Development and Application of OpenQuake, an Open Source Software for Seismic Risk Assessment

V. Silva
University of Aveiro, Portugal

H. Crowley, M. Pagani, R. Pinho
GEM Foundation, Pavia, Italy

D. Monelli
ETH, Zurich, Switzerland

SUMMARY

Risk analysis is a critical link in the reduction of casualties and damages due to earthquakes. Recognition of this relation has led to a rapid rise in demand for accurate, reliable and flexible risk assessment numerical tools and software. As a response to this need, the Global Earthquake Model (GEM) initiated the development of an open source platform called OpenQuake for calculating seismic hazard and risk at any scale. In this paper a description of the architecture of OpenQuake is provided, highlighting the current workflow of the main loss calculators. In addition, a case study is presented using the Marmara Region (Turkey) for the calculations.

Keywords: seismic risk, loss assessment, software, open source

1. INTRODUCTION

The OpenQuake project was initiated as part of the Global Earthquake Model (GEM [1]), a global collaborative effort that brings together state-of-the-art science and national/regional/international organizations and individuals with the aim of establishing uniform and open standards for calculating and communicating earthquake risk worldwide. In January 2009, GEM launched a pilot project, GEM1, in which a number of existing hazard and risk software applications were evaluated allowing the first scientific requirements of OpenQuake to be defined (Danciu *et al*. 2010; Crowley *et al*. 2010). Currently, OpenQuake is in its second year of development and it is comprised of a number of calculators, but just the three main ones will be described herein: a scenario risk calculator capable of computing losses and loss statistics due to a single event (deterministic scenario) for a collection of assets (e.g.: buildings, population); a probabilistic event-based risk calculator that estimates the probability of exceedance of certain levels of loss in a given time span based on stochastic event sets, and finally, a classical PSHA-based risk calculator that allows the computation of probability of losses and loss statistics for single assets, based on probabilistic hazard curves. Such functionalities are fundamental in order to support activities such as emergency management planning, raising societal awareness of risk, identification of areas with a high seismic risk or estimation of the expected economic or human losses for the up coming years. Other calculators (not described in detail in this paper) include a scenario damage calculator for estimating the distribution of damage for a scenario event and a retrofitting benefit-cost ratio calculator.

The case study that is presented herein is applied to the Marmara Region (Turkey). In this exercise, the distribution of losses has been computed for an event of moment magnitude M_w 7.1 with a location under the Sea of Marmara, as well as a probabilistic assessment for a 50 years time span, considering the reinforced concrete building stock of the Metropolitan Area of Istanbul.

2. OPENQUAKE: SEISMIC HAZARD AND RISK SOFTWARE

2.1. Foreword

OpenQuake is an open-source software written in Python and Java for calculating seismic hazard and risk at any scale. Its code is under a Affero General Public License (AGPL). It currently makes use of a number of other, independent, open-source projects such as OpenSHA (Field *et al.*, 2003), as well as Celery [2] and RabbitMQ [3]. It is hosted on GitHub [4], a web-based hosting service for software development projects that allows for collaborative development on the same code base. The main results currently produced by OpenQuake (version 0.7) are the following:

- *Hazard curves*: curves providing probabilities of exceedance in a given time span for given values of a ground motion parameter.
- *Hazard maps*: maps describing the geographic distribution of values of a ground motion parameter with a fixed probability of exceedance in a given time span.
- *Stochastic Event Sets*: sets of earthquake ruptures - occurring in a given time span - obtained through random sampling of an earthquake rupture forecast.
- *Ground-Motion Fields*: each ground-motion field describes the geographic distribution of a scalar ground-motion parameter obtained considering an earthquake rupture and a ground-motion prediction equation (GMPE); the spatial correlation of the intra-event residuals can be considered in the generation of these fields.
- *Seismic hazard disaggregation* in terms of longitude, latitude, magnitude, epsilon and tectonic region.
- *Loss exceedance curves*: curves describing losses versus probability of exceedance in a given time span; losses can refer to single assets or can be aggregated where multiple assets exist.
- *Loss maps*: maps describing the geographical distribution of values of loss with a fixed probability of exceedance in a given time span.
- *Mean loss maps*: maps describing the geographic distribution of mean loss within a given time span (e.g. the average annual loss).
- *Collapse maps*: maps describing the distribution of percentage/number of collapsed buildings.
- *Damage/Loss statistics*: per event or across all events (mean loss, standard deviation of loss, mean damage distribution etc.).
- *Benefit-Cost Ratios*: based on the average annual loss with and without retrofitting.

2.2 OpenQuake Risk Calculation Workflows

This paper focuses on three risk calculation workflows within OpenQuake: one computing losses due to a single event, and the other two computing seismic risk due to most or all of the possible events that might occur in a given region within a certain time span. The calculation workflows are comprised of a number of separate calculators. In order to run any of the calculation workflows, it is necessary to define the geographic coordinates of the region of interest, the type of calculations, the path to the input files, the type of results that are to be produced and several parameters necessary for the hazard calculations. Currently, a configuration file to be provided to OpenQuake incorporates this information.

With regards to the seismic hazard input, the first risk calculation workflow requires the definition of a finite earthquake rupture whilst for the other two calculation workflows, a PSHA input model is required. This latter input is comprised of two files: one describing the seismic source system (i.e. the combination of one or several initial seismic source models and a logic tree structure) the other specifying the GMPE logic tree.

All of the risk calculators require an exposure model and a vulnerability model. The first input contains the information about the location, value and taxonomy of the assets, while the second comprises a vulnerability function (probability of loss ratio for a set of intensity measure levels) for each type of asset.

A comprehensive description of the methodologies included in OpenQuake can be found in a document - called the OpenQuake Book (GEM Foundation, 2011). In the following sections, a summary description of the properties characterizing each risk calculation methodology is provided.

Scenario Risk Calculation Workflow
This calculation sequence is capable of computing losses and loss statistics due to a single, deterministic earthquake, for a collection of assets. Such analyses are of importance, for example, for emergency management planning and for raising societal awareness of risk.

The hazard input consists of a finite rupture and a single GMPE. By repeating the same rupture, and sampling the inter- and intra-variability from the GMPE each time, many ground motion fields can be computed to account for the aleatory variability in the ground motion. During the generation of each ground motion field, the spatial correlation of the intra-event variability can be considered (currently based on the model of Jayaram and Baker, 2009), so that assets located close to each other are likely to have similar ground motion levels (see e.g. Crowley *et al.*, 2008 for a summary of ground motion variability treatment in loss models). The set of ground motion fields is then provided to the Scenario Risk calculator, together with the vulnerability and exposure models, to compute the losses for each asset in the exposure model, per ground motion field. The mean or median value of losses across all ground motions fields can be found for a given asset, and the spatial variation of this value for a given asset typology can be plotted in a loss map. The losses to all assets across the region of interest can also be aggregated per ground motion field, to obtain a list of aggregated losses, which can then be used to compute the mean and standard deviation of the aggregated losses. The workflow in Figure 2.1 describes this procedure.
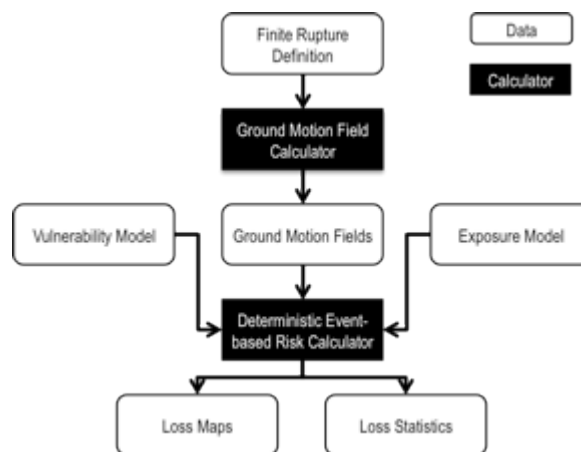

Figure 2.1. Workflow of the Scenario Risk calculator.

Probabilistic Event-Based Risk Calculation Workflow
This calculation workflow computes the probability of losses and loss statistics for a collection of assets, based on the probabilistic hazard. The losses are calculated with an event-based approach, such that the simultaneous losses to a set of assets can be calculated.

This workflow requires a number of calculators in order to calculate ground motion fields. Firstly, a Logic Tree Processor calculator uses information contained within the seismic source system together with a Monte Carlo approach to sample the logic tree structure and produce a seismic source model (SSM). Each seismic source model computed is used by the Earthquake Rupture Forecast (ERF) calculator to produce a list of all the possible ruptures occurring on all the sources in the SSM; each rupture is associated with a probability of occurrence in the time span specified by the user in the configuration file. Then, the Stochastic Event Set calculator uses the ERF to create one or several groups of ruptures. Each group represents a possible realization of the seismicity generated in the specified time span by the entire set of seismic sources included in the seismic source model. Afterwards, the Logic Tree Processor is again used to process the GMPEs system and provide the

ground motion relationship that shall be used by the Ground Motion Field calculator, together with each earthquake rupture, to compute the ground motion values at a set of sites. The spatial correlation of the intra-event residuals of the ground motion model can be considered if specified on the configuration file. As mentioned previously, in that case, sites that are closer are more likely to have similar levels of ground motion. This set of ground motion fields is combined with the exposure and vulnerability model in the Probabilistic Event-Based Risk calculator, to compute the losses for each asset per ground motion field. The list of losses per asset can be used to build a cumulative histogram which gives the number of losses in selected bins of loss over the time span, from which the loss exceedance curve is computed (loss versus probability of exceedance in a given time span). This approach can be used to compute a loss curve for each asset within the exposure model, or by aggregating all the losses throughout the region per ground motion field, an aggregated loss curve representative of the whole set of assets within the exposure file is obtained. The workflow in Figure 2.2 describes this procedure.
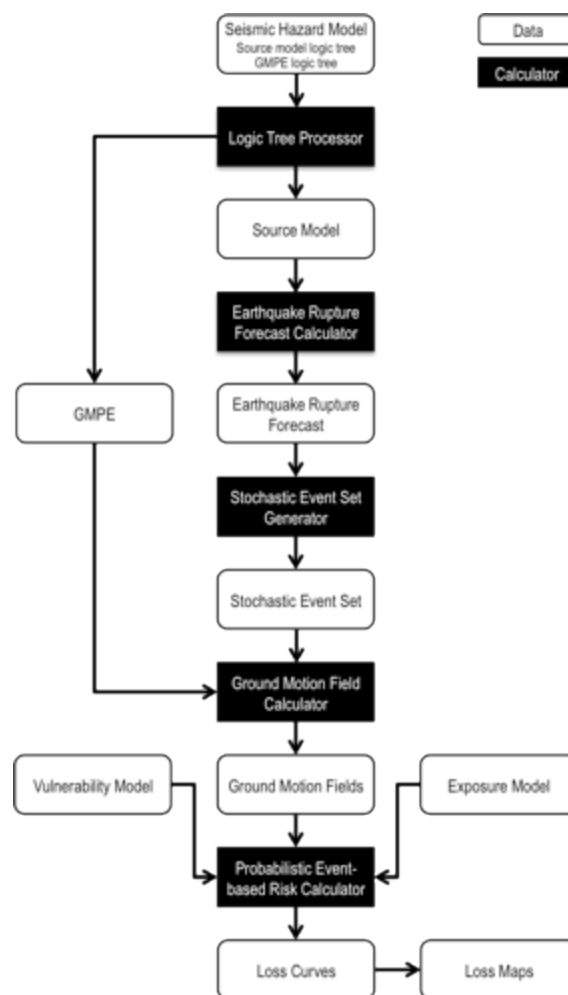

Figure 1.2. Workflow of the Probabilistic Event-Based Risk calculator.

Classical PSHA-Based Risk Calculation Workflow
This calculation workflow leads to the computation of the probability of losses and loss statistics for single assets, based on probabilistic hazard curves. The output of this calculator is useful for comparative risk assessment between assets at different locations.

This workflow has an initial architecture similar to the previous one, in which a Logic Tree Processor uses the structure defined in the Seismic Source System to provide the required parameters to the Earthquake Rupture Forecast (ERF) calculator, which produces a list of all the possible ruptures

occurring on all the sources included in the seismic hazard model. Then, using the GMPEs system, the Logic Tree processor states which GMPEs the Classical Hazard Curves calculator will use. This calculator uses the classical PSHA approach (Cornell, 1968, McGuire, 2004) following the methodology presented by Field *et al.* (2003) to compute a hazard curve at each site. This set of hazard curves is then provided, together with the vulnerability and exposure model to the Classical PSHA-based Risk calculator. Here, the first step is to convert each discrete vulnerability function into a loss ratio exceedance matrix (e.g. a matrix which describe the probability of exceedance of each loss ratio for a discrete set of intensity measure levels). Once these matrices are built, the values of each column are multiplied by the probability of occurrence of the associated intensity measure level. This probability is extracted from the previously computed hazard curves. Finally, the list of probabilities of exceedance of the loss ratio curve is obtained by summing all the values per loss ratio. This loss ratio curve is then converted into a loss curve by multiplying each loss ratio by the associated asset value. The workflow in Figure 2.3 describes the architecture of this calculator.
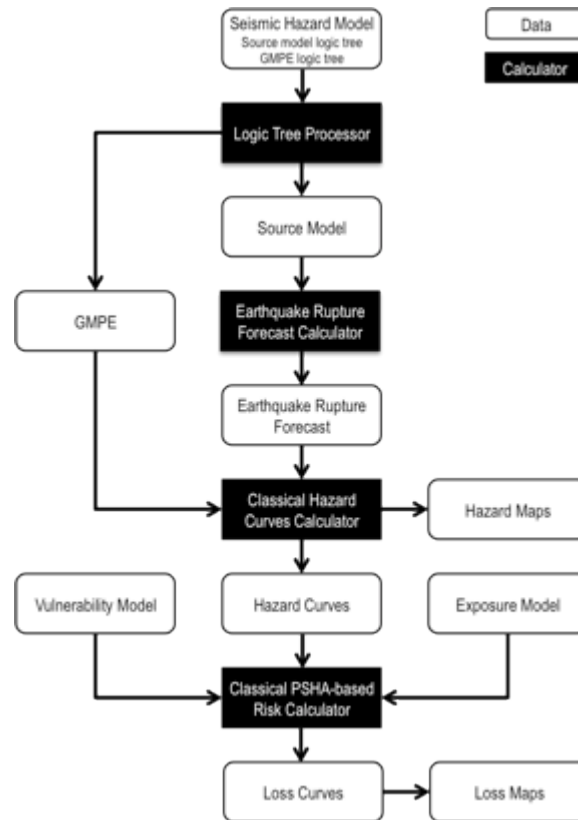

Figure 2.3. Workflow of the Classical PSHA-Cased Risk calculator.

3. CASE STUDY APPLICATIONS

3.2 PSHA model

The seismic hazard input data utilized to exercise the OpenQuake calculators comes from a preliminary seismic hazard model developed for Turkey (Demicioglu *et al.*, 2008). The PSHA model consists of a seismic source model based on two source typologies: faults (Figure 3.1) and areas (Figure 3.2). The ground motion model contains a logic tree consisting of three GMPEs: Boore and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008). All three GMPEs received an equal weight.
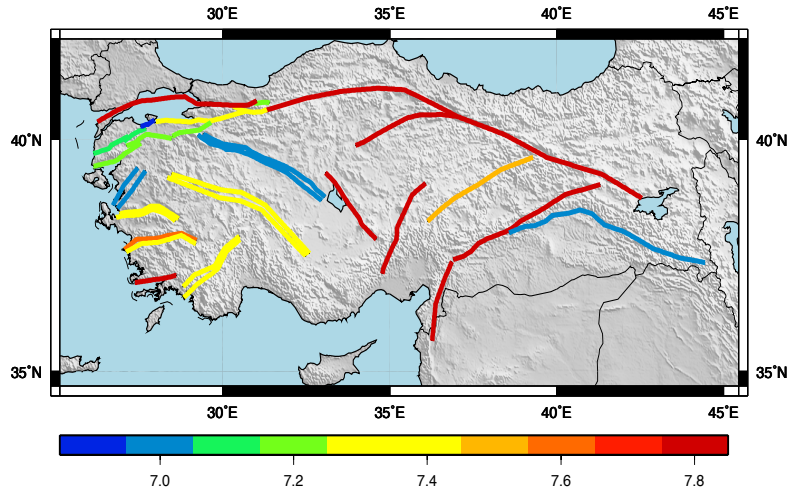
Figure 3.1 Fault source model for Turkey. Faults are assumed to be vertical, so only fault traces are shown. Colours represent maximum magnitude (M_w).
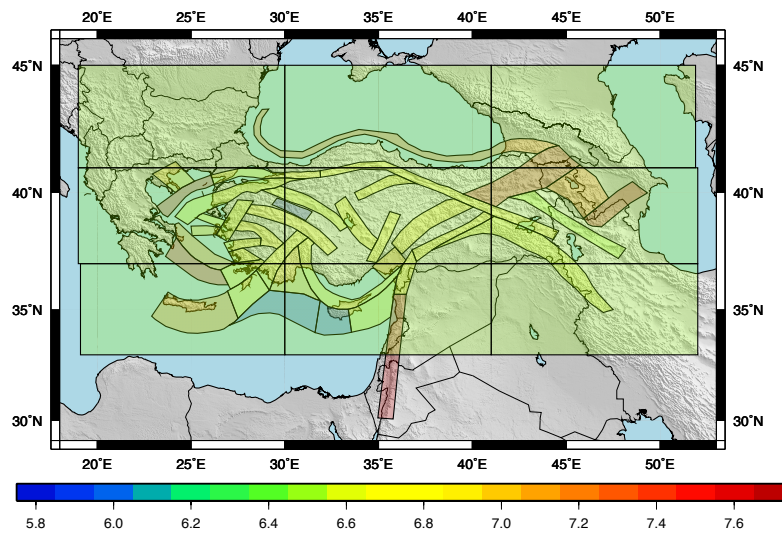

Figure 3.2. Area source model for Turkey. Large-scale rectangular background sources cover the entire country, whilst most of the small-scale area sources follow fault source geometries.

3.3 Scenario Rupture Model

The scenario rupture model for the city of Istanbul considers a single strike-slip rupture equivalent to a magnitude of 7.5. The rupture extends for 120 km along the North Anatolian fault, on a section close to the Bosphorus strait.

3.4 Building Exposure Model

The building inventory for the metropolitan area of Istanbul that was obtained had been created using a combination of data from aerial photos taken in 1995 and 1998 and census data from 2000 conducted by the Turkish State Statistics Institute (BU-ARC, 2002). For the purposes of the current application, only the RC frame buildings have been considered. The dataset uses an evenly spaced grid with a 0.005×0.005 decimal degree spatial resolution and for each grid cell, the number of buildings for each typology is provided. In order to understand the distribution of building value throughout the metropolitan area of Istanbul, the economic value of the buildings was aggregated per grid cell and the results are illustrated in Figure 3.3.
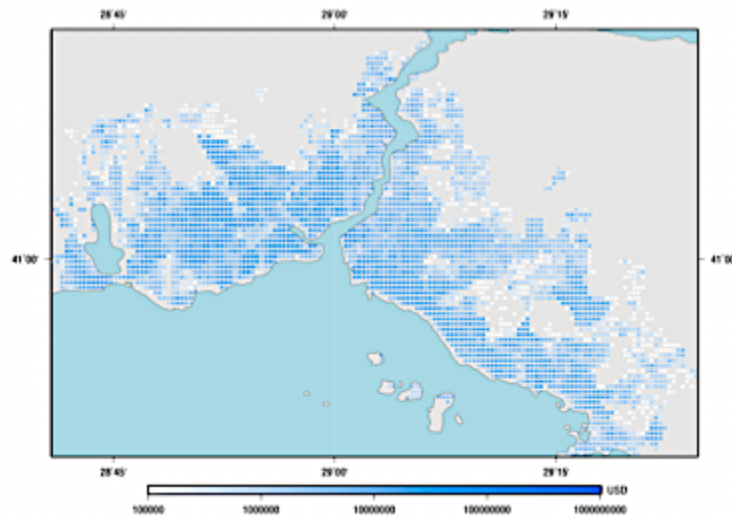
Figure 3.3. Distribution of the value of RC frame buildings in the metropolitan area of Istanbul.

3.5 Vulnerability Model

Only reinforced concrete frame buildings have been considered in the case study application as they constitute 75% of the building stock in Istanbul and this percentage is even higher when one considers the building value rather than count (BU-ARC, 2002). Several studies have been carried out in the past to calculate fragility functions of typical Turkish buildings (Akkar *et al*., 2005, Hancilar *et al*., 2006, Kirçil and Polat, 2006, Erberik, 2008, Ozmen *et al*., 2010), many of which can be employed in OpenQuake. The vulnerability model used in this exercise was developed by Silva *et al*. (2012) and it is composed by 6 vulnerability functions (one per asset typology).

3.6 Output

Seismic Hazard Map for Turkey
Hazard curves have been computed from 35.0 to 43.0 degree north, and from 25.0 to 47.0 degree east, every 0.05 degrees. A total of 71001 hazard curves have been derived for each GMPE. As can be seen, the hazard is mostly driven by fault sources, especially the North Anatolian fault, with levels of PGA of about 1.3g along the fault trace. Area sources surrounding fault sources also play an important role. Their effect it to widen the region of significant hazard around fault sources. Large-scale background area sources produce instead a rather stable value of PGA of about 0.2g in all locations that are far from small-scale area sources or fault sources.
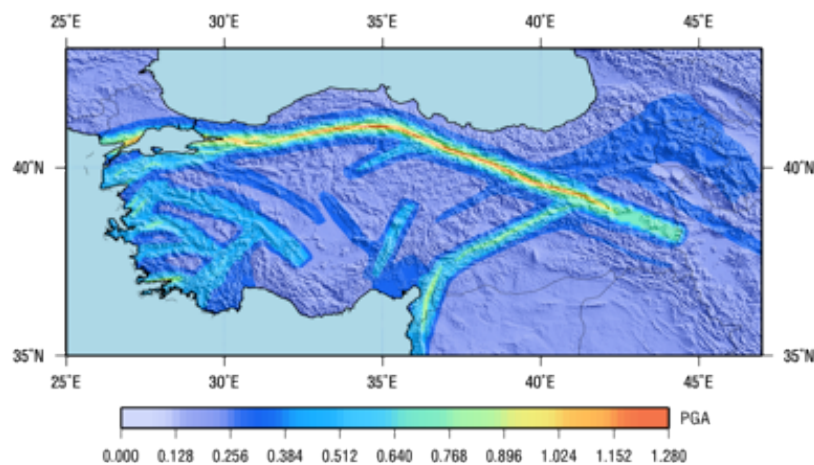

Figure 3.4. Mean hazard map for Turkey (corresponding to a probability of exceedance of 10% in 50 years).

Figure 3.5 presents a loss map with the distribution of mean economic losses (across all ground motion fields) for the reinforced concrete buildings. When many building typologies existed simultaneously in a given grid cell, the loss values for each typology were aggregated per event.
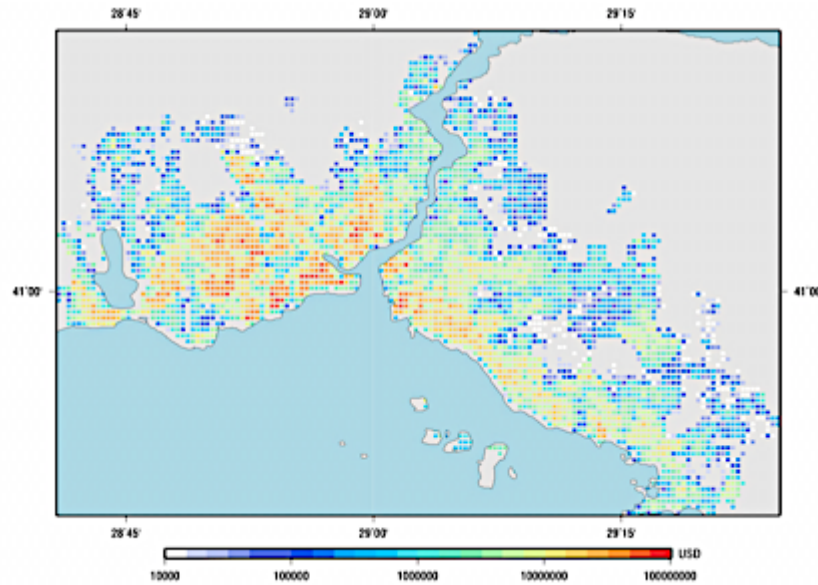

Figure 3.5. Loss map with the distribution of mean economic losses for reinforced concrete buildings.

For this seismic event, it was estimated that a total mean economic loss equal to 20.39 billion USD (with a standard deviation of 7.60 billion USD) would occur, representing about 26% of economic value of the reinforced concrete building stock.

Probabilistic risk assessment for Istanbul
In total, due to the different building typologies presented within the exposure model, about 28,488 loss curves were computed for 4,107 different locations. The spatial correlation of the intra-event variability was considered (using the model of Jayaram and Baker, 2009), allowing the possibility of aggregating the losses throughout the region and the calculation of a total loss exceedance curve representative of the whole RC building stock, as shown in Figure 3.6.
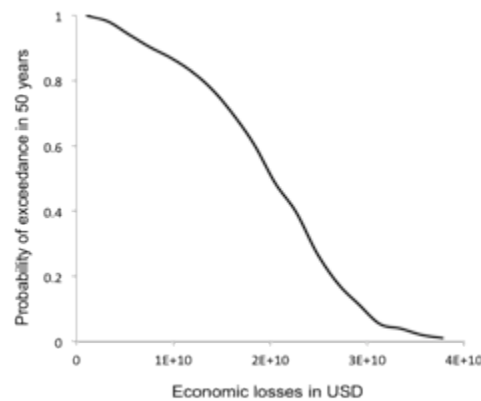

Figure 3.6. Total loss exceedance curve for RC buildings in Istanbul.

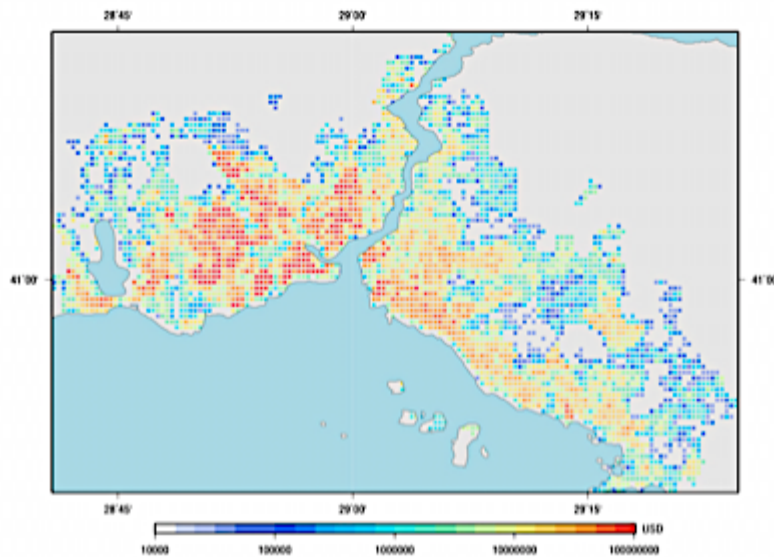In Figure 3.7, a loss map for a probability of exceedance of 10% in 50 years is presented.

Figure 3.7. Loss map for a probability of exceedance of 10% in 50 years.

It is recalled that the losses that have been presented in this section are purely exemplificative, and that they only refer to the RC building stock. Interested readers may refer to Erdik (2007) for a discussion on the expected losses in Istanbul, which shows that the values presented herein are within the range of estimates that have been previously made, though they are at the upper end of this range. The vulnerability model used herein needs further calibration and testing before being used in real applications.

4. CONCLUSIONS

In this paper, an open source software capable of computing seismic hazard and risk was presented, with focus given to three main risk calculation workflows that are currently supported to compute losses either due to a single event, or due to probabilistic hazard using both classical hazard-based and event-based methods.

Due to its transparent, modular and test-driven development philosophy, OpenQuake aims to be a community effort in which anyone can contribute with their own methods and formulae. This differs from traditional practice, where a closed "enterprise" development tends to be followed, even if the source code is eventually openly released. Any interested researcher or scientist that would like to include a methodology relevant to seismic risk calculation within the OpenQuake software should contact the developers [5].

Through the case study applications presented herein for Turkey, it has been demonstrated that OpenQuake – despite still being in a development phase – already offers several functionalities and a wide spectrum of tools for seismic risk assessment.

REFERENCES

Akkar, S., Sucuoglu, H., and Yakut, A. (2005) "Displacement-based fragility functions for low- and mid-rise ordinary concrete buildings", *Earthquake Spectra* 21(4), 901-927.
Boore, D. M. and G. M. Atkinson (2008). "Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s", Earthquake

Spectra 24, 99—138.

BU-ARC (2002). ''Earthquake Risk Assessment for Istanbul Metropolitan Area,'' Project Report, Boğaziçi University and American Red Cross, Boğaziçi University Publications, Istanbul, Turkey.

Campbell, K. W. and Bozorgnia, Y. (2008)."NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10s", *Earthquake Spectra,* 24(1), 139-171.

Chiou, B., and Youngs, R. R. (2008). "An NGA model for the average horizontal component of peak ground motion and response spectra". *Earthquake Spectra,* 24, 173–215.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58, 1583-1606.

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.

Danciu L., Monelli D., Pagani M. and Wiemer S. (2010). GEM1 Hazard: Review of PSHA software, GEM Technical Report 2010-2, GEM Foundation, Pavia, Italy

Demircioglu, M.B., K.Sesetyan, E.Durukal and M.Erdik (2007). "Assessment of Earthquake Hazard in Turkey" *Proceedings of the 4th International Conference on Earthquake Geotechnical Engineering*, June 2007, Thessaloniki, Greece.

Erberik, M., A. (2008) "Fragility-based assessment of typical mid-rise and low-rise RC buildings in Turkey" *Engineering Structures* 30(5), 1360-1374.

Erdik, M. (2007). Discussion of: ''Istanbul at the threshold: an evaluation of the seismic risk in Istanbul,'' by Pyper Griffiths, J.H., Irfanoglu, A, and Pujol, S., Earthquake Spectra, 23(1), 63–75.

Field, E.H., Jordan, T.H. and Cornell, C.A. (2003). OpenSHA: A Developing Community-Modeling Environment for Seismic Hazard Analysis, *Seismological Research Letters*, 74, 406-419.

GEM Foundation (2011) OpenQuake Book. Available from URL: http://openquake.org

Hancilar, U., E. Durukal, G. Franco, G. Deodatis, M. Erdik & A. Smyth. (2006). "Probabilistic Vulnerability Analysis: An Application to a Typical School Building in Istanbul". *Proceedings of 1st European Conference on Earthquake Engineering and Seismology* (1st ECEES), Geneva, Switzerland, 3-8 September 2006. Paper Number: 889.

Jayaram, N. and Baker, J. W. (2009). "Correlation model for spatially distributed ground-motion intensities". *Earthquake Engineering and Structural Dynamics* 38(15), 1687-1708

Kirçil, M., S. and Polat, Z. (2006) "Fragility analysis of mid-rise R/C frame buildings". *Engineering Structures*, Vol. 28(9), pp.1335-1345.

McGuire, R. R. (2004). Seismic hazard and risk analysis. Earthquake Engineering Research Institute Publication No. MNO-10, Second monograph series.

Ozmen, H., B., Inel, M., Meral, E. and Bucakli, M. (2010). "Vulnerability of Low and Mid-Rise Reinforced Concrete Buildings in Turkey", *Proceedings of the 14th European Conference of Earthquake Engineering*, Ohrid.

Silva, V., Crowley, H., Pagani, M., Monelli, D., Pinho, R. (2012). "Development and application of OpenQuake, an opensource seismic risk assessment software. *Earthquake Engineering and Structural Dynamics,* in press.

Web references
[1]   Global Earthquake Model - http://www.globalquakemodel.org
[2]   Celery project - http://celeryproject.org/
[3]   RabbitMQ project - http://www.rabbitmq.com/
[4]   OpenQuake repository - https://github.com/gem/openquake/tree/master/docs/schema
[5]   OpenQuake website – http://openquake.org