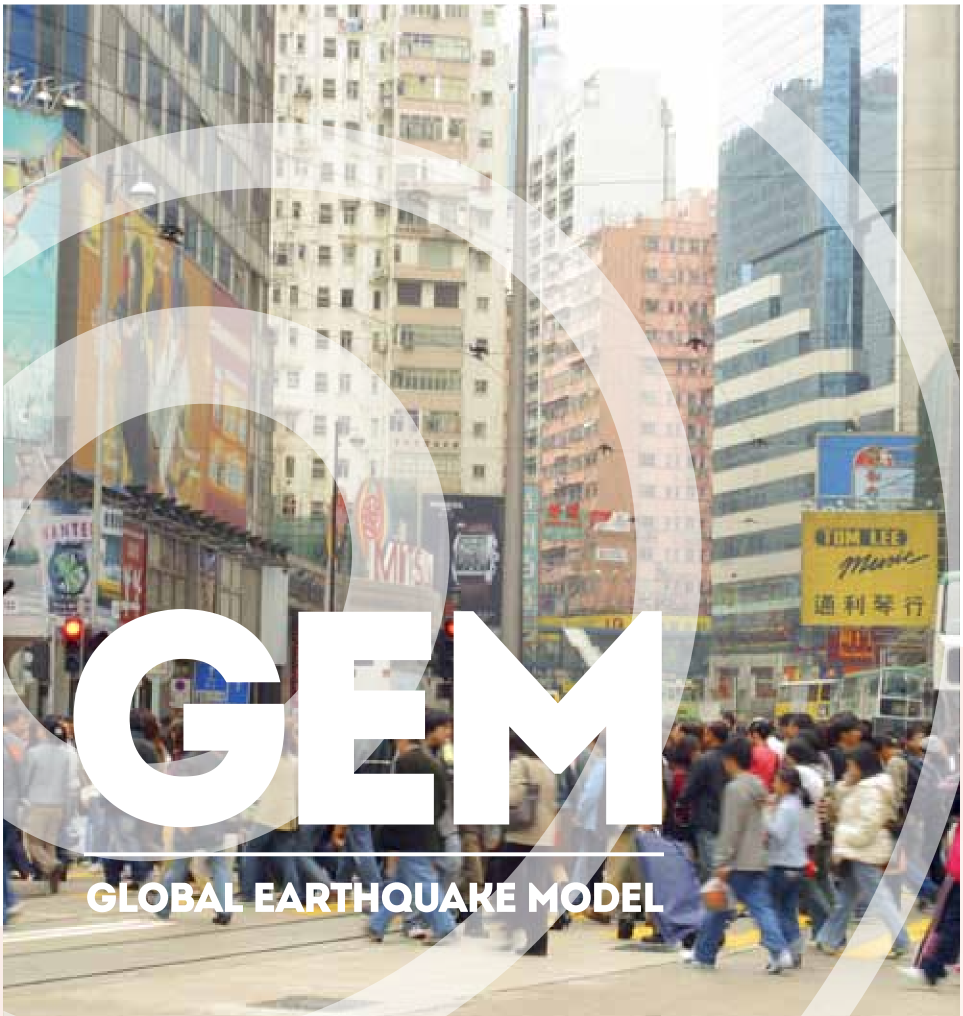
VULNERABILITY AND LOSS MODELLING

GEM TECHNICAL REPORT 2014-11 V1.0.0

Guidelines for Empirical Vulnerability Assessment

Rossetto, T., I. Ioannou,
D.N. Grant and T. Maqsood



Guidelines for Empirical Vulnerability Assessment

Technical Report 2014-11

Version: 1.0.0

Date: December 2014

Author(s)*: Rossetto, T., I. Ioannou, D.N. Grant and T. Maqsood

(*) Authors' affiliations:

Tiziana Rossetto, University College London

Ioanna Ioannou, University College London

Damian N Grant, Arup

Tariq Maqsood, Geoscience Australia

Citation advice

Rossetto, T., I. Ioannou, D.N. Grant and T. Maqsood (2014), Guidelines for empirical vulnerability assessment, GEM Technical Report 2014-08 V1.0.0, 140 pp., GEM Foundation, Pavia, Italy, doi: 10.13117/ GEM.VULN-MOD.TR2014.11.

http://www.globalquakemodel.org/

ACKNOWLEDGEMENTS

These guidelines are mainly based on current state-of-the-art techniques assembled from the literature. However, no current empirical vulnerability and fragility study has to date systematically compared statistical modelling techniques for different empirical datasets nor explored many of the issues raised in the document regarding the treatment of uncertainty. Hence, in order to make the recommendations that are presented in this document, a significant effort was dedicated to explore techniques, test and compare proposed methods, different parameter choices and assumptions for deriving the vulnerability and fragility curves for various post-earthquake survey databases. In this context, the Authors would like to greatly thank Professor Richard Chandler of the Department of Statistical Science at University College London, Professor Anne Kiremidjian and David Lallemant of the Department of Civil and Environmental Engineering at Stanford University, and Dr Hae Young Noh of Civil and Environmental Engineering at Carnegie Mellon University. Professor Jason Ingham of the Department of Civil and Environmental Engineering at University of Auckland, Professor Michael Griffith and Ms Lisa Moon of the School of Civil, Environmental and Mining Engineering at the University of Adelaide are thanked for providing the empirical data for the Christchurch 2010-2011 earthquake example. Geoscience Australia is thanked for providing the empirical data and example application in Appendix D of the “Guidelines for Empirical Vulnerability Assessment” to the Newcastle 1989 and Kalgoorlie 2010 earthquakes. Professor Anne Kiremidjian (Stanford University), Professor Keith Porter (University of Colorado at Boulder) and Dr David Wald (USGS) are also thanked for taking the time to review the Guidelines for Empirical Vulnerability Assessment. This research has been funded by the GEM Foundation and University College London.

ABSTRACT

These Guidelines provide state-of-the-art guidance on the construction of vulnerability relationships from post-earthquake survey data. The Guidelines build on and extend procedures for empirical fragility and vulnerability curve construction found in the literature, and present a flexible framework for the construction of these relationships that allows for a number of curve-fitting methods and ground motion intensity measure types (IMTs) to be adopted. The philosophy behind the design of the framework is that the characteristics of the data should determine the most appropriate statistical model and intensity measure type used to represent them. Hence, several combinations of these must be trialled in the determination of an optimum fragility or vulnerability curve, where the optimum curve is defined by the statistical model that provides the best fit to the data as determined by a number of goodness-of-fit tests. The Guidelines are essentially a roadmap for the process, providing recommendations and help in deciding which statistical model to attempt, and promote trialling of models and IMTs.

The Guidelines are targeted at analysts with Master's level training in a numerate subject that includes some level of statistics. The Authors recognise that the statistical analysis understanding of analysts varies. To accommodate for these differences, two levels of statistical approaches for constructing empirical fragility functions that include procedures of increasing complexity, are proposed.

All stages of the fragility and vulnerability curve construction are reviewed and presented in the Guidelines with practical advice given for the preparation of empirical data for use in the construction of these curves, for the identification of sources of uncertainty in the data and in the chosen intensity measures, and where possible, for uncertainty quantification and modelling.

To facilitate adoption of the Guidelines, the code and commands required for the implementation of the described statistical models are provided for the open source software R (2008). Appendices B to G also provide example applications of the guidelines, where each step of the guideline is illustrated for empirical datasets deriving from the 1980 Irpinia, Italy, earthquake, the 1978 Thessaloniki, Greece, Earthquake, the 1989 Newcastle and 2010 Kalgoorlie, Australia, earthquakes and for two earthquakes that affected the town of Christchurch New Zealand in 2010 and 2011. The fragility and vulnerability curves developed from these applications are all presented using a reporting template (presented in Appendix A) designed to facilitate the evaluation and inclusion of empirical fragility curves derived using these Guidelines into the Global Earthquake Model (GEM).

Keywords

Empirical vulnerability; empirical fragility; post-earthquake survey; uncertainty; model fitting.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Casualty Ratio:	The number of people ina geopgraphical unit who suffered a certain casualty level of above over the total exposed population in this unit.
Confidence Intervals:	Intervals expressing the confidence in the systematic component of a statistical model.
Damage Factor:	The cost of repairing one building over the cost of replacing this building. For a group of buildings, the mean cost of repair over the mean replacement cost.
Data Point:	A point expressing levels of the response variable and one or more explanatory variables.
Damage Ratio:	The number of buildings located in a geographical unit which sustained damage greater or equal to a given damage state over the total number of buildings in this unit.
Empirical Fragility Assessment:	The construction of a fragility function from the statistical analysis of field observations.
Fragility Function:	A relationship expressing damage as a function of an intensity measure type.
Geographical unit:	The smallest survey geographical area from which a sample size is obtained.
Intensity Measure:	A measure of the ground motion intensity at the site where the examined assets are located.
Intensity Measure Level:	A value of a given type of the intensity measure, e.g., peak ground acceleration=0.5g.
Intensity Measure Type:	Type of the ground motion intensity measures, e.g., peak ground acceleration, spectral displacement.
Loss:	Consequences of seismic damage.
Loss Measure:	Measures of loss expressed in terms of cost of direct damage, casualty and downtime.
Prediction Interval:	Intervals used to predict the response variable for future values of the explanatory variables.
Explanatory variable:	The ‘input’ to a statistical model, which is used to predict the level of the response variable. Here, explanatory variable is the intensity measure type.
Statistical Modelling Techniques:	Techniques which fit a parametric or non-parametric statistical model to the available data points.
Statistical Model:	A parametric or non-parametric model which allows for the prediction of damage or loss given an intensity measure type.
Response Variable:	The ‘output’ of the regression model.
Structural Unit:	A measure of the surveyed elements, i.e., buildings, dwellings or rooms.

Empirical Vulnerability Assessment:	The construction of vulnerability functions by fitting statistical models to field observations.
Vulnerability Function :	A relationship expressing seismic loss (i.e., economic loss of direct damage to buildings, casualties or downtime) as a function of an intensity measure type.

GLOSSARY OF ABBREVIATIONS

AIC	Akaike Information Criterion
CI	Confidence Interval
FC	Fragility Curve
GAM	Generalized Additive Model
GEM	Global Earthqake Model
GEMECD	GEM Earthquake Consequence Database
GEMGVM	GEM Global Vulnerability Estimation Methods
GKS	Gaussian Kernel Smoother
GLM	Generalized Linear Model
IML	Intensity Measure Level
IMT	Intensity Measure Type
VF	Vulnerability Function

LIST OF SYMBOLS

Symbol	Description:
b	Parameter of the Gamma distribution (GLMs).
$b(.)$	A basis function.
c	Parameter of the Gamma distribution (GLMs).
CD	Cook's Distance.
d	Parameter of the Inverse Gaussian distribution (GLMs).
DS	Damage considered as a response variable.
ds_i	A specific damage state.
$f(.)$	Probability density function of a variable.
$g(.)$	The link function which related the mean response to the linear predictor for GLMs and GAMs.
h_j	Leverage, i.e. the diagonal value *j* of the hat or projection matrix.
IM	Ground motion intensity measure (a typical explanatory variable).
$\overline{IM}$	'true' intensity measure.
J	Roughness penalty component.
$L(.)$	Likelihood function.
log	Natural logarithm (according to 'R')
M	Total number of data points.
N_X	The total number of explanatory variables.
n_{Class}	Total number of building classes.
n_i	Number of buildings in bin *i*.
n_T	Total number of buildings in the database.
q	The total number of smoothing parameters.
R	Seismic source-to-site distance.
S	Random variable for soil conditions
$s(.)$	Smoothing function.
$var(.)$	The variance of a variable.
wj	Is the weight for the continuous conditional distribution values of the response variable for bin j.
$\mathbf{X}$	A vector of explanatory variables.
x_i	A given level of an explanatory variable.
Y	The response variable.

y_i	A given level of a response variable.
y_N	A set of new response data.
$\hat{Y}_j$	The prediction from the GLM for observation *j.*
$\hat{Y}_{j(i)}$	The prediction for observation *j* from a refitted model which omits observation *i*.
r_{Pj}	Pearson residuals.
r_{Psj}	Standardized Pearson residuals.
B	Unknown parameter for GAM.
$\Gamma(.)$	Gamma function.
$E[.]$	The expected value of a variable.
ε	A normally distributed random variable with mean equal to 0 and standard deviation equal to 1.
$\boldsymbol{\theta}$	A vector of the parameters of a statistical model.
$\boldsymbol{\theta}^{\mathbf{opt}}$	The optimum estimate of a model's parameters.
λ	A non-negative parameter which adjust the degree of smoothing.
η	The linear predictor of a GLM or GAM model.
$\boldsymbol{\xi}$	Vector of parameters which describe the relationship of the true and the observed *IM*.
μ	The mean of a variable.
$\boldsymbol{\pi}$	Vector of the parameters of the prior distribution of the observed *IM.*
σ	Standard deviation.
σ_T	Total standard deviation of the GMPE.
Φ	Cumulative standard normal distribution function.
φ	Dispersion parameter (related to the GLMs and GAMs).

1 Introduction

1.1 Scope of the Guidelines

This document aims to provide a simple but flexible guide for the construction of vulnerability curves from post-earthquake damage and loss survey data. The guidelines attempt to provide a rational and statistically rigorous approach for the construction of empirical fragility and vulnerability curves that explicitly quantifies the uncertainty in the data, and where possible reduce the epistemic uncertainty.

The Guidelines build on existing literature on empirical fragility and vulnerability functions, and consider how best to develop empirical functions from post-earthquake survey data of diverse type and quality (see Rossetto et al. (2013), for a description of common approaches to constructing fragility and vulnerability functions and a detailed discussion of empirical data quality issues). The guidelines have been developed so as to be used in constructing fragility and vulnerability functions from post-earthquake survey data recorded in the GEM Earthquake Consequence Database (GEMECD)[1]. However, databases which suffer from severe sampling bias or cover a very narrow band of ground motion intensity levels may not result in meaningful empirical fragility or vulnerability functions.

Guidance is provided for the construction of empirical vulnerability functions for a defined buildings class, i.e., a discrete or continuous relationship between an intensity measure type and a loss measure, from field data. In the context of this document, the seismic loss measures considered are expressed in terms of direct cost of damage, fatalities and downtime. In addition, ground shaking is considered the only source of seismic damage to the building inventory. A 'direct' and an 'indirect' approach are proposed for the construction of vulnerability functions depending on the nature of the available data. If loss data are available, a 'direct' approach should be used in order to relate the loss measure to an intensity measure type through the use of statistical model fitting techniques. By contrast, if damage data are available, an 'indirect' approach is necessary, which constructs vulnerability curves in two steps. First, suitable statistical model fitting techniques are adopted to construct fragility curves. Next, the fragility curves are transformed into vulnerability curves through the use of appropriate damage-to-loss functions (e.g., fatality rates or damage factors conditioned on damage state).

These Guidelines present a framework for the construction of empirical vulnerability and fragility functions that allows for a number of curve-fitting methods and ground motion intensity measure types (IMTs) to be adopted. The philosophy behind the design of the framework is that the characteristics of the data should determine the most appropriate statistical model and intensity measure type used to represent them. Hence, several combinations of these must be trialled in the determination of an optimum fragility or vulnerability curve, where the optimum curve is defined by the statistical model that provides the best fit to the data as determined by a number of goodness-of-fit tests. The Guidelines are essentially a roadmap for the process, providing recommendations and help in deciding which statistical model to attempt, and promote trialling of models and IMTs. It is noted that both direct and indirect empirical vulnerability procedures need to address

[1] http://www.globalquakemodel.org/what/physical-integrated-risk/consequences-database/

a number of uncertainties, as summarized in Figure 1.1. The present document provides guidance on the identification of sources of uncertainty, and where possible, their quantification and modelling.

The guidelines are targeted at analysts with Master's level training in a numerate subject that includes some level of statistics. The Authors recognise that the statistical analysis understanding of analysts varies. To accommodate for these differences, two levels of statistical approaches for constructing empirical fragility functions that include procedures of increasing complexity, are proposed:

- Level 1 approaches are appropriate for use by any analyst with a basic understanding of statistical modelling.
- Level 2 approaches can be applied by analysts who are confident in using advanced parametric statistical models (e.g. Bayesian regression analysis) as well as fitting non-parametric models to data.

To facilitate use of the Guidelines, the code and commands required for the implementation of the described statistical models are provided for the open-source software 'R' (2008). Furthermore, example applications of the Guidelines to different post-earthquake survey datasets are provided in Appendices B to G.

1.2 Relationship to other GEM Guidelines and Work

In the present state-of-the-art, seismic fragility and vulnerability functions can be derived through three different approaches: empirical, analytical and via expert opinion (Porter et al., 2012). The purpose of the GEM Global Vulnerability Estimation Methods (GEMGVM)[2] working group is to develop guidelines for the construction of each of these three types of vulnerability function. In term of relationship, the four guidelines produced by the working group (i.e., the present report; D'Ayala et al., 2014; Porter et al., 2014; and Jaiswal et al., 2014) are complementary. As the construction of empirical vulnerability functions requires the fitting of statistical models to loss and excitation observations, the white papers produced during the GEMGVM project by Noh and Kiremidjian regarding the use of Bayesian analysis (Noh et al., 2011) [3] and the fitting of nonparametric models (Noh, 2011b[4]; Noh et al., 2011; Noh et al., 2013) have been incorporated into this report.

The strategy foreseen by the GEMGVM consortium is that when consistent empirical vulnerability functions are lacking, the gaps are filled first using the results from analytical methods, and then by using expert opinion. This assumes that empirical vulnerability functions are the most credible type (Porter et al., 2012) and post-earthquake data can be found in the Global Earthquake Consequences Databases (So and Pomonis, 2012)[5]. However, the reliability of empirical fragility or vulnerability functions is questionable if few data of poor quality are available. For these cases, a system for evaluating the reliability of existing empirical fragility and vulnerability functions is provided in Rossetto et al. (2013). A possible procedure for combining existing fragility and vulnerability functions is also provided in Rossetto et al. (2014).

For what concerns the nomenclature of building typology and sub-typology and building attributes, reference is made to the classification recommended by GEM Building Taxonomy v2 (Brzev et al. 2013)[6]. For what

[2] http://www.globalquakemodel.org/what/physical-integrated-risk/physical-vulnerability/
[3] http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/fragility-function-updating-using-bayesian-framewo.pdf
[4] http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/fragilityusingkernelsmoothing-haeyoungnoh.pdf
[5] http://www.globalquakemodel.org/what/physical-integrated-risk/consequences-database/
[6] http://www.globalquakemodel.org/what/physical-integrated-risk/building-taxonomy/

concerns the hazard and the seismic demand reference is made to the output of the Global Ground Motion Prediction Equation component (e.g., Stewart et al. 2013a,b)[7], while for data on typology distributions, exposure and inventory we will refer to the Global Exposure Database (Gamba et al. 2012)[8] and the GEM source for damage and loss data from past events is, as mentioned previously, GEMECD[1].

[7] http://www.globalquakemodel.org/what/seismic-hazard/gmpes/
[8] http://www.globalquakemodel.org/what/physical-integrated-risk/exposure-database/

Uncertainty

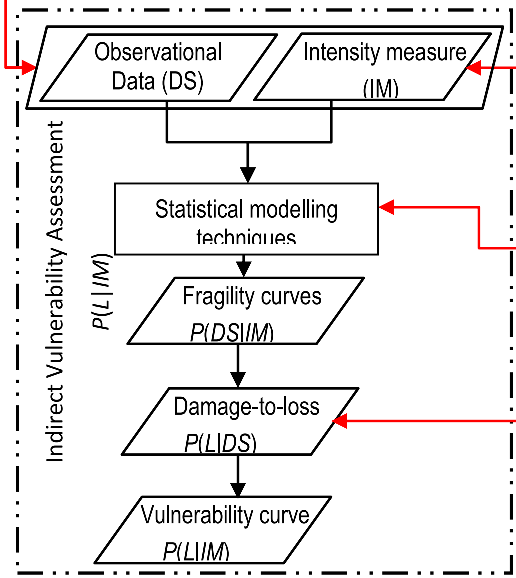
- Variation in response of similar structures for a given IM.
- Variation in geometric/material properties and seismic design of structures of a given class.
- Structural irregularities in structures of a given class.
- Incorrect classification of the observed damage.

For multiple databases:

- Uncertainty in converting between damage scales.

For a single database:

- Sampling errors- Limited number of observations for a given building class.
- Coverage errors.
- Non-response errors-Missing data.
- Processing errors when compiling the database.
- Incorrect identification of the hazards.



Uncertainty

- Variation of the estimated level of IM in a survey geographical unit.
- Lack of observed ground motion intensity levels.
- Uncertainty inherent in the GMPE.
- Selection of IM.
- Selection of GMPE.

Uncertainty

- Manipulation of data.
- Functional relationships.

Uncertainty

- Variation in the cost of a repair technique.
- Variation in the repair techniques used for the buildings of a given class.
- Variation in the extent of damage in the buildings of a given class.
- Variation in the level of replacement cost.
- Measurement error in the loss data.
- Bias/scatter due to limited field data.

Figure 1.1 Examples of sources of uncertainty associated with indirect empirical vulnerability assessment.

2 Proposed Framework for Vulnerability Assessment and Structure of the Guidelines

The proposed framework for the direct and indirect construction of empirical vulnerability functions consists of seven and eight main steps, respectively, as illustrated in Figure 2.1. The structure of this report (the Guidelines) essentially follows the main steps of the framework, and the relevant sections of the report are also referred to in Figure 2.1.

In Step 1, the quality and quantity of the available loss or damage data are assessed. If these are found to be acceptable then the data are prepared for use in the model-fitting procedure. In §3, minimum levels of information and number of buildings required for the development of reliable vulnerability and fragility functions are proposed. Guidance is also provided in identifying, quantifying and reducing the bias in a database, as well as how to combine multiple databases.

In Step 2, a set of intensity measure types for use in vulnerability function derivation is proposed. In §4 guidance is also given on procedures for evaluating their values for the locations of the available loss or damage data, and how to identify sources of uncertainty in the IM evaluation.

Vulnerability or fragility functions are stochastic relationships which express the loss or damage as a function of ground motion characteristics and buildings characteristics. These functions are constructed by fitting a parametric or non-parametric statistical model to the post-earthquake data. Three models are proposed here namely, the Generalized Linear Model (GLM), the Generalized Additive Model (GAM) and the Gaussian Kernel Smoothers (GKS). In Step 3, statistical models suitable for use in constructing the vulnerability or fragility functions are selected. In §5, guidance is provided for identifying the statistical models which best fits the data as well as estimating the level of confidence around the obtained vulnerability and fragility functions.

In Step 4, the statistical models selected in Step 3 are fitted to the post-earthquake data. §§6-8 describe 3 model fitting procedures of increasing complexity, which can be used to construct the vulnerability or fragility functions and their confidence intervals. Procedures for assessing the goodness-of-fit of the selected models to the data are also provided. Blue boxes translate procedures discussed in the text into the open-source statistical programing language and software environment called 'R' (R Development Team, 2008) to assist the analyst in performing the model fitting and goodness-of-fit assessments.

It is emphasised that the guidance provided is intentionally not prescriptive, and a set of statistical model fitting methods are outlined for the vulnerability or fragility assessment in Step 4. The reason is that the 'nature' of the available loss or damage and ground motion intensity data, (i.e., any issue regarding the quality and quantity of these data, see Rossetto et al., 2013), will influence the selection of statistical model, intensity measure type and fitting method adopted.

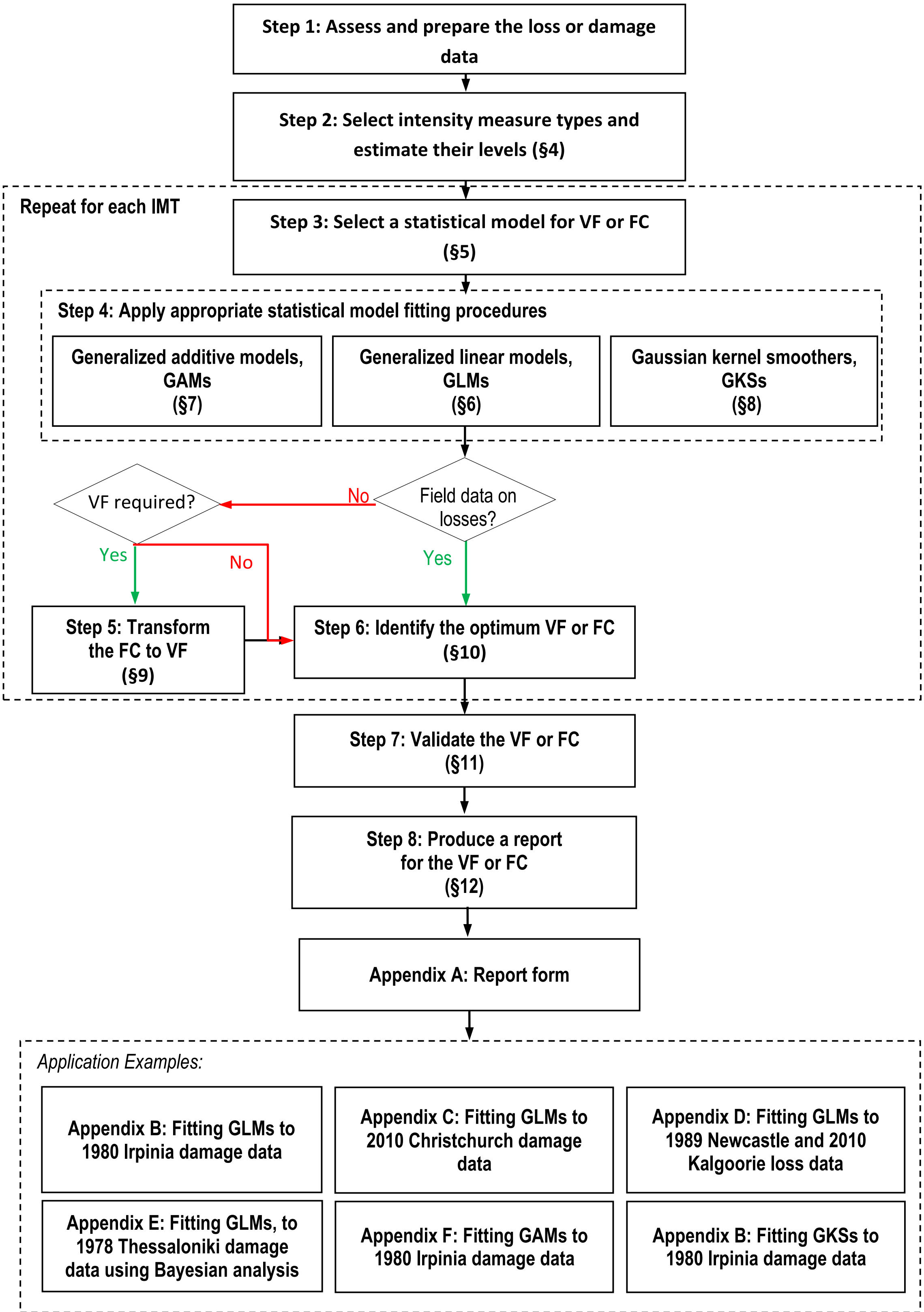


Figure 2.1 Illustration of the framework for direct and indirect empirical vulnerability assessment. Reference is made in the flowchart to relevant sections of the guidelines document (VC: vulnerability functions, FC: fragility functions).

The authors also recognise that the level of understanding of the analyst around the statistical model fitting techniques varies. For this reason, within each step of the framework, proposed procedures are grouped into two levels of difficulty:

- Level 1 approaches require a basic understanding of parametric statistical model fitting techniques. The analyst can simply copy the 'R' code provided and produce fragility or vulnerability curves following the supplied examples.
- Level 2 approaches require that the analyst is confident in the use of advanced parametric statistical model-fitting techniques and has a good understanding of non-parametric statistical models. For Level 2 approaches, general guidance is provided and it is left to the analyst to advance the complexity of the models and to identify the vulnerability or fragility function that best fits their data, based on the proposed goodness-of-fit tests and their experience.

Even in the case of experienced analysts, it is recommended that Level 1 approaches are attempted first, and more complex procedures adopted only when the simpler Level 1 approaches do not yield a satisfactory fit of the empirical data to the statistical model.

Step 5 is necessary only if damage data are available and an 'indirect' approach is used for vulnerability assessment. In this step the empirical fragility curves, obtained from Step 4, are transformed into vulnerability curves following the procedures presented in §9.

The framework envisages that Step 4 be repeated for several IMs, and that the choice of final vulnerability function be based on a consideration of which IM results in the empirical fragility or vulnerability function associated with the lowest uncertainty. Hence, in Step 6, the optimum empirical vulnerability or fragility function is identified. A procedure for this step is provided in §10.

In Step 7, the optimum empirical vulnerability function is assessed on whether it is fit for purpose. §11 outlines procedures for assessing the performance of the optimum empirical vulnerability relationship against a new database of empirical data, or subset of the empirical database that was not used for the construction of the fragility or vulnerability function.

Finally, in Step 8 it is recommended that a report is prepared, which summarises the results of the vulnerability assessment, the adopted procedures in each step and any unresolved issues. Appropriate reporting must accompany any vulnerability or fragility function that is to be included in the GEM Global Vulnerability Database. A template for this type of report is provided in Appendix A, and applications of it shown in Appendices B to G.

Six applications of the proposed guidelines are found in Appendices B–G. In particular, the damage data collected in the aftermath of the 1980 Irpinia (Italy) earthquake are adopted in order to illustrate the use of GLMs, GAMs and GKSs in Appendices B, F and G, respectively. The damage data from four successive strong events that affected Christchurch (New Zealand) in 2010 are used in order to illustrate the GLMs and how to construct fragility curves for successive strong events in Appendix C. Appendix D constructs vulnerability functions by fitting GLMs to the loss data from the 1989 Newcastle and the 2010 Kalgoorie earthquakes in Australia. Finally, Appendix E uses Bayesian analysis in order to fit a GLM model to the 1978 Thessaloniki damage data.

3 Step 1: Assessment and Preparation of Loss or Damage Data

3.1 Introduction

The quality of vulnerability and fragility curves strongly depends on the quality and quantity of observations in the adopted empirical database. The survey method, the sample size (i.e., the number of buildings surveyed as a proportion of the total population of buildings located in the affected area), and level of detail of the collected information affect the reliability of the data. These factors also affect the level of detail in the analyst's determination of damage scales and building classes for the construction of the vulnerability/fragility curves. A detailed discussion of the main sources of observational damage and loss data, and their associated characteristics can be found in Rossetto et al. (2013), and are summarised in Table 3.1.

Table 3.1 Database typologies and their main characteristics.

Type	Survey Method	Typical Sample Sizes	Typical Building Classes	Typical No. of Damage States	Reliability of observations	Typical issues
Damage	Rapid Surveys	Large	All buildings	2-3	Low	Safety not damage evaluations. Misclassification errors.
	Detailed "Engineering" Surveys	Large to Small	Detailed Classes	4-6	High	Unrepresentative samples.
	Surveys by Reconnaissance Teams	Very Small	Detailed classes	4-6	High	Unrepresentative samples.
	Remotely sensed	Very Large	All buildings in an area	2-3	Low	Only collapse or very heavy damage states may be reliable. Misclassification errors.
Economic Loss	Tax assessor data	Very large	All buildings/ Detailed classes	-	High	May include data on damaged buildings only.
	Claims data	Very large	All buildings	-	High	Concentrate on damaged and/or insured buildings only.

Casualties	Government Surveys	Very large	All buildings	1-2	Low	Unlikely association of building damage and causes of injury
	Surveys by NGOs/ hospitals	Varies	All buildings	1-5	Low	Possibility of unrepresentative samples.
	Detailed Casualty Surveys	Very Small	Detailed classes	3-5	High	Possibility of unrepresentative samples.

In this section, guidance on determining the minimum level of information and sample sizes necessary for the construction of reliable vulnerability or fragility curves is provided. In addition, procedures for dealing with biases in the databases (refer also to Rossetto et al., 2013 for discussion of data biases) or small sample sizes are also presented.

Guidance is then provided on how to transform data from post-earthquake surveys into a set of data points for the regression analysis of direct vulnerability and fragility curves (see Figure 3.1).

For the construction of fragility curves, a data point represents an intensity measure level and the corresponding damage ratio, i.e. the fraction of buildings of a given class that suffered damage reaching or exceeding a damage state, divided by the total number of buildings in the examined building class located in an area affected by a specified intensity measure level. Similarly, a data point used in the construction of direct economic loss curves is expressed in terms of an intensity measure level and a corresponding damage factor, i.e., the repair cost divided by the replacement cost of a single building (or group of buildings within the building class of interest) exposed to the specified intensity measure level. Finally, a data point for the construction of direct casualty curves is expressed as an intensity measure level and a corresponding casualty ratio, i.e., the fraction representing the number of people who died (or who were injured and died) divided by the total exposed population in an area affected by the specified intensity measure level.

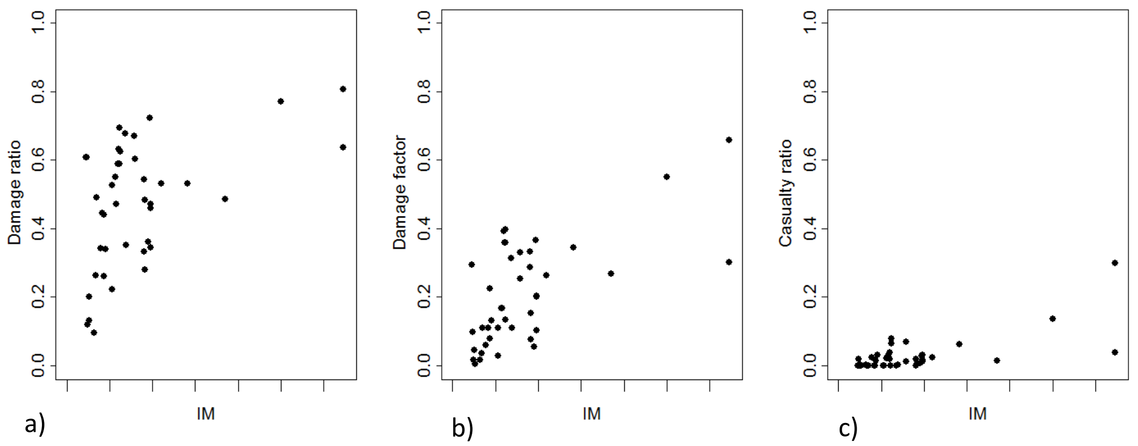


Figure 3.1 Fictitious data points for the direct construction of (a) fragility curves and (b) vulnerability curves for economic loss measured in terms of a damage factor and (c) vulnerability curves for human loss expressed in terms of a casualty ratio.

3.2 Preliminary Quality Assessment of a Single Database

Before beginning the process of data treatment for the construction of a fragility or vulnerability curve, the basic quality of a damage or loss database should be examined to ensure that the constructed curves are meaningful and fit for purpose. The quality of the database depends on the design and execution of the survey used to collect the data as well as the level of information collected for the buildings and the site where they are located. This section provides guidance on the basic characteristics of a high, medium and low quality database for use in the construction of fragility and vulnerability curves. It also provides guidance for identifying databases of an unacceptably low quality that are unlikely to yield reliable fragility or vulnerability functions. Figure 3.2 depicts the categories of information required to assess the quality of an empirical database.

A *high quality* database should derive from a reliable and detailed survey. A large sample (>100 buildings) of buildings in each building class is collected over a wide range of intensities measure levels, thus allowing the construction of a fragility or vulnerability curve. The sample should be representative of the total population of affected buildings, e.g., the sample should not contain data only on damaged buildings. In addition, the location (e.g., longitude-latitude) of each building and soil on which it is founded need to be well-defined. Ideally, the buildings should be located in the close vicinity of ground motion recording stations to allow a good estimation of the seismic excitation they were subjected to. Building classes should be described according to the construction material, the structural system, the height and the occupancy of the building, which are considered the minimum set of attributes for the definition of high quality. A detailed description of damage or loss should also be available, e.g., four or more damage states are used to describe the damage.

Medium quality databases are based on less detailed or reliable surveys. These databases contain a sufficient sample (>30 buildings) of buildings in each building class, collected over a wide range of intensities measure levels. However, the building classes may not be described in as much detail as for the high quality databases, e.g., the height of the buildings is not provided or the structural system is not clear. Similarly, the description of damage or loss may be less detailed, e.g., the three safety tagging system is used, which can ascribe a wide range of damage levels to each safety tag. In some cases, the exact location of the buildings may not be clear, e.g., aggregated information of damage or loss is provided at the level of a municipality. These databases can be used for vulnerability assessment but the resulting empirical functions might not be able to accurately capture the uncertainty due to the aggregated nature of the observational data. Moreover, medium quality databases can be characterised by the absence of ground motion stations from some or all the locations of interest. In this case, the measurement errors in the intensity measure levels should be modelled.

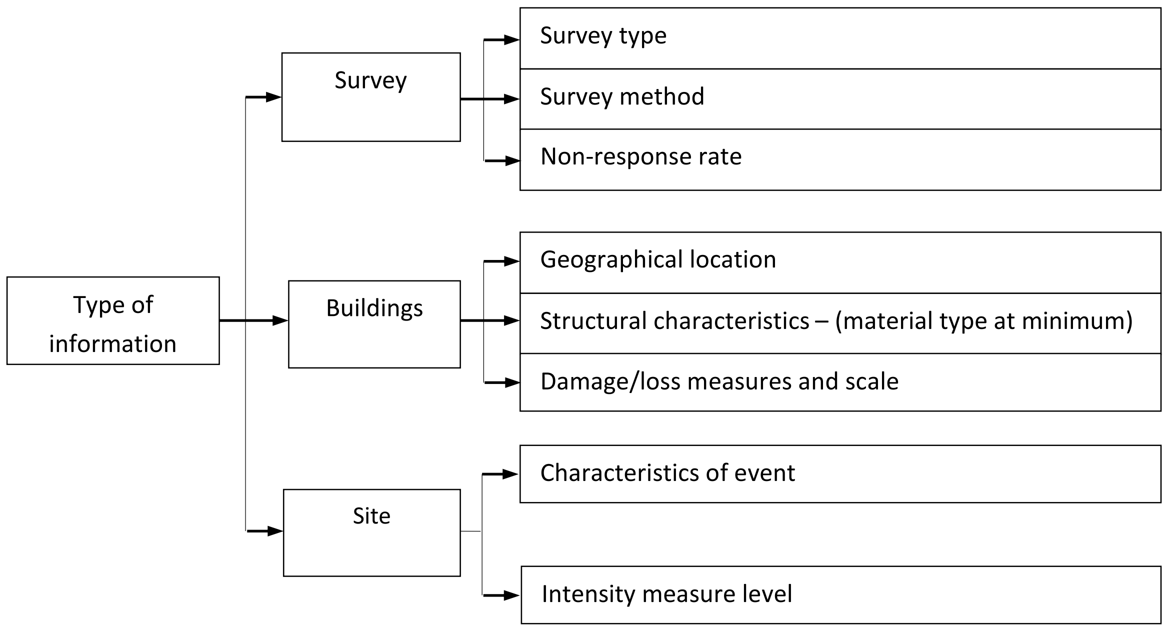


Figure 3.2 Minimum level of information for an empirical database.

A *low quality* observational damage or loss database is one that, in addition to the attributes of a medium quality database, is also associated with errors in the design or the execution of the surveys, e.g., only damaged buildings are surveyed or there is a known misclassification of building damage. In this case, the databases can be used for the construction of empirical vulnerability assessment only if these errors are reduced through the use of independent databases, as described in §3.3.

Finally, a database is of *unacceptably low quality* and should not be used for the construction of vulnerability or fragility functions if:

- The sample size of a buildings class is smaller than 30 buildings.
- The data are concentrated in single level of macroseismic intensity or a very narrow range of intensity measure levels and no additional databases are available that could expand the range of intensity measure levels.
- The database is heavily biased and no independent databases are available for reducing this bias

If the analyst is able to characterise their observational damage or loss database as high to low quality, the next step is to address any biases in the data and prepare the data for use in the statistical modelling.

3.3 Preparation of Loss or Damage Data from a Single Survey Database

Within the following sub-sections guidance is provided for the preparation of empirical damage or loss data for use in the construction of fragility or vulnerability functions, respectively. Here it is assumed that the empirical data being used derive from a single earthquake event and an in-field survey procedure. In particular, advice for the identification of potential sources of bias and uncertainty in the database is provided. It is emphasised that in cases where multiple surveys are adopted, each database should be checked for these sources of uncertainty.

3.3.1 Type of Data

Irrespective of the data source, the empirical damage or loss data are typically presented at two levels of detail:

- Building-by-Building data: when loss or damage data are recorded at an individual building basis. These data typically provide detailed information regarding the location of the affected buildings as well as the level of loss or damage they have sustained. Such data are more likely to be obtained from detailed surveys, specialist reconnaissance, or from the interpretation of aerial or satellite imagery.

  With regard to the construction of fragility curves, the data point based on this type of data can take two values:

$$\text{data point} = \begin{cases} 1 & \text{the building sustained damage } DS \geq ds_i \\ 0 & \text{the building sustained damage } DS < ds_i \end{cases} \tag{3.1}$$

  With regard to the construction of vulnerability curves for economic loss, the data point can be expressed in terms of the damage factor for each building:

$$\text{data point} = \frac{\text{repair cost for a given building}}{\text{replacement cost for a given building}} \tag{3.2}$$

- Grouped Data: when empirical distributions (e.g., histograms) of the loss or damage observations for a building class over defined areas or IM values are reported. These data contain less information and are typical of past earthquake surveys for which detailed data were not collected, where loss data have been aggregated to maintain anonymity or in cases where the original survey forms are unavailable. The construction of useful models may be difficult for highly aggregated databases due to the difficulty in exploring alternative (perhaps more complex) models indicated by the diagnostic procedures.

  With regard to the construction of fragility curves, the data point for a given geographical unit based on this type of data is expressed in the form:

$$\text{data point} = \frac{\text{No of building in a given unit sustained damage } DS \geq ds_i}{\text{Total No of buildings in a given unit}} \tag{3.3}$$

  With regard to the construction of vulnerability curves for economic loss, the data point for a given isoseismic unit can be expressed in terms of the damage factor for each building:

$$\text{data point} = \frac{\text{repair cost for all building in a given unit}}{\text{replacement cost for these building}} \tag{3.4}$$

A single survey database could contain both types of data. In this case it is recommended that the database be brought to a common form, i.e., that the building-by-building data be aggregated in a similar way to the grouped data present in the database.

3.3.2 Identification and treatment of uncertainty

The level of uncertainty associated with a single database depends on sampling as well as non-sampling errors, as presented in Figure 3.3 (UN, 1964). Sampling errors occur because only a subset of the population of buildings located in the affected area has been surveyed. Non-sampling errors represent the remaining sources of error occurring in the survey design, as well as errors in the collection and processing of the data.

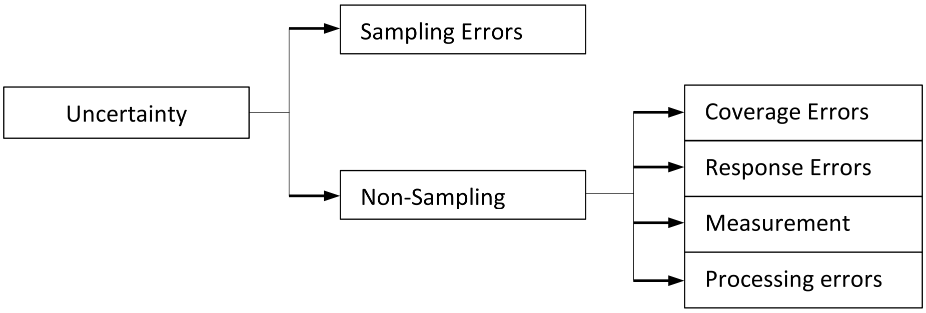


Figure 3.3 Errors in a single database.

Four main sources of non-sampling error have been identified (United Nations, 1982), namely: coverage, response, measurement and processing errors (see Figure 3.3). A brief description of these four types of errors is presented in Sections 3.2.2.1 to 3.2.2.5, where guidance is provided on how to reduce or eliminate them from the database. The proposed procedures depend on whether the identified large errors are random or systematic. The influence of the nature of these two types of errors on a variable *X* is illustrated in Figure 3.4. Figure 3.4a shows that random errors do not affect the 'true' mean of *X* but do affect the level of its uncertainty, with large random errors leading to increased spread in the density distribution of *X*. Random errors can be reduced mainly through the use of larger sample sizes. By contrast, Figure 3.4b shows that high systematic errors lead to estimates of mean that differ widely from the 'true' mean. The elimination of systematic errors requires procedures based on the analyst's understanding of the mechanism that generated the errors and through comparisons with independent databases.

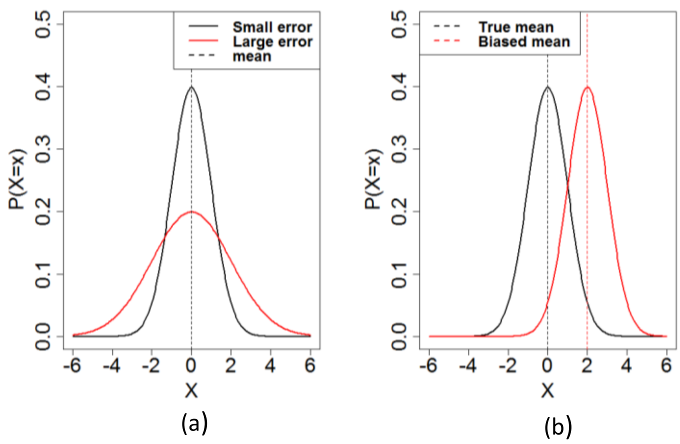


Figure 3.4 Impact of (a) random error and (b) systematic error on variable *X*.

In the following sections, common sources of errors found in seismic loss or damage databases are identified and the nature of the errors taken into account to suggest approaches for their reduction or elimination. Adjusting for non-sampling errors is a largely unexplored subject in empirical vulnerability assessment literature and limited guidance is provided below. The analyst may refer to the broader literature for more information, e.g., Lessler and Kalsbeek (1992).

3.3.2.1 Sampling error

Sampling error is introduced when the database contains only a subset of the total number of buildings located in the affected area. The sampling error is acceptable if the subset is shown to be representative of the target population of buildings. The required number of buildings required for the sample to be representative can be calculated from standardised procedures, which vary according to the sampling approach (see Levy and Lemeshow, 2008).

However, most post-earthquake surveys are not designed to collect a statistically representative sample of buildings in the affected area. Samples are often small and may lead to unacceptably wide confidence intervals around the mean vulnerability and fragility curves. In this case, the analyst should ideally increase the sample size with a new field survey designed using a well thought out sampling technique to ensure representativeness. Alternatively, sampling error can be, to some extent, reduced by combining the small database with one or more other surveys from the same event or from different events for similar tectonic environments, which affected similar building inventories (see Section 3.5).

3.3.2.2 Coverage error

Coverage errors is a systematic error which occurs in cases where the database of damage or loss observations is not representative of the target population of the considered asset, e.g., total number of buildings in the affected area. In other words, the database is biased and larger sample sizes cannot reduce this error. Typical coverage errors include missing data and duplicate data:

- Missing data introduce a "no-coverage" or "under-coverage" error. In general, the analyst should quantify this error by comparing the survey observations with data from independent sources. If the difference between the survey and independent database is found to be unacceptably high, e.g., >10%, then this error should be eliminated from the database. Typical examples of missing data found in the empirical vulnerability literature are presented in Table 3.2 together with suggested approaches to treating or eliminating these errors.
- Duplicate data introduce an "over-coverage" error. This error is found when data gathered from different sources overlaps (e.g. Eleftheriadou and Karampinis, 2008) or when field surveys are carried out with overlap occurring between the areas covered by different survey teams. If the analyst has access to the survey forms, then they should try to remove the duplicate entries from the database. If access to the survey forms is not possible, but there are strong indications of over-coverage, then the analyst should not use the total number of units in the database but they should randomly select a large sample, e.g. 50%, of available buildings, which can reduce impact of the over-coverage error.

Table 3.2 Identification and elimination of coverage errors found in the literature.

Error	Example	Treatment
Under-coverage	Damage data are collected by surveying the buildings only upon the building owner's request.	If it is safe to assume that almost all the damaged buildings have been surveyed in the affected area, then the total number of buildings can be estimated using a building counts from a census (compiled as close to the event as possible).
	The data regarding losses below the deductibles are absent from a database of insured loss data.	The analyst can use the existing database to construct direct vulnerability curves for loss values greater than the deductible. If, however, the zero losses or the losses below the deductibles are required, additional

		information should be obtained from other sources that may assist the analyst to adjust for this bias.

3.3.2.3 Response error

Response error occurs when for parts of the survey all desired attributes of a building, damage level or loss that are required for a fragility or vulnerability assessment are not collected. This can happen when the survey is conducted rapidly, by an inexperienced team, when there is poor supervision or when there are problems with survey questionnaire used e.g., it is too lengthy and/or unclear. How can response errors be identified? The analyst is expected to be able to contact the authority/organisation that collected the data and investigate in detail the survey method and identify the whether there is notable response error as well as its nature. Response errors can either be random or systematic in nature, and this needs to be identified through examination of the survey forms and any observable patterns. For example, systematic errors may occur when a particular survey team consistently omits to fill in that the surveyed buildings are undamaged in the survey form. In both random and systematic response error cases, the ratio of incomplete to complete survey forms (termed rate of non-response) should be estimated. The type of non-response should also be identified. Typically this is of two types: (1) total non-response errors, where survey forms for buildings are missing or inappropriately geo-referenced, such that the location of the building cannot be identified (2) partial non-response errors, where only certain fields of the survey forms for buildings are incomplete.

In the case of total non-response errors, depending on the rate of non-response, the database may still be used. For example, if damage data is grouped at the level of a municipality (with data available for multiple municipalities), and the total non-response rate in any one municipality is ≤5%, the analyst may assume that the database for this municipality is complete. If the non-response rate is >5% then the analyst needs to understand the characteristics of the non-surveyed buildings, either by the gathering more information on the survey or from independent surveys of the same area, to assess whether the error is random or systematic. If the error is random, higher non-response rates can be adopted but the analyst needs to perform sensitivity checks (e.g., see Rota et al 2008), to establish the level of error introduced by the assumed rate of completeness. If the error is systematic, the analyst should explore whether the data can be completed with additional census data (e.g. if the non-response errors pertain to undamaged buildings only), or whether smaller geographical units with lower non-response rates could be used.

Partial non-response error occurs when only some of the fields of the survey form are missing. If the missing data are ≥5% of the total collected information (e.g. Graham, 2009), then the analyst should try to identify whether this error is systematic or random.

The data can be missing completely at random, e.g., if one or more teams omitted accidentally some information from the survey form. In this case, the analyst can remove the buildings with missing data altogether. This will reduce the sample size but the database is unbiased.

The data can be missing at random. For example, the structural system of RC buildings is more likely to be missing for severe damage states due to the inability of the teams to access severely damaged buildings. In this case, the missing data can be completed by the observed data using a procedure termed multiple regression imputation. According to this procedure, the missing data of a given variable are completed by fitting a parametric statistical model using as response variable the variable whose missing values are estimated, and as explanatory variables all the other variables in the database. The procedure is performed for all variables. The user is referred to Schafer (1999) and Little and Rubin, (2002) for more information regarding this procedure, which has not been applied to the field yet. In most cases, it is not straightforward to identify whether the mechanism is missing completely at random or at random. In these cases, the analyst is advised to follow both procedures and compare the two models.

Finally, the data are not missing at random e.g., if one or more teams systematically fails to enter the slight damage in the survey forms. In this case, more information regarding the reason why they are missing is required in order to reduce this error.

3.3.2.4 Measurement error

Measurement error describes the deviance of the observed value of a variable from its ‘true’, but unobserved value. In the case of post-earthquake survey data such errors usually arise due to misclassification, i.e. a building is attributed to a damage state when it is really in another. This error commonly occurs due to the data collection method (e.g., remote sensing procedures are prone to large misclassification errors), the procedure followed in the field (e.g., rapid external assessments of building damage can result in misclassification errors due to ignorance of internal damage extent), the use of inexperienced engineers for the field surveys. The quantification of measurement errors is difficult and requires comparison with additional highly detailed survey data for the same area or for a representative subset of the area.

It is important to note that misclassification error does not only concern the damage or loss evaluation but also the hazard. In the latter case, bias is introduced by considering that the damage/loss results from the main shock, rather than also as a result of large aftershocks. In cases where secondary hazards (e.g., fire) have contributed to the recorded damage or loss this source of error should be quantified from independent surveys and the data should be appropriately adjusted or corrected.

In §6, a procedure for including misclassification error in the response variable in the model fitting procedure is proposed.

3.3.2.5 Processing errors

A damage/loss database is compiled from a number of survey forms. Processing the data may result in errors due to typing mistakes, misunderstanding of the nomenclature, etc. If the analyst has access to the original survey forms they should check whether the data are accurate. In general, processing errors of peer reviewed databases are considered negligible.

3.3.3 Construction of building classes

A vulnerability and fragility curve should be based on observations from buildings with very similar seismic performance (i.e. for a building class), which ideally minimises the uncertainty due to the seismic response (see Figure 1.1). This can be achieved by defining highly detailed buildings classes, constructed following the definitions of GEM Building Taxonomy v2 (Brzev et al. 2013). In practice, however, the level of detail for which the building classes can be defined depends on:

- The level of detail recorded in the post-earthquake building survey forms. Typically, detailed ‘engineering’ and reconnaissance surveys include detailed descriptions of the buildings structural characteristics. By contrast, remote sensed studies, rapid surveys and claims data include broad descriptions of the buildings.
- The total number of buildings belonging to each class. The total number of buildings in the database is affected by the type of survey (see Table 3.1). However, it is highlighted that even in the case of surveys that include large samples, the sample sizes of particular building classes that are not common in the affected area can be small.

- The distribution of data points resulting from a building classification across a range of IM values. If the data points resulting from a particular building classification are highly clustered, a wider range of IMs may be obtained from a coarser definition of the building class or through the combination of multiple surveys for similar building types (see §3.5 for the latter).

A reliable fragility or direct vulnerability curve depends on a minimum number of buildings and data points (see §3.4). Hence, the building classes are determined as a compromise between level of detail in the building class definition and the number of observations required to construct a vulnerability and fragility function with acceptable levels of confidence.

3.3.4 Determination of damage/loss scales/measures

When using existing damage/loss databases, the analyst is advised to use the same definitions for damage/human loss scales and economic loss measures as the data source in order to avoid the unnecessary introduction of uncertainty. However, if these definitions are not appropriate for the overall risk assessment, then the uncertainty introduced from conversion between parameters/scales should be considered and quantified if possible. If the error cannot be quantified, it should at least be acknowledged. Conversion between damage scales is covered in Rossetto et al., (2014).

If the analyst aims to collect new post-earthquake damage/loss data or if they need damage-to-loss relationships for the indirect assessment of vulnerability, then the guidance provided in Rossetto et al., (2014) should be followed in selecting the most appropriate damage/human loss scales and economic loss measures.

3.4 Minimum Number of Data

A database including *high quality data*, as defined in §3.2, can result in reliable mean vulnerability and fragility curves, (i.e., having very narrow confidence intervals), only if it contains a large number of observations, which result in data points that are distributed across a wide range of ground motion intensity values. In general, the minimum number of data points required to construct a vulnerability or fragility function depends on the level of uncertainty the analyst is willing to accept, with a smaller number of points resulting in larger uncertainties. However, a determination of the level of confidence around the mean vulnerability or fragility curves that is achievable by using different sample sizes for the construction of direct vulnerability and fragility curves requires a sensitivity analysis to be carried out and the presence of a very large number of available observations. This can be problematic if the analyst wants to decide whether a sample size is adequate before fitting their statistical models.

Rules of thumb for establishing minimum number of sample sizes necessary for the prediction of the trend in linear models have been proposed in the literature. For example, researchers (Miller and Kunce, 1973; Harrell et al, 1985; Bartlett et al, 2001, Babyak 2004) commonly advocate the used of 10 data points for a single explanatory variable. Other researchers increase the number of data points to 51 (e.g., Harris, 1975). In other cases 100 data points are suggested (see Green, 1991), whilst 200 data points are proposed by both Guadagnoli and Wayne (1988) and Nunnally and Bernstein (1994).

The user is advised to ensure a sample size of at least 100 observations spanning a wide range of IM values for the construction of vulnerability curves from data reported at the building-by-building level. In the case of fragility curves, a minimum of 100 buildings should be used and at least 30 of them should have reached or exceeded a given damage state (see also Noh etal., 2012)[9], again with the data points spanning a wide range of IM values.

In the case of grouped loss or damage data (typical of *medium* or *low quality* databases, see §3.2), a database comprising at least 200 building surveys (individual building observations) and resulting in at least 10 data points (each data point being grouped data based on a similar number of building observations), can be considered acceptable for the construction of vulnerability function or damage-state specific fragility functions. Data points obtained from observations (of loss or damage) made on less than 30 buildings can be used for the construction of vulnerability and fragility curves but the analyst should expect very large confidence intervals around the mean vulnerability or fragility, which will question their usefulness for the prediction of damage or loss for future events.

It is advisable to assess the feasibility of constructing vulnerability or fragility functions from databases early on in the process, in order to ensure the desired outcome is achievable within acceptable limits. It is recommended that not only the quantity of data points available for each damage/loss limit state be assessed, but also the spread of these data points across the IM range desired. For example, if the data are clustered in the low damage/loss state area and low IM value range (typical of empirical data collected from small to moderate earthquakes), vulnerability or fragility functions constructed from this data will not give a meaningful prediction for high damage/loss states at low IM values or for any damage/loss state at high IM values.

In the case of small or highly clustered datasets, it is recommended two different strategies. Firstly, the analyst is advised to use a Bayesian analysis, which informs existing vulnerability or fragility curves with the small available database. Alternatively, the analyst should consider adopting data from multiple surveys for the construction of vulnerability and fragility functions (see §3.5). In the case of datasets with some clustering it is possible to construct fragility functions, but the analyst is limited in the types of regression techniques this guideline would recommend (see §5).

3.5 Preparing Empirical Loss or Damage Data from Multiple Databases

Multiple databases of damage or loss from post-earthquake surveys can be adopted in the construction of single sets of vulnerability and fragility relationships. This may be desirable if the number of observations in one or more single databases is small and/or does not cover a wide range of *IM* values. However, care should be taken when considering which databases to combine, with only data from regions with similar tectonic environments and for the same building classes used. In the literature observational data from different databases has been typically been aggregated into a single database with little or no consideration of the relative reliability of each database (e.g. Rossetto and Elnashai, 2003; Rota et al, 2008; Colombi et al, 2008). In the present guidelines, the analyst is given the opportunity to construct vulnerability or fragility curves accounting for the different sources of the observations and their reliability. Guidance for this is provided in §6.3.6 and §7.3.6.

In the case of multiple surveys, the procedures set out in §3.3 and §3.4 for single surveys should still be carried out for each database, so as to determine data type, minimum number of observations per dataset, appropriateness of number of data points and their spread across the *IM* range, and sources of uncertainty associated with the database. In addition, if the available surveys correspond to the same earthquake, the

[9]http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/issues-in-empirical-fragility-functions.pdf

analyst is advised to use only the database with the higher quality. If both databases are of acceptable quality then oth can be used but occurrences of repeated data should be checked for and eliminated.

3.5.1 Harmonising data type

The types of data of multiple databases should be expressed in a common form in order to be adopted in the construction of vulnerability and fragility curves. If one database presents data at a building-by-building level and another contains grouped data, then the more detailed building-by-building data should be aggregated and transformed into the same form as the grouped data. The aggregation process should be carried out after a common building class and damage/loss scale has been determined and the data appropriately transformed (see §3.5.2 and §3.5.3).

The aggregation method should also follow that of the grouped database. For example, grouping over an area with similar IM may be done by aggregating buildings in a geographical area (e.g., zip-code or town), or by grouping all observational data for all areas deemed of constant IML (e.g., in an isoseismal). However, the number of observations in the new (combined) database should also be checked to satisfy the minimum number of data criteria set out in §3.4, and the new database assessed for new sources of uncertainty resulting from the aggregation process.

3.5.2 Harmonising building classes

The combination of multiple databases in empirical vulnerability and fragility assessment cannot be achieved unless observations in the individual databases refer to the same building class. However, each database may include different structural characteristics of the surveyed buildings. Once the GEM Building Taxonomy v2 has been followed to assign building class definitions to each database, the lowest level (coarsest) definition of building class present in the databases should be applied when describing the resulting vulnerability and fragility function.

3.5.3 Harmonising the damage/loss scales/measures

The least uncertainty is introduced in the empirical vulnerability assessment if the adopted databases adopt identical damage scales or loss measures. However, the differing purposes of surveys or changes over time in survey methods and tools lead to the use of different scales/measures even for earthquakes occurring in the same country. In these cases, the analyst is advised to select an appropriate damage/human loss scale or loss measure from those adopted by the databases being combined, and converting all the damage/loss data to this scale. Strong assumptions have to be made to transform a database presented in terms of few damage/loss states into one consisting of a larger number of damage/loss states. These would introduce large, and poorly quantifiable, uncertainties in the resulting vulnerability and fragility curves. Hence, the analyst is advised to adopt the damage/loss scale with the coarsest definition and lowest number of damage/loss states, and transform the more detailed survey data to this scale.

No guidance is here provided to convert between (economic or human) loss scales, as the measures used in their definition vary significantly between surveys. In the case of fragility functions, damage scale conversion tables are provided in Rossetto et al. (2014) for common reinforced concrete, masonry and timber types of construction. The list of damage scales present in these tables is not exhaustive but includes those damage scales that have been reported in the majority of past empirical fragility curves. These tables can be used to map the more detailed to the coarser damage definitions. Where the more detailed damage state definition appears to straddle two of the coarser damage states, it is recommended that observations from the detailed damage state in question be mapped onto the lower damage state of the coarser scale (it is not recommended that assumptions be made to distribute the observational data between the two damage states).

4 Ground Motion Intensity

Within this section a brief overview of the ground motion intensity measures (IM) most commonly adopted in the construction of empirical fragility and vulnerability functions is presented. Suggestions are then provided for the appropriate evaluation of the IMs at sites, where the surveyed buildings are located. As outlined in §2, these guidelines prescribe that several IMs should be used for the construction of empirical vulnerability curves. Ideally, the optimum of the aforementioned vulnerability curves is the one which yields the fitted vulnerability or fragility curve with the smallest confidence bounds.

The incorporation of the uncertainty in the ground motion intensity in the empirical vulnerability assessment is a relatively new subject. Straub and Der Kiureghian (2008) have been the first (and only to date) to construct empirical fragility functions (for electrical substations) that account for measurement error in the ground motion intensity. Recently Ioannou et al. (2014) highlighted the importance of this error in the shape of the fragility curve, thus it is believed by the authors of this Guideline that the incorporation of uncertainty in IM is important for an accurate evaluation of the vulnerability and fragility functions as well as their confidence and prediction intervals. Sources of uncertainty in IM determination are therefore highlighted in this section. A first attempt procedure is also proposed to incorporate uncertainty in *IM* in §6. It is highlighted that this proposed procedure has only been tested on a few datasets, and as it is still unclear whether including uncertainty in the *IM* of vulnerability and fragility curves is a better approach than the established procedure, which models the measurement error of IM in the hazard curve for the assessment of the overall risk, the proposed method for carrying out the former should be regarded as novel but optional. In any case, that care should be taken to avoid double counting the measurement error of IM in risk assessment

4.1 Introduction to Intensity Measures and Considerations for Empirical Vulnerability

There are two main categories of ground motion intensity measures that are used in empirical vulnerability and fragility assessment: (1) those based on macroseismic intensity scales (e.g., Modified Mercalli Intensity, MMI); (2) those based on instrumental quantities (e.g., peak ground acceleration, PGA). Here, a general description of these intensity measures is presented together with some considerations that should be taken into account when using them for the construction of empirical vulnerability curves.

Macroseismic intensity scales are based on how strongly the ground shaking is experienced in an area, as well as observations of building damage. This means that they are readily available for any site where there are humans present to observe, and buildings present to be damaged. In fact, the same damage survey that is used to collect fragility or vulnerability data can also be used to estimate the macroseismic intensity. In a sense, this means that macroseismic scales are really an intermediate quantity between fragility and intensity – they already include information about building fragility in the areas in which they have been calibrated. This should be taken into account when applying a macroseismic scale outside the area in which it was originally developed – for example, using the European Macroseismic Scale (EMS) outside Europe. We could expect that regressed fragility relationships based on macroseismic intensity scales would have tighter confidence intervals, since damage data are already implicitly included in the intensity measurement.

Moreover, Intensities are recorded as some average of effects over some spatial scale. This averaging is seen to reduce the standard deviation for Intensity Prediction Equations (IPE) with respect to Ground Motion Prediction Equations for other IMs that are evaluated at point locations (e.g., Peak Ground Acceleration).

With the exception of online questionnaire-based macroseismic intensities (Wald et al., 2011), macroseismic scales are discrete rather than continuous – fractional values have no meaning, and often Roman numerals are used to reflect this. They are monotonic (in the sense that VII generally relates to a stronger ground shaking than VI, for example), but nonlinear (each increment does not necessarily represent a constant increase in ground shaking). Ground motion intensity conversion equations (GMICE) indicate log-linear, often bilinear progressions of peak ground motions versus intensity. Fragility or vulnerability relationships that were consistent with these properties would be step-wise functions, and the mean estimates of damage for each intensity value would be constrained to be monotonically increasing. However, this is not usually done in practice, and a continuous functional form is fitted to the data. To a certain extent, the nonlinearity in the intensity measure can then be accommodated by judicious selection of the functional form – noting that a functional form that fits MMI data well, for example, may not be appropriate for other macroseismic intensity measures (e.g., MSK). The artificial continuousness is also carried through the overall risk calculation, as prediction equations also give continuous intensity estimates, rather than being constrained to discrete values. However, this is unlikely to introduce significant inaccuracy in damage or loss estimates.

Instrumental intensity measures are based on quantities that have been (or could be) calculated from strong ground motion recordings. The most commonly used instrumental measure in the vulnerability and fragility literature is peak ground acceleration (PGA), with peak ground velocity (PGV) also being used in several relationships. Compared with macroseismic scales, instrumental measures may be less correlated with damage, and regressed fragility relationships could be expected to have wider prediction intervals reflecting this increased variability.

Derived quantities such as spectral acceleration, spectral displacement, or Arias Intensity (Arias, 1970) have also been used in the literature, and may be classified as instrumental measures. Spectral ordinates are used to attempt to include the effects of frequency content of the seismic excitation. The spectral ordinate is evaluated at a period that is considered to be most relevant for damage in the building class – generally an estimate of the mean elastic fundamental period. In practice, building vibration periods are not available for damage survey data, or building inventories for loss estimation, and therefore an empirical relationship based on the height of the building and structural system is used. Damping is usually assumed to be 5%. Furthermore, a building class is likely to contain structures with different heights and configurations, and hence a non-constant fundamental period applies across the building class. The use of an empirical relationship for estimating the period of the buildings, the variation of structural periods in a building class, and the assumed damping, all introduce additional epistemic and aleatory uncertainties. These are not generally considered in empirical vulnerability and fragility regression analysis. It is also worth noting that the residual on the empirical period relationship (i.e., the amount by which the relationship over- or underestimates the real period) is potentially correlated with the fragility of the building. For example, a 10-m tall building that has a shorter period than average for its height may be stiffer due to a larger area of structural walls per floor than average, and therefore may also be less likely to be damaged in earthquake shaking of a given intensity. Again, this uncertainty and its influence on fragility estimation is not typically considered in loss assessment. If the uncertainty in structure period prediction is accounted for together with the uncertainty in the ground motion prediction equation, the perceived benefit in using a spectral ordinate rather than a peak ground motion value as IM (i.e., as it accounts for more characteristics of the earthquake), might be removed due to the larger resulting prediction intervals in the vulnerability and fragility relationships. However, it is not possible to determine this a priori.

When spectral ordinates are used, there is a choice between spectral acceleration and displacement (also velocity, although this is not used in any of the reviewed literature). Spectral acceleration has traditionally been used, presumably due to its role in the estimation of forces for seismic design. More recently there has been a movement towards the use of spectral displacement, based on the fact that damage is known to correlate better with displacement than with force. In any case, spectral displacements and accelerations are related by the building period squared; therefore for buildings of the same height, there is no additional predictive power introduced by switching from spectral acceleration to displacement. Amongst buildings with different heights (and periods), there is no strong argument that higher spectral displacements should be correlated with higher damage – the problem is that displacements are not necessarily directly linked to deformations, which are much more directly related to damage. For example, a two-storey moment frame building with 100 mm of roof displacement (not identical to spectral displacement, but neglect this difference for the moment) may suffer similar damage to a one-storey building with the same roof displacement if it deforms in a soft storey (column sway) mechanism in which all the ductility demand is concentrated over a single floor. On the other hand, if a beam sway mechanism develops (which will ideally be the case for capacity-designed frames), the damage would be similar to a one-storey frame with around 50 mm of roof displacement, as the ductility demand is distributed in the beams at both levels. Spectral displacement (or spectral acceleration) alone cannot distinguish between these different mechanisms, and therefore we rely upon the building classification to group together buildings that are likely to be damaged or fail in a similar manner – e.g., older, pre-capacity design, buildings are more likely to exhibit soft storey failures. If we also use building classes with a small height range and therefore relatively small variation in estimated periods, then there is little difference between the predictive power of different spectral measures.

Some instrumental measures, such as strong ground motion duration, may not be efficient intensity measures on their own, but may be a relevant second parameter to consider, along with some measure of peak motion or spectral response, in a vector intensity measure. Ground motion duration is especially important for degrading structures, such as unreinforced masonry (Bommer et al., 2004). Vector IMs would require significant architectural changes to GEM's OpenQuake-engine framework (Pagani et al., 2014; Silva et al., 2013), and have not been considered further here.

4.2 Selection of IMs for Construction of Empirical Vulnerability Curves

The choice of IMs for use with an empirical dataset depends on the location of the damage or loss survey and availability of locally applicable GMPEs. These guidelines recommend that a number of parameters for IM be tried, with empirical vulnerability and fragility functions constructed for each.

In general terms, due to the absence of a single macro-seismic intensity scale that is adequate for all locations across the world, instrumental measures are preferred, with PGA, PGV, $Sa(T1)_{5\%}$ and $Sd(T1)_{5\%}$ suggested. However, this is not prescriptive, and further instrumental IMs or macro-seismic intensities can be used. It is noted that within the context of seismic risk assessment in GEM, the final choices for IM measures may be influenced by GEM guidance for hazard assessment at a particular location (e.g., the availability of ShakeMaps (Wald et al., 1999b), recordings or reliable GMPEs for different parameters at the assessed site).

4.3 Evaluation of IM Values

This section focuses on estimating intensity measure values associated with damage survey data for the development of vulnerability and fragility relationships. In this section, reference is made to GEM documents, in particular the OpenQuake-engine (Pagani et al., 2014; Silva et al., 2013) and the GEM GMPE pre-selection guide (Douglas et al., 2013). Conversions between measures can be used, with relationships available in the literature (e.g., Trifunac and Brandy, 1975; Wald et al., 1999; Faenza and Michelini, 2010), but these conversions introduce a significant amount of uncertainty into the loss estimates (whether accounted for or not), and should be avoided if possible. At the time of writing this document, GEM's OpenQuake-engine and platform are still under development, and therefore the processes here reported for hazard evaluation in GEM may differ from the final procedures.

Within the context of the empirical vulnerability guidelines, once a candidate set of IM measures has been chosen, values of the IM must be evaluated at each survey site. The method for IM evaluation depends strongly on the amount of information available on the earthquake event causing the damage, its source, survey location site, building characteristics and the availability of local recordings of the ground motion.

Assigning a macroseismic intensity value to each set of damage survey data is relatively straightforward, as it may be assigned directly based on damage observed in the survey. However, this introduces a direct interdependence between the axes of the derived fragility curves. It is therefore preferred that the macroseismic intensity be assigned to sites from field observations that also account for the other factors influencing macroseismic intensity (for example, human response and geotechnical failures). Of course, there is subjectivity in the assignment of macroseismic intensity values, and therefore uncertainty about the "true" value. Ideally, this uncertainty could be estimated by having multiple engineers assign intensity values independently.

Instrumental measures (e.g., PGA, PGV, $Sa(T1)_{5\%}$ and $Sd(T1)_{5\%}$) can be evaluated directly from ground motion recordings at post-earthquake survey sites. Ideally, damage survey data would be collected in the vicinity of an accelerometer that recorded the earthquake ground motion (e.g., Liel and Lynch, 2009), but in most cases, this would impose a large restriction on the quantity of data available, as accelerometers are not common (especially in developing countries). In any case, even over a distance of hundreds of metres, local soil conditions can lead to large variations in ground shaking values. Therefore, until we have accelerometers in the basement of every building a degree of uncertainty on our estimates of instrumental values is inevitable.

In the more general case where few or no recordings, local macroseismic intensity observations or a ShakeMap are available in the area of study, ground motion prediction equations (GMPEs) can be used to estimate the intensity measure value at surveyed sites from knowledge of the earthquake characteristics, site conditions and the distance of the site from the earthquake source. Uncertainty in the IM values is introduced when GMPEs are used rather than on-site recordings/observations, and these are discussed further in §4.4.1-§4.4.2.

An alternative estimate of the spatial distribution of ground shaking intensity for historical events can be obtained from the USGS ShakeMaps system (http://earthquake.usgs.gov/shakemap/). The GEM Earthquake Consequences Database (Pomonis and So, 2012) will provide IMs standard for the USGS ShakeMap system (PGA, PGV, Sa(T), and MMI) for nearly one-hundred earthquakes with reported loss data. The ShakeMap Atlas (Allen et al., 2008) and its update (Garcia et al., 2012) provide ShakeMaps for hundreds of additional earthquakes with archived loss data. The advantage of using the gridded ShakeMap IM values is that they are intelligently interpolated using all available source, site, macroseismic and ground motion data and employ

GMPEs bias-corrected to remove the inter-event aleatory uncertainty as best possible (Allen et al., 2008). ShakeMap values also provide uncertainty estimates at each point (Worden et al., 2010). This spatial distribution of both ground shaking level and uncertainty could be used to correlate observed levels of damage in the event with ground motion intensity.

4.4 Sources of Uncertainty in IM

These guidelines provide a possible procedure for incorporating uncertainty in IM in empirical vulnerability and fragility functions. This procedure is presented in §6, and assumes the uncertainty is quantifiable and represented by a variance value. The following sections provide a discussion of the sources of uncertainty that may contribute to this variance value.

The discussion provides some pointers on what to consider when quantifying these uncertainties and advice on potential ways to reduce them. Full guidance on IM uncertainty evaluation is not provided and falls outside the scope of these Guidelines. However, it is noted that IM uncertainty evaluation should be carried out with a full appreciation of how the hazard is calculated in the overall risk analysis, so as to avoid the double counting of sources of uncertainty.

4.4.1 Uncertainty due to choice of GMPE

Complete guidance on selecting GMPEs for different regions around the world is provided by GEM (Stewart et al., 2013). However, if dealing with empirical data from past earthquakes, and some recordings of ground motion were available, a further criterion for the selection of GMPE would be the inclusion of these recordings in the equation.

Stewart et al. (2013b) suggests that it is good practice to choose and combine multiple GMPEs to estimate the IM values. This can be done through the use of a logic tree approach. Such an approach can quantify the epistemic uncertainty in the GMPE choice. Guidance for this is provided in Devalaud et al. (2012). It is noted that the OpenQuake-engine (Pagani et al., 2014; Silva et al., 2013) provides a scenario-based earthquake hazard based on a deterministic event that can be used to calculate IM values at sites of interest by inputting the earthquake parameters. This tool has the functionality to allow multiple GMPEs to be used and also allows the source characteristics to be varied, if these have not been appropriately constrained (e.g., for earthquakes in the significant past).

Where some recordings are available, a method for constraining inter-event variability is to adjust the IM estimates from the GMPEs (e.g., by selecting an appropriate value of standard deviation from the median prediction) so as to provide a good fit to the recorded ground motion values, see 4.4.2.

4.4.2 Uncertainty due to measurement error in GMPE

The measurement error of IM estimates from GMPEs is significant. The standard deviation on ground motion prediction equations, when calculated on the natural logarithm of an intensity measure, is generally of the order of 0.5–0.7, meaning that one standard deviation on either side of the median can represent a range of half to double of the median estimate. This uncertainty is not reduced when multiple GMPEs are combined.

Although there is significant uncertainty in estimates of instrumental intensity made in this way, there are many additional sources of information that could be used to reduce the uncertainty, and potentially adjust median estimates of intensity. A likely source is when ground motion recordings are available from the earthquake in which damage data were collected. Even if these were recorded some distance from the site of interest, the measurements give an indication of whether the event was typical of an earthquake with the same magnitude, distance and source characteristics, or whether it produced higher or lower levels of ground shaking. Most modern ground motion prediction equations separate the standard deviation (σ_T) into "inter-event" (τ) and "intra-event" (σ) components (Joyner and Boore, 1993):

$$\sigma_T^2 = \tau^2 + \sigma^2 \tag{4.1}$$

The first term, the inter-event term, represents the event-to-event variation, while the second term, the intra-event, represents the variation in intensities measured from a single event (record-to-record variability).

The inter-event standard deviation is typically of the order of three times smaller than the intra-event term, and therefore the benefit from removing τ from Eq. (4.1) may be small. For example, if $\tau = 0.2$ and $\sigma = 0.6$, then $\sigma_T = 0.63$, so very minimal benefit is obtained by fixing the number of τ. That said, Jayaram and Baker (2010) recently showed that the inter-event term is underestimated by the traditional approaches used for the development of ground motion prediction models, and therefore there may be some added benefit from removing it from Eq.(4.1). ShakeMaps do remove the inter-event term where it is knowable (Worden et al., 2010).

If we have recordings reasonably near our site of interest, we could reduce the total uncertainty further by taking into account spatial correlation of ground motions. For example, if a single recording of an event is available a few kilometres from where the damage survey data was collected, the spatial correlation of the residuals can be taken into account. If this recording is higher than expected (positive residual with respect to the median of the GMPE), then it is likely that the site where damage data were collected also experienced greater-than-expected shaking. The probability distribution for the intensity at the site conditioned on the relatively nearby recording has a higher mean and is narrower (lower standard deviation) than the *a posteriori* distribution for the intensity without any recordings. When more than one recording is available, this distribution can be refined further. These considerations, and the resulting spatially-varying effect on uncertainties, are explicitly included in the ShakeMap methodology (Worden et al., 2010).

Spatial correlation has been taken into account in risk assessment calculations, but has not currently been considered in the development of fragility or vulnerability relationships, except coarsely in the collection of survey data only within a specified radius (say 100 feet of a recording station, King et al., 2005). The latter essentially assumes that ground motion intensity levels are identical and fully correlated within this radius, and sufficiently uncorrelated outside the radius that the damage data are not worth collecting. This is of course an approximation, and also provides a basis for randomising the locations of survey areas (assuming accelerometers were placed somewhat randomly, and not, for example, because high or low ground shaking was expected in that location). Currently this remains an area for further research.

4.4.3 Uncertainties from spatial variation across geographical survey units

Many existing databases of observed damage and loss present the data in aggregated form, the aggregation being by geographical area. The survey units can be of varied size (e.g., ZIP-Codes, communes, villages, towns, cities) but invariably are assigned a single value of IM in existing vulnerability and fragility studies (see Rossetto et al., 2013). Unless the survey unit area is very small, there will be a variation in the IM values affecting buildings within it. This variation derives from a number of factors, which can include differences in distance and site conditions across the area. The variation is more likely to be larger for large survey units, and can vary across datasets from single surveys if the size of survey units differs (e.g., ZIP-Codes may have different areas).

This source of uncertainty can be evaluated by treating each survey unit as an area, rather than a point, when determining the IM values as per §4.3. Considering variations in earthquake source distance and site conditions across the area, it is thus possible to obtain a mean estimate and variance for the mean *IM* associated to each survey unit. It should be noted that this source of uncertainty is in addition to those highlighted in 4.4.1 and 4.4.2. Reference is made to Stafford (2012) where an approach for reducing the variance of the ground motion (as compared to the value given by the GMPE) is proposed, when the latter is modelled over an area, rather than a single point.

In the case of high-quality building-by-building damage or loss data, the intensity measure levels are evaluated at the location of each individual building observation. This type of empirical data is therefore not affected by this additional source of uncertainty.

4.4.4 Uncertainty in estimation of spectral values for building classes

An additional source of uncertainty arises (especially for aggregated damage or loss data) if a spectral ordinate is used as the intensity measure type. In existing empirical fragility or loss functions, the spectral ordinate is commonly evaluated at a structural period thought to be representative of the natural period of a typical structure for the building class being assessed. However, a building class contains a variation in building configurations, heights, seismic design and materials that may not be known in detail. This is especially true when using data from past surveys for which the building classification is coarse and the original survey forms are unavailable. Consequently any building class will contain a variation in structural periods, which will result in a variation in the spectral ordinates for any dataset.

This is an area that has not been researched in the past, and hence no guidance can be given as to what type of distribution in structural periods is typical of different building class definitions.

In the case of analytical vulnerability or fragility functions, as the structural models are known in detail, their natural period can be evaluated and variation across the building class (set of known structural models) included in their performance evaluation (e.g., Rossetto and Elnashai, 2005). Hence, this source of uncertainty may not be an issue. In the case of empirical fragility functions , this source of uncertainty can be reduced if building-by-building data are available, obtained from detailed on-the-ground surveys that have captured the structural properties of each building in detail. It can also be reduced through the use of detailed building classes.However, it is still likely to remain an issue.

5 Introduction to Statistical Models and Model Fitting Techniques

5.1 Introduction

In general, statistical-model fitting techniques aim to determine the probability distribution characteristics of a response variable Y, i.e., damage or loss, conditioned on one or more explanatory variables $\mathbf{X}=[X_1, X_2,\ldots, X_{N_x}]$, e.g., ground motion intensity. They do this by fitting a statistical model to the empirical data. Hence, the proposed procedure for constructing empirical fragility and vulnerability functions from data points essentially includes three main steps:

1. Selection of a statistical model
2. Selection of a model fitting technique
3. Assessment of the model's goodness of fit

The Guidelines suggest a series of statistical models and associated model fitting techniques be adopted in the development of empirical fragility and vulnerability relationships, with the simpler techniques (Level 1) attempted first and more complex techniques (Level 2) attempted only if the simpler approaches yield unsatisfactory goodness-of-fit. Goodness-of-fit can be assessed using a series of diagnostic tools described for each statistical model in §6 to §8. Hence, steps 1 to 3 above may be repeated several times.

This section presents an overview of the likely forms that response and explanatory variables can take in the case of empirical data. It then introduces the three main statistical models considered in these Guidelines, which are described in operational detail in §6 to §8. Finally, it provides an overview of which combinations of statistical models and model fitting techniques should be used according to the characteristics of the empirical data adopted.

5.2 Understanding Empirical Data in terms of Response and Explanatory Variables

Before embarking on the construction of empirical fragility and vulnerability functions, it is important to understand the characteristics of the observational survey data being used and the resulting form of the predictor and response variables that will define the data points $(x_{j1}, x_{j2}, \ldots, x_{jN_x}, y_j)$ for the statistical-model fitting.

In the case of empirical vulnerability and fragility relationships, the main explanatory variable is the ground motion intensity (*IM*). Depending on the choice of an intensity measure type, the explanatory variable can be modelled either as discrete (e.g., for macroseismic intensities) or continuous (e.g., for instrumental measures). The nature of the intensity measure type does not affect the construction of the statistical model. These guidelines recommend that the analyst try various ground motion intensity measure types with the option to add more explanatory variables, e.g. soil conditions, as part of the diagnostics tools (see §6.3.2 and §7.3.3) in order to examine whether the model can be improved by adding more relevant variables.

Within the following sections, where vulnerability functions for economic loss are to be constructed, economic loss is modelled as a continuous non-negative variable, which can be positively skewed (see Wesson et al., 2004). In this case, data points (x_j,y_j) are introduced in 'R' as presented in Figure 5.1.

```
### Loss data
> Dat<-data.frame(loss=c(1000,50000,100000),IM=c(0.1,0.4,0.5))
> Dat     # example of a database of economic loss in $ and PGA in g.
    loss    IM
1 1e+03    0.1
2 5e+04    0.4
```

Figure 5.1 Form of damage data points in 'R'.

For vulnerability functions for human loss or fragility functions are to be constructed, human loss or damage are typically recorded in the survey forms in terms of a discrete variable that increases in intensity, e.g., No damage (ds_0), Slight damage (ds_1), Moderate damage (ds_2), Severe Damage (ds_3) and Collapse (ds_4). Hence, the proposed statistical model fitting techniques in §6-§8 accommodate the modelling of response in terms of a binary variable. For example, the binary response variable, *Y*, for building-by-building damage data is expressed in terms of:

$$Y = \begin{cases} 1 & DS \geq ds_i \\ 0 & DS < ds_i \end{cases} \tag{5.1}$$

In this case, the data points (x_j,y_j) express the ground motion intensity levels, that affect a structural unit *j* and its binary response y_j, which equals 1 if the building has suffered damage equal to or exceeding ds_i and 0 otherwise (see Figure 5.2).

If grouped data are available, the data points (x_j,(y_j,n_j-y_j)) contain y_j counts of buildings with $DS \geq ds_i$ and n_j-y_j counts of buildings with $DS<ds_i$ for bin *j* with intensity measure level x_j (see Figure 5.2).

```
### Form of building-by-building damage data assuming Y is expressed by Eq.(5.1).
Dat<-data.frame(Y=c(0,1,1),Damage=c(0,3,4),IM=c(0.1,0.4,0.5))  # Damage refers to the damage state dsi=0-4,
                                                                  #  IM is in terms of  PGA in g,
                                                                  # Y is the indicator obtained by Eq(5.1) for e.g. ds1.
> Dat
   Y Damage    IM
1  0       0    0.1
2  1       3    0.4
3  1       5    0.5
…

### Form of grouped damage data.
> Dat<-data.frame(DSi=c(20,30,90),noDSi=c(80,70,10),IM=c(0.1,0.4,0.5))  # assuming 100 buildings in each bin
> Dat
  DSi noDSi  IM
```

Figure 5.2 Form of damage data points in 'R'.

5.3 What are Statistical Models and which should I use?

The data points based on the damage or loss databases are used in to construct fragility or direct vulnerability curves by fitting an appropriate statistical model.

A statistical model is a stochastic relationship between the seismic damage or loss and ground motion intensity. Practically, a statistical model consists of two components: random and systematic:

- The random component expresses the probability distribution of the response variable (e.g. the counts of buildings reaching or exceeding a damage state or the economic loss for each building) given the explanatory variable (i.e. the ground motion intensity).
- The systematic component expresses the mean response as a function of the explanatory variable. In fragility assessment, the systematic component typically represents the fragility curve, i.e., the probability that a damage state is reached or exceeded given ground motion intensity. In loss assessment, the systematic component is termed vulnerability curve, e.g., the mean loss given intensity measure levels.

Three types of statistical model are proposed in these Guidelines for expressing direct vulnerability and fragility relationships:

1. Generalised linear models (GLM). GLM are robust well-established models that depend on a small number of model parameters. Their systematic component is a strictly monotonic function. These models also have strong predictive capacity provided that the assumptions regarding their random and systematic component are met. They have been used for the construction of empirical fragility curves for bridges (e.g. Shinozuka et al., 2000) and steel tanks (e.g. O'Rourke and So, 2000) and buildings (Ioannou et al., 2012).
2. Generalised additive models (GAM). These are novel in the field of empirical fragility and vulnerability assessment. They differ from GLMs in the level of flexibility allowed in the systematic component, i.e., the trend is non-strictly monotonic. Within GAM the systematic component can be used to capture a non-strictly monotonic trend in the data. Like GLM a number of assumptions regarding their two components have to be satisfied for them to result in reliable fragility and vulnerability relationships.
3. Gaussian Kernel Smoothers (GKS). GKSs do not make any assumptions regarding the systematic and random components and hence are the most flexible of the three statistical model types. They can fit the data very closely and like GAM can accommodate a non-strictly monotonic trend in the data. Their use in the field is limited (see Noh et al., 2013).

Guidance for fitting these three models is provided in §6-§8. The reader may refer to the classic books of McCullagh and Nelder (1989) and Hastie and Tibshirani (1990) for a detailed introduction to linear and additive models, respectively. The reader is also referred to the book by Wand and Jones (1995) for a better understanding of Gaussian Kernel Smoothers. The reader is referred to the paper of Ioannou and Rossetto (2015) for a detailed comparison of the three models.

5.4 Which Combination of Statistical Model and Model Fitting Technique should I use?

The present Guidelines account for the varying degree of understanding of statistics by the analysts as well as the objective needs of the data by proposing two Levels of complexity of the statistical model and model-fitting procedures.

As a first step, the analyst is advised to determine building classes and construct empirical vulnerability or fragility curves for these classes according to Level 1 approach using the available data. This step does not distinguish between data obtained from a single event or multiple events, the use of different sample sizes or the presence of measurement errors in the data. Nonetheless, the analyst is advised to plot the cumulative proportion of buildings in each IM bin given the corresponding levels of intensity in order to have an idea of whether the data are clustered or more or less uniformly distributed in the available range of intensity measure levels. The Level 1 approach includes the construction of a GLM model, which relates the damage or loss with an intensity measure type. The GLM is fitted to the damage or loss and intensity data by maximising the likelihood function, *L*(.), as:

$$\boldsymbol{\theta}^{opt} = \arg\max\left[L\left(\boldsymbol{\theta}; y, \mathbf{x}\right)\right] = \arg\max\left[\prod_{j=1}^{M} f\left(y_j \mid x_j, \boldsymbol{\theta}\right)\right] \tag{5.2}$$

where *f*(.) is the probability density distribution of the discrete (or continuous) response variable conditioned on an intensity measure type. This approach assumes that the GLM parameters, **θ**, are fixed but unknown. The analyst should then examine whether the selected systematic and random component capture the trend and the uncertainty in the data satisfactorily or if a better GLM with different assumptions regarding these two components can be identified.

The proposed GLM model is acceptable if the following conditions are met:

(1) If the data is collected from a single event,
(2) negligible measurement error can be assumed in the response and explanatory variables
(3) the proposed GLM is found to fit these data well following a goodness of fit check

The user is warned that for empirical fragility assessment the selected random component is often unable to capture the larger uncertainty in the grouped data. In this case, the user should ensure that the model is modified in order to capture this larger uncertainty.. If the systematic or random component of a GLM do not provide a very good fit and there are more explanatory variables in the database (e.g., soil conditions or subgroups of building characteristics), the user is advised to add them to the systematic component in order to examine whether the fit is improved.

The conditions which require the use of Level 2 procedures are presented below. The user is reminded that Level 2 approaches construct models which capture more sources of uncertainty and for this reason they appear to be more complex. However, this does not mean that there this extra complexity is required for each database. The analyst is advised to critically compare these models to their simpler counterparts constructed by the Level 1 approaches in order to justify the need for this complexity.

Level 2 approaches are necessary if:

- There is a measurement error in the response variable.
- There is a measurement error in the ground motion intensity levels.
- The available sample size is small (e.g. <100 buildings).

For these three conditions, the fitting of a GLM is advised through a Bayesian analysis. This procedure has the advantage of estimating the vulnerability or fragility curves by updating existing curves for similar building classes with the available data. Bayesian analysis estimates the posterior distribution of the model parameters, *f*(**θ**|*y*,*x*), from Bayes' theorem:

$$f\left(\boldsymbol{\theta} \mid y, x\right) = \frac{L\left(\boldsymbol{\theta}; y, x\right) f\left(\boldsymbol{\theta}\right)}{f\left(y\right)} = \frac{L\left(\boldsymbol{\theta}; y, x\right) f\left(\boldsymbol{\theta}\right)}{\int L\left(\boldsymbol{\theta}; y, x\right) f\left(\boldsymbol{\theta}\right) d\theta} \tag{5.3}$$

where *f*(**θ**) is the prior distribution of the model parameters accounting for existing prior knowledge.

Level 2 approaches are also necessary if data collected from multiple events are available. It should be noted that the models proposed here are novel to the field and the analyst is urged to consult existing literature, e.g., Gelman and Hill (2007) or Zuur et al. (2009). The philosophy behind the Level 2 approach proposed here is that data for each event produce a different fragility curve and that the events available in the database are randomly selected from a very large population of events. So GLM models, termed generalised linear mixed models, are used in order to account for this random effect of the earthquake.

Level 2 approaches should also be used if the sample size is large (>100 buildings) and the fragility curves cross (this leads to meaningless results in the construction of the probability matrices). For these cases, the analyst is advised to fit GAM or GKS models.

Finally, GAM or GKS models can be used if the data have a non-strictly monotonic trend, which cannot be captured by the GLM models. However, the use of the non-parametric models in this case requires densely distributed data points in the available IM range (>100 data points) in order to justify that this trend is significant. The analyst can consult the plot of the cumulative proportions of buildings for each IM in order to decide whether the distribution of the data is sufficiently dense.

In what follows, a collection of procedures for statistical model fitting corresponding to different levels of analyst's experience is presented. This includes established 'R' packages which can be used directly for the construction of simple and mixed models and the assessment of their goodness of fit. Guidance for more advanced, i.e., Bayesian, methods is also provided for the construction of complex models.

6 Fitting Generalised Linear Models (GLM)

6.1 Introduction

Generalised linear models (GLM), in the context of vulnerability or fragility assessment, consist of three main components (Nelder and Wedderburn, 1972):

- The random component, $f(Y|IM=im)$: A continuous or discrete probability density function (of the exponential family) that expresses the variability in the response variable given ground motion intensity levels. In the case of economic loss, the conditional probability can be expressed by three, positively skewed, continuous distributions, namely: gamma, inverse Gaussian and lognormal (McCullagh and Nelder, 1989). The three distributions can treat only variables having values greater than zero (Y>0). With regard to discrete damage (or human loss), the expression of the random component depends on the type of the data. For grouped damage data (*low* to *medium quality*), the binomial distribution is selected. The Bernoulli distribution, a special case of the binomial, is considered if building-by-building data (*high quality*) are available.
- The systematic component, which relates the mean response variable to the selected intensity measure type. This component is expressed in terms of the linear predictor, η, in the form:

$$\eta = \vartheta_0 + \vartheta_{1i}\, im \tag{6.1}$$

  where ϑ_0, ϑ_{1i} are the model parameters corresponding to *the damage state* or casualty state i and N is the total number of explanatory variables.
- The link function: This is a monotonic and differentiable parametric function that relates the mean response variable with the linear predictor.

$$g\left(E\left[Y\,|\,X\right]\right) = g\left(\mu\right) = \eta \tag{6.2}$$

  where $g(.)$ is the link function, such as the probit function mostly used for the expression of fragility curves.

A detailed procedure for how to select the three GLM components and then estimate the fragility or vulnerability function parameters is provided in this Section. Within this context, the maximum likelihood and Bayesian statistical-model fitting techniques are considered and appropriate diagnostic tools are provided to assess the goodness-of-fit of the fitted model to the empirical data.

6.2 Construction of a Parametric GLM (Level 1)

Parametric generalised linear models have the generic form:

$$\begin{aligned} Y \sim f\left(y\,|\,im\right) \qquad & \text{with } E\left[Y\,|\,im\right] = \mu = g^{-1}\left(\eta\right) \\ & \operatorname{var}\left[Y\,|\,im\right] = \varphi \operatorname{var}\left(\mu\right) \end{aligned} \tag{6.3}$$

where $f(\mathrm{y}|im)$ is a probability density function (pdf) of the response variables given one or more explanatory variables; φ *is the* dispersion parameter associated with the variance of the pdf; $\operatorname{var}(\mu)$ is the variance function of μ. From Eq.(6.3), three main elements of the GLMs can be identified, namely:

- the linear predictor.
- the probability density function of the response variable conditioned on the explanatory variables.
- the link function that relates the mean response to the linear predictor.

6.2.1 Determination of the linear predictor

The simplest linear predictor, η, can be constructed to have the form:

$$\eta = \vartheta_0 + \vartheta_{1i}\, im \tag{6.4}$$

The diagnostics, presented in §6.3.2, determine whether this model fits the empirical loss or damage data satisfactorily. If the model does not fit the data then the intensity measure is transformed, typically into its natural logarithm (i.e., log(IM)) or in some cases a higher order polynomial (e.g., IM^2). Failure of these transformations to improve the fit of the parametric model might indicate that the nonlinear structure in the data cannot be depicted by a global parametric model. Therefore, a nonparametric model should be constructed (see §7.2).

6.2.2 Determination of a probability density function for the response variable

Table 6.1 summarises a set of continuous and discrete probability density functions (pdfs) that can be used to represent loss and damage, and shows the mean response variable as a function of the linear predictor and its variance. A brief description of these distributions and guidance on their selection is presented here.

Table 6.1 Characteristics of the continuous probability density distribution functions of the response variable (McCullagh and Nelder, 1989).

	pdf	$f(y\mid im)=$	Parameters	$\mu=$	$\mathrm{var}[Y\mid im]=$	$\varphi=$
Economic loss	Inverse Gaussian	$\sqrt{\left(\frac{\sigma^2}{2\pi y^3}\right)}\exp\left(-\frac{\sigma^2(y-\mu)^2}{2\mu^2 y}\right)$	$\mu>0$	$g^{-1}(\eta)$	$\varphi\mu^3$	σ^2
	Gamma	$\frac{\exp\left(-\frac{y}{b}\right)}{b\Gamma(c)}\left(\frac{y}{b}\right)^{c-1}$	$b>0, c>0$		$\varphi\mu^2$	$\frac{1}{b}$
	Normal	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right)$	$\mu,\sigma>0$		φ	σ^2
Damage or Human Loss	Bernoulli	$\mu^{y_j}\left[1-\mu\right]^{1-y_j}$	$\mu\in[0,1]$		$\varphi\mu(1-\mu)$	1
	Binomial	$\binom{n_j}{y_j}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j}$				

The variability of economic loss for a given level of the linear predictor can be expressed by three, positively skewed, continuous distributions, namely: gamma, inverse Gaussian and lognormal (see Table 6.1). In the literature, Wesson et al., (2004) used the gamma distribution in the construction of their vulnerability functions for economic loss. Nevertheless, analysts are advised to follow their intuition and use the diagnostics tools in §6.3.2 to determine whether their selected distribution is adequate. It is noted that if the loss is expressed in terms of repair over replacement ratio, which is bounded in [0,1], then a beta regression might be used. This model is new in the field and it is not covered in these Guidelines. The analyst is referred to the package 'betareg' in R (2013) for more information.

Table 6.1 shows that the mean loss, expressed as a function of the linear predictor, is identical for the gamma, inverse Gaussian and lognormal distributions. Differences are noted in the variance of loss. In particular, the variance is proportional to the squared or cubed mean if a gamma or inverse Gaussian distribution is selected, respectively. By contrast, the variance of the natural logarithm of loss is assumed constant for all levels of the linear predictor. Both the gamma and lognormal distributions assume that the coefficient of variation for the various levels of the linear predictor is constant.

The selection of discrete probability density functions for damage or human loss is determined by the type of empirical data used:

- The Bernoulli distribution, a special case of the binomial, should be adopted if building-by-building data (*high quality*) are available.
- The binomial distribution should be adopted if grouped data are available (*low to medium quality*).

It should be noted that the mean of both distributions is essentially the probability that a level of damage is reached or exceeded given a level of ground motion intensity, and that the dispersion parameter is unity, $\varphi = 1$.

6.2.3 Selection of a link function

A link function is a strictly monotonic and differentiable function that relates the mean response variable with the linear predictor. The analyst is advised to examine the fit of all appropriate link functions presented in Table 6.2 and to follow the diagnostic procedures presented in §6.3.2 in order to identify which link function results in the best fit of the statistical model to the empirical data.

Table 6.2 Appropriate link functions for each distribution of the exponential family.

pdf	Link function, η =						
	identity	inverse	log	'1/mu^2'	logit	probit	complementary log-log
	μ	$1/\mu$	$\log(\mu)$	$1/\mu^2$	$\log[1/(\mu^{-1}-1)]$	$\Phi^{-1}(\mu)$	$-\log[-\log(1-\mu)]$
Inverse Gaussian	°	°	°	°			
Gamma	°	°	°				
Lognormal	°	°	°				
Bernoulli/Binomial			°		°	°	°

6.3 Fitting a Parametric GLM: The Maximum Likelihood Method (Level 1)

6.3.1 Estimation of the GLM parameters

The parameters of the GLM's systematic component (e.g., ϑ_0 and ϑ_1) are considered to have a given true value which is unknown. Their values are estimated by maximising the likelihood function (see Eq.(5.2)). The estimation is carried out numerically through an iterative reweighted least squares algorithm (Nelder and Wedderburn, 1972).

The summary of the outcomes of the parametric statistical fitting method in 'R' includes the asymptotic (i.e., approximate) standard errors of the model parameters. The levels of these errors depend on the number of available observations. This means that they can be reduced if more observations become available. The effect of the available number of observations on the confidence of the systematic component should be assessed as presented in Figure 6.1. These intervals are estimated by considering that the values of the linear predictor, η, are normally distributed for each level of intensity. Therefore, the e.g., 90% confidence interval of a mean curve can be approximately estimated by adding ± 2* standard error to the mean fitted linear predictor, η. The required intervals are then estimated by transforming the aforementioned values of the linear predictor into the vulnerability or fragility curves through the link function. The smaller the number of existing observations leads to wider constructed confidence intervals. Nonetheless, these intervals are approximate and assume that the parameters as well as the linear predictor are normally distributed. Appendices B-D illustrate the shape of these intervals for fragility as well as direct vulnerability curves.

If the analyst desires an alternative, numerical, evaluation of the width of the confidence intervals, the bootstrap technique (Chandler and Scott, 2013, pp 117-118) should be adopted (see Figure 6.2). This technique is based on multiple realisations of the following 4 steps:

1. A GLM is determined by selecting a pdf from Table 6.1, a link function from Table 6.2 and the linear predictor from Eq.(6.4). The model's parameters (i.e. ϑ_0, ϑ_1) are estimated.
2. New values of the response Y are generated for the available intensity measure levels. To do this, the response for the given intensity measure levels follow the fitted GLM.
3. A GLM with the same pdf and link function is fitted to the generated data points and the procedure is repeated a large number of times (e.g., 1000 times).
4. The obtained response values for each intensity measure level are then ranked, and the specified prediction intervals are obtained.

With regard to the empirical fragility assessment, the binomial distribution captures uncertainty significantly smaller than the variability in the grouped data. This over-dispersion can be addressed by the use of the quasibinomial distribution (see McCullagh and Nelder, 1989). The use of this distribution leads to identical mean fragility curves with the binomial distribution but the confidence intevrals are in this case wider, following closer the uncertainty in the data points.

Alternatively, the user can use a bootstrap analysis (Chandler and Scott, 2013, pp 117-118) to construct the confidence intervals (also see Figure 6.3). This technique is very similar to the one mentioned above with the only difference being the determination of the new data points:

1. A GLM is determined by selecting a pdf from Table 6.1, a link function from Table 6.2 and the linear predictor from Eq.(6.4). The model's parameters (i.e. ϑ_0, ϑ_1) are estimated.
2. New values of the response Y are generated for the available data points by sampling them with replacement. This means that each data point can be repeated from no times to multiple times for every iteration.
3. A GLM with the same pdf and link function is fitted to the generated data points and the procedure is repeated a large number of times (e.g., 1000 times).
4. The obtained response values for each intensity measure level are then ranked, and the specified prediction intervals are obtained.

The user is advised to use the second bootstrap technique, based on the sampling from the raw data with replacement in order to appropriately capture the uncertainty in them.

```
### Generalised linear model for loss data (Y>0).
fit<-glm(Loss~IM, family=Gamma(link=c('identity','inverse','log'))
fit<-glm(Loss~IM, family= inverse.gaussian (link=c('identity','inverse','log','1/mu^2'))
fit<-glm(log(Loss)~IM, family=gaussian(link=c('identity','inverse','log'))

### Generalised linear model assuming that the building-specific damage data follow a Bernoulli distribution.
fit<-glm(Y~IM, family=binomial(c('logit', 'probit', 'log', 'cloglog'),data=Data)

### Generalised linear model assuming that the grouped damage data follow a binomial distribution.
fit<-glm(D~IM, family=binomial(c('logit', 'probit', 'log', 'cloglog'))

fit<-glm(D~IM, family=quasibinomial(c('logit', 'probit', 'log', 'cloglog'))

summary(fit)   # provides a summary of the outcomes of the analysis, including the values of the regression
               parameters and the dispersion.
## Construction of approximate confidence intervals of the systematic component assuming a binomial distribution
with a logit link ##function.
fit.pred<- predict(fit,type='link', se.fit=TRUE)     # the standard error and the mean values of the linear predictor is
provided
f.upper<-1/(exp(fit.pred$fitted.values-2*fit.pred$se.fit)+1)  # the 90% confidence interval.
f.lower<-1/(exp(fit.pred$fitted.values+2*fit.pred$se.fit)+1)  # the  5% confidence interval.
```

Figure 6.1 Generalised linear models for correlating the loss or damage with an intensity measure (See appendices B-D for an illustration of these intervals).

```
####### Bootstrap grouped damage data, ignoring the over-dispersion.
DATA<-data.frame(IM=IM.observed,DS=Collapse.Observed,Tot=TotalBuildingsinBins)
N<-1000 # number of iterations
Nim<-100 # the number of new intensity measure levels for which the confidence intervals are estimated.
newdata<-data.frame(IM=seq(min(DATA$IM,max(DATA$IM),length=Nim))
fit<-glm(DS~log(IM),family=binomial('probit'),data=DATA)
fit.pred<-predict(fit,type="response")
frg<-array(NA,c(N,Nim))
ord.FC<- array(NA,c(N,Nim))
mean.FC<- array(NA, Nim)
for(i in 1:N){
  newDATA<- data.frame(IM=IM.observed,DS= rbinom(Nim, size=DATA$Total,
prob=fit.pred),Tot=TotalBuildingsinBins)
 # sample from the binomial distribution.
  fitBoot <- glm(cbind(DS,Total-DS)~log(IM), family=binomial("probit"),data=newDATA)
  frg[i,]<-predict(fitBoot,newdata=newdata,type="response")
}
for (j in 1:Nim){
  ord.FC[,j]<-sort( frg[,j])
}
for (j in 1:Nim){
  mean.FC[j]<-mean(ord.FC[,j])
}
Up<-N*95/100; Lw<-N*5 /100;  # 95% and 5% confidence intervals
Up.FC<- ord.FC[Up,] # Fragility curves for the 95%.
Lw.FC<- ord.FC[Lw,] # Fragility curves for the 5%.
```

Figure 6.2 Bootstrap technique resampling from the GLM model.

```
####### Bootstrap accounting for over-dispersion for grouped damage data.
DATA<-data.frame(IM=IM.observed,DS=Collapse.Observed,Tot=TotalBuildingsinBins)
N<-1000 # number of iterations
Nim<-100 # the number of new intensity measure levels for which the confidence intervals are estimated.
newdata<-data.frame(IM=seq(min(DATA$IM,max(DATA$IM),length=Nim))
fit<-glm(DS~log(IM),family=binomial('probit'),data=DATA)
frg<-array(NA,c(N,Nim))
ord.FC<- array(NA,c(N,Nim))
mean.FC<- array(NA, Nim)
for(i in 1:N){
  newDATA<-DATA[sample(length(DATA[,1]),replace=T),]  # sample with replacement
  fitBoot <- glm(cbind(DS,Total-DS)~log(IM), family=binomial("probit"),data=newDATA)
  frg[i,]<-predict(fitBoot,newdata=newdata,type="response")
}
for (j in 1:Nim){
  ord.FC[,j]<-sort( frg[,j])
}
for (j in 1:Nim){
  mean.FC[j]<-mean(ord.FC[,j])
}
Up<-N*95/100; Lw<-N*5 /100;  # 95% and 5% confidence intervals
Up.FC<- ord.FC[Up,] # Fragility curves for the 95%.
Lw.FC<- ord.FC[Lw,] # Fragility curves for the 5%.
```

Figure 6.3 Bootstrap technique resampling from the raw data.

6.3.2 Diagnostics – checking the goodness of fit

Having estimated the unknown parameters for the selected model (see Figure 6.1), the goodness of fit of this model should be assessed. A number of significance tests and graphical assessment procedures are available and could be used. However, the present guidelines recommend the use of a few graphical procedures, which assess the adequacy of the mean and variance function to describe the data, and highlight ways in which the model can be improved.

The proposed diagnostic tools which assess the adequacy of the mean and variance function are based on the study of Pearson residuals, which are equivalent to the raw residuals of the least squares regression analysis:

$$r_{Pj} = \frac{y_j - \mu}{\sqrt{\mathrm{var}(\mu)/w_j}} \tag{6.5}$$

where r_{Pj} is the Pearson residual for the data point j; w_j is a weight for the continuous distributions of the response variable, which w_j =1 if the distribution of the response variable is continuous and w_j = n_j (n_j is the number of buildings in bin j) when this distribution is binomial.

Standardised Pearson residuals r_{Psj} are calculated as per Eq.(6.6) (see Figure 6.3 for the 'R' code used for this calculation) and are adopted in the diagnostics presented in the following sections.

$$r_{Psj} = \frac{r_{Pj}}{\sqrt{\varphi\left(h_j - 1\right)}} \quad (6.6)$$

where h_j is the leverage, i.e. the diagonal value j of the hat matrix, this value indicates to what extent the predicted value for an observation is determined by the observed value for that observation. The standardised Pearson residuals follow a normal distribution with mean equal to 0 and variance equal toφ. For the damage categorical data which are typically assumed to follow a binomial distribution, standardised Pearson residuals follow a normal distribution with mean equal to 0 and variance equal to 1.

The procedures proposed for the assessment of the goodness of fit of the constructed generalised linear models are described in §6.3.2.1-§6.3.2.2. Figure 6.5 presents the procedure in terms of a flowchart. This rather mechanistic approach adopted for the description of the diagnostics is meant to highlight commonly adopted strategies that can be followed in order to construct a model that represents the available data. The suggested tools are not exhaustive, and the needs of some databases may require the tools to be applied in a different order than the one presented in Figure 6.5.

```
### Pearson residuals.
res<-residuals(fit,type='pearson')        # Pearson residuals.
#and
res<-glm.diag$rp                        # Standardised Pearson residuals.
#alternatively
res<-rstandard(fit)
```

Figure 6.4 Calculating the Pearson residuals in 'R'.

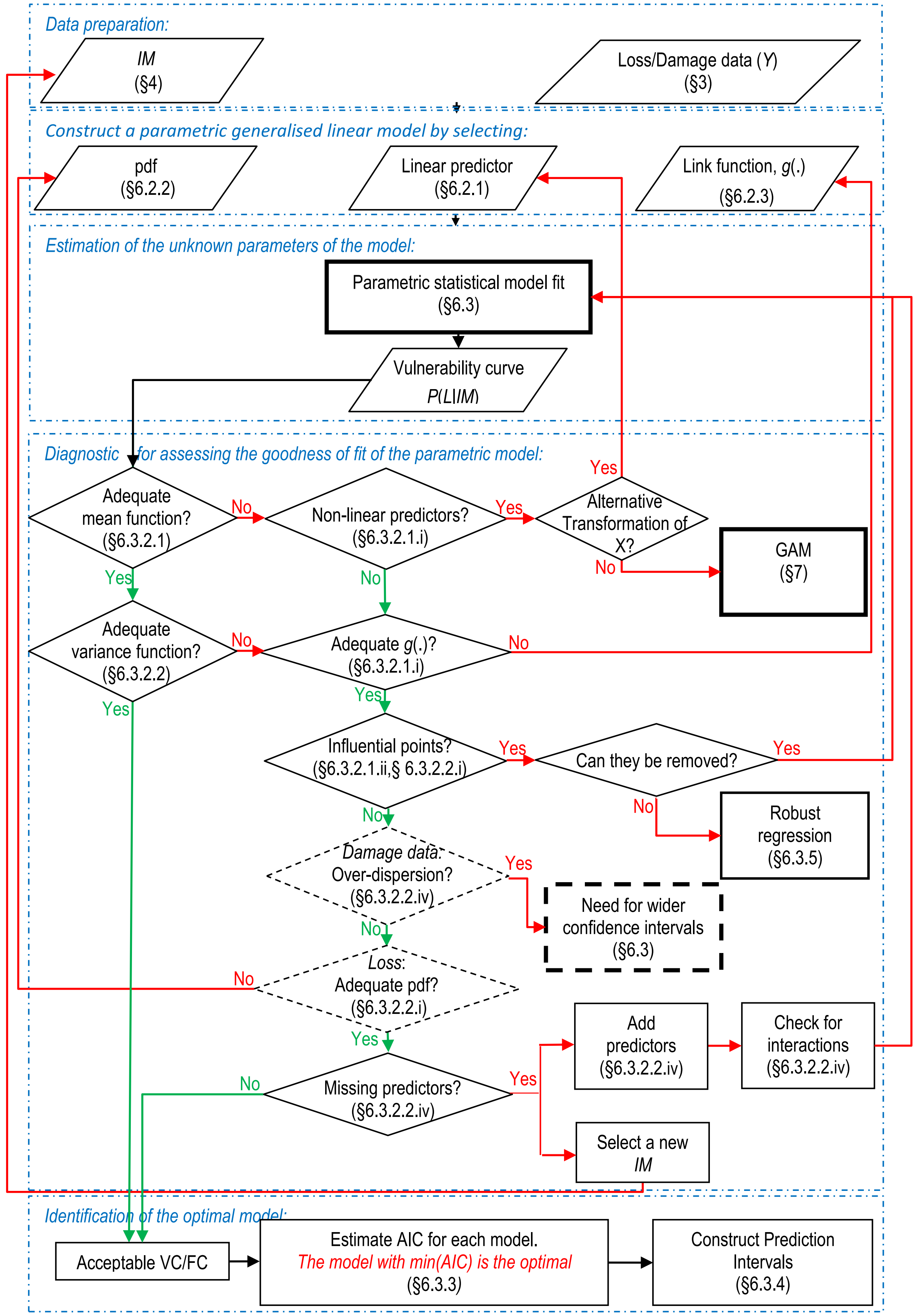


Figure 6.5 Flowchart of the parametric regression analysis for the construction of direct vulnerability curves and individual fragility curves for a given measure of ground motion intensity.

6.3.2.1 *Assessing the adequacy of the mean function*

The adequacy of the mean function can be assessed by examining whether the Pearson residuals have mean zero. This can be tested by plotting the residuals against the linear predictor (see Figure 6.6). The trend of the mean residuals is then identified by fitting a nonparametric mean curve. If the trend is approximately horizontal around zero, the linear predictor appears to be a good fit.

```
### Plots of residuals against linear predictor.
f<-predict(fit, type='link'); plot(f, res);          # Plot of the residuals against the linear predictor.
lines(lowess(f, res), col='red',lty=1,lwd=2)   # smooth mean curve.
abline(0,lty=2);   identify(fitted(fit), res)        # Identify potential outliers.
```

Figure 6.6 Plots of Pearson residuals against linear predictor (See Appendices B-D for illustrations of these plots).

The presence of a pattern in the plot of the Pearson residuals versus the linear predictor (e.g., the smoothed mean curve is parabolic) is an indication that there is:

i. *A nonlinear contribution of an individual explanatory variable*. To check this, the Pearson residuals should be plotted against *IM* (see Figure 6.7). The study of the patterns of the nonparametric mean curve in these latter graphs can identify potential means of improvement, e.g., the use of a higher order for the intensity measure *IM*, e.g., IM^2, or the natural logarithm of this variable, e.g., log(*IM*).

```
### Plot of residuals against each explanatory variable.
plot(IM, res)                                   # Plot of the residuals against the explanatory variable.
lines(lowess(IM, res), col='red',lty=1,lwd=2)
abline(0,lty=2)                                  # An auxiliary horizontal curve with y=0.
```

Figure 6.7 Plot of residuals against each explanatory variable.

ii. *Influential points***.** The poor fit of the selected model may be caused by the presence of influential points, i.e., points whose removal causes significant difference in the fit of the mean vulnerability and fragility curve. Influential points appear to be either unexpectedly distant from the average response (termed outliers) or from the average of intensity measure levels (termed points of high leverage). The outliers indicate unexpected loss or damage suffered by one or more buildings. This might arise due to a processing error or due to some special conditions that lead these buildings to be particularly susceptible to high levels of loss/damage. The points of high leverage typically are the result of the sparseness of data for high levels of intensity.

Outliers can be identified by checking for abnormally large residuals in the plots of the standardised Pearson residuals against the linear predictor (Figure 6.6).

Influential points can also be detected (see Figure 6.8) by examining the influence of deleting each data point in the set of model parameters. This is achieved by plotting the Cook's distance against the index of observations. Cook's distance is estimated as:

$$\mathrm{CD_i} = \frac{\sum_{j=1}^{M}\left(\hat{Y}_j - \hat{Y}_{j(i)}\right)^2}{\left(N_x + 1\right)\mathrm{var}\left(\mu\right)} \tag{6.7}$$

where $\hat{Y}_j$ is the prediction from the GLM for observation *j*; $\hat{Y}_{j(i)}$ is the prediction for observation *j* from a refitted model which omits observation *i*; CD is Cook 'distance.

```
## Index plot of Cook's distance
plot(fit, which=4)
or
plot(cooks.distance(fit)); identify(cooks.distance(fit))
```

Figure 6.8 Graphical assessment for outliers and influential data points (The reader is referred to Appendix D for an illustration of these plots).

Data points identified as potentially influential via the aforementioned procedures can then be removed and the same statistical model fitted to the remaining data points. The new fit is compared to the old fit. If the difference is large, then these points are truly influential and should be reported.

Note: Influential points or outliers, especially if there are more than one, can be removed ONLY if the analyst can justify that these points are not representative of the population, e.g., they are grouped damaged data based on a small sample of buildings (<20). If the removal of the influential points cannot be justified, parametric robust regression analysis should be adopted instead (see §6.3.5).

iii. *An inadequate link function*. The suitability of the link function can be investigated by the use of the Akaike Information Criterion (see §6.3.3) values for each model (see Figure 6.9).

```
## an example of the use of AIC.
Fit1<-glm(D~IM, family=binomial('logit')
Fit2<-glm(D~IM, family=binomial('probit')
Fit3<-glm(D~IM, family=binomial('log')
Fit4<-glm(D~IM, family=binomial('cloglog')

AIC(Fit1); AIC(Fit2); AIC(Fit3); AIC(Fit4);
```

Figure 6.9 Assessing the link function (The reader is referred to Appendix B for an illustration for this sensitivity analysis).

If the damage or loss data appear not to increase strictly monotonically and neither the transformation of the explanatory variables nor the presence of influential points can explain the trend, a nonparametric model should be fitted to the data according to the provisions outlined in §7 and §8.

6.3.2.2 Assessing the adequacy of the variance function

The ability of the selected variance function to capture the variability in the damage or loss data can be assessed by examining whether the variance of the residuals is constant (homoscedasticity check). The homoscedasticity of the residuals can be assessed by studying plots of the standardised Pearson residuals against the linear predictor (see Figure 6.5). If the scatter of these residuals appear to increase or decrease systematically with an increase in the linear predictor, then the residual variance cannot be considered constant, which means that the variance function is inadequate.

The homoscedasticity of the standardised Pearson residuals can be alternatively be assessed by a scale-location plot, i.e., the square root of the absolute values of the standardised residuals are plotted against the linear predictor (see Figure 6.10). A nonparametric mean curve is fit to the values of the y-axis. If the curve is horizontal, the homoscedasticity assumption holds. By contrast, a strong correlation between the two variables indicates that the assumption is violated.

```
### Scale-Location plot
library(boot); plot(predict(fit, type='response'), sqrt(abs (res) ))                    # the scale-location plot
lines(lowess(predict(fit, type='response'), sqrt(abs (res),col='red',lty=1,lwd=2)  # a nonparametric mean curve
identify(predict(fit, type='response'), sqrt(abs (res))                              # identify potential outliers.
```

Figure 6.10 Scale-location plot.

Heteroskedasticity of the residuals indicates that the variance of the selected distribution is unable to capture the variability in the data. An inadequate variance function could be caused by a number of reasons:

i. Inappropriate link function (see §6.3.2.1.iii).

ii. *Influential points* (see §6.3.2.1. ii).
Note: Outliers can also be identified in the scale location plot (Figure 6.9).

iii. *An inappropriate probability density distribution of the response variable (for loss data)*. The distribution of loss which fits the data best can be identified by a sensitivity analysis using the AIC values of the models.

iv. *Over-dispersion (for grouped data)*. The variance of the standardised Pearson residuals of damage data is unity if the damage data follow a binomial distribution. However, a large number of residuals outside the 99% confidence interval [-3,3] in Figure 6.6 or non-constant scatter of the residuals indicate that the data do not follow a binomial distribution. Over-dispersion can be taken into account in the construction of the confidence intervals following the procedure outlines in §6.3.1 (See Appendix B and C).

v. *Missing explanatory variables*. Whether important explanatory variables are missing can be assessed by adding explanatory variables, (e.g., soil conditions, building typology), and assessing their significance by a likelihood ratio test, which compares the more elaborate model with the simpler one (Figure 6.11).

```
## Examine whether the selected intensity measure type is a significant explanatory variable. We compare
##the model with the intensity measure type (fit) against the simpler model having only an intercept.

### Likelihood-ratio-test of the hypothesis: 'The extended model is not different to the reduced model'.
anova(fit,test= 'Chisq')

##If p<0.05 then there is strong evidence that the hypothesis is rejected, therefore the selected intensity
##measure type is an important explanatory variable.

## Accounting for more explanatory variables, e.g. soil conditions, in the model.
fit.ext<-glm(D~IM+S, family=binomial('probit')

## Likelihood-ratio-test of the hypothesis: 'The extended model is not different to the reduced model'.
anova(fit,fit.ext,test= 'Chisq')
```

Figure 6.11 Likelihood ratio test for nested models.

If multiple explanatory variables are included, interactions can exist between these variables, e.g., the soil conditions affect the levels of intensity. This interaction should be accounted for in the model and their significance should also be tested by comparing the difference between the models with and without the interaction (see Figure 6.12).

```
## Accounting for possible interactions in the model.
fit.int<-glm(D~IM+S+IM*S, family=binomial('probit'))

## Likelihood-ratio-test of the hypothesis: 'The model with the interaction is not different to the simpler
##model'.
anova(fit.ext,fit.int)

##If p<0.05 then there is strong evidence that the hypothesis is rejected, therefore the interactions are
##important.
```

Figure 6.12 Likelihood ratio test assessing for the significance of interactions.

Note: The incorporation of additional significant variables associated with the seismic characteristics indicates that the ground motion intensity measure is insufficient. In these cases, the analyst should perhaps use a different intensity measure following the provisions of §4.
Note: The building subclasses can be modelled by a dummy variable ($class$=C_1, C_2, C_3, ... ,Cn_{Class}).
Note: In some cases, especially if the curves appear to be almost horizontal the analyst is advised to examine whether the explanatory variable is significant to detect the trend in the data or it is not significantly different than a horizontal curve which cannot depict ay trend in the data. In this case, they are advised to use the likelihood ratio test outlines in Figure 6.11.

6.3.3 Identification of the optimal model

The diagnostics described in §6.3.2 highlight the goodness of fit (i.e., plots of residuals without patterns) of each chosen model to the data. Where more than one model has been tried for a given *IM* and is found to satisfy the goodness of fit tests, the analyst should select one to carry forward in the risk assessment by adopting either of the following procedures:

- Select the simpler model. For example, if the choice is between a gamma model with squared explanatory variable and a linear predictor using an inverse Gaussian distribution, then the second should be adopted.
- Use the AIC introduced by Akaike (1974) to determine which model provides the optimal fit to the data. The general function of the AIC is:

$$AIC = -2\log(likelihood) + 2(\text{number of parameters}) \tag{6.8}$$

The value of AIC is provided in the summary of the outcomes of the regression analysis. The regression model which provides the optimal fit to the data is the model with the smallest AIC value.

6.3.4 Construction of prediction intervals

Prediction intervals account for both the uncertainty in the mean estimate of the direct vulnerability curve ($\mu_j - \hat{\mu}_j$), presented by the confidence intervals (see §6.3.1), as well as the uncertainty of the mean curve predicting the observed data ($Y_j - \hat{\mu}_j$).

If an acceptable parametric vulnerability model is constructed following the procedures outlined above, its prediction intervals for the available intensity measure levels can be estimated by the bootstrap technique. The bootstrap technique can be found in Chandler and Scott (2013, pp 117-118).

It is important to note that the prediction of the response variable is valid only for the range of the available values of predictors (i.e., range of IM values). Predictions for values of the explanatory variables outside this range should be avoided.

6.3.5 Robust regression

Robust regression analysis should be used in cases where influential points have been identified by the diagnostics (§6.3.2) but cannot be removed. This method essentially reduces the influence of the influential points on the mean curve. The 'R' code for carrying out robust regression is presented in Figure 6.13.

```
### Robust regression
library(robust)
Rfit<-glmRob(D~IM, family=binomial).
### Diagnostics
plot(Rfit, which.plot=2) # deviance residuals vs fitted values
```

Figure 6.13 Robust regression.

6.3.6 Generalised linear mixed models (GLMM) for data from multiple events

Generalised linear mixed models should be used when data from multiple earthquakes are present. For these models, earthquake is considered an explanatory variable. However, we are not interested in the effect of the individual events, e.g., 1978 Thessaloniki, on the vulnerability or fragility curves. Instead, mixed models consider that the available earthquakes are randomly selected from a large population of seismic events and they model the random effect of earthquake to the construction of these curves.

The construction and fit of the mixed models is presented in Figure 6.14. These models can be compared with their GLM counterparts through the use of the AIC.

```
library(glmmML)

### Generalised linear model assuming that the building-specific damage data follow a Bernoulli distribution
fit_mix<-glmmML(Y~IM, family=binomial, cluster=event, data=Data)

library(glmmPQL)

### Generalised linear model assuming that the building-specific damage data follow a Bernoulli distribution
fit_mix<-glmmPQL(Y~IM, family=binomial, random=~event, data=Data)

summary(fit_mix)
```

Figure 6.14 Generalised linear mixed models.

6.4 Fitting a Parametric GLM: The Bayesian Approach (Level 2)

Bayesian statistical model fitting approaches require a strong statistical background. The main difference with the maximum likelihood approach (see §6.3) is that the GLM parameters are considered as random variables whose distribution is determined by the available data as well as any prior knowledge, obtained from existing vulnerability or fragility curves of similar building classes. Bayesian approaches are useful for the estimation of the GLM parameters for small sample sizes (i.e. $30 < N_T < 100$) as well as the parameters of increased complexity GLMs which account for measurement error in response or intensity measure levels as well as the random effects introduced by the use of databases from multiple earthquakes. Appendix E illustrates a simple application of this approach.

6.4.1 Estimation of the GLM parameters

Bayesian analysis estimates the posterior distribution of the model parameters, instead of their point estimate (as is done in the maximum likelihood approach). The posterior distribution of these parameters is proportional to the likelihood function and to their prior distribution:

$$f\left(\boldsymbol{\theta} \mid y,\mathbf{x}\right) \propto f\left(\boldsymbol{\theta}\right) L\left(\boldsymbol{\theta};y,\mathbf{x}\right) = f\left(\boldsymbol{\theta}\right) \prod_{j=1}^{M} f\left(y_j \mid \mathbf{x}_j,\boldsymbol{\theta}\right) \tag{6.9}$$

The denominator in the expression of Bayes theorem (see Eq.(5.3)) is a normalising constant component, and its determination is not required for the Bayesian model fitting analysis. Discrete approximations of the posterior distribution are obtained by using the Gibbs sampling method (for more info see Kruschke, 2011). According to this method, the values of the posterior distribution are generated by sampling from the conditional probabilities of the variables present in the right hand side of Eq.(6.9). This is achieved by a Markov Chain Monte Carlo (MCMC) algorithm, based on the construction of one or more chains of sampled values, where a realisation *j* depends on the previous realisation *j*-1. Initial values for the model parameters need to be provided (preferably close to the values of the posterior distribution), and a large number of iterations should be allowed for to obtain convergence.

An example of the Bayesian regression procedure, as performed in R, is provided in Figure 6.15. It should be noted, that in this example a weakly informative prior distribution has been used. For large numbers of observations, the standard error obtained is very close to the asymptotic values obtained by the maximum likelihood approach in §6.3.1.

```
### Construction of the posterior distribution of the parameters of a logit model:
library(R2OpenBUGS)
library(BRugs)

## Construction of the model
BayesModeldef <- function() {

  for (i in 1:nData) {                              # Likelihood Function
   y[i] ~ dbin( p[i],m[i] )
   logit(p[i]) <-  th0.star + th1*(x[i]-mean(x[])) # Standardising the predictor around the mean in order to reduce the
                                                    ## aucorrelation between the parameters and ensure faster
                                                    ##convergence.
   }                                                # Prior probabilities
  th0<-th0.star-th1*mean(x[])
  th0.star ~ dnorm( 0 , 1.0E-12 )
  th1 ~ dnorm( 0 , 1.0E-12 )    #  The variability of the prior distribution of the parameters is modelled in terms of the
                                #   precision=1/variance
write.model(BayesModeldef,"BayesModel.txt")
modelCheck( "BayesModel.txt" )

## Observed Grouped Data
nData<-length(D$IM)
nPredictors<-(1)
zy<-D$DSi
zx<-D$IM
tot<-(D$DSi+D$NoDSi)

## Initial values for the four chains. Note that the values differ significantly for each chain.
BayesInits <- list(list(th0.star= rnorm(1,0,12) ,th1=rnorm(1,0,12) ))

##Get the data in BUGS:
BayesModelData <-list(x=zx,m=tot,y=zy,nData=n)

## Run OpenBugs
a <- bugs(data=BayesModelData,
          n.chains=1,                              # 1 Marcov chain.
           inits=BayesInits,                       #  Initial values for each chain.
            n.burnin=500,                           # 500 burn-in points.
             n.iter=5000,                          # Total number of iterations.
              parameters.to.save=c("th0","th1","mu"),     # parameters for which the mean values and inference is
##required.
               model.file="BayesModel.txt",
                n.thin=1,          # Thinning, i.e., save the value of the posterior distributions every 100th iterations for
##each chain.
                 debug=TRUE)

plot(a)
```

Figure 6.15 Bayesian model fitting analysis.

6.4.1.1 Assessing the performance of the numerical algorithm used in the Bayesian procedure

The performance of the Markov Chain Monte Carlo algorithm should be examined in order to establish the convergence of the numerical procedure and the absence of autocorrelation in the chains. The sensitivity of the procedure to the prior distribution should also be examined.

i. The validity of the results of a Bayesian procedure depends on the convergence of its numerical procedure. The convergence can be assessed by plotting the value of the GLM parameters against their corresponding iterations (see Figure 6.16). The convergence is typically highly dependent on the initial values, which are not always close to the values of the posterior distribution. Therefore, the convergence rate can be improved by reducing the sensitivity to the initial values through assuming a burn-in period, i.e., through determining the number of initial samples in each chain which should be ignored (see Figure 6.15). The sensitivity to the initial values of the parameters can be examined by considering 3 or 4 chains with widely different initial values.
ii. The MCMC algorithm produces samples of the posterior distribution of the model parameters that depend on previous values. Autocorrelation is caused when the model parameters are highly correlated. This raises the question on whether autocorrelation in each chain, leads to unrepresentative estimates of the posterior distribution. This can be assessed by the autocorrelation plots (see Figure 6.16). High autocorrelation is present if a number of successive positive or negative values are noted in adjoining lags. The influence of autocorrelation can be typically reduced by the re-parameterisation of the GLM. An example of this technique can be found in Figure 6.15.
iii. The shape of the posterior distribution of the regression parameters often depends on the characteristics of the prior distribution. The effect of the prior distribution on the results can be assessed by comparing this distribution with the posterior. If the two appear to be very close, the data are not very informative. In these cases, the selection of the prior is important and the analyst should explain the rationale behind the selection of this distribution. Ideally, a prior distribution should adequately reflect the belief the analyst has on the values of the parameters. This distribution can be non-informative (very large uncertainty) if the analyst has no prior knowledge and informative (if the analyst has some prior knowledge). The shape of the selected distribution can also affect the results. For example, the selection of a uniform prior assumes that the probability of selecting values outside the predetermined range is zero. This could have a profound impact on the posterior distribution in certain cases. In addition, the use of prior distribution which allow for negative values although the parameter is expected to be positive is also problematic.

```
### Bayesian Regression Diagnostics.

## Testing the convergence.
par(mar=c(4,3,2,2))
plot(as.mcmc.list(a),smooth=FALSE)

## Testing the autocorrelation.
plot(acfplot(as.mcmc.list(a)))
```

Figure 6.16 Bayesian model fitting approach diagnostics (see Appendix E for an illustration of these diagnostics).

6.4.1.2 Assessing the goodness of fit of the model

Apart from the above diagnostics presented for assessing the MCMC algorithm, the analyst is advised to examine the whether the fit is acceptable. This can be assessed by plotting the observed response against the predicted and compare their trend compared to the 45 degree line. If the plotted data appear to be randomly distributed and reasonably close to the 45 degree line, the model is acceptable.

6.4.2 Identification of the optimal model

The proposed parametric model fitting techniques may result in more than one adequate model for a given IM. In this case, the analyst should select one, by using the Bayesian factor (BF) to compare the models:

$$BF = \frac{f\left(D|M_A\right)}{f\left(D|M_B\right)} = \frac{\int f\left(D|\boldsymbol{\theta}_A, M_A\right) f\left(\boldsymbol{\theta}_A|M_A\right) d\boldsymbol{\theta}_A}{\int f\left(D|\boldsymbol{\theta}_B, M_B\right) f\left(\boldsymbol{\theta}_B|M_B\right) d\boldsymbol{\theta}_B} \tag{6.10}$$

where M_A, M_B is the model A and B respectively.

6.4.3 Construction of prediction intervals

The predictive probability of the fitted curve for new data can be estimated as:

$$f\left(y_N|y, im\right) \propto L\left(\boldsymbol{\theta}; y, im\right) f\left(im\right) \tag{6.11}$$

where y_N a set of new response data.

6.4.4 Combining multiple databases of post-earthquake survey data

The Bayesian approach can be used for combining multiple databases of post-earthquake survey data, e.g. databases E1, E2, E3 etc. Assuming that the observations in these databases are independent, multiple Bayesian model fitting analyses can be performed in order to obtain the final vulnerability and fragility relationship. This is done through the following steps:

Step 1: Direct vulnerability and fragility curves are constructed using Database E1 assuming appropriate prior distributions of the parameters. Non-informative distributions should be selected if the analyst has no prior knowledge and information otherwise.

Step 2: New direct vulnerability and fragility curves are constructed by using the distributions of the model parameters obtained in Step 1 as prior distribution in this step and updating them using the observed data from Database E2.

Step 3: The procedure is repeated for all available databases.

Within this process it is advised that the largest database be used for Step 1. Small databases with non-informative prior distributions may result in problems with the convergence of the MCMC algorithm.

6.4.5 Modelling measurement error

Measurement error in response (see §3.3.2.4) and explanatory variables (see §4.4.2) can potentially affect the shape of the mean vulnerability and fragility curves as well as their confidence intervals. If they are significant, the analyst is advised to account for these errors in the construction of GLMs and compare the obtained vulnerability or fragility curves with their corresponding obtained without explicitly account for the measurement error. Brief guidelines are provided here regarding the modelling of the misclassification error in the damage assessment and the measurement error in the intensity measure levels.

6.4.5.1 Modelling misclassification error in empirical damage data

The misclassification error in damage (see §3.3.2.4) is nontrivial if the damage observations derive from assessments made using remote sensing techniques. Misclassification errors can also be found, (although no study has to date attempted their quantification), in databases deriving from rapid surveys due to the speed with which the surveys are conducted or misunderstanding of the survey form. The method presented below for modelling this error in the fragility assessment assumes that this error is non-differential, i.e., independent from the ground motion intensity.

The proposed procedure is based on expressing the observed damage in a binary form, *Y* (see Eq.(5.1)). The analyst is expected to have an idea about the size of the error, e.g., by comparing the contaminated database with independent accurate surveys and to have assessed the performance of the former through the estimation of its specificity (i.e., the probability that the observed response is 1 given that the 'true' response is 1) and sensitivity (i.e., the probability that the observed response is 0 given that 'true' response is 0), as per Eq.(6.11). The misclassification error in the observations can then be written as:

$$\begin{aligned} P(\text{false positive}) &= P(y_j=1|z_j=0) = 1-P(y_j=1|z_j=1) = 1-\text{specificity} \\ P(\text{false negative}) &= P(y_j=0|z_j=1) = 1-P(y_j=0|z_j=0) = 1-\text{sensitivity} \end{aligned} \tag{6.12}$$

The values of the probabilities in Eq.(6.12) are considered known and they can be incorporated in the likelihood function as:

$$L(\boldsymbol{\theta};y,im) = \prod_{j=1}^{M} \mu_j^{y_j}\left(1-\mu_j\right)^{1-y_j} \tag{6.13}$$

where the mean μ is obtained from:

$$\begin{aligned} \mu_j = E\left[y_j|im_j\right] &= P(y_j=1|im_j) = P(y_j=1|z_j=1)P(z_j=1|im_j) + P(y_j=1|z_j=0)P(z_j=0|im_j) = \\ &= \left[1-P(y_j=0|z_j=1)\right]P(z_j=1|im_j) + P(y_j=1|z_j=0)\left[1-P(z_j=1|im_j)\right] \end{aligned} \tag{6.14}$$

and where the probability $P(z_j=1|im_j)$:

$$g\left[P(z_j=1|im_j)\right] = \vartheta_0 + \vartheta_1 im_j \tag{6.15}$$

Assuming statistically independent observations and negligible measurement error in intensity measure levels, a Bayesian procedure can be used to estimate the inference for the regression parameters. The analyst is referred to the work of McInturff et al. (2004).

6.4.5.2 Modelling the measurement error in intensity measure levels

The measurement error in intensity measure levels, as described in §4.4.2, typically occurs in cases where, (in the absence of ground motion records), the intensity is determined through ground motion prediction equations. This error is considered non-differential, i.e., the error does not depend on the response variable, and is incorporated in the construction of vulnerability and fragility curves through a Bayesian approach. A classic model is selected to express the measurement error where the observed intensity level, *IM,* is estimated as a function of the 'true' $\overline{IM}$ in the form:

$$\ln\left(IM\right)=\ln\left(\overline{IM}\right)+\sigma_T\varepsilon=f\left(M,R,S\right)+\sigma_T\varepsilon \tag{6.16}$$

where σ_T is the standard deviation of the GMPE and ε follows a normal distribution N(0,1). The new model is obtained by expanding the simple model, presented in Figure 6.16, in order to account for the error model as:

Regression Model: $f\left(y_j \mid x_j,\boldsymbol{\theta}\right)$

Measurement Error Model: $f\left(x_j \mid \bar{x}_j,\boldsymbol{\xi}\right)$

Prior Model: $f\left(\bar{x}_j \mid \boldsymbol{\pi}\right)$

where y the response variable; x is the ground motion intensity; $\boldsymbol{\theta}=[\vartheta_0, \vartheta_1, \ldots]$ is the vector of regression parameters; $\boldsymbol{\xi}$ is a vector of parameters which describe the relationship of the true and observed *IM*; $\boldsymbol{\pi}$ is the vector of the parameters of the prior distribution of the observed *IM*. Therefore, the posterior distribution is estimated by the MCMC algorithm which models the following joint distribution (see Richardson and Gilks, 1993):

$$f\left(\boldsymbol{\theta},\boldsymbol{\xi},\boldsymbol{\pi} \mid y\right)\propto f\left(\boldsymbol{\theta}\right)L\left(\boldsymbol{\theta};y,x\right)=f\left(\boldsymbol{\theta}\right)f\left(\boldsymbol{\xi}\right)f\left(\boldsymbol{\pi}\right)\prod_{j=1}^{M}f\left(y_j \mid x_j,\boldsymbol{\theta}\right)\prod_{j=1}^{M}f\left(x_j \mid \bar{x}_j,\boldsymbol{\xi}\right)\prod_{j=1}^{M}f\left(\bar{x}_j \mid \boldsymbol{\pi}\right) \tag{6.17}$$

6.4.6 Modelling the random effect from multiple earthquakes

Instead of using the generalised linear mixed models from the 'R' package (see §6.3.6), the analyst could construct mixed models using a Bayesian framework. The latter approach is more flexible as it can be expanded to incorporate the measurement error in the response or the intensity measure. This procedure is novel and the user is referred to the tutorials in OpenBugs (Lunn et al., 2009) for more information.

7 Fitting Generalised Additive Models (Level 2)

7.1 Introduction

Generalised additive models (GAM) are recommended in §5 for use when the trend in damage or loss data is not-strictly monotonically increasing with the ground motion intensity. These models are an extension of the GLMs (Hastie and Tibshirani, 1986), outlined in §6, the difference being that the assumption of linearity is relaxed.

Similar to GLMs, generalised additive models are constructed by:

- Selecting the probability distribution function of the response variable conditioned on the explanatory variables, following similar procedures to GLM (see §6).
- Determining the systematic component, which can be written as:

$$g(\mu)=\eta=s_0+\sum_{i=0}^{N}s_i(X_i) \tag{7.1}$$

where s_0, $s_j(.)$ are smoothing functions. Guidance is provided below for the determination of the parametric form and the degree of smoothness of these functions required for the estimation of the shape of the generalised additive models. The smoothing functions are expressed in the form:

$$s_i(X_i)=\sum_{j=0}^{q}b_{ij}(X_i)\beta_{ij} \tag{7.2}$$

where q is the total number of the smoothing parameters; $b_{ij}(.)$ is a basis function and β_{ij} is an unknown parameter. The basis function is expressed here in terms of a spline, due to its ability to account for dependencies in the damage or loss observations (Wood, 2006). The splines are made of sections of a function (typically 3rd degree polynomial) and are joined together at specified points, termed knots, in order to form a continuous regression curve. The parameters β_{ij}, which are different in each section, are estimated from a penalised likelihood method ($l_p(.)$), as:

$$\boldsymbol{\beta}^{opt}=\arg\max l_p(\boldsymbol{\beta})=\arg\max\left[l(\boldsymbol{\beta})-\frac{1}{2}\lambda\mathbf{J}\right] \tag{7.3}$$

where λ_j is a non-negative parameters which adjusts the degree of smoothing, **J** is a roughness penalty component (for further information read Hastie and Tibshirani, 1990).

In what follows, the analyst is provided with guidance regarding the determination of the random and systematic component of the GAM, the estimation of its parameters and diagnostic tools to assess its goodness of fitted as depicted in Figure 7.1. The provided guidance is based on the 'mgcv' package (Wood, 2014). An example application is provided in Appendix F.

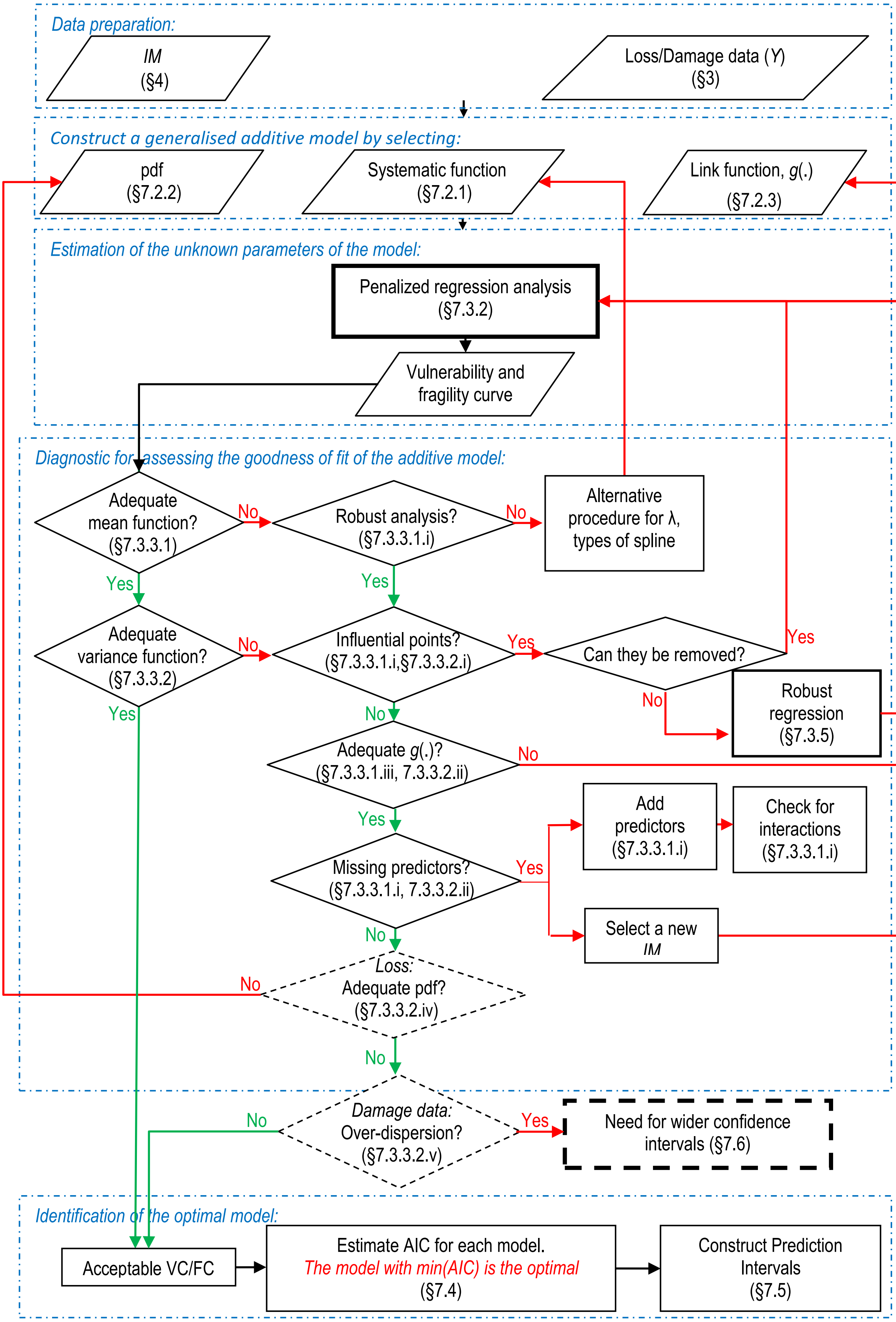


Figure 7.1 Flowchart of the GAM regression analysis for the construction of direct vulnerability curves and individual fragility curves for a given measure of ground motion intensity.

7.2 Construction of a Generalised Additive Model, GAM

7.2.1 *Determination of the systematic component*

Similar to GLMs, the simplest form of the systematic component, *η,* of GAMs is a function of the intensity measure IM, expressed in the form:

$$\eta = \theta_0 + s\left(IM\right) = \theta_0 + \sum_{j=0}^{k-1} \theta_j b_j\left(IM\right) \tag{7.4}$$

where $b_j(IM)$ is the smoothing basis function, expressed in terms of a cubic regression spline, favoured in the literature (e.g., Wood, 2006). These curves consist of piecewise 3rd degree polynomials which are joined at points termed knots. The dimension *k* of the basis function determines the number of knots, (and thus the number of piecewise curves fitted to the data points). This effectively determines the number of parameters required to fit the constructed GAM. In general, the smaller the number of knots selected by the analyst, the smoother the resulting curve. However, the analyst is advised to examine the impact of *k* on the fit. If the number of parameters is very close to the number of observations, selected model is over-fitted.

7.2.2 *Selection of a probability distribution for the response variable*

The probability distributions of loss or damage presented in §6.2.2 can also be used in the case of the generalised additive models.

7.2.3 *Selection of link function*

The link functions presented in §6.2.3 are also suitable for the GAM models, and the reader is referred to this section.

7.3 Estimation of the Parameters in the Nonparametric Model

The estimation of the model parameters, ϑ_j, depends on the prior estimation of the smoothing parameter, λ.

7.3.1 *Estimation of the smoothing parameter, λ*

The value of the smoothing parameter, λ, is important for the determination of the GAMs. For $\lambda \to \infty$, the likelihood function (Eq.(7.3)) is maximised within the roughness penalty component equal to zero (Green, 1987). Therefore the additive model is effectively turned into a GLM model, which is unable to capture the non-strictly monotonic trends in the data. By contrast, $\lambda \to 0$ results in curves which are very wiggly indicating perhaps over-fitted models. The optimum value of this parameter is estimated by data-dependent procedures which are included in the adopted R package (see Figure 7.2). The selected procedure depends on the distribution of the response given levels of intensity is known of not. If distribution has known dispersion (i.e. for the binomial or Bernoulli distribution) then the procedure of un-biased risk estimator is adopted. Alternatively, the cross-validation or generalised cross-validation procedure should be selected. The user is referred to Wood (2006) for more information regarding the two procedures.

7.3.2 *Estimation of the parameters of the GAM*

Having determined the smoothing parameter λ, the nonparametric mean vulnerability curve and fragility curve (for each damage state) are constructed by the numerical estimation of the unknown model parameters, ϑ_j, of Eq.(7.1). This is done by maximising the penalised likelihood function using a penalized iterative least squares algorithm (see Figure 7.2).

The confidence intervals around the systematic component can be estimated from the standard errors (see Figure 7.2) provided by 'R' (R Development Team, 2008).

```
### Penalised maximum likelihood determination of GAMs, using a cubic regression spline.
library(mgcv)

fit<-gam(Loss~s(IM,bs='cr'), family=Gamma(link=c('identity','inverse','log'),method= 'GCV.Cp')
fit<-gam(Loss~ s(IM,bs='cr'), family= inverse.gaussian (link=c('identity','inverse','log','1/mu^2'),method= 'GCV.Cp')
fit<-gam(log(Loss)~ s(IM,bs='cr'), family=gaussian(link=c('identity','inverse','log'),method= 'GCV.Cp')

### Generalised additive model assuming that the building-specific damage data follow a Bernoulli distribution.
fit<-gam(Y~ s(IM,bs='cr'), family=binomial(c('logit', 'probit', 'log', 'cloglog'), method= 'UBRE')

### Generalised additive model assuming that the grouped damage data follow a binomial distribution.
fit<-gam(D~ s(IM,bs='cr'), family=binomial(c('logit', 'probit', 'log', 'cloglog'),method= 'UBRE')

summary(fit)  # provides a summary of the outcomes of the analysis.

## Construction of Bayesian confidence intervals of the mean regression curve.
fit.pred<- predict(fit,type='link', se.fit=TRUE)     # the standard error and the mean values of the linear predictor is provided if the
                                                     # logic link function is selected.
f.upper<-1/(exp(fit.pred$fitted.values-2*fit.pred$se.fit)+1)  # the 90% confidence interval.
f.lower<-1/(exp(fit.pred$fitted.values+2*fit.pred$se.fit)+1)  # the  5% confidence interval.
```

Figure 7.2 ***Penalised maximum likelihood determination of GAM models and their confidence intervals*** **(Wood, 2006).**

7.3.3 Assessing the goodness of fit of the GAM

Similar to the case for generalised linear models, the assessment of the goodness of fit of the generalised additive models is performed here through, mainly, graphical diagnostics tools. The diagnostics tools are mostly, but not exclusively, based on the study of the Pearson residuals (Eq.(6.5) see Figure 7.3), which should have zero mean and constant standard deviation if the selected GAM model is a good description of the available data. Figure 7.1 presents (a rather mechanistic) flowchart of the procedure that can be adopted in order to assess the fit of the selected generalised additive model. Recommendations regarding the assessment of the mean and variance function of the fitted model follow.

```
### Pearson residuals.
library(mgcv)
res<-residuals(fit,type=c('pearson'))        # Pearson residuals.
```

Figure 7.3 Calculating the Pearson residuals in 'R' (see Appendix F for an illustration of these residuals).

7.3.3.1 Assessing adequacy of the mean function

The goodness of fit of the mean function can be assessed by:

i. *Examining the robustness of the model to the procedure adopted for the determination of λ and the dimension k.* The robustness of the smoothness selection procedure should be examined (see Figure 7.4). A sensitivity analysis is performed by selecting different procedures depending on whether the dispersion in unknown or known. The model is considered robust to smoothness selection if the changes in the effective degree of freedom, which is a function of the smoothing parameter λ, are small.

```
### Robustness of model to smoothness selection procedure.
fit<-gam(Loss~s(IM),family=Gamma(link=log),method= 'GCV.Cp')
```

Figure 7.4 Checking the robustness of model to smoothness selection procedure.

In order to check the robustness of the number of knots *k,* the GAM model is re-fit by considering a larger (e.g., twice as large) value for *k*. The model is considered robust if the changes in the results are small, (see Figure 7.5.).

```
### Checking the robustness of model to the number of knots.
fit<-gam(y~s(IM, k=c(10,20) ,family=Gamma(link=log), method= 'GCV.Cp')
```

Figure 7.5 Checking the robustness of model to the number of knots.

Note: The number of regression parameters, *k*, should not be as large as the number of observations in order to avoid over-fitting.

ii. *Examining whether the residuals have mean zero*. This can be tested by plotting the residuals against the systematic component, η (see Figure 7.6). However, this plot may not be as informative as for the assessment of the generalised linear models, given that GAM generally follow the trend in the data.

iii. *Plotting the observed data points against their corresponding fitted values*. If the points are evenly spread and close to the diagonal trend, without highlighting a pattern or influential points (see §6.3.2), the model can be considered satisfactory (see Figure 7.6). *Note:* If the points appear to be non-randomly scattered, then more explanatory variables needs to be added to the model.

```
### Plot of observed data against the corresponding fitted values.
plot(fitted(fit), y)
lines(c(0,max(y)),c(0,max(y)))      # diagonal curve
```

Figure 7.6 Plot of the observed data points against their corresponding fitted values.

Unsatisfactory mean functions can be attributed to:

i. *The presence of influential points.* If potential influential points are identified in the residuals in Figure 7.3, the analyst should remove them and repeat the model fitting procedure. If the two curves (i.e., with and without the potentially influential points) appear to deviate significantly, then the points are indeed influential. The analyst may remove them only if there is enough information to justify this action. If no such information exists, robust regression analysis should be used (see §7.3.5).

ii. *Inadequate link function.* A sensitivity analysis using different functions (see Table 6.2) should be performed in order to identify a potentially better GAM model.

iii. *Missing explanatory variables*. In general, adding more explanatory variables, e.g., soil conditions, building typology, improves the fit of a model. Whether this improvement is significant is assessed by preforming a likelihood ratio test, which compares the more elaborate model with the simpler one (see Figure 7.7).

```
## ANOVA test of hypothesis: 'The GAM model, fit_nl, is not significantly different than the GLM, fit.'
anova(fit_nl, fit, test='Chisq')

#If p<0.05 then there is strong evidence that the hypothesis is rejected, therefore the GAM model fits the
#data better than its GLM counterpart.
```

Figure 7.7 Chi-square test for comparing a nested GAM models.

If more than one explanatory variable are included, interactions between these variables should be considered (e.g., the soil conditions affect the levels of intensity). This interaction should be accounted for in the model and their significance should also be tested by comparing the difference between the models with and without the interaction (see Figure 7.8). *Note*: The incorporation of additional significant explanatory variables associated with the seismic characteristics indicates that the ground motion intensity measure is insufficient. In these cases, the analyst should perhaps use a different intensity measure following the provisions of Chapter 4.

```
## Accounting for possible interactions in the model.
fit.int<-gam(D~s(IM)+S+IM*S, family=binomial('probit'))

## Testing the hypothesis: 'The model with the interaction is not different to the reduced model'.
anova(fit_int, fit, test='Chisq ')
```

Figure 7.8 Likelihood ratio test assessing for the significance of interactions.

7.3.3.2 Assessing adequacy of the variance function

The ability of the selected variance function to capture the variability in the damage or loss data can be assessed by examining whether the variance of the residuals is constant (i.e., homoscedasticity assumption). The homoscedasticity of the residuals can be assessed by the study of the plots of the residuals against the systematic component (see Figure 7.3). If the scatter of the residuals appears to increase or decrease systematically with an increase of the linear predictor, then the residual variance cannot be considered constant. In addition, if for grouped damage data a large number of residuals lies outside the 99% confidence interval [-3,3] in Figure 7.3 and/or there is a non-constant scatter of the residuals, then the data may not follow a binomial distribution.

The aforementioned undesirable cases indicate that the variance of the selected distribution is unable to capture the variability in the data. An inadequate variance function could be caused by a number of reasons:

i. *Influential points* (see § 7.3.3.1.v).
ii. *Inappropriate link function* (see § 7.3.3.1.vi).
iii. *Missing* explanatory *variables* (see § 7.3.3.1.vii).
iv. *Inappropriate probability density distribution of the response variable (for loss data)*. The distribution of loss which fits the data best can be identified by a sensitivity analysis where the three distributions are selected and the goodness of fit of each model is assessed.

v. *Over-dispersion (for grouped damage data)*. This can be attributed either to the presence of too many response values with y_j=zero values, or to potential spatial dependencies among the buildings (see the discussion in §6.3.2.2). The latter can be taken into account by the use of mixed models, presented in §7.6.

vi. *Heteroskedasticity (for a continuous response variable)*. This can be improved through the use of mixed models, presented in §7.3.6.

If the aforementioned strategies fail to produce a satisfactory variance function, then the use of a "kernel smoothing procedure", which does not depend on a specific variance function, is recommended (see §8).

7.3.4 Identification of the optimal model

The diagnostics of §7.3.3 may highlight a good fit (i.e., plots of residuals without patterns) of more than one GAM model for a given intensity measure type. Similarly to the procedures proposed for the generalised linear models, the analyst should select the model which:

- Has the smallest AIC (see §6.3.3).

The importance of the nonlinearity captured by the optimal GAM model compared with a corresponding linear GLM model should also be examined by (see Figure 7.9):

- Comparing the AICs of the GLM and GAM models fit. The model with the lower value of AIC is considered the best.
- If the GAM model includes the GLM (i.e. both models have the same link function and the same distribution of the response for given IMLs), a Chi-square test is performed, which tests the hypothesis that the GAM model is not significantly different than the GLM. If $p<0.05$, there is sufficient evidence to reject the hypothesis. Therefore, in this case, the GAM fits the data best.

```
## AIC
AIC(fit_nl,fit)

##Chi-square test of hypothesis: 'The nonlinear GAM model, fit_nl, is not significantly different than the linear model,
##fit.'
anova(fit_nl, fit, test='Chisq ')
```

Figure 7.9 Comparing a GLM and GAM models fit to the same database.

7.3.5 Robust Regression

If the diagnostics in §7.3.3 identify the presence of outliers that cannot be removed, robust regression should be performed. Available procedures for robust regression on empirical damage and loss data are limited. An algorithm for robust regression for discrete damage data recently published by Alimadad and Sabidian-Barrera (2011) can be found in the package 'rgam' (Sabidian-Barrera et al., 2014). However, the application of this algorithm produces only a mean fit. The residuals based on this method are plotted against the fitted values in order to establish the improvement of the fit.

7.3.6 Generalised additive mixed models, GAMM

Similar to GLMs, the random effect of an earthquake in the data collected from multiple events is taken into account by the construction of generalised additive mixed models (GAMMs). The analyst is warned that these methodologies, especially for modelling the spatial correlation, are not well established and therefore should be used with care. It is also reminded that the construction of mixed models is one of the strategies dictated by the poor diagnostics of the variance function: i.e., over-dispersion of residuals for discrete response variables and heteroskedasticity for economic loss. Figure 7.10 provides the ‘R’ code for fitting a mixed model assuming that ‘event’ is a random effect. The plots of residuals presented in Figure 7.3 should also be used here in order to establish the level of improvement obtained by fitting the mixed model to the data.

```
## Estimating the parameters of the GAMM model expressed by Eq.(5.6).
library(mgcv)
fit_mix<-gamm(D~IM, random=list(event=~1), family=binomial)

## AIC for models comparison
AIC(fit_mix$lme)
```

Figure 7.10 Construction and diagnostics of GAMMs.

If more than one GAMM model provides a good fit to the data, the optimum can be identified as the one having the smallest AIC. The overall improvement in the description of data of the selected model is also assessed by comparing its AIC with that obtained from the GAM model that does not account for the random effect of the earthquake.

The construction of prediction intervals for mixed models can be achieved by using a Bayesian analysis.

The analyst should also examine the presence of spatial correlation following the diagnostics outlined in §6.3.6. The proposed ‘R’ package is capable of modelling a spatial structure. This, however, is a subject of current research and the analyst is referred to the examples illustrated by Zuur et al (2009) for more information.

7.3.7 The Bayesian approach

Similar to the recommendations for GLMs, a Bayesian analysis is necessary in order to fit the complex GAMs which accounts for the measurement errors in the response and explanatory variables as well as model mixed effects.

8 Fitting Gaussian Kernel Smoothers (Level 2)

Gaussian Kernel Smoothers (GKS) can be used in order to capture a non-strictly monotonic trend in the data. These models are more flexible than GAMs as they do not require an assumption regarding their random or systematic component. They could be used for the modelling of the measurement error in the intensity measure levels, although this is still an area of active research. An introduction to the models and the procedures used to construct the non-parametric fragility/direct vulnerability curves can be found in the GEM report by Noh, (2011).

9 Transformation of Fragility Curves into Vulnerability for Specific Measures of Intensity

Indirect vulnerability assessment requires the transformation of fragility curves into vulnerability through the total probability theorem:

$$P\left(L>l\,|\,IM\right)=\sum_{i=1}^{n}P\left(L>l\,|\,ds_i\right)p\left(ds_i\,|\,IM\right) \tag{9.1}$$

where n the number of damage states, $p(ds_i\ |IM)$ is the probability of a building sustaining a damage state ds_i given intensity IM; $P(L>l\ |\ ds_i)$ is the probability that loss, L, exceeds a value l given a damage state given ds_i, termed complementary cumulative distribution of the loss given ds_i; $P(L>l\ |\ IM)$ is the complementary cumulative distribution of loss given a level of intensity IM.

The transformation expressed by Eq.(9.1) requires the construction of damage probability matrices from the fragility curves for specific levels of ground motion intensity:

$$p\left(ds_i\,|\,IM\right)=\begin{cases}1-P\left(DS\geq ds_i|IM\right) & i=0\\ P\left(DS\geq ds_i|IM\right)-P\left(DS\geq ds_{i+1}|IM\right) & 0<i<n\\ P\left(DS\geq ds_i|IM\right) & i=n\end{cases} \tag{9.2}$$

For a given IM, this matrix represents the distance between two successive fragility curves, as presented in Figure 9.1.

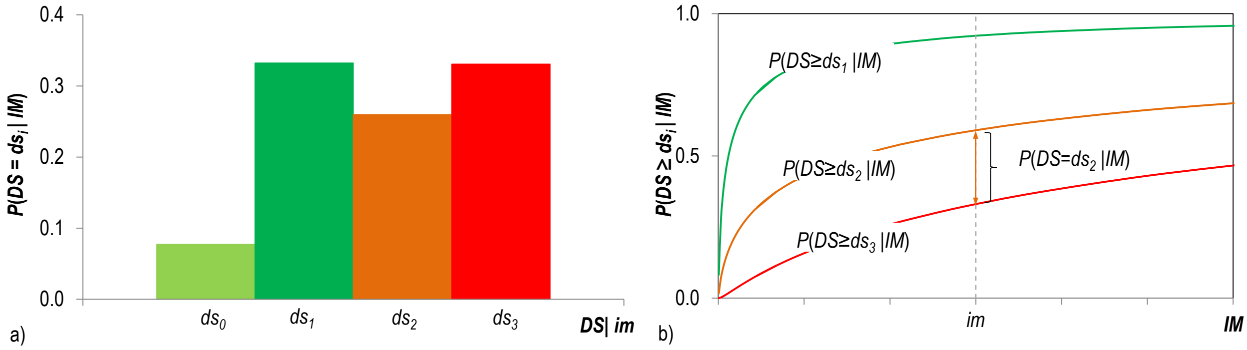


Figure 9.1: Illustration of a) a column of a DPM for given intensity measure level *im*, b) fragility curves corresponding to *n=3* damage states for the same building class.

In what follows, methods of varying complexity are presented in order to estimate the loss for given intensity measure levels from Eq.(9.1). Choice of which method to use depends on the quality of available information regarding the damage-to-loss functions as well as the analyst's requirements.

- *Method 1*: can be adopted if the loss given levels of damage is a random variable with known mean and variance. In this case, the mean and variance of the loss for given intensity measure levels can be obtained through the following closed form solutions:

$$E\left(L\mid im\right)=\sum_{i=1}^{n}E\left(L\mid ds_i\right)p\left(ds_i\mid im\right) \tag{9.3}$$

$$\text{var}\left(L\mid im\right)=\sum_{i=1}^{n}\left[\text{var}\left(L\mid ds_i\right)+E^2\left[L\mid ds_i\right]\right]p\left(ds_i\mid im\right)-E^2\left[L\mid im\right] \tag{9.4}$$

Method 1 results in the estimation of hte first two moments, namely: the mean and the variance of the loss for given intensity measure levels.

- *Method 2:* can be used if the conditional distribution of loss, $P(L>l \mid ds_i)$, is known and the mean fragility curves are considered. In this case, the nonparametric distribution of the vulnerability can be numerically estimated from Eq.(9.1) through a Monte Carlo procedure as proposed by Porter et al. (2001), and shown in Figure 9.2. According to this procedure, for each iteration:

Step 1. A number u ~ [0,1] is randomly generated from a uniform distribution.
Step 2. The corresponding damage state is obtained: $ds_i=F^{-1}(DS\mid im)$.
Step 3. Then, a number y is randomly generated from the distribution of loss given the corresponding damage state ($l = F^{-1}(L\mid DS= ds_i)$).
Step 4. Steps 1-3 are repeated a large number of times, e.g., 10,000.

Method 2 estimates the mean, variance as well as the shape of probability distribution of the loss for given intensity measure levels.

In some cases, the width of the confidence intervals around hte mean fragility curves may be notably wide. For these cases, the analyst is advised to propagate this uncertainty to the loss for given intensity measure level. This can be achieved by an advanced Monte Carlo procedure (for more details see Ioannou and Chrysanthopoulos). This procedure leads to a family of probability distributions for the loss conditioned on a intensity level instead of a single probability distributions estimated by Method 2.

```
### Estimating the distribution of the loss given a level of intensity.
## Assuming that a logistic regression presented in Figure was successfully fitted to the data.

ds<-numeric(i)    # vector of damage states
u<-numeric(i)     # random variable in (0,1)
F<-numeric(i)     # fragility curves
L<-numeric(i)     # loss variable
lmu<-numeric(i)   # lognormal mean of the lognormal distribution of the loss given a damage state.
lsgm<-numeric(i) # lognormal standard deviation of the lognormal distribution of the loss given a damage
state.
xnew=im           # A specific level of intensity
dat<-data.frame(x=xnew)
Nit=10,000   # number of iterations
N=3 # number of fragility curves
## Fragility curves
F[1]=predict(fit1,newdata=dat,type='response');F[2]=predict(fit2,newdata=dat,type='response');F[3]=predi
ct(fit3,newdata=dat,type='response')
u<-runif(Nit, min=0, max=1) # random generation of Nit values u in (0,1)
for (i in 1:Nit) {
 if (u[i]>F[1]){
 ds[i]=0
   }
else {
 ds[i]=1
  }
for (j in 1:N) {   # N the total number of damage states
 if (u[i]<F[j]){
 ds[i]=j
   }
  }

if (ds[i]==0) {
 L[i]=0.0
 }
else{
 k=ds[i]
 y<-rlnorm(1,meanlog=0,sdlog=1)
 L[i]= lmu[k]+y* lsgm[k]
}
}
print(L)
```

Figure 9.2 Algorithm for the estimation of the distribution of loss given a level of intensity from parametric fragility curves, using lognormal distributed damage-to-loss functions.

10 Identification of the Optimum Vulnerability Curve

The present guidelines urge the analyst to try a number of ground motion intensity measures. Therefore, in some cases, vulnerability and fragility curves corresponding to more than one measure may be found acceptable, i.e., meet the criteria outlined in the goodness of fit sections presented in §6 and §7. The analyst should then identify the optimum curve. Ideally, this can be achieved by integrating the vulnerability and fragility curve with the hazard and selecting the intensity measure which results in the smallest overall uncertainty. Given the difficulty of this procedure, which is under research, the analyst is advised to compare the statistical models with different *IM*s and select the model corresponding to the *IM* that best fits the data. The fit of models containing the same number of parameters, e.g., two GLMs, can be compared through assessment of their AIC values (see §6.3.3).

11 Validation

The procedures presented so far are not capable of assessing the predictive capacity of the optimum fragility or vulnerability curve. The analyst is advised to validate the constructed vulnerability and fragility curves using independent new observations obtained either from the same event or from events corresponding to similar tectonic environments and building inventory. These observations should be of high quality (see §3.2) not less than 30, with levels of intensity in the range of the constructed curves. In the absence of new data, the analyst can comment on the predictive capacity of the model by using cross-validation techniques.

Validation of a fragility or vulnerability curve using new data is performed by plotting the new data together with the mean vulnerability and fragility curve and their 90% prediction intervals. The predictive capacity of the model is established if the new data fall between these intervals.

In cross validation techniques, data points from the existing database are separated into two groups: one group is used to construct the vulnerability or fragility curve and the one is used for the validation of this curve. The size of each group depends on the number of the available observations as well as on the purpose of validation. If the analyst wants to validate the prediction of the model in a single location, then the "leave-one-out" approach is perhaps the best. In this approach, the analyst excludes a single point and fits the selected statistical model to the remaining data. Then, the analyst estimates what is considered a good measure of prediction, e.g., the square prediction error, and repeats the procedure for a number of points. By contrast, if the validation of several locations is required, then a 50-50 or 60-40 (depending on the total number of available points) split of the two groups is needed, and the aforementioned procedure is repeated.

12 Presentation of Constructed Vulnerability and Fragility Functions

The presentation of vulnerability and fragility functions constructed according to the present guidelines should ensure their reproducibility and highlight their reliability. Thus, a comprehensive form (see Appendix A) is provided, which enables the analyst to summarise the characteristics of the constructed vulnerability and fragility functions, as well as establish their reliability by providing important information regarding the quality and quality of the empirical data used, the complexity of the statistical model and the statistical model fitting procedure adopted.

The form consists of two parts:

- Part 1 summarises the main characteristics of the fragility or vulnerability functions, the quality and quantity of the adopted empirical databases and the main characteristics of the intensity measure used.
- In Part 2, the analyst is invited to justify the reliability of the constructed vulnerability and fragility functions by providing information on the adopted database(s) and briefly presenting the main assumptions and techniques adopted for preparing the observations and conducting the statistical modelling. An illustrative example is presented in Appendix A.

13 Final Comments

This guideline document presents a framework for constructing vulnerability and fragility functions by fitting appropriate statistical models to databases of post-earthquake loss or damage observations. It provides a roadmap for undertaking statistical modelling of increasing complexity (when needed) to obtain vulnerability and fragility functions that provide a good fit to the empirical data. Diagnostic tools are provided to aid the analyst determine the goodness of fit of the statistical models to the data, however, it is highlighted that some level of interpretation and subjectivity enter this process. It is advised that multiple ground motion intensity measures be used as explanatory variables, and guidance is provided for the selection of an optimum vulnerability and fragility relationship amongst these. A procedure for validating the optimum relationship using new data is provided, as is a form for reporting the constructed vulnerability and fragility curves. Example applications are also provided in the Appendices that demonstrate a range of issues encountered when dealing with existing empirical damage and loss databases. The applications are limited in simple level 1 approaches which seem to result in models which fit the databases adequately. However, the user should not conclude that there is no place for complex methods in empirical fragility or vulnerability assessment. The need for complex methods is evident in cases where the assumptions such as negligible uncertainty in IM or high data quality need to be relaxed. More research is needed in order to highlight the advantages of complex methods.

Throughout the report, sources of potential uncertainty are identified and discussed. Where possible, methods have been suggested to incorporate these uncertainties in the developed empirical fragility and vulnerability functions. Amongst these is an optional procedure for the incorporation of uncertainty in IM into the vulnerability and fragility relationships. It is acknowledged that the sources of uncertainty considered are not exhaustive, and procedures have not been suggested for dealing with, for example, for over-dispersion in the Bayesian analysis or construction of empirical functions using data affected by multiple hazards (e.g., liquefaction, ground shaking as well as secondary hazards such as landslides or fire). These are areas for further research. It is also acknowledged that the guidance on minimum number of data for a reliable fragility and vulnerability assessment is rather arbitrary and would also benefit from further research.

REFERENCES

Akaike, H. [1974] "A new look at the statistical model identification", Automatic Control, IEEE Transactions on, Vol. 19, No. 6, pp. 716-723.

Alimadad A., Sabidian-Barrera M. [2011] "An outlier-robust fit for generalized additive models with applications to disease outbreak detection", Journal of the American Statistical Association, Vol. 106, No. 494, pp. 719-731.

Allen T. I., Wald D.J., Hotovec A.J., Lin K., Earle P.S., Marano K. [2008] ''An Atlas of ShakeMaps for selected global earthquakes '', Report, U.S. Department of the Interior, U.S.G.S.

Arias A. [1970] ''A measure of earthquake intensity'', Seismic Design for Nuclear Power Plants'', Hansen R.J., MIT Press, Cambridge, Massachusetts.

ATC-13 [1985] ''Earthquake damage evaluation data for California'', Report, Redwood City, Palo Alto, California.

Babyak M.A. [2004] "What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models", Psycosomatic Medicine, Vol. 66, No. 3, pp. 411-421.

Barlett J.E., Kortlik J.W., Higgins C.C. [2001] "Organizational research: determining appropriate sample size in survey research", Information Technology, Learning and Performance Journal Vol. 19, No. 1, pp. 43-50.

Bommer J. J., Magenes G., Hancock J., Penazzo P. [2004] "The influence of strong-motion duration on the seismic response of masonry structures", Bulletin of Earthquake Engineering, Vol. 2, No. 1, pp. 1-26.

Braga F., Dolce M., Liberatore D. [1982] ''Southern Italy November 23, 1980 earthquake: a statistical study of damaged buildings and an ensuing review of the M.S.K.-76 scale'', Report, Rome, Italy.

Brzev S., C. Scawthorn, A.W. Charleson, L. Allen, M. Greene, K. Jaiswal, and V. Silva (2013), GEM Building Taxonomy Version 2.0, GEM Technical Report 2013-02 V1.0.0, 188 pp., GEM Foundation, Pavia, Italy, doi: 10.13117/GEM.EXP-MOD.TR2013.02.

CEQID [2013] Cambridge Earthquake Impact Database, Available from URL: http://www.ceqid.org/CEQID/Home.aspx.

Chandler R., Scott M. [2011] ''Statistical methods for trend detection and analysis in the environmental science'', John Wiley & Sons, London, UK.

Colombi M., Borzi B., Crowley H., Onida M., Meroni F., Pinho R. [2008] "Deriving vulnerability curves using Italian earthquake damage data", Bulletin of Earthquake Engineering, Vol. 6, No. 3, pp. 485-504.

D' Ayala D., Meslem A., Vamvatsikos D., Porter K., Rossetto T., Crowley H., Silva V. [2014] ''Guidelines for analytical vulnerability assessment '', Report, GEM Foundation, Pavia, Italy, Available from URL:

http://www.nexus.globalquakemodel.org/gem-vulnerability/files/gem-vam--guidelines--2014-03-06--final.pdf.

Devalaud E., Cotton F., Akkar S., Scherbaum F., Danciu L., Beauval C., Drouet S, Douglas J., Basili R., Sandikkaya M.A., Segou M., Faccioli E., Theodoulidis N. [2012] "Towards a ground-motion logic tree for probabilistic seismic hazard assessment in Europe", Journal of Seismology, Vol. 16, No. 3, pp. 451-473.

Dhu T., Jones T. [2002] ''Earthquake risk in Newcastle and Lake Macquarie.'' Geoscience Australia Record 200/15, Geoscience Australia, Canberra.Dolce M., Kappos A., Masi A., Penelis G., Vona M. [2006] "Vulnerability assessment and earthquake damage scenarios of the building stock of Potenza (Southern Italy) using Italian and Greek methodologies", Engineering Structures, Vol. 28, No. 3, pp. 357-371.

Douglas J., Cotton F., Abrahamson N.A., Akkar S., Boore D.M., D. A. C. [2013] ''Pre-selection of ground motion prediction equations'', Report, GEM, Pavia, Italy (http://www.nexus.globalquakemodel.org/gem-gmpes/posts/gmpe-project-report-on-pre-selection-of-gmpes).

Edwards M., Griffith M., Wehner M., Lam N., Corby N., Jakab M., Habili N. [2010] ''The Kalgoorie earthquake of the 20th April 2010'', Proceedings of Australian Earthquake Engineering Society Conference, Perth, Australia.Eleftheriadou A.K., Karabinis A.I. [2008] ''Damage probability matrices derived from earthquake statistical data'', Proceedings of 14th World Conference on Earthquake Engineering, Beijing, China.Faenza, L., A. Michelini [2010] "Regression analysis of MCS intensity and ground motion parameters in Italy and its application in ShakeMap", Geophysical Journal International, Vol. 180, No. 3, pp. 1138-1152.

Gamba P., Crowley H., Keller N. [2012] ''Building a global exposure database within the framework of the collaborative Global Earthquake Model (GEM)'', Report, Pavia Foundation, Pavia, Italy.

Garcia D., Wald D.J., Hearne M.G. [2012] "A global earthquake discrimination scheme to optimise ground-motion prediction equation selection", Bulletin of Seismological Society of America, Vol. 102, No. 1, pp. 185-203.

Gelman A., Hill J. [2007] ''Data analysis using regression and multilevel/hierarchical models '', Cambridge University Press, Cambridge, UK.

Graham, J.W. [2009] ''Missing data analysis: making it work in the real world'', Annual Review of Psychology, Vol.60, 549-576.

Green, P. J. [1987] "Penalized Likelihood for General Semi-Parametric Regression Models", International Statistical Review / Revue Internationale de Statistique, Vol. 55, No. 3, pp. 245-259.

Green S.B. [1991] "How many subjects does it take to do a regression analysis?", Multivariate Behavioral Research, Vol. 26, No. 3, pp. 499-510.

Guadagnoli E., Wayne F.V. [1988] "Relation to sample size to the stability of component patterns", Pcychological Bulletin, Vol. 103, No. 2, pp. 265-275.

Harrell F.E., Lee L.K., Matchar D.B., Reichert T.A. [1985] "Regression models for prognostic prediction: Advantages, problems and suggested solutions", Cancer Treatment Reports, Vol. 69, No. 10, pp. 1071-1077.

Harris R.J. [1975] ''A primer of multivariate statistics'', Academic Press, New York, USA.

Hastie T., Tibshirani R. [1986] "Generalised additive models", Statistical Science, Vol. 1, No. 3, pp. 297-318.

Hastie T., Tibshirani R. [1990] ''Generalized additive models'', Chapman and Hall, London, UK.

Ioannou I., Cryssanthopoulos M.K. [2011] ''Probabilistic estimates of seismic loss through a two-stage Monte Carlo approach'', Proceedings of 11th International Conference on Applications of Statistics and Probability in Civil Engineering, Zurich, Swicherland.

Ioannou, I., J. Douglas, T. Rossetto [2015] "Assessing the impact of ground-motion variability and uncertainty on empirical fragility curves", Soil Dynamics and Earthquake Engineering, Vol. 69, No. 0, pp. 83-92.

Ioannou I., Rossetto T. [2014] "Application of statistical models for empirical fragility assessment of buildings." (Under review at Structural Safety)

Ioannou I., Rossetto T., Grant D.N. [2012] ''Use of regression analysis for the construction of empirical fragility curves'', Proceedings of 15th World Conference on Earthquake Engineering, Lisbon, Portugal.

Jaiswal K., Wald W. D., Perkins D, Aspinall W.P., Kiremidjian A.S. [2013] ''Estimating structural collapse fragility of generic building typologies using expert judgement'', Proceedings of 11th International Conference on Structural Safety & Reliability, New York, USA.

Jayaram N., Baker J.W. [2010] "Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment", Earthquake Engineering & Structural Dynamics, Vol. 39, No. 10, pp. 1109-1131.

Joyner W.B., Boore D.M. [1993] "Methods for regression analysis of strong-motion data", Bulletin of Seismological Society of America, Vol. 83, No. 2, pp. 469-487.

Karababa F.S., Pomonis A. [2010] "Damage data analysis and vulnerability estimation following the August 14, 2003 Lefkada Island, Greece, Earthquake", Bulletin of Earthquake Engineering, Vol. online.King S., Kiremidjian A., Sarabandi P., Pachakis D. [2005] ''Correlation of observed building performance with measured ground motion'', Report, Department of Civil and Environmental Engineering, Stanford University USA.

Kruschke J.K. [2011] ''Doing bayesian data analysis A tutorial with R and BUGS'', Elsevier, Oxford, UK.

Lesser V.M., Kalsbeek W.D. [1992] ''Non-sampling error in surveys'', John Wiley & Sons.

Levy P.S., Lemeshow S. [2008] ''Sampling of Populations: Methods and Applications, 4th edition'', John Wiley & Sons, New Jersey, USA.

Liel A.B., Lynch K.P. [2009] "Vulnerability of reinforced concrete frame buildings and their occupants in the 2009 L'Aquila", Natural Hazards Review, Vol. In press.Little R.J.A., B. Rubin D. [2002] ''Statistical analysis with missing data'', Willy Intescience, New Jersey, USA.

Lunn, D., D. Spiegelhalter, A. Thomas, N. Best [2009] "The BUGS project: Evolution, critique and future directions", Statistics in Medicine, Vol. 28, No. 25, pp. 3049-3067.

Margaris B., Skarlatoudis A., Savvaidis A., Theodulidis N., Kalogeras I., Koutrakis S. [2011] ''Strong-Motion networks in Greece and their efficient use in the derivation of regional ground-motion predictive models'', *Earthquake Data in Engineering Seismology*", G. P. Akkar S., Van Eck T.,, Springer, London, UK.

McCullagh P., Nelder J. A. [1989] ''Generalized linear models'', Chapman and Hall, London, UK.

McCullagh P., Nelder J.A. [1989] ''Generalised Linear Model'', Chapman & Hall/CRC.

McInturff, P., W. O. Johnson, D. Cowling, I. A. Gardner [2004] "Modelling risk when binary outcomes are subject to error", Statistics in Medicine, Vol. 23, No. 7, pp. 1095-1109.

Miller L.E., Kunce J.T. [1973] "Prediction and statistical overkill revisited", Measurement and Evaluation in Guidance, Vol. 6, No. 3, pp. 157-163.

Nelder, J. A., R. W. M. Wedderburn [1972] "Generalized Linear Models", Journal of the Royal Statistical Society. Series A (General), Vol. 135, No. 3, pp. 370-384.

Noh H.Y. [2011] ''Fragility function updating using Bayesian framework '', Report, GEM Foundation, Pavia, Italy, Available from URL: http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/fragility-function-updating-using-bayesian-framewo.pdf.

Noh H.Y. [2011b] ''Fragility function estimation using kernel smoothing'', Report, GEM Foundation, Pavia, Italy, Available from URL: http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/fragilityusingkernelsmoothing-haeyoungnoh.pdf .

Noh H. Y., Lignos D.G., Nair K., Kiremidjian A. [2011] "Development of fragility functions as a damage classification/prediction method for steel moment-resisting frames using a wavelet-based damage sensitive feature", Earthquake Engineering & Structural Dynamics, Vol. 10.1002/eqe.1151.

Noh H-Y., Kiremidjian A.S., Lallemant D. [2012] ''Issues related to the development of empirical feragility functions'', Report, GEM Foundation, Pavia, Italy, Available from URL: http://www.nexus.globalquakemodel.org/gem-vulnerability/files/uncertainty/issues-in-empirical-fragility-functions.pdfNoh H-Y., Kiremidjian A.S., Lallemant D. [2013] ''Development of empirical fragility functions using Gaussian Kernel smoothing methods '', Proceedings of 11th International Conference on Structural Safety & Reliability, New York, USA.

Nunnally J.C., Bernstein I.H. [1994] ''Psychometric Theory '', McGraw-Hill, New York, USA.

O'Rourke M.J., So P. [2000] "Seismic fragility curves for on-grade steel tanks", Earthquake Spectra, Vol. 16, No. 4, pp. 801-815.

Pagani M., Monelli D., Weatherill G., Danciu L., Crowley H., Silva V., Henshaw P., Butler L., Nastasi M., Panzeri L., Simionato M. and Vigano D. [2014] "OpenQuake Engine: An open hazard (and risk) software for the Global Earthquake Model," Seismological Research Letters, Vol. 85, No. 3, pp 692-702

Pomonis A., So E. [2012] ''Guidelines for the collection of consequence data'', Report, GEM Foundation, Pavia, Italy, Available from URL:.

Porter K.A., Farokhnia K., Cho I.H., Rossetto T., Ioannou I., Grant D., Jaiswal K., Wald D., D' Ayala D., Meslem A., So E., Kiremidjian A.S., Noh H-Y. [2012] ''Global vulnerability estimation methods for the global earthquake model'', Proceedings of 15th World Conference on Earthquake Engineering.

Porter K., Farokhnia K., , Vamvatsikos D., Cho I. [2014] ''Analytical derivation of seismic vulnerability functions for highrise buildings'', Report, GEM Foundation, Pavia, Italy, Available from URL: http://www.nexus.globalquakemodel.org/gem-vulnerability/files/analytical-structural-vulnerability.

Porter K.A., Kiremidjian A., LeGrue S. [2001] "Assembly-based vulnerability of buildings and its use in performance evaluation", Earthquake Spectra, Vol. 17, No. 2, pp. 291-312.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

Richardson S., Gilks W.R. [1993] "Conditional independence models for epidemiological studies with covariate measurement error", Statistics in Medicine, Vol. 12, No. 18, pp. 1703-1722.

Rossetto T., D' Ayala D., Ioannou I., Meslem A. [2014] ''Evaluation of Existing Fragility Curves'', *SYNER-G: Typology definition and fragility functions for physical elements at seismic risk*", K. C. Pitilakis, H.; Kaynia, A.M., Springer.

Rossetto T., Elnashai A. [2003] "Derivation of vulnerability functions for European-type RC structures based on observational data", Engineering Structures, Vol. 25, No. 10, pp. 1241-1263.

Rossetto T., Elnashai A. [2005] "A new analytical procedure for the derivation of displacement-based vulnerability curves for populations of RC structures", Engineering Structures, Vol. 27, No. 3, pp. 397-409.

Rossetto T., Ioannou I., Grant D.N. [2013] ''Existing empirical fragility and vulnerability relationships: Compendium and guide for selection '', Report, GEM Foundation, Pavia, Italy, Available from URL: http://www.nexus.globalquakemodel.org/gem-vulnerability/posts/existing-empirical-vulnerability-and-fragility-relationships-compendium-and-guide-for-selection.

Rota M., Penna A., Strobbia C.L. [2008] "Processing Italian damage data to derive typological fragility curves", Soil Dynamics and Earthquake Engineering, Vol. 28, No. 10-11, pp. 933-947.

Rynn J.M., Brennan E., Hughes P.R., Pedersen I.S., Stuart H.J. [1992] "The 1989 Newcastle, Australia, Earthquake: The facts and the misconceptions", Bulletin of New Zealand National Society of Earthquake Engineering, Vol. 25, No. 2, pp. 77-144.Sabidian Barrera M., Cubranic D., Alimadad A. [2014] ''Robust Generalized Additive Models'', Report, CRAN.

Schafer J.L. [1999] "Multiple imputation: a primer", Statistical Methods in Medical Research, Vol. 8, pp. 3-15.

Shinozuka M., Feng M.Q., Lee J., Naganuma T. [2000] "Statistical Analysis of Fragility Curves", Journal of Engineering Mechanics, Vol. 126, No. 12, pp. 1459-1467.

Silva V., Crowley H., Pagani M., Monelli D., Pinho R. [2013] “Development of the OpenQuake engine, the Global Earthquake Model’s open-source software for seismic risk assessment” Natural Hazards, DOI: 10.1007/s11069-013-0618-x.

So E., Pomonis A. [2012] "Events in the GEM earthquake consequences database." GEM Technical Report 2011 – 1, GEM Foundation, Pavia, Italy, Available from URL:

Stafford, P.J. [2012] “Evaluation of structural performance in the immediate aftermath of an earthquake: a case study of the 2011 Christchurch earthquake,” International Journal of Forensic Engineering, Vol. 1, No. 1, pp. 58-77

Stewart J.P., Douglas J., Javanbarg M., Abrahamson N.A., Bozorgnia Y., Boore D.M., Campbell K.W., Delavaud E., Erdik M., Stafford P.J. [2013a] "Selection of ground motion prediction equations for the global earthquake model", Earthquake Spectra, Vol. in press.

Stewart J.P., Douglas J., Javanbarg M., Di Alessandro C., Bozorgnia Y., Abrahamson N.A., Boore D.M., Campbell K.W., Delavaud E., Erdik M. and Stafford P.J. [2013b] “PEER 2013/22 - GEM-PEER Task 3 Project: Selection of a Global Set of Ground Motion Prediction Equations”, Available from URL: http://www.globalquakemodel.org/resources/publications/technical-reports/selection-global-GMPEs/.

Straub D., Der Kiureghian A. [2008] "Improved seismic fragility modeling from empirical data", Structural Safety, Vol. 30, No. 4, pp. 320-336.

Trifunac, M. D., A. G. Brady [1975] "On the correlation of seismic intensity scales with the peaks of recorded strong ground motion", Bulletin of the Seismological Society of America, Vol. 65, No. 1, pp. 139-162.

United Nations [1964] ''Recommendations for the Preparation of Sample Survey Reports (Provisional Issue)'', Report, Statistical Papers, Series C, ST/STAT/SER.C/1/Rev.2, New York, USA.

United Nations [1982] ''National household survey capability programme. Non-sampling errors in household survyes: sources, assessment and control'', Report, Department of Technical Co-Operation for Development and Statistical Office, New York, USA.

Wald D.J., Quitoriano V., Worden C.B., Hopper M., Dewey J.W. [2011] "USGS 'Did you feel it?' Internet-based Macroseismic intensity maps", Annals of Geophysics, Vol. 54, No. 6, pp. 688-709.

Wald D. J.,Quitoriano V., Heaton T. H., Kanamori H. [1999] "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., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, C. B. Worden [1999b] "TriNet “ShakeMaps”: Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California", Earthquake Spectra, Vol. 15, No. 3, pp. 537-555.

Wand M.P., Jones M.C. [1995] ''Kernel Smoothing'', Chapman & Hall.

Wesson R.L., Perkins D.M., Leyendecker E.V., Roth R.J., Petersen M.D. [2004] "Losses to single-family housing from ground motions in the 1994 Northridge, California, earthquake", Earthquake Spectra, Vol. 20, No. 3, pp. 1021-1045.Wood S. [2006] ''Generalized Additive Models: an introduction to R'', Chapman and Hall/CRC, USA.

Wood S. [2014] ''Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation'', Report, CRAN.

Worden C. B., Wald D. J., Allen T. I., Lin K., Garcia D., Cua G. [2010] "A Revised Ground-Motion and Intensity Interpolation Scheme for ShakeMap", Bulletin of the Seismological Society of America, Vol. 100, No. 6, pp. 3083-3096.

Zuur A.F., Ieno E.N., Walker N.J., Saveliev A.A., Smith G.M. [2009] ''Mixed effect models adn extensions in ecology with R'', Springer, New York, USA.

APPENDIX A Form for Presenting and Evaluating New Empirical Vulnerability and Fragility Functions

The form which should be used to present the results is presented in Table A.1.

Table A.1 Presentation form.

Vulnerability and Fragility Specification for [Building Class] [Name]	
Developer, Affiliation and Date:	
Statistical Package:	
Selected Building Class:	
Type of Assessment:	☐ Direct Vulnerability ☐ Indirect Vulnerability ☐ Fragility
Number of Buildings per Class:	
Sources of Data:	
Overall Rating of Quality of Data (per Source):	☐ High ☐ Moderate ☐ Poor ☐ Other(____)
Definition of Loss Parameter:	
Intensity Measure (IM):	
Range of IM:	
Evaluation of IM:	☐ Ground motion records (Number) ☐ GMPE (Ref) ☐ ShakeMap ☐ Other(____)
Mean Direct or Indirect Vulnerability Relationship	
Function of GLM:	
Parameter (ϑ_0):	
Parameter (ϑ_j):	
Parameter (ϑ_N):	
Shape of Nonparametric Curve:	(Attach table with (*x*,*y*) values and plot it in Figure A.5)
Confidence and Prediction Intervals of the Vulnerability Relationship	
Confidence Intervals for Mean Curve:	(Attach table with (*x*,*y*) values)
Prediction Intervals for Model:	(Attach table with (*x*,*y*) values)
Direct or Indirect Vulnerability Assessment	
Statistical Model:	☐ GLM ☐ GAM ☐ Kernel Smoother
Model Fitting Procedure:	☐ Maximum Likelihood ☐ Robust Maximum Likelihood ☐ Bayesian ☐ Other(____)
Type of Data:	☐ A (building-by-building) ☐ B (grouped data)
Number of Data Points:	
Grouped Data: Definitions of	

aggregated unit:					
Grouped Data: Min Number of Buildings | Data Point:					
Statistical Assumptions:	☐ Independence of observations ☐ Measurement error in IM ☐ Measurement error in response				
Goodness of Fit Assessment (GLM or GAM):	☐ Acceptable mean function ☐ Acceptable variance function				
Procedure for the construction of:					
Confidence Intervals:	☐ Asymptotic ☐ Bootstrap ☐ Bayesian (______)				
Prediction Intervals:	☐ Asymptotic ☐ Bootstrap ☐ Bayesian (______)				
Indirect Vulnerability Assessment - Fragility Curves (FC)					
Damage Scale:					
Damage State (DS):	***ds_1***	**…**	***ds_n***	**…**	***ds_n***
Description of DS:					
Function of Parametric FC:					
Parameter (λ):					
Parameter (ζ):					
Shape of Nonparametric FC:					
Confidence Intervals for mean FC:					
Statistical Model:					
Model Fitting Procedure:					
Data Type:	☐ A (building-by-building) ☐ B (grouped data)				
Number of Data Points:					
Grouped Data: Definition of the aggregation unit:					
Grouped Data: Min Number of Buildings / Data Point:					
Fitting Assumptions:					
Measurement Error in IM:					
Measurement Error in Response:					
Other(____________________):					
Goodness of Fit Assessment:					
Mean function:					
Variance function:					
Procedure: Confidence Intervals:	☐ Asymptotic ☐ Bootstrap ☐ Bayesian	☐ Asymptotic ☐ Bootstrap ☐ Bayesian	☐ Asymptotic ☐ Bootstrap ☐ Bayesian	☐ Asymptotic ☐ Bootstrap ☐ Bayesian	☐ Asymptotic ☐ Bootstrap ☐ Bayesian

Indirect Vulnerability Assessment - Damage-to-Loss Functions						
Source of Damage to Loss Functions:						
Damage Scale:						
Damage State (*DS*):	ds_0	ds_1	…	ds_i	…	ds_n
Shape of damage to loss function distribution:						
Parameter ϑ_1 of the damage-to loss distribution:						
Parameter ϑ_2 of the damage-to loss distribution:						
Parameter ϑ_n of the damage-to loss distribution:						
Procedure:						
Discussion:						
***Data Quality Assessment*:**						
Data Form:						
Data Collection Technique:						
Data preparation (for each data source):						
Procedure:						
Assumptions:						
Systematic errors eliminations:						
Quality of each data source:						
Vulnerability Assessment Procedure:						
Statistical Model Fitting Technique:						
Data points:						
Goodness of fit:						

The form should include the following figures and tables:

- Figure A.1 Pie chart showing the % of total number of buildings of the examined class per source.
- Figure A.2 Histograms showing the number of buildings per damage state per source.
- Figure A.3 Cumulative distribution of the proportion of the examined buildings in each data point against the corresponding intensity measure.
- Figure A.4 Fragility curves with, e.g., 90%, confidence intervals.
- Figure A.5 Vulnerability function for given intensity measure levels with confidence and prediction intervals (e.g., 90%).

Table A.2 Nonparametric confidence intervals for fragility curves

Intensity	DS1		DS2		DS3		DS4		DS4	
	5%	95%	5%	95%	5%	95%	5%	95%	5%	95%
			…	…	…	…	…	…	…	…
…	…	…								

Table A.3 Nonparametric confidence and prediction intervals for the vulnerability curve(s).

Loss (%)	Intensity Measure Level	
	5%	95%
0.01	…	…
0.02		
…		

APPENDIX B Fitting GLM to damage data from the 1980 Irpinia Earthquake, Italy

Vulnerability Curve for Italian Field Stone Masonry Buildings with Wooden Floors.					
Developer and Date:	I. Ioannou, UCL EPICentre, 01/04/12				
Statistical Package:	R (R Development Team, 2008)				
Selected Building Class:	Field stone masonry buildings with wooden floors				
Type of Assessment:	☐ Direct Vulnerability ■ Indirect Vulnerability ■ Fragility				
Number of Buildings per Class:	8,859				
Sources of Data:	1980 Irpinia Earthquake (reported in Braga et al (1982) and obtained from CEQID (2013))				
Overall Rating of Quality of Data (per Source):	☐ High ☐ Moderate ☐ Poor ☐ Other(____)				
Definition of Loss Parameter:	Repair divided by replacement cost				
Intensity Measure (IM):	PGV in m/s				
Range of IM:	(0.08-1.3) m/s				
Evaluation of IM:	☐ Ground motion records (Number) ☐ GMPE (Ref) ■ ShakeMap ☐ Other(____)				
Mean Indirect Vulnerability Relationship					
Shape of Nonparametric Curve:	Mean damage factor against PGV plotted in Figure B.7.				
Confidence Intervals of the Vulnerability Relationship					
Confidence Intervals:	Mean damage factor plus one sigma against PGV presented in Figure B.7. The uncertainty in the estimation of the mean fragility curves is ignored.				
Indirect Vulnerability Assessment - Fragility Curves (FC)					
Damage Scale:	MSK-76				
Damage State (DS):	***ds***$_1$	***ds***$_2$	***ds***$_3$	***ds***$_4$	***ds***$_5$
Description of DS:	**Insignificant**	**Slight**	**Moderate**	**Severe**	**Collapse**
Function of Parametric FC:	Eq.(B.3)	Eq.(B.3)	Eq.(B.3)	Eq.(B.3)	Eq.(B.3)
Parameter (ϑ_0):	1.97982	0.80121	0.19432	-0.67497	-0.8513
Parameter (ϑ_1):	0.39727	0.45440	0.60559	0.47808	0.7458
Confidence Intervals for FC:	√	√	√	√	√
Statistical Model:	■ GLM	■ GLM	■ GLM	■ GLM	■ GLM
Model fitting Procedure:	■ ML	■ ML	■ ML	■ ML	■ ML
Data Type:	☐ Building-by-Building ■ Grouped data				
Number of Data Points:	41	41	41	41	41
Grouped Data: Definition of the aggregation unit:	Municipality (42 km^2 on average)				
Grouped Data: Min Number of	3	3	3	3	3

Buildings / Data Point:						
Model Assumptions:						
Measurement Error in IM:	-		-	-	-	-
Measurement Error in Response:	-		-	-	-	-
Other(____________________):	-		-	-	-	-
Goodness of Fit Assessment:						
Mean function:	√		√	√	√	√
Variance function:	No		No	No	No	No
Procedure:						
Confidence Intervals:	□ Asymptotic		□ Asymptotic	□ Asymptotic	□ Asymptotic	□ Asymptotic
Indirect Vulnerability Assessment - Damage-to-Loss Functions						
Source of Damage to Loss Functions:	Dolce et al (2006) Damage to Loss Functions					
Damage Scale:						
Damage State (*DS*):	ds_0	ds_1	ds_2	ds_3	ds_4	ds_5
Mean of damage-to-loss distribution (μ):	0.005	0.035	0.145	0.305	0.800	0.950
Standard deviation of damage-to-loss distribution (σ):	0.035	0.043	0.056	0.111	0.113	0.060
Procedure:	Damage to Loss functions obtained from observational data from different Italian building classes.					

Discussion:

An indirect vulnerability assessment procedure is adopted here in order to estimate the economic losses caused by direct damage to the building class of field stone masonry with wooden floors. The procedure requires six steps. In the first three steps, a set of empirical fragility curves for the selected building class is identified from the damage data obtained from the 1980 Irpinia Earthquake, Italy. An existing set of Italian damage-to-loss functions are then used in order to transform the fragility curves to vulnerability functions for the range of the selected intensity measure.

1st Step: Preparation of the damage data

***Data Quality Assessment*:**

Grouped damage data, aggregated per building type into 41 damage distributions each corresponding to a different affected municipality, were obtained from CEQID (2013). The original completed survey forms were not available but the survey method is reported in the literature. The 1980 Irpinia Earthquake database was constructed by a one-stage cluster sampling method (Levy and Lemeshow, 2008); i.e., the total number of buildings from 41 municipalities (out of more than 600 affected by the event) in the Campania-Basilicata area were surveyed (Braga et al., 1982). With regard to non-sampling errors, the comments on the data collection found in CEQID (2013) raise concerns on whether the total number of buildings in each commune has been surveyed. For the needs of this application we assume that this error is negligible. Overall, the quality of the database is considered moderate due to the aggregated type of data.

Data preparation per source:

The building classes account for the material of the vertical as well as the lateral load resisting structural components. Fragility curves are constructed using the largest and most vulnerable building class in this database, which consists of 8,859 field stone masonry buildings with wooden floors. The observed damage is classified into six discrete states, varying from no damage to collapse, according to MSK-76. This is the original damage classification used by Braga et al.

(1982) in collecting the data. Figure B.1 highlights the significant (~25%) percentage of buildings which suffered heavy damage or collapse.

The construction of fragility curves using 'R' (R Development Team, 2008) requires the transformation of the grouped damage data into data points (x_j,(y_{ij},n_j-y_{ij})), where y_{ij} is the count of buildings which suffered $DS \geq ds_i$ and n_j-y_{ij} is the count of buildings which sustained $DS<ds_i$ for municipality j with intensity measure level x_j. Thus, 41 data points are obtained for each of the five damage states ds_i (i=1-5). The number of the buildings surveyed in each municipality varied widely from 3 to 1205. Six data points are seen to be based on very small numbers of buildings (<20). These six points are also included in the analysis and the goodness of fit diagnostics will determine whether they should be removed.

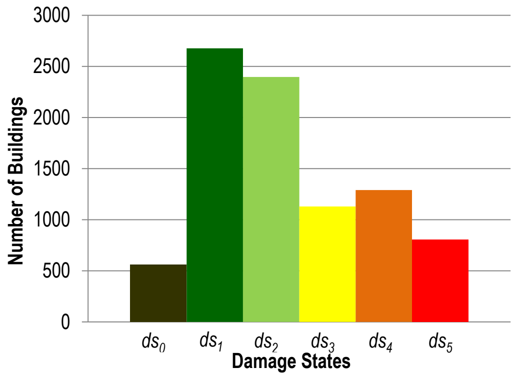


Figure B.1 Number of field stone masonry buildings with wooden floors that suffered damage in the 1980 Irpinia Earthquake.

2nd Step: Selection and Estimation of the Intensity Measure

Two intensity measure types, namely PGV and MMI, have been selected. Their levels are estimated by a ShakeMap for the earthquake and are also provided in the CEQID (2013). The intensity measure values are assumed to have a single constant value within each municipality. This is considered a reasonable assumption given the relatively small surface area of each municipality (on average 24km^2). Nonetheless, the measurement error associated with these estimated intensity levels is not known, and therefore the measurement error in the intensity measure level estimated for each commune is ignored.

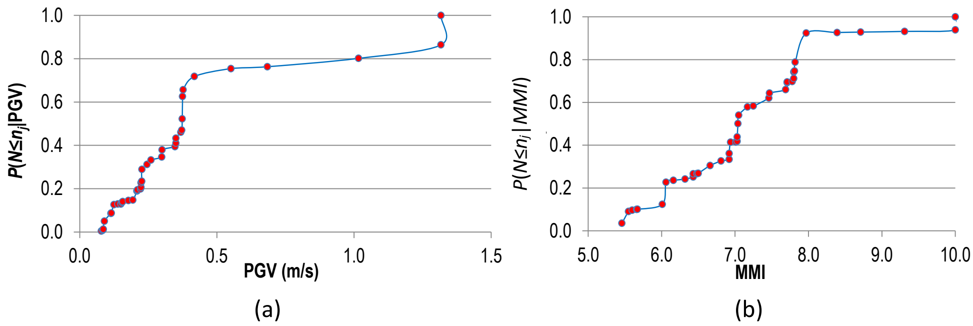


Figure B.2 Cumulative distribution of the proportion of the examined buildings exceeding the selected intensity measure values.

3rd Step: Selection of Statistical Model

The distribution of the data points in the range of the intensity measure levels can be used to determine an acceptable statistical model. For this reason, the cumulative proportion of buildings (for the 41 data sets) and their corresponding intensity measure levels in shown Figure B.2a,b. These figures show that the majority of buildings are clustered in low-to-intermediate intensity measure levels (i.e., pgv<0.5 m/s or MMI<8). Given the relatively small number of data points and especially their sparseness over the higher intensity measure levels, a non-parametric regression model is expected to depict trends specific for this database (more details in Ioannou et al. 2012). Hence, in an attempt to capture the overall trend of the data points, a set of GLMs is selected according to the recommendations of §6.2.

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\binom{n_j}{y_j}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \quad \text{where, } \mu=P\left(DS\geq ds_i \mid IM\right)=\frac{1}{1+\exp\left(\vartheta_0+\vartheta_1 IM\right)} \tag{B.1}$$

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\binom{n_j}{y_j}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \quad \text{where, } \mu=P\left(DS\geq ds_i \mid IM\right)=\Phi\left(\vartheta_0+\vartheta_1 IM\right) \tag{B.2}$$

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\binom{n_j}{y_j}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \quad \text{where, } \mu=P\left(DS\geq ds_i \mid IM\right)=\Phi\left(\vartheta_0+\vartheta_1 \log\left(IM\right)\right) \tag{B.3}$$

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\binom{n_j}{y_j}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \quad \text{where, } \mu=P\left(DS\geq ds_i \mid IM\right)=1+\exp\left[-\exp\left(\vartheta_0+\vartheta_1 IM\right)\right] \tag{B.4}$$

The selected GLM models assume that the counts of buildings suffering a given damage state or above for a specific intensity measure level follow a binomial distribution, and that the mean of this distribution is related to the intensity measure through the three link functions (see Table 6.2): namely: logit (Eq.(B.1)), probit (Eq.(B.2) and Eq.(B.3)) and complementary log-log (Eq.(B.4)).

4th Step: Statistical Analysis

Estimation of the GLM model parameters

Statistical model fitting is performed using 'R' (R Development Team, 2008). The parameters of this model are estimated numerically by maximising their likelihood function (see §6.3.1). The curves fit for the 5 damage states and two intensity measure types, for each link function are depicted in Figure B.3a,b. In Figure B.3a, the three models, expressed by Eq.(B.1), Eq.(B.2) and Eq.(B.4), appear to have negligible differences, indicating that the selection of link function is not important. By contrast, there is a notable difference in the fit of the model expressed by Eq.(B.3). This highlights the significant influence of the transformation of the intensity measure (i.e., from PGV to log(PGV)). We consider that the Eq.(B.3) provides a more realistic model as its systematic component provides probability estimates which tend to zero for as the intensity measure levels tend to zero. For MMI, all four models appear to have negligible differences indicating that the fit is not sensitive to the shape of the link function or the transformation of the explanatory variable.

Goodness of fit checks:

A set of graphical plots is used here to diagnose the ability of the selected statistical models to fit the available data. For PGV, the plots of the Pearson residuals for ds_4, obtained for the probit models using Eq.(B.2) and Eq.(B.3), against the fitted values are plotted in Figures B.4a,b, respectively. In general, a statistical model can be considered acceptable if the Pearson residuals have zero mean and a constant variance equal to 1. The mean function, μ, as expressed by Eq.(B.3) appears to be a better fit of the data, as the smoothing curve is reasonably close to zero. By contrast, for all models there is a significant number of residuals which are outside the 99% confidence interval [-3,3] for both models, indicating that the variance of the Pearson residuals is greater than 1. This over-dispersion highlights that the assumption of the binomial distribution is grossly violated in the models.

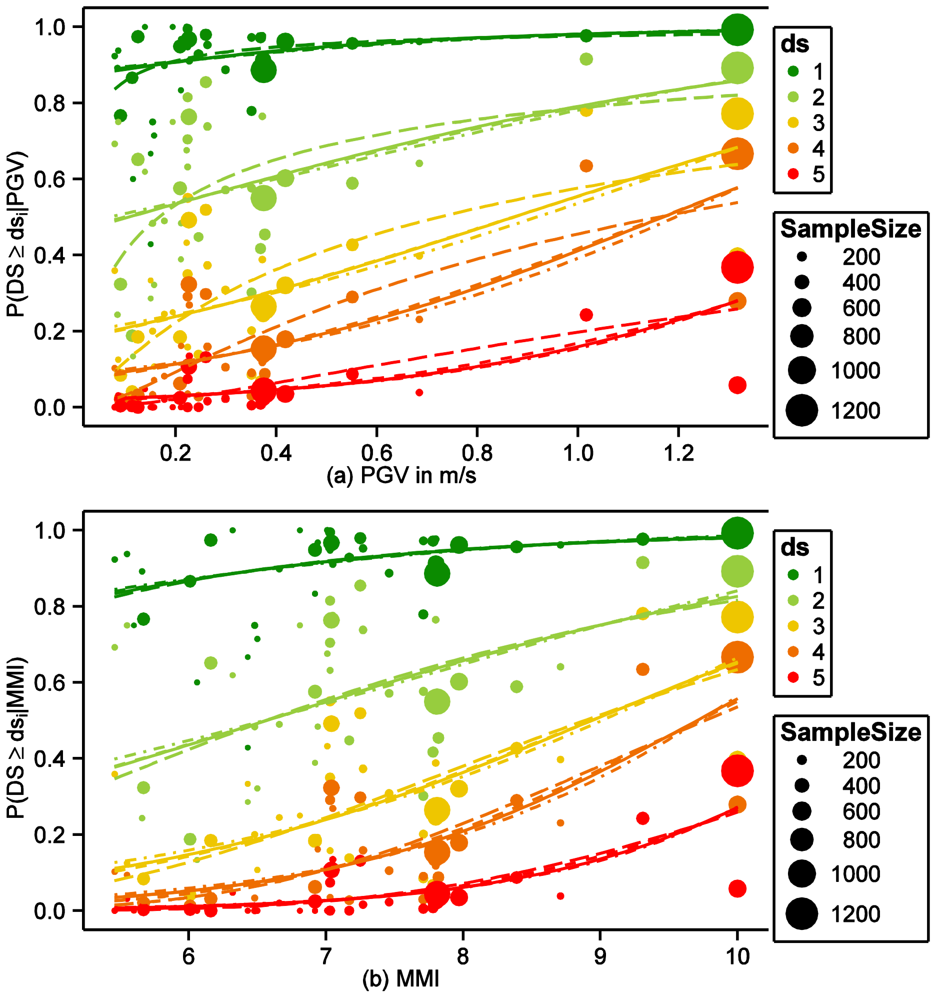
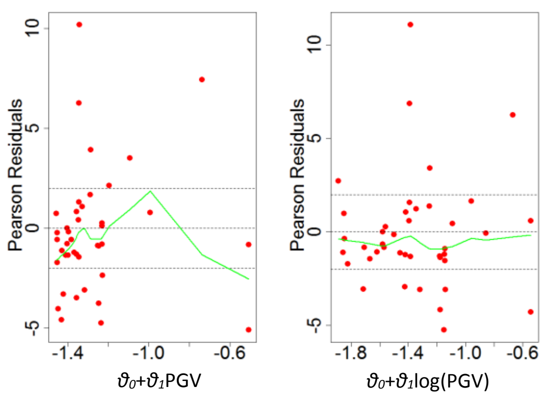


Figure B.3 Fragility curves (continuous: Eq.(B.1), dashed Eq.(B.2), longdash: Eq.(B.3), dotdash: Eq.(B.4)) corresponding to the 5 damage states expressed in terms of the five regression models for (a) PGV and (b) MMI.



(a) (b)

Figure B.4 Pearson residuals against the fitted values for fragility curves corresponding to *ds_4* for PGV using the probit link function expressed by (a) Eq.(B.2) and(b) Eq.(B.3).

The observed over-dispersion could be addressed by:

- *Removing the influential points.* Potential influential points are identified by plotting the Cook's distance against the data point. Then, these points are removed and the plotted again.

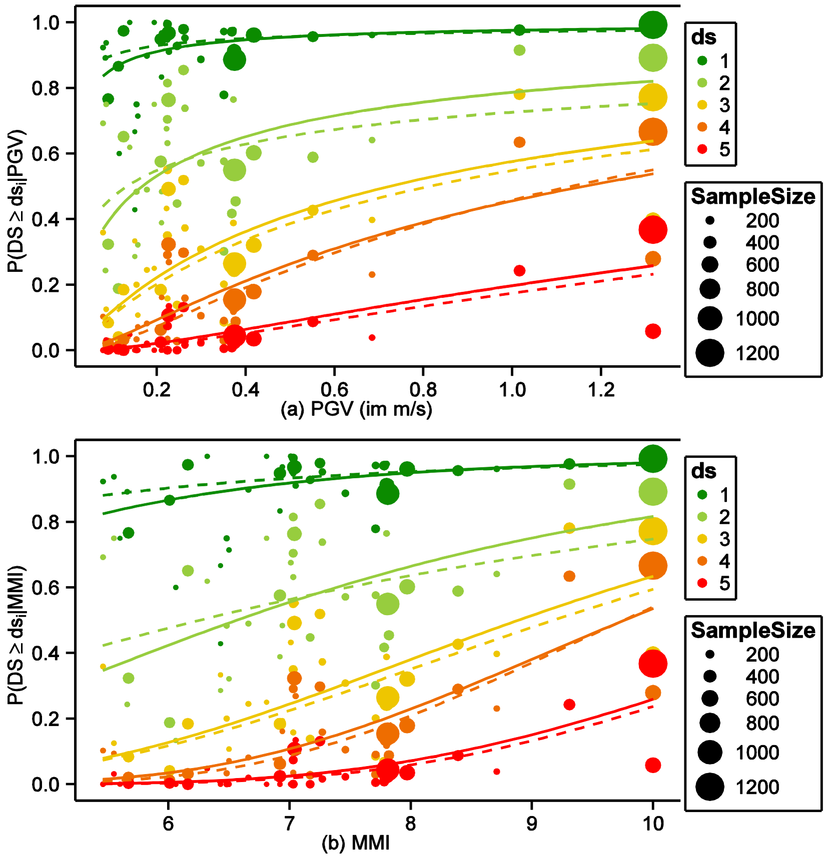


Figure B.5 Fragility curves (continuous lines) for (a) PGV and (b) MMI using Eq.(3) if the potential outliers are not removed (solid lines) and removed (dashed lines).

Figure B.5 shows that the removal of the 3 points (not necessarily the same for all curves) with the higher Cook's Distance does not lead to significant differences in the fragility curves.

- *Adding missing explanatory variables*. However, this cannot be addressed given the lack of additional independent explanatory variables in the database.

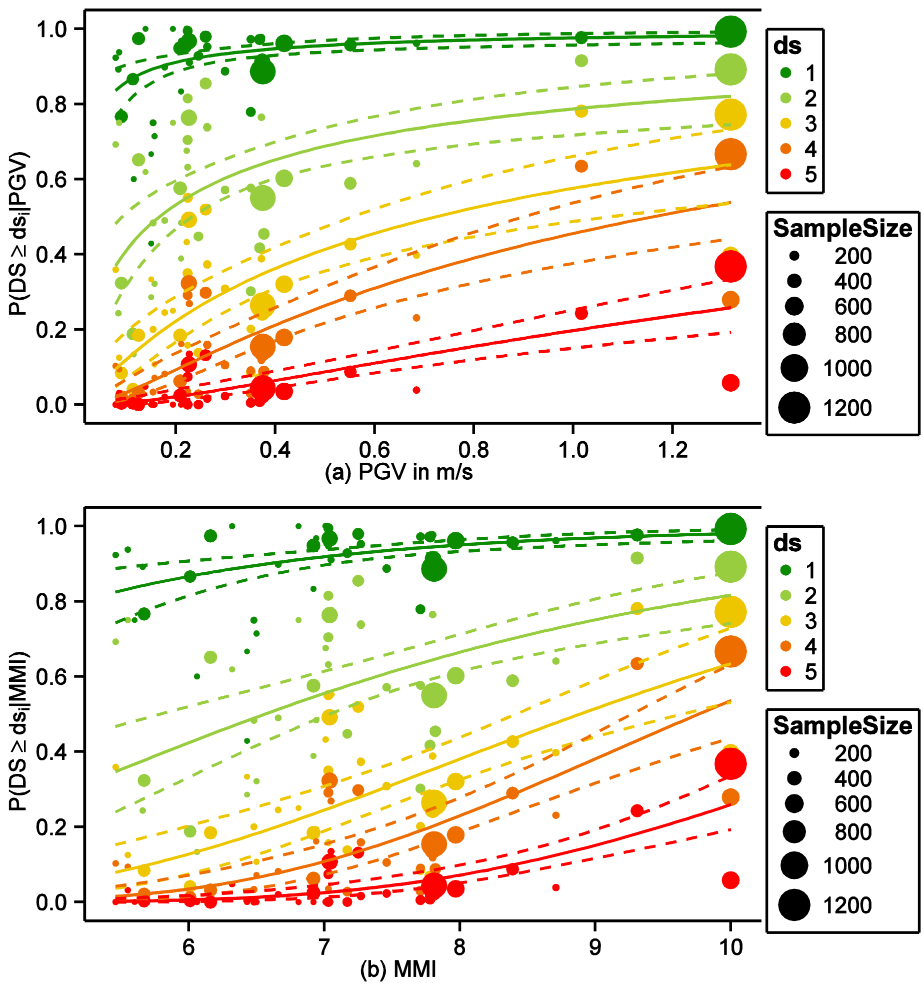


Figure B.6 Fragility curves (continuous lines) for (a) PGV and (b) MMI using Eq.(3) and their corresponding 90% confidence intervals (dashed lines).

Given our inability to improve the fit of the selected model by increasing its complexity, over-dispersion is addressed by the use of the quasi-binomial distribution (see §6.3.1). The use of this distribution leads to the same mean fragility curves but produces wider confidence intervals which reflect the uncertainty in the grouped data. Figure B.6a,b depicts the mean fragility curves as well as the 90% confidence intervals of the fragility curves corresponding to the five damage states for the two intensity measure types. It can be seen that the 90% intervals of the curves corresponding to ds_3 and ds_4 overlap for higher values of intensity measure levels that might have an impact on the propagation of the overall uncertainty to the vulnerability. It is ignored for the purposes of this application. The values for the upper and lower limit for the fragility curves corresponding to the 5 damage states can be seen in Table B.2.

Table B.1 AIC values for the fitted statistical models.

Model	ds_1		ds_2		ds_3		ds_4		ds_5	
	PGV	MMI	PGV	MMI	PGV	MMI	PGV	MMI	PGV	MMI
Eq.(B.3)	480	481	1092	1101	1226	1245	962	983	607	613

The model expressed by Eq.(B.3) has been considered the be the most realistic model. Which of the two intensity measure types provide the best fit? According to the recommendations of section 10, the AIC for the models corresponding to the two IMs for ach damage states are compared. The AIC of the fragility curves for PGV is smaller than their counterparts for MMI, therefore the set of fragility curves for PGV is considered optimum (see Table B.1).

5th Step: Selection of Appropriate Damage-to-Loss Functions

The transformation of the fragility curves in vulnerability functions requires the selection of damage-to-loss functions. The mean and variance of the damage factor for 6 damage states proposed by Dolce et al. (2006) is selected here.

6th Step: Identify the optimum Vulnerability Curve

The first two moments of the damage factor for the range of PGV are estimated using Eq.(9.3) and Eq.(9.4) having estimated the damage probability matrices from Eq.(6.2). The mean and mean plus one standard deviation vulnerability curves are plotted in Figure B.7. This figure depicts the substantial uncertainty that exists in the damage factor.

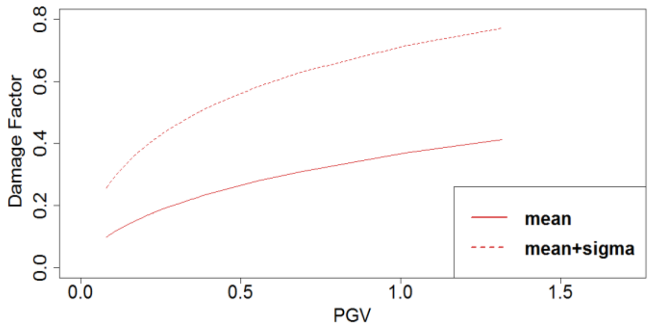


Figure B.7 Mean and mean plus one standard deviation vulnerability curve for PGV.

Table B.2 Nonparametric confidence intervals for the 5 fragility curves corresponding to the 41 data points.

IM	$P(DS \geq ds_1 \mid IM)$		$P(DS \geq ds_2 \mid IM)$		$P(DS \geq ds_3 \mid IM)$		$P(DS \geq ds_4 \mid IM)$		$P(DS \geq ds_5 \mid IM)$	
m/s^2	95%	5%	95%	5%	95%	5%	95%	5%	95%	5%
0.09	84.64%	91.61%	41.32%	13.88%	25.67%	25.67%	5.13%	11.98%	0.98%	3.43%
0.09	84.75%	91.65%	41.49%	55.73%	13.99%	25.77%	5.19%	12.06%	0.99%	3.46%
0.10	84.89%	91.69%	41.70%	55.85%	14.13%	25.90%	5.27%	12.16%	1.01%	3.50%
0.10	85.01%	91.73%	41.88%	55.96%	14.25%	26.01%	5.34%	12.24%	1.03%	3.53%
0.11	85.10%	91.76%	42.01%	56.04%	14.34%	26.09%	5.39%	12.30%	1.04%	3.55%
0.13	85.73%	91.97%	43.01%	56.64%	15.03%	26.71%	5.77%	12.79%	1.13%	3.72%
0.13	85.93%	92.04%	43.32%	56.83%	15.24%	26.90%	5.89%	12.94%	1.16%	3.77%
0.14	86.12%	92.10%	43.62%	57.01%	15.46%	27.10%	6.02%	13.10%	1.19%	3.83%
0.16	86.56%	92.26%	44.36%	57.46%	15.99%	27.57%	6.33%	13.47%	1.27%	3.96%
0.17	86.84%	92.37%	44.84%	57.76%	16.34%	27.88%	6.53%	13.72%	1.32%	4.05%
0.17	86.84%	92.37%	44.84%	57.76%	16.34%	27.88%	6.53%	13.72%	1.32%	4.05%
0.17	87.00%	92.42%	45.10%	57.92%	16.53%	28.05%	6.65%	13.86%	1.35%	4.10%
0.17	87.01%	92.43%	45.12%	57.93%	16.55%	28.07%	6.66%	13.88%	1.35%	4.11%
0.18	87.23%	92.51%	45.50%	58.17%	16.83%	28.32%	6.83%	14.08%	1.39%	4.18%
0.21	87.95%	92.81%	46.82%	59.00%	17.83%	29.20%	7.45%	14.81%	1.56%	4.45%
0.21	87.96%	92.81%	46.83%	59.01%	17.85%	29.22%	7.45%	14.82%	1.56%	4.45%
0.23	88.32%	92.96%	47.51%	59.45%	18.38%	29.69%	7.79%	15.21%	1.65%	4.60%
0.23	88.42%	93.01%	47.70%	59.57%	18.53%	29.82%	7.88%	15.32%	1.67%	4.64%
0.24	88.63%	93.10%	48.12%	59.85%	18.86%	30.12%	8.10%	15.57%	1.73%	4.73%
0.24	88.66%	93.12%	48.18%	59.89%	18.92%	30.16%	8.13%	15.60%	1.74%	4.74%
0.25	88.75%	93.16%	48.37%	60.01%	19.07%	30.29%	8.23%	15.72%	1.77%	4.78%
0.25	88.80%	93.18%	48.45%	60.07%	19.14%	30.36%	8.28%	15.77%	1.78%	4.80%
0.25	88.82%	93.19%	48.49%	60.09%	19.17%	30.39%	8.30%	15.79%	1.78%	4.81%
0.27	89.32%	93.44%	49.54%	60.79%	20.04%	31.15%	8.87%	16.44%	1.94%	5.06%
0.28	89.42%	93.49%	49.75%	60.94%	20.21%	31.31%	8.99%	16.58%	1.98%	5.11%
0.32	90.17%	93.89%	51.41%	62.10%	21.66%	32.61%	9.98%	17.69%	2.26%	5.54%
0.32	90.21%	93.92%	51.50%	62.17%	21.74%	32.68%	10.04%	17.75%	2.28%	5.56%
0.36	90.96%	94.40%	53.37%	63.55%	23.45%	34.24%	11.27%	19.12%	2.65%	6.09%
0.37	91.03%	94.45%	53.54%	63.69%	23.62%	34.40%	11.40%	19.25%	2.69%	6.15%
0.37	91.03%	94.45%	53.54%	63.69%	23.62%	34.40%	11.40%	19.25%	2.69%	6.15%
0.38	91.27%	94.62%	54.18%	64.19%	24.23%	34.97%	11.85%	19.76%	2.84%	6.35%
0.39	91.36%	94.68%	54.42%	64.38%	24.47%	35.19%	12.03%	19.95%	2.89%	6.43%
0.39	91.41%	94.73%	54.58%	64.50%	24.62%	35.33%	12.15%	20.08%	2.93%	6.48%
0.39	91.43%	94.74%	54.61%	64.53%	24.66%	35.36%	12.17%	20.11%	2.94%	6.49%
0.39	91.44%	94.75%	54.65%	64.56%	24.69%	35.39%	12.20%	20.14%	2.95%	6.50%
0.44	92.06%	95.25%	56.46%	66.08%	26.53%	37.16%	13.64%	21.73%	3.42%	7.14%
0.57	93.38%	96.55%	60.92%	70.40%	31.54%	42.41%	17.94%	26.60%	4.99%	9.20%
0.74	94.69%	97.97%	66.13%	76.59%	38.29%	50.77%	24.59%	34.86%	7.86%	13.04%
1.00	96.09%	99.20%	72.46%	84.93%	47.64%	64.27%	35.19%	49.90%	13.45%	21.60%
1.29	97.23%	99.76%	78.30%	91.77%	57.30%	77.99%	47.36%	67.48%	21.19%	35.48%
1.29	97.23%	99.76%	78.30%	91.77%	57.30%	77.99%	47.36%	67.48%	21.19%	35.48%

Lessons

The aggregation of damage data, even over reasonably small geographical units, may lead to significant loss of information. In the absence of more information which can increase the complexity of the model, the resulted over-dispersion can be easily modelled by the generalised linear models.

References:

Braga F., Dolce M., Liberatore D. [1982] ''Southern Italy November 23, 1980 earthquake: a statistical study of damaged buildings and an ensuing review of the M.S.K.-76 scale'', Report, Rome, Italy.

Dolce M., Kappos A., Masi A., Penelis G., Vona M. [2006] "Vulnerability assessment and earthquake damage scenarios of the building stock of Potenza (Southern Italy) using Italian and Greek methodologies", Engineering Structures, Vol. 28, No. 3, pp. 357-371.

Ioannou I., Rossetto T., Grant D.N. [2012] ''Use of regression analysis for the construction of empirical fragility curves'', Proceedings of 15th World Conference on Earthquake Engineering, Lisbon, Portugal.

Levy P.S., Lemeshow S. [2008] ''Sampling of populations: Methods and applications'', Wiley.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

APPENDIX C Fitting GLMs to damage data from two earthquakes affecting Christchurch, New Zealand, in 2010-2011

Fragility Curves for New Zealand Unreinforced Masonry Buildings.	
Developer and Date:	I. Ioannou, UCL EPICentre, J. Ingham, Auckland University, M. Griffith and L. Moon, University of Adelaide, 9/012/13
Statistical Package:	R (R Development Team, 2008)
Selected Building Class:	Unreinforced masonry (URM)
Type of Assessment:	☐ Direct Vulnerability ☐ Indirect Vulnerability ☒ Fragility
Number of Buildings per Class:	262
Sources of Data:	Survey for the needs of a PhD project
Overall Rating of Quality of Data (per Source):	☐ High ☐ Moderate ☐ Poor ☒ Other(*Questionable*)
Definition of Loss Parameter:	-
Intensity Measure (IM):	PGA in g
Range of IM:	1st event: (0.18-0.41)g, 2nd event: (0.09-1.0) g
Evaluation of IM:	☒ Ground motion records (Number) ☐ GMPE (Ref) ☐ ShakeMap ☐ Other(____)

Fragility Curves for September event (FC)					
Damage Scale:	ATC-13				
Damage State (DS):	***ds₁***	***ds₂***	***ds₃***	***ds₄***	***ds₅***
Description of DS:	**Insignificant**	**Moderate**	**Heavy**	**Severe**	**Destroyed**
Function of Statistical Model:	Eq.(C.3)	Eq.(C.3)	Eq.(C.3)	Eq.(C.3)	Eq.(C.3)
Parameter (ϑ_0):	4.49	-	-	-	-
Parameter (ϑ_1):	2.53	-	-	-	-
Confidence Intervals for mean FC:	√	x	x	x	x
Statistical Model:	☒ GLM	☒ GLM	☒ GLM	☒ GLM	☒ GLM
Model Fitting Procedure:	☒ ML	☒ ML	☒ ML	☒ ML	☒ ML
Data Type:	☐ Building-by-Building ☒ Grouped data				
Number of Data Points:	11	11	11	11	11
Grouped Data: Definition of the aggregation unit:	Units with the same PGA level				
Grouped Data: Min Number of Buildings / Data Point:	3	3	3	3	3
Model Assumptions:					
Spatial Independence of Data Points:	-	-	-	-	-
Measurement Error in IM:	-	-	-	-	-
Measurement Error in Response:	-	-	-	-	-

Other(____________________):	-	-	-	-	-
Goodness of Fit Assessment:					
Mean function:	√	No	No	No	No
Variance function:	√	No	No	No	No
Procedure:					
Confidence Intervals:	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic
Fragility Curves for February event conditioned on insignificant damage from the September event (FC)					
Damage Scale:	ATC-13 (1985)				
Damage State (DS):	***ds_1***	***ds_2***	***ds_3***	***ds_4***	***ds_5***
Description of DS:	**Insignificant**	**Moderate**	**Heavy**	**Severe**	**Destroyed**
Function of Statistical Model:	Eq.(C.2)	Eq.(C.2)	Eq.(C.2)	Eq.(C.2)	Eq.(C.2)
Parameter (ϑ_0):	-	1.06	0.21	-	-
Parameter (ϑ_1):	-	0.79	0.47	-	-
Confidence Intervals for mean FC:	x	√	√	x	x
Statistical Model:	☐ GLM	■ GLM	■ GLM	■ GLM	■ GLM
Model Fitting Procedure:	☐ ML	■ ML	■ ML	■ ML	■ ML
Data Type:	☐ Building-by-Building	■ Grouped data			
Number of Data Points:	15	15	15	15	15
Grouped Data: Definition of the aggregation unit:	Units with the same PGA level				
Grouped Data: Min Number of Buildings / Data Point:	-	1	2	1	1
Model Assumptions:					
Spatial Independence of Data Points:	-	-	-	-	-
Measurement Error in IM:	-	-	-	-	-
Measurement Error in Response:	-	-	-	-	-
Other(____________________):	-	-	-	-	-
Goodness of Fit Assessment:					
Mean function:	No	√	√	No	No
Variance function:	No	√	√	No	No
Procedure:					
Confidence Intervals:	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic	☐ Asymptotic

Discussion:

The empirical fragility assessment procedure is adopted here to assess the probability of a damage state being reached or exceeded given intensity levels from the four main events which affected Christchurch. The procedure requires three steps. In the first step, the quality of the database is discussed. In the second step, the quality of the excitation observation is assessed. The third step involves the identification of a statistical model which best fits the empirical data.

1st Step: Preparation of the empirical damage data

***Data Quality Assessment*:**

The Christchurch database includes data from 627 URM, mostly commercial, buildings, collected over 32 suburbs of Christchurch. These suburbs were affected by four successive earthquake events, i.e. September 2010, February 2011, June 2011 and December 2011. The data has been collected by Auckland and Adelaide Universities. The sampling technique adopted is not clear. The damage has been classified into six damage states according to the ATC-13 damage scale (1985). Figure C.1 depicts the distribution of the buildings amongst each of the 6 damage states for the four earthquake events. It can be seen that for the first event, more than half of the buildings were not attributed a damage state. The percentage of missing data reduces for the February 2011 earthquake, but is substantial for the June and December 2011 events. This very high proportion of missing data is ignored in this illustrative example application. Implicit in this decision is the assumption that the missing data is randomly distributed across the surveyed area, and that by disregarding them we are simply reducing the sample size without introducing bias in the remaining sample. Due to the very small samples provided for the June and December 2011 earthquake events, the focus of this study is on constructing fragility curves using the data from the first two earthquakes.

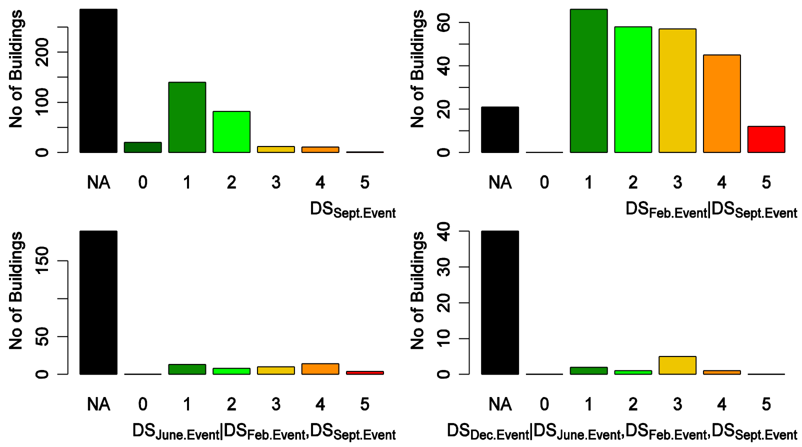


Figure C.1 Number of buildings suffered each of the six states of the observed damage for the four events.

It should be noted that the damage data are grouped into bins corresponding to a given intensity measure level, irrespective of the suburb in which they are located. This results in 10 data points per damage state for September event. The fragility curves for the February event are constructed considering that the buildings may have already experienced some damage from the September event. Figure C.2 depicts the distribution of the damage experienced by buildings from the February event given the damage state experienced by the buildings by the September. Figure C.2 depicts that the biggest sample size corresponds to the buildings that were damaged insignificantly by the September event.

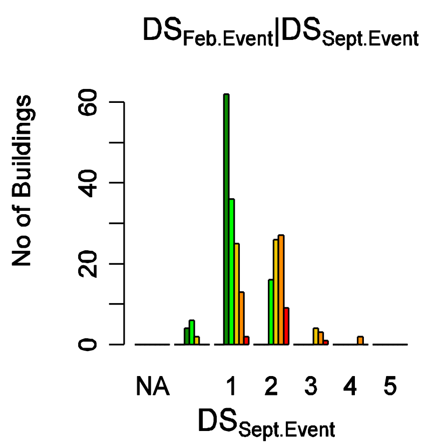


Figure C.2 Distribution of buildings affected by the February 2011 earthquake.

Overall, the quality of the adopted data is considered questionable and more research is required in order to establish its quality.

Data preparation per source:

The unreinforced masonry (URM) buildings considered in the database are mostly built using clay bricks. The data includes 7 sub-classes of the surveyed URMs based on their height and whether they stand independently or are joined to other buildings. Due to the small sample sizes for each sub-class, fragility curves are constructed for the generic URM class ignoring the sub-categories. The construction of the fragility curves requires the transformation of the grouped damage data found in the database into data points (x_j,(y_{ij},n_j-y_{ij})), where y_{ij} is the count of buildings that suffered $DS \geq ds_i$ and n_j-y_{ij} is the count of buildings that sustained $DS<ds_i$ for the bin j with intensity measure level x_j, where j=1:11 for the September 2010 event and j=1:15 for the February 2011 earthquake.

2nd Step: Selection and Estimation of the Intensity Measure

Peak ground acceleration (PGA) is selected as the ground motion intensity type. The PGA values determining intensity level contours used for the damage data aggregation are obtained from accelerograms at the ground motion stations in the area. Given this, we consider the measurement error in PGA as negligible.

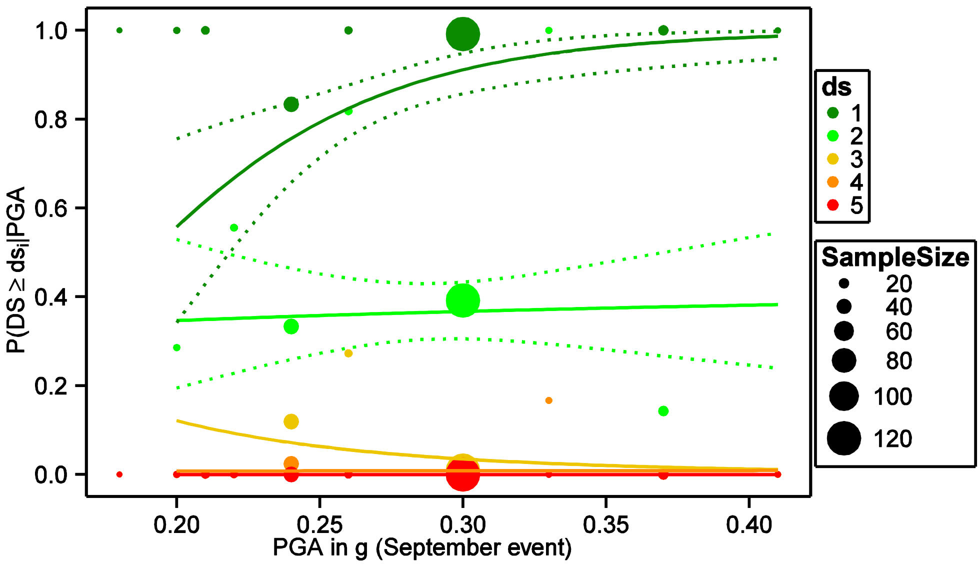


Figure C.3 Number of buildings suffered each of the six states of the observed damage for the four events.

3rd Step: Selection of the statistical model

Plots of the cumulative proportion of buildings for the 11 and 15 data points of the first and second earthquake events, against the intensity measure levels are shown in Figure C.4a,b. These figures show that for the September 2010 earthquake, the majority of the buildings are clustered in the low-to-intermediate range of ground motion intensity levels, whilst a wide range of PGA values is covered by the February 2011 earthquake data (i.e., 0.09g-1.0g). In addition, most data appear to be concentrated in the bin with PGA_{Sept}=0.3g or PGA_{Feb}= 0.7g.

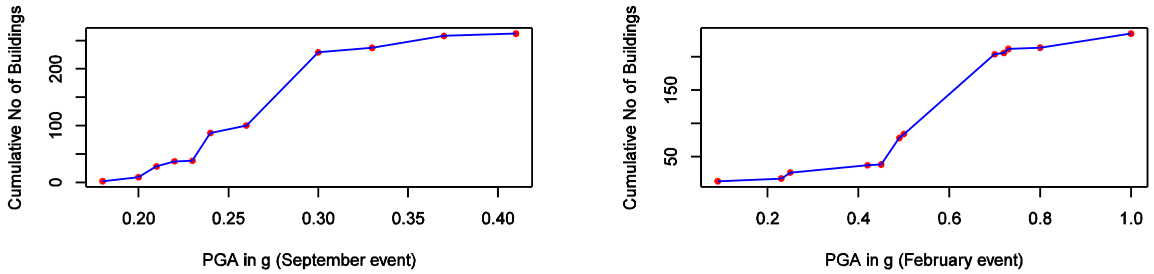


Figure C.4 Cumulative distribution of the proportion of the examined buildings in each data point against the corresponding intensity measure.

GLM models are chosen for the analysis. These models assume that the damage conditioned by an intensity measure type follows a binomial distribution and the mean of this distribution is related with the PGA through the three link functions (see Table 6.2), namely: logit (Eq.(C.1)), probit (Eq.(C.2) and complementary log-log (Eq.(C.3)). The statistical models tested in this exercise are presented here:

$$Y \sim f\left(y \mid IM, \boldsymbol{\theta}\right) = \begin{pmatrix} n_j \\ y_j \end{pmatrix} \mu^{y_j} \left[1-\mu\right]^{n_j - y_j} \quad \text{where, } \mu = P\left(DS \geq ds_i \mid IM\right) = \frac{1}{1+\exp\left(\vartheta_0 + \vartheta_1 \log\left(IM\right)\right)} \tag{C.1}$$

$$Y \sim f\left(y \mid IM, \boldsymbol{\theta}\right) = \begin{pmatrix} n_j \\ y_j \end{pmatrix} \mu^{y_j} \left[1-\mu\right]^{n_j - y_j} \quad \text{where, } \mu = P\left(DS \geq ds_i \mid IM\right) = \Phi\left(\vartheta_0 + \vartheta_1 \log\left(IM\right)\right) \tag{C.2}$$

$$Y \sim f\left(y \mid IM, \boldsymbol{\theta}\right) = \begin{pmatrix} n_j \\ y_j \end{pmatrix} \mu^{y_j} \left[1-\mu\right]^{n_j - y_j} \quad \text{where, } \mu = P\left(DS \geq ds_i \mid IM\right) = 1+\exp\left[-\exp\left(\vartheta_0 + \vartheta_1 \log\left(IM\right)\right)\right] \tag{C.3}$$

4th Step: Statistical Analysis

Estimation of the statistical parameters

The fitting of the three statistical models is performed using 'R' (2008). The parameters of these models are estimated numerically by maximising their likelihood function (see section 6.3.1). The five fitted fragility curves for the URM building class based on the data collected following the September 2010 earthquake are depicted in Figure C.5. The curves corresponding to ds_2-ds_5 appear to be either flat or have a negative slope, indicating that perhaps the statistical sample is not capable to depict the trend in the data correctly. A likelihood ration test is adopted in order to assess whether PGA is a statistically significant explanatory variable. This test identifies that the PGA is an important explanatory variable only for ds_1. By contrast, it indicates that there is not enough evidence to show that the presence of PGA leads to a model significantly better than a horizontal line. Therefore, the four fragility curves are not considered reliable. This can be attributed to the very small sample sizes especially for the three most extreme levels of damage, and raises questions on the validity of the assumptions made regarding the missing data.

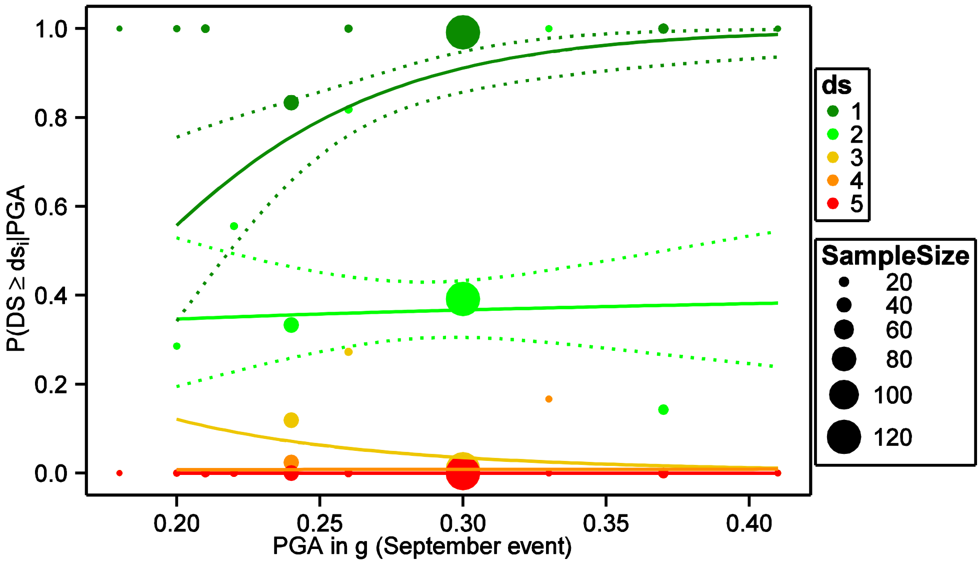


Figure C.5 Fragility curves for the 5 ATC-13 damage states constructed from data for the September 2010 earthquake.

Goodness of fit checks:

A set of graphical plots is used here to diagnose the ability of the selected statistical model to fit the available empirical data. Plots of Pearson residuals for *ds₁*, against the fitted values are shown in Figure C.4. In general, a statistical model can be considered acceptable if the Pearson residuals have zero mean and a constant variance equal to 1. This practically means that most residuals should be included in the [-3,3] interval and be randomly distributed. From Figure C.6, it is seen that the residuals appear to be in the expected interval. Their distribution does not appear to be random. However, the small number of available points cannot result in conclusive observations regarding their distribution. All three GLM model is therefore considered acceptable.

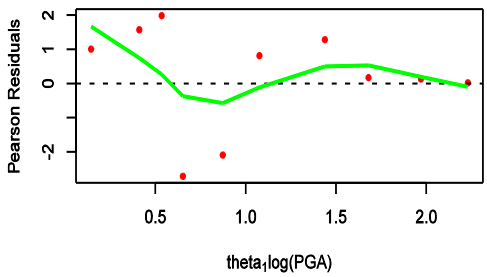


Figure C.6 Pearson residuals against the fitted values for fragility curves corresponding to ds_1 for Eq.(C.3).

The AIC test shows that the model using the 'cloglog' link function is the optimum model.

Table C.1 AIC values for the three models expressing the fragility curves corresponding to ds_1.

Model	AIC
Eq.(C.1)	32.34
Eq.(C.2)	33.55
Eq.(C.3)	***31.18***

The construction of fragility curves based on the damage data from the September 2010 event are based on the realistic consideration that the buildings were undamaged prior to the event. However, the fragility curves for buildings affected by the February event are constructed using the buildings which suffered insignificant (ds1) damage by the September event.

The fragility curves from the 4 damage states are depicted in Figure C.7. A likelihood ratio test carried out on the fitted model shows that the PGA can be considered a significant improvement explanatory variable only for the ds_1 and ds_2. However, the small sample size (235) leads to confidence intervals which overlap at lower PGA values, indicating that the reliability of the curves is questionable.

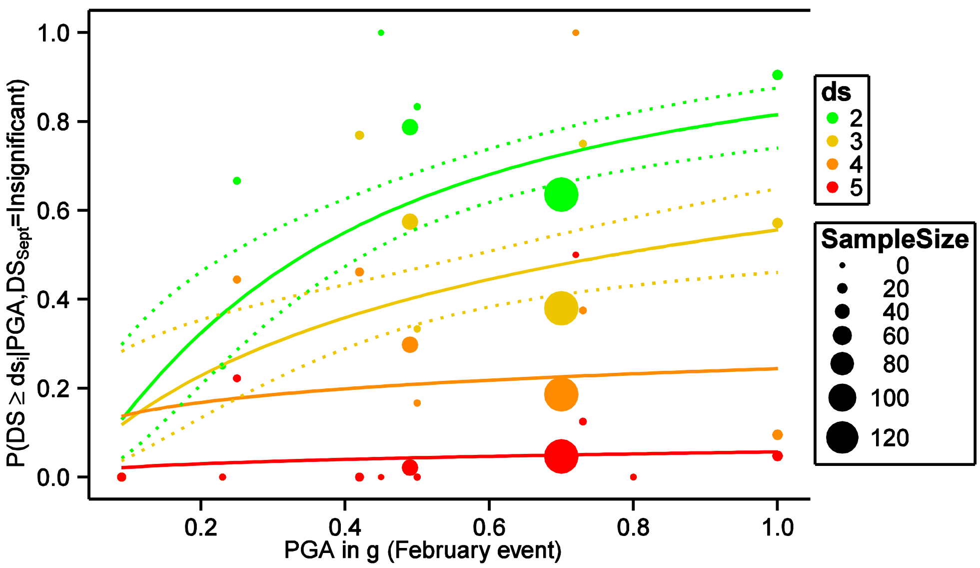


Figure C.7 Fragility curves for February event given that buildings were insignificantly (*ds_1*) damaged in the September earthquake.

Similar to the analysis of the data from the September event, Figure C.8 depicts the plots of the standardised Pearson residuals against the linear predictor. All three models might suffer from over-dispersion.

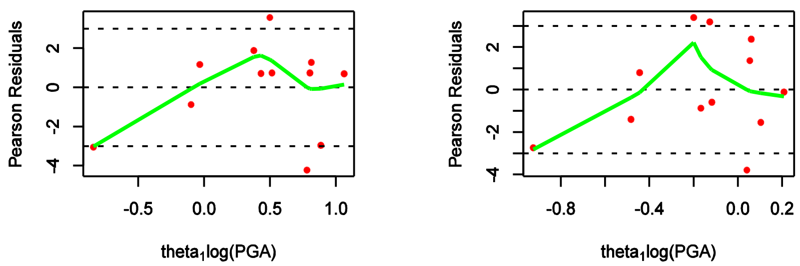


Figure C.8 Pearson residuals against the fitted values for fragility curves corresponding to *ds_1* (left) and *ds_2* (right) for Eq.(C.2).

The use of AIC (see Table C.2) identifies the model expressed by Eq.(C.2) as the best fit.

Table C.2 AIC values for the three models expressing the fragility curves corresponding to ds_1 and ds_2.

Model	AIC (ds_1)	AIC (ds_2)
Eq.C1	72.99	72.99
Eq.C2	***66.12***	***72.73***
Eq.C3	68.86	73.7

Lessons:

The empirical data adopted in this study is rare in that it reports damage to the same buildings following two earthquake events. A novel procedure is demonstrated for accounting for previous damage in the buildings in the construction of fragility curves for the second earthquake event. The questionable quality combined with small quantities of surveyed buildings results in fragility curves of questionable reliability. This highlights the importance of a thorough planning of the sampling technique. It is also observed that constructing fragility curves for each damage state following the proposed procedure might require different expressions of the link function for each fragility curve , other than the probit, which is overwhelmingly adopted in the literature.

References:

ATC-13 [1985] ''Earthquake damage evaluation data for California'', Report, Redwood City, Palo Alto, California.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

APPENDIX D Fitting GLMs to Data from two Australian Earthquake Events

Vulnerability Curves for Australian Unreinforced Masonry and Timber Buildings	
Developer, Affiliation and Date:	T. Maqsood and M. Edwards, GeoScience Australia 10/12/13
Statistical Package:	R (R Development Team, 2008)
Selected Building Class:	Brick (unreinforced masonry) buildings and timber buildings
Type of Assessment:	☒ Direct Vulnerability ☐ Indirect Vulnerability ☐ Fragility
Number of Buildings per Class:	3,796 Brick + 5,330 Timber (Newcastle earthquake) and 400 brick (Kalgoorlie earthquake)
Sources of Data:	1989 Newcastle earthquake insurance loss data and 2010 Kalgoorlie earthquake field survey loss data
Overall Rating of Quality of Data (Newcastle):	☐ High ☒ Moderate ☐ Poor ☐ Other(____)
Overall Rating of Quality of Data (Kalgoorlie):	☒ High ☐ Moderate ☐ Poor ☐ Other(____)
Def on of Loss Parameter:	Damage Factor: Ratio of repair cost to replacement cost or ratio of adjusted claims to adjusted cover
Int y Measure (IM): Damage Factor	MMI
Rar of IM:	V to VIII
Eva ion of IM:	☐ Ground motion records (Number) ☐ GMPE (Ref) ☐ ShakeMap ☒ Other(Isoseismal maps)
Mean Vulnerability Relationship	
Shape of Nonparametric Curve:	Mean damage factor against MMI plotted in Figure D.13.
Confidence Intervals of the Vulnerability Relationship	
Confidence Intervals for Mean Vulnerability Curve:	Mean damage factor plus minus one sigma against MMI presented in Figure D.13.
Direct Vulnerability Assessment - Loss Functions	
Developer, Affiliation and Date:	Tariq Maqsood and Mark Edwards, GeoScience Australia, 10/12/13
Damage Scale:	xDamage Factor (ratio of repair cost to replacement cost) against MMI plotted in Figure D.13.
Procedure:	Damage to Loss functions obtained from insurance data for URM buildings and Timber buildings.

Discussion:

A direct vulnerability assessment procedure is adopted here in order to estimate the economic loss caused by direct damage to masonry and timber buildings. Two datasets are prepared resulting from the 1989 Newcastle and 2010 Kalgoorlie earthquakes. Both datasets have positive and negative aspects. The former provides insurance claim information for more than 9,000 buildings but without street addresses. The latter provides detailed building attributes and damage information but only for 400 buildings. Also, the intensity range is quite limited for both datasets. Data is only available for three intensity levels (MMI VI to VIII) in the case of the Newcastle event and for two intensity levels (MMI V to VI) in the case of the Kalgoorlie event. Nevertheless, these two events provide the best available information for the study of earthquake vulnerability in Australia.

1st Step: Preparation of the loss data

***1989 Newcastle earthquake data*:**

A M_L5.6 earthquake occurred in Newcastle on 28 December 1989 causing heavy damage and the loss of 13 lives (Dhu and Jones, 2002). Unfortunately, there are no strong motion recordings of the earthquake close to the heavily damaged areas. NRMA insurance (now the Insurance Australia Group, IAG) records are obtained from the Newcastle City Council to estimate the cost of damage to buildings due to the 1989 Newcastle earthquake. There are approximately 14,000 NRMA claims in total that include for each claim, the suburb, the value insured, the pay-out and whether the claim was for a brick building, a timber building or the contents. For the study region, NRMA data gives total building claims of approximately $86 million (1989 US dollars) and total insured value (buildings) of $8,981 million (1989 US dollars). However, this data is a biased sample of building loss as it does not include the buildings for which claims were not made. Furthermore, there is uncertainty as to what percentage of the buildings in the study region was insured by NRMA, and what the underinsurance factor and excess fees were.

To address these issues and prepare a damage database Geoscience Australia consulted the IAG. During the consultation, claim rates for brick and timber buildings were estimated for each shaking intensity level; and an underinsurance factor and a typical deductible value was also evaluated. Demand surge or post-event inflation, which can distort the claims, is believed to be minor and hence neglected for rest of the analysis. The Newcastle Earthquake occurred at a time of softening demand in the building industry and the Kalgoorlie Earthquake was not of a severity that could cause significant demand surge inflation.

As the insurance claim data does not provide street addresses for each claim, the claims are aggregated at the suburb level (114 suburbs). By using the outcomes of a survey of more than 6,000 properties conducted by Geoscience Australia in Newcastle in 1999, an indicative age (pre-1945 and post-1945) is attributed to each suburb to differentiate the older building stock from the relatively new one. For each suburb the claims are sub-sampled based on building type (brick or timber) and age category (pre 1945 and post 1945) and the number of buildings and total cover in the suburb is expanded to a notional portfolio by using an agreed claim rate for each of the four categories and intensity levels. Then, adjustments are made for underinsurance and deductibles. In the final step to prepare the damage database, the Damage Factor (DI) is calculated as a ratio of adjusted claim to adjusted cover for each suburb.

***2010 Kalgoorlie earthquake data*:**

The M_L 5.0 earthquake shook Kalgoorlie and neighbouring areas on 20 April 2010. The resultant ground motion was found to vary markedly across the town (Edwards et al., 2010). Geoscience Australia conducted an initial reconnaissance and captured street-view imagery of 12,000 buildings within Kalgoorlie by using a vehicle mounted camera system. The subsequent foot survey collected detailed information from nearly 400 buildings in Kalgoorlie and Boulder. The shaking caused widespread damage to pre-World War One unreinforced masonry buildings. More modern masonry buildings also experienced some damage in the vicinity of Boulder. The Damage Factor for each surveyed building is calculated by firstly recording damage to different building elements and assigning a damage state in terms of None, Slight, Moderate, Extensive and Complete to match the HAZUS damage states. Secondly, a percentage damage is assigned to each element and lastly percentage loss for a building is determined as the sum over all building elements of: (% of building cost contributed by the element)x(%damage)x(%of element so damaged). The Kalgoorlie data provides just three data points for two suburbs at MMI V and VI.

Data preparation:

Figure D.1 presents the damage data for brick (unreinforced masonry) buildings and Figure D.5 shows the damage data for timber buildings for each suburb in the study region from Newcastle and Kalgoorlie events. The number of claims per suburb varies from a few to more than 400 per suburb. In order to develop vulnerability functions from a statistically significant database, suburbs with fewer than 20 claims are eliminated from the database. Figures D.2 and

D.6 show the damage data for brick and timber buildings respectively from suburbs with more than 20 entries.

In order to differentiate between the vulnerability of older legacy buildings from relatively newer construction, the data is sub-divided, for brick and timber buildings, into two age categories, i.e., pre 1945 and post 1945, using the notional age of each suburb. Figure D.3 and D.4 illustrates the data for pre- and post-1945 brick buildings, respectively. Figure D.7 and D.8 show the data for pre- and post-1945 timber buildings, respectively. Overall, the quality of the Newcastle database is considered moderate due to the aggregated type of data. The quality of the Kalgoorlie database is believed to high due to building by building field surveys.

2nd Step: Selection and Estimation of the Intensity Measure

Even though the guideline requires a number of intensity measures to be selected, due to a lack of strong motion recordings, the only available intensity measure for this study is the Modified Mercalli Intensity (MMI). Rynn et al., (1992) have produced a local intensity map for the Newcastle and Lake Macquarie area with MMI range of VI to VIII. Each suburb is therefore assigned an MMI value from the intensity map. An averaged intensity is assigned where a suburb has two or more isoseismal contours according to the intensity map prepared by Rynn et al., (1992) and number of claims within the suburb. For the Kalgoorlie Earthquake, the MMI values derive from direct observation and interviews with residents. The estimated MMI in Kalgoorlie and Boulder were V and VI, respectively (Edwards et al., 2010).

3rd Step: Selection of GLM

For this study, generalised linear models (GLM) are selected to develop vulnerability functions for brick and timber buildings.

4th Step: Statistical Model Fitting Procedure

Statistical model fitting is performed using 'R'. The parameters of this model are estimated numerically by maximising their likelihood function (see §6.3.1). As mentioned in §5.3, the GLM consist of three main components i.e., a probability distribution, a linear predictor and a link function. For each of the four derived databases (pre/post 1945 Brick and pre/post 1945 Timber), three distributions (Gamma, Inverse Gaussian and Lognormal) and three link functions (Identity, Log, and Inverse) were used to derive the vulnerability curves. Then, goodness of fit checks are carried out to select the optimum vulnerability curve.

Goodness of fit checks:

A set of graphical plots is used to diagnose the ability of the selected regression models to fit the available data. The plots of the Pearson residuals against the fitted values, scale location plot and influential data points are plotted in Figures D.9 to D.12. In general, a statistical model can be considered acceptable if the Pearson residuals have zero mean and constant variance equal to φ. The smoothing curve in these Figures appears to be strongly influenced by certain residuals and deviate considerably from the expected zero value. This, however, can be attributed to the strong influence of very small number of residuals and is considered acceptable.

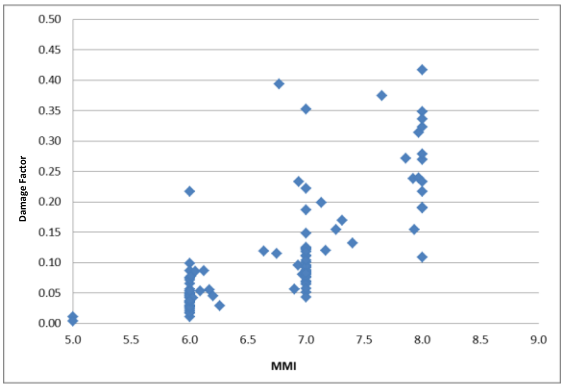


Figure D.1 Damage Factor vs MMI for each suburb for brick (unreinforced masonry) buildings.

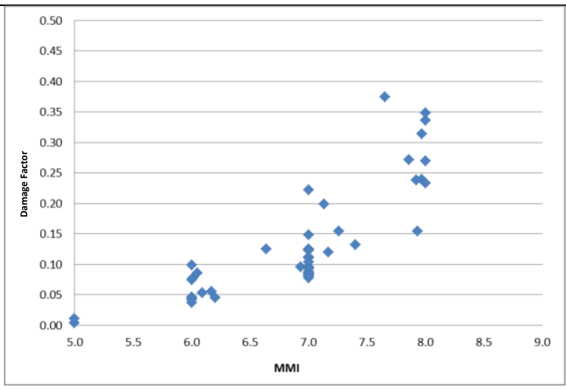


Figure D.2 Brick buildings: Damage Factor vs MMI for each suburb for with >20 claims.

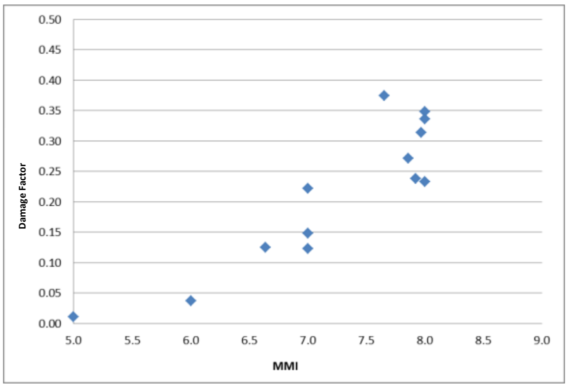


Figure D.3 Pre 1945 Brick buildings: Damage Factor vs MMI for each suburb with >20 claims.

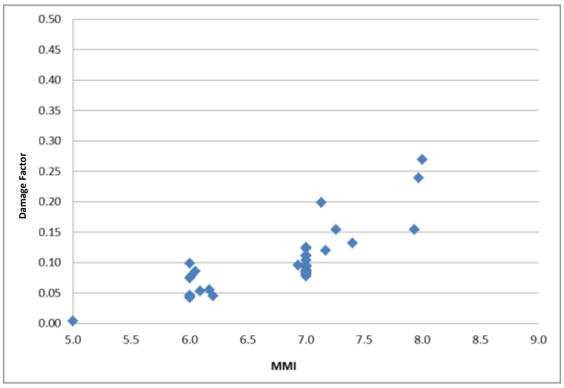


Figure D.4 Post 1945 Brick buildings: Damage Factor vs MMI for each suburb with >20 claims.

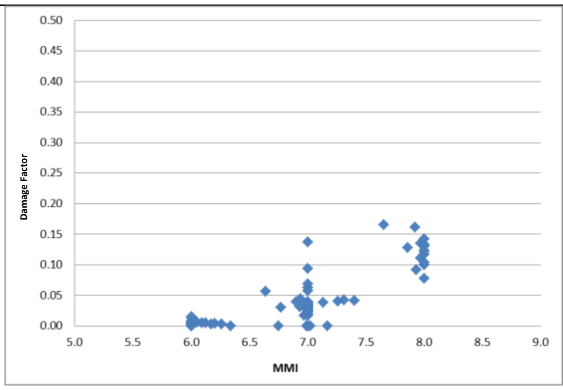


Figure D.5 Damage Factor vs MMI for each suburb for timber buildings.

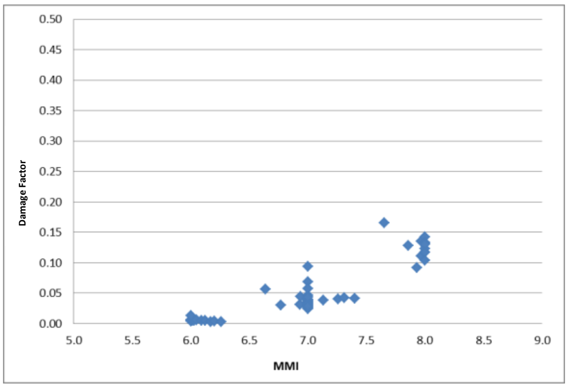


Figure D.6 Timber buildings: Damage Factor vs MMI for each suburb for with >20 claims.

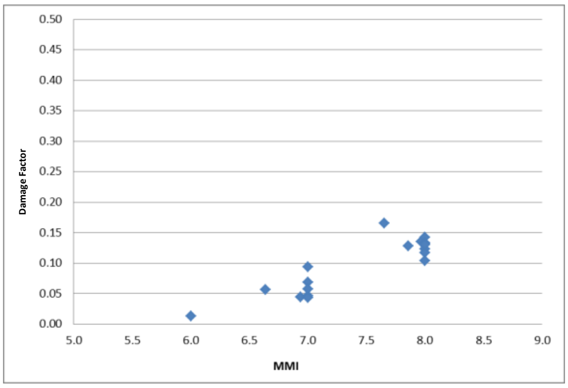


Figure D.7 Pre 1945 Timber buildings: Damage Factor vs MMI for each suburb with >20 claims.

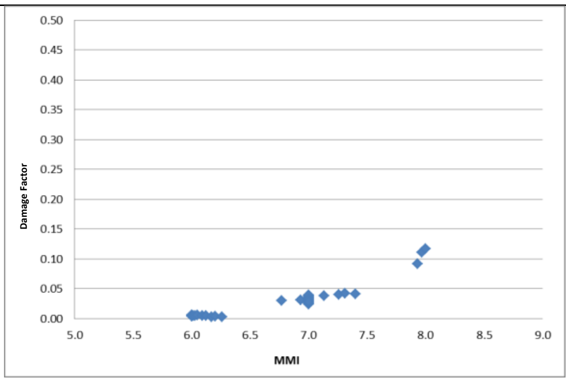


Figure D.8 Post 1945 Timber buildings: Damage Factor vs MMI for each suburb with >20 claims.

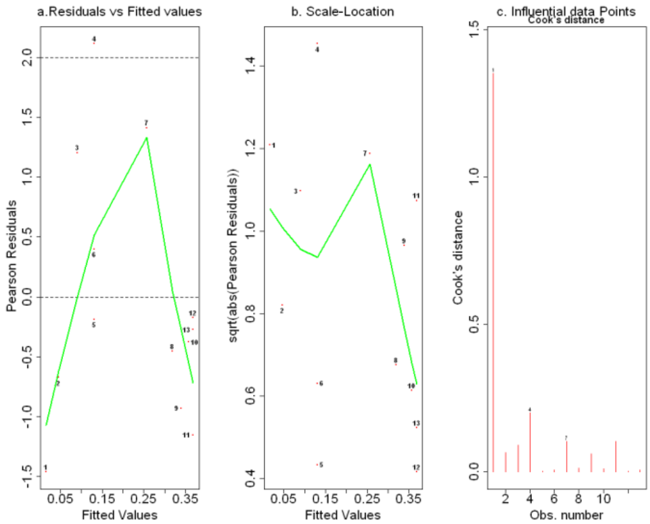


Figure D.9 (a) Pearson residuals against the fitted values, (b) Scale-location plot, (c) Plot of influential points for brick pre 1945 build

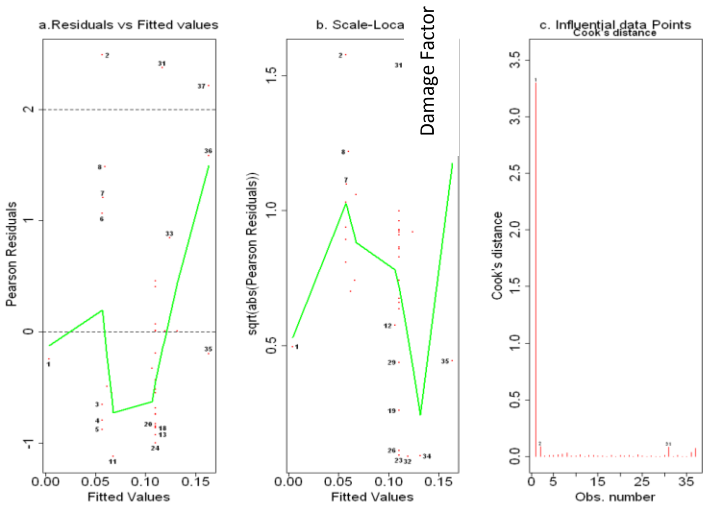


Figure D.10 (a) Pearson residuals against the fitted values, (b) Scale-location plot, (c) Plot of influential points for brick post 1945 buildings.

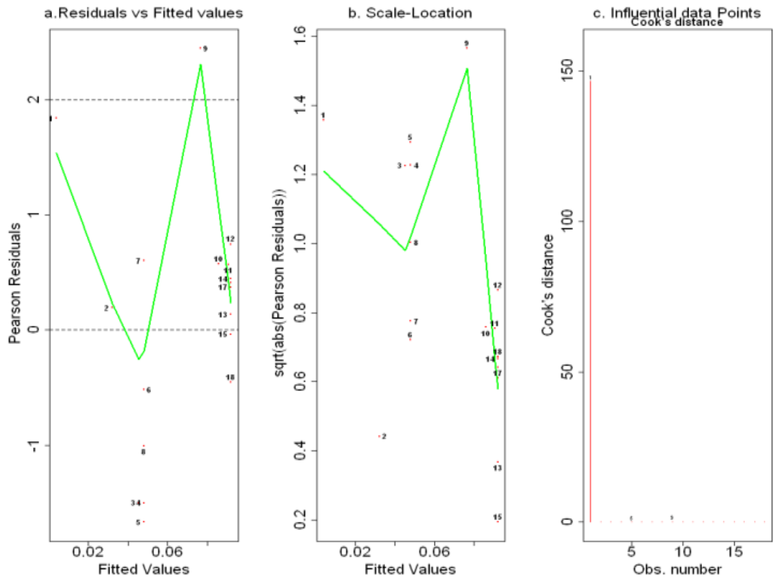


Figure D.11 (a) Pearson residuals against the fitted values, (b) Scale-location plot, (c) Plot of influential points for timber pre 1945 buildings.

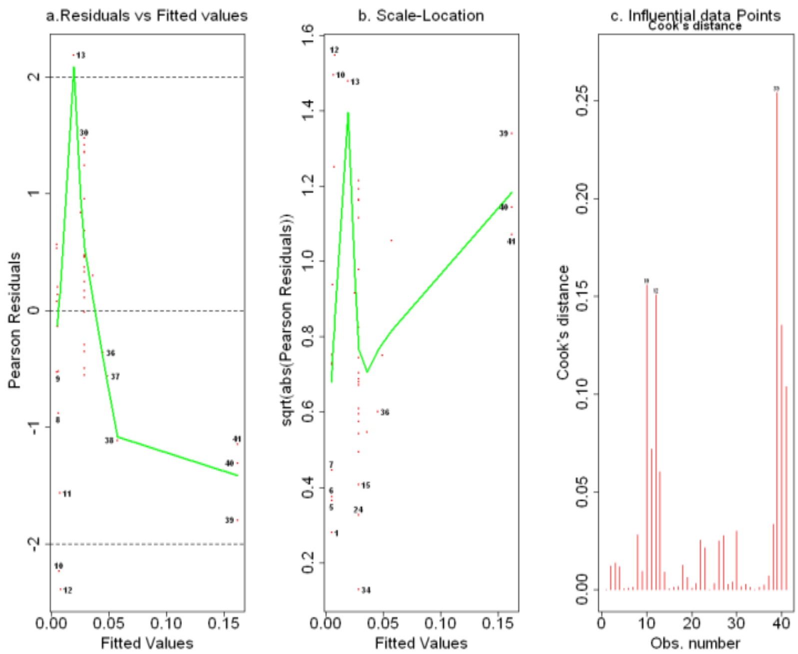


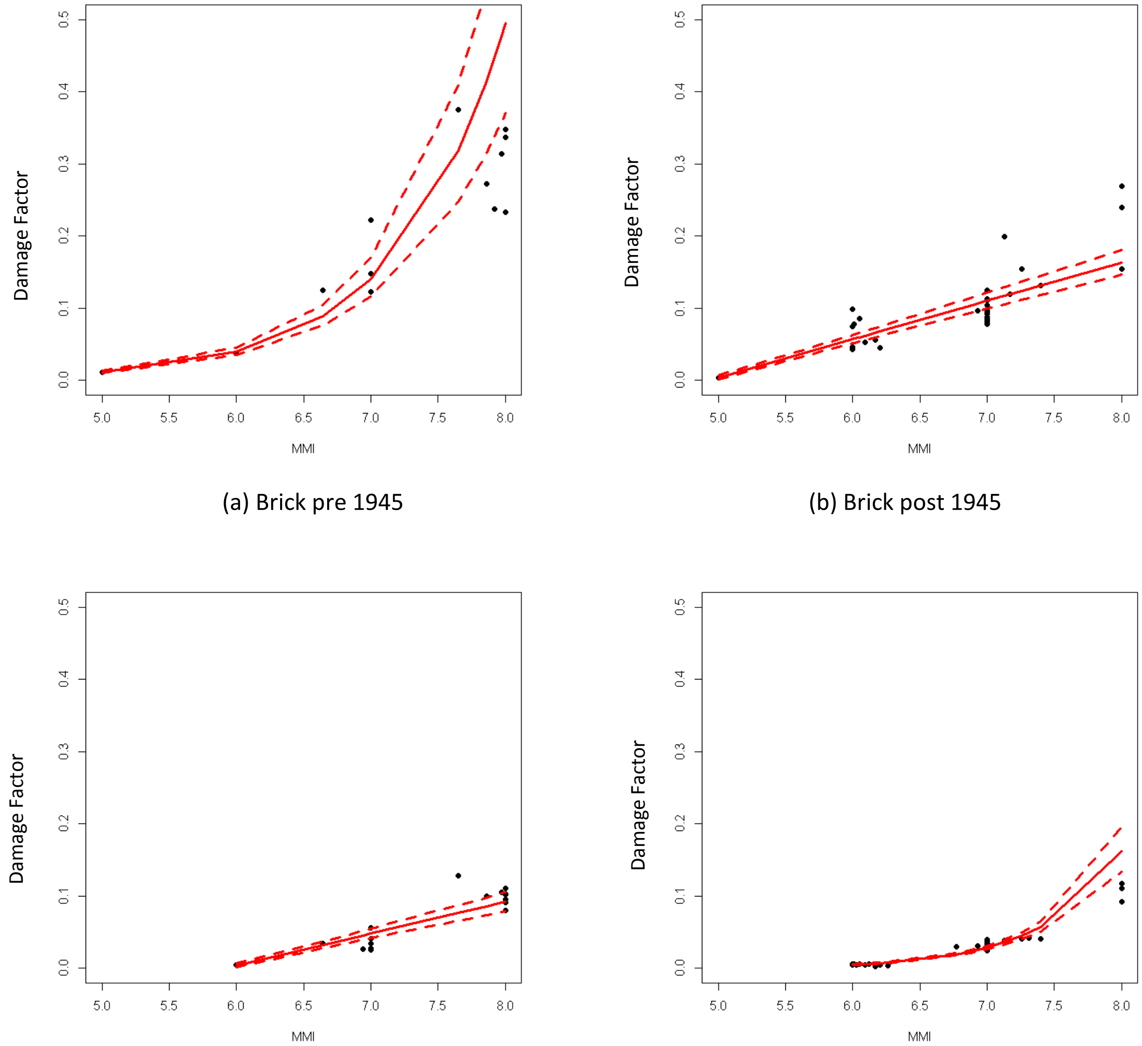
Figure D.12 (a) Pearson residuals against the fitted values, (b) Scale-location plot, (c) Plot of influential points for timber post 1945 buildings.

The diagnostics of brick pre 1945 building data showed that the Inverse Gaussian distribution with Log link function reasonably relates the loss with MMI (refer to Figures D.9). Figures D.10 presents the diagnostics resulting from the use of Gamma distribution with Identity link functions for brick post 1945 data. Similarly Figures D.11 show the diagnostics resulting from the use of Gamma distribution with Identity link functions for timber pre 1945 data. Figures D.12 presents diagnostics resulting from the use of Gamma distribution with Log link functions for timber post 1945 data.

5th Step: Identify the optimum vulnerability curve

According to the recommendations of §6.3, the AIC for the different models representing various distributions (Normal, Gamma, Inverse Gaussian) and link functions (Identity, Inverse, Log) corresponding to the MMI are compared. The model with smallest AIC value is considered to be the optimum according to the guidelines in §6.3.

In this case of brick pre-1945 buildings the Inverse Gaussian distribution with the Log link function produces the smallest AIC value. For brick post-1945 buildings the model which best fits the data and has the smallest AIC value is the Gamma distribution with the Identity link function. For timber pre 1945 building data, the optimum model is the Gamma distribution with the Identity link function; and for timber post 1945 buildings it is Gamma distribution with the Log link function. The mean vulnerability curve as well as the mean plus one standard deviation are plotted in Figure D.13.



(a) Brick pre 1945

(b) Brick post 1945

(c) Timber pre 1945

(d) Timber post 1945

Figure D.13 Mean vulnerability curve and mean plus minus one standard deviation for the range of MMI levels for (a) brick pre 1945, (b), brick post 1945, (c) timber pre 1945 and (d) timber post 1945.

Lessons

From an Australian perspective, it is concluded that the document provides detailed and comprehensive guidelines to prepare an empirical damage/loss database and to perform sophisticated regression analysis. However, at the same time it requires a large sample of observations to generate reasonable vulnerability functions, which may not be available. It can be seen from the results that sometimes even the best fit function fails to capture the observed trend at higher intensities due to limited data points. An attempt to fit the curve to observations at higher intensities led to a significant deviation at the lower intensity, a value which is more critical as even a small percentage difference for more likely ground shaking intensities can result in significant change in annualised losses.

The NRMA insurance data provides claim and cover information for more than 9,000 buildings in 114 suburbs impacted from the 1989 Newcastle earthquake. Geoscience Australia conducted a detailed survey of more than 5,000 buildings in Newcastle which captured detailed building attributes. Geoscience Australia also has access to the Newcastle City Council damage survey data which categorises building damage into red, amber, blue and green depending upon the severity of damage. The three databases have the potential to be utilised for detailed vulnerability assessment, however, it requires a common attribute such as street address to link the databases and augment the captured information. At present this link is missing which forced the building by building claim data to be aggregated at suburb level.

The aggregation of data and the very small number of data points may have led to significant loss of information and caused an increased uncertainty in the model. However, there is potential to improve the details of the loss database and building categorisation based on structural system, wall material, age and number of storeys by utilising the above mentioned three databases. Thereafter, more refined vulnerability curves can be developed for more building types.

References:

Dhu T., Jones T. [2002] ''Earthquake risk in Newcastle and Lake Macquarie.'' Geoscience Australia Record 200/15, Geoscience Australia, Canberra.

Edwards M., Griffith M., Wehner M., Lam N., Corby N., Jakab M., Habili N. [2010] ''The Kalgoorie earthquake of the 20th April 2010'', Proceedings of Australian Earthquake Engineering Society Conference, Perth, Australia.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

Rynn J.M., Brennan E., Hughes P.R., Pedersen I.S., Stuart H.J. [1992] "The 1989 Newcastle, Australia, Earthquake: The facts and the misconceptions", Bulletin of New Zealand National Society of Earthquake Engineering, Vol. 25, No. 2, pp. 77-144.

APPENDIX E Fitting GLMs to damage data from the 1978 Thessaloniki Earthquake, Greece, using Bayesian analysis

Fragility Curve for Greek Masonry Buildings.				
Developer, Affiliation and Date:	I. Ioannou, UCL EPICentre, 09/12/13			
Statistical Package:	R (R Development Team, 2008)			
Selected Building Class:	Masonry			
Type of Assessment:	□ Direct Vulnerability □ Indirect Vulnerability ■ Fragility			
Number of Buildings per Class:	28,559			
Sources of Data:	Survey by local authorities with the aim of safety assessment.			
Overall Rating of Quality of Data (per Source):	□ High ■ Moderate □ Poor □ Other(____)			
Definition of Loss Parameter:	-			
Intensity Measure (IM):	PGA in g			
Range of IM:	0.05g-0.25g			
Evaluation of IM:	□ Ground motion records (No) ■ GMPE (Margaris et al 2011) ShakeMap □ Other(____)			
Fragility Curves				
Damage Scale:	OASP (Green-Yellow-Red)			
Damage State (DS):	***Yellow***	***Red***	***Yellow***	***Red***
Description of DS:				
Function of Statistical Model:	Eq.(E.2)	Eq.(E.2)	Eq.(E.2)	Eq.(E.2)
Parameter (ϑ_0):	0.153	-0.433	0.172	-0.426
Parameter (ϑ_1):	0.254	0.314	0.262	0.317
Confidence Intervals for mean FC:	√	√	√	√
Statistical Model:	■ GLM	■ GLM	■ GLM	■ GLM
Model Fitting Procedure:	■ ML	■ ML	■ Bayesian	■ Bayesian
Data Type:	□ Building-by-Building ■ Grouped data			
Number of Data Points:	73	73	73	73
Grouped Data: Definition of the aggregation unit:	Municipality			
Grouped Data: Min Number of Buildings / Data Point:	13	13	13	13
Statistical Assumptions:				
Spatial Independence of Data	-	-	-	-

Points:	-	-	-	-
Measurement Error in IM:	-	-	-	-
Measurement Error in Response:	-	-	-	-
Other(____________________):				
Goodness of Fit Assessment:				
Mean function:	√	No	No	No
Variance function:	√	No	No	No
Procedure:				
Confidence Intervals:	☐ Asymptotic	☐ Asymptotic	■ Bayesian	■ Bayesian

Discussion:

The empirical fragility assessment procedure is adopted here to assess the probability of a safety state being reached or exceeded given intensity levels from the safety assessment data of masonry buildings affected by the 1978 Thessaloniki earthquake. The procedure requires three steps. In the first step, the quality of the database is discussed. In the second step, the quality of the excitation observations is assessed. The third step involves the identification of a statistical model which fits the data best.

1st Step: Preparation of the damage data

***Data Quality Assessment*:**

The database was obtained by UCL EPICentre from the organization responsible for recording and repairing seismic damage in Northern Greece. The 1978 Thessaloniki database contains damage data of mainly reinforced concrete and masonry residential buildings located in 16 urban and 13 rural municipalities. The surveyed buildings are assigned one of three safety levels, namely: Green, Yellow and Red according to the provisions of Earthquake Planning and Protection Organization (1997). By comparing the total number of buildings, irrespective of their use or construction material, in each municipality with the number of buildings recorded in the 1981 census, a large non-coverage error has been identified for 7 urban municipalities and all 13 rural municipalities (i.e. ≤70% of the buildings in a municipality have been surveyed).

With regard to the urban areas, the 9 municipalities with non-coverage error ≤30% are adopted for the construction of fragility curves. By contrast, the non-coverage error in the 13 rural municipality is reduced by estimating the total number of residential buildings as the average of the number of residential buildings in each rural municipality reported in the census of 1971 and 1991 (due to the absence of detailed information regarding the number of residential buildings in each municipality in the 1981 census). Implicit in this, is the assumption that all residential buildings in rural municipalities are of masonry construction. Given that the buildings were surveyed only at the owner's request, it is also considered that non-surveyed buildings have suffered no damage, i.e. they are assigned the Green safety level.

The improved database contains damage data regarding 28,559 (instead of the total 30,783) masonry buildings located in 73 urban and rural postcodes. The damage data used has been collected during two rounds of post-earthquake seismic safety assessment. We consider that the second round produced a more accurate assessment, and that the buildings that were not been surveyed twice were assigned the correct safety level in the first survey. This is perhaps an unrealistic assumption to address the misclassification error but it is adopted for illustrative purposes.

Overall, the database is considered of moderate quality.

For the prior distributions, the 2003 Lefkada database (CEQID, 2013) is used. This database contains data from 4,793 masonry buildings aggregated in 39 municipalities. The damage is classified according to the six-state EMS-98 damage scale ($ds_{i=0\text{-}5}$). Given that there were no collapsed buildings, the damage scale is reduced to five-states, ranging from no-damage to heavy damage. The intensity is considered constant within each municipality. Its value is estimated from USGS ShakeMaps in terms of PGA (in g) and ranges from 0.009g to 0.023g. Contrary to the first database, the 2001 census has been used to estimate the total number of buildings and the non-surveyed buildings have been considered undamaged. Nonetheless, the damage is systematically overestimated, by an unknown degree, since it contains the overall damage accumulated after the main event as well as a strong aftershock.

Fragility curves corresponding to four damage states (i.e. ds_1-ds_4) for three sub-classes of low-rise residential masonry buildings are constructed by fitting cumulative lognormal distribution functions using an ordinal probit regression analysis. The three subclasses include buildings built a. before 1919, b. between 1919 and 1945 and c. after 1945 according to the classification of Karababa and Pomonis (2010). These four states are harmonized to the three safety levels adopted in this study according to the recommendations of Rossetto and Elnashai (2003). Therefore, the fragility curves corresponding to Yellow and Red safety levels are approximately equivalent to:

$$\begin{aligned} P\left(DS_{True} \geq Yellow \mid im_j\right) &\approx P\left(DS_{True} \geq ds_3 \mid im_j\right) \\ P\left(DS_{True} \geq Red \mid im_j\right) &\approx P\left(DS_{True} \geq ds_4 \mid im_j\right) \end{aligned} \quad \text{(E.1)}$$

2nd Step: Selection and estimation of the Intensity Measure

The intensity measure type selected is PGA and its value is evaluated at the centre of each municipality using the GMPE proposed by Margaris et al. (2011). This introduces a measurement error in the intensity measure levels, but this error is ignored for this application.

Figure E.1 depicts that the majority of the data are urban and correspond to a narrow range of PGAs between 0.07.-0.09g.

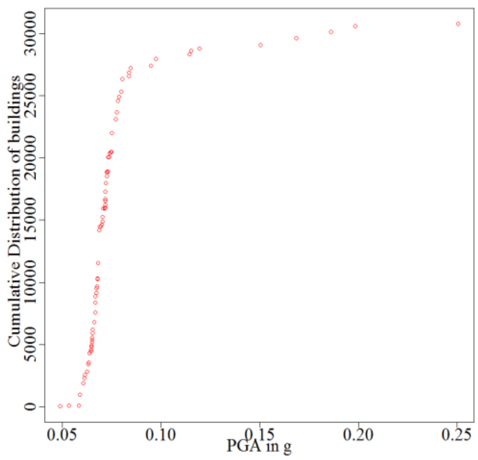


Figure E.1 Cumulative distribution of the proportion of the examined buildings in each data point against the corresponding intensity measure.

3rd Step: Selection of the statistical model

Maximum likelihood approach

A probit model is fit to the safety assessment (see Eq.(E.2)) data.

$$Y \sim f\left(y \mid IM, \boldsymbol{\theta}\right) = \left(\frac{n_j}{y_j}\right) \mu^{y_j} \left[1-\mu\right]^{n_j - y_j}$$

$$\text{where,} \quad \mu = P\left(DS \geq ds_i \mid IM\right) = \Phi\left(\vartheta_0 + \vartheta_1 \log\left(IM\right)\right) \quad \text{(E.2)}$$

The fragility curve for the yellow and red safety levels as well as their corresponding 90% confidence intervals are constructed by fitting the model expressed by Eq.(E.2) using a maximum likelihood approach.

Bayesian approach

A Bayesian approach is also adopted in order to estimate the model parameters accounting for prior knowledge regarding their distribution based on the 2003 Lefkada database. Figure E.2 depicts the algorithm adopted in R. It should be noted that the linear predictor of the adopted GLM is centred around the mean of the natural logarithm of IM:

$$Y \sim f\left(y \mid IM, \boldsymbol{\theta}\right) = \binom{n_j}{y_j} \mu^{y_j} \left[1-\mu\right]^{n_j - y_j}$$

$$\text{where,} \quad \mu = P\left(DS \geq ds_i \mid IM\right) = \Phi\left(\vartheta_0 + \vartheta_1\left(\log\left(IM\right) - E\left[\log\left(IM\right)\right]\right)\right) = \tag{E.3}$$

$$= \Phi\left(\vartheta_0 - \vartheta_1 E\left[\log\left(IM\right)\right] + \vartheta_1 \log\left(IM\right)\right) =$$

$$= \Phi\left(\vartheta_0 * + \vartheta_1 \log\left(IM\right)\right)$$

This is a common practice introduced for faster convergence of the Markov chains.

The next step involves the determination of the prior distribution for the two parameters:

- ϑ_1 (th1): This is expected to be positive (negative values of the slope of the fragility curves means that the probability of damage decreases with the increase of intensity measure levels). For this reason this parameter is assigned a gamma distribution, with parameters b and c.
- ϑ_0^* (th0.star): is a normal distribution described by its mean and the precision (1/standard deviation ^2).

The parameters of the two distributions are determined by fitting the models expressed by Eq.(E.3) to the 2003 Lefkada data. The fragility curve corresponding to "red" for LBSM2 is dismissed due to negative values of the slope, indicating that the damage decreases with an increase in intensity measure level, which is not expected. The mean and variance of the five values of ϑ_1 and ϑ_0^* determine the parameters of the distributions of these two parameters.

Table E.1 Parameters of models fitting to 2003 Lefkada data and the resulting parameters of the prior distributions.

Class	**Safety Level**	**ϑ^*_0**	**ϑ_1**
LBSM1	Yellow	-1.04250	1.16192
	Red	-1.85695	2.02338
LBSM2	Yellow	-1.05792	0.56047
	Red	-1.79286	-0.82059
LBSM3+4	Yellow	-1.29482	0.63767
	Red	-2.21538	0.80459
Priors - Yellow			
mean		-1.13	0.79
St.Dev.		0.141	0.327
Gamma parameters for ϑ_1		**b**	**c**
		5.78	7.34
Priors - Red			
mean		-2.04	1.41
St.Dev.		0.253	0.862
Gamma parameters for ϑ_1		**b**	**c**
		2.69	1.90

```
OrdinalProbitLn<-function() {

  for(i in 1:nBinsIM){

    DS.true[i] ~ dbin(FC.true[i],N[i])

    probit(FC.true[i])<- th0.star + th1*(log(IM.true[i])-lnIM.mean)

    }
  th1~dgamma(b,c)
  th0.star~dnorm(mutheta0.star,prec.theta0)
  th0<-th0.star - th1*lnIM.mean
}

write.model(OrdinalProbitLn,"OrdinalProbitLn.txt")
modelCheck("OrdinalProbitLn.txt")

##Get the data in BUGS:

lnIM.mean<-mean(log(IM.Observed))

# Yellow

BayesModelData <-list(nBinsIM=length(IM.Observed),DS.true=N.YellowOrAbove,
                      IM.true=IM.Observed,lnIM.mean=lnIM.mean,
                      mutheta0.star=-1.13,prec.theta0=(1/0.141)^2,
                      b=5.78,c=7.34,
                      N=N.buildings)

bugs.data(BayesModelData,data.file="DataBUGS.txt")

BayesInits <-  list(list(th0.star=-1.0,th1=1.0),
                    list(th0.star=-2.0,th1=2.0),
                    list(th0.star= 1.0,th1=0.5))

Yellow<- bugs(data="DataBUGS.txt",n.chains=3,
              inits=BayesInits,n.burnin=5000,n.iter=10000,OpenBUGS.pgm=NULL,
              parameters.to.save=c("th0","th1"),
              model.file="OrdinalProbitLn.txt",n.thin=1,debug=TRUE)
#### Red
BayesInits <-  list(list(th0.star=-1.0,th1=1.0),
                    list(th0.star=-2.5,th1=2.0),
                    list(th0.star=-5.5,th1=0.5))
```

Figure E.2 Bayesian analysis in R.

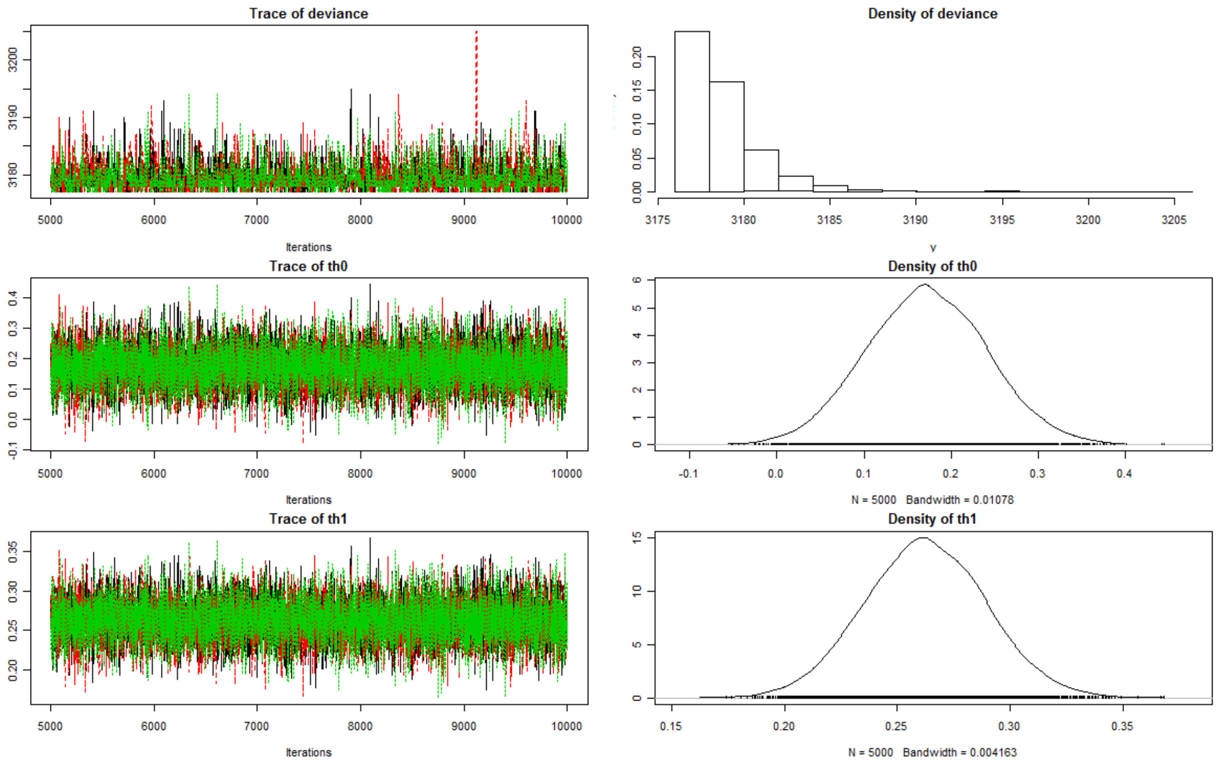


Figure E.3 Diagnostics for Yellow (The mix of the three chains appears to be adequate).

4th Step: Statistical Analysis

Estimation of the statistical parameters

The model expressed by Eq.(E.3) is fitted to the data by a Bayesian analysis using OpenBugs (Lunn et al., 2009) and R (R Development Team. 2008). A total of 10,000 runs are adopted and the results from the first 5,000 runs are ignored. The performance of the algorithm is examined in Figure E.3-4. Figure E.3 depicts the values of the GLM parameters and the deviation against the corresponding iterations. The shape of the graph indicates that all three chains converged. The autocorrelation in the sampled parameter values is tested in Figure E.4. High autocorrelation can be depicted if there are successive negative or positive values. Figure E.4 illustrates that with the exception of the deviance of a chain autocorrelation is not an issue.

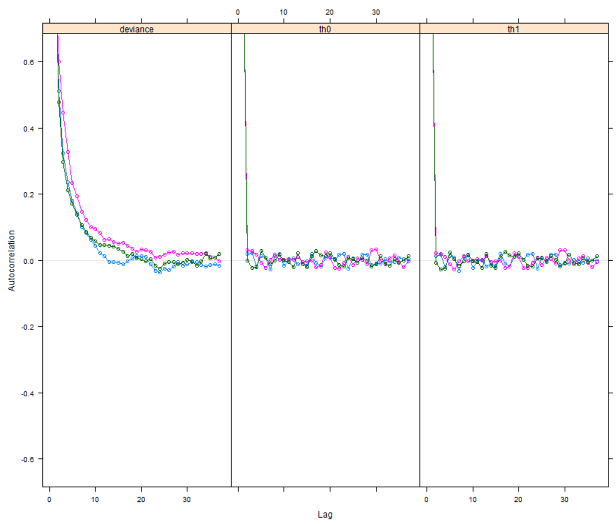


Figure E.4 Diagnostics for Yellow.

The obtained fragility curves are depicted in Figure E.5. The mean curves appear to overlap with the ones produced by the maximum likelihood method. This indicates that the presence of a large number of buildings dominates the results and that the prior knowledge regarding the distribution of the GLM's parameters does not influence the results. Small differences in the 90% intervals produced by the two procedures can be noted, but can be considered negligible. The narrow confidence intervals reflect the large sample size especially in the urban area of Thessaloniki. It should be noted that the over-dispersion in the grouped data is not taken into account in this application and is the subject of future work.

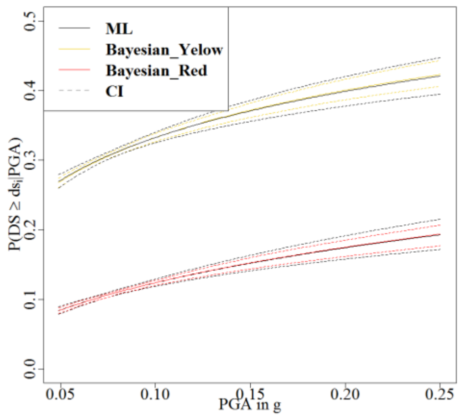


Figure E.5 Fragility curves and corresponding confidence intervals estimated by the maximum likelihood approach and the Bayesian analysis (It should be noted that in both cases, the binomial distribution is wrongly assumed to capture the uncertainty in the data).

Lessons:

The use of Bayesian analysis will not necessarily lead to fragility curves that are different from those resulting from a maximum likelihood approach in cases where a large database is available. The model should be expanded in order to capture the over-dispersion in the data.

References:

Lunn, D., D. Spiegelhalter, A. Thomas, N. Best [2009] "The BUGS project: Evolution, critique and future directions", Statistics in Medicine, Vol. 28, No. 25, pp. 3049-3067.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

APPENDIX F Fitting GAMs to damage data from the 1980 Irpinia Earthquake, Italy

Fragility Curve for Italian Field Stone Masonry Buildings with Wooden Floors.		
Developer, Affiliation and Date:	I. Ioannou, UCL EPICentre, 01/04/12	
Statistical Package:	R (R Development Team, 2008)	
Selected Building Class:	Field stone masonry buildings with wooden floors	
Type of Assessment:	☐ Direct Vulnerability ☐ Indirect Vulnerability ■ Fragility	
Number of Buildings per Class:	8,859	
Sources of Data:	1980 Irpinia (Braga et al (1982) from CEQID (2013))	
Overall Rating of Quality of Data (per Source):	☐ High ☐ Moderate ☐ Poor ☐ Other(____)	
Definition of Loss Parameter:	Repair over replacement cost	
Intensity Measure (IM):	PGV in m/s	
Range of IM:	(0.08-1.3) m/s	
Evaluation of IM:	☐ Ground motion records (Number) ☐ GMPE (Ref) ■ ShakeMap ☐ Other(____)	
Indirect Vulnerability Assessment - Fragility Curves (FC)		
Damage Scale:	MSK-76	
Damage State (DS):	***ds_3***	***ds_3***
Description of DS:	**Moderate**	**Moderate**
Function of Non-parametric FC:		√
Function of Parametric FC:	Eq.(F.1)	
Parameter (ϑ_0):	0.19432	
Parameter (ϑ_1):	0.60559	
Confidence Intervals for mean model:	-	-
Statistical Model:	■ GLM	■ GAM
Statistical Model Fitting Procedure:	■ ML	■ ML
Data Type:	☐ Building-by-Building ■ Grouped data	
Number of Data Points:	41	41
Grouped Data: Definition of the aggregation unit:		
Grouped Data: Min Number of Buildings / Data Point:	3	3
Model Assumptions:		

Measurement Error in IM:	-	-
Measurement Error in Response:	-	-
Other(____________________):	-	-
Goodness of Fit Assessment:		
Mean function:	√	√
Variance function:	No	No
Procedure:		
Confidence Intervals:		

Discussion:

The empirical fragility assessment procedure is adopted here to estimate fragility curves for Italian stone masonry buildings with wooden floors. The adopted database has also been adoted in Appendix B where the use of GLM models has been illustrated. The aim of this application is to highlight issues with the use of the generalized additive models when the data are aggregated and sparse.

1st Step: Preparation of the damage data

***Data Quality Assessment*:**

Grouped damage data, aggregated per building type into 41 damage distributions each corresponding to a different affected municipality, were obtained from the CEQID (2013) . The original completed survey forms were not available but the survey method is reported in the literature. The 1980 Irpinia Earthquake database was constructed by a one-stage cluster sampling method (Levy and Lemeshow, 2008); i.e., the total number of buildings from 41 municipalities (out of more than 600 affected by the event) in the Campania-Basilicata area were surveyed (Braga et al., 1982). With regard to non-sampling errors, the comments found in the CEQID (2013) raise concerns on whether the total number of buildings in each commune has been surveyed. For the needs of this application we assume that this error is negligible. Overall, the quality of the database is considered moderate due to the aggregated type of data.

Data preparation per source:

The building classes account for the material of the vertical as well as the lateral load resisting structural components. Fragility curves are constructed using the largest and most vulnerable building class in this database, which consists of 8,859 field stone masonry buildings with wooden floors. The observed damage is classified into six discrete states, varying from no damage to collapse, according to MSK-76. This is the original damage classification used by Braga et al. (1982) in collecting the data. Figure F.1 highlights the significant (~25%) percentage of buildings which suffered heavy damage or collapse.

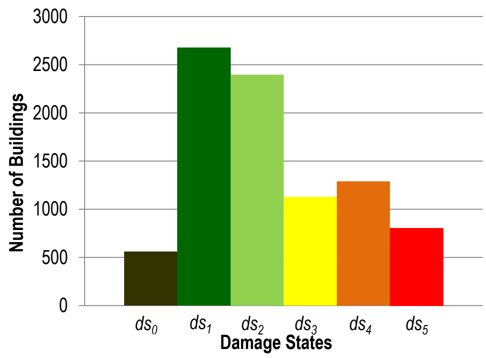


Figure F.1 Number of field stone masonry buildings with wooden floors that suffered damage in the 1980 Irpinia earthquake.

The construction of fragility curves using 'R' (R Development Team, 2008) requires the transformation of the grouped damage data into data points (x_j,(y_{ij},n_j-y_{ij})), where y_{ij} is the count of buildings which suffered $DS \geq ds_i$ and n_j-y_{ij} is the count of buildings which sustained $DS<ds_i$ for municipality j with intensity measure level x_j. Thus, 41 data points are obtained for each of the five damage states ds_i (i=1-5). The number of the buildings surveyed in each municipality varied widely from 3 to 1205. Six data points are seen to be based on very small numbers of buildings (<20). These six points are also included in the analysis and the goodness of fit diagnostics will determine whether they should be removed.

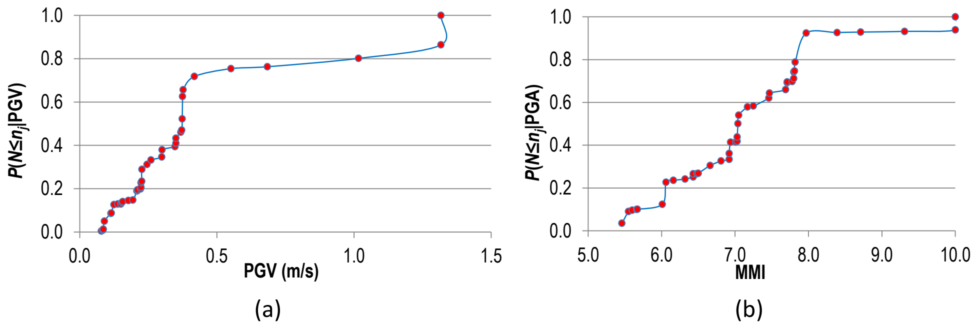


Figure F.2 Cumulative distribution of the proportion of the examined buildings exceeding the selected intensity measure values.

2nd Step: Selection and Estimation of the Intensity Measure

Two intensity measure types, namely PGV and MMI, have been selected. Their levels are estimated by a ShakeMap for the earthquake and are also provided in the CEQID (2013). The intensity measure values are assumed to have a single constant value within each municipality. This is considered a reasonable assumption given the relatively small surface area of each municipality (on average 24km^2). Nonetheless, the measurement error associated with these estimated intensity levels is not known, and therefore the measurement error in the intensity measure level estimated for each commune is ignored.

3rd Step: Selection of Statistical Model

The distribution of the data points in the range of the intensity measure levels can be used to determine an acceptable statistical model. For this reason, the cumulative proportion of buildings (for the 41 data sets) and their corresponding intensity measure levels in shown Figure F.2. In Appendix B, the relatively small number of data points and especially their sparseness over the higher intensity measure levels, lead the analyst to decide that a non-parametric model would be inappropriate. In this application, we examine the validity of this decision. For this reason, the fit of a GLM model is compared here to the fit of a GAM model.

The GLM model is expressed in the form:

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\begin{pmatrix} \dfrac{n_j}{y_j} \end{pmatrix}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \text{ where, } \mu=P\left(DS\geq ds_i \mid IM\right)=\Phi\left(\vartheta_0+\vartheta_1\log\left(IM\right)\right) \quad \text{(F.1)}$$

The GAM model uses the same link functionis expressed in the form:

$$Y \sim f\left(y \mid IM,\boldsymbol{\theta}\right)=\begin{pmatrix} \dfrac{n_j}{y_j} \end{pmatrix}\mu^{y_j}\left[1-\mu\right]^{n_j-y_j} \text{ where, } \mu=P\left(DS\geq ds_i \mid IM\right)=\Phi\left(\vartheta_0+\sum_{j=0}^{k-1}\vartheta_j b_j\left(IM\right)\right) \quad \text{(F.1)}$$

Three GAM models are fit to the data using the same link function and linear predictor as for the GLM:

- GAM.5.1 : for 5 knots and gamma=1.
- GAM.5.120: for 5 knots and gamma=120.
- GAM.15.1 : for 15 knots and gamma=1.

4th Step: Statistical Model Fitting Procedure

Estimation of the model parameters

The model is fit to the damage data using 'R'. Figure F.3 depicts the four curves. GAM.5.1 and GAM.15.1 are non-monotonic curves. It can also be noted that the higher the number of knots the more wobbly the fragility curve. This has a positive effect on the over-dispersion as depicted by the residual plots in Figure F.4b,d, which shows that the residuals of the GAM.15.1 appear to vary between [-3,3].

Does this mean that this model fits the data best? The answer is complicated. The fit might be indeed better but the predictive ability of the model is compromised as the curve is influenced by the rather sparkly distributed data. The trend depicted by GAM.15.1, however, is specific to the particular database and is unlikely to be the case if more data points are available.

Does this mean that the GAM.5.1 provides a better trend of the data that the GLM as well as GAM.15.1? Again, the non-monotonic trend in the data is a result of the particular database. In other words, the decrease in the probability of damage for increasing intensity measure levels can be attributed to the particular arrangement of the few data points, instead of an indication that the increase of PGA for low intensity measure levels is expected to decrease the probability of damage. For this reason, the author does not think can be considered realistic.

A non-strictly monotonic curve is fitted by the smoother curve of GAM.5.120. This appears to be intuitively better, however, it leads to a model that suffers from over-dispersion (see Figure 4c) and its trend is also highly influenced by the particular dataset and this challenges its ability to predict the probability of damage from new events. For these reason, the analysis believes that the GAM is any of the examined form is not better than its corresponding GLM for the proposed database.

Lessons

The use of generalised additive models is questionable for sparse aggregated data.

References:

Braga F., Dolce M., Liberatore D. [1982] ''Southern Italy November 23, 1980 earthquake: a statistical study of damaged buildings and an ensuing review of the M.S.K.-76 scale'', Report, Rome, Italy.

Levy P.S., Lemeshow S. [2008] ''Sampling of populations: Methods and applications'', Wiley.

R Development Core Team [2008] ''R: A language and environment for statistical computing'', Report, R Foundation for Statistical Computing, Vienna, Austria.

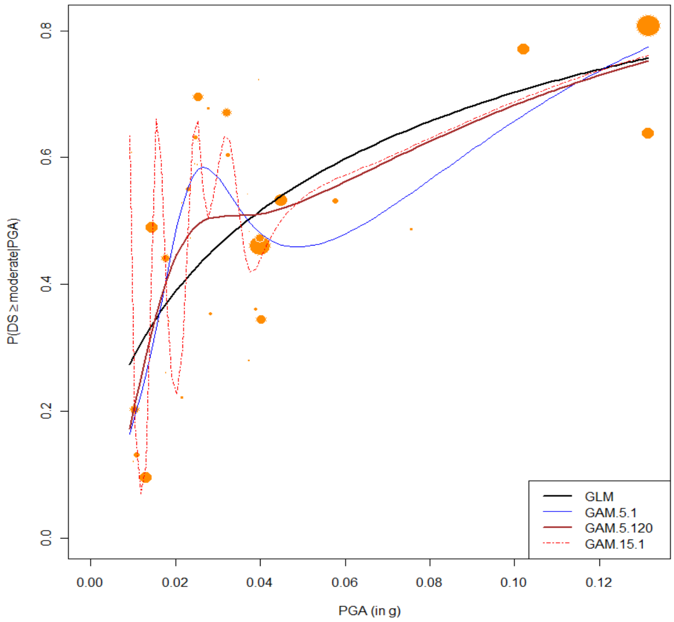


Figure F.3 Fragility curves corresponding to moderate damage constructed by GLM and GAMs.

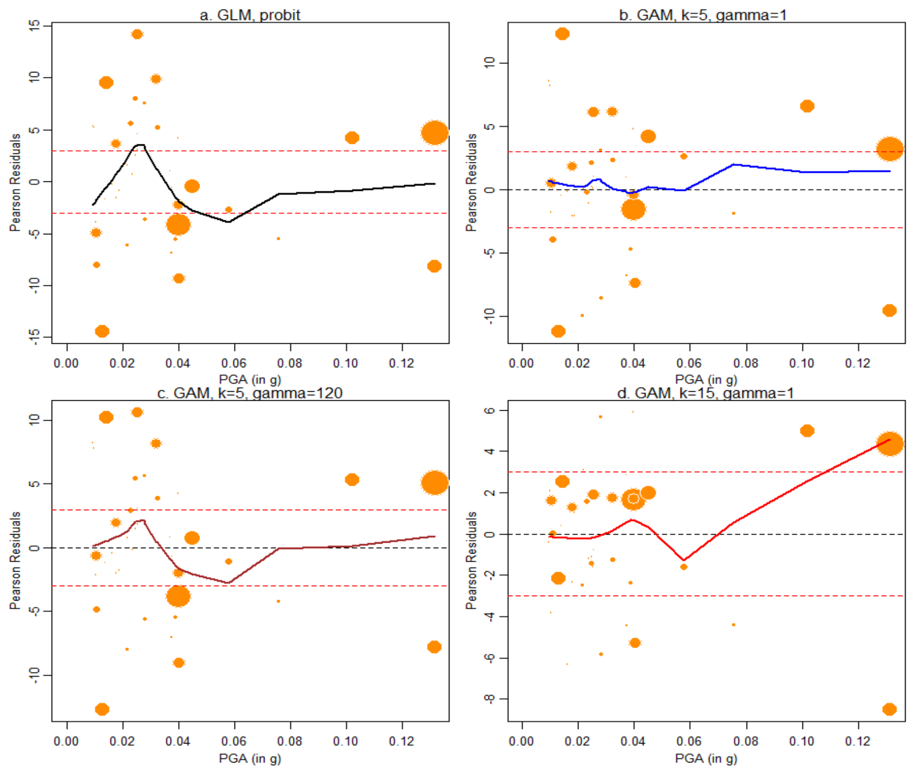


Figure F.4 Pearson residuals against the PGA for the GLM and GAMs.

APPENDIX G Fitting GKSs to damage data from the 1980 Irpinia Earthquake, Italy

The analyst is referred to ICOSSAR13 paper by Noh et al. (2013) for an illustration of fitting Gaussian kernel smoothers to post-earthquake damage data.

THE GLOBAL EARTHQUAKE MODEL

The mission of the Global Earthquake Model (GEM) collaborative effort is to increase earthquake resilience worldwide.

To deliver on its mission and increase public understanding and awareness of seismic risk, the GEM Foundation, a non-profit public-private partnership, drives the GEM effort by involving and engaging with a very diverse community to:

- Share data, models, and knowledge through the OpenQuake platform
- Apply GEM tools and software to inform decision-making for risk mitigation and management
- Expand the science and understanding of earthquakes.

GEM Foundation

Via Ferrata 1
27100 Pavia, Italy
Phone: +39 0382 5169865
Fax: +39 0382 529131
info@globalquakemodel.org
www.globalquakemodel.org

DECEMBER 2014

GLOBAL EARTHQUAKE MODEL
working together to assess risk