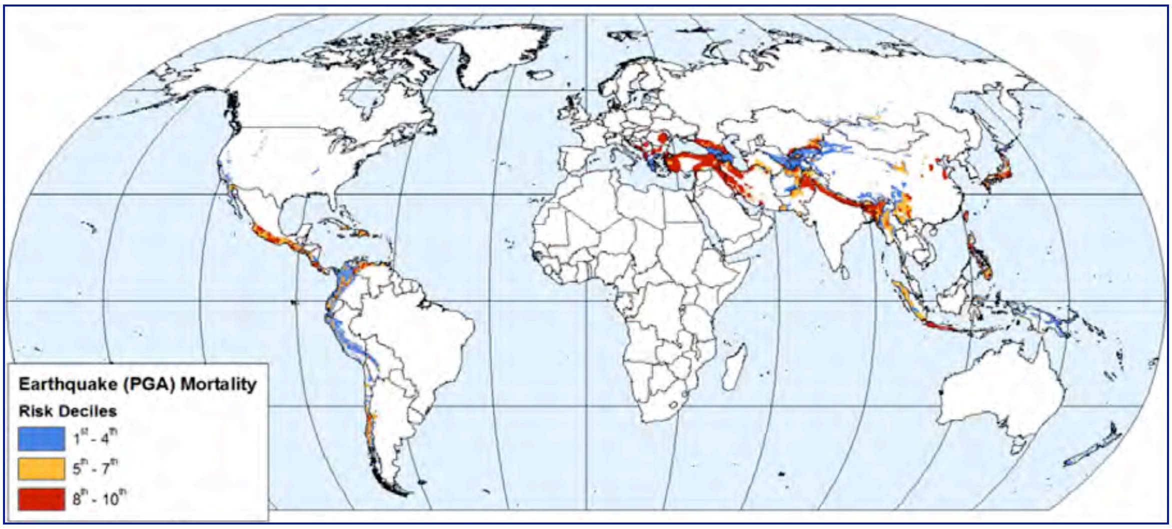


Figure 12.2 Hotspots: Global distribution of earthquake risk (mortality)

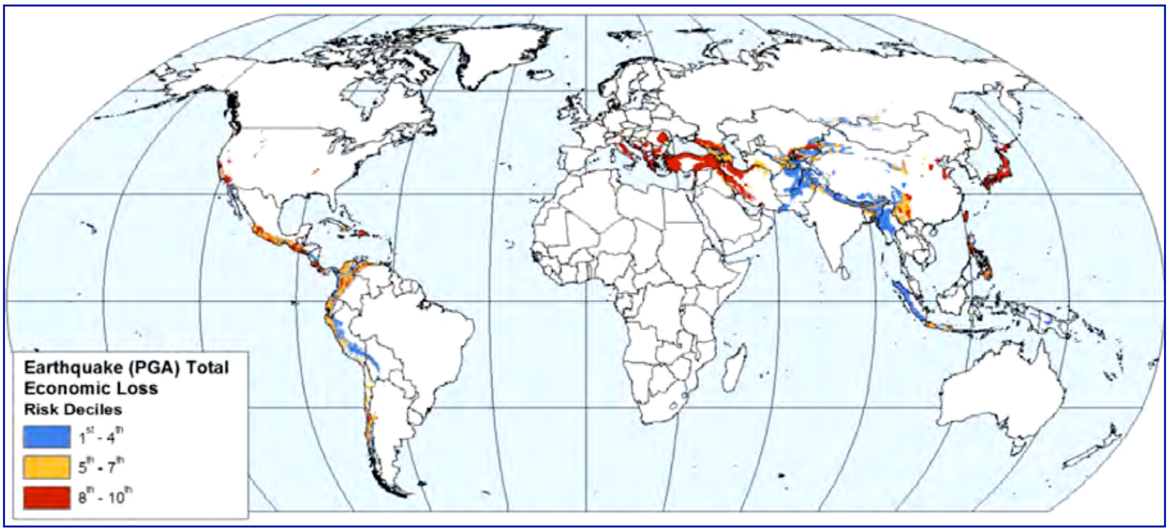


Figure 12.3 Hotspots: Global distribution of earthquake risk (total economic loss)

12.3 Global Assessment Report on Disaster Risk Reduction 2009

The 2009 Global Assessment Report (GAR) on Disaster Risk Reduction [ISDR, 2009] was a collaborative effort of the International Strategy for Disaster Reduction (ISDR) [8]. A global disaster risk analysis was carried out in order to assess global mortality and economic loss risk for natural hazards including tropical cyclones, floods, landslides, earthquakes and tsunamis.

The application of the GAR risk model involved the following steps for earthquake hazard:

1. Compilation of geographical and physical information on past earthquake locations and magnitudes from 1976 to 2002.

2. Determination of the footprint or area of impact (the ShakeMap Atlas V.1 from USGS was used [9], with four groups of hazard level identified by combining intensity levels).

3. For each impact area, computation of exposure as the number of people and economic assets within that area.

4. Linking of available loss information for each hazard event (sourced from EM-DAT) to the hazard event information (hazard severity and exposure).

5. Adding of information on vulnerability. Since global data on direct vulnerability factors such as building quality are unavailable, this analysis uses country-level indicators for the year in which the event occurred, such as government accountability or per capita income.

6. Estimation of empirical loss functions that relate event mortality or economic losses to risk factors (hazard characteristics, exposure and vulnerability) using statistical regression techniques.

7. Derivation of an estimate of expected average annual losses and exposure. Annual frequency grids have been processed for MMI categories by summing the number of times each pixel has been affected by a given category of intensity (according to the ShakeMap Atlas) and dividing the total by the length of the dataset period. The estimated loss functions are used to compute disaster outcomes for all recorded events, whether or not a loss estimate is available in EM-DAT or not. This is done using data on exposure and vulnerability for 2007 such that annualized average estimates reflect current conditions.

8. Application of estimates to all pixels in a geographic grid. The loss estimates are aggregated at different levels (1 km x 1 km cells; sub-national administrative areas; countries) allowing the identification of geographic concentrations of risk. Mortality risk is classed in deciles using a logarithmic index with values ranging from 1 = negligible to 10 = extreme risk. Economic loss risk is calculated for World Bank regions and country income groups.

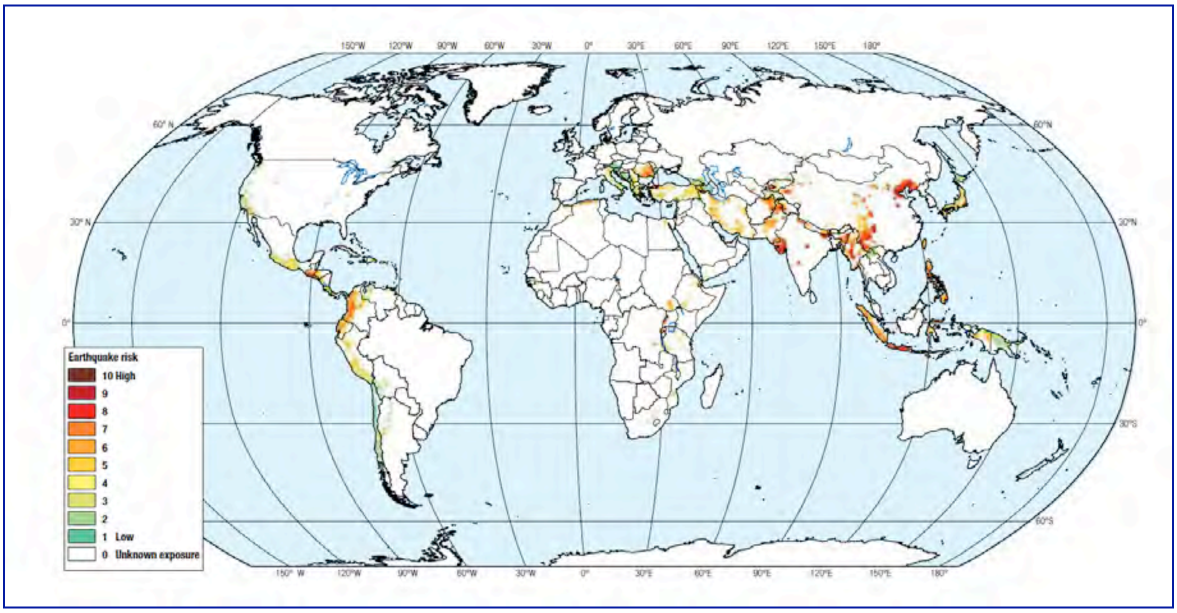


Figure 12.4 Global distribution of mortality risk associated with earthquakes from the GAR report

13 Conclusions

13.1 General Summary of Software

A number of seismic risk codes were reviewed and tested during GEM1 by the Risk team. These codes were studied in detail and test-bed applications were run, mainly to improve further the team's knowledge of the codes. Running the software allowed for the identification of various pros and cons, many of which were discussed with the developers of these codes. The use of test-bed applications which were not provided with the codes was most helpful for the understanding of the potential and limitations of the software. The Risk team was able to compare and thus validate some of the codes under the same conditions of input and methodology, and such results were shared with the developers to help them understand where and why differences were arising. As most of the codes were still under development during this period, this feedback was very well received by the developers. However, the majority of the software run with different approaches (i.e. seismological based or engineering based, empirical or analytical), and different analytical methods (CSM, MADRS, etc.). For this reason most of the codes were not directly comparable and thus attempts at further validation were not carried out.

Some of the limitations highlighted in this report might since have been superseded, so it is worth noting the version of the software that was reviewed herein (see Table 1.2). For example, in QLARM it was not currently possible for a user to upload their own inventory and hazard data. For the time being, the RISKSCAPE hazard tool which allows the user to upload their own hazard, was not working. The OpenRisk applications could only be run with hard-coded vulnerability functions. However, the developers of these software have all assured the Risk team that these limitations are on their planned development list.

Due to time constraints, only deterministic scenario case studies could be carried out during the GEM1 effort. ELER and CAPRA were the only codes capable of running end-to-end probabilistic events-based analyses, whereas only OpenRisk was based on probabilistic hazard curves. The other codes which were based on probabilistic hazard were only able to read and provide losses using a single probabilistic seismic hazard map at a time. Towards the end of GEM1 a new hazard input format for CAPRA composed completely by ASCII files, along with software for preparing it, was shared with the group. Furthermore, due to the fact that the GEM1 risk team decided to focus on just testing the damage calculation of different software (and not the hazard), the EQRM team developed a secondary interface to allow the GEM Risk team to by-pass all of the hazard components of the EQRM (i.e. event based source generation, ground-motion forecasting and site amplification) so that GEM1's own ground-motion fields could be used as the basis for calculations. This means that EQRM was not exactly tested in the manner that it was designed and thus the significant work that has been invested in optimising this style of calculation for large scale portfolio risk assessment has not yet been fully reviewed. The intention of the Risk group to run probabilistic events-based loss analyses using GEM-defined hazard for the Marmara region, would have been possible with CAPRA and EQRM based on the aforementioned updates. Unfortuntely due to time constraints, the group were not able to carry out this task. However, it is recommended by the team that the aforementioned software are fully reviewed and tested in the future when probabilistic events-based analyses need to be generated with the GEM risk engine.

Thanks to the efforts described within this report and the interactions with the software developers, the GEM1 Risk team has learnt a great deal about seismic risk software which has allowed the initial design and coding of a global risk engine to commence, as described in more detail in part 2 of this report. The characteristics of this risk engine have been based on the positive features identified in this review, as summarised further in the next section.

13.2 Optimal Features of Risk Software

All of the software reviewed herein have both positive and not so positive features. The latter have been identified by the GEM1 Risk team and were in many cases already acknowledged by the developers. These included the need for proprietary software to run the code, the lack of detailed documentation for users, the use of extensive hard coding and lack of flexibility for user-defined input, extensive formatting and preparation of input data, lack of user-interface, limitations on the type of analyses and output.

Rather than underlining the drawbacks of each of the software, the GEM1 Risk team has decided to focus on the positive features that have been found within each software and which it is believed should be included in an optimal risk engine.

IT Details

- Professional looking, user-friendly GUI with possibility to view input data (for “sanity checks”) and output data (e.g. MAEviz, RiskScape);
- Open source software which can be downloaded from the web (e.g. SELENA, EQRM, MAEviz), ideally through an open source software developers website (e.g. EQRM);
- No need for proprietary software to develop or run the software (e.g. EQRM, MAEviz, LNECLoss, SELENA, OpenRisk, RiskScape, QLARM);
- Web-based online software to allow users to access the most recent version (e.g. QLARM);
- Possibility to run in different operating systems i.e. platform independent (e.g. EQRM, MAEviz, ELER, OpenRisk, RiskScape, QLARM);
- Tools to transform an input file in the format needed by the software (e.g. ELER);
- Tools to aggregate the results at different levels of resolution (e.g. MAEviz).

Scientific Details - General

- Use of a logic tree approach to account for epistemic uncertainty (e.g. SELENA);
- Applicability to user-defined regions (e.g. SELENA, CEDIM, CAPRA, LNECLoss, MAEviz);
- Updated, clear user manual (e.g. SELENA);
- Full flexibility in terms of user-defined input (e.g. SELENA, LNECLoss).

Scientific Details - Exposure

- Different types of asset within the exposure model such as buildings, bridges, lifelines, etc. (e.g. RiskScape).

Scientific Details - Hazard

- Multi-hazard software including ground-shaking, inundation, storm, volcanic ashfall, etc. (e.g. RiskScape, CAPRA);
- Various ground-shaking parameters such as response spectrum and macroseismic intensity (e.g. CAPRA, ELER, LNECLoss);
- Inclusion of local site effects through a 1D SHAKE-type analysis (e.g. LNECLoss).

Scientific Details - Vulnerability

- Combination of seismological and engineering-based methods (e.g. ELER, LNECLoss);
- Large number of analytical methods to choose between (e.g. ELER);
- Large number of empirical methods to choose between (e.g. LNECLoss);
- Inclusion of both structural and non-structural damage (e.g. MAEviz, EQRM);
- Extensive modelling of economic losses (e.g. MAEviz);
- Extensive modelling of social losses (e.g. MAEviz).

The GEM1 Risk team thus recommends to GEM that the key characteristics of its risk engine, based mainly on the positive features described above, should comprise the following:

Open-source software development

The source code of the risk software should be available to any user and the development of the code should be a product of the efforts of a community, and not just limited to a working group.

The aim would thus be to create an interactive environment in which the scientific community can share and develop its knowledge about seismic risk. An open source tool will allow the seismic risk community to dynamically interact and communicate, and it should facilitate the collaboration between different end-users such as academia, governments and industry. The engine will have to interact with living databases and for this purpose, local and national governments will be called to keep the databases related with their own country up-to-date. The creation of a network of knowledge related to seismic risk could initiate the development of the GEM risk engine, and this will help GEM achieve its goals related to raising awareness of the benefits of seismic risk modelling.

Flexible and Dynamic

As stated above, the GEM engine needs to be developed with the purpose of creating a platform for risk assessment, instead of another static risk application. This platform should allow users to evaluate risk due to seismic hazard, but also allow other types of hazard such as floods or hurricanes, to be considered.

Moreover the engine has to be able to run risk analyses with deterministic events, probabilistic events, and probabilistic seismic hazard maps and curves. The user should be able to upload their own hazard data or hazard data from the GEM hazard engine but it should also be able to carry out risk assessment using hazard data calculated on-the-fly.

The engine should be as flexible as possible to allow different methodologies and different kinds of data to be implemented into the software. For the sake of flexibility, external tables and user-defined objects are strongly encouraged, with tools available to help users prepare their input in the required format. The calculator should allow users to update their results based on newer models, datasets or hazard inputs.

Finally, as some of the existing software offer, a multiple choice of visualization of results should be provided through the engine (e.g., *.dbf, *.shp, *.pdf, *.kml, etc.). The user should be able to graphically visualize the results in real time but also save them for post-processing with their own tools.

Modular and Expandable

This risk calculator should be developed in a way that any user can easily implement and combine different methodologies. To make this attribute possible, an object-oriented philosophy should be adopted.

Different vulnerability, fragility and consequence functions should be input and run within the engine, allowing users to estimate damage to the buildings (structural and non structural), as well as direct economic losses and social losses, and compare results from the different approaches.

Uncertainty propogation

The engine should allow for the propagation of uncertainties, both aleatory and epistemic, at all stages of the risk assessment. A number of approaches (e.g. Monte Carlo or FORM for aleatory uncertainty and logic trees for epistemic uncertainty) should be accommodated within the engine.

Scalable

This tool should allow one to perform risk assessment at different levels of resolution from an urban level to a global scale. If only crude data at a national level is available, an empirical approach might be adequate to estimate risk for that country. However, if in some parts of the country data at census tract or even building-by-building level is available, an analytical (engineering-based) approach can be followed to assess the risk. This might mean that different results for a given area of the world are available, and it should be possible to visualise these different results as the user zooms to higher levels of resolution. Tools to aggregate and disaggregate the results for visualization at different levels of resolution (e.g. when viewing results globally as opposed to when viewing a given city) would also be beneficial.

To conclude, the GEM1 Risk team believes that its efforts in the testing, analysing and running of several seismic risk software have enabled the identification of shortcomings to be avoided and, more importantly, the features to strive for, in the development of a risk engine for GEM.

13.3 Lessons Learned from Global Seismic Risk Initiatives

The global risk assessments presented in Chapter 12 comprise some of the first attempts that have been made to globally assess the risks from natural hazards, including earthquakes. All of the earthquake risk assessments are based on empirical data, for both the hazard and the vulnerability. Whilst an empirical approach to the definition of vulnerability has many advantages related to the use of real observed loss data that is correlated to recent past events of different levels of intensity, the same cannot be said of seismic hazard. In some cases, data from a limited number of years (20-40 years) has been used to define the seismic hazard; such an observation length is too short for low frequency events such as earthquakes. There will be countries which have not experienced a large disaster in this period but which may nevertheless have a high probability of occurrence of a large magnitude event; these countries will not be identified by many of the risk assessments presented in Chapter 12.

Another limitation of the approaches considered in previous efforts is related to the lack of consideration of the built environment in the vulnerability assessment; this could have been considered even without full engineering-based calculations by considering countries with building construction similarities/differences or code compliance.

The GEM1 calculations presented Part 2 of this report represent a major step forward in the assessment of global seismic risk due to the use of probabilistic seismic hazard at a range of return periods. That is, the probability of occurrence of a range of ground-motion intensities is calculated on each node of a grid considering all possible sources of earthquakes. In these preliminary calculations, which are more focused towards a proof-of-concept of end-to-end global loss assessment, empirical vulnerability functions are being used which only implicitly consider the characteristics of the buildings; however, the future developments required to take into account more explicitly the built environment within the proposed framework are also laid out herein.

Part 2. Development of GEM1 Risk Engine

By H. Crowley, A. Cerisara, K. Jaiswal, N. Keller, N. Luco, M. Pagani, K. Porter, V. Silva, D. Wald, B. Wyss

14 Introduction

At the Canberra GEM1 Kick-off meeting in March 2009, it was unanimously agreed that the GEM1 Risk team should evaluate a number of existing software for seismic risk assessment, and that this should be done through a number of test-bed applications. The aim of this study was to identify the potential of existing tools in order to verify if there was a software that could fulfil the requirements of a global risk engine for GEM. Part 1 of this report presents the results of this evaluation and critical review, a preliminary version of which was presented to GEM's Model Advisory Group (MAG) in November 2009. There was a general consensus at this first MAG review that none of the codes were adequate for the GEM risk engine. The MAG endorsed "*the suggested approach of developing a new Risk Engine that addresses the specific requirements prescribed by the Risk Team, with the Risk Engine being based upon components identified from the reviews undertaken to-date*". The aforementioned requirements for the software that have been proposed by the Risk Team can be summarised as follows:

- Open-source software development: the source code of the global risk calculator should be available to any user and the development of the code should be a product of the efforts of a community, and not just limited to a working group.
- Platform independent: this tool should be able to be used in any operative system.
- Flexible: this code needs to be developed with the purpose of creating a platform for risk assessment, instead of another static risk application. This platform should allow users to evaluate risk due to seismic hazard, but also allow other types of hazard such as floods or hurricanes, to be considered.
- Dynamic: this calculator should allow users to update their results based on newer models, datasets or hazard inputs.
- Modular and expandable: this risk calculator should be developed in a way that any user can easily implement and combine different methodologies. To make this attribute possible, an object-oriented philosophy should be adopted.
- Scalable: this tool should allow one to perform risk assessment at different levels of resolution from an urban level to a global scale.

The MAG thus recommended to the GEM1 Risk group that:

- *"All endeavours be made to develop the new Risk Engine within the GEM1 framework and duration as it is an essential feature of being able to demonstrate complete hazard-to loss computation within the prototype GEM1 model. It is acknowledged that this may result in effort being redirected from the existing model evaluation process.*
- *The GEM1 risk model be kept simple, with allowances for tracking uncertainties in building characteristics and fragility/vulnerabilities if possible, but avoiding trying to populate or rationalise these within GEM1 – time for that later!*
- *That the Risk/IT linkage be strengthened so as to facilitate the development of a suitable risk engine with GEM1, one that enables the concept to be demonstrated."*

Hence, following the MAG review, a change of direction and focus of part of the Risk team was made in order to develop the requested risk engine. The EUCENTRE team, where the coordinator of the Risk group is located, began immediately to plan this development and decided to leverage existing expertise within GEM1 for what concerns the hazard engine, which already featured many of the aforementioned attributes suggested for the risk engine. Furthermore, it was felt that by working closely with the hazard engine development, the hazard and risk engines would become more consistent in their design and language, which would help to ensure a more homogeneous OpenGEM system.

This part of the GEM1 Risk Report describes the development of the GEM1 risk engine up until March 2010. Since the completion of GEM1, and following an IT review in June 2010, there have been many developments related to the engine (e.g. porting from Java to Python, setting up of open source development practices, use of Agile development philosophy and more); more information can be found on the GEM website.

15 GEM1 Risk Engine

15.1 Summary of Software

This Chapter summarises the GEM1 risk engine in a similar manner to that used for the seismic risk software studied and presented in Part 1 of this report.

The GEM1 risk engine implementation currently covers the following asset, loss and vulnerability types:

- Asset types: Population, Buildings
- Loss types: Fatality, Repair Cost
- Vulnerability: Discrete intensity measure (e.g. MMI) based

where Asset is defined as anything of value; Loss is a decrease in Asset value, and Vulnerability is the probability density function of a loss ratio given a scalar intensity measure level (IML), where the distribution of the mean loss ratio with IML can be discrete or continuous.

The software is currently able to carry out the following types of calculations for the asset and loss types described above:

Deterministic Event-based Analyses

- Maps of mean loss (or loss ratio) and standard deviation of loss (or loss ratio) for scenario earthquakes, for single assets;

Classical PSHA-based Analyses

- Loss (or loss ratio) exceedance curves (i.e. loss (or loss ratio) vs. probability of exceedance in a given time span) for single assets;
- Conditional loss (or loss ratio) maps (i.e. loss (or loss ratio) maps for a given probability of exceedance in a given time span), for single assets;
- Mean loss (or loss ratio) maps in a given time span, for single assets.

15.2 Methodology

The initial development of the GEM1 risk engine uses the methodology proposed by the PAGER global vulnerability model [Jaiswal *et al.*, 2009]; this is a direct empirical vulnerability approach as shown in Figure 1.1. This model was used as it provides global coverage and it was developed by one of the partners of the GEM1 Risk team. This approach uses a set of vulnerability curves that are based on historical data that contained fatalities due to earthquakes since 1973. This is an empirical model that uses ground shaking intensity (using the Modified Mercalli Intensity scale) to predict a fatality ratio. Each vulnerability function is described in a discrete way and each intensity measure level is associated with a loss ratio and a respective coefficient of variation. A similar vulnerability model to assess repair cost due to building damages was also implemented on the engine. This model (ATC-13) also follows an intensity-based methodology and it was developed by the Applied Technology Council [ATC, 1986].

It is important to understand the differences in the methodology when computing loss results for a deterministic scenario as opposed to using classical PSHA. Although the engine might be using the same vulnerability model in both cases, the way the hazard input is processed, the necessary calculations and the type of outputs that can be produced are considerably different. However, it is noted that due to the object-oriented design of the engine (see Section 15.3), a number of common pieces of code are used for both the deterministic event and classical PSHA-based analyses.

The current GEM1 risk engine uses ground-motion fields (also known as ShakeMaps) to perform deterministic event-based loss assessment and hazard curves to perform classical PSHA-based loss assessment. The format that these hazard descriptions take might differ significantly depending of which resource is being used and hence, several interfaces were created to make the engine suitable for a number of different hazard sources. Concerning the vulnerability, the collection of curves within the model needs to be defined at a discrete number of seismic intensity levels. The uncertainty of each value can be expressed as a standard deviation or a coefficient of variation for both the seismic hazard (where it is given at each site in the ground-motion fields) and the vulnerability model (where it is given for each intensity measure level). If the user doesn't want to consider the uncertainty, these parameters can simply be set to zero and all of the calculations are carried out considering the respective values as deterministic. Figure 15.1 shows the main differences and similarities between these two approaches.

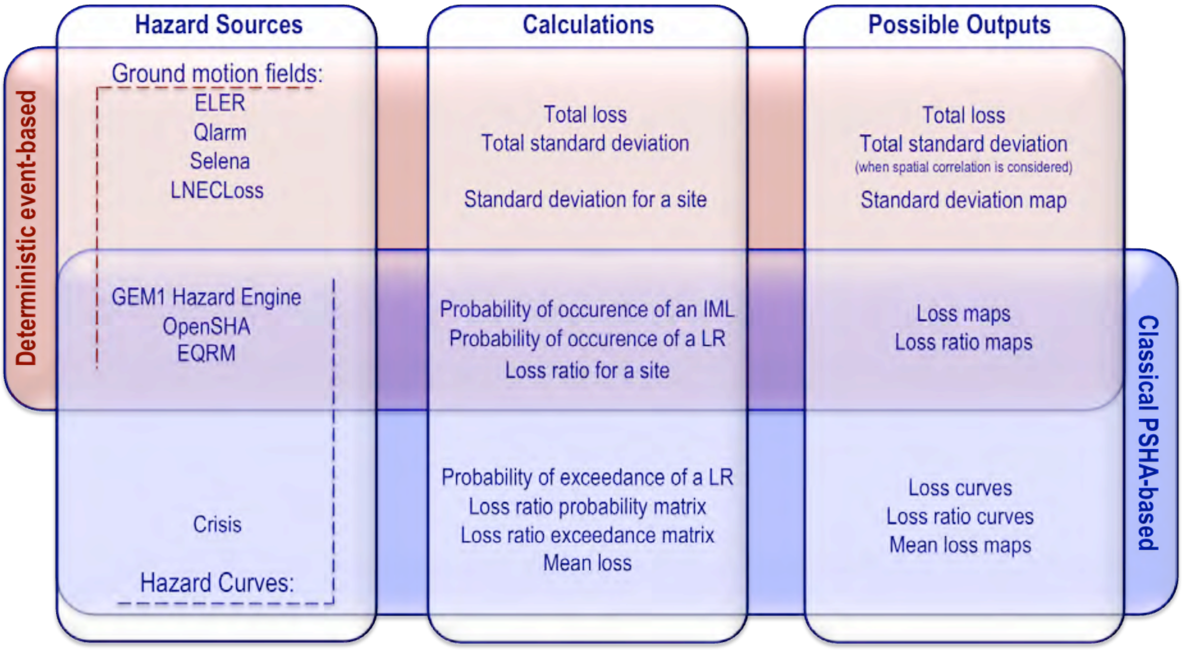


Figure 15.1 Differences between the features of deterministic event and classical PSHA-based risk assessment.

15.2.1 Deterministic Event-Based

The capability of predicting fatalities for a given scenario earthquake for any region in the world was the first achievement of the GEM1 risk calculator. To do so, the mean ground motion and respective uncertainty for every site of the region of interest need to be previously computed and inserted as an input into the calculator. As previously mentioned, different sources provide this kind of data in various formats. An optimal file format proposed within this document allows the intensity levels for every location, and information related to how the ground-motion fields were produced, to be stored. This tool was designed taking into consideration the differences between using ground-motion fields from real-events and using ground-motion fields simulated for a hypothetical event. This is an important feature since significant differences are expected regarding the distribution of the uncertainty for each site for these two cases.

After the hazard data have been read, the risk engine may need to adjust the intensity values to the requirements of the vulnerability model, such as excluding the intensity levels that are not within the vulnerability model range or the rounding of values to the intensity measure levels where the vulnerability functions are defined.

The main outputs of the calculator, for deterministic events, are loss and loss ratio maps, for both the mean and standard deviation. For loss maps, it is necessary to have access to a population exposure database. Several databases of world population distribution are already available and for this initial phase, the GEM1 risk engine is using LandScan™ [18] and GRUMP [19]. Since for earthquake scenarios the region of interest might be only a small portion of a certain country, it is also possible to provide a population distribution based on a national census or any other more detailed source. Further information will be given later within this document for each file format. Figure 15.2 illustrates the basic components of the calculator for this type of assessment.

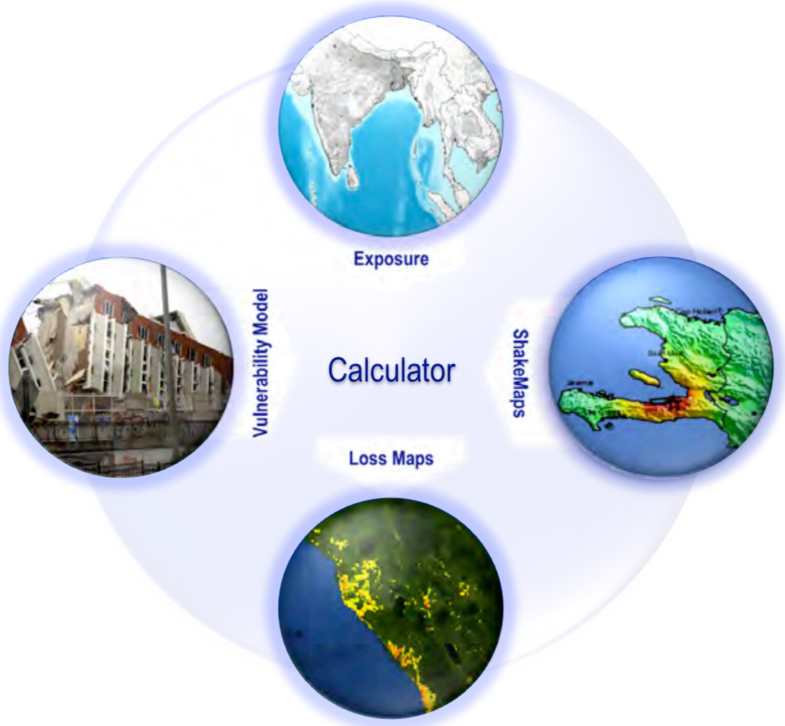


Figure 15.2 GEM1 Risk Engine components: deterministic event-based analysis

Although loss and loss ratio maps are considered the main outputs of this calculator, other results can be obtained such as plots of population exposed versus intensity measure level, losses versus intensity measure level, standard deviation maps, and aggregated mean loss (i.e. summed over all sites). Other parameters such as probability of exceedance of a loss ratio or expected loss ratio for a given probability are not currently featured on the risk calculator for deterministic events, however, due to the modular architecture of the code, these procedures can be easily incorporated using the methods that have already been developed for the classical-PSHA based calculations (see following section), should users express a need for such output.

The GEM1 risk engine can also currently be used for the assessment of economic losses (repair cost) due to building damage. The evaluation of this type of loss raises several issues regarding the need for a reliable building exposure database or the lack of vulnerability models that contain curves for every building typology. Currently, the calculator is using the vulnerability functions from ATC-13 [ATC, 1985] that were created based on United States construction characteristics. Due to these difficulties, the main objective was to extend the engine for damage and economic loss estimation in the most flexible way possible. This means that once a global building exposure database and respective vulnerability curves are available (which is the focus of two of the GEM Risk Global Components), the engine can easily incorporate them in the calculations. On the other hand, a user can also provide his/her own exposure database and vulnerability model and run the engine for the region that is covered by this data. The details of the calculations for deterministic event-based risk analyses are given in Appendix B.

15.2.2 Classical PSHA-Based

The next step in the development of the GEM1 risk engine was to extend the model to consider seismic hazard from classical PSHA. This type of assessment is considerably more complex than the previous one not only due to the fact that the seismic hazard is represented by curves of intensity measure level versus probability of occurrence/exceedance, but also because loss ratio probability matrices and loss ratio exceedance matrices are required and computed “on the fly”. Starting with the description of the hazard, once again these data need to be previously computed and inserted as an input to the calculator. A user can either insert the hazard curves using probability of occurrence or probability of exceedance although, if he/she chooses to provide probability of exceedance for each intensity measure level, then a minor procedure needs to be performed within the engine in order to always transfer probability of occurrence to the calculator. This feature is quite important since the majority of the hazard resources/codes tend to provide probability of exceedance instead of occurrence. Another challenge regarding the usage of hazard curves, was the need to come up with a way of storing this set of values for each location in a way that they could be easily read by the engine and that the file size wouldn’t be too large. A file format with such attributes and that also supports information regarding the process that was used to generate these curves is presented within this document.

For classical PSHA-based risk assessment, matrices are needed in order to obtain the loss ratio for each location for a given probability of exceedance. These matrices are computed taking into account the probabilistic distribution of the vulnerability functions and therefore, this parameter needs to be provided along with the mean loss ratio and coefficient of variation for each intensity measure level. At the moment, the risk calculator can support vulnerability functions whose uncertainty distribution is either normal or lognormal, but additional distributions such as Beta will need to be added in the future. For the assessment of human losses, the PAGER vulnerability model has a global coverage and hence all of these calculations can be done with no limitations. When computing economic losses due to building damage, there is currently no vulnerability model that covers a global taxonomy of building typologies. However, the goal of GEM’s Global Vulnerability Estimation Methods global component will be to begin to fill in such gaps. Figure 15.3 illustrates the basic components of the calculator for classical PSHA-based assessment.

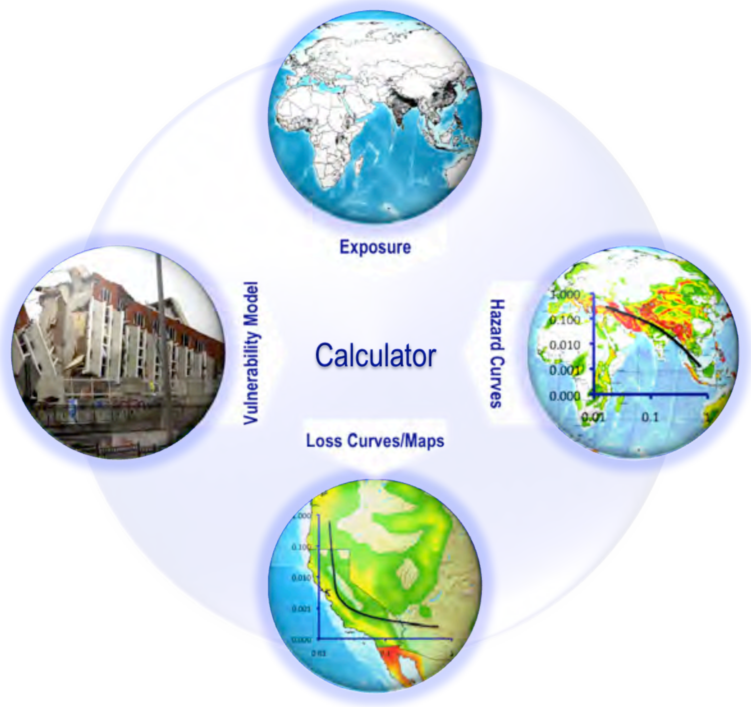


Figure 15.3 GEM1 Risk Engine components: classical PSHA-based assessment

The first results computed by this calculator are loss ratio curves (also known as loss ratio exceedance curves) that can be easily converted into loss curves by multiplying every value of the curve by the respective exposure value. Using these data and simple interpolation procedures, any loss ratio map or loss map at a given probability of exceedance within a certain time span can be created. Currently, the risk calculator is processing hazard curves that were computed for a time span of 50 years and producing loss/loss ratio maps for probabilities of exceedance of 1%, 2%, 5% and 10%. The mean loss for the chosen time span is also computed for each location. Examples of these types of output are presented in this report. The details of the calculations for classical PSHA-based risk analyses are given in Appendix B.

15.3 IT Details

Programming Language

The core framework was developed in the Java programming language. Java was chosen because it meets the requirements of being object oriented, platform independent, web enabled, and potentially distributed over the internet. However, the hazard engine (which is based on OpenSHA), that is also programmed in Java, uses data and/or modelling components that are implemented in other programming languages. These components are “wrapped” with an appropriate Java interface so they can plug into the framework.

Code Architecture

The GEM1 risk engine code has been divided up into a number of layers, as shown in Figure 15.4. Each layer represents a different package within the code repository.

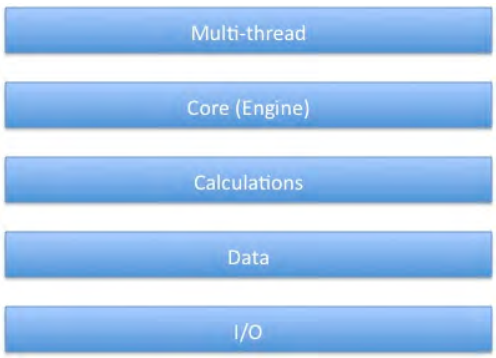


Figure 15.4 Layers of the code architecture

I/O Layer

The bottom layer “I/O” is a layer that is used to read input data from external sources and write output data to external sources. The engine is structured in such a way that technology-related details can be easily changed, due to the use of interfaces. For example, in the I/O layer one will find objects that are used for reading data from ASCII files or Raster files with different formats. One can thus easily change the input file format as this will not modify the core risk engine, but will just require a new object to be added that can read the data from this file. Currently, this layer is only used for input data but will soon be extended for the output data.

Data

The data layer contains the data read by the implementation classes in an "object-oriented form". It includes classes such as DiscreteVulnerabilityFunction, HazardCurve and Asset.

Calculations

The calculations layer includes the objects which carry out calculations, such as the calculation of loss exceedance probability matrices, of mean losses, of loss curves etc. using data from the data layer.

Core (Engine)

The core package includes the engine which processes assets over a given region using an event-based architecture.

Multi-thread

This layer has not yet been coded, but this package would have different instances of the engine which run and compute a specific subset of the entire region. Every instance would run inside a different thread to parallelize the computation.

The aforementioned layer structure divides the IT concepts and domain concepts into different packages; for example, the Implementation, Core and Multi-thread are IT packages whist the Data and Calculations are packages that are related to the risk domain.

Event-Based Architecture

The risk engine developed in GEM1 uses an event-based architecture with the following *events* identified at present: Start, Asset Loaded, Asset Valid Ready, Asset Invalid Ready, Asset Computed, Stop. The philosophy behind the engine is that it is a big loop that iterates on all sites, loading the asset of each site, as shown in the sketch in Figure 15.5.

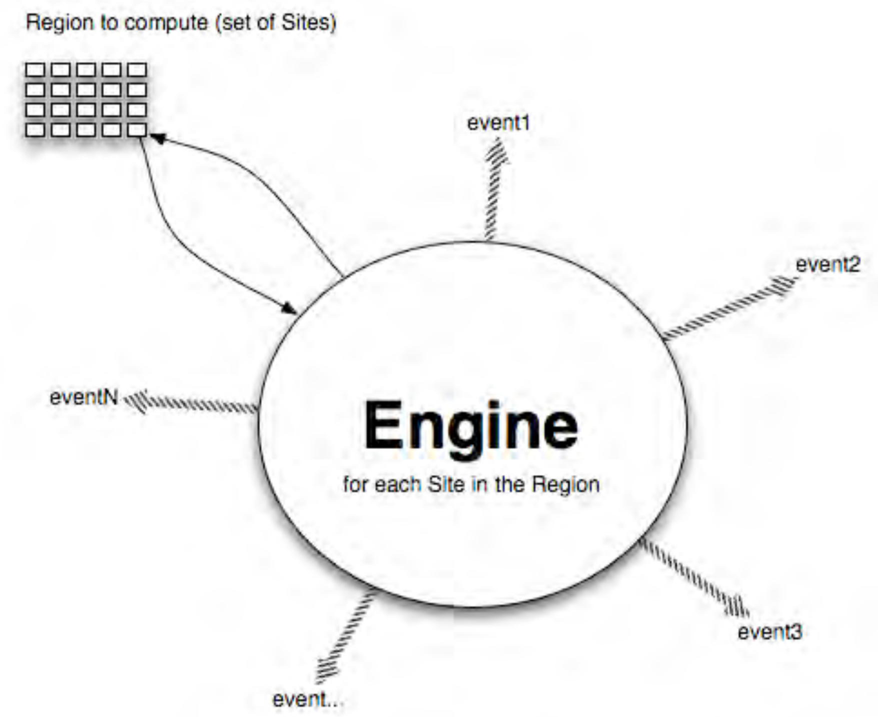


Figure 15.5 GEM1 Risk engine: event based architecture

At each event, a number of *listeners* can be attached and the engine will call them in order, using a *pipe and filter architecture*[1] so that data and results can be shared between listeners (Figure 15.6). For example, once the Asset Valid Ready event occurs (which means that the engine has loaded an asset and the asset is valid in that all necessary data is available), a vulnerability function listener will be waiting to see which vulnerability function should be loaded, then a distribution listener will wait to see which distribution the uncertainty in the vulnerability function is described with and so on. It is noted that the listeners themselves can also raise events, which brings higher flexibility to the engine.

[1] See e.g. http://www.dossier-andreas.net/software_architecture/pipe_and_filter.html

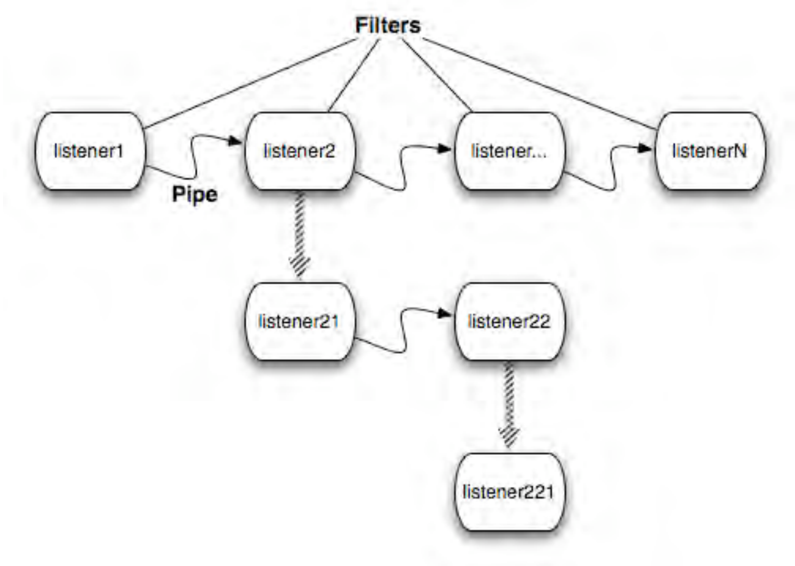


Figure 15.6 GEM1 Risk engine: pipe and filter architecture of the listeners

Caching and Performance

Events can be added to the computation that load and store computed data into the cache. At the moment the following implementation is used for caching: http://www.opensymphony.com/oscache/, but the engine can support different implementations because that logic is shielded behind an interface. The performance of the engine has been found to be linear, i.e. the time grows linearly with the growth of the number of cells to be computed. More information on the performance of specific applications of then engine is given in Chapter 16.

Command Line Client

The current code can be run using a command line client from which the user can:

- Specify an external administrative unit code file
- Specify an external population exposure file
- Specify a fixed CoV of IML when computing the deterministic event-based losses
- Specify a fixed mean IML value when computing the deterministic event-based losses
- Specify an external file with the spatial distribution of mean IML for scenario losses
- Specify the type of output (ASCII arc grid or a simple output with longitude latitude values)
- Specify the computation type (deterministic event loss, loss curve, loss map at a given return period)
- Specify an external file for the hazard curves
- Specify the region to compute
- Specify the cell size to use
- Specify the probability of exceedance at which to compute a loss map

Figure 15.7 shows a screenshot of the command line client with the aforementioned list of options. This is an alpha version which has limited functionality; further extensions and developments could be carried out in the future depending on whether these are considered necessary for GEM.

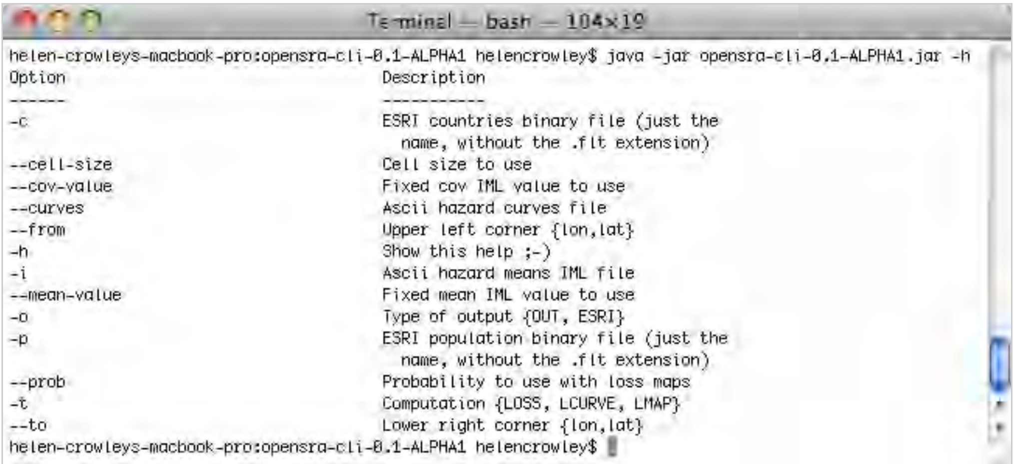


Figure 15.7 GEM1 Risk engine command line client screenshot

License

The GEM1 risk code that has been developed so far has been licensed with GNU Lesser General Public License Version 3, following the Governing Board decision on the licensing policy for GEM [Borgognoni and Milanesi, 2009].

In order to assign an open source license to the GEM1 code, the license was added to the top of each source file (copyright date, holder, and license and where full license text can be found) and a full license text was included with the code in a file called LICENSE.

15.4 Exposure Module

This section describes how the value of the different asset types is inserted into the calculator. Depending on the asset under consideration, different value types are expected. For instance, population count is expected when computing human losses, while for economic losses due to building damages, replacement/repair cost is required. The GEM1 risk engine has been designed to support global exposure databases although, should a user have more detailed data, local databases can also be used.

Since the calculator is using a country-based vulnerability model for predicting human losses, it is also necessary to rely on a world administrative limits database to relate every grid cell with the corresponding region. A full description and examples of these databases are presented within this section.

15.4.1 Population

For the implementation of the PAGER global vulnerability model, a database of the population distribution is required. Databases with such attributes already exist and in this initial phase, GEM1 has reviewed the following databases based on a report produced by the Food and Agriculture Organisation [Salvatore *et al.* 2005]: Gridded Population of the World (GPW, Version 3) [20], LandScan™ [18] and GRUMP [19]. After a critical evaluation of these sources, it was decided to make the risk engine suitable for the last two databases. General information about how each database was produced, along with the main advantages and disadvantages are given below.

LandScan™

The Oak Ridge National Laboratory produces this (proprietary) database and it can be obtained in an ESRI grid format or ESRI binary raster format. It has a 30 arc second spatial resolution (about 1km at the equator) and a great number of data sources were used to create this database such as: Digital Chart of the World (DCW), VMap1 (a map of major roads and rail networks, drainage networks, utility systems, elevation contours, coastlines, international boundaries and populated places), Nighttime lights, Global Land Cover Characterization (GLCC) [21] and high-resolution aerial photography and satellite imagery. The dataset provides population estimates based on aggregate data for second order administrative units from the US Census Bureau's International Program Center [22]. An interpolation method is used that assesses the likelihood of population occurrence in cells on the basis of road proximity, slope, land cover and Nighttime lights. There is no specific distinction made between urban and rural areas, though urban areas can be inferred by analyzing population density [Dobson *et al*., 2000]. The most recent version of LandScan™ is from 2008 and it has a world coverage that goes from North 84 degrees to South 90 degrees and West 180 degrees to East 180 degrees. The values of each cell are integers, which represent average population count.

GRUMP

The Global Rural Urban Mapping Project (GRUMP) dataset is produced by the Center for International Earth Science Information Network (CIESIN) is currently an alpha version with a 30 arc second resolution. A combination of the following datasets have been used to produce this database:

- Georeferenced human settlements database (55,000) with a population of more than 1000 according to census data, Gazetteer [23] and City Population [24].
- Nighttime lights and the DCW datasets: these are used to identify urban extents, and population values from settlement points within a 3km buffer are assigned to these extents.
- Administrative boundary databases containing total population for each administrative unit.
- United Nations (UN) National estimates (UN Population Division [25]): these are used to obtain the percentage of population in urban and rural areas and the total population.

The population is assigned to the cells considering the classification of the areas in urban or rural. After this, checks are made on the total population in each administrative unit according to UN estimates [Nachtergaele and Petri, 2007]. The urban extents are defined by a combination of datasets and not just Nighttime lights, which can miss small settlements in less developed countries and greatly overestimate urban extents for large settlements due to the “blooming” effect.

Critical Review of Population Databases

Although a few spot checks of population databases can be made, there will be some areas of the world where one dataset performs better than another and the opposite might be true in another area. A study by Gunasekera et al. [2009] has shown that in Istanbul the LandScan data was closer to the ground truth, but more comparisons are needed to understand which population database is more reliable. For this reason, considering that the population density is thus an epistemic uncertainty in loss calculations, both the GRUMP and LandScan databases were implemented in the risk engine.

One problem with LandScan concerns the road database. The model processes the input layers by country without taking into consideration the spatial continuity of the roads networks between them, resulting in uneven changes of population density at country boundaries. Another problem is that, owing to the way in which LandScan processing methods evolved, population comparisons between available revisions of the database is not possible. Also, the underlying models have not been published or few information can be found about them and hence, assumptions employed by LandScan to distribute population counts per cell are not known. Another problem that has been identified within this dataset is the fact that Nighttime lights can be more linearly correlated with GDP and electrification than population density and there is a clear “blooming” effect which means that the extent of urban areas is often overestimated and some small settlements are not clearly identified [Elvidge *et al*., 2004].

Regarding GRUMP, the main advantage is that it uses population data from the census, rather than predicting it based only on probability coefficient or lighted areas. Also, it makes use of other geographic information systems data to identify urban areas, compensating for the small settlements in poor countries that are not detected by the Nighttime lights. The resulting grid is a dataset of population distribution that takes into account the urban and rural areas. Some limitations of GRUMP are the fact that although it recognizes that applying a threshold would reduce the number of small settlements that are not frequently identified by Nighttime lights, due to the complexity of finding a single threshold that would work globally, no light threshold was applied. This limitation also produces overestimation of the urban extents in some parts of the World [Salvatore *et al.,* 2005]. In addition, this database lacks continuous updating and therefore, only the population distributions for three time periods are found: 1990, 1995 and 2000. Besides this fact, it is also important to understand that the lights factor refers to a time between 1994 and 1995 and hence, care should be taken when using this database in countries that experienced a fast growth during the last decade (e.g. Kuala Lumpur, Malaysia).

The so-called blooming effect, that affects both databases in different ways, is an overestimation of the real extents of urban areas or population distribution, and is believed to be dependent on the intrinsic characteristics of the devices that are used to capture the Nighttime lights. Early efforts to threshold the lights globally proved to be inappropriate since usually different results are obtained for cities with different sizes and often several small settlements are lost in the process if they are not frequently illuminated. Regarding the GRUMP database, some studies showed that even though a 10% threshold could reduce the blooming effect without significantly attenuating many individual small settlements for the 1994/1995 dataset, this detection frequency threshold does not provide a globally consistent basis for reconciling lighted areas to the urban environment. For this reason, and because detection is more of a concern than blooming for the global database, no thresholds were applied. Using instruments from NASA's Terra and Aqua satellites, as well as the joint U.S. Geological Survey-NASA satellite LandSat, researchers created land-use maps distinguishing urban surfaces for some zones in the globe. Figure 15.8 shows the difference between the urban limits produced using Nighttime lights and data from the LandSat Program for the city of Quito in Ecuador. In this case, the area provided by the Nighttime lights is about 10 times bigger than the urban limits produced by the second resource.

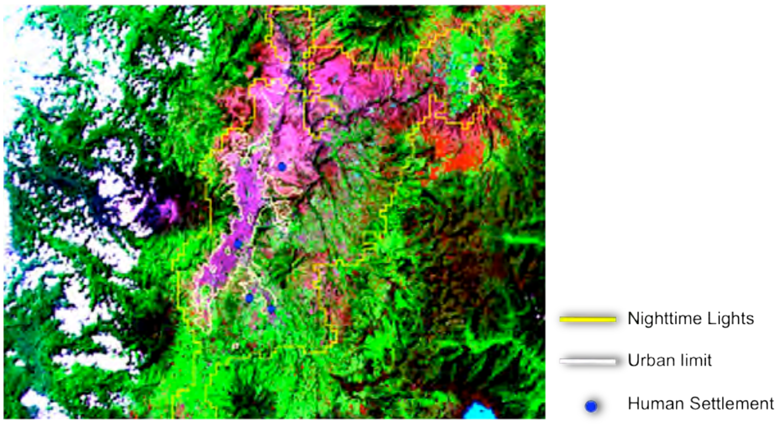


Figure 15.8 Blooming effect shown for the city of Quito, Ecuador.

One of the main conclusions that can be taken from this discussion is that it is important to consider the uncertainty in the distribution of both human and building populations.

15.4.2 Buildings

A spatial distribution database of buildings at a global scale does not currently exist, but methods to estimate the built area/number of buildings from census data and remote sensing will be developed and applied within GEM together with crowd data collection methods (in the Global Exposure Database and Inventory Data Capture Tools global components); these methods would theoretically allow for the future population of such databases. In order to overcome the current lack of such data, one solution could be to use the building area/number information which has been modelled at level 0 (national) or level 1 (sub-national) administrative boundaries by PAGER, and to distribute this to the grid cells based on population and/or GDP distribution and the land use/cover characterisation.

An alternative of the use of global building inventories is to create national building databases that might be enough if a user only wants to obtain data from a certain region. Such databases have been created already for some large cities exposed to regular seismic activity such as Istanbul, Lisbon or Los Angeles and have been used by the GEM1 Risk team in the software evaluation study.

PAGER Building Inventory Approach

The PAGER group has already taken the first steps in the development of a global building inventory database [Jaiswal and Wald, 2008]. This inventory consists of estimates of the fractions of building types observed in each country, their functional use and their average day and night occupancy. Four tables, each reflecting a combination of rural or urban and residential or non-residential categories essentially comprises this database. The fraction of building types or dwellings and their occupancy characteristics have been collated for each country to represent a country-based distribution using the PAGER structure taxonomy. This construction type classification was a product of an evaluation of several sources that classify buildings according to attributes such as structural system, load transfer mechanisms, predominant construction material, performance during past earthquakes, etc. The PAGER group recognized that none of the existing sources could provide an adequate classification of the buildings since most of them required building-specific information that was not available for most of the inventory data. Hence, it was necessary to adopt a classification based on the material used for the construction of the walls and roofs. This classification is similar to the one used by most of the housing census and surveys carried out for a large number of countries in the past.

In order to produce this building inventory, a great number of data sources were studied and rated according to the process used to gather the data. This classification can give an idea of how reliable these sources are. The following list describes the three levels used to rate the quality of the data:

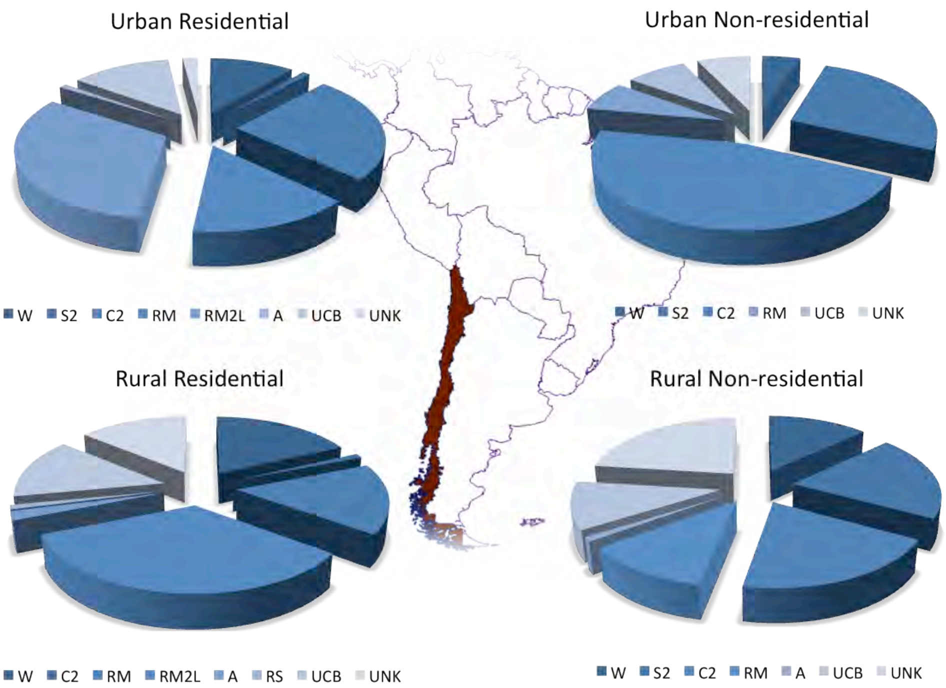
- Low: data compiled by non-engineering agencies and is not specifically meant for engineering loss analyses;
- Medium: data compiled by general field survey and assignments generally not based on engineering standards;
- High: data compiled after engineering or telephonic surveys; field visits for ground-truth or data compilations from local engineering experts;

The following table describes all the sources that were used, the coverage of each one and the assigned quality rating:

Table 15.1 Building inventory data sources.

Source of data	Quality Rating	Global Coverage
World Housing Encyclopaedia	Medium	110 residential construction types in 37 countries. Exact fraction of each housing type in a given country is not known. The day and night time occupancy by construction type is available.
UN Database	Low	44 countries with construction type description based on external walls and 96 countries with type of housing units. About 110 countries with the average occupancy estimated based on total building stock.
Census of Housing	Medium	197 countries conducted housing census in 1990. Several countries do not publish housing statistics even though housing census was conducted.
Published Literature	High	About 10 countries have been identified that contains high quality information based on the conducting survey and the verification of other published information such as census/tax assessor's data. The day and night occupancy by construction type is not available.
WHE-PAGER Survey	High	Inventory information for about 22 countries has been gathered in the first phase of WHE- PAGER expert opinion survey. In order to facilitate their judgment, country-specific inventory information gathered from general internet research and housing censuses was provided to these experts.

The amount of data that can be obtained from this building inventory database is presented in Figure 15.9 using Chile as an example:



Label	Description	Label	Description
W	Wood Building	A	Adobe Block Walls
S2	Steel Braced Frame	RS	Ruble Stone Mansory
C2	Reinforced Concrete Shear Walls	UCB	Unreinforced concrete block mansory
RM	Reinforced Mansory	UNK	Unknown Category
RM2L	Reinforced Mansory Low-Rise		

Figure 15.9 Building distribution per type of area/use for Chile.

The described database will be constantly updated as more data become available and therefore, it is expected that some parts of the inventory will be replaced by better quality data. The current database can be downloaded from the following URL: http://pubs.usgs.gov/of/2008/1160/downloads/PAGER_database/.

15.4.3 Administrative Limits

Since the PAGER global vulnerability model followed by this engine gives different relationships between fatality rate and intensity measure level according to country, a database that specifies the limits of each administrative region was required. This database will allow the calculator to choose the vulnerability function for each cell, taking into consideration where this cell is located. For the implementation of the PAGER vulnerability model, the vulnerability functions only change from country to country, with the exception of the United States of America where the state of California has its own vulnerability function. In the future, a user might want to establish his/her own set of vulnerability functions that might only be applied to a certain region and not the whole country and hence, smaller administrative levels were also considered.

Several databases that can provide different levels of the administrative boundaries already exist such as GRUMP [19], LandScan™ [18], GADM [27], GLC2000 [28] or GAUL [26]. Currently, the GEM1 risk engine uses the latter as its administrative limits database since it is believed to be one of the most reliable sources for spatial information.

The Global Administrative Unit Layers (GAUL) is an initiative implemented by FAO within the EC-FAO Food Security Programme funded by the European Commission. GAUL aims at compiling and disseminating the most reliable spatial information on administrative units for all countries in the world. GAUL keeps track of administrative units that have been changed, added or dismissed in the past for political reasons. Changes implemented in different years are recorded in the GAUL on different layers. For this reason the GAUL product is not a single layer but a group of layers, named the "GAUL Set", in which each layer represents the administrative units of the world at a given point in time. In its current form, each layer of the GAUL set includes three levels of administrative boundaries and units:

- Level 0 – International limits;
- Level 1 – First level of spatial organization;
- Level 2 – Second level of spatial organization;

If France were used as an example, the spatial organization would be as shown in Figure 15.10.

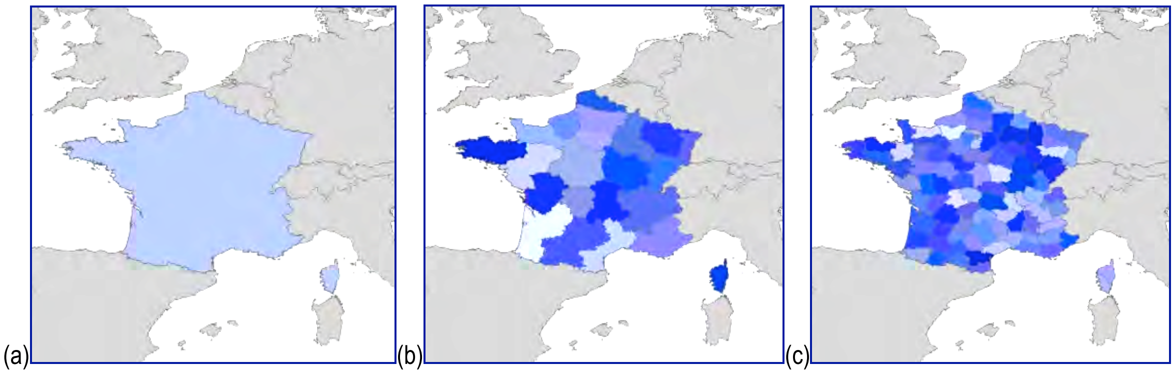


Figure 15.10 a) Level 0 - International limits, b) Level 1 - Administrative regions limits, c) Level 2 - Departments limits.

It is also important to understand that the nomenclature used to define the name of these smaller regions might differ from country to country (e.g. USA – states and counties, Italy – regions and provinces, Portugal – districts and counties). These three levels were built using the Global Administrative Unit Layer (GAUL), though further improvements have also been implemented. For instance, for small countries such smaller levels of spatial organization did not exist in GAUL (e.g. Guam, Switzerland, Aruba), and therefore it was necessary to add them. In order to perform this procedure, shapefiles with the desired spatial organization can be downloaded from sources such as DIVA-GIS [29], and incorporated in the main administrative layer.

It could also happen that the chosen vulnerability model requires an administrative limits database that doesn't fit into the previous mentioned levels since it might be a mixture between many levels. For instance, the PAGER vulnerability model is basically country-based (level 0) but it has a different vulnerability function for California. Because of this exception, a customized database was created where this state would be differentiated from the rest of the country. From a computational point of view, this database simply provides a code per grid cell to the calculator, which allows it to relate that cell with the designated vulnerability function for that country. In order to reach a standard way of using these codes, the system shown in Table 15.2 was created:

Table 15.2 World administrative limits code system.

Administrative limits	Code size	Description
Level 0	AAA	ISO number of the country, alphabetically ordered.
Level 1	AAA BBB	ISO number of the country followed by the number of the second administrative region, alphabetically ordered.
Level 2	AAA BBB CCC	ISO number of the country followed by the number of the second administrative region, followed by the number of the third administrative region both alphabetically ordered.

Figure 15.11 clarifies how this code would be computed for every region in the world.

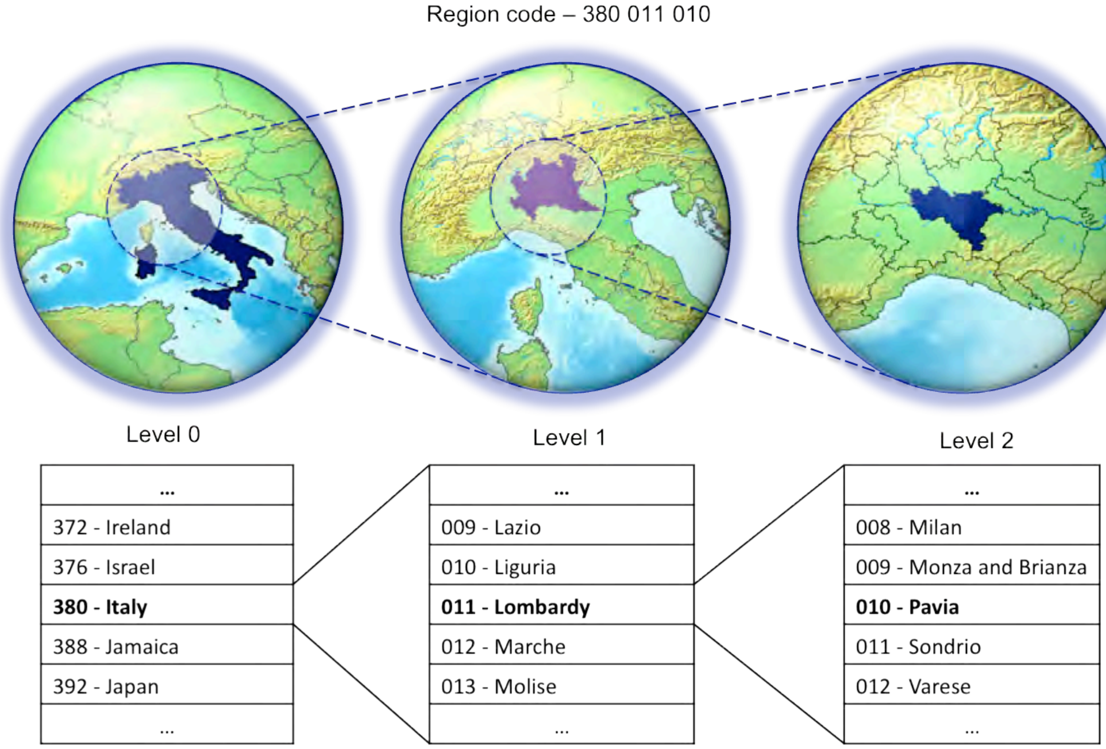


Figure 15.11 Interpretation of the administrative code system.

Along with these three layers, tabular databases with the names of all the countries, respective second administrative boundaries and finally the third administrative boundaries were also created. All of these regions have been identified by their names in English and arranged in alphabetical ascendant order.

15.4.4 Global Urban/Rural Classification

A database that can classify each area in the world between urban and rural might be necessary since it is common to use different methods to distribute human population within each type of aforementioned area. Also, it was mentioned previously that a global building inventory might also be developed taking into consideration the different distributions of buildings within these two classifications. Due to these reasons, an urban/rural database was also studied and implemented on the risk engine. This database was produced within the GRUMP project and it delineates urban extents associated with human settlements globally. The Nighttime Lights dataset for the period 1994–1995, DCW Populated Places, and cities from the Tactical Pilotage Charts (standard charts produced by the Australian Defence Imagery and Geospatial Organization) were used to obtain the maximum possible coverage for each country [Salvatore *et al.,* 2005]. Figure 15.12 shows the distribution of urban and rural areas in Europe.

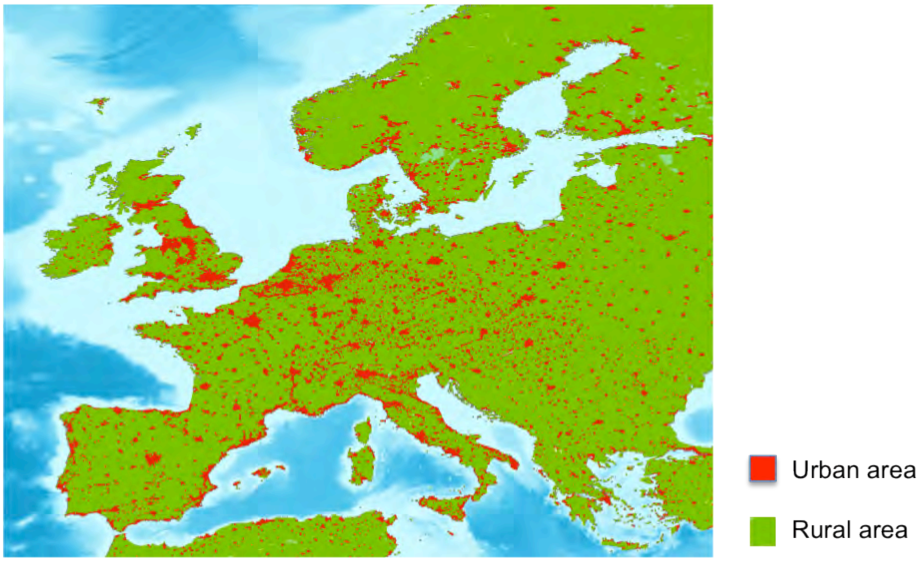


Figure 15.12 Distribution of urban and rural areas in Europe.

The definition of urban areas can vary significantly from country to country and in some of the cases it might not be directly related with the amount of population living in that area but with other parameters associated with the infrastructures and services available in the region. For instance, according to the Demographic Yearbook produced by the United Nations [2008], the definition of urban area for Peru is a region with more then 100 dwellings while for Pakistan is an area where the municipal corporation or town committee reside. Even for countries whose urban classification is directly related with the amount of population in the area, this quantity of people might differ considerably. For instance, in South American countries such as Argentina or Bolivia, urban implies populations of greater than 2000 while for Japan or Turkey more then 20,000 people are required.

15.4.5 Irregularities between Databases

The GEM1 risk engine relies on global databases produced by different sources (e.g. LandScan™, GRUMP, GAUL), which might lead to issues regarding the location of the administrative limits between countries and location of the coastline. Country boundaries might change frequently in zones with active political conflicts (e.g. Middle East) and therefore, different international limits might be found between databases produced at different times. In order to overcome this issue, care should be taken regarding the period when the databases were produced.

Another source of error can be the location of the coastline between databases. For instance, in order to predict human losses globally, the risk engine needs to access the global administrative limits database to understand the location of each cell and attribute a vulnerability function and then, access a second database in order to retrieve the population count for the respective cells. In both databases, a specific value is retrieved if the engine reaches grid cells that are not located in land (a *NoData* value is returned which is usually -9999). These particular grid cells are the ones located on the oceans, lakes, seas and rivers where there is (in general) zero exposure. While developing the risk engine, it was noticed that this *NoData* values were often not retrieved simultaneously by both databases for grid cells located near the coastline. A closer look at the geographic location of this line between the different sources revealed that there is not a common definition of the position of this limit. Two cases can be found:

- Case A: In this situation, the administrative limits database extend further into the ocean and include cells in land that according to the population distribution database, are already off limits. Basically, the risk engine would know what fatality rate should be applied to those locations, but wouldn't have any information about the affected population.

- Case B: In this case the opposite happens, which means, the population distribution database assigns population to grid cells that are not included by the administrative limits. This means that the engine would recognize that there is population in those places, but wouldn't know which vulnerability function should be used.

Figure 15.13 presents these two cases.

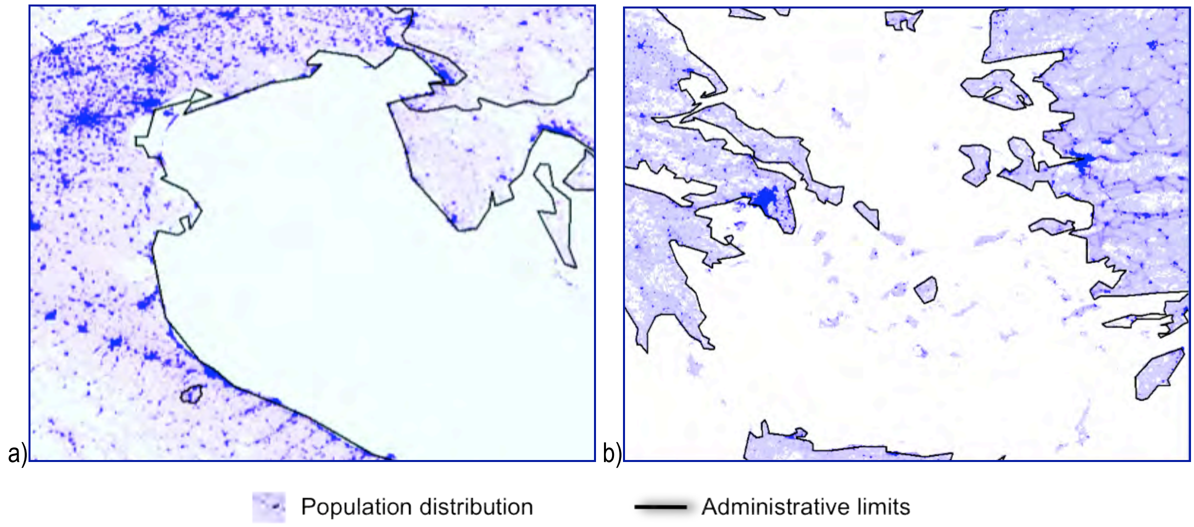


Figure 15.13 Different location of the coastline between databases in a) Case A: Gulf of Venice, b) Case B: Greek islands.

In order to avoid this type of errors it was decided to discard the grid cells every time that any of the databases retrieves a *NoData* value. Obviously the occurrence of case B implies ignoring population that could be affected by significant ground motion. These situations show that it is important to harmonise databases to ensure that the international limits and coastlines are similar.

15.4.6 Exposure File Format

Currently all of the previously described data is being stored in binary raster files. This format is simple and can be used to transfer raster data between various applications. It consists of two files, the IEEE floating-point file and a supporting ASCII header file. The header file must have the same name as the data file, but with a .hdr file extension. The header data includes the following keywords and values:

- ncols - number of columns in the data set.
- nrows - number of rows in the data set.
- xllcenter or xllcorner - x-coordinate of the centre or lower-left corner of the lower-left cell.
- yllcenter or yllcorner - y-coordinate of the centre or lower-left corner of the lower-left cell.
- cellsize - cell size for the data set.
- nodata_value - value in the file assigned to cells whose value in unknown. This keyword and value is optional. The *NoData* defaults to -9999.
- byteorder - the byte order of the binary cell values. You can choose between two keywords, msbfirst or lsbfirst. Msbfirst is used for cell values written with the most significant bit first. Lsbfirst is used for cell values written with the least significant bit first.

15.5 Hazard Module

The current GEM1 risk engine uses ground-motion fields to perform deterministic event-based loss assessment and hazard curves to compute results for classical PSHA-based losses. Both hazard types are represented in different formats and can be obtained from distinct sources. Currently, the seismic hazard needs to be pre-computed and inserted in the risk calculator in a specific format that is described below. This section also explains how the uncertainty is being considered, the procedures in order to adapt the hazard input to the requirements of the vulnerability model, and how this risk engine has been developed in a way that allows multiple hazard sources to be used.

15.5.1 Ground-motion fields

This type of seismic hazard format contains the ground motion produced by a single seismic event. Several applications can produce this kind of data such as OpenSHA [30], GEM1 hazard engine, ELER [Cagnan *et al.*, 2009]. This ground motion is usually provided in Peak Ground Acceleration (PGA) or Spectral Acceleration (SA) however, the GEM1 risk engine is currently using intensity-based methodologies that require macroseismic intensity such as Modified Mercalli Intensity (MMI). Due to this requirement, conversion equations can be used to obtain the ground motion into the desired format [Allen and Wald, 2009, Cua *et al.*, 2010]. Besides this conversion procedure, it is also important to understand that the vulnerability models are being defined in a discrete mode and, therefore, it is necessary to round the ground motion value according to the intensity measure levels where these functions are defined. The following example shows how this whole procedure is done using the earthquake of January 2010 in Haiti.

1 – The ground-motion field is obtained using Peak Ground Acceleration as the intensity measure type:

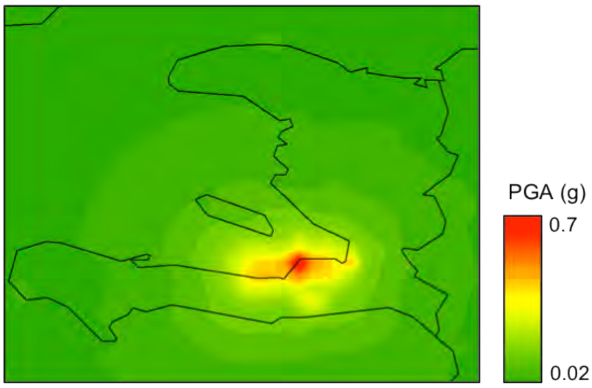


Figure 15.14 Ground-motion field in PGA for the Haiti earthquake (12/01/2010).

2 - Assuming that human losses are going to be computed using the PAGER global vulnerability model, it is necessary to use a conversion equation to convert PGA to MMI. For this example, the equation proposed by Wald *et al.* [1999b] was used:

$$I_{mm} = 3.66\log(PGA) - 1.66 \tag{15.1}$$

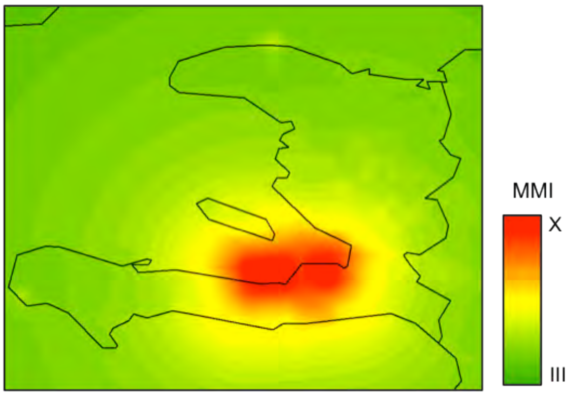


Figure 15.15 Ground-motion field in MMI for the Haiti earthquake (12/01/2010).

3 - Since the vulnerability model is bounded between 5 (V) and 10 (X) with increments of 0.5, the intensity measure levels were rounded using a round half up method and keeping the first decimal place:

$$I_{mm} = Round\_half\_up(I_{mm}) \qquad (15.2)$$

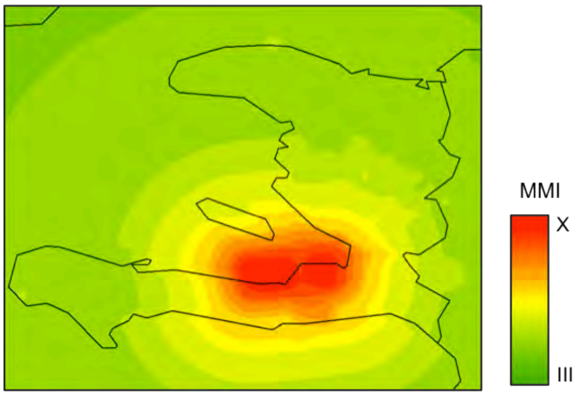


Figure 15.16 Ground-motion field in MMI (rounded) for the Haiti earthquake (12/01/2010).

Uncertainty

Even though many of the common tools used to compute ground-motion fields tend to ignore the uncertainty, it was decided that the GEM1 risk engine should be developed in a way that this source of error is considered. The development of this part of the calculator was a product of a joint effort between the risk group and the hazard group, since it was necessary to understand how this uncertainty was going to be provided to the risk engine. It was concluded that the uncertainty on post-event ground-motion fields (real scenarios) and pre-event ground-motion fields (hypothetical scenarios) should be different. For the latter, the same uncertainty is expected in every grid cell since the values are computed based on a ground-motion prediction equation, which has a specific logarithmic standard deviation. This source of uncertainty might increase should a conversion equation be used, and in this case, the standard deviation of this equation is also considered. On the other hand, for post-event ground-motion fields, the uncertainty of the values depends of the standard deviation of the ground-motion prediction equation, but also of other parameters such as the proximity to a ground motion observation point, the magnitude of the event and the knowledge of the dimensions of the source [Wald *et al.*, 2008]. Since a large range of uncertainty is expected for this latter scenario, it was decided that a second map would be used to record these values where, instead of the intensity measure levels, the spatial distribution of standard deviation is stored.

File Format

The file format proposed herein was a product of a critical review of an initial study for GEM1 carried out by Porter and Scawthorn [2009] with the purpose of developing several data exchange formats that would fulfil the requirements of GEM. The proposed file format to store ground-motion fields can be separated in two parts: the header and the gridded data. The header is composed by two rows, where the first one has information regarding the seismic event, and the second row has the information needed to relate the ground motion with each cell. The following list describes these parameters:

First row:

- ID – this is an integer that represents the ground-motion field on the calculator;
- IMT – the intensity measure type represented on the ground-motion field;
- ERF, Source, and Rupture – These parameters are related with how the ground motion was computed. This set of attributes might not be adequate to reflect this process for zones where an earthquake rupture forecast model might not be well defined or in case a user just wants to define a pair of coordinates and a moment magnitude to generate a hypothetical seismic event. Therefore, more developments in this area will be done when further feedback has been received from the hazard group;
- Magnitude – magnitude of an event assumed to be moment magnitude, M_w.
- GMPE – This parameter represents the ground-motion prediction equation used to compute the intensity measure levels. Using a single code to represent this attribute might not be enough since more than one GMPE might be used for the same ground-motion field. Furthermore, it could also be relevant to store the conversion equations that might be used in case macroseismic intensity is required. For this reason, further information might be added to this parameter.

In order to achieve a way to represent these parameters in a standard way, it was decided to adopt the codes proposed by Porter and Scawthorn [2009] for the IMT, ERF and GMPE.

Second row:

- Number of columns – Number of columns of the grid;
- Number of rows – Number of rows of the grid;
- Longitude – Value of the longitude of the bottom left cell in decimal degrees;
- Latitude – Value of the latitude of the bottom left cell in decimal degrees;
- Cell size – Width of the cells in decimal degrees.

After these two rows, the gridded data is presented. Every line represents a row of the grid and the columns are separated by tabs. Figure 15.17 presents an example of this file.

34	MMI	[ERF]	[SRC]	[RUP]	7.8	CB2003						
78	36	27.225	41.575	0.05								
6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.5	5.5	5.5	5.5	5.5	5.5
5.5	5.5	5.5	5.5	5.5	5.0	5.0	5.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	5.5	5.5	5.5
5.5	5.5	5.5	5.5	5.5	5.5	5.0	5.5	5.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.0	5.0	5.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Figure 15.17 File format proposed to store ground-motion fields.

The uncertainty for post-event ground-motion fields are stored in a similar way, where instead of having intensity measure levels, one will have standard deviation. Traditionally, tools capable of producing seismic hazard tend to provide this data in a "longitude, latitude – value" format, which increases the amount of information needed to be stored since the coordinates for every grid cell are recorded. Storing these values in a gridded format has important advantages such as the fact that the risk engine can process the data faster and the file size is reduced considerably. Although it is acknowledged that the size of the files are not a concern for deterministic scenarios where only a small region might be affected by significant ground motion, this does not happen when assessing probabilistic hazard, hence, these measures were taken in order to enhance the way of storing this information.

15.5.2 Hazard Curves

Curves of probability of occurrence/exceedance as a function of intensity measure level are used as the hazard input for classical PSHA-based assessment within the GEM1 risk engine. The data used while testing the engine and computing results was produced by the GEM1 hazard engine and therefore, it was possible to request some specifications regarding the type of probability that would be provided and the file format where these data would be stored. These curves have been computed for the intensity measure levels required by the PAGER global vulnerability model and using probability of occurrence. This aspect spared the risk engine additional calculations that would delay the computing time considerably. However, it is important to prepare the calculator to also be capable of performing procedures that would convert this type of data to fit the requirements of any vulnerability or hazard model.

File Format

To define a hazard curve, a set of intensity measure levels and corresponding probabilities of exceedance/occurrence within a certain time span are necessary. The following list presents all the parameters that are important to fully describe how the hazard curve was computed:

- Time span – This parameter describes the time span within which the seismic events are assumed to happen (in years);
- IMT – Intensity measure type represented on the hazard curves;
- ERF – This parameter describes the model used to predict all possible earthquake ruptures, and their associated probabilities of occurrence, throughout a region for a specified time span;
- Type of curve – Many types of curves can be provided such as: median, mean, particular percentiles or curves related with a specific branch of the logic tree.
- GMPE – Name of the ground-motion prediction equation(s) that was used to compute the ground motion at each site;
- CE – Name of the conversion equation that might have been used to convert ground motion from one Intensity measure type to another.

The first four parameters will be inserted on the first row of the file and therefore, it is assumed that all the curves within this file follow these same parameters. The other two parameters may differ within the region of interest since zones might exist with different seismologic characteristics (e.g. active crust, subduction zone or stable continent). For this reason, a system was created to relate every site with a GMPE and CE. Further information will be given below about this method.

Currently, the probabilities of occurrence are calculated for the intensity measure levels (IML) where the vulnerability functions are defined and therefore, it is necessary to provide a list of these values to the hazard engine. This aspect avoids the use of interpolation methods that would reduce the accuracy of the model, however it is important to understand that in the future the intention is to make the risk engine flexible enough to handle hazard curves that might not be defined at the same intensity measure levels as the vulnerability model. This will allow the risk engine to accept hazard curves produced by other applications.

The set of IML's that the risk engine will provide depends on whether the hazard engine will compute probability of exceedance (PE) or probability of occurrence (PO). For the latter option, the values of IML's should be exactly the ones defined on the vulnerability model. To compute the probability of occurrence of a certain IML, it is necessary to establish an upper and lower limit for that level. Assuming that the PAGER global vulnerability model is being used and since the intensity measure levels are defined between 5 and 10 with 0.5 increments, each range will be: IML $\pm$ 0.25. So for instance, the probability of occurrence of an intensity measure level of 6, will be the probability of occurrence of any intensity between 5.75 and 6.25. In general, this range can be given by the following expression:

$$\left(\frac{IML_{n-1} + IML_n}{2}\right) \leq IML_n \leq \left(\frac{IML_n + IML_{n+1}}{2}\right) \tag{15.3}$$

On the other hand, if the hazard engine provides a list of probabilities of exceedance, then the IML's where these probabilities are going to be computed should not be those from the vulnerability model, but the limits between the IML's, as presented in the following figure:

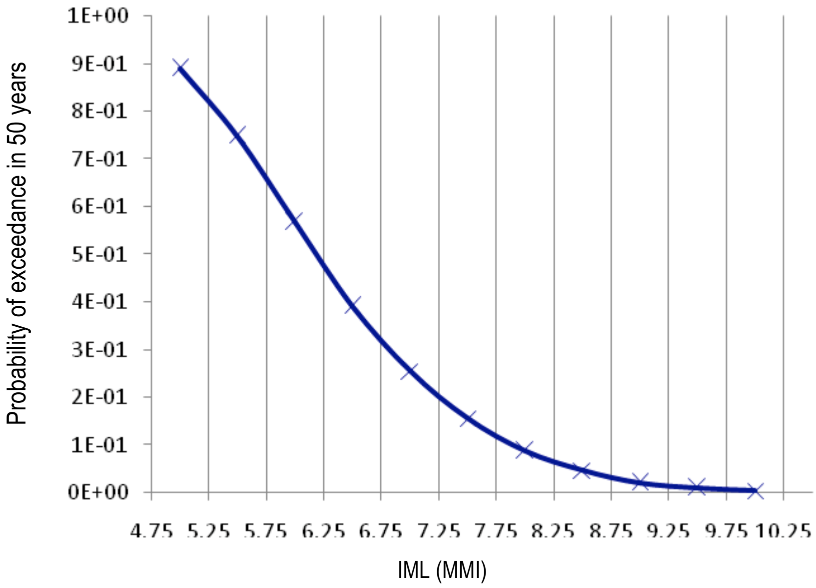


Figure 15.18 Hazard curve of probability of exceedance versus intensity measure levels.

Considering what has been established above, the probability of occurrence can be easily computed using the following formula:

$$PO(IML_n) = PE\left[IML_n + \left(\frac{IML_{n+1} - IML_n}{2}\right)\right] - PE\left[IML_n - \left(\frac{IML_n - IML_{n-1}}{2}\right)\right] \tag{15.4}$$

For the time being, this list of intensity measure levels for which the probabilities of exceedance/occurrence have been computed are being stored in the second row of the seismic hazard file.

An example explaining the necessity in attributing a different code to relate every site with the procedure used to compute MMI will now be presented. A region of Western United States has been assumed where different GMPE's might be used.

Figure 15.19 Western region of United States (California and Nevada).

Let us assume that the following grid can describe the GMPE's and CE's used for each site for this region.

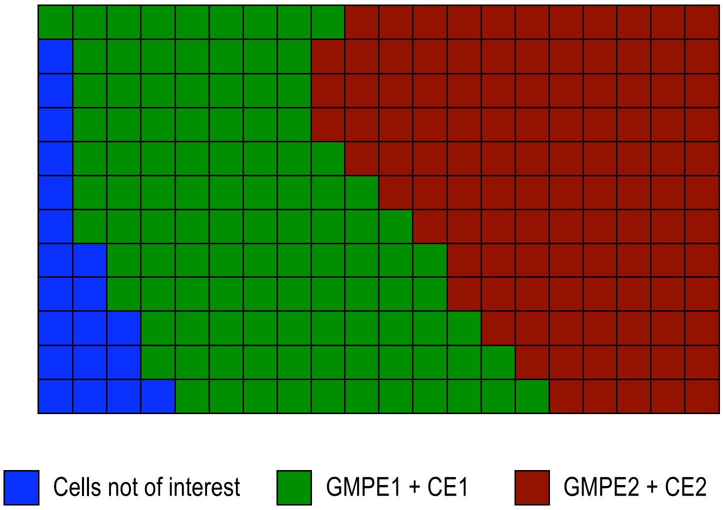


Figure 15.20 Gridded representation of the method used to compute the intensity levels.

Assuming this distribution, it is possible to set a code for every combination and insert this information on the header of the hazard file. The following table presents an example of this combination system:

Table 15.3 Codes and respective methods to compute the intensity levels.

Code	Information for the header	Description
0	No_data	There might be certain zones within the region of interest such as oceans, rivers, lakes or even deserts where there is no relevance in computing the hazard. These sites can be flagged with a 0 so the risk engine can skip these cells.
1	CY2008+WEA1999	For the GMPE Chiou and Youngs [2008] was used and for the conversion equation Wald *et al*. [1999] was used.
2	AB2006+AK2007	For the GMPE Atkinson and Boore [2006] was used and for the conversion equation Atkinson and Kaka [2007] was used.

Note that the above codes to describe the GMPE are the ones proposed by Porter and Scawthorn [2009]. For each site there is a "combination code" and a list of probabilities. This set of values will be contained in the hazard file in a gridded format. This way, there is no need to specify the latitude and longitude of every site where the hazard curve was computed, decreasing the file size considerably. To make this gridded format possible to use, it is necessary to add an extra line to the header defining the following parameters:

- Number of columns (ncol)– Integer number that specifies the number of columns of the grid;
- Number of rows (nrow) – Integer number that specifies the number of rows of the grid;
- Longitude (lon) – Value of the longitude of the bottom left cell in decimal degrees;
- Latitude (lat) – Value of the latitude of the bottom left cell in decimal degrees;
- Cell size (csize) – Width of the cells in decimal degrees.

This attempt to keep track of the formulae used to compute the ground motion at every site might not be possible in some regions since the computed intensity might be the outcome of the contribution of several ground-motion prediction equations. It is likely that only in zones classified as “Stable Continental Region” (e.g. east Canada), will one ground-motion prediction equation be used.

Regarding which conversion equations should be used, within GEM1 a study was carried out by Cua *et al.* [2010] where the globe was divided in three zones according to their tectonic activity, and specific equations are proposed for each region. However, the results within this report were computed from ground-motion fields and hazard curves that used only Wald *et al.* [1999] for the conversion of the ground motion into macroseismic intensity.

At the moment, the gridded region has a spatial resolution of 0.1 decimal degrees. The following figure presents an example of a hazard input for classical PSHA-based loss assessment:

```
50      MMI     Mean    UCERF2
6.0 7.0 8.0 9.0 10.0
0 No_Data       1 CY2008+WA01999        2 AB2006+AK2007
48      30      27.225  41.575  0.1
0       0       1 0.20747 0.18165 0.13784 0.06782 0.02568       1 0.14532 0.07906 0.05256 0.02544 0.00528
0       0       1 0.27247 0.14815 0.10173 0.03852 0.00990       2 0.20697 0.12140 0.11702 0.06194 0.03111
0       1 0.20894 0.21798 0.16408 0.08184 0.03081       2 0.18165 0.09877 0.06782 0.02568 0.00660       2
0       1 0.16348 0.08889 0.06138 0.02112 0.00594       2 0.19982 0.10865 0.07460 0.02825 0.00726       2
```

Figure 15.21 File format proposed to store hazard curves.

- First row: Time span, IMT, Type of curve, ERF;
- Second row: List of the intensity measure levels;
- Third row: Procedure used to compute the intensities. Every combination is identified by an integer followed by the code for the GMPE and the code for the CE;
- Fourth row: ncol, nrow, lon, lat, csize;

Remaining rows: Every set of parameters for each cell starts with the code of the procedure used to compute the intensity level (if the risk engine is supposed to skip this cell, then 0 should be inserted). Following this code, the list of probabilities should be inserted.

15.6 Vulnerability Module

In this initial phase, the vulnerability model is being represented in a discrete way, which means, for each vulnerability function a list of discrete intensity measure levels, with associated mean loss ratios (LR) and respective coefficients of variation (CV) are required. Currently, the GEM1 risk engine is only using intensity-based methodologies to compute human and economic losses. Within this section the two vulnerability models that are being used are described as well as how the uncertainty is being considered and lastly the proposed file format to store the vulnerability curves is explained.

15.6.1 PAGER Global Vulnerability Model

This vulnerability model was developed by the PAGER group at USGS [Jaiswal *et al.,* 2009], and further developed by PAGER in GEM1, and it is an empirical global vulnerability model capable of predicting human losses. In order to develop this model, earthquake mortality rates for more than 4,500 worldwide earthquakes that occurred since 1973 were studied and a set of vulnerability functions for specific countries or regions were built. This earthquake fatality rate can be defined as the ratio of the total number of shaking-related fatalities to the total population exposed at a given shaking intensity, using Modified Mercalli as the intensity measure type.

The PAGER collection of vulnerability functions that return a fatality rate for a given intensity measure level, can be expressed in terms of a two-parameter lognormal cumulative distribution. For countries where at least 4 fatal seismic events have occurred (earthquakes where casualties occurred) a specific vulnerability function was computed. However, only a few countries have experienced > 4 fatal earthquakes since 1973 and therefore, it was necessary to aggregate some countries in regions in order to have enough fatal earthquakes for this catalogue period. For this reason, a new global regionalization scheme was proposed based on the likelihood of suffering similar losses in future earthquakes. This spatial organization was done based primarily on geography, building inventory and social-economic similarities. The indicators that were used to make these choices were:

- Human development index – This index combines normalized measures of life expectancy, literacy, education and gross domestic product per capita worldwide and it can be used to rank countries by "human development" level. These socio-economic conditions can significantly affect building construction quality and maintenance practices. For instance, in well developed countries such as United States, Japan or New Zealand, significant improvements on their building stock with consistent maintenance and retrofitting of weaker structures proved to reduce considerably building collapses, and consequently, the amount of human losses. However, in developing countries such as Pakistan, China or India, poorer socio-economic conditions tend to affect negatively the way people build and maintain their buildings and hence, a large number of building collapses due to earthquakes are common in these countries.
- Climate classification – For countries with similar human development index but different climate classification, further consideration might be needed. Local climate strongly influences architectural characteristics like wall thickness, roof height, number of openings and construction materials which tap into natural resources such as heat and sun light. For instance, in hot climates, buildings tend to have their outer walls thicker in order to keep the strong heat outside, while in cold areas, it is the inner walls that tend to be thicker in order to insulate and keep the heat in.

Figure 15.22 presents the regionalization proposed by this methodology. Countries with a specific vulnerability function have their own colour and are represented by their ISO code, while groups of countries are shown with the same colours and a particular code for that region.

According to the proposed regionalization, the vulnerability model is composed of 27 functions for particular countries and 23 functions for groups of countries. These expression which give fatality rate (ν) as a function of shaking intensity (S) are expressed as a lognormal distribution function as follows:

$$\nu(S) = \phi\left[\frac{1}{\beta}\ln\left(\frac{S}{\theta}\right)\right] \tag{15.5}$$

Where ϕ is the standard normal cumulative distribution function, S is the discrete value of the intensity measure level and β and θ are parameters of the distribution. This vulnerability model is bounded between intensity V to X and it is expressed in numerical values with 0.5 increments. The two parameters of the distribution were computed in a way that the residual error would be minimized for large and smaller earthquakes simultaneously. Several approaches were studied to compute this error, since it was important to maintain a balance between the computed residual errors when considering small and large earthquakes. In order to test the goodness of fit of the way the residual error was being computed, a special case of the Kolmogorov-Smirnov test (Lilliefors) was used, and it was concluded that the chosen approach was reasonable to model this error.

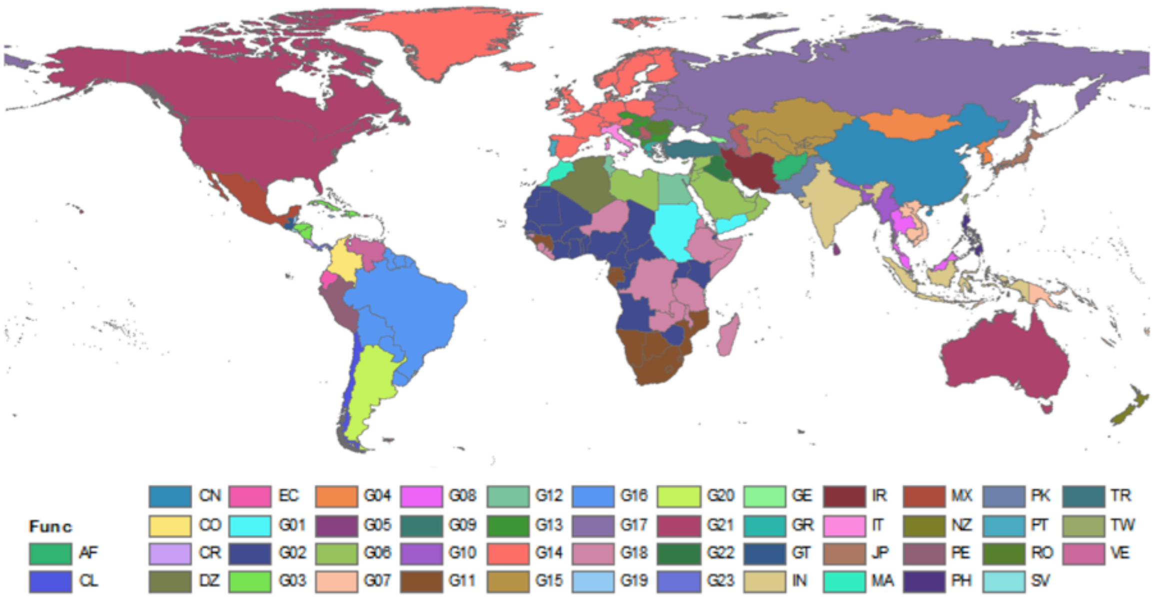


Figure 15.22 Regionalization scheme proposed for the PAGER global vulnerability model.

In order to represent the vulnerability functions in a discrete way, Equation (15.5) and the parameters computed for each region were used and the fatality rates were calculated for a discrete number of intensity measure levels. Figure 15.23 presents the vulnerability functions of the eight countries with the highest fatality rate.

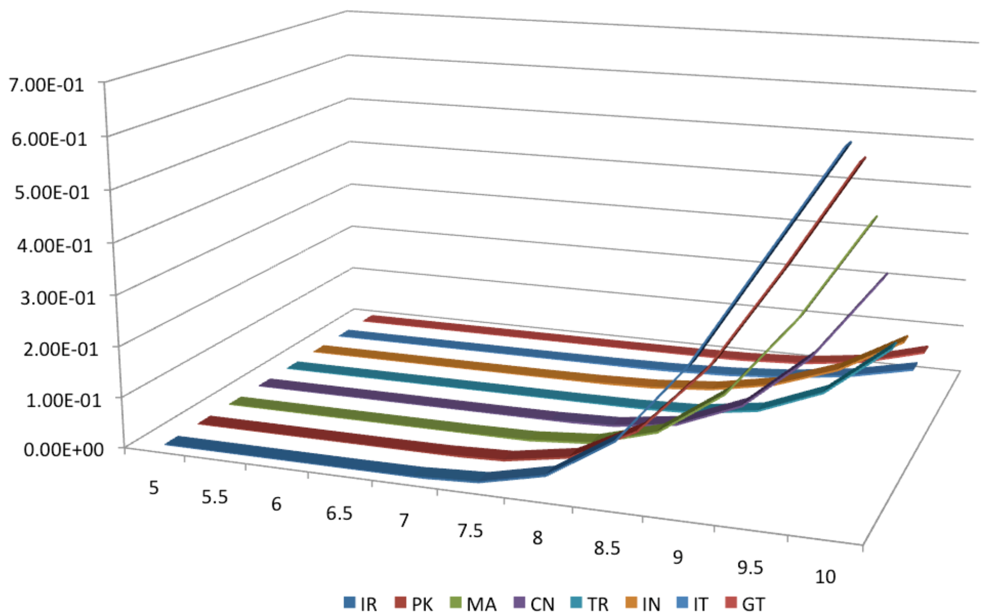


Figure 15.23 PAGER vulnerability functions for the eight countries with highest fatality rate(where IR – Iran; PK – Pakistan; MA – Morocco; CN – China; TR – Turkey; IN – Indonesia; IT – Italy; GT – Guatemala).

By the end of the GEM1 project, another version (2.0) of this vulnerability model was released by PAGER. This new version offers a more detailed regionalization of the world (especially in Africa and South America) based on other parameters such as 2008/2009 gross domestic product (GDP) and losses due to important earthquakes that occurred in the meantime (e.g. Haiti 14/01/2010). The way the total population exposed to each intensity measure level was calculated at high intensities was also slightly changed. For high intensities, the population exposed to intensities of IX and X were aggregated and assumed to be affected only by intensity IX.

The new parameters required to compute the vulnerability curves were recently provided to the GEM1 risk team though it is noted that the plots shown in this report (which are for illustrative purposes only) were produced with version 1.0 of the PAGER vulnerability model.

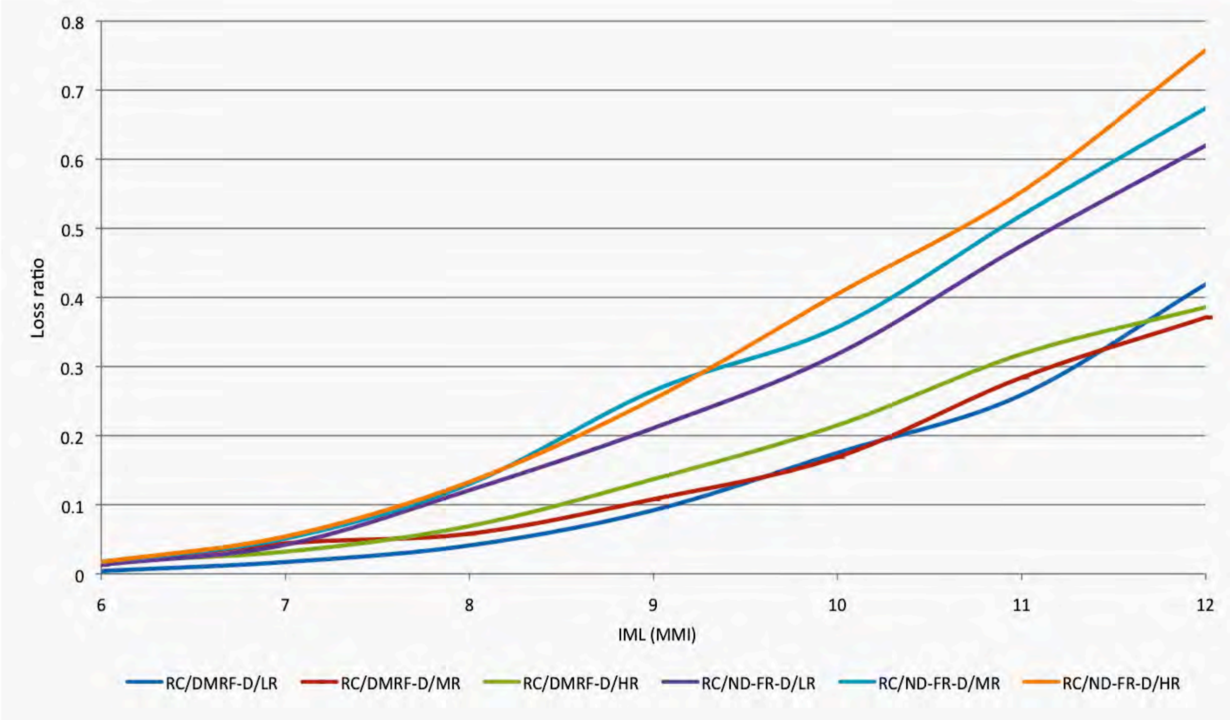
15.6.2 ATC-13 Building Vulnerability Model

This vulnerability model is part of a report produced by Applied Technology Council (funded by the Federal Emergency Management Agency) in 1985 entitled “Earthquake Damage Evaluation Data for California”. The following features can be found in this document:

- Expert-opinion on ground motion - damage relationships, presented in the form of damage probability matrices, for 78 classes of structures, including buildings (40 classes) and lifeline structures (electrical, water, sanitary sewer, natural gas and telephone components and systems);
- Methods and data for estimating damage and loss resulting from collateral hazards, including fault rupture, ground failure, inundation, and fire;
- Estimates of the time required to restore damaged buildings and lifeline structures to their pre-earthquake functionality;
- Methodology and data for estimating deaths and injuries.

Contrarily to what was previously seen, this vulnerability model was not meant to be applied globally but only in California since it was developed considering earthquake historical data and building characteristics from that state. Regarding the building typology, it supports 40 types of building that are organized according to attributes such as structural behaviour, construction materials and height of the buildings. No distinction is made regarding the type of code or existence of seismic retrofit. This model was used in tests of the GEM1 risk engine but large scale analyses were not carried out as these functions are not valid outside of the US. Should an intensity-based vulnerability model for buildings exist with world coverage and a wide list of building typologies, it could be easily incorporated in the current GEM1 Risk engine.

This vulnerability model is defined at a discrete number of intensity measure levels, from 6 to 12 with increments of 1 and using Modified Mercalli Intensity. For each IML, a loss ratio and corresponding coefficient of variation is defined. This loss ratio is the relationship between the cost of repair over the cost of replacement. A lognormal distribution of the loss ratio for each IML was assumed. The following figure presents the vulnerability function for moment-resisting concrete frame buildings based on ATC-13.



RC/DMRF-D/LR	Moment-resisting non-ductile concrete frame low rise	RC/ND-FR-D/LR	Moment-resisting ductile concrete frame low rise
RC/DMRF-D/MR	Moment-resisting non-ductile concrete frame med rise	RC/ND-FR-D/MR	Moment-resisting ductile concrete frame med rise
RC/DMRF-D/HR	Moment-resisting non-ductile concrete frame high rise	RC/ND-FR-D/HR	Moment-resisting ductile concrete frame high rise

Figure 15.24 ATC – 13 vulnerability functions (mean loss ratios) for moment-resisting concrete frame buildings.

15.6.3 Vulnerability Model File Format

The way information is being stored slightly differs between each vulnerability model. There is a common file that is used by both methodologies where all of the parameters of every function are stored, and a second file required only for the PAGER global vulnerability model. This second file contains the information that allows the risk engine to relate every country to its designated vulnerability function. The development of the proposed file formats was carried out taking into consideration the report produced by Porter and Scawthorn [2009].

Starting with the first file, here are the parameters:

- Asset type – It is fundamental to create this file format in a way that fits not just for buildings but also population, contents and other assets and, therefore, it is necessary to specify the assets that can be related with the collection of vulnerability functions. This parameter can represent the following asset types:
  - Population;
  - Buildings;
  - Contents;
  - Occupants;
  - Economic output;
- Loss type – This parameter is obviously related with the type of asset and reflects the loss measure produced by the vulnerability functions contained within the file. This parameter can assume the following types:
  - Fatality;
  - Collapse;

- Repair cost;
- Cost of lost of use;

Further loss measures are being considered such as those presented by Porter and Scawthorn [2009]. These two parameters (Asset type and Loss type) will be featured in the first row of the file and hence, it is assumed that all the vulnerability functions within the file are applicable to the same asset type and will produce the same loss. The next parameters are specific attributes of each vulnerability function:

- ID – This is a unique code within the file and it is used to identify the respective function inside of the risk engine. It does not need to be an integer value, it can be a string such as G01 or ITALY02;
- IMT – The intensity measure type that is required by the vulnerability model and represented according to the recommendations of Porter and Scawthorn [2009].
- Type of distribution – For each intensity measure level, a mean loss ratio and respective coefficient of variation are defined, and in order to take into consideration this uncertainty, it is necessary to define the type of probabilistic distribution of the values. At the moment the risk engine supports three different distributions, which are represented by the following code:
  - Uniform – UN;
  - Normal – NO;
  - Lognormal – LN.
- IML, LR, and COV – Following the previous parameters, a set of values that define the function and the uncertainty at each point need to be inserted. These values are always introduced in the following sequence: IML -> LR -> COV. This sequence is repeated until the whole function has been introduced.

The next figure presents how a single vulnerability function has been stored following the proposed file format:

```
Population  Deaths
US045   MMI   LN    5    0.001  1.3   5.5   0.002  1.1   6    0.005  1.0
6.5     0.007 0.9   7    0.011  0.8   7.5   0.019  0.7   8    0.038  0.6
8.5     0.093 0.6   9    0.120  0.5
```

Figure 15.25 File format proposed to store the vulnerability functions.

The second file is only required for the PAGER vulnerability model and it relates every region in the world with a vulnerability function as previously described. Using the Global Administrative Unit Layer database, the risk calculator retrieves an administrative code for every grid cell in the world (spatial resolution of 30 arc-second) and then, it uses this second file to relate every code with a vulnerability function. The following parameters are being stored in this file:

- Administrative code – This code is being represented by the ISO number of each country (integer with 3 numbers) in particular for the PAGER vulnerability model. However, other codes were also developed that will allow one to relate vulnerability functions not just with countries but also with smaller administrative levels.
- Vulnerability code –This code identifies the vulnerability function and it is represented as described in the previous file format (first code of every row).
- Region Name – This parameter is a string that contains information about the name of the region where the cell is located.

The following figure presents an example of this file:

004	AF	Afghanistan
008	G12	Albania
012	DZ	Algeria
016	G07	American Samoa
020	G13	Andorra
024	G08	Angola
660	G05	Anguilla
010	G07	Antarctica
028	G05	Antigua And Barbuda
032	G17	Argentina
051	G14	Armenia
533	G05	Aruba
036	G18	Australia

Figure 15.26 File format proposed to relate every vulnerability function to its region.

15.7 Output

15.7.1 Background Research

The GEM1 Risk team has carried an extensive research into open source web-based solutions which could be used for the presentation of risk results in a web-based user interface. Other portals and similar initiatives from which ideas and inspiration have been derived include:

- OpenGeo's GeoNode, a Web 2.0 Portal and Spatial Data Infrastructure [31, 32].
- FAO's Global Information and Early Warning System (GIEWS) Workstation Portal [33].
- French National Geographical Institute's Geoportail [34].
- Earth Atlas, 2D and 3D [35].
- OneGeology Portal [36].

Some of the features that the group explored further for the risk user interface included:

- A workspace that auto fits to the browser, and does not require a scroll bar;
- A user-friendly layout, with resizable/collapsible side panels for extended map area and accordion menu items and tabs in the side panel for increased user-input area;
- Pop-up windows for maps or user-defined selections/input etc.;
- Layers with scale-dependent visibility (which turn on and off as one zooms to different levels of resolution) (e.g. see GeoPortail);
- Use of tabs to multiply the number of workspaces;
- Possibility to export maps (as .pdf, .jpg, .kml etc.), plots (as .jpg) and data;
- Possibility to view and even modify data in tables on the interface;
- Both 2D and 3D views of the world (e.g. Earth Atlas);
- "Architecture of Participation" through Web 2.0 features (e.g. rating, commenting, tagging) and features for users to style their own maps (e.g. see GeoNode).

One of the open source tools that many of these user interfaces use to obtain the aforementioned user-oriented features is Ext JS, a cross-browser JavaScript library for rapidly building rich internet applications. Other open source libraries and tools that have been found in these portals and some of which have been used in the demo interface for the risk results include:

- OpenLayers: a pure JavaScript library for displaying map data in most modern web browsers, with no server-side dependencies. OpenLayers implements a (still-developing) JavaScript API for building rich web-based geographic applications, similar to the Google Maps and MSN Virtual Earth APIs, but OpenLayers is a Free Software, developed for and by the Open Source software community.

- GeoServer: GeoServer is an open source software server written in Java that allows users to share and edit geospatial data. Designed for interoperability, it publishes data from any major spatial data source using open standards.
- MapServer: an open source platform for publishing spatial data and interactive mapping applications to the web.
- MapServer and GeoServer publish data in WMS, WFS and WCS formats. WMS (Web Map Service) is a standard protocol for serving georeferenced map images over the internet that are generated by a map server using data from a GIS database). WFS (Web Feature Service) is a standard interface allowing requests for geographical features (which can be thought of as the source code behind the map, not just the image which is obtained with WMS) across the web using platform independent calls. Web Coverage Service Interface Standard (WCS) defines a standard interface and operations that enables interoperable access to geospatial "coverages". The term "grid coverages" typically refers to content such as satellite images, digital aerial photos, digital elevation data, and other phenomena represented by values at each measurement point.
- GeoExt: a combination of the geospatial knowhow of OpenLayers with the user interface ability of Ext JS to allow powerful desktop style GIS applications to be built on the web with JavaScript.
- Ext JS extension for Google Earth API: an open source project integrating Ext JS with the Google Earth API [37].
- MySQL or PostgreSQL: open source databases.

15.7.2 GEM1 Risk Demo Interface

The risk demo interface prepared by the GEM1 Risk group has been developed as an open source tool. The tool and its capabilities were developed considering the scope of the project, flexibility, ease of use and expandability.

Client Side:

The GEM1 Risk mapping interface uses the ExtJS JavaScript library. This library uses DOM (Document Object Model) and Ajaxs (asynchronous JavaScript and XML). DOM and Ajaxs are used to allow JavaScript to dynamically interact with web applications. This library is very well suited for this type of interface as it offers many flexible user interfaces such as:

- Accordion side panels (regions), which allow for multiple 'layers' on information in a compact format
- Grids (tables) with both read only and edit modes, storable data
- List boxes and combo-boxes
- Radio and checkbox controls
- Tree controls
- Tab panels
- Charts

In addition to the ExtJS feature listed above, the library also offers many extensions. The extensions deployed in this application are GeoEXT, ExtJS Google Earth API, whilst AeroSQL was also considered (this allows databases to be manipulated from the interface). GeoExt allows for an ExtJS interface with OpenLayers, and it provides a number of geographical interfaces. GeoExt can integrate into the tool a WMS layer store interface with GeoServer, allowing the user to interact with multiple layers in a single session. The ExtJS Google Earth API is used to view and interact with Google Earth in an ExtJS module, and allows for Google Earth layer control, KML & KMZ file import, and location search. A future development of the mapping tool could also use the Thematic Mapping API, which is a JavaScript thematic mapping library that can be used to display data is a 3D KML format.

Server Side:

ExtJS is very flexible in connecting with external databases. ExtJS is compatible with XML, JSON, ColdFusion and other interchange technologies. GeoExt is designed specifically to interact with data served from GeoServer. GeoServer is used as the geographic data server for sharing and editing data in the mapping tool. GeoServer's research and development documentation discusses plans to incorporate in future versions of GeoServer a feature that will allow synchronizing metadata between GeoServer and GeoNetwork. This type of catalog service web solution would allow for greater flexibility in cross agency cooperation. GeoWebCache has been used to speed up the performance of the interface. PostgreSQL/PostGIS may also be used as database servers, but these have not been used in the current risk demo interface as a quick-fix solution has currently been sought to show how data and plots might be shown in the future. The combination of these tools was selected based on their capabilities, their compatibility with each other, and usability.

The tools used in the user interface and those required behind the scenes to publish the results and data on the interface are shown in Figure 15.27. The user interface and application server tools have been installed and used in the current demo, whereas the database servers have not been set up, but the structure would be fully compatible with open source solutions such as PostgreSQL/PostGIS.

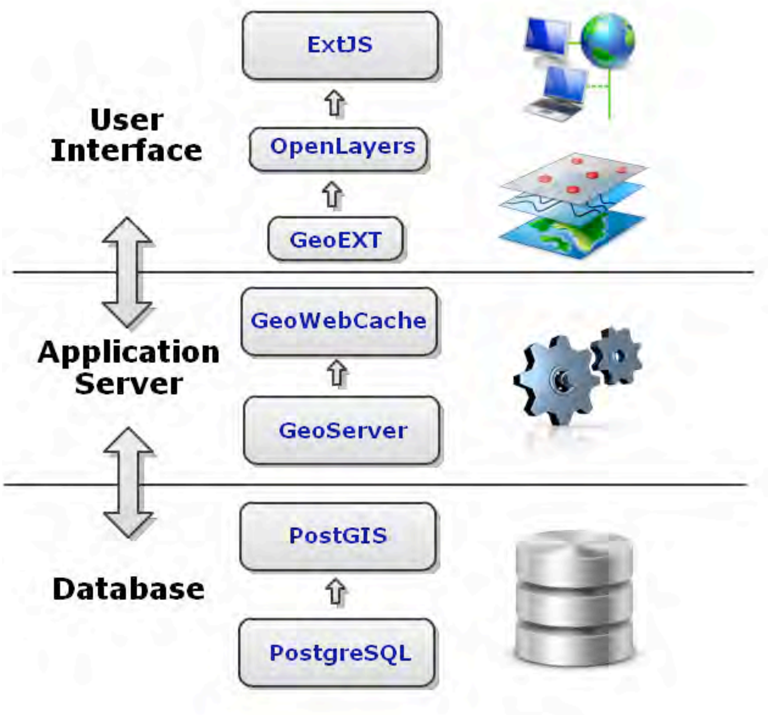


Figure 15.27 GEM1 Risk demo interface structure and tools and illustration of how they could be linked with a database in the future

15.7.3 Web Interface: Use Cases

The GEM1 Risk demo interface can be thought of as a sandbox of tools that the GEM1 Risk team has been "playing" with in order to explore the possibilities for the presentation and calculation of risk results. This interface is completely static, and does not have any dynamic functionality e.g. the selection buttons do not work, there are just there for demonstration purposes. However, by adding such buttons to the interface, the Risk team has been able to trace the work process for different users, to identify the best way of presenting such information, and ways of optimising space and so on. It has also helped to understand the different resolutions which should be used for plotting risk maps at a global scale.

A number of users/consumers have been considered within the demo interface; it is expected that there would be three main types of risk users/consumers:

- Basic consumers who want to view maps of particular scenario events or global loss maps (e.g. mean loss within a given time span) by simply adding layers from the WMS Layer Store to the layer menu item.

- Expert/power users who want to view the aforementioned results, in addition to loss curves, and can make decisions on the type of input and output. For example, they would choose between different input hazard models, vulnerability models, exposure models, and different output fractiles (e.g. 16 percentile, median).
- Expert/power users who want to be able to run their own risk calculations with their own input data.

The following screenshots show some possibilities that have so far been envisaged.

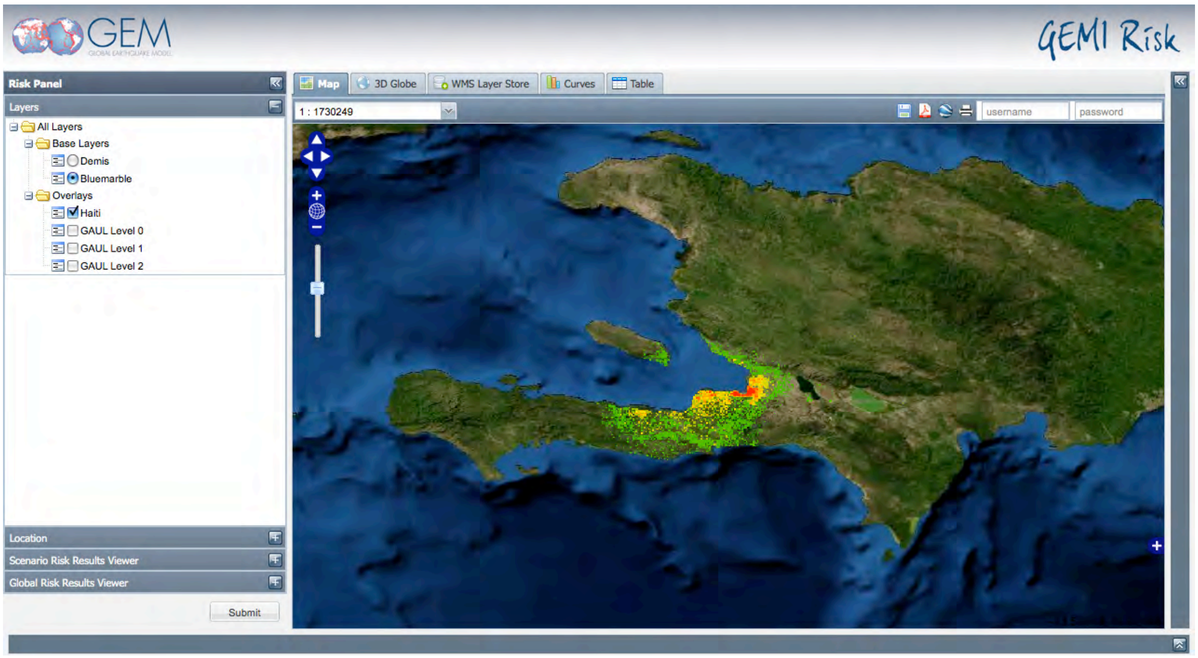


Figure 15.28 Basic Consumers: Visualise loss (ratio) map layers in 2D

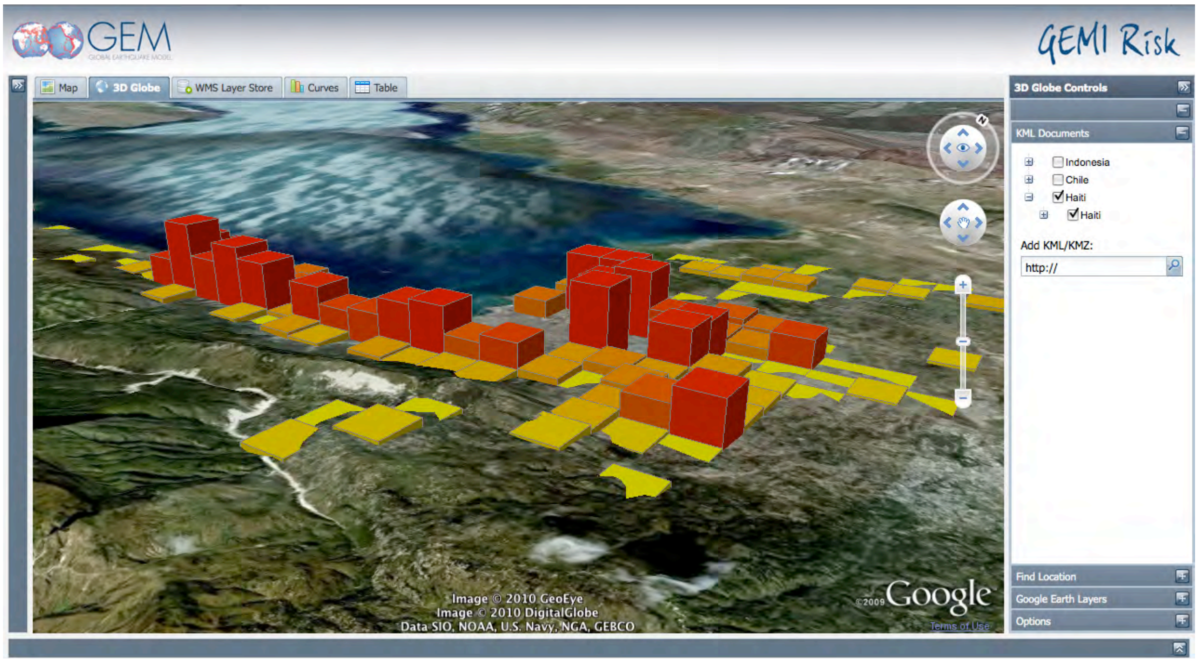


Figure 15.29 Basic Consumers: Visualise loss (ratio) map layers in 3D

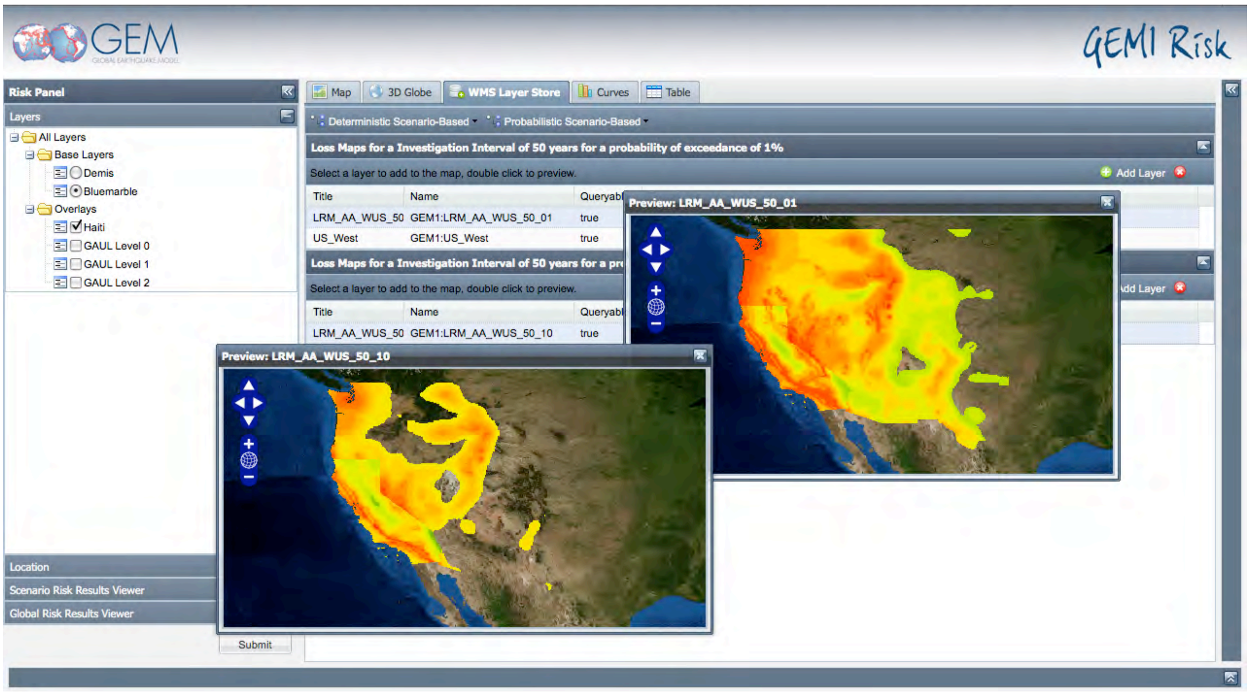


Figure 15.30 Basic Consumers: Preview and compare loss (ratio) map layers in 2D

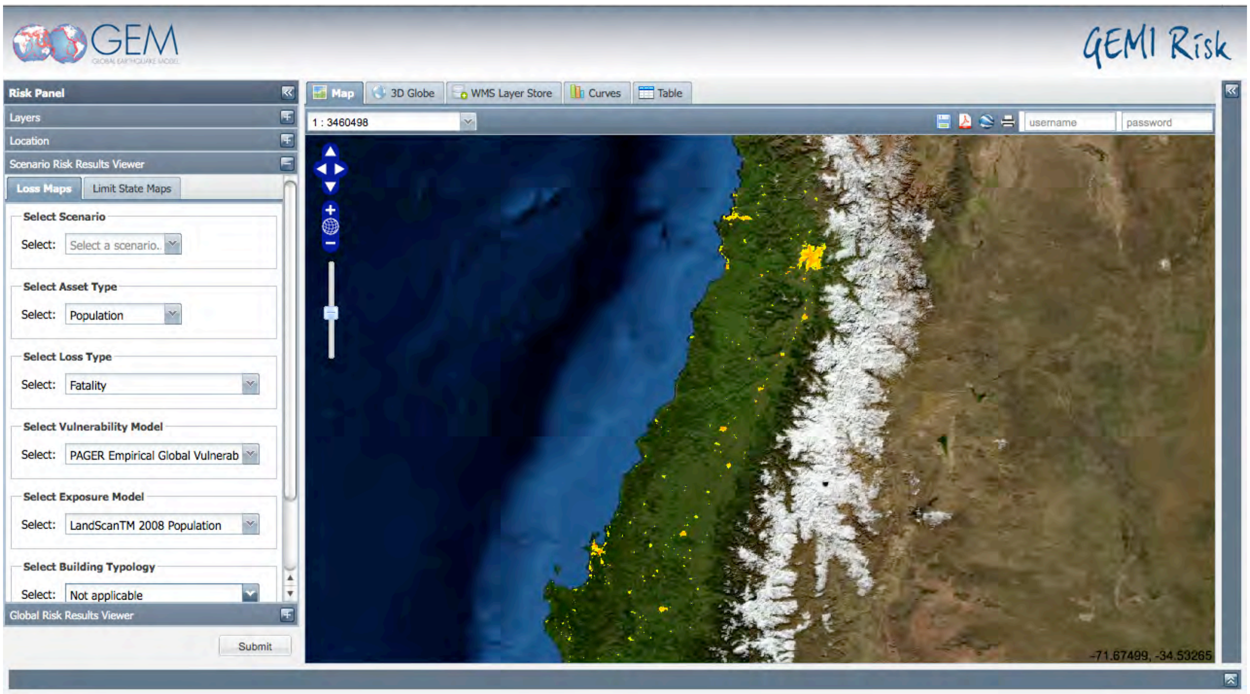


Figure 15.31 Expert Users: Select and visualise a specific loss (ratio) map layer in 2D

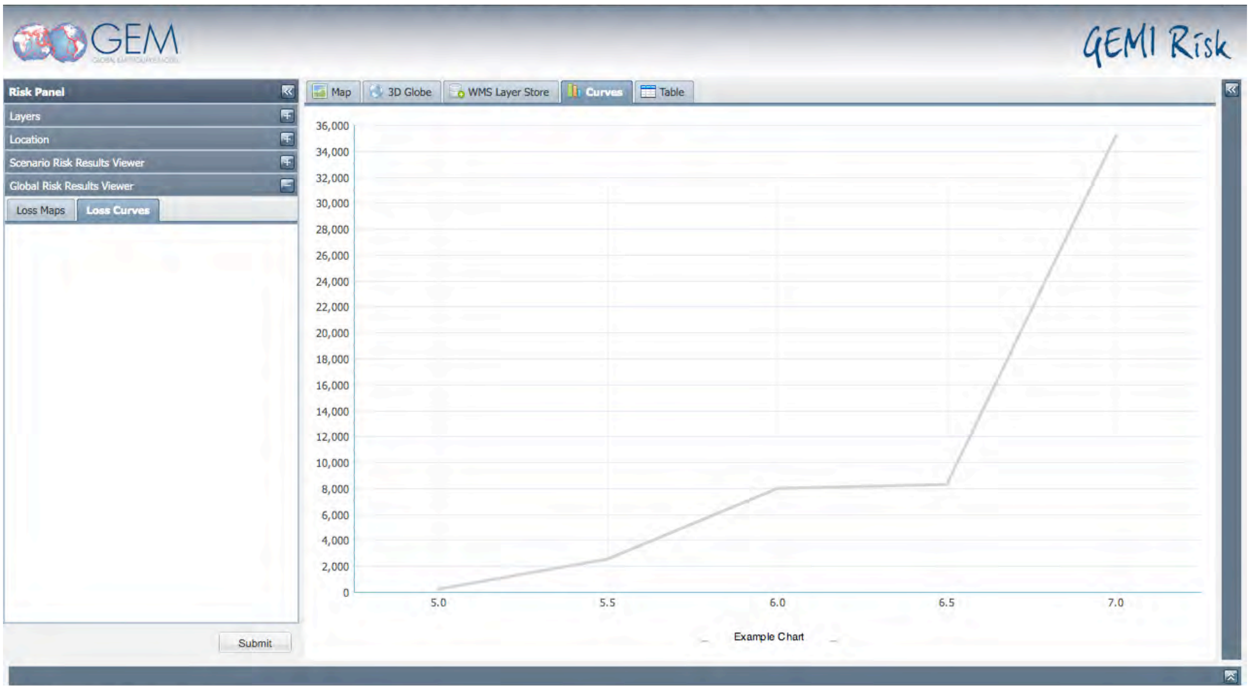


Figure 15.32 Expert Users: Select and visualise a specific curve (note: example curve, could be MMI vs number of fatalities)

16 Demonstration Applications

16.1 Deterministic Event-Based Applications

16.1.1 Human Losses

To demonstrate the results produced by the GEM1 risk engine for a single event, the earthquake of 27th February 2010 that occurred in the region of Chile was chosen. This seismic event had a magnitude of 8.8 M_w and lasted for 90 seconds. The epicentre of the earthquake was offshore from the Maule Region, approximately 11 km southwest of Curanipe and 100 km north-northeast of Chile's second largest city, Concepción.

One of the reasons that influenced choosing this scenario was the fact that large towns were struck with strong ground motions such as Arauco and Coronel with an intensity level of VIII and Talca, Valparaiso and Santiago with an intensity level of VII and it is believed that 80% of Chileans felt the earthquake. The death toll due to ground shaking has been reported to be between 480 and 550 casualties (e.g. http://en.wikipedia.org/wiki/2010_Chile_earthquake).

Since Chile experienced several fatal earthquakes in the past, the PAGER global vulnerability model used in the GEM1 risk engine provides a specific function for this country. The following figure shows the curve computed for Chile. Although this vulnerability model doesn't provide information regarding the uncertainty for each fatality rate, in order to test all the functionalities of the risk engine, it was decided to assume a coefficient of variation of 0.30 in this demonstration application.

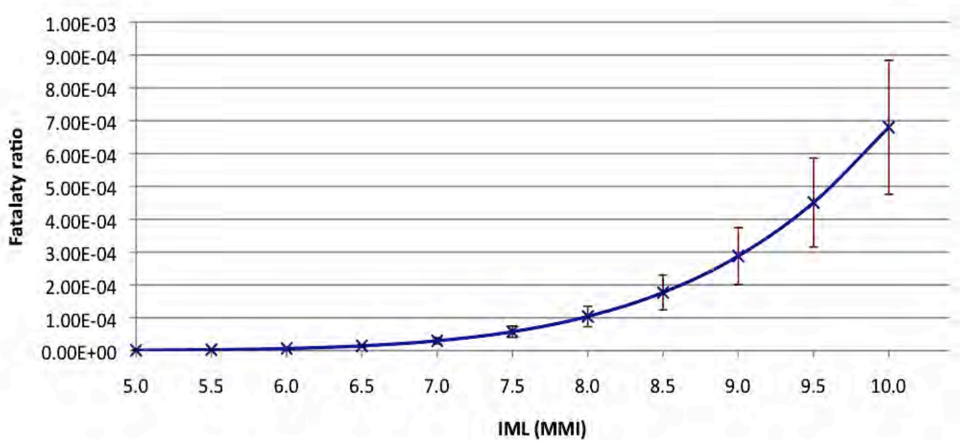


Figure 16.1 Discrete vulnerability function for Chile.

The ground-motion field (in this case a USGS "ShakeMap") with the distribution of the intensity levels was extracted from the archive of the U.S. Geological Survey (alert version 8). The following figures shows the intensity levels felt in the region that are within the vulnerability model (from V to X), the distribution of the fatality ratios, the absolute human losses and finally the standard deviation for both fatality ratios and absolute human losses for each grid cell. The LandScan 2008TM dataset was used for these calculations [38]. Although there was information available regarding the uncertainty of the ground shaking, it was decided to assume no uncertainty on the ground-motion field for this application. A study of the influence that the standard deviation on the ground motion has on the final results is also presented.

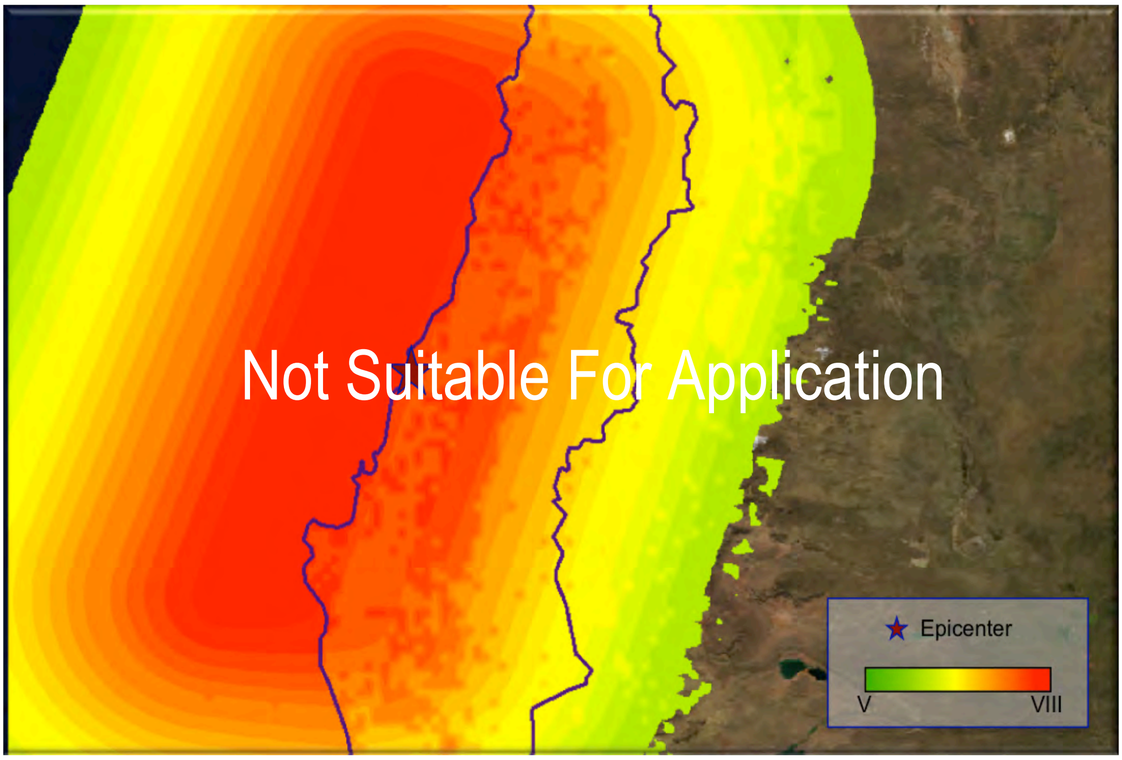


Figure 16.2 Intensity ShakeMap for the earthquake of 27/02/10 in Chile.

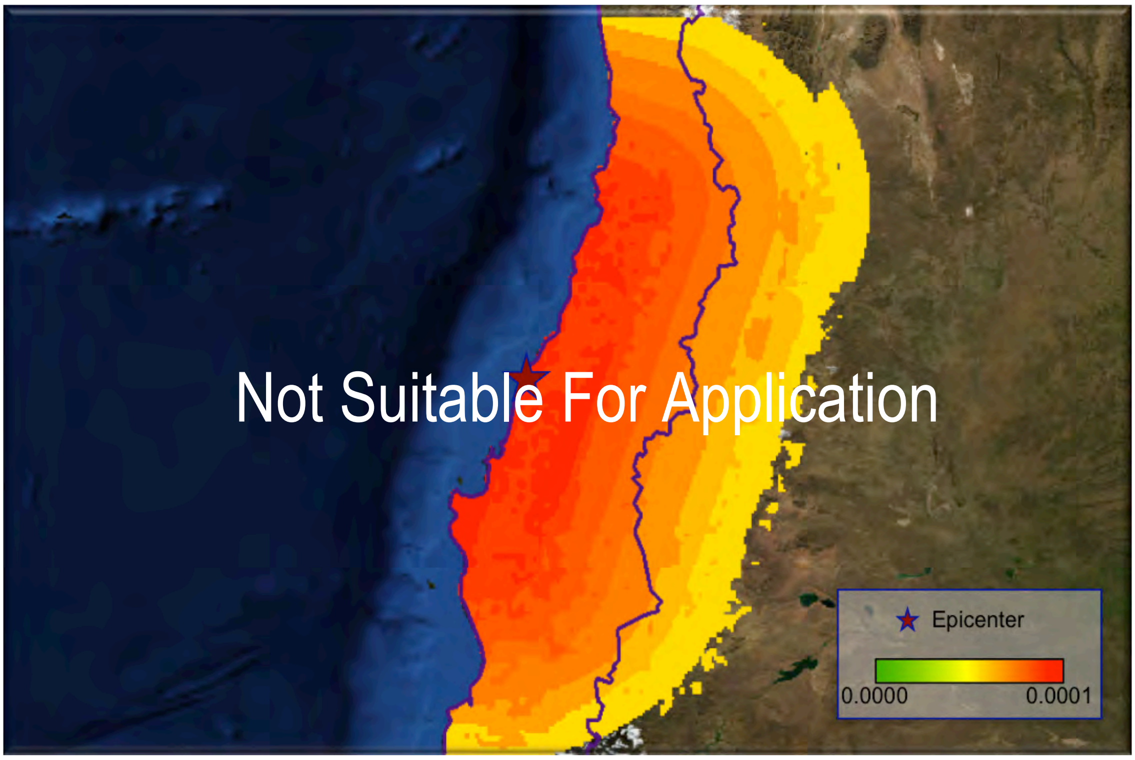


Figure 16.3 Fatality ratio map for the earthquake of 27/02/10 in Chile.

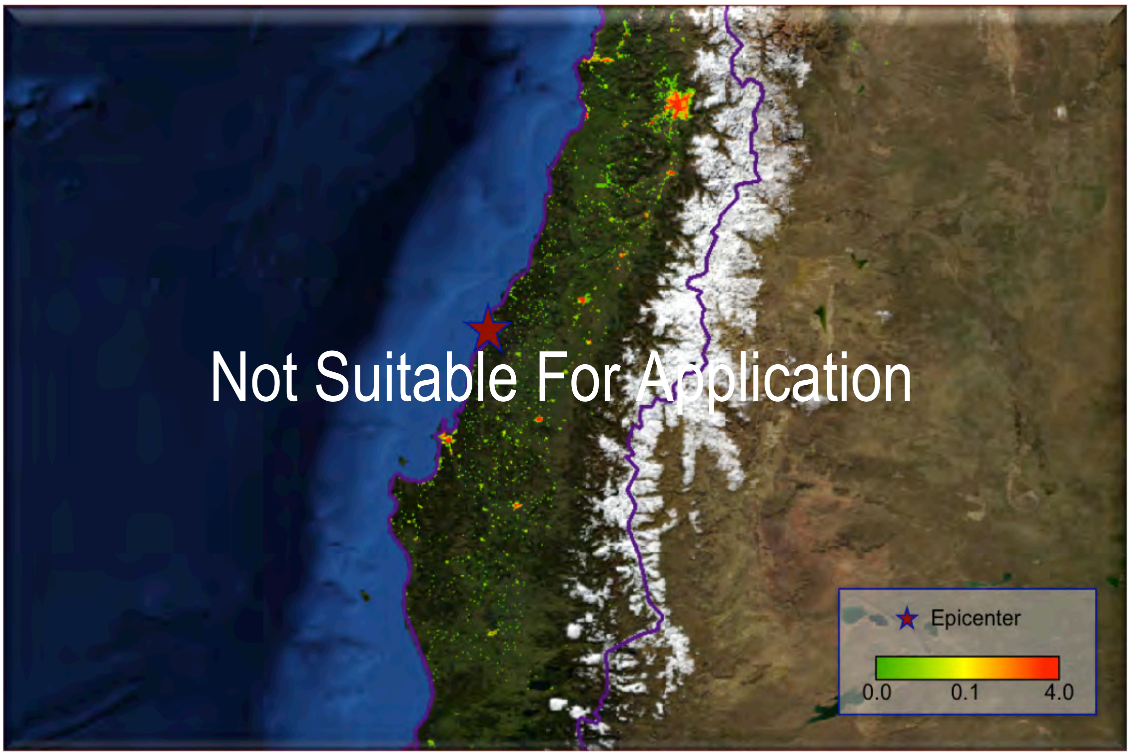


Figure 16.4 Human losses map for the earthquake of 27/02/10 in Chile.

Figure 16.5 Distribution of the standard deviation of the loss ratio for the earthquake of 27/02/10 in Chile.

Figure 16.6 Distribution of the standard deviation of the losses for the earthquake of 27/02/10 in Chile.

The GEM1 risk engine has also the capability of calculating the amount of population exposed to each intensity measure level. The results for this scenario are presented in the following figure.

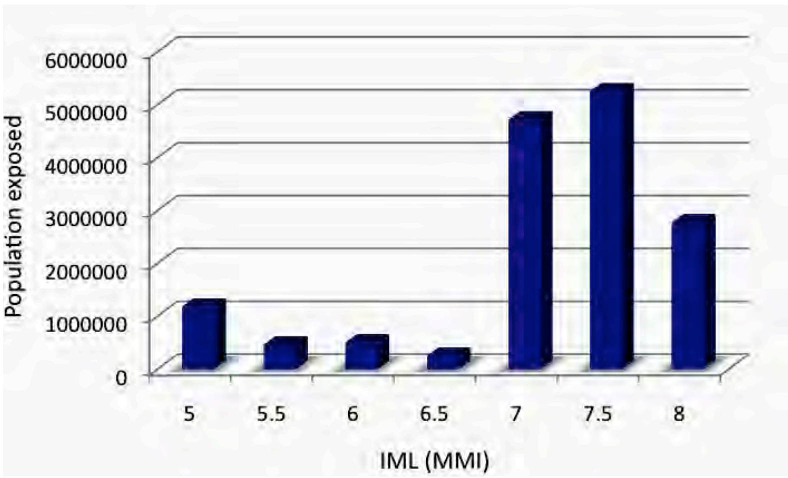


Figure 16.7 Population exposed versus IML for the earthquake of 27/02/10 in Chile.

In order to evaluate the influence of the uncertainty on the ground shaking on the results, the same seismic scenario was run assuming different constant standard deviation (SD) values to the ground-motion fields. The results are presented in the following figure and show the influence that the modelling of uncertainty has on the predicted number of fatalities.

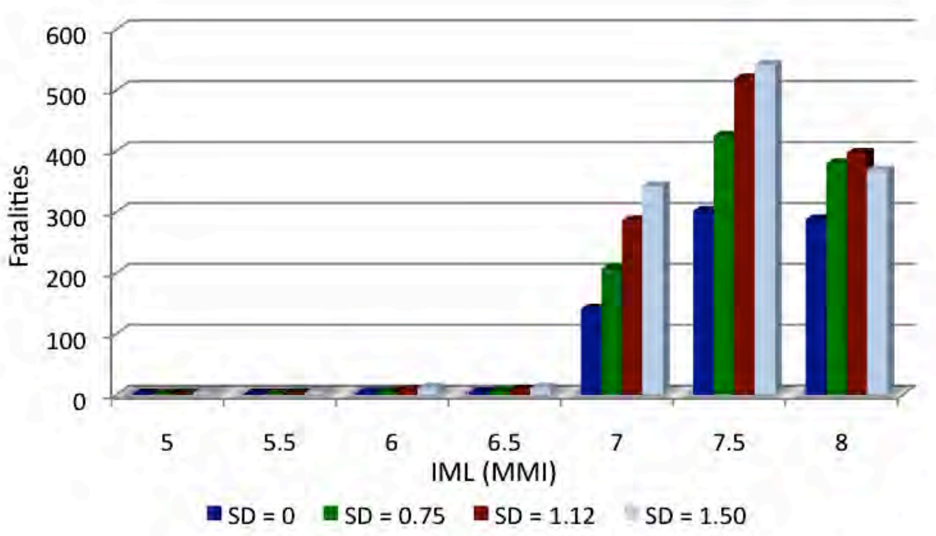


Figure 16.8 Fatalities versus IML for the earthquake of 27/02/10 in Chile with different levels of ground motion uncertainty (SD)

16.2 Classical PSHA-Based Applications

16.2.1 Progress of the Results

In order to perform the classical PSHA-based calculations, the hazard group provided the GEM1 risk group with seismic hazard curves for specific locations of the globe that are created based on the regional tectonic characteristics. These files contain the hazard curves for each grid cell with a spatial resolution of 0.1 decimal degrees and a time span of 50 years. The risk results have been computed according to the seismic hazard coverage of each file. The seismic hazard input model can cover nations or regions depending on the context within which they were developed/created. For regions with strong seismic activity such as Japan, Turkey, Indonesia, New Zealand, Mexico or Iran, it is common that a specific country-based hazard model exists. Figure 16.9 presents the regions where risk results were computed in GEM1.

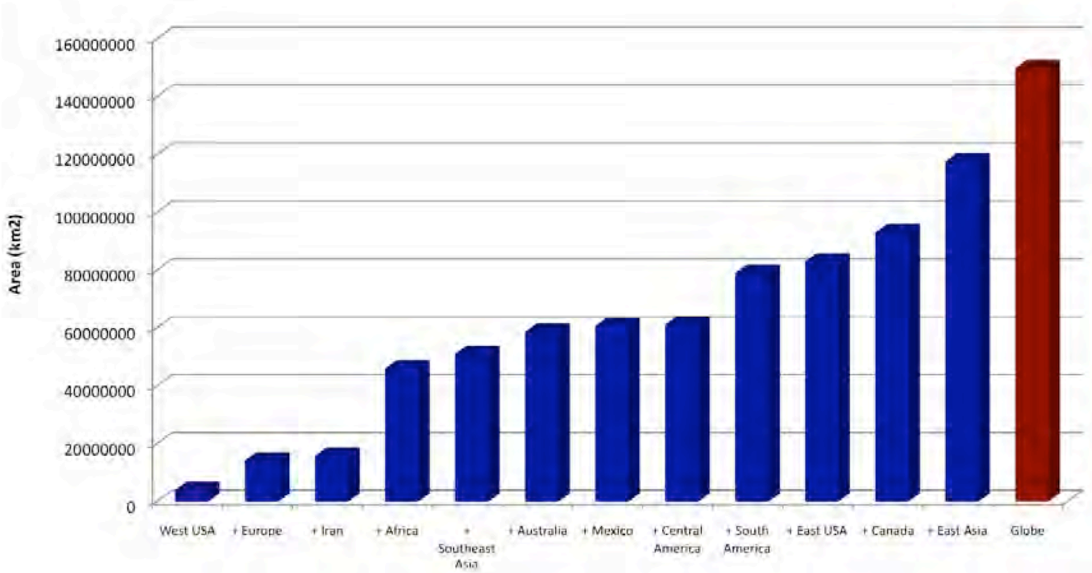


Figure 16.9 Progress of the global seismic risk calculations to in GEM1

Currently the risk results have been computed for approximately 78 % of the globe. For each region, the following outputs have been calculated:

- Loss ratio maps (with probabilities of exceedance of 1%, 2%, 5% and 10% in 50 years);
- Loss maps (with probabilities of exceedance of 1%, 2%, 5% and 10% in 50 years);
- Mean loss ratio (in 50 yrs) map;
- Mean loss (in 50 yrs) map;
- Loss ratio curve for each grid cell (spatial resolution of 30 arc seconds);
- Loss curve for each grid cell (spatial resolution of 30 arc seconds);

16.2.2 Human losses

In order to compute the mean global human losses within a time span of 50 years the PAGER vulnerability model version 1.0 was used together with the LandScan 2008™ dataset [38]. For this exercise, no uncertainty was considered in the vulnerability model. The following figure shows the results for the regions that were considered in GEM1, noting that this is for illustrative purposes only, as a proof-of-concept of the GEM1 global risk engine capabilities.

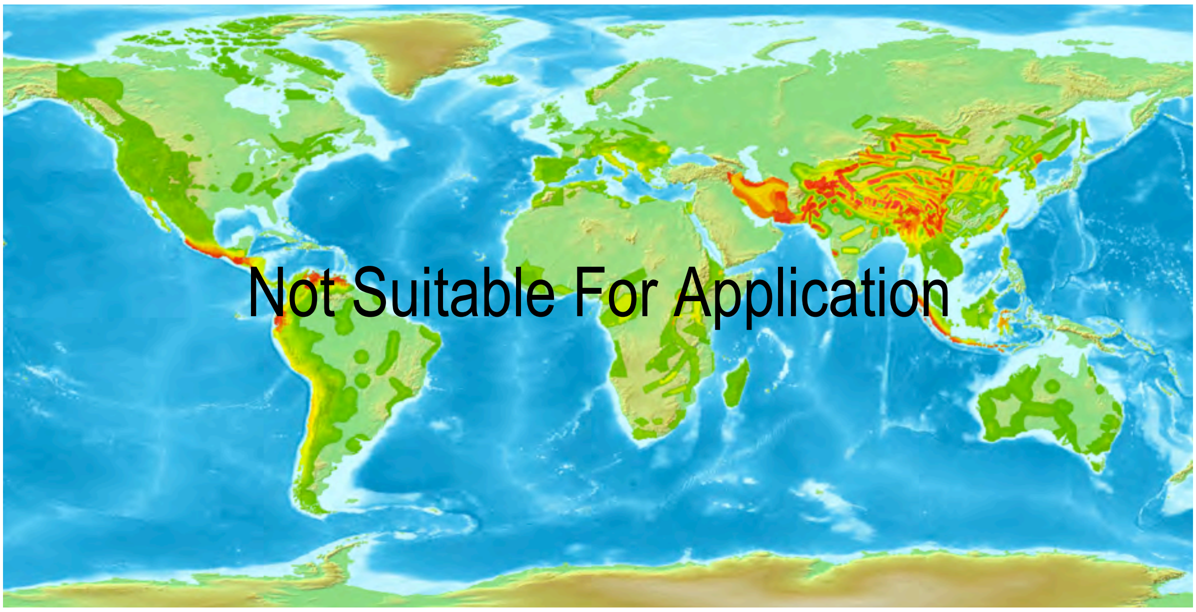


Figure 16.10 Current status of the global map of mean fatality ratio in 50 years

16.3 Conclusions

This chapter has presented the results of some preliminary risk calculations based on both ground-motion fields (deterministic event-based analysis) and hazard curves (classical PSHA-based analysis) using the GEM1 risk engine. Regarding the deterministic event-based calculations, the current GEM1 risk engine is capable of producing loss and loss ratio maps (both mean and standard deviation) considering uncertainty on both the vulnerability model and ground motion input. Further results can be obtained such as total mean losses for each intensity measure level. For the classical PSHA-based calculations, besides conditional loss/loss ratio maps, mean loss and mean loss ratio maps for a given time span (50 years) are also available. The calculator also supports curves of loss and loss ratio versus probability of exceedance.

The global classical PSHA-based loss results presented herein have been produced using the output of the GEM1 hazard engine as input to the risk engine. Other sources of hazard have been used for the deterministic event-based calculations such as ground-motion field output from OpenSHA and ELER. This has been undertaken to highlight that the risk engine is not tied to the hazard engine of GEM1, but that it can be used by any user that has his/her own hazard data.

The main emphasis of these calculations has been on human fatalities, but in order to show the flexibility of the engine some preliminary economic loss calculations based on building damage have been undertaken. These latter results are not as reliable as the former and are not available on a global scale due to the lack of a global exposure database and a global vulnerability model for buildings (the development of such databases will be the focus of some of GEM's global components). The emphasis has thus currently been on a proof of concept, i.e. that end-to-end loss calculations from hazard to loss can be carried out with the GEM1 engines. Nevertheless, the preliminary global risk results presented herein in terms of human fatalities do constitute a step forward with respect to some of the previous global initiatives due to the consideration of probabilistic seismic hazard curves using the most up-to-date regional seismic hazard models from around the world.

17 Future Developments of GEM1 Risk Engine

This Chapter presents future developments of the risk engine, focusing on the risk domain related developments following the recommendations of the MAG. The IT developments will be guided by the IT review that was held in Zurich in June 2010 (as described on the GEM website).

Vulnerability

A continuous vulnerability function class will be added to the risk engine (currently only discrete vulnerability functions can be input). Furthermore continuous and discrete fragility functions will also be added, where fragility is defined as the probability density (or mass) function of damage state levels given a scalar intensity measure level (IML). The distribution of the probability of exceedance of a given limit state with IML can be discrete or continuous. The following limit state types could then be included: damage levels from none to complete, injury levels from none to death, production levels from not producing to producing at 100%.

Further extensions of the code will include calculating the vulnerability and fragility on-the-fly (for example using non-linear static procedures where the response of the structure is estimated from the convolution of capacity curves and response spectra and this is then related to damage exceedance probabilities), and with discrete consequence functions (i.e. functions defining the mean and standard deviation of loss given a damage state). The priority in terms of methodologies that the software should support will depend on the recommendations of the Global Vulnerability Estimation Methods global component consortium.

Exposure

The engine will need to be modified to read exposure input that comprise portfolios of buildings at discrete locations (i.e. not on a grid). Furthermore, once the Global Exposure Database consortium propose the structure of the global inventory database, the engine will need to be developed to read and process this information in the loss calculations.

Convolution of hazard and risk

When performing risk analyses, various outputs can be obtained as a function of the following parameters:

- Number of Events: this parameter distinguishes the deterministic event-based analyses (one event) from classical PSHA or probabilistic event-based analyses (many events).
- Number of locations/Assets: one asset at one site, many assets at one site or many assets at many sites (also known as portfolio analysis) can be considered.
- Spatial correlation of intra-event variability: when considering many closely-spaced sites, the correlation of the ground motions at those sites should be accounted for in the generation of the ground-motion fields.

Using different combinations between the above variables, it has been possible to achieve 8 different cases. This set of combinations can be represented using the diagram in Figure 17.1. In order to understand this diagram, readers should start from the centre and proceed from tier to tier until the most peripheral layer (which contains the case type) is achieved. A description of each of the cases (and recommendations on when one case might be used rather than another) are provided below. The aim of immediate future developments is to ensure that the risk engine can consider all of these cases.

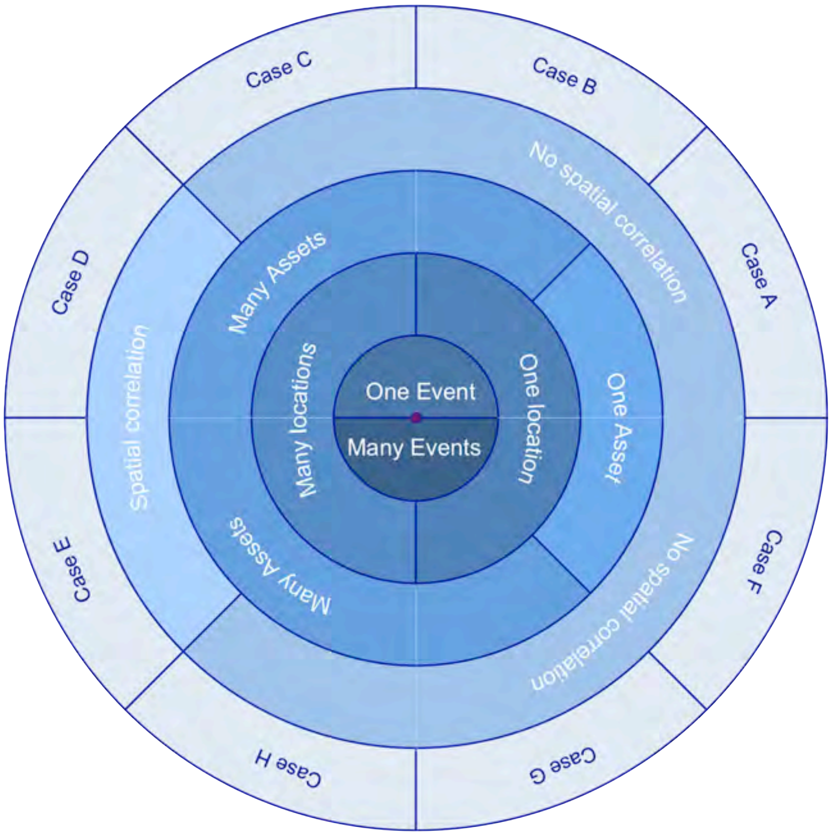


Figure 17.1 Possible cases of risk assessment to be modelled in the risk engine

Case A

Deterministic event-based analysis where only 2 ground-motion parameters are calculated, the mean and the standard deviation of the ground shaking intensity (scalar intensity measure level, IML) at the location of the asset, leading to a mean loss and standard deviation of loss.

Case B

As case A (i.e. just 2 ground-motion parameters), with the loss calculations repeated for each asset at the site. There could be a need to consider the correlation of the asset vulnerability functions at the site in order to aggregate the loss distributions for each asset, but this is probably too theoretical for implementation within the risk engine and there is insufficient data upon which to model such correlation.

Case C

As case A but with the mean and standard deviation of the ground-motion parameter (scalar IML) calculated at the location of each asset. The mean loss and standard deviation of loss of each asset can be calculated and then the aggregate[2] mean loss of all assets can be computed. Cases A, B and C are likely to be used for rapid post-earthquake loss assessment to obtain summary statistics for the event and for emergency planning for selected scenario earthquakes to understand the spatial distribution of mean loss, due to the possibility to produce maps.

Case D

Deterministic event-based analysis where multiple ground-motion fields (GMFs) of scalar IMLs are calculated (e.g. a number of maps showing the spatial distribution of PGA for a given earthquake). The event is repeated a number of times and each realisation considers a random sample of the inter-event epsilon for the event, and a random sample of the intra-

[2] This refers to the sum of the losses for all assets.

event epsilon at each site, considering the spatial correlation of epsilon at pairs of sites. A preliminary version of the calculator for these stochastic ground-motion fields has already been prototyped [Pagani *et al*. 2010]. Mean aggregate loss and total standard deviation of loss can be calculated. Case D is fairly time consuming and has the only advantage of allowing the total standard deviation of loss to all assets for a given event to be calculated, which is often not of interest for communication purposes.

Case E

As case D, but repeated for all the events within stochastic event sets[3]. In this case, loss exceedance curves can be generated considering each asset separately, or considering the combined value of all assets; the latter is of particular importance when the risk to a portfolio of closely spaced assets needs to be calculated. Maps of loss can also be produced, either for losses at a given probability of exceedance or for the mean loss within a given time span.

Case F

Classical PSHA-based risk analysis requiring a single hazard curve (of scalar IML) at a single site, leading to a single site loss exceedance curve. The same results can be obtained in a less efficient manner using the ground motions generated from stochastic event sets.

Case G

Classical PSHA-based risk analysis repeated for many assets at a given site. This suffers from the same considerations on vulnerability correlation as case B. The same results can be obtained in a less efficient manner using the ground motions generated from stochastic event sets.

Case H

Classical PSHA-based analysis requiring hazard curves (for scalar IMLs) at the location of each asset. The mean losses from the different assets can be aggregated, but the loss exceedance curves can only be aggregated if the inter- and intra-event variability are separated in the calculations (for example, by using the recent methodology proposed by Wesson *et al*. [2009], though it is noted that this method does not currently include spatial correlation of intra-event epsilon).

In order for vector-based intensity measures to be considered adequately in the risk calculations, their correlation should be modelled. Should such partial correlation be ignored (for example, scenario response spectra required for the Capacity Spectrum Method might be generated considering all spectral ordinates as fully correlated) then the scalar intensity measures described in cases A to E above can be replaced by vector-based intensity measures. For the same example of response spectra, cases F to H would need to be modified by an event-based approach (in order to avoid the use of the Uniform Hazard Spectrum), similar to cases A to C, respectively, and repeated for each event within the stochastic event sets.

There are likely to be other requirements for the calculations that are not accounted for in the above cases, and GEM's risk engine development team will strive, together with the risk community, to include these into the code as they arise.

Community Development

The development of GEM's risk engine has been proposed herein to be undertaken as an open source project. So far, the GEM1 risk code has been given an open source license, but until it is placed on a public repository and the development is carried out in the open, it cannot be claimed to be open source. GEM's IT team is currently revising the risk code following the IT review that was held in June 2010 and integrating it within a single engine together with the hazard, and is planning to release the code for open source development in January 2010.

---

[3] A stochastic event set is a set of ruptures generated by randomly sampling an earthquake rupture forecast (ERF); it can be considered a possible realization within a specified time span of the seismicity described by the ERF, where the latter is a collection of all the possible ruptures, and their probabilities of occurrence, occurring on all the sources within a PSHA input model.

REFERENCES

Document References

Allen, T. and Wald, D., [2009]. " Evaluation of ground-motion modeling techniques for use in Global Shakemap – A critique of instrumental ground-motion prediction equations, peak ground motion to macroseismic intensity conversions, and macroseismic intensity predictions in different tectonic settings". Technical Report 2009-1047, U.S. Geological Survey, Golden, United States of America.

Allen, T. I., and Wald, D. J. [2007] "Topographic slope as a proxy for global seismic site conditions (V_{S30}) and amplification around the globe", *U.S. Geological Survey Open-File Report,* 2007-1357, 69 p.

Allen, T., Wald, D., Earle, P. S., Marano, K., Hotovec, A., Lin, K., Hearne, M., [2009] "An Atlas of ShakeMaps and population exposure catalogue for earthquake loss modelling". *Bulletin of Earthquake Engineering*, Vol. 7, No. 3, pp. 701-718.

Ambraseys N.N. [1985] "Intensity-attenuation and magnitude intensity relationships for Northwestern European earthquakes" *Earthquake Eng. Struct. Dyn.,*Vol. 13, pp. 733–778.

ATC, Applied Technology Council, [1985]. ATC-13, Earthquake Damage Evaluation Data for California, Redwood City, United States of America.

ATC [1996] *Seismic evaluation and retrofit of concrete building,*. ATC 40, Applied Technology Council, Redwood City, California.

ATC [2005] *Improvement of nonlinear static seismic analysis procedures*, FEMA 440, Applied Technology Council, Washington D.C., USA.

Atkinson, G.M., and Boore D.M. [2006] "Earthquake ground-motion prediction equations for eastern North America" *Bulletin of the Seismological Society of America,* Vol 96, pp. 2181—2205.

Atkinson, G. M., and S. I. Kaka [2007] "Relationships between felt intensity and instrumental ground motion," *Bulletin of the Seismological Society of America*, Vol. 97, pp. 497–510.

BIS [2002] Indian standard—criteria for earthquake resistant design of structures, part 1—general provisions and buildings, Technical report, Bureau of Indian Standards (BIS), ICS 91.120.25.

Bommer, J.J., Elnashai, A.S., Chlimintzas, G.O., and Lee, D. [1998] "Review and development of response spectra for displacement - based seismic design", *ESEE Research Report nº 98-3*, Civil Engineering Department, Imperial College, London, United Kingdom.

Bommer,J.J., Stafford, P.J., and Alarcón, J.E. [2007] "The Influence of Magnitude Range on Empirical Ground-Motion Prediction," *Bulletin of the Seismological Society of America*, Vol. 97; No. 6; pp. 2152-2170.

Boore, D. M., Joyner, W. B., and Fumal, T.E. [1997] "Equations for Estimating Horizontal Response Spectra and Peak Accelerations from Western North American Earthquakes: A Summary of Recent Work," *Seismological Research Letters*, 68, 128-153.

Bramerini F., Di Pasquale G., Orsini A., Pugliese A., Romeo R., Sabetta F. [1995] “Rischio sismico del territorio italiano. Proposta per una metodologia e risultati preliminari,” *Rapporto tecnico del Servizio Sismico Nazionale SSN/RT/95/01*, Roma (In Italian).

Cagnan, Z., Demircioglu, M.B., Durukal, E., Erdik, M., Hancilar, U., Harmandar, E., Sesetyan, K., Tuzun, C., Yenidogan, C, Zulfikar, A., [2009] “Development of ELER (Earthquake Loss Estimation Routine) Methodology: Vulnerability Relationships”. Technical report, KOERI, Istanbul, Turkey,

Carvalho A., [2007] “Modelação estocástica da acção sísmica em Portugal Continental”, *Phd Thesis*, IST, Portugal.

Carvalho A., Zonno G., Franceschina G., Bilè Serra J., Campos Costa A. [2008] “Earthquake shaking scenario for the metropolitan area of Lisbon”, *Soil Dynamics and Earthquake Engineering*, Vol 28, pp.347-364.

CEN [2003] *Eurocode 8: Design of Structures for Earthquake Resistance*, Draft No. 6, European Committee for Standardization, Brussels.

Cheu, D.H. [1995] “*Northridge Earthquake – January 17, 1994: The Hospital Response,”* Universal City, CA, FEMA, 1-15, Northridge Earthquake: One Year Later.

Chiou, B. and Youngs, R.R. [2008] “An NGA model for the average horizontal component of peak ground motion and response spectra,” *Earthquake Spectra* Vol. 24, No. 1 , pp. 173-216.

Coburn A. and Spence R. [1992] “Factors Determining Human Casualty Levels in Earthquakes: Mortality Prediction in Building Collapse,” *Proceedings of the 10th World Conference on Earthquake Engineering*, Madrid, Spain, pp. 5989-5994.

Coburn, A. and Spence R. [2002] *Earthquake Protection* (Second Edition), John Wiley and Sons Ltd., Chichester, England.

Cua, G., Wald D. J., Allen T. I., Garcia D., Worden C. B., Gerstenberger M., Lin K., Marano K. [2010] "Best practices" for using macroseismic intensity and ground motion-intensity conversion equations for hazard and loss modes in GEM1,” GEM1 Technical Report 2010-4, GEM Foundation, Pavia, Italy.

Demircioglu M.B., Erdik M., Hancilar U., Sesetyan K., Tuzun C., Yenidogan C., Zulfikar A.C. [2009] “Earthquake Loss Estimation Routine ELER v1.0 – Technical Manual”, NERIES.

Di Pasquale, G. and Orsini, G. [1997] “Proposta per la valutazione di scenari di danno conseguenti ad un evento sismico a partiredai dati ISTAT” *Proceedings of VIII Convegno Nazionale L’Ingegneria Sismica in Italia*, Palermo, Italy.

Dilley M., Chen R.S., Deichmann U., Lerner-Lam A.L., Arnold M., Agwe J., Buys P., Kjekstad O., Lyon B. and Yetman G. [2005] Natural Disaster Hotspots: A Global Risk Analysis. Synthesis Report. Available from URL: http://sedac.ciesin.columbia.edu/hazards/hotspots/synthesisreport.pdf

Dobson, J., Bright, E., Coleman, P., Durfee, R., Worley, B., [2000] “LandScan: A Global Population Database for Estimating Populations at Risk*”. Photogrammetric Engineering and Remote Sensing*, Vol. 66, No. 7, pp. 849-857.

Durkin M. E., and Thiel C. C. [1993] “Towards a Comprehensive Regional Earthquake Casualty Modeling Process,” *Proc. of National Earthquake Conference*. Vol I, Central U.S. Earthquake Consortium, pp. 557-566.

ECOS [2006] *Earthquake Catalogue of Switzerland.* Swiss Seismological Service, Zürich, Switzerland.

Elvidge, C., Safran, J., Nelson, I., Tuttle, B., Hobson, V., Baugh, K., Dietz, J., Erwin, E., [2004]. “Area and position accuracy of DMSP nighttime lights data”. Remote Sensing and GIS Accuracy Assessment, CRC Press, p. 281-292.

Erdik M. [1993] “Seismicity and Strong Ground Motion,” *in Damage Report on 1992 Erzincan Earthquake, Turkey* pp.2-1,7, Architectural Institute of Japan, Japan Society of Civil Engineers and Boğaziçi University, Tokyo, Japan.

Erdik, M. and Aydınoglu, N., [2002] “Earthquake performance and vulnerability of buildings in Turkey.” *The World Bank Group Disaster Management Facility Report*.

Fajfar P. [2000] “A non linear analysis method for performance-based seismic design,” *Earthquake Spectra*, Vol. 21, No. 3, pp. 573-591.

FEMA & NIBS [1999] HAZUS99 user and technical manuals, Federal Emergency Management Agency Report: HAZUS 1999, Washington D.C.,USA.

FEMA [2003] HAZUS-MH Technical Manual . Federal Emergency Management Agency, Washington, DC, U.S.A.

Giovinazzi S. and Lagomarsino S. [2004] “A Macroseismic Model for the vulnerability assessment of buildings,” *Proceedings of the 13th World Conference on Earthquake Engineering. Vancouver*, Canada.

Giovinazzi S. and Lagomarsino S. [2005] “Fuzzy-Random Approach for a Seismic Vulnerability Model”, *Proceedings of ICOSSAR*, Rome, Italy.

Grünthal, G. [1998] “European Macroseismic Scale 1998,” Centre European de Geodynamique et de Seismologie, Luxembourg.

Gunasekera, R., Toyos, G., Saito, K., [2009] “Sensitivity analysis and application of remotely derived global population datasets in risk assessment of natural phenomena”, Remote Sensing and Photogrammetry Society Annual Conference 2009, Leicester, United Kingdom.

Huo J. and Hu Y. “Study on attenuation laws of ground motion parameters” *Earthquake Engineering and Engineering Vibration*, Vol12, pp. 1–11.

IBC [2006] *International Building Code*, International Code Council, USA.

ISDR [2009] Global Assessment Report on Disaster Risk Reduction. Risk and poverty in a changing climate. United Nations. Geneva. Switzerland. Available from: http://www.preventionweb.net/english/hyogo/gar/report/index.php?id=9413

Jaiswal, K., Wald, D., [2008]. “A Global Building Inventory for Earthquake Loss Estimation and Risk Management”. Technical Report 2008-1160, U.S. Geological Survey, Golden, United States of America.

Jaiswal, K., D. Wald, Hearne, M., [2009]. “Estimating casualties for large worldwide earthquakes using an empirical approach“. Technical Report 2009-1136, U.S. Geological Survey, Golden, United States of America.

Kamer Y., Abdulhamitbilal E., Demircioglu M.B., Erdik M., Hancilar U., Sesetyan K., Yenidogan C., Zulfikar A.C. [2009] “Earthquake Loss Estimation Routine ELER v1.0 – User Manual” NERIES.

KOERI (Department of Earthquake Engineering) [2002] “Earthquake Risk Assessment for Istanbul Metropolitan Area”, *Report prepared for American Red Cross and Turkish Red Crescent,* Bogazici Univeristy, Istanbul, Turkey.

Miranda, E. [1999]. “Approximate Seismic Lateral Deformation Demands in Multistory Buildings,” *Journal of Structural Engineering*, Vol. 125, No. 4, 1999, pp. 417-425.

Molina, S. Lang, D.H., Lindholm, C.D [2009a] SELENA v.4.0 User and Technical Manual v4.0, NORSAR, Universitat d'Alacant, ICG (International Center fro Geohazars and SAFER).

Molina, S., Lang, D.H., Lindholm, C.D., Lingvall, F. [2009b] User Manual for the Earthquake Loss Estimation tool: SELENA (v.5), NORSAR, Universitat d'Alacant, ICG (International Center fro Geohazars and SAFER).

Müller, M., Vorogushyn, S., Maier, P., Thieken, A.H., Petrow, T., Kron, A., Büchele, B., and Wächter, J. [2006] "CEDIM Risk Explorer – a map server solution in the project ´Risk Map Germany´", Natural Hazards and Earth System Sciences, Vol. 6, pp. 711–720.

Murakami H. O. [1992] "A simulation model to estimate human loss for occupants of collapsed buildings in an earthquake," *Proceedings of the 10th World Conference in Earthquake Engineering*, Madrid, Spain, pp.5969 -5974.

Nachtergaele, F., Petri, M., [2007] "Mapping Land Use Systems at global and regional scales for Land Degradation Assessment Analysis", Version 1.0, internal report, LADA project.

Newmark N.M. and Hall W.J. [1982] "Earthquake spectra and design" *Technical report*, Earthquake Engineering Research Institute (EERI), Oakland, CA, EERI.

Pagani, M., D. Monelli, H. Crowley, L. Danciu, E. H. Field, S. Wiemer, D. Giardini [2010] "GEM1 Hazard: Description of input models, calculation engine and main results," GEM Technical Report 2010-3, GEM Foundation, Pavia, Italy.

Porter, K. and Scawthorn, C. [2010] "Data Interchange Formats for the Global Earthquake Model (GEM)" GEM Technical Report 2010–E3, GEM Foundation, Pavia, Italy.

Porter, K.A. [2005] "A Taxonomy of Building Components for Performance-Based Earthquake Engineering," PEER Report No 2005/03, Pacific Earthquake Engineering Research (PEER) Center, Berkley, CA. http://www.sparisk.com/pubs/Porter-2005-Taxonomy.pdf

Porter, K.A. [2009a] "Cracking an open safe: HAZUS vulnerability functions in terms of structure-independent spectral acceleration," *Earthquake Spectra* Vol. 25, No. 2, pp. 361-378

Porter, K.A., [2009b] "Cracking an open safe: more HAZUS vulnerability functions in terms of structure-independent spectral acceleration," *Earthquake Spectra* Vol. 25, No. 3, pp. 607-618.

Porter, K.A., Kennedy, R.P., Bachman, R.E. [2006] "Developing Fragility Functions for Buildings Components," A Report to ATC-58, Applied Technology Council, Redwood City, CA.

Porter, K.A., Kennedy, R.P., Bachman, R.E. [2007] "Creating fragility functions for performance-based earthquake engineering," *Earthquake Spectra*, Vol. 23, No. 2, pp 471-489.

RGELFE (Research Group for Estimating Losses from Future Earthquakes) [1995] "Estimating Losses form Earthquakes in China in the Forthcoming 50 years," State Seismological Bureau, Seismological Press, Beijing.

Sachs J.D., Mellinger A.D., and Gallup J.L. [2001] "The geography of poverty and wealth." *Scientific American* Vol. 284, No. 3, pp. 70-75.

Salvatore, M., Pozzi, F., Ataman, E., Huddleston, B. and Bloise, M. [2005] "Mapping global urban and rural population distributions," Food and Agriculture Organization of the United Nations, Rome, 2005.

Samardjieva, E. and Badal J. [2002] "Estimation of the expected number of casualties caused by strong earthquakes," *Bulletin of the Seismological Society of America*, Vol. 92, No. 6, pp. 2310-2322.

Shiono, K., Krimgold, F. and Ohta, Y. [1991] "A Method for the Estimation of Earthquake Fatalities and its Applicability to the Global Macro-Zonation of Human Casualty Risk," *Proceedings Fourth International Conference on Seismic Zonation*, Stanford, CA, Vol. III, pp. 277 - 284.

Sousa, M.L. [2006] "Risco Sísmico em Portugal Continental", *Phd Thesis*, Universidade Técnica de Lisboa – Instituto Superior Técnico, Portugal.

Sousa, M.L. [2007] "Lisbon Case Study". Dissemination Workshop on Earthquake Disaster Scenario Predictions and Loss Modelling for Urban Areas. – Application, LESSLOSS – SP10, Kandili Observatory and Earthquake Research Institute, Bogazici University, Turkey.

Stojanovski, P., and Dong, W. (1994). Simulation Model for Earthquake Casualty Estimation. Proc. 5th U.S. National Conference on Earthquake Engineering, Paper No. 00592, Chicago, Illinois.

Tiedemann, H. [1992] "Earthquakes and volcanic eruptions: A handbook on risk assessment". Swiss Reinsurance Company, Zurich, Swiss.

Tyagunov, S., Grünthal, G., Wahlström, R., Stempniewski, L., Zchau, J. [2006a] "Seismic risk mapping for Germany," *Natural Hazards and Earth System Sciences (NHESS)*, Vol. 6, 4, pp. 573-586.

Tyagunov, S., Grünthal, G., Wahlström, R., Stempniewski, L., Zchau, J. [2006b] "Building stock vulnerability modeling for earthquake damage and loss assessment," *First European Conference on Earthquake Engineering and Seismology*, Geneva, Switzerland.

UNDP [2004] "Reducing Disaster Risk. A Challenge for Development." United Nations Development Programme, Bureau for Crisis Prevention and Recovery

United Nations [2008] "Demographic Yearbook 2005", United Nations Publication, 57° edition, New York, United States of America.

Vacareanu R., Lungu D., Arion C., Aldea A. [2004], WP7 *Report: Seismic Risk Scenarios*, Risk-UE Project, An advanced approach to earthquake risk scenarios with applications http://www.risk-ue.net.

Wald, D. J., Quitoriano, V., Heaton, T. H., Kanamori, H., Scrivner, C. W., and Worden, C. B. [1999a] "TriNet ShakeMaps: Rapid Generation of Instrumental Ground Motion and Intensity Maps for Earthquakes in Southern California," *Earthquake Spectra*, Vol 15, 537-556.

Wald, D., Quitoriano, V., Heaton, T., Kanamori, H., [1999b] "Relationships between Peak Ground Acceleration, Peak Ground Velocity and Modified Mercalli Intensity in California". Earthquake Spectra, Vol. 15, No 3, pp. 557-564.

Wald, D.J., Worden, B., and J. Goltz. [2003] ShakeMap — A Tool for Earthquake Response, U.S. Geological Survey Fact Sheet 087-03.

Wald, D. J., Quitoriano, V., and Dewey, J. W. [2006] "USGS "Did you feel it?" community internet intensity maps: macroseismic data collection via the internet," *First European Conference on Earthquake Engineering and Seismology*, Geneva, Switzerland.

Wald, D., Lin, K., Quitoriano, V., [2008]. "Quantifying and Qualifying USGS Shakemap Uncertainty". Technical Report 2008-1238, U.S. Geological Survey, Golden, United States of America.

Wesson, R.L., Perkins, D.M., Luco, N., Karaca, E. [2009] "Direct Calculation of the Probability Distribution for Earthquake Losses to a Portfolio," *Earthquake Spectra*, Vol 25, No. 3, pp. 687-706

Wyss, M. and Trendafiloski, G. [2009]. "Trends in the casualty ratio of injured to fatalities in earthquakes," *Proceedings of the Second International Workshop on Disaster Casualties*, Cambridge.

Youngs R.R., Chiou S. J., Silva W. J. and Humphrey, J. R. [1997] "Strong Ground Motion Attenuation Relationships for Subduction Zone Earthquakes" *Seismological Research Letters*, Vol. 68, No. 1, 58-73.

Zuccaro, G. and Papa, F. [2002] "Method of seismic vulnerability and exposure assessment at national scale – the Italian case" *Proceedings of XII European Conference on Earthquake Engineering,* London, United Kingdom.

Website references

1. Servizio Sismico Nazionale (SSN).
http://www.serviziosismico.it/PROG/1998/RISCHIO/start.html
2. USGS website
http://www.usgs.gov/
3. LandScan web site
http://www.ornl.gov/sci/landscan/
4. QLARM web site
http://qlarm.ethz.ch/
5. MAEviz web site
http://mae.cee.uiuc.edu/software_and_tools/maeviz.html
6. RISKSCAPE web site
http://www.riskscape.org.nz/
7. SELENA web site
http://www.norsar.no/pc-35-68-SELENA.aspx
8. AGORA web site
http://www.risk-agora.org/
9. United Nations Development Programme
http://www.undp.org
10. United Nations Environment Programme
http://www.unep.org
11. Division of Early Warning and Assessment of UNEP
http://www.unep.org/dewa/index.asp
12. Global Resource Information Database – Europe
http://www.grid.unep.ch
13. Centre for Research on the Epidemiology of Disasters (CRED)
http://www.cred.be
14. ProVention Consortium
http://www.proventionconsortium.org
15. Global Seismic Hazard Assessment Program
http://www.seismo.ethz.ch/GSHAP
16. United Nations International Strategy for Disaster Risk Reduction
http://www.unisdr.org
17. USGS ShakeMap Atlas
http://earthquake.usgs.gov/eqcenter/shakemap/atlas.php
18. LandScan
http://www.ornl.gov/sci/landscan
19. GRUMP (Global Rural-Urban Mapping Project)
http://sedac.ciesin.columbia.edu/gpw
20. Gridded Population of the World
http://sedac.ciesin.columbia.edu/gpw
21. GLCC (Global Land Cover Characterization)
http://edc2.usgs.gov/glcc/glcc.php
22. US Census Bureau's International Program Center
http://www.census.gov/ipc/www/idb

23. Gazetteer
http://www.world-gazetteer.com
24. City Population
http://www.citypopulation.de
25. UN Population Division
http://www.un.org/popin
26. GAUL (Global Administrative Unit Layers)
http://www.fao.org/geonetwork/srv/en/main.home
27. GADM
http://www.gadm.org
28. GLC2000
http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php
29.Diva-GIS
http://www.diva-gis.org/gData
30. OpenSHA
http://www.opensha.org
31. GeoNode
www.geonode.org
32. GeoNode (CAPRA)
http://geonode.capra.opengeo.org
33. FAO Global Information and Early Warning System
http://lprapp08.fao.org/fenix-portal
34. French National Geographical Institute's Geoportail
www.geoportail.fr
35. Earth Atlas
http://earthatlas.info
36. OneGeology portal
http://portal.onegeology.org
37. Google Earth API
http://code.google.com/p/ext-js-google-earth-api

Other references

38. LandScan Citation
This product was made utilizing the LandScan 2008™ High Resolution global Population Data Set copyrighted by UT-Battelle, LLC, operator of Oak Ridge National Laboratory under Contract No. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government has certain rights in this Data Set. Neither UT-BATTELLE, LLC NOR THE UNITED STATES DEPARTMENT OF ENERGY, NOR ANY OF THEIR EMPLOYEES, MAKES ANY WARRANTY, EXPRESS OR IMPLIED, OR ASSUMES ANY LEGAL LIABILITY OR RESPONSIBILITY FOR THE ACCURACY, COMPLETENESS, OR USEFULNESS OF THE DATA SET.

APPENDIX A Summary Comparison of Risk Software

In the following, a list of tables with the main features of the tested codes are provided. These tables are reported in order to give an immediate idea of the software and they are useful to give a general overview and comparison between them. The tables are sub-divided using the same main themes followed within the body of the report: general aspects, IT details, exposure module, hazard module, vulnerability module and output.

Table A.1 Summary table of the software: General Information

General Information				
ELE Software	Owner or developer	Development status	Availability Status	Applicability Region
SELENA	NORSAR	- Version 4: Matlab - Version 5: C	Open Source	User defined
EQRM	Geoscience Australia	Version 1.0svn1393	Open Source	- Australia (default) - User defined
ELER	KOERI (NERIES project)	Version 2.0	Standalone application available	- Euro-Mediterranean region - User defined
QLARM	WAPMERR – SED - ETHZ	Version 1.1.7	Source Code provided upon request	Worldwide
CEDIM (CREST)	CEDIM		Source Code provided upon request	User defined
CAPRA	World Bank	Version 0.9.8.0	Source Code provided upon request	User defined
RiskScape	FRST – GNS Science - NIWA	Version 0.1.4	Standalone application provided	New Zealand
LNECLoss	LNEC		Source Code provided upon request	- Portugal (default) - User defined
MAEviz	MAE Center	Version 3.1.1	Open Source	- Mid America (default) - User defined
OpenRisk	SPA Risk LLC		Source Code provided upon request	Southern California

Table A.2 Summary table of the software: IT details

IT details					
ELE Software	Software requirements	Programming Language	Format of input/output	Visualisation of results	GUI
SELENA	- Matlab - C compiler - RISe: compatible GIS software	- Matlab - C	ASCII-table files	Yes	Yes
EQRM	Python, Numpy, Scipy, Numeric, Shapely, minGW	Python	ASCII-table files, *.csv	No	No
ELER	None – standalone application	Matlab	*.xml, *.mat, *.shp, *kml, *dbf, manual input using a GUI	Yes	Yes
QLARM	Internet connection because it is a web application	Java	HTML, *.xml, *.csv, *.pdf, *.jpg, manual input using a GUI	Yes	Yes
CEDIM (CREST)	ArcGIS	Visual Basic	*.dbf, *.shp	Yes	Yes
CAPRA	None – standalone application	Visual Basic.Net	*.ame, *shp, *.fvu, *.dat, *.dbf	Yes	Yes
RiskScape	None – standalone application	Java	*.kml, *.shp, *.xls, *.pdf, manual input using a GUI	Yes	Yes
LNECLoss	- Fortran compiler - Optional ArcGIS	Fortran	ASCII-table files, *.dbf, *.shp	Yes	Yes
MAEviz	None – standalone application	Java	*.dbf, *.shp	Yes	Yes
OpenRisk	None – standalone application	Java	ASCII-table files, manual input using a GUI	Yes	Yes

Table A.3 Summary table of the software: Scientific details, Exposure

Scientific Details - *Exposure*				
ELE software	Building Inventory	Economic Inventory	Population inventory	Other inventory
SELENA	yes (user-defined)	yes (user-defined)	yes (user-defined)	no
EQRM	yes (user-defined)	yes (user-defined)	no	no
ELER	yes (hard-coded or user-defined)	no	yes (hard-coded or user-defined)	no
QLARM	yes (hard-coded)	no	yes (hard-coded)	no
CEDIM (CREST)	yes (user-defined)	yes (user-defined)	yes (user-defined)	no
CAPRA	yes (user-defined)	yes (user-defined)	yes (user-defined)	Different types of assets, as long as the user has the necessary vulnerability function
RiskScape	yes (hard-coded)	yes (hard-coded)	yes (hard-coded)	Electricity cables, Telecommunication cables, Network Junction Points, Roads, Agriculture, Waterways, Pipelines
LNECLoss	yes (user-defined)	yes (user-defined)	yes (user-defined)	no
MAEviz	yes (hard-coded or user-defined)	yes (hard-coded or user-defined)	yes (hard-coded or user-defined)	Bridges, Lifelines
OpenRisk	yes (hard-coded)	yes (hard-coded)	no	no

Table A.4 Summary table of the software: Scientific details, Hazard

ELE software	Typology of hazards	Ground-motions parameters	Hazard Types supported	Local Site effects
SELENA	Ground-shaking	Response Spectrum	- Deterministic event-based (including real time) - Probabilistic seismic hazard maps	Yes
EQRM	Ground-shaking	Response Spectrum or macroseismic intensity	- Deterministic event-based - Probabilistic events-based	Yes
ELER	Ground-shaking	Response Spectrum or macroseismic intensity	- Deterministic event-based	Yes
QLARM	Ground-shaking	Macroseismic Intensity or instrumental parameters (e.g. PGA, PGV)	- Deterministic event-based	Yes
CEDIM (CREST)	Ground-shaking	Macroseismic Intensity	- Deterministic event-based - Probabilistic seismic hazard maps	Yes
CAPRA	Multi-hazard: ground-shaking and other different natural hazard	User-defined	- Deterministic event-based - Probabilistic events-based	Yes
RiskScape	Multi-hazard: ground-shaking, inundation related to river flooding, inundation related to tsunami, storm and volcanic ashfall	Macroseismic Intensity	- Deterministic event-based	Yes
LNECLoss	Ground-shaking	Macroseismic Intensity or response spectrum	- Deterministic event-based	Yes
MAEviz	Multi-hazard: ground-shaking and liquefaction	Response Spectrum	- Deterministic event-based - Probabilistic seismic hazard maps	Yes
OpenRisk	Ground-shaking	Response Spectrum	- Classical PSHA-based	Yes

Table A.5 Summary table of the software: Scientific details, Vulnerability and Output

Scientific Details – *Vulnerability and output*				
ELE software	Vulnerability	Building damage	Economic Losses	Social Losses
SELENA	Indirect, analytical approach: CSM, MADRS, DCM method	Structural damage	yes	yes
EQRM	Indirect, analytical approach: CSM	Structural and non structural damage	yes	no
ELER	- Indirect analytical approach: CSM, MADRS, RFM, DCM - Direct and indirect empirical approaches (seismological and engineering-based)	Structural damage	no	yes
QLARM	Indirect empirical approach (engineering-based)	Structural damage	no	yes
CEDIM (CREST)	Indirect empirical approach (engineering-based)	Structural damage	yes	yes
CAPRA	Indirect approach, user-defined vulnerability functions (can be empirical or analytical)	Structural and non structural damage	yes	yes
RiskScape	Indirect empirical approach (engineering-based)	Structural damage	yes	yes
LNECLoss	- Indirect analytical approach: CSM - Direct and indirect empirical approaches (four methods are implemented, seismological and engineering-based)	Structural damage	yes	yes
MAEviz	Indirect analytical approach	Structural and non structural damage	yes	yes
OpenRisk	Indirect approach, user-defined vulnerability functions (can be empirical or analytical)	Structural and non structural damage	yes	no

APPENDIX B Calculations of the GEM1 Risk Engine

A large part of the efforts while developing this risk engine were focused on having reliable results and keeping a good balance between accuracy and computational performance. The present section will describe the main assumptions that were made and possible sources of error whose mitigation could affect considerably the computational performance. This section is organized into three main parts: the probabilistic distributions that are currently being considered for both approaches, the calculations inherent to deterministic scenarios and finally, the procedures carried out for classical PSHA-based risk assessment. In order to better explain these methods, several examples are also presented.

B.1 Probabilistic Distributions

Normal Distribution

In statistics the normal distribution or Gaussian distribution is a continuous probability distribution that describes data that are dispersed around a mean value. The graph of the associated probability density function (PDF) has a bell shape with a peak at its mean and its dispersion will be as high as its standard deviation. The general formula for the probability density function (PDF) of a normal distribution is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{B.1}$$

Where μ stands for the mean of the distribution and σ stands for the standard deviation. The cumulative distribution function (CDF) of normal distributions cannot be represented through a formula and therefore, it needs to be computed numerically using the PDF. The following figure shows several lognormal cumulative distribution curves for different standard deviations:

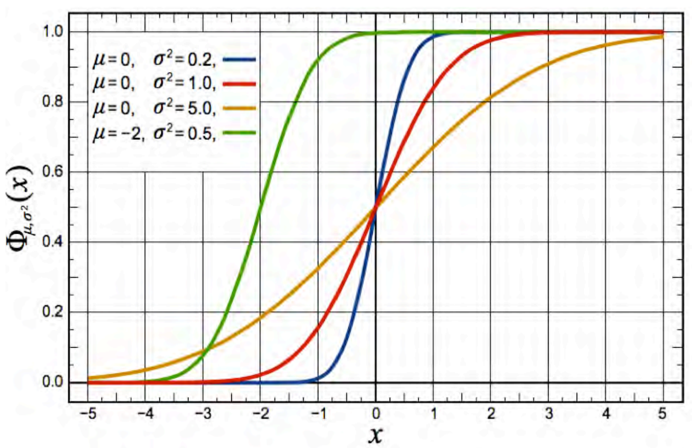


Figure B.1 Curves of cumulative distribution functions of a normal distribution.

Lognormal Distribution

A lognormal distribution is a probability distribution of a random variable whose logarithm is normally distributed. If Y is a random variable with a normal distribution, then $X=exp(Y)$ has a lognormal distribution. This type of distribution can also be referred to as the Galton distribution. The general formula for the probability density function (PDF) of a lognormal distribution (two parameters only) is:

$$f(x) = \frac{1}{x\sqrt{2\pi\zeta^2}} \exp\left[-\frac{(\ln x - \lambda)^2}{2\zeta^2}\right] \tag{B.2}$$

To calculate the logarithmic mean λ, and logarithmic standard deviation ζ the following formulas can be used:

$$\zeta = \sqrt{\ln\left(1 + \frac{\sigma^2}{\mu^2}\right)} = \sqrt{\ln\left(1 + COV^2\right)} \tag{B.3}$$

$$\lambda = \ln \mu - \frac{1}{2}\zeta^2 \tag{B.4}$$

Where μ stands for the mean of the distribution and σ stands for standard deviation. The cumulative distribution function (CDF) of a lognormal distribution can either be computed numerically using the previous probability density function or using the cumulative distribution function of a normal distribution. The following figure shows several lognormal cumulative distributions curves for different standard deviations:

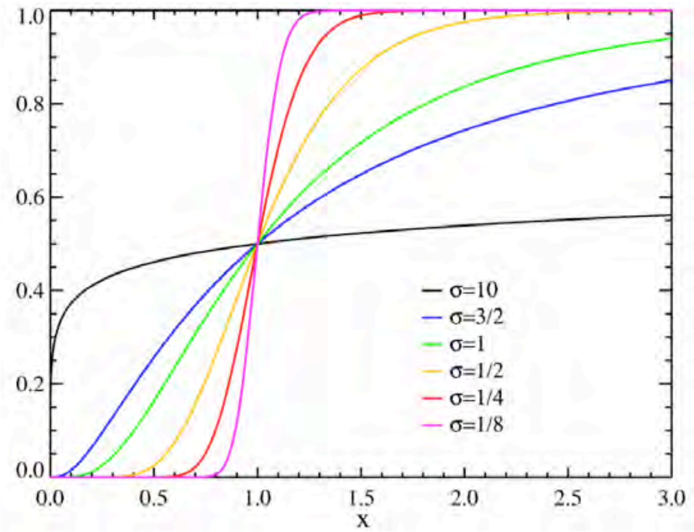


Figure B.2 Curves of cumulative distribution functions of a lognormal distribution.

The median, mode and variance of this type of distribution can be calculated using the following expressions:

$$M = e^{\lambda} \tag{B.5}$$

$$\sigma^2 = (e^{\zeta^2} - 1)e^{2\lambda + \zeta^2} \tag{B.6}$$

Uniform Distribution

In statistics, a continuous uniform distribution is a family of probability distributions such that for each member of the family, all intervals of the same length on the distribution are equally probable. The function is usually defined by two parameters, *a* and *b*. The general formula for the probability density function (PDF) of a uniform distribution is:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & for\ x \in [a,b] \\ 0 & for\ x \notin [a,b] \end{cases} \tag{B.7}$$

Where *a* and *b* represent the minimum and maximum values respectively. The cumulative distribution function (CDF) of a uniform distribution can be obtained through the following formula:

$$F(x)=\begin{cases}0 & for\ x \leq a\\ \dfrac{x-a}{b-a} & for\ x \in [a,b]\\ 0 & for\ x \geq b\end{cases} \tag{B.8}$$

The median, mode and variance of this type of distribution can be calculated using the following expressions:

$$M=\frac{1}{2}(a+b) \tag{B.9}$$

$$\sigma^2=\frac{1}{12}(b-a)^2 \tag{B.10}$$

Further details about the methods that can be used to solve these functions will be given in the following section.

Numerical Methods

In order to compute probabilities of occurrence or exceedance of certain events it is necessary to implement procedures within the risk engine that are capable of solving cumulative distribution functions. As previously mentioned, this type of function might not have a numerical expression and therefore, numerical methods are required.

The following figure presents a lognormal distribution of loss ratios for a certain intensity measure level.

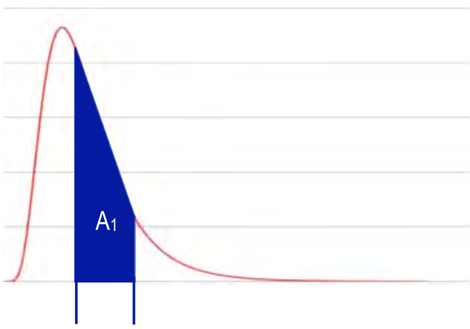


Figure B.3 Lognormal distribution of loss ratios for a given intensity measure level.

The probability of occurrence of any loss ratio bounded between LR_1 and LR_2 is given by the area delimitated by the curve of the graph and the horizontal axis, as demonstrated in the previous figure (area A_1) :

$$P(LR_1 < x < LR_2) = A_1 \tag{B.11}$$

To compute this area the risk engine could simply assume a trapezoid shape and apply the respective area formula. This assumption could generate significant numerical errors since the upper limit of the trapezium is slightly different from the curve. To avoid this problem, this method should be applied repeatedly to small ranges (trapezium rule). Similar methods such as the Simpson rule were also studied. An alternative to these numerical methods is to use statistical libraries that already support methods to solve cumulative distribution functions. Currently, the GEM1 risk engine is using a statistical function called Error Function (erf) or Gauss Error Function featured on a Java library (Apache Commons Mathematics library). This function can be expressed by the following formula:

$$erf(x)=\frac{2}{2\pi}\int_0^x e^{-t^2}.dt \tag{B.12}$$

Its graph has a sigmoid shape and it is presented in the following figure:

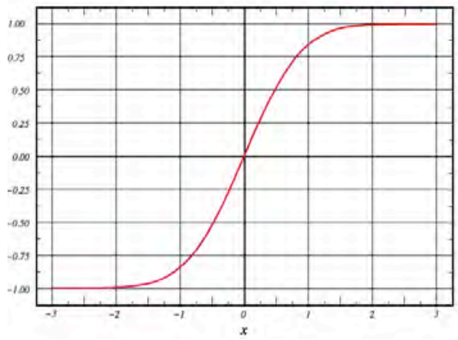


Figure B.4 Graph of the error function.

In the current risk engine, the x value is provided and the value of this statistical function is returned, and then, using this result, the value of cumulative distribution function can be computed for normal and lognormal distributions using the following formulas:

- For normal probabilistic distributions:

$$F(x,\mu,\sigma)=\frac{1}{2}\left[1+erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \tag{B.13}$$

- For lognormal probabilistic distributions:

$$F(x,\lambda,\zeta)=\frac{1}{2}\left[1+erf\left(\frac{\ln(x)-\lambda}{\zeta\sqrt{2}}\right)\right] \tag{B.14}$$

These formulas always return the cumulative value between - ∞ and x (LR). Therefore, to compute the probability of occurrence of a range of loss ratios, it is necessary to apply the following expression:

$$P(LR_1 < x < LR_2) = F(LR_2,\mu,\sigma) - F(LR_1,\mu,\sigma) \tag{B.15}$$

B.2 Deterministic Event-Based Analysis

To compute economic or human losses for single scenario events, the GEM1 risk engine uses “ground-motion fields”, which provide the ground motion and respective standard deviation for each cell within the region of interest. The required calculations to compute the loss ratio for every zone considering the uncertainty of the hazard and vulnerability model are explained within this section. A demonstration of how to compute the mean loss ratio and the respective standard deviation using the total probability theorem is shown.

Mean Loss Ratio

For this type of assessment, it is important to understand the intensity measure levels that might occur in each cell. Using the two “ground-motion fields” provided to the risk engine (mean intensity levels and standard deviation), the risk engine builds a probabilistic distribution of the intensities assuming a lognormal dispersal of the values and then, computes the mean loss ratio for the considered grid cell. In order to better explain how this procedure is being done, an example is presented for a specific location in Turkey (Marmara region) where an IML of VII (Modified Mercalli Intensity) was felt.

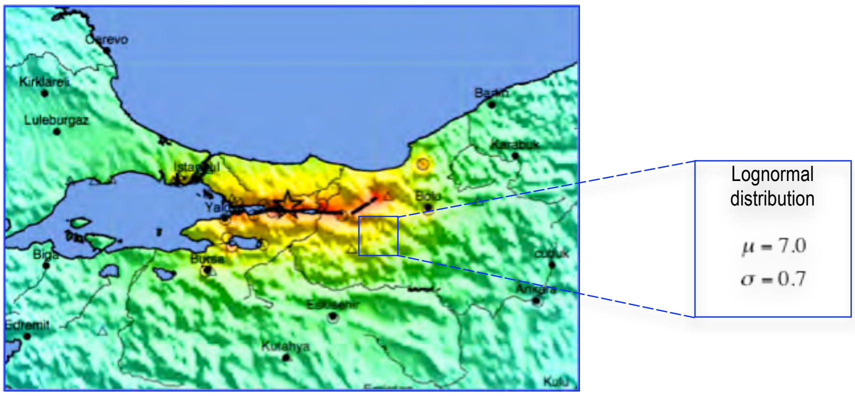


Figure B.5 Intensity level and standard deviation for a certain grid cell in Turkey.

Human losses are the loss type considered in this example and, therefore, the PAGER vulnerability model was used. The following table presents the vulnerability function produced for Turkey:

Table B.1 PAGER vulnerability function for Turkey (where μ is mean and COV is coefficient of variation)

IML	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
μ(LR)	1.96E^{-15}	2.53E^{-12}	8.00E^{-10}	8.31E^{-08}	3.52E^{-06}	7.16E^{-05}	7.96E^{-04}	5.37E^{-03}	2.39E^{-02}	7.51E^{-02}	1.77E^{-01}
COV	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.4	0.4	0.3

Although the hazard input indicates a mean IML of 7, it is important to understand that there is an uncertainty, and therefore, one needs to take into consideration the influence of the other IML's that might occur in that zone. The following figure shows the distribution of the IML's for this specific cell:

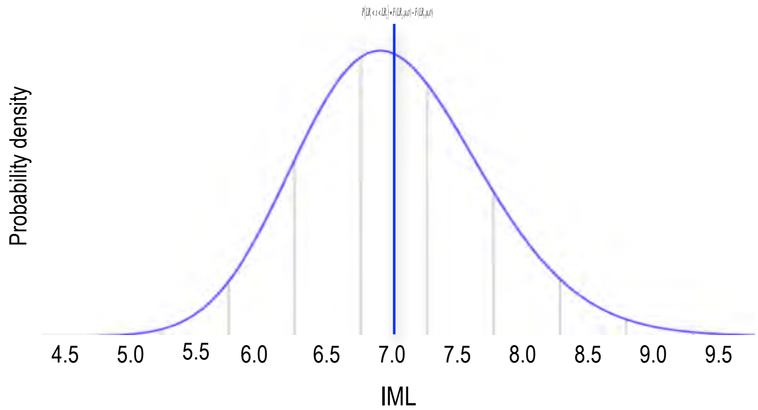


Figure B.6 Probability density curve of the intensity levels.

To compute the probability of occurrence of a certain IML, it is necessary to establish an upper and lower limit for that value. The figure above presents how the probability density curve was divided over the different values of IML. In this case, the list of IML's is defined between 5 and 10 with 0.5 increments, and therefore, each range will be: IML ± 0.25. The following expression shows how these limits can be computed for any collection of intensity levels.

$$\left(\frac{IML_{n-1}+IML_n}{2}\right) \le IML_n \le \left(\frac{IML_n+IML_{n+1}}{2}\right) \qquad \text{(B.16)}$$

Special care needs to be taken for the limit values of the intensity measure levels list, since there will not be a previous or following value to compute one of these bounds. The following is being assumed to compute the probability of occurrence in these special cases:

– First value of the list:

$$PO(IML_0) = PO\left(IML_0 \leq IML \leq \frac{IML_0 + IML_1}{2} \right) \tag{B.17}$$

– Last value of the list:

$$PO(IML_{n+1}) = PO\left(IML \geq \frac{IML_n + IML_{n+1}}{2} \right) \tag{B.18}$$

Using what has been established above and the respective cumulative probability function, the risk calculator uses the following formula to compute the probability of occurrence of a certain IML:

$$PO(IML_n) = F\left(\frac{IML_{n+1} + IML_n}{2}, \lambda, \zeta \right) - F\left(\frac{IML_n + IML_{n-1}}{2}, \lambda, \zeta \right) \tag{B.19}$$

Considering the probability of occurrence of every IML, the conditional loss ratio can be computed through the following formula:

$$LR = \sum_{n=1}^{m} PO(IML_n) \times LR_n \tag{B.20}$$

The following table shows the results for the previous example. Such results have been calculated in Excel for each step of all calculations for comparison and testing of the output of the risk engine.

Table B.2 Results for the calculation of the conditional loss ratio.

IML_n	$PO(IML_n)$	LR_n	$PO(IML_n) \times LR_n$
5.0	0.002	1.96E-15	4.39E-18
5.5	0.025	2.53E-12	6.31E-14
6.0	0.111	8.00E-10	8.91E-11
6.5	0.238	8.31E-08	1.98E-08
7.0	0.280	3.52E-06	9.84E-07
7.5	0.202	7.16E-05	1.44E-05
8.0	0.097	7.96E-04	7.76E-05
8.5	0.034	5.37E-03	1.81E-04
9.0	0.009	2.39E-02	2.12E-04
9.5	0.002	7.51E-02	1.39E-04
10.0	0.000	1.77E-01	5.67E-05
		$\sum_{n=1}^{m} PO(IML_n) \times LR_n$	6.82E-04

If the IML's follow a normal distribution, this procedure can be enhanced by only including IML's that are within 3 standard deviations since outside this range the probability of occurrence is not significant. Once this conditional loss ratio is computed for every grid cell on the region of interest, loss ratio maps and loss maps can be produced. In order to create the latter it is also necessary to obtain the value of the asset for every cell (which can be population count, replacement cost, number of buildings, etc) and multiply it by the computed loss ratio.

Standard Deviation

In order to compute the standard deviation of the loss ratio, the following formula is being used:

$$SD[LR]^2 = E\left[LR^2\right] - E[LR]^2 \tag{B.21}$$

Where E[LR] stands for the mean loss ratio previously computed (simply called LR) and the E[LR^2] can be calculated using the following formula:

$$E\left[LR^2\right]=\sum_{n=1}^{m} PO(IML_n)\times E\left[LR_n^2\right] \quad (B.22)$$

In order to compute every E[LR $_n^2$] the total probability theorem can be used assuming that:

$$E\left[LR_n^2\right]=SD\left[LR_n\right]^2+E\left[LR_n\right]^2 \quad (B.23)$$

Where E[LR$_n$] refers to LR$_n$. The following table presents the results for the previous example:

Table B.3 Results for the calculation of the standard deviation of the mean loss ratio.

IML_n	$PO(IML_n)$	LR$_n$	SD[LR$_n$]	E[LR $_n^2$]	$PO(IML_n)\times E\left[LR_n^2\right]$
5.0	2.00E-03	1.96E-15	1.37E-15	5.74E-30	1.28E-32
5.5	2.50E-02	2.53E-12	1.77E-12	9.50E-24	2.37E-25
6.0	1.11E-01	8.00E-10	5.60E-10	9.52E-19	1.06E-19
6.5	2.38E-01	8.31E-08	4.99E-08	9.39E-15	2.23E-15
7.0	2.80E-01	3.52E-06	2.11E-06	1.68E-11	4.71E-12
7.5	2.02E-01	7.16E-05	4.30E-05	6.97E-09	1.41E-09
8.0	9.70E-02	7.96E-04	3.98E-04	7.93E-07	7.72E-08
8.5	3.40E-02	5.37E-03	2.69E-03	3.61E-05	1.22E-06
9.0	9.00E-03	2.39E-02	9.56E-03	6.62E-04	5.87E-06
9.5	2.00E-03	7.51E-02	3.00E-02	6.54E-03	1.21E-05
10.0	0.00E+00	1.77E-01	5.32E-02	3.43E-02	1.10E-05
				E[LR^2] =	3.03E-05

Using expression 6.23, the standard deviation for the mean loss ratio is equal to:

$$SD[LR]=\sqrt{3.03E^{-05}-(6.82E^{-04})^2}$$
$$SD[LR]=0.0055$$

B.3 Classical PSHA-Based Analysis

As previously mentioned, the procedures that use classical PSHA-based analysis require a larger amount of calculations since loss ratio probability/exceedance matrices are computed for each vulnerability function. The introduction of the hazard input is also more complex since for every location a set of intensity measure levels and respective probability of occurrence/exceedance is going to be provided, instead of just a fixed value as is the case with deterministic events. The development of this part of the code was carried out keeping a constant communication with the hazard group, with the purpose of understanding how the hazard input should be computed and provided. Within this section, how the previously mentioned matrices are being computed, the main formulas that are being used and the various types of outputs that can be produced are described. It is noted that the main difference between the previous method and this hazard curve based method is that the former uses just two parameters to describe the probability of occurrence of IML (mean and standard deviation) whereas this method uses the full distribution

Loss Ratio Probability Matrix (LRPM)

A loss ratio probability matrix contains the probability of occurrence of each range of loss ratio for a set of intensity measure levels. For every IML, a probability density function of loss ratio needs to be computed, and the area between its curve and the horizontal axis needs to be distributed among the different ranges of loss ratio. In the following example, this procedure is performed for a specific intensity level.

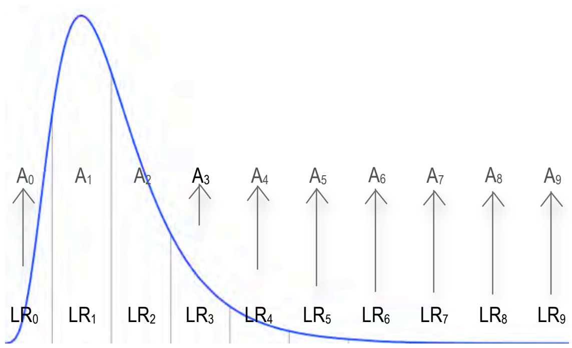


Figure B.7 Division of the probability density curve among the loss ratios for LRPM.

An upper and lower limit once again needs to be defined for each loss ratio similarly to what has been explained for the distributions of the intensity measure levels. The following table shows how the different areas are being organized within the loss ratio probability matrix.

Table B.4 Sample of a loss ratio probability matrix.

	IML			5.5
LR_0	A_0		0.000	0.093
LR_1	A_1		0.001	0.420
LR_2	A_2		0.003	0.301
LR_3	A_3		0.006	0.122
LR_4	A_4		0.018	0.049
LR_5	A_5	Hypothetical values ->	0.035	0.010
LR_6	A_6		0.068	0.004
LR_7	A_7		0.096	0.001
LR_8	A_8		0.120	0.000
LR_9	A_9		0.167	0.000

The following formula can be applied to compute each probability of occurrence, using the cumulative distribution function:

$$PO(LR_n) = F\left(\frac{LR_{n+1} + LR_n}{2};\lambda,\xi\right) - F\left(\frac{LR_n + LR_{n-1}}{2};\lambda,\xi\right) \qquad \text{(B.24)}$$

The exact same procedure would then be repeated for the remaining IML's. This means the LRPM will have a number of columns equal to the number of IML's. On the other hand, the number of rows in a LRPM does not need to be necessarily the amount of loss ratios defined by the vulnerability model. In the previous example, the dispersal of the probabilities was smooth since there were significant probabilities of occurrence in a large number of ranges of loss ratio within the same IML. However, if for instance there were a smaller number of loss ratios ranges or a distribution of the values with a smaller standard deviation, the probability of occurrence would be high in one of the loss ratio ranges, and then close to zero on the rest of the intervals. This aspect reduces the influence of the uncertainty of the vulnerability model and decreases the accuracy when computing probabilities of exceedance/occurrence of a certain loss ratio. An example of this issue will be shown within the following section.

Loss Ratio Exceedance Matrix (LREM)

A loss ratio exceedance matrix contains the probability of exceedance of each loss ratio for a set of intensity measure levels. The following formula gives this probability for any loss ratio:

$$PE(LR_n) = 1 - F(LR_n; \lambda, \zeta) \qquad (B.25)$$

Once one of the previous matrices is computed, the second can be calculated using the latter through interpolation methods, which is usually faster. However, this procedure brings some numerical errors that consequently decrease the exactness of the matrix. For this reason, it was decided to choose accuracy over computational performance and continue using the cumulative distribution functions to build these matrices.

Segmentation of the Loss Ratio Probability Exceedance Matrix

As previously mentioned, the GEM1 risk engine takes into account the fact that a small number of rows on the loss ratio probability/exceedance matrix might decrease the accuracy of the results. The following example shows the difference that might occur when not considering a proper number of loss ratio intervals. A lognormal distribution of loss ratios for a certain IML with a mean of 0.07 and coefficient of variation of 1.0 was assumed.

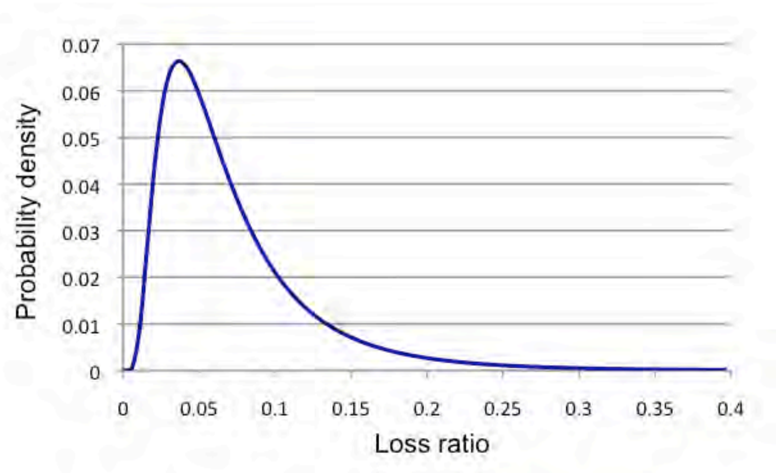


Figure B.8 Lognormal distribution of the loss ratios for a certain IML.

A minimum loss ratio of 0.0 and a maximum of 0.36 were assumed. Increments of 0.03 and 0.06 were used in case 1 and case 2, respectively. After the segmentation of the loss ratios, the probability of exceedance of each loss ratio was computed and the results are shown in the following table.

Assuming that the probability of exceedance of a loss ratio of 0.1 is required, the results would be:

- Case 1:

$$PE(LR = 0.1) = \frac{(0.12 - 0.10) \times 0.239 + (0.10 - 0.09) \times 0.136}{0.12 - 0.09}$$
$$PE(LR = 0.1) = 0.205$$

- Case 2:

$$PE(LR = 0.1) = \frac{(0.12 - 0.10) \times 0.435 + (0.10 - 0.06) \times 0.136}{0.12 - 0.06}$$
$$PE(LR = 0.1) = 0.236$$

Table B.5 Fraction of the loss ratio exceedance matrix for case 1 and 2.

Case 1

0.00	1.000
0.03	0.780
0.06	0.435
0.09	0.239
0.12	0.136
0.15	0.081
0.18	0.050
0.21	0.032
0.24	0.021
0.27	0.014
0.30	0.010
0.33	0.007
0.36	0.005

Case 2

0.00	1.000
0.06	0.435
0.12	0.136
0.18	0.050
0.24	0.021
0.30	0.010
0.36	0.005

The exact solution was also calculated for comparison purposes and a value of 0.197 was achieved. This represents an error of 3.8% for the first case and 19.6% on the second case. This error can increase for distributions with smaller standard deviations or in loss ratio matrices with a smaller number of rows. It can be concluded that a great number of rows should be used in these matrices. Usually the distribution of loss ratios for a particular IML is not equally distributed. It tends to have different spaces between smaller loss ratios and higher loss ratios. For this reason, it was decided to respect the distribution of loss ratios from the vulnerability function (assuming that the vulnerability has been defined in a discrete way), and only introduce intermediate intervals between those limits, for instance:

– Loss ratio distribution from the vulnerability function:

IML	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
LR	0.003	0.011	0.043	0.070	0.090	0.107	0.121	0.133

– Loss ratio distribution on the extended loss ratio matrix:

LR	0.000	0.0015	0.003	0.007	0.011	0.027	0.043	0.057	0.070	0.080	0.090	0.099	0.107	0.114	0.121	0.127	0.133

The segmentation of the loss ratios from the vulnerability function greatly influences the computational performance of the risk engine and therefore, it is important to keep a balance between the time required to obtain the results and reliability of the values. In order to evaluate the influence of this parameter, several tests were done using Iran as the region of interest. The computing time required to calculate a loss ratio map using different numbers of intervals is represented on the following figure:

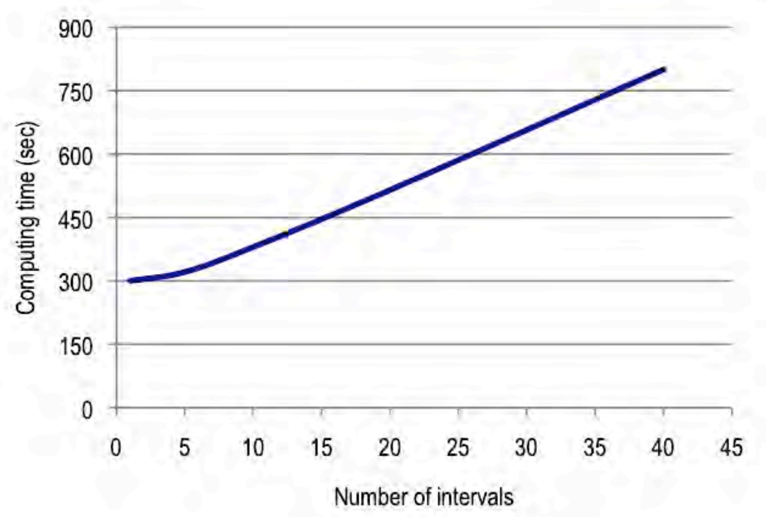


Figure B.9 Lognormal distribution of the loss ratios for a certain IML.

The progression of the curve is clearly linearly and therefore, the use of more efficient hardware resources can shorten the time required to obtain results. At the moment the risk engine is using 5 intermediate steps between consecutive loss ratios, which means that the LRPM will have a number of rows equal to 5 times the number of loss ratios defined within the vulnerability function.

Loss Ratio Curve (LRC)

A LRC gives the relation between a list of loss ratios and the corresponding probabilities of exceedance within a certain time span, which is the same as that used to compute the hazard curve. To compute this curve, the risk engine uses the loss ratio exceedance matrix and the probability of occurrence of each intensity measure level defined in the vulnerability model. In the following example a vector with the probability of occurrence of each level of intensity and a LREM defined at 4 IML's are assumed.

Table B.6 Probability of occurrence of each IML in 50 years.

IML	5.0	6.0	7.0	8.0
PO(IML)	0.138	0.099	0.068	0.041

Table B.7 Results for the calculation of the probability of exceedance of every loss ratio.

LR	**LREM**					**LREM x PO(IML)**					**PE(LR)**
0.00	1.000	1.000	1.000	1.000	⇨	0.138	0.099	0.068	0.041	⇨	0.346
0.03	0.266	0.510	0.841	0.999		0.037	0.050	0.057	0.041		0.185
0.06	0.099	0.270	0.587	0.980		0.014	0.027	0.040	0.040		0.120
0.09	0.039	0.143	0.370	0.908		0.005	0.014	0.025	0.037		0.082
0.12	0.017	0.078	0.223	0.781		0.002	0.008	0.015	0.032		0.057
0.15	0.008	0.044	0.133	0.626		0.001	0.004	0.009	0.026		0.040
0.18	0.004	0.026	0.080	0.475		0.001	0.003	0.005	0.020		0.028
0.21	0.002	0.015	0.048	0.346		0.000	0.002	0.003	0.014		0.019
0.24	0.001	0.010	0.029	0.246		0.000	0.001	0.002	0.010		0.013
0.27	0.001	0.006	0.018	0.171		0.000	0.001	0.001	0.007		0.009
0.30	0.000	0.004	0.011	0.117		0.000	0.000	0.001	0.005		0.006

LR	LREM				LREM x PO(IML)				PE(LR)
0.33	0.000	0.003	0.007	0.080	0.000	0.000	0.000	0.003	0.004
0.36	0.000	0.002	0.005	0.054	0.000	0.000	0.000	0.002	0.003
0.39	0.000	0.001	0.003	0.037	0.000	0.000	0.000	0.002	0.002
0.42	0.000	0.001	0.002	0.025	0.000	0.000	0.000	0.001	0.001
0.45	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Several ranges of loss ratio were considered in order to have a smooth distribution of the probability of exceedance. Considering what has been established above, the probability of exceeding a certain loss ratio within a time span can be given by the following formula:

$$PE(LR_n) = \sum_{i=1}^{n} PE\left[LR(IML_n)\right] \times PO(IML_n) \tag{B.26}$$

The following figure shows the curve computed for the previous example (noting that in this case the time span is 50 years, but it could be any length of time depending on the time span considered in the hazard calculations):

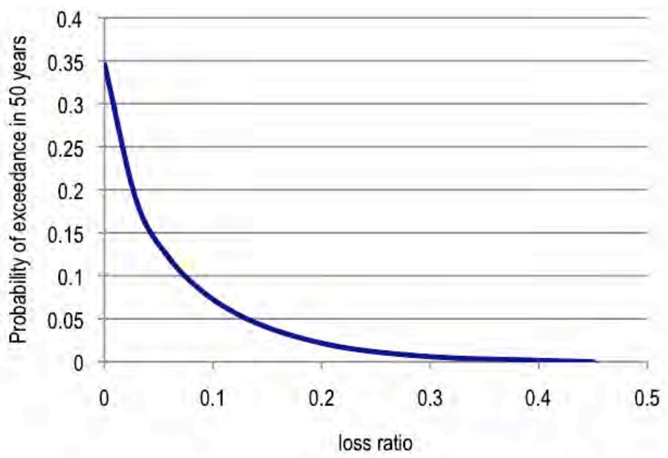


Figure B.10 Loss ratio exceedance curve.

Logarithmic scales are often applied in both axes in order to better view the relation between these two parameters.

Loss Ratio Map (LRM)

Using the loss ratio curve, it is possible to compute conditional loss ratio maps that provide the estimated loss ratio at every site conditional on a given probability of exceedance within a certain time span. The following figure presents the loss ratio curves of a simple region with just three different sites:

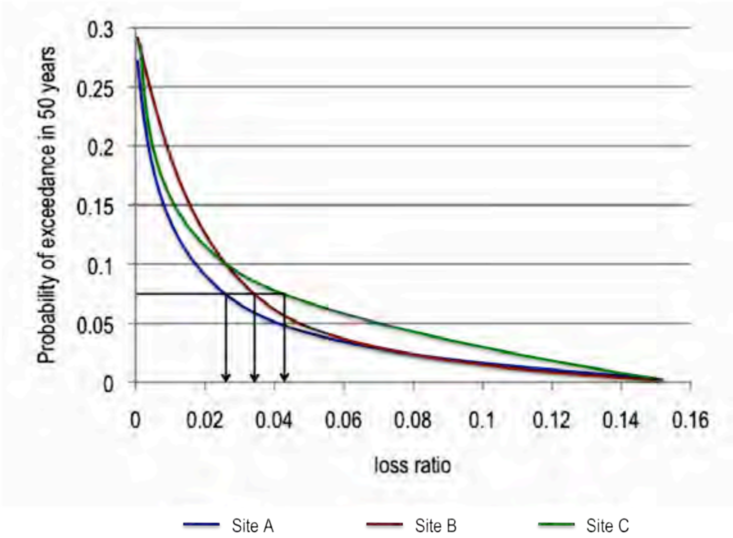


Figure B.11 Loss ratio curves for the region of interest.

To compute this type of map it is necessary to calculate the loss ratio for each site at the required probability of exceedance. It is quite probable that the probabilities where the loss ratio curve is defined, will not be equal to the one specified to build the map and therefore, linear interpolation methods need to be adopted. The following formula presents this procedure:

$$L_n = \frac{(PE_{n-1} - PE_n) \times L_{n+1} + (PE_n - PE_{n+1}) \times L_{n-1}}{PE_{n-1} - PE_{n+1}} \tag{B.27}$$

Where PE_n represents the probability of exceedance for which the loss map is being created. The following example shows how the loss ratio for a 10% probability of exceedance in 50 years is being computed:

LR	PE(LR)	
0.21	0.131	
0.24	0.108	
0.27	0.089	⇦10%
0.30	0.066	

$$L_{10\%} = \frac{(0.108 - 0.100) \times 0.27 + (0.100 - 0.089) \times 0.24}{0.108 - 0.089}$$

$$L_{10\%} = 0.25$$

Loss Curve (LC) and Loss Map (LM)

Relating the value of the elements exposed at each site with the previous parameters, one can compute loss curves and loss maps. A LC gives the probability of exceedance of a given loss within a certain time span. To convert a loss ratio curve into a loss curve, the risk engine only multiplies the exposure value for the respective loss ratios as expressed in the following formulas:

$$PE(L_n) = PE(LR_n) \tag{B.28}$$

Where:

$$L_n = LR_n \times Exposure \tag{B.29}$$

Expression B.29 is also being used to convert loss ratio maps into loss maps. The following figures show the difference between these two types of maps:

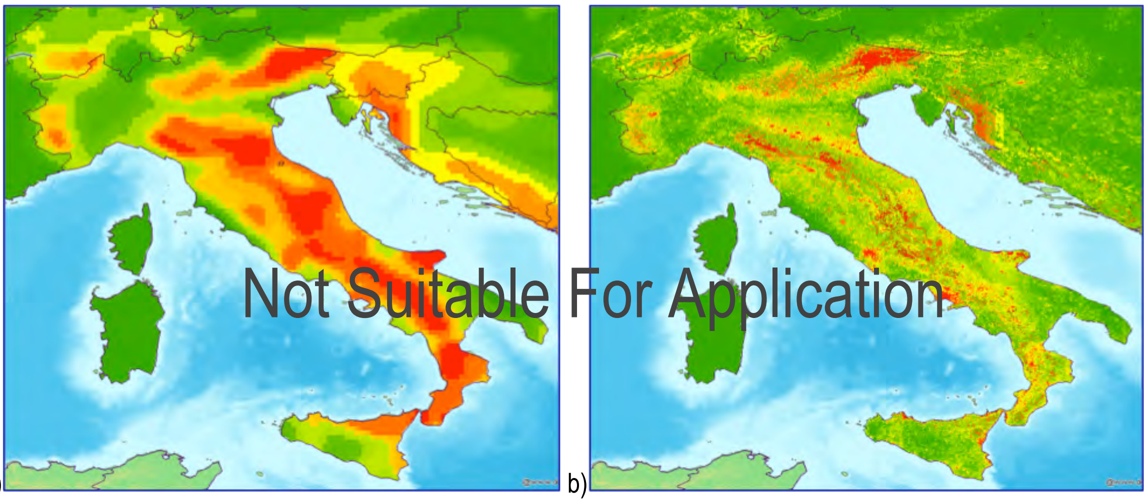
a) b)


Figure B.12 a) Loss ratio map b) loss map for a time span of 50 years with a 10% PE.

Mean Loss (ML)

Another important parameter used in probabilistic risk analysis is the mean loss within a certain time span for a particular site. To compute this value, it is necessary to calculate the probabilities of occurrence of a set of levels of losses. As previously mentioned, when computing probabilities of occurrence of a certain value, it is necessary to define an upper and lower bound. The accuracy of this procedure depends considerably into how many intervals the list of losses will be divided. The following formula can be used to compute the ML:

$$ML = \sum_{n=1}^{m} L_n \times PO(L_n) \tag{B.30}$$

The probability of occurrence of a particular range of losses can be computed using the values from the loss curve. However, it is important to understand that this curve provides probability of exceedance and not probability of occurrence and therefore, a procedure to convert the former to the latter is required. The following example explains how the probability of occurrence of a particular loss (L_n) can be computed from a LC:

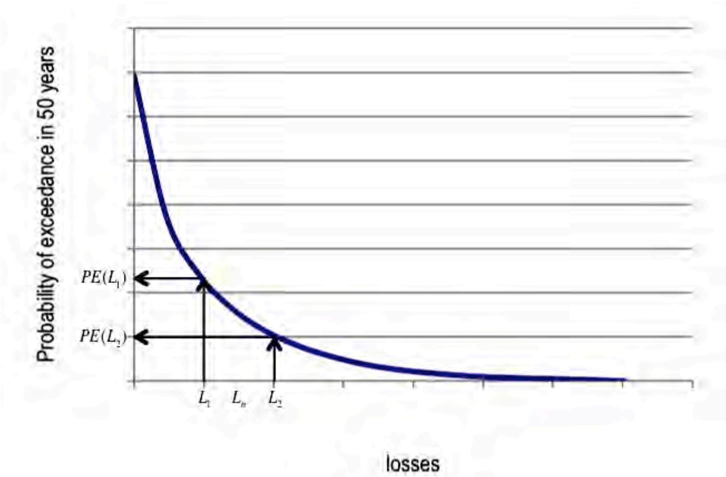


Figure B.13 Procedure to compute the probability of occurrence of an interval of loss.

$$PO(L_n) = PE(L_1) - PE(L_2) \tag{B.31}$$

To increase the accuracy of this procedure, several ranges of losses should be considered and hence, the probability for every loss limit needs to be interpolated. The following expression (based on linear interpolation) can be used to compute the PE of any loss:

$$PE(L_n) = \frac{(L_{n+1} - L_n) \times PE(L_{n-1}) + (L_n - L_{n-1}) \times PE(L_{n+1})}{L_{n+1} - L_{n-1}} \tag{B.32}$$