Development of Open Source Seismic Risk Modeling Framework

Prepared for

U.S. Geological Survey
12201 Sunrise Valley Dr MS 905
Reston VA 20192

Under award no
07HQAG0002

Draft submitted to Nico Luco on 6 Feb 2009
Final submittal date 12 Mar 2009

By
Keith Porter and Charles Scawthorn
SPA Risk LLC
2501 Bellaire St, Denver CO 80207
(626) 233-9758, www.sparisk.com

SPA Risk Project number 10006-01-06-01

Version history

Version	Date	Comment
0.9	6 Feb 2009	Initial draft for review
1.0	12 Mar 2009	Final Technical Report

EXECUTIVE SUMMARY

To manage catastrophe risk and to understand the potential impacts of new disaster science or policy, one typically must have access to analytical and computer risk models. These models are constantly in flux as science, engineering, and disaster social science develop, and many researchers and practitioners lack risk-integration tools and methods needed for an overall understanding of risk. They must either re-develop existing integrative software or abandon potentially fruitful study. Current end-to-end risk models have been developed in a hierarchical paradigm and cannot nimbly respond to emerging knowledge and data. In response, a new paradigm of open risk analysis (ORA) is emerging, promoted by AGORA (the Alliance for Global Open Risk Analysis), and which includes development of a number of open-source seismic-risk-related software codes (OpenSHA, OpenRisk, OSRE, MIRISK; see www.risk-agora.org). OpenRisk is a family of general catastrophe-risk modeling software for use by researchers and practitioners. In the work discussed here, several initial software applications and databases were developed:

- A single-site earthquake loss exceedance calculator
- A single-site earthquake risk-mitigation benefit-cost ratio calculator
- A portfolio file import tool and editor
- A portfolio earthquake expected annualized loss calculator
- A calculator designed to create fragility functions from damage data (sponsored by others).
- A gridded wind hazard database covering the conterminous United States and Alaska
- A set of seismic vulnerability functions relating indoor casualty rates to spectral acceleration
- A set cost vulnerability functions relating mean damage factor to spectral acceleration

These products are available for download from a website maintained by the Alliance for Global Open Risk Assessment (AGORA) at www.risk–agora.org.

CONTENTS

1 Introduction ... 5
1.1 Overview of cat modeling ... 5
1.2 Objectives and scope of openrisk ... 8
1.3 Organization of this report ... 10
1.4 Acknowledgments ... 10
2 Development approach ... 12
3 OpenRisk products ... 13
3.1 Design Document ... 13
3.2 BCR calculator ... 14
3.3 Loss-exceedance-frequency calculator ... 15
3.4 Fragility function calculator ... 17
3.5 Portfolio database import tool ... 17
3.6 Portfolio EAL calculator ... 20
3.7 Wind hazard database ... 21
3.8 Vulnerability function databases ... 26
4 Summary and conclusions ... 29
5 References cited ... 31

INDEX OF FIGURES

Figure 1. Analytical stages of a cat model ... 6
Figure 2. Input panes for BCR calculator beta version ... 14
Figure 3. Output from BCR calculator beta version ... 15
Figure 4. Input panes for loss-exceedance-frequency calculator ... 16
Figure 5. Output from loss-exceedance-frequency calculator beta version ... 16
Figure 6. Fragility function calculator ... 17
Figure 7. Initial screen of portfolio database import tool ... 18
Figure 8. Portfolio expected annualized loss calculator ... 21
Figure 9. ASCE 7-05 basic windspeed map for the southeastern United States (ASCE 2005) ... 23

INDEX OF TABLES

Table 1. Portfolio contents ... 19
Table 2. Spot-check tabulated windspeeds vs ASCE 7-05 ... 23
Table 3. Peterka et al. (1998) conversion factors $f(T)$ for other recurrence intervals ... 24
Table 4. Wind hazard data files ... 25
Table 5. Contents of casualty vulnerability function database ... 25
Table 6. Contents of casualty vulnerability function database ... 27
Table 7. Contents of building repair cost vulnerability function. ... 28
Table 8. Timeline for development of OpenRisk ... 29

1 INTRODUCTION

1.1 OVERVIEW OF CAT MODELING

Catastrophe risk modeling (often called cat modeling) refers here to the use of mathematical models encoded in computer software to estimate the performance of assets such as buildings subjected to various hazards. Performance is expressed in terms of economic costs, human safety, and loss of use ("dollars, deaths, and downtime"). Cat models are typically used for decision-making: by estimating risk, one can decide how to manage it, such as in the decision of whether and how much insurance or reinsurance to buy, or whether and how to mandate risk mitigation. A cat model can be seen as comprising four analytical stages (Figure 1):

Asset analysis characterizes the assets exposed to environmental excitation, given tabulated asset data. It can assign a geographic location and site characteristics such as NEHRP site soil classification to an asset based on its street address. It may estimate asset replacement cost based on square footage and occupancy type. Ideally it could quantify uncertainties in location, value, structure type, etc.

Hazard analysis characterizes the probabilistic environmental excitation imposed on those assets, such as windspeeds or earthquake shaking intensities. For example, one might estimate the mean frequency in events per year with which each asset experiences each of several levels of 1.0-second 5%-damped elastic spectral acceleration response.

Loss analysis estimates the physical damage or ground-up loss to the assets under consideration. For example, one might estimate the mean annual repair cost to the assets due to earthquake shaking.

Financial analysis relates the financial or human loss or other impacts experienced by a stakeholder in the assets. For example, one might estimate the portion of the annual repair costs borne by the assets' owner, after earthquake insurance claims are paid.

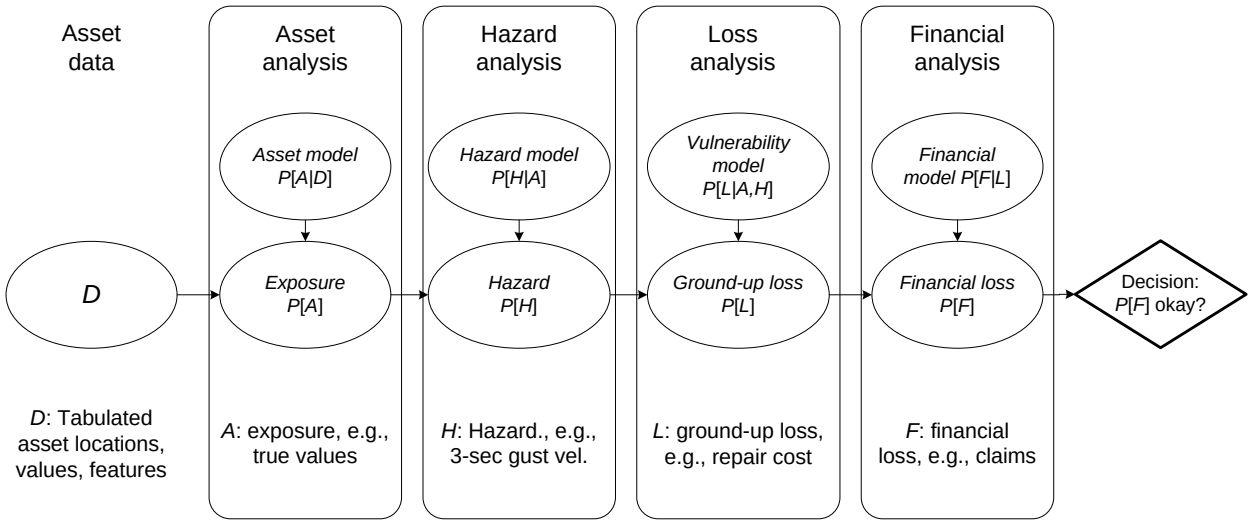


Figure 1. Analytical stages of a cat model

Most commonly the hazards considered in cat models include earthquake, hurricane, and flood; less commonly landslide, tornado, hail, winter storms, volcano, or blast are modeled. Scawthorn (2006) offers a summary of the history of computerized cat models, dating flood and hurricane risk models to the mid-to-late 1960s (the US Army Corps of Engineers flood model and Friedman's hurricane model for the Travelers Insurance Co), and seismic risk software to the early 1970s (the EPER model; Wiggins et al. 1976).

Cat models are typically used in the finance, insurance, and real estate industry to inform decisions related to mortgage underwriting (whether a lender should require earthquake insurance), insurance and reinsurance transactions (whether and how much insurance or reinsurance to buy, and at what price), and the creditworthiness of insurers and reinsurers (whether the investor in such a firm is likely to lose an investment because of the insurer's liability after a natural disaster). They are also used in the public sector to inform decisions about emergency planning and disaster response and mitigation. For example, MMC (2005) describes numerous examples of public mitigation efforts funded after assessing their probabilistic cost effectiveness; the study used the HAZUS software (NIBS and FEMA 2003) along with other tools to confirm for the US Congress that natural hazard mitigation saves more than it costs, on average $4 for every $1 spent in the period 1993-2003 under various FEMA programs.

More recently, researchers have developed new approaches to cat modeling to assess the socioeconomic implications of new science. The Southern California Earthquake Consortium (SCEC) has developed finite element models of the earth's crust and peta-scale computing

facilities capable of modeling seismic waves as they propagate from seismic sources (fault ruptures) to the earth's surface (Graves and Somerville 2006). These capabilities complement sophisticated open-source seismic hazard analysis software (www.opensha.org) that implements earthquake rupture forecast models, seismic attenuation relationships, and surficial geology maps to quantify hazard (Field et al. 2005). A pilot study leveraged these resources to demonstrate the ability to perform physics-based modeling of rupture, seismic wave propagation, and structural response of buildings throughout Southern California. The study (Porter et al. 2007b) compared structural response under physics-based modeling of ground motion with an approach that uses seismic attenuation relationships. The results suggest that risk analysis with physics-based modeling may be able to reduce uncertainty in earthquake insurance risk, as well as avoiding systematic bias in regions affected by ground motion directionality.

Cat-modeling software tends to involve sophisticated sub-models from several fields of expertise. For example, earthquake cat models involve seismology, geotechnical and structural engineering, and actuarial science, and can use sophisticated geographic information systems and computational engines requiring extensive programming expertise. They therefore tend to be costly to produce. For example, to develop HAZUS-MH (e.g., NIBS and FEMA 2003) has cost the US Federal Emergency Management Agency on the order of $50 to $100 million to develop. Several commercial cat models have been developed by firms specializing in the topic, most notably RMS, EQECAT, and AIR. Licenses to each of these models cost on the order of $1 million per year per seat, and the overall market for cat modeling is probably on the order of $1 billion annually.

Largely because of its value as intellectual property, the source code for these models is closed, i.e., unavailable to users. Although the principles and data involved in the HAZUS-MH software are largely public information, its source code too is unavailable for examination and modification. Commercial cat modelers allow outsiders to view their source code only under carefully controlled conditions such as in response to regulatory requirements (e.g., the Florida Commission on Hurricane Loss Projection Methodologies; http://www.sbafla.com/methodology/).

Because of the closed-source nature of cat models, researchers use them only under limited circumstances, usually without the ability to modify their underlying methods. Consumers of

commercial cat models likewise are unable to see or modify for themselves the underlying methods, and though the cat modelers go to great efforts to explain their models to their clients, the closed nature of commercial cat models can lead users to have serious concerns about the models' dependability, especially when different models produce dramatically different estimates of risk from the same input data, or when an update to a given model produces dramatically different estimates of risk from the prior version.

Kishi (2007) shows how the three principal cat modeling firms estimated industry losses from Hurricane Katrina that varied between them on the order of a factor of 3 at any given time, and changed by up to a factor of 5 over the space of a few weeks. The commercial cat models also change over longer timespans, with periodic new releases that incorporate the modeler's perception of the best new science and data. The modifications sometimes result in dramatic changes to modeled risk for a given portfolio. For reasons of commercial competitiveness and because of the effort involved, modelers sometimes limit their explanation about these modifications. Furthermore the new science can take some time to find its way into the models. For all these reasons, users of cat models commonly express the desire for new models whose methods and software are open to inspection and rapid modification, referred to here as *open risk analysis*.

Among several nearly simultaneously developed efforts to meet this need are OSRE and MIRISK, developed by Kyoto University researchers (see e.g., www.risk-agora.org/downloads.html); the US Geological Survey's PAGER (Prompt Assessment of Global Earthquakes for Response, http://earthquake.usgs.gov/eqcenter/pager/alert/); Geoscience Australia's Earthquake Risk Model (EQRM, Robinson et al. 2006); and *OpenRisk*, which is discussed here.

1.2 OBJECTIVES AND SCOPE OF OPENRISK

OpenRisk's mission is to develop object-oriented, web- and GUI-enabled, open-source, and freely available software code for conducting multihazard risk analysis. The resulting body of code and applications is also referred to as OpenRisk. What makes it different from other existing open-source cat models is its intimate linkage with the US Geological Survey's OpenSHA software. Developed largely by Field and his colleagues at the USGS and Southern

California Earthquake Center (SCEC; see for example Field et al. 2005), OpenSHA is a suite of small focused software applications to perform seismic hazard analysis with any combination of several state-of-the-art earthquake rupture forecasts and ground motion prediction equations. The initial OpenRisk software embeds OpenSHA software classes and adds risk-analysis elements. Other OpenRisk applications and data, as will be described later, are unrelated to seismic hazard analysis.

OpenRisk's development began with the creation of a design document, discussed below. It is currently archived (and periodically updated) at www.risk-agora.org. The design document was produced by SPA Risk LLC during 2006-2007 with funding from the Southern California Earthquake Center. Actual working software was created in 2007-2008 with funding from the US Geological Survey. Our objectives for OpenRisk in 2007-2008 were:

- To create a framework for the development, collection, and dissemination of open-source seismic risk modeling, and eventually multihazard risk modeling.
- To make all code and data available to users.
- To foster the use and development of OpenRisk among the earthquake community.
- To integrate OpenRisk with the USGS's OpenSHA, ResRisk, and other risk-related projects currently underway at the USGS in Pasadena, CA and Golden, CO.
- To produce seed objects sufficient to perform deterministic and probabilistic seismic risk assessments.
- To have OpenRisk complement (not compete with) HAZUS, commercial software, and USGS hazard and risk products.

As noted above, OpenRisk largely builds upon the USGS and SCEC's OpenSHA effort (www.opensha.org), including several components of its seismic hazard analysis software. OpenRisk also comprises three databases: one of gridded US wind hazard, the other two of US seismic vulnerability functions—see Porter 2009a and b for the latter. The suite of software includes one general-purpose fragility function calculator (implementing the ATC-58 fragility methodologies presented in Porter et al. 2007), and three applications related to Southern California earthquake: a single-site benefit-cost ratio calculator, a single-site loss-exceedance-

curve calculator, and two applications to read and edit portfolio databases and to calculate the expected annualized loss for a portfolio of assets.

The fragility calculator was prepared by a Caltech graduate student paid by the Applied Technology Council; the two single-site calculators were programmed by USGS and SCEC developers; and the portfolio editor and EAL calculator were prepared by Instrumental Software Technologies, Inc. (ISTI), under a subaward from SPA Risk LLC.

OpenRisk's seismic-risk software is currently largely limited to the Southern California earthquake peril (the fragility function calculator is not specific to any peril or geography). We hope that OpenRisk will soon extent to North American earthquake and to any arbitrary gridded hazard such as hurricane or straight-line wind, tornado, etc., anywhere in the world, as long as the user can supply the hazard and vulnerability data. OpenRisk is distributed by the Alliance for Global Open Risk Analysis (AGORA), a nonprofit virtual organization begun in early 2007 and currently comprising approximately 100 scholars and professionals in Japan, the US, and Europe who are themselves working on open-source risk software.

1.3 ORGANIZATION OF THIS REPORT

This chapter has introduced cat modeling and briefly summarized the motivation and objectives for creating new open-source risk software called OpenRisk. Chapter 2 presents a development approach. OpenRisk products are presented in detail in Chapter 3. Chapter 4 contains and summary and conclusions. Chapter 5 lists references cited.

1.4 ACKNOWLEDGMENTS

OpenRisk's initial design was funded by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for version 1.0 of this document is 1063. The research was also supported by the U.S. Geological Survey (USGS), Department of the Interior, under USGS award no 07HQAG0010, which supported the development of risk software and wind hazard information. Some of the vulnerability functions were developed under sponsorship of the US geological Survey's Prompt Assessment of Global Earthquake Risk (PAGER) project. The fragility function calculator was developed for the Applied Technology Council as part of the ATC-58 project, with the programming assistance of Mr. Xin Xu. The views and conclusions

contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either express or implied, of the U.S. Government or Applied Technology Council.

OpenRisk benefited greatly from the advice and efforts of Drs Edward Field and Nico Luco of the United States Geological Survey, and the programming assistance of Mssrs Nitin Gupta, Vipin Gupta, and Eric Martinez. Sid Hellman and his colleagues at Instrumental Software Technologies Incorporated (ISTI) performed some of the recent software development under a subaward from SPA Risk LLC, and provided valuable advice on the development process. The authors thank all these contributors.

2 DEVELOPMENT APPROACH

Development began with a design document that lays out a roadmap and a number of algorithms for common loss-estimation problems. It suggests a number of software classes required to carry them out, imagines system use cases (illustrations of how the software would actually be run), and provides a set of vulnerability functions to serve as seed data. More detail is provided later.

After creating the design document, we selected a number of initial applications for implementation. We began with applications that would be useful yet easy to implement, leveraging to the greatest extent possible already-existing software and data, especially OpenSHA code and seismic vulnerability functions. In each case we implemented risk algorithms that had been previously published or accepted for publication in peer-reviewed, scholarly literature.

For programming effort, we collaborated with USGS and SCEC researchers, and in some cases employed graduate students and consultants. We simultaneously set up AGORA. The OpenRisk software is created, tested, and uploaded to an AGORA web page (www.risk-agora.org), in the hope that other researchers will find it useful and seek to build on it.

3 OPENRISK PRODUCTS

As of this writing, OpenRisk comprises a design document, five software applications and three electronic databases, listed below and described in detail in the following pages. In many cases the detail is provided by reference to other publications. The products are:

- A design document
- A single-site benefit-cost-ratio (BCR) calculator
- A single-site loss-exceedance-frequency calculator
- A portfolio database import and editing tool
- An expected annualized loss calculator (EAL)
- A fragility-function calculator
- A wind hazard database
- A casualty vulnerability function database
- A building repair-cost vulnerability function database

3.1 DESIGN DOCUMENT

The initial development effort laid out the purpose and overall design of the software in a design document (Porter and Scawthorn 2007) that contains the following sections.

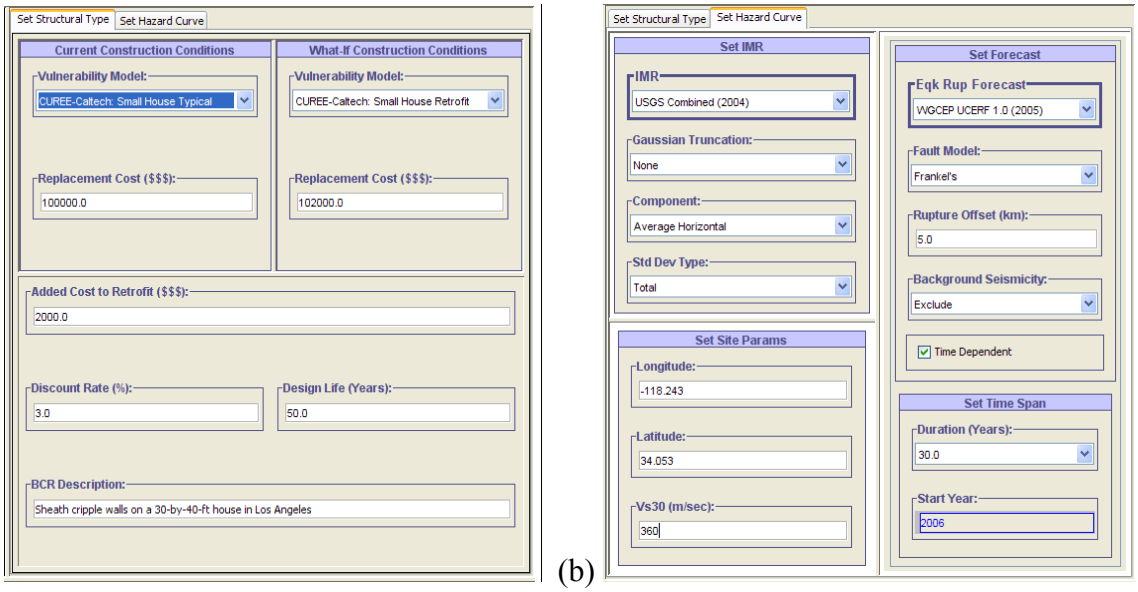
1. ***Introduction***, documenting the motivation and objectives of OpenRisk;
2. ***Development roadmap***, including a brief history of loss-estimation software, the OpenRisk mission statement, and a sequence of development goals copied to Table 8.
3. ***Single-site risk-calculation algorithms.*** These define the math required to calculate risk for a single asset. Math is provided on how to calculate single-site expected annualized loss (EAL); benefit-cost ratio (BCR) of a design alternative or retrofit; a loss-exceedance frequency (LE) curve; or a fragility function (a relationship that gives the probability of an undesirable outcome given input excitation).
4. ***Portfolio risk-calculation algorithms***. These currently include portfolio EAL and simultaneous operational failure of two or more facilities.
5. ***Software system use cases and UML***. This section details proposed applications including system use cases and UML-like summaries of classes to be created. A system use case describes on a step-by-step basis how a user would actually run the software—the screens the user would see, data he or she would enter, buttons to push, etc., and what the software would do. On the advice of Sid Hellman and others at Instrumental Software Technologies Inc., we found that a system use case helps to crystallize the details of a proposed application in the designer's mind and also helps to communicate the software requirements to the programmer.
6. ***Seed data.*** These currently include building vulnerability functions from ATC-13 (1985) and the CUREE-Caltech Woodframe Project, and fatality risk vulnerability functions developed from HAZUS (i.e., functions that give mean fraction of building occupants

killed as a function of 5%-damped site-soil-class-adjusted spectral acceleration response at either 0.3 sec or 1.0 sec period).

7. ***Design principles and development guidelines.*** These include guidance on the programming language, registration of code, software license, documentation, commenting, verification and validation, etc.
8. ***References cited***. Where practical, these include URLs for online documents.
9. ***Two appendices***: a glossary of terms and a brief summary of HAZUS.

3.2 BCR CALCULATOR

The BCR calculator is illustrated in Figure 2 and Figure 3: the first shows the input panes for the asset before and after the proposed change, and the second shows the output screen (which also echoes the input parameters). In the "Set Structural Type" pane, the user selects a vulnerability model to represent the as-is and what-if conditions, enters 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). The user also selects a discount rate and planning period for purposes of calculating the present value of the reduction in future losses. In the "Set Hazard Curve" pane, the user selects the intensity measure relationship (i.e., attenuation relationship) to be used, the site location (latitude and longitude; site soil class can be set automatically from a web service), and the earthquake rupture forecast (a model of the seismic sources and their seismicity). The user then presses "Compute," which performs the calculations shown in the design document and produces the output shown in Figure 3.



(a) (b)

Figure 2. Input panes for BCR calculator beta version

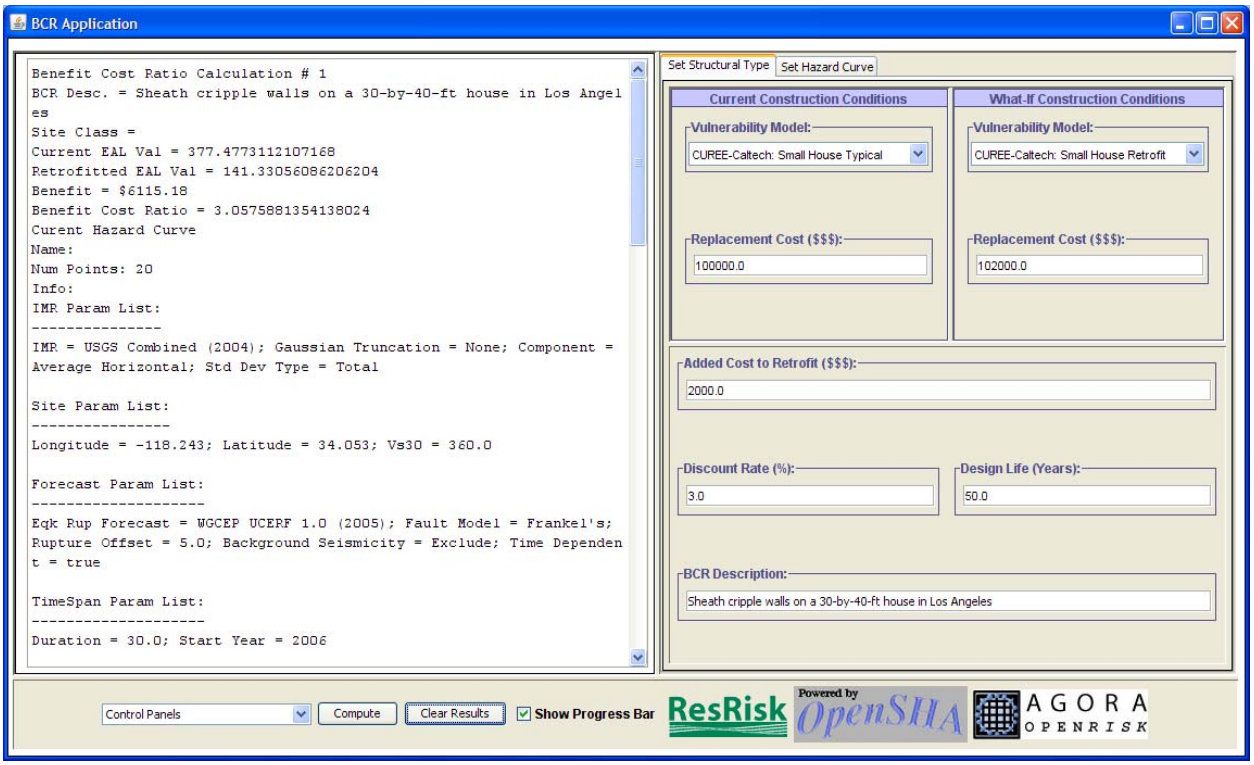
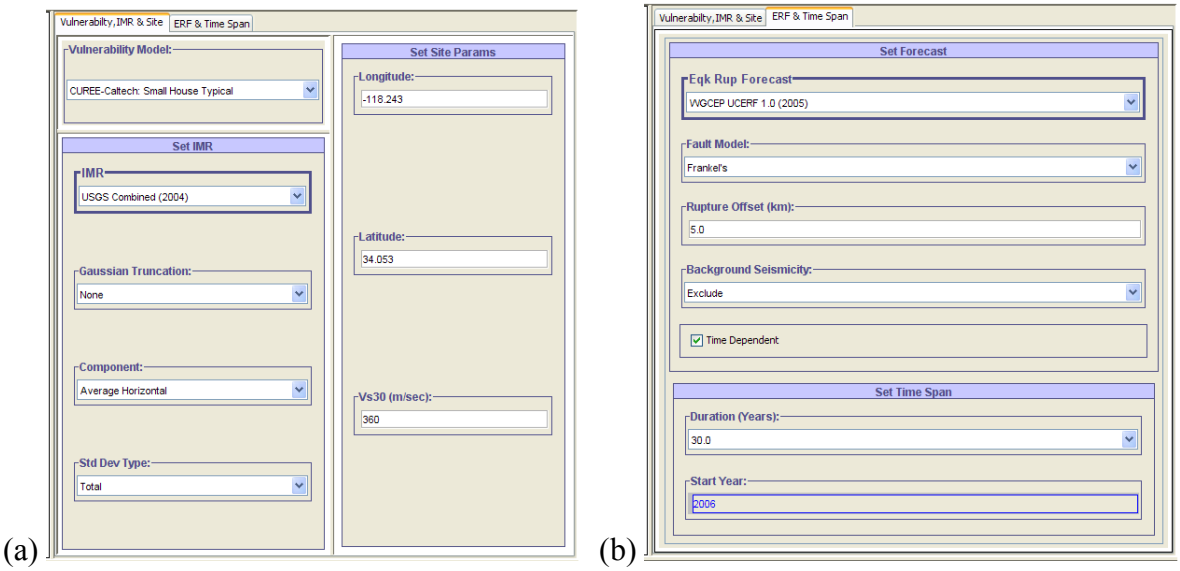


Figure 3. Output from BCR calculator beta version

3.3 LOSS-EXCEEDANCE-FREQUENCY CALCULATOR

The LEF calculator is illustrated in Figure 4 and Figure 5. The first shows the input panes for vulnerability model, the intensity-measure relationship, the site parameters, and the earthquake rupture forecast. The second shows the output screen: a plot of the mean annual exceedance frequency at various levels of damage factor. The input data are essentially the same as in the BCR calculator. The user can customize the labels, axes, and other aspects of the output plot.



(a) (b)

Figure 4. Input panes for loss-exceedance-frequency calculator

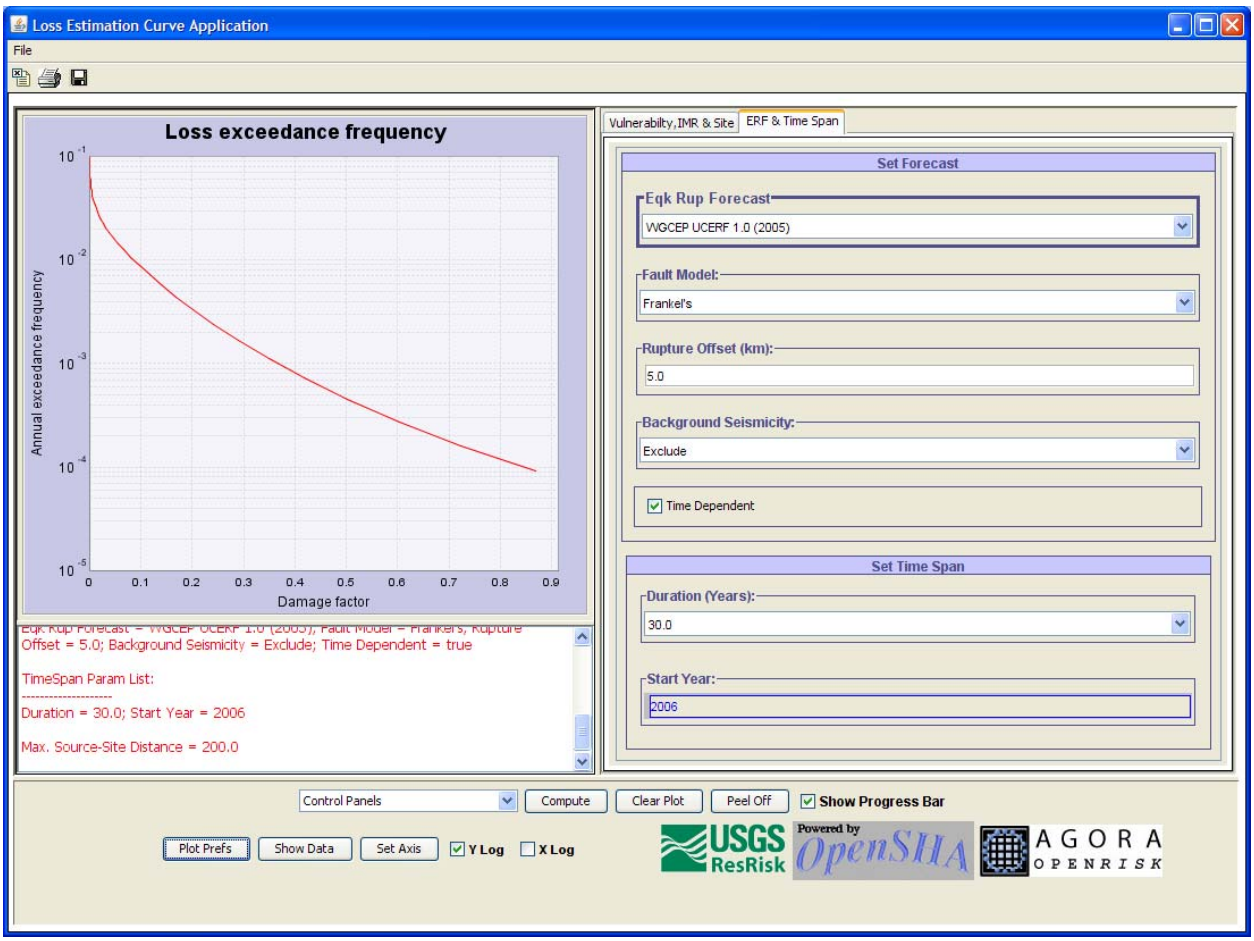


Figure 5. Output from loss-exceedance-frequency calculator beta version

3.4 FRAGILITY FUNCTION CALCULATOR

This tool is illustrated in Figure 6. The latest version can be found at www.risk-agora.org. It applies the procedures presented in Porter et al. (2007a). The calculator takes as input the excitation and observed damage state of each of several of specimens, along with an explanation of the category of component being tested, the specimens tested, the excitation to which they were subjected, the damage measure, and how the damage was observed. It performs the calculations referred to in the design document and produces as output the parameters of a fragility function, which it can submit to an Internet-accessible database for use by others. Such a database has been created with basic functionality, but has not been placed on a server.

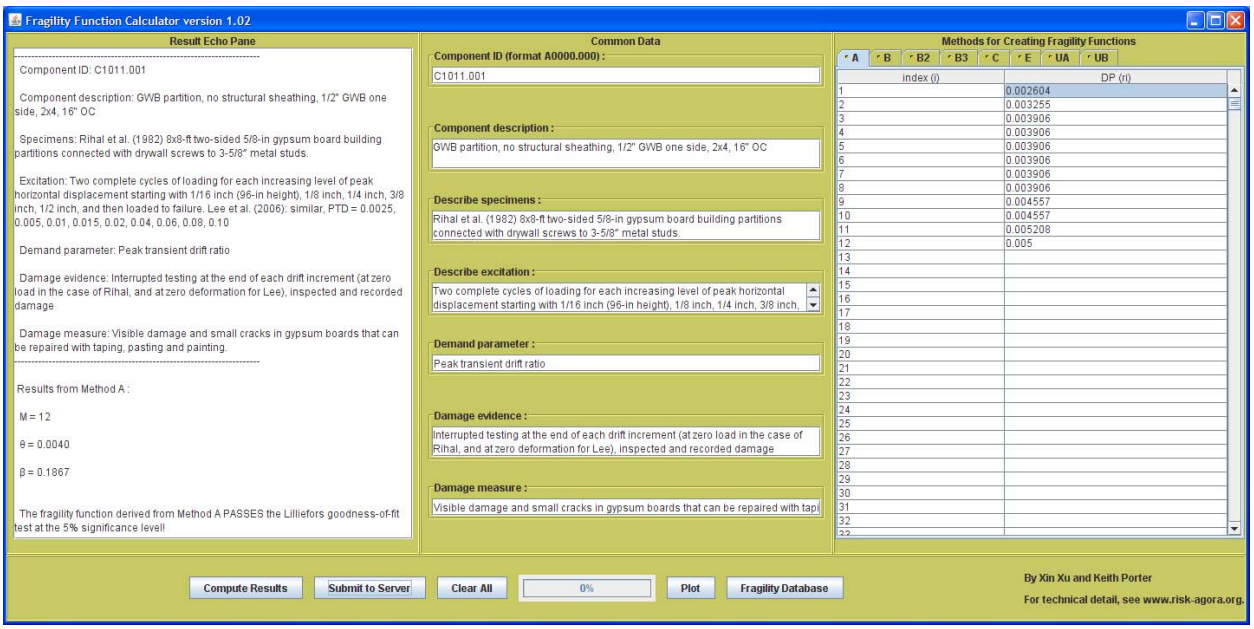


Figure 6. Fragility function calculator

3.5 PORTFOLIO DATABASE IMPORT TOOL

Its purpose is to create an electronic portfolio database that is readable by portfolio risk calculators developed later. Here, portfolio means a set of one or more assets exposed to earthquake, wind, or flood hazard. The database can be created by importing a human-readable text file, by on-screen data entry, or by editing the contents of an existing portfolio database. The tool starts with blank screen shown in Figure 7.

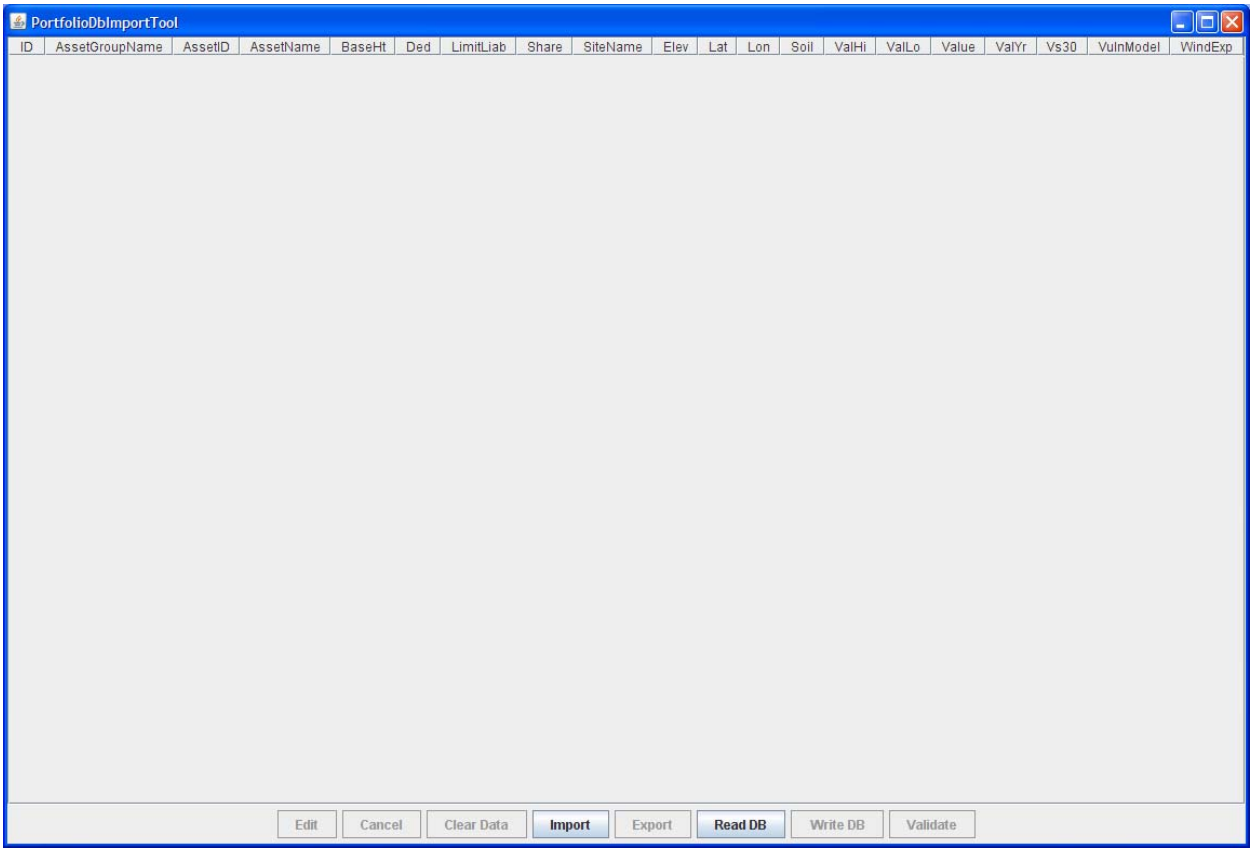


Figure 7. Initial screen of portfolio database import tool

Most often, a portfolio is imported from computer file, in this case a comma-and-quote delimited text file. The file has a header line and an arbitrary number of data lines. The header line contains the following text:

> "AssetID", "AssetName", "SiteID", "SiteName", "AssetGroupID", "AssetGroupName", "Lat", "Lon", "Value", "VulnModel", "Soil", "Vs30", "WindExp", "Elev", "ValHi", "ValLo", "ValYr", "BaseHt", "LimitLiab", "Ded", "Share"

Each data line contains information about one point asset, i.e., an asset located at a single point. A sequence of data items about each asset appears in the line, in the same order as the header line. The data items are detailed in Table 1.

Table 1. Portfolio contents

Field	Meaning	Type	Comment
AssetID	Asset identifier	Integer	A unique ID, e.g., 1
AssetName	Asset name	Text	A label for the asset, e.g., "House 1"
SiteID	Site identifier	Integer	A potentially non-unique identifier for the location, e.g., 1
SiteName	Site name	Text	A label for the site, e.g., "769 N Michigan Ave, Pasadena CA 91104"
AssetGroupID	Group identifier	Integer	Non-unique ID for group to which the asset belongs, e.g., 1
AssetGroupName	Group name	Text	Non-unique name for group to which asset belongs, e.g., "Houses"
Lat	Latitude	Double	Decimal degrees N, within ±90.00
Lon	Longitude	Double	Decimal degrees E, within ±180.00
Value	Value at risk	Double	Best est of value at risk, monetary or number or people, ≥ 0.00.
VulnModel	Vulnerability function	Text	Name of the vulnerability function to use for this asset, e.g., "CUREE small house as-is," restricted to a list of available models
Soil	NEHRP soil class	Text	NEHRP site soil classification, in {A, B, C, D, E, F}
Vs30	Shearwave vel, m/sec	Double	Mean shearwave velocity in top 30 m of soil, m/sec, >0.00
WindExp	Wind exp category	Text	Exposure category per IBC 2006 Sec 1609.4.3; in {B, C, or D}
Elev	Base elevation, m	Double	Elevation of threshhold of main entrance, m above MSL
ValHi	Upper bound of value at risk	Double	Value is uncertain or varies over time. Upper bound here is 96th percentile, ≥ Value
ValLo	Lower bound value at risk	Double	Lower bound is 4th percentile, ≤ Value
ValYr	Year in which Value is based	YYYY	Reserved for future use. Value can increase or decrease over time. ValYr can be used to update Val to present year.
BaseHt	Base elev above gr, m	Double	Elevation of lowest floor relative to threshhold of main entrance
LimitLiab	Limit of liability	Double	Same units as Value, applicable for coverage-specific limits.
Ded	Deductible, same units as Val	Double	0 ≤ Ded ≤ Value. Applicable for coverage-specific deductible. Also use as attachment point for fac insurance
Share	Pro-rata share	Double	Fraction of limit, 0 ≤ Share ≤ 1.00

3.6 PORTFOLIO EAL CALCULATOR

This tool is illustrated in Figure 8. It calculates the expected annualized loss (EAL) for each property in the portfolio and reports the sum. The method for calculating the EAL for an individual property is reported in Porter et al. (2006). It estimates the EAL for an arbitrary portfolio of assets, and an arbitrary peril. It takes as input a portfolio file, as provided by the Portfolio Database Import Tool. It calculates the EAL for each asset separately on the fly. There are 3 main panes to the interface: The I/O pane for output after calculations, the parameters pane, with two tabs, and the control pane, with control options and logos.

The I/O pane starts out blank. It will contain text in RTF format after a calculation is complete, and will become blank again when the "Clear Results" button is pressed. The parameters pane has two tabs: "Portfolio and Vulnerability" and "Set Hazard Curve." It starts in the former. There is a blank text box labeled "Enter Portfolio URL," and a button labeled "Open Portfolio." To select a portfolio, one clicks on the button and chooses a file from the file selector. The name and path of the file is displayed in the text box, which is uneditable. In the "Set Hazard Curve" tab, there are three sections: "Set IMR," "Set Forecast," and "Set Site Params." In "Set IMR," one selects all of the different values from drop down boxes. Some IMRs are not supported: Field (2000), Abrahamson (2000), ShakeMap (2003). In "Set Forecast," one also selects values from drop-down boxes or manually. Two forecasts are not supported: “Point Source ERF” and “Point 2 Mult Vertical SS Fault ERF.” In "Set Site Params," there are parameters based on the IMR that was selected. Longitude, latitude, and Vs30 are all uneditable because they are defined for each asset in the portfolio file. There are tooltips for each parameter in each section briefly viewable via mouseover.

In the control pane, there are 4 different selectable objects: the "Control Panels" drop-down box, the "Compute" button, the "Clear Results" button, and the "Show Progress Bar" check box. In the "Control Panels" drop down box, there are 3 options, "Max-site source distance," which is used to set the farthest distance away to worry about in the calculations, "Portfolios of Interest," which will allow one to select a pre-defined portfolio, and "Set Site Params from Web Services," which allows one to set the site parameters from a web service. The "Compute" button will start calculations after a portfolio is selected, and print information to the I/O Pane. (Note: When one presses compute, a new button pops up, "Cancel", and the "Compute" and "Clear Results" button

become disabled. When the calculation is done, or one selects the "Cancel" button, the Control Pane goes back to its original state.). The "Clear Results" button clears the I/O Pane of text. The program will remember how many calculations have already been done though, and will still display accordingly. If the "Show Progress Bar" check box is selected, a progress bar will show for each calculation when the "Compute" button is pressed. If "Show Progress Bar" is not selected, no progress bar will show up.

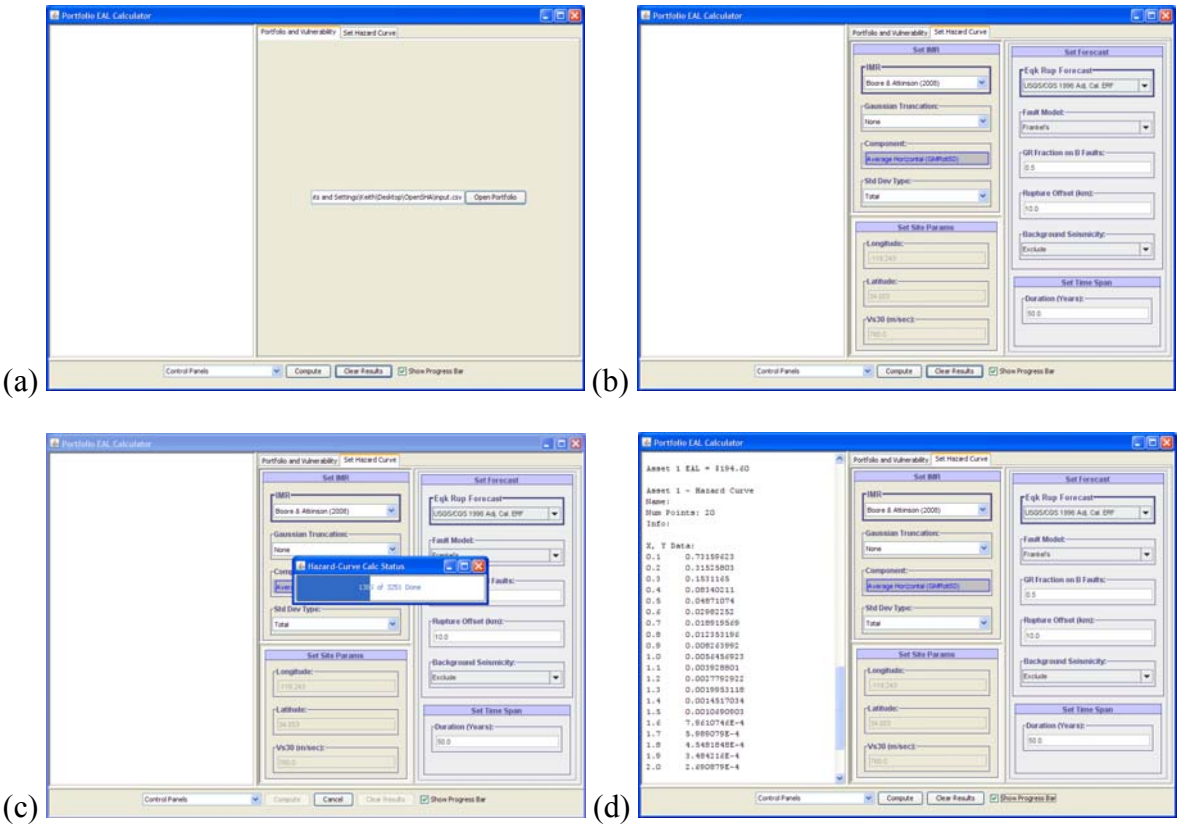


Figure 8. Portfolio expected annualized loss calculator

3.7 WIND HAZARD DATABASE

In the seismic domain, hazard information is commonly available throughout the United States and in many other parts of the world. Of particular note are the gridded seismic hazard maps offered by Frankel and Leyendecker (2002): on an approximately 0.05-degree grid throughout the United States, they offer mean annual exceedance frequencies for between 18 and 20 levels of seismic excitation, measured in terms of 5%-damped elastic spectral acceleration response at each of several reference periods, on soil at the boundary between NEHRP categories B and C. Using these maps, along with soil maps and vulnerability functions derived elsewhere,

one can evaluate the benefit of a mitigation measure anywhere in the United States (see e.g., Porter et al. 2006).

There is no public analog in the wind domain for Frankel and Leyendecker (2002) gridded seismic hazard maps. Commercial vendors RMS, EQECAT, AIR, and others offer proprietary hazard and risk models, especially for the portions of the US exposed to hurricane winds, but the cost of their use is typically prohibitive for research purposes and one generally does not have access to the wind hazard information. The FEMA-funded emergency planning software HAZUS-MH incorporates a similar hurricane hazard model, but it requires the purchase of commercial GIS software. HAZUS-MH is not open source software, meaning its source code is not accessible to outsiders, and consequently access to its underlying databases of hazard and vulnerability is limited.

As part of an effort to develop open-source risk modeling software and data, we have therefore developed a public gridded wind hazard model analogous to the Frankel and Leyendecker (2002) seismic hazard model. That is, we developed $G(s)$ for the continental US and Alaska, measuring s in terms of 3-second peak gust velocity at 10-meter elevation on ASCE 7 Roughness Category C, and $G(s)$ as the mean exceedance frequency of s for several reference values of s.

It was impractical under the present study to generate wind hazard analytically from first principles, or empirically from long-term anemometer records. Instead, wind hazard was extrapolated from the ASCE 7-05 Basic Windspeed Map (e.g., Figure 9) as follows. The windspeed map was digitized for the entire United States by a commercial GIS firm, Geovect, and evaluated at each point on a 0.05-degree grid. The grid includes 561,644 locations on land in the United States between 67.00W and 169.95W, and between 25.3N and 71.2N (44.40N, 67W). Its values, denoted here by V_{50}, were taken as the 10-m elevation, exposure-C, 3-sec peak gust velocity (in miles per hour) with mean recurrence interval of 50 yr. We spot-checked the digitized values of basic windspeed in the locations listed in Anchorage, San Francisco, Galveston, Omaha, Chicago, Miami, and Portland, ME, as summarized in Table 2.

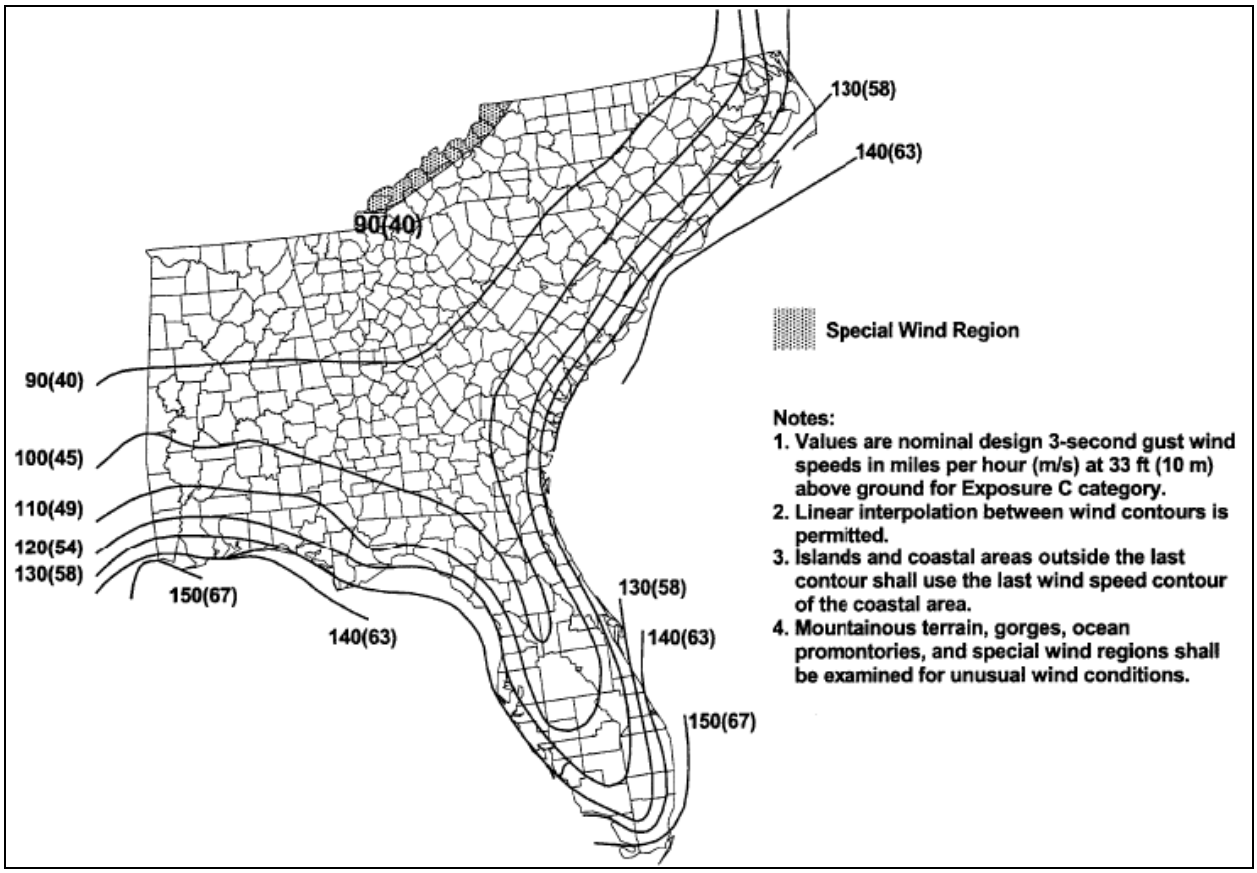


Figure 9. ASCE 7-05 basic windspeed map for the southeastern United States (ASCE 2005)

Table 2. Spot-check tabulated windspeeds vs ASCE 7-05

Location	Lat N	Lon W	V_{50}	Agrees?
Near Anchorage AK	60.00	148.00	130.00	OK
Near San Francisco CA	38.00	122.00	87.39	OK
Near Houston TX	30.00	95.00	106.04	OK
Near Omaha NE	41.00	96.00	90.00	OK
Near Chicago IL	42.00	87.00	90.00	OK
Near Miami FL	25.75	80.25	147.59	OK
Portland ME	43.75	70.25	98.15	OK

The extrapolation to other recurrence periods was performed as follows. In a paper that underlies ASCE 7-05 map, Peterka et al. (1998) offer conversion factors for 7 intervals: 500, 100, 50, 25, 10, 5, and 1 yr (the conversion factor for 50 yr being 1.0). Table 3 recaps these conversion factors, denoted here by $f(T)$, where T denotes mean recurrence interval in years.

$$V_T = V_{50} f\left(T\right) \tag{1}$$

Table 3. Peterka et al. (1998) conversion factors *f*(*T*) for other recurrence intervals

T (yr)	$85 \le V_{50} \le 100$	$V_{50} > 100$ (hurricane)	Alaska
500	1.23	1.36	1.18
100	1.07	1.105	1.06
50	1.00	1.00	1.00
25	0.93	0.90 (84 mph minimum)	0.94
10	0.84	0.77 (76 mph minimum)	0.87
5	0.78	0.68 (70 mph minimum)	0.81
1	0.61	0.48 (55mph minimum)	0.67

ASCE 7-05 (ASCE 2005) Table 6-3 offers exposure coefficients to estimate V_{50} at sites with Roughness Categories other than C, and at heights other than 10 m. Interpolating in the ASCE table for 10-m height, results in the exposure coefficients of $K_h = 0.72$ for roughness category B and $K_h = 1.18$ for roughness category D. Thus, for any particular location whose V_{50} is known from the ASCE 7-05 basic windspeed map, one can estimate the peak gust velocity for a given roughness category *R* and recurrence interval *T*, as follows:

$$s(R,T) = V_{50} \cdot K_h(R) \cdot f(T) \tag{2}$$

Values of *s*(*R*,*T*) were calculated for each gridpoint in the database, each return interval *T* in {1, 5, 10, 50, 100, 500}, and for each roughness category *R* in {B, C, D}. Note that Equation (2) does not give the hazard information in the desired format, which was the mean recurrence interval associated with various reference values of *s*; rather it gives an inverse of *G*(*s*): the values of *s* associated with various reference values of 1/*G*. However, the final conversion to *G*(*s*) is straightforward. The factors in Table 3 are approximately linearly related to ln(1/*T*), meaning that ln(*G*) is approximately linear with *s*. Hence, for a given gridpoint, roughness category *R*, and reference value s_i

$$G(s_i) = \exp\left(-\ln T_0 + \left(\frac{s_i - s_0}{s_1 - s_0}\right)(\ln T_0 - \ln T_1)\right) \tag{3}$$

where s_0 and s_1 denote the values of *s* nearest to s_i, and T_0 and T_1 denote the values of *T* associated with s_0 and s_1, respectively. Recall that $G = 1/T$ and $\ln(1/T) = -\ln T$, so Equation (3) is simple linear interpolation or extrapolation in the space of ln(*G*) versus *s*.

We calculated hazard at the reference windspeeds (70, 80, 90, … 170 mph), three roughness categories (B, C, and D), and calculated *G*(s) on two rectangular grids: one for the conterminous United States and one for Alaska. The corners of the grid for the continental US are (49.00N, 124.80W) and (24.50N, 66.95W). The grid corners for Alaska are (169.95W, 71N) and (129.50W, 53.00N). Results are tabulated in six text files, available for public download at www.risk-agora.org, listed in Table 4.

Table 4. Wind hazard data files

Region	Roughness	Revision	Data file
Conterminous US	B	0	SPA-2008-US-WS-B-R0.txt
Conterminous US	C	0	SPA-2008-US-WS-C-R0.txt
Conterminous US	D	0	SPA-2008-US-WS-D-R0.txt
Alaska	B	0	SPA-2008-AK-WS-B-R0.txt
Alaska	C	0	SPA-2008-AK-WS-C-R0.txt
Alaska	D	0	SPA-2008-AK-WS-D-R0.txt

Each data file begins with an explanatory header, followed by one data line per gridpoint. Each data line contains the information summarized in Table 4. Records are arranged to begin in the upper left-hand corner of the grid, proceed from left to right (i.e., west to east) at given latitude, then top to bottom (i.e., north to south). The organization of the data files mimics that of Frankel and Leyendecker's (2002) gridded seismic hazard data.

Table 5. Contents of casualty vulnerability function database

Field name	Type	Contents
Lat	Double	Latitude (degrees N) of gridpoint, in {24.50, 24.55, … 49.00} for the conterminous US, or {53.00, 53.05, … 71.00} for Alaska
Lon	Double	Longitude (degrees E) of gridpoint, in {-124.80, -124.75, … -66.95} for conterminous US, or {-169.96, 169.90, … -129.50} for Alaska
G75	Double	Mean frequency (yr^{-1}) with which 3-second gust velocity at 10-m elevation exceeds 75 mph
G100	Double	Ditto, 100 mph
G125	Double	Ditto, 125 mph
G150	Double	Ditto, 150 mph
G175	Double	Ditto, 175 mph
G200	Double	Ditto, 200 mph

3.8 VULNERABILITY FUNCTION DATABASES

Two vulnerability function databases have been created that provide relationships between loss and spectral acceleration for US construction. The first provides fraction of indoor occupants at each of 4 injury severity levels as a function of 0.3-second or 1-second, 5%-damped spectral acceleration response, conditioned on structure type (using the standard FEMA structure type classifications of, e.g., NIBS and FEMA 2003), magnitude, distance, NEHRP site soil classification, and seismic domain (plate boundary or continental interior). The methodology and calculations are detailed in Porter (2009a), and for brevity are not repeated here. The vulnerability function table contains the data listed in Table 6.

Table 6. Contents of casualty vulnerability function database

Field name	Type	Contents
MBT	Text	HAZUS-MH model building type (e.g., W1, meaning small woodframe) plus a character to indicate code era (e.g., h, meaning high code)
Domain	Text	Western US ("WUS"), or plate boundary or central and eastern US ("CEUS"), or continental interior
M	Integer	Magnitude, in {5, 6, 7, 8}
R	Integer	Fault distance, in {10, 20, 40, 80} km
Siteclass	Text	NEHRP site soil classification, in {A, B, C, D, E}
S_d	Double	Spectral displacement at performance point, inches
S_a	Double	Spectral acceleration at performance point, g
SA03	Double	Site-soil amplified 5%-damped elastic spectral acceleration response at 0.3-sec period
SA10	Double	Site-soil amplified 5%-damped elastic spectral acceleration response at 1.0-sec period
IM	Text	Better intensity measure to use to at this intensity: "SA03" means that the performance point lies on the constant-acceleration portion of the idealized response spectrum, and hence it is better to use SA03. "SA10" means that the performance point lies on the constant-velocity portion of the response spectrum, and hence it is better to use SA10.
L1	Double	Fraction of indoor occupants with injuries requiring basic medical aid that could be administered by paraprofessionals. These types of injuries would require bandages or observation. Some examples are: a sprain, a severe cut requiring stitches, a minor burn (first degree or second degree on a small part of the body), or a bump on the head without loss of consciousness. Injuries of lesser severity that could be self treated are not included.
L2	Double	Fraction of indoor occupants with injuries requiring a greater degree of medical care and use of medical technology such as x-rays or surgery, but not expected to progress to a life threatening status. Some examples are third degree burns or second degree burns over large parts of the body, a bump on the head that causes loss of consciousness, fractured bone, dehydration or exposure.
L3	Double	Fraction of indoor occupants with injuries that pose an immediate life threatening condition if not treated adequately and expeditiously. Some examples are: uncontrolled bleeding, punctured organ, other internal injuries, spinal column injuries, or crush syndrome.
L4	Double	Fraction of indoor occupants instantaneously killed or mortally injured

The second database gives mean damage factor (repair cost as a fraction of replacement cost new) as a function of the same measures of spectral acceleration response, given structure type, occupancy type (using the standard FEMA occupancy classifications of NIBS and FEMA 2003), magnitude, distance, NEHRP site soil classification, and seismic domain. Its layout is

summarized in Table 7. The methodology employed and the calculations are detailed in Porter (2009b). Both databases are contained in comma-separated-value text files. The casualty table is named "SPA-2008-Casualty-R1.txt," where R1 refers to revision number 1. The mean-damage-factor table is named "SPA-2008-MDF-R1.txt." Both are offered at www.risk-agora.org.

Table 7. Contents of building repair cost vulnerability function.

Field name	Type	Contents
MBT	Text	HAZUS-MH model building type, e.g., W1, meaning small woodframe, plus a character to indicate code era, e.g., h, meaning high code
Occupancy	Text	HAZUS-MH occupancy type, e.g., RES1, meaning single family dwelling
Domain	Text	Western US ("WUS"), or plate boundary or central and eastern US ("CEUS"), or continental interior
M	Integer	Magnitude, in {5, 6, 7, 8}
R	Integer	Fault distance, in {10, 20, 40, 80} km
Siteclass	Text	NEHRP site soil classification, in {A, B, C, D, E}
S_d	Double	Spectral displacement at performance point, inches
S_a	Double	Spectral acceleration at performance point, g
SA03	Double	Site-soil amplified 5%-damped elastic spectral acceleration response at 0.3-sec period
SA10	Double	Site-soil amplified 5%-damped elastic spectral acceleration response at 1.0-sec period
IM	Text	Better intensity measure to use to at this intensity: "SA03" means that the performance point lies on the constant-acceleration portion of the idealized response spectrum, and hence it is better to use SA03. "SA10" means that the performance point lies on the constant-velocity portion of the response spectrum, and hence it is better to use SA10.
MDF	Double	Mean damage factor, i.e., repair cost as a fraction of replacement cost new

4 SUMMARY AND CONCLUSIONS

A design document, five initial applications, and three databases have been developed and made freely available. The applications and databases address hazard, vulnerability, and risk from earthquake and wind excitation. OpenRisk executable software, documents, and databases are available to risk researchers, insurers and reinsurers, and others who register with AGORA. Source code is offered to anyone who agrees to contribute any new or modified source code under the same terms. The development history and intended timeline for future work is presented in Table 8.

Table 8. Timeline for development of OpenRisk

Event	Time	Comment
Initial design	2007	A design document
AGORA Workshop	2007	Workshop Feb 27-28, Caltech
OpenRisk suite 1	2007	Single-site earthquake BCR and LE applications, a fragility function calculator
AGORA meeting	Sept 2007	Italy
OpenRisk suite 2	2008	Enhanced fragility function calculator, portfolio import tool, portfolio earthquake EAL calculator, enhanced vulnerability catalog, gridded US wind hazard
AGORA meeting	2008	Beijing; Directors meeting
OpenRisk suite 3	2009-2010	Earthquake portfolio loss-exceedance calculator, extend single-site seismic risk calculations to continental US
AGORA meeting	2009	Cambridge University
OpenRisk suite 4	2010	Full actuarial quality portfolio seismic loss analysis capability

We expect that insurers will find various ways to benefit from OpenRisk: (1) to gain insight into the components of the commercial or vendor loss models (those of RMS, AIR, EQECAT, etc.), (2) to anticipate the impacts of new science on risk, and (3) potentially to supplement the vendor models with new perils, geographic regions, or asset classes of interest. To explain:

Insight. Commercial cat models often provide limited access to and control over their analytical components. As a result, it can be hard to see the intermediate steps: only what goes in and what comes out, and not how the input was manipulated in between. With OpenRisk, the user can see and understand every step, because the algorithms and data are documented,

because the source code is open, and because the applications permit a great deal of control over how the calculations are performed.

New science. The vendors of commercial cat models work on a development cycle of a year or two, so it can be many months between the publication of new science and its uptake in vendor models. With OpenRisk, researchers can implement the new science more quickly. Indeed as shown by Field et al. (2005), OpenSHA has been an integral part of creating the new science, and we have used the fragility function calculator in the preparation of new fragility functions for ATC-58 publications, so OpenRisk users can explore implications of new science—estimate its effects on risk—long before the vendors adopt it.

New perils and geographies. We hope shortly to develop a single-site loss-exceedance calculator that accepts an arbitrary gridded hazard data file and user-defined vulnerability functions, which will allow users to model new perils or perils in new geographic locations, as long as they can supply the required hazard and vulnerability data. And of course insurers or researchers could create such an application themselves by extending existing OpenSHA or OpenRisk source code.

The OpenRisk design document and several software applications and databases extend the capabilities of OpenSHA to catastrophe loss estimation. They offer the promise to insurers and researchers of greater insight and control over the calculation of risk. We hope that by making applications and source code available to others, they will find it easier to build on existing source code than to reinvent it themselves, and will contribute their efforts back to the community to the benefit of all catastrophe researchers.

5 REFERENCES CITED

(ASCE) American Society of Civil Engineers, 2006. *Minimum Design Loads for Buildings and Other Structures, SEI/ASCE 7-05*, Reston, VA, 388 pp.

(ATC) Applied Technology Council, 1985. *ATC-13, Earthquake Damage Evaluation Data for California,* Redwood City, CA, 492 pp.

Field, E.H., N. Gupta, V. Gupta, M. Blanpied, P. Maechling, and T.H Jordan, 2005. Hazard calculations for the WGCEP-2002 forecast using OpenSHA and distributed object technologies. *Seismological Research Letters* 76, 161-167

Graves, R.W. and P.G. Somerville, 2006. Broadband ground motion simulations for scenario ruptures of the Puente Hills Fault. *Proc. 8th National Conference on Earthquake Engineering*, ~18-21 Apr 2006, San Francisco CA

Kishi, N., 2007. Insurance Industry Perspectives. *1st International Workshop on Open-Source Risk Software*, California Institute of Technology; February 27th and 28th, 2007. www.risk-agora.org/downloads.html

(MMC) Multihazard Mitigation Council, 2005. *Natural Hazard Mitigation Saves: An Independent Study to Assess the Future Savings from Mitigation Activities*, National Institute of Building Sciences, Washington, DC. http://www.nibs.org/MMC/MitigationSavingsReport/natural_hazard_mitigation_saves.htm

(NIBS and FEMA) National Institute of Building Sciences and Federal Emergency Management Agency, 2003. *HAZUS-MH MR1 Advanced Engineering Building Module Technical and User's Manual,* Federal Emergency Management Agency, Washington, DC, 119 pp., http://www.fema.gov/hazus/dl_aebm.shtm

Peterka, J.A., and S. Shahid, 1998, "Design Gust Wind Speeds in the United States," *Journal of Structural Engineering*, American Society of Civil Engineering, Reston, VA, February 1998, 207-214

Porter, K., R. Graves, E. Reis, and P. Somerville, 2007b. *Index Woodframe Houses and their Response to Puente Hills Scenario Earthquakes*. USC-SCEC/CEA Technical Report #9 for Milestone 3d, Southern California Earthquake Consortium, Los Angeles CA., http://www.sparisk.com/publications.html

Porter, K.A., 2009a (expected). Cracking an open safe: more HAZUS vulnerability functions in terms of structure-independent spectral acceleration. Accepted for publication in *Earthquake Spectra* November 2008.

Porter, K.A., 2009b (expected). Cracking an open safe: HAZUS vulnerability functions in terms of structure-independent spectral acceleration. Accepted for publication in *Earthquake Spectra* November 2008.

Porter, K.A., and C.R. Scawthorn, 2007. OpenRisk: Open-Source Risk Estimation Software. SPA Risk, Pasadena, CA, 107 pp., http://www.risk-agora.org/downloads.html

Porter, K.A., C.R. Scawthorn, and J.L. Beck, 2006. Cost-effectiveness of stronger woodframe buildings. *Earthquake Spectra* 22 (1), February 2006, 239-266

Porter, K.A., R.P. Kennedy, and R.E Bachman, 2007a. Creating fragility functions for performance-based earthquake engineering. *Earthquake Spectra*. 23 (2), May 2007, pp. 471-48

Robinson, D., T. Dhu, and J. Schneider, 2006. Practical probabilistic seismic risk analysis: a demonstration of capability. Seismological Research Letters, 77 (4): 452-458

Scawthorn, C.R., 2006. History of seismic risk assessment. *Workshop on Strategic Directions for (Seismic) Risk Modeling and Decision Support*, Boulder CO, July 2006. Mid-America Earthquake Center.

Wiggins, J.H., J. Slosson, and J. Krohn, 1976. *Natural Hazards: Earthquake, Landslide, Expansive Soil Loss Models*. J.H. Wiggins Company, Redondo Beach, CA