Advancements in the Seismic Risk Assessment of Mid-rise Reinforced Concrete Buildings

by

Stylianos Minas

A Thesis submitted for the degree of Doctor of Engineering
Department of Civil, Environmental and Geomatic Engineering
University College London

December 2018

Declaration

I, Stylianos Minas, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has

been indicated in the thesis.

Signature:

Date:

Abstract

Reinforced concrete (RC) buildings constitute a significant part of the existing stock in many areas exposed to high seismic risk, such as in urban centres in the Mediterranean region. Recent earthquakes in Italy confirmed the high vulnerability of RC buildings and highlighted the need for more reliable models to estimate the expected losses and quantify the possible risks. Such models are widely used in insurance industry and are based on the three main components of the risk equation, namely hazard-vulnerability-exposure.

This Thesis investigates the hazard and fragility/vulnerability components of the risk equation and systematically assesses the effect of these on the economic losses estimated for populations of mid-rise RC buildings within a catastrophe risk framework.

Regarding hazard, a variety of intensity measures (IM), including advanced and conventional IMs, are tested to determine the optimal IM to represent seismic hazard. Optimal IMs reduce uncertainty as they are better correlated with the buildings' seismic response, and allow a smaller number of analyses to be performed. Regarding fragility and vulnerability, catastrophe models require the assessment of large populations of structures; however, simplifications are needed to this procedure due to computational limitations. To this aim, the simplified analysis approach FRACAS bridging capacity spectrum assessment and fragility curve generation is presented alongside with a new software interface. Additionally, a Bayesian emulator-based approach (BEA) to efficiently compute the probabilistic seismic response and develop fragility functions is also proposed. The BEA is based on fewer assumptions and is shown to achieve better uncertainty characterization compared to standard methods.

Several combinations of the developed knowledge are applied to a probabilistic seismic risk assessment case-study to determine their influence on the resultant damage/loss prediction. It is

shown that the choice of IM and fragility function have a significant impact on the losses, with

the produced advanced IM and fragility models yielding the best results.

Impact Statement

Seismic risk assessment of reinforced concrete buildings is a widely studied topic for academic and commercial research within the field of earthquake engineering. Estimating the impact of a potential future seismic event in terms of human and economic loss is a crucial topic for the mitigation of earthquake risk, and is of interest to various stakeholders, including insurers/reinsurers, insurance brokers, engineers, policy makers and governments. This work contributes to better seismic risk and resilience assessment prediction and can be used for assessing the effect of potential earthquake events on the social, economic, political, and cultural environment. The impact of this study can be divided into economic, societal and academic impact.

Regarding economic impact, the outcomes of this study may provide guidance for the development of a more robust seismic risk assessment framework, which can be implemented within insurance/reinsurance applications (such as catastrophe models), but also for consultancy projects. This improved framework will result in a better seismic risk characterization, which will allow insurance and consultancy firms to develop more sophisticated tools and offer competitive services to the global market.

The advancements presented in this study may also have a societal impact. Such advancements can be ultimately used for the development of effective emergency planning, through the allocation of available resources, and disaster risk response, through the implementation of prompt seismic response assessment tools. In addition, an advanced framework will better identify the various risks associated with the area of interest and will potentially highlight the need for seismic retrofitting of buildings and risk mitigation strategies.

With respect to academic impact, this study shed light on the limitations associated with the components of seismic risk assessment framework and proposed new methods and tools for the enhancement of future risk models. This research work resulted in three published journal papers with another one under review.

Acknowledgments

my EngD programme. To Dr. Fatemeh Jalayer, my MPhil examiner, for the constructive feedback and the suggestions which helped me to improve my final Thesis. I would like to thank Dr. Michalis Fragkiadakis, for the scientific advice and the insightful conversations we had during the time I spent at NTUA.

To my colleagues and dear friends from the EPICentre family, Alexandra, Arash, Athanasios, Chen, Crescenzo, David, Ioanna, Ivy, Omar, Pierre, Sam, Sascha, Silvia and Valentina, I would like to express my sincere appreciation for the all support and the moments we shared together: the interesting scientific (or not so scientific) discussions, the lunch breaks, the "surprise" parties.

Special thanks to Dr. Dimitris Katsikadakos, my best companion, my flatmate and recently my "Koumparos". You know you are the main reason I am writing these lines, after all you are probably the person who knows me best. To my "brother" Dr. Kostas Zarogoulidis, my true soulmate, thank you for everything you did for me "Manako".

To all the wonderful friends I made in London, Bahman, Aris, Evi, Ilias, Panagiotis and Vasilis. Thank you for all the amazing times we had together, the many laughs and the great nights out (or at home).

To my best friends from Greece, Andreas, Argyris, Eleni and Kostis, thank you for being by my side all these years, even though you were a few thousand kilometres away. You always found the right ways to take my stress away and make me smile.

I am eternally grateful to Tania, my beautiful girlfriend, whose love and support have been my biggest motivation in these years. This achievement would have been meaningless without the chance of sharing all the happy moments with you.

Finally, I would like to dedicate my work to my family, in particular to my parents, Giorgos and Alexandra, and my sister, Dimitra. Thank you for your unconditional love and support throughout these long years. I wouldn't make it without you.

London, December 2018

Stelios

Table of Contents

List of Figures

20

List of Tables

List of Abbreviations

Abbreviation	Meaning
AAL	Average annualized loss
ADRS	Acceleration displacement response spectrum
ANN	Artificial neural network
AP	Analysis points
APO	Adaptive pushover
BEA	Bayesian emulation-based approach
BN	Bayesian network
CDF	Cumulative distribution function
CMS	Conditional Mean Spectrum
CSM	Capacity spectrum method
DCH	High ductility class
DM	Damage measures
DoE	Design of experiments
DS	Damage state
DSHA	Deterministic seismic hazard assessment
EDP	Engineering demand parameter
EERI	Earthquake Engineering Research Institute
EMS-98	European Macroseismic Scale 1998
EP	Exceedance probability
EPH	Elastic-plastic with positive strain-hardening capacity idealization model
EPP	Elastic-perfectly plastic capacity idealization model
ESD	European Strong-motion Database
FBCC	Full-block cross correlation model
FEMA	Federal Emergency Management Agency
FRACAS	FRAgility through Capacity spectrum ASsessment

FRM	Force-resisting mechanism
GLM	Generalized linear model
GMPE	Ground motion predicting equation
GP	Gaussian process
GUI	Graphical user interface
HRC	Homogenized Reinforced Concrete damage scale
IDA	Incremental dynamic analysis
IM	Intensity measure
ITACA	ITalian ACcelerometric Archive
ISSARS	Integrated System for Structural Analysis and Record Selection
LS	Least squares
MDoF	Multi degree of freedom
MDR	Median damage ratio
MIDR	Maximum inter-storey drift ratio
MLE	Maximum likelihood estimation
MRF	Moment resisting frame
MSA	Multiple Stripes Analysis
MSE	Mean squared error
NC	No correlation model
NDA	Nonlinear dynamic analysis
NSP	Nonlinear static procedure
OSM	Open Street Map
PAGER	Prompt Assessment of Global Earthquakes for Response
PBEE	Performance-Based Earthquake Engineering
PCE	Polynomial chaos expansion
PDF	Probability density function
PEER	Pacific Earthquake Engineering Research center
PO	Pushover

PP	Performance point
PSHA	Probabilistic seismic hazard assessment
PSRA	Probabilistic seismic response analysis
RC	Reinforced concrete
RD	Roof drift
RDR	Roof drift ratio
ReLUIS	Rete dei Laboratori Universitari di Ingegneria Sismica
SDoF	Single degree of freedom
SEAOC	Structural Engineers Association of California
SIMBAD	Selected Input Motions for displacement-Based Assessment and Design
SRSS	Square-root-of-sum-of-squares
SSCM	Simple spectral correlation model
TL	Trilinear capacity idealization model
TRI-PO	Triangular pushover
UHS	Uniform Hazard Spectrum
UNI-PO	Uniform pushover
VGI	Volunteered Geographic Information
WHE	World Housing Encyclopedia

List of Symbols

Symbol	Meaning
AL_i	Annual loss for the year number i
$AvgSA$	Averaged logarithm of spectral accelerations over a range of periods
a_i	Spectral acceleration at the ith analysis point
a_u	Spectral acceleration at ultimate point
a_y	Spectral acceleration at yielding point
a, b	Power-law regression parameters
a_1, b_1 and b_2	Power-law bilinear regression parameters
$a(T)$, $b(T)$	Period-dependent model parameters associated with the attenuation level of the spatial correlation with distance
$\mathbf{C}$	Positive-definite correlation matrix
c	Calibration coefficient associated with the advanced IM S_a^c
D	Source-to-site distance
DS_i	The ith damage state
d_i	Spectral displacement at the ith analysis point
d_u	Spectral displacement at ultimate point
d_y	Spectral displacement at yielding point
EDP	Conditional median of the demand given the IM
e	Zero-mean random variable representing the variability of ln(EDP) given the IM
edp_i	EDP calculated through FRACAS corresponding to the i-th ground motion
$\mathbf{F}$	Basis matrix of the surrogate model
$f(.)$	Simulator function
$\hat{f}(.)$	Statistical emulator function
G	Vector of simulated values used for the evaluation of kriging predictions
$\hat{G}$	Vector of predicted values by the kriging model

Symbol	Meaning
$g(.)$	Link functions used in GLM model
g	Acceleration of gravity (9.81 m/s^2)
h	Separation distance between two different locations used for the estimation of $\rho_h(T)$
I_A	Arias intensity
I_{Np}	Spectral-shape based IM proposed by Bojórquez and Iervolino (2011)
$I_{T_{min}<0.25}$	Indicator function equal to one when $T_{min}<0.25$, or equal to zero otherwise
$I(EDP|IM_2|IM_1)$	Relative sufficiency of a second IM (IM_2), with respect to a first one (IM_1)
IM_{ij}	Ground motion intensity value for an event i at a site j
IM_0	The IM threshold of the demand regime change (used in bilinear cloud regression)
$IM_{1I\&2E}$	Displacement-based combination type IM proposed by Luco and Cornell (2007)
k	Confinement factor
$\mathbf{L}$	Lower triangular matrix obtained from Cholesky decomposition, such that $\mathbf{LL}^T = \mathbf{C}$
L	Metric of loss
L_i	Repair cost assigned to each one of the defined damage states DS_i
$LL95_i$	Lower 95% limit of the ith interval
M^*	Effective modal mass for the fundamental mode of vibration
M_i	Magnitude of an event i
M$_w$	Moment magnitude
ME_i	Mean percent error
$m(t)$	Term representing the systematic variation of the output with the IM, as a Gaussian process
N	Total number of floors
N_{DS}	Number of damage states

Symbol	Meaning
N_{GM}	Total number of ground motions
N_p	Spectral-shape proxy parameter proposed by Bojórquez and Iervolino (2011)
N	Number of prediction intervals
n_m	Number of the observations $\mathbf{x}$ - $\mathbf{y}$ used to formulate the surrogate model
n_s	Number of year simulations of seismicity
n_x	Dimension of surrogate model input vector
n_y	Dimension of surrogate model output vector
PGA	Peak ground acceleration
PGD	Peak ground displacement
PGV	Peak ground velocity
$\mathbf{R}$	Correlation matrix of the surrogate model
R	Epicentral distance
$R_{T1,T2}$	Spectral shape parameter equal to the ratio of $S_a(T_2)$ over $S_a(T_1)$
R_{ij}	Epicentral distance from seismic source for an event i at a site j
RD_i^2	Coefficient of determination
$S_a(T_1)$	Spectral acceleration at natural period T_1
$S_a(T_1)^*$	Target spectral acceleration at natural period T_1 used in CHS
$S_a(T_2)$	Spectral acceleration at the second period
$S_{agm}(T_i)$	Geometric mean of spectral acceleration values at different periods, proposed by (Kazantzi and Vamvatsikos 2015)
S_{a12}^* , S_{a123}^*	A weighted average of spectral accelerations at the first two or three periods respectively, proposed by Shome and Cornell (1999)
S_a^c	Average spectral acceleration IM proposed by Cordova et al. (2001)
$S_d(T_1)$	Spectral displacement at natural period T_1
$S_v(T_1)$	Spectral velocity at natural period T_1

Symbol	Meaning
SI	Spectrum intensity proposed by Housner (1952)
$\mathbf{s}$	Surrogate model parameter vector
s	Power law regression's standard deviation
$t(\mathbf{x})$	Intensity measure term as a function of ground motion sequence
T_1	Natural period of vibration
T_C	Corner period of design response spectrum
T_{max}	Maximum period of vibration
T_{min}	Minimum period of vibration
$UL95_i$	Upper 95% limit of the ith interval
u_N	Top floor (N -floor) displacement
V_b	Base shear force
VSI	Velocity spectrum intensity proposed by Von Thun et al. (1988)
$\mathbf{X}$	Input matrix of the surrogate model
$\mathbf{x}$	Vector of inputs of the surrogate model (e.g., representing a ground motion sequence)
y_i	Binary indicator to assess the performance in GLM
$\hat{y}_k(\mathbf{x})$	Kriging predictor mean
Y	Response variable
$\mathbf{Y}$	Output matrix of the surrogate model
$\mathbf{Y}_G$	Vector of multivariate Gaussian distributed random field
$\mathbf{y}$	Vector of outputs of the surrogate model
$\mathbf{Z}$	Vector of independent Gaussian distributed random variates
$z(t)$	Approximation error function
$\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}^*$	Parameters for the approximation of kriging mean
α	Calibration coefficient associated with the advanced IMs S_a^c and I_{Np}

Symbol	Meaning
β	Standard deviation
β_c	Standard deviation linked to the structural capacity
β_e	Standard deviation of the residual of the simple linear regression model
β_d	Standard deviation associated with uncertainty in demand
β_{DS}	Standard deviation in the threshold defining a specific damage level
$\beta_{\ln EDP|IM}$	Standard deviation of the natural logarithm of *EDP* given *IM*
$\beta_{\ln imDS}$	Standard deviation of the natural logarithm of intensity
Γ	Modal participation factor
ε	Number of standard deviations by which an observed logarithmic spectral acceleration differs from the mean logarithmic spectral acceleration of a GMPE
$\varepsilon(.)$	Discrepancy term
ε_k^g	Standard Gaussian variable with zero mean and a variance of one
θ	Mean of the cumulative lognormal distribution
θ_{ij}	Parameter associated with the faulting types, site classes and other predictors for an event *i* at a site *j*
$\mu(t)$	Term representing the systematic variation of the output with the IM
$\mu_{EDP|IM}$	Median estimate of the natural logarithm of the *EDP* given *IM*
μ_{Fr}	Conditional mean response representing the fragility curve in GLM
μ_{imDS}	Median damage state intensity
ξ	Damping ratio
$\rho(h_{i,j})$	Coefficient of correlation of residuals at two locations separated by a distance $h_{i,j}$
$\rho_h(T)$	Coefficient of correlation of the intra-residuals at two different locations
$\rho_\tau(T_1, T_2)$	Coefficient of spectral correlation in the inter-event residual
σ^2	Variance of the Gaussian process term $m(t)$

Symbol	Meaning
$\sigma_k(\mathbf{x})$	Kriging predictor standard deviation
σ_k^2	Diagonal elements of the matrix $\left(\mathbf{Y} - \mathbf{F}\boldsymbol{\alpha}^*\right)^T \mathbf{R}^{-1}\left(\mathbf{Y} - \mathbf{F}\boldsymbol{\alpha}^*\right)/n_m$
$\sigma\varepsilon_{ij}$	Intra-event variability of the GMPE model for an event i at a site j
τ^2	Variance of the term $\mu(t)$
$\tau\nu_i$	Inter-event variability of the GMPE model for an event i
Φ	Standardized Gaussian CDF
φ	Parameter controlling the rate of decay of the correlation
ϕ_j	The jth floor element of the fundamental mode shape (ϕ_1)

1. Introduction

1.1 Preamble

In the last few decades, major earthquakes have been responsible for an annual death toll of over 60,000 people and heavy economic losses, significantly impacting the world's economy (Silva 2013). Such an impact becomes even more devastating in the case of large, densely-populated cities. Reinforced concrete (RC) buildings constitute a significant part of the existing stock in many areas exposed to high seismic risk, such as in urban centres in the Mediterranean region (Ricci 2010). Despite major advancements in seismic design standards, building codes and practices, recent earthquakes in Italy, Turkey and Greece have resulted in extensive damage to the built environment and numerous human losses. In particular, recent events in these countries confirmed the high vulnerability of RC buildings and highlighted the need for more reliable models to estimate the expected losses and quantify the possible risks (e.g., Rossetto et al. 2009, 2012).

Loss models are widely used in the insurance industry (the so-called catastrophe risk models) for performing (seismic) risk analysis. These models estimate the expected losses at a given set of location/portfolio of buildings for a given return period by integrating the seismic hazard, the vulnerability of the built environment and the exposed value, i.e. the three main components of the risk equation.

Quantitative assessment of seismic risk is a very broad field, including seismology and geotechnical engineering to quantify the shaking that a structure might experience at its base (seismic hazard), structural engineering to quantify the structure's response and the resulting damage (vulnerability), as well as finance and public policy to quantify social and economic consequences of earthquake induced damage (exposure). Several sources of uncertainty are identified within all these aspects, requiring that the assessment is made in probabilistic terms, adding a further layer of complexity.

Different approaches for analytical seismic risk assessment have been proposed by several authors. The seismic risk assessment approach proposed by Giovinazzi et al. (2014) is comprised of six steps. The first step to classify the building stock based on an existing or an *ad hoc* taxonomy. Second step is to define an appropriate damage scale for the chosen building stock. The third step is to determine a seismic hazard parameter (i.e. IM) to evaluate damage. The fourth step is to generate the hazard scenarios to be used in the seismic risk assessment. The fifth step is to assign damage to buildings by carrying out a probabilistic seismic response analysis (PSRA) and consequently derive sets of fragility functions. Finally, the sixth step involves estimating the direct losses and impacts related to the seismic-induced physical damage to buildings.

The Pacific Earthquake Engineering Research (PEER) center has developed a framework for assessing the seismic risk to single structures. This is a widely used framework for seismic risk assessment, known as the Performance-Based Earthquake Engineering (PBEE) model (Deierlein et al. 2003), and is described here in order to introduce some of the terminology later. Though developed for the assessment of individual structures, the PBEE framework has been extended to applications for large urban areas and constitutes the basis for some earthquake catastrophe models (Liel and Deierlein 2012). The PBEE framework consists of four basic elements, which are associated with quantifying the seismic ground motion hazard, assessing the structural response, estimating the damage to building and contents and resulting consequences. These four independent components, which are described in detail in the following paragraph, are linked together using intermediate output variables, namely intensity measures (IMs), engineering demand parameters (EDPs) and damage measures (DMs). The IM defines, in a probabilistic sense, the features of the ground motion hazard that affect structural response. EDP describes the structural response in terms of deformations, accelerations, or other response quantities calculated by simulations of the building to earthquake ground motions. The DMs describe the physical condition of the structure and its components as a function of the imposed EDPs. Once the DMs are determined, the final step in the PBEE framework is to

calculate decision variables (DV), which translate DMs into quantities that relate to risk management decisions concerning economic loss and life safety. The reference equation that can summarize the concept of PBEE:

$$\lambda(DV) = \iiint \left| G\langle DV \mid DM \rangle dG\langle DM \mid EDP \rangle dG\langle EDP \mid IM \rangle d\lambda(IM) \right| \qquad (1.1)$$

Equation (1.1) is an application of total probability theorem, where uncertainties at each step of the process are described in terms of independent conditional probabilities.

Specifically, the first component of PBEE equation is the hazard analysis, where the mean annual rate of exceedance for a given ground motion intensity measure, $\lambda(IM)$, is evaluated for a site, taking into account the seismic sources and soil conditions. The second component is associated with the probabilistic response $G\langle EDP \mid IM \rangle$ which is determined through a set of analyses of a structural model (or multiple structural models). The third component is the damage analysis, where fragility functions are used to compute the conditional probability, $G\langle DM \mid EDP \rangle$, that a structure's component (or multiple components) will reach or exceed a specific damage state, as defined by DM. This conditional probability, when represented graphically, is known as fragility curve. A fragility curve represents the probability that an engineering demand parameter exceeds a predefined threshold as a function of a ground motion IM. Finally, the fourth component of PBEE equation focuses on the conditional probabilistic losses, $G\langle DV \mid DM \rangle$.

The seismic hazard component is widely perceived to be the one characterized by the largest aleatory and epistemic uncertainties within the seismic risk assessment framework. This is particularly evident for the case of structure-specific applications; however, several recent studies have shown that probabilistic response analysis and fragility assessment significantly contribute to the total uncertainty of seismic loss estimates, particularly when considering portfolio-level applications (e.g., Silva (under review)). The focus of this Thesis is on the

probabilistic response and fragility analysis components of the seismic risk assessment framework.

1.2   Research questions and scope

As described above, analytical seismic risk assessment requires substantial computational effort and involves large uncertainty within each step of this process. This Thesis aims to scrutinize the fragility component of the seismic risk with particular application to the assessment of populations of mid-rise RC buildings. The research questions investigated are:

1.   How to compute reliably the seismic response over a population of buildings without prohibitive computational expense?

2.   Which IM is best suited for using in fragility analysis of mid-rise RC buildings?

3.   What approach could be used to better capture the epistemic uncertainty in the EDP|IM component of analytical fragility functions?

4.   What is impact of different input assumptions and methods used in the estimation of loss within seismic risk framework?

In order to investigate the research questions addressed above, the first step is to critically review each component within the seismic risk in order to highlight the limitations of current methods. Based on the identified limitations, new advancements and techniques are proposed herein, which better account for uncertainty propagation at each step of the risk framework, and consequently produce more reliable estimates of risk. Within the scope of this research, mid-rise RC building models representing typical European and Mediterranean construction of different vintages are considered for the illustration of the proposed techniques. Finally, a probabilistic seismic risk assessment case study is carried out where the developed knowledge is applied in several combinations to determine their influence in the resultant damage and economic loss prediction. It is noted that the proposed techniques can be potentially applied to other structural types through appropriate building and uncertainty modelling within the analytical seismic risk framework.

From an investigation of the Research Questions, the main contributions of this Thesis to the state-of-the-art are summarized as:

- To propose an alternative analysis method to calculate seismic response within catastrophe risk framework. To this aim, the variant of capacity spectrum method (CSM), FRAgility through Capacity spectrum ASsessment (henceforth FRACAS) is introduced. FRACAS overcomes some of the limitations of CSM currently used in catastrophe models, but maintains the advantage of rapid analysis times. A user-friendly graphical user interface (GUI) version of FRACAS tool is also developed.

- To reduce the uncertainty within the hazard component through the implementation of an optimal IM parameter. Optimal IMs reduce uncertainty as they are better correlated with the buildings' seismic response, and therefore are more efficient, as they allow a smaller number of analyses to be carried out.

- To propose an advanced statistical approach to compute efficiently the probabilistic seismic response and develop fragility functions. A Bayesian emulator-based approach (hereafter BEA) is adopted to compute the mean estimates, and the associated variances, for the conditional distribution of EDP|IM relationship. This approach relies on fewer assumptions, and offers better uncertainty characterization in structural response to earthquakes comparing to standard methods.

- To determine the influence of different input assumptions, in terms of hazard and vulnerability modelling options in the estimation of seismic loss assessment of RC buildings. A probabilistic regional seismic loss case study is performed utilizing several combinations of modelling options (including the developments proposed in this Thesis and methods used in current catastrophe modelling practice).

The outputs of the present Thesis (including research publications, software and research material) are summarized in Appendix A: List of publications and outputs.

1.3   Thesis structure

The present Thesis consists of seven additional Chapters.

Chapter 2 provides a detailed review of the technical concepts involved within the seismic risk assessment framework based on the extended procedure proposed by Giovinazzi et al. (2014).

Chapter 3 presents the three mid-rise RC case study buildings which are generated for the scope of this Thesis. The three buildings, which represent typical Italian buildings corresponding to different vulnerability classes (i.e. designed according to different seismic provision levels), are analysed with the aim of simplified and advanced analysis to determine their corresponding seismic response, and consequently construct the associated fragility functions. These buildings are then considered for the illustration of techniques and sensitivity tests proposed in the Chapters to follow.

Chapter 4 introduces the simplified analytical method for fragility assessment FRACAS. This Chapter begins with a thorough description of FRACAS methodology, including the performance point determination and the derivation of fragility curves. The results from FRACAS are then evaluated against the results obtained from nonlinear time-history analysis . The effect of record-to-record variability on fragility curves generated by FRACAS is also investigated. Last, the FRACAS tool -a user-friendly GUI version of FRACAS- is introduced. This Chapter intends to answer Research Question 1.

Chapter 5 investigates the selection of optimal IM for fragility analysis, with particular focus on the simplified analysis method FRACAS. A number of conventional IMs (commonly used for fragility and catastrophe modelling purposes) and advanced IMs are considered in this Chapter. The performance of the chosen IMs is then assessed with the aim three IM selection criteria, namely (1) efficiency, (2) sufficiency/ relative sufficiency and (3) hazard computability, in order to determine the optimal IM. The aim of this Chapter is to answer Research Question 2.

Chapter 6 discusses the development of a new advanced statistical approach (i.e. BEA) to efficiently compute the probabilistic seismic response and develop fragility functions. An overview of the Bayesian emulation framework is first provided, extended to be applied in the context of fragility estimation. A number of sensitivity tests are then performed to highlight the sensitivity of the BEA to several input assumptions and reveal the range of applicability of the approach. In the last part of this Chapter, the process for deriving fragility functions using BEA is presented. This Chapter provides an answer to Research Question 3.

Chapter 7 investigates the impact of hazard and vulnerability modelling options to the estimation of probabilistic regional seismic loss assessment of RC buildings. A synthetic building stock representing a typical Mediterranean city (i.e. Avellino city in the Campania region of Southern Italy) is utilised as a case study, and the details of the selected hazard, vulnerability and exposure models are discussed. The sensitivity of seismic losses to input assumptions (including: choice of IM, analysis method, fragility derivation method) is then assessed. All the developments presented in previous Chapters are put into test against commonly used practices in seismic risk in order to determine their influence in the resultant damage and economic loss prediction. This Chapter answers Research Question 4.

Finally, Chapter 8 summarizes the main findings of the Thesis and concludes by recommending future directions of investigation.

Appendix A presents the list of publications and outputs of this research work, Appendix B gives an overview of existing damage scales, and Appendices C, D and E provide additional figures and tables for Chapter 3, Chapter 4 and Chapter 7 respectively.

2. Literature Review

2.1 Introduction

The aim of this Chapter is to provide a review of the state-of-art in numerical seismic risk assessment of mid-rise RC building populations. Each step of the seismic risk assessment framework for this building type is described, highlighting research developments in this field. The existing literature is critically reviewed and current technical gaps and shortcomings are identified having in mind large scale risk assessment applications, such as catastrophe models. Particular emphasis is put on the methods for estimating structural response, the selection ground motion parameters to describe hazard intensity and the statistical methods used to determine the relationship between a building's response and the hazard intensity.

2.2 Background

Although the seismic performance of a single building can be assessed by a detailed structural analysis, the assessment of seismic-induced damage and loss to populations of buildings (i.e. hundreds or thousands of buildings within an urban or regional area) relies on probabilistic procedures. This is necessary in order to account for the many uncertainties associated with the quantification of seismic hazard, the identification of different building types and their characteristics within the building population, the relationship between seismic hazard and physical damage, and the relationship between the physical damage and socioeconomic consequences.

According to Giovinazzi et al. (2014), the seismic risk assessment procedure for buildings constitutes of six steps (**Figure 2.1**).
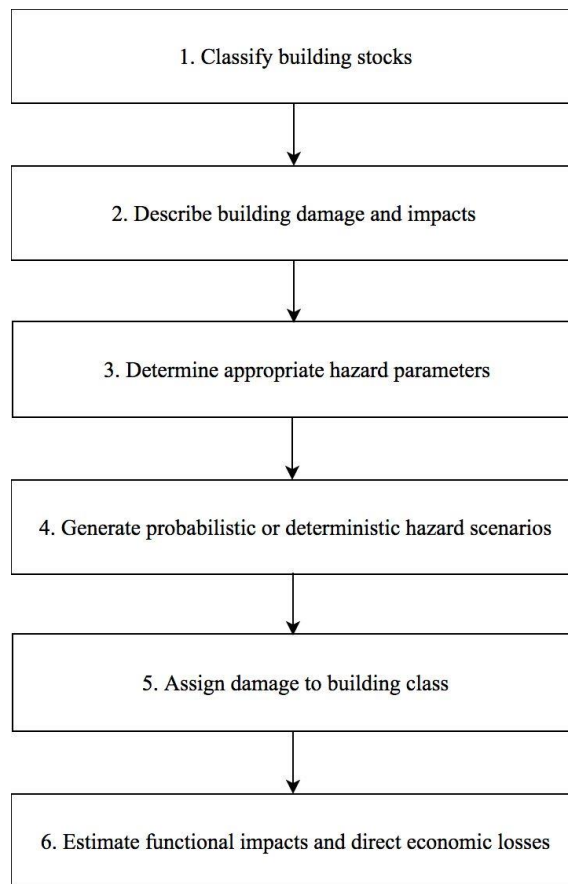
Figure 2.1 - Seismic risk assessment steps for buildings (adapted from: Giovinazzi et al. 2014).

The first step is to classify the building stock, so that buildings that are expected to have a similar seismic performance. To this aim, it is necessary to select an existing building taxonomy or define an *ad hoc* building taxonomy, classify the building within the taxonomy, and count the buildings in each one of the identified building classes. The second step is to define the appropriate scales for classifying damage to building structural and nonstructural components. The third step is to determine the appropriate earthquake hazard parameters against which damage should be evaluated. The fourth step is to generate the hazard scenarios (e.g., ground motions and land deformations) for which the building damage needs to be assessed. The fifth step is to identify the appropriate hazard-damage relationship and assign damage to buildings. The most commonly used tool to relate building damage to seismic hazard probabilistically is the fragility function, (also known as fragility curve when presented graphically). For a given level of ground motion intensity, fragility functions determine the probability that a building or a group of buildings will be in, or exceed a certain damage state. Finally, the sixth step is to

estimate the direct losses and functional impacts related to the earthquake-induced physical damage to buildings.

In the next section each step of the procedure is described in more detail in the context of existing literature.

2.3 Classification of buildings within a building population

The first step to predict expected damage to buildings, in the event of an earthquake, is to develop an accurate classification of the building stock. The purpose of classifying building stock is to group together buildings that can be expected to behave similarly, sustaining a comparable level of damage, when subjected to comparable levels of earthquake-induced ground shaking. To this aim, relevant data and information on the building stock should be collected. A non-exhaustive overview of the steps needed to classify building stocks for seismic risk analysis purposes is presented in the following subsections.

2.3.1 Building taxonomies

Building taxonomies classify buildings into different groups based on their material of construction, lateral load resisting, geometry, age/maintenance conditions, level of code enforcement, height and many other factors that can influence their seismic behaviour. Additionally, building taxonomies group buildings by their occupation and social function (e.g., residential, commercial, industrial). In fact, after the seismic-induced physical damage to building, the number of casualties, the impact on content and machineries, the direct economic losses, and the business interruption aspects are strictly correlated with the building occupation and social function.

First building taxonomies grouped building typologies according to their construction material only. For instance, Medvedev–Sponheuer–Karnik (MSL-64) macroseismic scale (Medvedev and Sponheuer 1969) identified three different building classes based their material (**Table 2-1**).

Table 2-1 - Building classes according to MSL-64 macroseismic scale (Adapted from: Medvedev and Sponheuer 1969).

Building class	Description
A	Buildings with dry stone or clay, adobe, or mud walls
B	Buildings with brick walls, mortar blocks, masonry and mortar, stone block, timber frame
C	Buildings with metal structure or reinforced concrete

The need for a deeper diversification of building behaviours, when subjected to an earthquake loading led to the definition of more elaborated classification systems, such as the one proposed by HAZUS (NIBS 2003) and European Macroseismic Scale EMS-98 (Grünthal 1998) for the United States and Europe respectively. In addition to the construction material, the aforementioned taxonomy systems also considered the building structural system, the design code level, the building height (low-, mid- and high-rise) and the presence of nonstructural elements.

Recent international projects and initiatives worldwide have developed and defined more refined taxonomy systems in order to better represent the variety of building types around the world. Pitilakis et al. (2013) proposed a modular and flexible taxonomy for classifying European masonry and reinforced concrete buildings, where buildings are classified according to some main categories, including the force-resisting mechanism (FRM), material of FRM, plan (regular or irregular), elevation (regular or irregular), cladding, detailing (ductile, non-ductile, with tie-rod beam, without tie-rod beam), floor system, roof system, height level, code level, and material. Subcategories (e.g., FRM, material of FRM, cladding characteristics, floor system material, roof system material, number of stories) were introduced for some of the aforementioned categories, to provide a more detailed classification, when more accurate data are available for the building stock under analysis. The Prompt Assessment of Global Earthquakes for Response (PAGER) project (Porter 2005) developed a comprehensive building taxonomy with a global coverage including structural types common in developing countries (i.e. adobe, clay, poor quality wood constructions). The Earthquake Engineering Research

Institute (EERI) developed a World Housing Encyclopedia (WHE) (EERI 2000) collecting information related to housing construction practices in seismically active areas worldwide. The Global Earthquake Model (GEM) developed a global building taxonomy aiming to represent all the building types around the globe (Brzev et al. 2013; Gallagher et al. 2013) based on the findings from WHE and on a comprehensive review of existing taxonomies (Charleson 2011). The effectiveness of existing taxonomy to represent building typologies worldwide was checked against different criteria (Porter 2005) aiming to define a simple, versatile, and detailed taxonomy, international in scope, user-friendly, and able to distinguish differences in seismic performance, but extensible to future needs (e.g., to other natural hazards). Following this comprehensive and critical review, the SYNER-G taxonomy (Pitilakis et al. 2013) was selected as the benchmark to define the GEM building taxonomy. The GEM taxonomy is organized as a series of expandable tables, describing various attributes at increasing levels of detail.

Further to taxonomies for building structures, specific building taxonomies have been extended to incorporate building contents and nonstructural components. Specifically, HAZUS (NIBS 2003) provided a classification for architectural components (e.g., veneer and finishes, cantilever elements and parapets, etc.), mechanical and electrical components (e.g., piping systems, elevators, etc.), and contents (e.g., file cabinets, bookcases etc.). NISTIR 6389 (Charette and Marshall 1999) provided taxonomy for shell, interiors, services, equipment, furniture, and special constructions. Porter (2005) proposed an expendable and flexible taxonomy for nonstructural elements based on the review, combination, and improvement of the existing ones, including NISTIR 6389 (Charette and Marshall 1999).

2.3.2 Building stock inventories and classification

Preparation of a building stock inventory is a critical part of a damage and loss estimation study, and can be both time and resource heavy. More refined taxonomy approaches require immense amount of data to describe the characteristics of a building typology and consecutively a building stock, therefore, they cannot commonly be used for large scale risk assessments. As a

result, one needs to compromise between the time and cost required for higher-quality data on the building stock and the uncertainties arising from low-quality data.

The first step to develop an "affordable" (in terms of time and computational cost) inventory is to collect data from all the available sources, including databases belonging to state, regional, and local authorities, private sector sources, as well as inventory information arising from previous loss or available within seismic risk assessment tools or platforms, e.g. HAZUS (NIBS) 2003). For example, census data can provide aggregated data on the size, the number, and the age of residential buildings, while databases prepared for insurance or civil defence planning purposes can provide information on commercial and/or strategic buildings. It is noteworthy to mention that these data are often overlapping and incomplete and that additional work might be required to reconcile multiple sources and missing data.

Apart from the ready-to-use data sets discussed above, different methods exist for collecting, processing, and archiving building inventory data. According to Tenerelli and Crowley (2013), the most widely used approaches include (1) ground screening procedures (e.g., field reconnaissance, sidewalk survey, street-front direct visual observation, interpretation of street-front photographs, etc.), (2) satellite remote sensing approaches, or (3) a combination of ground screening procedures and satellite remote sensing approaches. Ground screening procedures collect building attributes through direct observations, usually via (a) paper forms, *ad hoc* formatted to quickly record relevant building characteristics and attributes (e.g., AeDES form - Baggio et al. 2007; FEMA 154 2015), and (b) handheld equipment and software (e.g., tablet, phone) that enables user-defined electronic inventory data. For both the approaches, a critical aspect is the definition of the survey form that has to be coherent with the adopted building taxonomy (see Section 2.3.1) and with the adopted approach to assign damage to buildings as a function of the ground motion. A proper training for the staff performing the ground screening survey is essential since identifying structural types and other building attributes from the street can be challenging. Satellite and airborne remote sensing techniques may provide information on building stocks, e.g., building footprint and height information, in a timely and cost effective

manner. However, these techniques must be closely calibrated to the particular region under study, and the analysis must be supported by expert judgment, coupled with some field reconnaissance. Finally, an alternative approach for accessing building inventory data is through open-access datasets and Volunteered Geographic Information (VGI). One of the most commonly used VGI is OpenStreetMap, hereafter OSM, a freely accessible geographic database of the entire world (OpenStreetMap 2017). OSM provides supplementary or alternative sources of spatially-based building information and can be used for the development of exposure models for catastrophe risk assessment (Sousa et al. 2017).

The last step within the building classification process is to identify the total number of buildings composing the building stock by the assumed classification system. As discussed above, a lot of information about the building stock is readily available at regional level, and such data sets can be further enriched by employing different approaches to collecting additional data. However, the resultant data sets might be insufficient for a direct classification of the building in terms of the construction type and occupancy class identified by an assumed taxonomy, therefore, inferences may need to be utilized. The assumed inferences should be verified on the basis of the "expert opinion" of local engineers and building officials in order to verify that they reflect the construction features of the region. A ground screening procedures can be used to develop or check the inference rules, used to characterize the building stock into categories, as well to assess the accuracy of the available information (Giovinazzi et al. 2014).

2.4 Building damage description

Earthquake-induced damage to building structures and nonstructural elements is usually described in a scale of multiple increasing damage limit states (simply called damage states, hereafter DSs), which range from no damage to total collapse. Damage scales exist for different building types, however, in this Section will only focus on damage scales relevant to RC buildings.

The FEMA multi-hazard risk assessment method HAZUS (NIBS 2003; FEMA 2015) adopts a damage scale which consists of five DSs, namely 'no damage', 'slight damage', 'moderate damage', 'extensive damage', and 'collapse'. This damage scale was developed explicitly for seismic loss estimation and describes extensively the expected observable damage for each building type including failure mechanism descriptions. **Table B 1** in Appendix B: presents the description of each DS for a RC moment resisting frame (MRF) building type. The drawback of this damage scale is that building types are based on the US standards and are not always relevant in a European context (Hill 2011).

EMS-98 (Grünthal 1998) also incorporates a damage scale which consists of six DSs, namely 'no damage', 'negligible to slight', 'moderate', 'substantial to heavy', 'very heavy' and 'destruction'. The EMS-98 damage scale relies on a qualitative description and some simple graphics, as shown in **Table B 2** in Appendix B: for the case of RC buildings. It is noted that this damage scale is developed for European standards and applications. Several studies built upon EMS-98 damage and proposed modified damage scales that are more relevant to physical behaviour or loss estimation (e.g., Blong 2003; Lang and Bachman 2004).

The Structural Engineers Association of California's (SEAOC) 'Vision 2000' performance-based design methodology provides another damage scale (SEAOC 1995). This is a US-centric scale, which is mainly descriptive but has some reference to physical parameters. Performance-based design methodologies provide a more practical human and engineering relevance than DSs that use abstract names such as 'slight', 'moderate' or 'heavy damage'. This is because they use performance objectives. For example a level at which the building's functions remain operational or where life is preserved (life safety). **Table B 3** in Appendix B: illustrates the damage states/performance levels definitions for 'Vision 2000' damage scale.

FEMA 356 (FEMA 2000) is a US code of practice for the seismic rehabilitation of buildings, which proposes a performance-based damage scale, similar to 'Vision 2000'. FEMA 356 improves on 'Vision 2000' by also including more information on building's top drift,

associated with each building type. The detailed structural performance levels according to FEMA 356 for concrete frames are presented in **Table B 4** in Appendix B:.

Rossetto and Elnashai (2003) proposed damage scales specific for RC buildings (referred to as homogenized reinforced concrete, (HRC)) and also provided conversion/correspondence tables mapping the HRC damage scales to performance objectives, namely, serviceability, damage control, and collapse prevention. The DS definitions are based on very detailed damage descriptions and a parameter of physical behaviour (inter-storey drift), which is calibrated against data from experimental large scale tests. This damage scale is related to existing damage scales and adopted in the empirical fragility curve generation with the aim of producing a single set of 'homogeneous' or 'general' fragility relationships that are applicable to all lateral-load resisting systems. Detailed DS descriptions of the HRC scale for various RC building types can be found in **Table B 5** in Appendix B.

Another damage scale derived for Europe is from the RISK-UE project (Milutinovic and Trendafiloski 2003). The damage scale is aligned to EMS/HAZUS levels and for a two-tier loss estimation methodology, leading to different data sets for each case study location in the project. In this study, two damage scales are derived. The first scale (LM1) uses a semi-empirical vulnerability approach that is analogous with a macroseismic intensity damage scale. The second scale (LM2) uses an analytical vulnerability approach that is analogous with FEMA's HAZUS damage scale. **Table B 6** in Appendix B presents the definition of DSs used in Risk-UE project.

The Crowley et al. (2004) damage scale includes a description and a parameter of physical performance namely strain limits for RC frame buildings (**Table B 7** in Appendix B:). The scale is another example of a damage scale developed for European vulnerability studies and loss estimation. The main deficiency is that whilst deformation is closely related to damage there is no information given on how many structural members would reach the limits or where they would occur in the building before the DS is reached. Also, it is not stated where the damage

focus should be, i.e. whether reaching the strain limit constitutes entering the damage level or whether other descriptive aspects of the damage level will constitute reaching the indicative strain rates (Hill 2011).

Other simpler damage scales exist in the literature, including the tagging system proposed by the Applied Technology Council Report (Rojahn et al. 1988), which is used for post-earthquake safety assessment. It noted that different definitions of damage tagging and usability tagging are used in the US and Italy; for more details one should refer to Comerio (2006) and De Luca et al. (2015) respectively. More detailed reviews of existing damage scales, including their potentialities and shortcomings for their use in seismic risk analysis and for loss assessment applications, can be found in Hill (2011) and Rossetto and Novelli (2016).

2.5   Selection of hazard parameters for fragility analysis

The selection of earthquake hazard parameters, commonly referred to as IM, that characterize the strong ground motion (e.g., amplitude, duration, frequency content, energy content) and correlate well with the building response is of high importance within the seismic risk framework. One of the key questions arising from fragility/vulnerability derivation process is how suitable an adopted IM is for representing ground motion uncertainty. This in turn will also affect the required number of analyses and the accuracy of the resulting seismic demand estimates. The development of fragility functions requires the choice of an IM, that is able to predict seismic demands involved in the considered performance objectives with the smallest scatter and providing a significant amount of information with respect to other ground motion features (Jalayer et al. 2012).

The most common parameter is the peak ground acceleration, PGA, which denotes the maximum ground acceleration (absolute value) observed on the accelerometric time series, usually evaluated for each component of the ground motion. It can be evaluated directly either from raw or corrected signals of ground motions, since PGAs change slightly with respect to the processing procedure. The peak ground velocity, PGV, and peak ground displacement, PGD,

are similarly defined. However, unlike PGA, PGV and PGD can be read only after processing the accelerometric record, and their values are strictly related to the band-pass filtering. Although the PGA is widely used in earthquake engineering, as it is directly related to the inertia force, it poorly characterizes the ground motion (Pacor and Luzi 2014). Records with the same PGA can have completely different features in the other ground motion parameters, depending on the magnitude of the event, the source-to-site distance and the geological condition of the site where the recording station is located.

For engineering purposes, the most important parameter to represent seismic motion is the response spectrum, in terms of relative displacement, relative velocity or absolute acceleration, i.e. $S_d(T_1)$, $S_v(T_1)$ and $S_a(T_1)$. The response spectrum is defined as the maximum response amplitude of a single degree of freedom (SDoF) system, subject to the acceleration time series, as a function of the system period T_1. The response spectrum can be calculated for different damping ratios, ξ (usually a value equal to 5 % of the critical damping is assumed, which is roughly applicable to most structures).

A number of IMs have been proposed in the recent years, and can be categorized into two main groups, i.e. acceleration-related and displacement-related groups. Acceleration-based IMs, which are developed on the basis of the spectral acceleration at the fundamental period of the structure ($S_a(T_1)$), can be further divided into three sub-groups, i.e. combination-type, vector-valued and integral-type. The effects of higher vibration modes and period elongation are considered in the combination-type IMs. The IMs proposed by Shome and Cornell (1999), i.e. S^*_{a12} and S^*_{a123}, are expressed as the weighted average of spectral accelerations at the first two or three periods respectively, where the weighted average coefficients are based on modal mass participation factors. In the two-parameter IM suggested by Cordova et al. (2001), S^c_a, the so-called period elongation effect is reflected by the softening period (equal to $2T_1$) and the power exponent (calibration coefficient) α. Vamvatsikos and Cornell (2005) and Lin et al. (2011), also proposed IMs that account for period elongation effects (employing similar combination forms

as Cordova et al. (2001)) , but using different power exponents and softening period definitions. Bojórquez and Iervolino (2011) proposed the advanced scalar IM, I_{N_p}, based on $S_a(T_1)$ and the spectral-shape proxy parameter N_p. In I_{N_p}, α is suggested to be 0.4, while softening period lies within the range of 2 and 2.5 times the value of T_1. Kazantzi and Vamvatsikos (2015) investigated the effectiveness of IM, $S_{agm}(T_i)$, which is based on the geometric mean of central values (mean or median) of the first-mode and second-mode periods of all index buildings selected to represent a building class. It is noted that the abovementioned IMs, which are based on the averaged logarithm of spectral accelerations over a range of periods, are in some cases referred to as *AvgSA* (e.g., Kohrangi et al. 2017a, b). Other acceleration-based IMs, which also constitute a variation of S_a^c introduced above, are the IM_{SR} and IM_{CR} proposed by Mehanny (2009). IM_{SR} and IM_{CR} consider a demand dependent structural parameter, i.e. the nominal relative lateral strength of the SDoF.

In the vector-valued sub-group, multiple ground motion parameters are used in tandem to form the so-called vector-valued IMs. The most common form of vector-valued IMs is when different indicators characterizing spectral shape, such as ε, $R_{T1,T2}$ and N_p, are combined together with $S_a(T_1)$. The indicator, ε, corresponds to the number of standard deviations by which an observed logarithmic spectral acceleration differs from the mean logarithmic spectral acceleration of a GMPE. The predictor $R_{T1,T2}$ is measure of spectral shape defined as the ratio of spectral acceleration at the second period ($S_a(T_2)$) over $S_a(T_1)$. Some of the vector-valued IMs frequently found in the literature include $<S_a(T_1)$, ε $>$, $<S_a(T_1)$, $R_{T1,T2}>$ proposed by Baker and Cornell (2005) to explore the effect of pulse-like near-fault ground motions, and $<S_a(T_1)$, $N_p>$ proposed by Bojórquez et al. (2012) to reflect different spectral shapes for ranges of periods. Numerical integration of spectral acceleration over the interval between the fundamental period and softening period is adopted in the integral-type IMs (e.g., Kadas et al. 2011; Biasio et al. 2014).

Most displacement-related IMs are combination-type IMs and based on spectral displacement at the fundamental period, $S_d(T_1)$. The IM $IM_{1E\&2E}$ proposed by Luco and Cornell (2007) is equivalent to the estimate of a peak storey drift angle that uses the first two modes and the square-root-of-sum-of-squares (SRSS) rule of modal combination. The $IM_{1E\&2E}$ can be simply converted to $IM_{1I\&2E}$, which is an IM that accounts for the contribution of both the first and second mode, and also the effects of nonlinearity (Luco and Cornell 2007). Yahyaabadi and Tehranizadeh (2011) recommended a flexible $S_d(T_1)$-based IM for the consideration of softening period or higher modes for various structures. This IM has similar expression to $IM_{1I\&2E}$, however, differences between the two lie in the number of terms combined and the coefficients of spectral displacements.

In contrast to the numerous research studies conducted on the spectral-acceleration-based and spectral-displacement-based IMs, research on spectral-velocity-based IMs have not gained much attention in recent years. However, the few IMs based on spectral velocity proposed decades ago still have an important reference value. The spectrum intensity (denoted as SI) proposed by Housner (1952) is defined as the area enclosed by spectral pseudo-velocity curve over the period range of 0.1 to 2.5s. More than three decades later, an IM with similar integral form and velocity spectrum intensity (VSI) was introduced by Von Thun et al. (1988). Van Cao and Ronagh (2014) have demonstrated that structural seismic damage has been found to have strong correlation with SI and VSI for low-rise RC buildings.

Another type of IMs are the duration-based IMs. These IMs attempt to identify the portion of record in which ground motion amplitude can potentially cause damage to a structure. Several definitions for "duration" exist in literature, including: (a) Bracketed duration (the time interval between the first and the last exceedance of a predefined acceleration threshold – typically 0.05g or 0.1g (Bolt 1973)), (b) Significant duration (the time interval over which the integral of the square of the ground acceleration is within a given range of its total value - usually this range is between 5 % and 95 % or between 5 % and 75 % (Trifunac and Todorovska 2001)).

The Arias intensity (I_A) can be derived from the significant duration. I_A is an integral parameter adopted to measure the severity of ground motion. Additional IM types have been proposed in many other studies. However, these IMs have not been used in seismic risk studies, and a comprehensive overview of this previous research is not within the scope of this work.

The selection of an optimal IM for conditioning seismic demand models is not a trivial matter and has been extensively investigated. Several studies have been conducted to determine the quantitative and qualitative characteristics of IMs that better describe aspects of ground motions that are responsible for the corresponding damage to particular elements or the system as a whole (e.g., Padgett et al. 2008; Ebrahimian et al. 2014b; Kohrangi et al. 2016a). To this aim, numerous selection criteria exist in the literature to assess the appropriateness of a candidate IM for fragility analysis. Some the most commonly used IM selection criteria include:

- efficiency
- sufficiency
- practicality
- proficiency
- scaling robustness
- effectiveness and
- hazard computability

Efficiency is the most commonly used quantitative criterion for the determination of optimal IMs, and is related to variation of demand estimations for different values of a studied IM (Padgett et al. 2008). Specifically, more efficient IMs result in reduced dispersion of the median EDP estimations conditional to a given IM, therefore, less analysis runs are required to narrow down the confidence interval. Sufficiency characterises the level of IMs' statistical independence conditional to specific earthquake characteristics, such as source to source distance (*D*) and magnitude (*M*). Practicality provides with evidence regarding the level of correlation between the demand estimation obtained from the analysis and a particular IM. For

instance, a practical IM indicates that there is a significant dependence of the levels of EDP responses to the associated IM levels. Proficiency is a two component IM criterion recently proposed by Padgett et al. (2008), linking the composite effect of efficiency and practicality. In other words, proficiency is a characteristic providing an indication of uncertainty introduced in the analysis by utilising a specific IM. Scaling Robustness means that the amplitude (linear) scaling of records does not induce bias in the estimation of the seismic demand. The effectiveness of an IM is determined by its ability to evaluate, in a closed form, the mean annual frequency of a DV exceeding a given limiting value. Hazard computability refers to the availability of estimates of the expected IM from hazard studies. Numerous hazard maps and Ground Motion Prediction Equations (GMPEs), also known as attenuation relationships, exist for more commonly used IMs, namely PGA and spectral ordinates at given periods (representing sometimes a restricted range of possible discrete periods), making these IMs more favourable from the hazard computability perspective.

There are several recent papers explicitly addressing the selection of an optimal seismic IM intended to be used in performance-based assessment of existing structures. For example, Mollaioli et al. (2013) assess the capability of twenty seven IMs to predict the seismic response of RC base-isolated buildings, using efficiency and sufficiency criteria. Bojórquez and Iervolino (2011) investigated the performance of spectral-shape based and vector-valued IMs in estimating the response of steel buildings. The efficiency of five different IMs (including scalar and vector-valued IMs) in predicting the response of a 3D RC building subjected to a bi-directional excitation, is examined by (Lucchini et al. 2011). Padgett et al. (2008) conducted a study on the selection of optimal IMs for highway bridges portfolios. To this aim, ten scalar IMs are tested using five selection criteria. Kazantzi and Vamvatsikos (2015) utilized a number of spectral acceleration-based IMs (averaged spectral accelerations, using a different number and range of periods) to carry out analytical vulnerability assessment. Efficiency and sufficiency were used for assessing the seismic losses of an RC building class. Kohrangi et al. (2016a, b, c) extensively described the use of scalar and vector-valued IMs within PBEE framework for

assessing building response of RC frames for both loss and collapse estimation. Ebrahimian et al. (2015) performed a comparative study to rank thirty eight scalar and vector-valued IMs, evaluated with the regard to efficiency and relative sufficiency criteria. This study showed that most informative IMs (i.e. vector-valued and advanced IMs) are more suitable for predicting the seismic response and eventually the fragility of a structure; nevertheless, $S_a(T_1)$ is the most favourable conventional IM as it is easy to derive and performs well under most IM selection criteria.

It is worth noting that all of the above studies were based on nonlinear dynamic approaches (see Section 2.7.1) and Incremental Dynamic Analysis (IDA) (see Section 2.7.3.2) for carrying out the seismic performance assessment of structures. However, none of those past studies has focused on simplified demand assessment (and fragility analysis) based on nonlinear static (pushover-based) procedures. In addition, only in recent years, advanced IMs (i.e. spectral-shape based) and innovative selection criteria (e.g., the relative sufficiency measure based on the information theory (Jalayer et al. 2012)) have been introduced. Although some tests exists in separate studies (and for different structural types), no one has attempted to convolve all these state-of-the-art concepts and identify optimal IMs to be used in probabilistic seismic demand analysis through nonlinear static methods (these methods are described in detail in Section 2.7.2).

2.6   Estimating ground motions intensities

There are two primary approaches to estimate ground motion intensities for seismic risk analysis, namely, the probabilistic seismic hazard assessment (PSHA) and the deterministic seismic hazard assessment (DSHA). The DSHA allows for more realistic representation of individual event, i.e. use of scenario earthquakes, by considering each seismogenetic source separately and assuming the occurrence of an earthquake of specified size at a specified location. However, a single scenario earthquake does not account for the uncertainties in seismic hazard. To overcome this limitation, a large scale simulation technique (such as Monte

Carlo simulation or Latin hypercube sampling) can be used to represent the distribution of uncertainties by undertaking analysis for multiple scenario earthquakes (Giovinazzi et al. 2014). For each scenario, the attributes are determined probabilistically from the historic earthquake catalogue for the area of interest.

On the other hand, PSHA combines the contributions from all the relevant seismogenetic sources related to the area of study and allows one to determine the rate at which particular levels of ground motion intensities occur. In PSHA the uncertainties associated with the determination of time, location, depth, and magnitude of future earthquakes can be addressed by probabilistically (Bommer 2002). Through PSHA the ground motion intensity values associated with particular exceedances probabilities can be evaluated at any site considering all possible seismic resources of the region. When the procedure is repeated for multiple sites across a studied region probabilistic seismic hazard maps of ground motion intensity can be produced. The maps of IM values constructed from a PSHA represent the possible ground motion fields, i.e. the possible levels of ground shaking (logarithmic mean and residual) with the associated probability of being exceeded within a return period, giving the rate at which particular levels of ground motion might occur. For example, **Figure 2.2** presents the seismic hazard map for Italy, which shows PGA estimates characterized by a 10% exceedance probability during the service life of a common residential building (i.e. 50 years), corresponding to a return period of 475 years. Alternatively, results from a PSHA can be represented in terms of hazard curves, providing the probabilities of exceedance within a given time period for a continuous range of ground shaking intensities at each location. Ground motion fields obtained via DSHA represent at each site the level of ground shaking (logarithmic mean and residual) resulting from a single, deterministic earthquake (e.g., the maximum historical event from a specific seismic source or the maximum earthquake compatible with the known tectonic framework) (Giovinazzi et al. 2014).

According to Bommer (2002), there is not a mutually exclusive dichotomy between DSHA and PSHA and different approaches combining DSHA and PSHA have been proposed in the

literature. One approach is to select the event that is most likely to contribute to the hazard for a given IM measure at a site, from the disaggregation of PSHA outputs (Bazzurro and Cornell 1999). Bazzurro and Luco (2005) proposed the use of Monte Carlo simulations to reproduce a series of single earthquake scenarios within PSHA.

Regardless of the choice of ground motion intensity estimation approach, the first step is to identify the seismogenetic sources affecting the region of interest and classify them in terms of location, geometry, seismic activity, and possibly released energy. The second step is to represent the propagation of the ground motion by the use of a suitable GMPE. GMPEs, also referred to as "attenuation" relationships, are empirical equations derived from observed strong motion records and provide a means for predicting the level of ground shaking in terms of the desired IM, and its associated uncertainty at any given site or location, from knowledge on the earthquake magnitude, M, source-to-site distance, R, local soil conditions, fault mechanism, etc. Due to the empirical nature of GMPEs, the ground motion at each site is estimated by the median, or logarithmic mean, of the ground motion plus two random variables, referred to as residuals (Giovinazzi et al. 2014). These residuals represent the aleatory variability, namely, the difference between the predicted and observed ground motions in the regression analysis, occurring intra-event (i.e. within a single event), and inter-event (i.e. between separate events) (Douglas 2018). For this reason, when GMPEs are applied to a large area, adjacent predicted ground motions do not correlate with each other and the resulting ground motion field is random. This does not have an effect in the case of spatially clustered analysis (i.e. analysis at a single site). However, when the analysis is targeting spatially distributed building stocks and the simultaneous generation of ground motions at multiple sites is required, the ground motion field should be spatially correlated to more realistically represent the earthquake scenario. Such observation was confirmed by several past studies which showed that the spatial correlation of ground motion field can significantly influence the outputs of a seismic loss analysis for distributed portfolios, when simulating a single earthquake scenario. A number of spatial correlation models are available in the literature. A comprehensive discussion and description of

different methods to produce spatially correlated ground motion fields and to account for site amplifications can be found in Weatherill et al. (2011).
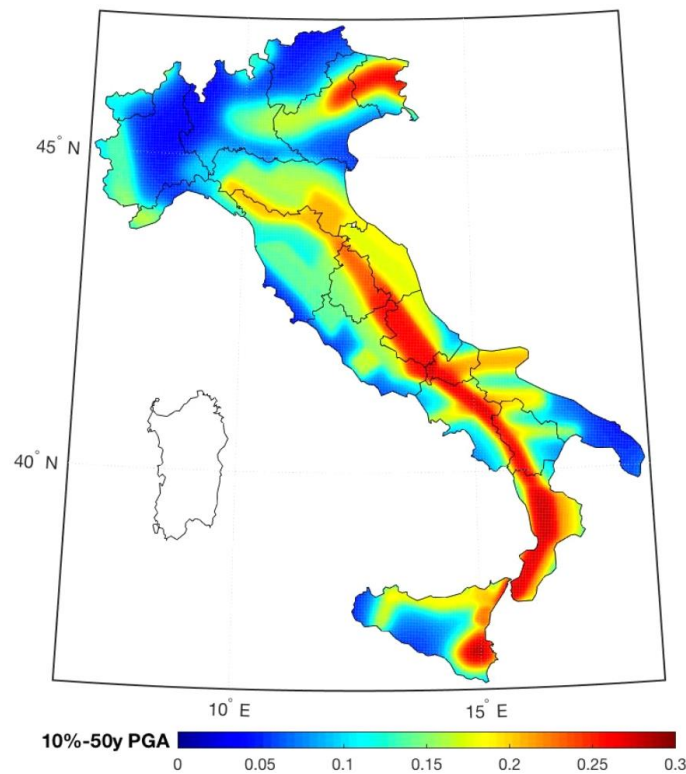
Figure 2.2 - Seismic hazard map for Italy, in terms of 10%-50y PGA (http://esse1-gis.mi.ingv.it).

2.7   Assigning damage to buildings - Review of analytical methods for fragility assessment

The level of damage caused to a building or building stock following a seismic event can be estimated through the use of fragility functions. Fragility functions can be derived using (a) empirical methods (actual damage observation), (b) expert judgement, (c) analytical methods, or a combination of the three abovementioned approaches, also known as hybrid methods. This Thesis mainly focuses on analytical methods for seismic risk assessment, therefore, for the sake brevity, the other three approaches are not discussed herein.

In this subsection the existing analytical methods for fragility assessment are reviewed (e.g., nonlinear dynamic and nonlinear static analysis methods). The PSRA framework is then addressed, with a particular focus on the cloud analysis approach. Next, the derivation of

fragility functions is discussed, including the alternative methods for PSRA and the statistical procedures to fit fragility. Finally, the implementation of surrogate models and kriging metamodels within analytical fragility process is investigated.

2.7.1 Building response calculation - Nonlinear dynamic analysis

Nonlinear dynamic analysis (NDA) is attributed to be the most reliable and accurate method to determine the seismic performance of individual structural models. NDA implements very detailed structural models, which account for the hysteretic behaviour of the structure, the plasticity model (lumped- or distributed-plasticity models), the equivalent damping, the integration algorithms etc. (Silva et al. 2014a). A series of natural or simulated ground motions of varying intensities is utilised as an input for the so called nonlinear time-history analysis. In particular, these records subject the structure of interest to a virtual shaking, which results in a better capturing of higher mode effects in conjunction with the identification of various failure mechanisms. All structural elements of the model are analysed in order to compute the forces and deformations developed and eventually used to evaluate the global response measures (e.g., maximum peak inter-storey drifts, storey accelerations and forces, roof displacement etc.), also known as EDPs.

The use of ground motion time-histories from multiple earthquakes allows the model to obtain not only an estimate of the mean response given a certain level of ground shaking intensity, but also allows it to account for the variability in the buildings' nonlinear response generated by different records of the same IM (e.g., Jalayer & Cornell, 2009; Vamvatsikos & Cornell, 2002).

The analysis of a structure model under multiple ground motions of increasing intensity is most commonly conducted with three approaches: Cloud analysis, IDA and Multiple Stripes Analysis (MSA). The first method (i.e. cloud analysis) attempts to represent the variation in response through the use of a wide selection of many ground motions, grouped into bins. The latter two methods achieve the same by stepwise increment of a few ground motion records. The above methods are discussed in more detail in Section 2.7.3.

The selection of the ground motion records that are used as input for the dynamic analyses is of paramount importance for the reliable evaluation of the seismic response. The quantity and the distribution of IMs in the sample of records have indeed a great influence on the fragility parameters (both the standard deviation and the median) (Baker et al. 2014). As fragility curves are commonly derived for structures in a particular geographical area, earthquakes time-histories can be selected based on specified intervals of magnitude, source-to-site distance and possibly other scenario characteristics, such as focal depth and mechanism that characterise the seismic activity of the area (e.g., Bommer and Acevedo 2004). For this purpose, special software for the selection of strong ground motion recordings from European and international databases can be used, such as REXEL (Iervolino et al. 2009) (see Section 2.7.1.1). In the record selection and analysis processes it is important to consider records with possible special features, such as near-source directivity pulses. Such records must be appropriately accounted for, since the structural response they induce can be significantly different than that for records from distant sources. When soil foundation- structure interaction (SSI) is taken into consideration, modelling both soil and structure in a coupled system, the input motion is normally introduced at the seismic bedrock and therefore it should refer to rock conditions, as the SSI model directly captures site effects (Karapetrou 2015).

Detailed guidance on the implementation of NDA methods in fragility analysis can be found in Deierlein et al. (2010) and D'Ayala et al. (2013a) amongst others. A flow chart describing the NDA procedure is illustrated in **Figure 2.3**.
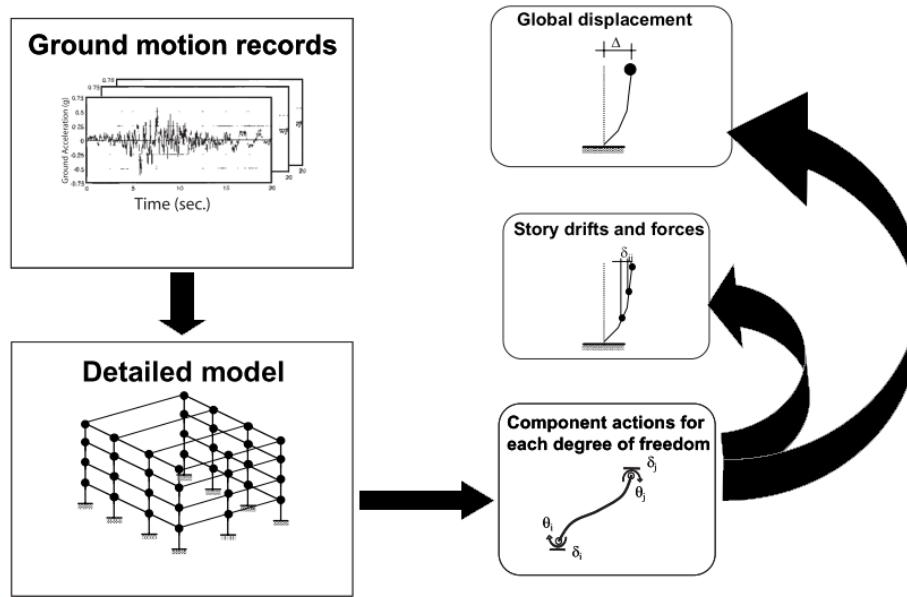
Figure 2.3 - Flow chart depicting the NDA process to obtain the structure's response (image retrieved from: FEMA, 2005).

2.7.1.1 Ground motion selection for nonlinear time-history analysis

Past research has shown that amongst all possible sources of uncertainty, including structural and soil properties, modelling and analysis assumptions as well as the choice of earthquake-induced ground motion, the latter is one of the main contributors in the resultant variability in the structural response (Shome et al. 1998; Padgett and DesRoches 2007). Therefore, the selection of a "reliable" suite of earthquake ground motions to represent the variability in response of a structure at one IM value due to aleatory variation in ground motion constitutes an important step within the structural analysis procedure as a whole.

Along these lines, numerous computational methods and tools have been developed for (a) selecting suites of earthquake records from available strong ground motion record databases (b) generating synthetic and artificial ground motions, or (c) modifying (scaling) existing ground motions until they present desirable target characteristics (Sextos 2014). Some of the most commonly used methodologies for ground motion selection and scaling include the following:

- Selection by M and R and scaling to a target IM: Ground motion sets are selected to satisfy a given range of M and R for specific seismotectonic or soil class criteria. Once eligible ground motions are selected, their accelerations are multiplied by a scaling

66

factor to match a target IM, typically being PGA or $S_a(T_1)$. As discussed in Section 2.5, IMs have also been proposed, involving the spectral shape and the structural characteristics, which are expected to result in a more accurate prediction of the seismic demand (Baker 2005; Luco and Cornell 2007). This ground motion selection and scaling approach has been used for the derivation of fragility functions in several studies, including Lagomarsino and Giovinazzi (2006), Polese et al. (2013) and others.

- Selection and scaling to a target Uniform Hazard Spectrum (UHS): Ground motion sets are selected so that their response spectra match with a target, UHS (Kramer 1996; ASCE 2005). The UHS is determined either from a ground motion prediction relationship, a seismic hazard assessment for the site of interest, or from the seismic code provisions. In this case, ground motions are scaled with a single scaling factors (as in the case of Eurocode 8) or individual scaling factors (e.g., FEMA 2009) so that their individual spectrum, or the mean of their response spectra closely match and in fact exceed, in terms of "shape," the ordinates of the target UHS. The matching is commonly carried out over a given range of periods around the fundamental period of the structure (Sextos 2014). An example study where UHS was used for fragility function derivation is Cimellaro et al. (2011).

- Selection and scaling to a target Conditional Mean Spectrum (CMS): A CMS (Baker 2011) represents the expected mean response spectrum, conditioned on the occurrence of a target spectral acceleration at the period of interest (typically the fundamental period of the structure T_1). The target CMS is calculated for the scenario $S_a(T_1)$, M, and R, together with other criteria such as fault type or soil, and represents the expected response spectrum for the defined ground motion scenario, which is based on a target $S_a(T_1)$ at a single structural period. To develop this new target spectrum, PSHA can be used to determine the spectral acceleration $S_a(T_1)$ that corresponds to the target probability of exceedance at the site of interest, denoted as $S_a(T_1)^*$. A detailed

explanation of the CMS matching process can be found in Sextos (2014). This approach was used for fragility function derivation in Cardone et al. (2018) amongst others.

The use of site specific UHS is recognized to be a conservative approach for both fragility function derivation and for the estimation of aggregate measures of risk/loss assessment as it implies that large-amplitude spectral values will occur at all periods within a single ground motion time history (Baker 2011), but it has nevertheless been traditionally accepted due to the lack of a better tool. UHS corresponds to an envelope of spectral ordinates with equal probabilities of exceedance at different periods. As a result, they do not correspond to any particular earthquake scenario. On the other hand, the CMS overcomes this shortcoming by considering a single or multiple earthquakes obtained from seismic hazard disaggregation (Cardone et al. 2018). Previous studies (Liel et al. 2011; Haselton et al. 2011) have shown that using records compatible with UHS tends to underestimate the collapse capacity of a structure compared with the same structure subjected to a set of records compatible with CMS.

Ground motion records can be accessed from several ground motion databases in Japan, Taiwan, and Europe (European Strong Ground Motion database, www.isesd.hi.is), the PEER-NGA Next Generation Attenuation strong motion database amongst others. Given the above extensive repository of earthquake records and the fact that the most common earthquake record selection procedures involve spectral matching, recent work evolved to develop computational tools for quantifying and/or optimizing spectrum compatibility to prescribed codes. REXEL (Iervolino et al. 2009) was the first, all-in-one, software introduced for this purpose. It facilitates the search for suites of waveforms compatible to target spectra that are either user defined or automatically generated according to Eurocode 8 and the Italian seismic code. More recently, the Integrated System for Structural Analysis and Record Selection (ISSARS) software has been developed (Katsanos and Sextos 2013). This software retrieves ground motions from the PEER-NGA database, to form suites of records that not only comply with specific criteria but also ensure, through numerical analyses of the structure that run at the background, a target level of dispersion of structural response quantities.

2.7.2 Building response calculation - Nonlinear static procedures

The high computation effort and the convergence issues associated with NDA, renders this method largely impractical for portfolio of structures and highlights the need of using more rapid and simple approaches, such as nonlinear static procedures (NSPs). However, the applicability of NSPs for the estimation of the precise response of structures (e.g., in seismic assessment) requires further investigation. This is mainly because NSPs rely on simplifying assumptions. For instance, the response of a complex multi degree of freedom (MDoF) system is assumed to be represented by an equivalent SDoF system, while in most NSPs standardized design spectra (e.g., code-based) are used to characterise the seismic demand. As a result, the response of structures dominated by higher mode effects (i.e. high-rise buildings) is not sufficiently captured by such procedures. Other limitations associated with most NSPs are that they fail to identify complex failure mechanisms and they do not account for record-to-record variability (Fragiadakis et al. 2014). To summarize, Fragiadakis et al. (2014) constructed a list of criteria for the applicability of NSP methods, which is presented in tabulated form in **Table 2-2**.

Table 2-2 - Criteria for applicability of NSPs (Fragiadakis et al. 2014).

Criterion	Description
Selection of EDP:	Displacement-based EDPs are recommended (i.e. roof displacement, roof- and inter-storey drift). NSPs perform well in association with displacement, but yield unreliable estimations for storey shears and overturning moments (FEMA 440, 2005)
Levels of inelasticity:	NSP is recommended for light and moderate levels of inelasticity

Period range:	The threshold of period is taken as $2T_C$ [1]
Structural system:	NSPs are less efficient when combined with shear wall structural systems
Post-elastic strength:	NSPs may lack accuracy when significant P-Δ effects or post-elastic strength softening occur
Inelastic mechanism:	When the inelastic mechanism is dominated by higher modes, NSP may fail to capture the right damage mechanism

The step-by-step procedure to determine a structure's seismic response using NSP is illustrated in **Figure 2.4**.
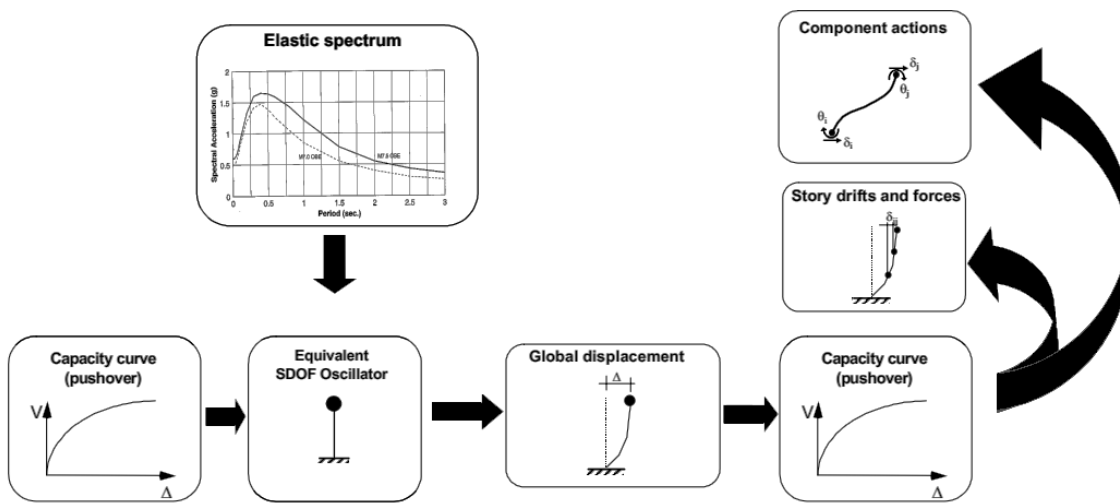
Figure 2.4 - Step-by-step process of NSPs to determine a structure's seismic response (image retrieved from: FEMA, 2005).

The first step of NSP is to perform a simple nonlinear static analysis, known as pushover analysis (PO) to estimate the response of the structure. This response is expressed in terms of a EDP (most commonly displacement based EDP), for a given seismic input in terms of IM. Increments of static lateral loads are imposed on the structure at locations of mass concentration

---

[1] Where T_C represents the corner period of standardized code-based design spectrum.

(typically at each floor level) over a predefined loading phase (target displacement control). The applied lateral load pattern is representative of that which would be experienced by the structure if it deflected under its first mode of vibration. The outcome of this procedure is a force-deformation relationship, expressed at a global level in terms of base shear and roof displacement, and is known as the pushover curve (or simply PO curve). The PO curve characterises the lateral resistance of the structure and varies according to the choice of lateral load pattern applied in the PO analysis.

In conventional PO procedures the lateral load can be distributed uniformly along the buildings height (UNI-PO) or following an inverse triangular pattern (TRI-PO), corresponding to floor masses and storey heights respectively. Alternative more advanced PO procedures also exist, such as non-adaptive multi-modal PO approach (i.e. Chopra & Goel, 2002), adaptive displacement based PO approach (i.e. Antoniou & Pinho, 2004) amongst others.

Detailed comparison of different PO methods, including single and multi-modal PO methods, exist for various structural systems (e.g., Onem, 2008; Diotallevi et al., 2008; Kalkan & Kunnath, 2007 to name a few). These studies tend to agree that first-mode PO methods provide good predictions of peak floor displacements for low- and mid-rise buildings, while they tend to overestimate the responses of high-rise buildings. On the other hand, multi-mode methods are seen to perform better in most cases. However, none of the tested PO approaches was consistently reliable for all types of buildings and EDPs used (Fragiadakis et al. 2014). A review of some of the most commonly used NSPs can be found in NIST (2011) and Causevic & Mitrovic (2011).

Next, an "equivalent SDoF" is defined from the PO, which maintains the same dynamic properties as the initial model, namely natural period, damping and stiffness. By so doing, the applied shear force and the associated displacement can be converted to spectral acceleration and spectral displacement respectively, to generate the so-called capacity curve. The seismic demand is represented by elastic or inelastic spectra derived from ground motion records or

design spectra as suggested by building codes, which can be plotted in the same format as the capacity curve. In some approaches (e.g,. N2 method, originally proposed by Fajfar 2000) instead of using real records and spectra, design spectra, as defined in codes of practice, are used. The performance of the building of interest is obtained at the unique intersection of the demand and the capacity curves, known as the performance point (PP). The PP is essentially the global top displacement (spectral displacement) of the SDoF under the considered seismic loading. By checking when this global top displacement occurs in the structure's PO, it is possible to evaluate the selected EDPs for a given intensity level (EDP|IM).

2.7.2.1 Summary of existing NSP methods

In the past few years several different NSP methods have been proposed aiming to realistically estimate the seismic response of structures. Some of these methods earn additional credibility as they are recommended by various design codes around the globe. Most notably, the N2 method (Fajfar 2000) is suggested by the Eurocode 8 (EN 1998-1 2004), the coefficient method which is included in ASCE/SEI 41-06 (ASCE, 2006) and in FEMA-356 (FEMA, 2000a), the capacity spectrum method (CSM) from FEMA-440 (FEMA, 2005) and ATC-40 (ATC, 1996), are the nonlinear static approaches one can find in European and American building codes.

These approaches usually require a standardized design spectrum (e.g., code-based) and the use of a corner period (usually named T_C) to identify acceleration- and velocity-sensitive segments of the demand spectrum. Therefore, these standardized earthquake spectra are commonly defined as smooth functions that do not account for the variability present in natural spectra. Some recent studies (e.g., Faella et al. 2008; De Luca et al. 2013b) show that natural (unsmoothed) response spectra can be used in the N2 method thereby accounting for record-to-record variability in the performance-point evaluation.

Another simplified approach that also accounts for the record-to-record variability in the estimation of seismic response of a structure is the FRACAS method, which is recommended by GEM guidelines (D'Ayala et al. 2013), and is extensively described and tested in Chapter 4.

FRACAS uses acceleration time histories to compute both elastic and inelastic spectra which are then used to find the PP. It is acknowledged that response spectra do not capture the entire variability in earthquake ground motions, e.g., response spectra fail to account for the duration of shaking. However, several studies have shown that the amplitudes and shape of the elastic response spectra have a key influence on the inelastic structural response, particularly at high nonlinearity levels (e.g., Gehl et al. 2014) and collapse occurrence (Rossetto et al. 2014a). Moreover, response spectra of earthquake ground motions do show considerable variability, and these differences will be reflected in the fragility curves derived using this CSM.

2.7.3   Probabilistic seismic response analysis (PSRA)

PSRA is used to estimate the seismic response of structures when subjected to ground motion of varying intensity taking into consideration the different sources of uncertainty. PSRA constitutes an essential step to extract information required (i.e. EDP|IM statistics) for the determination of fragility, which is discussed in Section 2.7.4. Different methods exist to estimate the probabilistic seismic response as a function of seismic hazard with aim of NDA tools. Some of these methods include:

i.   Regression analysis on analysis response data based on unscaled or uniformly scaled IM inputs (the so-called cloud analysis).
ii. Fitting a probability distribution on analysis data obtained from records scaled to certain target levels of intensity.

As discussed in detail in Section 2.5, the seismic hazard is most commonly expressed in terms of single scalar IMs; however, different approaches exist that implement vector-valued IMs. In general, the implementation of vector-valued IMs is computationally more demanding without necessarily resulting in significant improvements in estimating the seismic response. As a result, only scalar IMs are presented for PSRA in this study.

The most commonly used seismic response approaches are the cloud analysis method, the IDA and MSA. Each method employs a different way to select ground motions to represent the seismic hazard (e.g., scaled or unscaled records).

In the following subsections the cloud analysis method is presented, alongside with alternative seismic response approaches to determine the seismic fragility, such as IDA and MSA. Next, the various statistical procedures used for the fragility fitting are discussed.

2.7.3.1 Cloud analysis

The cloud analysis (or cloud method – the two terms are used in this Thesis interchangeably) is one of the most commonly used approaches for characterizing the relationship between EDP and IM (Bazzurro et al. 1998; Luco and Cornell 1998; Jalayer 2003). Within this method, computer simulations are used to determine the seismic response of a structure to a series of ground motion time-histories, each of which has an IM value. The simulations are usually based on NDA , with the building response expressed in terms of EDP. It is noted that the incapability of some NSPs to capture record-to-record variability precludes their use within the cloud analysis framework. However, the recently proposed variant of the capacity spectrum method FRACAS, can be implemented in cloud analysis as it utilizes spectra from ground motion inputs to perform the structural assessment.

The resultant peak values of EDP for given IM levels form a scatter of points, the so-called "cloud". Least-squares regression is then used to fit a simple model to the cloud of data points. Typically, the EDP is considered to vary as a power-law of the form aIM^b such that, after taking logarithms, the relationship can be expressed as in Eq. (2.1):

$$\ln(EDP) = \ln(a) + b\ln(IM) + e \qquad (2.1)$$

where EDP is the conditional median of the demand given the IM , a , b are the parameters of the regression, and e is a zero mean random variable representing the variability of $\ln(EDP)$ given the IM . An example of the simple linear regression model described above is shown in **Figure 2.5**. However, some situations exhibit substantial heteroscedasticity (i.e. non-constant variance), which needs to be modelled explicitly within the fragility analysis framework (Modica and Stafford 2014); for example by performing linear regressions locally in a region of

IM values of interest. The use of logarithmic transformation assumes that the EDPs are lognormally distributed conditional upon the values of the IMs. This is a common assumption that has been confirmed as reasonable by many past studies (e.g., Cornell et al. 2002; Ibarra and Krawinkler 2005; Porter et al. 2007).
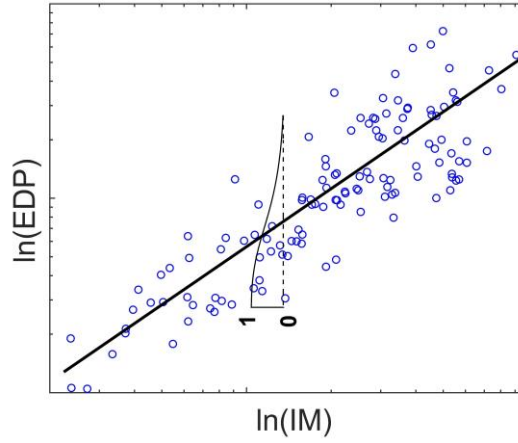
Figure 2.5 - Cloud of EDP|IM data points fitted by a linear regression model, and a conditional CDF of the EDP|IM.

In some special cases the seismic demand cannot be represented properly by a single linear logarithmic regression model (Ramamoorthy et al. 2006). Therefore, bilinear regression models are utilised, as shown in **Figure 2.6**.
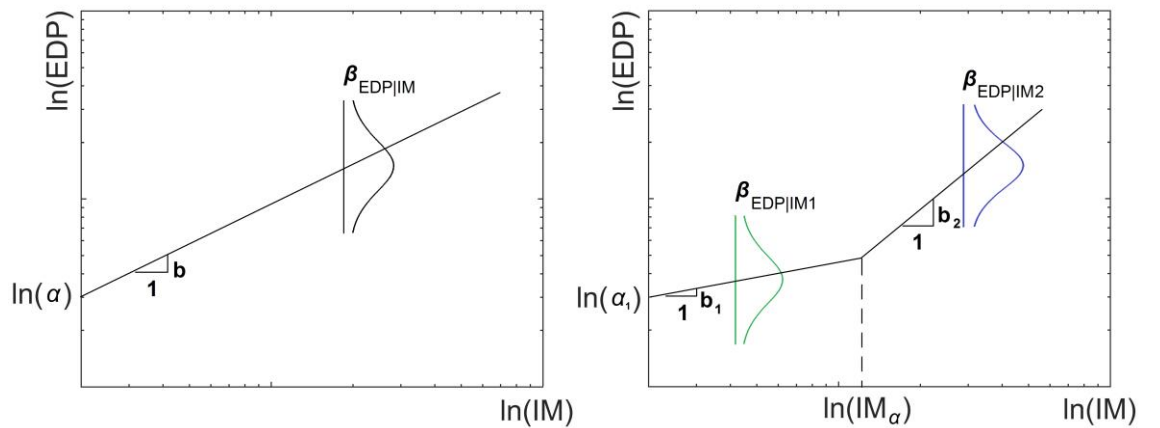
Figure 2.6 - Linear and bilinear regression fitting of the EDP and IM pairs (adapted from: Jeon et al. 2015).

In the case of bilinear regression, the Eq. (2.1) will be modified as shown below:

$$\ln(EDP) = \begin{cases} \ln(a) + b_1 \ln(IM) & , \quad IM \leq IM_0 \\ \ln(a_1) + b_1 \ln(IM_0) + b_2 [\ln(IM) - \ln(IM_0)], & IM > IM_0 \end{cases} \qquad (2.2)$$

where a_1, b_1 and b_2 are the parameters of the bilinear regression model and IM_0 the threshold of the demand regime change. IM_0 can be computed as the minimum of the residual sum of squares. The dispersion of each demand regime is calculated separately.

The simplicity and rapidity of cloud analysis make this method attractive for fragility assessment, and is recommended for safety-checking in guidelines, such as FEMA-356, SAC (FEMA 2000; Cornell et al. 2002).

In its original form cloud analysis was mainly used to capture the uncertainty associated with record to record variability. However, several researchers have attempted to incorporate the structural modelling uncertainties within the cloud analysis framework. Specifically, several studies are proposed to account for uncertainty in the capacity of components as well as in material mechanical properties and structural detailing (Jalayer et al. 2007, 2011). Some recent studies have also proposed to combine cloud analysis with Latin Hypercube Sampling (LHS) method (Lu et al. 2008) and with Monte Carlo within a Bayesian framework (Jalayer et al. 2014), in order to consider the effect of modelling uncertainty. The incorporation of capacity and modelling uncertainty (i.e. epistemic uncertainty) is translated as an increase in the dispersion in the exceedance probability of a damage limit state.

The cloud approach has the advantages of simplicity and rapidity over alternative fragility assessment methods. However, it also has some restrictions. Firstly, an assumption is made that the relationship between IM and EDP is represented by a linear model in the log space. This assumption may be valid for a limited range of IM and EDP combinations but not for the entire cloud response space. Additionally, the cloud method requires a large number of earthquake records to be used as an input, and the accuracy of the approach is highly dependent on the record selection process followed (Jalayer et al. 2014).

2.7.3.2 Incremental Dynamic Analysis (IDA)

IDA was originally proposed by Vamvatsikos and Cornell (2002), to estimate different limit-state capacities of a structure subjected to a series of GMs scaled to multiple levels of IM (**Figure 2.7a**). The outcomes of IDA are implemented to determine the probability of a limit-state to be reached or exceeded for a predefined IM threshold, im. A lognormal distribution function is utilised to fit the empirical cumulative distribution of IMs and eventually describe the probability of collapse (**Figure 2.7b**).
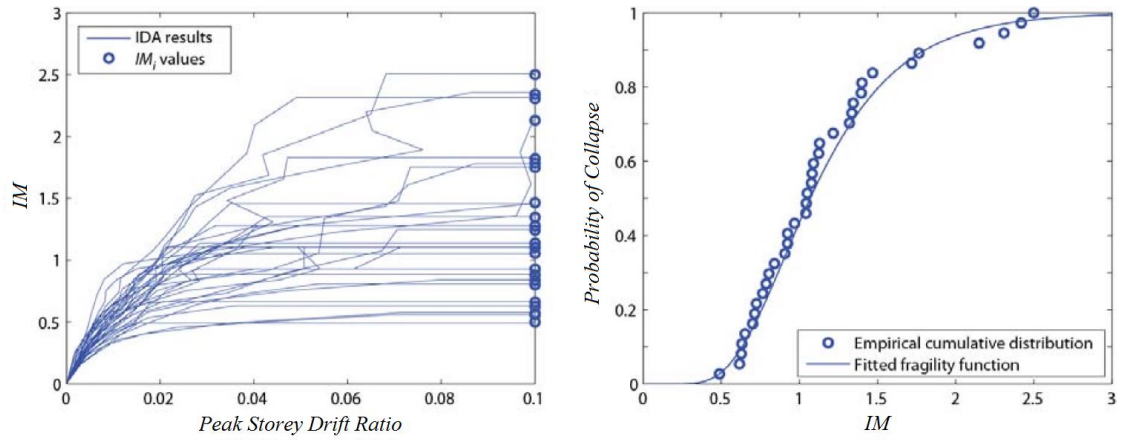
Figure 2.7 - Incremental dynamic analysis results example (Left panel); Collapse fragility curve using IDA and observed fractions of collapse as a function of IM (Right panel) (excerpted from: Baker, 2015).

In order to plot the associated fragility function, the mean and standard deviation (μ and β respectively) need to be determined as per Equations (2.3) and (2.4):

$$\ln \hat{\mu} = \frac{1}{N_{GM}} \sum_{i=1}^{N_{GM}} \ln IM_i \tag{2.3}$$

$$\beta = \sqrt{\frac{1}{N_{GM}-1} \sum_{i=1}^{N_{GM}} \left(\ln(IM_i / \hat{\mu}) \right)^2} \tag{2.4}$$

where N_{GM} is the number of considered ground motion records and IM_i is an IM value of a record i that causes a certain limit-state to be reached or exceeded. This is a method of moments estimator, as $\ln \mu$ and β are the mean and standard deviation (or moments) of the normal

distribution representing the $\ln IM$ values (Baker 2015). Note that the mean of $\ln IM$ is equal to the median of IM in the case that IM is lognormally distributed, which is why using the sample mean in this manner produces an estimate of μ. The method of moments estimator is a widely used fitting procedure for the derivation of analytical fragility curves and is performed to obtain the parameters of the lognormal distribution of the input sample, i.e. mean and standard deviation (**Figure 2.7b**). This approach is denoted as "Method A" in Porter et al. (2007). A related alternative is to use counted fractiles of the IM_i values, rather than their moments, to estimate μ and β (Vamvatsikos and Cornell 2004).

The main advantage of IDA is that it produces estimates of higher accuracy comparing to cloud method, but is significantly more computationally demanding. This limitation effectively precludes this approach when analysing large populations or portfolios of buildings, for example for catastrophe modelling purposes. Another shortcoming of IDA is that in some cases, earthquake records may need to be scaled up heavily to push the structure to collapse. The main concern about this is how heavily scaled GMs can represent the effect of a natural extreme earthquake (Baker and Cornell 2005). To overcome the aforementioned problem, truncated IDA can be utilised. In this case, IDA is carried out up to a defined IM limit and with the aim of maximum likelihood method the best fitted fragility function is obtained from the likelihood of observing the data that were observed (Baker 2015).

2.7.3.3 Multiple Stripes Analysis (MSA)

The MSA is described extensively in several studies (e.g., Bazzurro et al. 1998; Jalayer and Cornell 2009). Like IDA, MSA aims to characterise the relationship of EDP and IM over a broad range of IM levels. An individual "stripe" of the analysis corresponds to NDAs using ground motion records scaled to reach a predefined value of IM, based on the rationale that each suite of records represents the earthquake hazard for the specific IM value. In the past, the same suite of ground motions was utilised to carry out the structural analyses along each "stripe" and form a single MSA (Luco and Cornell 1998; Vamvatsikos and Cornell 2006 amongst others).

On the contrary, recent studies suggest that different suites of ground motions should be employed to carry out the structural analyses of different "stripes" (Bradley 2010; Lin et al. 2013). This is because the ground motion target properties change at each target IM level and thus so do the representative ground motions. With this approach, the analysis need not be performed up to IM amplitudes where all ground motions cause collapse (Baker 2015). It is noteworthy to mention that in some cases, increasing IM values may not necessarily result in strictly increasing fractions of collapse, although it is expected that the true probability of collapse is increasing with IM. This is due to the fact that different ground motions used at each IM level. In MSA collapse limit state is identified at locations where a ground motions of same IM level result in a radical increase of EDP, as can be observed in **Figure 2.8a**.

The fitting method to derive the fragility curve in MSA is maximum likelihood estimation, which is described in the following Sections.
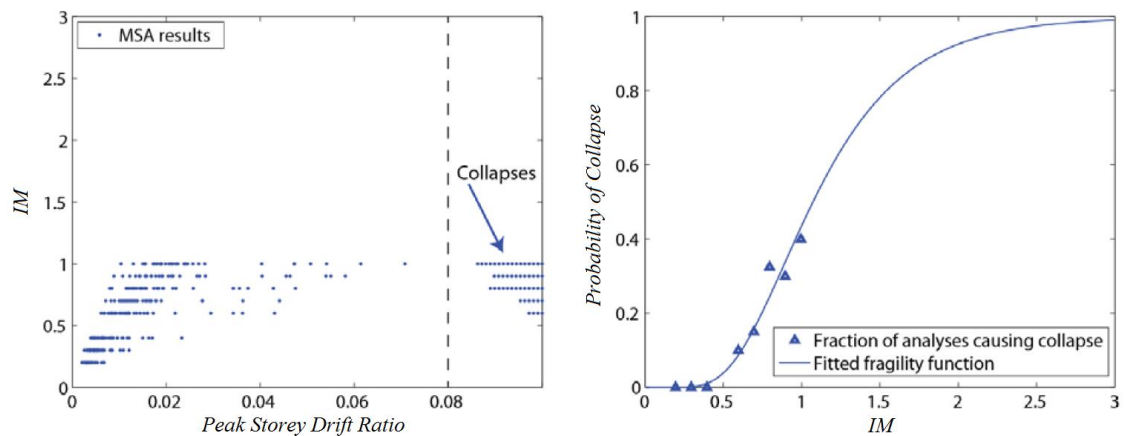
Figure 2.8 - Multiple stripes analysis results example. Collapse EDP threshold at peak storey drift ratio equal to 0.08 (Left panel); Collapse fragility curve using MSA and observed fractions of collapse as a function of IM (Right panel) (excerpted from: Baker, 2015).

2.7.4   Determination of fragility

Following the estimation of seismic response of structures within the PSRA procedure, the next step is to derive the seismic fragility functions (also referred to as fragility curves). Fragility functions constitute an essential part of the seismic risk assessment framework, as they provide a probabilistic tool to characterise the vulnerability of a structural system (Calvi et al. 2006).

For the derivation of a fragility function, one needs to select carefully the following parameters:

a. An IM that represents well the ground motion characteristics that influence seismic structural response.

b. A damage scale consisting of DSs that can capture the potential mechanisms causing failure to the structure(s) under consideration.

c. A numerical/statistical model that describes the probability of reaching or exceeding a DS as a function of the selected IM.

As discussed extensively in Section 2.5, the seismic hazard is commonly expressed in terms of IM; however, the choice of IM is of high importance. The adopted IM should be able to predict the response of the system with the smallest scatter and providing a significant amount of information to predict the responses quantities involved in the performance objectives (Jalayer et al. 2013).

A damage scale, as described before, is a metric that describes the physical state that a structure can reach when subjected to a seismic excitation. Each physical state, also referred to in the literature as DS, is associated with a specific level of damage, e.g. light, moderate, extensive damage and collapse. The threshold that distinguishes two consecutive DSs is known as damage threshold (or damage level) and is expressed in terms of EDPs. These thresholds can be defined at a global (structure) level (i.e. EDP expressed as drift ratio or displacement (e.g., Akkar et al. 2005; Erberik 2008)), or in a rigorous manner at a local (member) level (i.e. EDP expressed as member strain of a distinct building, (e.g., Crowley et al. 2004). The aim of the present study is to derive fragility functions of a single or multiple buildings within a simplified analysis framework. Thus, global level damage thresholds (expressed as thresholds maximum inter-storey drift ratios) are adequate to define the successive fragility curves.

A fragility curve can be expressed as a conditional probability assuming a standard normal cumulative distribution function (CDF), Φ. The lognormal CDF is a simple parametric model used to represent a fragility function. The main advantage of this model is the fact that is bounded between 0 and 1 in the y-axis, which represents the probability of damage. On the other hand, the x-axis, representing the IM can only take non-negative values (zero is the lower

bound), meaning that for 0 IM value no damage is expected. It is also mathematically convenient, that when a random variable is assumed to be lognormally distributed and is multiplied by factors (e.g., safety factors), which are also uncertain and lognormally distributed, the end product (i.e. fragility curves) will also follow a lognormal distribution (Shinozuka et al. 2000). The fragility curve can be mathematically expressed as the probability that an EDP will exceed a certain damage state value DS_i for a given intensity IM:

$$P(EDP \geq DS_i \mid IM = im) = 1 - \Phi\left(\frac{\ln DS_i - \ln \mu_{EDP|IM}}{\beta_{\ln EDP|IM}} \right) \tag{2.5}$$

where $\Phi\left(\frac{\ln DS_i - \ln \mu_{EDP|IM}}{\beta_{\ln EDP|IM}} \right)$ represents the standard normal distribution function, $\mu_{EDP|IM}$ and $\beta_{\ln EDP|IM}$ are the estimated median and the standard deviation of the natural logarithm of the EDP ($\ln(EDP)$) given IM. The above definition is suitable when the limit state is based on a threshold value of an EDP. In the case of an IM-based definition of a limit state, which is common for the collapse limit state, the fragility function can be defined as follows:

$$P(DS_i \mid IM = im) = \Phi\left(\frac{\ln im - \ln \mu_{imDS}}{\beta_{\ln imDS}} \right) \tag{2.6}$$

where $P(DS_i \mid IM = im)$ is the probability of exceeding the DS_i if the intensity measure assumes a value equal to im, μ_{imDS} is the median damage state intensity, and $\beta_{\ln imDS}$ is the corresponding standard deviation of the natural logarithms. The fragility function is therefore defined by two parameters, the median value ($\mu_{EDP|IM}$ or μ_{imDS}) and the corresponding standard deviation ($\beta_{\ln EDP|IM}$ or $\beta_{\ln imDS}$) originated from the uncertainty within the process of determining a DS for a given EDP or IM.

This uncertainty can be divided into two types, the aleatory uncertainty (due to inherent randomness) and the epistemic uncertainty (due to lack of knowledge) (Jalayer and Cornell 2009). In general, the total standard deviation of the fragility function can be associated with

several sources of uncertainty, including the ground motion variability, the structural capacity among others. D'Ayala et al. (2013a) recommends a relationship that encapsulates the different sources of uncertainty within the calculation of the total standard deviation (β):

$$\beta = \sqrt{\beta_c^2 + \beta_d^2 + \beta_{DS}^2} \qquad (2.7)$$

where β_c is the standard deviation linked to the structural capacity, i.e. uncertainties in the modelling, in material properties, in geometrical characteristics etc.; β_d is the standard deviation associated with uncertainty in demand, i.e. uncertainty in ground motion variability; β_{DS} is the standard deviation in the threshold defining a specific damage level.

2.7.5   Statistical procedure to fit fragility

In the existing literature of fragility analysis, the main focus is laid on the different methods for obtaining damage data (e.g., analytical and empirical methods); however, the choice of statistical models, and consequently statistical fitting procedures, also play a key role in increasing the accuracy of fragility analysis frameworks.

In the following subsections some of the most commonly used statistical procedures adopted in the derivation of fragility curves are discussed, namely:

- Least squares
- Maximum likelihood estimation
- Generalised linear model

A complete review of these statistical fitting procedures, emphasizing the strengths and shortcoming of each method, was carried out by Lallemant et al. (2015). In addition, recent studies have also investigated the sensitivity of different fitting techniques to the number of ground motions used (Gehl et al. 2015). It is worth noting that the simplified fragility derivation procedure FRACAS (extensively described in Chapter 4) supports all three statistical fitting procedures presented below.

2.7.5.1 Least Squares

The Least Squares (LS) regression, also referred to as "regression on a cloud" in some studies, is a variant of method of moments. As discussed in Section 2.7.3.1, a linear regression model is fitted on the structural analysis results, expressed as the logarithms of IM and EDP values, under the assumption that a power-law model described the aforementioned relationship.

The response residuals of the regression are then computed, and β_e denotes the standard deviation of the residuals. For a given threshold representing a specific damage state, DS_i, the parameters of the fragility are determined by the expressions:

$$\begin{cases} \mu = \exp\left(\ln\left(\dfrac{DS_i}{b} \right) \Big/ a \right) \\ \beta = \dfrac{\beta_e}{a} \end{cases} \qquad (2.8)$$

where μ and β are the median demand and the standard deviation respectively of each fragility curve, and a, b are the parameters of the regression. In contrast to other methods discussed below, LS implements the actual EDP estimates. The strength of the LS approach is that it only requires a few points to characterise the IM and EDP relationship, as it takes advantages of all the information enclosed in the dynamic analysis results. On the other hand, the limitation of this method is that the standard deviation is estimated for the whole range of IM levels, resulting in a constant dispersion in fragility curves across the IM range. One simple way to overcome this shortcoming is to perform piecewise regressions for different IM intervals (Cărăuşu and Vulpe 1996).

It is highlighted that the method of moments is only valid for analytical data and experimental data of specimens tested to failure, but is neither applicable in empirical methods (where it is not always possible to obtain accurate IM values for the associated field damage observation) nor MSA (NDA carried out a certain number of IM levels), (Lallemant et al. 2015).

2.7.5.2 Maximum Likelihood Estimation

Another statistical fitting approach that has been used in fragility studies is Maximum Likelihood Estimation (MLE), proposed by Shinozuka et al. (2000). In this approach, the EDPs are classified to specific DSs by implementing a binary scheme. For example, EDP values are compared to damage thresholds associated with to specific DS and a value of 0 or 1 is allocated for the cases the threshold is not exceeded or exceeded respectively.

Similarly to LS method, MLE also assumes that the EDP|IM results follow a lognormal distribution. As its name indicates, MLE determines the parameters of the distribution (median and standard deviation) that maximise the likelihood that the corresponding distribution produces the observed data.

In general, MLE approach is considered to be more appropriate than LS for deriving fragilities from empirical data as it classifies information into two bins (damaged, not-damaged). However, several recent studies have used MLE to develop analytical fragility functions (e.g., Gehl et al. 2013). The main shortcoming of applying MLE to analytical data is that a substantial amount of information is not taken into account (i.e. actual seismic responses). This, however, can be considered as an advantage for near-collapse and collapse estimates, where most numerical approaches lack precision or can experience convergence issues (Baker 2015). In contrast to the LS, MLE is not assigned any idealisation model (e.g., power-law) to describe the IM and EDP relationship.

2.7.5.3 Generalised Linear Model

The generalised linear regression, also called Generalised Linear Model (GLM), is a variant of ordinary linear regression, where the performance predictor variables are interrelated linearly to the response by implementing a link function (Lallemant et al. 2015).

GLM consists of three components:

- A probability distribution of the exponential family
- A linear predictor and

- A link function

It is noted that GLMs also implement a binary scheme and utilise MLE to fit the fragility model. In particular, a binary indicator y_i is used to assess the performance (i.e. DS) of the given EDP values:

$$y_i = \begin{cases} 1 & EDP_i \geq EDP_{th} \\ 0 & EDP_i < EDP_{th} \end{cases} \tag{2.9}$$

A Bernoulli distribution is then employed to represent the relation between the indicator y_i and the intensity of the ground motion im, as shown below:

$$Y\Big|im \sim \begin{pmatrix} 1 \\ y_i \end{pmatrix} \mu_{Fr}{}^{y_i} \left[1 - \mu_{Fr}\right]^{1-y_i} \tag{2.10}$$

where μ_{Fr} is the conditional mean response that represents in this case the fragility curve.

The fragility curve for a specific DS is given by the following expression:

$$g(\mu) = g\left(P\left(DS \geq ds \mid im\right)\right) = \theta_1 \ln(im) + \theta_0 \tag{2.11}$$

where θ_0 and θ_1 are the unknown parameters of the GLM model and $g(.)$ the link function. Some of the most commonly used link functions in GLMs, which are also implemented in FRACAS, are the probit (equivalent to MLE approach described above), the logistic and the complementary log-log link functions:

$$g(\mu) = \begin{cases} \Phi^{-1}\left(\mu_{Fr}\right) & \text{probit} \\ \log\left(\dfrac{\mu_{Fr}}{\mu_{Fr} - 1}\right) & \text{logit} \\ \log\left[\log\left(1 - \mu_{Fr}\right)\right] & \text{complentary loglog} \end{cases} \tag{2.12}$$

For fragility curves consisting of multiple DSs an ordered GLM may be used, to ensure that there will be no intersection amongst fragility curves for consecutive DSs. This procedure is thoroughly described in (D'Ayala et al. 2013).

A general description of GLM can be found at McCullagh and Nelder (1989); for detailed comparisons between different GLM link functions and GLM fitting fragility applications, one may refer to Lallemant et al. (2015) and Baker (2015) respectively.

This statistical method tends to yield unstable values when data points are scarce, while the regression on a cloud would still be able to provide stable fragility estimates; however, it can be argued that the GLM regression prevents the modeller from using erroneous results, since the statistical model fails before proposing poorly constrained estimates (Gehl 2017).

2.7.6 Surrogate metamodels

Propagating uncertainties usually requires a large number of repeated calls to the model for different values of the input parameters, for instance, through a Monte Carlo simulation procedure. Such an approach requires thousands to millions of analysis runs, which is not realistic in many practical cases even when high-performance computing is used. To tackle this problem, surrogate models may be used, which replace the original computational model by an easy-to-evaluate function (Hastie et al. 2001). These surrogate models, also known as response surfaces or metamodels, (these terms are used in this Thesis interchangeably), are capable of quickly predicting responses to new input realizations. Whenever the quantity of interest is computationally expensive to evaluate for a set of input points, one can replace this quantity with a surrogate model, to provide an approximate relationship between the quantity and the input that is fast to evaluate (Jin et al. 2001). This allows large numbers of model evaluations to be carried out, such as simple uncertainty propagation, structural reliability analysis and design optimization, in a reasonable time.

Several types of surrogate models exist in literature, including Artificial Neural Networks (ANN), Bayesian Networks (BN), Polynomial Chaos Expansions (PCE) and Kriging (also

known as Gaussian process modelling) to name a few. Each of these techniques has strengths and limitations regarding the nonlinear behaviour of the model, the dimensionality, the data sampling approach etc., making each technique more appropriate for particular types of problems. For detailed comparisons of the various surrogate modelling techniques one may refer to Jin et al. (2001), Kroetz and Beck (2015) and Schobi (2017).

2.7.6.1 Overview of surrogate models in earthquake engineering applications

In earthquake engineering and structural reliability several researchers have attempted to implement surrogate modelling techniques to improve computational efficiency (e.g., Fragiadakis et al. 2015; Gidaris 2015; Giovanis and Papadopoulos 2015; Kroetz and Beck 2015; Jia and Taflanidis 2013, to name a few). However, in most cases the focus of these metamodel-based approaches is laid on dealing with the epistemic uncertainty.

Specifically, Giovanis & Papadopoulos (2015) and Fragiadakis et al. (2015) utilised ANN combined with Monte Carlo simulation (MCS) in order to account for the uncertainty associated with the mechanical properties of the structure (i.e. using random phase angles or random variables respectively to describe the uncertain structural parameters). Ferrario et al. (2015) used ANN to replace the computationally demanding finite element model for the simulation of seismic response of masonry buildings.

Gehl and D'Ayala (2016) developed a BN-based procedure for the derivation of multi-hazard fragility functions, which allows the integration of multiple risks at the vulnerability level. This approach also allows the derivation of component fragility curves for different hazard loadings and failure mechanisms, with different methods (e.g., analytical or empirical). However, the use of multiple methods may raise some issues regarding the uncertainty treatment, as different methods may use different uncertainty characterization strategies.

Gidaris et al. (2015) implemented a kriging-based framework to deal with both the uncertainties related to the structural properties and the hazard (i.e. seismicity characteristics), treating all uncertain parameters as random variables. It is noteworthy to mention that the description of

hazard used in this framework was based on ground motion (stochastic) simulations. An interesting advancement would be to use real earthquake records to describe seismic hazard and have a more realistic quantification of the aleatory uncertainty.

2.7.6.2 Kriging metamodels

Among the most popular surrogate modelling techniques is kriging surrogate modelling (Sacks et al. 1989), which originates from geographical data in mining (Krige 1952) and today is also known as Gaussian process modelling (Rasmussen and Williams 2006). The Kriging metamodel is interpreted as a realization of a Gaussian process (GP). Kriging merits its popularity to the ability to easily generate confidence intervals for predictions; a very attractive feature for problems involving quantification of uncertainty. Here kriging is described in further detail as it is later referred to in Chapter 6.

Let $\mathbf{y}$ denote the n_y dimensional vector of the output quantities to be approximated by the surrogate model, and $\mathbf{x}$ the n_x dimensional input parameter vector for the kriging formulation. To form the surrogate model, a database of n_m observations is obtained, which provides information about the pair $\mathbf{x}$-$\mathbf{y}$. This is a very common process used, and is known as design of experiments (DoE). To this aim, n_m samples for $\left\{ \mathbf{x}^j, j=1,...,n_m \right\}$, also known as support points (or training points), are created , following usually a Latin-hypercube grid (McKay et al. 1979) over the expected range of possible values for each $x_i, i=1,...,n_x$.

The above dataset is used to formulate the kriging model providing the kriging predictor. The kriging predictor, which is a GP, has a mean $\hat{y}_k(\mathbf{x})$ and standard deviation $\sigma_k(\mathbf{x})$ (Sacks et al. 1989), and its response output is calculated as:

$$y_k(\mathbf{x}) = \hat{y}_k(\mathbf{x}) + \varepsilon_k^g \sigma_k(\mathbf{x}) \tag{2.13}$$

where ε_k^g is a standard Gaussian variable with zero mean and a variance of one. The fundamental building blocks of kriging are the n_p dimensional basis vector, $\mathbf{f}(\mathbf{x})$, and the

correlation function, $R\left(\mathbf{x}^l, \mathbf{x}^m\right)$, defined through the parameter vector $\mathbf{s}$. More details on the selection of both $R\left(\mathbf{x}^l, \mathbf{x}^m\right)$ and $\mathbf{s}$ can be found in (Gidaris 2015).

Next, for the set of training point observations n_m, with an input matrix $\mathbf{X} = \left[\mathbf{x}^1 ... \mathbf{x}^{n_m}\right]^T$ and output matrix $\mathbf{Y} = \left[\mathbf{y}\left(\mathbf{x}^1\right) ... \mathbf{y}\left(\mathbf{x}^{n_m}\right)\right]^T$, the corresponding basis matrix $\mathbf{F}$ and the lm-element correlation matrix $\mathbf{R}$ are defined. In addition, for every new input $\mathbf{x}$, the correlation vector $\mathbf{r}(\mathbf{x}) = \left[\mathbf{R}\left(\mathbf{x}, \mathbf{x}^1\right) ... \mathbf{R}\left(\mathbf{x}, \mathbf{x}^{n_m}\right)\right]^T$ is defined between the input and each of the elements of $\mathbf{X}$. As a result, the mean kriging approximation is obtained as (Lophaven et al. 2002):

$$\hat{\mathbf{y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\alpha}^* + \mathbf{r}(\mathbf{x})^T \boldsymbol{\beta}^* \tag{2.14}$$

where parameters $\boldsymbol{\alpha}^*$ and $\boldsymbol{\beta}^*$ are defined as:

$$\boldsymbol{\alpha}^* = \left(\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F}\right)^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{F} \tag{2.15}$$

$$\boldsymbol{\beta}^* = \mathbf{R}^{-1}\left(\mathbf{Y} - \mathbf{F}\boldsymbol{\alpha}^*\right) \tag{2.16}$$

The calibration of parameters $\mathbf{s}$ of the kriging correlation function can result in efficient approximations of complex functions. To this aim, the MLE principle is used for the estimation of optimal parameters $\mathbf{s}$. According to Lophaven et al. (2002), maximizing the likelihood with respect to $\mathbf{s}$ leads to an optimisation of the problem:

$$\mathbf{s}^* = \arg\min\left[|\mathbf{R}|^{1/n_m} \sum_{k=1}^{n_y} \hat{\sigma}_k^2\right] \tag{2.17}$$

where $|.|$ represents the determinant of a matrix and $\hat{\sigma}_k^2$ correspond to the diagonal elements of the matrix $\left(\mathbf{Y} - \mathbf{F}\boldsymbol{\alpha}^*\right)^T \mathbf{R}^{-1}\left(\mathbf{Y} - \mathbf{F}\boldsymbol{\alpha}^*\right) / n_m$. The prediction error variance $\sigma_k^2(\mathbf{x})$ for y_k and input $\mathbf{x}$ is computed as:

$$\sigma_k^2(\mathbf{x}) = \sigma_{\hat{k}}^2 \left[1 + \mathbf{u}^T \left(\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F} \right)^{-1} \mathbf{u} - \mathbf{r}(\mathbf{x})^T \mathbf{R}^{-1} \mathbf{r}(\mathbf{x}) \right] \qquad (2.18)$$

where $\mathbf{u} = \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}(\mathbf{x}) - \mathbf{f}(\mathbf{x})$. The performance of the surrogate model can be validated directly by the process variance $\sigma_{\hat{k}}^2$, or by calculating different error statistics for each component of the output vector $\mathbf{y}$, such as the coefficient of determination RD_i^2 or the mean percent error ME_i, using a leave-one-out cross-validation approach (Kohavi 1995). According to this approach, each observation is removed sequentially from the dataset, and the remaining training points are used to predict the output of that one and then evaluate the error between the predicted and the real responses. The validation statistics are then obtained by averaging the errors corresponding to all observations. There are two main strategies to enhance the performance of a surrogate model: (i) to increase the number of training points n_m, or (ii) to carefully select the location of training points n_m (Picheny 2009; Picheny et al. 2010).

2.8   Estimation of damage impacts and direct economic loss

Earthquake-induced structural and nonstructural damage to buildings is the main cause of direct economic and life losses (Shome 2014). Physical damage can be translated into monetary loss, (i.e. estimation of repair and reconstruction cost), through damage-to-loss functions. Similarly, fragility functions, can be converted to vulnerability functions, which express the exceedance probability of a measure of loss, L, as a function of IM (see Rossetto et al. 2014b). L, is defined using numerical scales (for example the ratio of repair cost to the component replacement value, also known as the mean damage ratio), which allows its direct use in probabilistic risk and loss calculations (Yamin et al. 2017). Usually a Beta probability distribution function is assumed for the calculation of losses (ATC 1985). There are two main approaches for the development of vulnerability curves, depending on data availability, project objectives, and analyst's expertise (D'Ayala et al. 2013):

- Building-based vulnerability assessment approach: Vulnerability functions are obtained by convolving building level fragility curves with the cumulative cost of a given DS (damage-to-loss functions). This is the most commonly used approach in the literature (especially for large scale projects such as insurance applications), as it requires less detailed data availability and significantly less computational resources comparing to the following approach.

- Component-based vulnerability assessment approach: Vulnerability functions are constructed by correlating the components level-based drifts directly to loss (Aslani 2005,ATC 2012a, b). In general, this approach is appropriate when performing loss analysis for single buildings, or when the majority of the economic losses are related to content and non-structural components.

To transform a fragility function into a vulnerability function, the following total probability relation can be used:

$$E\left(L \mid IM = im\right) = \sum_{i=0}^{N_{DS}} L_i \cdot P\left(DS_i \mid im\right)$$

(2.19)

where $E\left(L \mid IM = im\right)$ is the mean loss given a certain ground motion intensity im, N_{DS} is the number of damage states, L_i is the repair cost assigned to each damage state DS_i (as obtained from damage-to-loss functions described below) and $P\left(DS_i \mid im\right)$ is the discrete probability of reaching a DS_i. The equation (2.20) below gives the loss variance given a certain ground motion intensity $\sigma^2(L \mid IM = im)$:

$$\sigma^2(L \mid IM = im) = \sum_{i=0}^{N_{DS}} \left(L_i - E\left(L \mid im\right)\right)^2 \cdot P\left(DS_i \mid im\right)$$

(2.20)

Various empirical relationships, conforming with the seismic risk framework, exist in the literature to translate structural and nonstructural damage into percentage of losses. These relationships are known as damage-to-loss functions and are expressed in terms of mean

damage ratio (MDR) to derive direct financial losses to buildings based on their assigned damage. MDR is defined as the ratio of repair cost over replacement cost. Damage-to-loss functions are generally determined from 'expert' judgment or empirical post-earthquake studies of loss data, and in many cases are proposed as part of damage scales (Hill 2011). It is noted that in existing literature MDR is sometimes referred to with different terms, such as damage factor (e.g., Di Pasquale and Goretti 2001), cost ratio (e.g., Crowley et al. 2005), replacement cost ratio (e.g., Erdik and Durukal 2008). However, to avoid confusion, only the term MDR is adopted in this Thesis. Exhaustive reviews of existing damage to loss functions can be found in Hill (2011) and Rossetto and Novelli (2016).

2.9   Research gaps

Regarding the analytical seismic risk assessment of RC buildings, the literature review has identified a wide range of approaches and studies, which demonstrate the significant technical developments that have contributed to this topic. The majority of studies focus on a specific component within this approach and attempt to evaluate the impact of each research development towards the resultant outputs (e.g., effect of IM selection on fragility analysis, or impact of spatial correlation models in the estimation of loss). Despite this, very few studies have attempted to quantify at the same time the impact of multiple components within the seismic risk framework. It is a widely held view that the uncertainty arising from the hazard component has the highest effect on the variability observed in the structural response (e.g., Shome et al. 1998; Padgett and DesRoches 2007), however, the fragility/vulnerability component also affects significantly the resultant response and consequently the losses. Therefore, understanding the impact of the different input assumptions in the resultant loss estimates and risk quantification will not only contribute towards the development of a more robust framework, but will also serve as a guide to prioritize the potential improvements of future catastrophe models.

Existing analysis methods for seismic response calculation include advanced (NDA) and simplified (NSP) analysis approaches, where each method is associated with some advantages and shortcomings as discussed in Sections 2.7.1 and 2.7.2. Specifically, NDA represents the most reliable and accurate approach for computing seismic response of structures. However, the excessive computational cost precludes the use of NDA for the case of large scale seismic risk assessment projects, such as catastrophe models. On the other hand, NSPs, such as CSM, are widely used for seismic response analysis in the current catastrophe modelling practice, due to their ability to perform nonlinear analysis rapidly. CSM-based methods rely on simplifying assumptions, therefore they can fail to identify alternative failure mechanisms and they cannot account for record-to-record variability, neglecting some major sources of uncertainty (Fragiadakis et al. 2014). A gap is identified between the two analysis methods, highlighting the need for a new method that overcomes some of the limitations of NSPs but maintaining the advantage of swift analysis times.

The selection of hazard parameters (i.e. IMs) for fragility and vulnerability analysis is one of the most extensively studied topics in the field of seismic risk. Numerous studies have been conducted, which assess the performance of scalar (e.g., conventional, spectral shape based) and vector-valued IMs within various seismic risk applications. To this aim, a number of IM selection criteria (including innovative selection criteria based on the information theory) have been proposed in the literature for assessing the appropriateness of a candidate IM for fragility analysis. It is worth noting that the above studies rely on NDA or NDA-based approaches (e.g., IDA) for calculating seismic response of structures, in order to determine an optimal IM. Despite this, none has attempted to convolve all these state-of-the-art concepts and identify optimal IMs to be used in probabilistic seismic demand analysis through NSPs.

With regard to the implementation of surrogate models, several metamodels have been proposed for earthquake engineering applications, as presented in Section 2.7.6, however, the following points are rarely discussed:

- Surrogate models are rarely used to fit the IM|EDP relationship: As discussed in Section 2.7.6.1, several applications exist in earthquake engineering literature, however, most of these applications focus on the fragility component and/or the uncertainty associated with the mechanical properties of the structure.

- The use of real earthquake records to describe seismic hazard: In most surrogate model applications the seismic hazard is described by using stochastic/synthetic ground motions (Gidaris et al. 2015; Wang et al. 2018). However, the use of real ground motion records will result in a more realistic quantification of the aleatory uncertainty.

- Uncertainty in the training data sample: The efficiency of the predictions of a surrogate model significantly depends on the training process of the metamodel (see Section 2.7.6.2). In most cases, the data set utilized to train a surrogate model is obtained from high fidelity analyses, i.e. computationally expensive simulations (e.g., NDA). However, the majority of studies found in the literature utilise such data sets without accounting for the uncertainty of the high fidelity simulations.

The following Chapters look to explore the identified research gaps and answer the Research Questions introduced in Chapter 1. This is done by advancing the development of a capacity spectrum approach for fragility assessment of buildings (Chapter 4), investigating IM selection suitability for use in NSP-based fragility functions (Chapter 5), and proposing a new approach based on surrogate modelling for the construction of analytical fragility functions (Chapter 6). The influence of each of these finding on risk evaluation is finally investigated in Chapter 7.

3. Structural modelling and analysis approach

3.1 Introduction

The present Chapter introduces the three mid-rise RC MRF case study buildings, which are selected as structural inputs to illustrate the developments presented throughout this Thesis. Each of the three buildings is considered to be a central index building that represents populations of buildings corresponding to three different vulnerability classes, depending on their period of construction. The buildings are analysed with the aim of analysis methods of different complexity (advanced and simplified) in order to derive sets of fragility functions.

This Chapter begins with a detailed description of the three case study buildings, providing in depth information about the design specifications of each building, and stating the modelling assumptions considered for the development of the computer models. Next, the buildings are analysed with the aim of advanced and simplified analysis methods (NDA and FRACAS respectively), and the resultant seismic response is presented in terms of different hazard representations (i.e. IMs). Finally, a tailored damage scale is developed specifically for each case study building and used for the derivation of distinct sets of fragility functions. The building models and the analysis outputs presented in this Chapter are considered for the illustration of techniques and sensitivity tests proposed in the Chapters to follow.

3.2 Case study buildings

Three regular RC 4-storey, 4-bay bare frames are modelled and analysed for the purposes of this research. Each building is associated with a different level of seismic design, and together they constitute the population of structural models used in this study to represent the mid-rise RC building class. Specifically, the selected case-study structures share the same geometry (bay widths and storey heights) but are characterized by different material properties, element cross-sections and reinforcement detailing. The first frame is designed to only sustain gravity loads

following the Royal Decree n. 2229 (1940) that regulated the design of RC buildings in Italy up to 1971, hereafter Pre-code building; the second frame was designed according the Decreto Ministeriale (1972), hereafter Low-code building; and the third frame was designed according to the recent Italian seismic code, NIBC08 (Decreto Ministeriale del 14/01/2008 2008), following the High Ductility Class (DCH) rules, hereafter Special-code building. Further information regarding the design of Pre-code and Special-code buildings is available in De Luca et al. (2009). In this Thesis the ''archetype analysis model'' approach is used, based on the archetype framework described in Haselton (2007). The ''archetype analysis model'' refers to the simplest model which is able to capture the associated seismic response of the building of interest, in this case mid-rise RC MRF bare frame buildings. Specifically, the two-dimensional frame analyses here are considered to be representative of the three-dimensional building seismic behaviour, based on the simplifying assumption that the building is symmetric; therefore, the same structural frame can be used to simulate the seismic response in both directions (Haselton 2007). Inter-storey heights, span of each bay and cross-sections dimensions for each case-study building are reported in **Figure 3.1**. The considered frames are regular in plan and in elevation.

In the case of the Pre-code building, concrete with characteristic compressive strength f_{ck} = 19MPa and reinforcement of characteristic yield stress f_{yk} = 360MPa are used. In the case of the Low-code building, concrete with characteristic compressive strength f_{ck} = 25MPa and reinforcement of characteristic yield stress f_{yk} = 380MPa are used. Finally, the characteristic compressive strength of concrete and the characteristic yield stress of steel reinforcement building used for Special-code are f_{ck} = 29MPa and f_{yk} = 450MPa, respectively. The characteristic compressive strength of concrete and characteristic yield stress of steel reinforcement for each case study building are presented in tabulated form in **Table 3-1**.

All frames are modelled using the finite element platform SeismoStruct (Seismosoft 2007). The effect of confinement is taken into account by implementing the confinement model proposed by Mander et al. (1989). An insufficient level of confinement is observed in all sections of the

Pre-code building: the confinement factor, k, is defined as the confined-unconfined concrete compressive stress ratio and ranges from 1.01 to 1.05.
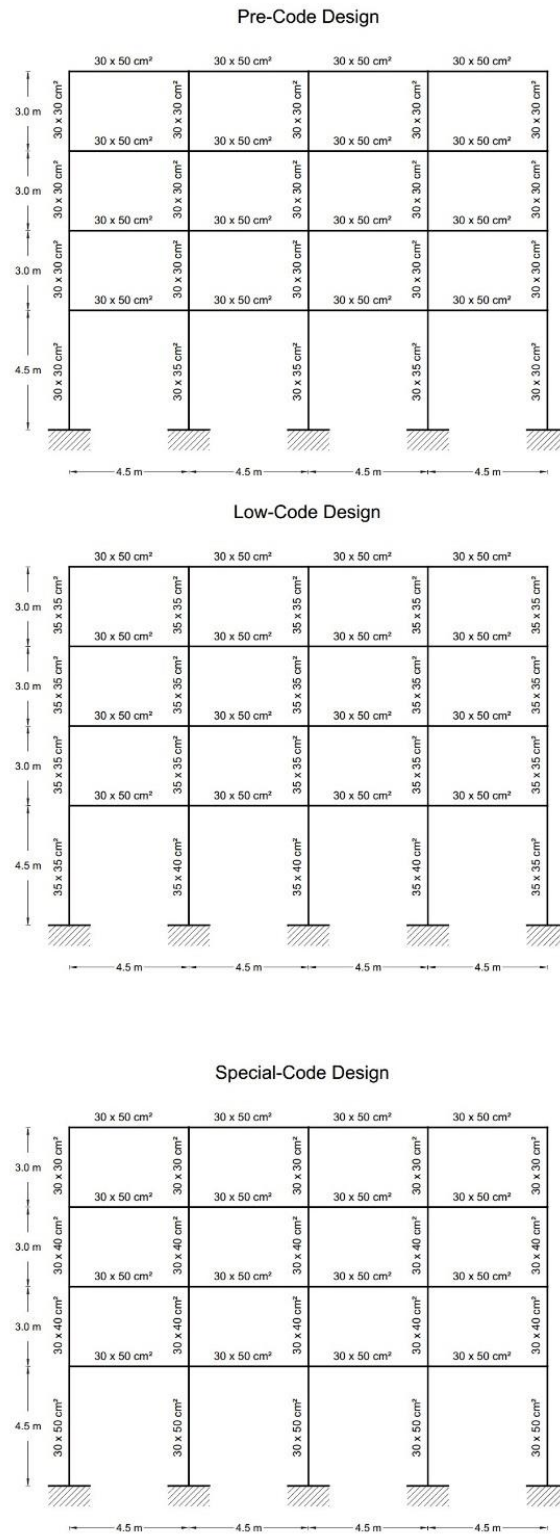
Figure 3.1 - Elevation dimensions and members cross-sections of the considered RC frames.

Table 3-1 - Characteristic compressive strength of concrete and characteristic yield stress of steel reinforcement for each case study building.

Case study building	f_{ck} [MPa]	f_{yk} [MPa]
Pre-code	19	360
Low-code	25	380
Special-code	29	450

The uniaxial hysteretic stress-strain relation proposed by Menegotto and Pinto (1973) is used to represent the reinforcement steel behaviour with the parameters proposed by Filippou et al.(1983) for the inclusion of isotropic strain-hardening effects. To account for material inelasticity, a distributed plasticity approach is used. Thus each RC section consists of a total of 150 steel, confined concrete and unconfined concrete fibres.

3.2.1 Nonlinear static pushover analysis

Two sets of static PO analyses are carried out with different applied lateral load distributions, namely uniform and triangular. Lateral loads are incrementally applied to the side nodes of the structure. These lateral loads are proportionally distributed with respect to the local masses at each floor level (uniform distribution) and the inter-storey heights (triangular distribution). It is noted that according to ASCE/SEI41-17 (ASCE 2017), TRI-PO load distribution is recommended to carry out static PO analysis. However, UNI-PO was preferred for the case of Pre- and Low-code buildings as it resulted more conservative PO curve estimates, while TRI-PO was utilized for the case of Special-code building, in accordance with EC8 recommendation. In particular, UNI-PO analysis is used for the cases of Pre- and Low-code frames, as a local mechanism is considered to be more representative for buildings of these vulnerability classes; a soft-story failure mechanism at the ground level is observed due to brittle shear failure of ground level columns. TRI-PO is deemed appropriate to capture the global failure mechanism is observed for this building type. In both cases, the PO analysis is carried out

until a predefined target displacement is reached, corresponding to the expected partial collapse state (as defined in **Table 3-3**).

Table 3-2 summarizes the structural and dynamic properties associated with each of the case-study building models, namely the type of PO analysis, the total mass of the system m and the fundamental period T_1 (based on Seismostruct estimations).

Table 3-2 - Structural and dynamic information of each case study building.

Case study building	PO Analysis Type	Total mass [t]	T_1 [s]
Pre-code	UNI	172.9	0.902
Low-code	UNI	177.6	0.673
Special-code	TRI	178.3	0.506

Figure 3.2 presents the static PO curves for the three tested case-study buildings. The curves are reported in terms of top centre of mass displacement divided by the total height of the structure (i.e. the roof drift ratio, RDR) along the horizontal axis of the diagram, and base shear divided by the building seismic weight along the vertical axis (i.e. base shear coefficient). This figure shows the capability of the structural model to directly simulate response up to collapse.
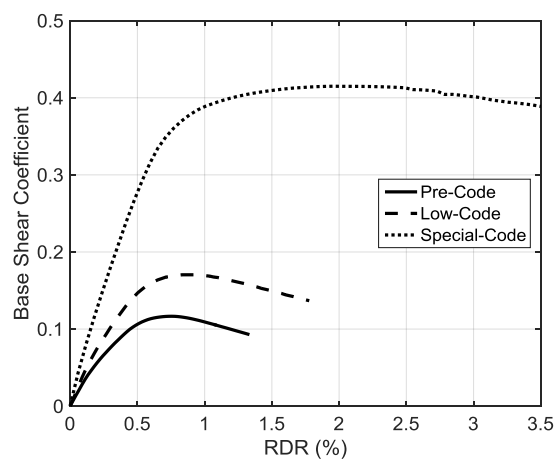
Figure 3.2 - Static PO curves for the case-study buildings.

It is noted that the generated static PO curves are used as input for the variant of capacity spectrum method FRACAS, presented in Chapter 4.

3.2.2 Nonlinear dynamic analysis

The response of case study buildings is calculated by implementing NDA. NDA is used to represent an analysis type of higher precision and will be referred to as high fidelity procedure. NDA is the state-of-the-art of seismic analysis when high accuracy is the objective; however, it is a time-consuming and computationally expensive procedure. In NDA, detailed structural models are subjected to a compendium of ground motions (scaled or unscaled, natural or artificial) in order to estimate responses and failure mechanisms more reliably. This procedure is extremely sensitive to the selection of ground motion records, as discussed in detail in Sections 2.7.1.1 and 3.2.2.1.

3.2.2.1 Ground motion selection

Over 900 real (i.e. recorded during past events) GMs records are selected from the recent-developed SIMBAD database (Selected Input Motions for displacement-Based Assessment and Design; Smerzini et al., 2013) in order to perform NDA. The SIMBAD database was developed in the framework of ReLUIS[2] 2010-2013 Project, task Displacement Based Approaches for Seismic Assessment of Structures, as a strong ground motion database suitable for displacement-based design and assessment. A compendium of 467 records, consisting of two horizontal (X-Y) and one vertical (Z) components (1401 recordings), generated by 130 seismic events (including mainshocks and aftershocks) that have occurred worldwide, form the aforementioned database. These accelerograms were assembled from various ground motion databases derived for different regions of the world following the selection criteria addressed below:

1. Shallow crustal earthquakes worldwide with moment magnitude (M_w) ranging from 5 to 7.3 and epicentral distance $R \leq 35$ km. This ensures to provide strong ground motion records of engineering relevance for most of the design conditions of interest that can be used without introducing large scaling factors.

---

[2] The laboratories university network of seismic engineering (Rete dei Laboratori Universitari di Ingegneria Sismica)

2. Good quality at long periods, so that only records for which the high-pass cut-off frequency used by the data provider is below 0.15 Hz were considered. Therefore, most records are from digital instruments (about 80%), while from analogue instruments only those records with a good signal to noise ratios at long periods, typically from large magnitude earthquakes, were retained.

3. Availability of site class information based on quantitative criteria.

Figure 3.3 (Left panel) shows the distribution of magnitude and distance for the acceleration records compiled in the database. Most of the records come from Japan (47%), Italy (18%), USA (9%) and New Zealand (16%) with few contributions from European and Middle East countries, Turkey, Iran and Greece (10%), as shown in **Figure 3.3** (Right panel).

A subset of 150 records is considered as the ground motion input to provide a significant number of strong-motion records of engineering relevance for the applications presented in here. These records are selected by first ranking the 467 records in terms of their PGA values (by using the geometric mean of the two horizontal components) and then keeping the component with the largest PGA value (for the 150 stations with highest mean PGA). The same set of ground motions is also used to carry out FRACAS analyses, as discussed in Chapter 4. Details about the 150 ground motion records included in the subset used herein can be found in **Table C 1** in Appendix C.

It is worth noting that no site specific record selection has been performed for the scope of this study, for instance by using the conditional spectra approach (Lin et al. 2013) at different intensity levels. This is consistent with the cloud analysis approach used here; Jalayer et al. (2017) showed that cloud analysis can lead to very good fragility estimates under certain conditions which are met in this study (e.g., the suite of records covers a wide range of seismic intensity levels and the significant portion of records lead to demand to capacity ratios greater than one).
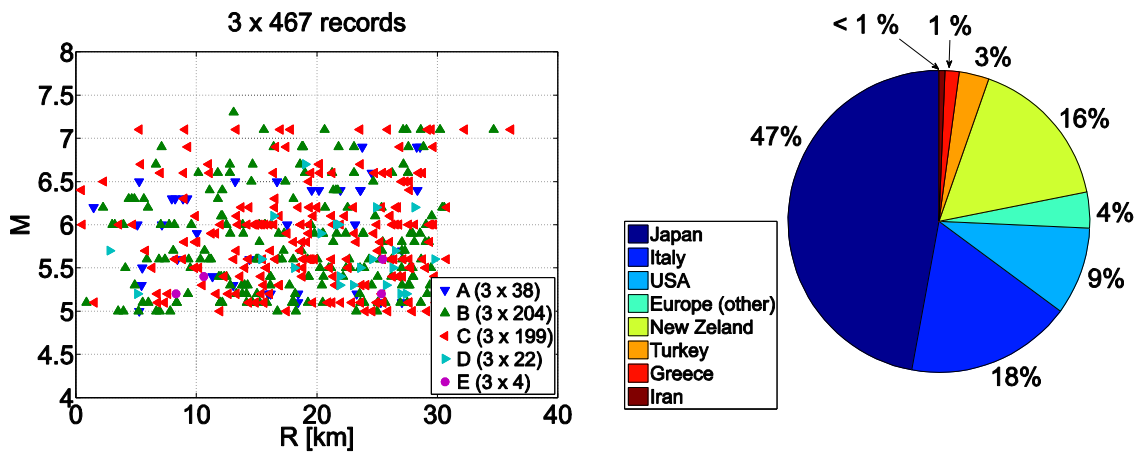
Figure 3.3 - Magnitude versus epicentral distance distribution for SIMBAD dataset. The records are grouped by site class according to Eurocode 8 (EC8; EN 1998-1 2004) classification (Left panel); distribution of the records of SIMBAD by the countries they come from (Right panel) (retrieved from: Smerzini et al. 2013).

3.2.2.2 Seismic response of mid-rise RC buildings in Italy

The results of NDA of the three case study buildings for the subset of the unscaled ground motions discussed in Section 3.2.2.1 are shown in **Figure 3.4** to **Figure 3.6**. These figures plot the peak values of EDP for given IM levels, showing a scatter of points, i.e. the so-called "cloud".

In this study, the deformation-based EDP maximum (over all storeys) peak inter-storey drift ratio (denoted as MIDR) is adopted. The plots show the EDP response with respect to the most commonly used IMs for seismic response and fragility analysis, namely the peak ground acceleration (PGA) and the spectral acceleration at the fundamental period of the each frame ($S_a(T_1)$). In addition to the above IMs, the spectral acceleration at a fixed period level T_1 equal to 1 second ($S_a(1.0s)$) is also tested for all case study buildings, as this is the representative natural period for mid-rise building class in catastrophe modelling.
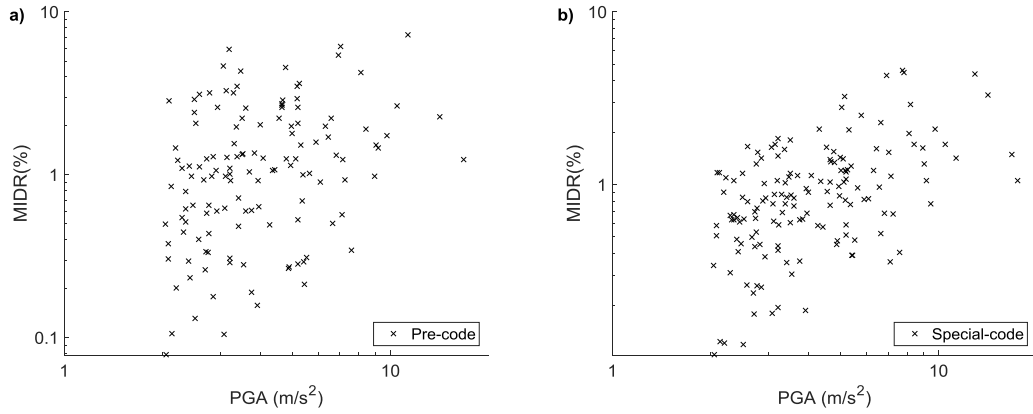
Figure 3.4 - Seismic response results expressed in terms of PGA and MIDR(%) for: a) Pre-code and b) Special-code buildings.
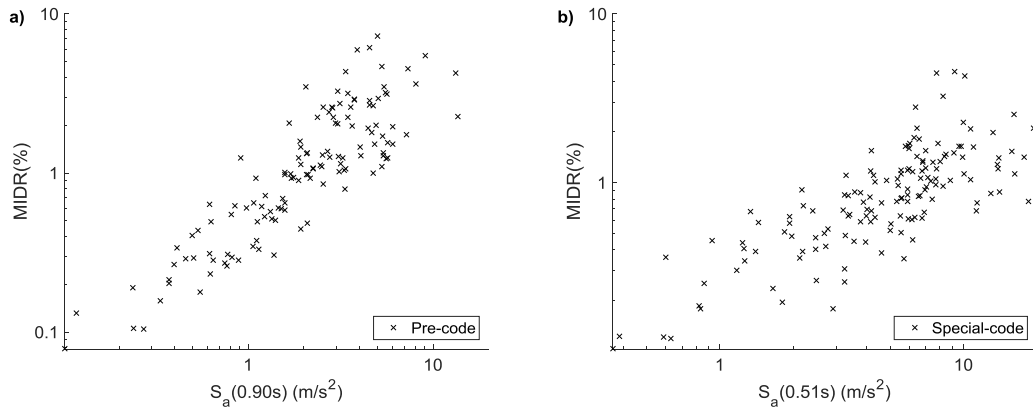
Figure 3.5 - Seismic response results expressed in terms of *Sa(T₁)* and MIDR(%) for: a) Pre-code and b) Special-code buildings.
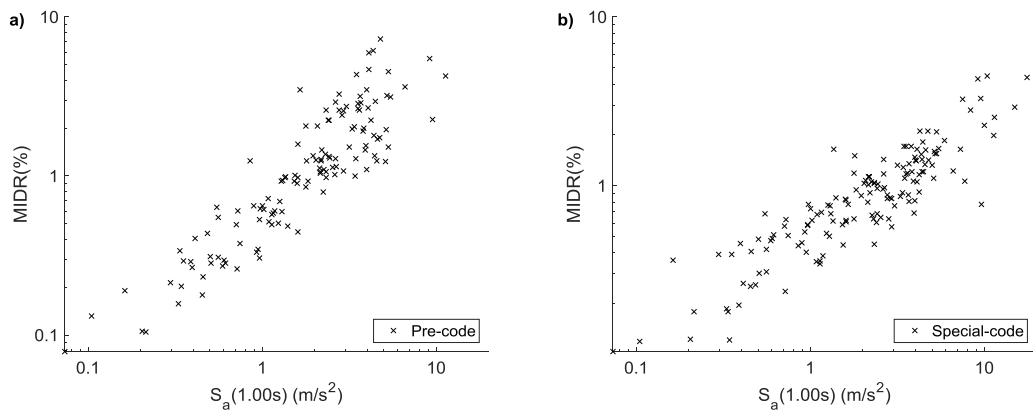
Figure 3.6 - Seismic response results expressed in terms of *Sa(1.0s)* and MIDR(%) for: a) Pre-code and b) Special-code buildings.

Figure 3.4 to **Figure 3.6** illustrate the seismic response of the Pre-code and Special-code building expressed in terms of each of the three tested IMs (i.e. PGA, $S_a(T_1)$ and $S_a(1.0s)$)

103

respectively) against MIDR. The seismic responses of the Low-code building are also computed resulting consistent observations as obtained for all the case study buildings. For the shake of brevity, the seismic responses of the Low-code building are presented in **Figure C 1** to **Figure C 3** in Appendix C.

A close observation of **Figure 3.4** to **Figure 3.6** highlights that the choice of IM has a significant impact on the seismic response scatter, which can affect substantially the construction of the fragility curves. For example, the seismic response scatter associated with the spectral accelerations (**Figure 3.5** and **Figure 3.6**) is narrower than one obtained when PGA is used (**Figure 3.4**). The effect of the choice of IM in fragility analysis is investigated thoroughly in Chapter 5.

3.3   Fragility analysis of mid-rise RC buildings in Italy

There is a number of procedures available in the literature for estimating the parameters of analytical fragility functions. As defined before, fragility functions are continuous relationships between the ground motion IM and the probability that a specified asset will reach or exceed predefined DS in accordance to their structural response. The structural response of buildings, as obtained from the NDA analysis method discussed in the previous Section, is expressed in terms of EDPs, i.e. MIDR, and compared to EDP thresholds associated with each DS. Statistical methods are employed to characterize this response probabilistically as a function of seismic hazard, and build the conditional distribution of an EDP given the IM. In this Section, cloud analysis is used for the generation of the fragility curves. For more details regarding cloud analysis and the least-square regression model employed in this application one may refer to Section 2.7.3.1.

Other well established methods for fragility assessment also exist, such as IDA (Vamvatsikos and Cornell 2002) and Multiple Stripes Analysis (Jalayer and Cornell 2009), but are not considered here due to the excessive computational resources required for the nature of applications such as catastrophe models.

In the following set of fragility curves and in the rest of this Thesis, fragility functions are derived by adopting thresholds of MIDR to define three DSs. The structure response characteristics associated with each DS description are summarized in **Table 3-3** and are based on a re-interpretation of the HRC damage scale of Rossetto and Elnashai (2003) and the damage scale in Dolšek and Fajfar (2008). The DS descriptions of the HRC scale are used but Dolšek and Fajfar (2008) are adopted to define MIDR thresholds for RC MRF designed to modern seismic codes. The MIDR thresholds associated with each DS are then derived from observations of when one of the identified response characteristics first occurs in the building's PO analysis. In this way the DS definitions are tailored to each building (see **Table 3-3** for the thresholds used here for the two model structures). It is noted that for the case of definition of DS3 shear capacity and chord rotation checks are performed according to the expressions defined in Eurocode 8, part 3 (EN 1998-3 2004).

These DS definitions are used to generate all the fragility functions presented in this Chapter and throughout this Thesis. **Figure 3.7** presents the locations at which DSs are reached along the static PO curves for the three case-study buildings.
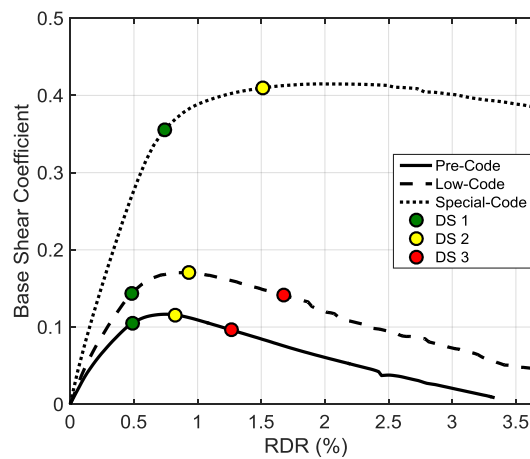
Figure 3.7 - Static PO curves for the case-study buildings with associated damage thresholds.

Table 3-3 - Description of damage states and damage state thresholds used in this Thesis.

HRC Damage State	DS1 Moderate	DS2 Extensive	DS3 Partial Collapse[3]
Observed Damage	Cracking in most beams & columns. Some yielding in a limited number. Limited concrete spalling.	Ultimate strength is reached in some elements.	Failure of some columns or impending soft-storey failure.
Response Characteristics (Threshold defined by the first occurrence of any of these)	Global yield displacement, as obtained by the idealized curve.	Maximum moment capacity of a supporting column is reached.	- There is a drop in strength to 80% of the maximum global capacity; - Shear failure of one element; - The rotation capacity of a critical column is reached.
MIDR Threshold Pre-code structure (%)	0.49	1.53	3.00
MIDR Threshold Low-code structure (%)	076	1.89	4.27
MIDR Threshold Special-code structure (%)	0.95	2.11	5.62

In this Section, only two case study buildings, namely the Pre- and Special-code buildings, are used to demonstrate the derivation of fragility curves for the chosen IMs. Low-code fragility functions are reported in **Figure C 4** to **Figure C 6** in Appendix C. Figure **3.8** illustrates the median fragility curve sets expressed in terms of PGA for Pre-code (left panel) and Special-code (right panel), and their associated 95% confidence intervals estimated using the bootstrap technique. In brief, a large number of bootstrap samples (1,000 in this case) are simulated and a fragility function is drawn for each bootstrap iteration and consequently the 95% bootstrap confidence intervals are computed. For more details about the derivation of bootstrap confidence intervals one may refer to Rossetto et al. (2014). Similarly to **Figure 3.8**, **Figure 3.9**

---

[3] The HRC defined "Partial Collapse" DS corresponds to the Dolšek and Fajfar (2008) "Near Collapse" DS and the Silva et al. (2014a) "Complete" DS.

and **Figure 3.10** present the median fragility curve sets and their associated 95% confidence intervals expressed in terms $S_a(T_1)$ and $S_a(1.0s)$ respectively.
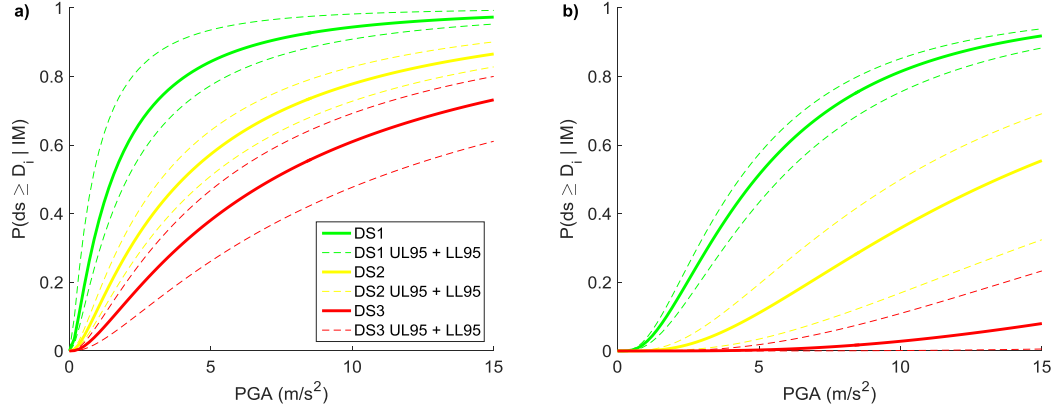
Figure 3.8 - Fragility curves and associated 95% confidence intervals expressed in terms of PGA for: a) Pre-code and b) Special-code building.
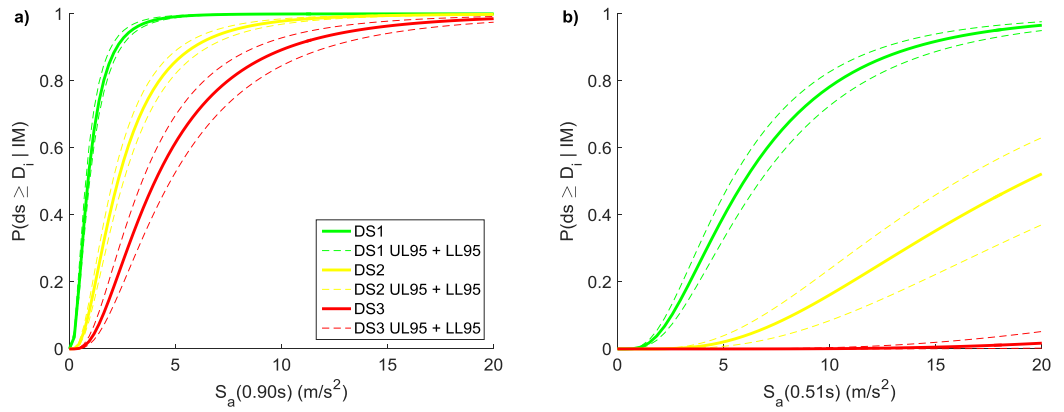
Figure 3.9 - Fragility curves and associated 95% confidence intervals expressed in terms of *Sa(T₁)* for: a) Pre-code and b) Special-code building.
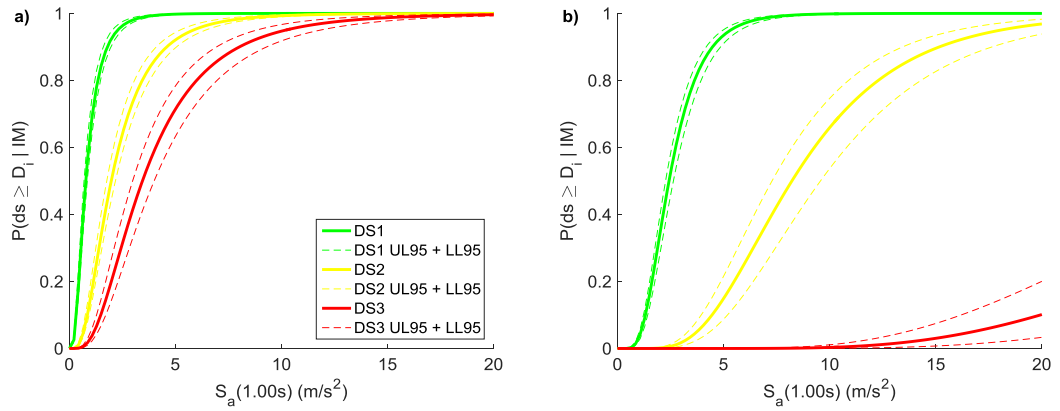
Figure 3.10 - Fragility curves and associated 95% confidence intervals expressed in terms of *Sa(1.0s)* for: a) Pre-code and b) Special-code building.

As expected, the fragility curves expressed in terms of PGA (**Figure 3.8Figure 3.8**) result in significantly wider confidence intervals, for both case study buildings, as compared to the ones obtained when spectral acceleration is used as an IM (**Figure 3.9** and **Figure 3.10**). This is because spectral acceleration is a better predictor of nonlinear structural response, as confirmed in several past studies (e.g., Minas et al. 2015 ; Ebrahimian et al. 2015).

Significant difference is also observed when inspecting the fragility curves expressed in terms of $S_a(T_1)$ and $S_a(1.0s)$, (**Figure 3.9** and **Figure 3.10** respectively). Although T_1 equal to 1 second is assumed to be a representative natural period for mid-rise buildings (in insurance/catastrophe-modelling applications), it does not always match closely the natural period of buildings constituting this particular class. This becomes apparent primarily in the case of Special-code case-study (but also Low-code), where the natural period of building substantially deviates from the assumed period level of 1 second (as shown in **Table 3-2**). Thus, the resultant fragility curves expressed in spectral ordinates at fixed period levels may not be representative for some building corresponding to mid-rise class.

3.4   Conclusions

This Chapter introduces in detail the three case study buildings that are used throughout all the applications and methods presented in this Thesis. All buildings are analysed with the aim of standard nonlinear analysis and fragility derivation methods that are commonly used in seismic assessment applications. The outcomes of this analysis are then used for validation of the methods proposed in the following Chapters.

The three buildings, which fall within the mid-rise RC bare frame building typology, correspond to distinct vulnerability classes depending on their design-code level, namely Pre-, Low- and Special-code. First, the buildings are analysed with the aim of nonlinear static PO analysis using different load patterns. A soft-storey at ground floor was identified as the failure mechanism in the case of Pre- and Low-code buildings, as a result of the brittle shear failure of ground level columns. A global mechanism was observed in the case of Special-code building. The product

of PO analysis, i.e. PO curves, is used as an input for capacity spectrum methods, such as FRACAS which is presented in Chapter 4. Next, the nonlinear seismic response of all building is simulated for a large number of unscaled GMs with the aim of NDA. The resultant peak responses are plotted against IM levels associated with each ground motion input to form a scatter of points, the so-called "cloud". Three cloud analysis sets are generated for each building, where IM is expressed in terms of *PGA*, $S_a(T_1)$ and $S_a(1.0s)$. Finally, the abovementioned outputs combined with the building-specific damage scale proposed herein, are used as the input for the construction of fragility functions. The resultant fragility curves show that the choice of IM can have a significant impact in the uncertainty characterization, as the use of *Sa* (either for T=T_1 or T=1.0s) reduces substantially the confidence intervals' size comparing to the associated *PGA* ones.

4. Simplified analytical methods for fragility assessment

4.1 Introduction

This Chapter introduces an alternative simplified analytical method for fragility assessment, FRACAS, and also introduces the new GUI of the FRACAS tool. FRACAS, unlike other CSMs utilises natural (unsmoothed) spectra in order to account for record-to-record variability in the evaluation of performance of a single or population of structures. Ground motion acceleration time-histories are directly used within FRACAS procedure to compute the elastic and inelastic spectra and eventually estimate the PPs, which can then be used for the construction of fragility curves. This characteristic makes the methodology more reliable and efficient comparing to other NSPs, but also computationally more demanding yet faster than NDA. Therefore, FRACAS bridges the gap identified in the literature review between common NSPs and NDA methods, as it strikes a favourable compromise between speed and accuracy.

An in-depth discussion of FRACAS methodology is presented in the first section of Chapter 4 (Section 4.2). Within this Section, the FRACAS procedure is described in steps, which are discussed in detail in the subsections to follow (Sections 4.2.1 to 4.2.6). Next, FRACAS performance is compared to NDA in terms of seismic response and fragility estimation, using the case study buildings introduced in Chapter 3. The effect of record-to-record variability in the resultant fragility curves using FRACAS, is then investigated. Finally, the FRACAS tool -a user-friendly GUI version of FRACAS- is introduced.

The outcomes of this Chapter demonstrated that FRACAS, in comparison to NDA, provides reasonably good estimates of seismic response for the buildings corresponding to the tested typology (i.e. mid-rise RC buildings). This Chapter also showed the ability of FRACAS to capture the inelastic record-to-record variability and properly translate it into the resulting fragility curves. It is noteworthy to mention that the choices of capacity idealisation and

statistical fitting method have a significant impact in the derivation of fragility curves. This Chapter aims to answer the Research Question 1.

4.2   FRACAS methodology

FRACAS is a procedure for fragility curve generation that builds on and improves the modified CSM first developed by Rossetto and Elnashai (2005). FRACAS takes the basic methodology proposed in Rossetto and Elnashai (2005) and, within new software tool builds upon to it, allows more sophisticated capacity curve idealizations, the use of various hysteretic models for the SDoF in the inelastic demand calculation, and the construction of fragility functions through several statistical model fitting techniques. The proposed approach is highly efficient and allows for fragility curves to be derived from the analysis of a specific structure or a population of frames subjected to a number of earthquake records with distinct characteristics. In this way, the method is able to account for the effect of variability in seismic input and structural characteristics on the damage statistics simulated for the building class, and evaluate the associated uncertainty in the fragility prediction.

FRACAS approach presented in this Chapter is a result of collaborative work and is largely based on a journal publication (Rossetto et al. 2016a). The author of this Thesis developed the building models, performed all evaluation tests and scripted additional components of FRACAS code, including the FRACAS GUI.

The FRACAS procedure is based on the following steps (**Figure 4.1**):

1. Mathematical models of a population of buildings are generated by selecting a representative building, termed ''index building", and generating variations of the index building with differing structural or geometrical properties. See D'Ayala et al. (2013a) for recommendations on how to generate the model population to represent a building class. Alternatively, large sets of structures can be generated stochastically based on statistical models of geometric and material properties (e.g., Silva et al., 2014).

2. The computational models of the index building and its variations are analysed with static PO analysis or static adaptive PO analysis (APO, e.g., Elnashai 2001).

3. The PO curve is transformed into a capacity curve in an acceleration displacement response spectrum (ADRS) space, through the use of relative floor displacements and floor masses (see Section 4.2.1).

4. An idealized shape is fit to the capacity curve making various choices regarding the selection of the yielding and ultimate points, the number of segments (bilinear or multilinear) and the presence of increased strength post-yield (e.g., **Figure 4.1a** and **Figure 4.2**).

5. The idealized curve is discretized into a number of analysis points (APs) (**Figure 4.1b**) each representing an inelastic SDoF with the elastic stiffness, ductility and post-elastic properties shown by the capacity curve up to the considered AP.

6. At each AP, the response of the corresponding SDoF under the selected ground motion record is assessed through the Newmark-beta time-integration method. In particular, the elastic response is calculated for APs preceding yield and the inelastic response for those on the inelastic branch of the capacity spectrum (e.g., **Figure 4.1c** and **5.1d**).

7. Using both elastic and inelastic parts of the response spectrum, the PP is estimated by the intersection of the capacity curve and response curve. No iterative process is required.

8. The selected EDP is determined from the PP by re-visiting the results of the PO analysis at the corresponding capacity curve point. MIDR is adopted as the EDP in FRACAS, but others can be determined if required, for example the roof drift (RD). Different IMs associated with the given accelerogram used in the assessment are also calculated and stored.

9. Steps 6–8 are repeated for each capacity curve producing PPs (with associated IM and EDP) at different ground motion intensity levels. This can be done by either scaling up the selected accelerogram(s) to cover a range of intensities (similarly to the IDA procedure) or by using several accelerograms selected to represent different intensities

of ground shaking (similarly to the cloud procedure; e.g., Jalayer and Cornell, 2009). The number of PPs generated equals the product of the number of structural models, number of accelerograms and number of scaling factors used.

10. Fragility curves are constructed from the set of IM and EDP pairs through an appropriate statistical curve fitting approach.
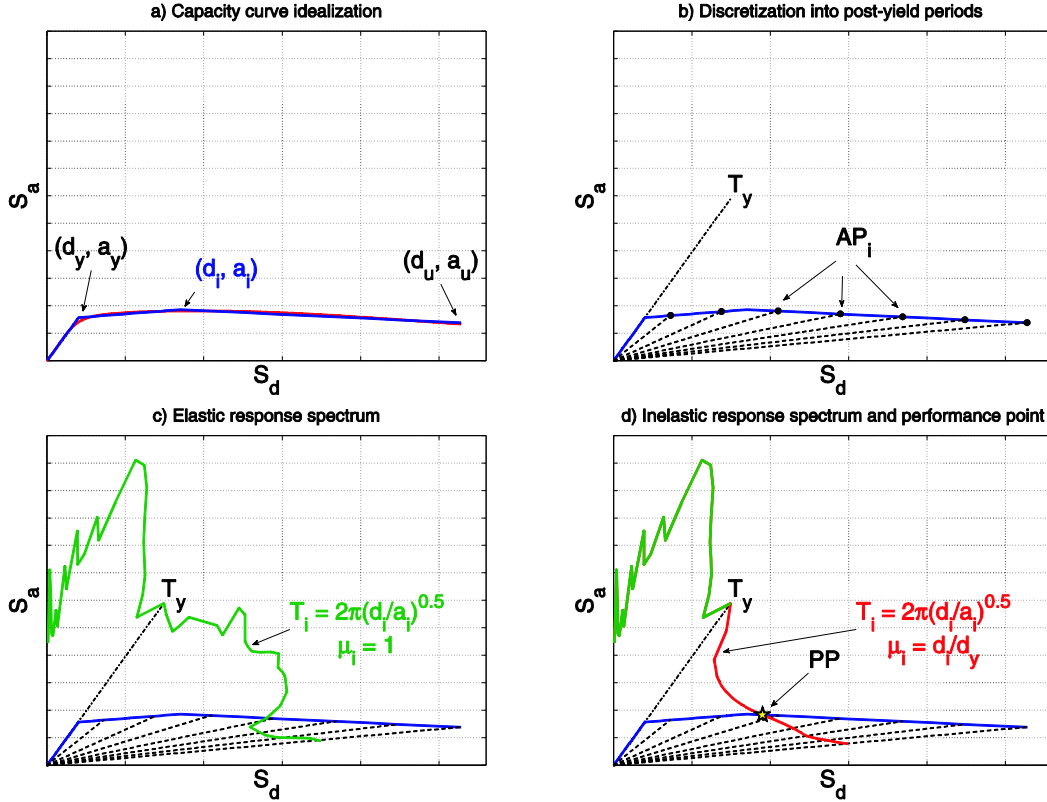
Figure 4.1 - Main steps of FRACAS for the derivation of the performance point (PP) using the trilinear idealization model. (a) shows the fitting of the idealised trilinear curve to the structure capacity curve; (b) shows the identification of Analysis Points (AP), (c) compares the elastic demand spectrum with the capacity curve at the point of intersection of the demand curve with the line representing the yield period of the structure; (d) shows the determination of the Performance Point (PP).Evaluation of FRACAS to NDA.

It is important to note that, in contrast to other CSMs, FRACAS does not rely on reduction factors or indices to estimate the inelastic spectrum from the elastic one. Instead, it carries out, for each AP (with target ductility and period), a simplified dynamic analysis on the idealized nonlinear SDoF model corresponding to the capacity curve. This process proves to be more time-consuming than the commonly-used static approaches but it remains faster than performing full time history analyses on finite element models of full structures. This feature

also has the advantage of permitting the use of various natural, artificial or simulated accelerograms that generate unsmoothed spectra as opposed to standardized design spectra. Therefore, the record-to-record variability can be directly introduced and the resulting cloud of PPs leads to fragility curves that account for the natural variability in the seismic demand. The key steps of the FRACAS approach are explained in more detail below.

4.2.1 Transformation of the pushover curve to ADRS space

In the case of traditional static PO analysis the transformation of the base shear-top (i.e. roof) drift curve (PO curve) to ADRS space is done in FRACAS using the modal participation factors, Γ, and effective modal weight ratios, M^*, determined from the fundamental mode of the structure, using the following equations:

$$\Gamma = \frac{\sum_{j=1}^{N} m_j \phi_j}{\sum_{j=1}^{N} m_j \phi_j^2} \; ; \; S_a = \frac{V_b}{M^*} \; ; \; S_d = \frac{u_N}{\Gamma \phi_N} \; ; \; M^* = \frac{\left(\sum_{j=1}^{N} m_j \phi_j\right)^2}{\sum_{j=1}^{N} m_j \phi_j^2} \qquad (4.1)$$

Where N is the total number of floors, u_N is the top floor displacement, V_b is the base shear force, ϕ_j is the jth floor element of the fundamental mode shape (ϕ_1), m_j is the lumped mass at the jth floor level and M^* is the effective modal mass for the fundamental mode of vibration.

In the case of APO, the transformation must account for the combined effect of multiple response modes. A single transformation is not valid for the adaptive PO curve; this is because the relative contribution of each mode is varying for the different adaptive load increments. Therefore, an approximate method for the transformation is adopted, where the instantaneous displaced shape and storey forces at each increment step of the adaptive PO are used to transform the force displacement curves into ADRS space. The same expressions as for the SDoF transformation are adopted (Equations (4.1)), with the current displaced shape of the structure normalised to the top displacement (ϕ_N) replacing the fundamental mode shape (ϕ_1). Therefore, ϕ_1 is now the component of ϕ_N corresponding to the jth storey. Consequently, u_N

and V_b are the top displacement and base shear at the current load increment, respectively. The reasoning behind this transformation method is that the force distribution and resulting displacement distribution implicitly incorporate the modal combinations. This assumption may not be theoretically justified, but it is observed to provide reasonable assessment results (see Rossetto 2004).

4.2.2   Capacity curve idealization

In FRACAS, the capacity curve obtained from the PO is directly idealized as a multi-linear curve that: (a) is used to represent the capacity curve when it is compared to the demand values in the determination of the PPs and (b) is used to define the inelastic backbone curve of an inelastic SDoF system for the demand calculation explained in Section 4.2.3. Different curve shapes can currently be used to model the capacity curve of the structures: an elastic-perfectly plastic model (EPP), a non-degrading elasticplastic with positive strain-hardening (EPH) or a trilinear model (TL). These are illustrated in **Figure 4.2**. The choice of model depends on the type of structure and shape of the resulting capacity curve, with, for example, EPH being better suited to steel frames without infill and TL to RC frames with infill. Further models will be considered in the future.

Various curve fitting options are possible within FRACAS. In particular, an automated identification of the successive segments of the idealized models is provided, with three different options available for defining the yield point (d_y, a_y) (see **Figure 4.2**):

1. First deviation from the initial stiffness (i.e. evolution of the tangent slope of the capacity curve with respect to the initial gradient– absolute deviation – or with respect to the previous gradient – relative deviation).

2. Intersection of the initial stiffness line with the maximum spectral acceleration of the capacity curve.

3. Coordinates of the nominal value of the capacity curve (i.e. secant stiffness line)

Similarly, three different options are available to define the ultimate point (d_u, a_u) (see **Figure 4.2**):

1. The spectral displacement corresponding to the collapse drift (i.e. the last limit state in the considered damage scale).

2. The spectral displacement corresponding to the last point of the capacity curve.

3. The spectral displacement corresponding to a 20% drop of the spectral acceleration with respect to the maximum capacity.

Finally, an equal-energy criterion may also be used, where the spectral ordinate of the yield point (or the intermediate point in the case of TL) is adjusted in order to obtain the same areas under the idealized and actual capacity curves.
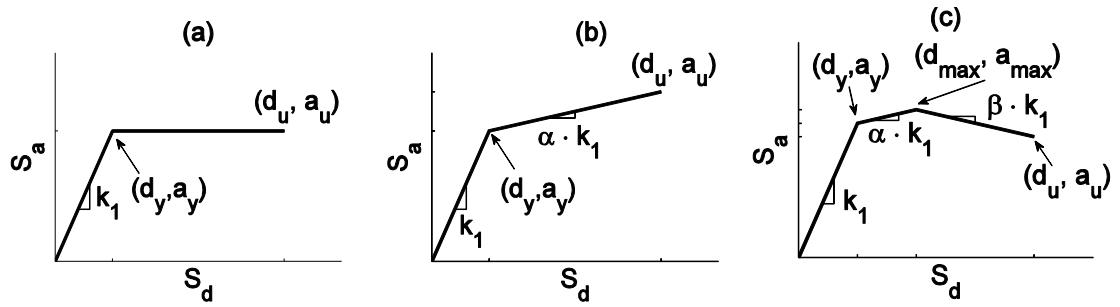
Figure 4.2 - Capacity curve idealization models used in FRACAS: a) elastic-perfectly plastic model (EPP), b) elastic-plastic with positive strain-hardening (EPH) and c) trilinear model (TL).

In addition, the manual selection of the global yield point, ultimate point and, in the case of TL, the point of the second change in slope, is also allowed. The aim in fitting the model to the capacity curve is to select the parameters of the models such that they reproduce the capacity curve as closely as possible. The different modelling options that are offered to the user are useful in emphasizing various aspects of the studied structure: one can decide whether the idealized curve must closely fit the elastic period, the maximum strength, the yield deformation or the energy dissipation capacity of the structural system. It is noted that De Luca et al. (2013) have found that using the equivalent-energy criterion may lead to large biases in the prediction of the structural response, especially when the curve fitting induces a significant change in the initial stiffness.

4.2.3 Discretization of the idealized capacity curve and definition of a suite of SDoF systems for inelastic demand analysis

In order to improve the efficiency of the PP calculation, FRACAS discretizes the capacity curve into a number of pre- and post-yield periods, which are used APs. In FRACAS the number of APs is user-defined, but it is recommended that a minimum of 5 points pre-yield and 25 points post-yield (evenly distributed along the post-yield branches of the idealized curve) be used. In addition, points defining changes of slope in the idealized curve (e.g., the yield point) should always be adopted as APs. Each AP is characterized by its spectral coordinates (i.e. d_i and a_i in **Figure 4.1**), and a ductility value, defined by the spectral displacement of the analysis point (d_i) divided by that of the global yield of the structure (d_y). Together with the elastic period of the idealized curve, this ductility value is used to define a SDoF system from which the inelastic demand is calculated. The inelastic backbone curve of the SDoF system is also defined by the shape of the idealized curve up to the analysis point.

4.2.4 Inelastic demand calculation

For a given earthquake record (which could be scaled to a certain IM level), the inelastic seismic demand corresponding to each AP is calculated through the analysis of the SDoF system associated with that AP (see above). The earthquake record used in the analysis is discretized into time increments smaller than (1/50)th of the smallest vibration period of interest to ensure stability of the Newmark-Beta time integration (i.e. $\Delta t/T \leq 0.55$; e.g., Clough and Penzien 1993). The acceleration record is applied in these time steps to the SDoF system and the Newmark-beta time-integration method for linear acceleration is used to solve the dynamic nonlinear equilibrium equation for the evaluation of the SDoF response. The successive loading cycles follow a plain hysteretic curve with parallel unloading and reloading paths whose slope is the original elastic behaviour of the structure, which do not currently account for pinching or for degradation of unloading stiffness. More advanced models will be considered in the future. The maximum response from the entire record defines the spectral displacements and accelerations

used to characterize the demand at the AP. It is noted that the inelastic dynamic analysis only needs to be carried out on a SDoF system under the applied accelerogram at each AP, increasing the rapidity of the assessment.

4.2.5   Determination of the PP and EDPs

In ADRS space the AP on the capacity curve and the inelastic demand calculated for the matching inelastic SDoF (with elastic period and ductility determined by the idealized capacity curve to the AP) lie on a diagonal that passes through the origin and the AP. Although not used in the analysis, this diagonal theoretically represents the effective period of an equivalent linear SDoF. The inelastic demand and capacity curve can be directly compared along this diagonal, as they have the same ductility. If these points match, then the PP is reached. Exact matching is difficult to achieve from the predefined APs, which are spaced at subjective intervals along the capacity curve. Hence, it can be beneficial to draw a "response curve" by joining together the inelastic demand values of S_a and S_d calculated at each analysis point, as the PP can be efficiently determined from the intersection of the capacity curve with this demand curve (see **Figure 4.1d**).

In order to determine the EDPs corresponding to each PP, the capacity curve coordinates at the PP are used to determine the corresponding load step of the nonlinear static analysis file, and relevant response parameters (e.g., MIDR) are read from this file. Damage thresholds of EDP can be determined from an appropriately selected damage scale for the structure being analysed (as discussed in Section 3.3).

4.2.6   Fragility curve construction

With regard to fragility curve generation, the capacity spectrum assessment is repeated for each structural model subjected to ground motions of increasing intensity, either by scaling each earthquake record or adopting a range of earthquake records with increasing intensity. A statistical curve fitting method is then adopted to fit a fragility curve shape from the IM - EDP

cloud generated. Within FRACAS either a LS approach or a GLM can be used for the curve fitting, and confidence bounds derived using a bootstrap analysis of the data points. The former approach (LS) is more commonly used in the fragility literature but the method assumptions may be violated by the data, as shown by Rossetto et al. (2014).

4.3   Comparison of FRACAS with NDA

NDA provides a benchmark against which to test the performance of simplified capacity spectrum approaches like FRACAS. Hence, in this Section the differences in EDP estimates obtained using FRACAS and NDA are investigated over a wide range of IM values. To this aim, two RC MRF buildings corresponding to Pre- and Special-code vulnerability class are used here as case-studies. A detailed description and the modelling assumptions of the two case study buildings are presented in Chapter 3. The resulting differences in derived fragility curves are also assessed.

The two case study buildings are analysed with the aim of uniform and triangular PO analysis, which are then adopted as the capacity input for FRACAS analysis. Although APO approaches are generally perceived to provide better estimates of structure response than conventional static PO, particularly when higher modes and structural softening are important (as shown in many previous studies, such as Antoniou and Pinho, 2004), it is decided not to adopt APO in the current comparison study. Inclusion of APO in FRACAS is computationally very expensive when dealing with a large number of unscaled accelerograms (as in the current study), as an APO needs to be developed for each accelerogram used. The static PO curves for the two case-study buildings (uniform and triangular lateral load distributions for Pre- and Special-code buildings respectively) are presented in **Figure 3.2** of Chapter 3.

The same set of 150 unscaled ground motion records from the SIMBAD database also used to carry out the NDA analysis in Chapter 3, is adopted here as a hazard input for FRACAS. A comparison of the MIDR values obtained from NDA for the two frames with those estimated by FRACAS using the two PO analyses as capacity input are shown in **Figure 4.3**.
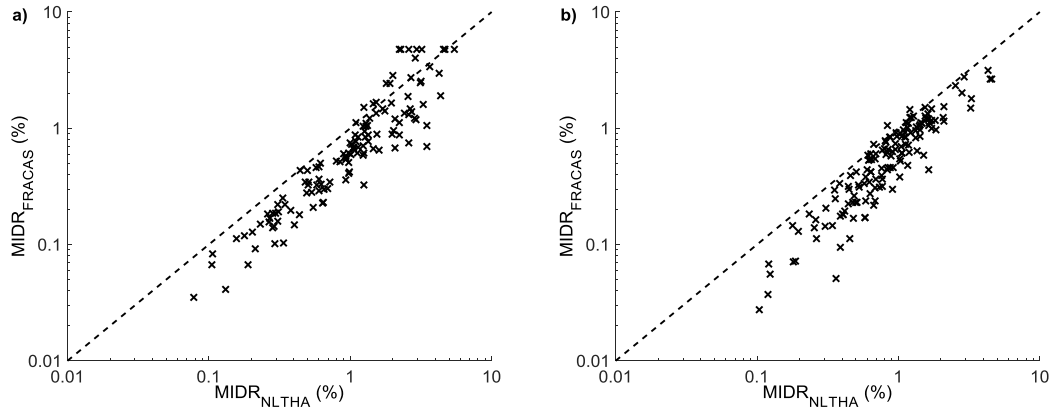
Figure 4.3 - Comparison of FRACAS with NDA in terms of MIDR: a) Pre-code building, TRI-PO, and TL and (b) Special-code building, TRI-PO, and EPP model.

Results show that for both case study structures FRACAS provides a reasonable estimate of the MIDR values predicted by NDA across the 150 ground motion records (average error is around 25% across the considered models), particularly when nonlinear structural response is considered (average error is around 15% across the considered models). FRACAS generally tends to under-predict the MIDR values across the various IM levels and this is expected given the non-inclusion in the simplified method of (1) effect of higher modes (even in the elastic range of response) and (2) hysteresis models incorporating cyclic strength and stiffness deterioration. A non-negligible bias is observed for lower EDP due to the idealization of the capacity curve. In fact, the elastic branch of the idealized capacity model is obtained by directly connecting the origin to the first yield point, thus resulting in an initial stiffness (i.e. fundamental period) that is different from the one found for the actual structure. Therefore the structural response in the elastic range may also be strongly influenced by the chosen idealization strategy.

There are indications in the capacity spectrum assessment literature that the choice of capacity curve idealization affects the resulting PPs (e.g., Rossetto and Elnashai 2005). This is also observed by the authors. However, as discussed previously, one of the features of FRACAS is the ability to adopt different models for the capacity curve idealization. For instance, both TL and EPP idealizations were trialled to represent the response of the Pre-code structure, which displays a (monotonic) degrading response curve after its maximum capacity. It is observed that

the TL idealization results in a better approximation of the MIDR predicted by NDA than the EPP idealization, particularly near collapse. Furthermore, in carrying out this assessment a high sensitivity to the selected shape of the TL curve was observed. Despite this, the choice of capacity curve idealization does not significantly affect the resulting FRACAS fragility curves, as seen in **Figure 4.4**, particularly when the GLM approach is employed in the statistical fitting. A greater effect may be observed for structures with infills where the PO curve presents different successive phases due to the failure of infill panels and where the EPP idealization provides a very poor fit to the PO curve. This will be a subject of a future work.
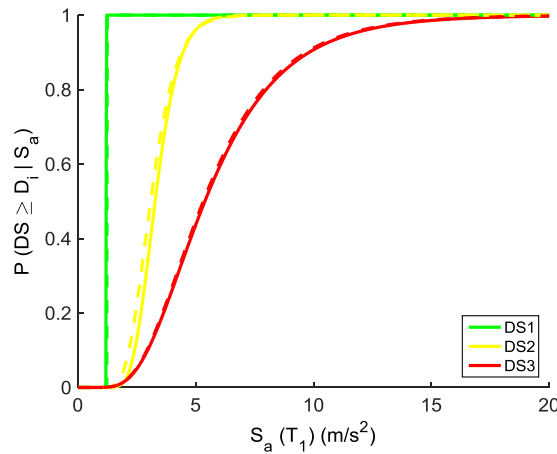
Figure 4.4 - Fragility functions derived by FRACAS using TL (solid lines) and EPP (dashed lines) for Pre-code building and TRI-PO.

On the other hand, it is highlighted here that the choice of statistical model fitting technique may significantly influence the shape of the resulting fragility function. In this paper, fragility curves are fit to the analysis data using the GLM approach with a Probit link function (see Rossetto et al. 2014). As mentioned in Section 4.2.6 the GLM model is theoretically more valid than LS, which assumptions are violated by the data used for the fragility assessment (Rossetto et al. 2014b). As an example, **Figure 4.5** shows the fragility functions obtained using GLM and LS for the Special-code building assessed for the 150 unscaled records with FRACAS. It can be observed that the GLM and LS approaches result in large discrepancy between the fragility curves derived for DS2, with the GLM approach showing a greater variability in the results. The GLM better captures the fact that there are only a limited number of data points (observations)

available for the higher DSs. It is noted, that despite the intensity of the chosen records being significant, too few data are available to derive the collapse DS curve for the Special-code building (without scaling the records). Hence, this DS curve is not presented in the following carried out comparisons.
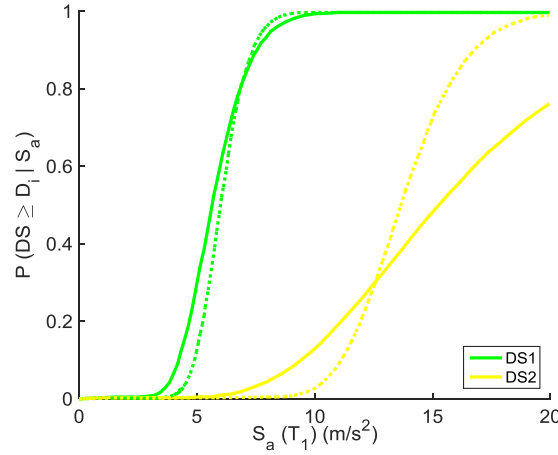
Figure 4.5 - Fragility functions derived by FRACAS using the GLM (sold lines) and the LS fitting (dashed lines) for the Special-code building, TRI-PO, and EPP model.

Comparisons between fragility functions developed using NDA as compared to simplified assessment methods is not often presented in the literature, and so it is discussed here. As expected, the under-prediction in EDPs by FRACAS is seen to translate directly into a lower fragility prediction compared to NDA (see **Figure 4.6**). Despite this, it is observed that the two assessment methods provide similar fragility functions for all DSs of the Pre-code building. However, larger discrepancies are observed for the Special-code building, especially for the second DS (D2). These discrepancies are actually observed to arise from the sensitivity of the fragility curve fitting (in both approaches) to small numbers of observations rather than from the ability of FRACAS to simulate the NDA response of the building. For example, only 8 out of 150 records result in damage DS2 in the Special-code building when NDA is employed, compared to only 5 in the case of FRACAS. These number further decrease when looking at damage DS3. The development of general guidelines as to how many analyses are required to create stable fragility functions, particularly when cloud-type approaches are used, is a subject of active research by several authors (e.g., Gehl et al. 2015) and not investigated here, however

the importance of considering this is highlighted by this example. Overall, FRACAS is observed to predict well the EDP response observed in NDA to failure in the case of the Pre-code building, and hence only this structure is adopted in the following Sections of this Chapter where a study of the ability of FRACAS to capture the effect of record-to-record variability on fragility functions is presented.
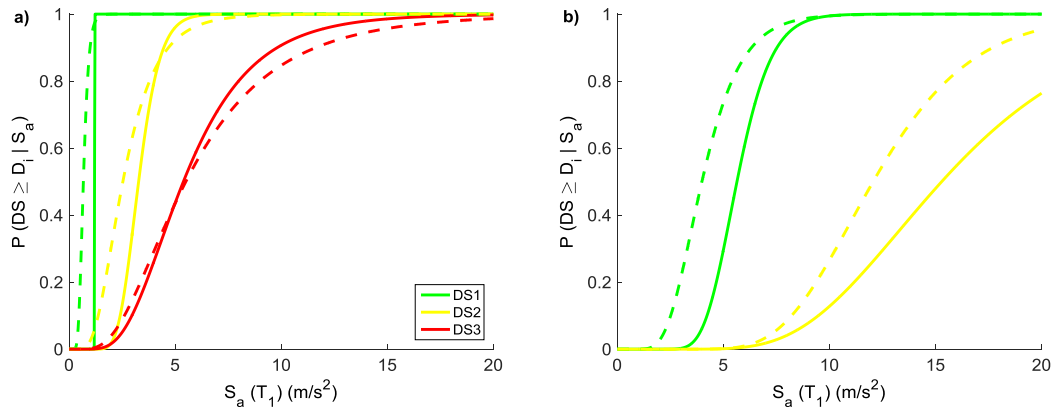
Figure 4.6 - Fragility functions derived by FRACAS (solid line) and NDA (dashed line): a) Pre-code building, TRI-PO, and TL and b) Special-code building, TRI-PO, and EPP model.

4.4 Investigation of the effect of record-to-record variability on fragility curves

To investigate the effect of record-to-record variability on fragility curves, the Pre-code building presented above is assessed under different suites of appropriately selected accelerograms using FRACAS. As the aim is to show how FRACAS captures the response spectral variability and translates it into fragility curves, natural accelerograms (i.e. recorded during past earthquake events) are selected and modified records are generated, based on the match of their spectra to a target spectrum. Although the current best practice does not require the compatibility or the matching with a given target spectrum (contrary to code-based procedures for single structures; e.g., NIST 2011) neither does it recommend the use of a single hazard level (i.e. corresponding to just one return period of the seismic hazard) for fragility analysis (e.g., Lin and Baker 2013), the approach followed here is deemed appropriate to investigate whether FRACAS is able to capture record-to-record variability.

124

4.4.1 Earthquake spectrum and input accelerograms for the structural assessment

The base spectrum chosen to carry out the structural assessment in FRACAS is the Type 1 EC8 (EN 1998-1 2004) spectrum for soil class B (stiff soil), with a PGA of 0.17 g for a 475-year return period (i.e. 10% probability of exceedance in 50 years). This PGA is taken directly from the detailed PSHA of Stucchi et al. (2011) for Italy and corresponds to a site located in Naples, Southern Italy, representative of moderate-to-high seismicity regions. Sets of accelerograms are chosen such that their mean spectrum matches the base spectrum over structural response periods 0.05–2 s with a lower limit tolerance of 10%. As EC8 does not provide any restrictions on the higher limit tolerance of the selected records, a maximum higher limit tolerance of 30% is arbitrarily selected (e.g., Iervolino et al. 2009). EC8 (Section 3.2.3.1.2) does instead provide guidance on the relevant range of structural periods over which to carry out the matching, specifying this range in terms of the structural fundamental period of vibration (T_1) as 0.2 T_1 to 2 T_1. In the case of the considered case-study structure (Pre-code building), the latter period range lies well within the adopted period range for matching. In the case of the natural records, the software REXEL (Iervolino et al. 2009) is used to select unscaled accelerograms from the three databases included in the software, namely the SIMBAD (Smerzini et al. 2013), the European Strong-motion Database (ESD) (Ambraseys et al. 2004) and the Italian Accelerometric Archive (ITACA) (Pacor et al. 2011). It is worth noting that the EC8 approach for ground motion selection is not sufficiently conservative for the derivation of fragility functions and this guideline has not been developed for this purpose. However, as discussed above, the main aim of the simple exercise carried out in this Section is to demonstrate FRACAS' ability to capture record-to-record variability.

In particular, three suites of input accelerograms are adopted to assess the effect of record-to-record variability on the fragility curve produced for the Pre-code building:

1. Twenty natural accelerograms (unscaled) to be compatible with the target spectrum over the period range 0.05–2 s (thereafter referred to as REAL-EC8:475).

2. The same accelerograms as in REAL-EC8:475 adjusted using wavelets so that their spectra better match the target (MATCHED-EC8:475). The program SeismoMatch (Seismosoft 2013) was adopted to adjust earthquake accelerograms to match a specific target response spectrum, using the wavelets algorithm proposed by Hancock et al. (2006).

3. Twenty accelerograms generated using SIMQKE (Gasparini and Vanmarcke 1976), a method (and associated software) that can generate response spectrum-compatible statistically-independent synthetic motions showing very little dispersion in their spectra and matching the target closely (SIMQKE-EC8:475).

Table 4-1 presents basic information on the ground motions used within the first two record sets. Within FRACAS these sets of accelerograms are scaled several times for the capacity spectrum assessment at increasing IM levels.

The elastic spectra for the three suites of accelerograms are shown in **Figure 4.7**a–c together with their average, and the target spectrum. Stafford and Bommer (2010) postulate that, when deriving fragility curves accounting for ground motion variability, the peak-to-trough variability in their response spectra should not be too small. They define the peak-to-trough variability as the standard deviation of the natural logarithm of spectral ordinates over a number of records and a range of response periods defined by the 'bandwidth', i.e. the range of periods surrounding a central period. This bandwidth roughly corresponds to the degree of structural nonlinearity that is expected and the contribution of higher mode effects. With the CSM used here the structure is not affected by spectral ordinates with periods shorter than the natural period (equal to 0.902 s for the Pre-code structure). Therefore, we modify the definition of the peak-to-trough variability to account for only those periods longer than the natural period. The peak-to- trough variability for each suite of accelerograms is plotted in **Figure 4.7**d as function of an elongated period (equal to $k \cdot T_1$) rather than bandwidth.

Table 4-1 - Summary of record data returned by REXEL for the REAL-EC8:475 set.

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class	Database
147y	Friuli (aftershock)	9/15/1976	6	thrust	14	B	ESD
198x,y	Montenegro	4/15/1979	6.9	thrust	21	A	ESD
333x,y	Alkion	2/24/1981	6.6	normal	20	C	ESD
879y	Dinar	10/1/1995	6.4	normal	8	C	ESD
1726y	Adana	6/27/1998	6.3	strike slip	30	C	ESD
103y	Friuli (aftershock)	9/15/1976	5.9	thrust	16	A	ITACA
171y	Irpinia	11/23/1980	6.9	Normal	19	B	ITACA
381x	Umbria-Marche (aftershock)	9/26/1997	6	Normal	6	D	ITACA
22x	W Tottori Prefecture	10/6/2000	6.6	strike-slip	19	B	SIMBAD
146x	S Suruga Bay	8/10/2009	6.2	reverse	25	B	SIMBAD
411y	Hyogo - Ken Nanbu	1/16/1995	6.9	strike-slip	17	C	SIMBAD
437x	Parkfield	9/28/2004	6	strike-slip	10	B	SIMBAD
438x,y	Parkfield	9/29/2004	6	strike-slip	15	B	SIMBAD
443x	Imperial Valley	10/15/1979	6.5	strike-slip	25	B	SIMBAD
449x	Superstition Hills	11/24/1987	6.6	strike-slip	20	C	SIMBAD
458y	Northridge	1/17/1994	6.7	reverse	11	C	SIMBAD
459y	Northridge	1/18/1994	6.7	reverse	20	C	SIMBAD

Figure 4.7 shows that, as expected, the MATCHED and SIMQKE records show similar but considerably less spectral variability than the natural accelerograms. The variability shows a similar behaviour to that shown by Stafford and Bommer (2010) but with higher absolute values.

4.4.2   Fragility assessment

For the construction of fragility functions the three suites of accelerograms presented in Section 4.4.1 are used to define the demand spectra and are combined with the capacity curves obtained from the PO analysis of the structure. FRACAS analyses are carried out on the Pre-code building with each of the ground motion sets, scaled to varying spectral accelerations (Sa). Fragility curves are derived for each set of accelerograms (300 data points each, i.e. 20 original records scaled 15 times) using a GLM statistical curve fitting approach. The resulting curves follow a cumulative lognormal distribution with mean θ and standard deviation β as the fragility parameters (i.e. the parameters of the associated normal distribution).

Figure 4.8 shows that the fragility curves derived for MATCHED- EC8:475 and SIMQKE-EC8:475 differ in both median and standard deviation from the fragility curves derived for the REAL-EC8:475 set, especially at the higher DSs. The initial record-to-record variability is efficiently translated to the final fragility curves, as shown by the different values of the standard deviation β. The use of scaled accelerograms allows the computation of the standard deviation of the MIDR within each bin of IM (i.e. set of 20 ground motions scaled according to $S_a(T_1)$), as shown in **Figure 4.9**. It is noted that the variability in the structural response increases with the imposed intensity level: this observation emphasizes the role of the specific nonlinear computations that are performed in FRACAS during the estimation of the inelastic response spectrum. Once the yield limit is reached, the relation between the IM and the structural response shows high heteroscedasticity. The relative variability between the three sets of accelerograms also follows a similar trend to the peak-to-trough variability. It is highlighted that the dispersion in the structural response tends to stabilize or even decrease for higher

intensities due to a peculiarity of FRACAS, which considers PPs exceeding the ultimate point as ''collapse'' events, and thus sets their value to the last point of the curve. The fragility curve parameters for all the functions are presented in **Table D 1** and **Table D 2** in Appendix D.
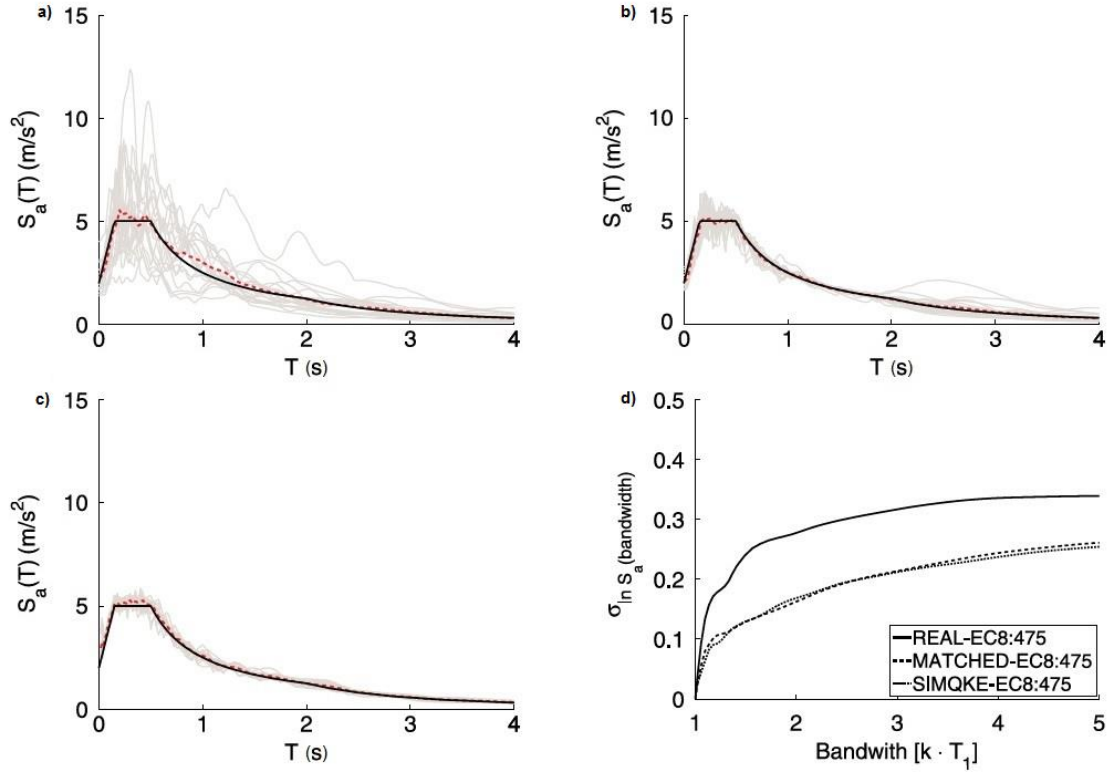
Figure 4.7 - a)-c) Elastic response spectra of the three sets of accelerograms (in grey), corresponding average (in red) and target spectrum used for the selection (in black): a) REAL-EC8:475, b) MATCHED-EC8:475, c) SIMQKE-EC8:475; and d) peak-to-trough variability of accelerograms.
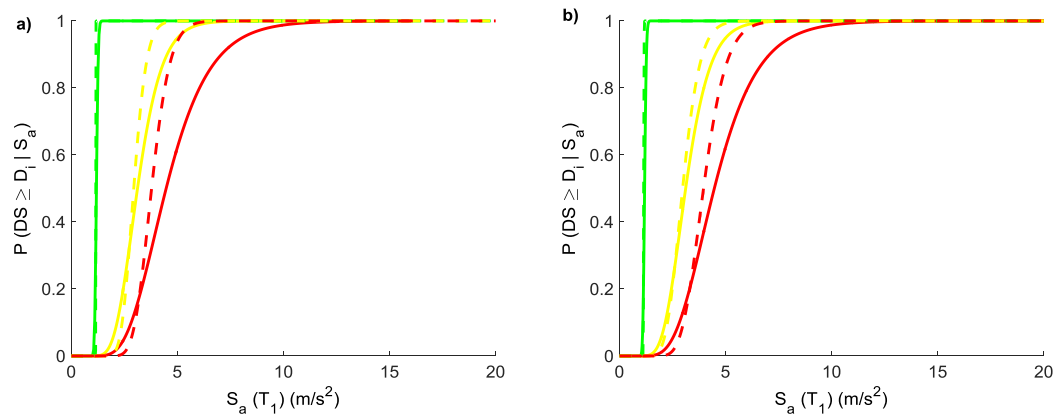
Figure 4.8 - Fragility functions by FRACAS using the three sets of accelerograms for the Pre-code building, TRI-PO, and TL: a) REAL-EC8:475 (solid lines) versus SIMQKE-EC8:475 (dashed lines); b) REAL-EC8:475 (solid lines) versus MATCHED-EC8:475 (dashed lines).
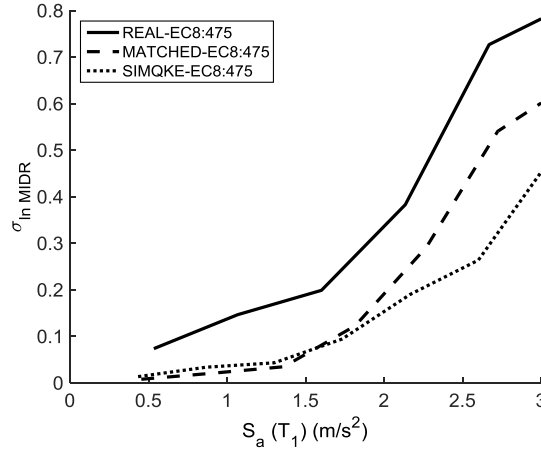
Figure 4.9 - Standard deviation of ln MIDR as a function of spectral acceleration for the Pre-code building, TRI-PO, and TL.

The reason for the different medians (θ) of the MATCHED-EC8:475 and SIMQKE-EC8:475, compared to the REAL-EC8:475 fragility curves, particularly for DS3, could be related to the bias introduced by spectral matching recently evidenced by Seifried and Baker (2014). These authors studied the reason for the observed un-conservative bias in fragility curves when closely matched spectra are used (i.e. the effect that is observed when comparing REAL-EC8:475 and MATCHED-EC8:475). They show that this bias is solely due to the loss of extreme spectral ordinates (i.e. peaks) in closely matched spectra, which are usually responsible for large deformations in the structural system: the nonlinear relation between the IM and EDP value (i.e. with higher IMs leading to a much larger EDPs variation than proportionally smaller IMs) coupled with the spectrum variability, therefore, explains the loss of higher EDP values when using closely matched spectra. This observation raises the question as to whether the natural record-to-record variability, which is originally not present for matched records, should be added back in during the final steps.

4.5 FRACAS tool

FRACAS has been developed into a Matlab-based automated tool, which is freely available at https://www.ucl.ac.uk/epicentre/resources/software or from the authors. To further simplify the use of FRACAS methodology, a GUI is developed, as shown in **Figure 4.10**. The FRACAS

GUI is divided in three main parts: (1) the "Model Preparation" panel, (2) the visualisation of the "Capacity Curve Idealization" panel and (3) the FRACAS analysis button. In addition to the above, a Help section is also included in the GUI providing useful links (to demonstration videos and other information) and access to the FRACAS manual.

Regarding operation, FRACAS tool does not provide access to the main code and all the necessary analyses are performed externally. The FRACAS tool consists of three individual modules that function together to perform FRACAS analysis, namely (1) Capacity module, (2) Hazard module and (3) Fragility module.

Before performing FRACAS analysis the required capacity and demand input text files (.txt) need to be input using the right format in the FRACAS application folder (".../EPICentre/FRACAS/application" folder). The capacity input file (or PO input file), provides the base shears values and the associated displacements at every floor as obtained at each incremental step of the PO analysis. The demand input file(s) provide(s) the earthquake record(s) in the acceleration time-history format.
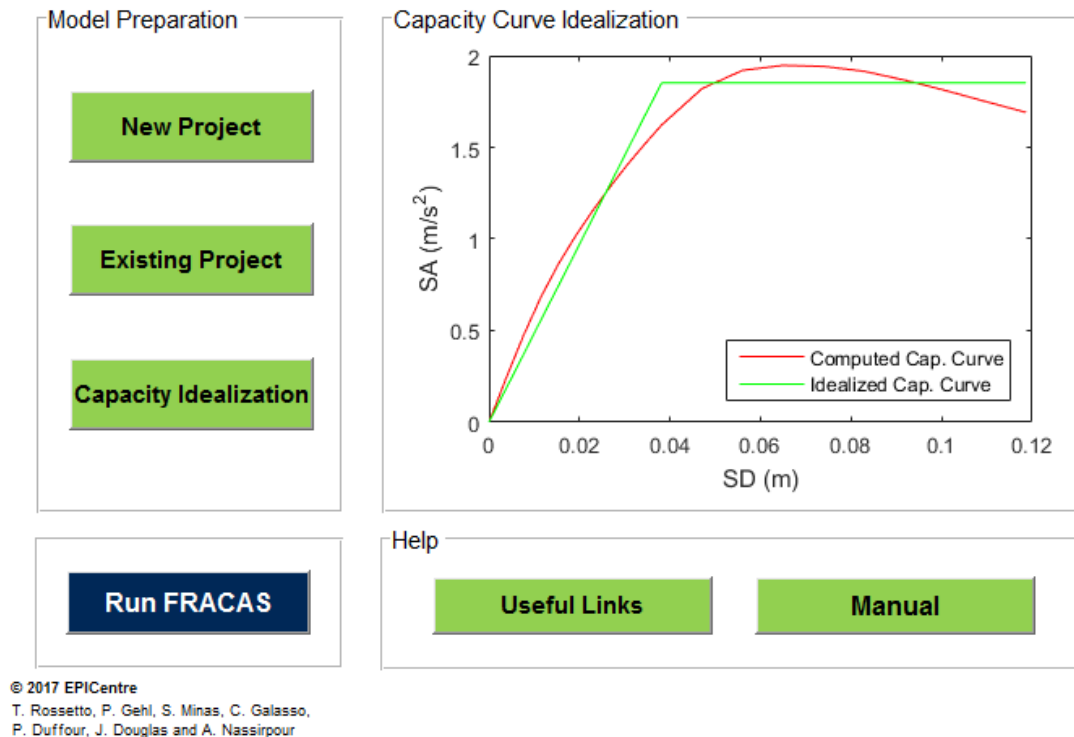
Figure 4.10 - Screenshot of FRACAS tool GUI.

For the model preparation, the user can choose between a new or an existing past project in order to call a controller spreadsheet to define the input options for the FRACAS analysis. The input controller spreadsheet is divided into three tabs, where each tab controls the inputs associated with each of the FRACAS modules.

In the capacity tab, the user can specify the building details (including the label of the capacity input file, the storey masses and inter-storey heights), the damage limit states options and the capacity idealization models. In the demand tab, the user can enter the label(s) of the demand input file(s), define whether record scaling is used and specify the scaling options.

Regarding the fragility tab, the user can set if the generation of the fragility curves is required or not, and consequently select the fragility derivation method and link function, and activate or deactivate the confidence bounds.

The user can confirm the input options by simply saving and closing the input controller spreadsheet, and can then return to the FRACAS tool GUI. A capacity curve idealization check can be performed by clicking on the "Capacity Idealization" button in "Model preparation" panel (**Figure 4.10**). If the user is not satisfied with the capacity curve idealization plot (as shown in "Capacity Curve Idealization" panel, **Figure 4.10**), they can refer back to controller spreadsheet by clicking "Existing Project" and change the capacity idealization options in the capacity tab.

To proceed to the FRACAS analysis, one can click on "Run FRACAS" button. The output of each run is saved in a separate folder labelled based on time and date format within the output folder (".../EPICentre/FRACAS/application/output" folder). Each folder contains a series of figures and data files, including the following:

- Fragility curves figure

- Engineering demand parameter versus intensity measure graph figure

- Performance point calculation for each capacity analysis figure

- Performance points in ADRS space figure

- Performance points details data file.

A flowchart describing the FRACAS tool steps is presented in **Figure 4.11**. For more details regarding the FRACAS software tool one may refer to the FRACAS manual (Rossetto et al. 2016b).
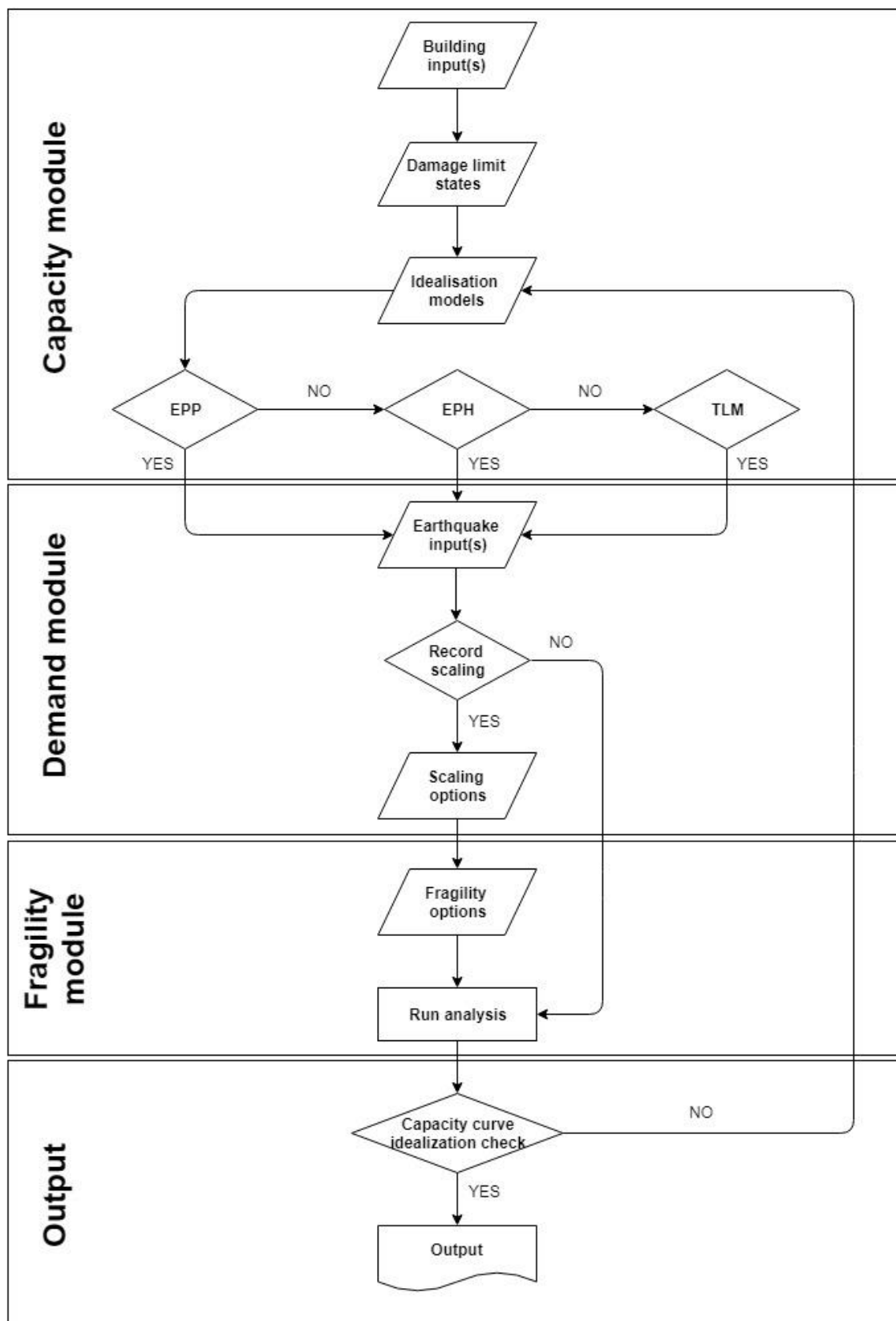
Figure 4.11 - FRACAS analysis steps flowchart.

4.6   Conclusions

This Chapter presents a new simplified approach for the derivation of fragility curves, named FRACAS, and also introduces the new GUI of the FRACAS tool. FRACAS adapts the capacity spectrum assessment method and uses inelastic response spectra derived from earthquake accelerograms to construct fragility curves. A comparison of the predicted MIDR response obtained by FRACAS and NDA is carried out for two case-study 4-storey RC frames assessed under 150 accelerograms. FRACAS is seen to represent well the response of both case study structures when compared to NDA. The case study application also highlights the sensitivity of the FRACAS EDP predictions to the adopted capacity curve idealization, but shows an insensitivity of the derived fragility function to the idealization model choice as long as the idealization model provides a reasonable fit to the real capacity curve. The statistical model used to fit the fragility function is seen to have a significant influence on the resulting curves, and the Chapter highlights the sensitivity of the fragility function to the number of analysis runs.

This Chapter also shows how FRACAS is able to capture the inelastic record-to-record variability and properly translate it into the resulting fragility curves. In particular, through an example application, it is shown that the variability in spectral ordinates for periods beyond the natural period of the undamaged structure is directly correlated to the standard deviations of the fragility curves. A variant of the peak-to-trough measure of the variability in the input spectra (accounting only for periods longer than the natural period) is proposed and is seen to provide a useful measure of this variability. Consequently, it is concluded that differences between fragility curves derived using static PO approaches can be partially explained by differences in the input spectra, even if the mean target spectra are similar.

Overall, it is demonstrated that FRACAS is able to represent the effects of record-to-record variability in fragility curves, and has the advantage of simplicity and rapidity over other methods that use accelerograms directly.

5. Selection of optimal intensity measures for simplified fragility analysis of mid-rise RC buildings

5.1   Introduction

The selection of IM to characterize seismic hazard is a key aspect within the fragility component of the seismic risk framework. Several recent research studies have attempted to determine an appropriate IM to be used in performance-based assessment of existing buildings (a comprehensive overview of the existing literature can be found in Section 2.5); however, it is worth noting that all the aforementioned studies focused on fragility assessment based on NDA or NDA-based approaches (e.g., IDA), while no studies have been conducted based on NSPs. Hence this Chapter looks to answer Research Question 2, by identifying which IM is best suited to the representation of structural response for RC MRF buildings in fragility analyses when NSP are used.

One common, well-acknowledged, limitation of NSPs is their inability of capturing ground motion spectral shape (e.g., Silva et al.(under review)), which has been shown to have a significant influence on the estimated seismic response and resulting structural fragility. However, this limitation does not apply to the simplified approach FRACAS, which was introduced in Chapter 4, as it utilizes spectra from actual (real or simulated) ground motion records to perform the structural assessment. To this aim, the effect of implementing different combinations of IMs and EDPs in simplified fragility analysis by FRACAS is explored herein. In addition, only in recent years, advanced IMs (i.e. spectral-shape based) and innovative selection criteria (e.g., the relative sufficiency measure based on the information theory (Jalayer et al. 2012)) have been introduced and tested in separate studies (and for different structural types). The present Chapter convolves all these state-of-the-art concepts and aims to identify optimal IMs to be used in probabilistic seismic demand analysis through NSPs. Specifically,

this Chapter aims to shed light in comparing different IM/EDP combinations for the fragility analysis of mid-rise RC buildings by using FRACAS, using the building models described in Chapter 3.

This Chapter begins with an introduction of the six IMs considered for this research work, including four conventional and two advanced IMs. Next, the methodology steps followed in study are summarized. Specifically, Section 5.4 describes the statistical regression techniques implemented to determine the IM that better predicts each of the considered EDPs. Section 5.5, outlines the most commonly used IM selection criteria, with particular focus on the ones utilized in this Chapter, namely efficiency, sufficiency/relative sufficiency and hazard computability. Finally, the fragility functions generated from two different IMs are compared.

In agreement with previous studies performed with NDA, the results of this assessment show that advanced IMs perform best in all criteria (including hazard computability criterion) and yield improved fragility predictions compared to conventional IMs.

5.2 Considered intensity measures

To identify the ground motion features that best influence the nonlinear response of the structures of interest, several types of IMs are tested. Conventional IMs (**Table 5-1**), including the peak ground responses (PGA, PGV and PGD) and spectral acceleration at the first period, $S_a(T_1)$, for 5% damping, are the most commonly used IMs. In general, PGA and $S_a(T_1)$ poorly predict the structural response of mid- to high-rise MRFs, although the latter IM can sufficiently capture the elastic behaviour of first-mode dominated MDoF systems, especially in the case of low-to-moderate fundamental periods (Shome et al. 1998). However, the behaviour of highly nonlinear structures or structures dominated by higher-mode periods (less than T_1) is not very well captured by utilizing $S_a(T_1)$ due to the lack of information on the spectral-shape provided by this IM. Therefore, it is essential to implement advanced IMs (**Table 5-2**) that can account for the elongated periods and/or consider nonlinear demand dependent structural

parameters (e.g., Mehanny 2009). Kazantzi and Vamvatsikos (2015) and Kohrangi et al. (2016) amongst others have investigated the adequacy of numerous advanced scalar IMs measures that take into consideration the aforementioned parameters.

The first advanced scalar IM considered here is S_a^c (proposed by Cordova et al. (2001)), which utilizes spectral-shape information (period elongation), and is expressed as:

$$S_a^c = S_a(T_1) \left[\frac{S_a(cT_1)}{S_a(T_1)} \right]^\alpha \tag{5.1}$$

where c and α are coefficients assumed to be $c = 2$ and $\alpha = 0.5$ respectively, based on the calibration carried out by the authors in the original study.

Bojórquez and Iervolino (2011) also proposed the advanced scalar IM, I_{N_p}, which is based on $S_a(T_1)$ and the parameter N_p, defined as:

$$I_{N_p} = S_a(T_1) N_p^\alpha \tag{5.2}$$

where α parameter is assumed to be $\alpha = 0.4$ based on the tests conducted by the authors and N_p is defined as:

$$N_p = \frac{S_{a,avg}(T_1,...,T_N)}{S_a(T_1)} = \frac{\left[\prod_i^N S_a(T_i) \right]^{1/N}}{S_a(T_1)} \tag{5.3}$$

T_N corresponds to the maximum period of interest and lays within a range of 2 and $2.5T_1$, as suggested by the authors. In this Chapter T_1 is assumed to be equal to 1 second for all the considered buildings, as this is the representative natural period for mid-rise building class in catastrophe modelling, T_N is assumed to be equal to $2T_1$ (i.e. 2 seconds), and only one intermediate period is considered (i.e. 1.5 seconds). Ten different values, from 0.1 to 1, for the

α- parameter are considered here in order to identify the optimal value for α, as discussed in the following Sections.

Table 5-1 - Conventional scalar intensity measures.

Intensity Measure	Name	Reference Study
$PGA = \max\left(\left\|a(t)\right\|\right)$	Peak ground acceleration	-
$PGV = \max\left(\left\|v(t)\right\|\right)$	Peak ground velocity	-
$PGD = \max\left(\left\|d(t)\right\|\right)$	Peak ground displacement	-
$S_a(T_1)$	Spectral acceleration at T_1	-

Table 5-2 - Advanced scalar intensity measures.

Intensity Measure	Name	Reference Study
$S_a^c = S_a(T_1)\left[\dfrac{S_a(cT_1)}{S_a(T_1)}\right]^{\alpha}$	-	Cordova *et al.* (2000)
$I_{N_p} = S_a(T_1)N_p^{\alpha}$	-	Bojórquez and Iervolino (2011)

5.3 Case-study application

The three regular RC 4-storey, 4-bay bare frames, presented in Chapter 3 are selected to illustrate the evaluation of the considered IMs. It is reminded that the three buildings, namely Pre-, Low- and Special-code buildings, represent different vulnerability classes based on the building codes used for their design. Specifically, the selected case-study structures share the same geometry (bay widths and storey heights) but are characterized by different material properties, cross-section dimensions and reinforcement detailing.

The buildings' response is calculated with the aim of simplified analysis FRACAS. First, the buildings are subjected to static PO analysis using different applied lateral load distributions, namely uniform distribution for the case of Pre- and Low-code and triangular distribution for

Special-code building. The resultant PO curves combined with the a set of ground motion time histories (as obtained from SIMBAD database, discussed in Section 3.2.2.1) are utilized to compute the capacity curves the elastic/inelastic spectra respectively, in order to estimate the PPs. The results obtained from theses analyses are then expressed in data sets of IM and EDP. **Figure 5.1** shows the PPs in the ADRS space computed by FRACAS using the GMs records described in Section 3.2.2.1 and the distributions of lateral loads associated to each case-study building.
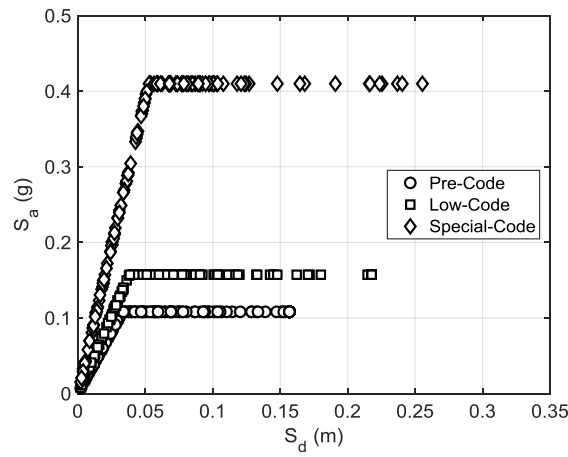
Figure 5.1 - Performance points generated by FRACAS using Elastic Perfectly Plastic (EPP) idealization model in the spectral acceleration-displacement space.

5.4 Methodology

Statistical regression techniques are implemented here to determine the IM that better predicts each of the considered EDPs. To determine the statistical properties of the cloud response (Jalayer and Cornell 2009), the linear least squares is applied on EDPs versus IMs for the suite of ground motion (unscaled) in order to estimate the conditional mean and standard deviation of EDP given IM. The simple power-law model presented in Section 2.7.3.1 (Eq. (2.1)) is used here, where α and b are the parameters of the regression. The regression's standard deviation (β) is assumed to be constant with respect to IM over the range of IMs in the cloud and equal to:

$$\beta = \sqrt{\frac{\sum\limits_{i=1}^{N}\left[\ln\left(edp_i\right) - \ln\left(a\ im_i^b\right)\right]^2}{N_{GM} - 2}} \qquad (5.4)$$

In Eq. (5.4), edp_i and im_i are the EDP and IM values corresponding to the i-th ground motion - edp_i being calculated through FRACAS; N_{GM} is the total number of ground motions. Also, in the case of having vector IMs, a multivariate linear regression model can be easily applied. More advanced statistical approaches also exist, such as the Bayesian emulator-based approach introduced in Chapter 6 (Minas et al. 2018). However, the use of a less complex approach is chosen here as it adequately serves the purposes of this study.

The power-law model can be simply re-written as shown below in Eq. (5.5), as a linear expression of the natural logarithm of the EDP and the natural logarithm of the IM:

$$\ln\left(EDP\right) = \ln\left(a\right) + b\ln\left(IM\right) + e \qquad (5.5)$$

where e is a zero-mean random variable representing the variability of ln(EDP) given the IM. The use of logarithmic transformation indicates that the EDPs are assumed to be conditionally lognormally distributed (conditional upon the values of the IMs); this is a common assumption that has been confirmed as reasonable in many past studies. In the current study, the focus is laid on displacement-based EDPs, which are listed below:

1. peak (over time) inter-storey drift ratio, as the largest difference between the lateral displacements of two adjacent floors, divided by the height of the stoery (denoted as *IDRi* for storey *i*);

2. maximum (over all stories) peak inter-storey drift ratio (i.e. *MIDR*);

3. ratio of the peak lateral roof displacement to the building height (i.e. *RDR*).

The abovementioned EDPs have demonstrated to be well correlated to both structural and non-structural damage, which contribute a major share of the total loss in an earthquake.

5.5   Optimal IM selection criteria

The choice of a given IM may have a significant effect on the uncertainty associated with the probabilistic seismic demand model and the resulting fragility curves. Therefore, the selection of optimal IMs is of high importance within the entire probabilistic seismic risk assessment process and consequently, raised the need for defining quantitative and qualitative selection criteria in order to facilitate these procedures. As discussed above, several studies have been carried out to investigate aspects of ground motions that are responsible for the corresponding damage to particular elements or the system as a whole.

The most commonly used criteria for the determination of an optimal IM found in the literature include efficiency, sufficiency/relative sufficiency, practicality, proficiency, scaling robustness and hazard computability. However, in this Chapter our main focus will be on the criteria listed below and reviewed in greater detail in the subsections 5.5.1-5.5.3:

- Efficiency

- Sufficiency/Relative Sufficiency

- Hazard computability.

5.5.1   Efficiency

Efficiency is the most commonly used quantitative criterion for the determination of optimal IMs, and is related to variation of demand estimations for different values of a studied IM (Padgett et al. 2008). Specifically, more efficient IMs result in a reduced dispersion of the median EDP estimates conditional to a given IM. Thus, less analysis runs are required to narrow down the confidence intervals for those estimates.

The most efficient IM, that best predicts the EDP, is the one that provides the largest value of the coefficient of determination, R^2, among those considered or, equivalently, the one with the smallest value of standard deviation, s. R^2 is the proportion of variability in the EDP that is accounted for by the statistical model.

5.5.2 Sufficiency/Relative Sufficiency

With regard to sufficiency, different approaches exist for the assessment of this metric. In particular, Padgett et al. (2008), among others, define sufficiency as a criterion that characterizes the level of IM statistical independence conditional to specific earthquake characteristics, such as source-to-site distance (D) and magnitude (M). According to this approach, sufficiency is quantified based on the residuals dependence on M and D obtained through regression analysis (Luco and Cornell 2007). Therefore, the estimations obtained from a sufficient IM are not correlated with the ground motion parameters M and D. However, the possible p-value analysis from such an approach provides only a binary evaluation of the IM (i.e. sufficient or insufficient) without offering any indication concerning the degree of insufficiency, when this is detected or, more in general, it doesn't explicitly quantify the relative performance of different, sufficient or insufficient, candidate IMs.

An alternative definition of sufficiency, known as relative sufficiency, was recently proposed by Jalayer et al. (2012). In particular, this measure investigates the relative sufficiency of a second IM, i.e. IM_2, with respect to a first one, IM_1. This measure is derived on the basis of information theory concepts and quantifies the suitability of one intensity measure relative to another. Specifically, the relative sufficiency measure, denoted herein as $I(EDP|IM_2|IM_1)$, is equal to the average difference between the information gained (measured in terms of Shannon entropy or simply entropy (Cover and Thomas 1991)) about the performance variable EDP given IM_1 and IM_2 and that gained given IM_1 only. If $I(EDP|IM_2|IM_1)$ is positive, this means that on average IM_2 provides more information about EDP than IM_1; hence, IM_2 is more sufficient than IM_1. Similarly, if $I(EDP|IM_2|IM_1)$ is negative, IM_2 is less sufficient than IM_1. This is numerically expressed as:

$$I\left(EDP \mid IM_2 \mid IM_1\right) \cong \frac{1}{N_{GM}} \sum_{i=1}^{N} \log_2 \frac{p\left[EDP = edp_i \mid IM_1\right]}{p\left[EDP = edp_i \mid IM_2\right]} \qquad (5.6)$$

In Eq. (5.6), the probability density function (PDF) $p\left[EDP_i \mid IM\right]$, considering a lognormal distribution with the parameters defined through the cloud analysis, is calculated as follows:

$$p\left[EDP = edp \mid IM\right] = \frac{1}{edp \cdot \beta} \Phi\left(\frac{\ln edp - \ln\left(a\ im^b\right)}{\beta}\right)$$

(5.7)

where $\Phi(\bullet)$ is the standardize Gaussian PDF. Hence, the relative sufficiency measure can finally be expressed as:

$$I\left(EDP \mid IM_2 \mid IM_1\right) \cong \frac{1}{N_{GM}} \sum_{i=1}^{N} \log_2 \frac{\beta_1}{\beta_2} \frac{\Phi\left[\left(\ln edp_i - \ln\left(a_2\ im_2^{b_2}\right)\right)/\beta_2\right]}{\Phi\left[\left(\ln edp_i - \ln\left(a_1\ im_1^{b_1}\right)\right)/\beta_1\right]}$$

(5.8)

where $\left(a_1, b_1, \beta_1\right)$ are the cloud parameters corresponding to IM_1 while $\left(a_2, b_2, \beta_2\right)$ are the cloud parameters corresponding to IM_2, and N_{GM} real ground motion records. More details on the derivation of Eq. (5.8) are provided in Jalayer et al. (2012). Relative sufficiency is estimated for each cloud analysis performed and is measured in units of bits of information. According to Jalayer et al. (2012) the relative sufficiency measure provides a preliminary ranking of candidate IM_2 with respect to the reference IM_1. This approximation can be used for a fast screening of various candidate IMs, however, in some cases leads to inaccurate measures. To account for these inaccuracies, Jalayer et al. (2012) proposed a refined method, which uses a stochastic ground motion model combined with disaggregation of the seismic hazard at the site, and then employs Monte Carlo simulation to estimate the expectation involved in the relative sufficiency measure's definition. It is noteworthy to mention that the estimates of the refined method generally show good agreement with the rough preliminary estimates, although in some cases refined modelling results in different rankings of the candidate IMs. However, the use of the approximate formula shown in Eq. (5.8), is still considered to be an adequate option for comparing the IMs without the need to use the refined method (Jalayer et al. 2012).

5.5.3 Hazard computability

According to the definition given in Giovenale et al. (2004), hazard computability describes the process to obtain the earthquake hazard at a given site in terms of a considered IM. Numerous hazard maps and GMPEs exist for the most commonly used IMs, namely PGA and spectral ordinates at given periods (representing sometimes a restricted range of possible discrete periods), making these IMs more favourable from the hazard computability perspective; whereas, other IMs (spectral ordinates at the actual fundamental and/or elongated periods and advanced IMs) may require interpolation or supplementary structural or dynamic information, making the computation of the hazard a more time-consuming process.

5.6   Results and discussion

In this Section all the considered IMs are assessed based on the selection criteria discussed in Section 5.5. Specifically, the overall performance of each considered IM, for each considered EDP (computed through FRACAS), is obtained based on the comparison of the quantitative parameters associated with each testing criterion. For sake of brevity, only two individual case studies, namely the Pre- and Special-code buildings subjected to uniform and triangular PO loads respectively, are selected to show the optimal IM selection process discussed in this study, and therefore determine the optimal IM for the fragility analysis of each particular building class. However, the same process has also been applied to the Low-code case-study building and the results of the analysis are essentially consistent with the observations obtained from the Pre-Code case study building. It is also worth noting that, to incorporate modelling uncertainties in the assessment of seismic demand assessment of the case-study buildings, samples of the material, geometrical, and mechanical parameters for each structural model could be generated and analysed subjected to the considered suite of earthquake ground motion records. However, the focus here is on investigating the use of record-to-record variability in nonlinear static capacity approaches for fragility assessment.

As shown in **Figure 3.2**, the two selected structures behave highly nonlinearly over a threshold of approximately 0.40% and 0.65% RDR for the Pre- and Special-code buildings respectively. The ground motion records not "pushing" the considered structures into the nonlinear range are discarded from the IM-EDP datasets used herein. It is noted that FRACAS, as a capacity spectrum-based method, utilizes an equivalent SDoF system to represent the behavior of the modelled building and, by definition, $S_a(T_1)$ represents the response of a linear elastic SDoF. As result, the seismic intensity and the estimated FRACAS response are perfectly correlated within the elastic regime. Based on this, the actual number of ground motions that pushed the frame into the nonlinear range is relatively small but still significant, corresponding to approximately 15% and 25% of the total number of records used, for the case of Special- and Pre-code buildings respectively (i.e. 68 and 108 GMs, as shown in **Figure 5.1**).

Figure 5.2 and **Figure 5.3** illustrate the scatter plots of the structural seismic demand in terms of MIDR versus the scalar candidate IMs for the considered ground motion subset and for each selected case-study structure. Note that for the calculation of I_{N_p} values presented in these figures, a single α-parameter is used for each case study building, based on the optimal calibration process described later in this Section. These figures also show the median and two standard deviations above and below the median, respectively, from the logarithm regression model fitted to the data. With regard to efficiency, the visual inspection of **Figure 5.2** and **Figure 5.3** confirms that displacement-based EDPs, such as MIDR, is better correlated with the spectral shape parameter I_{N_p} for all studied cases (lowest value of standard deviation, *s*). In contrast, the peak ground parameters, namely PGA and PGD, are confirmed to be poor predictors of the nonlinear structural response of mid-rise MRFs (highest values of standard deviation), while PGV performs reasonably well in the case of the Special-code building and poorly in the case of the Pre-code case-study building. As expected, $S_a(T_1)$ appears to be the most efficient conventional IM in all cases, closely matching the performance of I_{N_p} for the

Special-Code case study. Finally, S_a^c performs consistently well in terms of efficiency for all tested case studies.
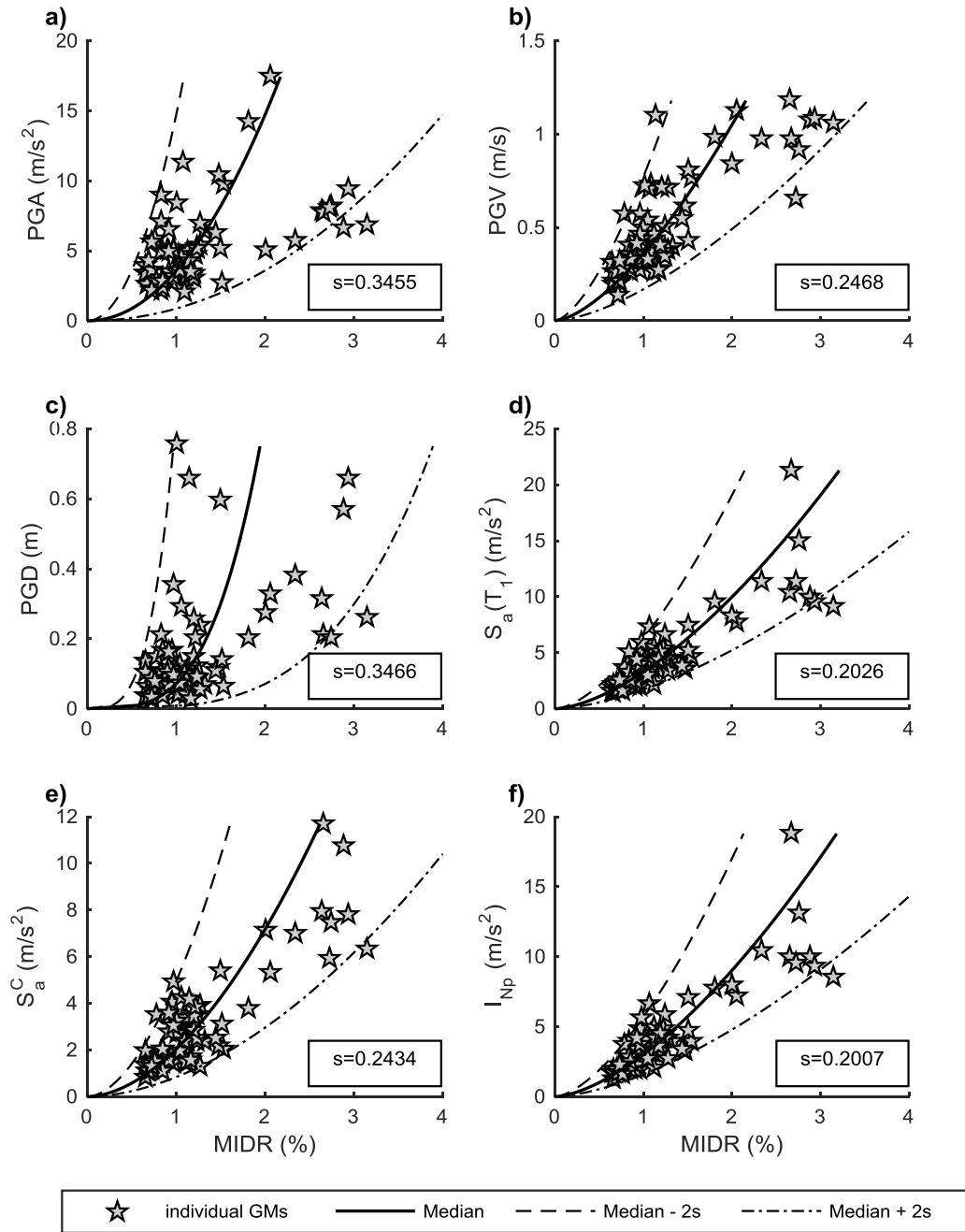
Figure 5.2 - Scatter plots of the adopted IMs versus MIDR for the considered subset of ground motion records (Special-code building, Triangular PO).
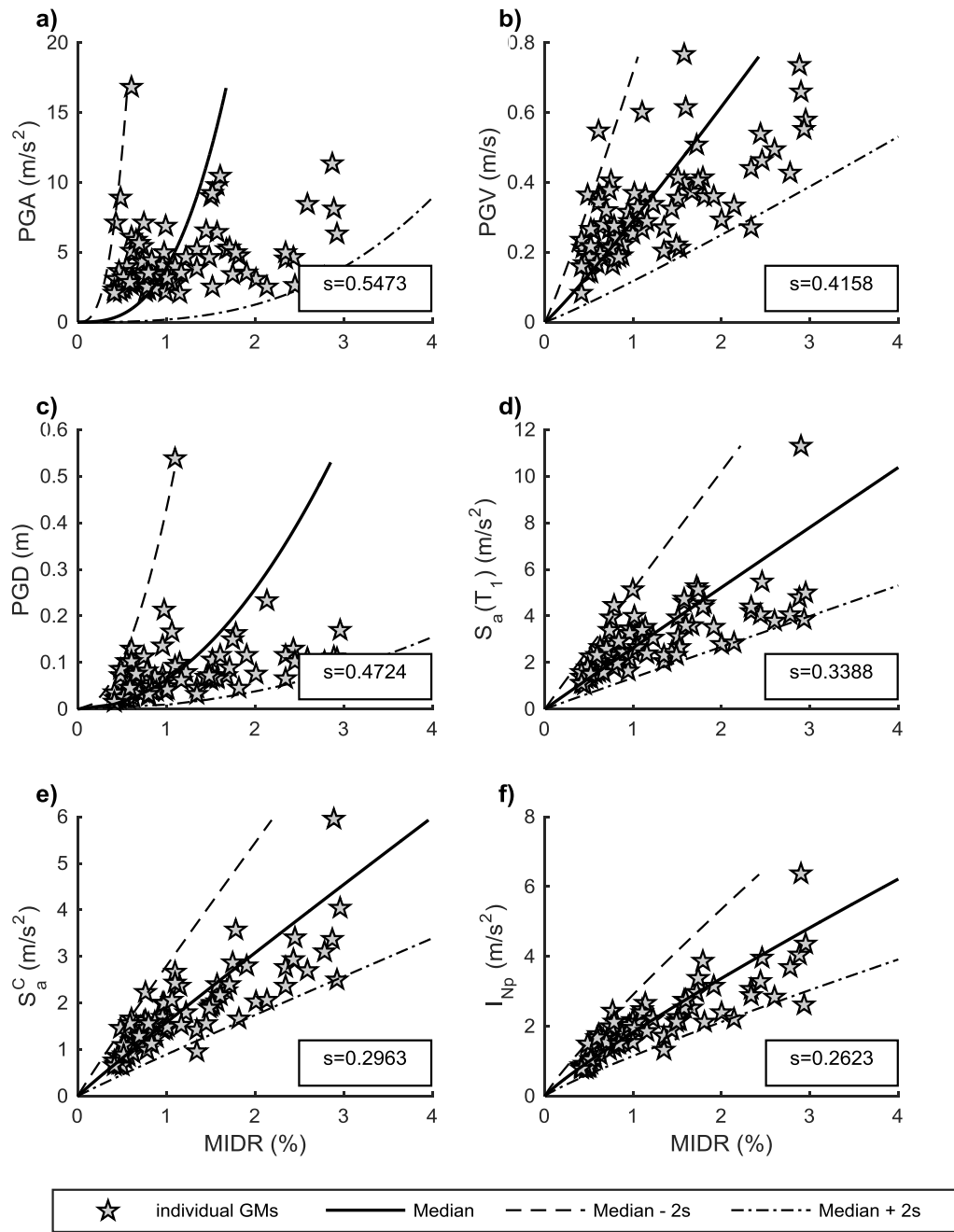
Figure 5.3 - Scatter plots of the adopted IMs versus MIDR for the considered subset of ground motion records (Pre-code building, Uniform PO).

As expected, the superiority of the advanced IMs becomes more evident for the case study buildings experiencing highly nonlinear demands, such as Pre- and Low-code buildings, as these IMs can account for the effect of period elongation. On the other hand, the superiority of advanced IMs is less apparent for the Special-code building, although it is nominally designed to provide higher nonlinear capacity/ductility than the Pre- and Low-code buildings. This is due

to the high strength level (e.g., in terms of base shear coefficient) characterizing the Special-code case-study structure. In fact, only a small number of the selected ground motion set actually "pushes" the building into the nonlinear range (68 records), making more difficult to exploit the nonlinearity assumed in the design.

A potential improvement in the performance of the advanced IM, I_{N_p} may be achieved by calibrating the α-parameter in its definition for the specific case study structures rather than using the values suggested by other researchers for different case study structures. To this aim, ten different sets of α-values are tested, ranging from 0.1 to 1.0, and the associated I_{N_p} - MIDR relationship dispersions are computed and compared to dispersion of the other candidate IMs for all three buildings, as shown in **Figure 5.4**.
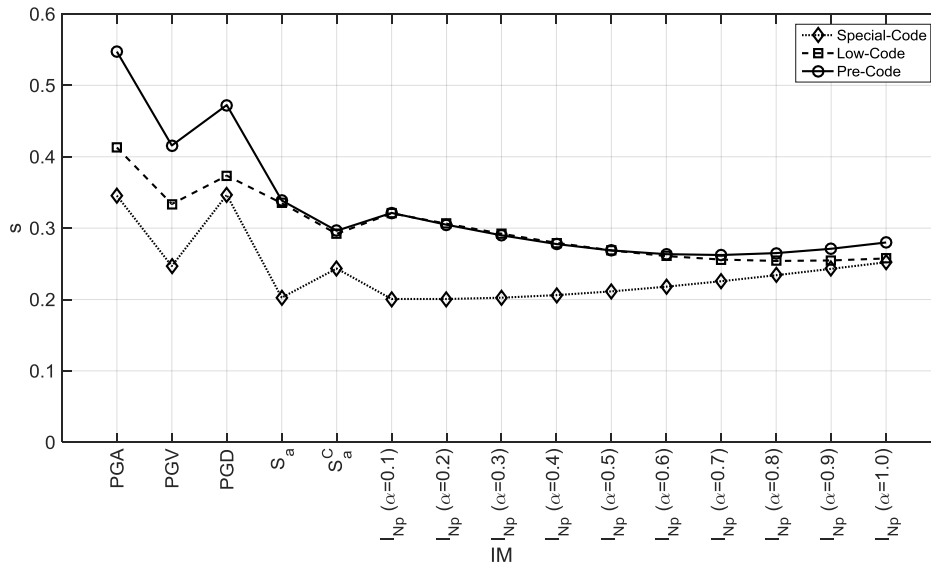
Figure 5.4 - Standard deviation (dispersion) of residuals of MIRD for the considered IMs and each case-study building.

When the α-parameter approaches values close to zero, I_{N_p} approximates $S_a(T_1)$, which also becomes apparent when comparing the standard deviations of the EDP- $I_{N_p}(\alpha = 0.1)$ and EDP-$S_a(T_1)$ models (**Figure 5.4**), resulting in almost identical s-values. Similarly, when the α-parameter values tend to one, I_{N_p} approximates the average spectral acceleration, *AvgSA*, over

three periods, namely 1, 1.5 and 2s. The latter case practically means that spectral accelerations corresponding to periods greater than T_1 (i.e. 1.5 and 2s – elongated periods) contribute equally in the estimation of MIDR. As discussed above, the Special-code building is characterized by higher nominal inelastic displacement capacity and ductility; however, the use of cloud analysis coupled with the specific, yet realistic, ground motion set selected herein do not allow one to exploit those. While this is a well-known limitation of the cloud analysis approach as opposite to IDA (see results in Kazantzi and Vamvatsikos (2015)), the ground motion set used here consists of the strongest records in the SIMBAD database and the corresponding IM values are hazard-consistent, even at very higher return periods, as it will be shown in the next Section. This explains why the optimal α-parameter for the case of Special-code building is equal to 0.2. In contrast, the optimal α-parameter for Pre- and Low-code buildings is equal to 0.7 and 0.8 respectively, as a bigger number records "push" the building to the nonlinear range of behaviours where period elongation occurs.

The relative sufficiency measure for MIDR and the candidate IMs is presented in **Figure 5.5** for all three buildings. The reference IM_1 is chosen here as the one corresponding to the lowest s value from the regression (**Figure 5.4**). The results in **Figure 5.5** confirm the results in terms of efficiency (**Figure 5.4**). The IMs resulting in the highest efficiency are also characterized by the highest relative sufficiency.

Next, this Chapter investigates the sensitivity of the dispersion of residuals of MIDR for the considered IMs to the number of periods used to compute I_{N_p}. Ten equally spaced successive periods between 1 and 2s are utilized to recalculate I_{N_p}. **Figure 5.6** shows the new dispersion estimates for the recalculated I_{N_p} (grey lines) and how they compare the previous dispersion estimates, where only three periods (i.e. 1, 1.5 and 2s) are used (black lines). The results in **Figure 5.6** confirm the results shown in **Figure 5.5**, regarding the optimal α-parameter calibration. Additionally, this test reveals that the use of a bigger number of periods does affect the resultant performance of I_{N_p}, but not necessarily in a positive way. The use of more periods

results in a reduction of the dispersion of MIDR residuals for the cases of Pre-(-0.7 to -4.4%) and Low-code buildings (-1.0 to -10.3%), while it has a counter effect in the case of Special-code, as the associated dispersions tend to increase (+0.2 to +4.2%). This is again attributed to small number of records actually "pushing" the Special-code building deep in the nonlinear regime. As a result, the effect of the additional spectral accelerations corresponding to higher periods not only is not adding any additional information but is also having an adverse impact in the resultant standard deviation estimates.
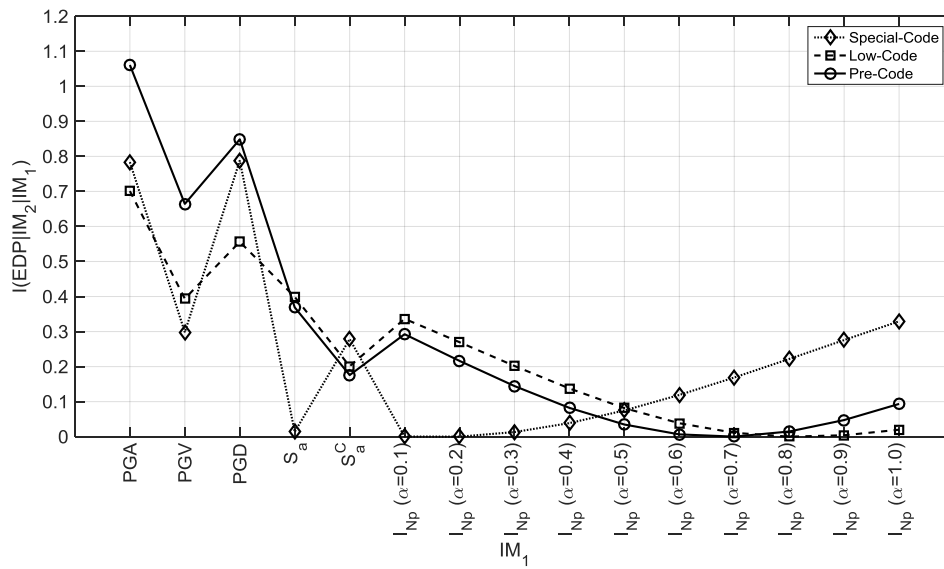
Figure 5.5 - Relative sufficiency measure for alternative IMs with respect to the IM with the lowest dispersion (Figure 7) for each case-study building.

It is worth noting that, given the complex relationship between ground motions and nonlinear structural response and the difficulty of capturing relevant ground motion features with a single parameter, it is unlikely that a scalar IM will render nonlinear structural response conditionally independent of all other ground motion parameters. This is also true for a vector IM that uses only a small number of parameters. In the context of practical applications, as those presented in this Chapter, a sufficient IM is the one minimizing the influence of the ground motion set that is used to estimate seismic demands. This also implies that the careful record selection and/or modification required to obtain a good estimate of the structural performance may not be required or may be significantly reduced if advanced IMs are used.
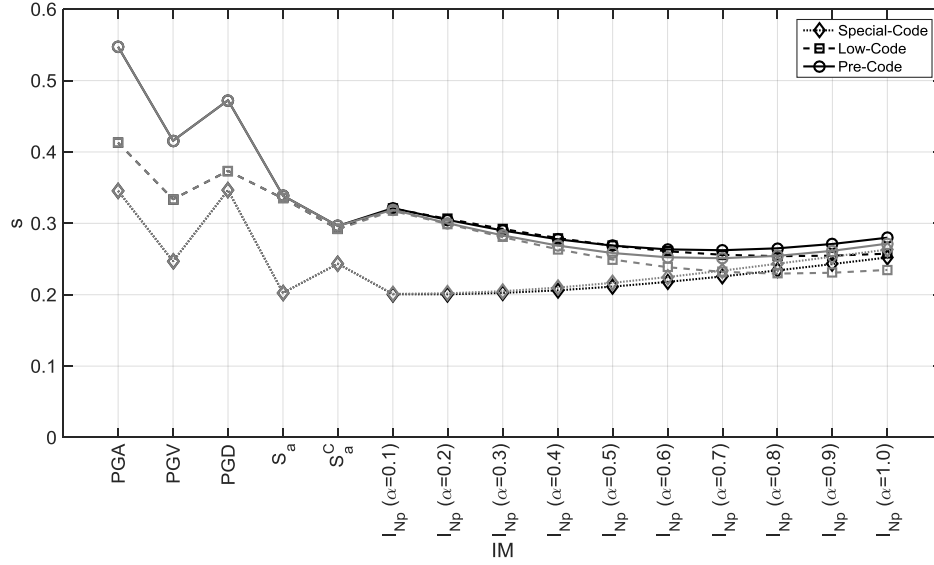
Figure 5.6 - Sensitivity of the standard deviation (dispersion) of residuals of MIRD for the considered IMs and each case-study building to the number of periods used to compute (grey lines correspond to 10 equally-spaced periods between 1s and 2s).

The last criterion for the determination of an optimal IM is the hazard computability. For this criterion, conventional IMs have a significant advantage over advanced IMs, as numerous GMPEs and hazard maps exist particularly for PGA, PGV and PGD, and some spectral ordinates for specific ranges of periods. On the other hand, it is still possible to derive GMPEs for some of the advanced IMs, namely I_{N_p} and S_a^c, using either *direct* or *indirect* methods. In particular, Kohrangi et al. (2018), derived empirical GMPEs for the average spectral acceleration, *AvgSA*, defined as the geometric mean of spectral acceleration values over a range of periods (i.e. for I_{N_p} with α = 1). As pointed out by the authors (and also in Bojórquez and Iervolino (2011)), one of the advantages of the advanced, spectral-shape-based IMs, stands on the assumption that their distribution is computable from the available GMPEs for spectral acceleration, GMPE-*SA*, (indirect method) without the need for deriving new specific GMPEs for *AvgSA*, GMPE-*AvgSA*, (direct method). The results in Kohrangi et al. (2018) show that the indirect approach yields median *AvgSA* estimates that are identical to those of the direct approach. However, the estimates of *AvgSA* variance of the two methods are identical only if

both the GMPE-*SA* and their empirical correlation coefficients among different *SA* ordinates are derived from the same record dataset.

5.6.1 Example of hazard computability in terms of I_{Np}

An illustrative example of site-specific PSHA in terms of I_{N_p} with $\alpha = 0.7$ is presented here. Such a PSHA is carried out by using a Monte Carlo simulation-based approach (e.g., Assatourians and Atkinson 2013), accounting for uncertainty in all the factors affecting ground motions at a given site. Specifically, the town of Avellino in the Campania region of Southern Italy is used here as a case-study site. Avellino is characterized by high seismicity, with two major events occurred in the last 90 years (i.e. the M6.7 on 27/7/1930, and M6.9 on 23/11/1980). To perform PSHA for Avellino, a synthetically generated set of potential earthquakes, with their temporal and geographical distribution, is developed by drawing random samples from the assumed PSHA model components (and related probability distributions), i.e. source-zone geometries and magnitude-recurrence parameters and maximum magnitude. The official Italian seismogenetic zonation, named ZS9 (Meletti et al. 2008), is used in this study; the calculation is limited to events with source-to-site distance up to 100 km (**Figure 5.7**). Gutenberg-Richter parameters implemented for generating each record are adapted from Barani et al. (2010).

The resulting synthetic catalogue has a duration of 5,000 years; each record of the synthetic catalogue contains the following fields: time (in decimal years), coordinates (latitude and longitude) and magnitude of earthquake, source zone number and corresponding fault-style. In fact, ZS9 assigns a prevalent mechanism of faulting – interpreted as the mechanism with the highest probability of generating future earthquakes – to all its source zones for use in the GMPEs.

The considered IM is evaluated for each seismic event contained in the catalogue by using the indirect approach and models presented in Kohrangi et al. (2018), assuming type B ground. 500 realizations of random numbers drawn from the standard normal distribution is multiplied by

the given sigma value (variability of the GMPE model) and added to the median log-ground motions (from the GMPE) to model the aleatory variability in ground motions. The resulting site-specific hazard curves for each realization (light grey colour) as well as the median, 16th and 84th hazard curves (blue colour) are shown in **Figure 5.8**.
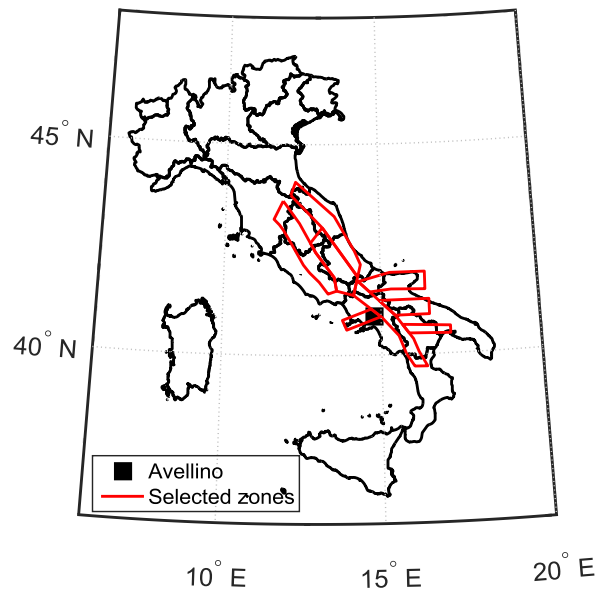
Figure 5.7 - Case-study location (Avellino) and considered ZS9 seismic sources.

The actual possibility of computing hazard analysis is crucial for the usefulness of the proposed optimal IMs. It is worth noting, as discussed in Kazantzi and Vamvatsikos (2015), that within a risk assessment framework, the reduction in response dispersion gained by a more efficient IM, does not reduce the overall risk variability. A different IM is simply 'a different partitioning of the sample space for applying the total probability theorem on which the risk calculation is based upon' (e.g., Kazantzi and Vamvatsikos (2015)). Thus, part of the variability is merely shifted to a different level within the risk assessment, and in particular to the seismic hazard curve definition. By definition a more efficient IM is more structure specific, therefore, higher dispersions are obtained when trying to define an appropriate GMPE. In other words, no matter the adopted IM, as long as sufficiency is maintained, the same overall variability at the end of the risk assessment, i.e. after convolving the vulnerability with the hazard, should be observed. However, shifting a part of the variability from the response analysis level towards the

probabilistic hazard analysis level has a major advantage, a significant reduction in the computational costs associated to the structural analysis.
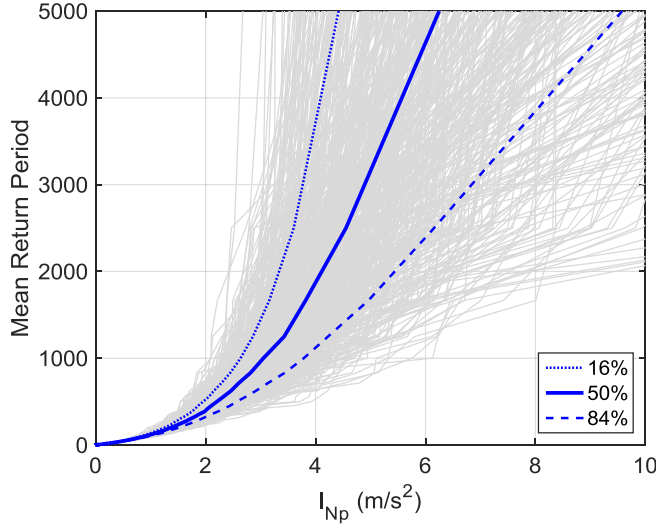
Figure 5.8 - Site-specific hazard curves for Avellino in terms of I_{Np} with $\alpha = 0.7$.

5.7 Fragility assessment

In this Section, two sets of fragility functions are presented, one set for the conventional IM, $S_a(T_1)$, and one for the advanced IM, I_{N_p}, using the optimal α-parameter as calibrated for each vulnerability class. These IMs are chosen to represent the conventional and advanced IMs respectively, due to their overall performance in the tests conducted in Section 5.6. The fragility curves are derived from the FRACAS analysis results by adopting thresholds of MIDR to define three DSs. The descriptions and the thresholds associated with each DS, which are used for the derivation of the fragility curves can be found in **Table 3-3** (Chapter 3). In addition, the locations of DSs along the static PO curves for the three case-study buildings is depicted in **Figure 3.7** (Chapter 3).

As in Section 5.6, only two case-study buildings, namely the Pre- and Special-code buildings, are used to demonstrate the derivation of fragility curves for the selected optimal IMs. **Figure 5.9** and **Figure 5.10** illustrate the median fragility curve sets expressed in terms of $S_a(T_1)$ (left

panel) and I_{N_p} (right panel), and their associated 95% confidence intervals estimated using the bootstrap technique for Pre- and Special-code buildings respectively. In brief, a large number of bootstrap samples (1,000 in this case) are simulated and a fragility function is drawn for each bootstrap iteration and consequently the 95% bootstrap confidence intervals are computed. For more details about the derivation of bootstrap confidence intervals one may refer to Rossetto et al. (2014). It is noted that only the first two DSs are shown in **Figure 5.10**, as the partial collapse damage state (DS3) is not reached for this case study for the particular suite of records used herein.

Visual inspection is considered to be an acceptable means of comparison of fragility curves due to the difficulty of addressing a numerical goodness of fit measure (Bojórquez et al. 2012). A steep fragility curve indicates a significant explanatory power of the IM in question, while the width of confidence intervals describes the level of dispersion.

As a result, the visual comparison of the two fragility curves sets shown in both **Figure 5.9** and **Figure 5.10** reveal an apparent superiority of the I_{N_p} comparing to the $S_a(T_1)$) fragility curves. The use of the advanced IM, I_{N_p}, results in steeper fragility curves for both DS1 and DS2, and also narrower confidence intervals. This improvement is less obvious for the case of Special-code building for the reasons extensively discussed in Section 5.6. In addition, $S_a(T_1)$ showed to perform considerably well, almost matching the performance of I_{N_p}, for cases where the impact of nonlinearity is limited or not fully exploited.
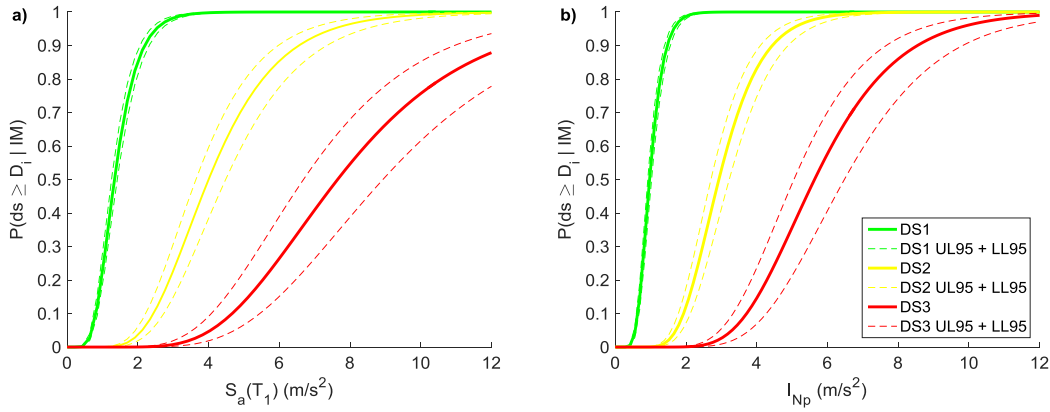
Figure 5.9 - Fragility curves and their 95% confidence intervals derived for Pre-Code building, expressed in terms of: a) $S_a(T_1)$ and b) I_{Np}.
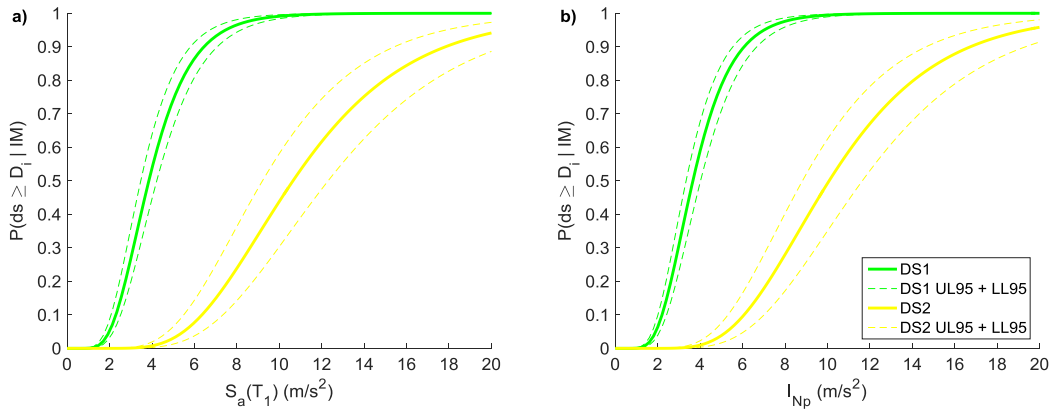
Figure 5.10 - Fragility curves and their 95% confidence intervals derived for Special-Code building, expressed in terms of: a) $S_a(T_1)$ and b) I_{Np}.

5.8   Conclusions

This Chapter presented an investigation aiming at identifying the ground motion IMs that are better correlated with displacement-based response parameters for simplified fragility analysis of mid-rise RC buildings. The selection of an efficient, sufficient, and hazard computable IM is not a trivial task towards developing analytical vulnerability curves and consequently assessing the seismic losses for a building class. For this purpose, three case-study existing RC frame buildings, typical of the Italian-Mediterranean inventory of different vintages were considered. To establish correlations between IMs and EDPs describing the nonlinear performance of the case-study structures, a comprehensive set of ground motions and an innovative CSM that uses

inelastic response spectra derived from actual earthquake accelerograms were used. Six alternative ground motion IMs, including advanced IMs accounting for spectral shape, have been considered. Efficiency, sufficiency and relative sufficiency, and hazard computability were recognized as the main criteria for judging the adequacy of each candidate IM in this study and selecting the optimal IM. The outcomes of the present work are consistent with previous investigations conducted by the authors and other researchers on selecting optimal IMs (scalar or vector-valued) for predicting structural response by using computationally demanding, full nonlinear dynamic analysis. In fact, it has been demonstrated that a successful IM might be formed by specifying an appropriate period range that includes periods above the mean first mode period (thus accounting for inelastic period elongation) and combines the associated spectral acceleration values via the geometric mean. In particular, an IM that uses only three such periods was found to perform best in terms of efficiency and relative sufficiency across the entire practical range of local (i.e. at storey level) and global EDP for each case-study structure In general, this type of advanced IMs comfortably satisfy all the selection criteria, including the hazard computability criterion. Indeed, this work showed how probabilistic seismic hazard analysis for I_{N_p} can be easily performed with existing tools.

6. Surrogate Modelling for Probabilistic Seismic Response Analysis (PSRA)

6.1   Introduction

As discussed in Section 2.7.6 of the literature review, surrogate models are becoming increasingly popular within the context of seismic fragility and vulnerability assessment because they are able to replace the time-consuming analysis applications. However, although they have been utilized in various different applications, they have been rarely used to fit the IM|EDP relationship.

Several approaches exist for the analytical fragility assessment of buildings, offering the user the flexibility to choose from different analysis methodologies, various structural parameters, characteristics of input ground motions etc. Large numbers of analysis runs are required in order to account for uncertainties in the earthquake ground motions (and structure characteristics if looking at building classes), hence analytical approaches can be computationally very demanding, especially when nonlinear dynamic analyses are used. To reduce this computational time, past studies have made simplifications to the structural model, used a reduced number of ground motions or adopted a simplified structural analysis approach. This Chapter instead draws upon the gaps identified in the existing literature, and proposes a BEA that allows a reduced number of analyses to be carried out to compute the probabilistic seismic response and fragility of buildings. A main advantage of the BEA is its flexibility, as it does not impose a fixed mathematical form on the EDP|IM relationship (unlike other approaches such as the standard cloud method). It is highlighted that past risk models were only able to incorporate parametric expressions of EDP|IM, however, modern risk codes are able to accommodate nonparametric relationships. In addition, BEA makes fewer assumptions than standard methods, and provides

improved characterization of uncertainty, including the uncertainty in the EDP|IM representation and uncertainties in model parameter estimation.

The Chapter starts with an overview of statistical emulation and BEA. Next, the procedure followed in this research work for the evaluation of BEA performance versus that of a standard approach (i.e. cloud method) is described, including the detailed definition of the experimental design and the sensitivity tests carried out on BEA input. Finally, the process for deriving fragility functions using BEA is presented and evaluated. This Chapter aims to provide an answer to Research Question 3.

6.2 Developing EDP versus IM relationships through Bayesian emulation

Complex mathematical models of real-world processes exist in all scientific and engineering areas. These mathematical models are usually translated into computer codes and may require significant computational resources. The mathematical functions and the related computer codes may be referred to as simulators (O'Hagan 2006). A simulator, as a function $f(.)$, takes vectors of inputs $\mathbf{x}$ and generates unique outputs, $y = f(\mathbf{x})$ for each given input. In the context of analytical fragility estimation, the input $\mathbf{x}$ represents a ground motion sequence and the output $f(\mathbf{x})$ represents the simulated EDP for that sequence. A simulator will produce outputs that have an associated uncertainty, for example, due to uncertainties in the inputs and in the simulator construction itself (Kennedy and O'Hagan 2001). In analytical fragility functions this uncertainty is due to differences in ground motions associated with any IM value, limitations in numerically reproducing physical phenomena of the structure response and simplified assessment. To assess the effect of these uncertainties, one option is to use uncertainty and sensitivity analysis methods requiring many simulator runs, as discussed in Saltelli et al. (2000). However, this is impractical for complex simulators that are computationally demanding to run (O'Hagan 2006). To overcome this, an approximation $\hat{f}(\mathbf{x})$ of the simulator $f(\mathbf{x})$, known as a statistical emulator, may be introduced to act as a surrogate. If $\hat{f}(\mathbf{x})$ is a good proxy of the

simulator, it can be used to carry out uncertainty and sensitivity analysis but with significantly less effort. Additionally, $\hat{f}(\mathbf{x})$ can be also used to estimate the values of $f(.)$, at input configurations where the simulator has not been run.

A small number of data points, obtained by running the simulator at carefully chosen configurations of the inputs, is required to train an emulator (and provide values of standard deviations to quantify the approximation error). Following common practice in this type of problem (e.g., Kennedy and O'Hagan 2001), the present study adopts a Bayesian approach with an emulator based on a GP model.

In the standard GP approach to the modelling of simulator outputs, the output function $f(\mathbf{x})$ is regarded as a realized value of a random process such that the outputs at distinct values of $\mathbf{x}$ jointly follow a normal (or Gaussian) distribution. In most applications, $f(\mathbf{x})$ is a 'smooth' function in the sense that a small variation in the input $\mathbf{x}$ will result in a small perturbation in the output $f(\mathbf{x})$. Smoothness of a GP is ensured by specifying an appropriate structure for the covariance between process values at distinct values of $\mathbf{x}$. For example, the use of a Gaussian covariance model (not to be confused with the "Gaussian" in the GP, which relates to the distribution, rather than the covariance structure), results in smooth realizations that are infinitely differentiable (Tarantola 2005).

As noted above, in the context of fragility estimation, the input $\mathbf{x}$ represents a ground motion sequence. However, fragility estimates are rarely presented as functions of an entire ground motion sequence: rather, they are presented as a function of an intensity measure $t = t(\mathbf{x})$, say, that is used as a proxy for the complete sequence. In practice, the relationship between any such scalar proxy measure and the EDP is imperfect so that the simulator outputs, regarded as a function of t rather than $\mathbf{x}$, are no longer smooth. As an example of this, two different ground motion records characterized by the same value of t (i.e. having the same IM) will not in general generate identical outputs. Consequently, the standard emulation approach must be modified slightly in order for it to be applied in the context of fragility estimation.

The required modification is still to regard $f(\mathbf{x})$ as the realized value of a random process, but now such that the conditional distributions of $f(\mathbf{x})$ given $t(\mathbf{x})$ themselves form a GP. Specifically, write $f(\mathbf{x})$ as:

$$f(\mathbf{x}) = \mu(t) + \varepsilon(\mathbf{x}) \tag{6.1}$$

where $\mu(t)$ represents the systematic variation of the output with the IM and $\varepsilon(\mathbf{x})$ is a discrepancy term with $E[\varepsilon(\mathbf{x})] = 0$, representing uncertainty due to the fact that t does not fully capture the relevant information in $\mathbf{x}$. In the first instance it is convenient to assume that $\varepsilon(\mathbf{x})$ is normally distributed with a constant variance, τ^2 (this assumption can be checked when applying the methodology). It is also assumed that $\varepsilon(\mathbf{x})$ is independent of $\mu(t)$, and that $\varepsilon(\mathbf{x}_1)$ is independent of $\varepsilon(\mathbf{x}_2)$ when $\mathbf{x}_1 \neq \mathbf{x}_2$. Note that in equation (6.1), the term $\varepsilon(\mathbf{x})$ is deliberately used instead of $\varepsilon(t)$. The reason for doing this is to allow two simulator outputs with different $\mathbf{x}$ but same value of t, to be considered as separate, hence, allowing for scatter about the overall mean curve in a plot of $f(\mathbf{x})$ against $t(\mathbf{x})$. Equation (6.1) then immediately yields the conditional distribution of EDP given IM as normal, with expected value $\mu(IM)$ and variance τ^2.

To exploit GP emulation methodology, a final additional assumption is made, whereby $\mu(t)$ varies smoothly with t and hence can itself be considered as a realization of a GP. Specifically, we denote the expected value of $\mu(t)$ by $m(t)$, its variance by σ^2 and we specify a correlation function (such as the Gaussian) that ensures smooth variation with t. Under the Gaussian correlation function, the covariance between $\mu(t_1)$ and $\mu(t_2)$ is:

$$\sigma^2 \exp\left[-\frac{(t_1 - t_2)^2}{2\varphi^2} \right] \tag{6.2}$$

where φ is a parameter controlling the rate of decay of the correlation between function values at increasingly separated values of t. As an illustrative example, the mean function $m(t)$ could be specified as linear, i.e. $m(t) = \beta_0 + \beta_1 t$: this represents a first approximation to the function $\mu(t)$ in equation (6.1), with the correlation structure allowing for smooth variation in the "approximation error" $\mu(t) - m(t) = z(t)$, say.

Combining all of the above, in the case where the mean function $m(t)$ is linear we find that values of $f(\mathbf{x})$ at distinct values of $\mathbf{x}$ are jointly normally distributed such that:

- The expected value of $f(\mathbf{x})$ is: $\beta_0 + \beta_1 t$

- The variance of $f(\mathbf{x})$ is equal to: $\text{Var}\big[\mu(t)\big] + \text{Var}\big[\varepsilon(\mathbf{x})\big] = \sigma^2 + \tau^2$

- For $\mathbf{x}_1 \neq \mathbf{x}_2$, the covariance between $f(\mathbf{x}_1)$ and $f(\mathbf{x}_2)$ is:

$$\text{Cov}\big[f(\mathbf{x}_1), f(\mathbf{x}_2)\big] = \text{Cov}\big[\mu(t_1) + \varepsilon(\mathbf{x}_1), \mu(t_2) + \varepsilon(\mathbf{x}_2)\big]$$
$$= \text{Cov}\big[\mu(t_1), \mu(t_2)\big] = \sigma^2 \exp\left[-\frac{(t_1 - t_2)^2}{2\varphi^2}\right] \tag{6.3}$$

where $t_1 = t(\mathbf{x}_1)$ and $t_2 = t(\mathbf{x}_2)$.

These expressions depend on the full input vector $\mathbf{x}$ only through the IM $t(\mathbf{x})$. As a result, the simulator outputs can be analyzed as though they are functions of t alone and, specifically, subjected to a standard GP analysis to estimate the systematic variation $\mu(t)$. The only difference between this formulation and a standard GP emulation problem is that when $t_1 = t_2$, the covariance between the simulator outputs is $\sigma^2 + \tau^2$ rather than just σ^2. The additional τ^2 term allows for additional variance associated with the discrepancy term $\varepsilon(\mathbf{x})$ in equation (6.1).

Having formulated the problem in terms of GPs, the simulator can be run a few times at different values of $\mathbf{x}$, and the resulting outputs can be used to estimate the parameters in the GP model which, as set out above, are the regression coefficients β_0 and β_1, the variances σ^2 and

τ^2, and the correlation decay parameter φ. A simple way to do this is to use ordinary least-squares regression to estimate the regression coefficients and then to carry out a least-squares fit on the variogram of the residuals to estimate the remaining parameters as is standard in geostatistics (David 1977); The resulted β_0 and β_1 estimates are then plugged into the relevant formulae in order to carry out the interpolation and assess its uncertainty. However, this approach fails to account properly for the uncertainty in the parameter estimates. From this perspective a Bayesian approach is preferable: among other things, for any IM value this provides samples from the posterior predictive distribution of the associated EDP. This predictive distribution automatically accounts for parameter uncertainty: percentiles of the distribution can be used to provide prediction intervals (we use 95% intervals below), and the posterior predictive probability of exceeding any threshold of interest (e.g., corresponding to a given DS so as to obtain a fragility curve) can be calculated directly. For an introduction to the calculation of predictive distributions in Bayesian analyses, see Gelman et al. (2013). In our work, the Bayesian approach is implemented using the krige.bayes() routine from the geoR package in R (Ribeiro Jr and Diggle 2001).

Any Bayesian approach requires the specification of prior distributions for each parameter estimated: following the implementation in krige.bayes(), we adopt standard non-informative priors where possible for all parameters in order to ensure that the results reflect the information content of the simulator runs to the maximum extent possible. The model parameters and their prior distribution used in our BEA model are listed below:

- Regression coefficients, β_0 and β_1 (intercept and slope). Here, non-informative "flat" prior distributions are chosen.

- Partial sill, σ^2. A prior distribution proportional to $1/\sigma^2$ ("reciprocal") is used here.

- Range, φ. An improper "uniform" prior distribution over the range (0,∞) is selected.

It is noteworthy that the krige.bayes() function does not allow a Bayesian treatment of the variance τ^2 so this is fixed at the value obtained from a conventional kriging analysis as defined above. The effect will be that uncertainty in the predictions will still be underestimated, albeit much less than in most standard analyses. The results of the pseudo-reality experiment reported below will help to determine whether this is likely to be a problem in practice.

Having estimated the parameters, the results can be used to interpolate optimally between the existing simulator runs so as to construct an estimate of the predicted EDP|IM curve $\mu(t)$ and associated standard deviation. The optimal interpolation relies on standard results for calculating conditional distributions in a Gaussian framework. The approach is used extensively in geostatistics to interpolate between observations made at different spatial locations, where it is often referred to as "kriging" (Sacks et al. 1989). In geostatistics, the term τ^2 is referred to as a 'nugget' and accounts for local-scale variation or measurement error. Nonetheless, in the present context the above argument shows that it can be derived purely by considering the structure of the probabilistic seismic demand problem.

It is noteworthy to mention that similarly to the cloud method, the BEA analysis is also carried in terms of $\ln(EDP)$, due to the normality and heteroscedasticity assumptions, which are better justified on the log scale.

6.3 BEA methodology

The adequacy and the performance of the proposed Bayesian approach (BEA) for seismic response is assessed through its comparison with the cloud based approach. A roadmap of the methodology followed herein is shown in **Figure 6.1**. As a first step, two case study buildings, representing distinct vulnerability classes, are analysed at two levels of analysis fidelity (i.e. using nonlinear dynamic and static analysis methods, representing high and low fidelity respectively), in order to construct the associated EDP|IM relationships (steps 1-3, Section 6.3.1). Each of the resultant data sets of IM and EDP is used to generate three "pseudo-realities"

based on different assumptions, which are each characterized by a mean and standard deviation value. The reason for implementing three different "pseudo-reality" scenarios is to investigate the prediction capability of the tested methods under favourable or less favourable conditions. A large number of data sets of IM and EDP are then simulated, using the same mean and standard deviation values as calculated for each "pseudo-reality" (step 4, discussed in Section 6.3.2). Different sampling configurations are employed, including two sampling types and three sampling sizes to construct the data sets (step 5, Section 6.3.3), which are used to train the tested methods (step 6). The effect of selecting different covariance structures within BEA is also investigated. Finally, the results of the tested methods are compared and evaluated against the "pseudo-realities" with the aim of three different metrics (step 7, Section 6.4). A computer code implementing the proposed framework is scripted in R(R Core Team 2014). A code supplement to this study can be accessed online through EPICentre website (http://www.ucl.ac.uk/epicentre/resources/software). Note that although the scripted BEA procedure is more complex than the standard method, it does not practically increase the computational expense.

This methodology process aims to shed light on the sensitivity of the BEA to several input assumptions, and sampling configurations, in the estimation of the mean of the EDP|IM distribution and the uncertainty characterization. The comparisons presented here are carried out in terms of EDP|IM predictions, however, a single case-study example is presented in Section 6.4.7 where the two methods are compared at the fragility level.

Case Study Structures		①
Special-Code Building	Pre-Code Building	

↓

Analysis Fidelity		②
High Fidelity: NLTHA	Low Fidelity: FRACAS	

↓

IM – EDP Data Sets		③
Advanced IM: I_{Np}	Standard IM: $Sa(T_1)$	

↓

Generation of "Pseudo-Realities" ④

Reality 1: Favors: Cloud | Reality 2: Favors: BEA | Reality 3: Favors: None

Compute of mean and standard deviation for each pseudo-reality

Simulate 1000 IM-EDP Data sets using same mean and standard deviation

↓

Sampling Configuration ⑤

Sampling method: Stratified | Sampling method: Random

Sampling size: Full (100%) | Sampling size: Large (75%) | Sampling size: Medium (50%)

↓

Training of Tested Methods ⑥

BEA Method | Cloud Method

Select Covariance Structure

↓

Results of Comparison – Evaluation against "Pseudo-Realities" ⑦

Coverage Probability 95% | Mean Squared Error | Average Length

Figure 6.1 - Schematics of the methodology roadmap.

6.3.1 Analysis setup

In this Section, two regular RC MRFs, representing a Pre- and Special-code vulnerability classes are employed as case studies. These structures are analysed at two levels of fidelity, namely low fidelity and high fidelity (using FRACAS and NDA respectively), to estimate their associated seismic response. The detailed description of the analysis procedure is presented in Chapters 3 and 4.

The results obtained from the various analyses are then expressed in data sets of IM and EDP. In each data set, which were generated using the 150 strongest ground motions of SIMBAD database, records that did not manage to push the structures to the nonlinear range are discarded. This is because FRACAS, as a CSM-based method, draws a direct correlation between the seismic intensity and the estimated demand parameters within the elastic regime; therefore, bias is introduced (Minas et al. 2015). For consistency, the same records are also discarded from the NDA analysis. After this amendment, the actual number of ground motions that pushed the frame into the nonlinear range is relatively small but still significant, corresponding to 15% and 25% of the total number of records used, for the case of Special- and Pre-code buildings respectively (i.e. 68 and 108 GMs). **Table 6-1** - Number of ground motions used for analysis of the case-study buildings based on sample size. summarizes the number of ground motion records used for analysis of the case-study buildings based on three sample sizes, namely full-, large- and medium-sample.

Table 6-1 - Number of ground motions used for analysis of the case-study buildings based on sample size.

Building Type	Ground motion sample size		
	Full sample (100%)	Large sample (75%)	Medium sample (50%)
Pre-code	108	81	54
Special-code	68	51	34

In this Section, the deformation-based EDP , MIDR is adopted. With regard to IM inputs, two scalar IMs are used herein, namely the spectral ordinate $S_a(T_1)$ and the advanced IM I_{Np} (Bojórquez and Iervolino 2011). The selection of the adopted IMs is based on the outcome of Chapter 5 regarding the selection of optimal IMs for the fragility assessment of mid-rise RC buildings.

6.3.2 Defining pseudo-realities

Three different scenarios are considered herein, each corresponding to a different "reality" represented by an artificially generated EDP|IM relationship, constructed as described below. This pseudo-reality approach is used to provide an experimental setup that mimics real applications as closely as possible. In each relationship, the expected value of some function of the EDP is given by a mean function $\mu(t)$ as in equation (6.1); where $\mu(t)$ is the sum of the deterministic mean function $m(t)$ and an approximation error $z(t)$:

$$\mu(t) = m(t) + z(t)$$

$$(6.4)$$

The deterministic mean function $m(t)$ is represented by a simple regression model, which is trained using the sets of EDP|IM data generated as described in Section 6.3.1. Standard cloud analysis corresponds to a situation in which $\mu(t)$ is the expected value of $\ln(EDP)$, the mean function $m(t)$ is a linear function of $\ln(t)$ and the approximation error $z(t)$ is zero. However, the framework considered here allows exploration of the capability of the tested methods to predict EDP|IM relationship under favourable or less favourable situations.

Our three chosen "pseudo-realities" are as follows:

a) Reality 1: the expected value of $\ln(EDP)$ is an exact linear function of $\ln(IM)$ with no approximation error (i.e. with $z(t) = 0$ in equation (6.4)). This "pseudo-reality" is exactly the situation for which the cloud method is designed. The coefficients in the linear function are derived from a sample of real analysis data, as described in Section

6.3.1: this ensures that as a test case, this represents the kind of scenario that might conceivably be encountered in a real earthquake engineering situation.

b) Reality 2: this is as Reality 1 except that a non-zero approximation error $z(t)$ is included in equation (6.4) so that $\mu(t)$ is no longer an exact linear function of $\ln(t)$. The approximation error is generated as a Gaussian random field (but making sure that the resultant "true" function is monotonically increasing) – again, derived from the real analyses of Section 6.3.1. The choice of this "reality" favours the BEA methodology.

c) Reality 3: the function $\mu(t)$ is constructed in the same way as for Reality 2, but now it represents the expected value of EDP rather than $\ln(EDP)$. This case can be used as an example where both models (cloud and BEA) are wrong, and will shed light on the capability of these two models to estimate the EDP|IM relationship.

As an example, the three pseudo-realities generated for the Special-code building analysed at high fidelity is illustrated in **Figure 6.2**.
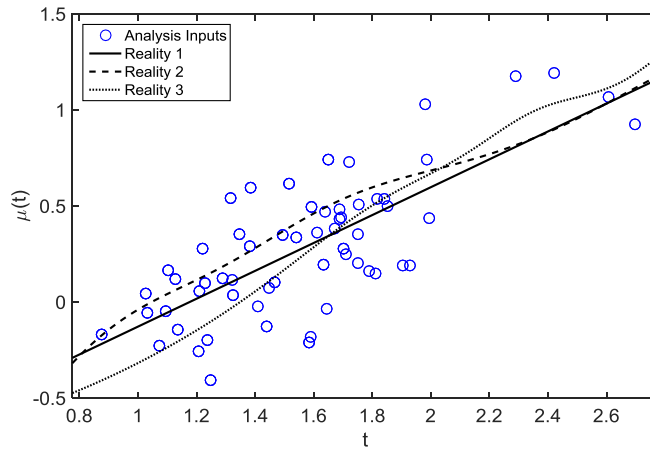
Figure 6.2 - Three artificially generated "pseudo-realities" from the real high-fidelity analyses (analysis inputs) of the Special-code case study.

For each "pseudo-reality", an artificial analysis dataset can be simulated by adding random "residuals" to the fitted mean curve, at the given IM locations of the original analysis set. In each simulation experiment reported below, 1000 such datasets were generated. A subset of the generated EDPs is utilized to calibrate both the cloud and BEA methods. These approaches are

then assessed for their ability to predict the target "pseudo-reality". The predictive performance of the tested methods is not only assessed in terms of estimation of the mean "pseudo-reality", but also in terms of evaluating the uncertainty of the entire distribution, as both parameters are required for the fragility analysis. As well as predicting the mean "pseudo-reality" function itself, 95% prediction intervals are computed as a way of quantifying the uncertainty in the predictions.

The coverage probability of the 95% prediction intervals (hereafter called coverage), the average length of these intervals and the mean squared error (MSE) are the metrics used to assess the predictive performance of the BEA and cloud approaches. The coverage measures the accuracy of the uncertainty assessment, by calculating the proportion of intervals containing the actual EDP. If the uncertainty assessment is accurate, the coverages from a method should be equal to their nominal value of 0.95. The average interval length assesses the amount of uncertainty in the prediction intervals and, for any collection of n prediction intervals, is computed as:

$$\text{Average Length} = \sum_{i=1}^{n} \left(\frac{UL95_i - LL95_i}{n} \right) \tag{6.5}$$

where $UL95_i$ and $LL95_i$ are respectively the upper and lower limits of the ith interval. Ideally, prediction intervals should be as small as possible subject to the correct coverage.

The MSE assesses empirically the quality of the vector of n predictions, $\hat{G}$, with respect to a vector of simulated EDPs, G, and is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{G}_i - G_i \right)^2 \tag{6.6}$$

6.3.3   Sampling configuration – Training of the BEA and the cloud method

The predictive ability of the BEA and cloud approaches are checked when different sampling configurations are adopted for the selection of EDP|IM training data sets.

Two sampling configurations are considered here, one focusing on the sample size and one on the sampling method. With regard to the sample size, three different cases are investigated, one using the full training input (N and M samples respectively for the Pre-code and Special-code buildings), and two utilizing large and medium training samples, which correspond to 75% and 50% of the full samples respectively. The use of smaller sample sizes results in a reduction of the computational time. This is particularly true for high fidelity analyses, which each have a runtime of between a few minutes to two hours of CPU time on a personal computer with a quad-core processor (2.70GHz – 16GB RAM) running on a single core, depending on the complexity of the structural model and the duration of the GM.

Two sampling methods are considered in determining the reduced training data sets from the full sample: random sampling, and stratified sampling. For the stratified sampling, the full input sample is divided into 5 strata (bins) of equal intensity measure width, and then random sampling is applied within each stratum to select the user-specified number of inputs.

Following the sampling procedure, the selected inputs are used to train the cloud and BEA. In the case of BEA, as stated in Section 6.2, the setup of the emulation model allows one to implement a regression model to describe the mean function $\mu(t)$ and is suitable for the nature of the problem of interest. In this study, a power-law regression model is adopted to enable a comparison with the cloud method. However, alternative models, such as linear or higher order polynomials may also be used.

The results of this exercise shed light on the capability of the BEA to compute the probabilistic seismic response of buildings using a reduced number of analyses. This is a first step towards the development of an optimal experimental design for BEA, which will be a subject of future work.

6.4   Results of comparison – Evaluation against pseudo-realities

All the steps discussed in the Section 6.2 are applied here in order to build the BEA. BEA accounts for the uncertainty of the model's parameters, and provides estimates, and the associated variances, for the conditional distribution of the EDP|IM relationship. The outcomes of the BEA approach are here compared to the cloud method outcomes. A series of tests, which are described in detail in Section 6.3, are carried out to facilitate the evaluation process, highlighting the sensitivity of the BEA to several input assumptions and revealing the range of applicability of the approach proposed herein. In addition to these tests, the sensitivity of BEA predictions to the choice of the covariance structure is explored. Last, a Section on fragility analysis using BEA is presented.

6.4.1   Performance assessment for the 2 case study buildings analysed using NDA

In the first test-case investigated, the high fidelity analysis data sets, expressed in terms of MIDR and I_{Np}, are used to train both the BEA and the cloud method models. Following the procedure described in the methodology Section, the two data sets, corresponding to Pre- and Special-code buildings respectively, are sampled to three sample subsets following the stratified sampling process. Based on these test samples, the BEA and the cloud processes are used to predict the "three assumed realities". The performance of the emulation and cloud approach is then assessed using the coverage probability, MSE and average interval length metrics. A detailed summary of results for this test-case is presented in **Table E 1** in the Appendix E. The first metric to check is the coverage, which in both the cloud and the BEA case closely matches the nominal coverage probability of 95% (never more than +1.1% difference). Having this condition satisfied, next step is to assess the performance of the two tested approaches in terms of MSE and average length, as shown in **Figure 6.3**.

The MSE estimates for the BEA are smaller than those for the Cloud approach for "pseudo-realities 2 and 3", whilst they are comparable for "pseudo-reality 1". This observation highlights the superiority of the BEA method over the Cloud method in predicting the mean EDP|IM

function, under different "pseudo-reality" scenarios. The same observation also applies to the average length estimates, although the improvement in the BEA's performance is not as significant as in the case of the MSE metric.

With regard to the training subsets, it is observed that the quality of predictions of the mean depends, as expected, on the size of the training subsets. In particular, the performance of the BEA and the cloud models becomes less accurate as the training subset size reduces. However, the same observations discussed above are also present in all instances of reduced training subsets, i.e. the BEA shows improved performance with respect to cloud.

In some cases, it is observed that the BEA can generate predictions that better capture the mean and maintain the same level of accuracy as the cloud method, but requiring a smaller training subset size. For example, though to a lesser extent, when BEA is trained with a large subset of the Special-code data set, it provides better MSE estimates, while matching the level of accuracy resulting from using the cloud method with the full training set (e.g., **Figure 6.3** top panels, **Table E 1**). This reduction in 25% of the required NDAs translates into significant savings in computational time.

Similar benefits of BEA over the cloud method are also seen in the case of the Pre-code building, but are not presented here for brevity.

6.4.2   Sensitivity to analysis fidelity level

The tests presented in Section 6.4.1 show that the BEA performs well for high fidelity data, at times outperforming the cloud approach. In this Section, the BEA is tested for its performance when low fidelity analysis data are used. The use of low fidelity data for fragility analysis is appealing as it significantly reduces the analysis time. However, this is at the expense of accuracy when compared to the high fidelity data set. **Figure 6.4** and **Table E 2** are generated following the same procedures discussed in the Section 6.4.1.

Similar observations to those made for the high fidelity analysis case are made. The coverage closely matches the 95% threshold, (±1.9%, see **Table E 2**). Improvement of MSE estimates is observed when the BEA is employed compared to cloud, especially for "pseudo-realities 2 and 3". In the case of Special-code building, this improvement can be translated to BEA MSE estimates reduced up to 21.65% and 30.68% for "pseudo-realities 2 and 3", respectively. A substantial reduction of average length is also observed when comparing BEA and cloud estimates. This reduction ranges between 4.73% to 8.37%, for "pseudo-realities 2 and 3" for the case of Special-code building, as shown in the **Figure 6.4** top panels. The aforementioned cases also show the capability of the BEA, under certain conditions, to provide predictions that significantly reduce MSE and average length estimates, even when utilizing smaller training sample sizes. With regard to "pseudo-reality 1", BEA again produces similar estimates to the cloud, which is the optimal method for this scenario. Improvements can be also observed in the case of the Pre-code building, however, as in the case of the high fidelity sample, the benefits are less. It is noteworthy, that BEA always outperforms cloud in terms of the MSE and average length estimates when using medium training subsets. However, it is known from previous studies, that cloud method is not recommended to be used with small data sets.

6.4.3 Sensitivity to selection of IM

Previous studies have shown that advanced IMs (such as I_{Np}) appear to be better correlated to displacement-based EDPs (Ebrahimian et al. 2015; Minas et al. 2015) resulting in EDP|IM data sets that have less scatter. In this test-case, the performance of the BEA is investigated when trained with a more scattered input data set, due to the use of the conventional IM, $S_a(T_1)$, for both high and low fidelity analysis data.

The outcomes of this test can be visually inspected in **Figure 6.5**, while a detailed summary of results is provided in **Table E 3**. For the sake of brevity, only the results for the Special-code case-building are shown here. However, the same process is applied to the Pre-code building and similar trends are observed but are seen to be less significant than for the Special-code.
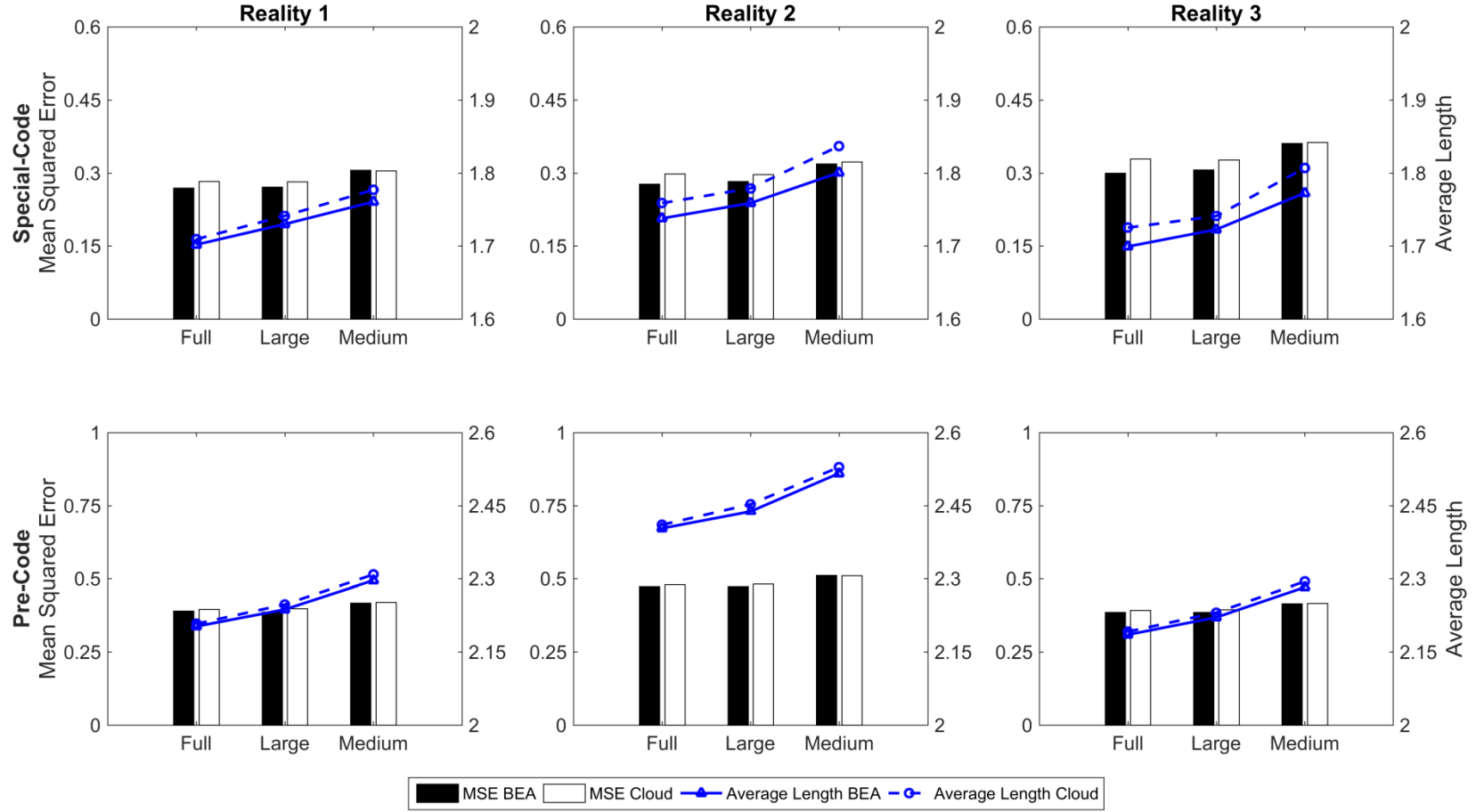
Figure 6.3 - Comparison of BEA and Cloud method predictions in terms of Mean Squared Error (left y-axis of each graph) and Average Length (right y-axis of each graph) using different sample sizes. Each column represents the different "pseudo-realities", and each row the tested building types: Special-code (top panels) and Pre-code (bottom panels). In this test-case high fidelity analysis is used alongside with advanced IM I_{Np} and stratified sampling process.
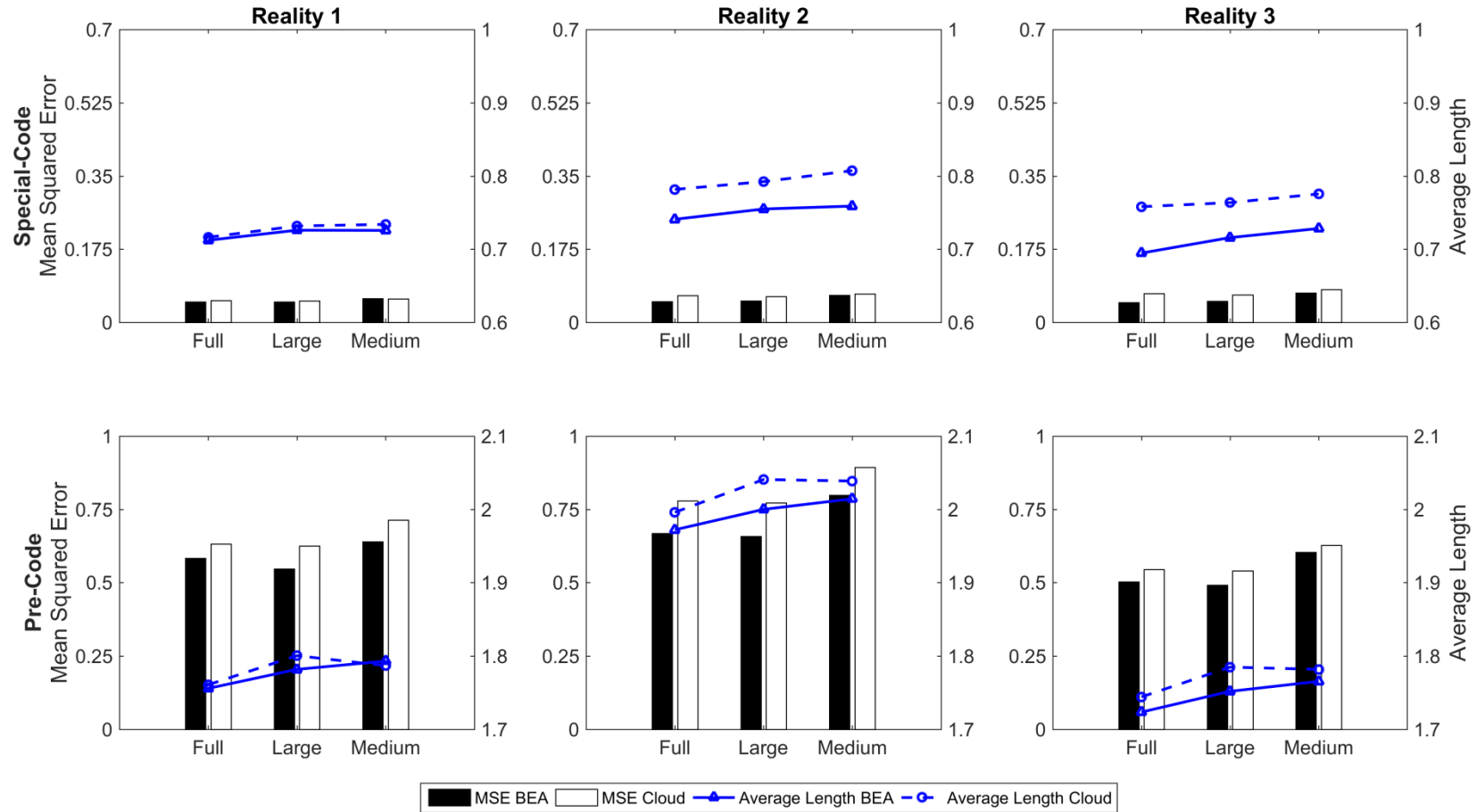
Figure 6.4 - Comparison of BEA and Cloud method predictions in terms of Mean Squared Error (left y-axis of each graph) and Average Length (right y-axis of each graph) using different sample sizes. Each column represents the different "pseudo-realities", and each row the tested building types: Special-code (top panels) and Pre-code (bottom panels). In this test-case low fidelity analysis is used alongside with advanced IM I_{Np} and stratified sampling process.

As before, both approaches satisfy the coverage check. A close observation of the results reveals that the superiority of the BEA over cloud method in capturing the mean (i.e. considerable reduction of the BEA MSE estimates as shown in **Table E 3**) for both high and low fidelity inputs, even when $S_a(T_1)$, is used. The only exceptions are seen to be the cases of the medium sized high and low fidelity samples for "pseudo-reality 2" (**Figure 6.5**, middle panels), where cloud MSE predictions are better than those of BEA. The average length obtained from the BEA is always narrower, excluding the case mentioned above. Furthermore, BEA appears to perform significantly better when trained with low fidelity inputs, for "pseudo-realities 2 and 3".

6.4.4 Sensitivity to the sampling approach

The same case study arrangement as in Section 6.4.3 is used here in order to investigate the effect the random sampling procedure. Following a visual inspection of **Figure 6.6**, it can be concluded that the observations between the BEA and cloud estimates are essentially consistent with what it is shown in Section 6.4.3. Nevertheless, it is interesting to evaluate directly the sensitivity of predictions when using random sampling instead of stratified sampling.

The results presented in **Table E 3** and **Table E 4** reveal that BEA predictions (as well as cloud predictions) show some improvement in both MSE and average length when random sampling is used instead of stratified sampling. Regarding the comparison of BEA and cloud performance metrics, it is observed that the choice of sampling does not have significant effects on capturing the mean function, resulting in similar differences for all fidelity levels and "pseudo-realities". The BEA with random sampling, maintains its superiority over cloud method even for the low fidelity cases using medium sample sizes. Moreover, random sampling shows slight improvement in terms of average length metric, resulting narrower lengths comparing to ones calculated using stratified sampling.

In theory, one would expect stratified sampling to perform better than random sampling, as the former approach ensures that all parts of the IM range are sampled reasonably well and hence to capture local-scale features in the predicted EDP|IM curves if present. In our pseudo-realities

however, there are no such prominent local features so, at first sight, it appears that stratified sampling has few advantages. However, stratified sampling does have an additional advantage that is not quantified by the metrics described above. Specifically, we find that stratified sampling ensures monotonicity of the predicted EDP|IM curve in the vast majority of the cases. In contrast, random sampling is more likely to provide a non-monotonically increasing predicted EDP|IM curve, especially for the cases where reduced training samples are utilized. The proportion of non-monotonically increasing predicted EDP|IM curves obtained from random sampling can reach up to 15%, for the case of reduced samples (i.e. large and medium) expressed in terms of I_{Np}. This is high compared to the proportion for stratified sampling, which is less than 4%.

6.4.5 Sensitivity of BEA to the selection of covariance model

The last test conducted as part of the BEA evaluation process is to investigate the sensitivity of BEA to the selection of the covariance model. It is reminded that the Gaussian covariance model is used to produce the "pseudo-realities 2 and 3", however, the user does not know this in practice when choosing the covariance model for BEA. Thus, this test is partly designed to explore the effect of choosing a wrong covariance model. For the sake of consistency, the Special-code building is used here alongside with the IM $S_a(T_1)$, high fidelity analysis and stratified sampling approach.

The Gaussian model is the default choice ensuring the EDP|IM curves are very smooth (infinitely differentiable). A number of commonly used parametric models are studied here to explore the performance of the BEA when the wrong covariance model is used. Specifically, the 9 covariance models tested herein are listed below:

- Gaussian   - Cubic   - Gneiting
- Matérn   - Exponential   - Circular
- Spherical   - Cauchy   - Pure-nugget

The above mentioned models, which are available within the cov.spatial function of the geoR package, are described in detail in its documentation (Ribeiro Jr and Diggle 2001).
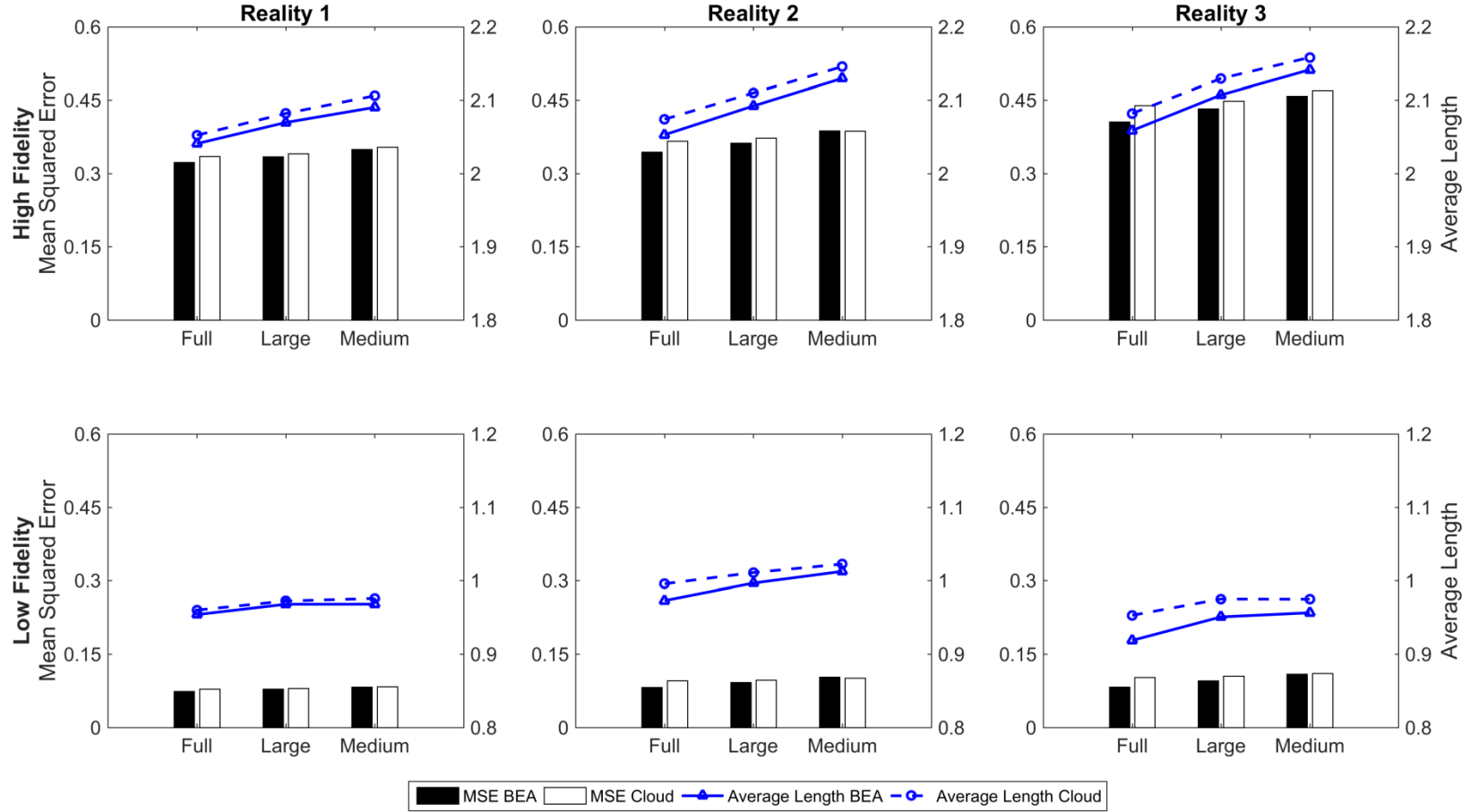
Figure 6.5 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error (left y-axis of each graph) and Average Length (right y-axis of each graph) using different sample sizes. Each column represents the different "pseudo-realities", and each row the level of analysis fidelity: High Fidelity (top panels) and Low Fidelity (bottom panels). In this test-case Special-code building is used alongside with the IM $Sa(T_1)$ and stratified sampling process.
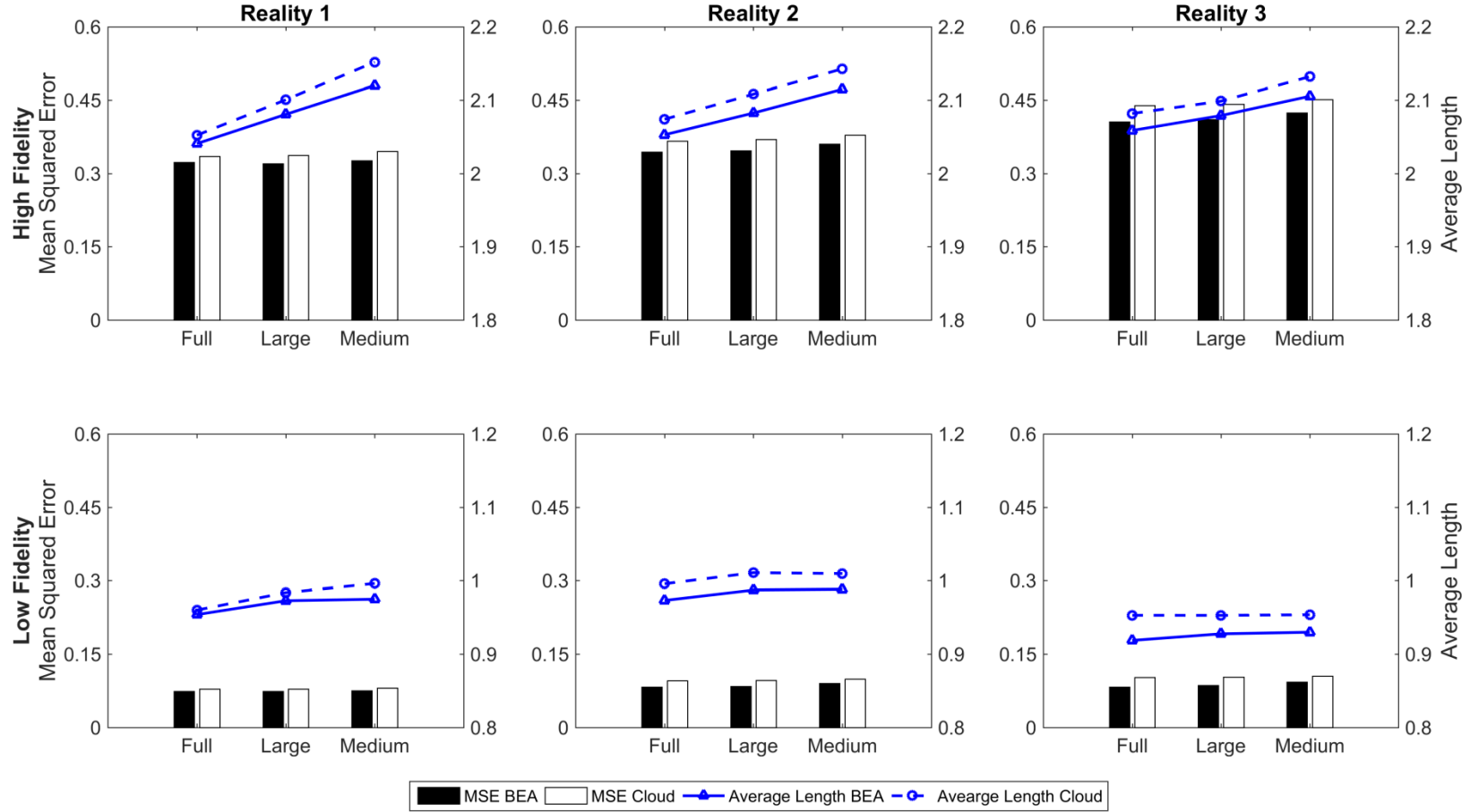
Figure 6.6 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error (left y-axis of each graph) and Average Length (right y-axis of each graph) using different sample sizes. Each column represents the different "pseudo-realities", and each row the level of analysis fidelity: High Fidelity (top panels) and Low Fidelity (bottom panels). In this test-case Special-code building is used alongside with the IM $Sa(T_1)$ and random sampling process.

Figure 6.7 shows as a bar chart the BEA's predictions when using the various covariance models assessed in terms of the MSE and the average length metrics (left and right panel respectively) for different sample subsets. The associated predictions for the cloud method, illustrated as dashed lines, are also provided as a point reference. A detailed summary of the results is presented in **Table E 5**. The coverage probability for all tested covariance structures matches or slightly exceeds the nominal coverage probability.
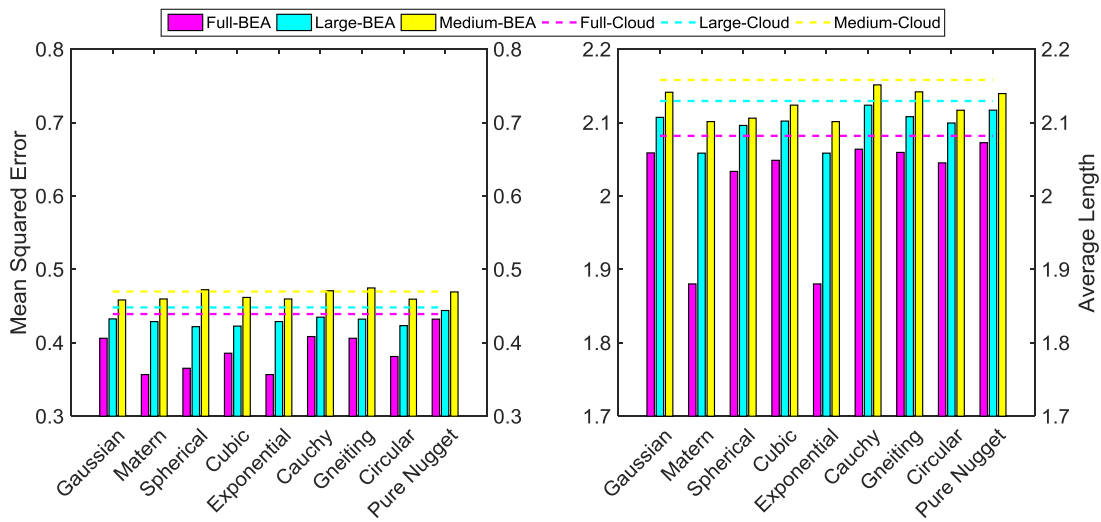
Figure 6.7 - Sensitivity of BEA predictions when using different covariance models and comparison with cloud method predictions, in terms of mean squared error (left panel) and average length (right panel).

Close inspection of **Figure 6.7** reveals that irrespective of the choice of the covariance model, the BEA always outperforms the cloud in both MSE (significant improvements) and average length (slight improvements). Regarding the average length predictions (**Figure 6.7**, right panel), it is observed that Matérn and Exponential models perform better than remaining models, especially for the full and large sample cases. Gaussian, Spherical, Cubic, Gneiting, Circular and Pure nugget models perform effectively in a consistent manner, regardless of the subset size. The worst performance is recorded for the Cauchy model, where the resultant predictions are similar to the cloud's predictions.

The most favorable models for reducing the average length, namely Matérn and Exponential, are also seen to be the two best options for estimating the MSE (**Figure 6.7**, left panel) for all training subsets. Circular, Cauchy and Gneiting models using medium subset are characterized

by poor performance. The rest of the tested models perform well in reducing the MSE under all sampling sizes, showing small differences between them.

This test shows that the BEA predictions are sensitive to the selection of different covariance models. Matérn and Exponential models result in significant reductions in the estimates of both MSE and average length. Gaussian and the remaining models are a good compromise in improving both MSE and average length estimates. Cauchy and Pure Nugget covariance models are not recommended.

However, another consideration when choosing a covariance model is that the resultant mean estimate curve should be monotonically increasing in order to have a physical meaning. To this aim, an illustrative example is presented in **Figure 6.8**, where BEA mean estimates are obtained using three different covariance structures, namely Gaussian, Matérn and Cubic models. For this example, each version BEA is trained using a large training subset as obtained from the Special-code building analysed at low fidelity level.

The results obtained from this example are in a good agreement with the tests presented in **Figure 6.7**, where Matérn and Cubic covariance structures contribute to a significant reduction of both the MSE and the average length estimates comparing to the cloud method. However, **Figure 6.8** also reveals the oversensitivity of the latter covariance structures at locations where training data are sparse, i.e. middle range of IMs.

This oversensitivity is translated into a 'jump' of the curves (also present in the case of Gaussian model) followed by a substantial 'drop' (**Figure 6.8**, middle and right panels), resulting in mean estimates that are no longer monotonically increasing. This limitation suggests that the Gaussian model is the most appropriate to describe the covariance structures for problems of this nature. Further investigation is required to explore the applicability of the different covariance structures.
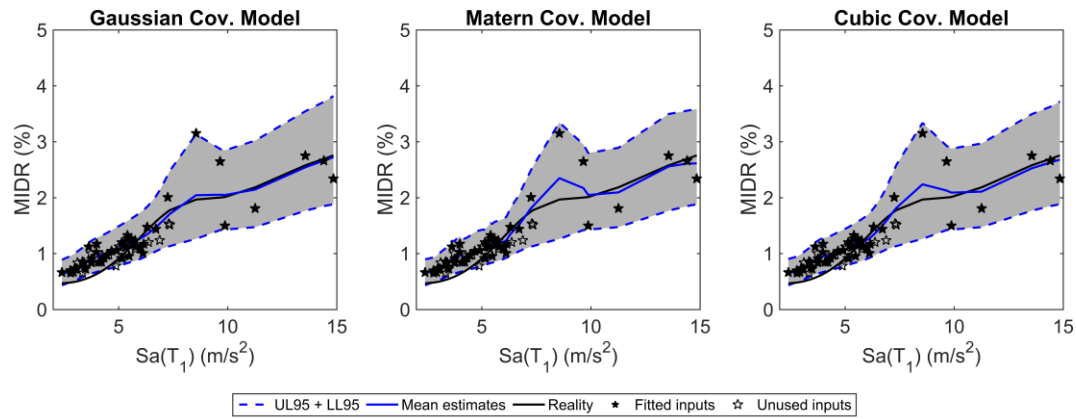
Figure 6.8 - BEA mean estimates using Gaussian (left panel), Matérn (middle panel) and Cubic covariance models (right panel).

Table 6-2 illustrates the sensitivity of BEA predictions to the choice of the covariance structure,

and how these compare to the cloud predictions.

Table 6-2 - Predictions of BEA for different covariance structures and comparison with cloud method predictions in terms of MSE, Average Length and Coverage probability, for the case of a single case-study building. In this test-case Special-code building is used alongside with the IM I_{Np} and stratified sampling process.

Covariance Model	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
Gaussian	0.039 (0.044)	10.42	0.814 (0.828)	1.71	98.41	98.41
Matérn	0.042 (0.044)	4.17	0.756 (0.828)	8.70	100	98.41
Cubic	0.040 (0.044)	8.81	0.783 (0.828)	5.41	100	98.41

6.4.6 Discussion - Case-study example utilizing original analysis data set

The results reported in Sections 6.4.1 to 6.4.5 were averaged over 1000 simulation data sets,

however, it is important to illustrate what those average performance measures mean in practice

for a single data set. To this aim, the case study introduced in Section 6.4.5 is studied in detail to

benchmark the capability of the BEA to estimate the mean and the associated variance, for the

conditional distribution of EDP|IM relationship.

The Gaussian model is employed to describe the covariance structure of the BEA. The resultant data outputs are then used to train both the BEA and cloud approach, in order to predict the "pseudo-reality 3". For the sake of brevity, only one sampling-case is shown per sample size. The reason for choosing this case-study example is because it depicts some of the capabilities of the BEA, and its superiority over the cloud method, following the observations discussed in Sections 6.4.1 to 6.4.5. However, consistent observations arise when testing alternative case studies.

Figure 6.9 shows the mean estimates and the 95% confidence intervals for both the BEA and the cloud, when employing three different sampling sizes. **Table 6-3** shows the sample size, the metrics used to assess the performance of the BEA and the cloud approach as well as the respective differences, for the three different sample sizes investigated. The coverage probability for both the BEA and the cloud approach is always exceeding the nominal coverage probability.

Table 6-3 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error, Average Length and Coverage probability, for the case of a single case-study building. In this test-case Special-code building is used alongside with the IM I_{Np} and stratified sampling process.

Sample Size	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
Full	0.039 (0.042)	7.14	0.743 (0.748)	0.67	98.41	95.24
Large	0.039 (0.044)	10.42	0.814 (0.828)	1.71	98.41	98.41
Medium	0.051 (0.055)	5.74	0.922 (0.948)	2.74	98.41	98.41

A visual inspection of **Figure 6.9** reveals the capability of the BEA to identify the 'jump' occurring in the mid-range of the "pseudo-reality". This "jump" cannot be captured by the cloud method due to the fixed mathematical model used by the latter approach. The flexibility of the BEA is also highlighted by the reduction in MSE, which ranges between 5.7-10.4%, depending

on the training subset used. Regarding, the average length metric, small improvements are also observed when using the BEA, and can be translated to 0.7-2.7% difference comparing to cloud's average length.
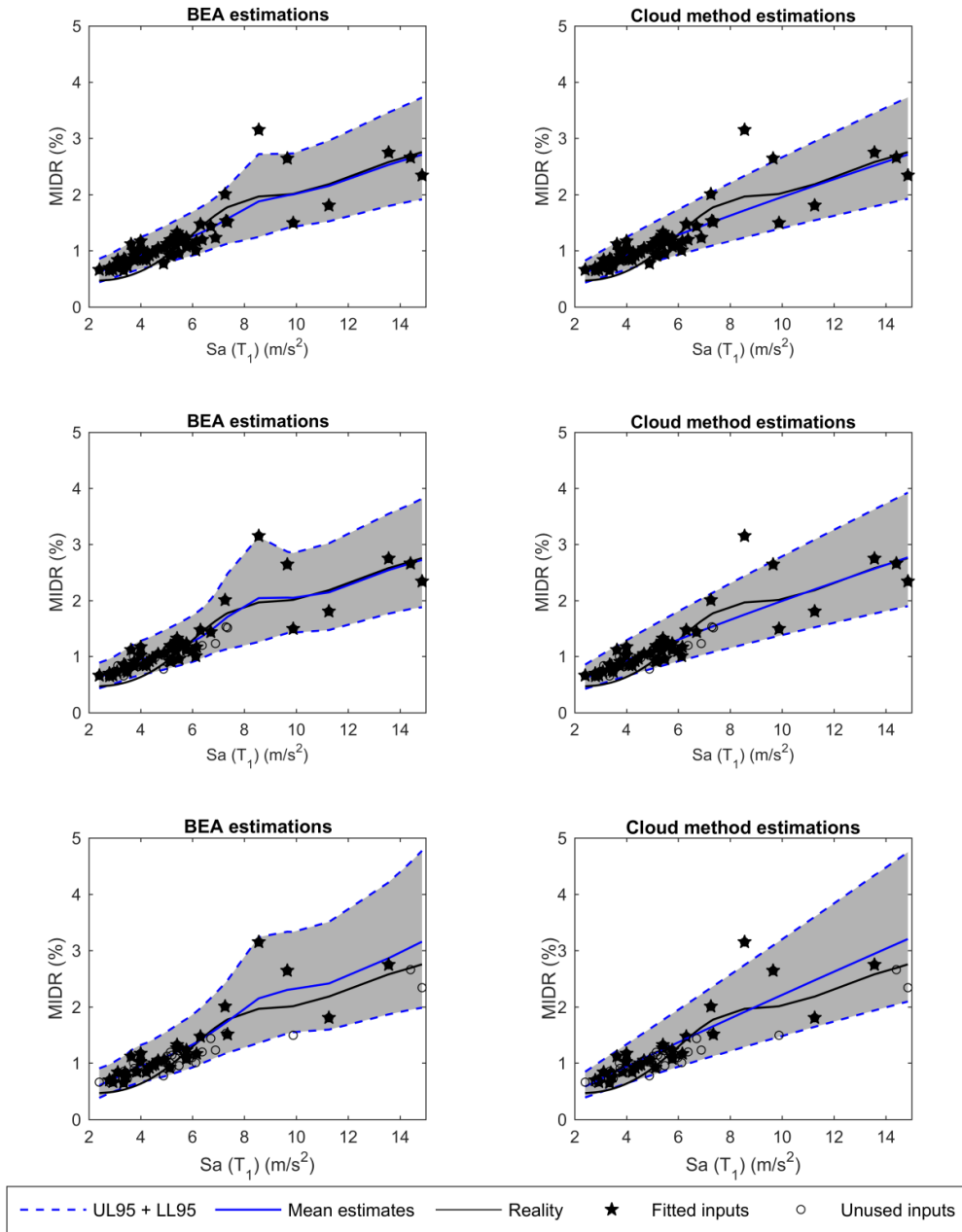
Figure 6.9 - Mean estimations and associated 95% confidence interval of BEA approach (left panels) and the cloud method for full (top row), large (mid row) and medium sample size (bottom row). Case study building: Special-code; Analysis fidelity: Low; Reality: 3; IM: INp; Sampling method: Stratified sampling; Covariance model: Gaussian.

6.4.7   Fragility analysis using BEA

In this Section, the process for deriving fragility functions using BEA is presented. A fragility function represents the probability that a *EDP* will exceed a certain level of capacity damage state DS_i given an earthquake intensity *IM* . For a traditional cloud analysis this probability is usually derived under the assumption that *EDP* is a lognormal variable, and the parameters of its distribution are fixed at the best estimates obtained from the predicted EDP|IM curves (e.g., (Jalayer and Cornell 2009)):

$$P(EDP > DS_i \mid IM) = 1 - \Phi\left(\frac{\ln DS_i - \ln \mu_{EDP|IM}}{\beta_{\ln EDP|IM}}\right) \tag{6.7}$$

where Φ is the standardized Gaussian CDF, $\mu_{EDP|IM}$ and $\beta_{\ln EDP|IM}$ are the estimated median and the standard deviation of $\ln(EDP)$ given *IM* . The BEA can improve upon this, however, both in its more flexible representation of the predicted EDP|IM curve itself and in its ability to account for parameter uncertainty. As described in Section 6.2, in the Bayesian setting the fragility for a given IM is just the posterior predictive probability of exceeding DS_i , and can be obtained directly from the output of the krige.bayes() function using the geoR package.

The case-study introduced in Section 6.4.5 is used as an illustrative example for the comparison of BEA and cloud method at the fragility level. The damage scale (description and damage thresholds) used herein is discussed in detail in Chapter 3 (Section 3.3).

In consistency with the process described in Section 6.2, a large number of EDPs is simulated (1000 simulations) at the given locations of IM, using the same mean and standard deviation values as computed for each "pseudo-reality". The suite of the 1000 fragility functions is then generated for both BEA and cloud method (median and 16-84th percentiles), and compared to the median "real fragility" as shown in **Figure 6.10**.

In **Figure 6.10** an apparent superiority of BEA over the cloud method in predicting the fragility is observed. Specifically, the median predictions of BEA match almost perfectly the "real

fragility" for both DSs, even when 25% fewer analyses are used to train BEA (large and full samples are used for BEA and cloud respectively). This practically means that BEA not only improves the accuracy but also captures the variability of the fragility curves, showing varied widths in the 16th and 84th percentile bounds across the IM range, as illustrated in **Figure 6.10**. It is highlighted that the third DS (DS3) is not reached for the Special-code building using the chosen suite of ground motions.
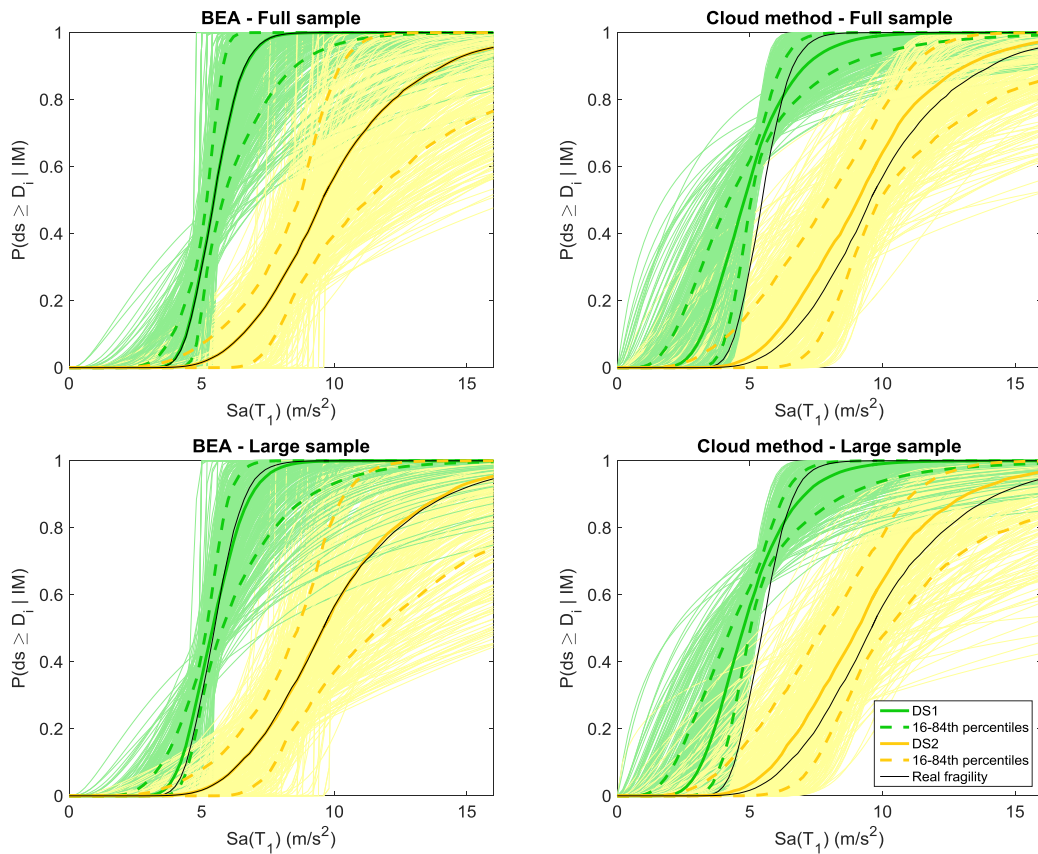
Figure 6.10 - A suite of 1000 fragility curves (median and 16-84th percentile) generated from sets of simulated analyses, for BEA (left panels) and cloud method (right panels), and comparison with median "real fragility", utilizing: full sample size (top panels) and large sample size (bottom panels).

Similar observations are essentially obtained for "pseudo-realities 2 and 3" of all tested case-studies, while both BEA and cloud manage to predict well the "real fragility" in all the scenarios related to "pseudo-reality 1".

6.5 Conclusions

This Chapter presents a new Bayesian emulation approach, which is a version of kriging, for estimating the mean and the associated variance of the conditional distribution of EDP|IM relationship, accounting for the uncertainty of the BEA model's input parameters. The capabilities of this new approach, which overcomes some of the limitations of existing approaches, are investigated when trying to predict different "pseudo-realities" and the results are compared against standard cloud analysis.

Two RC mid-rise building representing the Special- and Pre-code vulnerability classes of the Italian building stock are used as case-study structures. These structures are analysed at two levels of fidelity, namely high and low fidelity, and their seismic performance is expressed as a function of standard and advanced IMs, i.e. $S_a(T_1)$ and I_{Np} respectively. Three "pseudo-realities" (based on different mathematical models) are fitted to each of the resultant 8 data sets, and are then used as a target to evaluate the performance of new emulation-based approach. Next, 1000 artificial sets of EDPs are generated at given IM levels utilizing each "pseudo-reality" as a mean function and the standard deviation as calculated from the sample of real analysis data. MSE, average length and coverage probability are the metrics used to assess the performance of the BEA.

A number of tests are conducted to highlight the sensitivity of the proposed emulation-based method to several input assumptions and show how this method compares to the cloud approach. The main conclusions of this study can be summarized as follows:

The BEA has this ability to better capture the mean function of each "pseudo-reality" for all tested case studies. This observation is clearly reflected in the MSE predictions of the BEA, which are smaller than those of the cloud analysis for all tested cases. A reduction of the average length also results when the BEA is used compared to the cloud method. This minor reduction, can be translated as a reduction of the uncertainty of the resultant predictions. BEA outperforms cloud for both metrics in the vast majority of cases. However, more significant

improvements are obtained for the "pseudo-realities 2 and 3". Regarding, "pseudo-reality 1", the cloud method is known to be optimal, and therefore the results presented here suggest that BEA is indistinguishable from the optimal method in this situation.

The BEA can essentially result in more accurate mean predictions than cloud method with less uncertainty. The best estimates of the BEA are attained when combined with Special-code vulnerability class, low fidelity analysis and advanced IM I_{Np}.

Under certain conditions, the BEA can reduce the overall computation time by decreasing the analysis inputs required for training. In particular, the BEA can generate predictions that better capture the mean function, closely matching the level of accuracy as in the cloud method, but requiring a smaller number of training inputs. This capability of the BEA is mainly observed when a large training subset is combined with the favourable inputs mentioned above (namely Special-code vulnerability class, low fidelity analysis and advanced IM I_{Np}).

The estimates of both the BEA and the cloud approach are, as expected sensitive to the size of the training subsets. Smaller training subsets, result in worse estimates for both MSE and average length metrics. The use of medium subset resulted in some cases non reliable outputs, therefore it is not recommended to be used for the training of the BEA.

The choice of sampling approach does not yield significant differences in the resultant MSE and average length metrics, however, stratified sampling is a more suitable sampling method, comparing to random sampling, for the experimental design of the BEA. This is because stratified sampling is better able to capture local features in the predicted EDP|IM curve, and is also better able to preserve important global features such as monotonicity.

Some sensitivity is witnessed when different parametric models are used to describe the covariance structure of BEA. Gaussian model constitutes a good all-around choice for improving both MSE and average length estimates but also it ensures the smoothness of the GP realizations, which is essential for the nature of this problem. Finally, BEA drastically improves

the prediction of the median fragility for all DSs, matching perfectly the "real fragility" in most cases, and also capturing realistically the variability of the fragility curves at a smaller computational expense.

Overall, although BEA is more complex as compared to the cloud method, it presents significant advantages over the latter for predicting nonlinear realities and for representing uncertainty. It is observed to perform well when high fidelity analysis and larger data samples are used. However, it also outperforms the cloud method when predicting nonlinear realities with smaller data samples. The cloud method remains a suitable option when predicting a linear EDP|IM relationship (linear reality).

7. Probabilistic regional seismic loss assessment considering hazard and vulnerability modelling options

7.1 Introduction

Previous Chapters of this Thesis shed light on some of the shortcoming of existing seismic risk assessment approaches, and proposed new advancements and techniques focusing particularly on the improvement of the fragility/vulnerability component. The impact of the different modelling options within the regional seismic loss assessment framework is explored here considering a fictitious building portfolio which represents the town of Avellino in South Italy. Understanding the impact of the different components in the resultant loss estimates will contribute towards the development of a more robust seismic risk framework, and will also serve as a guide to prioritize the potential improvements of future catastrophe models. This Chapter therefore contributes to answering Research Question 4.

In order to investigate the sensitivity of probabilistic regional seismic loss assessment of RC buildings to different hazard and vulnerability modelling options, this study systematically assesses the sensitivity of losses to (1) ground motion intensity characterization (through selection of different types of IMs), (2) spatial scale resolution (considering two levels of spatial resolution), (3) ground motion correlation models (including spatial, spectral and cross-correlation models), (4) assumptions in structural analysis (NSPs versus NDA)and (5) fragility derivation methods (standard cloud method and BEA). The resultant loss exceedance curves and average annualized losses (AAL) for all the tested combinations are compared, and recommendations are drawn regarding the optimal modelling approaches.

7.2 Probabilistic seismic loss assessment methodology

The probabilistic loss assessment of spatially distributed building portfolios requires the definition of hazard, exposure and fragility/vulnerability components. The following subsection

addresses the methodology used in the probabilistic assessment of hazard (Section 7.2.1) and fragility/vulnerability (Section 7.2.2). The dataset of exposed buildings investigated in this study is further presented in Section 7.3.1.

7.2.1 Hazard assessment

7.2.1.1 Spatial correlation and spectral cross correlation modelling

Recent studies have demonstrated that strong ground motion intensities for a specific seismic scenario are spatially correlated (e.g., Boore et al. 2003; Sokolov et al. 2010), and also that the distances where substantial correlations may occur are usually greater for the long period characteristics of the ground motion (e.g., Jayaram and Baker 2009). Based on these points, several researchers have also attempted to create spatial cross-correlation models to spatially characterize the correlation of ground motion IMs fields (e.g., Goda and Atkinson 2009; Loth and Baker 2013). The inclusion of such models into the seismic risk framework for building portfolios has a significant impact, resulting in an increase in the median loss estimates some cases, compared to the estimates obtained when no correlation or cross-correlation models are used. More in general, considering correlation allows one to capture more extreme loss cases, both higher and lower than average.

When dealing with heterogeneous building portfolios where different IMs may be used in the fragility assessment of different building classes, the spatial cross-correlation of these IMs is typically considered in the simulation of correlated random variates assuming a joint normal distribution (Jayaram and Baker 2008). The same assumption of joint normality hold for IM values at different location (Weatherill et al. 2015).

Existing GMPEs account for the uncertainty in the estimation of IMs via the aleatory variability terms incorporated in the generic GMPE function, as shown below:

$$\log(IM_{ij}) = f(M_i, R_{ij}, \theta_{ij}) + \tau \nu_i + \sigma \varepsilon_{ij}$$

(7.1)

where IM_{ij} is the ground motion intensity value at a site j, which is situated at a distance R_{ij} from the seismic source, for an event of magnitude M_i; θ_{ij} may represent other covariates, such as the faulting types, site classes and other predictors (e.g., basin effect, etc.); and τv_i and $\sigma \varepsilon_{ij}$ components represent the inter-event and intra-event variability of the GMPE model respectively. The random variables v_i and ε_{ij} are in most cases assumed to be independent, normally distributed with zero mean and standard deviation τ and σ, respectively.

As discussed above, it is now well-established in the literature that the intra-event variability component is spatially correlated, i.e. the coefficient of correlation, $\rho_h(T)$, of the intra-event residuals at two different locations is a function of the separation distance, h, between the two locations. An exponential functional form is used in most spatial correlation models (e.g., Goda and Atkinson 2009; Jayaram and Baker 2009), as shown below:

$$\rho_h(T) = \exp\left[a(T)h^{b(T)}\right]$$

(7.2)

where $a(T)$ and $b(T)$ are period-dependent model parameters associated with the attenuation level of the spatial correlation with distance. Additionally, the intra-event residual term is also known to be cross-correlated; the coefficient of correlation of the intra-event residual term of spectral acceleration at two different periods at the same site will decay as the period difference increases (Baker and Cornell 2006; Baker and Jayaram 2008).

The standard decomposition method (Davis 1987) is used here to simulate the spatially correlated random fields of GMPE residual values that are not conditioned to any observation (e.g., those coming from recording stations). The multivariate Gaussian distributed random field, $\mathbf{Y}_G$, which is a vector consisting of a set of ground motion residual values for N sites, is assumed to follow a joint log-normal distribution, and is defined as:

$$\mathbf{Y}_G = \mu + \mathbf{LZ}$$

(7.3)

In Equation (7.3), μ is a zero N-element vector, $\mathbf{Z}$ is a vector of independent Gaussian distributed random variates, and $\mathbf{L}$ is the lower triangular matrix as obtained from Cholesky decomposition, where $\mathbf{C}$ is the N by N positive-definite correlation matrix:

$$\mathbf{C} = \mathbf{L}\mathbf{L}^T = \begin{bmatrix} 1 & \rho(h_{1,2}) & \text{L} & \rho(h_{1,N}) \\ & 1 & \text{L} & \rho(h_{2,N}) \\ & & \text{O} & \text{M} \\ sym & & & 1 \end{bmatrix} \tag{7.4}$$

In Equation (7.4), $\rho(h_{i,j})$ represents the coefficient of correlation of residuals at two locations separated by a distance $h_{i,j}$. For the case of homogenous spatially distributed building portfolios, where a single IM is used (e.g., PGA or Spectral acceleration at a single period), one can use Equation (7.3) directly to determine the IM spatial correlation at different locations. Next, the logarithmic IM is computed as shown in Equation (7.1), i.e. the residual values are multiplied by the aleatory variability term (σ), as obtained from the GMPE and added to the GMPE expected value. In the case of heterogenous building portfolios, the aforementioned approach becomes more complex, as the spatial cross-correlation structure needs to be maintained while co-simulating the Gaussian random fields of IMs. To this aim, a number of spatial cross-correlation approaches have been proposed in the existing literature, such as Oliver 2003; Iervolino et al. 2010; Loth and Baker 2013. A thorough review of such approaches can be found in Weatherill et al. (2015). This study showed that the methodologies which account for spatial cross-correlation and inter-event residual correlation are more favourable to calculate accurately the losses for a heterogeneous building portfolio. For this reason, the full-block cross-correlation (FBCC) methodology (Oliver 2003) that incorporates both spatial cross-correlation and inter-event residual correlation is used herein to represent the state-of-the-art approach. In addition to the advantages discussed above, the adoption of this approach also provides the flexibility to utilize correlation models that may encompass regional characteristics of the site of interest. The FBCC is combined with the period-dependent spatial correlation model proposed by Jayaram and Baker (2009):

$$\rho(h) = \exp\left(\frac{-3h}{b(T)}\right) \quad \text{where} \quad b = \begin{cases} \begin{cases} 8.5 + 17.2T & \text{"Case 1"} & T < 1s \\ 40.7 - 15.0T & \text{"Case 2"} & T < 1s \end{cases} \\ 22.0 + 3.7T & & T \geq 1s \end{cases} \tag{7.5}$$

where b is the spatial length scale of the correlation, which is the distance where the correlation coefficient drops below 0.05. "Case 1" refers to sites characterized by different soil conditions, while "Case 2" refers to sites that show clustering, i.e. the soil conditions are similar. With regard to the spectral correlation approach, Goda and Atkinson (2009) model is used:

$$\rho_\tau(T_1, T_2) = \frac{1}{3}\left(1 - \cos\left\{\frac{\pi}{2} - \left[\theta_1 + \theta_2 I_{T_{min} < 0.25} \times \left(\frac{T_{min}}{T_{max}}\right)^{\theta_3} \log_{10}\left(\frac{T_{min}}{0.25}\right)\right]\log_{10}\left(\frac{T_{max}}{T_{min}}\right)\right\}\right)$$
$$+ \frac{1}{3}\left\{1 + \cos\left[-1.5\log_{10}\left(\frac{T_{min}}{T_{max}}\right)\right]\right\} \tag{7.6}$$

where T_{min} and T_{max} denote the minimum and maximum periods, θ_1, θ_2 and θ_3 are the model parameters (taken as 1.374, 5.586 and 0.728 respectively), and $I_{T_{min} < 0.25}$ is an indicator function equal to one when $T_{min} < 0.25$, or equal to zero otherwise.

In this Chapter, three correlation modelling approaches are tested:

  i.   The full-block cross-correlation (FBCC) model proposed by Oliver (2003)

  ii. The simple period-dependent spatial correlation model (SSCM) proposed by Jayaram and Baker (2009) and

  iii. No correlation model (NC)

It is noted that apart from FBCC other models including cross-correlation and inter-event residual correlation exist, such as Loth and Baker (2013). However previous studies showed that two methodologies are similar in terms of computational demand and yield very similar loss estimates (Weatherill et al. 2015). Thus, there is not necessarily a clear case for adopting one over the other in application. For brevity, just the FBCC case is considered here.

7.2.1.2 Considered IMs

In most building codes and modern catastrophe risk models, seismic hazard is defined in terms of ground-motion IMs. As discussed in detail in Chapter 5, conventional IMs, such as peak ground responses (PGA, PGV and PGD) and $S_a(T_1)$, poorly predict the structural response of mid- to high-rise MRFs, although $S_a(T_1)$ can sufficiently capture the elastic behaviour of first-mode dominated structures, characterized by low-to-moderate fundamental periods (Shome et al. 1998) and low levels of nonlinearity. On the other hand, advanced IMs perform better over a range of IM selection criteria, therefore are more suitable for predicting the seismic response of structure, and eventually the seismic fragility and loss. Chapter 5 concluded that $S_a(T_1)$ and I_{Np} are the optimal conventional and advanced IMs respectively, for this reason are used herein as benchmark.

In this study T_1 for both IMs is assumed to be equal to 1 second for all the considered buildings, as this is the representative natural period for the mid-rise building class in catastrophe modelling, and T_N (for the definition of I_{Np}) is assumed to be equal to $2T_1$ (i.e. 2 seconds). Additionally, the α parameter of I_{Np} is calibrated in accordance to the building vulnerability class of interest (see Section 5.6).

7.2.2 Fragility and vulnerability assessment

7.2.2.1 Considered analysis methods

An essential step towards the development of analytical fragility functions and consequently vulnerability functions is the choice of the analysis method to calculate seismic demands. Although conventional linear elastic analysis is appropriate for design purposes, inelastic approaches are preferred for fragility and loss assessment as they better depict the inelastic seismic behaviour of structures, particularly under severe ground motions. Nonlinear methods are therefore characterized by lower uncertainty; the degree of uncertainty within inelastic analysis methods varies substantially, depending on the level of detailing of the structural model

and the representation of the input seismic loading. Two main types of inelastic analysis exist and are widely used in practice, namely the NDA and NSP types, such as CSM, N2, FRACAS etc. A more detailed background and review of the different analysis methods is presented in Chapter 2 (Section 2.7).

In consistency with previous Sections, NDA and the nonlinear static approach FRACAS are used here for analysing the seismic response of the buildings of interest.

7.2.2.2 Development of fragility curves

As discussed in Chapter 2 (Section 2.7), a number of procedures is available in the literature for estimating the parameters to construct analytical fragility functions. To recap, fragility functions are defined as continuous relationships between the ground motion IM and the probability that a specified asset will reach or exceed predefined DSs in accordance to their structural response. The structural response of buildings, as obtained from the chosen analysis methods, is expressed in terms of EDPs, e.g., MIDR, and compared to EDP thresholds associated with each DS. Statistical methods are employed to characterize this response probabilistically as a function of seismic hazard, and build the conditional distribution of an EDP given the IM.

Cloud method and BEA are utilized for the estimating the probabilistic seismic response and consecutively for estimating fragility function parameters. BEA is still a cloud-based approach which utilizes an advanced statistical regression for the EDP|IM fitting, in contrast to the simple linear regression fitting used in conventional cloud method. Other well-established analysis strategies for fragility assessment also exist, such as IDA (Vamvatsikos and Cornell 2002) and MSA (Jalayer and Cornell 2009), but are not considered here as they require challenging record selection and scaling procedures especially in the case of building portfolio applications.

Twenty four sets of fragility functions are presented herein. These twenty four fragility curves, which are illustrated in **Figure 7.1** and **Figure 7.2**, correspond to:

i)   3 building types, representative of different vulnerability classes:

         a. Pre-code

         b. Low-code

         c. Special-code

ii)   2 structural analysis types:

         a. NDA

         b. FRACAS

iii)   2 types of IMs:

         a. $S_a(T_1)$

         b. I_{Np}

iv)   2 fragility derivation methods:

         a. Conventional cloud method

         b. BEA
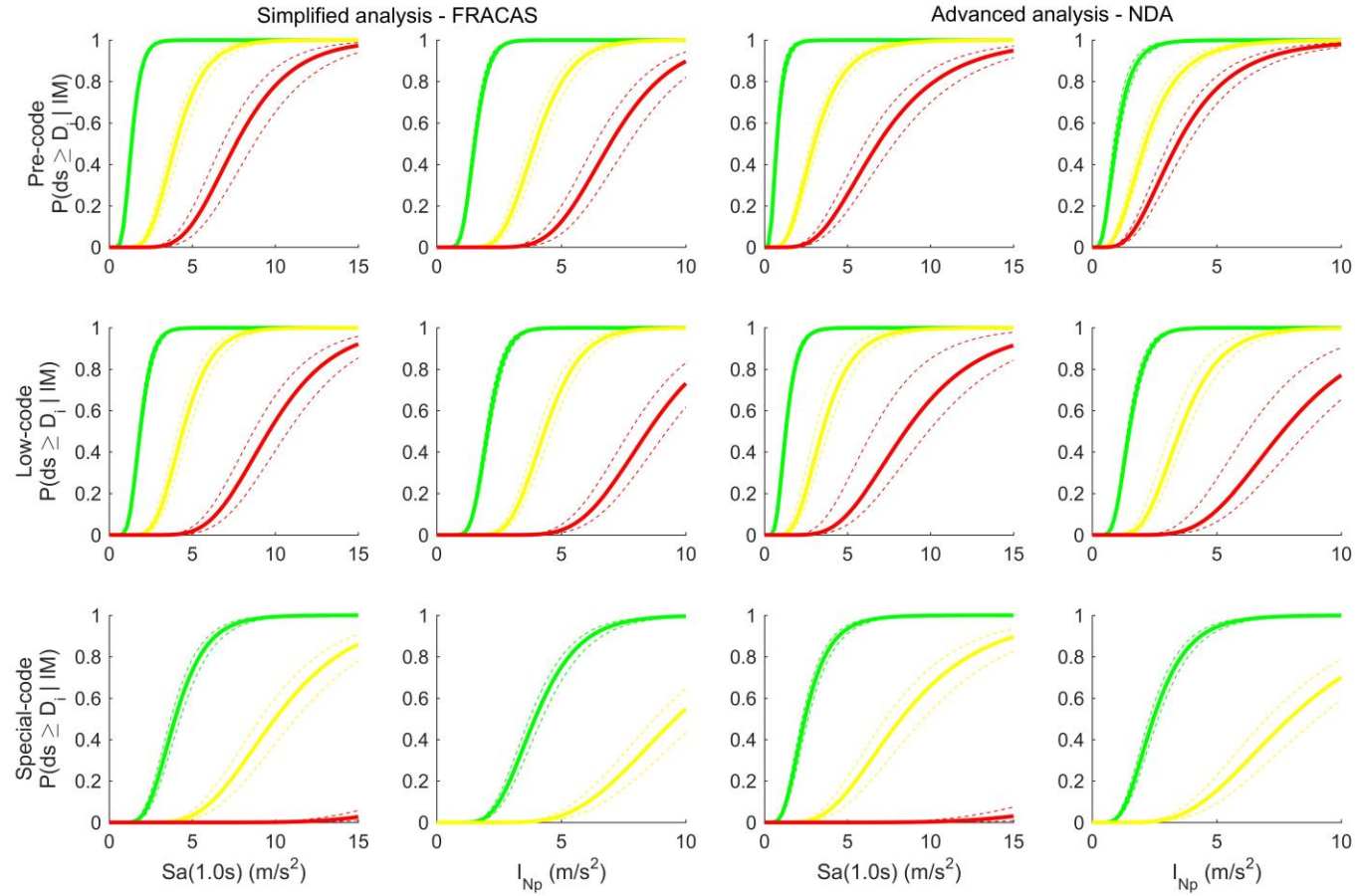
Figure 7.1 - Fragility curves derived using Cloud method for the following twelve combinations: three vulnerability classes, namely Pre-, Low- and Special-code (corresponding to first, second and third row respectively); two analysis types, namely FRACAS and NDA (corresponding to columns 1-2 and 3-4 respectively); two types of IMs, $S_a(1.0s)$ and (corresponding I_{Np} to columns 1,3 and 2,4 respectively).

Figure 7.2 - Fragility curves derived using BEA method for the following twelve combinations: three vulnerability classes, namely Pre-, Low- and Special-code (corresponding to first, second and third row respectively); two analysis types, namely FRACAS and NDA (corresponding to columns 1-2 and 3-4 respectively); two types of IMs, $S_a(1.0s)$ and (corresponding I_{Np} to columns 1,3 and 2,4 respectively).

7.3 Case study application

7.3.1 Exposure model

A synthetic exposure model representing the town of Avellino in the Campania region of Southern Italy is used here. This is a well-studied area which is chosen because its building stock mainly consists of RC buildings, and is characterized by moderate-to-high seismicity (two major events occurred in the last 90 years; i.e. the Mw 6.72 on 27/7/1930 and Mw 6.89 on 23/11/1980).

According to recent field surveys (Ricci 2010; Del Gaudio 2015), the predominant construction material in the town is RC, representing about 80% of total building inventory (**Figure 7.3**), the majority of which fall within mid-rise category (~63% RC), i.e. four to six storeys (**Figure 7.4**). For simplicity, it is assumed for this example that the building stock only consists of mid-rise RC buildings, an assumption deemed appropriate for the nature of this study. The buildings are spatially distributed according their exact locations, as they are extracted directly from the OSM database (OpenStreetMap 2017), as illustrated in **Figure 7.5**.
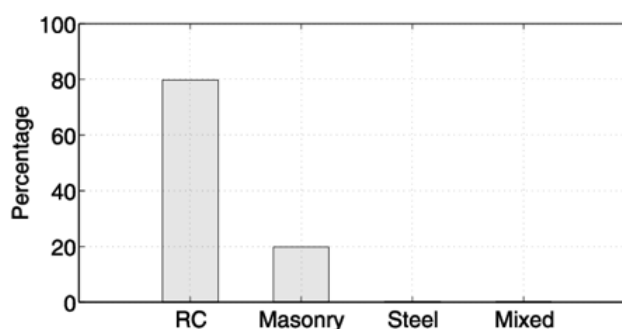
Figure 7.3 - Distribution of building typologies (Del Gaudio 2015).
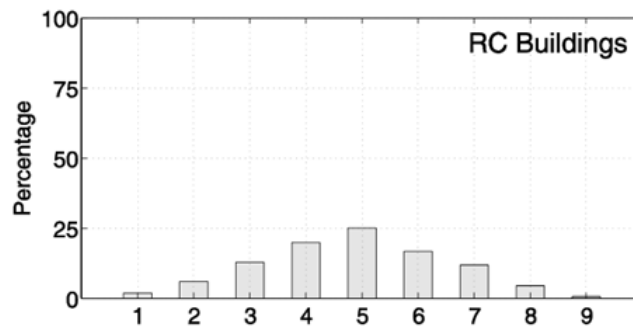
Figure 7.4 - Distribution of number of storeys of RC buildings (Del Gaudio 2015).
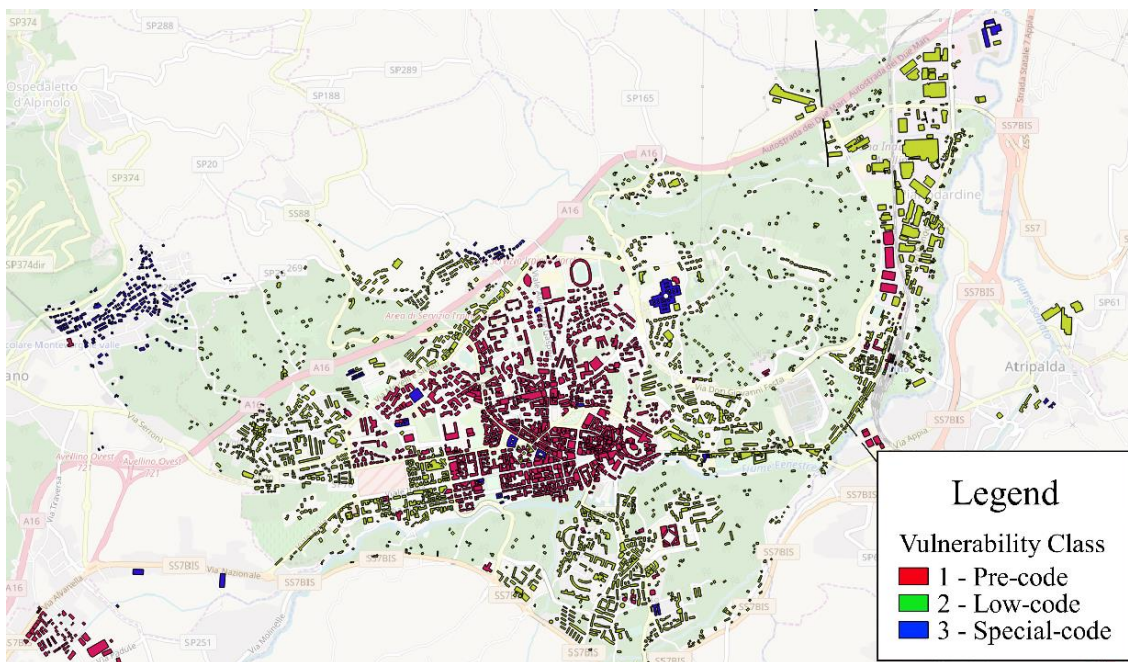
Figure 7.5 - The map of Avellino town, and the associated building portfolio (OpenStreetMap 2017).

The buildings are organized based on their year of construction, and associated design-code level into three main classes, namely Pre-code (constructed before 1972, representing 32% of the total population), Low-code (erected between 1973 and 2002, representing 60% of the total population) and Special-code buildings (built after 2003, representing 8% of the total population). The OSM databases do not include any information about the year of construction of each building, therefore, a large concentration of Pre-code buildings is assumed at the centre of Avellino encircled by Low-code buildings, while Special-code buildings are scattered in the outskirts of the town. This is a typical Mediterranean city urban layout, where the city is built

around the historic centre, which mainly consists of older buildings, and is surrounded by layers of buildings erected in later periods.

7.3.2 Hazard model and ground motion simulation

With regard to the hazard estimation, the output of the Earthquake Model for the Pan-European Region developed by AIR-worldwide[4] (further referred to as AIR) is used. Seismicity in the Pan-European region can be attributed to crustal earthquakes on known faults, large subduction interface earthquakes, large subduction zone earthquakes, and shallow background and deep seismic activity. In the AIR Earthquake Model for the Pan-European Region, large earthquakes associated with subduction zones and active crustal faults are modelled as characteristic events. The seismicity of smaller magnitude earthquakes on subduction zones and crustal faults and the seismicity of regions not captured by these sources are represented by background seismicity. A smoothed background-seismicity distribution, implemented on a three-dimensional grid, is used to capture the small-to-moderate events that occur on as yet unknown or unmapped faults. The model domain is divided into two depth layers based on historical activity and the tectonics. Shallow events such as those affecting the region investigated occur at depths of 0-25 km.

The hazard component of the Pan-European Earthquake Model was used to develop a stochastic catalogue of events with 10,000 year simulations of seismicity, incorporating a total of 817,251 simulated events with Mw $\geq$ 5, covering thirty European countries. Each event in the stochastic catalogue is characterized by an epicentre, a focal depth, a moment magnitude, rupture geometry (length, width, azimuth and dip angle,) and rupture mechanism. Both the spatial and the depth distributions are consistent with the earthquakes included in a historical catalogue of events for the European region. The aforementioned catalogue, which consists of 12,500 historical earthquakes, was generated from local catalogues compiled by research groups and government organizations that include both instrumental earthquakes recorded by local and regional seismic networks and pre-instrumental events determined from various intensity

---

[4]   http://www.air-worldwide.com/publications/brochures/documents/air-earthquake-model-for-the-pan-european-region

reports. It should be noted that a 10,000-year catalogue may not be long enough to obtain a stable exceedance probability curve (in terms of hazard and loss). For this reason, AIR developed a 100,000-year catalogue to which an optimization procedure was applied, creating an optimized 10,000-year sample of the 100,000 year catalogue that ensures the accuracy of hazard and loss results, while optimizing the computational effort.

The ground motion footprint of three different IMs, namely $S_a(0.3s)$, $S_a(1.0s)$ and $S_a(2.0s)$ were generated for each event in the catalogue over a grid resolution of 250x250m^2. It is noted that each of the three periods used for the calculation of each spectral acceleration is assumed to represent the natural period of low-, mid- and high-rise buildings respectively; this is a typical assumption used in most catastrophe models. For this example, spectral acceleration for natural period of one second is used ($S_a(1.0s)$) (representing the natural period of mid-rise RC buildings in catastrophe models), while all three available spectral accelerations are used to calculate the corresponding value of I_{Np}. In this way, $S_a(0.3s)$ can capture the effect of the higher modes (characterized by periods less than T_1), which are explicitly modelled when performing NDA, while $S_a(2.0s)$ can capture the period elongation effect. The AIR model incorporates a weighted combination of GMPEs designated for Italy, European and Mediterranean region obtained from a variety of studies for active, stable, and subduction/deep earthquakes.

7.3.3 Vulnerability model

The three RC MRF buildings introduced in Chapter 3, namely Pre-, Low- and Special-code buildings, are used here to populate the vulnerability model of this case study. The nonlinear response of the buildings is estimated with the aim of NDA and FRACAS analysis methods over a large set of ground motions suitable for displacement-based design and assessment, selected from the SIMBAD database, following the approach presented in detail in Chapter 3 (Section 3.2.2.1).

The twenty four sets of fragility functions presented in Section 7.2.2 are then combined with a damage to loss model in order to derive the vulnerability for each of the twenty four combinations. In this example, the damage to loss model designated for Italy by Di Pasquale and Goretti (2001) is employed. The Di Pasquale and Goretti (2001) model is based on EMS-98 damage scale (Grünthal 1998), which consists of 6 DSs, namely None; Negligible to Slight; Moderate; Substantial to Heavy; Very Heavy; Destruction. The description of EMS-98 DSs is provided in Appendix B: (**Table B 2**). **Table 7-1** illustrates the median damage ratio values (MDR) and the associated standard deviation for each DS included in Di Pasquale and Goretti (2001) model (assuming a beta distribution). It is noted that MDR in the aforementioned model is defined as the ratio of repair cost over replacement cost, and no nonstructural damage is considered.

Table 7-1 - Damage ratios and associated standard deviations for each damage state of Di Pasquale and Goretti (2001) model.

	D0	**D1**	**D2**	**D3**	**D4**	**D5**
DS name	None	Negligible to Slight	Moderate	Substantial to Heavy	Very Heavy	Destruction
MDR	0.003	0.041	0.218	0.410	0.781	0.814
SD	0.033	0.054	0.158	0.211	0.142	0.254

In contrast to EMS-98, the damage scale used herein (Section 3.3, **Table 3-3**) consists of four DSs (including also "No Damage" DS); therefore, assumptions are introduced to approximate the equivalence between the two damage scales. The equivalence of different damage scale is a common practice and has been used in several past studies (including Hill 2011; Rossetto and Novelli 2016). Correspondence between the DSs of the two different scales is drawn from interpretations of the damage scale descriptions alone (Rossetto and Novelli 2016). However, both damage scales provide limited damage descriptions, as shown in **Table 3-3** and **Table B 2**) making such equivalence a difficult process. For simplicity, it is assumed that each DS of the scale used here corresponds directly to DS of the EMS-98 scale. **Table 7-2** below shows the equivalence of the damage scale proposed herein (current damage scale) and the EMS-98 scale.

Table 7-2 - Damage scale equivalence table.

EMS-98 Dam. scale	D0	D1	D2	D3	D4	D5
DS name	None	Negligible to Slight	Moderate	Substantial to Heavy	Very Heavy	Destruction
Current Dam. Scale	DS0		DS1	DS2	DS3	
DS name	No Damage		Moderate	Extensive	Partial Collapse	
MDR	0.003		0.218	0.410	0.781	
St. dev.	0.033		0.158	0.211	0.142	

Figure 7.6 and **Figure 7.7** show two sets of twelve vulnerability functions which are derived based on the fragility functions shown in **Figure 7.1** and **Figure 7.2**, and the MDRs of the current damage scale (**Table 7-2**).

The level of damage can be converted into a fraction of loss by using the MDR and assuming a beta distribution to account for the possible range of MDRs within each DS.

The last step before performing the seismic loss estimation is to convert the MDRs computed above to losses per unit area. This is done by merely multiplying MDR to values of replacement cost corresponding to building typologies constituting the portfolio of interest, i.e. mid-rise RC buildings located in Avellino city. The average construction costs for structures in Italy is estimated to be to 1300 €/m^2 (CMDC 2011), and is nearly uniform for all the Italian territory. In this study, a replacement cost rounded up to 1500 €/m^2 is assumed for all three case study buildings, in order to account for the damages of the building content (Asprone et al. 2013).
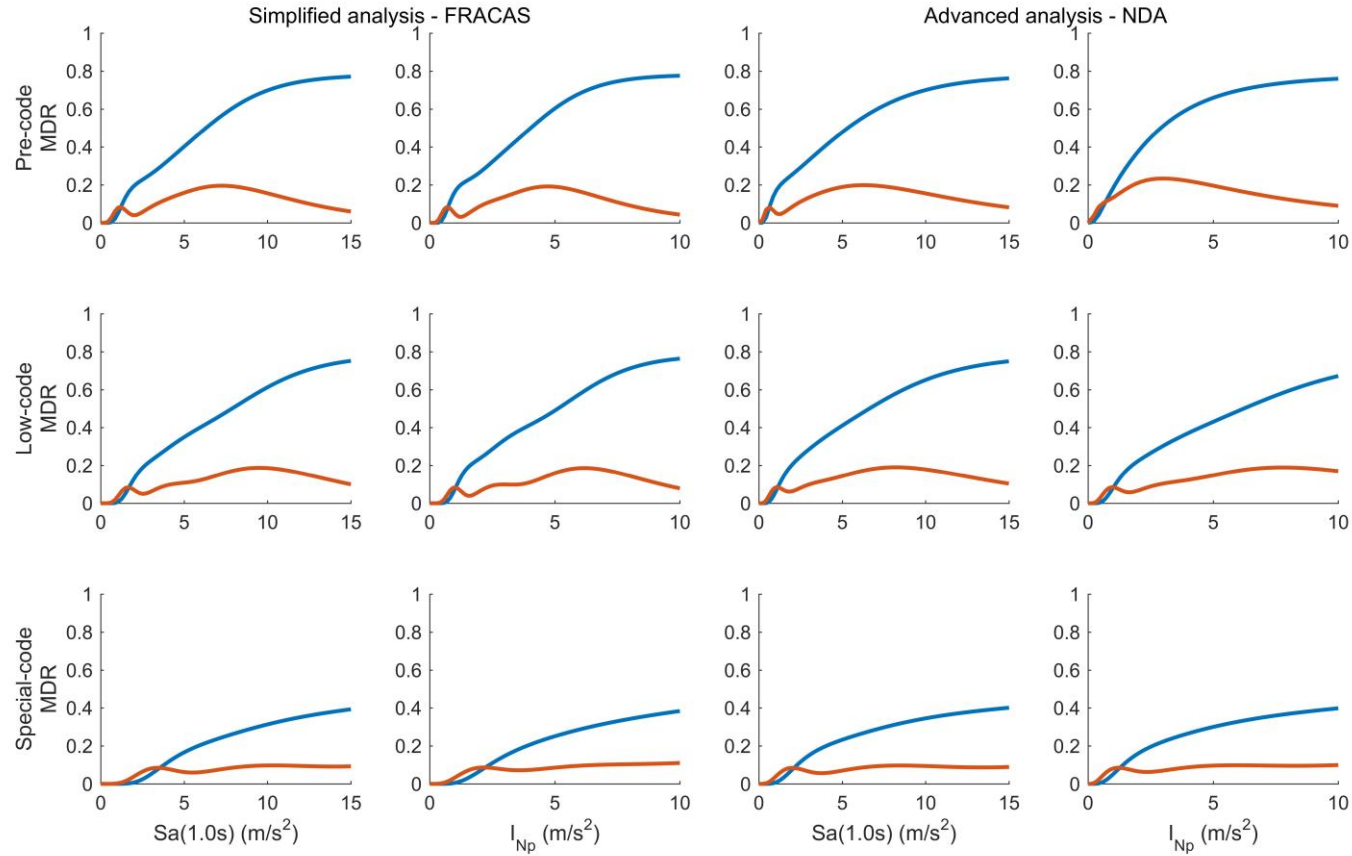
Figure 7.6 - Median vulnerability curves (blue lines) and associated variance (red lines) derived using Cloud method for the following twelve combinations: three vulnerability classes, namely Pre-, Low- and Special-code (corresponding to first, second and third row respectively); two analysis types, namely FRACAS and NDA (corresponding to columns 1-2 and 3-4 respectively); two types of IMs, $S_a(1.0s)$ and (corresponding I_{Np} to columns 1,3 and 2,4 respectively).
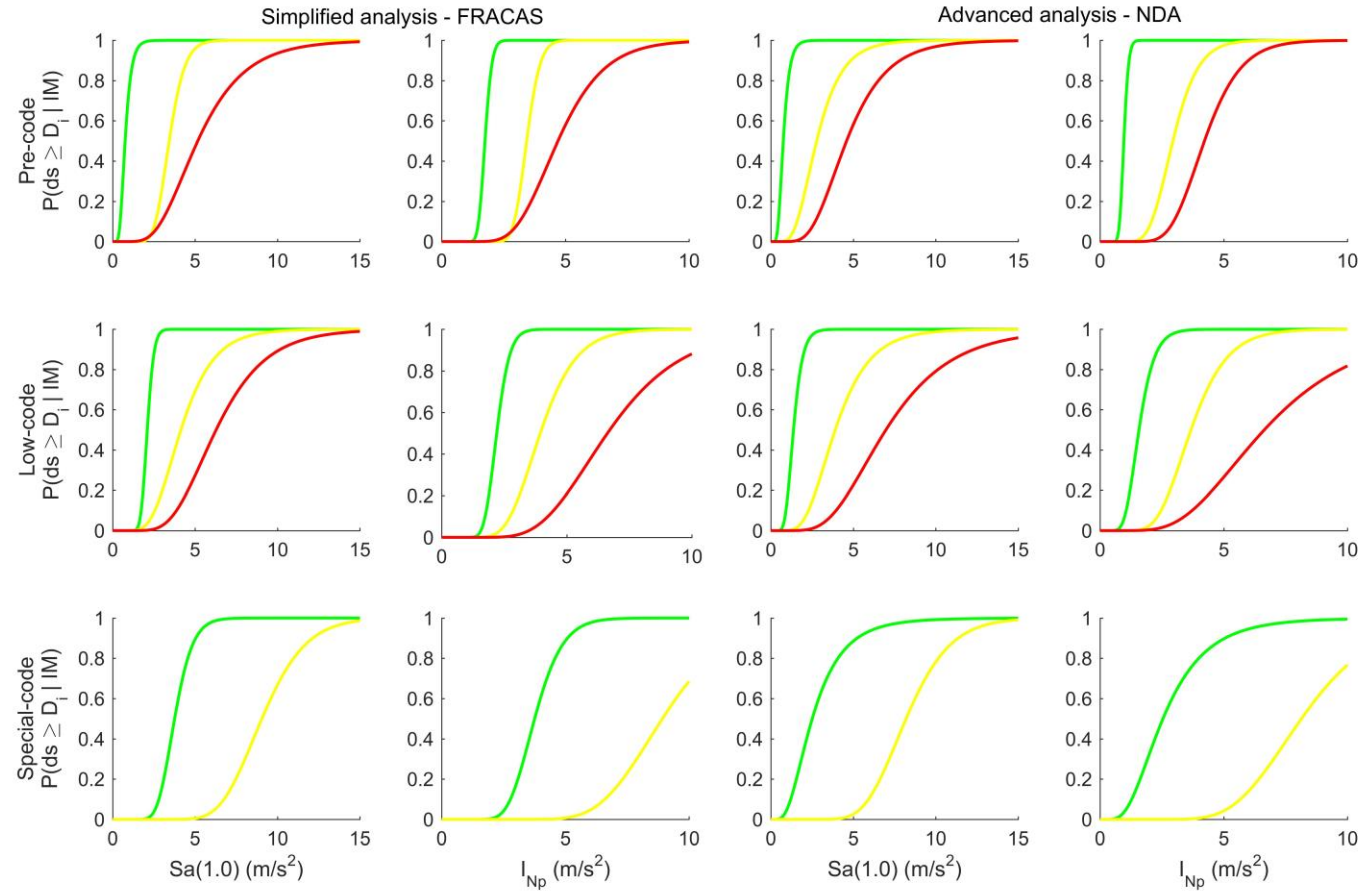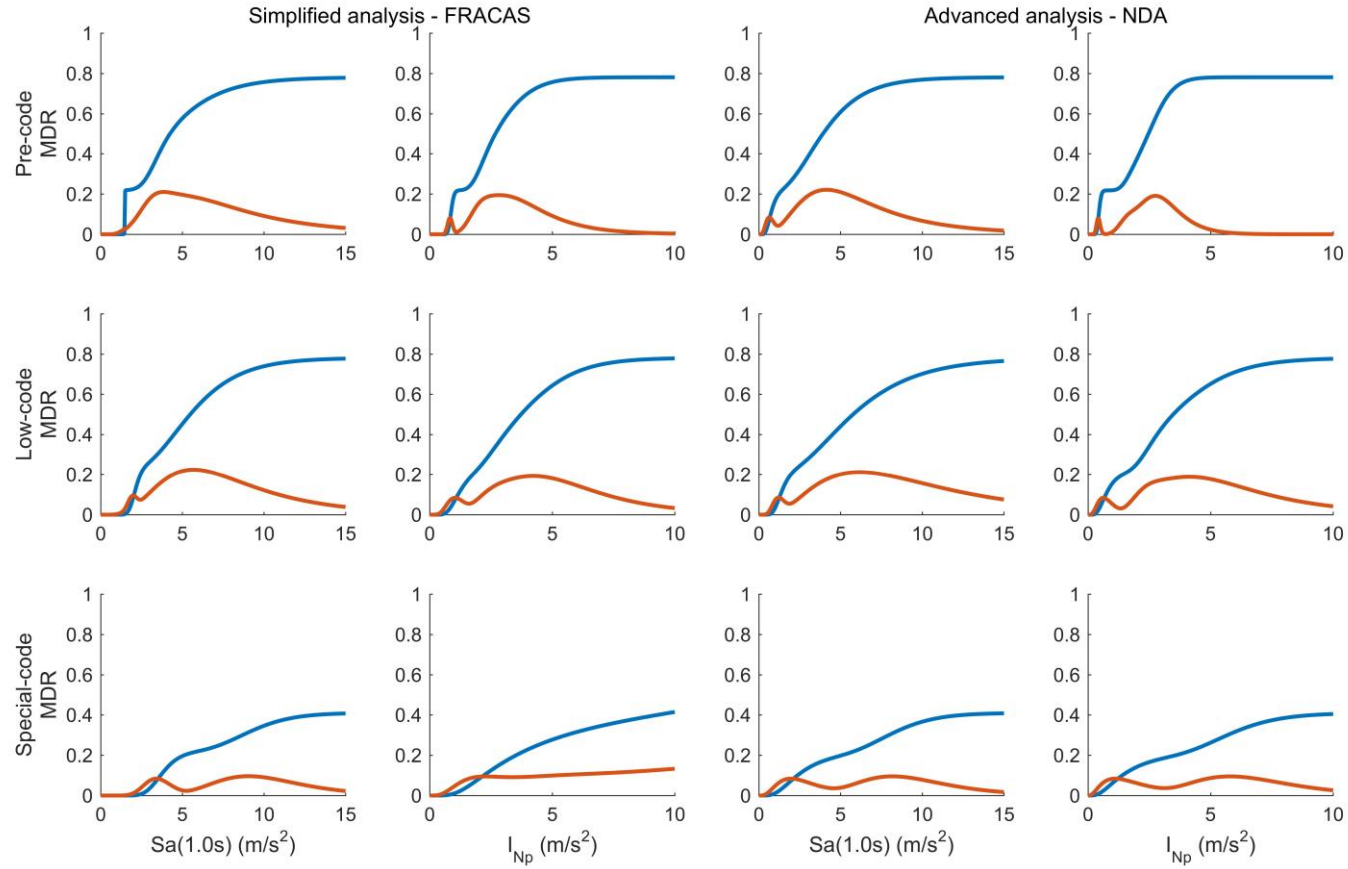
Figure 7.7 - Median vulnerability curves (blue lines) and associated variance (red lines) derived using BEA method for the following twelve combinations: three vulnerability classes, namely Pre-, Low- and Special-code (corresponding to first, second and third row respectively); two analysis types, namely FRACAS and NDA (corresponding to columns 1-2 and 3-4 respectively); two types of IMs, $S_a(1.0s)$ and (corresponding I_{Np} to columns 1,3 and 2,4 respectively).

7.4 Sensitivity of seismic loss estimation to input assumptions

To understand the impact of the different input assumption on the risk analysis, a comparison of the estimated seismic losses for each modelling option is carried out. Eight modelling options are considered here as obtained from combining two different types of IMs discussed above (conventional IM – $S_a(1.0s)$, and advanced IM – I_{Np}), two types of analysis (simplified analysis – FRACAS, and advanced analysis – NDA), and the two PSRA approaches for fragility derivation (standard Cloud method, and advanced BEA), as summarized below:

- Case 1: Vulnerability function derived from NDA analysis and ground motion intensity expressed in terms of $S_a(1.0s)$

- Case 2: Vulnerability function derived from NDA analysis and ground motion intensity expressed in terms of I_{Np}

- Case 3: Vulnerability function derived from FRACAS analysis and ground motion intensity expressed in terms of $S_a(1.0s)$

- Case 4: Vulnerability function derived from FRACAS analysis and ground motion intensity expressed in terms of I_{Np}

Table 7-3 below presents all the different configuration cases used in this case study example. It is noted for each of the abovementioned cases there are two subcases depending on the PSRA approach for fragility derivation used.

Table 7-3 - Configuration cases implemented for the computation of seismic losses.

Case No	IM	Analysis Method	Fragility Method
Case 1	$S_a(1.0s)$	NDA	Cloud
Case 2	I_{Np}	NDA	Cloud
Case 3	$S_a(1.0s)$	FRACAS	Cloud
Case 4	I_{Np}	FRACAS	Cloud
Case 1a	$S_a(1.0s)$	NDA	BEA

Case 2a	I_{Np}	NDA	BEA
Case 3a	$S_a(1.0s)$	FRACAS	BEA
Case 4a	I_{Np}	FRACAS	BEA

The abovementioned modelling configurations are also used to assess the sensitivity of losses the inclusion of spatial and spectral correlation models, as well as the effect of spatial resolution grid of the exposure.

The seismic losses are then computed for each of these modelling options. A probabilistic event-based approach was used to evaluate the seismic losses to each of the portfolios, using the aforementioned stochastic catalogue of events with 10,000-year simulations of seismicity. Aggregated losses were evaluated for each simulated earthquake event. The resulting set of portfolio event losses is commonly referred to as the "event loss table", which was used to evaluate the annual probability of achieving or exceeding any possible level of loss. This information is used to build the so-called "loss exceedance probability curve" (or simply EP curve). EP curves describe the probability that various levels of loss will be exceeded over a specific timeframe (usually one year) and are obtained by transforming the "event loss table" into a "year loss table". More specifically, the total loss for each year is computed as the sum of the losses for the events simulated in that year. As a result, it is possible to evaluate the annual frequency of exceedance of any loss by ranking the year loss values in descending order and dividing each rank by the length of the catalogue (e.g., the highest year loss has rank equal to one, and a frequency of exceedance of 1/10,000, whereas the lowest year loss has a rank equal to 10,000 and an annual exceedance frequency of 10,000/10,000 = 100%). This process is repeated for 100 independent realisations of all the random variables involved (e.g., ground motion IM, fragility and damage-to-loss ratio). This allows one to account for the various sources of uncertainty discussed above, and to calculate the median EP curve.

Another important metric used in catastrophe modelling either for establishing insurance premiums or for fund distribution for risk reduction at a national scale, is the AAL. AAL can be

214

calculated from the numerical integration of the EP curve. Alternatively, AAL can be merely computed as the sum of annual losses throughout the entire catalogue of events, divided by the number of year simulations of seismicity, as shown below:

$$AAL = \frac{1}{n_s}\sum_{i=1}^{n_s} AL_i \qquad (7.7)$$

where AL_i is the annual loss for the year number i, and n_s is the total number of year simulations of seismicity.

The model using the advanced IM, the advanced analysis method and the advance fragility derivation method (i.e. I_{Np} combined with NDA and BEA respectively – Case 2a) combined with the FBCC correlation model, is assumed to be the *Benchmark* as it represents the best available model/practice for this work and in the literature. **Figure 7.8** shows the median EP of the *Benchmark* case alongside with the 16th and 84 percentiles plotted against the aggregated loss per unit area. The *Benchmark* model is used as a reference case for the evaluation of the different configurations for the estimation of seismic losses presented in the following Sections.
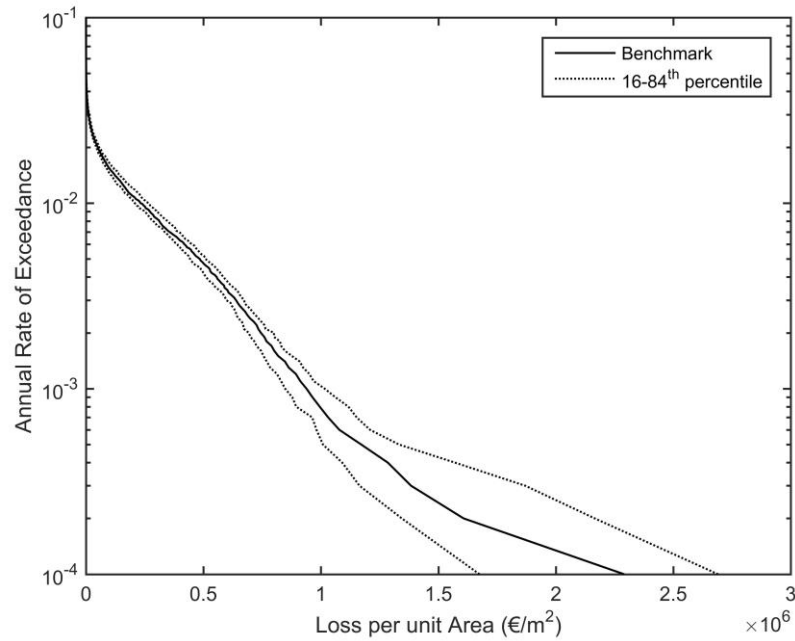
Figure 7.8 - Exceedance probability curves for the median *Benchmark* case (continuous line) and the 16-84th percentiles (dotted lines).

215

7.4.1 Impact of hazard assessment assumptions in seismic risk analysis

7.4.1.1 Sensitivity to the selection of IMs and fragility derivation method

In the first test-case Section, the effect of IM selection and fragility derivation method to the resultant seismic loss estimation is assessed. Following the procedure described above, EP curves and AAL are calculated considering four configuration approaches, namely Case1, Case2, Case1a and Case2a (see **Table 7-3**). All four configuration cases employ the same analysis type and correlation model (i.e. NDA and FBCC), but use a different combinations of IM type, and fragility derivation method. It is noted that Case2a combined with the FBCC model corresponds to the *Benchmark* case. The left panel of **Figure 7.9** illustrates the EP curves expressed in terms of loss per unit area, while right panel shows these curves normalized using the *Benchmark* model.
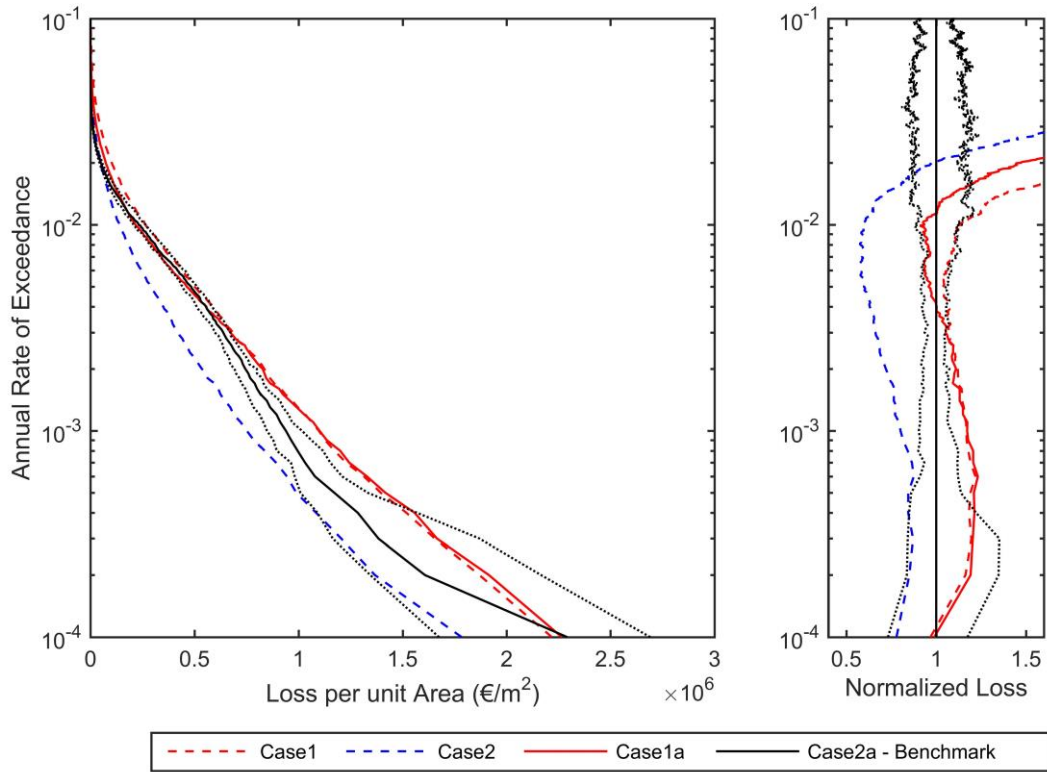
Figure 7.9 - Exceedance probability curves using FBCC model generated in terms of loss per unit area (left panel), and in terms of loss normalized to *Benchmark* loss (right panel) for the median Case1(red dashed line), Case2(blue dashed line), Case1a (red continuous line), and Case2a - *Benchmark* (black continuous line) and the associated 16-84th percentiles (black dotted lines).

Regarding the impact of IM choice, the EP curves obtained using $S_a(1.0s)$ (Case1 and Case1a) show a substantial discrepancy comparing to the associated I_{Np} curves (Case2 and Case2a), with the losses of $S_a(1.0s)$ in most cases exceeding the associated I_{Np} ones along the various EP levels. For the configurations Case1 and Case1a, this discrepancy ranges between -10% and +20% for rare events (EP rates between 10^{-4} to $2 \cdot 10^{-2}$) and most of times lies within or nearly within the 16-84th percentiles of the *Benchmark* case. The normalized losses above annual rates of exceedance of $2 \cdot 10^{-2}$ show very large discrepancies. This is due to the fact that the losses obtained for higher probability levels (i.e. more frequent events) are close to zero; therefore, a small variation in the loss can cause a significant deviation in the computation of normalized losses. Similar deviations are also observed in test-cases presented in the following tests.

The choice of a given IM is of high importance, as significant difference in the loss estimation is observed when different types of IMs are used. As opposed to $S_a(1.0s)$, I_{Np} can account for the spectral shape in a range of periods, potentially resulting in more accurate structural response and damage/loss estimates. This discrepancy is further amplified by the fact that the ground motion records used in the derivation of fragility functions by NDA (and also by FRACAS as shown in the following Sections) have not been selected based on the actual site-specific hazard. Specifically, the IM footprints used in this study were generated combining the stochastic catalogue and a set of GMPEs for Italy and Mediterranean region, while the fragility curves were developed utilizing 150 unscaled records arbitrarily selected from the SIMBAD database (recorded worldwide); see Rossetto et al. (2016) for details. This observation is consistent with the study of Kohrangi et al. (2017), which shows a non-negligible influence of the site-specific hazard (and corresponding ground motion selection) on the fragility/vulnerability derivation, the degree of which depends on the IM adopted for assessment. This study also showed that advanced IMs can considerably reduce such dependence.

The choice of fragility derivation method also has an effect in the seismic loss estimation, as shown in the comparison between Case1 and Case1a, and Case2 and Case2a. In the former case, (where $S_a(1.0s)$ is used), a small impact is observed on normalized losses for EP rates greater than $3 \cdot 10^{-3}$ (frequent events), where the configuration using BEA (Case1a) is matching more closely the *Benchmark* model. The normalized losses for EP rates lower than $3 \cdot 10^{-3}$ (rare events) are almost identical. However, this effect becomes more evident in the latter case, where I_{Np} is used (Case2 and Case2a). This is mainly because, BEA as an advanced fragility method, performs better when combined with I_{Np}, taking advantage of the additional information provided by an advanced IM, confirming the observation discussed in Chapter 5.

Table 7-4 presents the AAL and the associated standard deviation (σ_{AL}) for the four considered modelling cases, in conjunction with the associated differences (%) to the *Benchmark* model. AAL estimates and the variability for both $S_a(1.0s)$ configurations (Case 1 and Case1a) are greater than the *Benchmark* model (Case2a), while Case2 configuration results in lower estimates. Regarding configurations Case1 and Case2, the AAL and standard deviation estimates differ substantially from the target *Benchmark*, which may lead to unrealistic insurance premiums or fund distribution for risk reduction. On the other hand, although differences obtained from Case1a are not very significant, they still have an impact to percentiles of annual losses which may affect the decision-making process.

Table 7-4 - Average annualized losses (AAL) and associated standard deviation (σ_{AL}) considering 2 modelling cases for the FBCC correlation model

Case No	AAL	Benchmark difference (%)	σ_{AL}	Benchmark difference (%)
Case 1	2547.47	31.9%	39847.41	11.3%
Case 2	1442.83	-25.3%	26512.18	-25.9%
Case 1a	2177.94	12.8%	38919.08	8.7%
Case 2a - *Benchmark*	1931.58	0.0%	35788.66	0.0%

7.4.1.2 Sensitivity to spatial and spectral correlation models

This sub-section investigates the sensitivity of spatial and spectral correlation models choice in the calculation of seismic losses. For consistency, the same four configuration approaches that were used in Section 7.4.1.1 (Case1, Case2, Case1a and Case2a - **Table 7-3**) are also tested here but this time combined with different spatial and spectral correlation models. **Figure 7.10** (left panel) depicts the EP curves expressed in terms of loss per unit area, generated using the simple spectral correlation model (SSCM). These loss curves are normalized using the results from *Benchmark* model in the right panel. In the same way, **Figure 7.11** shows the curves generated using no correlation model (NC). **Figure 7.12** and **Figure 7.13** show the loss curves for all tested correlation models, for Case1a and Case2a respectively. **Table 7-5** and **Table 7-6** show the AAL and the associated standard deviation (σ_{AL}) for the four considered configurations for the SCCM and NC models respectively.

Figure 7.10 - Exceedance probability curves using SSCM model generated in terms of loss per unit area (left panel), and in terms of loss normalized to Benchmark loss (right panel) for the median Case1(red dashed line), Case2(blue dashed line), Case1a (red continuous line), Case2a(blue continuous line), and Benchmark (black continuous line) and the associated 16-84th percentiles (black dotted lines).

Figure 7.11 - Exceedance probability curves using NC model generated in terms of loss per unit area (left panel), and in terms of loss normalized to Benchmark loss (right panel) for the median Case1(red dashed line), Case2(blue dashed line), Case1a (red continuous line), Case2a(blue continuous line), and Benchmark (black continuous line) and the associated 16-84th percentiles (black dotted lines).

A close inspection of **Figure 7.10**, confirms the observations from **Figure 7.9** regarding the choice of IM type and fragility method in the seismic loss computation. The use of a model that accounts for spatial-correlation, such as SSCM, results in losses that follow similar trends to the ones presented in the case of FBCC (**Figure 7.12** and **Figure 7.13**). Specifically, the losses estimated using $S_a(1.0s)$ (Case1 and Case1a) are greater than the ones obtained using I_{Np} (Case2 and Case2a), for the whole range of this EP levels. The observations regarding the impact of fragility derivation method are also in good agreement with the outcomes presented in Section 7.4.1.1, i.e. the normalised losses for the cases of $S_a(1.0s)$ (Case1 and Case1a) are not very much affected by the choice of fragility method, while the associated losses for the cases of I_{Np} (Case2 and Case2a) differ substantially. On the other hand, the use of NC results in significant underestimation of seismic losses comparing to the *Benchmark* model, as shown in **Figure 7.11**, an observation which is in agreement with previous studies (e.g., Weatherill et al.

220

2015). Non-inclusion of spatial correlation diminish the tails of the distribution meaning that for each event the probability of sampling very high values in the right tail is reduced, resulting in increased losses. This is also confirmed visually in **Figure 7.12** and **Figure 7.13**, where NC EP curves are well below the Benchmark curve, lying in most cases outside the 16th-84th percentile space.

With regard to AAL and standard deviation, all results using SSCM model (**Table 7-5**) appear to be shifted down comparing to FBCC model (**Table 7-4**). This shift, becomes more apparent in the cases where the advanced IM, I_{Np} is used (Case2 and Case2a), resulting a difference between -9.9 to -13.3% in AAL (-6.7% to -7.7% in σ_{AL}). In contrast, when the conventional IM, $S_a(1.0s)$, is used this drop almost negligible, resulting a difference between -1.0 to -1.7% in AAL (-0.2% to -1.2% in σ_{AL}).

Non-inclusion of correlation model (i.e. NC) also has a significant impact in the AAL estimation (**Table 7-6**). Specifically, the employment of NC model results in substantially lower AAL estimates comparing to the *Benchmark* case, ranging between -31,3% to -54.6% (-22.5% to -47.4% in σ_{AL}). Neglect of correlation effect generates unrealistically low losses. This observation is in agreement with observations from previous studies (Weatherill et al. 2015; Silva (under review)); therefore, it is not recommended for probabilistic seismic loss assessment applications.

Table 7-5 - Average annualized losses (AAL) and associated standard deviation (σ_{AL}) considering 4 modelling cases for the SCCM correlation model

Case No	AAL	Benchmark difference (%)	σ_{AL}	Benchmark difference (%)
Case 1	2508.17	29.9%	39396.63	10.1%
Case 2	1248.28	-35.4%	24134.21	-32.6%
Case 1a	2145.02	11.1%	38830.69	8.5%
Case 2a	1673.99	-13.3%	33025.03	-7.7%

Table 7-6 - Average annualized losses (AAL) and associated standard deviation (σ_{AL}) considering 4 modelling cases. NC model is considered here.

Case No	AAL	Benchmark difference (%)	σ_{AL}	Benchmark difference (%)
Case 1	1327.31	-31.3%	25712.11	-28.2%
Case 2	877.78	-54.6%	18826.52	-47.4%
Case 1a	1036.85	-46.3%	24398.27	-31.8%
Case 2a	1180.23	-38.9%	27730.86	-22.5%
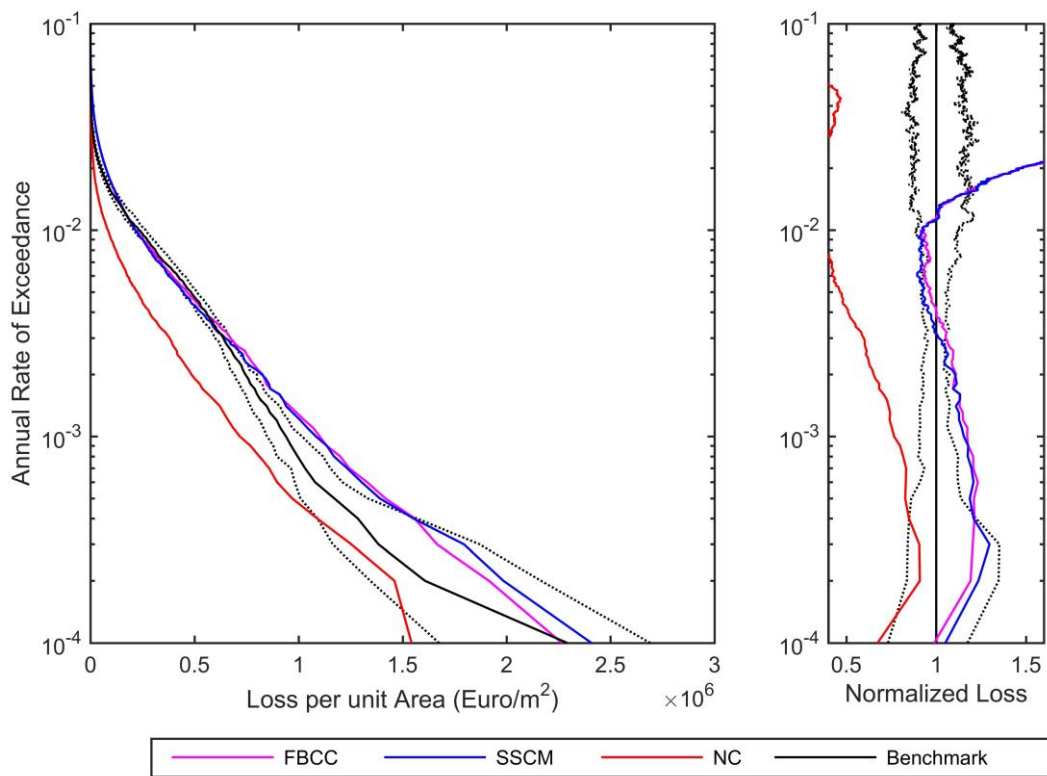
Figure 7.12 - Exceedance probability curves for Case 1a configuration generated in terms of loss per unit area (left panel), and in terms of loss normalized to Benchmark loss (right panel) using: NC (red line), SSCM (blue line) and FBCC/Benchmark (black line) and the associated 16-84th percentiles (black dotted lines).
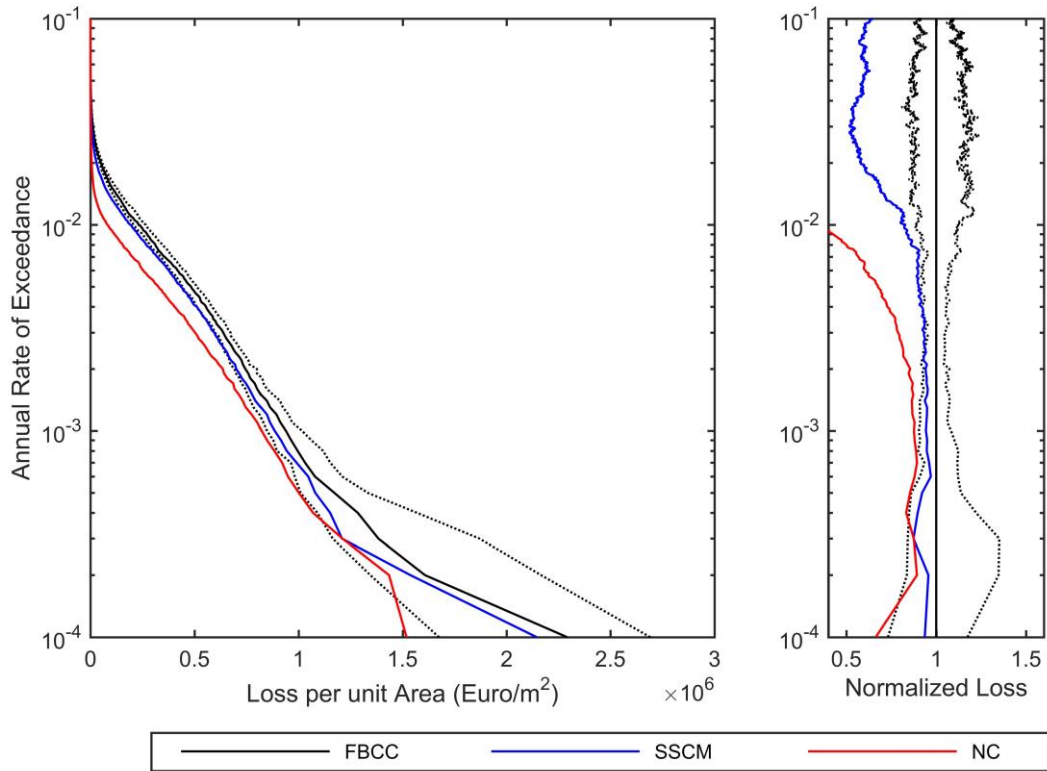
Figure 7.13 - Exceedance probability curves for Case 2a configuration generated in terms of loss per unit area (left panel), and in terms of loss normalized to Benchmark loss (right panel) using: NC (red line), SSCM (blue line) and FBCC/Benchmark (black line) and the associated 16-84th percentiles (black dotted lines).

7.4.1.3 Sensitivity to spatial resolution

Another option that influences the seismic loss estimation is the selection of spatial resolution of the exposure for the considered portfolio. As discussed in Section 7.3.2, the spatial resolution used in this case study is $250x250m^2$. However, in most cases catastrophe models employ lower resolution spatial scales (e.g., most commonly $1.0x1.0km^2$, or $500x500m^2$). This Section aims to explore the sensitivity of seismic losses when a lower spatial resolution grid of IM footprints, i.e. $500x500m^2$, is utilized. The EP curves for the same configuration cases as in previous Sections, combined with FBCC and a $500x500m^2$ spatial resolution are presented in **Figure 7.14**. The same arrangement is also used for the AAL results which are tabulated in **Table 7-7**. The same process has also been applied to SSCM and NC correlation configurations and the

results of the analysis are essentially consistent with the observations shown for the case of FBCC. For brevity reasons, the results of SSCM and NC correlation configurations are omitted.
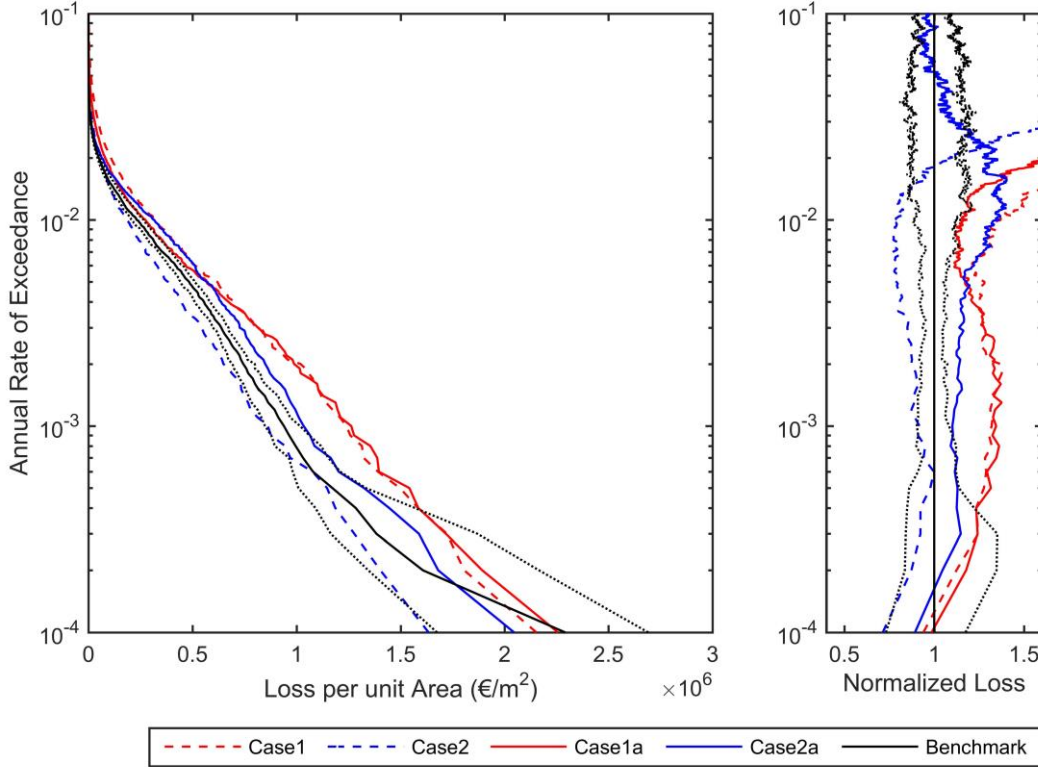
Figure 7.14 - Exceedance probability curves for the FBCC correlation model combined with a lower resolution spatial scale (500x500m^2) generated in terms of loss per unit area (left panel), and in terms of loss normalized to *Benchmark* loss (right panel) for the median Case1(red dashed line), Case2(blue dashed line), Case1a (red continuous line), and Case2a - *Benchmark* (black continuous line) and the associated 16-84th percentiles (black dotted lines).

As shown in **Figure 7.14**, all configuration cases deviate substantially from the *Benchmark* model along the majority of EP rates. This can be clearly observed in the normalized loss graph (**Figure 7.14** – right panel), where all configuration curves lie outside the 16th-84th percentiles of the *Benchmark* model, except for the losses corresponding to rare events, i.e. for EP rates lower than 10^{-3} for Case2 and Case2a, and $6.5 \cdot 10^{-3}$ for Case1 and Case 1a. This practically means that although the use of lower spatial resolution grid yields good loss estimates for rare events, it also produces unrealistic losses for frequent events. The use of a lower resolution grid results in shift of the curves to higher loss levels. **Table 7-7** confirms the aforementioned observation, an

AAL *Benchmark* difference increase is observed when comparing the associated AAL *Benchmark* difference for higher resolution (**Table 7-4**).

Table 7-7 - Average annualized losses (AAL) and associated standard deviation (σ_{AL}) considering four modelling cases for the FBCC correlation model combined with a lower resolution spatial scale (500x500m^2).

Case No	AAL	Benchmark difference (%)	σ_{AL}	Benchmark difference (%)
Case 1	2864.67	48.3%	44633.68	24.7%
Case 2	1724.15	-10.7%	30422.51	-15.0%
Case 1a	2515.14	30.2%	44068.87	23.1%
Case 2a	2318.16	20.0%	40331.20	12.7%

7.4.1.4 Sensitivity to structural analysis type

The tests presented in previous Sections were performed using data obtained from advanced structural analysis (i.e. NDA). In this Section, the sensitivity of seismic loss estimates is investigated when simplified structural analysis (i.e. FRACAS) is used instead. As discussed extensively in previous Chapters, the employment of simplified analysis is an appealing option in fragility analysis, as it significantly reduces the analysis time. However, the use of simplified analysis yields lower accuracy comparing to advanced counterpart.

Figure 7.15 shows the loss curves obtained when for configuration approaches that employ simplified analysis, namely Case3, Case4, Case3a and Case4a (see **Table 7-3**). In consistency with **Figure 7.9**, the four configuration cases employ the same correlation model (i.e. FBCC) and anaylsis method (i.e. FRACAS), but use a different combinations of IM type, and fragility derivation method, and are presented alongside with the *Benchmark* case. AAL estimates and the associated standard deviation (σ_{AL}) for the same four considered configuration cases is illustrated in **Table 7-8**.

The use of simplified analysis underestimates the resultant losses as shown in **Figure 7.15**. This becomes evident when comparing Case4a with the *Benchmark* case; the two cases use the same

IM and fragility method, but a different analysis method. The loss curves stemming from Case4a are always below the ones from the *Benchmark* case. The same trend also appears for the other cases that use the same IM and fragility combinations but different analysis method.
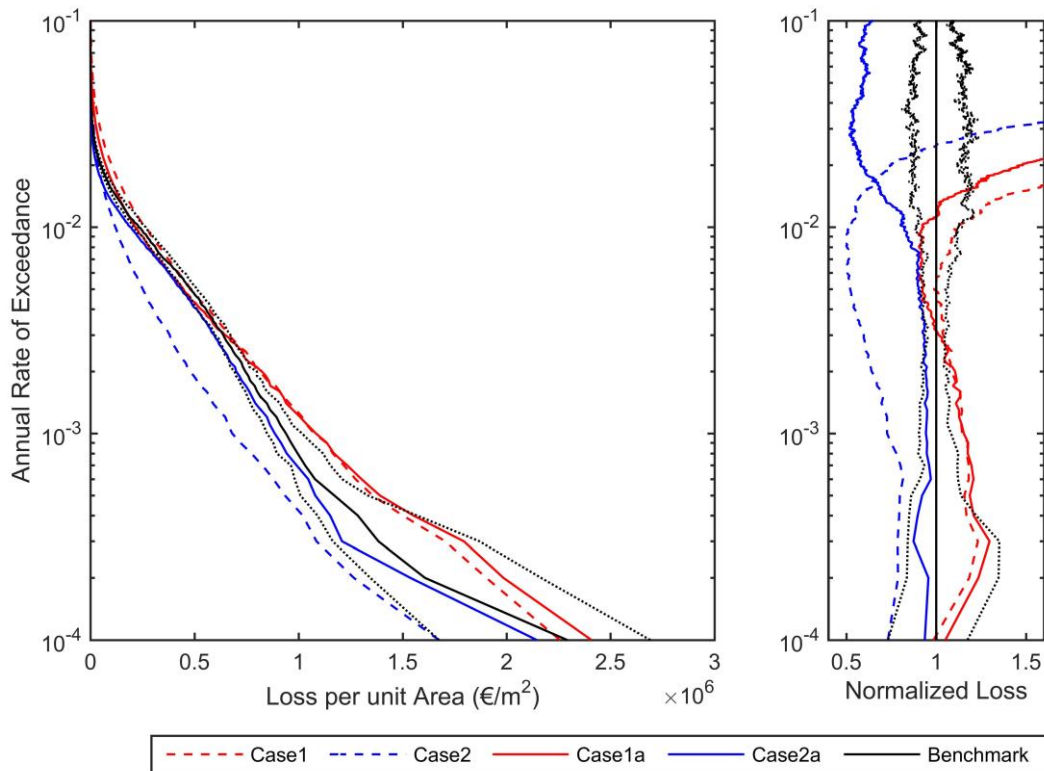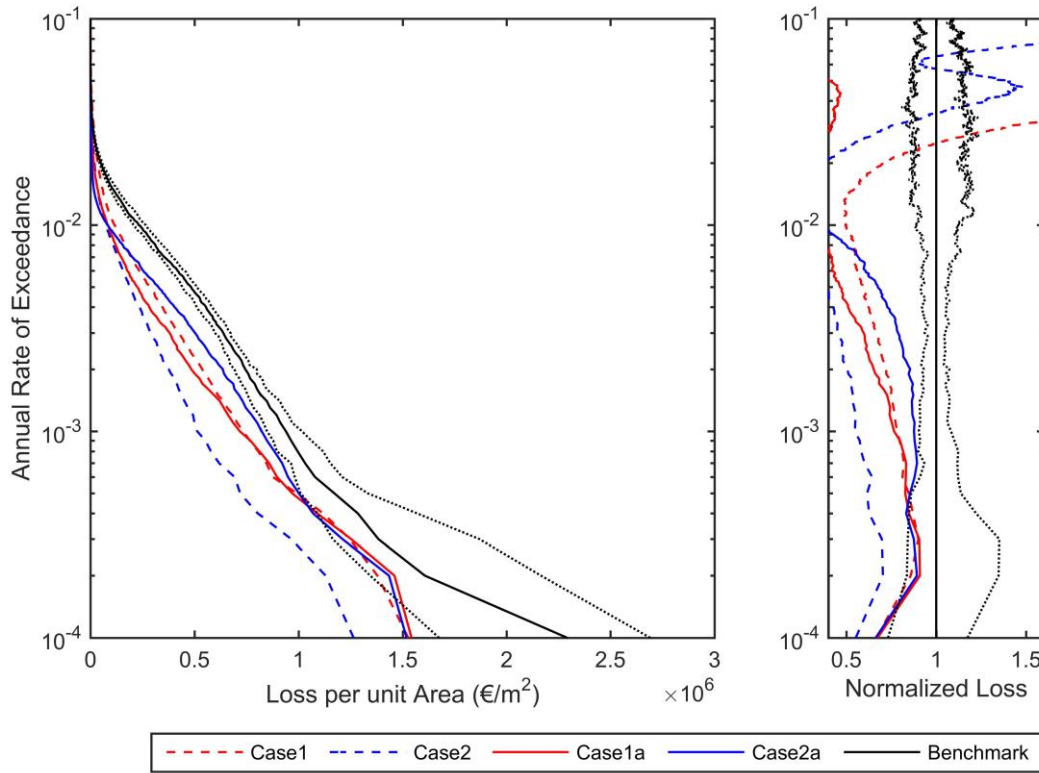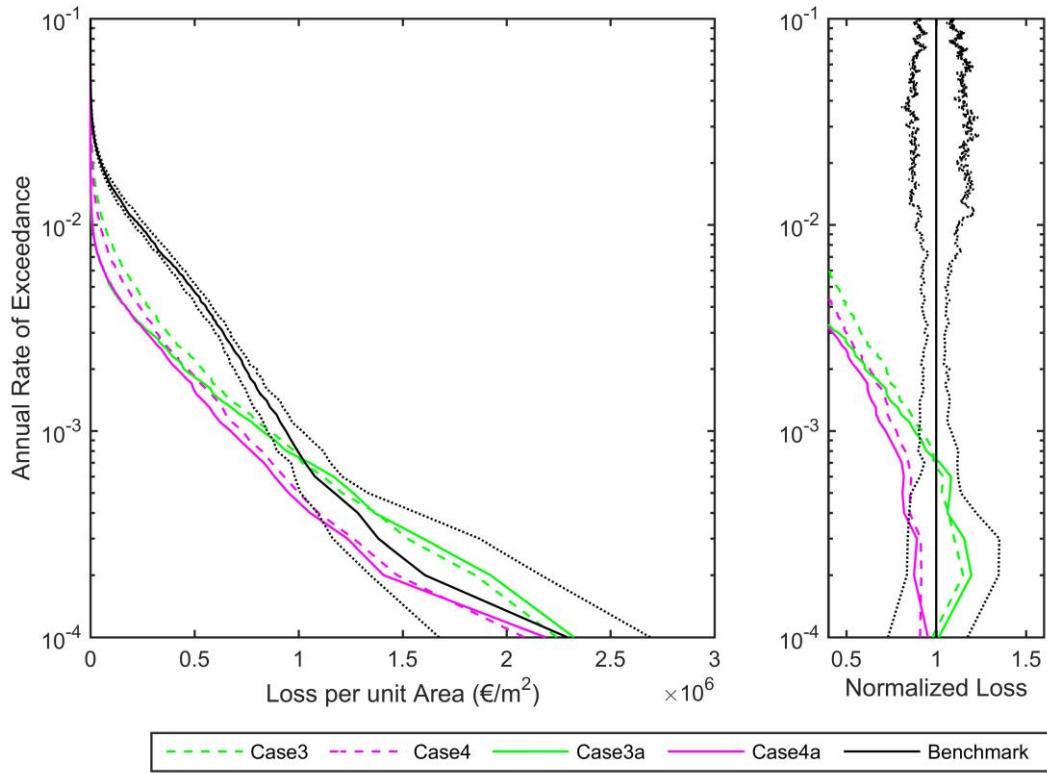

Figure 7.15 - Exceedance probability curves using FBCC model generated in terms of loss per unit area (left panel), and in terms of loss normalized to *Benchmark* loss (right panel) for the median Case3 (green dashed line), Case4 (magenta dashed line), Case3a (green continuous line), Case4a (magenta continuous line), and *Benchmark* (black continuous line) and the associated 16-84th percentiles (black dotted lines).

This drop in loss estimation can be attributed to the limitations of the pushover-based approaches, including FRACAS, thoroughly discussed in Sections 2 and 4. For instance, simplified methods cannot account for the contribution of the higher modes. In addition, FRACAS utilizes simple hysteretic models and capacity curve idealizations, and models failure modes in a simplified manner. However, the normalised loss estimates for rare events (EP rates lower than 10^{-3}) are within a range of -25% to +20% comparing to the *Benchmark* model, and in most cases included in the *Benchmark* 16th-84th percentile area. It is noteworthy to mention that the sensitivity of losses to IM and fragility type, follow similar patterns to the ones observed in previous Sections.

Regarding AALs, large differences are observed for all simplified analysis cases when compared to the *Benchmark* model, reaching in some cases reductions greater than 60% (**Table 7-8**). However, such differences are largely affected by the limited ability of simplified method to predict damage mechanisms within specific EP rates (frequent events), and do not reflect the good performance of FRACAS for lower EP rates.

Table 7-8 - Average annualized losses (AAL) and associated standard deviation (σ_{AL}) considering the four modelling cases associated with simplified analysis for the FBCC correlation model.

Case No	AAL	Benchmark difference (%)	σ_{AL}	Benchmark difference (%)
Case 3	1068.75	-44.7%	29115.89	-18.6%
Case 4	866.24	-55.2%	24847.94	-30.6%
Case 3a	750.91	-61.1%	28143.43	-21.4%
Case 4a	656.00	-66.0%	23479.84	-34.4%

7.5 Conclusions

This Chapter investigates the impact of hazard and vulnerability modelling options to the estimation of probabilistic regional seismic loss assessment of RC buildings.

A synthetic, heterogeneous portfolio of buildings located in Avellino, Southern Italy, is chosen as a case-study. A site-specific hazard analysis accounting for uncertainty in the factors affecting ground motions is carried out, allowing the calculation of IM footprints for different types of IMs, including conventional and advanced IM representations. The exposure database consists of three regular mid-rise RC bare frame buildings corresponding to three distinct vulnerability classes, which are spatially distributed to realistically simulate the typical Italian exposure. These buildings are analysed using NDA and the simplified analysis procedure FRACAS to generate sets of EDP|IM sets, where the IM is expressed in terms advanced and conventional intensity measures. Fragility functions are generated from this data, for the three building typologies, using advanced and standard fragility derivation methods (i.e. BEA and cloud respectively). The probabilistic regional economic earthquake losses are estimated and

227

presented in the form of EP curves and AAL for all tested hazard and vulnerability combinations (i.e. forty eight combinations in total).

The loss analyses show that the choice of IM type is of high importance due to the ability of advanced IMs to capture more information than conventional IM, e.g. spectral shape in a range of periods. In addition, it is observed that the use of advanced IMs can significantly reduce the effect of the ground motion characteristics to the predicted structural response, and consequent loss estimates, which is consistent with recent studies. On the other hand, the choice of fragility derivation method appears to have a smaller influence on the loss estimation. The advanced method BEA performs better when combined with I_{Np}, taking advantage of the additional information provided by an advanced IM, confirming the observation discussed in Chapter 5.

Results from this work highlight that the choice of the analysis method type has the most significant impact on the loss estimation. Use of a simplified rather than advanced structural analysis method is expected to introduce added uncertainty in the fragility functions (as seen in Chapter 6) and hence, to have an effect on the loss calculation, which accounts for this uncertainty. The NDA can account for the effect of higher modes and more advanced/realistic hysteretic behaviour, and can determine the failure modes more precisely. This results in estimated losses that are substantially higher than those predicted using simplified analyses procedures.

This study demonstrates the importance of including spatial cross-correlation in the estimation of seismic losses, an observation which is in agreement with other studies in this field (Weatherill et al. 2015; Silva (under review)). It is seen that the use of a simple spatial correlation model results in losses that follow similar trends to those predicted with spatial cross-correlation, but result in lower loss levels. This effect is further amplified when SSCM is combined with an advanced IM. Neglect of correlation effect generates unrealistically low losses; therefore, it is not recommended for probabilistic seismic loss assessment applications.

The effect of spatial resolution in the exposure and the corresponding simulated footprints is also investigated. A higher spatial resolution ensures a more refined intensity representation of footprints; therefore, a more realistic representation of losses is assumed. Tests conducted using a lower spectral resolution produce a shift of losses to significantly higher levels.

Based on the findings of this Chapter, the choice of structural analysis approach appears to have higher impact than IM choice or fragility derivation method choice in the loss calculation. This is an important observation that is not explicitly discussed in the existing literature. This observation also raises concerns as to the reliability of loss estimates made from fragility functions that adopt the N2 method, which is considered to be the standard practice in many catastrophe models. It is noteworthy to mention that some of the modelling options above result in overestimations of seismic losses (e.g., lower spatial resolution, choice of conventional IM), while other options cause a reduction in the loss estimated (e.g., simplified analysis method, implementation of simpler correlation model). As a result, combinations of the above modelling options are seen to cancel out the effect of each of them, yielding results that closely match the *Benchmark* case, almost by chance rather than due to the reliability and adequacy of the hazard and vulnerability input combination.

In conclusion, the *Benchmark* case is expected to provide the best loss estimates comparing to the other tested modelling combinations. Modelling combinations utilizing NDA for all IM and fragility options (i.e. Case 1, Case 2, Case 1a), are shown to provide reasonably good estimates, comparing to the *Benchmark* case; therefore, are highly recommended when high computational resources are available. The use of BEA fragility derivation method, when combined NDA (i.e. Case 2a and Case 1a), yields the best loss estimates for both IM types. However, the benefits of using BEA are not evident when it is combined with FRACAS.

The use of FRACAS, (which is an example of an NSP), consistently provides an underestimation of the losses, which cannot be redressed by IM choice or fragility method. Nevertheless, if an NSP is to be used, due to computational limitations, then the results of this

study would suggest that it can be used with any of the tested IMs and fragility methods, with little variation in the overall loss.

It is worthy mentioning that of existing CSM approaches, FRACAS is one of the most advanced, and has been shown to outperform the commonly adopted N2 approach (Rossetto et al. 2014a; Nassirpour 2018). From a catastrophe modelling point of view, this pronounced influence of analysis method raises concerns as to the reliability of loss estimates made from fragility functions that adopt the N2 method, which is considered to be the standard practice in many existing catastrophe models.

8. Conclusions and future work

8.1 Research summary and concluding remarks

Estimating the impact of a potential future seismic event in terms of human and economic loss is a crucial topic for the mitigation of earthquake risk, and is of interest to various stakeholders including insurers/reinsurers, insurance brokers, engineers and governments. Such estimation can be done through rigorous treatment of uncertainty and the consideration of different models within the components of the seismic risk framework. Seismic risk assessment of mid-rise reinforced concrete (RC) buildings is not a new topic for academic and commercial research within the field of earthquake engineering. Despite the major advancements in all the technical concepts involved within the seismic risk assessment of these buildings, several fallacies and shortcomings still exist, as made evident in the literature review presented in Chapter 2. This Thesis' aim was to propose some advancements in seismic risk assessment, with particular focus on the fragility/vulnerability component of risk, and evaluate their impact in the resulting loss estimation. This research makes a significant contribution to both academic and commercial understanding of seismic risk assessment for RC buildings, and is expected to be of particular interest to disaster resilience experts and catastrophe modellers.

First, the buildings used as case studies in this Thesis are presented in Chapter 3. The three central index buildings are all characteristic of the mid-rise RC moment resisting frame (MRF) building class, and represent typical Mediterranean structural designs over three different periods (i.e. before 1972, between 1973 and 2002 and after 2003) and associated vulnerability levels. The buildings were analysed using both advanced and simplified analysis methods (nonlinear dynamic analysis (NDA) and FRACAS respectively), and distinct sets of fragility function were derived. The fragility functions were constructed based on a tailored damage scale developed specifically for each case study building, and presented in terms of three IMs commonly used catastrophe risk applications, i.e. PGA, $S_a(T_1)$ and $S_a(1.0s)$. This study has produced thirty new sets of fragility functions that can be used directly in seismic risk and loss

assessment applications. However, it is important to recognise that the abovementioned fragility functions should be implemented with caution, appreciating the limitations of using functions derived based on single buildings.

Recognising that advanced analysis methods require significant resources in terms of computational effort and time, the simplified analytical tool to assess seismic fragility of buildings, FRACAS, was developed (Chapter 4). FRACAS builds upon the modified capacity spectrum method (CSM) first developed by Rossetto and Elnashai (2005) and captures the record-to-record variability, allowing the construction of fragility curves considering a number of sophisticated options. Some of these options include various capacity curve idealizations, hysteretic models for the inelastic demand calculation of a single degree of freedom (SDoF), and statistical techniques to fit fragilities. The performance of FRACAS was evaluated against NDA results. This comparison showed that FRACAS estimates well the seismic response of the tested buildings. Last, a free user-friendly graphical user interface (GUI) version of FRACAS, i.e. the FRACAS tool, was introduced. It is noted that FRACAS, as a simplified method, should only be employed for low- and mid-rise buildings of regular shape, as it is not capable of capturing higher-mode effects and torsional phenomena. FRACAS overcomes some of the limitations of CSM, which is typically used in some catastrophe models, and therefore, renders improved response estimates without significantly aggravating analysis times.

Seismic demand uncertainty highly depends on the type of ground motion intensity measure (IM), therefore the choice of an IM which reduces such uncertainty is of paramount importance. In Chapter 5, the performance of various conventional and advanced scalar IMs was tested for the determination of optimal IMs for fragility analysis through nonlinear static procedures (NSPs). An optimal IM estimates structural response (in terms of engineering demand parameter (EDP)) with minimum record-to-record variability, largest amount of information on ground motion features and is hazard computable. The three case study buildings presented in Chapter 3 were analysed using the simplified procedure FRACAS to determine their seismic response, over a large set of real ground motion records. For the representation of seismic

hazard, four conventional (PGA, PGV, PGD, and $S_a(T_1)$) and two advanced IMs (S_a^c and I_{Np}) were considered, and their performance was assessed with the aim of three IM selection criteria, namely efficiency, sufficiency/relative sufficiency and hazard computability. In agreement with previous studies performed with NDA, Chapter 5 showed that advanced IMs (and particularly I_{Np}) perform best in all criteria (including hazard computability criterion) and yield improved fragility predictions. Therefore, these IMs are considered to be optimal for developing analytical vulnerability functions and assessing the seismic losses. Regarding conventional IMs, $S_a(T_1)$ was shown to be the most efficient option for all tested criteria, closely matching in some cases the performance of I_{Np}.

Evaluation of the probabilistic seismic response of structures is an essential, but time consuming part within the fragility/vulnerability component of the seismic risk assessment process. An advanced Bayesian emulation approach (BEA), which overcomes some of the limitations of existing approaches (i.e. Cloud method) and computes efficiently the probabilistic seismic response of buildings was proposed in Chapter 6. The BEA, which is a version of kriging, uses a mean function as a first approximation of the expected EDP|IM distribution and then models the approximation errors as a Gaussian process (GP). In this Chapter, an overview of the Bayesian emulation framework was introduced, and was then extended to be applied in the context of fragility estimation. Artificial sets of EDPs at given IM locations were generated from the analysis data of the chosen case study buildings (as obtained from Chapter 3), and were used for assessing the sensitivity of the BEA to several input assumptions. It was shown that BEA required a smaller number of analyses to define a fragility curves and was more effective in capturing the particularities of highly nonlinear seismic response. In addition, BEA offered better characterization of the uncertainty associated with the curves and the BEA parameter estimation. The code of BEA process is available online through the EPICentre website (http://www.ucl.ac.uk/epicentre/resources/software).

Last, a probabilistic regional seismic loss assessment of RC building populations is carried out in Chapter 7. This Chapter investigated the impact of hazard and vulnerability modelling options, some of which were developed in previous Chapters of this Thesis, to the estimation of probabilistic seismic loss. With regard to the exposure model, a synthetic building stock of RC buildings (based on distributions of buildings introduced in Chapter 3) representing the town of Avellino in the Campania region of Southern Italy, was employed for this exercise. A site-specific hazard analysis accounting for uncertainty in the factors affecting ground motions was carried out, allowing the calculation of IM footprints for different types of IMs (i.e. the optimal conventional and advanced IMs, as determined in Chapter 5). Regarding the fragility/vulnerability component, sets of fragility and vulnerability functions were generated for all combinations of analysis types (NDA and FRACAS), fragility derivation methods (Cloud method and BEA), and IM representations ($S_a(T_1)$ and I_{Np}). Additionally, the effects of correlation of ground motion characteristics, and also the spatial resolution of the simulated footprints were also investigated. The probabilistic regional economic earthquake losses were calculated and presented in the form of exceedance probability curves (EP curves) and average annual losses (AAL) for all tested modelling input combinations. Chapter 7 showed that the choice of analysis method (primarily) and IM type (secondarily) have the most significant impact in the estimation of losses, with the estimates produced using advanced IM and NDA providing the best results. The choice of fragility derivation method also had a significant influence on the seismic losses, however, this influence was not as emphatic as in abovementioned cases. Finally, the inclusion of spatial-cross correlation of ground motion characteristics, and the use of higher resolution spatial grids for the simulated footprints produced more realistic loss estimates, which observations confirmed findings of previous studies.

8.2 Future research opportunities

The limitations of this Thesis also present opportunities for potential future research.

With regard to the structural models presented in Chapter 3, a number of limitations may be identified. In this Thesis, three central index single buildings were used to represent populations of buildings corresponding to three different vulnerability classes. Although this assumption was deemed appropriate for introducing the various research advancements presented herein, a more detailed representation of the exposure would lead to more realistic outcomes. This can be addressed by generating additional building models within each building class via simulating the buildings' properties as random variables, and adding new properties, such as the modelling of the infills. A more comprehensive range of structural classes and structural response will better represent the variability of structural capacity and will result in improved quantification of structural response uncertainty. In addition, a site-specific ground motion selection should be conducted in order to provide a more realistic representation of hazard and make the analysis results more predictive of building behaviour at that site of interest.

Regarding the simplified analytical tool for fragility assessment, FRACAS, a number of strategies could be considered to upgrade future versions of the FRACAS procedure and tool. A potential improvement would be to implement additional hysteresis models, and enhance FRACAS to account for pinching and for degradation of unloading stiffness. This can be done by creating an interface between FRACAS and another open source finite element application software such as OPENSEES (McKenna et al. 2006), and access libraries of materials, hysteresis models etc. Another potential improvement for FRACAS procedure would be to provide new capacity curve idealization models (e.g., the 10% fit approach proposed by De Luca et al. 2013a, multilinear idealization model applicable on infilled frames proposed by FEMA 2005). An upgraded version of FRACAS tool, which will incorporate some of the advancements presented in this Thesis is also part of the future activities. Specifically, the new FRACAS tool will utilize: (1) alternative conventional and advanced IMs, (2) advanced fragility derivation methods (e.g., BEA), (3) incorporate damage to loss function for the calculation of vulnerability and consequently loss.

In the investigation of optimal IM selection for fragility analysis presented in Chapter 5, only a small number of IM were considered. A potential future research opportunity would be to extend this optimal IM selection study to test a wider variety of IMs, also using additional IM selection criteria, such practicality, proficiency and scaling robustness. It is noted that the use of simplified analysis procedures precludes the use of integral (i.e. duration-based) IMs, as the effect of ground motion duration cannot be properly captured within the response spectrum inputs used in NSPs. The calibration of α parameter of I_{Np} based on a bigger sample of mid-rise RC buildings can be also a topic for future investigation.

Further computational benefits in the performance of BEA approach presented in Chapter 6 can be accomplished by developing an optimal experimental design for the BEA. As a part of a future research development is to formulate a strategy for carefully selecting locations to run the response analysis to optimise BEA estimates. Within this experimental formulation, a minimum number of analysis runs will be proposed to achieve certain level of accuracy requirements. The use of a different covariance structure may yield better performance of BEA, therefore, investigating the impact of using different covariance models within BEA will be a subject of future research.

Finally, further evaluation of seismic loss estimates obtained in Chapter 7 is required. Specifically, a more realistic case study (in terms of building exposure and vulnerability, and hazard conditions representation) needs to be implemented in order to carry out the probabilistic seismic loss assessment sensitivity test. This will allow to investigate the validity of the selected input assumptions within the context of the presented probabilistic seismic loss framework. Finally, the loss predictions estimated from the aforementioned detailed case study will be evaluated against real loss data from recent seismic events. To this aim, the seismic damage and loss data collected following the Central Italy earthquake sequences in October-November 2016 and January 2017 may be utilized, taking also into consideration the effect of cumulative damage by sequential seismic events.

References

Akkar S, Sucuoğlu H, Yakut A (2005) Displacement-based fragility functions for low- And mid-rise ordinary concrete buildings. Earthq Spectra 21:901–927 . doi: 10.1193/1.2084232

Ambraseys N, Smit P, Douglas J, Margaris B, Sigbjornsson R, Olafsson S, Suhadolc P, Costa G (2004) Internet-site for european strong-motion data. Boll di Geofis Teor ed Appl 45:113–129

American Society of Civil Engineers (ASCE) (2005) Seismic design criteria for structures, systems, and components in nuclear facilities, Structural Engineering Institute, Working Group for Seismic Design Criteria for Nuclear Facilities, ASCE/SEI 43-05. Reston, VA. ISBN: 9780784407622

American Society of Civil Engineers (ASCE) (2006) Seismic Rehabilitation of Existing Buildings, ASCE Standard ASCE/SEI 41-06. Reston, VA. ISBN: 9780784408841

American Society of Civil Engineers (ASCE) (2017) Seismic evaluation and retrofit of existing buildings (41-17). Reston, VA. ISBN: 9780784414859

Antoniou S, Pinho R (2004) Development and verification of a displacement-based adaptive pushover procedure. J Earthq Eng 643–661 . doi: 10.1080/13632460409350504

Applied Technology Council (1996) Seismic Evaluation and Retrofit of Concrete Building. Report (ATC-40). California, USA.

Applied Technology Council (1985) ATC-13. Earthquake damage evaluation data for California. Redwood City, California, USA.

Applied Technology Council (2012a) FEMA P-58-1: Seismic Performance Assessment of Buildings. Volume 1 – Methodology. California, USA. doi: 10.4231/D3ZW18S8N

Applied Technology Council (2012b) FEMA P-58-2: Seismic Performance Assessment of

Buildings. Volume 2 – Implementation Guide. California, USA. doi: 10.4231/D3ZW18S8N

Aslani H (2005) Probabilistic earthquake loss estimation and loss disaggregation in buildings. Stanford University

Asprone D, Jalayer F, Simonelli S, Acconcia A, Prota A, Manfredi G (2013) Seismic insurance model for the Italian residential building stock. Struct Saf 44:70–79 . doi: 10.1016/j.strusafe.2013.06.001

Assatourians K, Atkinson GM (2013) EqHaz: An Open-Source Probabilistic Seismic-Hazard Code Based on the Monte Carlo Simulation Approach. Seismol Res Lett 84:516–524 . doi: 10.1785/0220120102

Baggio C, Bernardini A, Colozza R, Corazza L, Bella M, Di Pasquale G, Dolce M, Goretti A, Martinelli A, Orsini G, Papa F, Zuccaro G (2007) Field manual for post-earthquake damage and safety assessment and short term countermeasures (AeDES). JRC Sci Thechnical Reports 1–100

Baker JW (2005) Vector-valued ground motion intensity measures for probabilistic seismic demand analysis. PhD Thesis, Stanford University

Baker JW (2011) Conditional Mean Spectrum: Tool for Ground-Motion Selection. J Struct Eng 137:322–331 . doi: 10.1061/(ASCE)ST.1943-541X.0000215

Baker JW (2015) Efficient analytical fragility function fitting using dynamic structural analysis. Earthq Spectra 31:579–599 . doi: 10.1193/021113EQS025M

Baker JW, Cornell CA (2005) A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. Earthq Eng Struct Dyn 34:1193–1217 . doi: 10.1002/eqe.474

Baker JW, Cornell CA (2006) Correlation of response spectral values for multicomponent

ground motions. Bull Seismol Soc Am 96:215–227 . doi: 10.1785/0120050060

Baker JW, Jayaram N (2008) Correlation of spectral acceleration values from NGA ground motion models. Earthq Spectra 24:299–317 . doi: 10.1193/1.2857544

Baker JW, Lin T, Haselton CB (2014) Ground Motion Selection for Performance-Based Engineering: Effect of Target Spectrum and Conditioning Period. In: Fischinger M. (ed) Performance-Based Seismic Engineering: Vision for an Earthquake Resilient Society, 32nd edn. Springer Netherlands, Dordrecht. doi: 10.1007/978-94-017-8875-5_28

Barani S, Scafidi D, Eva C (2010) Strain rates in northwestern Italy from spatially smoothed seismicity. J Geophys Res Solid Earth 115: . doi: 10.1029/2009JB006637

Bazzurro P, Cornell CA (1999) Disaggregation of Seismic Hazard. Bull Seismol Soc Am 89:501–520

Bazzurro P, Cornell CA, Shome N, Carballo JE (1998) Three proposals for characterizing MDOF nonlinear seismic response. J Struct Eng 124:1281–1289 . doi: 10.1061/(Asce)0733-9445(1998)124:11(1281)

Bazzurro P, Luco N (2005) Accounting for uncertainty and correlation in earthquake loss estimation. In: Augusti G, Schueller G, Ciampoli M (eds) Proceedings 9th international conference structural safety and reliability (ICOSSAR). Millpress, Rome, pp 2687–2694

Biasio M De, Allain F, Petre-Lazar I (2014) A Simple and Efficient Intensity Measure to account for Nonlinear Structural Behavior. Earthq Spectra 30:1403–1426 . doi: 10.1193/010614EQS006M

Blong R (2003) A new damage index. Nat Hazards 30:1–23 . doi: 10.1023/A:1025018822429

Bojórquez E, Iervolino I (2011) Spectral shape proxies and nonlinear structural response. Soil Dyn Earthq Eng 31:996–1008 . doi: 10.1016/j.soildyn.2011.03.006

Bojórquez E, Iervolino I, Reyes-Salazar A, Ruiz SE (2012) Comparing vector-valued intensity

measures for fragility analysis of steel frames in the case of narrow-band ground motions. Eng Struct 45:472–480 . doi: 10.1016/j.engstruct.2012.07.002

Bolt B (1973) Duration of strong ground motion. In: 5th world conference on earthquake engineering. Rome, pp 1304–1313

Bommer JJ (2002) Deterministic vs. Probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy. J Earthq Eng 6:43–73 . doi: 10.1080/13632460209350432

Bommer JJ, Acevedo AB (2004) The use of real earthquake accelerograms as input to dynamic analysis. J Earthq Eng 8:43–91 . doi: 10.1080/13632460409350521

Boore DM, Gibbs JF, Joyner WB, Tinsley JC, Ponti DJ (2003) Estimated ground motion from the 1994 Northridge, California, earthquake at the site of interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California. Bull Seismol Soc Am 93:2737– 2751 . doi: 10.1785/0120020197

Bradley BA (2010) Analysis of the sources of uncertainty for portfolio • level earthquake loss estimation. Earthq Eng Struct Dyn 39:1321–1342 . doi: 10.1002/eqe.995

Brzev S, Scawthorn C, Charleson AW, et al (2013) GEM Building Taxonomy Version 2.0. GEM Tech Rep 02:188 . doi: 10.13117/GEM.EXP-MOD.TR2013.02.

Calvi G, Pinho R, Magenes G, et al (2006) Development of seismic vulnerability assessment methodologies over the past 30 years. Earthquake 43:75–104

Cao V Van, Ronagh HR (2014) Correlation between seismic parameters of far-fault motions and damage indices of low-rise reinforced concrete frames. Soil Dyn Earthq Eng 66:102– 112 . doi: 10.1016/j.soildyn.2014.06.020

Cărăuşu A, Vulpe A (1996) Fragility estimation for seismically isolated nuclear structures by high confidence low probability of failure values and bi-linear regression. Nucl Eng Des 160:287–297

Cardone D, Perrone G, Piesco V (2018) Developing collapse fragility curves for base-isolated buildings. Earthq Eng Struct Dyn 1–25 . doi: 10.1002/eqe.3126

Causevic M, Mitrovic S (2011) Comparison between non-linear dynamic and static seismic analysis of structures according to European and US provisions. Bull Earthq Eng 9:467–489 . doi: 10.1007/s10518-010-9199-1

Charette RP, Marshall HE (1999) UNIFORMAT II Elemental Classification for Building Specifications, Cost Estimating, and Cost Analysis. NISTIR 6389. Washington, D.C., USA

Charleson AW (2011) Review of existing structural taxonomies. GEM Tech Rep

Chopra AK, Goel RK (2002) A modal pushover analysis procedure for estimating seismic demands for buildings. Earthq Eng Struct Dyn 31:561–582 . doi: 10.1002/eqe.144

Cimellaro GP, Reinhorn AM, D'Ambrisi A, De Stefano M (2011) Fragility Analysis and Seismic Record Selection. J Struct Eng 137:379–390 . doi: 10.1061/(ASCE)ST.1943-541X.0000115

Clough R, Penzien J (1993) Dynamics of structures. McGraw Hill, Inc., New York, USA. ISBN: 9780071132411

Comerio MC (2006) Estimating downtime in loss modeling. Earthq Spectra 22:349–365 . doi: 10.1193/1.2191017

Cordova PP, Deierlein GG, Mehanny SSF, Cornell CA (2001) Development of a two-parameter seismic intensity measure and probabilistic design procedure. In: 2nd U.S.-Japan Workshop on PBEE Methodology for Reinforced Concrete Building Structures

Cornell CA, Jalayer F, Hamburger RO, Foutch DA (2002) Probabilistic Basis for 2000 SAC Federal Emergency Management Agency Steel Moment Frame Guidelines. J Struct Eng 128:526–533 . doi: 10.1061/(ASCE)0733-9445(2002)128:4(526)

Cover TM., Thomas JA (1991) Elements of information theory. Wiley, New York, USA. ISBN: 9780071132411

Cresme il mercato delle costruzioni (2011) 2011–2015 Rapporto congiunturale e previsionale

Crowley H, Bommer JJ, Pinho R, Bird J (2005) The impact of epistemic uncertainty on an earthquake loss model. Earthq Eng Struct Dyn 34:1653–1685 . doi: 10.1002/eqe.498

Crowley H, Pinho R, Bommer JJ (2004) A probabilistic displacement-based vulnerability assessment procedure for earthquake loss estimation. Bull Earthq Eng 2:173–219 . doi: 10.1007/s10518-004-2290-8

D'Ayala D, Meslem A, Vamvatsikos D, Porter K, Rossetto T, Crowley H, Silva V (2013) Guidelines for analytical vulnerability assessment - Low/Mid-Rise. GEM Tech Rep 08:162 . doi: 10.13117/GEM.VULN-MOD.TR2014.12

David M (1977) Geostatistical ore reserve estimation. Elsevier Ltd, Amsterdam. ISBN: 9780444597618

Davis M (1987) Production of conditional simulation via the LU decomposition of the covariance matrix. Math Geol 192:91 –98 . doi: 10.1007/BF00898189

De Luca F, Elefante L, Iervolino I, Verderame GM (2009) Strutture esistenti e di nuova progettazione : comportamento sismico a confronto. In: Anidis 2009 XIII Convegno - L' ingegneria Sismica in Italia. Bologna, Italy

De Luca F, Vamvatsikos D, Iervolino I (2013a) Near-optimal piecewise linear fits of static pushover capacity curves for equivalent SDOF analysis. Earthq Eng Struct Dyn 42:523– 543 . doi: 10.1002/eqe.2225

De Luca F, Verderame GM, Gómez-Martínez F, Pérez-García A (2013b) The structural role played by masonry infills on RC building performances after the 2011 Lorca, Spain, earthquake. Bull Earthq Eng 12:1999–2026 . doi: 10.1007/s10518-013-9500-1

De Luca F, Verderame GM, Manfredi G (2015) Analytical versus observational fragilities: the case of Pettino (L'Aquila) damage data database. Bull Earthq Eng 13:1161–1181 . doi: 10.1007/s10518-014-9658-1

Decreto Ministeriale del 14/01/2008 (2008) Norme Tecniche per le Costruzioni. Gazzetta Ufficiale della Repubblica Italiana, 29., Rome, Italy

Decreto Ministeriale del 30/05/1972 (1972) Norme tecniche alle quali devono uniformarsi le costruzioni in conglomerato cementizio, normale e precompresso ed a struttura metallica. (in Italian). Suppl. Ordin. Gazz. Uff.

Deierlein GG, Krawinkler H, Cornell CA (2003) A framework for performance-based earthquake engineering. Pacific Conf Earthq Eng 273:1–8 . doi: 10.1061/9780784412121.173

Deierlein GG, Reinhorn AM, Willford MR (2010) Nonlinear Structural Analysis For Seismic Design: A Guide for Practicing Engineers - NIST GCR 10-917-5. Natl Inst Stand 36

Del Gaudio C (2015) Seismic fragility assessment of RC buildings at large scale. PhD Thesis, University of Naples Federico II

Di Pasquale G, Goretti A (2001) Vulnerabilità funzionale ed economica negli edifici residenziali colpiti da recenti eventi sismici nazionali. Proc X Ital Conf Earthq Eng (in Italian)

Diotallevi PP, Landi L, Pollio B (2008) Evaluation of conventional and advanced pushover procedures for regular and irregular RC frames. In: 14th World Conference on Earthquake Engineering. Beijing, China

Douglas J (2018) Ground motion prediction equations 1964 2018. Glasgow, UK. Available at: http://www.gmpe.org.uk/gmpereport2014.pdf

Ebrahimian H, Jalayer F, Lucchini A, Mollaioli F, De Dominicis R (2014) Case studies on

relative sufficiency of alternative intensity measures of ground shaking. In: 2nd European Conference on Earthquake Engineering and Seismology, Istanbul 25-29, 2014. Istanbul, Turkey p 12

Ebrahimian H, Jalayer F, Lucchini A, Mollaioli F, Manfredi G (2015) Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. Bull Earthq Eng 13:2805–2840 . doi: 10.1007/s10518-015-9755-9

EEFIT (2009) The L'Aquila, Italy Earthquake of 6 April 2009: A field report. London, UK

EERI (2000) World housing encyclopedia. Available at: www.world-housing.net

Elnashai AS (2001) Advanced inelastic static (pushover) analysis for earthquake applications. Struct Eng Mech 12:51–69 . doi: 10.12989/sem.2001.12.1.051

EN 1998-1 (2004) Eurocode 8: Design of structures for earthquake resistance – Part 1: General rules, seismic actions and rules for buildings. The European Union Per Regulation 305/2011, Directive 98/34/EC, Directive2004/18/EC, Brussels, Belgium

EN 1998-3 (2004) Eurocode 8: Design of structures for earthquake resistance - Part 3. Assessment and retrofitting of buildings. The European Union Per Regulation 305/2011, Directive 98/34/EC, Directive2004/18/EC, Brussels, Belgium

Erberik MA (2008) Fragility-based assessment of typical mid-rise and low-rise RC buildings in Turkey. Eng Struct 30:1360–1374

Erdik M, Durukal E (2008) Earthquake risk and its mitigation in Istanbul. Nat Hazards 44:181–197 . doi: 10.1007/s11069-007-9110-9

Faella C, Lima C, Maritnelli E (2008) Non-linear static methods for seismic fragility analysis and reliability evaluation of existing structures. In: 14th World Conference on Earthquake Engineering. Beijing, China

Fajfar P (2000) A Nonlinear Analysis Method for Performance-Based Seismic Design. Earthq

Spectra 16:573–592 . doi: 10.1193/1.1586128

Federal Emergency Management Agency (FEMA) (2015) Rapid Visual Screening of buildings for potential seismic hazards: a handbook (FEMA P-154). Washington, D.C., USA

Federal Emergency Management Agency (FEMA) (2000) Prestandard and Commentary for the Seismic Rehabilitation of Buildings. Publication n°356 (FEMA 356). The American Society of Civil Engineers for the Federal Emergency Management Agency, Washington, D.C., USA

Federal Emergency Management Agency (FEMA) (2005) Improvement of nonlinear static seismic analysis procedures, FEMA 440. Washington, D.C., USA

Federal Emergency Management Agency (FEMA) (2009) FEMA P-750/2009 NEHRP Recommended Seismic Provisions. Washington, D.C., USA

Ferrario E, Pedroni N, Zio E, Lopez-caballero F (2015) Application of metamodel-based techniques for the efficient seismic analysis of structural systems. In: Podofillini L, Sudret B, Stojadinovic B, et al. (eds) Proceedings of the 25th European safety and reliability conference, ESREL 2015, Zürich, Switzerland, 7–10 September 2015. CRC Press

Filippou F, Popov E, Bertero V (1983) Effects of Bond Deterioration on Hysteretic Behavior of Reinforced Concrete Joints. Report EERC 83-19

Fragiadakis M, Giovanis DG, Papadopoulos V (2015) Assessment of epistemic uncertainty using trained neural networks and incremental dynamic analysis. In: SECED 2015 Conference. pp 1–10

Fragiadakis M, Vamvatsikos D, Aschheim M (2014) Application of nonlinear static procedures for the seismic assessment of regular RC moment frame buildings. Earthq Spectra 30:767–794 . doi: 10.1193/111511EQS281M

Gallagher H, Farmer B, Mendoza C, Lee C, Dickson H, Greene M (2013) GEM building

taxonomy v2.0: Evaluation and testing report, GEM building taxonomy global component. GEM Tech Rep

Gasparini D, Vanmarcke E (1976) SIMQKE: A Program for Artificial Motion Generation. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, USA

Gehl P (2017) Bayesian Networks for the Multi-Risk Assessment of Road Infrastructure. PhD Thesis, University College London

Gehl P, D'Ayala D (2016) Development of Bayesian Networks for the multi-hazard fragility assessment of bridge systems. Struct Saf 60:37–46 . doi: 10.1016/j.strusafe.2016.01.006

Gehl P, Douglas J, Rossetto T, Macabuag J, Nassirpour A, Minas S, Duffour P (2014) Investigating the use of record-to-record variability in static capacity approaches. In: 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM2014) & 6th International Symposium on Uncertainty Modelling and Analysis (ISUMA2014). 13-16 July. Liverpool

Gehl P, Douglas J, Seyedi DM (2015) Influence of the Number of Dynamic Analyses on the Accuracy of Structural Response Estimates. Earthq Spectra 31:97–113 . doi: 10.1193/102912EQS320M

Gehl P, Seyedi DM, Douglas J (2013) Vector-valued fragility functions for seismic risk evaluation. Bull Earthq Eng 11:365–384 . doi: 10.1007/s10518-012-9402-7

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian Data Analysis, Third Edition. CRC Press. ISBN: 9781439840955

Gidaris I (2015) Risk assessment and optimal design of seismic protective systems through surrogate and reduced ordered modelling. PhD Thesis, University of Notre Dame

Gidaris I, Taflanidis AA, Mavroeidis GP (2015) Kriging metamodeling in seismic risk

assessment based on stochastic ground motion models. Earthq Eng Struct Dyn 2377–2399 . doi: 10.1002/eqe.2586

Giovanis DG, Papadopoulos V (2015) Spectral representation-based neural network assisted stochastic structural mechanics. Eng Struct 84:382–394 . doi: 10.1016/j.engstruct.2014.11.044

Giovenale P, Cornell CA, Esteva L (2004) Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses. Earthq Eng Struct Dyn 33:951–979 . doi: 10.1002/eqe.386

Giovinazzi S, Kongar I, Bocchini GM, Ottonelli D (2014) Damage to Buildings: Modeling. In: Encyclopedia of Earthquake Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–20

Goda K, Atkinson GM (2009) Probabilistic characterization of spatially correlated response spectra for earthquakes in Japan. Bull Seismol Soc Am 99:3003–3020 . doi: 10.1785/0120090007

Grünthal G (1998) European Macroseismic Scale 1998. Luxembourg. ISBN: 2879770084

Hancock J, Watson-Lamprey J, Abrahamson N, et al (2006) An improved method of matching response spectra of recorded earthquake ground motion using wavelets. J Earthq Eng 10:67–89 . doi: 10.1080/13632460609350629

Haselton CB (2007) Assessing seismic collapse safety of modern reinforced concrete moment frame buildings. PhD Thesis, Stanford University

Haselton CB, Liel AB, Deierlein GG, Dean BS, Chou JH (2011) Seismic Collapse Safety of Reinforced Concrete Buildings. I: Assessment of Ductile Moment Frames. J Struct Eng 137:481–491 . doi: 10.1061/(ASCE)ST.1943-541X.0000318

Hastie T, Tibshirani R, Friedman J (2001) The Elements of Statistical Learning: Data Mining,

Inference and Prediction. Springer, New York, USA

Hill MP (2011) A Framework for the Seismic Risk Assessment of Existing Buildings in Europe. PhD Thesis, University College London

Housner GW (1952) Spectrum Intensities of Strong-Motion Earthquakes. Symp Earthq Blast Eff Struct 20–36

Ibarra LF, Krawinkler H (2005) Global Collapse of Frame Structures under Seismic Excitations Global Collapse of Frame Structures under Seismic Excitations. Berkeley, California, USA

Iervolino I, Galasso C, Cosenza E (2009) REXEL: computer aided record selection for code-based seismic structural analysis. Bull Earthq Eng 8:339–362 . doi: 10.1007/s10518-009-9146-1

Iervolino I, Giorgio M, Galasso C, Manfredi G (2010) Conditional hazard maps for secondary intensity measures. Bull Seismol Soc Am 100:3312–3319 . doi: 10.1785/0120090383

Jalayer F (0003) Direct probabilistic seismic analysis: implementing non-linear dynamic assessments. PhD Thesis, Stanford University

Jalayer F, Beck JL, Zareian F (2012) Analyzing the Sufficiency of Alternative Scalar and Vector Intensity Measures of Ground Shaking Based on Information Theory. J Eng Mech 138:307–316 . doi: 10.1061/(ASCE)EM.1943-7889.0000327

Jalayer F, Cornell CA (2009) Alternative non-linear demand estimation methods for probability-based seismic assessments. Earthq Eng Struct Dyn 38:951–972 . doi: 10.1002/eqe.876

Jalayer F, Cornell CA (2002) Alternative non-linear demand estimation methods for probability-based seismic assessments. Earthq Eng Struct Dyn 38:951–972 . doi: 10.1002/eqe.876

Jalayer F, De Risi R, Elefante L, Manfredi G (2013) Robust fragility assessment using Bayesian parameter estimation. In: Vienna Congress on Recent Advances in Earthquake Engineering and Structural Dynamics 2013 (VEESD 2013) 28-30 August 2013, Vienna, Austria. pp 28–30

Jalayer F, De Risi R, Manfredi G (2014) Bayesian Cloud Analysis: efficient structural fragility assessment using linear regression. Bull Earthq Eng 13:1183–1203 . doi: 10.1007/s10518-014-9692-z

Jalayer F, Ebrahimian H, Miano A, Manfredi G, Sezen H (2017) Analytical fragility assessment using unscaled ground motion records. Earthq Eng Struct Dyn 46:2639–2663 . doi: 10.1002/eqe.2922

Jalayer F, Elefante L, Iervolino I, Manfredi G (2011) Knowledge-Based Performance Assessment of Existing RC Buildings. J Earthq Eng 15:362–389 . doi: 10.1080/13632469.2010.501193

Jalayer F, Franchin P, Pinto P (2007) A scalar decision variable for seismic reliability analysis of RC frames. Earthq Eng Struct Dyn 36:2059–2079 . doi: 10.1002/eqe704

Jayaram N, Baker JW (2009) Correlation model for spatially distributed ground-motion intensities. Earthq Eng Struct Dyn 38:1687–1708 . doi: 10.1002/eqe.922

Jayaram N, Baker JW (2008) Statistical tests of the joint distribution of spectral acceleration values. Bull Seismol Soc Am 98:2231–2243 . doi: 10.1785/0120070208

Jeon J, Park J, Desroches R (2015) Seismic fragility of lightly reinforced concrete frames with masonry in fi lls. Earthq Eng Struct Dyn 44:1783–1803 . doi: 10.1002/eqe.2555

Jia G, Taflanidis AA (2013) Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. Comput Methods Appl Mech Eng 261–262:24–38 . doi: 10.1016/j.cma.2013.03.012

Jin R, Chen W, Simpson TW (2001) Comparative Studies of Metamodeling Techniques Under Multiple Modeling Criteria. Struct Multidiscip Optim 23:1–13 . doi: 10.2514/6.2000-4801

Kadas K, Yakut A, Kazaz I (2011) Spectral Ground Motion Intensity Based on Capacity and Period Elongation. J Struct Eng 137:401–409 . doi: 10.1061/(ASCE)ST.1943-541X.0000084

Kalkan E, Kunnath SK (2007) Assessment of current nonlinear static procedures for seismic evaluation of buildings. Eng Struct 29:305–316 . doi: 10.1016/j.engstruct.2006.04.012

Karapetrou S (2015) Seismic vulnerability of reinforced concrete buildings considering aging and soil-structure interaction effects. PhD Thesis, Aristotle University of Thessaloniki

Katsanos EI, Sextos AG (2013) ISSARS: An integrated software environment for structure-specific earthquake ground motion selection. Adv Eng Softw 58:70–85 . doi: 10.1016/j.advengsoft.2013.01.003

Kazantzi AK, Vamvatsikos D (2015) Intensity measure selection for vulnerability studies of building classes. Earthq Eng Struct Dyn 44:2677–2694

Kennedy MC, O'Hagan A (2001) Bayesian Calibration of Computer Models. J R Stat Soc Ser B (Statistical Methodol 63:425–464 . doi: 10.1111/1467-9868.00294

Kohavi R (1995) A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Appear Int Jt Conf Articial Intell 5:1–7 . doi: 10.1067/mod.2000.109031

Kohrangi M, Bazzurro P, Vamvatsikos D (2016a) Vector and scalar IMs in structural response estimation, Part I: Hazard Analysis. Earthq Spectra 32:1507–1524 . doi: 10.1193/053115EQS080M

Kohrangi M, Bazzurro P, Vamvatsikos D (2016b) Vector and scalar IMs in structural response estimation, Part II: Building demand assessment. Earthq Spectra 32:1525–1543 . doi: 10.1193/053115EQS081M

Kohrangi M, Bazzurro P, Vamvatsikos D, Spillatura A (2017a) Conditional spectrum-based ground motion record selection using average spectral acceleration. Earthq Eng Struct Dyn 46:1667–1685 . doi: 10.1002/eqe.2876

Kohrangi M, Kotha SR, Bazzurro P (2018) Ground-motion models for average spectral acceleration in a period range: Direct and indirect methods. Bull Earthq Eng 16:45–65 . doi: 10.1007/s10518-017-0216-5

Kohrangi M, Vamvatsikos D, Bazzurro P (2017b) Site dependence and record selection schemes for building fragility and regional loss assessment. Earthq Eng Struct Dyn 46:1625–1643 . doi: DOI: 10.1002/eqe.2873

Kohrangi M, Vamvatsikos D, Bazzurro P (2016c) Implications of intensity measure selection for seismic loss assessment of 3-D buildings. Earthq Spectra 32:2167–2189 . doi: 10.1193/112215EQS177M

Kramer S (1996) Geotechnical earthquake engineering. Pearson Prentice Hall, Upper Saddle River, NJ, USA. ISBN: 9781439840955

Krige D (1952) A statistical approach to some basic mine valuation problems on the Witwatersrand. J Chem Metall Min Soc South Africa 52:119–139

Kroetz HM, Beck AT (2015) Performance of Surrogate Modeling Techniques in Structural Reliability. In: 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12. pp 1–7

Lagomarsino S, Giovinazzi S (2006) Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. Bull Earthq Eng 4:415–443 . doi: 10.1007/s10518-006-9024-z

Lallemant D, Kiremidjian A, Burton H (2015) Statistical procedures for developing earthquake damage fragility curves. Earthq Eng Struct Dyn 44:1373–1389 . doi: 10.1002/eqe.2522

Lang K, Bachman R (2004) On the seismic vulnerability of existing buildings: a case study of the city of Basel. Earthq Spectra 20:43–66 . doi: 10.1193/1.1648335

Liel AB, Deierlein GG (2012) Using Collapse Risk Assessments to Inform Seismic Safety Policy for Older Concrete Buildings. Earthq Spectra 28:1495–1521 . doi: 10.1193/1.4000090

Liel AB, Haselton CB, Deierlein GG (2011) Seismic collapse safety of reinforced concrete buildings. II: Comparative assessment of nonductile and ductile moment frames. J Struct Eng 137:492–502

Lin L, Naumoski N, Saatcioglu M, Foo S (2011) Improved intensity measures for probabilistic seismic demand analysis. Part 1: development of improved intensity measures. Can J Civ Eng 38:79–88 . doi: 10.1139/L10-110

Lin T, Baker JW (2013) Introducing adaptive incremental dynamic analysis: A new tool for linking ground motion selection and structural response assessment. In: ICOSSAR International Conference for Structural Safety and Reliability. New York, USA

Lin T, Haselton CB, Baker JW (2013) Conditional spectrum-based ground motion selection. Part I: Hazard consistency for risk-based assessments. Earthq Eng Struct Dyn 42:1847–1865 . doi: 10.1002/eqe2301

Lophaven SN, Søndergaard J, Nielsen HB (2002) DACE A Matlab Kriging Toolbox. IMM Informatiocs Math Model 1–28

Loth C, Baker JW (2013) A spatial cross-correlation model of spectral accelerations at multiple periods. Earthq Eng Struct Dyn 397–417 . doi: 10.1002/eqe

Lu D, Yu X, Pan F, Wang G (2008) Probabilistic Seismic Demand Analysis considering Random System properties by an improved cloud method. In: 14th World Conference on Earthquake Engineering. Beijing, China

Lucchini A, Mollaioli F, Monti G (2011) Intensity measures for response prediction of a torsional building subjected to bi-directional earthquake ground motion. Bull Earthq Eng 9:1499–1518 . doi: 10.1007/s10518-011-9258-2

Luco N, Cornell CA (2007) Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthq Spectra 23:357–392 . doi: 10.1193/1.2723158

Luco N, Cornell CA (1998) Seismic drift demands for two SMRF structures with brittle connections. Structural Engineering World Wide. Struct Eng World Wide; Elsevier Sci Ltd, Oxford, Engl paper T158-3

Mander J, Priestley M, Park R (1989) Theoretical stress-strain model for confined concrete. J Struct Eng 114:1804–1826

McCullagh P, Nelder J (1989) Generalized Linear Model, 2nd edn. Chapman & Hall/CRC, Boca Raton, Florida, USA. ISBN: 0412317605

McKay MD, Beckman RJ, Canover WJ (1979) A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. Technometrics 21:239–245

McKenna F, Fenves G, Scott M (2006) OpenSees: open system for earthquake engineering simulation.

Medvedev S V., Sponheuer W (1969) Scale of Seismic Intensity. Proc. IV World Conf. Earthq. Eng. 143–153

Mehanny SSF (2009) A broad-range power-law form scalar-based seismic intensity measure. Eng Struct 31:1354–1368 . doi: 10.1016/j.engstruct.2009.02.003

Meletti C, Galadini F, Valensise G, Stucchi M, Basili R, Barba S, Vannucci G, Boschi E (2008) A seismic source zone model for the seismic hazard assessment of the Italian territory.

Tectonophysics 450:85108

Menegotto M, Pinto P (1973) Method of analysis for cyclically loaded RC plane frames including changes in geometry and non-elastic behaviour of elements under combined normal force and bending. In: Proceedings of Symposium on the Resistance and Ultimate Deformability of Structures anted on by well defined loads. Lisbon, Portugal

Milutinovic ZV, Trendafiloski GS (2003) Vulnerability of current buildings, work-package 4 of RISK_UE Project

Minas S, Chandler RE, Rossetto T (2018) BEA: An efficient Bayesian emulation-based approach for probabilistic seismic response. Struct Saf 74: . doi: 10.1016/j.strusafe.2018.04.002

Minas S, Galasso C, Rossetto T (2015) Spectral Shape Proxies and Simplified Fragility Analysis of Mid- Rise Reinforced Concrete Buildings. In: 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12. Vancouver, Canada

Modica A, Stafford PJ (2014) Vector fragility surfaces for reinforced concrete frames in Europe. Bull Earthq Eng 12:1725–1753 . doi: 10.1007/s10518-013-9571-z

Mollaioli F, Lucchini A, Cheng Y, Monti G (2013) Intensity measures for the seismic response prediction of base-isolated buildings. Bull Earthq Eng 11:1841–1866 . doi: 10.1007/s10518-013-9431-x

Nassirpour A (2018) Performance based seismic assessment of masonry infilled steel frame structures. University College London

National Institute of Building Scinces (NIBS) (2003) Multi-hazard Loss Estimation Methodology - earthquake model. HAZUS MR4 Technical Manual. Washington, D.C., USA

National Institute of Standards and Technology (NIST) (2011) Selecting and Scaling Earthquake Ground Motions for Performing Response-History Analyses. NIST GCR 11-917-15. Gaithersburg, Maryland

O'Hagan A (2006) Bayesian analysis of computer code outputs: A tutorial. Reliab Eng Syst Saf 91:1290–1300 . doi: 10.1016/j.ress.2005.11.025

Oliver DS (2003) Gaussian cosimulation: Modelling of the cross-covariance. Math Geol 35:681–698 . doi: 10.1023/B:MATG.0000002984.56637.ef

Onem G (2008) Evaluation of practice-oriented nonlinear analysis methods for seismic performance assessment. PhD Thesis, Bogazisi University

OpenStreetMap (2017) Map data © OpenStreetMap contributors. CC_BY_SA. Available at: https://www.openstreetmap.org/

Pacor F, Luzi L (2014) Engineering Characterization of Earthquake Ground Motions. In: Encyclopedia of Earthquake Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–18

Pacor F, Paolucci R, Luzi L, Sabetta F, Spinelli A, Gorini A, Nicoletti M, Marcucci S, Filippi L, Dolce M (2011) Overview of the Italian strong motion database ITACA 1.0. Bull Earthq Eng 9:1723–1739 . doi: 10.1007/s10518-011-9327-6

Padgett JE, DesRoches R (2007) Sensitivity of Seismic Response and Fragility to Parameter Uncertainty. J Struct Eng 133:1710–1718 . doi: 10.1061/(ASCE)0733-9445(2007)133:12(1710)

Padgett JE, Nielson BG, Desroches R (2008) Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. Earthq Eng Struct Dyn 37:711–725 . doi: 10.1002/eqe.782

Picheny V (2009) Improving Accuracy and Compesating for Uncertainty in Surrogate

Modeling. University of Florida

Picheny V, Ginsbourger D, Roustant O, Haftka RT, Kim NH (2010) Adaptive designs of experiments for accurate approximation of a target region. J Mech Des 132:071008 . doi: 10.1115/1.4001873

Pitilakis K, Argyroudis S, Kakderi K, Argyroudi A (2013) SYNER-G D8.8: Systemic Seismic Vulnerability and Risk Analysis for Buildings, Lifeline Networks and Infrastructures Safety Gain. Publications Office of the European Union, Luxembourg

Polese M, Di Ludovico M, Prota A, Manfredi G (2013) Damage-dependent vulnerability curves for existing buildings. Earthq Eng Struct Dyn 42:853–870 . doi: 10.1002/eqe.2249

Porter K (2005) A Taxonomy of Building Components for Performance-Based Earthquake Engineering. PEER Report 2005/03. Berkeley, CA

Porter K, Kennedy R, Bachman R (2007) Creating Fragility Functions for Performance-Based Earthquake Engineering. Earthq Spectra 23:471–489 . doi: 10.1193/1.2720892

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Ramamoorthy SK, Gardoni P, Bracci JM (2006) Probabilistic Demand Models and Fragility Curves for Reinforced Concrete Frames. J Struct Eng 132:1563–1572 . doi: 10.1061/(ASCE)0733-9445(2006)132:10(1563)

Rasmussen CE, Williams CKI (2006) Gaussian Processes for Machine Learning. MIT Press

Regio Decreto 16/11/1939 n. 2229 (1940) Norme per la esecuzione delle opere in conglomerato cementizio semplice e armato., G.U. n. 92

Ribeiro Jr P, Diggle P (2001) geoR: A package for geostatistical analysis. R-NEWS 1:15–18

Ricci P (2010) Seismic vulnerability of existing RC buildings. PhD Thesis, University of Naples

Federico II

Rojahn C, Poland C, Scawthorn C (1988) Rapid visual screening of buildings for potential seismic hazards: A handbook. Redwood City, California

Rossetto T (2004) Vulnerability Curves for the Seismic Assessment of Reinforced Concrete Building Populations. PhD Thesis, Imperial College

Rossetto T, Alexander D, Verrucci E, et al (2012) The 20th May 2012 Emilia Romagna Earthquake. London, UK

Rossetto T, Elnashai AS (2003) Derivation of vulnerability functions for European-type RC structures based on observational data. Eng Struct 25:1241–1263 . doi: 10.1016/S0141-0296(03)00060-9

Rossetto T, Elnashai AS (2005) A new analytical procedure for the derivation of displacement-based vulnerability curves for populations of RC structures. Eng Struct 27:397–409 . doi: 10.1016/j.engstruct.2004.11.002

Rossetto T, Gehl P, Minas S, Galasso C, Duffour P, Douglas J, Cook O (2016a) FRACAS: A capacity spectrum approach for seismic fragility assessment including record-to-record variability. Eng Struct 125: . doi: 10.1016/j.engstruct.2016.06.043

Rossetto T, Gehl P, Minas S, Duffour P, Douglas J, Galasso C, Nassirpour A (2016b) Fragility Through Capacity Spectrum Assessment v.1.0 User Manual. London, UK

Rossetto T, Gehl P, Minas S, Nassirpour A, Macabuag J, Duffour P (2014a) Sensitivity analysis of different capacity approaches to assumptions in the modeling, capacity and demand representations. In: 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM2014) & 6th International Symposium on Uncertainty Modelling and Analysis (ISUMA2014). 13-16 July. Liverpool, UK

Rossetto T, Ioannou I, Grant DN, Maqsood T (2014b) Guidelines for Empirical Vulnerability

Assessment Report produced in the context of the Vulnerability Global Component project. Pavia, Italy

Rossetto T, Novelli V (2016) Indirect Vulnerability Estimation Methods: Choice of Damage Scales and Conversion from Damage to Loss. GEM Technical Report 2016 - X. Pavia, Italy

Sacks J, Welch WJ, Mitchell JSB, Henry PW (1989) Design and Experiments of Computer Experiments. Stat. Sci. 4:409–423

Saltelli A, Chan K, Scott E (2000) Sensitivity analysis. Wiley series in probability and statistics. Wiley, Chichester

Schobi R (2017) Surrogate models for uncertainty quantification in the context of imprecise probability modelling. PhD Thesis, ETH Zurich

SEAOC (1995) Vision 2000 : Performance-based Seismic Engineering. Sacramento, California, USA

Seifried AE, Baker JW (2014) Spectral Variability and Its Relationship To Structural Response Estimated From Scaled and Spectrum-Matched Ground Motions. In: 10th U.S. National Conference on Earthquake Engineering, Frontiers of Earthquake Engineering, July 21-25. Anchorage, Alaska, USA

Seismosoft (2007) SeismoStruct: A computer program for static and dynamic nonlinear analyzes of framed structures

Seismosoft (2013) SeismoMatch v2.1 – A computer program for spectrum matching of earthquake records.

Sextos AG (2014) Selection of Ground Motions for Response History Analysis. In: Encyclopedia of Earthquake Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–10

Shinozuka M, Feng MQ, Lee J, Naganuma T (2000) Statistical Analysis of Fragility Curves. J Eng Mech 126:1224–1231 . doi: 10.1061/(ASCE)0733-9399(2000)126:12(1224)

Shome N (2014) Seismic Loss Assessment. In: Encyclopedia of Earthquake Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 1–23

Shome N, Cornell CA (1999) Probabilistic seismic demand analysis of nonlinear structures. Report No. RMS-35

Shome N, Cornell CA, Bazzurro P, Carballo JE (1998) Earthquakes, records, and nonlinear responses. Earthq Spectra 14:469–500 . doi: 10.1193/1.1586011

Silva V (2013) Development of open models and tools for seismic risk assessment: Application to Portugal. PhD Thesis, University of Aveiro

Silva V (2018) Uncertainty and correlation in seismic vulnerability functions of building classes. Earthq Spectra

Silva V, Akkar S, Baker JW, Bazzurro P, Castro JM, Crowley H, Dolšek M, Galasso C, Lagomarsino S, Monteiro R, Perrone D, Pitilakis K, Vamvatsikos D. Current Challenges and Future Trends in Analytical Vulnerability Modelling (under review). Earthq Spectra

Silva V, Crowley H, Varum H, Pinho R, Sousa R (2014a) Evaluation of analytical methodologies used to derive vulnerability functions. Earthq Eng Struct Dyn 181–204 . doi: 10.1002/ eqe.2337

Silva V, Crowley H, Varum H, Pinho R, Sousa L (2014b) Investigation of the characteristics of Portuguese regular moment-frame RC buildings and development of a vulnerability model. Bull Earthq Eng 1455–1490 . doi: 10.1007/s10518-014-9669-y

Smerzini C, Galasso C, Iervolino I, Paolucci R (2013) Ground motion record selection based on broadband spectral compatibility. Earthq Spectra 30:1427–1448 . doi: 10.1193/052312EQS197M

Sokolov V, Wenzel F, Jean WY, Wen KL (2010) Uncertainty and spatial correlation of earthquake ground motion in Taiwan. Terr Atmos Ocean Sci 21:905–921 . doi: 10.3319/TAO.2010.05.03.01(T)

Sousa L, Silva V, Bazzurro P (2017) Using Open-Access Data in the Development of Exposure Data Sets of Industrial Buildings for Earthquake Risk Modeling. Earthq Spectra 33:63–84 . doi: 10.1193/020316EQS027M

Stafford PJ, Bommer JJ (2010) Theoretical consistency of common record selection strategies in performance-based earthquake engineering. In: Fardis M (ed) Advances in Performance-Based Earthquake Engineering, Geotechnical, Geological and Earthquake Engineering. Springer Netherlands, p 49–58, Vol:13

Stucchi M, Meletti C, Montaldo V, Crowley H, Calvi GM, Boschi E (2011) Seismic Hazard Assessment (2003-2009) for the Italian Building Code. Bull Seismol Soc Am 101:1885–1911 . doi: 10.1785/0120100130

Tarantola A (2005) Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia, USA. ISBN: 0898715725

Tenerelli P, Crowley H (2013) SYNER-G D8.9: Development of inventory datasets through remote sensing and direct observation data for earthquake loss estimation. Publications Office of the European Union, Luxembourg

Trifunac M, Todorovska M (2001) Evolution of accelerographs, data processing, strong motion arrays and amplitude and spatial resolution in recording strong earthquake motion. Soil Dyn Earthq Eng 537–555. doi: 10.1016/S0267-7261(01)00013-6

Vamvatsikos D, Allin Cornell C (2006) Direct estimation of the seismic demand and capacity of oscillators with multi-linear static pushovers through IDA. Earthq Eng Struct Dyn 35:1097–1117 . doi: 10.1002/eqe.573

Vamvatsikos D, Cornell CA (2005) Developing efficient scalar and vector intensity measures for IDA capacity estimation by incorporating elastic spectral shape information. Earthq Eng Struct Dyn 34:1573–1600 . doi: 10.1002/eqe.496

Vamvatsikos D, Cornell CA (2002) Incremental dynamic analysis. Earthq Eng Struct Dyn 31:491–514 . doi: 10.1002/eqe.141

Vamvatsikos D, Cornell CA (2004) Applied incremental dynamic analysis. Earthq Spectra 20:523–553 . doi: 10.1193/1.1737737

Von Thun J, Rochim L, Scott G, Wilson J (1988) Earthquake ground motions for design and analysis of dams. Earthq Eng Soil Dyn II - Recent Adv Ground-Motion Eval Geotech Spec Publ 20:463–481

Wang Z, Pedroni N, Zentner I, Zio E (2018) Seismic fragility analysis with artificial neural networks: Application to nuclear power plant equipment. Eng Struct 162:213–225 . doi: 10.1016/j.engstruct.2018.02.024

Weatherill GA, Crowley H, Pinho R, Franchin P, Cavalieri F, Iervolino I, Esposito S (2011) SYNER-G D2.13: A review and preliminary application of methodologies for the generation of earthquake scenarios for spatially distributed systems. Publications Office of the European Union, Luxembourg

Weatherill GA, Silva V, Crowley H, Bazzurro P (2015) Exploring the impact of spatial correlations and uncertainties for portfolio analysis in probabilistic seismic loss estimation. Bull Earthq Eng 13:957–981 . doi: 10.1007/s10518-015-9730-5

Yahyaabadi A, Tehranizadeh M (2011) New scalar intensity measure for near-fault ground motions based on the optimal combination of spectral responses. Sci Iran 18:1149–1158 . doi: 10.1016/j.scient.2011.09.013

Yamin LE, Hurtado A, Rincon R, Dorado JF, Reyes JC (2017) Probabilistic seismic

vulnerability assessment of buildings in terms of economic losses. Eng Struct 138:308–

323 . doi: 10.1016/j.engstruct.2017.02.013

Appendix A: List of publications and outputs

Journals

Rossetto T, Gehl P, Minas S, Galasso C, Duffour P, Douglas J, Cook O (2016) FRACAS: A capacity spectrum approach for seismic fragility assessment including record-to-record variability. Eng Struct 125:337–348 . doi: 10.1016/j.engstruct.2016.06.043

Minas S, Chandler RE, Rossetto T (2018) BEA: An efficient Bayesian emulation-based approach for probabilistic seismic response. Struct Saf 74: . doi: 10.1016/j.strusafe.2018.04.002

Minas S, Galasso C (in press) Accounting for spectral shape in simplified fragility analysis of mid-rise reinforced concrete buildings. Soil Dyn Earthq Eng

Books & Reports

D'Ayala D, Galasso C, Minas S, Novelli, V (2015) Review of retrofitting methods to reduce seismic vulnerability of buildings, with particular reference to hospitals and medical facilities Review of methods to assess the seismic vulnerability of buildings, with particular reference to hospitals and medical facilities. Evidence on Demand, UK. 1–33. doi: 10.12774/eod_hdr.june2015.ddayalaetal

D'Ayala D, Galasso C, Minas S, Novelli, V (2015) Review of retrofitting methods to reduce seismic vulnerability of buildings, with particular reference to hospitals and medical facilities. Evidence on Demand, UK. 1–34. doi: 10.12774/eod_hdr.june2015.ddayalaetal1

D'Ayala D, Galasso C, Minas S, Novelli, V (2015) Review of the non-structural considerations for seismically retrofitting hospitals, impact on hospital functionality, and hospital selection. Evidence on Demand, UK. 1–30. doi: 10.12774/eod_hdr.june2015.ddayalaetal2

Rossetto T, Verrucci E, Minas S (2018) MOVER – Level 3 Data schema for Physical and Social Vulnerability Indicators, Indices, and Functions. Global Facility for Disaster

Reduction and Recovery (GFDRR). Report available online at: https://www.gfdrr.org/sites/default/files/publication/MOVER%20Level%203_4%20-%20Physical%20and%20Social%20vulnerability%20data%20schemas_UPDATED.pdf

Galasso C, Minas S (2017) Report on building-specific structural risk assessments of priority, High-risk school buildings in the Province of Surigao del Norte, Philippines Rapid Structural Assessment of Public Buildings. World Bank.

Galasso C, Minas S (2018) Report on building-specific structural risk assessments of priority, High-risk school buildings in Metro Manila and DepEd Central Office. World Bank.

Conferences

Gehl P, Douglas J, Rossetto T, Macabuag J, Nassirpour A, Minas S, Duffour P (2014) Investigating the use of record-to-record variability in static capacity spectrum approaches. In: 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM2014) & 6th International Symposium on Uncertainty Modelling and Analysis (ISUMA2014). 13-16 July. Liverpool, UK (Oral presentation)

Rossetto T, Gehl P, Minas S, Nassirpour A, Macabuag J, Duffour P, Douglas J (2014) Sensitivity analysis of different capacity approaches to assumptions in the modeling, capacity and demand representations. In: 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM2014) & 6th International Symposium on Uncertainty Modelling and Analysis (ISUMA2014). 13-16 July. Liverpool, UK (Oral presentation)

Minas S, Galasso C, Rossetto T (2014) Preliminary investigation on selecting optimal intensity measures for simplified fragility analysis of mid-rise RC buildings. In: 2nd European Conference on Earthquake Engineering and Seismology, August 25-29, 2014. Istanbul, Turkey (Poster presentation)

Minas S, Galasso C, Rossetto T (2015a) Assessing spectral shape-based intensity measures for simplified fragility analysis of mid-rise reinforced concrete buildings. In: SECED 2015

Conference: Earthquake Risk and Engineering towards a Resilient World, July 9-10, 2015, Cambridge, UK (Oral presentation)

Minas S, Galasso C, Rossetto T (2015b) Spectral shape proxies and simplified fragility analysis of mid-rise reinforced concrete buildings. In: 12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP 2015, Vancouver, Canada (Oral presentation – Recipient of CERRA student recognition award)

Minas S, Chandler RE, Rossetto T (2017) New emulation-based approach for probabilistic seismic demand. In: 16th World Conference on Earthquake Engineering, 16WCEE 2017. Santiago, Chile (Oral presentation)

Minas S, Sousa L, Galasso C, Rossetto T (2018) Sensitivity of probabilistic regional seismic loss to hazard and vulnerability modelling options. In: 16th European Conference on Earthquake Engineering, June 18-21, 2018, Thessaloniki, Greece (Oral presentation)

Future publications (in progress)

Minas S, Sousa L, Galasso C, Rossetto T (in progress) Probabilistic regional seismic loss assessment considering hazard and vulnerability modelling options.

Other outputs

Co-editor of the online course "Simplified Structural Assessment for Fragility Analysis", available online at: https://extend.ucl.ac.uk/ . This course presents state-of-the-art techniques for the construction of fragility curves.

Co-developer of FRACAS software: This is software for simplified analytical method for fragility assessment. A detailed review of FRACAS methodology and capabilities can be found in Chapter 4 of this Thesis. FRACAS is freely available to download at: https://www.ucl.ac.uk/epicentre/resources/software

Developer of FRACAS tool: This is an extension of FRACAS software, which provides a simple and easy to use GUI for the execution of FRACAS analysis.

Developer of the BEA code: This code implements the Bayesian emulation-based approach (BEA) to compute the probabilistic seismic response of buildings. The BEA is freely available to download at: https://www.ucl.ac.uk/epicentre/resources/software

Co-developer of the Multi-Hazard Open Vulnerability Platform for Evaluating Risk (MOVER), which assesses the physical and social vulnerability for African countries. Proposed a scoring system for the usability and applicability of vulnerability models in the geographical area of interest.

Co-developer of an automated procedure for seismic performance and vulnerability assessment of school buildings as a part of Global Library of School Infrastructure (GLoSI) data repository project, for the World Bank.

Appendix B: Damage scales

Table B 1 - HAZUS Damage scale for RC MRF building, referenced as C1 (adapted from: FEMA 2015).

Damage state		Description
	Slight	Some beams and columns have flexural or shear-type hairline cracks in, near, or within joints.
	Moderate	Most beams and columns exhibit hairline cracks. In ductile frames, some of the frame elements have reached yield capacity, indicated by larger flexural cracks and some concrete spalling. Nonductile frames may exhibit larger shear cracks and spalling.
	Extensive	Some frame elements have reached their ultimate capacity, indicated by large flexural cracks, spalled concrete, and buckled main reinforcement. Nonductile frame elements may have suffered shear or bond failures at reinforcement splices or broken ties or buckled main reinforcement in columns, which may result in partial collapse.
	Complete	Structure has collapsed or is in imminent danger of doing so because of brittle failure of nonductile frame elements or loss of frame stability.

Table B 2 - EMS-98 damage scale description of damage states for RC buildings (Grünthal 1998).
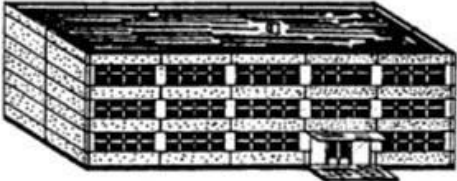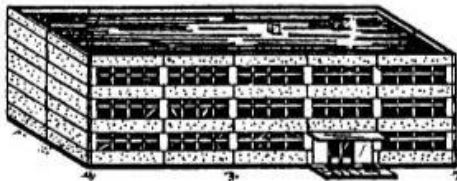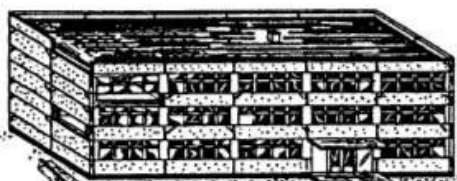
Classification of damage to buildings of reinforced concrete	
	Grade 1: Negligible to slight damage **(no structural damage,** **slight non-structural damage)** Fine cracks in plaster over frame members or in walls at the base. Fine cracks in partitions and infills.
	Grade 2: Moderate damage **(slight structural damage,** **moderate non-structural damage)** Cracks in columns and beams of frames and in structural walls. Cracks in partition and infill walls; fall of brittle cladding and plaster. Falling mortar from the joints of wall panels.
	Grade 3: Substantial to heavy damage **(moderate structural damage,** **heavy non-structural damage)** Cracks in columns and beam column joints of frames at the base and at joints of coupled walls. Spalling of conrete cover, buckling of reinforced rods. Large cracks in partition and infill walls, failure of individual infill panels.
	Grade 4: Very heavy damage **(heavy structural damage,** **very heavy non-structural damage)** Large cracks in structural elements with compression failure of concrete and fracture of rebars; bond failure of beam reinforced bars; tilting of columns. Collapse of a few columns or of a single upper floor.
	Grade 5: Destruction **(very heavy structural damage)** Collapse of ground floor or parts (e. g. wings) of buildings.

Table B 3 - Vision 2000 performance levels (SEAOC 1995).

System Description	Performance Level				
	10 Fully Operational 9	8 Operational 7	6 Life Safe 5	4 Near Collapse 3	2 Collapse 1
Overall building damage	Negligible	Light	Moderate	Severe	Complete
Permissible transient drift	< 0.2%+/-	< 0.5%+/-	< 1.5%+/-	< 2.5%+/-	> 2.5%+/-
Permissible permanent drift	Negligible.	Negligible.	< 0.5%+/-	< 2.5%+/-	> 2.5%+/-
Vertical load carrying element damage	Negligible.	Negligible.	Light to moderate, but substantial capacity remains to carry gravity loads.	Moderate to heavy, but elements continue to support gravity loads.	Partial to total loss of gravity load support.
Lateral Load Carrying Element damage	Negligible - generally elastic response; no significant loss of strength or stiffness.	Light - nearly elastic response; original strength and stiffness substantially retained. Minor cracking/yielding of structural elements; repair implemented at convenience.	Moderate - reduced residual strength and stiffness but lateral system remains functional.	Negligible residual strength and stiffness. No story collapse mechanisms but large permanent drifts. Secondary structural elements may completely fail.	Partial or total collapse. Primary elements may require demolition.
Damage to architectural systems	Negligible damage to cladding, glazing, partitions, ceilings, finishes, etc. Isolated elements may require repair at users convenience.	Light to moderate damage to architectural systems. Essential and select protected items undamaged. Hazardous materials contained.	Moderate to severe damage to architectural systems, but large falling hazards not created. Major spills of hazardous materials contained.	Severe damage to architectural systems. Some elements may dislodge and fall.	Highly dangerous falling hazards. Destruction of components.
Egress systems	Not impaired.	No major obstructions in exit corridors. Elevators can be restarted perhaps following minor servicing.	No major obstructions in exit corridors. Elevators may be out of service for an extended period.	Egress may be obstructed.	Egress may be highly or completely obstructed.

Table B 4 - Extract of structural damage states for concrete frames from FEMA 356 (FEMA 2000).

Table C1-3   Structural Performance Levels and Damage[1, 2, 3]—Vertical Elements

Elements	Type	Structural Performance Levels		
		Collapse Prevention S-5	Life Safety S-3	Immediate Occupancy S-1
Concrete Frames	Primary	Extensive cracking and hinge formation in ductile elements. Limited cracking and/or splice failure in some nonductile columns. Severe damage in short columns.	Extensive damage to beams. Spalling of cover and shear cracking (<1/8" width) for ductile columns. Minor spalling in nonductile columns. Joint cracks <1/8" wide.	Minor hairline cracking. Limited yielding possible at a few locations. No crushing (strains below 0.003).
	Secondary	Extensive spalling in columns (limited shortening) and beams. Severe joint damage. Some reinforcing buckled.	Extensive cracking and hinge formation in ductile elements. Limited cracking and/or splice failure in some nonductile columns. Severe damage in short columns.	Minor spalling in a few places in ductile columns and beams. Flexural cracking in beams and columns. Shear cracking in joints <1/16" width.
	Drift	4% transient or permanent	2% transient; 1% permanent	1% transient; negligible permanent

Table B 5 - HRC damage scale description of damage states for ductile, non-ductile and infilled RC moment resisting frames (Rossetto and Elnashai 2003).

DI$_{HRC}$	DAMAGE STATE	DUCTILE MRF	NON-DUCTILE MRF	INFILLED MRF
0	None	No damage	No damage	No damage
10	Slight	Fine cracks in plaster partitions/infills	Fine cracks in plaster partitions/infills	Fine cracks in plaster partitions/infills
20	Light	Start of structural damage	Start of structural damage	Cracking at wall-frame interfaces
30		Hairline cracking in beams and columns near joints (<1mm)	Hairline cracking in beams and columns near joints (<1mm)	Cracking initiates from corners of openings
40				Diagonal cracking of walls. Limited crushing of bricks at b/c connections
50	Moderate	Cracking in most beams & columns	Flexural & shear cracking in most beams & columns	Increased brick crushing at b/c connections
60		Some yielding in a limited number	Some yielding in a limited number	Start of structural damage
70		Larger flexural cracks & start of concrete spalling	Shear cracking & spalling is limited	Some diagonal shear cracking in members especially for exterior frames
80	Extensive	Ultimate capacity reached in some elements – large flexural cracking, concrete spalling & re-bar buckling	Loss of bond at lap-splices, bar pull-out, broken ties	Extensive cracking of infills, falling bricks, out-of-plane bulging
90		Short column failure	Main re-bar may buckle or elements fail in shear	Partial failure of many infills, heavier damage in frame members, some fail in shear
100	Partial Collapse	Collapse of a few columns, a building wing or single upper floor	Shear failure of many columns or impending soft-storey failure	Beams &/or columns fail in shear causing partial collapse. Near total infill failure
	Collapse	Complete or impending building collapse	Complete or isoft-storey failure at ground floor	Complete or impending building collapse

270

Table B 6 - Milutinovic and Trendafiloski (2003) RISK-UE damage scale and equivalences to FEMA/HAZUS.

Damage Grade	Damage Grade Label			Description
	LM1	LM2	FEMA/ NIBS (HAZUS)	
0 (D0)	None	None	None	No damage
1 (D1)	Slight	Minor	Slight	Negligible to slight damage
2 (D2)	Moderate	Moderate	Moderate	Slight structural, moderate nonstructural
3 (D3)	Substantial to heavy	Severe	Extensive	Moderate structural, heavy nonstructural
4 (D4)	Very heavy	Collapse	Complete	Heavy structural, very heavy nonstructural
5 (D5)	Destruction			Very heavy structural, total or near total collapse

Table B 7 - Crowley et al. 2004 damage state description for RC frames.

Structural damage band	Description
None to slight	Linear elastic response, flexural or shear type hairline cracks (<1.0 mm) in some members, no yielding in any critical section; hence limit state to damage band is structural yield point
Moderate	Member flexural strengths achieved, limited ductility developed, crack widths reach 1.0 mm, initiation of concrete spalling, limits to strains may be assumed as: $\varepsilon_c = 0.004$–0.005 $\varepsilon_s = 0.010$–0.015
Extensive	Significant repair required to building, wide flexural or shear cracks, buckling of longitudinal reinforcement may occur, limits to strains may be assumed as: Inadequately confined members: $\varepsilon_c = 0.005$–0.010 $\varepsilon_s = 0.015$–0.030 Adequately confined members: $\varepsilon_c = 0.010$–0.020 $\varepsilon_s = 0.040$–0.060
Complete	Repair of building not feasible either physically or economically, demolition after earthquake required, could be due to shear failure of vertical elements or excess displacement

Appendix C: Analysis of mid-rise RC MRF buildings – Additional figures and tables for Chapter 3
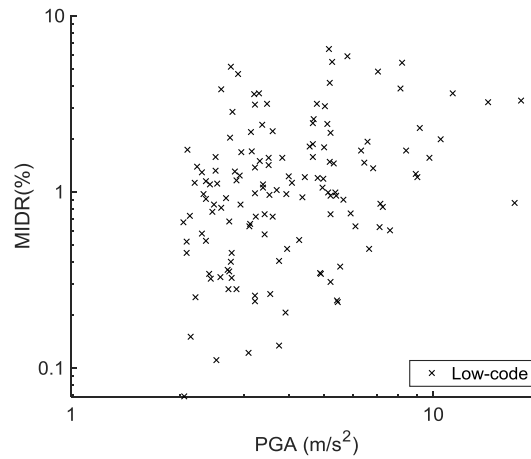
Figure C 1 - Seismic response results expressed in terms of PGA and MIDR(%) for Low-code building.
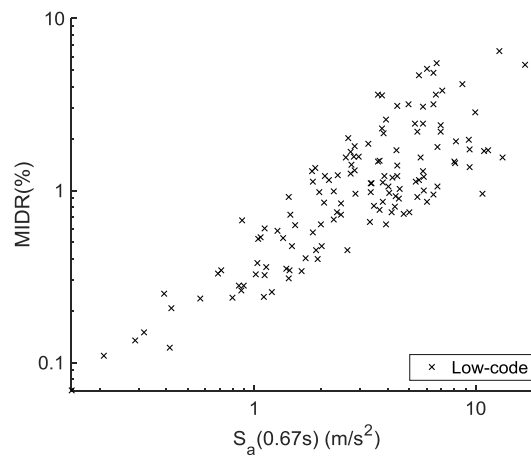
Figure C 2 - Seismic response results expressed in terms of $Sa(T_1)$ and MIDR(%) for Low-code building.
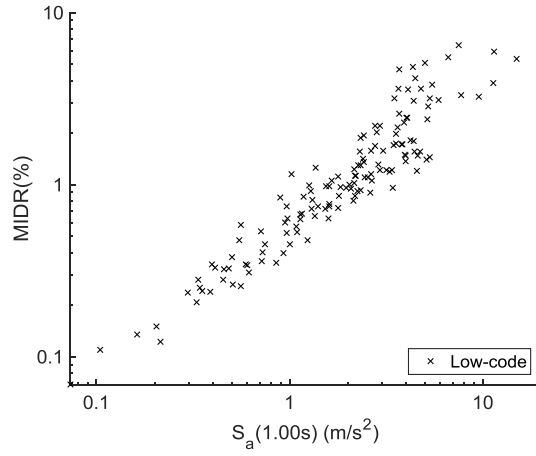
Figure C 3 - Seismic response results expressed in terms of *Sa(1.0s)* and MIDR(%) for Low-code building.
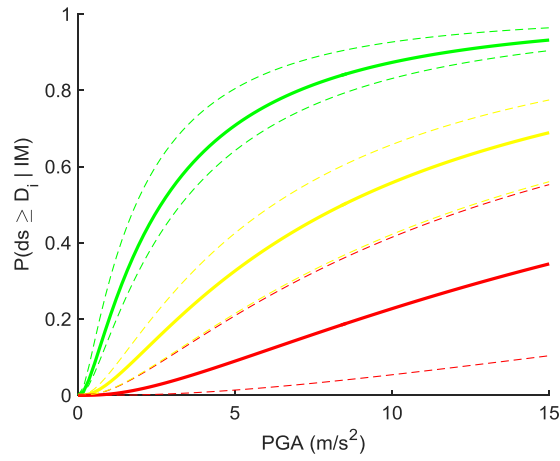
Figure C 4 - Fragility curves and associated 95% confidence intervals expressed in terms of PGA for the Low-code building.
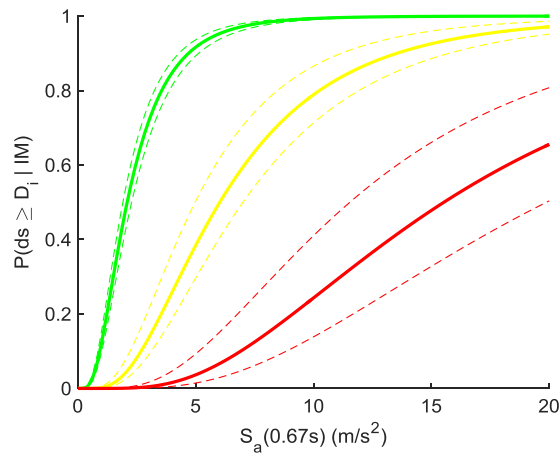
Figure C 5 - Fragility curves and associated 95% confidence intervals expressed in terms of *Sa(T₁)* for the Low-code building.

274
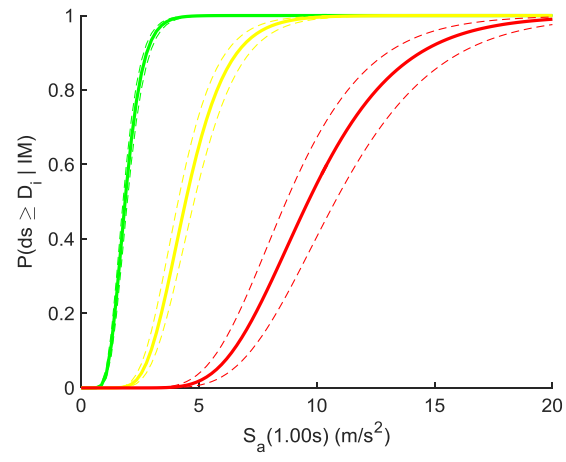
Figure C 6 - Fragility curves and associated 95% confidence intervals expressed in terms of *Sa(1.0s)* for the Low-code building.

Table C 1 - Subset of 150 strong-motion records obtained from SIMBAD database, used as a ground motion input in NDA and FRACAS analyses.

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
020y	W Tottori Prefecture	06/10/2000	6.6	strike-slip	11.78	B
021y	W Tottori Prefecture	06/10/2000	6.6	strike-slip	12.77	B
022x	W Tottori Prefecture	06/10/2000	6.6	strike-slip	19	B
023y	W Tottori Prefecture	06/10/2000	6.6	strike-slip	25.47	C
025x	Shimane Hiroshima Border	08/10/2000	5.1	strike-slip	6.8	C
032x	N Miyagi Prefecture	25/07/2003	5.5	reverse	18.12	D
039x	Mid Niigata Prefecture	23/10/2004	6.6	reverse	7.01	C
040y	Mid Niigata Prefecture	23/10/2004	6.6	reverse	11.09	C
041y	Mid Niigata Prefecture	23/10/2004	6.6	reverse	16.42	C
042y	Mid Niigata Prefecture	23/10/2004	6.6	reverse	21.25	B
043y	Mid Niigata Prefecture	23/10/2004	6.6	reverse	28.79	C
045y	Mid Niigata Prefecture	23/10/2004	6.3	reverse	8.99	C
046x	Mid Niigata Prefecture	23/10/2004	6.3	reverse	16.53	C
047x	Rumoi	14/12/2004	5.7	reverse	8.08	B
048x	Rumoi	14/12/2004	5.7	reverse	13.32	C
050x	NW Off Kyushu	20/03/2005	6.6	strike-slip	20.25	C
051y	NW Off Kyushu	20/03/2005	6.6	strike-slip	25.96	C
056y	Honshu	10/08/1996	5.9	reverse	13.89	B
060y	Kyushu	09/09/1996	5.7	normal	5.79	C
061x	Kyushu	09/09/1996	5.7	normal	17.24	B
068y	NW Kagoshima Prefecture	26/03/1997	6.1	strike-slip	12.05	B
069x	NW Kagoshima Prefecture	26/03/1997	6.1	strike-slip	12.15	B
071x	NW Kagoshima Prefecture	26/03/1997	6.1	strike-slip	19.69	C

Table C 1 – (continued)

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
081x	NW Kagoshima Prefecture	13/05/1997	6	strike-slip	15.42	C
082y	NW Kagoshima Prefecture	13/05/1997	6	strike-slip	16.19	B
083y	NW Kagoshima Prefecture	13/05/1997	6	strike-slip	15.7	B
087y	Yamaguchi Prefecture	25/06/1997	5.8	strike-slip	10.05	C
102x	Bam	26/12/2003	6.6	strike-slip	10.16	B
107y	Anza	12/06/2005	5.2	strike-slip	14.17	B
108x	Anza	12/06/2005	5.2	strike-slip	18.19	B
112y	South Iceland	17/06/2000	6.5	strike-slip	14.56	B
113y	South Iceland	17/06/2000	6.5	strike-slip	5.25	A
115x	South Iceland	17/06/2000	6.5	strike-slip	17.38	B
117y	South Iceland	21/06/2000	6.4	strike-slip	12.15	B
121y	South Iceland	21/06/2000	6.4	strike-slip	11.1	B
136x	Off Noto Peninsula	25/03/2007	6.7	reverse	6.64	B
137x	Off Noto Peninsula	25/03/2007	6.7	reverse	19.02	D
138y	Off Noto Peninsula	25/03/2007	6.7	reverse	27.17	B
139x	Southern Iwate Prefecture	13/06/2008	6.9	reverse	23.08	B
140y	Southern Iwate Prefecture	13/06/2008	6.9	reverse	18.82	B
141x	Southern Iwate Prefecture	13/06/2008	6.9	reverse	26.89	C
142x	S Suruga Bay	10/08/2009	6.2	reverse	26.71	C
143x	S Suruga Bay	10/08/2009	6.2	reverse	28.13	D
144y	S Suruga Bay	10/08/2009	6.2	reverse	22.54	B
145y	S Suruga Bay	10/08/2009	6.2	reverse	18.45	C
146x	S Suruga Bay	10/08/2009	6.2	reverse	25.38	B

Table C 1 – (continued)

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
150x	N Mie Prefecture	15/04/2007	5	reverse	11.95	C
153x	Southern Iwate Prefecture	14/06/2008	5.5	reverse	9.5	B
154x	Off S Niigata Prefecture	16/07/2007	6.6	reverse	24.79	C
155y	Off S Niigata Prefecture	16/07/2007	6.6	reverse	21.28	C
157x	Mid Niigata Prefecture	27/10/2004	5.8	reverse	23.27	C
158y	Mid Niigata Prefecture	27/10/2004	5.8	reverse	9.01	C
159y	App. Umbro-Marchigiano	06/10/1997	5.4	normal	10.62	E
165y	L'Aquila	06/04/2009	6.3	normal	4.63	B
166y	L'Aquila	06/04/2009	6.3	normal	4.39	B
167y	L'Aquila	06/04/2009	6.3	normal	5.65	B
168x	L'Aquila	06/04/2009	6.3	normal	4.87	B
184x	L'Aquila (aftershock)	07/04/2009	5.6	normal	16.81	B
208y	Duzce	12/11/1999	7.1	strike-slip	27.16	B
209y	Bingo'	01/05/2003	6.3	strike-slip	11.79	B
211x	Izmit (aftershock)	11/11/1999	5.6	normal	11.21	B
213x	Parkfield	28/09/2004	6	strike-slip	14.22	C
214x	Parkfield	28/09/2004	6	strike-slip	8.17	B
215y	Parkfield	28/09/2004	6	strike-slip	7.75	B
216x	Parkfield	28/09/2004	6	strike-slip	7.14	A
217x	Olfus	29/05/2008	6.3	strike-slip	8.89	A
218x	Olfus	29/05/2008	6.3	strike-slip	8.25	A
219x	Olfus	29/05/2008	6.3	strike-slip	7.97	A
220x	E Off Izu Peninsula	20/04/2006	5.6	strike-slip	24.33	B

Table C 1 – (continued)

ID	Earthquake Name	Date	M$_w$	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
221x	E Off Izu Peninsula	20/04/2006	5.6	strike-slip	8.85	C
241x	Western Fukushima Pref	29/09/2010	5.6	reverse	11.04	B
248x	Northern Nagano Pref	11/04/2011	5.4	strike-slip	21.37	C
251x	Eastern Fukushima Pref	12/04/2011	5.9	strike-slip	24.74	C
258x	E Off Fukushima Pref	22/04/2011	5.1	normal	18.09	B
264x	Northern Gifu Pref	26/02/2011	5.1	reverse	11.95	B
272y	Mid Niigata Pref	11/03/2011	6.2	reverse	20.66	B
274x	Mid Niigata Pref	11/03/2011	6.2	reverse	5.97	B
285x	Mt Fuji Region	15/03/2011	5.9	strike-slip	12.8	B
286x	Mt Fuji Region	15/03/2011	5.9	strike-slip	19.78	B
289x	Northern Ibaraki Pref	19/03/2011	5.8	normal	29.37	C
304y	Loma Prieta	18/10/1989	6.9	oblique	18.75	B
306y	Hyogo - Ken Nanbu	16/01/1995	6.9	strike-slip	16.6	C
307x	Duzce	12/11/1999	7.1	strike-slip	36.1	C
308x	Duzce	12/11/1999	7.1	strike-slip	32.26	C
311y	Emilia Pianura Padana	20/05/2012	6.1	reverse	13.36	C
312y	Emilia Pianura Padana	29/05/2012	6	reverse	3.58	C
316y	Emilia Pianura Padana	29/05/2012	6	reverse	16.04	C
317x	Emilia Pianura Padana	29/05/2012	6	reverse	16.43	C
319y	Emilia Pianura Padana	29/05/2012	6	reverse	21.34	C
329x	Darfield	03/09/2010	7.1	strike-slip	9.06	C
330y	Darfield	03/09/2010	7.1	strike-slip	13.31	C
331x	Darfield	03/09/2010	7.1	strike-slip	17.82	C

Table C 1 – (continued)

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
333y	Darfield	03/09/2010	7.1	strike-slip	16.97	C
335x	Darfield	03/09/2010	7.1	strike-slip	23.58	C
337x	Christchurch	21/02/2011	6.2	reverse	2.25	B
338x	Christchurch	21/02/2011	6.2	reverse	24.93	D
339y	Christchurch	21/02/2011	6.2	reverse	1.48	A
340x	Christchurch	21/02/2011	6.2	reverse	14.38	C
341x	Christchurch	21/02/2011	6.2	reverse	13.73	C
347y	Christchurch	22/02/2011	5.5	reverse	4.1	B
348x	Christchurch	22/02/2011	5.5	reverse	23.6	D
349x	Christchurch	22/02/2011	5.5	reverse	5.51	A
350x	Christchurch	22/02/2011	5.5	reverse	9.86	C
355y	Christchurch	22/02/2011	5.6	reverse	7.38	B
357x	Christchurch	22/02/2011	5.6	reverse	8.42	A
358x	Christchurch	22/02/2011	5.6	reverse	8.22	C
376y	Christchurch	13/06/2011	5.3	reverse	4.89	B
384y	Christchurch	13/06/2011	6	reverse	3.24	B
386x	Christchurch	13/06/2011	6	reverse	5.1	A
395x	Christchurch	21/06/2011	5.2	reverse	15.04	B
404x	Christchurch	21/06/2011	5.2	reverse	6.76	C
410x	Athens	07/09/1999	6	reverse	13.85	B
412x	Hector Mine	16/10/1999	7.1	strike-slip	28.61	B
419y	Umbria-Marche (1st shock)	26/09/1997	5.7	normal	2.81	D
422y	Friuli (1st shock)	06/05/1976	6.4	reverse	21.72	B

Table C 1 – (continued)

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
424x	Friuli (2nd shock)	11/09/1976	5.6	reverse	26.2	B
425y	Friuli (2nd shock)	11/09/1976	5.6	reverse	14.96	B
427y	Friuli (3rd shock)	15/09/1976	5.9	reverse	17.29	B
429y	Friuli (4th shock)	15/09/1976	5.9	reverse	16.83	B
431x	Dinar	01/10/1995	6.4	normal	0.47	C
432x	Parkfield	28/09/2004	6	strike-slip	19.5	B
433y	Parkfield	28/09/2004	6	strike-slip	3.02	B
434x	Parkfield	28/09/2004	6	strike-slip	12.49	B
436x	Parkfield	28/09/2004	6	strike-slip	3.95	B
439x	Parkfield	28/09/2004	6	strike-slip	7.03	B
440x	Gazli	17/05/1976	6.7	reverse	12.78	B
441x	Tabas	16/09/1978	7.1	reverse	20.63	B
442x	Imperial Valley	15/10/1979	6.5	strike-slip	2.31	C
444y	Imperial Valley	15/10/1979	6.5	strike-slip	19.33	C
445y	Imperial Valley	15/10/1979	6.5	strike-slip	27.03	C
446y	Imperial Valley	15/10/1979	6.5	strike-slip	27.68	C
447x	Imperial Valley	15/10/1979	6.5	strike-slip	27.35	C
450y	Loma Prieta	18/10/1989	6.9	oblique	9.3	C
451y	Loma Prieta	18/10/1989	6.9	oblique	7.1	B
452x	Loma Prieta	18/10/1989	6.9	oblique	28.57	A
453y	Loma Prieta	18/10/1989	6.9	oblique	29.66	C
454x	Loma Prieta	18/10/1989	6.9	oblique	28.83	B
455y	Loma Prieta	18/10/1989	6.9	oblique	16.41	B

Table C 1 – (continued)

ID	Earthquake Name	Date	M_w	Fault Mechanism	Epicentral Distance (km)	EC8 Site class
456y	Loma Prieta	18/10/1989	6.9	oblique	27.59	B
457x	Landers	28/06/1992	7.3	strike-slip	13.08	B
458x	Northridge	17/01/1994	6.7	reverse	11.02	C
459x	Northridge	17/01/1994	6.7	reverse	20.19	C
460y	Northridge	17/01/1994	6.7	reverse	23.62	C
461y	Northridge	17/01/1994	6.7	reverse	20.25	C
462x	Northridge	17/01/1994	6.7	reverse	5.39	C
463y	Northridge	17/01/1994	6.7	reverse	18.59	B
464x	Northridge	17/01/1994	6.7	reverse	14.67	B
465y	Erzincan	13/03/1992	6.6	strike-slip	8.97	C
466x	Duzce	12/11/1999	7.1	strike-slip	5.27	C
467x	Kozani	13/05/1995	6.5	normal	16.69	A

Appendix D: Additional tables for Chapter 4

Table D 1 - Fragility parameters (mean θ in m/s2 and standard deviation β). Fragility derivation method: Least Squares.

Building	IM	DS1		DS2		DS3	
		θ	β	θ	β	θ	β
Pre-code	PGA	1.499	1.193	4.007	1.193	7.167	1.193
	$Sa(T_1)$	0.873	0.732	2.284	0.732	4.033	0.732
	$Sa(1.0s)$	0.767	0.659	1.970	0.659	3.441	0.659
Low- code	PGA	2.639	1.166	8.443	1.166	23.899	1.166
	$Sa(T_1)$	2.084	0.633	6.017	0.633	15.533	0.633
	$Sa(1.0s)$	1.857	0.314	4.427	0.314	9.628	0.314
Special-code	PGA	4.862	0.807	13.420	0.807	46.677	0.807
	$Sa(T_1)$	5.963	0.663	19.269	0.663	81.329	0.663
	$Sa(1.0s)$	2.423	0.478	8.217	0.478	36.789	0.478

Table D 2 - Fragility parameters (mean θ in m/s2 and standard deviation β)

Building	Capacity model	Ground motions	Derivation method	DS1		DS2		DS3	
				θ	β	θ	β	θ	β
Pre-code	NDA	Unscaled	GLM	0.669	0.302	2.531	0.480	5.498	0.584
	TL	Unscaled	GLM	1.199	0.003	3.273	0.241	5.384	0.468
	EPP	Unscaled	GLM	1.230	0.002	3.084	0.277	5.317	0.467
	TL	REAL-EC8:475	GLM	1.179	0.056	3.115	0.289	4.466	0.359
	TL	SIMQKE-EC8:475	GLM	1.143	0.001	2.938	0.161	3.780	0.181
	TL	MATCHED-EC8:475	GLM	1.136	0.012	2.973	0.214	3.957	0.213
Special code	NDA	Unscaled	GLM	3.987	0.360	12.059	0.300	-	-
	EPP	Unscaled	GLM	5.639	0.225	15.291	0.374	-	-
	EPP	Unscaled	LS	5.995	0.159	13.677	0.159	-	-

Appendix E: Additional tables for Chapter 7

Table E 1 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error, Average Length and Coverage probability. In this test-case high fidelity analysis is used alongside with advanced IM I_{Np} and stratified sampling process.

Case Study Building	Reality	Sample Size	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
Special-code	1	Full	0.269 (0.283)	4.72	1.702 (1.710)	0.43	95.84	95.60
		Large	0.271 (0.282)	3.87	1.730 (1.741)	0.65	95.75	95.60
		Medium	0.306 (0.304)	-0.57	1.760 (1.777)	0.92	95.52	95.68
	2	Full	0.277 (0.299)	7.08	1.738 (1.759)	1.19	95.84	95.43
		Large	0.283 (0.297)	4.99	1.758 (1.779)	1.13	95.90	95.71
		Medium	0.319 (0.323)	1.19	1.801 (1.837)	1.98	95.59	95.56
	3	Full	0.299 (0.329)	9.01	1.699 (1.725)	1.50	96.06	95.49
		Large	0.307 (0.327)	6.15	1.723 (1.741)	1.06	95.81	95.54
		Medium	0.361 (0.363)	0.60	1.773 (1.807)	1.89	95.54	95.57
Pre-code	1	Full	0.390 (0.396)	1.54	2.203 (2.208)	0.24	95.87	95.86
		Large	0.387 (0.398)	2.74	2.237 (2.248)	0.46	96.10	96.07
		Medium	0.418 (0.420)	0.46	2.297 (2.309)	0.52	95.82	95.87
	2	Full	0.473 (0.479)	1.33	2.403 (2.411)	0.30	95.98	95.96
		Large	0.473 (0.483)	1.98	2.438 (2.453)	0.60	96.05	96.13
		Medium	0.511 (0.510)	-0.2	2.517 (2.529)	0.50	95.88	95.89
	3	Full	0.386 (0.392)	1.61	2.186 (2.192)	0.28	95.89	95.88
		Large	0.386 (0.395)	2.31	2.221 (2.231)	0.44	96.08	96.08
		Medium	0.415 (0.416)	0.19	2.283 (2.294)	0.50	95.86	95.87

Table E 2 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error, Average Length and Coverage probability. In this test-case low fidelity analysis is used alongside with advanced IM I_{Np} and stratified sampling process

Case Study Building	Reality	Sample Size	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
Special-code	1	Full	0.049 (0.052)	5.96	0.712 (0.716)	0.58	95.48	95.02
		Large	0.049 (0.052)	5.48	0.726 (0.732)	0.79	95.48	95.25
		Medium	0.057 (0.056)	-0.79	0.726 (0.734)	1.13	94.98	95.00
	2	Full	0.050 (0.064)	21.65	0.741 (0.782)	5.22	96.21	95.25
		Large	0.052 (0.062)	16.23	0.755 (0.792)	4.73	96.03	95.51
		Medium	0.065 (0.069)	5.37	0.759 (0.807)	6.00	94.97	95.14
	3	Full	0.048 (0.069)	30.68	0.695 (0.758)	8.37	96.90	95.54
		Large	0.051 (0.066)	23.28	0.716 (0.764)	6.26	96.41	95.68
		Medium	0.070 (0.078)	10.16	0.728 (0.776)	6.09	95.21	95.11
Pre-code	1	Full	0.583 (0.632)	7.80	1.756 (1.761)	0.28	95.06	94.77
		Large	0.547 (0.625)	12.58	1.782 (1.801)	1.04	94.91	94.80
		Medium	0.640 (0.714)	10.38	1.793 (1.787)	-0.35	94.15	94.10
	2	Full	0.668 (0.779)	14.28	1.972 (1.996)	1.20	95.05	94.90
		Large	0.658 (0.772)	14.83	2.000 (2.041)	2.00	94.98	94.93
		Medium	0.799 (0.894)	10.64	2.015 (2.039)	1.19	94.00	94.08
	3	Full	0.503 (0.545)	7.69	1.723 (1.744)	1.19	95.07	94.90
		Large	0.492 (0.541)	9.04	1.752 (1.785)	1.85	95.12	94.98
		Medium	0.603 (0.627)	3.85	1.765 (1.782)	0.92	94.11	94.02

Table E 3 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error, Average Length and Coverage probability. In this test-case Special-code building is used alongside with the IM $Sa(T_1)$ and stratified sampling process.

Fidelity	Reality	Sample Size	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
High fidelity	1	Full	0.323 (0.336)	3.76	2.041 (2.052)	0.56	95.90	95.58
		Large	0.335 (0.341)	1.83	2.070 (2.082)	0.59	95.80	95.57
		Medium	0.349 (0.354)	1.41	2.090 (2.106)	0.75	95.47	95.55
	2	Full	0.344 (0.367)	6.24	2.053 (2.074)	1.02	95.97	95.67
		Large	0.363 (0.372)	2.64	2.092 (2.110)	0.84	95.67	95.47
		Medium	0.388 (0.387)	-0.26	2.130 (2.146)	0.75	95.42	95.75
	3	Full	0.406 (0.439)	7.51	2.059 (2.082)	1.11	96.03	95.67
		Large	0.433 (0.448)	3.49	2.107 (2.129)	1.05	95.75	95.58
		Medium	0.458 (0.470)	2.44	2.141 (2.158)	0.78	95.44	95.60
Low fidelity	1	Full	0.074 (0.078)	5.66	0.954 (0.960)	0.64	95.58	95.30
		Large	0.078 (0.080)	2.23	0.968 (0.973)	0.46	95.58	95.25
		Medium	0.082 (0.083)	1.00	0.968 (0.976)	0.80	94.97	95.13
	2	Full	0.082 (0.095)	13.89	0.973 (0.996)	2.29	95.92	95.28
		Large	0.092 (0.097)	5.53	0.997 (1.011)	1.41	95.62	95.28
		Medium	0.103 (0.101)	-1.94	1.013 (1.023)	0.98	95.07	95.15
	3	Full	0.083 (0.102)	19.20	0.919 (0.953)	3.60	96.30	95.42
		Large	0.096 (0.105)	9.27	0.951 (0.975)	2.50	95.80	95.32
		Medium	0.109 (0.111)	1.16	0.956 (0.975)	1.90	95.00	95.02

Table E 4 - Comparison of BEA and cloud method predictions in terms of Mean Squared Error, Average Length and Coverage probability. In this test-case Special-code building is used alongside with the IM $Sa(T_1)$ and random sampling process. Full sample cases are not shown here as they coincide with the corresponding entries in **Table E 4**.

Fidelity	Reality	Sample Size	MSE BEA (Cloud)	Difference (%)	Average Length BEA (Cloud)	Difference (%)	Coverage (%) BEA	Coverage (%) Cloud
High fidelity	1	Large	0.320 (0.337)	5.07	2.081 (2.101)	0.95	96.12	96.05
		Medium	0.326 (0.345)	5.51	2.120 (2.152)	1.48	95.70	95.62
	2	Large	0.347 (0.370)	6.18	2.083 (2.108)	1.22	96.07	95.87
		Medium	0.360 (0.379)	4.92	2.115 (2.143)	1.29	95.70	95.58
	3	Large	0.411 (0.442)	7.07	2.079 (2.099)	0.94	96.12	95.82
		Medium	0.424 (0.452)	6.00	2.105 (2.132)	1.24	95.65	95.48
Low fidelity	1	Large	0.074 (0.079)	6.41	0.973 (0.984)	1.11	96.05	95.85
		Medium	0.075 (0.081)	6.91	0.975 (0.996)	2.13	95.50	95.27
	2	Large	0.084 (0.096)	12.51	0.987 (1.011)	2.32	96.07	95.50
		Medium	0.090 (0.099)	8.80	0.988 (1.010)	2.11	95.28	94.82
	3	Large	0.086 (0.103)	16.52	0.928 (0.953)	2.65	96.10	95.32
		Medium	0.093 (0.105)	11.85	0.930 (0.954)	2.50	95.32	94.77

Table E 5 - Sensitivity of BEA predictions when using different covariance models and estimation of the percent differences to the associated cloud results. Case study building: Special-code; Analysis fidelity: High; Reality: 3; IM: *Sa(T1)*; Sampling method: Stratified sampling.

Sample Size	Covariance Model	MSE	Difference (%)	Average Length	Difference (%)	Coverage (%)
Full	Gaussian	0.406	7.51	2.059	1.11	96.03
	Matérn	0.357	18.78	1.880	9.69	96.48
	Spherical	0.365	16.82	2.034	2.32	96.62
	Cubic	0.386	12.15	2.049	1.60	96.37
	Exponential	0.357	18.78	1.880	9.69	96.48
	Cauchy	0.409	6.98	2.064	0.87	95.95
	Gneiting	0.406	7.51	2.059	1.08	96.03
	Circular	0.381	13.17	2.045	1.77	96.55
	P. Nugget	0.432	1.58	2.073	0.45	95.83
Large	Gaussian	0.433	3.49	2.107	1.05	95.75
	Matérn	0.429	4.31	2.058	3.34	95.79
	Spherical	0.422	5.85	2.096	1.56	95.95
	Cubic	0.423	5.68	2.102	1.28	95.97
	Exponential	0.429	4.31	2.058	3.34	95.79
	Cauchy	0.435	2.96	2.124	0.27	95.71
	Gneiting	0.432	3.56	2.108	1.00	95.77
	Circular	0.423	5.54	2.099	1.41	96.18
	P. Nugget	0.444	0.94	2.117	0.58	95.63
Medium	Gaussian	0.458	2.44	2.141	0.78	95.44
	Matérn	0.460	2.16	2.101	2.63	95.03
	Spherical	0.472	-0.53	2.106	2.41	95.30
	Cubic	0.462	1.71	2.124	1.59	95.50
	Exponential	0.460	2.16	2.101	2.63	95.03
	Cauchy	0.471	-0.23	2.151	0.31	95.43
	Gneiting	0.475	-1.04	2.142	0.75	95.50
	Circular	0.459	2.21	2.117	1.91	95.17
	P. Nugget	0.469	0.11	2.140	0.86	95.48