SCALAR AND VECTOR PROBABILISTIC SEISMIC RISK ASSESSMENT PROCEDURES AND THEIR IMPLEMENTATION FOR A SIMPLIFIED MID-RISE RC TUNNEL FORM BUILDING

by

Haris Curić

B.S, Civil Engineering, Sakarya University, 2017

Submitted to the Kandilli Observatory and Earthquake Research Institute in partial fulfillment of the requirements for the degree of

Master of Science

Graduate Program in Earthquake Engineering

Boğaziçi University

2021

ACKNOWLEDGEMENTS

I would hereby like to thank my mentor, Professor Sinan Akkar for his invaluable support during my MSc studies and especially during my thesis studies. Other than being a great educator in the field of earthquake engineering, he has shown great patience and managed to keep me motivated and hardworking throughout the duration of my MSc studies at Boğaziçi University for which I am very grateful.

I would also like to thank my friend Uğur Özen for agreeing to collaborate in our theses studies and to Çağrı Dikmen for giving me constructive advices and indulging in interesting discussions about seismic risk assessment.

I am grateful to my friend and colleague Onur Güneş for giving me valuable support in final stages of writing of this thesis.

I am very grateful to my wife, my parents and my sister for supporting me and believing in me.

ABSTRACT

SCALAR AND VECTOR PROBABILISTIC SEISMIC RISK ASSESSMENT PROCEDURES AND THEIR IMPLEMENTATION FOR A SIMPLIFIED MID-RISE RC TUNNEL FORM BUILDING

This thesis implements various scalar and vector methods of probabilistic seismic risk assessment. Methods implemented differ by type of scaling and selection of ground motion records as well as by primary scalar intensity measure (IM) used as ground motion intensity metric. Two different ground motion scaling and selection methods are used, namely stripe scaling and cloud scaling. Conceptual differences between stripe scaling and cloud scaling as well as the differences they produce in risk assessment results are presented and discussed. Different scalar ground motion intensity measures used are fundamental period spectral acceleration ($SA(T_1)$) and two distinct *weighted average spectral accelerations*. Weighted average spectral acceleration is a novel intensity measure developed and presented within this thesis. It is shown to be a good predictor of structural response at high intensity levels where structure undergoes intense yielding. Weighted average spectral acceleration enables analyst to set weights to contributions of spectral accelerations at different periods according to their expected importance to structural response. Conditional spectra for ground motion selection are developed for all three considered scalar IMs. Multivariate normal distributions are presented as useful mathematical tools and they are used for various purposes throughout this thesis (conditional spectra, vector IM hazard, structural response modeling, etc.). Maximum likelihood estimation (MLE) based method is also developed for estimation of parameters used for modeling structural response using structural analyses data.

ÖZET

BASİTLEŞTİRİLMİŞ ORTA KATLI TÜNEL KALIP BETONARME BİNA İÇİN SKALER VE VEKTÖR OLASILIKLI SİSMİK RİSK DEĞERLENDİRME PROSEDÜRLERİ VE UYGULAMASI

Bu tez olasılıklı sismik risk değerlendirmenin çeşitli skaler ve vektörel yöntemlerini uygulamaktadır. Uygulanan yöntemler, yer hareket kayıtlarının ölçeklendirilmesi ve seçimi yöntemi ile yer hareketinin şiddetinin göstergesi olarak kullanılan ana şiddet ölçeğine (IM) göre farklılık göstermektedir. İki tane farklı yer hareket kayıtlarının ölçeklendirilmesi ve seçimi yöntemi kullanılmaktadır; şerit ölçekledirilmesi ve bulut ölçeklendirilmesi. İki tane farklı kayıt ölçeklendirme yönteminin kavramsal farkları ve risk sonuçlarına etkileri sunulup tartışıldı. Kullanılan skaler şiddet ölçekleri (IM) temel mod periyottaki spektral ivme ($SA(T_1)$) ve iki tane farklı *ağırlıklı ortalama spektral ivme*'dir. Ağırlıklı ortalama spektral ivme bu tezde geliştirilen ve sunulan yeni bir skaler şidder ölçeğidir (IM). Binanın yoğun olarak akma davranışı gösterdiği yer hareketlerinin yüksek şiddet seviyelerinde iyi bir yapısal cevap tahmin değişkeni olduğunu gösterildi. Ağırlıklı ortalama spectral ivme analiste farklı peryotlardaki spektral ivmelere yapısal cevabına beklenen etkilerine göre farklı ağırlıklar atamasına olanak sağlamaktadır. Üç farklı şiddet ölçeği kullanarak yer hareket kayıtlarının seçimi için kullanılan koşullu spektrumlar geliştirildi. Çok-değişkenli normal dağılımları kullanışlı matematiksel bir araç olarak sunuldu ve tezin içerisinde çeşitli amaçlar için kullanıldı (koşullu spektrumlar, vektör şiddet ölçekli deprem tehlikesi, yapısal cevap modellenmesi, vs.). Yapısal analiz dataları kullanarak yapısal modelleme için maksimum olasılık hesaba (maximum likelihood estimation) dayanarak yeni bir yöntem geliştirildi.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a	Parameter from Guttenberg – Richter recurrence model
$AvgSA$	Average spectral acceleration
$AvgSA\left(\vec{T}, \overrightarrow{W'}\right)$	Weighted average spectral acceleration
$AvgSA1$	Weighted average spectral acceleration type 1
$AvgSA2$	Weighted average spectral acceleration type 2
b	Parameter from Guttenberg – Richter recurrence model
b_i	Collapse binary logistic observed output from i'th structural analysis
B_i	Collapse binary logistic random variable from i'th structural analysis
$CMS_{IM}(T)$	Conditional mean spectrum
$CS_{AvgSA}(T)$	Conditional spectrum conditioned on average spectral acceleration
$CS_{AvgSA1}(T)$	Conditional spectrum conditioned on weighted average spectral acceleration type 1
$CS_{AvgSA2}(T)$	Conditional spectrum conditioned on weighted average spectral acceleration type 2
$CS_{SA(T_1)}(T)$	Conditional spectrum conditioned on fundamental period spectral acceleration
$\bar{C}$	Occurrence of non-collapse
$D_{IM \geq im}(k, i)$	Seismic hazard exceedance disaggregation data
$D_{IM=im}(k, i)$	Seismic hazard occurrence disaggregation data
EDP	Engineering demand parameter
EDP_i	Engineering demand parameter output from i'th non-collapse structural analysis
$\overrightarrow{EDP}$	Vector of engineering demand parameters
f	Probability density function
F	Cumulative distribution function
F_{base}	Base shear
F_y	Yield Force

$G_{IM\|M=m,R=r}\,(im)$	Probability that an intensity measure (IM) exceeds certain threshold value (im) given that certain seismic scenario occurs ($M = m, R = r$)
$G_{EDP\|IM}(y\|im)$	Probability that certain engineering demand parameter exceeds certain threshold level given intensity measure of ground motion is equal to specific value
$G_{\overrightarrow{EDP}\|IM}(\vec{y}\|im)$	Probability that vector of engineering demand parameters exceeds certain threshold levels given intensity measure of ground motion is equal to specific value
$G_{EDP\|\overrightarrow{IM}}(y\|\overrightarrow{im})$	Probability that certain engineering demand parameter exceeds certain threshold level given that vector of intensity measures of ground motion is equal to specific values
$G_{\overrightarrow{EDP}\|\overrightarrow{IM}}(\vec{y}\|\overrightarrow{im})$	Probability that vector of engineering demand parameters exceeds certain threshold level given a vector of intensity measures of ground motion
IM	Intensity measure
IM_i	Intensity measure of ground motion used for i'th structural analysis
$IM_{stripe-i}$	Intensity measure value at i'th stripe
$\overrightarrow{IM}$	Vector of intensity measures
k_d	Pinching behavior parameter; displacement ratio
k_s	Pinching behavior parameter; target strength reduction factor
K_e	Initial stiffness
$\overrightarrow{lnSAT}(rup, Vs30)$	Vector of logarithms of spectral accelerations at different periods representing certain seismic scenario spectrum distribution for certain site conditions
$LSNRD$	Limit states normalized roof displacement
m_0	Minimum considered magnitude in PSHA
M	Magnitude
M_{chr}	Characteristic magnitude
M_{eff}	Effective modal mass
$MRD_{IM}(im)$	Mean rate density function of an intensity measure
$MRD_{\overrightarrow{IM}}(\overrightarrow{im})$	Mean rate density function of a vector of intensity measures
$N(m)$	Mean annual rate of earthquakes exceeding with $M > m$

N_{C-i}	Number of collapses at i'th stripe
$P_{50\,years}(EDP > y)$	Probability of exceedance of certain threshold engineering demand parameter in 50 years exposure time
P_C	Collapse probability
$P_C(\overrightarrow{im})$	Vector intensity measure collapse fragility surface
P_{C-i}	Collapse probability at i'th stripe
PGA	Peak ground acceleration
R	Generic distance to site parameter
R_{hyp}	Hypo-central distance to site
R_{epi}	Epicentral distance to site
R_{JB}	Joyner – Boore distance to site
Rup	Rupture
S_a	Spectral acceleration
S_d	Spectral displacement
$SA(T)$	Spectral acceleration
$SA(T_1)$	Fundamental mode period spectral acceleration
SoF	Style of faulting
T	Period
T_1	Fundamental mode period
$\vec{T}$	Vector containing periods used to construct weighted average spectral acceleration
$\overrightarrow{T_1}$	Vector containing period set for weighted average spectral acceleration type 1
$\overrightarrow{T_2}$	Vector containing period set for weighted average spectral acceleration type 2
$\ddot{u}_{gx}(t)$	Non-rotated horizontal ground motion acceleration component 1
$\ddot{u}_{gy}(t)$	Non-rotated horizontal ground motion acceleration component 2
$\ddot{u}_g(t)$	Ground motion acceleration function (trace)
$\ddot{u}_{gx-rot\theta}(t)$	Rotated horizontal ground motion acceleration component 1
$\ddot{u}_{gy-rot\theta}(t)$	Rotated horizontal ground motion acceleration component 2
U_{roof}	Roof displacement

$Vs30$	Surface 30 meters average shear wave velocity
$\overrightarrow{W'}$	Vector containing weights used to construct weighted average spectral acceleration
$\overrightarrow{W'_1}$	Vector containing weights for weighted average spectral acceleration type 1
$\overrightarrow{W'_2}$	Vector containing weights for weighted average spectral acceleration type 2
α_c	Post-capping stiffness ratio
α_s	Post-yielding stiffness ratio
β	Parameter from Youngs and Coppersmith recurrence model
γ_a	Cyclic degradation parameter; energy capacity ratio for accelerated stiffness degradation
γ_c	Cyclic degradation parameter; energy capacity ratio for post-capping strength reduction
γ_k	Cyclic degradation parameter; energy capacity ratio for unloading stiffness reduction
γ_s	Cyclic degradation parameter; energy capacity ratio for strength reduction
Γ	Modal participation factor
$\frac{\delta_c}{\delta_y}$	Ratio between capping displacement and yield displacement
θ	Angle between structural principal horizontal axes and as-recorded input ground motion horizontal axes
λ	Residual strength ratio
λ_C	Rate of collapse
$\lambda_{EDP}(y)$	Rate of exceedance of certain engineering demand parameter threshold level
$\lambda_{\overrightarrow{EDP}}(\vec{y})$	Rate of exceedance of vector engineering demand parameter threshold level
$\lambda_{IM \geq im}$	Rate of exceedance of intensity measure im

$\lambda_{M>m}$	Rate of exceedance of magnitude m
$\mu_{C-ln(IM)}$	Logarithmic mean parameter used for functional modeling of collapse fragility
$\mu_{\ln(EDP)\|IM}(im)$	Mean of logarithm of engineering demand parameter conditioned on certain intensity measure
$\mu_{\ln(EDP)\|\overrightarrow{IM}}(\overrightarrow{im})$	Mean of logarithm of engineering demand parameter conditioned on a vector of intensity measures
$\mu_{\ln(IM)\|Rup}$	Mean of logarithm of intensity measure conditioned on rupture scenario
ρ	Correlation coefficients
$\sigma_{C-ln(IM)}$	Logarithmic standard deviation parameter used for functional modeling of collapse fragility
$\sigma_{\ln(EDP)\|IM}(im)$	Standard deviation of logarithm of engineering demand parameter conditioned on certain intensity measure
$\sigma_{\ln(EDP)\|\overrightarrow{IM}}$	Standard deviation of logarithm of engineering demand parameter conditioned on a vector of intensity measures
$\sigma_{\ln(IM)\|Rup}$	Standard deviation of logarithm of intensity measure conditioned on rupture scenario
ϕ_{roof}	Roof displacement component of modal shape vector

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
AFAD	Turkish Disaster and Emergency Management Presidency
ASB14	Ground Motion Prediction Equations Proposed by Akkar et al.
ASCE	American Society of Civil Engineers
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CEN	European Committee for Standardization
CS	Conditional Spectrum
EDP	Engineering Demand Parameter
GM	Ground Motion
GMPE	Ground Motion Prediction Equation
IM	Intensity Measure
LSNRD	Limit States Normalized Roof Displacement
MDOF	Multi-Degree of Freedom
MLE	Maximum Likelihood Estimation
MRD	Mean Rate Density
PDF	Probability Density Function
PEER	Pacific Earthquake Engineering Research Center
PMF	Probability Mass Function
PSHA	Probabilistic Seismic Hazard Assessment
PSRA	Probabilistic Seismic Risk Assessment
RC	Reinforced Concrete
SDOF	Single-Degree of Freedom
SSE	Sum of Squares of Errors
SSRA	Seismic Structural Response Assessment
TEBC	Turkish Earthquake Building Code
UHS	Uniform Hazard Spectrum
YC	Youngs and Coppersmith

1. INTRODUCTION

1.1. Historical Overview of Earthquake Engineering and Seismic Risk

This thesis describes methods of probabilistic seismic risk assessment of structures and demonstrates computer algorithms developed to perform these methods. Earthquake resistant structural design (earthquake engineering) has become one of the most important disciplines in civil engineering for the last 150 years and it continues to gain more importance. The reason behind this fact is simple and can be explained in everyday language as well as in terms of more professional earthquake engineering terminology.

In everyday language, the reason for the growing importance of earthquake engineering is the accelerated increase of worldwide population, which triggers the social sensitivity to natural disasters such as earthquakes. Historically, most of the facilities that mankind built for life support (houses, barns, silos, bridges etc.) are low rise (usually single story) buildings. Until only 500 years ago, most people lived in single-story village houses made of either wood or masonry. Most of the people had “jobs” in food production such as farming, livestock, hunting, fishing, etc. All of these occupations required no or only low-rise civil facilities. High-rise structures were mainly built for monumental, religious and royal power-showing purposes (e.g., Pyramids in Egypt). Being rare and over-financed, these high-rise facilities were mostly built in an over-designed manner that is controversial to modern engineering aspect. Even those that were not over-designed would later be rebuilt and made safe after some historical structural failure (Hagia Sofia in Istanbul). Considering these facts, humans, historically, did not consider earthquakes as important life endangering natural disasters. Earthquakes would occur causing no or very little damage to human lives and activities. For example, storms were more dangerous because they potentially could destroy farms and sentence the whole society (villages) to death by starvation.

As the population in large cities and urban areas keeps increasing in the last two centuries, there is a growing demand for housing, office, industry, manufacture, and other types of civil infrastructure, which inevitably tend to grow in height to satisfy the increasing demands of the people. At the same time, our living standards are increasing, which demands

for efficient production of products and services such as high-quality food, drinking water, electrical energy, etc. The tendency towards more complicated civil facilities makes them inevitably vulnerable to natural disasters such as earthquakes. Efficiency comes at a price of increased risk of some fatal failure. Take, for example, nuclear power plants. Nuclear energy is the most efficient energy production method, which makes it attractive to many states and large investors. However, this efficiency comes with a risk of endangering the human lives for large scale fatal events (Chernobyl Power Plant - 1986). Nuclear failure risk is caused by many factors among which seismic events represent one of the most important ones that was demonstrated once again by the Fukushima Nuclear Power Plant incident of 2007.

For the first time in history, earthquakes, which were benign events only important in folklore and mythological stories, became one of the most important and life endangering disasters, which wiped human lives from Earth on a large scale (e.g., nuclear power plants' fatal failures). In terms of earthquake engineering terminology, seismic hazard, which is, in simple language, the severity and rate of occurrence of strong ground motions, has remained approximately constant throughout the human history, but the seismic fragility of our life supporting facilities, which is in simple language, the sensitivity of our facilities to ground shaking, has increased pushing up the seismic risk scale. Although the seismic risks always existed, their frequency was historically so low that they were not even worth evaluating or mitigating. However, with the development of modern cities and modern efficient life supporting infrastructure these risks have increased, making earthquake engineering one of the most important disciplines in modern era. Assessment and mitigation of seismic risks have become key civil engineering endeavors in many countries located in seismically active areas (e.g., USA, Turkey, Japan, Italy, Switzerland, New Zealand etc.).

1.2. Modern Probabilistic Seismic Risk Assessment (PSRA)

Earthquake engineering and seismic risk assessment are very complex engineering endeavors. Considering that most of the factors affecting seismic risk are related to highly unpredictable natural and physical phenomena, which cannot be described deterministically, modern earthquake engineering tends toward developing methods of structural design that are based on statistics and probability theory. Among many publications that provide guidelines for PSRA, FEMA P58-1 summarizes its most important developments until year

2012 [1]. Keith Porter provides another very useful guideline document that gives statistical framework of entire PSRA procedure [2].

When it comes to factors affecting the seismic risk, although there are probably many unknowns to earthquake engineering community, the major ones are divided into two categories, namely *seismic hazard factors* and *seismic structural response factors*.

Probabilistic seismic hazard assessment (PSHA) is used in the seismic risk assessment, which has a goal of assessing the exceedance probability of a certain level of ground shaking intensity that will affect the facility whose seismic risk is being evaluated. PSHA accounts for seismic source factors, path factors and site factors. Even though there are some aspects of PSHA that are structure specific, it can be considered as a part of the seismic risk assessment that is ideally only location dependent. The goal of PSHA is to evaluate the rate of occurrence (or exceedance) of a specific ground motion intensity at a certain site. This is a necessary probabilistic input for risk assessment in order to evaluate the probability of some structural damage or some other undesired outcome that we need to. Being integral part of PSRA, PSHA is a subject of many publications and studies. Jack W. Baker provides clear explanation of basic PSHA methodology [3]. PSHA can be split into its two integral parts: *seismic source characterization* and *prediction of ground motion*. McGuire emphasizes this distinction between two mentioned parts of PSHA by explaining them in separate chapters of his monograph [4].

The next step in seismic risk assessment after PSHA is the seismic structural response assessment (SSRA) that has a goal to quantify the answers to the following questions:

- How sensitive is the structure being assessed to different levels of ground shaking intensity?
- What is the expected level of damage or loss given that structure becomes exposed to a certain level of ground shaking intensity?

Seismic structural response assessment is dealt with in many publications. Jack W. Baker demonstrates most widely used methods for SSRA in his paper titled “Probabilistic structural response assessment using vector valued intensity measures” [5]. One of these methods is incremental dynamic analysis also explained by Vamvatsikos and Cornell [6].

One of the key aspects of SSRA is scaling and selection of ground motions to be used for dynamic structural analyses. Baker and Cornell demonstrated hazard consistent method for ground motion scaling and selection based on using *conditional mean spectrum* as a target spectrum [7]. Conditional mean spectrum concept was later improved to become just *conditional spectrum* and it is dealt with in numerous publications [8, 9, 10, 11, 12, 13].

SSRA deals with the structural response when it is exposed to a specific shaking intensity. Ideally, it is supposed to be independent of structure's location and, thus, it only depends on structural properties of the facility whose seismic risk is being assessed. However, as it will be discussed later in the thesis, decoupling seismic hazard and structural response, respectively, into only location-dependent and only structure-dependent components of risk assessment is very hard to achieve so usually PSHA and SSRA depend on both location and structural properties of the facility being assessed.

After completing the seismic structural response assessment, seismic risk is assessed by integrating its two components together. That is, if occurrence rates of different ground shaking intensities (hazard) are known and if the expected undesired structural outcomes triggered by these intensities (structural response) are also known, the two can be decomposed under integral operation to evaluate the occurrence rate of undesired structural outcomes. The undesired outcome in this description can be some structural engineering demand parameter (e.g., maximum inter-story drift ratio) used by engineers to evaluate whether structural response complies with target performance. Besides, the financial loss that can be a decision metric for shareholders (or other authority) can also be estimated.

1.3. Thesis Objective, Content and Scope

This thesis illustrates different methods for seismic risk assessment by quantitatively evaluating seismic risk of a 6-story reinforced concrete (RC) wall building, typical of Turkish construction practice. The goal of this thesis is to give insight to the readers about advantages and disadvantages of these methods by comparing their theoretical background and results. This is achieved by developing seismic risk assessment algorithms in MATLAB [14].

The emphasis of this thesis is on the probabilistic methods for seismic risk assessment, so algorithms developed are mainly intended to demonstrate this emphasis. The structural model used for developing the structural response statistics is represented as nonlinear single-degree of freedom (SDOF) system that is simplified from the 6-story RC wall building. The details of this building and comprehensive structural analyses performed on this building (pushover, cyclic static push-pull and nonlinear response history analyses) are discussed in Özen's thesis [15].

One of the main objectives of this thesis is to compare and assess the PSRA results obtained by using different scalar intensity measures (IMs) to represent intensity of ground motion. Even though fundamental mode period spectral acceleration ($SA(T_1)$) has been shown as very efficient IM, late studies have found that one new IM, namely average spectral acceleration ($AvgSA$) is superior IM in terms of reduced uncertainty in seismic risk results [11]. This thesis compares results obtained by using $SA(T_1)$ and $AvgSA$ as IMs and provides explanations of their differences.

Another objective of this thesis is to demonstrate usefulness of random vectors and multivariate distributions as mathematical tools for development of probabilistic seismic risk assessment procedures. Complex vector IM-EDP procedure that utilizes these mathematical tools is proposed in Chapter 4.

This thesis is composed of 4 chapters. Chapter 2 of this thesis presents procedures undertaken to quantify seismic hazard at the chosen site in Istanbul. This chapter starts from definition of IMs used for probabilistic seismic hazard assessment (PSHA) and goes over source characterization, ground motion prediction, target spectra and finishes with advanced vector PSHA.

Chapter 3 of this thesis deals with seismic structural response assessment (SSRA). This chapter covers definition of building and its structural properties, ground motion scaling and selection, dynamic analyses to obtain response data and their statistical assessment.
Fourth chapter is the last chapter of this thesis and it summarizes obtained results and discusses potential future developments and improvements of the methods presented.

2. PROBABILISTIC SEISMIC HAZARD ASSESSMENT (PSHA)

This chapter explains simplified PSHA procedure used to acquire hazard curves, disaggregation data, uniform hazard spectra, conditional spectra and vector IM hazard data. This PSHA results are used as inputs for seismic structural response assessment (SSRA) presented in the following chapter.

This chapter contains 5 sections. First section represents introduction to PSHA, and it provides theoretical foundations necessary for the rest of the chapter. Second section explains the procedure to compute hazard curves and disaggregation data which are the output of PSHA. Third section explains the procedure to acquire uniform hazard spectrum. Section 4 explains hazard consistent conditional spectra and their mathematical formulation, and the last section of this chapter explains vector IM methods of PSHA.

2.1. Introduction to PSHA

Probabilistic seismic hazard assessment can be considered as part of the seismic risk assessment that is explained and discussed in many earthquake engineering related publications and literature. Baker [3] explains its essence in detail using simple examples and introduces PSHA in a way that is understandable to beginners. McGuire's monograph [4] also explains it in an easy and understandable way. In this section seismic hazard assessment procedures are summarized to provide necessary foundations to understand the hazard and risk related computations in the rest of the thesis. For more details on PSHA, the reader is referred to one of the two documents cited above.

The *seismic hazard* term represents rate (frequency) of exceedance or occurrence of a certain level of ground shaking intensity at a certain site. Sites that experience more frequent and more intense ground shakings are said to have higher seismic hazard. Even though this verbal explanation seems very straight-forward, seismic hazard is quite hard to quantify and describe mathematically. The reason for this is that the term *intensity of ground shaking* is very hard to define in a sense that is mathematically useful and efficient.

What is the intensity of ground shaking? What makes one shaking more intense than another?

Ground shakings, which are more commonly called *ground motions* according to earthquake engineering terminology, are represented by ground-motion records. Ground-motion records are measurements of ground acceleration in certain time intervals (usually with equal time intervals of 0.01 seconds to 0.005 seconds) during earthquakes. The measurement interval, also called as sampling interval, depends on the measurement frequency of device used to measure acceleration (accelerometer). Smaller sampling interval can be an indication of better record quality. Each ground-motion (GM) record is represented by three traces of acceleration measurements, also called acceleration traces. These three traces correspond to two horizontal components and one vertical component of the ground acceleration.

After being recorded, these traces are processed to minimize the errors, which appear because of the limitations in recording devices. Record processing is a complex subject, which is not one of the focuses of this thesis, so it is not discussed here. For more information on the subject, reader is referred to any signal processing textbook.
Seismic hazard at a specific site is evaluated by considering the following factors that affect it: source factors, path (attenuation) factors and local site factors. These will be discussed in the following sections.

Ground motion intensities are represented by numerous different parameters derived from acceleration traces. These parameters are called Intensity Measures (IMs).

2.1.1. Intensity Measures (IMs)

Intensity measures are parameters evaluated by various mathematical formulations using acceleration traces of ground-motion records. IMs can be used as indicators of severity of ground motions in terms of the damage they cause to structures.

What are efficient IMs?

Efficient IMs are ground-motion derived parameters that can be used to predict certain structural response quantities, also called as *engineering demand parameters* (EDPs), with high accuracy. The EDPs can be (among many others) the maximum roof displacement, a certain floor's inter-story drift ratio (maximum for the entire duration of shaking), maximum inter-story drift ratio (maximum among all the floors and for the entire duration of shaking), maximum floor acceleration (maximum among all the floors and for the duration of shaking), etc. In statistical terms, efficient IMs for predicting a certain EDP are the ones that show high correlation with that EDP.

There are many IMs widely used in earthquake engineering, so only a few of the most important ones will be mentioned here.

The simplest IM is the *peak ground acceleration* (*PGA*), which is the maximum absolute value of the ground acceleration trace:

$$PGA = \max\left(abs\left(\ddot{u}_g(t)\right)\right) \tag{2.1}$$

$\ddot{u}_g(t)$ in the above equation is the ground acceleration time history. Since there are three components of ground acceleration for every GM, there are also three *PGA* values as given below (the indices *x* and *y* denote the two mutually perpendicular horizontal components whereas *z* is the vertical component).

$$PGA_x = \max\left(abs\left(\ddot{u}_{gx}(t)\right)\right) \tag{2.2}$$

$$PGA_y = \max\left(abs\left(\ddot{u}_{gy}(t)\right)\right) \tag{2.3}$$

$$PGA_z = \max\left(abs\left(\ddot{u}_{gz}(t)\right)\right) \tag{2.4}$$

In general, horizontal and vertical components of any IM are used to predict different EDPs. Usually, the most important EDPs in terms of predicting the overall structural damage are

the ones computed from horizontal-component IMs. Vertical components are rarely used for this purpose because most structures are sensitive to horizontal ground shaking. However, vertical-component IMs are important for structures that have long consoles, beams and other horizontal elements. Since most of the structures are only sensitive to horizontal-component IMs, the vertical-component IMs are usually neglected, and the two horizontal components are considered jointly by constructing a new IM composed of some combination of the two. The combinations of horizontal IMs are to represent the directional uncertainty in GM and can be geometrical mean of the two, arithmetic mean of the two, square root of the sum of squares of the two, etc. [16]. The risk assessment in this thesis considers IMs established by geometrical mean IMs.

Pseudo-acceleration response spectrum ordinate given a specific period (hereafter spectral acceleration, designated as $SA(T)$) is the most frequently used IM. Given a SDOF system with period T , $SA(T)$ of a ground-acceleration trace (one of the horizontal components of ground motion) is computed by multiplying the corresponding displacement response spectral ordinate with $(2\pi/T)^2$. In order to compute the spectral ordinates (spectral displacement and pseudo-spectral acceleration) the Newmark's numerical method [17] is used in the MATLAB script that is given as part of the electronic supplement to this thesis. $SA(T_1)$ of a ground motion where T_1 is the fundamental mode period of the structure is a good indicator for the general response and damage in the structure. If a structural system behaves exactly like a linear SDOF system, then $SA(T_1)$ would be a perfect indicator of its response without any uncertainty. Unfortunately, this is not the case as the structural systems are best represented by nonlinear MDOF systems and their responses depend on many different parameters of the GM, or differently saying, they depend on many IMs. Simplifying the structural response prediction to a single IM increases the uncertainty because other important IMs that may influence the structural response can be overlooked. That is why, there are two different definitions of intensity measures in seismic risk assessment procedures: scalar and vector IM procedures. Scalar IM procedures are conceptually and computationally easy, but they increase the uncertainty in risk assessment since only a single IM is considered for the structural response estimation. Vector IM procedures consider joint hazards and joint effects of multiple IMs. They are better than scalar IM procedures because uncertainty is reduced, however, they are computationally (and conceptually) more demanding since joint hazard and joint fragility curves (surfaces) need to be evaluated and

integrated [5, 7]. When single $SA(T)$ is considered as the scalar IM, there is a procedure called *conditional spectrum record selection* [8, 11, 12] that can considerably mitigate shortcomings of scalar IM procedures. These topics will be discussed in following sections. Scalar and vector IM procedures will be demonstrated on two different examples.

One other way to reduce the shortcomings of scalar IM procedures is to construct new scalar IMs, which are averages of multiple (other) ordinary scalar IMs. For example, Kohrangi et al. [11] present a new IM, average spectral acceleration ($AvgSA$), which is the geometrical mean of multiple $SA(T)$ ordinates at different periods.

Weighted averages of scalar IMs can also be considered as new scalar IMs. These weights can be calibrated to represent the degrees of importance of individual IMs in estimating the target EDP. For example, this thesis introduces weighted average spectral acceleration, which is constructed by taking the weighted mean of $SA(T)$s at periods that correspond to the first few modal periods of the structure and at 1.5 times the first mode period to include nonlinear period lengthening effect. Weights are set according to the effective modal mass ratios, which are good indicators of the significance of modes contributing to the overall structural response. The weight of $SA(T)$ component at 1.5 times the fundamental period is set to be equal to the first mode weight because at this period $SA(T)$ gives a good estimate of structural response after yielding and period lengthening effects kick in. Detailed mathematical expression of this new IM will be given in following sections.

Now that IM concept has been explained, seismic hazard assessment can be described in a more precise form. As already indicated multiple times, PSHA estimates the exceedance rates of certain levels of IMs. The outcome of PSHA for a specific IM is a hazard curve, which shows the annual exceedance rates (given in y-axis) of a set of IMs (given in the horizontal x-axis) as shown in Figure 2.1 for $SA(T)$ at T = 0.96 seconds.

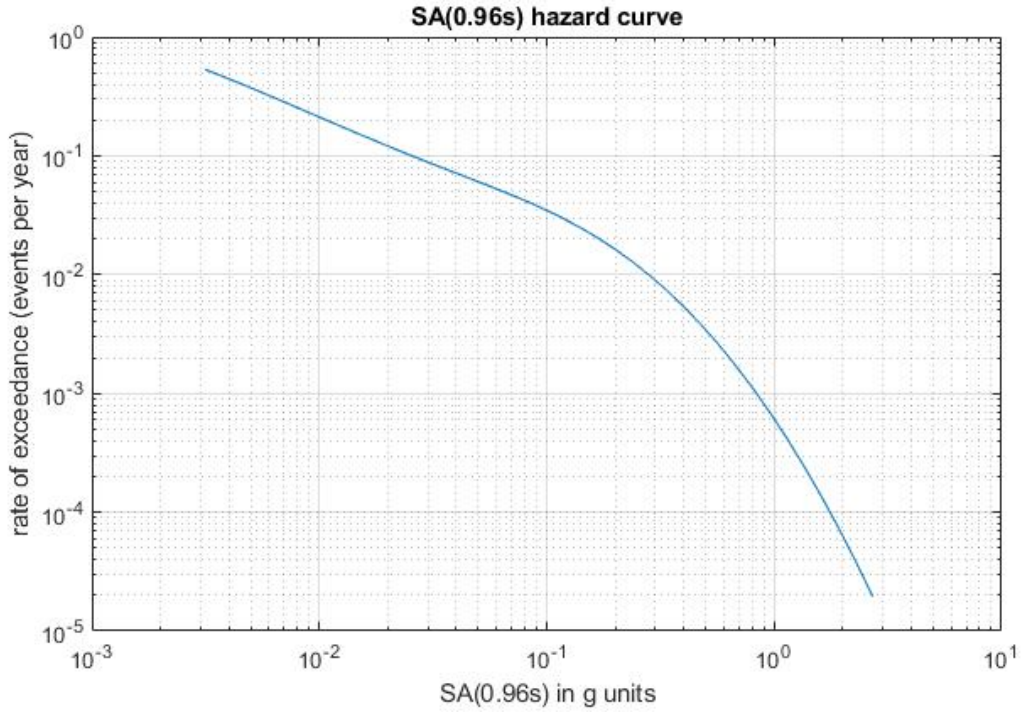

Figure 2.1: Example hazard curve of spectral acceleration at T=0.96s period

2.1.2. Coupling of Seismic Hazard Assessment and Seismic Structural Response Assessment

As mentioned before, seismic hazard assessment and seismic structural response assessment should ideally be decoupled. To open up this assertion, ideally one should be able to complete PSHA at a site without knowing any specific dynamic property of the structure, and likewise, SSRA of the structure should also be completed without knowing the location of the site. This should be the case because theoretically, seismic hazard is related to intensity and frequency of ground shaking at some site and does not depend on the future structures to be built on that site. Also, the seismic structural response is related to the dynamic properties of the structure under the influence of future shaking and does not depend on the location of the structure.

However, when using scalar IMs to represent GM intensity, the decoupling of PSHA and SSRA becomes very hard. It is only possible when using simple IMs such as PGA. But this comes at a price of high uncertainty in assessment results because PGA is a very poor estimator of structural response. GMs having same PGA can cause very different structural

responses. In order to decrease this uncertainty, IMs that are good estimators of structural response should be used, such as $SA(T_1)$ at the fundamental mode period T_1. This is where the decoupling problem starts. One should know the fundamental mode period T_1 of the structure to know which period $SA(T)$ hazard to compute. Also, as will be discussed later, one should know the $SA(T)$ hazard disaggregation to know which records to select for SSRA.

This is especially problematic from design perspective because properties of the structure to be designed are unknown in the beginning. One solution is to use a vector IM composed of $SA(T)$s at densely spaced periods covering the whole spectrum (for example from $T = 0.05s$ to $T = 4s$ in $dT = 0.05s$ increments). However, this is computationally a demanding procedure whose merits are questionable when compared to repeating scalar IM assessments multiple times.

The following sections give theoretic background and algorithms for evaluation of scalar and vector $SA(T)$ hazards.

2.2. Scalar SAT Hazard Computation

This section explains seismic hazard assessment procedure steps. This is done by demonstrating the computation of $SA(T)$ hazard curve for a site in Istanbul (site location coordinates: 40.982959N, 29.128342E).

The PSHA computation at a site can be conceptually divided into two parts. The first part deals with the occurrences rates of earthquakes (different seismic scenarios) that are likely to occur in the site vicinity and is usually addressed as Seismic Source Characterization. The second part studies the effects of those earthquakes on vibratory ground-motion intensity and aggregation of these effects to compute the hazard curve.

2.2.1. Seismic Source Characterization

The first step in PSHA is to detect the causes of the ground shakings (causes of the earthquakes). The causes of earthquakes are underground ruptures caused by movements of

tectonic plates. These ruptures occur usually at the boundaries of tectonic plates, which are called faults. When rupture occurs, some part of the energy release is directly converted to thermal energy, which heats up the surrounding rock and eventually dissipates, and some of the energy is released in the form of mechanical waves that spread through the earth medium and cause shaking of the surface, which we experience as earthquakes. Ruptures usually occur in places that they occurred before (on the tectonic plate boundaries) where the medium is already weakened by previous ruptures. However, sometimes new faults are also created depending on the developed stress levels. Locations (or regions) where underground ruptures occur are called as seismic sources in seismology and earthquake engineering. Each seismic source possesses certain activity rate, which is a rate at which it produces different sizes (magnitudes) of earthquakes. Although the seismic sources possess highly irregular geometrical shapes and non-uniform activity rates (different parts of a fault can have different seismic activities), in order to make PSHA procedures easier, their geometry is idealized as straight lines (as top view on the map) with a uniform activity rate. Fault idealization can be improved by subdividing the fault into multiple line segments each having its own activity rate. The PSHA for the chosen site in this thesis undertakes similar seismic source idealization by modeling a part of the North Anatolian Fault as a series of line sources with uniform activities.

Other than line source idealization, sources are sometimes idealized as area sources. This is a useful idealization when there are a lot of small faults spread over a large area, so practically the earthquakes can occur anywhere within the area. In this thesis, the PSHA does not incorporate any area source because the effect of North Anatolian Fault is dominant in the vicinity of Istanbul that can be successfully modelled as a line source. Figure 2.2 shows a representative map of the western segments of along the North Anatolian Fault (in the vicinity of Istanbul) and its idealization in this thesis.

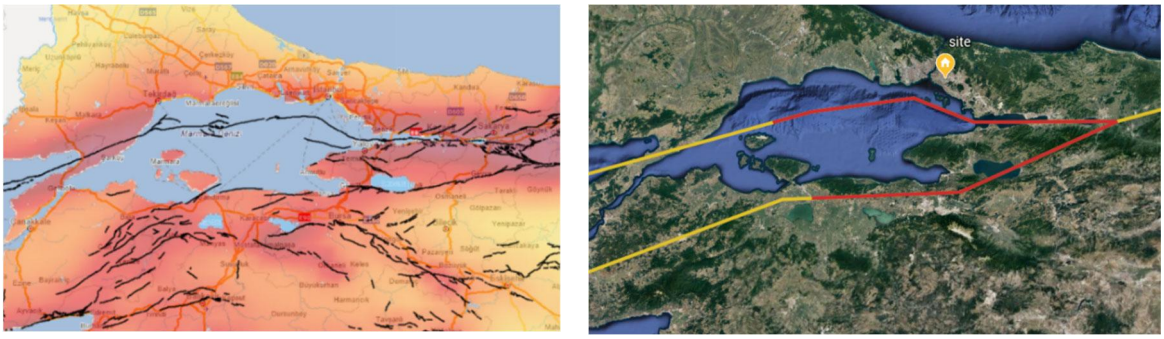

Figure 2.2: Left side: active fault map taken from AFAD's web site that display the Turkish Seismic Hazard maps (https://www.afad.gov.tr/turkiye-deprem-tehlike-haritasi). Right side: idealized line faults of the North Anatolian Fault segments used in this thesis

The North Anatolian Fault segments in the vicinity of Istanbul are modelled as two strand faults: northern and southern fault. Only parts of the fault within the 120 km radius from the site are modelled since it is assumed that the earthquakes having epicenters larger than 120 km have insignificant effect on the seismic hazard. This fact will be shown later from disaggregation results. The fault segments modelled as lines are shown in red in Figure 2.2. The northern strand is modelled with 4 piecewise linear segments (4 line sources) whereas southern strand consists of 2 piecewise linear segments. Figure 2.3 shows the line source segments and Table 2.1 gives beginning and ending coordinates of each line source.

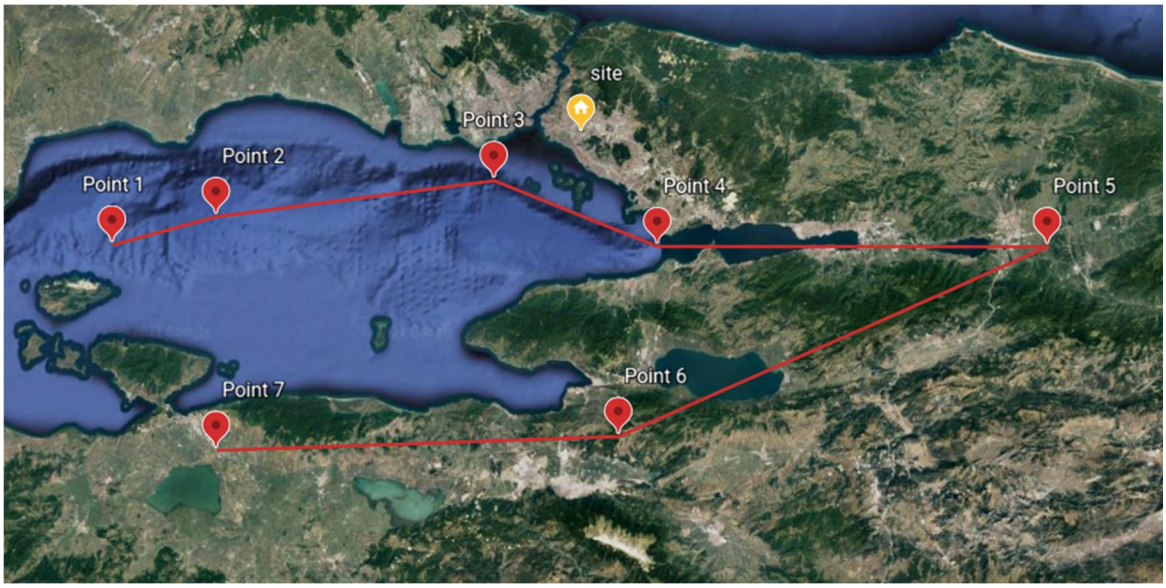

Figure 2.3: Points separating modelled piecewise linear fault segments

Table 2.1: Geographical coordinates of the site and the points separating idealized linear line sources (shown in Figure 2.3)

	Latitude	**Longitude**
Site	40.982959N	29.128342E
Point 1	40.7352143N	27.7429633E
Point 2	40.8000000N	28.0500000E
Point 3	40.8800000N	28.8700000E
Point 4	40.7300000N	29.3500000E
Point 5	40.7439887N	30.5234691E
Point 6	40.3075285N	29.2299190E
Point 7	40.2781255N	28.0534482E

The activity rates of fault segments are modelled by Youngs and Coppersmith composite model [18]. This model is a combination of Richter-Gutenberg exponential recurrence model and characteristic earthquake model [18]. It is shown to be a good representative of seismic activities for individual large crustal faults, such as San Andreas Fault and North Anatolian Fault [18]. The probability density function (PDF) of magnitude occurrences represented by Youngs and Coppersmith model is given in Eqn. 2.5-2.7:

$$f_M(m) = \frac{1}{1+c_2}\,\frac{\beta \exp\left(-\beta(m-m_0)\right)}{1-\exp\left(-\beta(M_{chr}-m_0-0{,}25)\right)} \qquad (2.5)$$

$$for \quad m_0 \le m \le M_{chr} - 0{,}25$$

$$f_M(m) = \frac{1}{1+c_2}\,\frac{\beta \exp\left(-\beta(M_{chr}-m_0-1{,}25)\right)}{1-\exp\left(-\beta(M_{chr}-m_0-0{,}25)\right)} \qquad (2.6)$$

$$for \quad M_{chr} - 0{,}25 < m \le M_{chr} + 0{,}25$$

$$c_2 = \frac{0{,}5\beta \exp\left(-\beta(M_{chr} - m_0 - 1{,}25)\right)}{1 - \exp\left(-\beta(M_{chr} - m_0 - 0{,}25)\right)} \tag{2.7}$$

The PDF model is defined by the parameters β and M_{chr}. M_{chr} determines the characteristic magnitude and β determines the negative slope of PDF in the pre-characteristic magnitude range. β is similar to b from Richter-Guttenberg model and conveys the same information by following the relation; $\beta = \ln(10)\, b$. m_0 is the minimum considered magnitude in the hazard assessment, and is chosen as $m_0 = 4$ in this thesis for PSHA.

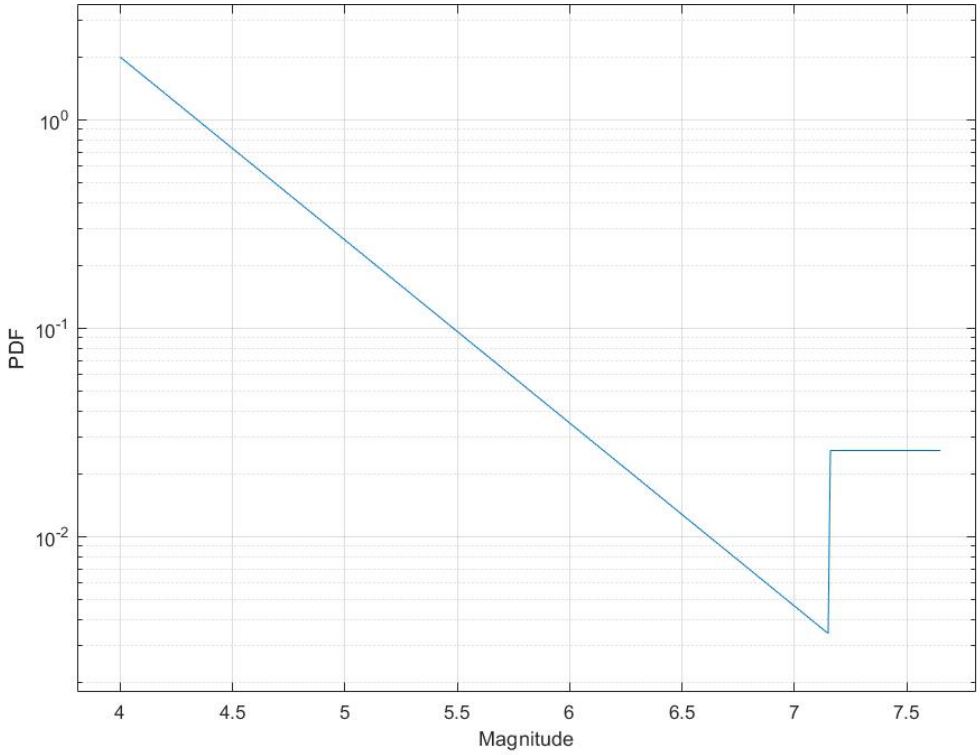

Figure 2.4: Young and Coppersmith magnitude PDF model for $M_{chr} = 7.4$, $b = 0.878$ and $m_0 = 4$.

Figure 2.4 displays the magnitude occurrence PDF example created from the Youngs and Coppersmith model (YC model). The magnitude PDF represents the yearly distribution of earthquake magnitudes, and it is displayed in logarithmic y-axis in Figure 2.4. It may not be obvious from this figure, but in reality, small earthquakes occur more frequently than large ones. The number of earthquakes usually decreases exponentially with increasing magnitude, which is demonstrated by the Guttenberg – Richter recurrence model. The

Guttenberg– Richter model agrees well while modelling the magnitude PDFs of large regions with a lot of small seismic sources. However, for individual large crustal faults it fails to predict large magnitude earthquake recurrence rates. Some large crustal faults tend to produce large earthquakes within a narrow magnitude range and this narrow range is centered at a magnitude called the characteristic magnitude (M_{chr}), and this recurrence model is called as pure characteristic earthquake recurrence model. Youngs and Coppersmith demonstrated that the Guttenberg–Richter model does not yield a good agreement for earthquake recurrences of large crustal faults. They showed that the activities of large crustal faults are best represented by the composite model with magnitude PDFs containing exponential decay and flat characteristic magnitude range. In Figure 2.4, the exponential decay range is between magnitudes 4 and 7.15, and characteristic flat magnitude range is between magnitudes 7.15 and 7.65. Earthquakes above 7.65 are assumed to be physically impossible for this fault to produce, so their PDF ordinate is zero.

Having defined the magnitude recurrence PDF model, a certain magnitude exceedance rate can be evaluated by integrating PDF between that magnitude and infinity and multiplying by the exceedance rate of the smallest considered magnitude in the hazard (in this thesis, only earthquakes of $m_0 = 4$ and above are considered and hence the minimum rate of exceedance is $\lambda_{M>4}$).

$$\lambda_{M>m} = N(m) = N(m_0) \int_m^{\infty} f_M(\tau) d\tau \tag{2.8}$$

$\lambda_{M>m}$ and $N(m)$ are the nomenclatures for an exceedance rate of magnitude m. $N(m_0)$ is a the rate of exceedance of the minimum magnitude considered (m_0). If f_M from YC model is substituted to the equation above, the expression given in Eqn. 2.9 – 2.10 is obtained.

$$N(m) = \frac{N(m_0)}{1 + c_2} \left(\frac{\exp\left(-\beta(m - m_0)\right) - \exp\left(-\beta(M_{chr} - m_0 - 0{,}25)\right)}{1 - \exp\left(-\beta(M_{chr} - m_0 - 0{,}25)\right)} + \frac{0{,}5\beta \exp\left(-\beta(M_{chr} - m_0 - 1{,}25)\right)}{1 - \exp\left(-\beta(M_{chr} - m_0 - 0{,}25)\right)} \right) \tag{2.9}$$

$$for \quad m_0 \leq m \leq M_{chr} - 0{,}25$$

$$N(m) = \frac{N(m_0)}{1 + c_2} \frac{\beta \exp\left(-\beta(M_{chr} - m_0 - 1{,}25)\right)}{1 - \exp\left(-\beta(M_{chr} - m_0 - 0{,}25)\right)} (M_{chr} - m + 0{,}25) \tag{2.10}$$

$$for \quad M_{chr} - 0{,}25 < m \leq M_{chr} + 0{,}25$$

These expressions are normalized by $N(m_0)$ and plotted for $M_{chr} = 7.4$, $b = 0.878$ and $m_0 = 4$ in Figure 2.5.

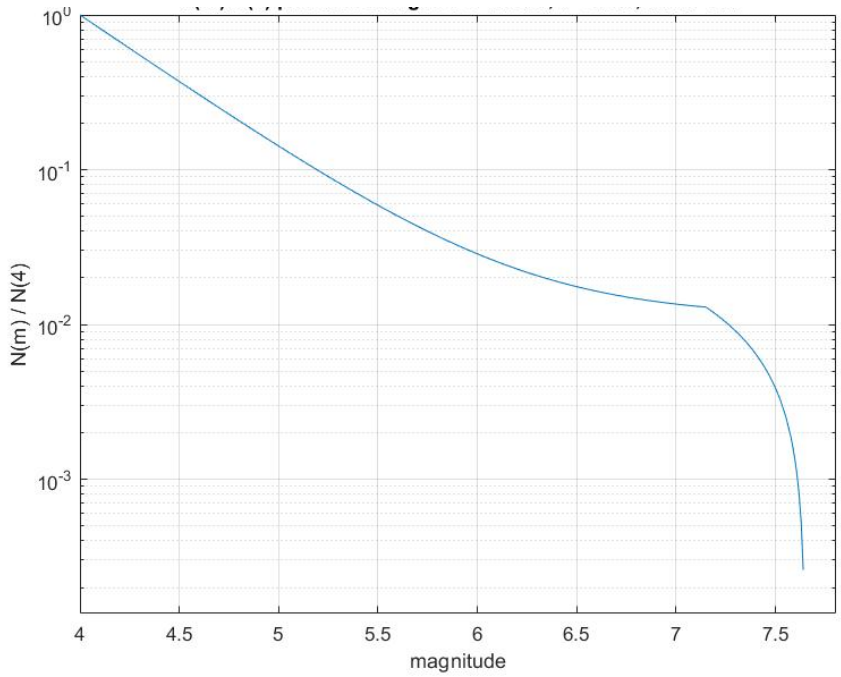

Figure 2.5: Magnitude exceedance rate normalized by total earthquake rate plot according to YC recurrence model for $M_{chr} = 7.4$, $b = 0.878$ and $m_0 = 4$.

In order to model different magnitude recurrences of the line sources using YC model, we need to determine 3 parameters for every line source: $N(m_0)$, β and M_{chr}. Calibrating these parameters for the best representation of seismicity is a difficult task and would require statistical studies on historical seismicity of the faults considered as well as the geological studies of the area surrounding the faults for more input in cases when the historical data are limited. Since the main focus of this thesis in not a detailed study of the seismicity but rather the demonstration of the PSHA and SSRA methods, these parameters are estimated from Utku's work [19]. In his paper, in Table 4, Utku [19] gives the estimated Guttenberg-Richter

a and b parameters for both northern and southern strands of the North Anatolian Fault in the Marmara region. The northern strand is defined as stretching between longitudes 24E and 30.5E, and the southern strand is between 24.95E and 30.5E [19]. The Guttenberg-Richter parameters from Utku [19] are given in Table 2.2:

Table 2.2: Guttenberg-Richter parameters given in Utku 2018 [19] for northern and southern strands of North Anatolian Fault in the Marmara region

Northern Strand	$a = 4.065, b = 0.878$
Southern Strand	$a = 3.614, b = 0.824$

$N(m_0)$ for northern and southern strands can be evaluated by the Guttenberg-Richter formulation as given in Eqn. 2.11 – 2.13:

$$N(m) = 10^{a-bm} \tag{2.11}$$

$$N_{northern}(4) = 10^{4.065-0.878x4} = 3.5727\, events/year \tag{2.12}$$

$$N_{southern}(4) = 10^{3.614-0.824x4} = 2.0797\, events/year \tag{2.13}$$

These are the activities of fault strands as defined in Utku 2018 [19]. Northern strand is between longitudes 24E and 30.5E, and its length is 569.45 km. This length is the measured length of the idealized fault model given in Figure 2.2 (right panel) between the above two longitudes. Southern strand is defined in Utku 2018 [19] as stretching between longitudes 24.95E and 30.5E, and its idealized length is measured as 507.96 km. The seismic activities are assumed to be uniform along the lengths of northern and southern strands. Therefore, $N(m_0)$ rates of individually modelled segments given in Figure 2.3 are estimated by scaling the rates of the whole strands by considering the respective lengths of each segment, as demonstrated in Eqn. 2.14:

$$N_{linear\ segmet}(4) = N_{strand}(4)\frac{L_{linear\ segment}}{L_{strand}} \tag{2.14}$$

$N_{linear\ segmet}(4)$ in the Eqn. (2.14) represents mean number of events (earthquakes) per year with magnitudes larger than 4 occurring at a specific idealized linear segment (Figure 2.3). $N_{strand}(4)$ represents mean number of events per year with magnitudes larger than 4 occurring at whole strand to which specific linear segment belongs. *L* terms represent lengths of idealized linear segment and the strand to which it belongs.

β is estimated by evaluating it from the b parameters given in Utku 2018, namely, for northern strand segments $\beta_{northern} = \ln(10)\, b_{northern} = \ln(10)\, x0.878 = 2.0217$ and for southern strand segments $\beta_{southern} = \ln(10)\, b_{southern} = \ln(10)\, x0.824 = 1.8973$.

M_{chr} is assumed to be 7.4 for all segments. This value corresponds to a 100-year return period magnitude for northern strand that is also estimated by Utku and is given in Table 4 of his paper [19]. Since the North Anatolian Fault is a strike-slip fault, although the certain segments of the same fault feature other styles-of-faulting, all line segments in this thesis are accepted as strike-slip segments only capable of producing strike-slip ruptures. Dip angles of faults are idealized as $90°$ even though in reality this is not the case. In reality, dip angles are around $75°$. The reason of this idealization is to simplify the characterization of the subject fault. The use of $75°$ dip angle would certainly produce accurate source-to-site distance metric computation. However, it would make the development of distance PDF function complicated, which is one of the integral parts of the PSHA algorithm developed for this study. The consideration of $90°$ dip angle enforces the epicenters to take place along the fault line, which is not the case for $75°$ dip angle where epicenter also depends on the depth of the rupture. On the other hand, this simplification results in a simpler distance PDF to be used in the PSHA integral.

It is clear that this approach in seismic source characterization is not very sophisticated and accurate. It is based on many approximations and assumptions, which are taken into account in order to avoid complicated models. As it is already stated in the previous paragraphs, PSHA is rather a tool in this thesis used to achieve the main objectives. Therefore, the above approximations and simplifications for the computation of PSHA are considered to be sufficient.

After the geographical orientation of seismic sources and defining their activities, this information can be used to evaluate occurrence rates of seismic scenarios. The seismic scenarios are defined in the PSHA terminology as the occurrences of ruptures having certain magnitudes *M* and certain source-to-site distances *R* (and other parameters describing the rupture such as style-of-faulting, etc.). Since the occurrence probability of any exact *M* and *R* values is zero from probabilistic aspect, scenarios in this thesis are represented by *M* and *R* bins. The *M* and *R* bins are narrow ranges of *M* and *R* values which are represented as single seismic scenarios. In order to make PSHA computationally feasible *M* and *R* ranges have to be grouped in finite number of bins and these bins should be represented as scenarios with non-zero probabilities of occurrence (or rates of occurrence).

If the seismic scenario is defined as the occurrence of *M* - *R* pairs, the rate of occurrence of a specific scenario equals to the rate of occurrence of any considered scenario ($N(m_0)$) multiplied by the magnitude and source-to-site distance probability density functions, as demonstrated in Eqn. 2.15:

$$\lambda_{M=m,R=r} = \lambda_{earthquakes}\, x\, P(M = m, R = r) = N(m_0) f_M(m) f_R(r) dmdr \qquad (2.15)$$

In the above equation, $\lambda_{M=m,R=r}$ is the occurrence rate of magnitude m and distance r. $\lambda_{earthquakes}$ is the occurrence rate of any event and it is equal to the exceedance rate of m_0, $N(m_0)$, since only events with magnitudes higher than m_0 are considered in PSHA. $f_M(m)$ and $f_R(r)$ are magnitude and distance-to-site PDFs, respectively. One important assumption in Eqn. (2.15) is that *M* and *R* are independent, so the joint PDF of *M* and *R* is just the product of their marginal PDFs. This may not hold true in reality because different parts of the faults may be prone to producing different magnitude earthquakes making magnitude dependent on the location hence source-to-site distance. However, since the uniform activity rate along the fault length assumption is made while modeling the activity of each line source, such a complication is avoided, and *M* and *R* distributions are accepted as independent.

$\lambda_{M=m,R=r}$ from Eqn. 2.15 is essentially equal to zero since the differential terms dm and dr are equal to zero. As discussed in previous text, the reason behind is that the occurrence of some exact *M* and *R* valued scenario has zero probability. In order to make PSHA computations numerically more feasible, the scenarios are represented as occurrences of

seismic events with *M* and *R* values being within certain *M* and *R* bins. According to this scenario description, their rate of occurrence is given in Eqn. 2.16:

$$\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}} = \lambda_{earthquakes}\, x\, P(m_i < M \leq m_{i+1}, r_k < R \leq r_{k+1}) = \\ = N(m_0)\big(F_M(m_{i+1}) - F_M(m_i)\big)\big(F_R(r_{k+1}) - F_R(r_k)\big) \tag{2.16}$$

In Eqn. 2.16, $\lambda_{m_i<M\leq m_{i+1},r_i<R\leq r_{i+1}}$ is the rate of occurrence of events with *M* within *i'*th and *R* within *k'*th bin. F_M and F_R are the cumulative distribution functions (CDFs) of magnitude and source-to-site distance, respectively.

Since scenarios are defined as events determined by their *M* and *R* values, the same scenario can occur at multiple locations as long as these locations have source-to-site distances within the same *R* bin. The same scenario can even originate from different sources, or from different line sources. In order to incorporate these cases, Eqn. 2.16 is modified as shown in Eqn. 2.17:

$$\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}} = \\ = \sum_{n=1}^{\#\ of\ sources} N_n(m_0)\big(F_M^n(m_{i+1}) - F_M^n(m_i)\big)\big(F_R^n(r_{k+1}) - F_R^n(r_k)\big) \tag{2.17}$$

Eqn. 2.17 sums the contributions of all sources to the scenario rates. $N_n(m_0)$ represents the rate of events having *M* larger than m_0 and occurring at source *n*. In the context of this thesis, it represents the minimum seismic activities of idealized line sources. Since there are six fault segments (Figure 2.3), the summation index *n* goes from 1 to 6. F_M^n and F_R^n represent, respectively, magnitude and source-to-site distance CDFs of *n'*th source.

The CDFs can be computed by integrating the PDFs from negative infinity to the target value. F_M is evaluated by integrating $f_M(m)$ from m_0 since lower magnitudes are not considered in PSHA – Eqn. 2.18.

$$F_M^n(m) = \int_{m_0}^{m} f_M^n(\tau) d\tau \tag{2.18}$$

Substituting f_M from Eqn. 2.5-2.7, one can obtain the closed form solution of F_M^n for Youngs and Coppersmith recurrence model.

$$F_M^n(m) = \frac{1}{1 + c_2^n} \frac{1 - \exp\left(-\beta_n (m - m_0)\right)}{1 - \exp\left(-\beta_n (M_{chr} - m_0 - 0.25)\right)} \tag{2.19}$$

$$for \qquad m_0 \leq m \leq M_{chr} - 0.25$$

$$F_M^n(m) = \frac{1}{1 + c_2^n} \left(1 + \frac{\beta_n \exp\left(-\beta_n (M_{chr} - m_0 - 1.25)\right)}{1 - \exp\left(-\beta_n (M_{chr} - m_0 - 0.25)\right)} (m - M_{chr} + 0.25) \right) \tag{2.20}$$

$$for \qquad M_{chr} - 0.25 < m \leq M_{chr} + 0.25$$

$$c_2^n = \frac{0.5 \beta_n \exp\left(-\beta_n (M_{chr} - m_0 - 1.25)\right)}{1 - \exp\left(-\beta_n (M_{chr} - m_0 - 0.25)\right)} \tag{2.21}$$

Different β parameters are used for the northern and southern strands segments, so β parameter depends on the segment, and, for clarification, the subscript *n* is associated with β. Since the same M_{chr} value is used for all sources, the subscript *n* is omitted for this parameter.

When it comes to F_R^n and f_R^n, the source-to-site distance CDFs and PDFs, respectively their computation involves planar geometry. Since the seismic activities of idealized linear fault segments are modelled as uniform along their lengths, $F_R^n(r)$ is simply a *normalized portion* of the segment *n* that resides within the circle of radius *r* with center at the site whose PSHA integral is being computed. The *normalized portion* in the previous statement is the ratio between the portion falling within the said circle of radius *r* and the total length of the

idealized linear fault segment. Figure 2.6 illustrates the above descriptions and Eqn. 2.22 represents the source-to-site distance CDF for the *n*'th segment.

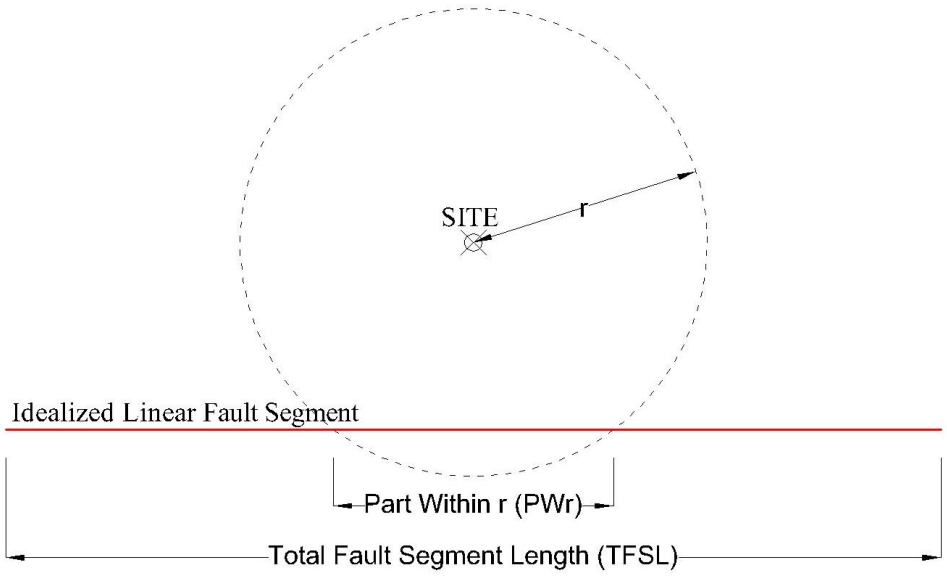

Figure 2.6: Idealized Linear Fault Segment and Part of the Same Within r Distance-From-Site

$$F_R^n = \frac{PWr}{TFSL} \tag{2.22}$$

In case, the activities of fault segments are not accepted as uniform along the segment length, the evaluation of F_R^n would be much more complex because the CDF of R would have to be constructed to reflect the heterogeneous variation of fault activity along the segment.

There are two distinct spatial orientations between the site and fault segment as demonstrated in Figure 2.7. The fault can be oriented such that the normal line drawn from the site intersects the fault line either within (Figure 2.7, left panel) or out (Figure 2.7, right panel) of the fault segment. These two distinct orientations have distinct derivations of F_R^n as given in Eqn. 2.23-2.29.

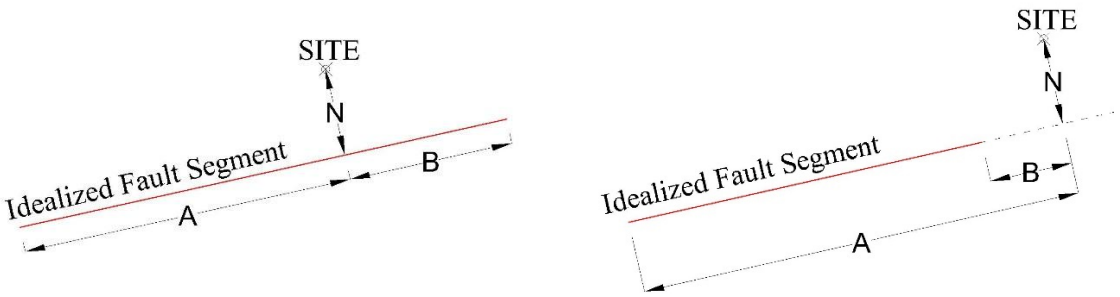

Figure 2.7: Two distinct orientations between fault segment and site location

For the orientation between the site and fault segment as displayed in the left panel of Figure 2.7, F_R can be derived as:

$$F_R(r) = 0 \quad for \quad 0 \le r \le N \tag{2.23}$$

$$F_R(r) = \frac{2\left(\sqrt{r^2 - N^2}\right)}{A + B} \quad for \quad N < r \le \sqrt{N^2 + B^2} \tag{2.24}$$

$$F_R(r) = \frac{B + \sqrt{r^2 - N^2}}{A + B} \quad for \quad \sqrt{N^2 + B^2} < r \le \sqrt{N^2 + A^2} \tag{2.25}$$

$$F_R(r) = 1 \quad for \quad r > \sqrt{N^2 + A^2} \tag{2.26}$$

For the orientation as shown in the right panel of Figure 2.7, F_R is given as:

$$F_R(r) = 0 \quad for \quad 0 \le r \le \sqrt{N^2 + B^2} \tag{2.27}$$

$$F_R(r) = \frac{\sqrt{r^2 - N^2} - B}{A - B} \quad for \quad \sqrt{N^2 + B^2} < r \le \sqrt{N^2 + A^2} \tag{2.28}$$

$$F_R(r) = 1 \quad for \quad r > \sqrt{N^2 + A^2} \tag{2.29}$$

Parameters A, B and N in the above equations can be evaluated from simple geometry if the coordinates of the site and fault end points are known. The hazard calculation algorithm presented in this thesis accomplishes this by ignoring Earth's curvature and converting geographical coordinates of these 3 points to Cartesian coordinate system by taking the site at its origin and unit distance being equal to 1km. Ignoring Earth's curvature is a good approximation for this case since source-to-site distances between the fault segments and the site are at the most 200 km, making Earth's curvature effect negligible. The flat Cartesian coordinate system is chosen such that one axis extends in the North-South, and other axis in the East-West direction. Coordinates of the fault end points on this planar coordinate are computed by evaluating the distances from site to the meridians and parallels passing through the fault endpoints (Figure 2.8).

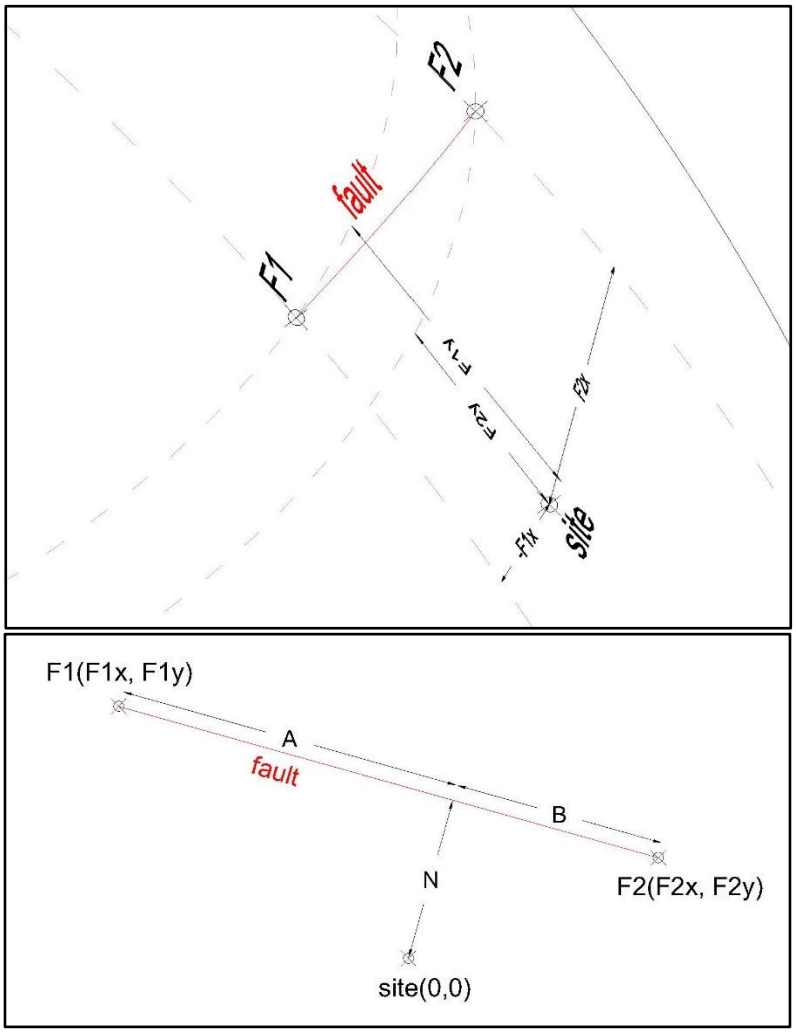

Figure 2.8: Transformation of geographical coordinates of the site and fault end points to a flat-plane Cartesian coordinate system

After the coordinates of fault end points *F1* and *F2* are established in kilometer units, A, B and N parameters are computed by simple geometrical expressions. which yield the following results displayed in Eqn. 2.30 – 2.34:

$$N = \sqrt{{X_I}^2 + {Y_I}^2} \tag{2.30}$$

$$A = \sqrt{(F1_x - X_I)^2 + \left(F1_y - Y_I\right)^2} \tag{2.31}$$

$$B = \sqrt{(F2_x - X_I)^2 + \left(F2_y - Y_I\right)^2} \tag{2.32}$$

$$X_I = \frac{\frac{F1_y - F2_y}{F1_x - F2_x} F1_x - F1_y}{\frac{F1_x - F2_x}{F1_y - F2_y} + \frac{F1_y - F2_y}{F1_x - F2_x}} \tag{2.33}$$

$$Y_I = \frac{\frac{F1_x - F2_x}{F1_y - F2_y} F1_y - F1_x}{\frac{F1_x - F2_x}{F1_y - F2_y} + \frac{F1_y - F2_y}{F1_x - F2_x}} \tag{2.34}$$

In Eqn. 2.30 – 2.34, parameters X_I and Y_I represent the coordinates of the nearest point on the fault line to the site (the point where perpendicular line form site to the fault line intersect each other). Having defined all terms in Eqn. 2.17, which are necessary to evaluate the occurrence rates of seismic scenarios, source characterization part of PSHA is completed. The next component of PSHA is estimation of the impacts of these scenarios on ground shaking intensity. This is done by employing probabilistic models that are called as ground-motion prediction equations (GMPEs) or ground-motion predictive models (GMPMs).

2.2.2. Aggregation of Different Seismic Scenario Effects

This section explains the effects of different seismic scenarios (earthquakes) on GM intensity measure at the site. These effects are studied in a probabilistic manner and are modelled by models called as GMPEs.

2.2.2.1. Ground Motion Prediction Equations (GMPEs)

Ground-motion prediction equations (GMPEs) are probabilistic models used to predict the GM intensity in terms of IMs given the certain seismic scenario has occurred. They are probabilistic models that consider large numbers of GM records from many earthquakes. Performing regression analyses using these GM records, GMPEs are developed to give the logarithmic mean and standard deviation of a specific IM given that some seismic scenario (*M*, *R*, Style of Faulting - *SoF*, etc.) has occurred (Eqn. 2.35)

$$\left[\mu_{\ln\,(IM)|Rup}, \sigma_{\ln\,(IM)|Rup}\right] = GMPE(Rup) \tag{2.35}$$

$\mu_{\ln\,(IM)|Rup}$ in the above equation is the logarithmic mean of a specific *IM* expected to occur at the site given that certain seismic scenario *Rup* has occurred. Seismic scenario *Rup* is usually defined by source related parameters such as *M*, *R* and *SoF* as well as site condition related parameters such as *Vs30*.

$IM|Rup$ is usually assumed to be lognormal random variable (because IMs do not attain negative values and have skewed distributions by nature), so its logarithmic mean and standard deviation ($\mu_{\ln\,(IM)|Rup}$ and $\sigma_{\ln\,(IM)|Rup}$) are sufficient to define its PDF and CDF:

$$ln(IM)|Rup \sim \mathcal{N}\left(\mu_{ln(IM)|Rup}, (\sigma_{ln(IM)|Rup})^2\right) \tag{2.36}$$

$$f_{ln(IM)|Rup}(im) = \frac{1}{\sigma_{ln(IM)|Rup}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{ln(im)-\mu_{ln(IM)|Rup}}{\sigma_{\ln(IM)|Rup}}\right)^2} \tag{2.37}$$

$$F_{IM|Rup}(im) = \Phi\left(\frac{\ln(im) - \mu_{\ln(IM)|Rup}}{\sigma_{\ln(IM)|Rup}}\right) \tag{2.38}$$

Exceedance probability of a certain IM level given that a specific seismic scenario has occurred is equal to the complementary cumulative distribution function (CCDF):

$$P(IM > im|Rup) = G_{IM|Rup}(im) = 1 - F_{IM|Rup}(im) \tag{2.39}$$

There are many different GMPEs for different IM parameters. The GMPE employed in this thesis is developed by Akkar et al. [20, 21] (hereafter ASB14). This GMPE can be used to estimate $SA(T)$ at any period for a given seismic scenario defined by magnitude, source-to-site distance, style-of-faulting and Vs30 (the average shear wave velocity within the top 30 meters of the soil column at the site). The ASB14 enables the use of three different types of source-to-site distance parameters, namely, Joyner-Boore distance (R_{JB}), epicentral distance (R_{epi}), hypocentral distance (R_{hyp}). Of these distance metrics, this thesis employs the epicentral distance, R_{epi}. The reason behind this decision is that the fault segments are modelled as linear strike-slip segments with 90° dip angle where ruptures are represented as events occurring at single locations, rendering R_{hyp} and R_{JB} not better than R_{epi}. This simplification is inaccurate for large magnitude ruptures as single-point idealization would fail to represent the actual path affect.

The ASB14 formulation [20] is repeated in Eqn. 2.40-2.42:

$$\ln(Y) = \ln[Y_{REF}(M_\omega, R, SoF)] + \ln[S(V_{s30}, PGA_{REF})] + \varepsilon\sigma \tag{2.40}$$

$$\ln(Y_{REF}) = \begin{cases} a_1 + a_2(M_\omega - c_1) + a_3(8.5 - M_\omega)^2 + [a^4 + a^5(M_\omega - c_1)]\ln\left(\sqrt{R^2 + {a_6}^2}\right) \\ +a_8F_N + a_9F_R \qquad for\ M_\omega \le c_1 \\ a_1 + a_7(M_\omega - c_1) + a_3(8.5 - M_\omega)^2 + [a^4 + a^5(M_\omega - c_1)]\ln\left(\sqrt{R^2 + {a_6}^2}\right) \\ +a_8F_N + a_9F_R \qquad for\ M_\omega > c_1 \end{cases} \tag{2.41}$$

$$\ln(S) = \begin{cases} b_1 \ln(V_{s30}/V_{REF}) + b_2 \ln\left[\dfrac{PGA_{REF} + c(V_{s30}/V_{REF})^n}{(PGA_{REF} + c)(V_{s30}/V_{REF})^n}\right] & for\ V_{s30} \le V_{REF} \\ b_1 \ln\left[\dfrac{min(V_{s30}, V_{CON})}{V_{REF}}\right] & for\ V_{s30} > V_{REF} \end{cases} \quad (2.42)$$

a_1 to a_9, b_1, b_2, c_1, V_{REF}, V_{CON}, c and n coefficients are provided for PGA and $SA(T)$ values at 63 different periods densely spaced between 0.01s and 4s. PGA_{REF} is the log mean PGA predicted for reference site condition ($V_{s30} = V_{REF} = 750\ m/s$). To compute the means of logarithms of IMs, PGA_{REF} is first calculated by employing another set of regression coefficients (Eqn. 2.41). In this case, $\ln(S) = 0$ (the site model) because $V_{s30} = V_{REF}$ = 750 m/s. After PGA_{REF} is computed, one can employ the site model (Eqn. 2.42) for actual site conditions of interest (for specific V_{s30} values) and compute the means and standard deviations of logarithms of $SA(T)$s for the period (or periods) of interest (Eqn. 2.40). For spectral periods that do not match with the 63 discrete period ordinates for computing $\mu_{\ln(SAT)|Rup}$ and $\sigma_{\ln(SAT)|Rup}$, one should employ interpolation to estimate the required $SA(T)$ value.

2.2.2.2. Aggregation

Now that the occurrence rates of different scenarios can be evaluated and that their effects on ground shaking intensity can be quantified, the exceedance rates at different levels of GM intensities can be computed. It is done by aggregating the contributions from all considered scenarios:

$$\lambda_{IM \ge im} = \sum_{i=1}^{M\ bins} \left(\sum_{k=1}^{R\ bins} \lambda_{m_i < M \le m_{i+1}, r_k < R \le r_{k+1}} G_{IM|M=\frac{m_i+m_{i+1}}{2}, R=\frac{r_k+r_{k+1}}{2}}(im) \right) \quad (2.43)$$

$\lambda_{IM \ge im}$ in the above equation is the occurrence rate of GMs at the site with IM ≥ *im*. In other words, it is the exceedance rate of a given IM level. It is achieved by summing the contributions from all considered seismic scenarios. These contributions are equal to the occurrence rates of considered seismic scenarios multiplied by the CCDF of IM provided

that the seismic scenario has occurred. $\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}}$ from the above equation is the rate of occurrence of scenario with M being in the i'th bin and R being in the k'th bin. $G_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)$ is the CCDF of IM provided that the seismic scenario has occurred. It represents probability of IM > im given that the seismic scenario has occurred. It can be evaluated by employing GMPEs as discussed in previous section. To evaluate mean and standard deviation by GMPE, the mid-values of M and R bins of scenarios are used, namely $M = \frac{m_i+m_{i+1}}{2}$ and $R = \frac{r_k+r_{k+1}}{2}$. The above equation can be evaluated for different im levels in order to acquire IM exceedance rate curve or IM hazard curve. The latter terminology (i.e., IM hazard curve) is more commonly used in earthquake engineering.

Different scenario contributions can also be aggregated to compute the occurrence rate of an IM level rather than its exceedance. The occurrence of the exact IM level is practically impossible, so under the term "rate of occurrence" two of the followings are referred: (a) mean rate density (MRD) function, and (b) occurrence rate of an IM range. The rate of occurrence of an IM range can be computed by finding the difference between the upper and lower limits of the exceedance rate in the range. The mean rate density (MRD) function of IM is a function whose integral over a specific range gives the occurrence rate of IM for that range. It is simply computed by replacing the complementary cumulative distribution function (CCDF) by probability density function (PDF) of IM in Eqn. 2.43 and it is given below in Eqn. 2.44:

$$MRD_{IM}(im) = \sum_{i=1}^{M\ bins}\left(\sum_{k=1}^{R\ bins}\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}} f_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)\right) \quad (2.44)$$

2.2.2.3. Disaggregation

Disaggregation refers to individual contributions of different seismic scenarios to various IM levels. Aggregation sums up the contributions of seismic scenarios to IM exceedance or occurrence rate whereas disaggregation provides the contributions of individual rates of all scenarios used in aggregation. To compute the contribution rate of a specific seismic scenario to overall hazard, its contribution is divided by the aggregated effect from all scenarios (Eqn. 2.45)

$$D_{IM\geq im}(k,i) = \frac{\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}} G_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)}{\lambda_{IM\geq im}} \quad (2.45)$$

In the above equation, $D_{IM\geq im}$ is the exceedance disaggregation function at hazard level im. It shows the contribution of seismic scenario $m_i < M \leq m_{i+1}$ and $r_k < R \leq r_{k+1}$ to overall exceedance rate IM ≥ im ($\lambda_{IM\geq im}$).

Disaggregation can also be computed for occurrence rate of some IM level:

$$D_{IM=im}(k,i) = \frac{\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}} f_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)}{MRD_{IM}(im)} \quad (2.46)$$

$D_{IM=im}$ is the occurrence disaggregation of IM at hazard level im.

2.2.3. Algorithm for Scalar Spectral Acceleration (SAT) Hazard Evaluation

The algorithm for computing $SA(T)$ hazard curve and disaggregation for any site in the vicinity of Istanbul was developed in MATLAB programing language [14]. The procedures described in the previous sections were implemented as a MATLAB script for scalar $SA(T)$ hazard computation and it is provided as an electronic supplement to this thesis.

The algorithm uses the site coordinates and local soil conditions (*Vs30* value) to compute the $SA(T)$ hazard curve at a specific period, mean rate density function, exceedance and occurrence disaggregation data. Figure 2.9 presents a sample $SA(T)$ hazard curve and corresponding disaggregation plot as example outputs of this script. The plots presented in Figure 2.9 are for $SA(T)$ hazard at $T = 0.50s$ at the chosen site (40.982959N, 29.128342E, $V_{S30} = 375\, m/s$). The disaggregation plot is computed for 475-year return period.

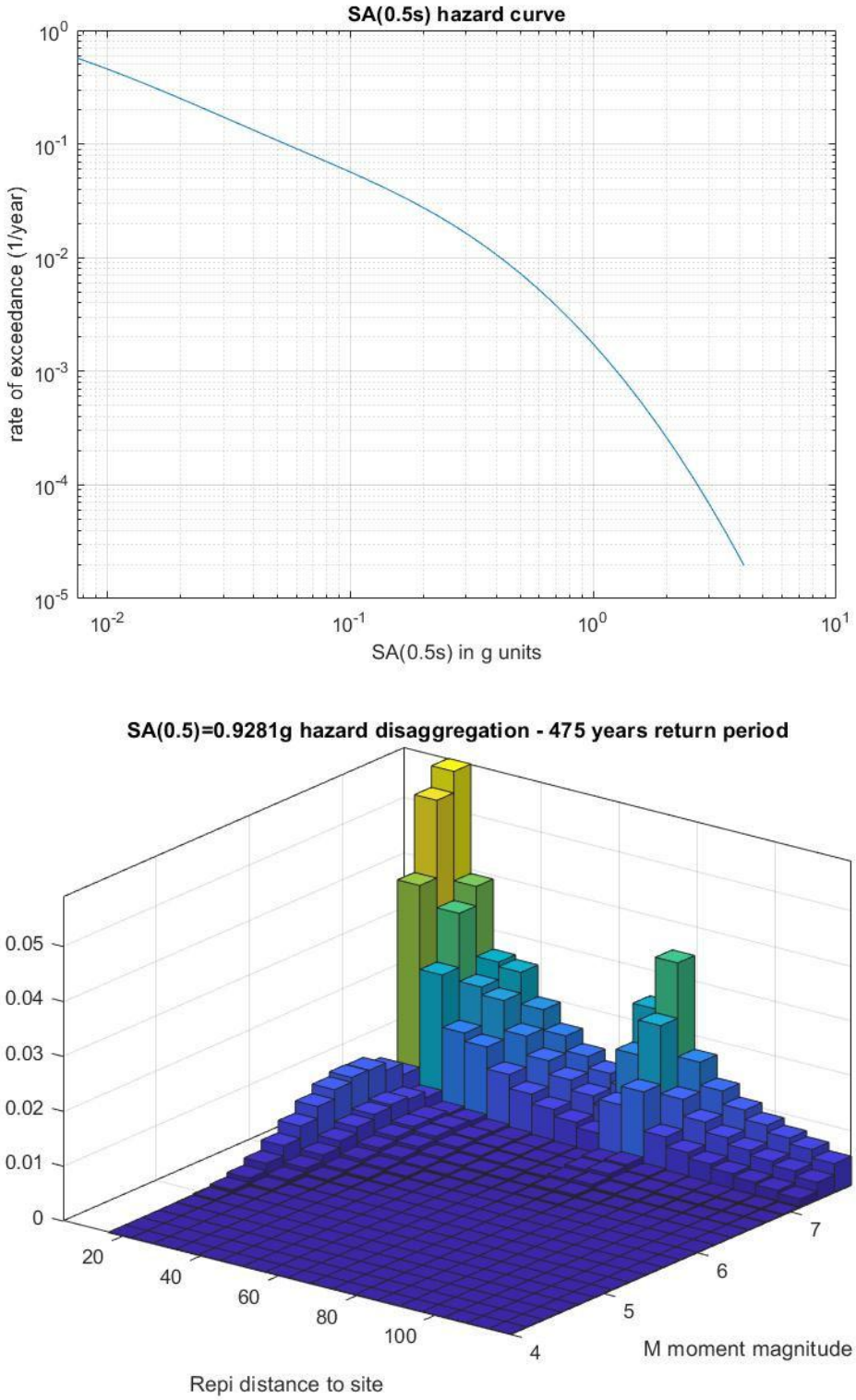

Figure 2.9: Upper plot shows SA(0.50s) hazard curve, lower plot represents occurrence disaggregation of SA(0.50) at 475 years return period hazard level. Disaggregation is displayed in 400 bins; 20 M and 20 R bins were considered.

2.3. Uniform Hazard Spectrum

The spectral ordinates of a uniform hazard spectrum (UHS) represent the same hazard level in terms of exceedance rate (or return period). UHS is frequently represented by acceleration spectrum. UHS ordinates are computed by performing scalar $SA(T)$ hazard calculation at many closely spaced periods and then combining the $SA(T)$ values from each hazard curve that correspond to the same exceedance rate. A high resolution UHS demands many scalar hazard computations at densely spaced periods. However, for all practical purposes UHS can be constructed by conducting a limited number of scalar $SA(T)$ hazard computations at periods spaced with increments of 0.1s.

In this thesis, the uniform hazard response spectra are computed at 54 discrete periods to span the important modal periods of the model building employed in the seismic structural response assessment. The period range accounts for the modal period shifts caused by the nonlinear response of the model building. Figure 2.10 shows an illustrative set of UHS at four different hazard levels for the site considered in the case study.

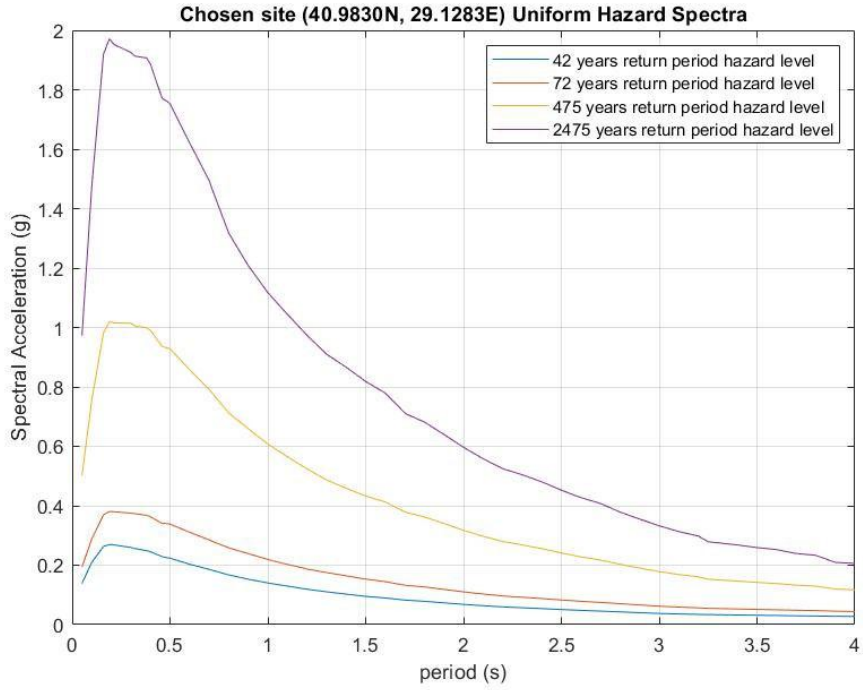

Figure 2.10: Uniform Hazard Spectra for 4 distinct hazard levels (represented by return periods)

The UHS can be used for ground-motion selection and scaling for SSRA [22]. The current seismic design code practice makes use of the UHS concept and provides seismic hazard maps to establish UHS compatible design spectra for designing and performance assessment of building systems (Turkish Earthquake Building Code 2018 [22]). Consequently, the UHS compatible design spectrum is used as the target spectrum for scaling the GM records.

For example, the aforementioned code, considers the 2475-year return period UHS (or its idealized design spectrum counterpart) to assess the performance of structures against non-collapse condition. The seismic performance of the building should be such that it should not collapse under the GM intensity that is expected to occur once in 2475 years. Assuming the earthquake phenomenon and ground motions follow Poisson process, the 2475-year return period indicates that the subject GM intensity level has an exceedance probability of 2% in 50 years. In order to show that the building will not collapse under this ground motion demand, the engineers perform either static or dynamic analyses using the representative UHS of 2475-year return period. For the latter case, the engineers use scaled GM records that fit the target UHS (or the design spectrum representative of UHS with 2475-year return period).

As pointed out by Baker and Cornell [7], UHS pools all possible seismic scenarios (already described in Section 2.2.2.2) but it overlooks the period-dependent correlation between the spectral ordinates. In other words, although spectral acceleration ordinates in UHS individually have the same exceedance rates, they do not have to collectively possess the same exceedance rate. The flipping coin analogy can help understanding this concept better: although every coin-flip has the same probability (rate) of producing certain outcome, it does not mean that throwing multiple coins at once have the same probability of producing all those outcomes simultaneously. Consequently, GM records scaled to match the target UHS (e.g., 2475-year return period) may yield very conservative seismic demands for SSRA.

The reason behind this theoretical drawback is that UHS establishes the entire spectral demand from individual scalar hazard computations (i.e., hazard curves) at discrete periods [3]. This is mathematically (as well as physically) not the consistent way of developing target spectrum corresponding to the target hazard level. The consistent way of achieving target spectrum requires joint seismic hazard assessment that accounts for the correlation of

spectral ordinates for the entire period range [23]. This concept is called as *vector-based hazard assessment* and seismic risk assessment procedures that employ vector-based hazard assessment are the most sophisticated methods to quantify the risk [23]. These probabilistic risk assessment procedures are referred to as *vector-based risk assessment* following their hazard counterpart.

There are simpler methods to overcome the mentioned bias in UHS that are called as conditional spectrum methods [8]. Both conditional spectrum methods and vector-based IM methods will be presented and discussed in this thesis in the following sections and chapters.

2.4. Conditional Spectrum

Conditional spectrum (CS) is another type of target spectrum used for record selection and scaling. There are many publications in the literature describing the background theory and implementation of conditional spectrum. The concept roots back to Baker and Cornell's article in 2006 [7] and was firstly proposed by Baker in 2010 [8]. Back then, this concept was only introduced as *conditional mean spectrum* which represented hazard consistent spectrum but failed to include variability in spectrum ordinates. This concept was later improved to become just *conditional spectrum* which, besides mean, also includes variability in spectrum ordinates. Lin et al. [9, 10] describe the conditional spectrum methodology in a very clear and straightforward way. This section summarizes the conditional spectrum methodology. For more detailed discussions on CS, the reader is referred to Lin et al. [9, 10] and several other publications [11, 13].

As explained in the previous section, UHS overestimates the actual spectral demand as it disregards the correlation between the spectral ordinates (see the relevant discussions in Baker and Cornell 2006 [7] and Baker 2010 [8]). The conditional spectrum concept is deemed to reduce this shortcoming. Considering the correlation between the spectral ordinates, this concept computes conditional spectral ordinates conditioned on a specific period T (usually chosen as the fundamental period of the structure) and hence the corresponding spectral acceleration $SA(T)$. Note that $SA(T)$ is computed from conventional (scalar) PSHA for a target hazard level (e.g., 475-year return period or 2475-year return period).

Spectrum conditioned on single period $SA(T)$ value is designated as $CS_{SA(T_1)}(T)$ in this thesis. $CS_{SA(T_1)}(T)$ is the acceleration spectrum value at period T of the conditional spectrum conditioned on the spectral acceleration value at period T_1. The mathematical background of $CS_{SA(T_1)}(T)$ is explained in the following section.

Conditioning spectral ordinate (or spectral period) in CS does not have to be a single spectral acceleration at a given period. Lately, conditional spectrum conditioned on the average spectral value along a range of spectral periods has become appealing. Kohrangi et al. [11] introduce this new CS and define the average spectral acceleration as $AvgSA$ in their paper. A slightly different nomenclature is used for this new spectrum in the thesis. The next coming sections also discuss this terminology.

2.4.1. Conditional Spectrum Conditioned on Single Period Spectral Acceleration

This concept is explained by including the basic theoretical background of multivariate normal distribution.

2.4.1.1. Multivariate Normal Distribution as a Representation of Scenario Spectrum

GMPEs provide the distribution of $SA(T)$ at period T given a seismic scenario. This distribution is usually assumed as lognormal, and it is described by two parameters: logarithmic mean and standard deviation of $SA(T)$. That is, $\mu_{\ln(SA(T))}(rup, Vs30)$ and $\sigma_{\ln(SA(T))}(rup, Vs30)$, respectively. The parameter *rup* represents the seismic scenario (or rupture) that is defined by a specific set of magnitude, source-to-site distance, style-of-faulting, etc. Some GMPEs also provide the correlation coefficients between the logarithms of spectral ordinates at discrete periods: $\rho_{\ln(SA(T1)),\ln(SA(T2))}(rup, Vs30)$. ASB14 (the GMPE used in this thesis) provides $\sigma_{\ln(SA(T))}$ and $\rho_{\ln(SA(T1)),\ln(SA(T2))}$ that are valid for the magnitude, source-to-site distance, Vs30, style-of-faulting and period range covered by the model. Given a seismic scenario and site condition, ASB14 defines the conditional logarithmic acceleration spectrum as a multivariate normal distribution vector given in Eqn.

2.47 – 2.54. The arrow sign designates the vector variables and the symbol $\mathcal{N}$ denotes normal distribution in the below expressions and in the rest of the text.

$$\overrightarrow{lnSAT}(rup, Vs30) \sim \mathcal{N}(\vec{\mu}_{\mathrm{lnSAT|rup,Vs30}}, \Sigma_{\mathrm{lnSAT}}) \tag{2.47}$$

$$\overrightarrow{lnSAT}(rup, Vs30) = \begin{bmatrix} \ln(SA(T_1)) \,|rup, Vs30 \\ \ln(SA(T_2)) \,|rup, Vs30 \\ \vdots \\ \ln(SA(T_n))|rup, Vs30 \end{bmatrix} \tag{2.48}$$

$$\vec{\mu}_{\mathrm{lnSAT}|rup,Vs30} = \begin{bmatrix} \mu_{\ln\,(SA(T_1))|rup,Vs30} \\ \mu_{\ln\,(SA(T_2))|rup,Vs30} \\ \vdots \\ \mu_{\ln\,(SA(T_n))|rup,Vs30} \end{bmatrix} \tag{2.49}$$

$$\Sigma_{\mathrm{lnSAT}} = \begin{bmatrix} Var(\ln(SA(T_1))) & Cov(\ln(SA(T_1)), \ln(SA(T_2))) & \cdots \\ Cov(\ln(SA(T_2)), \ln(SA(T_1))) & Var(\ln(SA(T_2))) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{2.50}$$

$$\Sigma_{\mathrm{lnSAT}} = \begin{bmatrix} \left(\sigma_{\ln(SA(T_1))}\right)^2 & \rho_{\ln(SA(T1)),\ln(SA(T2))}\,\sigma_{\ln(SA(T_1))}\sigma_{\ln(SA(T_2))} & \cdots \\ \rho_{\ln(SA(T1)),\ln(SA(T2))}\,\sigma_{\ln(SA(T_1))}\sigma_{\ln(SA(T_2))} & \left(\sigma_{\ln(SA(T_2))}\right)^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \tag{2.51}$$

$$\Sigma_{\mathrm{lnSAT}} = \boldsymbol{\sigma}_{\mathrm{lnSAT}} \boldsymbol{\rho}_{\mathrm{lnSAT}} \boldsymbol{\sigma}_{\mathrm{lnSAT}} \tag{2.52}$$

$$\boldsymbol{\sigma}_{\mathrm{lnSAT}} = \begin{bmatrix} \sigma_{\ln(SA(T_1))} & 0 & \cdots & 0 \\ 0 & \sigma_{\ln(SA(T_2))} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\ln(SA(T_n))} \end{bmatrix} \tag{2.53}$$

$$\boldsymbol{\rho}_{\mathrm{lnSAT}} = \begin{bmatrix} 1 & \rho_{\ln(SA(T_1)),\ln(SA(T_2))} & \cdots & \rho_{\ln(SA(T_1)),\ln(SA(T_n))} \\ \rho_{\ln(SA(T_2)),\ln(SA(T_1))} & 1 & \cdots & \rho_{\ln(SA(T_2)),\ln(SA(T_n))} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{\ln(SA(T_n)),\ln(SA(T_1))} & \rho_{\ln(SA(T_n)),\ln(SA(T_2))} & \cdots & 1 \end{bmatrix} \tag{2.54}$$

ASB14 provides correlation coefficient tables of $\rho_{\ln(SA(T_i)),\ln(SA(T_k))}$ between discrete periods of T_i and T_k for the period range 0.01 s ≤ T ≤ 4.0 s that can be used in the above expressions. If the spectral period of interest is between the two consecutive periods, one can implement interpolation to compute the corresponding correlation coefficient value. The interpolation should be repeated for other spectral parameters of concern as those given in Eqn. 2.47 – 2.54.

In essence, after computing $\mu_{\ln(SA(T))|rup,Vs30}$, $\sigma_{\ln(SA(T))}$ and $\rho_{\ln(SA(T_i)),\ln(SA(T_k))}$ for a given rupture scenario and site condition, the conditional acceleration spectrum specific to that scenario and site condition can be represented by the multivariate normal distribution through Eqn. 2.47 – 2.54.

2.4.1.2. Dominant Seismic Scenario of a Specific SA(T) Hazard Level

The procedure to compute conditional target spectrum conditioned on the spectral acceleration at the conditioning (desired) period, $CS_{SA(T_{conditioning})}(T)$, commences by identifying $SA(T_{conditioning})$ from the hazard curve at target hazard level (e.g., 2475-year return period). $T_{conditioning}$ is usually chosen as the fundamental period of the structure being assessed (T_1). Therefore $SA(T_1)$ is firstly computed from the corresponding hazard curve given the target hazard level. Occurrence disaggregation data for this hazard level is used to detect the dominant seismic scenario that significantly contributes to the hazard for $SA(T_1)$ level. Figure 2.11 shows $SA(0.96s)$ hazard curve and its occurrence disaggregation for 475-year return period (target) hazard level.

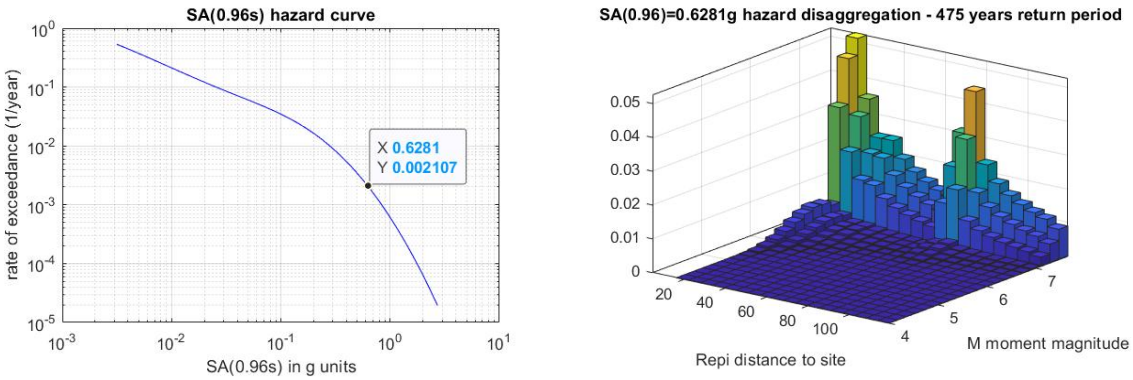

Figure 2.11: Left panel shows $SA(0.96s)$ hazard curve and right panel is the occurrence disaggregation for 475-year return period

As it can be seen from the disaggregation plot, multiple scenarios contribute to the 475-year return period hazard at $SA(0.96s)$. Pointing out a single one as the dominant scenario is not easy. Even though there are some methods to include multiple scenarios into conditional spectrum computation [13], most authors accept the mean scenario as the viable scenario representing the dominant single event. The mean scenario is evaluated by computing weighted means of the M and R from the disaggregation distribution as displayed in Figure 2.11 (M=7.33, R=54.08km).

After detecting the dominant scenario, the acceleration spectrum distribution can be obtained by employing the multivariate normal distribution formulations (Eqn. 2.47 – 2.54) through ASB14. Figure 2.12 displays the median spectrum (blue solid line) of the mean (dominant) scenario for 475-year return period, ±2 standard deviation about the median (blue dashed lines), and the 475-year return period $SA(0.96s)$ (red cross) determined from the corresponding hazard curve. The multivariate normal distribution includes the covariance values between different $SA(T)$ ordinates although they cannot be perceived adequately by the spectrum. As it can be seen from Figure 2.12, the 475-year return period $SA(0.96s)$ is above the median spectrum of the dominant scenario which indicates that there are other scenarios contributing to the 475-year hazard at $SA(0.96s)$. However, as discussed by Baker and Cornell 2006 [7] as well as the other published literature cited previously, these scenarios do not necessarily contribute to the spectral demands at other periods, which is the starting point of the conditional spectrum concept.

As it will be demonstrated in the next section, the conditional spectrum targets (conditions) the spectral demand at the 475-year $SA(0.96s)$ but computes the rest of the spectral ordinates conditioned on the target 475-year $SA(0.96s)$. The mathematical method for developing the conditional spectrum is explained in the following section.

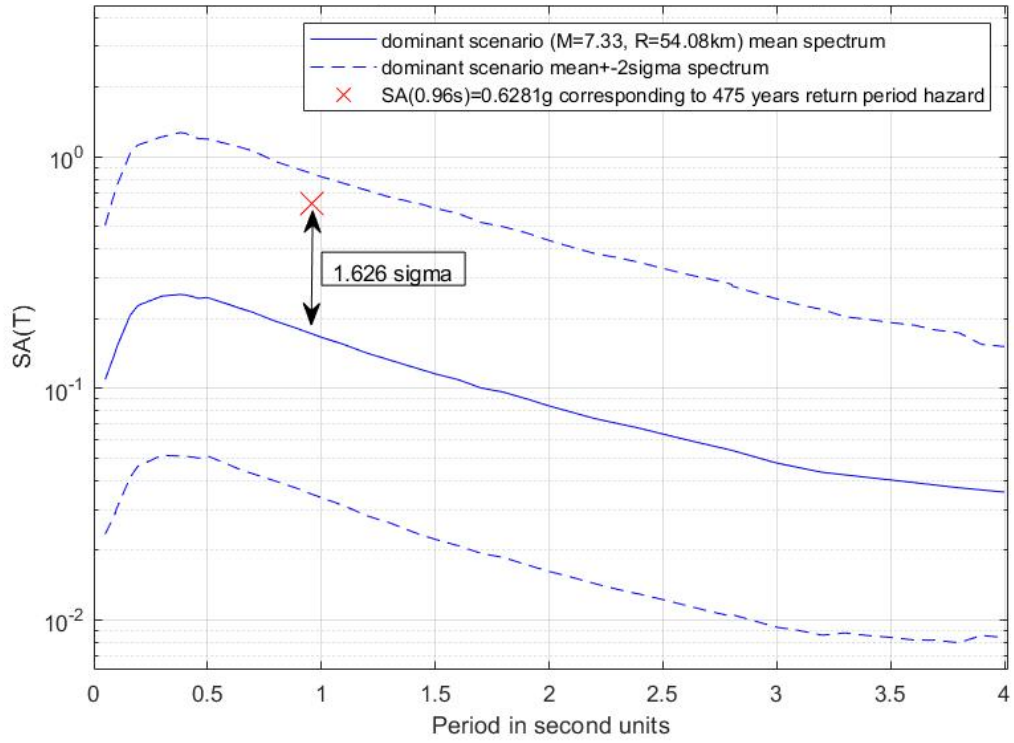

Figure 2.12: 475 years hazard level $SA(T_1)$ and dominant scenario spectrum distribution

2.4.1.3. Conditioning Dominant Scenario Spectrum on $SA(T_1)$

As mentioned in previous sections, ASB14 can be used to represent the dominant scenario spectrum as a multivariate normal distribution:

$$\overrightarrow{lnSAT}(rup-dominant, Vs30) \sim \mathcal{N}(\vec{\mu}_{\ln SAT|rup-dominant,Vs30}, \Sigma_{\ln SAT}) \qquad (2.55)$$

To compute the conditional spectrum, this scenario spectrum needs to be conditioned on the target spectral ordinate at 475-year return period $SA(0.96s)$.

Conditional multivariate normal distributions are explained and derived in many probability textbooks [24]. The mathematical formulations for conditional multivariate normal distributions are as follows:

$$\vec{Z} \sim \mathcal{N}(\vec{\mu}_Z, \Sigma_{\boldsymbol{Z}}) \tag{2.56}$$

$$\vec{Z} = \begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix} \tag{2.57}$$

$$\vec{\mu}_Z = \begin{bmatrix} \vec{\mu}_X \\ \vec{\mu}_Y \end{bmatrix} \tag{2.58}$$

$$\Sigma_{\boldsymbol{Z}} = \begin{bmatrix} \Sigma_{\boldsymbol{XX}} & \Sigma_{\boldsymbol{XY}} \\ \Sigma_{\boldsymbol{YX}} & \Sigma_{\boldsymbol{YY}} \end{bmatrix} \tag{2.59}$$

$$\vec{\mu}_{X|Y=y} = \vec{\mu}_X + \Sigma_{\boldsymbol{XY}} {\Sigma_{\boldsymbol{YY}}}^{-1} (\vec{y} - \vec{\mu}_Y) \tag{2.60}$$

$$\Sigma_{\boldsymbol{XX|Y=y}} = \Sigma_{\boldsymbol{XX}} - \Sigma_{\boldsymbol{XY}} {\Sigma_{\boldsymbol{YY}}}^{-1} \Sigma_{\boldsymbol{YX}} \tag{2.61}$$

$$\vec{\mu}_{Z|Y=y} = \begin{bmatrix} \vec{\mu}_{X|Y=y} \\ \vec{y} \end{bmatrix} \tag{2.62}$$

$$\Sigma_{\boldsymbol{Z|Y=y}} = \begin{bmatrix} \Sigma_{\boldsymbol{XX|Y=y}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{2.63}$$

$$\vec{Z}_{|Y=y} \sim \mathcal{N}\left(\vec{\mu}_{Z|Y=y}, \Sigma_{\boldsymbol{Z|Y=y}}\right) \tag{2.64}$$

The above formulations (Eqn. 2.56 – 2.64) demonstrate how to condition normally distributed random vector $\vec{Z}$ on some of its components (vector $\vec{Y}$). We first rearrange $\vec{Z}$ such that the conditioning components ($\vec{Y}$) and conditioned ones ($\vec{X}$) are partitioned (Eqn. 2.57 – 2.58). The covariance matrix is also rearranged to follow this partitioning format (Eqn. 2.59).

The rest of the expressions (Eqn. 2.60 – 2.64) provide the necessary formulations to develop the conditional distribution of $\vec{Z}$ conditioned on $\vec{Y}$.

For the specific case at hand, $\vec{Z}$ is analogous to $\overrightarrow{lnSAT}(rup - dominant, Vs30)$, $\vec{Y}$ is analogous to single component vector $\left[ln\left(SA(T_1)\right)\right]$ (scalar), and $\vec{X}$ is analogous to the vector consisting of the rest of $ln\left(SA(T)\right)$ members of $\vec{Z}$. After performing the steps described in Eqn. 2.56 – 2.64 on the dominant (mean) scenario spectrum $\overrightarrow{lnSAT}(rup - dominant, Vs30)$, the conditional spectrum distribution can be acquired: $\overrightarrow{lnSAT}(rup - dominant, Vs30)|_{SA(T_1)=sa(T_1)}$ or simply $CS_{SA(T_1)}(T)$.

Figure 2.13 displays $CS_{SA(T_1)}(T)$ spectrum distribution corresponding to 475-year return period hazard level for the case study under consideration. The corresponding UHS is also displayed in the figure.

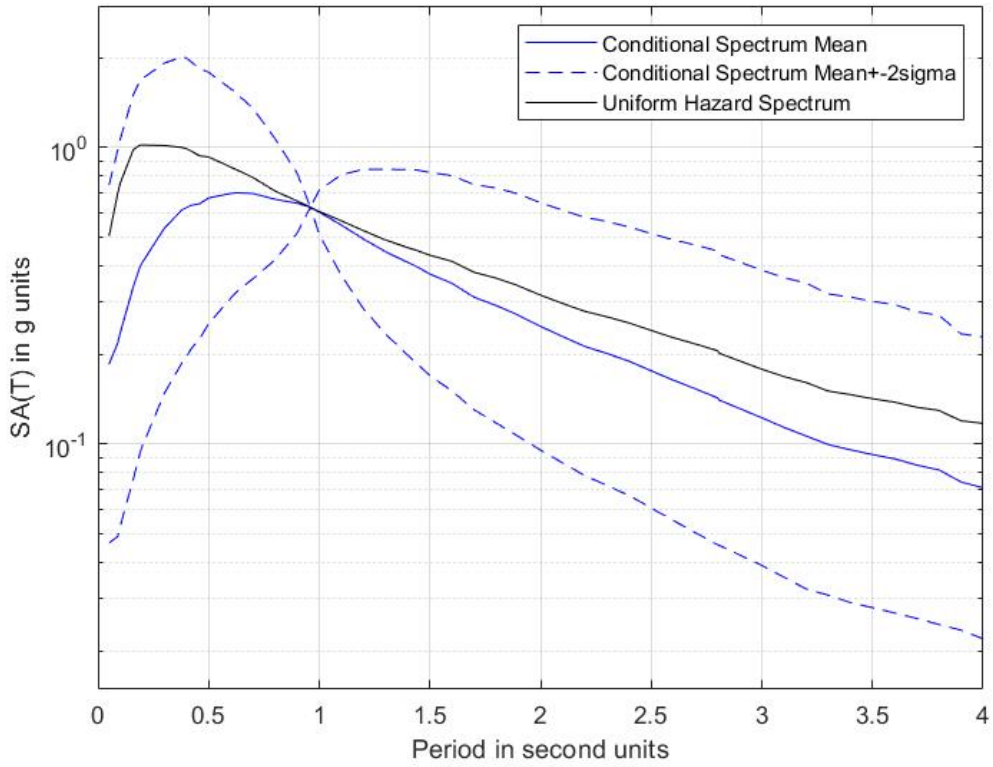

Figure 2.13: 475-year return period $\mathrm{CS}_{\mathrm{SA}(\mathrm{T}_1)}(\mathrm{T})$ spectrum and the corresponding UHS.

As can be seen in Figure 2.13, the mean of the $CS_{SA(T_1)}(T)$ spectrum, namely Conditional Mean Spectrum $CMS_{SA(T_1)}(T)$ and UHS give the same target spectral acceleration value result at $T = 0.96s$. However, as stated previously, for other periods UHS produces conservative $SA(T)$ levels. This conservative spectral demand estimation is reduced by $CS_{SA(T_1)}(T)$. The ±2 sigma spectra (blue dashed lines) are used to represent the likely aleatory variability of spectral acceleration demand about $CMS_{SA(T_1)}(T)$.

2.4.2. Spectrum Conditioned on Average of a Range of Spectral Accelerations

$CS_{SA(T_1)}(T)$ spectrum is a rational alternative as a target spectrum for hazard consistent record selection and scaling because it represents the spectral acceleration demands more realistically by considering the correlation between the conditioning spectral acceleration and the other spectral accelerations. The conditioning spectral acceleration ordinate is a scalar quantity (i.e., $SA(T_1)$) and generally corresponds to the fundamental building period. This choice seems to be rational as well because most of the times the scalar $SA(T_1)$ has a considerable effect on the structural response and thus dominates the structural behavior. Inherently, the conditional spectrum (as depicted in Figure 2.13) has a zero aleatory variability at T_1 because $SA(T_1)$ is the target hazard dictated by the target return period. The zero ground-motion variability at T_1 would influence structural response if it corresponds to the fundamental period of the structure. This could lead to some level of bias in SSRA if ground-motion scaling is based on $CS_{SA(T_1)}(T)$.

Another scalar IM that has become popular in the literature for SSRA is the so-called *average spectral acceleration* (*AvgSA*), as demonstrated by Kohrangi et al. [11]. The mathematical development of *AvgSA* and the derivation of corresponding conditional spectrum are given in following sections.

2.4.2.1. Average Spectral Acceleration as an Intensity Measure

The dynamic response of structural systems is best represented by considering multiple modes contributing the overall response. In many cases, not only the elastic fundamental mode, but also the shift in the fundamental mode due to post-elastic structural response or

higher modes other than the fundamental mode become prominent. For example, Kohrangi et al. propose considering the spectral period range between $0.2T_1 - 1.5T_1$ for more accurate structural response [11].

The starting point of *AvgSA* is to take such a period range, split it into closely spaced T increments and take the geometrical mean of $SA(T)$ at these T increments (Eqn. 2.65 – 2.66).

$$AvgSA(\vec{T}) = \sqrt[n]{\prod_{i=1}^{i=n} SA(T_i)} \tag{2.65}$$

$$\vec{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} \tag{2.66}$$

$AvgSA(\vec{T})$ in Eqn. 2.65-2.66 represents average spectral acceleration at periods given in vector $\vec{T}$. $\vec{T}$ is usually chosen to cover the structure sensitive period range at closely spaced steps. Therefore, $AvgSA$ represents the general level of acceleration spectrum within which the structural behavior is controlled. This fact makes it as a useful predictor IM in a general sense because, as discussed before, structural behavior is generally dominated by multiple modes represented by a range of $SA(T)$ ordinates.

The reason for defining $AvgSA$ as the geometrical mean of the $SA(T)$ ordinates is the fact that it corresponds to the arithmetic mean in log space as explained in Eqn. 2.67. The use of the transition between non-logarithmic and logarithmic spectral parameters are demonstrated in the following discussions.

$$AvgSA(\vec{T}) = \sqrt[n]{\prod_{i=1}^{i=n} SA(T_i)} \tag{2.67}$$

$$\ln\left(AvgSA(\vec{T})\right) = \ln\left(\sqrt[n]{\prod_{i=1}^{i=n} SA(T_i)}\right) \tag{2.68}$$

$$\ln\left(AvgSA(\vec{T})\right) = \frac{1}{n}\sum_{i=1}^{i=n} \ln(SA(T_i)) \tag{2.69}$$

There are few studies demonstrating the better performance of scalar $AvgSA(\vec{T})$ as the predictor of structural response when compared to $SA(T_1)$ [11]. These studies advocate that $AvgSA(\vec{T})$ captures the ground-motion intensity at several periods that dominate the structural response.

This thesis introduces an improvement to the definition of $AvgSA(\vec{T})$ that is given in the following expressions. Although structural response is sensitive to a wide range of $SA(T)$ ordinates, it does not indicate that all $SA(T)$ values within this range affect the structural response equally. Since the fundamental mode contribution to the structural response is generally higher than the higher mode contributions, it is clear that $SA(T_1)$ will most likely have a stronger impact on the overall response rather than, for example, $SA(T_2)$ or $SA(T_3)$. The idea behind the proposed improvement to $AvgSA(\vec{T})$ is to take the weighted average of $\ln(SA(T_i))$ ordinates rather than their arithmetic mean as shown previously.

$$\ln\left(AvgSA(\vec{T},\vec{W})\right) = \frac{1}{\sum_{i=1}^{i=n} W_i}\sum_{i=1}^{i=n} W_i \ln(SA(T_i)) \tag{2.70}$$

$$AvgSA\left(\vec{T},\vec{W}\right) = \prod_{i=1}^{i=n} (SA(T_i))^{\frac{W_i}{\sum_{i=1}^{i=n} W_i}} \tag{2.71}$$

$$\vec{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} \tag{2.72}$$

$$\overrightarrow{W} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} \tag{2.73}$$

$AvgSA(\vec{T}, \overrightarrow{W})$ in the above formulation represents weighted average of $SA(T_i)$ spectrum ordinates. $\overrightarrow{W}$ is the vector containing weights of $SA(T_i)$ contributions to $AvgSA(\vec{T}, \overrightarrow{W})$. Using this formulation $\overrightarrow{W}$ can be set such that it yields higher weight to the spectral ordinate at the fundamental mode. (The decision strategy for determining $\vec{T}$ and $\overrightarrow{W}$ for the structure used in the case study is discussed in following sections).

If $\overrightarrow{W}$ is normalized such that the sum of all of its components is 1, Eqn. 2.70 – 2.73 can be simplified in the following manner:

$$\overrightarrow{W'} = \overrightarrow{W} \frac{1}{\sum_{i=1}^{i=n} W_i} \tag{2.74}$$

$$\ln\left(AvgSA\left(\vec{T}, \overrightarrow{W'}\right)\right) = \sum_{i=1}^{i=n} W'_i \ln(SA(T_i)) \tag{2.75}$$

$$AvgSA\left(\vec{T}, \overrightarrow{W'}\right) = \prod_{i=1}^{i=n} (SA(T_i))^{W'_i} \tag{2.76}$$

2.4.2.2. Computation of Average Spectral Acceleration Hazard Curve

$AvgSA\left(\vec{T}, \overrightarrow{W'}\right)$ is a scalar IM just like $SA(T_1)$ and, conceptually, its hazard curve can be computed as presented in Eqn. 2.43 that is repeated as Eqn. 2.77 for convenience:

$$\lambda_{IM\geq im} = \sum_{i=1}^{M\ bins}\left(\sum_{k=1}^{R\ bins}\lambda_{m_i<M\leq m_{i+1},r_k<R\leq r_{k+1}}G_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)\right) \quad (2.77)$$

The difference between $SA(T_1)$ and $AvgSA\left(\vec{T},\overrightarrow{W'}\right)$ hazard curves lies in $G_{IM|M=\frac{m_i+m_{i+1}}{2},R=\frac{r_k+r_{k+1}}{2}}(im)$ term. This term is expressed as given in the below that is also discussed in Sections 2.2.2.1 and 2.2.2.2.

$$G_{IM|Rup}(im) = 1-\Phi\left(\frac{ln(im)-\mu_{ln(IM)|Rup}}{\sigma_{ln(IM)|Rup}}\right) \quad (2.78)$$

$$Rup\sim M = \frac{m_i+m_{i+1}}{2}, R = \frac{r_k+r_{k+1}}{2} \quad (2.79)$$

Terms $\mu_{ln(IM)|Rup}$ and $\sigma_{ln(IM)|Rup}$ can be directly computed from ASB14 in $SA(T_1)$,. However, these terms become $\mu_{\ln(AvgSA)|Rup}$ and $\sigma_{\ln(AvgSA)|Rup}$ for $AvgSA\left(\vec{T},\overrightarrow{W'}\right)$ whose calculation requires the linear transformation of multivariable normal distribution.

As discussed before and demonstrated in Eqn. 2.47 – 2.54, the seismic scenario spectrum can be represented as multivariate normal distribution $\overrightarrow{lnSAT}(rup,Vs30)$. This distribution has mean vector $\vec{\mu}_{\mathrm{lnSAT|rup,Vs30}}$ and covariance matrix Σ_{lnSAT}. Considering Eqn. 2.75, one can represent $\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)$ as the linear transformation of $\overrightarrow{lnSAT}(rup,Vs30)$.

$$\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)|_{rup,Vs30} = \boldsymbol{W''^{T}}\boldsymbol{lnSAT}(rup,Vs30) \quad (2.80)$$

In the Eqn. 2.80, the terms on the right represent matrix multiplication of two vectors $\overrightarrow{W''}$ and $\overrightarrow{lnSAT}(rup,Vs30)$. Vector $\overrightarrow{W''}$ is of the same size as $\overrightarrow{lnSAT}(rup,Vs30)$ and contains the same components as $\overrightarrow{W'}$ for the periods in $\vec{T}$ that consider the weighted average spectral acceleration $AvgSA\left(\vec{T},\overrightarrow{W'}\right)$ and zero at all other periods of $\overrightarrow{lnSAT}(rup,Vs30)$ spectrum.

Since $\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)|_{rup,Vs30}$ is a linear transformation of normally distributed vector $\overrightarrow{lnSAT}(rup,Vs30)$, it is also normally distributed variable with mean and sigma evaluated by the following relationships.

Let's say $\vec{X}$ is a normally distributed vector:

$$\vec{X}\sim \mathcal{N}(\vec{\mu}_X,\Sigma_{XX}) \tag{2.81}$$

Take $\vec{Z}$ as a normal transformation of $\vec{X}$;

$$\boldsymbol{Z}=\boldsymbol{AX} \tag{2.82}$$

$\boldsymbol{A}$ is a transformation matrix with number of columns equals to the size of the vector $\vec{X}$ and number of rows equals to the size of vector $\vec{Z}$.

$$\boldsymbol{\mu}_Z=E(\boldsymbol{Z})=E(\boldsymbol{AX})=\boldsymbol{A}E(\boldsymbol{X})=\boldsymbol{A\mu}_X \tag{2.83}$$

$$\begin{aligned}\Sigma_{ZZ}=var(\boldsymbol{Z})&=E\big((\boldsymbol{Z}-\boldsymbol{\mu}_Z)(\boldsymbol{Z}-\boldsymbol{\mu}_Z)^T\big)=E\big((\boldsymbol{Z}-\boldsymbol{\mu}_Z)(\boldsymbol{Z}-\boldsymbol{\mu}_Z)^T\big)\\&=E\big((\boldsymbol{AX}-\boldsymbol{A\mu}_X)(\boldsymbol{AX}-\boldsymbol{A\mu}_X)^T\big)\\&=E\Big((\boldsymbol{AX}-\boldsymbol{A\mu}_X)(\boldsymbol{X^TA^T}-\boldsymbol{\mu}_X{}^T\boldsymbol{A^T})\Big)\\&=E(\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu}_X)(\boldsymbol{X^T}-\boldsymbol{\mu}_X{}^T)\boldsymbol{A^T})\\&=\boldsymbol{A}E\Big((\boldsymbol{X}-\boldsymbol{\mu}_X)(\boldsymbol{X^T}-\boldsymbol{\mu}_X{}^T)\Big)\boldsymbol{A^T}=\boldsymbol{A}\Sigma_{XX}\boldsymbol{A^T}\end{aligned} \tag{2.84}$$

Using the above relationships (Eqn. 2.81 – 2.84) for $\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)|_{rup,Vs30}$, where $\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)|_{rup,Vs30}$ is analogous to $\boldsymbol{Z}$, $\boldsymbol{lnSAT}(rup,Vs30)$ is analogous to $\boldsymbol{X}$, and $\boldsymbol{W''^T}$ is analogous to $\boldsymbol{A}$, the following relations can be derived:

$$\mu_{\ln\,(AvgSA)|Rup} = \boldsymbol{W''}^{T}\boldsymbol{\mu}_{\mathrm{lnSAT|rup,Vs30}} \tag{2.85}$$

$$var\left(\ln\left(AvgSA\left(\vec{T},\overrightarrow{W'}\right)\right)_{|rup,Vs30}\right) = \boldsymbol{W''}^{T}\Sigma_{\mathrm{lnSAT}}\boldsymbol{W''} \tag{2.86}$$

$$\sigma_{\ln\,(AvgSA)|Rup} = \sqrt{\boldsymbol{W''}^{T}\Sigma_{\mathrm{lnSAT}}\boldsymbol{W''}} \tag{2.87}$$

Since now $\mu_{\ln\,(AvgSA)|Rup}$ and $\sigma_{\ln\,(AvgSA)|Rup}$ can be computed from Eqn. 2.85 – 2.87, Eqn. 2.77 – 2.78 can be used to compute the hazard curve of average spectral acceleration. Eqn. 2.44 and Eqn. 2.46 can be used to evaluate its occurrence disaggregation.

2.4.2.3. Computation of Average Spectral Acceleration Conditional Spectra

Since the average spectral acceleration is defined as a scalar IM, and its hazard curve as well as disaggregation can be computed, it can be used as a conditioning spectral ordinate for the development of hazard consistent $CS_{AvgSA}(T)$.

The methodology for development of $CS_{AvgSA}(T)$ is very similar to $CS_{SA(T_1)}(T)$. The target $AvgSA$ level is firstly located from its hazard curve. Disaggregation data corresponding to the occurrence of target $AvgSA$ is used to detect the dominant scenario. The dominant scenario spectrum distribution is developed from $\vec{\mu}_{\mathrm{lnSAT|rup-dominant,Vs30}}$ and covariance matrix Σ_{lnSAT}.

The last step is to condition the dominant scenario spectrum distribution on the target $AvgSA$ level. However, this last step is a little more complex than it is in $CS_{SA(T_1)}(T)$ because $\ln(AvgSA)$ is not a component of $\overrightarrow{lnSAT}$ but rather its linear transformation. So, computing $CS_{AvgSA}(T)$ requires conditioning multivariate normal distribution on its linear transformation. The fundamentals for performing this calculation is given below (Eqn. 2.88 – 2.97):

$$\vec{X} \sim \mathcal{N}(\vec{\mu}_X, \Sigma_{XX}) \tag{2.88}$$

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} \tag{2.89}$$

$$\boldsymbol{\mu}_Y = \boldsymbol{A}\boldsymbol{\mu}_X \tag{2.90}$$

$$\Sigma_{YY} = \boldsymbol{A}\Sigma_{XX}\boldsymbol{A}^T \tag{2.91}$$

$$\begin{aligned} Cov(\boldsymbol{X}, \boldsymbol{Y}) = \Sigma_{XY} &= E(\boldsymbol{X}\boldsymbol{Y}^T) - E(\boldsymbol{X})E(\boldsymbol{Y}^T) = E(\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{A}^T) - \boldsymbol{\mu}_X{\boldsymbol{\mu}_X}^T\boldsymbol{A}^T \\ &= E(\boldsymbol{X}\boldsymbol{X}^T)\boldsymbol{A}^T - \boldsymbol{\mu}_X{\boldsymbol{\mu}_X}^T\boldsymbol{A}^T = (E(\boldsymbol{X}\boldsymbol{X}^T) - \boldsymbol{\mu}_X{\boldsymbol{\mu}_X}^T)\boldsymbol{A}^T = \Sigma_{XX}\boldsymbol{A}^T \end{aligned} \tag{2.92}$$

$$\vec{Z} = \begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix} \tag{2.93}$$

$$\boldsymbol{\mu}_Z = \begin{bmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{A}\boldsymbol{\mu}_X \end{bmatrix} \tag{2.94}$$

$$\Sigma_{ZZ} = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix} = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XX}\boldsymbol{A}^T \\ \boldsymbol{A}\Sigma_{XX} & \boldsymbol{A}\Sigma_{XX}\boldsymbol{A}^T \end{bmatrix} \tag{2.95}$$

$$\boldsymbol{\mu}_{X|Y=y} = \boldsymbol{\mu}_X + \Sigma_{XX}\boldsymbol{A}^T(\boldsymbol{A}\Sigma_{XX}\boldsymbol{A}^T)^{-1}(y - \boldsymbol{A}\boldsymbol{\mu}_X) \tag{2.96}$$

$$\Sigma_{XX|Y=y} = \Sigma_{XX} - \Sigma_{XX}\boldsymbol{A}^T(\boldsymbol{A}\Sigma_{XX}\boldsymbol{A}^T)^{-1}\boldsymbol{A}\Sigma_{XX} \tag{2.97}$$

Using the formulation given above (Eqn. 2.88 – 2.97), $CS_{AvgSA}(T)$ can be evaluated by following formulation:

$$CS_{AvgSA}(T) \approx \overrightarrow{lnSAT}(rup - dominant, Vs30)_{|AvgSA(\vec{T},\overrightarrow{Wi})=avgsa} \tag{2.98}$$

$$\begin{aligned} &\overrightarrow{lnSAT}(rup - dominant, Vs30)_{|AvgSA(\vec{T},\overrightarrow{Wi})=avgsa} \sim \\ &\sim \mathcal{N}\left(\vec{\mu}_{\mathrm{lnSAT}|AvgSA(\vec{T},\overrightarrow{Wi})=avgsa}, \Sigma_{\mathrm{lnSAT}|AvgSA(\vec{T},\overrightarrow{Wi})=avgsa}\right) \end{aligned} \tag{2.99}$$

$$\begin{aligned}\boldsymbol{\mu}_{\ln SAT|AvgSA(\vec{T},\overrightarrow{W'})=avgsa} &= \boldsymbol{\mu}_{\ln SAT|rup-dominant} \\ &+ \Sigma_{\ln SAT}\boldsymbol{W'}(\boldsymbol{W'}^{T}\Sigma_{\ln SAT}\boldsymbol{W'})^{-1}(avgsa \\ &- \boldsymbol{W'}^{T}\boldsymbol{\mu}_{\ln SAT|rup-dominant})\end{aligned} \tag{2.100}$$

$$\Sigma_{\ln SAT|AvgSA(\vec{T},\overrightarrow{W'})=avgsa} = \Sigma_{\ln SAT} - \Sigma_{\ln SAT}\boldsymbol{W'}(\boldsymbol{W'}^{T}\Sigma_{\ln SAT}\boldsymbol{W'})^{-1}\boldsymbol{W'}^{T}\Sigma_{\ln SAT} \tag{2.101}$$

This thesis uses two different $AvgSA$ intensity measures with two different $\vec{T}$ and $\overrightarrow{W'}$ vectors. $\vec{T}$ and $\overrightarrow{W'}$ are given below. The reason for the selection of these alternative $\vec{T}$ and $\overrightarrow{W'}$ vectors are explained in the next chapter when the model building's specific properties are explained in detail.

$$AvgSA1 = AvgSA\left(\overrightarrow{T_1}, \overrightarrow{W'_1}\right) \tag{2.102}$$

$$\overrightarrow{T_1} = \begin{bmatrix} 0.05s \\ 0.096s \\ 0.42s \\ 0.63s \\ 0.96s \\ 1.44s \end{bmatrix} \tag{2.103}$$

$$\overrightarrow{W'_1} = \begin{bmatrix} 0.026 \\ 0.0992 \\ 0.088 \\ 0.088 \\ 0.3494 \\ 0.3494 \end{bmatrix} \tag{2.104}$$

$$AvgSA2 = AvgSA\left(\overrightarrow{T_2}, \overrightarrow{W'_2}\right) \tag{2.105}$$

$$\overrightarrow{T_2} = \begin{bmatrix} 0.96s \\ 1.44s \end{bmatrix} \tag{2.106}$$

$$\overrightarrow{W'_2} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \tag{2.107}$$

Figure 2.14 shows the 475-year return period $CS_{AvgSA1}(T)$ and $CS_{AvgSA2}(T)$ to perceive their difference.

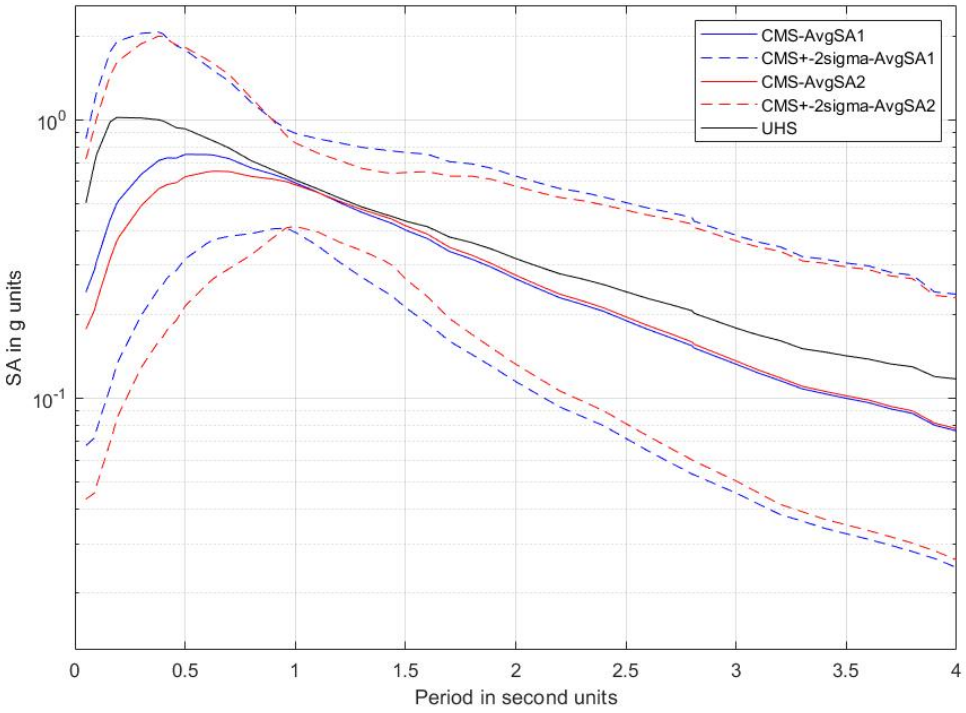

Figure 2.14: Alternative average spectral acceleration conditional spectra for 475-year return period hazard level and the corresponding uniform hazard spectrum

As it is depicted in Figure 2.14, the mean $CS_{AvgSA}(T)$ spectra (both alternatives) are lower than the corresponding UHS at all periods. However, at the designated periods where the contributions to the $AvgSA$ are expected to be significant, the two mean $CS_{AvgSA}(T)$ spectra as well as UHS attain closer values. Note that the standard deviations of $CS_{AvgSA}(T)$ are the lowest within the period range where the contributions to the $AvgSA$ are expected to be significant. $CS_{AvgSA2}(T)$ shows lower standard deviations than $CS_{AvgSA1}(T)$ in the period range between 1 s and 1.5 s. This is because at those periods (T=0.96s and T=1.44s) $AvgSA2$ impose higher weights than $AvgSA1$. $AvgSA1$ yields higher mean and lower standard deviation in shorter period range (< 0.9s) because of the selected periods (T=0.05s, 0.096s,

0.42s and 0.63s) fall within this short period range associated with higher contributions to $AvgSA1$. If $AvgSA$ is constructed to include only a single period T' with a weight of 1, then it would become $CS_{SA(T')}(T)$. Therefore, it turns out that $CS_{SA(T_1)}(T)$ is just a special case of $CS_{AvgSA(\vec{T},\overrightarrow{W'})}(T)$ where $\vec{T} = [T_1]$ and $\overrightarrow{W'} = [1]$.

Note that $AvgSA$ in literature [11] considers higher number of periods covering a wide spectral range but the resulting conditional spectrum will have high standard deviations within this range inflating the aleatory variability in the scaled ground motions. Consequently, it is better to give higher weights to the periods that contribute the most to the structural response because the resulting conditional spectrum will be associated with lower standard deviations that would produce lower dispersion in structural response statistics. This will lead to a reduction of uncertainty in the risk assessment.

2.5. Vector IM Seismic Hazard Assessment

Previous sections of this chapter demonstrate scalar IM seismic hazard assessment procedure as well as methodology to construct hazard consistent target spectra for record selection using results of scalar IM seismic hazard assessment (Conditional Spectra). Conditional spectra methodology represents statistically sound and coherent way to include multiple IMs (*SA(T)*s) that contribute to structural response by conducting only single scalar IM hazard assessment. However, there is a superior method to do this. It is called Vector IM Seismic Hazard Assessment and it produces results such as joint rate of occurrence function of multiple scalar IMs. These multiple scalar IMs represent IM vector ($\overrightarrow{IM}$).

Goal of vector IM seismic hazard assessment is to evaluate mean rate density function of $\overrightarrow{IM}$ - $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$. This function can be used to acquire joint rate of occurrence of all IM considered in $\overrightarrow{IM}$ within certain range - $\overrightarrow{im}_{low}$ and $\overrightarrow{im}_{high}$

$$\lambda_{\overrightarrow{im}_{low}<\overrightarrow{IM}<\overrightarrow{im}_{high}} = \int_{\overrightarrow{im}_{low}}^{\overrightarrow{im}_{high}} MRD_{\overrightarrow{IM}}(\overrightarrow{im})d\overrightarrow{im} = \int_{im_{low1}}^{im_{high1}} \int_{im_{low2}}^{im_{high2}} \cdots MRD_{\overrightarrow{IM}}(im_1, im_2 \dots) \dots dim_2 dim_1 \quad (2.108)$$

$MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ can be evaluated by following equation which is an analogous vector version of Eqn. 2.44:

$$MRD_{\overrightarrow{IM}}(\overrightarrow{im}) = \sum_{i=1}^{M\ bins} \left(\sum_{k=1}^{R\ bins} \lambda_{m_i<M\leq m_{i+1}, r_k<R\leq r_{k+1}} f_{\overrightarrow{IM}|M=\frac{m_i+m_{i+1}}{2}, R=\frac{r_k+r_{k+1}}{2}}(\overrightarrow{im}) \right) \quad (2.109)$$

Eqn. 2.44 and Eqn. 2.109 have almost identical right-hand side. Only different term in Eqn. 2.109 is $f_{\overrightarrow{IM}|M=\frac{m_i+m_{i+1}}{2}, R=\frac{r_k+r_{k+1}}{2}}(\overrightarrow{im})$ which represents multivariate PDF function of $\overrightarrow{IM}$ vector given that certain seismic scenario has occurred. $\overrightarrow{IM}|M = \frac{m_i+m_{i+1}}{2}, R = \frac{r_k+r_{k+1}}{2}$ usually takes the form of multivariate normal distribution. Statistical form of multivariate normal distribution PDF functions is given below (Eqn. 2.111):

$$\vec{X} \sim \mathcal{N}(\vec{\mu}_X, \Sigma_{XX}) \quad (2.110)$$

$$f_{\vec{X}}(\vec{x}) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_X})^T {\Sigma_{XX}}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_X})\right)}{\sqrt{(2\pi)^k|\Sigma_{XX}|}} \quad (2.111)$$

k in Eqn. 2.111 is the size of vector $\vec{X}$. So, in order to be able to evaluate $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$, one needs GMPE that provide $\vec{\mu}_{IM|rup,Vs30}$ and $\Sigma_{IM|rup,Vs30}$ for any seismic scenario and site conditions (rup and $Vs30$). In case $\overrightarrow{IM}$ is a vector composed of natural logarithms of $SA(T)$ spectral accelerations at multiple periods, ASB14 can be used to evaluate $\vec{\mu}_{\ln(SA(T))|rup,Vs30}$ and $\Sigma_{\ln(SA(T))}$ (discussed in Section 2.4.1.1).

This thesis evaluates $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ for $\overrightarrow{IM}$ composed of only two components: $\ln(SA(0.96s))$ and $\ln(SA(1.44s))$. In this case $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ represents a surface. It is displayed in Figure 2.15.

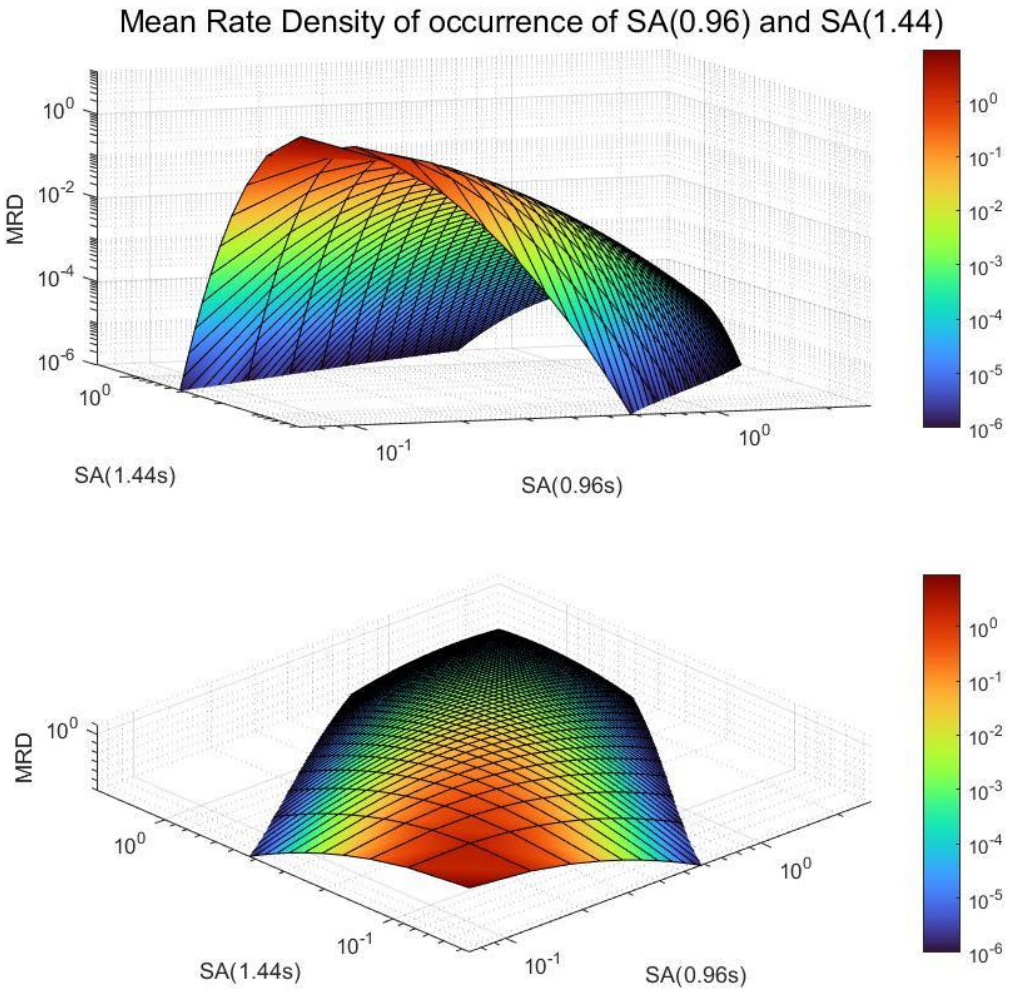

Figure 2.15: Mean Rate Density surface of SA(0.96s) and SA(1.44s) shown from two different viewing angles

Joint rate of occurrence of a range of $SA(0.96s)$ and $SA(1.44s)$ values can be obtained by evaluating the area below this surface corresponding to desired range (Eqn. 2.108). Figure 2.15 shows that mean rate density for low IM levels is high and drops for higher IM levels, which is as expected. It is also clear that mean rate density surface forms a diagonal of high

values and then drops abruptly for off-diagonal $\overrightarrow{IM}$ configurations. This is because events that produce high $SA(0.96s)$ and low $SA(1.44s)$ values and vice versa are extremely rare. Although it cannot be displayed in a single figure, $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ can be evaluated for $\overrightarrow{IM}$ of any size. Most sophisticated seismic hazard assessment method would be to evaluate $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ of vector $\overrightarrow{IM}$ composed of whole spectrum of $SA(T)$s at densely spaced periods T. However, this would entail a huge computational effort since $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ needs to be evaluated for all densely spaced $\overrightarrow{\imath m}$ configurations, and number of these configurations grows exponentially with the size of $\overrightarrow{\imath m}$. For example, in order to evaluate two-component $\overrightarrow{IM}$ mean rate density function as displayed in Figure 2.15, $SA(0.96s)$ and $SA(1.44s)$ axes were discretized into 50 values each, so $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ was computed at 2500 points. Each of these point $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ computations entails aggregation of 10000 seismic scenario (100 M and 100 R values) contributions (Eqn. 2.109). Each of these seismic scenario contributions entails evaluation of $\vec{\mu}_{\ln(SA(T))|rup,Vs30}$ by ASB14 ($\Sigma_{\ln(SA(T))}$ is constant for all scenarios) and $f_{\overrightarrow{IM}|rup,Vs30}(\overrightarrow{\imath m})$ by Eqn. 2.111. As such, this algorithm for evaluation of $MRD_{\overrightarrow{IM}}(\overrightarrow{\imath m})$ (where IM was two component vector of $SA(0.96s)$ and $SA(1.44s)$) was encoded into MATLAB and ran for couple of hours before completion. Of course, some optimization can be achieved by eliminating extreme "off-diagonal" $\overrightarrow{\imath m}$ configurations and accepting their mean rate density as zero. This would significantly improve computational efficiency for $\overrightarrow{IM}$ vectors of bigger size.

3. SEISMIC STRUCTURAL RESPONSE ASSESSMENT (SSRA)

The previous chapter introduced the background theory of PSHA, scalar and vector PSHA, and important target spectrum types, which use scalar and vector IMs for scaling ground motions that are used in SSRA. This chapter presents the theoretical aspects of seismic structural response assessment and later demonstrates the implementation of scalar and vector IM methods for probabilistic assessment of structural behavior under seismic excitation. As it is already stated a 6-story RC wall model building is used [15] for the implementation of the methods whose structural properties are explained in following section.

3.1. Structural Properties of 6-story RC Concrete Wall (Tunnel Form) Model Building

Figure 3.1 displays the 3D view of the model building and Figure 3.2 shows a typical floor plan from that model. The structural properties of the model building resemble the construction practice of tunnel form buildings in Istanbul (see Özen's thesis for details [15]). It is hypothetically located at a site (40.98N, 29.13E) in the Ataşehir district in Istanbul that can be considered as one of the most seismic prone locations in Istanbul populated with such residential buildings [15]. The structural properties such as modal periods, total building mass, effective modal masses, modal participation factors, etc. are computed and discussed in Özen's thesis [15]. Pushover and push-pull analyses corresponding to fundamental mode shapes in two orthogonal directions are also performed by Özen [15]. Özen [15] also identifies the damage limit states corresponding to *limited damage*, *controlled damage* and *collapse prevention* by considering the seismic design provisions of Turkish Earthquake Building Code (2018) [22], Eurocode 8 (CEN, 2004) [25] and ASCE/SEI 7-16 (2016) [26]. These studies are explained in detail by Özen [15] and this thesis uses them to idealize the building model as two inelastic SDOF systems to mimic the building response along the orthogonal directions (principal axes) of the building.

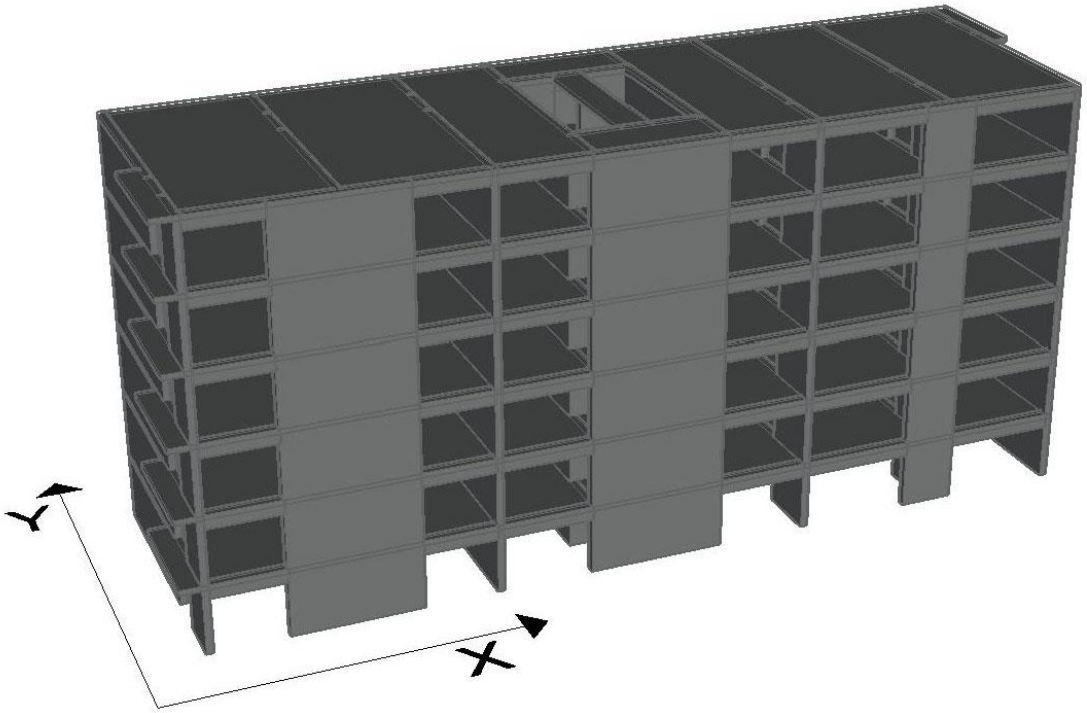

Figure 3.1: 3D view of the 6-story tunnel form model building used for seismic risk assessment procedures.

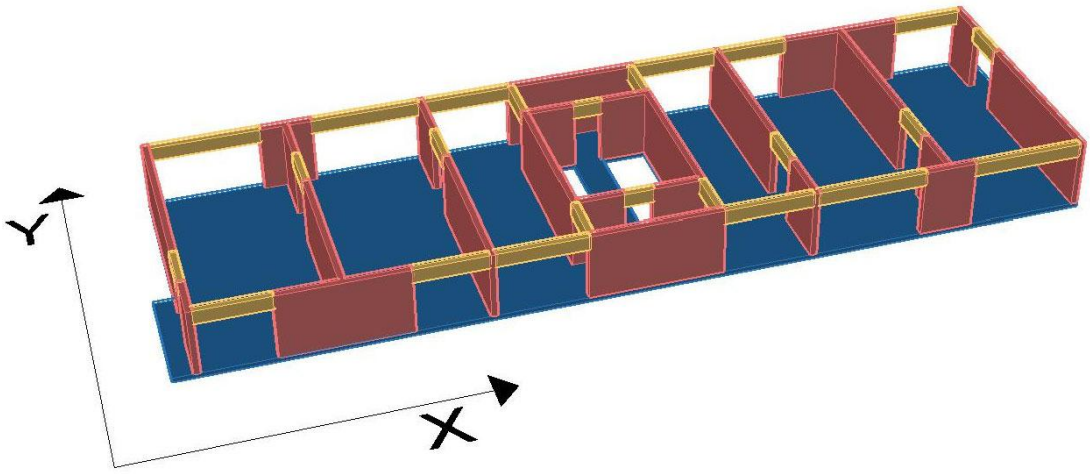

Figure 3.2: Typical floor plan of the model building

Pushover and push-pull curves as well as the limit state deformations of the model building are shown in Figure 3.3. The first four modal periods of the building, corresponding modal participation factors, effective modal mass ratios and the roof displacement components of modal shape vectors are given in Table 3.1.

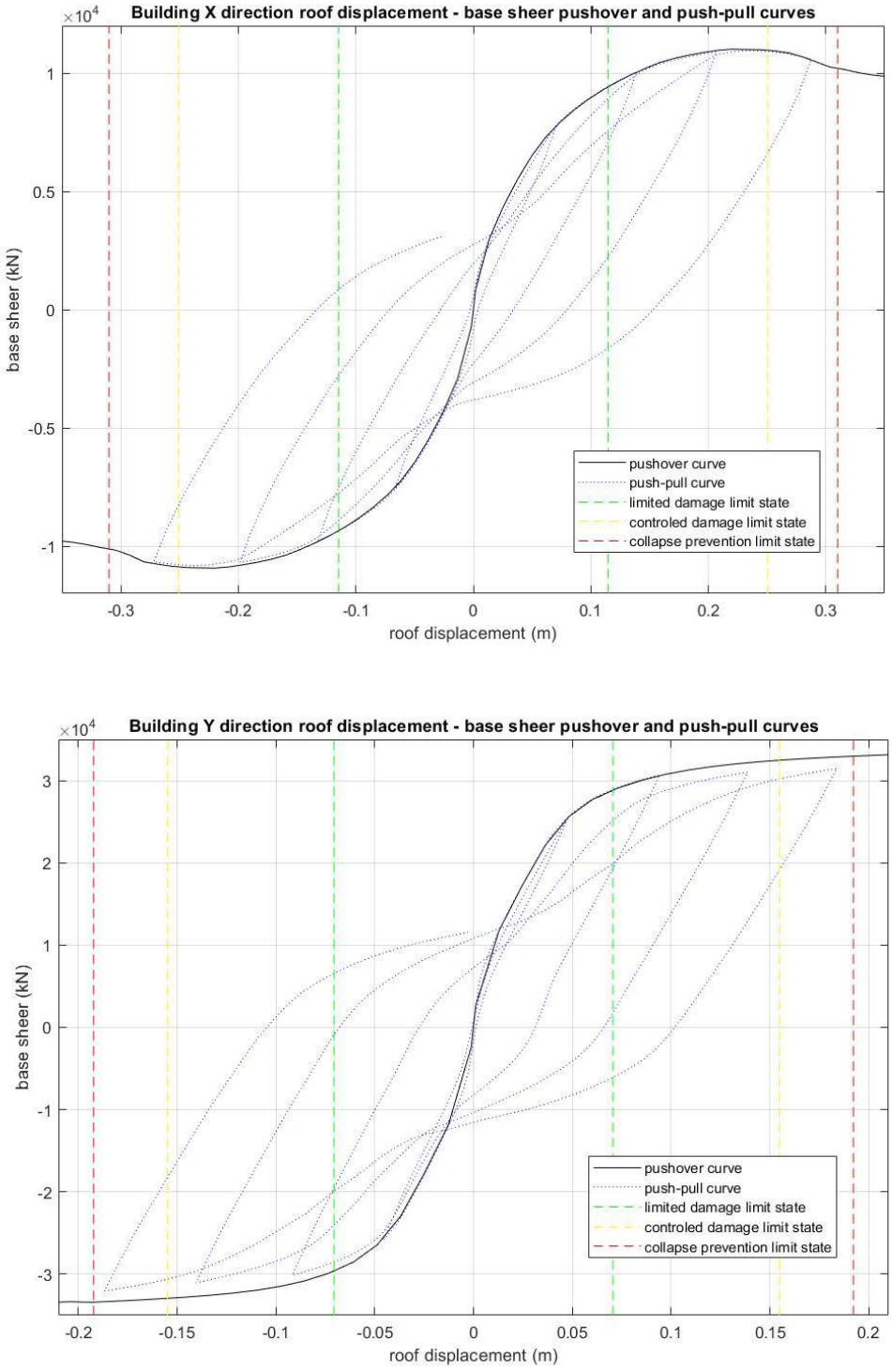

Figure 3.3: Building two orthogonal directions pushover and push-pull curves, and corresponding limit states computed in accordance with TEBC (2018) [22] provisions

Table 3.1: Modal Properties of 6 story building model used to model representative SDOF systems

Mode	Period	M_{eff-x}/M	M_{eff-y}/M	Γ_x	Γ_y	ϕ_{roof-x}	ϕ_{roof-y}
1	0.48s	0.705	~ 0	65.35	~ 0	0.021	~ 0
2	0.21s	~ 0	0.71	~ 0	66.30	~ 0	0.0206
3	0.096s	0.2	~ 0				
4	0.05s	~ 0	0.21				$M = 5052\ ton$

M_{eff-x} and M_{eff-y} represent modal effective mass in *x* and *y* directions respectively, Γ_x and Γ_y represent modal participation factors in *x* and *y* directions respectively, ϕ_{roof-x} and ϕ_{roof-y} represent roof displacement components of modal shape vector in *x* and *y* directions respectively, *M* is the total building mass

The information given in Table 3.1 is used to transform the pushover and push-pull curves from roof displacement - base shear ($U_{roof} - F_{base}$) to spectral displacement - spectral acceleration ($S_d - S_a$) pairs. The said transformation is done by the expressions given in Eqn. 3.1 and 3.2. The computed spectral capacity curves are shown in Figure 3.4 and Figure 3.5 for the orthogonal principal axes along x- and y-axis, respectively.

$$S_d = U_{roof}/(\Gamma\phi_{roof}) \tag{3.1}$$

$$S_a = F_{base}/M_{eff} \tag{3.2}$$

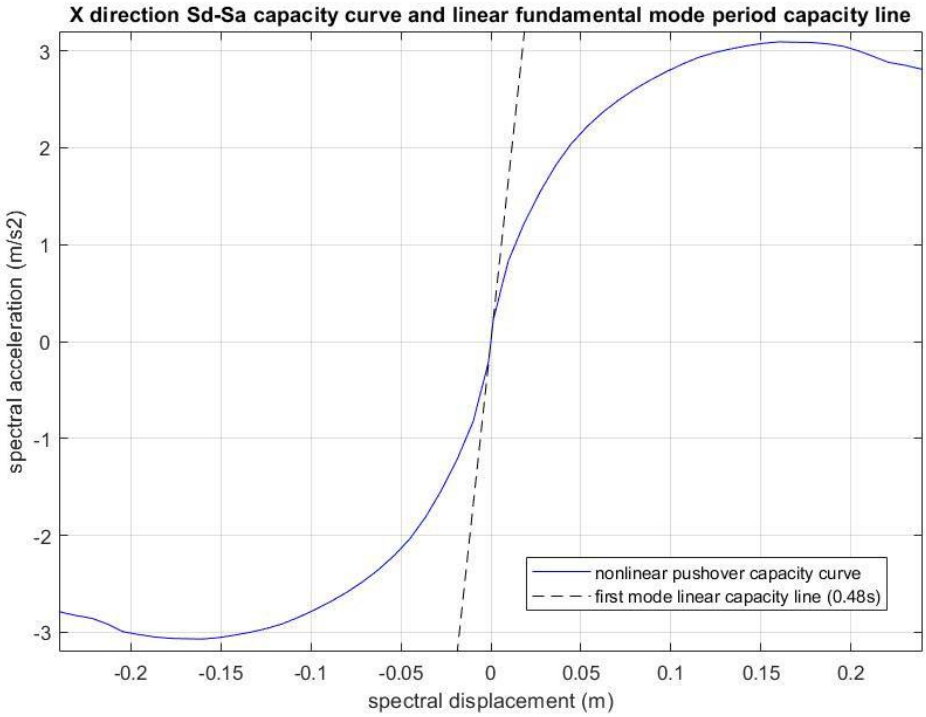

Figure 3.4: First mode (along X direction) $S_d - S_a$ nonlinear capacity curve and the tangent line to represent the first mode period computed from eigenvalue analysis

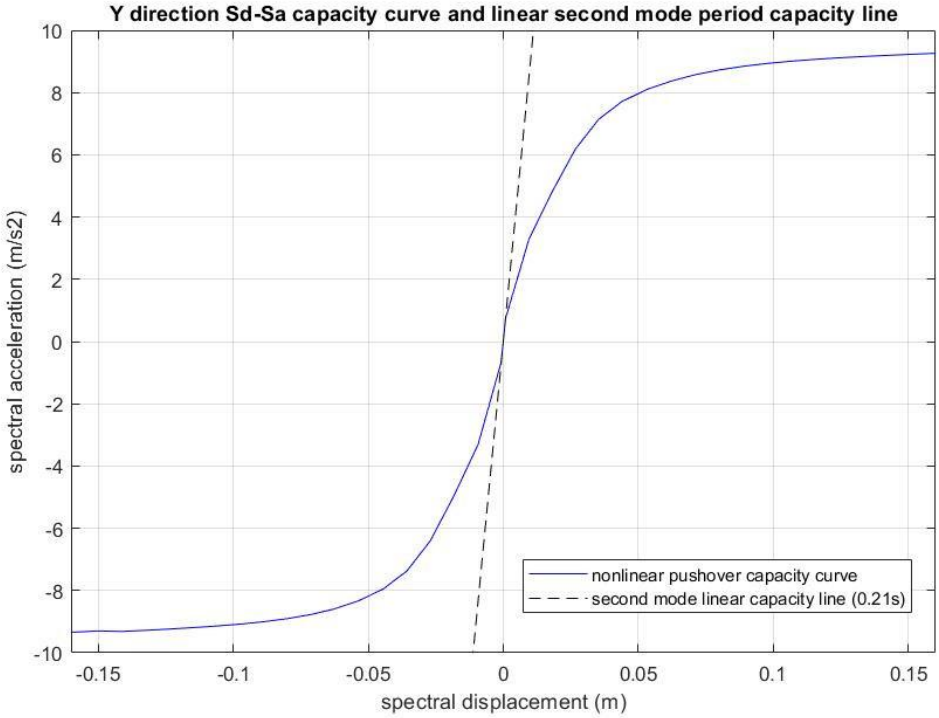

Figure 3.5: Second mode (along Y direction) $S_d - S_a$ nonlinear capacity curve and the tangent line corresponding to second mode period computed from eigenvalue analysis

Özen [15] evaluated three limit states of the building according to the provisions given in TEBC (2018) [22]. These limit states are given in Table 3.2 in term of roof displacement and spectral displacement for both orthogonal principal directions of the building.

Table 3.2: Three limit states given in terms of roof displacement and spectral displacement in two orthogonal directions. Limit states are obtained by TEBC (2018) provisions [22]. For more details about obtaining these limit states, check Özen's thesis [15].

Limit States		Limited Damage	Controlled Damage	Collapse Prevention
X Direction	U_{roof}	0.1147m	0.2506m	0.3100m
	S_d	0.0836m	0.1826m	0.2259m
Y Direction	U_{roof}	0.0706m	0.1548m	0.1922m
	S_d	0.0517m	0.1133m	0.1408m

3.2. Idealization of Building Seismic Response by Two Nonlinear SDOF Oscillators Corresponding to Two Orthogonal Building Directions

Since the model building is tunnel form, all vertical structural elements are RC walls (Figure 3.1 and Figure 3.2) and RC walls oriented in x-direction do not contribute to the stiffness and strength in y-direction, and vice versa. Besides, the structural damage in RC walls along one of the principal directions does not affect the building performance in the other principal direction. Therefore, one can assume that the building responses in two orthogonal directions are decoupled and the two principal axes can be treated (analyzed) separately. Note that this is not the case for common RC frame structures because columns have coupled response.

RC walls are much more flexible in out-of-plane direction compared to in-plane direction. Out-of-plane response is almost purely flexural while in-plane response incorporates shear behavior. In-plane response is much more rigid with decreased ductility. This means that RC walls parallel to lateral deformation will resist almost all the lateral load in that direction. Damage (energy dissipation) caused by this deformation will be almost completely concentrated in parallel walls. Perpendicular RC walls' contribution to load resisting and damage absorption (energy dissipation) is negligible because of their high flexibility in that direction. It can be assumed that RC walls only resist lateral loads in parallel direction to

their orientations. Their contribution to perpendicular load resistance can be neglected. This means that response of the building in two orthogonal directions can be decoupled with one group of vertical elements only resisting loads in one direction and other group in another.

The decoupling of model building's response in two orthogonal directions simplifies the response computations. This thesis represents the model building response as two idealized nonlinear SDOF oscillators along the principal directions. The properties of two SDOF systems are modelled according to the hysteretic model proposed by Ibarra et al. [27]. This hysteretic model incorporates trilinear monotonic backbone curve, energy dependent hysteretic strength and stiffness degradation, unloading stiffness degradation, pinching behavior, etc. Parameters that fully define the Ibarra et al. model are: K_e, F_y, α_s, $\frac{\delta_c}{\delta_y}$, α_c, λ, k_s, k_d, γ_s, γ_c, γ_k and γ_a. The first six parameters define the backbone curve, k_s and k_d define the pinching behavior, and γ parameters define energy dependent cyclic degradation. Detailed mathematical description of the model can be found in their paper [27]. An example SDOF response (right panel) to a protocol loading (left panel) is displayed in Figure 3.6.

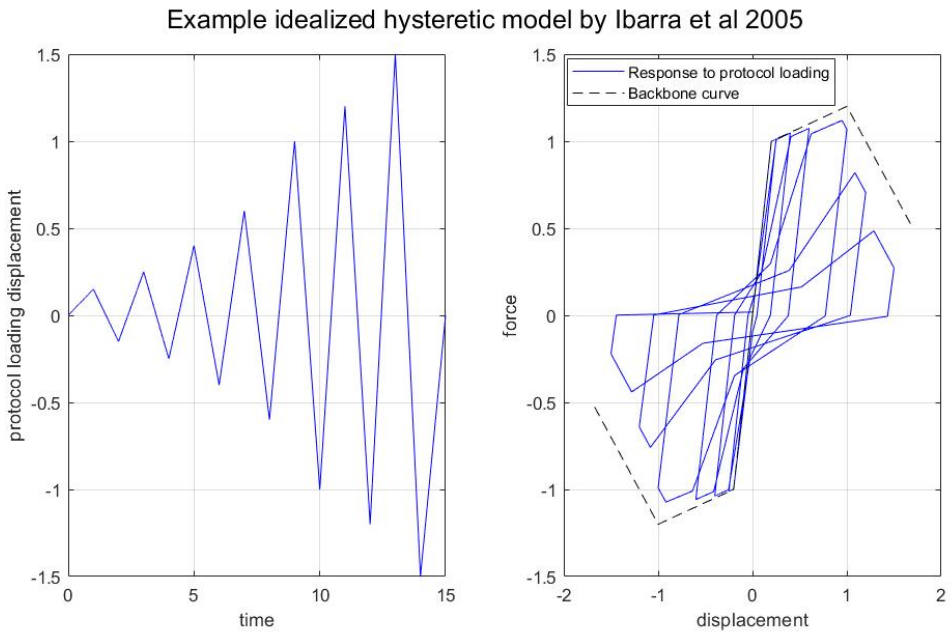

Figure 3.6: Example SDOF system response modelled by Ibarra et al [27]; $K_e = 5$, $F_y = 1$, $\alpha_s = 0.05$, $\frac{\delta_c}{\delta_y} = 5$, $\alpha_c = -0.30$, $\lambda = 0$, $k_s = 0.5$, $k_d = 0.5$, $\gamma_{s,c,k,a} = 100$, $c = 1$

The Ibarra et al. [27] hysteretic model is used to idealize $S_d - S_a$ pushover and push-pull curves corresponding to the first (along x-direction) and second (along y-direction) modes of the model building (Figure 3.4 and Figure 3.5 show the respective pushover curves). The hysteretic model parameters stated in the previous paragraph are optimized by minimizing the errors in total system energy and cumulative absolute energy (sum of the areas that fall between the idealized and real curves). The values to some of these parameters are given through visual inspection. The rest is computed through an optimization algorithm that computes the energy errors from a wide range of possible parameter configurations to detect the best fitting suit that minimizes the errors.

The K_e parameter represents the elastic stiffness of the idealized system, and it is computed by doubling the periods in x- and y-directions as shown in Eqn. 3.3 (K_e in x-direction; K_{e-x}) and Eqn. 3.4 (K_e in y-direction; K_{e-y}). The decision to reduce the elastic stiffness of the capacity curves was compulsory because, as it can be seen in Figure 3.4 and Figure 3.5, the initial stiffness lines corresponding to elastic modal periods (computed from eigenvalue analysis) are tangent to pushover capacity curves within a very narrow range and they shift away (change slope) after that narrow range. The reason behind the quick shift from initial tangent is the gradual yielding of beams long before the RC walls reach their ultimate strength capacity. Since the beams yield at very early stages during the pushover analysis, the entire system also softens at very early stages causing the capacity curve to shift away from the initial stiffness line in a rather fast way. Figure 3.7 shows the $S_d - S_a$ pushover curves for x- and y-directions along with the dashed lines representing the assumed initial stiffnesses of the idealized SDOF systems.

$$K_{e-x} = \left(\frac{2\pi}{2x0.48s}\right)^2 \tag{3.3}$$

$$K_{e-y} = \left(\frac{2\pi}{2x0.21s}\right)^2 \tag{3.4}$$

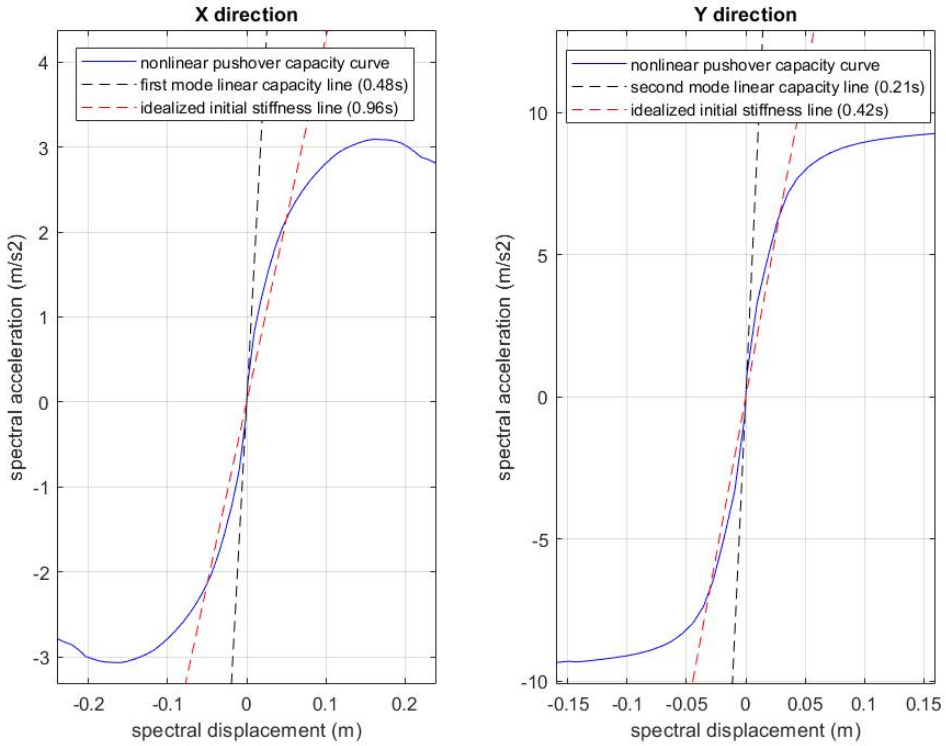

Figure 3.7: x- and y- direction (first and second modes, respectively) $S_d - S_a$ capacity curves. The initial stiffness lines obtained from linear modal analysis (black dashed lines) and idealization for the Ibarra et al. hysteretic model [27] (red dashed lines)

The parameters F_y, α_s, $\frac{\delta_c}{\delta_y}$ and α_c in Ibarra et al. model define the monotonic backbone curve of the SDOF system. These parameters are computed by an optimization algorithm that searches the optimum model configuration that minimizes the sum of errors between S_a ordinates of the pushover capacity curve. Also, this algorithm computes error in cumulative energy (area under the backbone) and accepts only configurations that keep this error under some low acceptable level (in this case 3%).

The parameters k_s, k_d, γ_s, γ_c, γ_k and γ_a are computed in the same manner but from the cyclic push-pull curve. These parameters define the cyclic hysteretic behavior of the idealized SDOF system. Figure 3.8 and Figure 3.9 displays two idealized SDOF system responses to the same push-pull protocol loading used to obtain the $S_d - S_a$ push-pull curves. Computed parameters of idealized SDOF systems are given in Table 3.3.

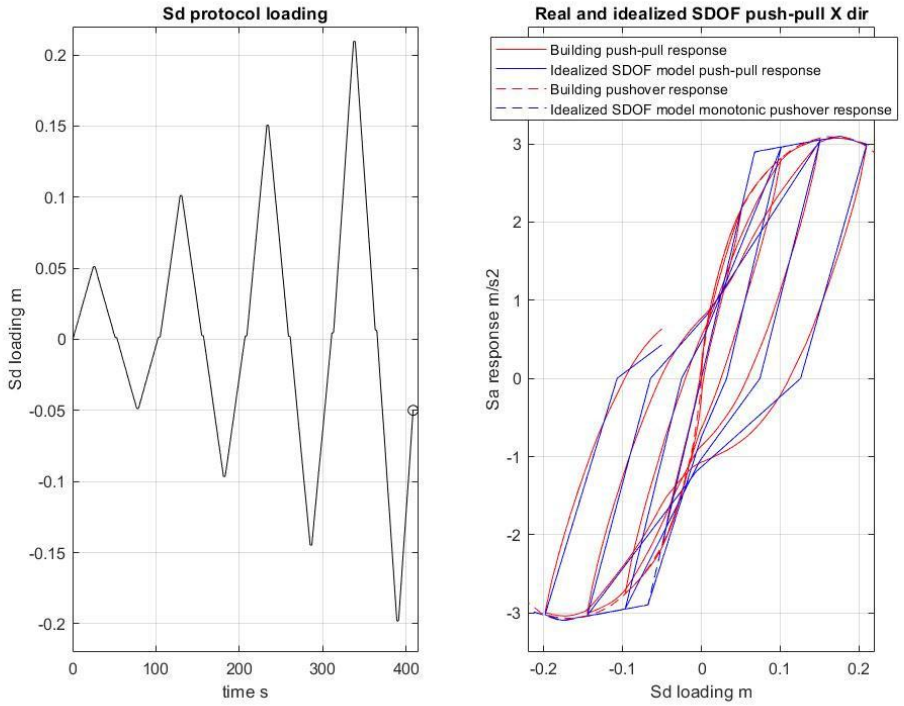

Figure 3.8: x-direction push-pull curve and the idealized SDOF system response to the same loading protocol (given on the left)

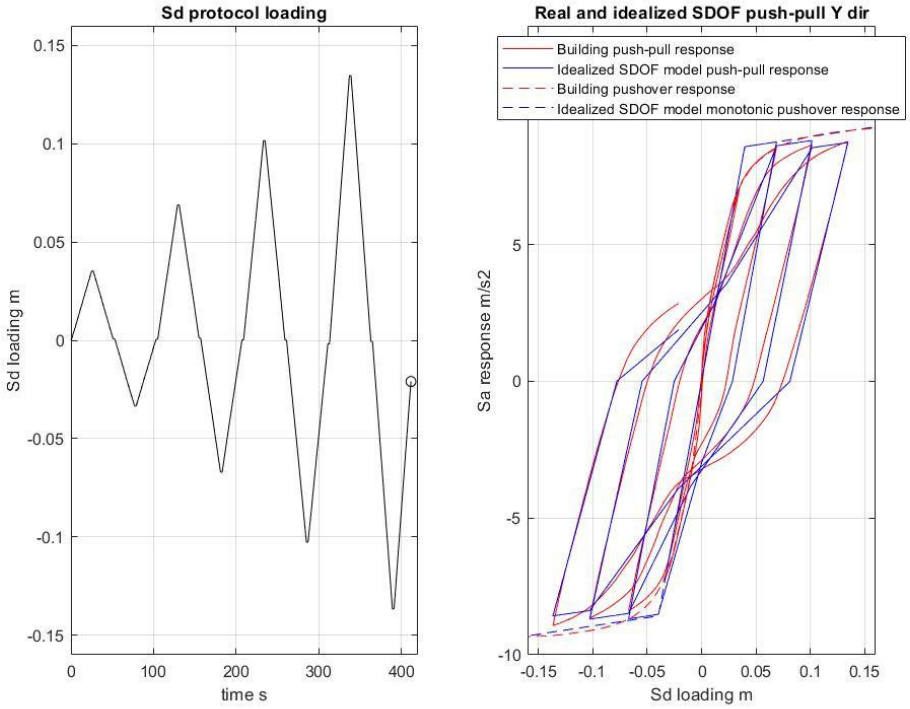

Figure 3.9: y-direction push-pull curve and the idealized SDOF system response to the same loading protocol (given on the left)

Table 3.3: Parameters used to model idealized SDOF systems by Ibarra et al. 2005 hysteretic model

	K_e	F_y	α_s	$\frac{\delta_c}{\delta_y}$	α_c	k_s	k_d	γ_s	γ_c	γ_k	γ_a
x-dir	42.84	2.90	0.0441	2.62	-0.076	0.72	1.00	2000	2000	30	2000
y-dir	223.80	8.58	0.0281	/	/	0.84	0.74	113	/	25	2000

Parameters K_e are computed by doubling modal periods from first two modes obtained by linear analysis (Eqn. 3.3 and 3.4). Parameters F_y, α_s, $\frac{\delta_c}{\delta_y}$ and α_c are computed by algorithm that searches their configuration that minimizes the sum of errors between idealized and real (obtained by MDOF analysis) push-over backbone curves. Post-caping behavior in y-direction was not modelled because collapse limit state spectral displacement is lower than capping displacement, so parameters $\frac{\delta_c}{\delta_y}$ and α_c are omitted in this direction. Parameters k_s, k_d, γ_s, γ_c, γ_k and γ_a are modelled by searching their configuration that minimizes sum of errors between idealized and real push-pull curves. In x-direction parameters γ_s, γ_c and γ_a are given high values (2000) because it was visually observed that cyclic strength deterioration (pre-capping and post-capping) and accelerated reloading stiffness deterioration are negligible. In y-direction this was done only for γ_a because cyclic strength deterioration was observed as significant and could not be neglected. Since post-capping behavior in y-direction was not modelled, γ_c parameter is omitted in this direction. Configurations of parameters that produced high errors (>3%) in cumulative energy absorbed by the system (cumulative area under the push-pull curve) were discarded even if they produced low sum of errors between S_a ordinates of idealized and real push-pull curves.

As can be seen from Figure 3.8 and Figure 3.9, the idealization is not perfect, but represents a fairly good estimation of actual building response. These two nonlinear idealized SDOF systems are used in the nonlinear analyses to estimate the actual building response to ground motions (i.e., GM records).

Section 2.4.2.3 proposes two different $AvgSA$ intensity measures that are referred to as $AvgSA1$ and $AvgSA2$. $AvgSA1$ represents the weighted geometrical mean of $SA(T)$ at six modal periods and $AvgSA2$ represents the weighted geometric mean from two modal

periods (Eqn. 2.103 - 2.104 and Eqn. 2.106 – 2.107). These periods and their weights are repeated below for convenience (Eqn. 3.5 – 3.8):

$$\overrightarrow{T_1} = \begin{bmatrix} 0.05s \\ 0.096s \\ 0.42s \\ 0.63s \\ 0.96s \\ 1.44s \end{bmatrix} \tag{3.5}$$

$$\overrightarrow{W'_1} = \begin{bmatrix} 0.026 \\ 0.0992 \\ 0.088 \\ 0.088 \\ 0.3494 \\ 0.3494 \end{bmatrix} \tag{3.6}$$

$$\overrightarrow{T_2} = \begin{bmatrix} 0.96s \\ 1.44s \end{bmatrix} \tag{3.7}$$

$$\overrightarrow{W'_2} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \tag{3.8}$$

The reason for selecting these two sets of periods and weights was not explained in Section 2.4.2.3 but it will be made clear in this section.

Since the idealized SDOF systems have initial stiffness moduli that correspond to the periods of 0.96s and 0.42s (twice the original periods from eigenvalue analysis), these two periods are now accepted as the idealized fundamental periods of vibration of the structure in two orthogonal directions. The periods of $T = 1.44s$ and $T = 0.63s$ are also accepted as important because they represent the 50% increased values of the two idealized fundamental periods (i.e., 1.5×0.96 and 1.5×0.42). The inclusion of these periods is important from the viewpoint of nonlinear period shift effects. When the 3rd and 4th modal periods (0.096s and 0.05s, respectively), which are the subsequent modal periods along x- and y- directions are included they constitute the six periods used in $AvgSA1$. Weights in $AvgSA1$ are chosen such that they correspond to effective modal masses for the first four modal periods. Periods that correspond to vibration in x-direction are firstly assigned a significance of 80% because

it was noticed that x-direction is the weak direction that determines the general behavior of the building. This weight of 0.8 is distributed between 3 periods for x-direction by considering modal effective mass ratios in x-direction. Weight of the period that corresponds to 1.5 times the fundamental mode period is set to be the same as fundamental mode period. The same was done for y-direction by distributing remaining weight of 0.2 to three y-direction periods. The weights are normalized such that their sum is equal to 1 and given in Eqn. 3.6.

After multiple trial SDOF analyses using the idealized SDOF system as given in Figure 3.8 and Figure 3.9, it is observed that the SDOF system representative of x-direction almost exclusively gives *higher* response than the SDOF system representative of y-direction. The adjective *higher* indicates that the SDOF system representing the x-direction is subject to higher limits states than the SDOF system representing the y-direction. For example, under some GM records while the SDOF system representing the structural behavior along x-direction reaches to *collapse prevention*, its counterpart representing the structural behavior along y-direction does not reach to *limited damage* limit state. This observation concludes that the model building is stronger in y-direction and the building behavior along x-direction will control the overall performance of the model building. As such, the x-direction fundamental period ($T = 0.96s$) and its %50 shifted version ($T = 1.44s$) mimicking the nonlinear response along x direction is used to define $AvgSA2$ (Eqn. 3.7 and 3.8).

3.3. Scalar IM Stripe Method Analyses

3.3.1. Record Selection and Scaling for Scalar IM Stripe Analyses

After describing the building model and the structural analysis methodology, this section explains the selection and scaling of ground-motion records, which constitutes another step in the risk assessment procedure. There are two widely used methods for record selection and scaling in the literature: stripe and cloud scaling. Baker (2007) [5] discusses stripe and cloud scaling methods for structural response assessment using both scalar and vector IMs. This section explains the stripe scaling method with a scalar IM as the metric of GM intensity.

Section 2.4 describes the methodology for computation of three different conditional spectra, namely $CS_{SA(T_1)}(T)$, $CS_{AvgSA1}(T)$ and $CS_{AvgSA2}(T)$. These three conditional spectra that are based on three distinct scalar IMs are used for record scaling and selection at 14 different intensity levels.

Baker and Lee (2017) [12] describe an algorithm for record selection to fit a target distribution. Their work improves the original version of this algorithm proposed by Jayaram et al. (2011) [28]. This thesis employs a very similar methodology for record selection and scaling as explained by Baker and Lee (2017) [12].

Record scaling and selection for scalar IM stripe analysis undertaken in this thesis is composed of following steps:

1) A record pool consisting of approximately 1300 records is assembled to represent a wide variety of GM records from different earthquakes. The GM records are downloaded from NGA West2 PEER Ground Motion Database [29].
2) Considering 3 different IMs, namely $SA(T_1)$, $AvgSA1$ and $AvgSA2$, their respective hazard curves are used to compute their values at 14 distinct hazard levels. For more details about computation of their hazard curves, check chapter 2 of this thesis. 14 hazard levels mentioned here correspond to following exceedance probabilities in 50 years: 90%, 70%, 50%, 30%, 10%, 7.5%, 5%, 4%, 2%, 1%, 0.75%, 0.5%, 0.2% and 0.1%. These hazard levels correspond to following mean return periods (earthquakes are assumed as Poisson processes): 22, 42, 72, 140, 475, 641, 975, 1225, 2475, 4975, 6642, 9975, 24975 and 49975 years.
3) Target conditional spectra ($CS_{IM}(T)$) are computed for all three mentioned IMs and for 14 hazard levels mentioned. So, in total, 42 $CS_{IM}(T)$ spectra are computed. Check Section 2.4 for details about computing $CS_{IM}(T)$.
4) For all three IMs and for all 14 hazard levels, all records in the record pool are scaled to equal corresponding IM values. Scaling factors are computed by Eqn. 3.9. IM_{target} in this equation represents IM value corresponding to specific hazard level (computed in step 2). IM_{record} represents original (unscaled) records' IM values.

$$s_{scaling\ factor} = \frac{IM_{target}}{IM_{record}} \tag{3.9}$$

5) Since $CS_{IM}(T)$ spectra computed in step 3 are just multivariate normal distributions (check Section 2.4) they can be used to simulate virtual spectra. This is done simply by random sampling from multivariate normal distribution. 20 virtual spectra are simulated for each $CS_{IM}(T)$
6) 20 simulated virtual spectra are used as target spectra for record selection. Scaled record pool (step 4) is used and 20 records that show the best fit to 20 simulated virtual records are identified. This is done by evaluating sum of the squares of errors (SSE) between record pool and virtual simulated spectra. Record set that produces lowest SSE values is selected. Mean and standard deviation of selected record set is compared with the mean and standard deviation of $CS_{IM}(T)$ and SSEs are computed between these two (between means and standard deviations).
7) Steps 5 and 6 are repeated multiple times. With every repetition, step 5 produces different set of simulated virtual spectra because they are produced by random sampling. This means that these repetitions do not produce the same results. In total, this process was repeated 2500 times. Among these repetitions, record set that produced minimum SSE between selected record set and $CS_{IM}(T)$ (in terms of both mean and standard deviation) was selected as final output of this selection process.

Figure 3.10, Figure 3.11 and Figure 3.12 show the 475-year return period spectra of the selected 20 records for $SA(T_1)$, $AvgSA1$ and $AvgSA2$ conditional spectra, respectively.

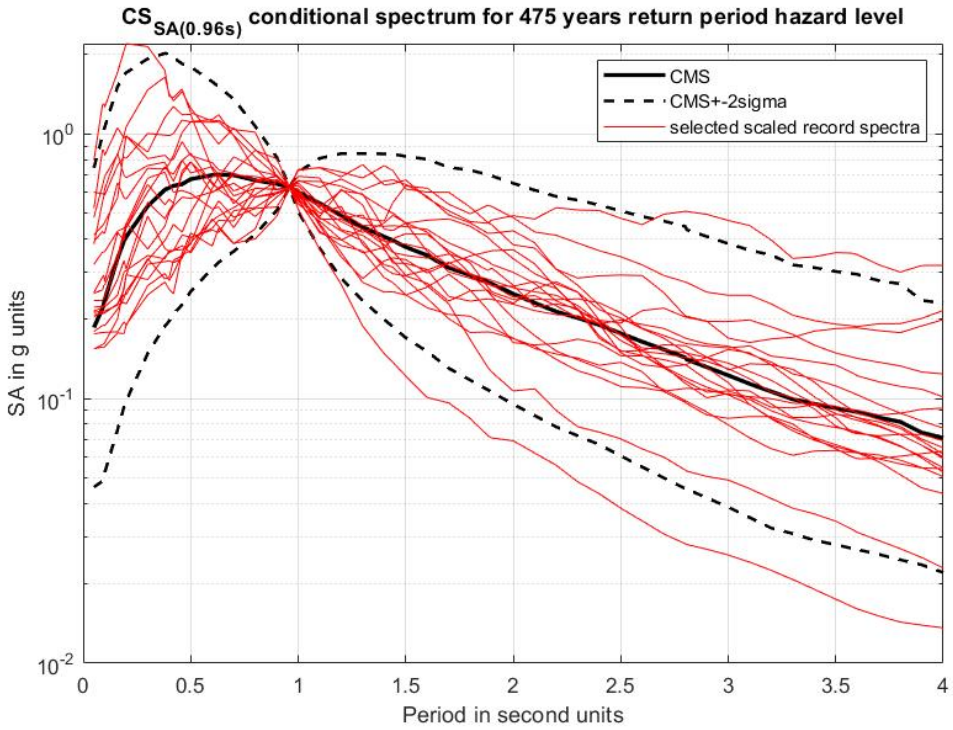

Figure 3.10: 475-year return period $CS_{SA(T_1)}(T)$ and the best matching records to this target

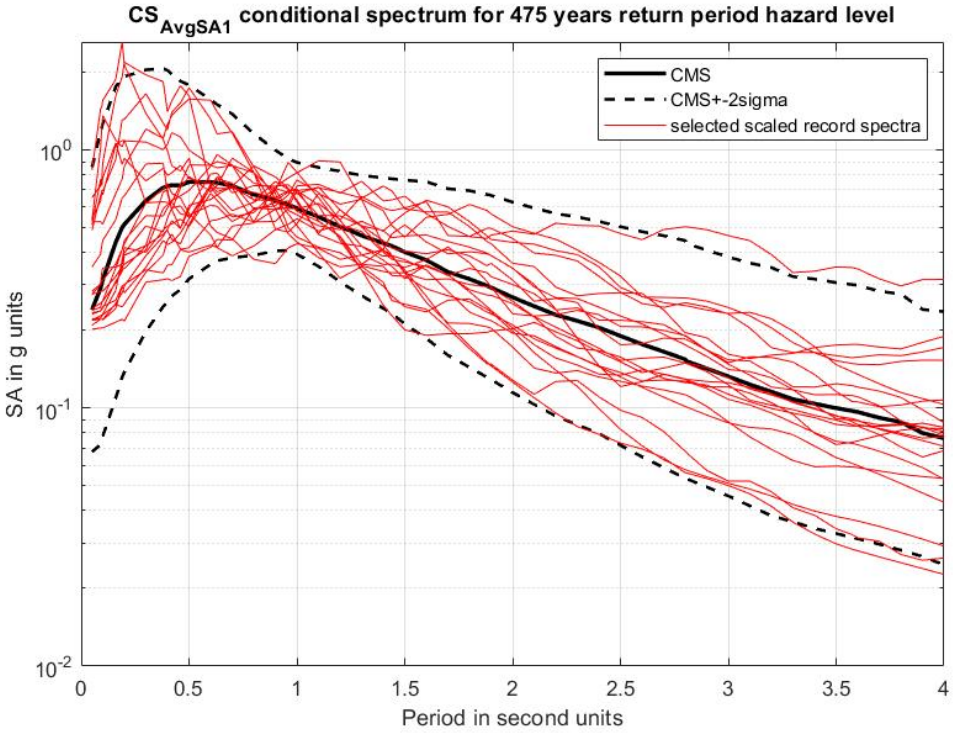

Figure 3.11: 475-year return period $CS_{AvgSA1}(T)$ and the best matching records to this target

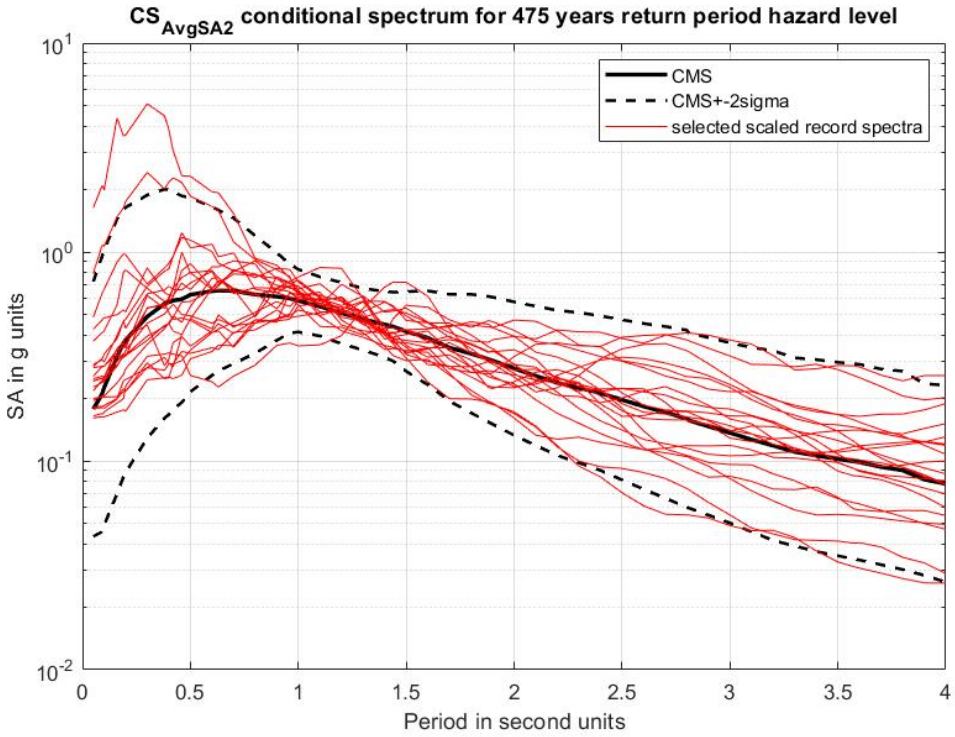

Figure 3.12: 475-year return period $CS_{AvgSA2}(T)$ and the best matching records to this target

3.3.2. Structural Response Data from SDOF Analyses

The records are selected and scaled by considering the geometric mean horizontal component definition as the target conditional spectra impose this horizontal component definition due to ASB14. At the first glance, everything can be considered as consistent because the horizontal components of scaled record, whose geometrical mean is used to scale the target conditional spectra, can be used to perform the response analyses of the two idealized nonlinear SDOF systems. On the other hand, this approach may introduce some bias in the SDOF response results because the analyses assume that the principal axes of the building model and the horizontal components of the scaled records have the same orientation. The occurrence probability of this assumption is almost negligible. Nievas et al. (2016) [30] propose a methodology to surmount this deficiency by considering multiple orientations of the two horizontal components of the scaled ground motions. For a given rotation angle θ, the respective orientations of the horizontal components are computed from Eqn. 3.10 and 3.11.

$$u_{gx-rot\theta}(t) = \cos(\theta)\, u_{gx}(t) + \sin(\theta)\, u_{gy}(t) \tag{3.10}$$

$$u_{gy-rot\theta}(t) = -\sin(\theta)\, u_{gx}(t) + \cos(\theta)\, u_{gy}(t) \tag{3.11}$$

$u_{gx}(t)$ and $u_{gy}(t)$ in Eqn. 3.10 and 3.11 represent as-recorded horizontal orthogonal components of ground motion. θ represents angle between as-recorded ground motion and building orientation. In other words, it is an angle between x-direction of as-recorded ground motion and x-direction of the building excited by that GM. Eqn. 3.10 and 3.11 evaluate components of ground motion in building x and y-directions. $u_{gx-rot\theta}(t)$ and $u_{gy-rot\theta}(t)$ represent GM components in building x and y-directions respectively.

Fifteen distinct non-redundant orientation angles are considered for each selected GM record. The angle θ is varied by $\Delta\theta = 12°$ increments from 0° to 168°. Consideration of $\theta = 180°$ and incrementing θ beyond this value would lead to redundancy in the results since such a process would only flip the sign of horizontal GM components duplicating the already computed structural responses.

Every projection of horizontal GM results in two horizontal components leading to two SDOF runs. One analysis uses $u_{gx-rot\theta}(t)$ as the acceleration input along x-direction and the corresponding idealized SDOF model, and the other uses $u_{gy-rot\theta}(t)$ along y-direction associated with the idealized SDOF system in the same direction.

A total of 25200 SDOF analysis runs were made for stripe scaling of scalar IMs for 15 angular rotations, 20 records, 2 horizontal components, 14 intensity levels and 3 different scalar IMs. These analyses yield 12600 SDOF response pairs giving $S_d - S_a$ response curves that can be transformed back to roof displacement – base shear coordinates by Eqn. 3.1 and 3.2.

Sample responses from different intensity levels are displayed in Figure 3.13, Figure 3.14 and Figure 3.15. The maximum roof displacements in x- and y-directions from 12600 SDOF analyses pairs are computed and stored for the risk assessment of the subject model building.

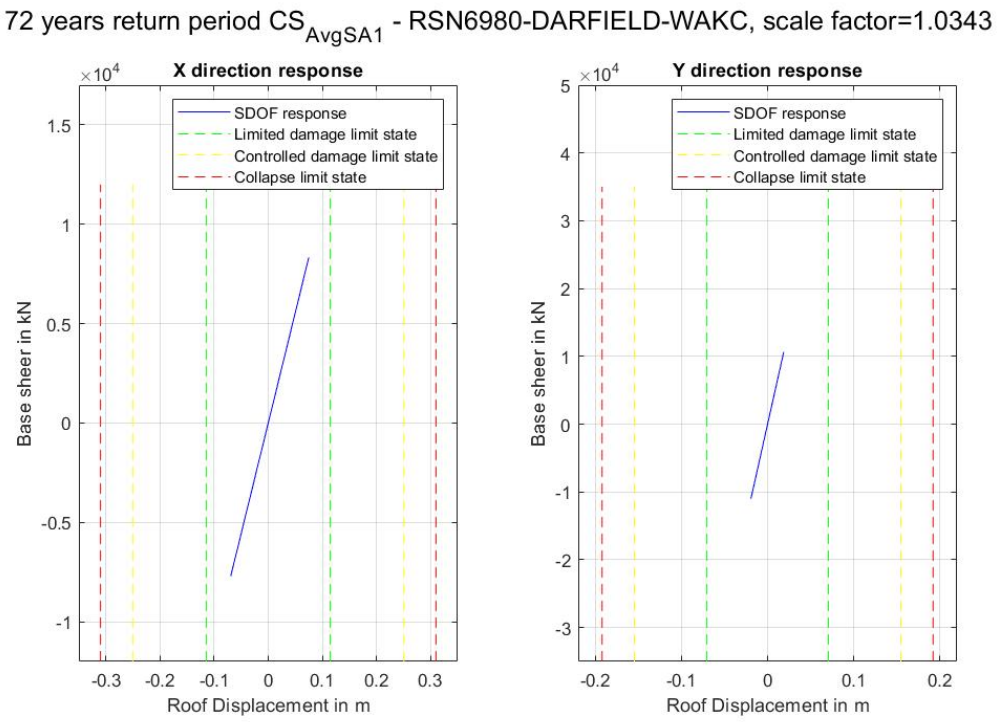

Figure 3.13: Idealized SDOF system responses to a selected GM record (non-rotated $\theta = 0°$) selected to fit 72-year return period hazard level for CS_{AvgSA1}

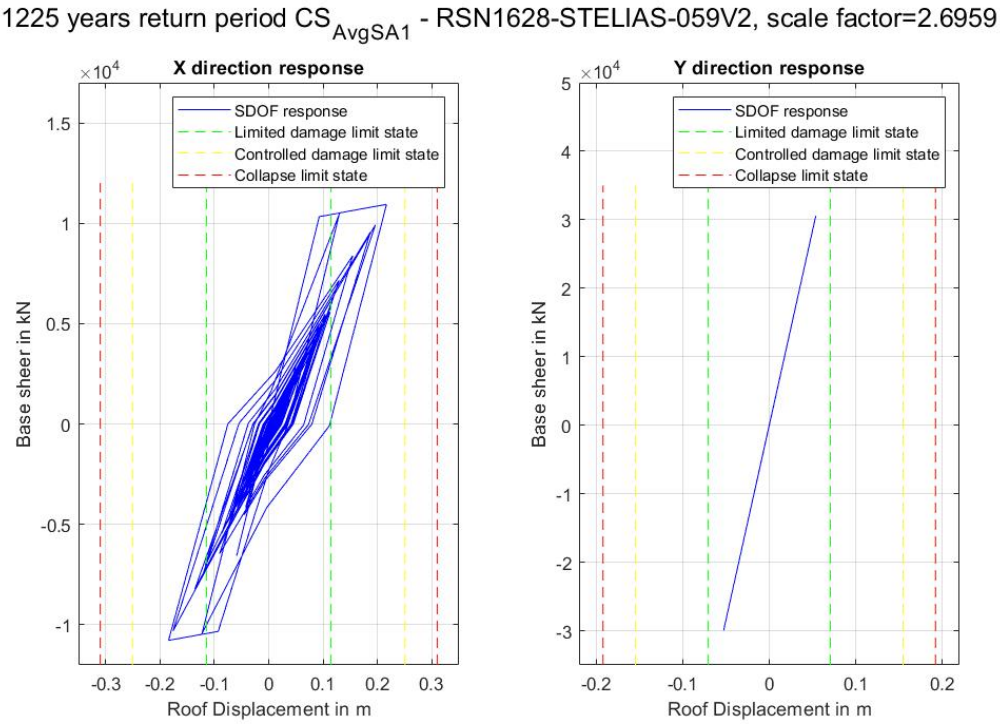

Figure 3.14: Idealized SDOF system responses to a selected GM record (non-rotated $\theta = 0°$) selected to fit 1225-year return period hazard level for CS_{AvgSA1}

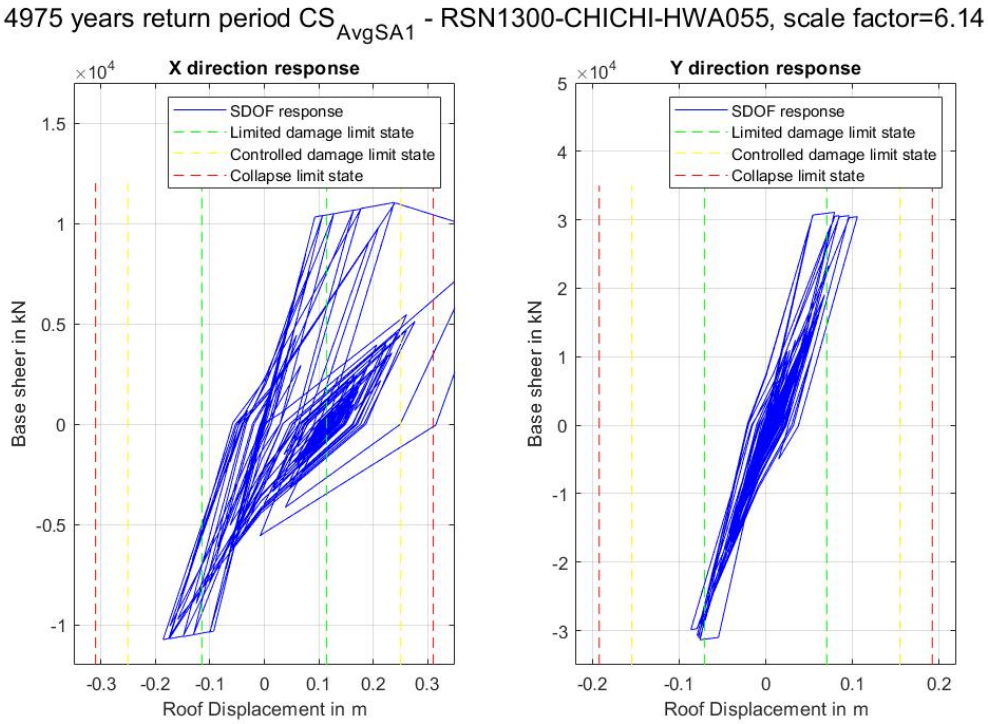

Figure 3.15: Idealized SDOF system responses to a selected GM record (non-rotated $\theta = 0°$) selected to fit 4975-year return period hazard level for CS_{AvgSA1}

3.3.3. Engineering Demand Parameter (EDP) Used for Risk Assessment

Engineering demand parameter is a response quantity of a structure chosen as a metric for general structural damage and loss. The goal of risk assessment of a structure is to evaluate at which rates the chosen EDP exceeds certain levels. This information can be used to estimate and assess financial and life risks by developing some statistical and probabilistic models between EDP and loss (financial or life losses). This thesis demonstrates the computation of EDP exceedance rate curves by different methods. Mathematical formulation for computation of EDP exceedance rate curves (also commonly known as *risk curves*) used in this thesis is given by Baker (2007) [5] and repeated below for completeness (Eqn. 3.12):

$$\lambda_{EDP}(y) = \int_{IM} G_{EDP|IM}(y|im) \left| \frac{d\lambda_{IM}(im)}{dim} \right| dim \tag{3.12}$$

$\lambda_{EDP}(y)$ represents the rate at which certain EDP parameter exceeds the threshold y. The EDP can be any structural response parameter (for example beam joint plastic rotation), however, it is desirable to choose the EDP as a parameter that is a good indicator of the

general damage of the building. Term $\left|\frac{d\lambda_{IM}(im)}{dim}\right|$ from Eqn. 3.12 represents the occurrence rate density of $IM = im$. It is computed by taking the absolute value of the derivative of the hazard curve as the hazard curve has negative slope. $G_{EDP|IM}(y|im)$ term represents the probability that EDP exceeds y given that $IM = im$ has occurred. This term is called as fragility and its modeling is presented as part of this chapter by making use of the structural response data.

Most buildings have many response parameters, and it is very hard to pick only one of them as a general representative of the building damage (loss). Mathematically, a sophisticated way of modeling this fact is to consider multiple structural response parameters. A suite of these parameters would constitute an *engineering demand parameter vector* ($\overrightarrow{EDP}$). In such a case Eqn. 3.12 is modified as given in Eqn. 3.13:

$$\lambda_{\overrightarrow{EDP}}(\vec{y}) = \int_{IM} G_{\overrightarrow{EDP}|IM}(\vec{y}|im)\left|\frac{d\lambda_{IM}(im)}{dim}\right| dim \tag{3.13}$$

$\lambda_{\overrightarrow{EDP}}(\vec{y})$ in Eqn. 3.13 represents the rate of all components in $\overrightarrow{EDP}$ being larger than specific thresholds in $\vec{y}$. $G_{\overrightarrow{EDP}|IM}(\vec{y}|im)$ in this case represents the probability that $\overrightarrow{EDP}$ components are larger than $\vec{y}$ given that $IM = im$. If sophisticated MDOF nonlinear dynamic analysis is used as structural analysis, $\overrightarrow{EDP}$ could be accepted as, for example, maximum inter-story drift ratio for all stories and for both x and y directions. The idealized SDOF analysis in this thesis focuses on only two independent structural response parameters, namely maximum roof displacement in x- and y-directions. Section 3.3.6 demonstrates how to evaluate $\lambda_{\overrightarrow{EDP}}(\vec{y})$ in such a case. The following section introduces a method to simplify the risk assessment by combining the maximum roof displacements in x- and y-directions as a single scalar EDP.

3.3.4. Limit State Normalized Maximum Roof Displacement as an EDP

This section introduces a novel scalar EDP that is supposed to simplify the risk assessment procedure by combining the x- and y-direction maximum roof displacements into single scalar parameter. The starting point of this novel scalar EDP was to consider the maximum

of the two orthogonal maximum roof displacements. However, since the model building has different limit states in two principal (orthogonal) directions, giving a reasonable decision on the one to consider in risk analysis would be difficult. For example, the collapse prevention limit state of maximum roof displacement in x-direction is 31cm whereas it is 19.22cm in y-direction for the subject model (check Özen's thesis [15])

To demonstrate the above fact, two fictitious building response results are given as example cases. In the first case, the maximum roof displacements in x- and y-directions (after nonlinear dynamic analysis) are computed as 25cm and 15cm, respectively. In the second case, let the maximum roof displacements in x- and y-directions be 25cm and 20cm, respectively. If maximum of the x- and y-direction maximum roof displacements is considered as the scalar EDP, it would take the value of 25cm in both cases. This EDP would correctly address the collapse performance of the model building for the first case as the collapse limits in x- and y-directions are not exceeded and the scalar EDP is below the critical maximum roof displacement limit for collapse state (31 cm in x-direction). However, the scalar EDP as 25cm would not correctly address the collapse prevention limit state in the second case as the building experiences collapse in the y-direction as the maximum roof displacement of 20cm exceeds the collapse limit state of the model building in y-direction (which is 19.22cm). So, the same scalar EDP results in different performance results, which suggests that consideration of maximum of maximum roof displacements in two orthogonal directions as a scalar EDP is not a viable choice.

To overcome this point, the novel EDP, referred to as *limit states normalized maximum roof displacement* (LSNRD) is introduced. The idea behind LSNRD is to take the maximum roof displacements in two orthogonal directions and see where they fall within the respective limit states. For example, the three limit states in x-direction for the subject building are given in Table 3.2 and repeated bellow for convenience (Eqn. 3.14 – 3.16):

$$MaxXRoofDisplacement_{limited\ damage} = 11.47cm \tag{3.14}$$

$$MaxXRoofDisplacement_{controled\ damage} = 25.06cm \tag{3.15}$$

$$MaxXRoofDisplacement_{collapse\ prevention} = 31.00cm \tag{3.16}$$

LSNRD in x-direction for these three limit states is 0.35 (~11.47/31), 0.80 (~25.06/31) and 1 (=31/31), respectively. The maximum roof displacement computed from a nonlinear dynamic analysis is evaluated by considering these three normalized limit states through interpolation. For example, if maximum roof displacement in x-direction is 15cm, $LSNRD_x$ is:

$$LSNRD_x = 0.35 + \frac{0.80 - 0.35}{25.06cm - 11.47cm}(15cm - 11.47cm) = 0.467 \tag{3.17}$$

If, for example, the maximum roof displacement exceeds the collapse prevention limit state then LSNRD is computed by dividing it to collapse prevention limit state. For example, assuming that the maximum roof displacement in x-direction is obtained as 40 cm from nonlinear dynamic analysis, $LSNRD_x$ is computed as:

$$LSNRD_x = \frac{40cm}{31cm} = 1.29 \tag{3.18}$$

The LSNRD value exceeding 1 indicates collapse.

The separate computation of LSNRD in x- and y-directions ($LSNRD_x$ and $LSNRD_y$) and considering the maximum of these two individual LSNRD values will yield the final LSNRD (Eqn. 3.19) for that nonlinear dynamic analysis:

$$LSNRD = max\left(LSNRD_x, LSNRD_y\right) \tag{3.19}$$

$LSNRD$, as a scalar EDP, is a good indicator of general damage level for a building. For example, for the subject model building, $LSNRD$ value of 0.80 means that the building has reached to controlled damage limit state in one of its two orthogonal directions. As already stated, $LSNRD > 1$ corresponds to collapse.

The proposed scalar EDP (LSNRD) enables the use of Eqn. 3.12 to compare the risk assessment results from different scalar IM methods. Needless to say, the information

provided by LSNRD for risk assessment is not as comprehensive as its vector $\overrightarrow{EDP}$ counterpart that employs the two orthogonal maximum roof displacements simultaneously.

3.3.5. Statistical Modeling of Structural Response

Since the scalar EDP is now established, Eqn. 3.12 can be employed to compute the EDP rate of exceedance curve (risk curve). The $G_{EDP|IM}(y|im)$ in Eqn. 3.12 can be modelled through discrete statistics of the structural response data. Figures 3.16 – 3.18 show these data for three different scalar IMs that are already explained: $SA(T_1)$, $AvgSA1$ and $AvgSA2$.

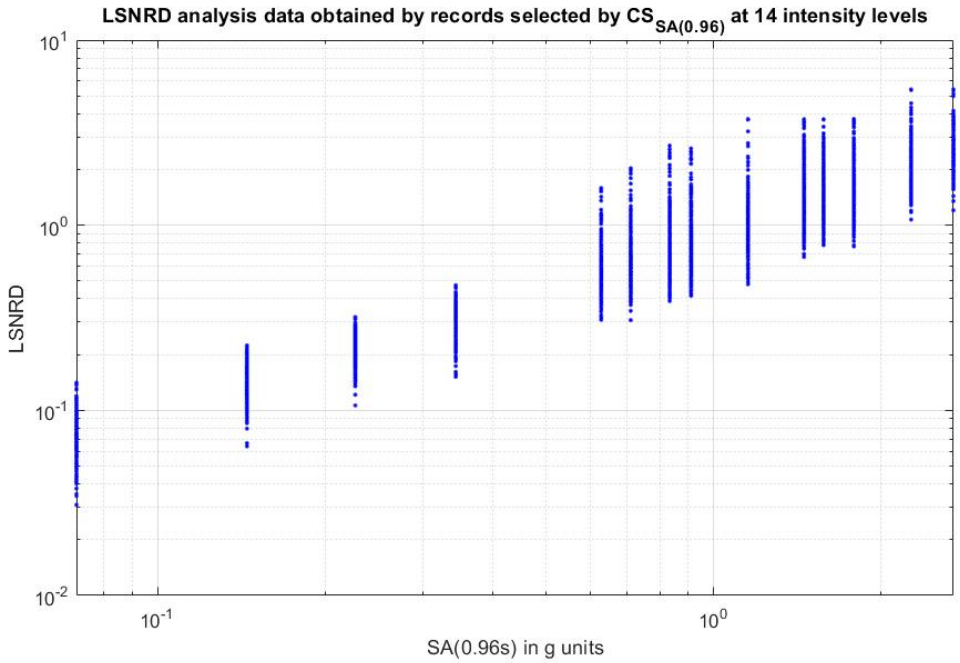

Figure 3.16: IM Stripe analysis data for $SA(T_1 = 0.96s)$ when $CS_{SA(0.96s)}$ is used for record scaling and selection at 14 discrete hazard levels

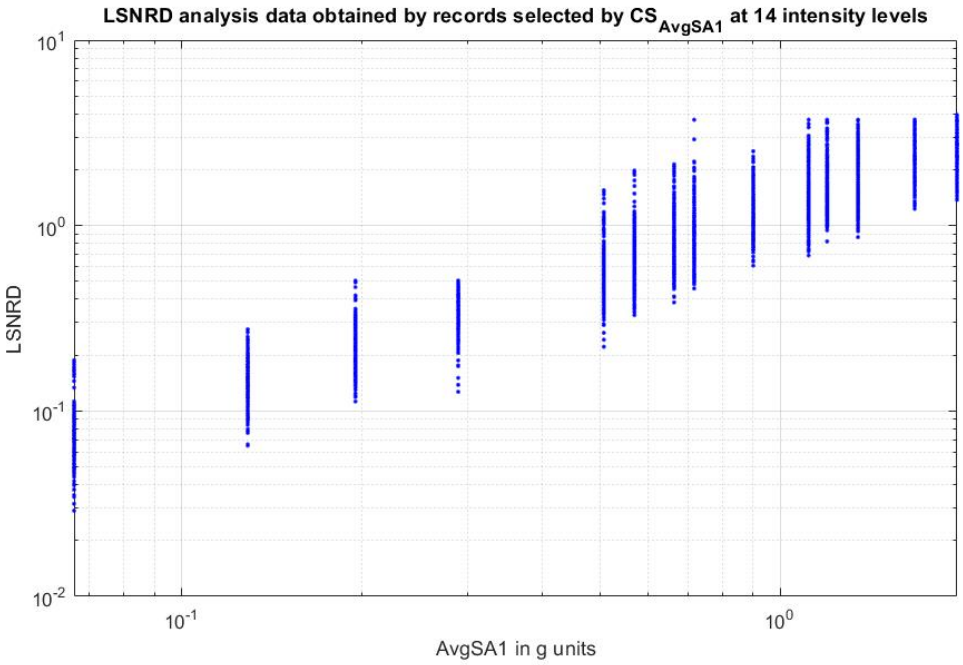

Figure 3.17: IM Stripe analysis data for $AvgSA1$ when CS_{AvgSA1} is used for record scaling and selection at 14 discrete hazard levels

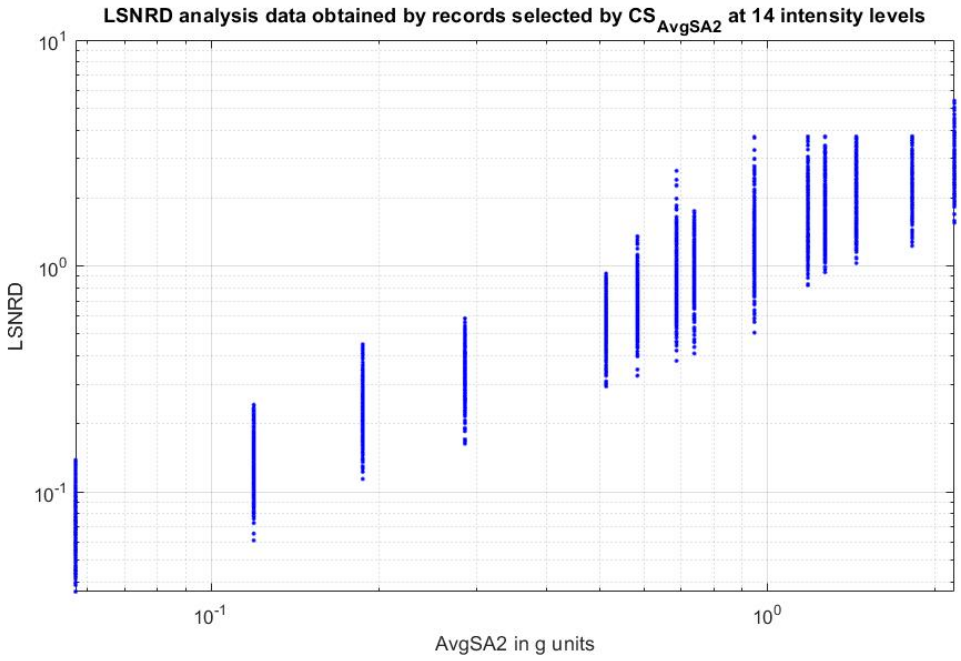

Figure 3.18: IM Stripe analysis data for $AvgSA2$ when CS_{AvgSA2} is used for record scaling and selection at 14 discrete hazard levels

The IM vs. EDP data as displayed in Figures 3.16 – 3.18 are grouped into 14 discrete IM stripes (origin of the terminology used for *stripe* scaling method) because the GM records used for the dynamic analyses are scaled to each one of these target IM levels. Each stripe

contains 300 data points corresponding to 20 records and 15 different orientations per record. The data in these figures are divided into two main parts: data with $LSNRD < 1$ and data with $LSNRD \geq 1$. the data having $LSNRD \geq 1$ represent collapse of the model building, and the other data group represents non-collapse results. If the number of collapse data in a given stripe is divided by the total data number in that stripe (i.e., 300), one can obtain the discrete collapse fragilities as given in Figures 3.19 – 3.21.

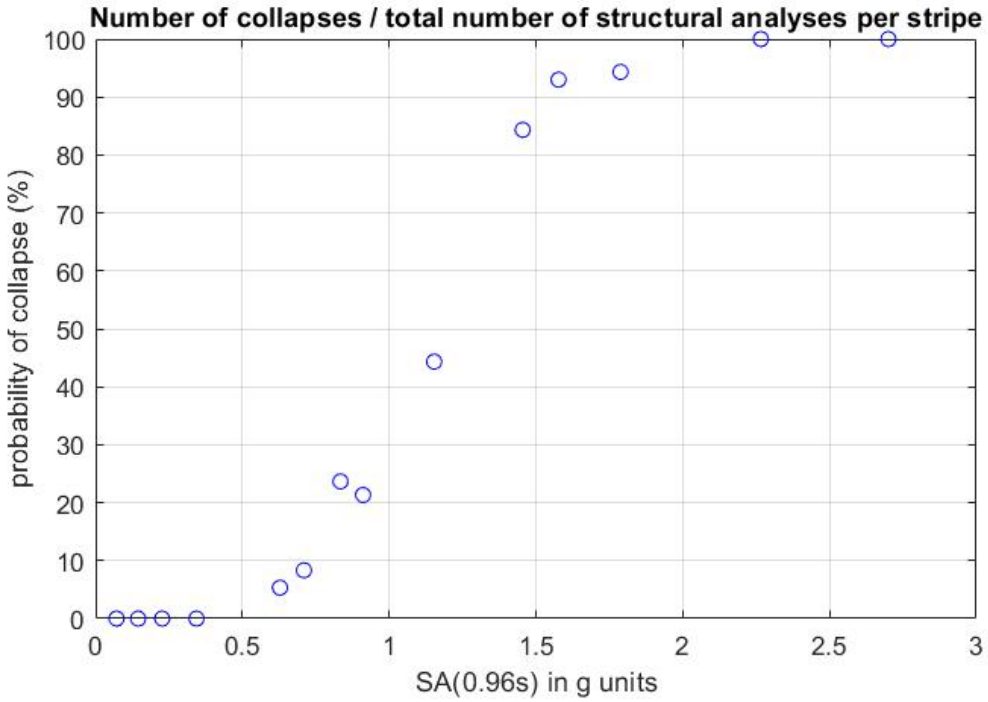

Figure 3.19: Discrete collapse probability data for $SA(T_1 = 0.96s)$ when $CS_{SA(0.96s)}$ is used for record scaling and selection at 14 discrete hazard levels

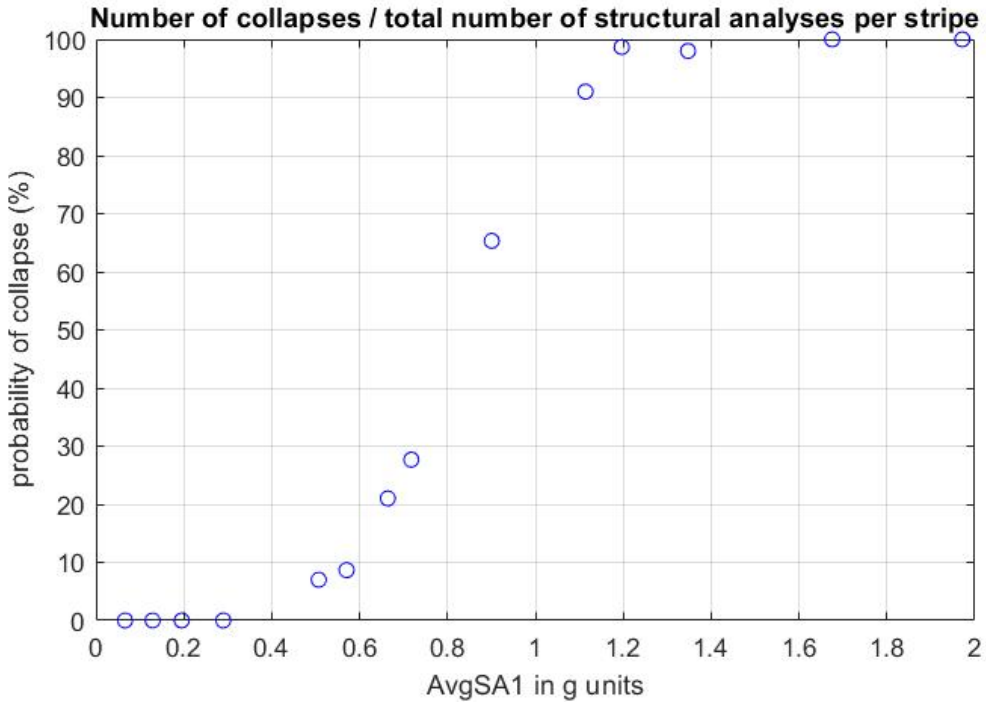

Figure 3.20: Discrete collapse probability data for $AvgSA1$ when CS_{AvgSA1} is used for record scaling and selection at 14 discrete hazard levels

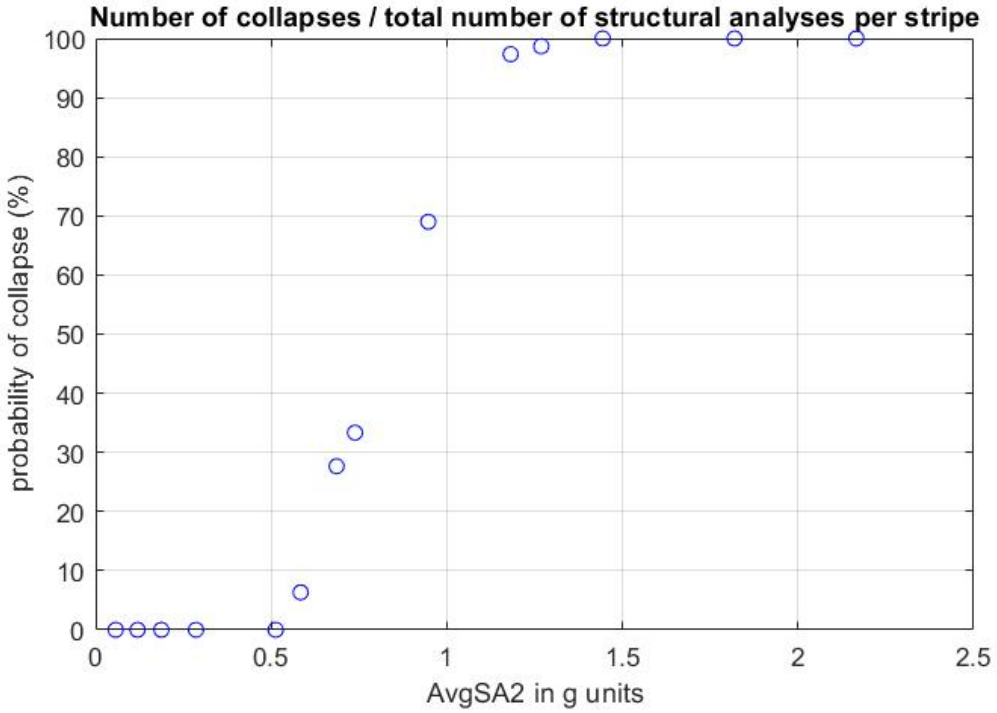

Figure 3.21: Discrete collapse probability data for $AvgSA2$ when CS_{AvgSA2} is used for record scaling and selection at 14 discrete hazard levels

The continuous collapse fragility curves are fitted to these data by assuming that the lognormal distribution prevails. Eqn. 3.20 shows this calculation using the lognormal CDF function:

$$P_C = \Phi\left(\frac{ln(IM) - \mu_{C-ln(IM)}}{\sigma_{C-ln(IM)}}\right) \tag{3.20}$$

$\mu_{C-ln(IM)}$ (logarithmic mean for collapse) and $\sigma_{C-ln(IM)}$ (logarithmic standard deviation for collapse) in Eqn. 3.20 are computed using the maximum likelihood estimation (MLE) method [31, 32]. In simple words, maximum likelihood estimation is a method of determining parameters of a certain model by finding their configuration that maximizes the probability of occurrence of observed data. Model that gives the highest probability of producing the observed data is the most likely model given the observed data. In this case, the observed data is number of collapses out of 300 structural response outcomes within each stripe. Modelled probability of collapse is given by Eqn. 3.20. Goal is to find configuration of parameters $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ of this model that maximizes probability of occurrence of observed data (numbers of collapses within stripes). This probability is also called likelihood function (it is a function of $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$). The likelihood function is established by assuming binomial distribution at each stripe and then by multiplying the probability mass functions of fourteen stripe outcomes (numbers of collapses within each stripe). Binomial distribution is explained in following paragraphs.

Let as imagine a binary process that has only two possible outcomes, *a* or *b*. Every trial of this process produces either *a* or *b*. In each trial, mentioned process has probability *p* to produce a outcome, and probability $q = 1 - p$ to produce *b* outcome. Given n trials of the process, binomial distribution represents the probability distribution of the number of a outcomes. If number of a outcomes is notified as N_a, then:

$$N_a \sim B(n, p) \tag{3.21}$$

$$f_{N_a}(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{3.22}$$

In Eqn. 3.21, B represents binomial distribution with n number of trials and p probability of outcome a in single trial. Eqn. 3.22 gives probability mass function (PMF) of N_a. $f_{N_a}(k)$ represents the probability that N_a is equal to k given N_a is binomially distributed variable as presented in Eqn. 3.21. For more details on binomial distribution reader is referred to any introductory statistics book or simply to one of many online resources on the subject [33].

Binomial distribution can be applied in stripe method when assessing collapse fragility. Since every stripe represents set of structural analyses results with two possible outcomes, namely collapse and non-collapse, numbers of collapses within each stripe can be represented by binomial distribution:

$$N_{C-i} \sim B(300, P_{C-i}) \tag{3.23}$$

In Eqn. 3.23, N_{C-i} is a random variable that represents number of collapses in *i*'th stripe (there are 14 stripes, so *i* can take values from 1 to 14) given that probability of collapse in *i*'th stripe is P_{C-i}. Every stripe contains results of 300 structural analyses, hence $n = 300$.

If k_i is number of collapses in *i*'th stripe obtained by structural analyses (Figures 3.19 – 3.21), steps for computing $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ numerically by MLE method are following:

1) $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ are assigned some candidate values. If IM value at stripe *i* is $IM_{stripe-i}$, P_{C-i} is computed for each stripe by Eqn. 3.24:

$$P_{C-i} = \Phi\left(\frac{ln\left(IM_{stripe-i}\right) - \mu_{C-ln(IM)}}{\sigma_{C-ln(IM)}}\right) \tag{3.24}$$

2) Since number of collapses at each stripe is assumed to be binomially distributed, probability of obtaining k_i observed number of collapses at *i*'th stripe is:

$$P(N_{C-i} = k_i) = \binom{300}{k_i} {P_{C-i}}^{k_i} (1 - P_{C-i})^{300-k_i} \tag{3.25}$$

3) Likelihood function (L) represents probability of occurrence of all the observed data, at all 14 stripes, so it can be computed by multiplying the probabilities of obtaining k_i observed number of collapses at all stripes:

$$L\left(\mu_{C-ln(IM)}, \sigma_{C-ln(IM)}\right) = \prod_{i=1}^{i=14} \binom{300}{k_i} {P_{C-i}}^{k_i} (1 - P_{C-i})^{300-k_i} \tag{3.26}$$

L is a function of $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ since they are parameters used to compute P_{C-i} (Eqn. 3.24).

4) Steps 1, 2 and 3 are repeated for wide ranges of densely spaced $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ candidate configurations. Configuration that maximizes L function is selected to represent MLE collapse fragility model.

Maximum likelihood collapse fragility curves are computed by algorithm described above and they are presented in Figures 3.22 – 3.24.

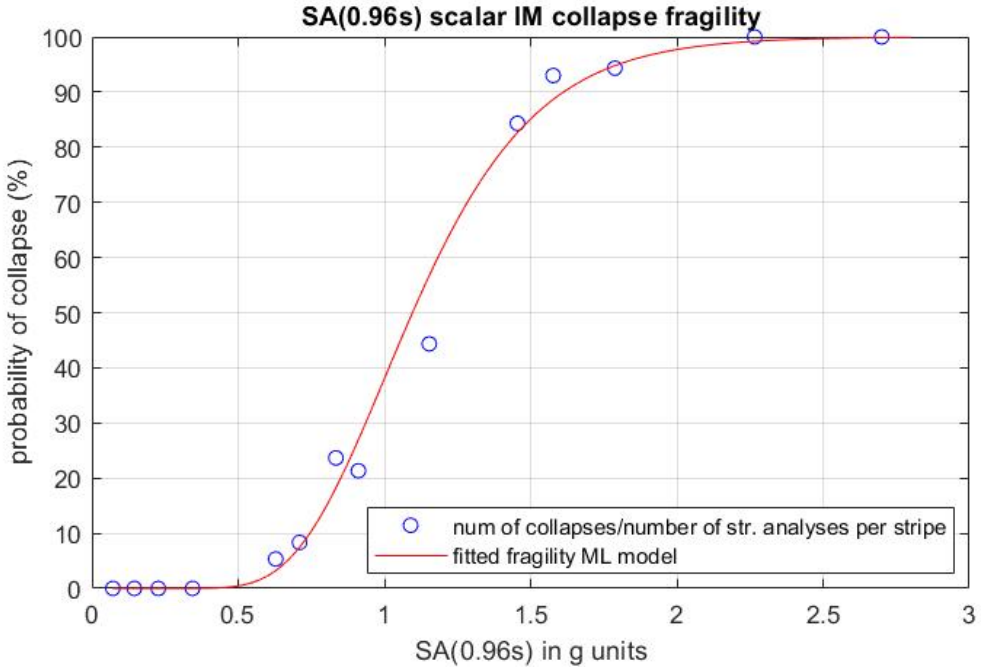

Figure 3.22: Collapse fragility fitted from maximum likelihood method for $SA(T_1 = 0.96s)$ when $CS_{SA(0.96s)}$ is used for record scaling and selection at 14 discrete hazard levels

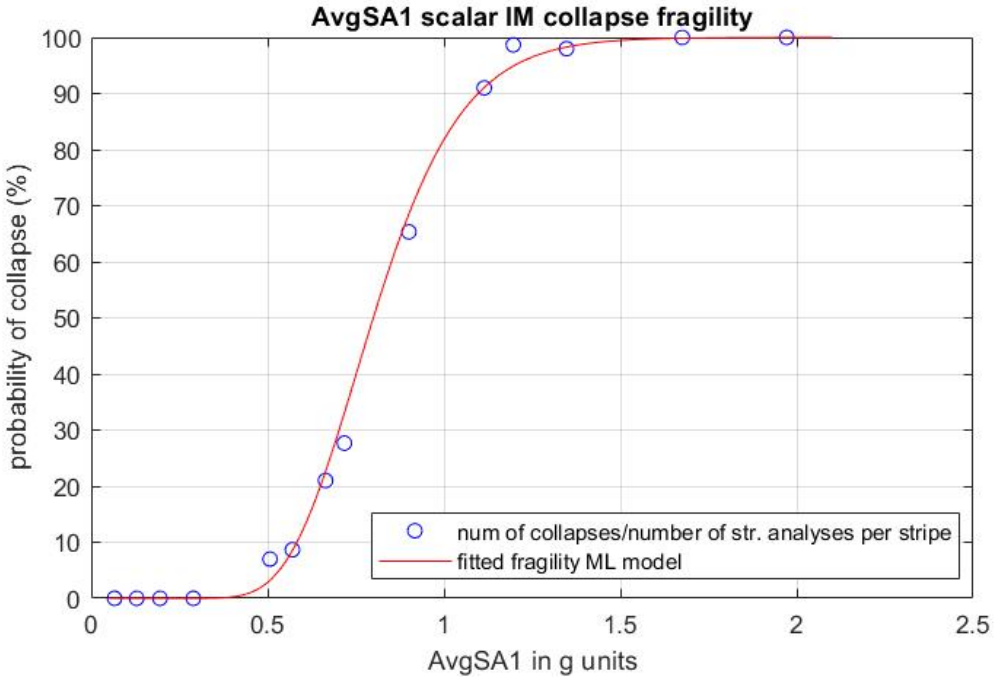

Figure 3.23: Collapse fragility fitted from maximum likelihood method for $AvgSA1$ when CS_{AvgSA1} is used for record scaling and selection at 14 discrete hazard levels

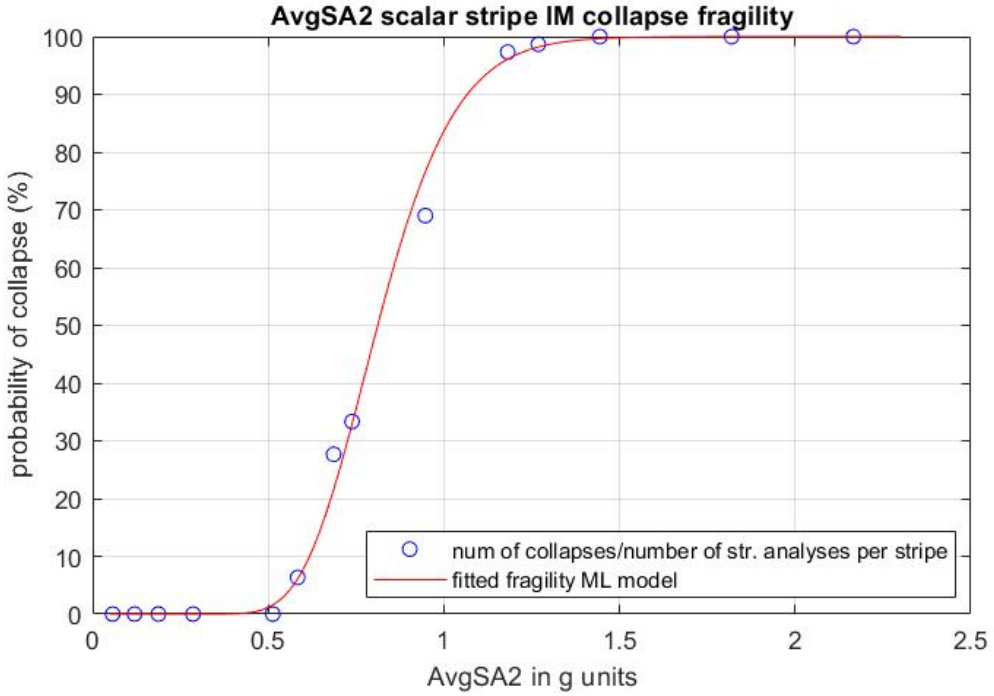

Figure 3.24: Collapse fragility fitted from maximum likelihood method for $AvgSA2$ when CS_{AvgSA2} is used for record scaling and selection at 14 discrete hazard levels

The EDP distribution for the non-collapse data at each stripe $(EDP|IM)$ is assumed to be lognormally distributed with mean $\mu_{\ln(EDP)|IM}(im)$ and standard deviation $\sigma_{\ln(EDP)|IM}(im)$.

These parameters can be computed for each stripe by only considering the non-collapse data. $\mu_{\ln(EDP)|IM}(im)$ and $\sigma_{\ln(EDP)|IM}(im)$ are assumed to vary linearly along the logarithmic IM axis. Given a stripe, if there is no non-collapse case, $\mu_{\ln(EDP)|IM}(im)$ attains value corresponding to collapse limit state ($ln(LSNRD = 1)$) and $\sigma_{\ln(EDP)|IM}(im)$ attains zero value indicating that collapse at that IM level is a sure event. Figures 3.25 – 3.27 display $EDP|IM$ distribution for non-collapse data for the three scalar IMs considered.

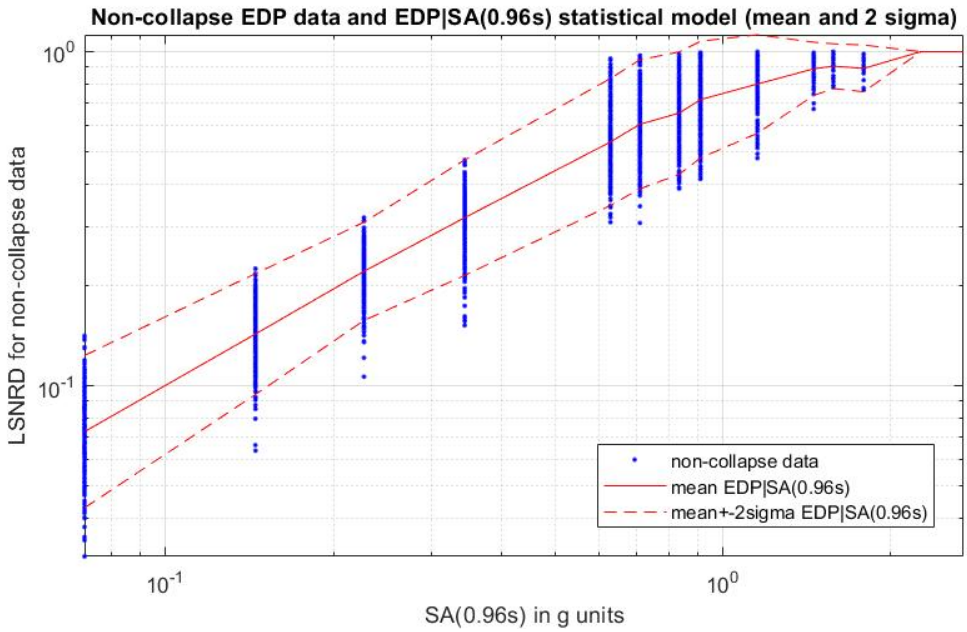

Figure 3.25: Non-collapse $EDP|IM$ data for $SA(T_1 = 0.96s)$ when $CS_{SA(0.96s)}$ is used for record scaling and selection at 14 discrete hazard levels.

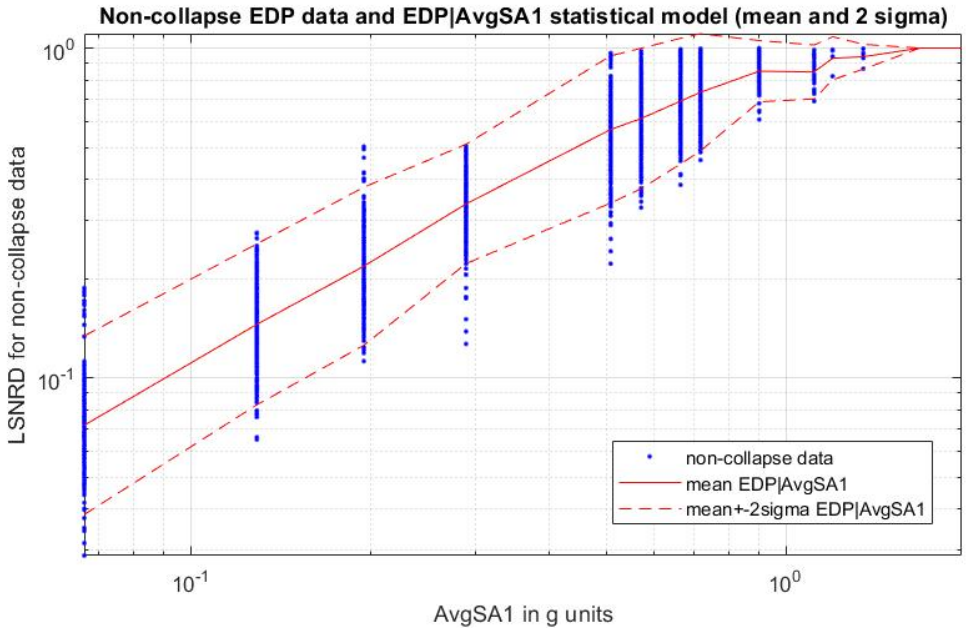

Figure 3.26: Non-collapse $EDP|IM$ data for $AvgSA1$ when CS_{AvgSA1} is used for record scaling and selection at 14 discrete hazard levels.

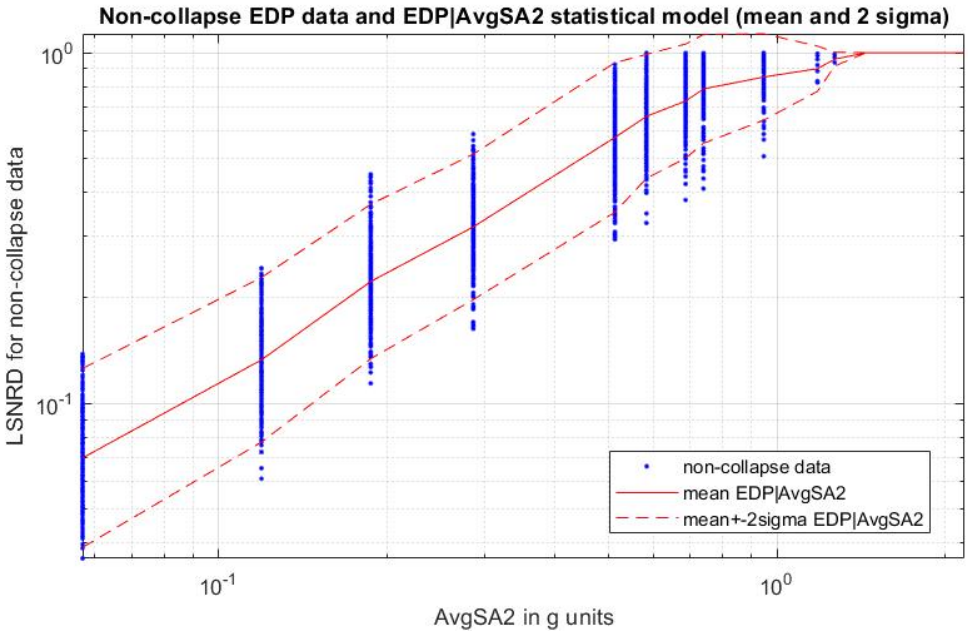

Figure 3.27: Non-collapse $EDP|IM$ data for $AvgSA2$ when CS_{AvgSA2} is used for record scaling and selection at 14 discrete hazard levels.

Since now the collapse fragility and non-collapse $EDP|IM$ probability models are established, $G_{EDP|IM}(y|im)$ can be represented by Eqn. 3.27:

$$G_{EDP|IM}(y|im) = P_C(im) + (1 - P_C(im))\left(1 - \Phi\left(\frac{\ln(y) - \mu_{\ln(EDP)|IM}(im)}{\sigma_{\ln(EDP)|IM}(im)}\right)\right) \quad (3.27)$$

$G_{EDP|IM}(y|im)$ represents the EDP exceedance probability for a threshold y ($EDP > y$) given the ground-motion intensity is $IM = im$. The first term, $P_C(im)$ on the right-hand side is the collapse probability at $IM = im$. In case of collapse, EDP is assumed to be ∞, so it exceeds any threshold y value. $P_C(im)$ is computed by Eqn. 3.20. The second term on the right-hand side represents the non-collapse case. The probability of non-collapse, $(1 - P_C(im))$, is multiplied by the probability that $EDP > y$ given non-collapse ($EDP > y|\bar{C}$). $P(EDP > y|\bar{C})$ is presented by the complementary cumulative distribution function of $EDP|IM$ shown in Figures 3.25 – 3.27.

3.3.6. Comparison of Structural Response Statistical Models by Different IMs

Figure 3.28 presents the comparisons of logarithmic standard deviations of non-collapse data $\sigma_{\ln(EDP)|IM}(im)$ obtained from three different scalar IMs: ($SA(T_1 = 0.96s)$, $AvgSA1$ and $AvgSA2$. The horizontal axis of this plot is in terms of return periods (reciprocal of annual exceedance rates) corresponding to the hazard levels considered in this study. Figure 3.29 shows the collapse probability as a function of return period for each of the three scalar IM methods considered in this study.

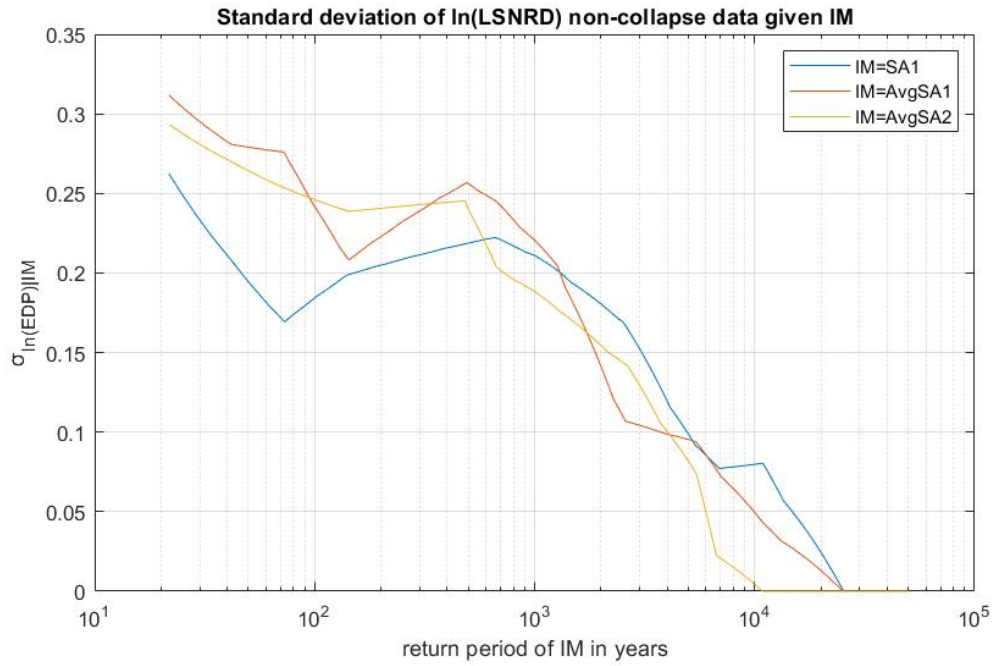

Figure 3.28: Return period dependent variation of logarithmic standard deviations of EDP given IM - $\sigma_{ln(EDP)|IM}$ for three different scalar IM metrics

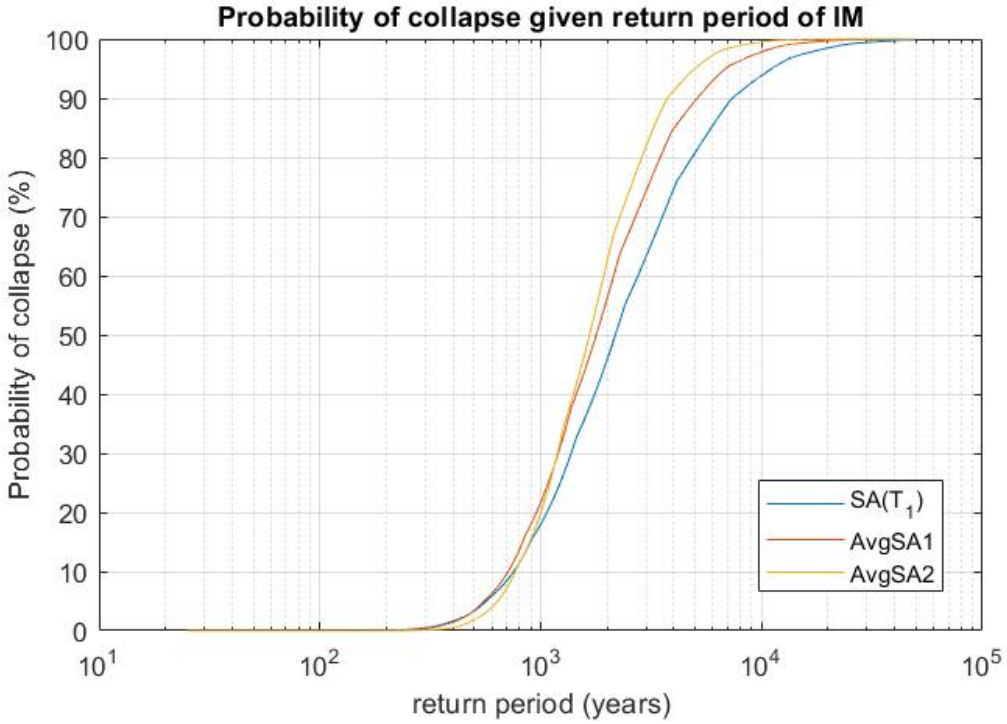

Figure 3.29: Collapse probability of the model building in terms of return period for three different scalar IM metrics

Figure 3.28 indicates that the scalar $SA(0.96s)$ produces the lowest standard deviation in the non-collapse EDP data for small return periods (hence low intensity levels). This is expected

since there is little or no yielding at these low intensity levels, so the structural response is mainly controlled by the elastic state of the structure (i.e., the initial stiffness controls). At large return periods (i.e., higher intensity levels) $AvgSA1$ and $AvgSA2$ yield lower standard deviations because the structural behavior shifts to post-elastic phase and elongated periods come in to play in structural behavior. This behavior results in better performances of the $AvgSA1$ and $AvgSA2$ scalar IM metrics. Note that $AvgSA2$ seems to perform better than $AvgSA1$. Needless to say, the variations in the EDP standard deviations computed from the three scalar IMs result in different collapse probabilities particularly at longer return periods (Figure 3.29).

3.3.7. Evaluation of EDP Exceedance Rate Curves

Since $G_{EDP|IM}(y|im)$ is now defined by Eqn. 3.27, $\lambda_{EDP}(y)$ can be evaluated by Eqn. 3.12. This is done by using the three different scalar IMs. The EDP exceedance rate curves (risk curves) are given in Figure 3.30.

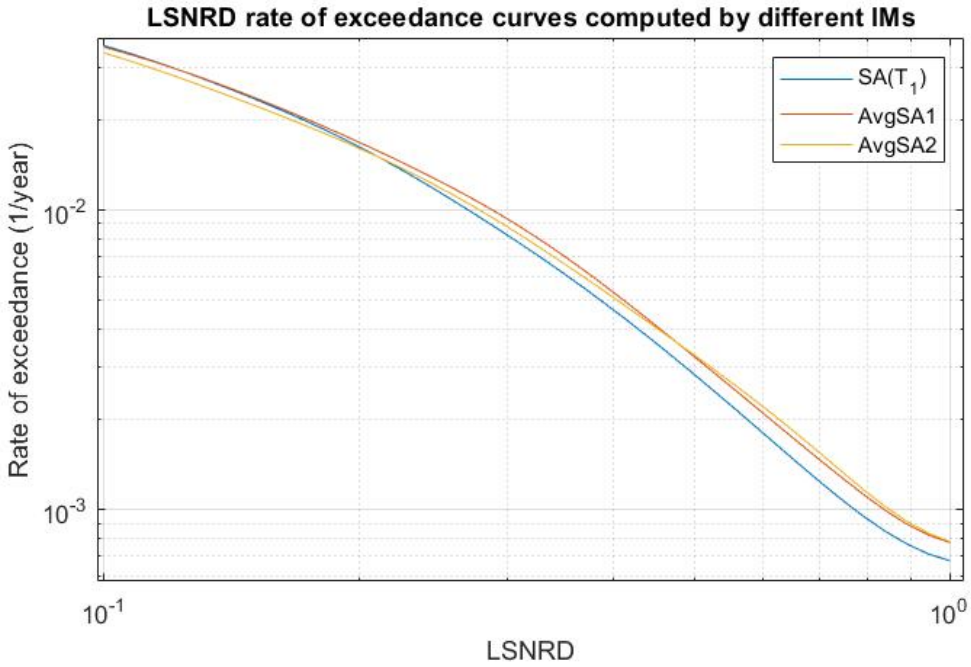

Figure 3.30: EDP exceedance rate curves computed from stripe scaling method and for three different scalar IMs

If the earthquake process as well as the ground motions are assumed to be Poisson processes, the following formulation (Eqn. 3.28) will give the exceedance probability of LSNRD in 50-year exposure time.

$$P_{50\ years}(EDP > y) = 1 - e^{-50\ years\ \lambda_{EDP}(y)} \tag{3.28}$$

Figure 3.31 shows the LSNRD exceedance probability curves in 50-year exposure period computed from stripe scaling method by using three different scalar IMs.

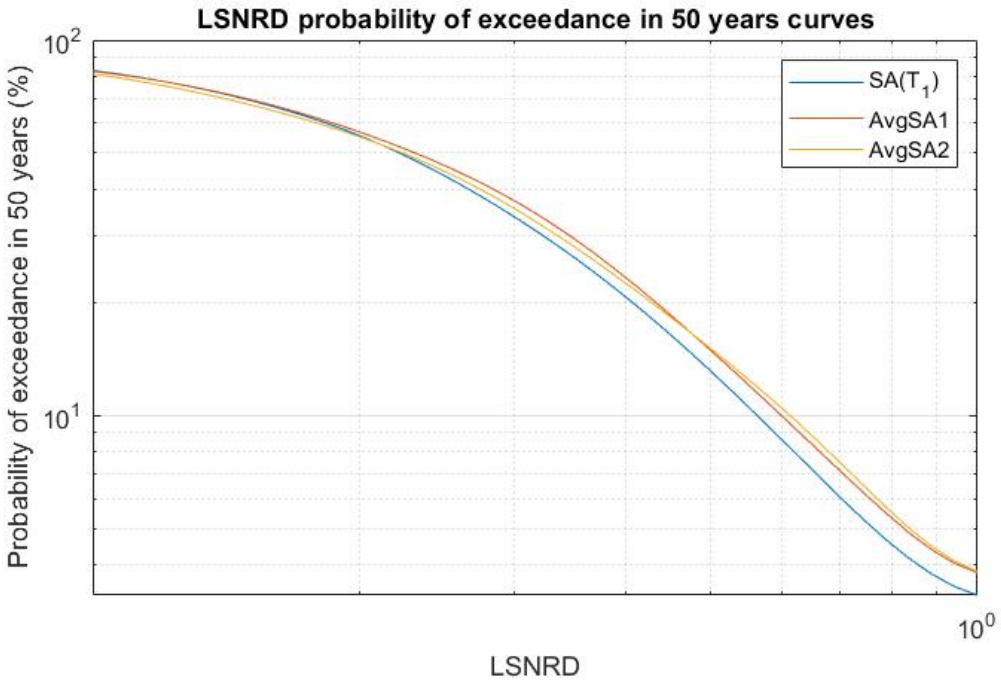

Figure 3.31: LSNRD exceedance probability curves for 50-year exposure period computed by stripe scaling and for three distinct scalar IMs

Figure 3.31 indicates that the exceedance probability of collapse ($LSNRD = 1$) in 50 years changes between 3% to 4% depending on the scalar IM used in stripe scaling. In other words, it is highly unlikely for this model building to experience collapse given the hazard levels and seismicity considered in this study. The exceedance probability of *limited damage* limit state (i.e., $LSNRD = 0.35$) changes between 24% and 28% whereas the exceedance probability of *controlled damage* limit state ($LSNRD = 0.8$) varies between 4.5% and 5.5%. One can also compute the rate of collapse by solving the below integral.

$$\lambda_C = \int_{IM} P_C(im) \left| \frac{d\lambda_{IM}(im)}{dim} \right| dim \tag{3.29}$$

3.3.8. Vector EDP Rate of Exceedance Surface Evaluation

This section illustrates the risk assessment procedure when EDP consists of multiple parameters (i.e., $\overrightarrow{EDP}$). The $\overrightarrow{EDP}$ is considered to account for the two components (x- and y-directions) of the maximum roof displacement. In this case, the $\overrightarrow{EDP}$ rate of exceedance is given in Eqn. 3.13, which is repeated once again as Eqn. 3.30 for convenience:

$$\lambda_{\overrightarrow{EDP}}(\vec{y}) = \int_{IM} G_{\overrightarrow{EDP}|IM}(\vec{y}|im) \left| \frac{d\lambda_{IM}(im)}{dim} \right| dim \tag{3.30}$$

$G_{\overrightarrow{EDP}|IM}(\vec{y}|im)$ represents the probability that $EDP_1 \geq y_1$ and $EDP_2 \geq y_2$. By making analogy with Eqn. 3.27, this term can be expressed as given in Eqn. 3.31.

$$G_{\overrightarrow{EDP}|IM}(\vec{y}|im) = P_C(im) + (1 - P_C(im))\big(P(EDP_1 \geq y_1, EDP_2 \geq y_2|IM = im, \bar{C})\big) \tag{3.31}$$

$P_C(im)$ term is the same as the one used in Eqn. 3.27. The term $P(EDP_1 \geq y_1, EDP_2 \geq y_2|IM = im, \bar{C})$ represents the probability that maximum roof displacements in x-and y-directions exceed the thresholds y_1 and y_2, respectively given $IM = im$, and provided that the building is non-collapsed. This term is modelled as a bivariate normal cumulative distribution function as given in Eqn. (3.32) to (3.37).

$$\overrightarrow{EDP}|IM, \bar{C} \sim \mathcal{N}\left(\vec{\mu}_{EDP|IM}(im), \Sigma_{EDP|IM}(im)\right) \tag{3.32}$$

$$P(EDP_1 \geq y_1, EDP_2 \geq y_2|IM = im, \bar{C}) = \int_{y_1}^{\infty} \int_{y_2}^{\infty} f_{\overrightarrow{EDP}|IM,\bar{C}}(\overrightarrow{edp}|im) d\overrightarrow{edp} \tag{3.33}$$

$$f_{\overrightarrow{EDP}|IM}\left(\overrightarrow{edp}|im\right) = \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{edp}-\boldsymbol{\mu}_{EDP|IM}\right)^{T}\Sigma_{EDP|IM}{}^{-1}\left(\boldsymbol{edp}-\boldsymbol{\mu}_{EDP|IM}\right)\right)}{\sqrt{(2\pi)^{k}\left|\Sigma_{EDP|IM}\right|}} \quad (3.34)$$

k in Eqn. 3.34 is the size of vector $\overrightarrow{EDP}$, in this case $k = 2$. In Eqn. 3.34 and following equations "$, \bar{C}$" is dropped for convenience in order to fit the equations in one line, however, it should be kept in mind that distribution described in Eqn. 3.32 – 3.37 corresponds to only non-collapse structural analyses results. Also, note that multivariate normal distribution described in Eqn. 3.32 – 3.37 is actually a distribution of $ln\left(\overrightarrow{EDP}\right)|IM$, however it is just notified as $\overrightarrow{EDP}|IM$ here in order to avoid congestion in equations. $\vec{\mu}_{EDP|IM}(im)$ is actually $\vec{\mu}_{ln(EDP)|IM}(im)$ and $\Sigma_{EDP|IM}(im)$ is $\Sigma_{ln(EDP)|IM}(im)$. Same analogy is valid for all other terms in Eqn. 3.32 – 3.37.

$$\vec{\mu}_{EDP|IM}(im) = \begin{bmatrix}\mu_{EDP_1|IM}(im)\\ \mu_{EDP_2|IM}(im)\end{bmatrix} \quad (3.35)$$

$$\Sigma_{EDP|IM}(im) = \begin{bmatrix}\left(\sigma_{EDP_1|IM}(im)\right)^{2} & Cov(EDP_1, EDP_2|IM = im)\\ Cov(EDP_1, EDP_2|IM = im) & \left(\sigma_{EDP_2|IM}(im)\right)^{2}\end{bmatrix} \quad (3.36)$$

$$Cov(EDP_1, EDP_2|IM = im) = \rho_{EDP_1,EDP_2|IM}(im)\,\sigma_{EDP_1|IM}(im)\,\sigma_{EDP_2|IM}(im) \quad (3.37)$$

Note that the terms $\mu_{EDP_1|IM}(im)$, $\mu_{EDP_2|IM}(im)$, $\sigma_{EDP_1|IM}(im)$, $\sigma_{EDP_2|IM}(im)$ and $\rho_{EDP_1,EDP_2|IM}(im)$ can be computed from the roof displacements along x- and y- directions determined from structural analyses using records selected for each target intensity level (for each stripe). As stated previously, these parameters are assumed to vary in a piecewise linear way on log scale (they are assumed to vary linearly in-between 14 target IM intensity levels in logarithmic scale). Figure 3.32 displays the maximum roof displacements in x- and y-directions computed from nonlinear SDOF dynamic analyses for 475-year return period intensity level. CS_{AvgSA1} conditional spectrum is used for scaling the ground-motion records for SDOF runs. The logarithmic means, standard deviations and correlation coefficient

parameters designated in the beginning of this paragraph are given in Table 3.4 for this particular case.

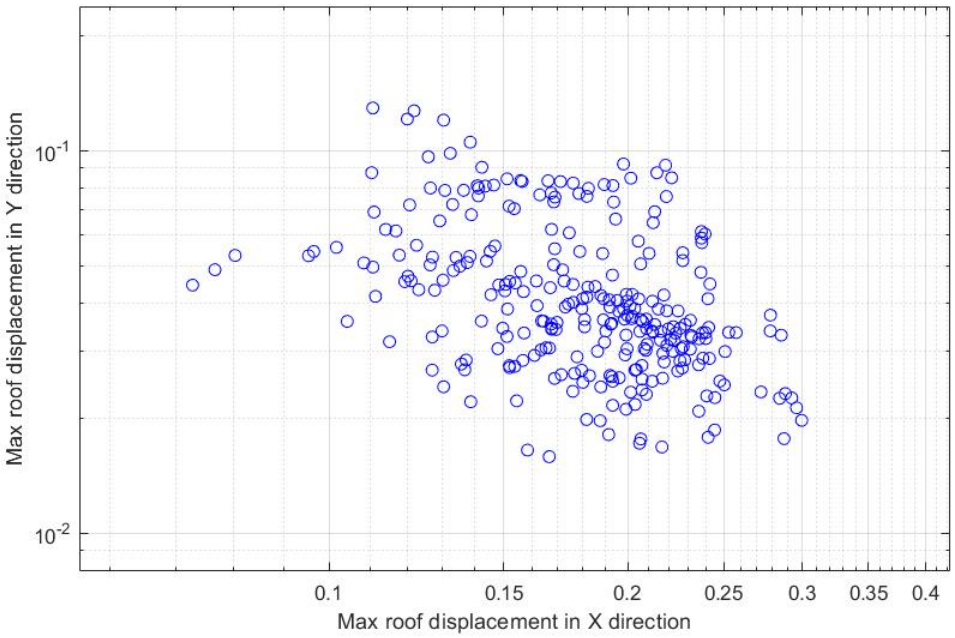

Figure 3.32: Maximum roof displacements along x- and y-directions for 475-year return period intensity level. The target spectrum used for stripe scaling to run nonlinear SDOF dynamic analyses is CS_{AvgSA1}

Table 3.4: Logarithmic means, standard deviations and covariance coefficients of maximum roof displacements in x and y directions for evaluated using structural analyses results from records corresponding to 475-year return period CS_{AvgSA1} spectrum

$\mu_{ln(MaxRD_x)\|AvgSA1}(avgsa1 = 0.5069g)$	-1.7239
$\mu_{ln(MaxRD_y)\|AvgSA1}(avgsa1 = 0.5069g)$	-3.2160
$\sigma_{ln(MaxRD_x)\|AvgSA1}(avgsa1 = 0.5069g)$	0.2530
$\sigma_{ln(MaxRD_y)\|AvgSA1}(avgsa1 = 0.5069g)$	0.4380
$\rho_{ln(MaxRD_x),ln(MaxRD_y)\|AvgSA1}(avgsa1 = 0.5069g)$	-0.4279

$MaxRD_x$ and $MaxRD_y$ represents maximum roof displacement in x and y directions respectively

These five parameters to define the bivariate normal distribution of $\overrightarrow{EDP}|IM$ can be computed at all target intensities after running the nonlinear SDOF analyses of stripe scaled ground motions.

Note that using Eqn. 3.30 – 3.37 in reverse order, one can compute $\lambda_{\overrightarrow{EDP}}(\vec{y})$ for any configuration of $\vec{y}$. Since $\vec{y}$ is a two-component vector, $\lambda_{\overrightarrow{EDP}}(\vec{y})$ represents a surface and it is displayed in Figure 3.33.

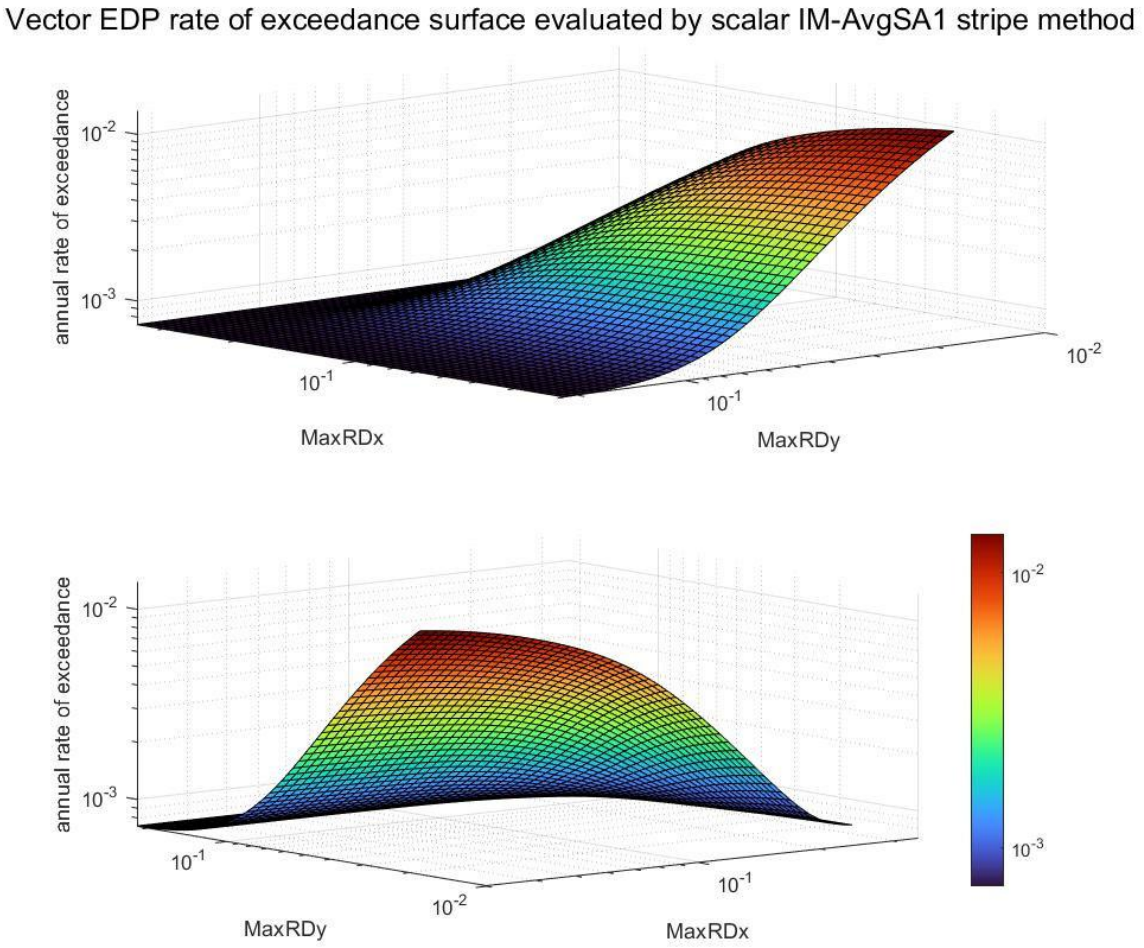

Figure 3.33: $\lambda_{\overrightarrow{EDP}}(\vec{y})$ annual exceedance rate surface for maximum roof displacements in x- and y- directions viewed from two different angles. *MaxRDx* and *MaxRDy* axis labels represent maximum roof displacements in x and y directions respectively

In Figure 3.33, ordinate value (Z coordinate) of every point on the displayed surface represents rate of exceedance of maximum roof displacements in x and y directions (two horizontal coordinates of the considered point). The surface becomes flat near collapse limit states. The annual exceedance rate at these levels is approximately 0.00073, which corresponds to 3.58% of collapse probability in a 50-year exposure time, which is almost identical to the result obtained from simplified scalar EDP (LSNRD) discussed and

presented in Section 3.3.7. Regardless the similarity of concluding outcomes between scalar and vector EDP methods, the latter is superior because loss is computed with lesser uncertainty when conditioned on multiple EDPs rather than a single scalar EDP.

3.4. Scalar IM - Cloud Scaling in the Analyses

3.4.1. Cloud Scaling and Selection of Records

In stripe-scaling, the selected records are scaled to a specific IM level (e.g., conditional acceleration spectrum value at T = 0.96 s) so that the EDPs computed from the nonlinear dynamic analyses are plotted as stripes along that specific target IM level [5]. This way, one can obtain the IM-EDP data. In cloud scaling, the records are not scaled to match some exact IM intensity level, but rather to be scattered across some desired range of IM [5, 8, 34]. To this end, the IM-EDP scatter plots look like a data cloud. This thesis introduces novel cloud scaling method that ensured that selected scaled record set is hazard consistent and has IM values that are uniformly distributed across desired IM range. Uniform distribution of record set across IM range is desired here in order to obtain IM-EDP dataset that represents all hazard levels equally. The cloud scaling methodology and corresponding record selection is explained in the following text.

Firstly, a total of 100 target hazard levels are established starting from 22-year return period (90% exceedance probability in 50 years) to 49975-year return period (0.1% exceedance probability in 50 years). The conditional mean spectra $CMS_{IM}(T)$ are computed for the 100 target hazard levels as explained in Section 2.4. The entire GM records in the pool are firstly scaled to the $CMS_{IM}(T)$ established for the largest target hazard level (i.e., 49975-year return period) and the one giving the minimum SSE for a pre-determined spectral acceleration range for that hazard level is selected and removed from the pool. The same process is continuously repeated for the other target hazard levels following the largest one in descending order. Each time, the GM record that gives the minimum SSE for the target hazard level of concern is removed from the pool. As already indicated, the SSE is always computed for the pre-determined spectral acceleration range. There are various alternatives to determine this spectral acceleration range [11] and in this study it is kept between 0.05s and 1.44s which correspond to lowest and highest period in $\overrightarrow{T_1}$ vector (check Section 3.2).

Note that this cloud scaling method ensures that selected records are hazard consistent by using CMS_{IM} as target spectra for a range of all IM levels (100 IM levels used for scaling). Figure 3.34 shows three cloud scaled records (out of 100) selected to fit $CMS_{SA(0.96s)}(T)$ at their respective IM target levels. As pointed before, there are 100 uniformly distributed IM target levels, as can be seen in Figure 3.34

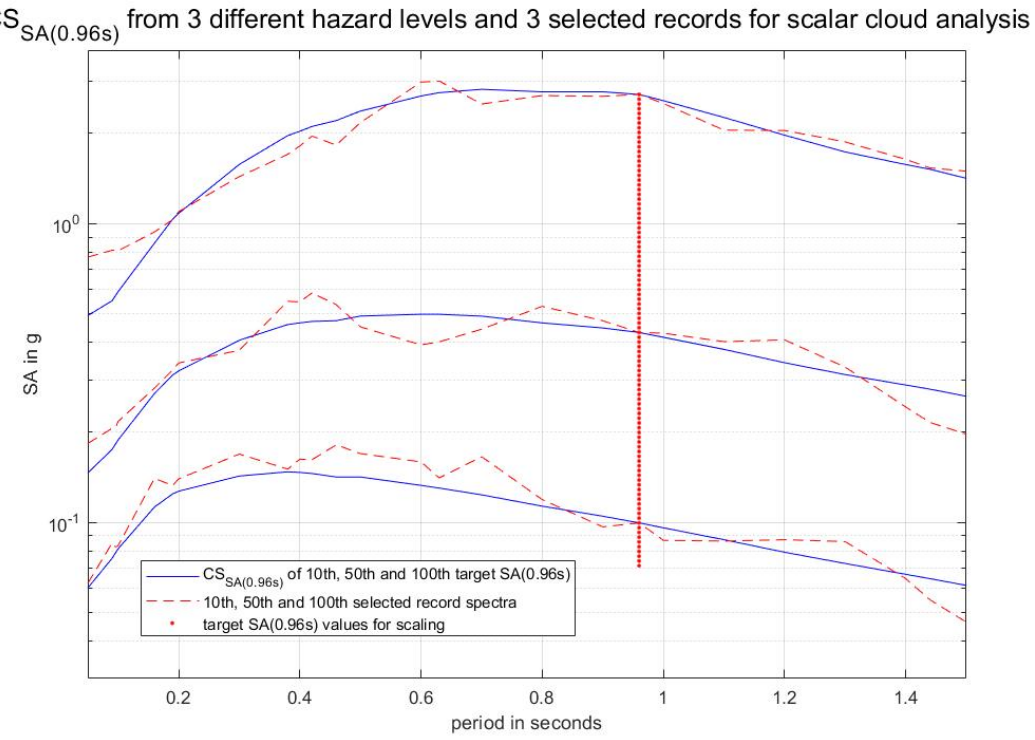

Figure 3.34: Three out of hundred cloud scaled records for a uniformly distributed target $SA(0.96s)$ values spanning wide range of hazard levels (from 22 to 49975-year return period). Records are selected to fit $CMS_{SA(0.96s)}$ spectra at each $SA(0.96s)$ level.

3.4.2. Structural Response Data from SDOF Analyses Using Cloud Scaling

The GM records scaled and selected by the procedure explained in the previous section are used for running SDOF response history analyses for three scalar IMs (i.e., $SA(0.96s)$, $AvgSA1$ and $AvgSA2$) to compute the scalar EDP along the two principal directions (x- and y-directions) of the model building. The scalar EDP is LSNRD as in the case of stripe scaling method. Each record results in 15 LSNRD data points corresponding to 15 angular rotations. The response data (EDP-IM) obtained from the nonlinear SDOF dynamic analysis are

presented in Figures 3.35 – 3.37 for the three scalar IMs that is being used throughout this thesis.

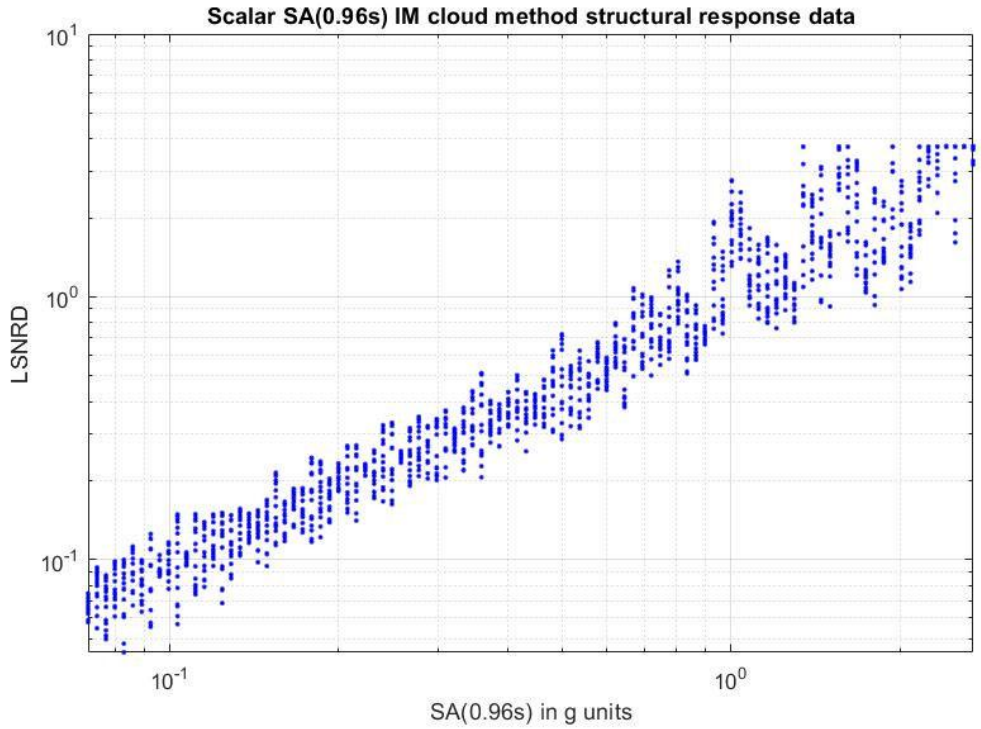

Figure 3.35: Structural response data for $SA(0.96s)$ IM by cloud scaling method

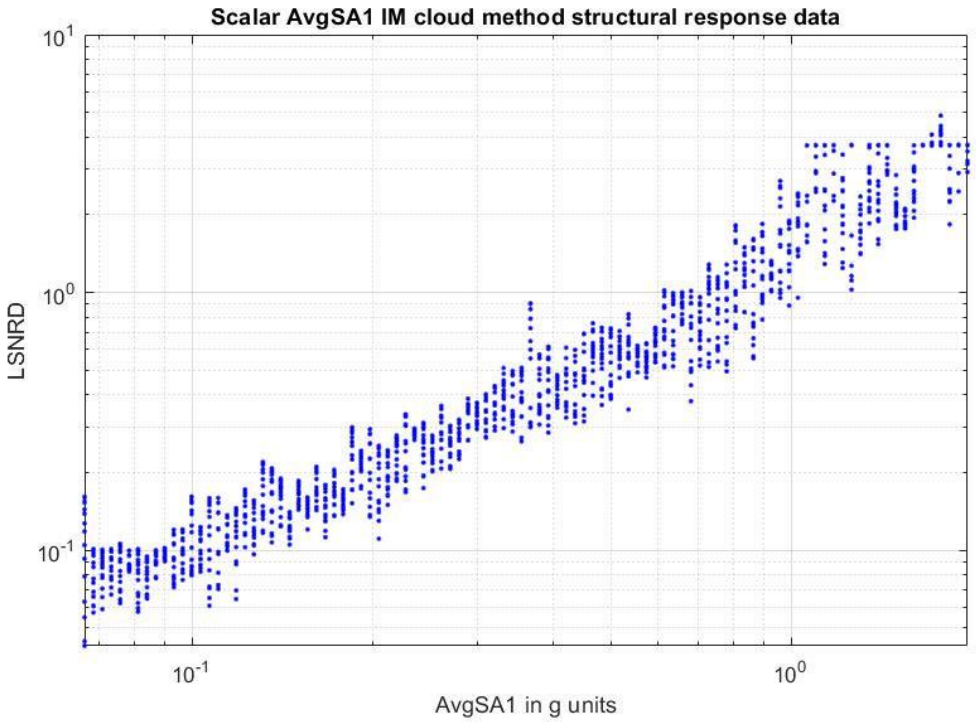

Figure 3.36: Structural response data for $AvgSA1$ IM by cloud scaling method

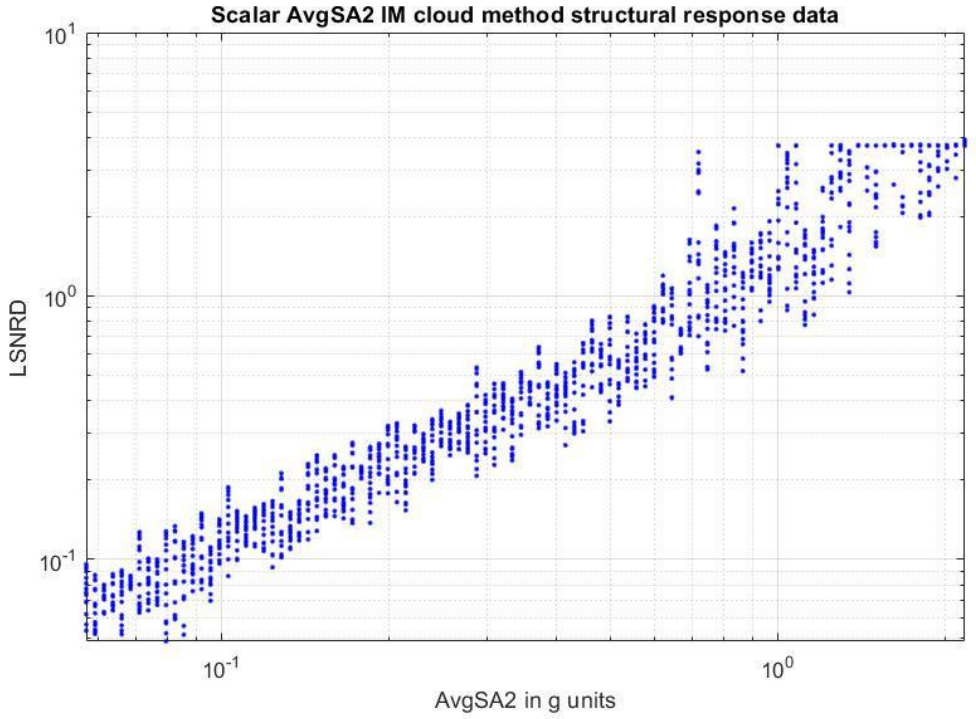

Figure 3.37: Structural response data for $AvgSA2$ IM by cloud scaling method

The response data obtained from cloud scaling are still in stripes because multiple rotations (for fifteen different angles) are done for each ground motion. As in the case of stripe scaling,

the data are separated into collapse data ($LSNRD \geq 1$) and non-collapse data ($LSNRD < 1$). Binary logistic value of 1 is assigned to collapse data points and zero is assigned to non-collapse data points. Figure 3.38 shows a sample plot for this process for $AvgSA1$ scalar IM.

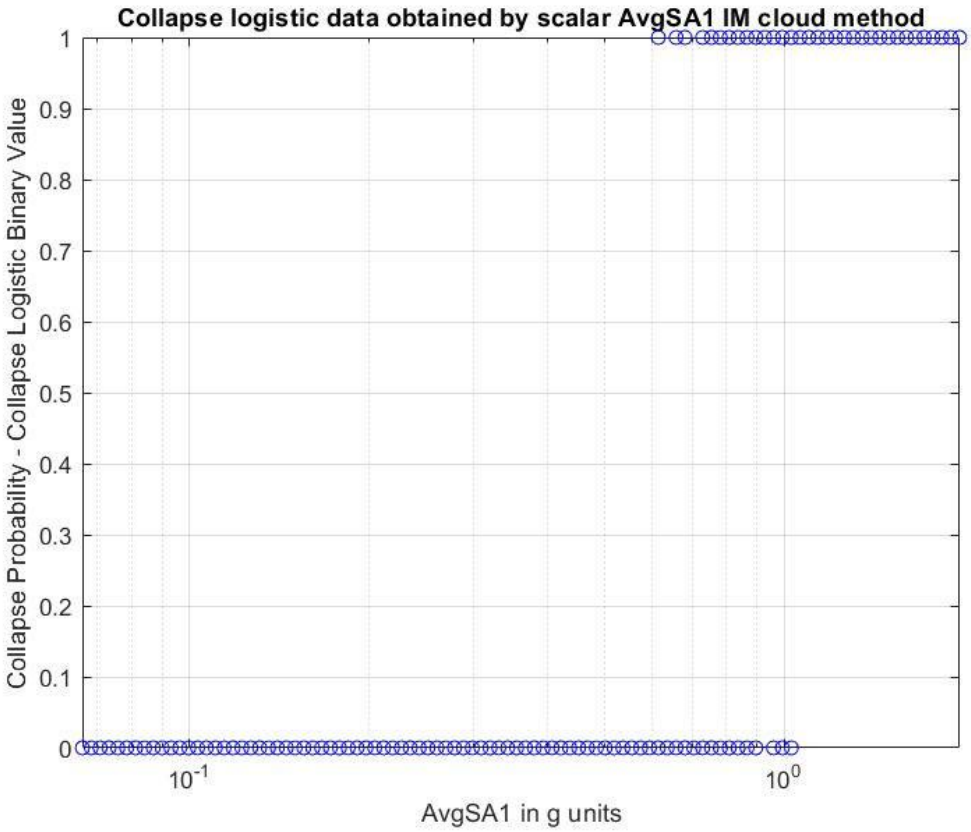

Figure 3.38: Collapse logistic binary data obtained from scalar $AvgSA1$ obtained from cloud scaling

The maximum likelihood method is used to fit the collapse fragility curve to these data. Note that there are a total of 1500 (100 records and 15 rotations per record) data points in Figure 3.38 but most of the data points overlap each other. The collapse probability $P_C(IM)$ is again assumed to have lognormal CDF functional form (Eqn. 3.38).

$$P_C(IM) = \Phi\left(\frac{ln(IM) - \mu_{C-ln(IM)}}{\sigma_{C-ln(IM)}}\right) \tag{3.38}$$

Again, collapse fragility curve is modelled by two parameters, $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$. As with the stripe scaling method, maximum likelihood estimation (MLE) method is undertaken

to estimate these parameters. Readers not familiar with maximum likelihood estimation method are advised to first read Section 3.3.5 of this thesis where the method is introduced. Modeling collapse fragility with results of structural analyses obtained by cloud scaling method is conceptually very similar as with stripe scaling method. However, there are some differences. Again, MLE method is used. Goal is to find configuration of $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ parameters that maximize the likelihood function. Again, likelihood function represents probability of obtaining observed data given the collapse fragility model (Eqn. 3.38). Difference is in observed data. Stripe scaling produces collapse data in the form of the numbers of collapses out of total structural analysis number for each stripe. Cloud scaling collapse data is in the form of logistic binary output for each structural analysis, as shown in Figure 3.38 (collapse or non-collapse – 1 or 0).

Since in cloud scaling method of this thesis 100 records are selected for structural analyses, and since for each record there are 15 structural analyses runs for different ground motion orientations, in total 1500 structural analyses results are obtained. When it comes to collapse, these results produce 1500 binary logistic data units – 1 if analysis outcome is collapse and 0 otherwise (Figure 3.38). If IM_i represents IM value of scaled record used for i'th structural analysis (i goes from 1 to 1500), if B_i represents binary random variable such that $B_i = 1$ if i'th structural analysis causes collapse and $B_i = 0$ otherwise and if b_i represents the observed value of B_i (Figure 3.38), steps to compute $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ parameters numerically by MLE method are:

1) Assume some candidate $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ configuration. Evaluate probability of collapse P_{C-i} for each IM_i by collapse fragility model given in Eqn. 3.38:

$$P_{C-i} = \Phi\left(\frac{ln(IM_i) - \mu_{C-ln(IM)}}{\sigma_{C-ln(IM)}}\right) \tag{3.39}$$

2) Probability of observing single structural analysis output $B_i = b_i$ is given by Eqn. 3.40:

$$P(B_i = b_i) = (P_{C-i})^{b_i}(1 - P_{C-i})^{1-b_i} \tag{3.40}$$

Note that $P(B_i = b_i)$ is equal to P_{C-i} if $b_i = 1$ and $(1 - P_{C-i})$ if $b_i = 0$.

3) Likelihood function represents probability of observing all b_i data:

$$L\left(\mu_{C-ln(IM)}, \sigma_{C-ln(IM)}\right) = \prod_{i=1}^{i=1500} (P_{C-i})^{b_i}(1 - P_{C-i})^{1-b_i} \tag{3.41}$$

4) Steps 1, 2 and 3 are repeated for a wide range of densely spaced $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ configurations. Configuration that produces the highest value of likelihood function L is chosen to represent collapse fragility model.

IM_i term from equations above represents IM value of scaled record used for i'th structural analysis and is not to be confused with IM corresponding to certain hazard level from stripe scaling method which was notified as $IM_{stripe-i}$

Figures 3.39 – 3.41 show the collapse logistic data and computed collapse fragility curves obtained by the maximum likelihood approach for three scalar IMs used in cloud scaling.

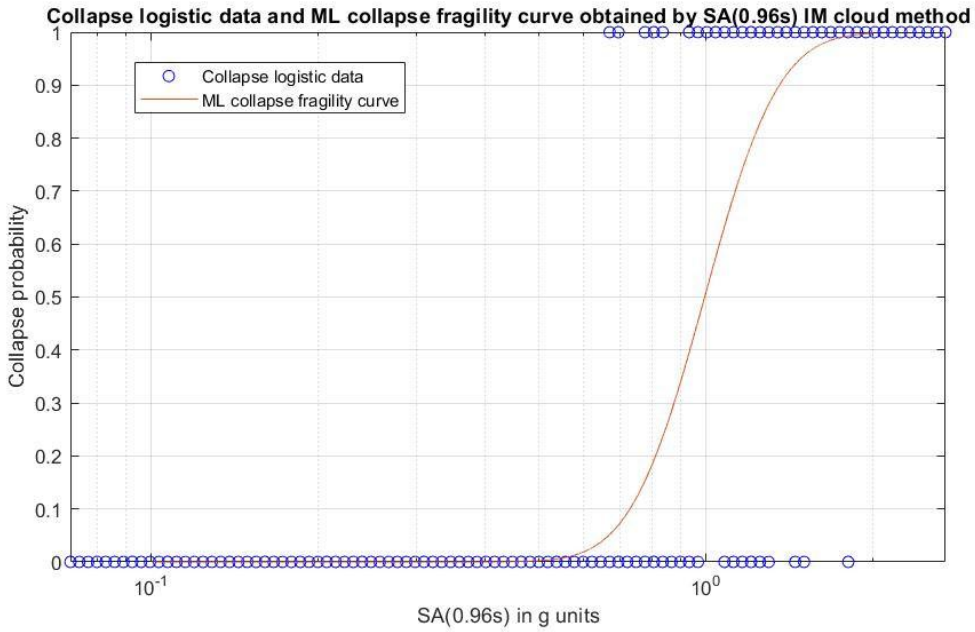

Figure 3.39: Collapse logistic data and the maximum likelihood collapse fragility curve for $SA(0.96s)$ scalar IM used for cloud scaling

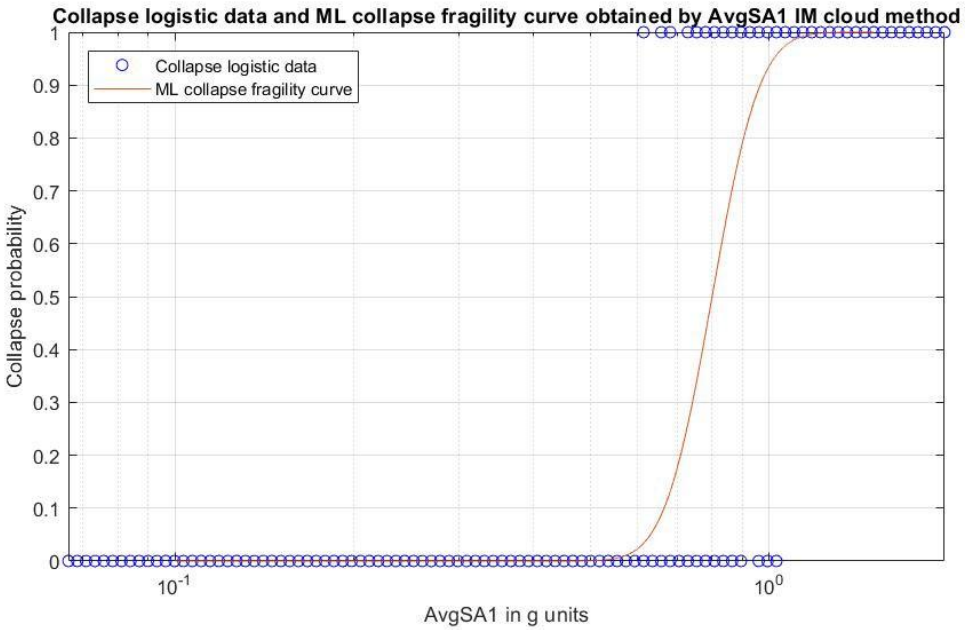

Figure 3.40: Collapse logistic data and the maximum likelihood collapse fragility curve for $AvgSA1$ scalar IM used for cloud scaling

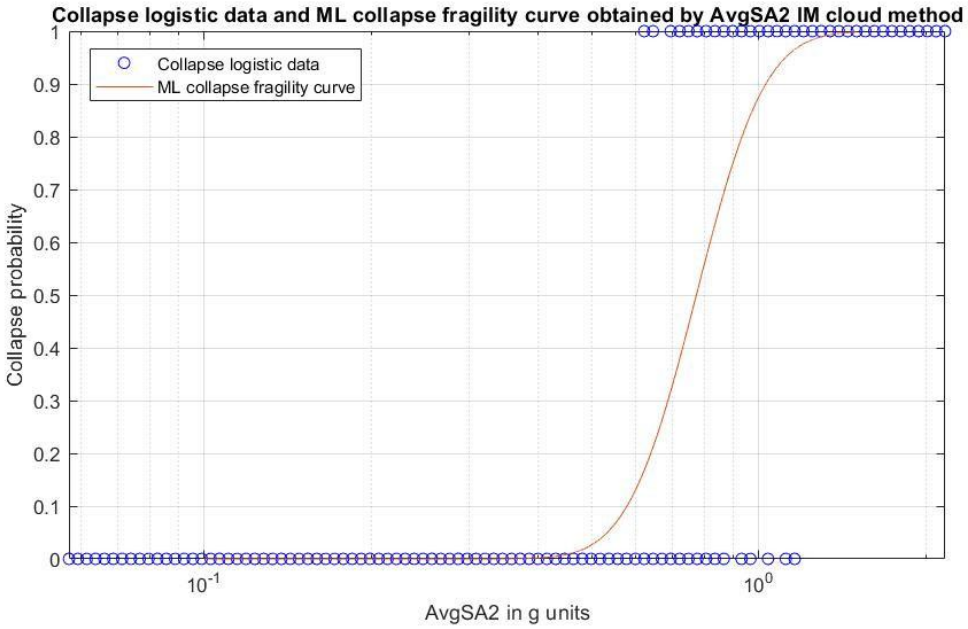

Figure 3.41: Collapse logistic data and the maximum likelihood collapse fragility curve for $AvgSA2$ scalar IM used for cloud scaling

Goodness of fit of these collapse fragility fits can be estimated by procedure explained in a paper by Shinozuka et at. [35]. This paper provides following condition (Eqn. 3.42 – 3.45) to assess the hypothesis that $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ values computed by MLE method are indeed true parameters of fragility curve:

$$P_{y^2} = \Phi\left(\frac{y^2 - \mu_{y^2}}{\sigma_{y^2}}\right) \leq 1 - \alpha \tag{3.42}$$

$$y^2 = \sum_{i=1}^{i=1500} (b_i - P_{C-i})^2 \tag{3.43}$$

$$\mu_{y^2} = \sum_{i=1}^{i=1500} P_{C-i}(1 - P_{C-i}) \tag{3.44}$$

$$\sigma_{y^2} = \sum_{i=1}^{i=1500} P_{C-i}(1 - P_{C-i})\,(1 - 2P_{C-i})^2 \tag{3.45}$$

If condition in Eqn. 3.42 is satisfied, $\mu_{C-ln(IM)}$ and $\sigma_{C-ln(IM)}$ cannot be rejected at the significance level α. In this case, P_{y^2} was computed as 0.598, 0.611 and 0.563 for $SA(0.96s)$, $AvgSA1$ and $AvgSA2$ intensity measure collapse fragilities respectively. This means that α can be accepted as 35% which is significantly higher that its levels proposed in Shinozuka et al. paper [35] (5% - 10%). This indicates that collapse fragility models obtained by MLE statistically represent good fit to observed collapse data.

When it comes to $EDP|IM, \bar{C}$ distribution for non-collapse data, the following probability model is implemented for cloud scaling [5]:

$$ln(EDP)|IM \sim \mathcal{N}\left(\mu_{ln(EDP)|IM}(im), \left(\sigma_{ln(EDP)|IM}\right)^2\right) \tag{3.46}$$

$$\mu_{ln(EDP)|IM}(im) = a_1 \ln(im) + a_0 \tag{3.47}$$

$$\sigma_{ln(EDP)|IM} = c \tag{3.48}$$

$$f_{\ln(EDP)|IM=im}(edp) = \frac{1}{\sigma_{\ln(EDP)|IM}\sqrt{2\pi}}\, exp\left(-\frac{1}{2}\left(\frac{\ln(edp) - \mu_{\ln(EDP)|IM}(im)}{\sigma_{\ln(EDP)|IM}}\right)^2\right) \tag{3.49}$$

"$, \bar{C}$" is dropped from Eqn. 3.46 – 3.49 in order to avoid congestion, however, it should be kept in mind that distribution presented represents non-collapse case. While stripe scaling method assumes parameters $\mu_{ln(EDP)|IM}(im)$ and $\sigma_{ln(EDP)|IM}(im)$ as piecewise linear functions of $ln(im)$, cloud scaling method assumes $\mu_{ln(EDP)|IM}(im)$ as a linear function of $ln(im)$ (Eqn. 3.47) and $\sigma_{ln(EDP)|IM}$ as constant (Eqn. 3.48).

Parameters a_1, a_0 and c from Eqn. 3.47 and 3.48 can be estimated by maximum likelihood method [31, 32]. This time observed data are IM – EDP pairs from non-collapse structural analyses. Let us say number of non-collapse structural analyses is n. This number is less than 1500 since 1500 is the total number of structural analyses and some of these ended in collapse. Let us represent IM value of a scaled record used for i'th non-collapse structural analysis as IM_i and the corresponding EDP as EDP_i. Steps to numerically compute parameters a_1, a_0 and c by maximum likelihood method are:

1) Assume some candidate configuration of a_1, a_0 and c. Using these parameters compute $\sigma_{ln(EDP)|IM}$ and $\mu_{ln(EDP)|IM}(IM_i)$ by Eqn. 3.48 and 3.47 respectively. $\mu_{ln(EDP)|IM}(IM_i)$ should be computed for all IM_i (i goes from 1 to n).
2) Compute probability density of occurrence of i'th observed data by PDF given in Eqn. 3.49:

$$f_{\ln(EDP)|IM_i}(EDP_i) = \frac{1}{\sigma_{\ln(EDP)|IM}\sqrt{2\pi}}\, exp\left(-\frac{1}{2}\left(\frac{\ln(EDP_i) - \mu_{\ln(EDP)|IM}(IM_i)}{\sigma_{\ln(EDP)|IM}}\right)^2\right) \quad (3.50)$$

3) Likelihood function can be computed by multiplying probability densities of all observed data:

$$L(a_1, a_0, c) = \prod_{i=1}^{i=n} f_{\ln(EDP)|IM=IM_i}(EDP_i) \quad (3.51)$$

4) Repeat steps 1, 2 and 3 for wide range of densely spaced a_1, a_0 and c configurations. Configuration that produces highest L value is chosen to represent $EDP|IM, \bar{C}$ model.

$EDP|IM$ non-collapse data and the maximum likelihood models as defined by Eqn. 3.46 – 3.49 are displayed in Figures 3.42 – 3.44 for the three scalar IMs.

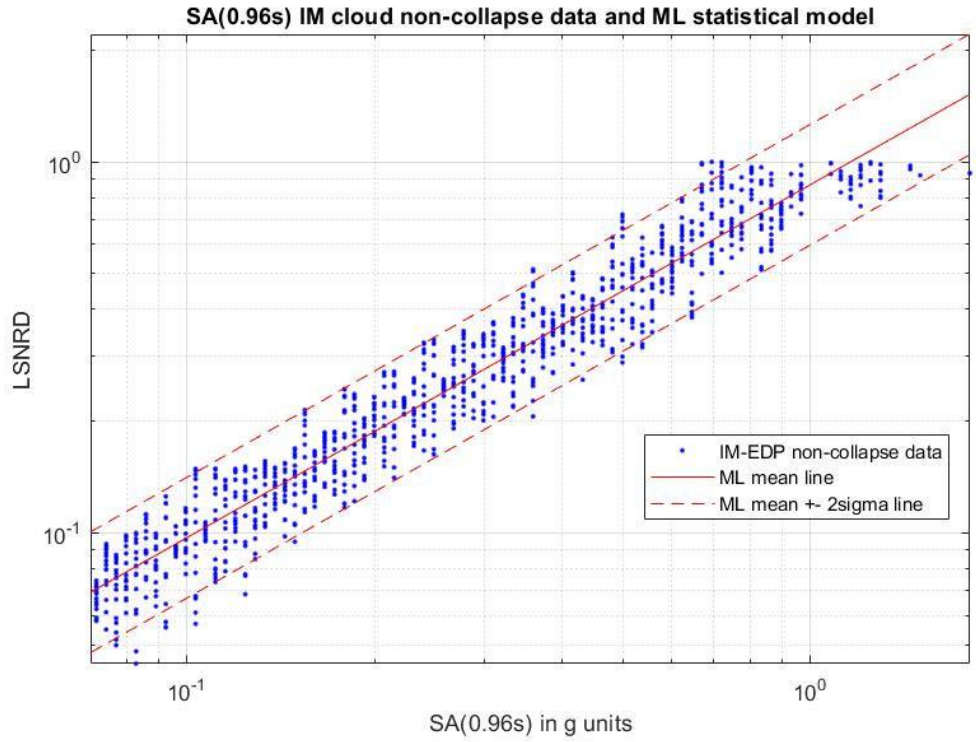

Figure 3.42: Non-collapse IM - EDP data from cloud scaling using $SA(0.96s)$ as IM and $EDP|IM$ maximum likelihood model

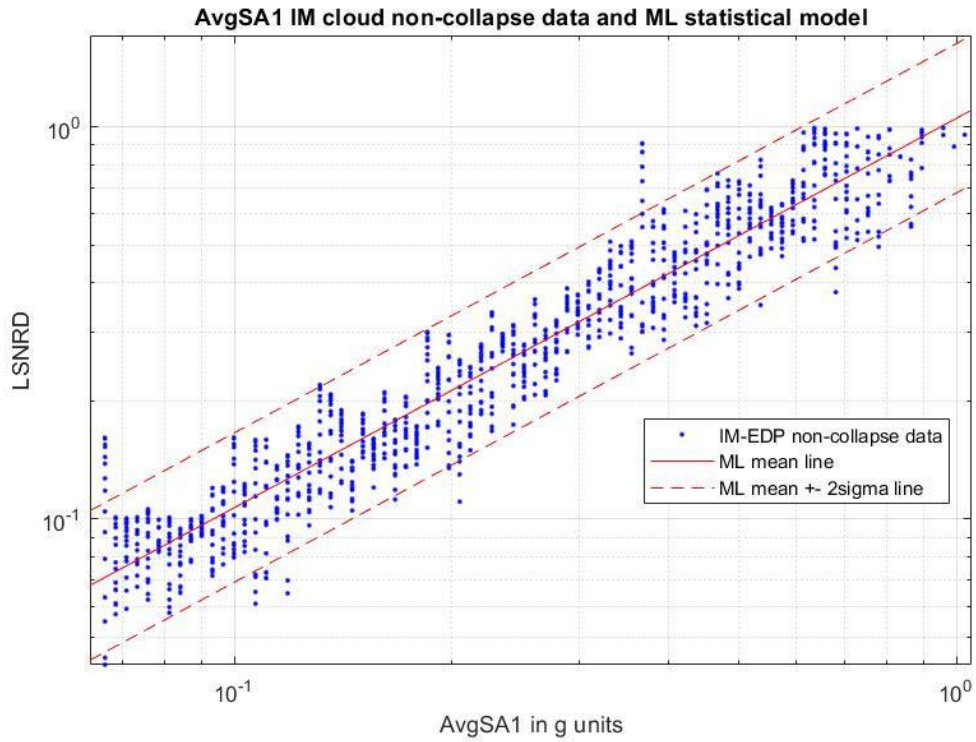

Figure 3.43: Non-collapse IM - EDP data from cloud scaling using $AvgSA1$ as IM and $EDP|IM$ maximum likelihood model

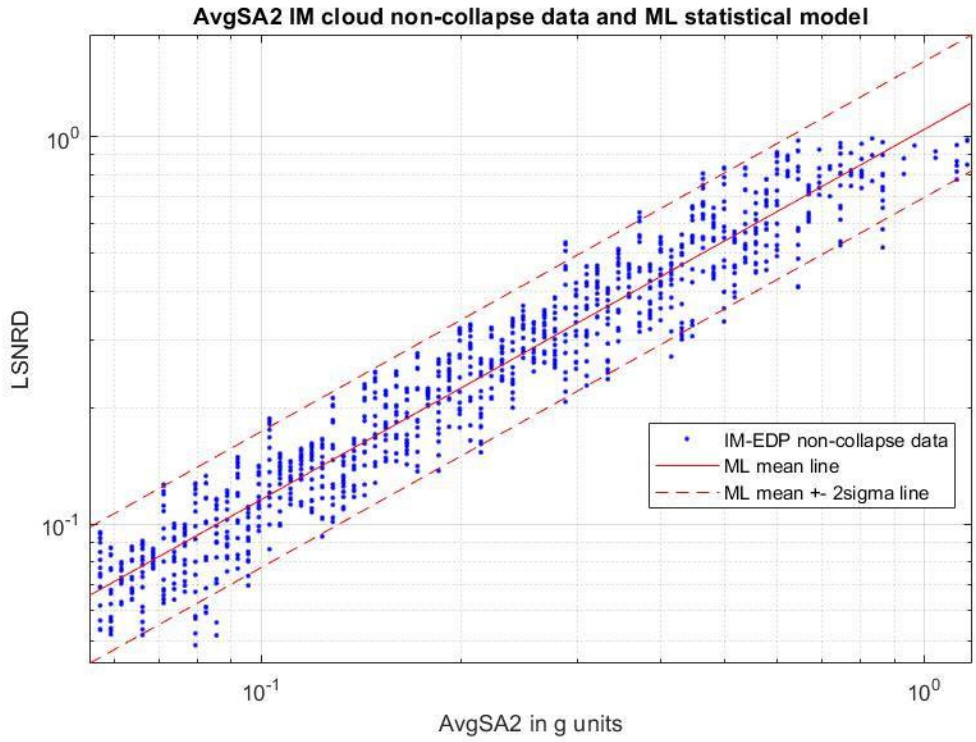

Figure 3.44: Non-collapse IM - EDP data from cloud scaling using $AvgSA2$ as IM and $EDP|IM$ maximum likelihood model

$\sigma_{\ln(EDP)|IM}$ values obtained by $SA(0.96s)$, $AvgSA1$ and $AvgSA2$ are 0.1872, 0.2191 and 0.2021 respectively. It seems that the scalar $SA(0.96s)$ is superior as it produces minimum dispersion for the non-collapse $EDP|IM$ data. The reason behind this observation is that $\sigma_{\ln(EDP)|IM}$ is accepted as constant in this case and the performance of the $SA(0.96s)$ scalar IM is superior for the IM ranges that produce little to no yielding, which is already discussed in the previous subsections. The reason behind the smaller $\sigma_{\ln(EDP)|IM}$ for this scalar could also be the fact that the GM records used for dynamic analyses are selected according to their agreement to the target CMS_{IM}, which omits the variance (dispersion) in the conditional spectrum, CS_{IM}. For a better understanding of these discussions, Figure 3.45 compares $CS_{SA(0.96s)}$ and CS_{AvgSA1} spectrum for 475-year return period hazard level.

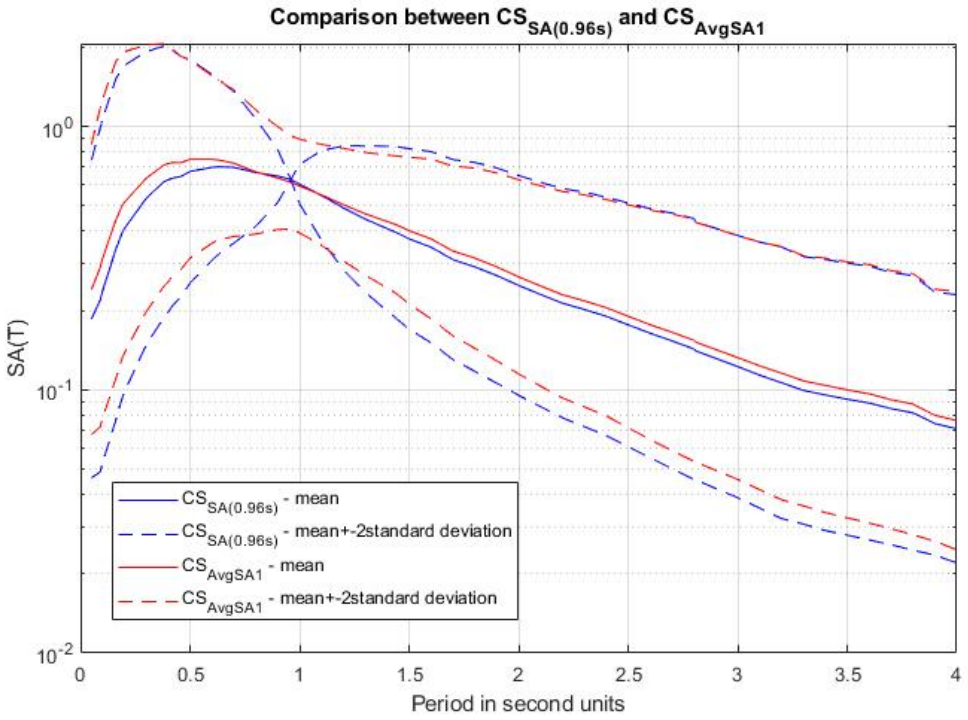

Figure 3.45: Comparison of $CS_{SA(0.96s)}$ and CS_{AvgSA1} for 475-year return period hazard level

As it can be seen from Figure 3.45, although $CS_{SA(0.96s)}$ has lower variance than CS_{AvgSA1} within narrow period range ($0.7 < T < 1.2$), CS_{AvgSA1} depicts lower level of variance (dispersion) within a wider range of periods ($T < 0.7s$ and $T > 1.2s$), which affect higher modes as well as post-yielding period shifted structural response. In stripe scaling method where variance of CS_{IM} is considered in record selection, this feature of $AvgSA$ intensity

measures makes them superior to $SA(T_1)$, in particular for intensity levels that cause high nonlinearity in structural response. As it is already stated, the cloud scaling disregards this variability by considering only CMS_{IM} for record selection. Disregarding dispersion in CS_{IM} takes away one of the main advantages of $AvgSA$ over $SA(T_1)$ which is lower dispersion of CS_{IM} over wider range of periods ($T < 0.7s$ and $T > 1.2s$).

Figure 3.46 shows the collapse probabilities plotted as a function of return period of the three scalar IMs in this thesis.

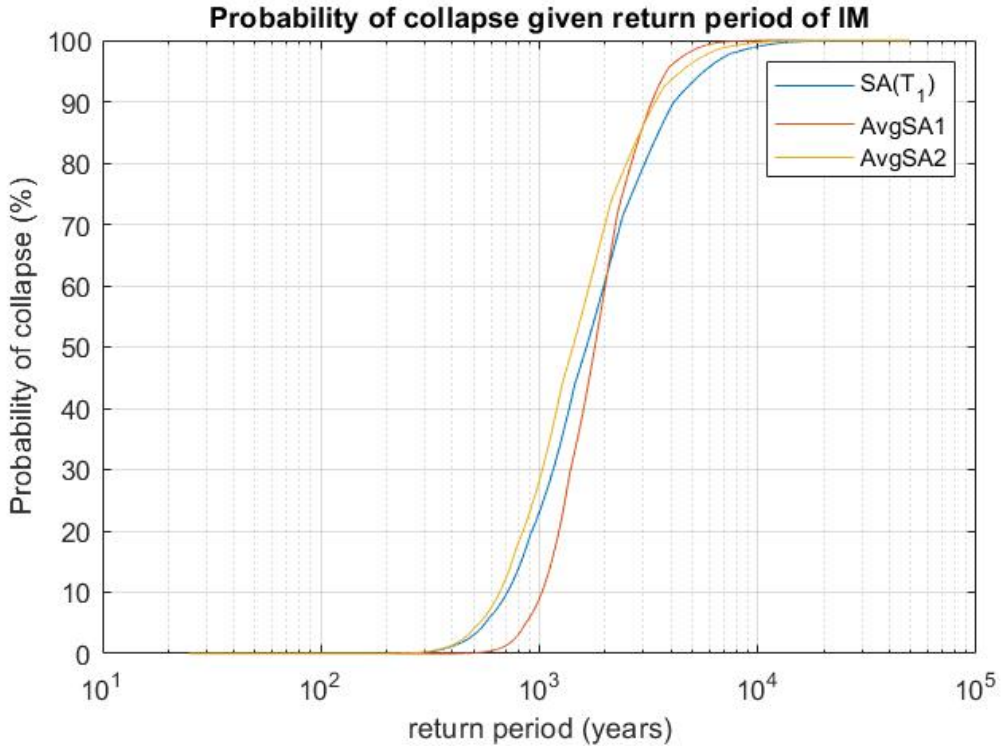

Figure 3.46: Collapse probability evaluated by scalar IM cloud method plotted against IM's return period

Figure 3.46 indicates that collapse is better predicted by $AvgSA1$ than $SA(0.96s)$ because the collapse probabilities computed by $AvgSA1$ raise from 0 to 100% probability within a shorter return period range than the collapse probabilities computed by $SA(0.96s)$ scalar IM. This means that structural collapse is more sharply "recognized" by $AvgSA1$ than by $SA(0.96s)$. Surprisingly, in this case $AvgSA1$ turns out to be superior to $AvgSA2$ which gives only slightly better results than $SA(0.96s)$. This was not the case in stripes analysis where $AvgSA2$ was shown to be the most superior of the three.

3.4.3. Evaluation of EDP Exceedance Rate (Risk) Curves

The EDP exceedance rate curves (risk curves) are computed by the same mathematical formulation as implemented in the stripe scaling method and it is repeated below for convenience.

$$\lambda_{EDP}(y) = \int_{IM} G_{EDP|IM}(y|im) \left| \frac{d\lambda_{IM}(im)}{dim} \right| dim \tag{3.52}$$

$$G_{EDP|IM}(y|im) = P_C(im) + (1 - P_C(im)) \left(1 - \Phi \left(\frac{\ln(y) - \mu_{\ln(EDP)|IM}(im)}{\sigma_{\ln(EDP)|IM}} \right) \right) \tag{3.53}$$

Computation of $P_C(im)$, $\mu_{\ln(EDP)|IM}(im)$ and $\sigma_{\ln(EDP)|IM}$ is already explained in the previous section. The computed $\lambda_{EDP}(y)$ are converted to exceedance probabilities for an exposure time of 50 years by Eqn. 3.28 (Poisson process assumption). These results are displayed in Figure 3.47.

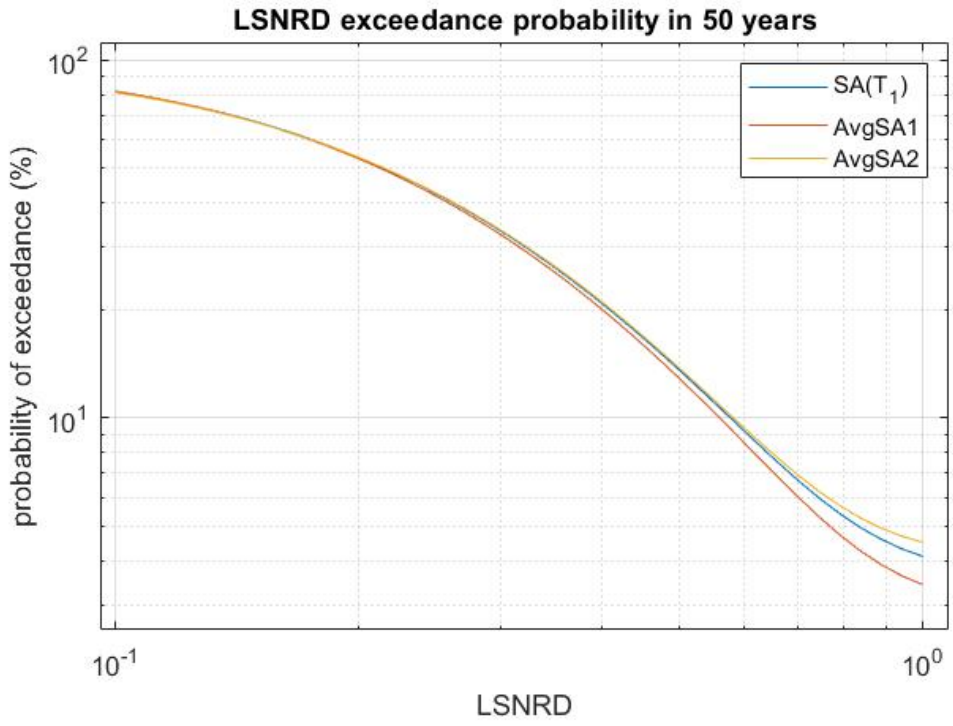

Figure 3.47: EDP exceedance probability curves (risk curves) for an exposure period of 50 years computed for three scalar IMs from cloud scaling method

Figure 3.47 shows that the three scalar IMs (hence the three risk curves) yield similar exceedance probabilities for small LSNRD levels (i.e., light-to-moderate damage levels). However, as the model building is reaching to collapse (LSNRD values near 1), the curves are diverging from each other. This is expected as the three scalar IMs predict slightly different collapse risk as it is noted from Figure 3.46. The collapse rates can be computed from Eqn. 3.29, which is repeated here once again for convenience.

$$\lambda_C = \int_{IM} P_C(im) \left| \frac{d\lambda_{IM}(im)}{dim} \right| dim \tag{3.54}$$

The λ_C is converted to collapse exceedance probability for a 50-year exposure time as 3.68%, 3.04% and 4.13% for $SA(0.96s)$, $AvgSA1$ and $AvgSA2$ scalar IMs, respectively (when cloud scaling is implemented for record selection and scaling). Note that these results are not exactly equal to those given in Figure 3.47 when LSNRD = 1. For example, probability of exceedance of $LSNRD > 1$ in 50 years exposure time computed by using $SA(0.96s)$ scalar IM cloud scaling is 4.02% (Figure 3.47) while it is computed as 3.68% by Eqn. 3.54. This is because Eqn. 3.52 becomes equivalent to Eqn. 3.54 only when y tends to ∞, not when it reaches 1. The term $(1 - P_C(im))\left(1 - \Phi\left(\frac{\ln(y) - \mu_{\ln(EDP)|IM}(im)}{\sigma_{\ln(EDP)|IM}}\right)\right)$ tends to zero when y converges to ∞. Figure 3.48 shows the same curves as Figure 3.47 but computed for $LSNRD$ values up to 3. Here, it can be seen that curves asymptotically approach to the collapse probabilities computed by Eqn. 3.54. Note that values of $LSNRD > 1$ do not have any physical meaning because structure collapses at $LSNRD = 1$. The purpose of this discussion is just to show why collapse probabilities obtained by Eqn. 3.54 (λ_C is converted to collapse probability in 50 years by Eqn. 3.28) are not exactly equal to rates of exceedance of $LSNRD > 1$ given in Figure 3.47 (computed by Eqn. 3.52 – 3.53).

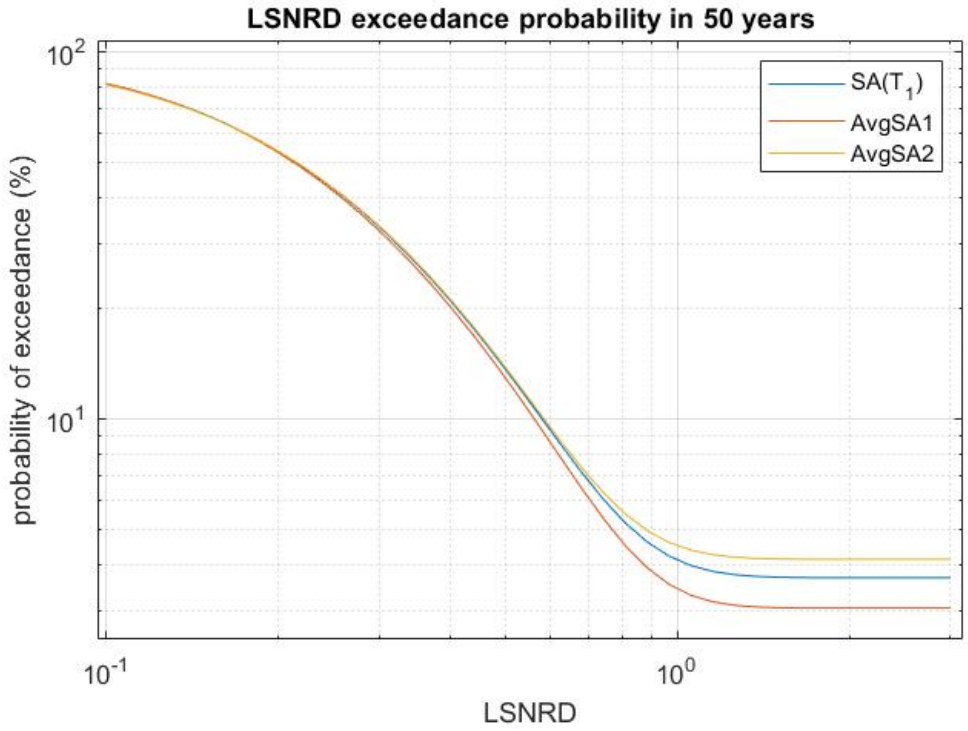

Figure 3.48: EDP probability of exceedance in 50 years curves evaluated by scalar IM cloud method for 3 IMs (evaluated to LSNRD=3)

This mathematical inconsistency is less obvious in stripe scaling method. Figure 3.49 compares the EDP exceedance probability curves for 50-year exposure time that are computed from stripe and cloud scaling methods (using the same IM in both cases - $AvgSA1$). As it can be seen, the risk curve computed from stripe scaling method reaches to an asymptote state earlier from the one computed by cloud scaling. In stripe scaling $\sigma_{\ln(EDP)|IM}(im)$ term becomes zero for high IM levels where all $EDP|IM$ data points collapse. This makes $\left(1-\Phi\left(\frac{\ln(y)-\mu_{\ln(EDP)|IM}(im)}{\sigma_{\ln(EDP)|IM}(im)}\right)\right)$ term equal to zero for $y>1$ at these high IM levels, forcing the exceedance probability curve to reach to the asymptote state faster.

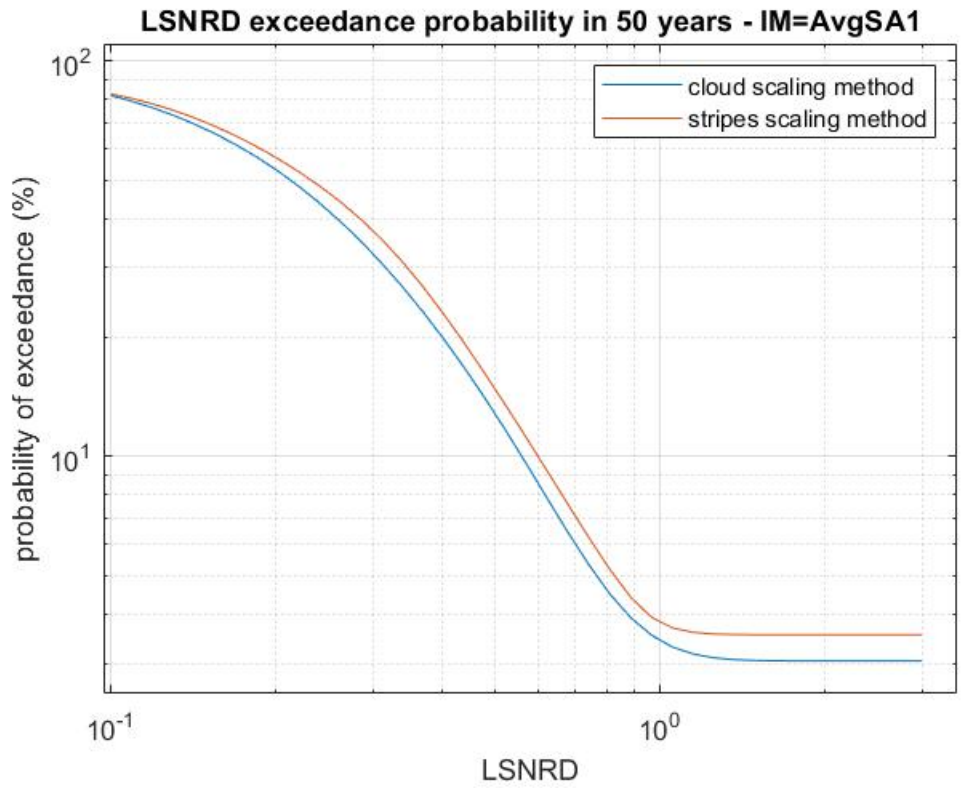

Figure 3.49: Comparisons between the EDP (LSNRD) exceedance probabilities for a 50-year exposure time that are computed by stripe and cloud scaling methods (scalar IM is AvgSA1)

3.5. Vector IM Cloud Method

The vector $\overrightarrow{IM}$ seismic risk assessment methods have already been described by Baker (2007) [5]. This part of the thesis demonstrates the implementation of vector $\overrightarrow{IM}$ for cloud scaling.

By making analogy to Eqn. 3.52, the mathematical formulation of risk curve for vector $\overrightarrow{IM}$ is given below in Eqn. 3.55.

$$\lambda_{EDP}(y) = \int_{\overrightarrow{IM}} G_{EDP|\overrightarrow{IM}}(y|\overrightarrow{im}) MRD_{\overrightarrow{IM}}(\overrightarrow{im}) d\overrightarrow{im} \tag{3.55}$$

$G_{EDP|\overrightarrow{IM}}(y|\overrightarrow{im})$ term represents the probability that $EDP > y$ given that ground motion intensity of $\overrightarrow{IM} = \overrightarrow{im}$ has occurred. $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ is analogous to $\left|\frac{d\lambda_{IM}(im)}{dim}\right|$ term in the scalar IM methods and represents the mean rate density of occurrence of ground motion intensity

of $\overrightarrow{IM} = \overrightarrow{im}$. Its computation through seismic hazard assessment has already been explained in Section 2.5.

The vector cloud scaling method mathematically models the term $G_{EDP|\overrightarrow{IM}}(y|\overrightarrow{im})$ as:

$$G_{EDP|\overrightarrow{IM}}(y|\overrightarrow{im}) = P_C(\overrightarrow{im}) + (1 - P_C(\overrightarrow{im}))\left(1 - \Phi\left(\frac{\ln(y) - \mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{im})}{\sigma_{\ln(EDP)|\overrightarrow{IM}}}\right)\right) \tag{3.56}$$

where $P_C(\overrightarrow{im})$ is vector-based collapse probability and it is given by:

$$P_C(\overrightarrow{im}) = P_C|\left(\overrightarrow{IM} = \overrightarrow{im}\right) = P_C|(IM_1 = im_1, IM_2 = im_2, \dots IM_n = im_n) \tag{3.57}$$

In the above expression n is the number of components in vector $\overrightarrow{IM}$. The logarithmic mean and standard deviation terms $\mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{im})$ and $\sigma_{\ln(EDP)|\overrightarrow{IM}}$ are modelled by expressions below:

$$\mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{im}) = a_0 + a_1 ln(im_1) + a_2 ln(im_2) + \cdots + a_n ln(im_n) \tag{3.58}$$

$$\sigma_{\ln(EDP)|\overrightarrow{IM}} = c \tag{3.59}$$

$P_C(\overrightarrow{im})$ term represents collapse fragility, and it is modelled by the functional form of multivariate normal cumulative distribution function of vector $\overrightarrow{IM}$:

$$P_C(\overrightarrow{im}) = F_{\overrightarrow{IM}_C}(\overrightarrow{im}) \tag{3.60}$$

$$ln\left(\overrightarrow{IM}_C\right) \sim \mathcal{N}\left(\vec{\mu}_{C-\ln(IM)}, \Sigma_{C-\ln(IM)}\right) \tag{3.61}$$

The subscript C is added to $\overrightarrow{IM}_C$ to signify that $\mathcal{N}\left(\vec{\mu}_{C-\ln(IM)}, \Sigma_{C-\ln(IM)}\right)$ is not distribution of $\overrightarrow{IM}$ by itself, but it is rather a functional form to represent $P_C(\overrightarrow{im})$. The components of $\vec{\mu}_{C-\ln(IM)}$ and $\Sigma_{C-\ln(IM)}$ are estimated by maximum likelihood method by using binary collapse logistic data from structural response results. This is explained briefly in the following text.

Let the number of data points (number of structural analyses) be k and let the number of components in vector $\overrightarrow{IM}$ be n. If $IM_{1/i}, IM_{2/i} \ldots, IM_{n/i}$ represent IMs of the ground motion used in the ith structural analysis ($\overrightarrow{IM_i}$) and if b_i represents a binary output of 1 if ith structural analysis caused collapse and zero otherwise (B_i is random binary collapse variable and b_i is observed data from structural analyses), steps to numerically compute components of $\vec{\mu}_{C-\ln(IM)}$ and $\Sigma_{C-\ln(IM)}$ by maximum likelihood function are:

1) Assume some candidate configuration of $\vec{\mu}_{C-\ln(IM)}$ and $\Sigma_{C-\ln(IM)}$. For example, if $\overrightarrow{IM}$ has two components, 5 parameters will have to be assumed (two components of $\vec{\mu}_{C-\ln(IM)}$, two standard deviations and one correlation coefficient to describe $\Sigma_{C-\ln(IM)}$). Compute collapse probability corresponding to all $\overrightarrow{IM_i}$ data points (i goes from 1 to k – number of structural analyses):

$$ln\left(\overrightarrow{IM_C}\right) \sim \mathcal{N}\left(\vec{\mu}_{C-\ln(IM)}, \Sigma_{C-\ln(IM)}\right) \tag{3.62}$$

$$P_C\left(\overrightarrow{IM_i}\right) = F_{\overrightarrow{IM_C}}\left(\overrightarrow{IM_i}\right) \tag{3.63}$$

2) Given $P_C\left(\overrightarrow{IM_i}\right)$, probability of occurrence of observed collapse data $B_i = b_i$ is:

$$P(B_i = b_i) = \left(P_C\left(\overrightarrow{IM_i}\right)\right)^{b_i} \left(1 - P_C\left(\overrightarrow{IM_i}\right)\right)^{1-b_i} \tag{3.64}$$

Compute $P(B_i = b_i)$ for all i's (from 1 to k).

3) Compute likelihood function by multiplying $P(B_i = b_i)$ for all i's:

$$L\left(\vec{\mu}_{C-\ln(IM)}, \Sigma_{C-\ln(IM)}\right) = \sum_{i=1}^{i=k} \left(P_C\left(\overrightarrow{IM_i}\right)\right)^{b_i} \left(1 - P_C\left(\overrightarrow{IM_i}\right)\right)^{1-b_i} \tag{3.65}$$

4) Repeat steps 1, 2 and 3 for wide range of densely spaced $\vec{\mu}_{C-\ln(IM)}$ and $\Sigma_{C-\ln(IM)}$ configurations (vary 5 parameters that describe them). Configuration that maximizes L function is selected to represent collapse fragility model from Eqn. 3.60 – 3.61.

When it comes to the modelling of $\mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{im})$ and $\sigma_{\ln(EDP)|\overrightarrow{IM}}$ from Eqn. 3.58 – 3.59, parameters describing them (a_0, a_1, $a_2 \dots a_n$ and c) are also computed by the maximum likelihood method using non-collapse $EDP|\overrightarrow{IM}$ data obtained from the structural analyses. The following text explains their computation.

Let the number of $EDP|\overrightarrow{IM}$ non-collapse data points (number of non-collapsed structural analyses) be j and let the number of components in vector $\overrightarrow{IM}$ be n. If $IM_{1/i}, IM_{2/i} \dots, IM_{n/i}$ represent IMs of the ground motion used in the ith non-collapsed structural analysis ($\overrightarrow{IM_i}$) and let the EDP_i represent corresponding EDP output from structural analysis. Steps to numerically compute parameters a_0, a_1, $a_2 \dots a_n$ and c by maximum likelihood estimation method are:

1) Assume some candidate configuration of a_0, a_1, $a_2 \dots a_n$ and c. Given this configuration, use Eqn. 3.59 and 3.58 to compute $\sigma_{\ln(EDP)|\overrightarrow{IM}}$ and $\mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{IM_i})$ for all $\overrightarrow{IM_i}$ vectors ($\overrightarrow{IM}$ vectors corresponding to scaled records used for non-collapse structural analyses):

$$\sigma_{\ln(EDP)|\overrightarrow{IM}} = c \tag{3.66}$$

$$\mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{IM_i}) = a_0 + a_1 ln(IM_{1/i}) + a_2 ln(IM_{2/i}) + \cdots + a_n ln(IM_{n/i}) \tag{3.67}$$

2) For each $EDP|\overrightarrow{IM}$ non-collapse data point, compute probability density to obtain observed data EDP_i from structural analyses:

$$f_{\ln(EDP)|\overrightarrow{IM}=\overrightarrow{IM}_k}(\ln(EDP_k)) = \frac{1}{\sigma_{\ln(EDP)|\overrightarrow{IM}}\sqrt{2\pi}}\, exp\left(-\frac{1}{2}\left(\frac{\ln(EDP_k) - \mu_{\ln(EDP)|\overrightarrow{IM}}(\overrightarrow{IM_k})}{\sigma_{\ln(EDP)|\overrightarrow{IM}}}\right)^2\right) \tag{3.68}$$

3) Compute the likelihood function by multiplying probability densities of all $EDP|\overrightarrow{IM}$ non-collapse observed data:

$$L(a_0, a_1, a_2 \dots a_n, c) = \prod_{i=1}^{i=j} f_{\ln(EDP)|\overrightarrow{IM}=\overrightarrow{IM}_k}(\ln(EDP_k)) \tag{3.69}$$

4) Repeat steps 1, 2 and 3 for wide range of densely spaced a_0, a_1, $a_2 \dots$ a_n and c configurations. Configuration that maximizes L function is chosen to represent $EDP|\overrightarrow{IM}$ non-collapse model (Eqn. 3.58 – 3.59).

After modeling the terms $P_C(\overrightarrow{im})$, $\mu_{\ln(EDP)|\overrightarrow{IM}}$ and $\sigma_{\ln(EDP)|\overrightarrow{IM}}$, Eqn. 3.55 and 3.56 can be used to obtain $\lambda_{EDP}(y)$.

If the components of $\overrightarrow{IM}$ are selected to cover $SA(T)$ at a wide range of densely spaced periods that are important for the subject building model (for example, from $0.2T_1$ to $2T_1$), the ground motion records used for structural analysis do not have to be selected by considering their consistency with CS_{IM}. This is because the term $MRD_{\overrightarrow{IM}}(\overrightarrow{im})d\overrightarrow{im}$ ensures hazard consistency. The GM records with response spectra disregarding hazard consistency will have low $MRD_{\overrightarrow{IM}}(\overrightarrow{im})d\overrightarrow{im}$ value that decreases their contribution to $\lambda_{EDP}(y)$. Note that the GM records should be chosen such that they cover as wide range of $\overrightarrow{IM}$ as possible in order to obtain the most reliable $P_C(\overrightarrow{im})$, $\mu_{\ln(EDP)|\overrightarrow{IM}}$ and $\sigma_{\ln(EDP)|\overrightarrow{IM}}$ models.

This thesis presents a sample vector $\overrightarrow{IM}$ cloud scaling method results for the model building and the seismic hazard computed for the fictitious site in Istanbul. As already explained in Section 2.5, $\overrightarrow{IM}$ is considered as $\overrightarrow{IM} = [SA(0.96s) \quad SA(1.44s)]^T$. The unscaled records were selected for structural analyses. The records are selected to span as wide range of $SA(0.96s)$ and $SA(1.44s)$ as possible. In other words, records are selected to span as wide area as possible on $SA(0.96s) - SA(1.44s)$ plane. Figures 3.50 – 3.52 display the computed results.

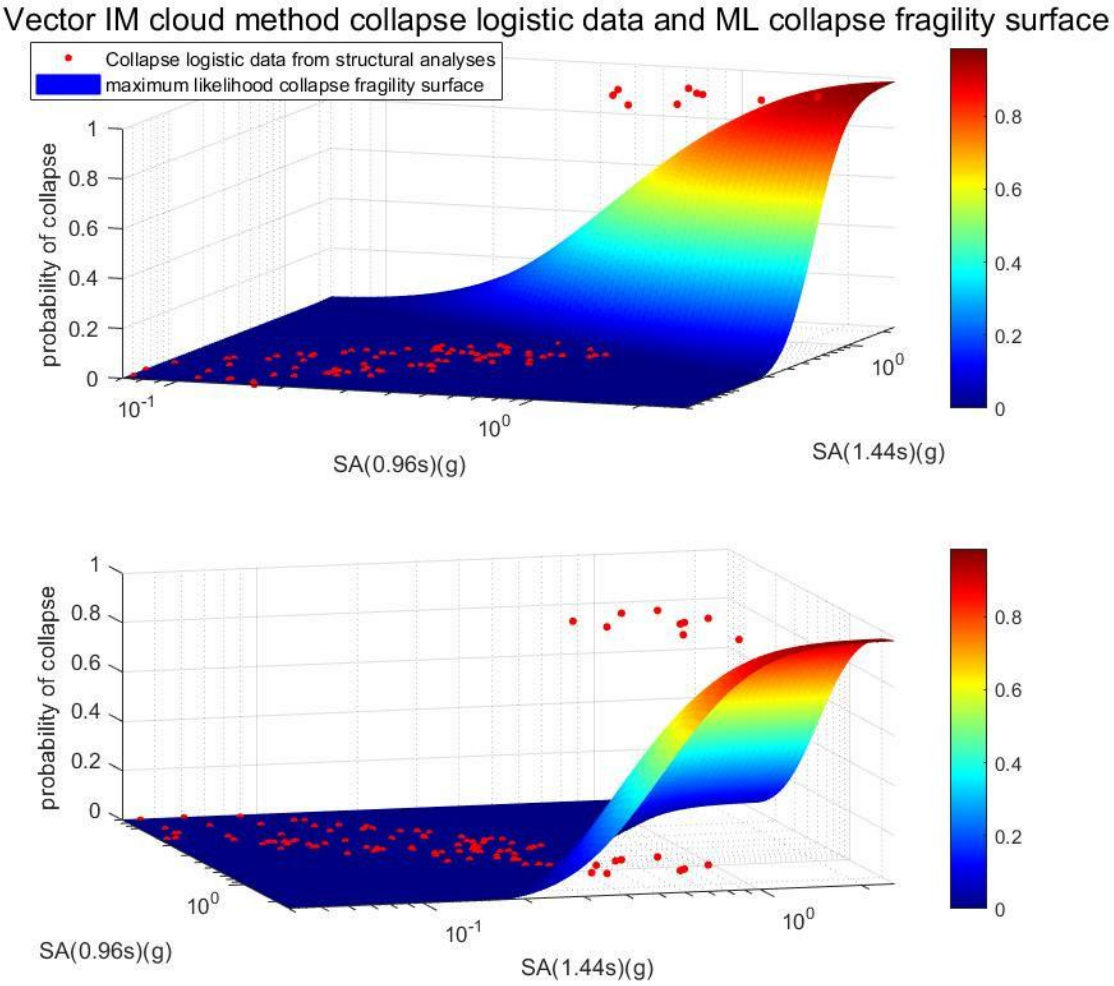

Figure 3.50: Vector $\overrightarrow{IM}$ cloud method from collapse logistic binary data and the maximum likelihood $P_C(\overrightarrow{im})$ collapse fragility surface

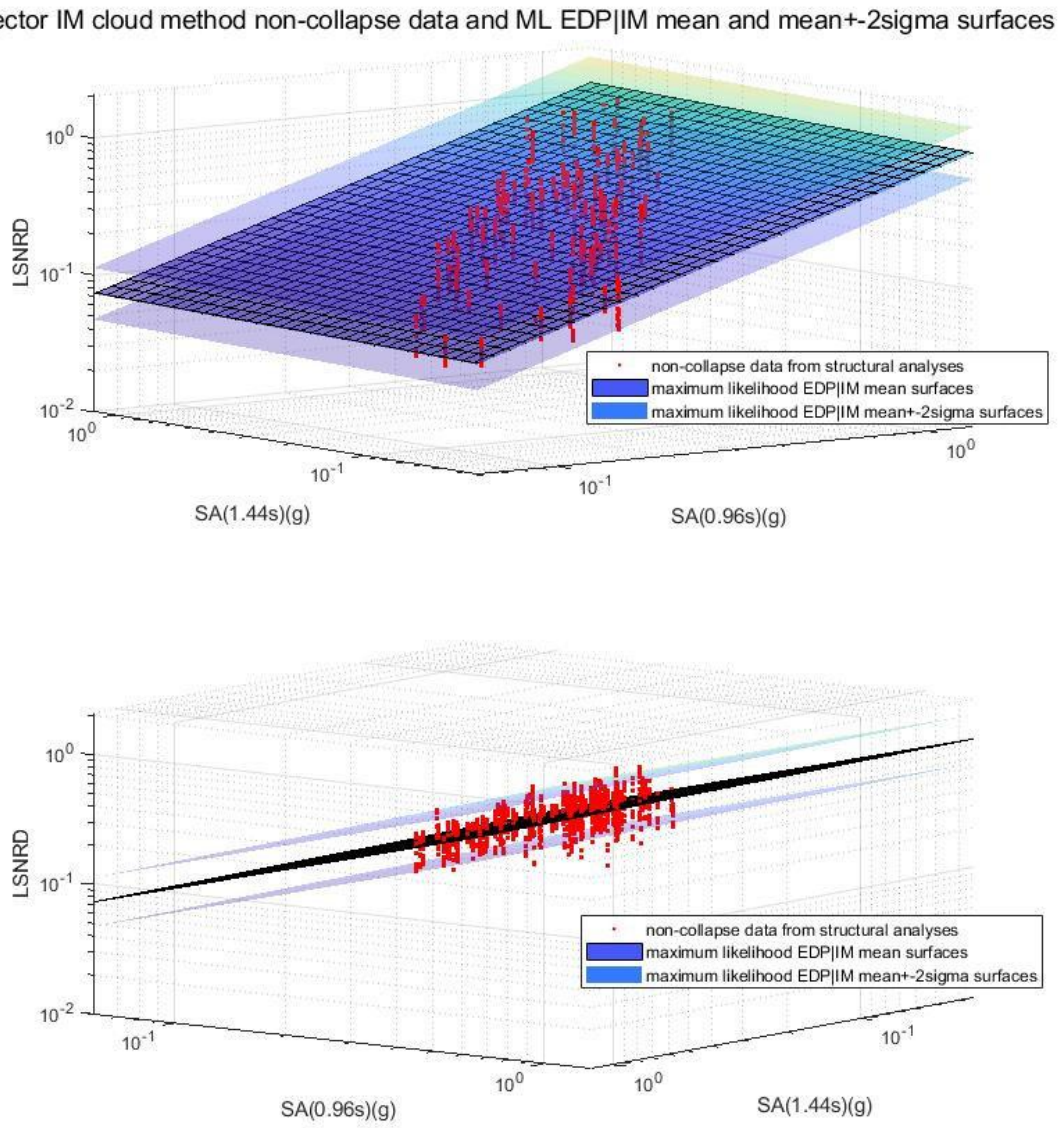

Figure 3.51: $EDP|\overrightarrow{IM}$ data points from non-collapse structural analyses and maximum likelihood $\mu_{ln(EDP)|\overrightarrow{IM}}$ and $\mu_{ln(EDP)|\overrightarrow{IM}} \pm 2\sigma_{ln(EDP)|\overrightarrow{IM}}$ surfaces

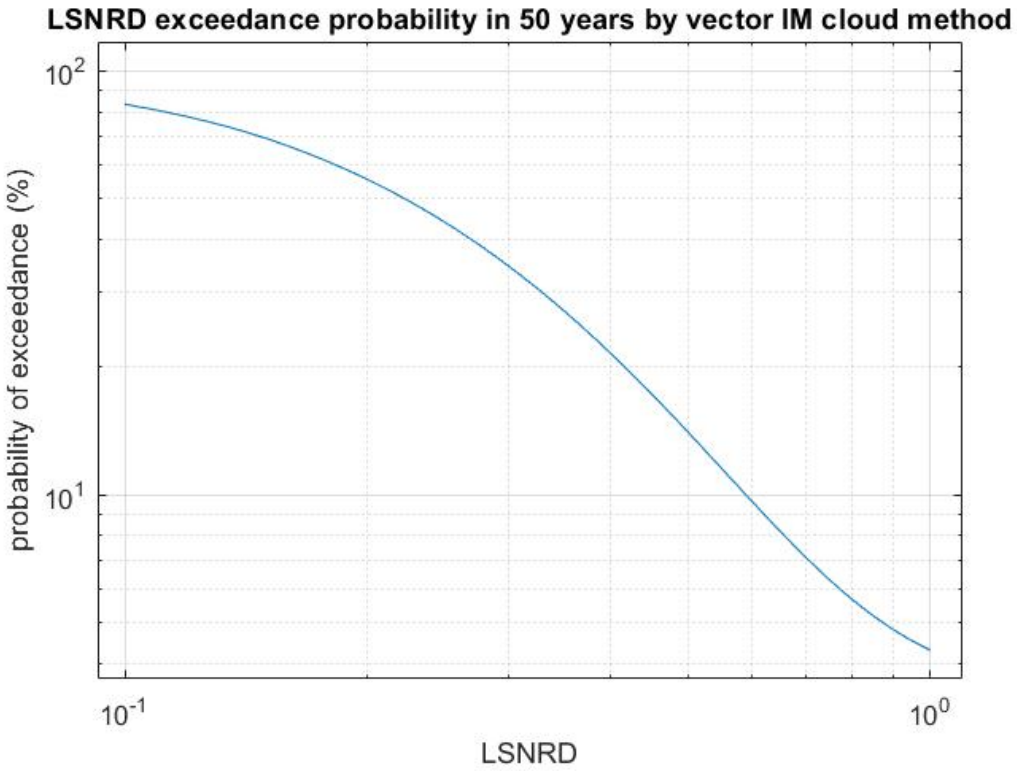

Figure 3.52: LSNRD exceedance probabilities for an exposure time of 50 years obtained from vector IM cloud method

Note that the rate of occurrence of collapse can be evaluated from the following relation:

$$\lambda_C = \int_{\overrightarrow{IM}} P_C(\overrightarrow{im}) MRD_{\overrightarrow{IM}}(\overrightarrow{im}) d\overrightarrow{im} \tag{3.70}$$

When λ_C rate of collapse is evaluated by Eqn. 3.70 and is transformed back to collapse probability for a 50-year exposure time, the resulting value is 3.48%. This result is similar to the ones obtained by scalar IM methods.

4. Summary, Conclusions and Future Works

4.1. Thesis Overview

This thesis represents the overview of whole seismic risk assessment procedure by demonstrating its implementation on an example mid-rise tunnel form building. Chapter 2 explains probabilistic seismic hazard assessment (PSHA) procedure implemented to acquire hazard curves for a site in Istanbul. Regional seismicity is represented by simplified source characterization model that idealizes two strands of North Anatolian Fault in Marmara region by a couple of idealized linear fault segments. After defining regional seismicity, this chapter also explains how to use ground motion prediction equations (GMPE) to compute hazard curves, uniform hazard spectra, conditional spectra and vector hazard surfaces (mean rate density of IM vector). Multivariate normal distributions are introduced and demonstrated as mathematical tools to implement these procedures.

Chapter 2 illustrates procedures undertaken to obtain PSHA results which are used as input for procedures presented in Chapter 3, which is a key chapter of this thesis. Chapter 3 presents simplified SDOF procedure to compute structural response to ground shaking excitation, statistical modelling of structural response and integration of these models with PSHA input to obtain risk curves. Statistical response models mentioned here represent collapse fragility curves and non-collapse $EDP|IM$ distribution models. These statistical models are developed by two different record scaling and selection methods (stripes and cloud methods) and three different scalar intensity measures (IMs). These intensity measures represent spectral acceleration at fundamental mode structural vibration period $SA(T_1)$ and two different average spectral acceleration intensity measures obtained by different averaging sets of spectral accelerations at different periods with different averaging weights.

Structural response models as well as final risk curve results obtained by different record scaling and selection methods as well as by different ground motion intensity measures (IMs) are compared and discussed. Results comparisons give valuable insights into advantages and disadvantages of different methods.

Chapter 3 also demonstrates robust vector methods of risk assessment that are capable of decreasing uncertainty of risk assessment procedure by considering joint effect of multiple ground motion intensity measures as well as multiple structural engineering demand parameters (EDPs).

4.2. Novelties, Observations and Conclusion

This thesis mainly adopts methodology of probabilistic seismic risk assessment that it well-established in earthquake engineering literature. However, there are a couple of novel methods proposed. One of them is a new ground motion intensity measure addressed as weighted average spectral acceleration (Section 2.4.2). Although average spectral acceleration is already well-established concept in literature, this thesis improves it by introducing weights of contributions of spectral accelerations at different periods to average spectral acceleration. The idea behind this new intensity measure is to give higher weights to periods that have highest effect on structural response (Section 3.2). It was observed that average spectral acceleration intensity measure is superior to fundamental mode period spectral acceleration at higher intensity levels where post-yielding period shift effects become important (Section 3.3.6). This better performance of average spectral acceleration at high intensity levels is even more emphasized when higher weights are given to higher periods which are important at high intensity levels because of period shift.

Another novel method introduced in this thesis is a new method of cloud scaling and record selection that ensures hazard consistency of selected records as well as produces uniform distribution of records across primary IM range that equally represents all intensity levels (Section 3.4.1). The idea behind this method is to scale ground motions to be uniformly distributed across primary IM range and to select one ground motion for each target IM level that best fits conditional mean spectrum conditioned on that target IM level. This method ensures that only records that are hazard consistent with respect to their intensity levels are selected.

One of important conceptual observations of this thesis is that stripe scaling method is superior to cloud scaling method for multiple reasons. Stripe scaling method considers dispersion in conditional spectrum while cloud scaling method only uses conditional mean

spectrum for record selection which ignores aleatory uncertainty in the target spectrum. Stripe method also produces superior non-collapse $EDP|IM$ distribution model that considers both the mean and standard deviation of EDP as piecewise linear functions of IM. Cloud method produces non-collapse $EDP|IM$ distribution model that assumes mean of EDP as a linear function of IM while standard deviation is modelled as a constant (Sections 3.3.1, 3.3.5, 3.4.1, 3.4.2 and 3.4.3).

This thesis also introduces mathematical algorithms for estimating parameters of statistical models for structural response based on maximum likelihood estimation method (Sections 3.3.5, 3.4.2 and 3.5). These mathematical tools are very useful because they can mitigate uncertainty that originates from errors in statistical modeling which would be inevitable had it been done by visual fitting.

When it comes to the performance of the example tunnel form building, it was observed that the collapse probability within 50 years exposure time varies between 3% and 4%. This is above 2% which is the upper limit proposed by most seismic design codes [22, 25, 26]. So, it can be concluded that mid-rise tunnel form RC buildings show low performance for collapse prevention criterium. However, when it comes to other performance criteria, namely immediate occupancy and life safety criteria, it was shown that exceedances of limited damage limit state and controlled damage limit state are approximately 26% and 5% respectively. These are nearly half of the corresponding performance limits given in codes (50% and 10%). So, when it comes to these two performance criteria, example building shows exceptionally good performance. This failure to satisfy collapse prevention criterium and exceptionality in immediate occupancy and life safety criteria can be explained by tunnel form buildings' high lateral force resisting strength but low ductility. Since RC walls provide high lateral strength, expected damage is very low at frequent low intensity ground motions. However, since the building has reduced ductility because of shear behavior of RC walls, collapse becomes more likely at high intensity ground motions. This reduces mid-rise tunnel form RC buildings' collapse prevention performance.

All the observations and innovations mentioned above indicate that there is a lot of potential for future developments and innovations in probabilistic seismic risk assessment field. Some of possible future development are discussed in following section.

4.3. Future Developments

There are two obvious potential future developments that can be conceptually guided by the methods proposed in this thesis. These are vector intensity measure hazard maps and statistically state-of-the-art vector IM-EDP risk assessment method.

4.3.1. Vector IM Hazard Maps

Vector $\overrightarrow{IM}$ methods are superior to scalar IM methods because they explicitly take into account multiple parameters that define intensity of ground shaking that facility being assessed might be exposed to. Since responses of structures depend on multiple IMs rather than one, vector $\overrightarrow{IM}$ methods reduce uncertainty in the procedure of seismic risk assessment. Vector $\overrightarrow{IM}$ methods have potential to decouple seismic hazard assessment (SHA) and seismic structural response assessment (SSRA). This could be done by constructing large $\overrightarrow{IM}$ vector that contains $SA(T)$ components corresponding to wide range of densely spaced periods T. This $\overrightarrow{IM}$ vector could be used for assessment of any structure because whatever the structure's fundamental modal period, $\overrightarrow{IM}$ will contain $SA(T)$ at period very close to it. This means that universal $\overrightarrow{IM}$ hazard maps could be developed and used for any potential structure's seismic risk assessment. These maps would contain $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ data (check Section 2.5) for any location within the map. For example, $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ could be evaluated for a grid of densely spaced points on the map and interpolated between them. Also, record set to be used for structural analyses could be the same for any structure. No special care needs to be taken to select hazard consistent record set for structural analyses since $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ term will ensure hazard consistency as discussed before.

Also, it is interesting to think about potential seismic risk assessment software development. In case of evaluation of $\overrightarrow{IM}$ hazard maps mentioned in previous paragraph, $MRD_{\overrightarrow{IM}}(\overrightarrow{im})$ data could be stored in a database. Software could be developed that first supplies record set to the user which would be used for structural analyses. The user would then input results of these structural analyses and location of the structure to the software and software would use

them and $\overrightarrow{IM}$ hazard map database to evaluate EDP rate of exceedance curves or even loss rate of exceedance curves in case statistical model of $Loss|EDP$ is supplied to the software.

4.3.2. Vector IM – EDP Methods of Seismic Risk Assessment

This thesis demonstrated two separate seismic risk assessment examples where vector $\overrightarrow{IM}$ (Sections 2.5 and 3.5) and vector $\overrightarrow{EDP}$ (Section 3.3.8) were utilized. But how about a most sophisticated method that would consider both IM and EDP as vectors?

Considering both IM and EDP as vectors would reduce uncertainty in assessment of seismic risk because, intuitively, structural damage depends on many structural response parameters (EDPs) which depend on many ground motion intensity parameters (IMs). Method that would incorporate multiple IMs as well as multiple EDPs into seismic risk assessment procedure would represent state-of-the-art statistical method of probabilistic seismic risk assessment. Multivariate statistics and maximum likelihood estimation-based procedures presented in this thesis could be used to develop this method.

REFERENCES

[1] Applied Technology Council, "Seismic Performance Assessment of Buildings Volume 1 - Methodology," Federal Emergency Management Agency, Washington, D.C., 2012.

[2] K. Porter, "A Beginner's Guide to Fragility, Vulnerability and Risk," University of Colorado Boulder and SPA Risk LLC, Denver CO USA, 2021.

[3] J. W. Baker, "Introduction to Probabilistic Seismic Hazard Analysis - White Paper Version 2.0.1," 2013.

[4] R. K. McGuire, "Seismic Hazard and Risk Analysis," Risk Engineering, Inc. Boulder,, Boulder, Colorado, 2004.

[5] J. W. Baker, "Probabilistic structural response assessment using vector valued intensity measures," *Earthquake Engineering and Structural Dynamics,* vol. 36, no. 13, pp. 1861-1883, 2007.

[6] D. Vamvatsikos and A. C. Cornell, "Incremental Dynamic Analysis," *Earthquake Engineering and Structural Dynamics,* vol. 31, no. 3, pp. 491-514, 2002.

[7] J. W. Baker and A. C. Cornell, "Spectral shape, epsilon and record selection," *Earthquake Engineering and Structural Dynamics,* vol. 35, no. 9, pp. 1077-1095, 2006.

[8] J. W. Baker, "Conditional Mean Spectrum: Tool for Ground-Motion Selection," *Journal of Structural Engineering,* vol. 137, no. 3, 2010.

[9] T. Lin, C. B. Haselton and J. W. Baker, "Conditional spectrum-based ground motion selection. Part I: Hazard consistency for risk-based assessments," *Earthquake Engineering and Structural Dynamic,* vol. 42, no. 12, pp. 1847-1865, 2013.

[10] T. Lin, C. B. Haselton and J. W. Baker, "Conditional spectrum-based ground motion selection. Part II: Intensity-based assessments and evaluation of alternative target spectra," *Earthquake Engineering and Structural Dynamics,* vol. 42, no. 12, pp. 1867-1884, 2013.

[11] M. Kohrangi, P. Bazzurro, D. Vamvatsikos and A. Stillatura, "Conditional spectrum-based ground motion record selection using average spectral acceleration," *Earthquake Engineering and Structural Dynamics,* vol. 46, no. 10, pp. 1667-1685, 2017.

[12] J. W. Baker and C. Lee, "An Improved Algorithm for Selecting Ground Motions to Match a Conditional Spectrum," *Journal of Earthquake Engineering,* vol. 22, no. 12, 2017.

[13] T. Lin, S. C. Harmsen, J. W. Baker and N. Luco, "Conditional Spectrum Computation Incorporating Multiple Causal Earthquakes and Ground-Motion Prediction Models," *Bulletin of the Seismological Society of America,* vol. 103, no. 2A, pp. 1103-1116, 2013.

[14] *MATLAB. (2020). version 9.8.0 (R2020a),* Natick, Massachusetts: The MathWorks, Inc, 2020.

[15] U. Özen, "A Study on the Fragility Modeling of Mid-rise Tunnel Form RC Buildings," Boğaziçi University, Istanbul, 2021.

[16] D. M. Boore, J. Watson-Lamprey and N. Abrahamson, "Orientation-Independent Measures of Ground Motion," *Bulletin of the Seismological Society of America,* vol. 96, pp. 1502-1511, 2006.

[17] A. K. Chopra, in *Dynamics of Structures: Theory and Applications to Earthquake Engineering*, Prentice Hall, 1981, pp. 164-170.

[18] R. R. Youngs and K. J. Coppersmith, "Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates," *Bulletin of the Seismological Society of America,* vol. 75, no. 4, pp. 939-964, 1985.

[19] M. Utku, "North Anatolian Fault Zone, Northern Turkey: Empirical Equations and the Changes in Some Kinematic Characteristics," *World Journal of Research and Review,* vol. 6, no. 3, pp. 38-61, 2018.

[20] S. Akkar, A. M. Sandikkaya and J. J. Bommer, "Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East," *Bulletin of Earthquake Engineering,* vol. 12, pp. 359-387, 2014.

[21] S. Akkar, A. M. Sandikkaya and B. Ö. Ay, "Compatible ground-motion prediction equations for damping scaling factors and vertical-to-horizontal spectral amplitude ratios for the broader Europe region," *Bulletin of Earthquake Engineering,* vol. 12, p. 517–547, 2014.

[22] *TEBC: Turkish Earthquake Building Code: Specifications for building design under earthquake effects,* Ankara: Resmi Gazete, 2018.

[23] P. Bazzurro and A. C. Cornell, "Vector-valued probabilistic seismic hazard analysis (VPSHA)," *Seismological Research Letters,* vol. 72, 2002.

[24] Y. L. Tong, The Multivariate Normal Distribution, New York: Springer, 1990.

[25] "Eurocode 8: Design of structures for earthquake resistance - Part 1 : General rules, seismic actions and rules for buildings," European Comittee for Standardization (CEN), Brussels, 2004.

[26] "Minimum Design Loads and Associated Criteria for Buildings and Other Structures ASCE/SEI 7-16," American Society of Civil Engineers (ASCE), 2016.

[27] L. F. Ibarra, R. A. Medina and H. Krawinkler, "Hysteretic models that incorporate strength and stiffness deterioration," *Earthquake Engineering and Structural Dynamics,* vol. 34, no. 12, pp. 1489-1511, 2005.

[28] N. Jayaram, T. Lin and J. W. Baker, "A Computationally Efficient Ground-Motion Selection Algorithm for Matching a Target Response Spectrum Mean and Variance," *Earthquake Spectra,* vol. 27, no. 3, pp. 797-815, 2011.

[29] Pacific Earthquake Engineering Research Center, "PEER Ground Motion Database," [Online]. Available: https://ngawest2.berkeley.edu/.

[30] C. I. Nievas and T. J. Sullivan, "Accounting for directionality as a function of structural typology in performance-based earthquake engineering design," *Earthquake Engineering and Structural Dynamics,* vol. 46, no. 5, pp. 791-809, 2016.

[31] J. Brooks-Bartlett, "Probability concepts explained: Maximum likelihood estimation," Towards Data Science Inc., 4 January 2018. [Online]. Available: https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1. [Accessed 26 July 2021].

[32] The Pennsylvania State University, "PennState - Eberly College of Science - STAT 415 - Introduction to Mathematical Statistics - Maximum Likelihood Estimation," [Online]. Available: https://online.stat.psu.edu/stat415/lesson/1/1.2. [Accessed 26 July 2021].

[33] B. G. Teo, "Binomial Distribution — Practical Intro with Drive-Thru Business Analytic," Towards Data Science, 30 November 2020. [Online]. Available: https://towardsdatascience.com/binomial-distribution-practical-intro-with-drive-thru-business-analytic-27390e5d0918. [Accessed 27 July 2021].

[34] P. Giovenale, A. C. Cornell and L. Esteva, "Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses," *Earthquake Engineering and Structural Dynamics,* vol. 33, no. 8, p. 951–979, 2004.

[35] M. Shinozuka, M. Q. Feng, J. Lee and T. Naganuma, "Statistical Analysis of Fragility Curves," *Journal of Engineering Mechanics,* vol. 126, no. 12, pp. 1224-1231, 2000.