Title
A Maximum Likelihood Prediction Model for Building Seismic Response

Author
Shokrabadi, Mehrdad

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Maximum Likelihood Prediction Model for Building Seismic Response

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Mehrdad Shokrabadi

2018

ABSTRACT OF THE THESIS

A Maximum Likelihood Prediction Model for
Building Seismic Response

by

Mehrdad Shokrabadi
Master of Science in Statistics
University of California, Los Angeles, 2018
Professor Ying Nian Wu, Chair

Earthquakes in seismically-active regions present a significant human and financial risk to communities. This study focuses on proposing a prediction model that could serve as a means for quantifying the impact of the main characteristics of earthquakes on the seismic risk of reinforced concrete moment frame structures. The seismic hazard due to sequential earthquakes is examined under mainshock-aftershock seismic sequences. A two-step maximum likelihood regression approach was adopted in order to propose a relationship that would link the seismic response of the studied buildings to the key characteristics of an earthquake. The accuracy of the proposed prediction equation was examined by comparing its outcomes with what was expected from physics-based models. Both the magnitude of an earthquake as well as the distance from the building's location to the rupture plane of the building were found to be among the parameters with the most notable impact on the buildings' seismic response.

The thesis of Mehrdad Shokrabadi is approved.

Qing Zhou

Henry J. Burton

Ying Nian Wu, Committee Chair

University of California, Los Angeles

2018

To my family ...

For their endless love, support and encouragement

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The research presented in this paper is supported by the United States Geological Survey (USGS) EHP Award Number G16AP00006. The views and conclusions presented here are those of the authors and do not represent the policy of the US Government.

CHAPTER 1

Introduction

1.1 Motivation

One of the most consequential natural hazards in terms of both human and financial losses is earthquake. Despite the recent advances in the design and construction of earthquake-resistance buildings, recent seismic events in China, New Zealand [1] and Japan [2] have highlighted the importance of the pre-planning for major seismic events. In recent years, designing buildings to withstand major seismic events has become increasingly importnat in the engineering community. The introduction of performance-based earthquake engineering [3] has marked a turning point is utilizing probabilistic frameworks for seismic design of civil structures. Performance-based earthquake engineering has provided engineers with the tools that are necessary in order to incorporate the uncertainties in the response of buildings to seismic events as well as the uncertainties in the occurrence of earthquakes themeselves. This is achieved through total probability theorem and assuming certain probability distribution for the response of buildings and the characteristics of earthquakes (e.g., magnitude and source-to-site distance). Building response is usually assumed to follow a lognormal distribution [4], the occurrence of earthquakes with certain magnitudes is usually modeled through an exponential distribution with a rate calculated using the relationship suggested by Gutenberg and Richter [5] and the distance from where the earthquake has occurred to the site where the building is located (source-to-site distance) is usually modeled using an empirical distribution fitted to the distance data of each earthquake.

A prohibitive factor in using the performance-based earthquake engineering's framework for building seismic design is the complexity and computational cost associated with eval-

uating seismic response of buildings that is a main requirement of the performance-based earthquake engineering's approach. To obtain an accurate estimate of how a building would respond to future earthquakes, first one needs to build a numerical model of the building. Buildings, depending on their size and configuration, can contain hundreds of different individual components (e.g., beams, columns, shear walls, foundation, bearing walls) with each component usually showing a range of complex behaviors when subjected to earthquakes with different intensities. Moreover, as some of these components get damaged in an earthquake, the distribution of the seismic loads accross the remaining components changes drastically that could have a major impact on the building seismic response. As such, not only constructing an accurate model of the building's individual components is important when evaluating seismic response, but it is crucial to accurately represent how these building elements would interact with each other in the numerical model.

Earthquakes are a compilation of random waves and therefore, even when having similar magnitudes, they could have substantially different impacts on the buildings that they are applied to. As such, an important step in numerically predicting the seismic response of a building is to subject its numerical model to a large set of earthquakes with distinct characteristics. This would allow for obtaining a comprehensive picture of the building's response to future seismic events. However, an accurate numerical model of the building could be significantly complex as described before and subjecting such a complex model to a large number of earthquakes could be very computationally demanding and time-consuming. A solution to this problem is to use a prediction equation for building response. Such a prediction equation would take as inputs the characteristics of an earthquake that have the most significant impacts on the building response as well as the most important properties of the building (e.g., number of stories, dynamic period of response) and yield an estimate of the expected response of the building to such a seismic event.

1.2 Objectives

The main objective of this study is to use a large dataset of mainshock-aftershock ground motions to propose an equation that can be used to predict the seismic response of buildings when they are subjected to a sequence of mainshock-aftershock ground motions. Aftershocks can play a significant role in the financial and human costs of major mainshock seismic events. Aftershocks have been shown to exacerbate the damage caused by mainshocks and in some cases, have caused significant human and financial losses [6] [7]. While aftershocks are usually smaller in magnitude than their causative mainshock, structures can be particularly vulnerable to aftershocks due to their high rate of occurrence and the reduction in the lateral load-carrying capacity caused by mainshock-damage. In the 1999 Kocaeli earthquake, several buildings that survived the mainshock, which had a magnitude of 7.4, collapsed during a magnitude 5.9 aftershock, which occurred one month later, killing seven people and injuring more than two hundred [8]. The aftershocks that followed the 2008 Wenchuan earthquake damaged 196 dams and claimed more lives. The 2010 magnitude 7.1 Darfield earthquake was followed by two magnitude6.2 and 6.0 aftershocks, which resulted in 185 fatalities and damaged approximately 100,000 buildings in the city of Christchurch [1]. The five aftershocks with magnitudes over 7 that followed the 2011 Tohoku earthquake caused additional damage to infrastructure, liquefaction and loss of lives [2]. As such, it is crucial for a building response prediction equation to account for the contribution of both mainshocks and aftershocks to the demand that a building would experience during earthquakes.

1.3 Organization

This study seeks to propose a prediction equation for nonlinear response of reinforced concrete moment frame buildings under mainshock-aftershock seismic sequences. The main body of the thesis consists of 3 chapters. Chapter 1 discusses the advantages that a prediction for building seismic response can have as well as the importance of accointing for both mainshocks and aftershocks in the developement process of the equation.

Chapter 2 describes the steps involved in building a numerical model of the reinforced concrete moment-frame structures that are used for developing the prediction equations and the method used for obtaining the response of buildings under mainshock-aftershock seismic events. Additionally, the set of mainshock-aftershock ground motions, which are all from real as-recorded seismic sequences, are discussed in this chapter.

Chapter 3 discussed the steps involved in developing the mixed-effect prediction equation. The outputs of the building response prediction equation are in the form of maximum relative displacement between the building floors recorded during a seismic event divided by the height of the story between the two floors. We will refer to this measure of building seismic response as maximum story drift ratio (SDR).

The main focus of Chapter 4 is placed on discussing different aspects of the prediction equation developed in Chapter 3. First, we would show the unbiasedness of the prediction equation by examining its residuals and then we will discuss the accuracy of the prediction equation by comparing its predicted outputs with the outputs that are expected based on physical models.

CHAPTER 2

Building numerical model and input ground motions

In the first part of this chapter, we will briefly discuss the steps that are taken in assembling a numerical model of the buildings. The outputs of this numerical model when subjected to a set of earthquake ground motions will comprise a crucial part of the dataset that will be used in Chapter 3 to propose a prediction equation for building seismic response. In the second half of the current chapter, we will describe the set of earthquake ground motions that the numerical models of the buildings are subjected to.

2.1 Building Model

Five modern, code-conforming reinfi buildings are used for this study. The building models are adopted from the set of models developed by Haselton [9]. The buildings are designed following the latest seismic provisions of the design and loading codes in the USA; namely ACI 318-02 and ASCE 7-05 ([10]; [11]). The design requirements of Chapter 21 of ACI 318-02 for the seismic design of special moment frames are also adopted. Figure 2.1 shows the location of the buildings as well as the locations of the faults that contribute the most to the seismic hazard at the location of the buildings. To investigate the effect of building height on mainshock-aftershock seismic response, five 2-, 4-, 8-, 12- and 20-story buildings are included in the study. The buildings are designed for a high-seismicity site in Southern California. The variation in the building heights height would provide a means to more thoroughly study the relationship between the building configuration and seismic response. The buildings were chosen to incorporate a broad period-range ($0.66s - 2.63s$) to evaluate whether the effects of alternate record-pairs varies across structure periods. The seismic

response of the 2-story, and to a lesser extent the 4-story, structure is mostly influenced by the high frequency energy of ground motions due to their low first-mode period (T_1). A broader range of frequencies influence the response of taller buildings because of their high T_1 and the presence of significant higher-mode effects.

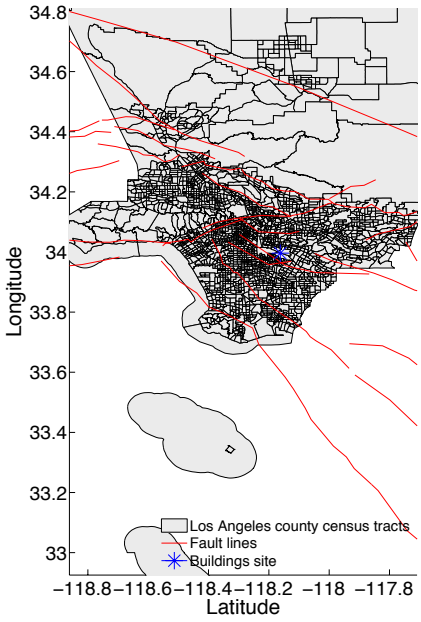

Figure 2.1: The site location of the buildings and the faults that contribute the most to the seismic hazard at the location of the buildings

Two-dimensional nonlinear models of the buildings are constructed using one of the most powerful platforms used for numerical modeling of civil structures calles OpenSees [12]. These numerical modols will be utilized to perform nonlinear dynamic analyses under a large set of ground motions. Figure 2.2 shows a schematic layout of the 4-story building's model. The remaining building models follow a similar layout. Table 2.1 summarizes the design information as well as the periods of the first two modes of each building. The design base shear coefficient in Table 2.1 is the ratio of the building's total weight used in designing the building the building to resist the earthquake loads and the yield base shear coefficient is the ratio of the force required for building to experience permanent, non-restorable deformations normalized by the building's total weight. The first mode periods are obtained using eigen-value analysis where the inputs are the building's mass and stiffness matrices.

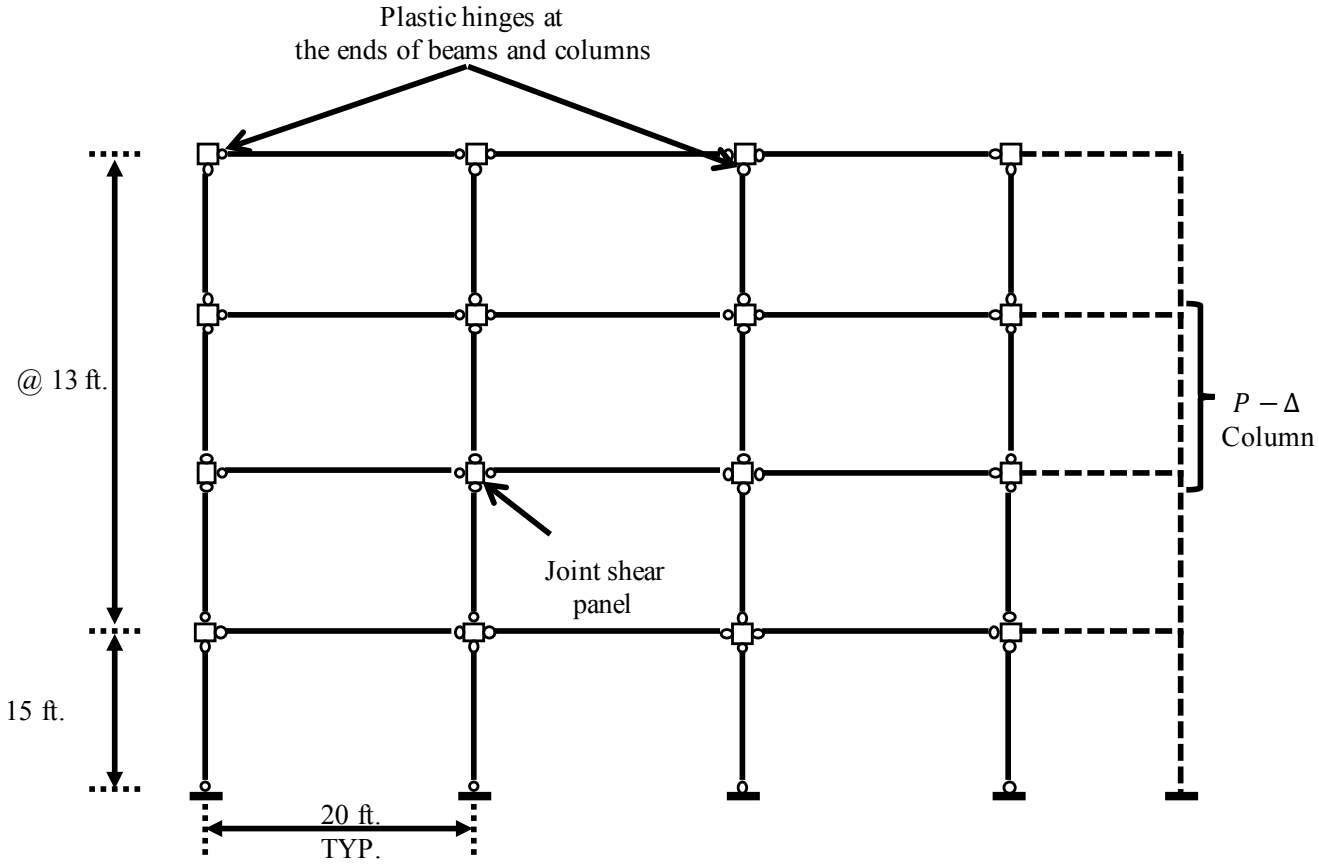

Figure 2.2: Schematic illustration of the numerical model of the archetypes used for nonlinear dynamic analyses

Table 2.1: Summary of building properties

Stories	*Design base shear coefficient*	*Yield base shear coefficient*	*First mode period*
2	0.125	0.392	0.66
4	0.092	0.143	1.12
8	0.050	0.077	1.71
12	0.044	0.075	2.01
20	0.044	0.070	2.63

Geometric destabilization effects are incorporated with a leaning column. Beams and columns are represented with flexural plastic hinges at the member-ends connected through an elastic element. The nonlinear behavior of the flexural hinges is modeled using the trilinear backbone curve developed by Ibarra et al. [13], which is capable of capturing both strength and stiffness deterioration as the building experiences consecutive loading and unloading cycles during an earthquake. The properties of the plastic hinges are obtained through the empirical relationships developed by Haselton [9]. The points of connections of beams and columns are modeled with an elastic element as the detailing requirements for special moment frames are expected to prevent shear failure in the joints. It is worth noting that the use of a two-dimensional model to simulate the seismic response of a building structure has its limitations. Three dimensional effects such as bi-axial bending in columns and accidental torsion are not captured. Moreover, ignoring the contribution of gravity frames could potentially underestimate the strength and stiffness of the structure [9]. More detailed information on the design and structural modeling approach and its limitations can be found in Haselton [9].

2.2 Mainshock-aftershock sequences

As noted previously, one of the objectives of this work is to investigate the impact of mainshock-aftershock ground motions on structural response. In order to achieve this objective, we need to compile a set of mainshock-aftershock ground motions recorded from past earthquakes that we would later applt to the numerical models of the buildings discussed in Section 2.1 to obatin their seismic response. One of the most comprehensive ground motion collections is the NGA-West2 database complied by the Pacific Earthquake Engineering Center (PEER) at University of California, Berkeley with more than 21000 records from 599 earthquakes [14].

The set that we used in developing the building response prediction equation contains 1240 mainshock-aftershock records from 13 events. We use actual mainshock-aftershock sequences as predicting building response under Synthetic records could be erroneousas re-

ported in a number of past studies. Goda [15] compared the collapse performance of a 2-story wood-frame building under both mainshock-mainshock and mainshock-aftershock ground motion pairs. The mainshock-mainshock sequence used the same records in the second event as in the first. The mainshock-mainshock sequence produced higher collapse probabilities than mainshock-aftershock. Ruiz-Garca [16] did a similar study using two low- and mid-height steel frames and reached the same conclusion. In developing a ground motion prediction equation, Abrahamson et al. [17] found that the median of spectral values of aftershocks at short periods are smaller than those from similar mainshocks, whereas at longer periods ($\geq 0.75 \quad sec$) the aftershock spectral ordinates were larger. Chiou and Youngs [18] also reached a similar conclusion. As such, to avoid any loss in accuracy of the building response prediction, we are using actual as-recorded mainshock-aftershock pairs. The mainshock-aftershock classification for the ground motions selected from the PEER NGA-West2 database is based on the time and distance windowing algorithms developed by Wooddell and Abrahamson [19]. A magnitude-dependent time window and a distance threshold of 40 km measured in terms of the centroidal Joyner-Boore distance [19] is used to identify the aftershock ground motions following a mainshock event. The magnitudes of the events that produced the selected ground motions range from 6.18 to 7.62 for mainshocks and 5.80 to 7.28 for aftershocks. The source-to-site distances are within 0 and 94.50 km window for mainshcoks and 0 to 510.00 km range for aftershocks. The response spectra of all the mainshock and aftershock ground motions as well as the median spectra are shown in Figures 2.3 and 2.4.

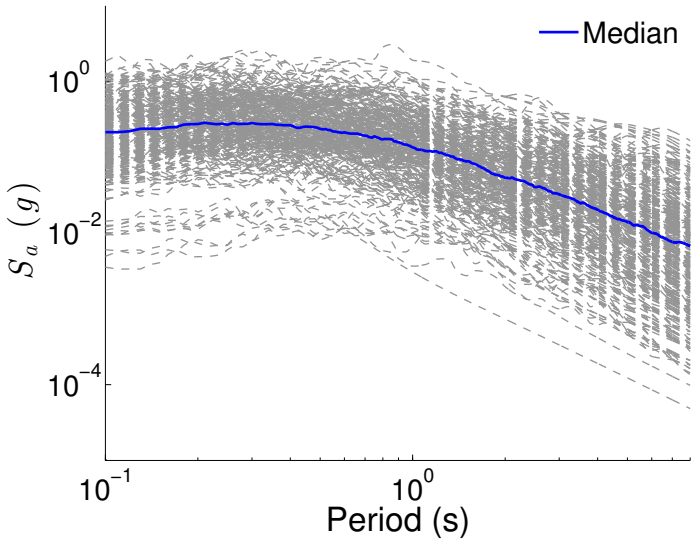

Figure 2.3: Response spectra of mainshock ground motions

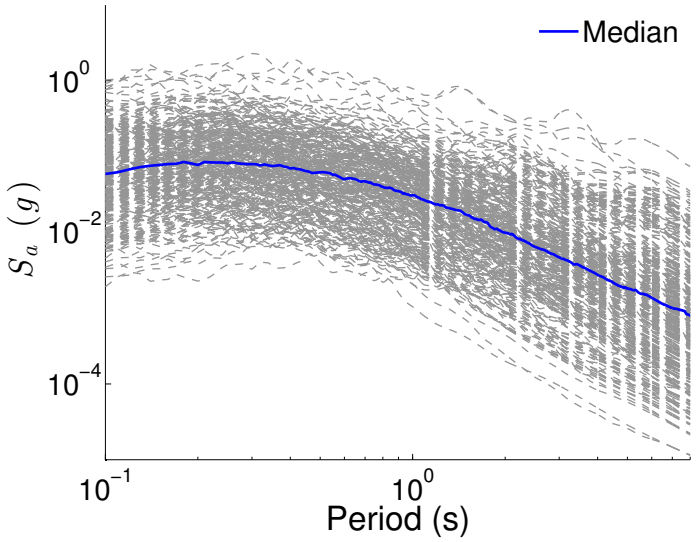

Figure 2.4: Response spectra of aftershock ground motions

CHAPTER 3

Developing a prediction equation for building seismic response

Ground motion prediction equations have been used in past studies to predict the response of simple, linear structures. Usually when an earthquake happens on a fault, the epicenter of the fault is far from the location of the structure whose response we are interested in. As the waves that comprise an earthquake pass through the sites located between the building and the earthquake source, they would get damped or amplified depending on the condition of the soil that they are passing through. A ground motion prediction equation uses such as magnitude and fault type as well as the distance from the earthquake's occurrence location to the building's site and the soil condition at the building's location to give an estimate of the intensity of the earthquake at the building's location. While there are a number of such equations available in the literature (e.g., [20][21][22]), the majority of ground motion prediction equation in the past have been developed for very simple structures whose responses to earthquakes were linear. While the equations proposed for linear systems could give a preliminary idea of how a strcuture responds to a ground motion and how the seismic response is influenced by the different characteristics of earthquakes, it is a well known fact that the response of a building even to a moderate earthquake would be highly nonlinear. Despite this shortcoming of the current prediction equations, most of the efforts in developing ground motion prediction equations have been directed towards linear systems. There are only a handful of precition equations developed for nonlinear systems and even then most of these equation are developed for the simplest forms of nonlinear systems, namely single-degree-of-freedom systems which only have one component and there is no interaction between different components or various dynamic modes of response in them. In fact, to

the author's knowledge there is only one prediction equation developed specifically for a complex multi-degree-of-freedom structure [23]. What distinguishes the current study from the Hancock et al.'s study is the incorporation of both mainshocks and aftershocks in the development steps of the prediction equation proposed in this study. As discussed before, aftershocks could have a significant impact on the seismic response of buildings. Ground motion prediction equations are historically developed by using statistical regression on large sets of data. In the early stages of developing ground motion prediction equations, ordinary least square (OLS) regression was the primary method utilized for performing regression on the earthquake ground motion data. However, Fukushima and Tanaka [24] showed that using OLS could result in significant inaccuracies in predicting response due to the strong correlation between earthquake magnitude and distance. To overcome this error, Fukushima and Tanaka [24] suggested a two-stage regression method to decouple the magnitude regression step from the regression on distance step. Random-effect models were introduced by Brillinger and Preisler [25]. Their models included two major sources of variation. The first component was to account for the variation from an earthquake to another earthquake. This variation mostly occurs due to the differences in the magnitude and fault style of different earthquakes. The second source of variation was the record-to-record variability. This element of variation accounts for the various ground motions that are recorded from a single earthquakes at different stations. The record-to-record variability mainly occurs due to the different soil conditions of the sites where accelerograms record the ground motions from a seismic event. At each of the two steps of regression on magnitude and distance, a maximum likelihood approach (MLE) is adopted in order to obtain the parameters of the regression equation. The prediction equation adopted in this study has the following form based on the physics-based model of ground motions and building response.

$$ln(SDR) = c_1(M - M_h) + c_2 SS + c_3 NS + c_4 RS + c_5 ln(\sqrt{R_{jb}^2 + h^2} - 1) + c_6 ln(V_{s30}/760) + \epsilon_r + \epsilon_e \tag{3.1}$$

Ground motions are usually recorded in two perpendicular directions. In this study, for each ground motions we applied its recorded time series in two perpendicular directions to

the building and took the geometric average of the building response under the two perpendicular ground motion directions as the response variable. As such, in Equation 3.1 SDR is the geometric mean of the maximum story drift ratio obtained for the two perpendicular components of the ground motion and c_1 to c_6 and h are coefficients computed through regression. M in Equation 3.1 is the magnitude of the earthquake under which we are predicting the building's response and R_{jb} is the shortest distance from the building's location to the surface projection of the rupture plane of the fault. The regression was performed using the two-step procedure of Joyner and Boore (1993) and Joyner and Boore (1994), which provides for an event-specific mean misfit, which is similar to an event term in a ground motion prediction equation developed using mixed effects regression. SS, NS and RS are dummy variables that define the fault type of each ground motion. For instance, if the fault type of ground motion i is SS, then $SS_i = 1$ and $NS_i = RS_i = 0$. V_{s30} is the parameter that defines the soil condition at the location of the building and is defined as the average shear wave velocity of in 30-meter long soil profile. ϵ_r is the record-to-record residual that would be different from a ground motion time series to another time series. ϵ_e is the earthquake-to-earthquake residual that would vary between different earthquakes but it will remain constant among all the time series that belong to the same earthquake.

Equation 3.1 has a nonlinear coefficient term $ln(\sqrt{R_{jb}^2 + h^2} - 1)$. To simplify the fiting steps by using a form of it where all the regression coefficients are linear, we start by calculating the Taylor expansion of Equation 3.1 at a realization of h which we call $h^{'}$ as shown in Equation 3.2 ([26]). The second order terms are ignored.

$$\begin{aligned} ln(SDR) =& c_1(M - M_h) + c_2 SS + c_3 NS + c_4 RS + c_5 ln(\sqrt{R_{jb}^2 + h^{'2}} - 1) \\ &+ c_6 ln(V_{s30}/760) + c_5 \frac{\frac{1}{2} 2h(R_{jb}^2 + h^{'2})^{\frac{1}{2}}}{\sqrt{R_{jb}^2 + h^{'2}} - 1}(h - h^{'}) \\ &+ \epsilon_r + \epsilon_e \end{aligned} \tag{3.2}$$

Equation 3.2 can be written in the matrix format described in the following.

$$Y = \begin{pmatrix} ln(SDR_1) \\ ln(SDR_2) \\ \vdots \\ ln(SDR_n) \end{pmatrix} \tag{3.3}$$

$$X = \begin{pmatrix} M_1 - M_h & SS_1 & NS_1 & RS_1 & ln(\sqrt{R_{rup_1}^2 + h'^2} - 1) & ln(V_{s30_1}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_1}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_1}^2 + h'^2} - 1} \\ M_2 - M_h & SS_2 & NS_2 & RS_2 & ln(\sqrt{R_{rup_2}^2 + h'^2} - 1) & ln(V_{s30_2}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_2}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_2}^2 + h'^2} - 1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_n - M_h & SS_n & NS_n & RS_n & ln(\sqrt{R_{rup_n}^2 + h'^2} - 1) & ln(V_{s30_n}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_n}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_n}^2 + h'^2} - 1} \end{pmatrix} \tag{3.4}$$

$$\beta = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ \delta h \end{pmatrix} \tag{3.5}$$

$$\epsilon = \begin{pmatrix} \epsilon_{r_1} + \epsilon_{e_1} \\ \epsilon_{r_2} + \epsilon_{e_2} \\ \vdots \\ \epsilon_{r_n} + \epsilon_{e_n} \end{pmatrix} \tag{3.6}$$

In the matrix of Equation 3.5, $\delta h = h - h^{'}$ and n is the total number of the time series used in the regression. In each step of the iteration, the new value of h is calculated as $h^t = h^{t-1} + \delta h$. The matrices of Equations 3.3, 3.4, 3.5 and 3.6 together form the matrix equation of Equation 3.7.

$$Y = X\beta + \epsilon \tag{3.7}$$

As discussed before, in this study we are interested in performing the regression in two stages. Performing the regression in two stages would allow us to differentiate the two main sources variation in the response of the structure; namely variation in the parameters that vary from one earthquake to another (M, SS, NS and RS) and the parameters that vary within the different time series recorded from the same earthquake (R_{jb} and V_{s30}). To achieve this goal, the first stage of the regression is performed only on M, SS, NS and RS as shown in Equation 3.8 ([26]).

$$P_i = c_1(M_i - M_h) + c_2 SS_i + c_3 NS_i + c_4 RS_i + \epsilon_{e_i} \tag{3.8}$$

We need to re-write the X and β matices of Equations 3.4 and 3.5 in the form shown in Equations 3.9 and 3.10.

$$X_r = \begin{pmatrix} ln(\sqrt{R_{rup_1}^2 + h'^2} - 1) & ln(V_{s30_1}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_1}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_1}^2 + h'^2} - 1} & E_{11} & E_{12} & \dots & E_{1N_e} \\ ln(\sqrt{R_{rup_2}^2 + h'^2} - 1) & ln(V_{s30_2}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_2}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_2}^2 + h'^2} - 1} & E_{21} & E_{22} & \dots & E_{2N_e} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ ln(\sqrt{R_{rup_n}^2 + h'^2} - 1) & ln(V_{s30_n}/760) & c_5' \frac{\frac{1}{2} 2h (R_{rup_n}^2 + h'^2)^{\frac{1}{2}}}{\sqrt{R_{rup_n}^2 + h'^2} - 1} & E_{n1} & E_{n2} & \dots & E_{nN_e} \end{pmatrix} \tag{3.9}$$

$$\beta_r = \begin{pmatrix} c_5 \\ c_6 \\ \delta h \\ P_1 \\ \vdots \\ P_{N_e} \end{pmatrix} \tag{3.10}$$

$$\epsilon_r = \begin{pmatrix} \epsilon_{r_1} \\ \epsilon_{r_2} \\ \vdots \\ \epsilon_{r_n} \end{pmatrix} \tag{3.11}$$

In Equation 3.9, N_e is the total number of earthquakes that are included in the regression. E_{ij} is 1 if the time series ij comes from earthquake i and it is zero if the time series ij does not belong to earthquake j. Equations 3.9, 3.10 and 3.11 together give the following matrix equation.

$$Y = X_r \beta_r + \epsilon_r \tag{3.12}$$

If we assume that ϵ_{r_i} follow a normal distribution a mean value of zero and a standard deviation taken as σ_r, then the vector of coefficients β_r can be calculated using the familiar relationship shown in Equation 3.13.

$$\beta_r = (X_r^T X_r)^{-1} X_r^T Y \tag{3.13}$$

The next step is to calculate the remaining coefficients c_1, c_2, c_3 and c_4 in Equation 3.1. In order to do so, we add and subtract $P_i^{'}$ terms in Equation 3.9 and move P_i to the right side of the equation. Recall that the $P_i^{'}$ terms are obtained using Equation 3.13 in the previous step ([26]).

$$P_i^{'} = c_1(M_i - M_h) + c_2 SS_i + c_3 NS_i + c_4 RS_i + \epsilon_{e_i} + P_i^{'} - P_i \tag{3.14}$$

Now we can re-arrange Equation 3.14 in the following matrix form.

$$Y_e = \begin{pmatrix} P_1^{'} \\ P_2^{'} \\ \vdots \\ P_n^{'} \end{pmatrix} \tag{3.15}$$

$$X_e = \begin{pmatrix} M_1 - M_h & SS_1 & NS_1 & RS_1 \\ M_2 - M_h & SS_2 & NS_2 & RS_2 \\ \vdots & & & \\ M_n - M_h & SS_n & NS_n & RS_n \end{pmatrix} \tag{3.16}$$

$$\beta_e = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \tag{3.17}$$

$$\epsilon_e = \begin{pmatrix} P_1^{'} - P_1 + \epsilon_{e_1} \\ P_2^{'} - P_2 + \epsilon_{e_2} \\ \vdots \\ P_n^{'} - P_n + \epsilon_{e_n} \end{pmatrix} \tag{3.18}$$

Which can be written in the matrix form of Equation 3.19.

$$Y_e = X_e \beta_e + \epsilon_e \tag{3.19}$$

The variance in the estimate of Y_e can be calculated as shown in Equation 3.20 ([26]):

$$Var[Y_e] = Var[P^{'} - P] + \sigma_e^2 I \tag{3.20}$$

Note that, in Equation 3.20 the correlation between $P^{'} - P$ and $\sigma_e^2 I$ iz zero. This is because $P^{'} - P$ is a linear combination of the error terms ϵ_r obtained through Equation **??** which themselves were independent normal random variables. The variance in the estimate of β_r in Equation 3.13 can be calculated using the relationship of Equation 3.21.

$$Var[\beta_r] = (X_r^T X_r)^{-1} \sigma_r^2 \tag{3.21}$$

We can write the likelihood function of the linear regression in Equation 3.19 can be written as shown in Eqaution 3.22.

$$L_e = (2\pi)^{\frac{N_e}{2}} |Var[Y_e]|^{-\frac{1}{2}} exp(\frac{1}{2}(Y_e - X_e\beta_e)^T Var[Y_e]^{-1}(Y_e - X_e\beta_e)) \tag{3.22}$$

Equation 3.22 represents a generalized least square problem whose soultion is as shown in Equation 3.23.

$$\hat{\beta}_e = (X_e^T Var[Y_e] X_e)^{-1} X_e^T Var[Y_e]^{-1} Y_e \tag{3.23}$$

Equations 3.23 and 3.13 provide the basis for calculating all the unknown coefficients of Equation 3.1. However, a main prblem with Equation 3.23 is that σ_e^2 in the $Var[Y_e]$ component is unknown. To calculate σ_e^2, we can use the following recursive relationship.

$$E[(Y_e - X_e\hat{\beta}_e)^T Var[Y_e]^{-1}(Y_e - X_e\hat{\beta}_e)] = n - 4 \tag{3.24}$$

In Equation 3.24, $n - 4$ is the number of the degrees of freedom of the regression in Equation 3.20. To calculate the coefficients c_1, c_2, c_3 and c_4 in Equation 3.1, we start with Equation 3.23 with an initial trial value for σ_e. For each trial value of σ_e, we need to solve Equation 3.20 and Equation 3.23 until the value of σ_e converges.

The variance-covariance matrix $Var[Y_e]$ is a diagonal matrix where each diagonal term can be calculated as shown in Equation 3.25. The N_i term is the number of time series used in the regression that belong to earthquake i. The $Var[Y_e]$ matrix is diagonal since all the variation in the second stage of the regression is due to the variations in the properties of earthquake that vary from one event to another and not the properties that would vary within the time series that belong to the same earthquake ([26]).

$$Var[Y_e]_i = \frac{1}{\frac{\sigma_r^2}{N_i} + \sigma_r^2} \tag{3.25}$$

CHAPTER 4

Outcomes of prediction equation for building seismic response

In this section, we will discuss the outcomes of the seismic response prediction equation discussed in Chapter 3. We use the prediction relationship of Equation3.1 tp predict the maximum relative displacement between the stories of the buildings described in Section 2.1 given certain characterisitcs of a future earthquake. These characteristics are magnitude, source-to-site distance, the type of the soil at the building's location and fault type. We develop an individual equation for each of the five concrete frame buildings to better reflect the unique dynamic characteristics of each building and their impact on the seismic response.

In order to find the regression coefficients of Equation 3.1, first we need to calculate the SDR values which would serve as the input observations of the equation. In order to do so, we need to subject the numerical models of the buildings described in Section 2.1 to the set of mainshock-aftershock ground motions discussed in Section 2.2 and record the maximum story drift values. The next step is to follow the regression steps discussed in Chapter 3 to calculate the regression coefficients. Equation 3.1 has an M_h term that defines the point where there will be a shift in the relationship between the building response and magnitude of the earthquake. We included this term in the regression equation since a number of past studies have suggested that the relationship between building response and earthquake magnitude is different for large-magnitude and small-magnitude earthquakes ([21, 20, 22]). Based on the analysis of the relationship between the seismic response of the five buildings examined in this study and the magnitudes of the considered earthquakes, we selected a value of 6.7 for M_h. The plots between the exponentials of ϵ_e from the first stage of regression and earthquake magnitude are used to calibrate M_h (each ϵ_e serves as an estimate of $ln(SDR)$

for the entire time series available from one earthquake). The exponentials of ϵ_e versus magnitude plots for the five studied buildings are shown in Figure 4.2 to demonstrate how the M_h calibration is done. As can be seen, there is a clear change in the relationship between magnitude and the exponentials of ϵ_e at the M_h values in the neighborhood of 6.7. As such, a bilonear regression with the breaking point at $M_h = 6.7$ is likely to represent the magnitude scaling component of the regression more accurately compared to a simple linear regression.

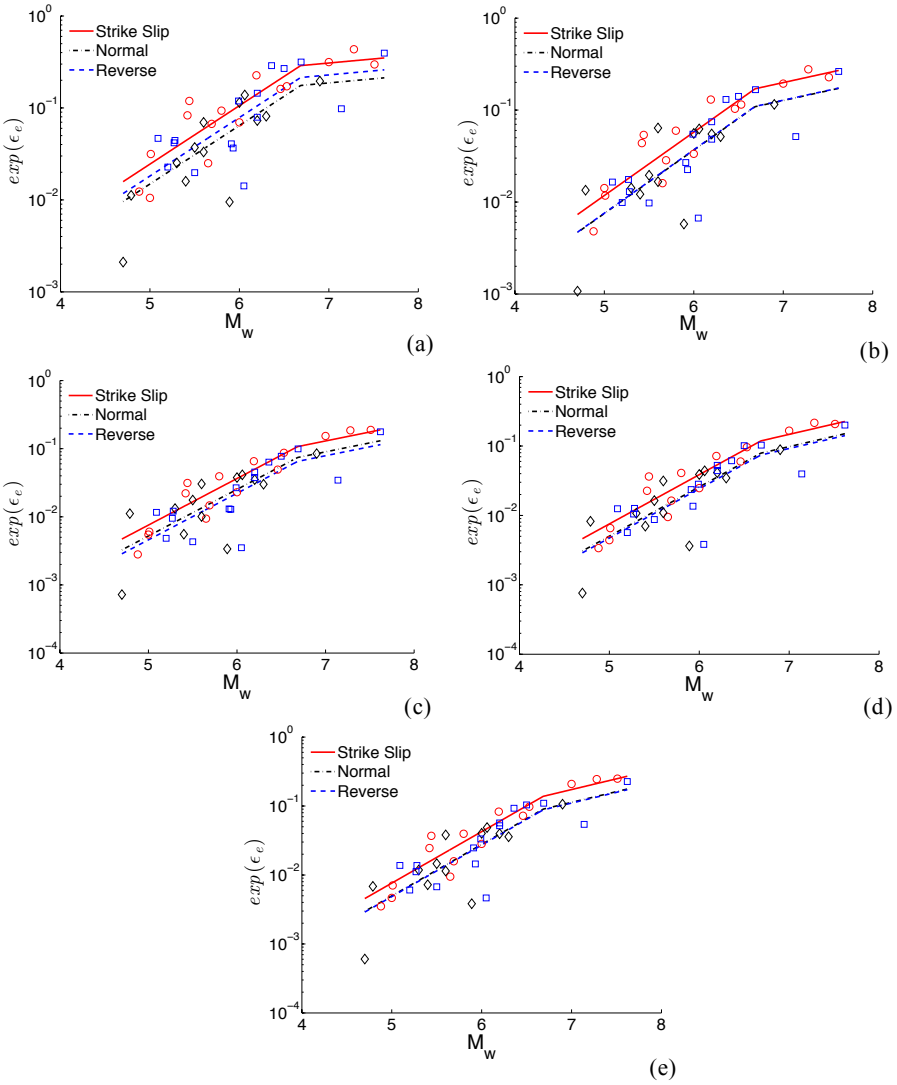

Figure 4.1: Calibration of M_h for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

Table 4.1 summarizes the coefficients of the regression performed using Equation 3.1

for each of the five studied buildings. The c_1 coefficient, which establishes the relationship between the earthquake magnitude and the maximum building demand is positive as expected; indicating that as the magnitude of the earthquake increases, the earthquake-induced demand that the buildings would experience increases too. The c_2, c_3 and c_4 are dummy variables that would take on a value of zero or one depending on the fault type of the earthquake being examined. as such, the serve as the intercept of the regression equation for each fault type. The c_5 coefficient has a negative value for all the building types. This suggests that there is an inverse relationship between the distance from the rupture surface of the fault to the site where the building is located and the seismic demand on the building. The value of h does not vary significantly among the five studied buildings and prevents the $c_5 ln(\sqrt{R_{rup}^2 + h^2} - 1)$ from becoming singular for very small values of R.

Table 4.1: Regression coefficients of Equation 3.1

Building	$c_1(M \leq M_h)$	$c_1(M > M_h)$	c_2	c_3	c_4	c_5	h	c_6
2-Story	1.46	0.21	-1.24	-1.74	-1.54	-1.25	11.81	-0.64
4-Story	1.59	0.49	-1.77	-2.22	-2.21	-1.14	9.15	-0.78
8-Story	1.57	0.61	-2.23	-2.60	-2.73	-1.10	8.02	-0.96
12-Story	1.63	0.69	-2.13	-2.54	-2.60	-1.21	9.08	-1.02
20-Story	1.72	0.72	-1.98	-2.41	-2.43	-1.31	10.67	-1.02

A metric for measuring the unbiasedness of regression is the distribution of its residuals. If residuals are distributed quite randomly about the horizontal axis that represents one of the predictors in the regression such that their mean value is close to zero across a range of values for that predictor, then it can be concluded that the prediction equation is unbiased with respect to that particular predictor. Figure 4.2 shows how the distribution of the earthquake-to-earthquake residuals (ϵ_e) obtained in the first stage of the regression vary as the magnitude of the earthquakes used in the regression changes. The ϵ_e are shown for all the five considered building cases. Recall that the first stage of the regression is performed only on the predictors that would vary from one earthquake to another earthquake. As such, the predictor selected for the plots of Figure 4.2 is magnitude which is the main parameter

that varies among different earthquakes. The means and standard deviations of the ϵ_e values are shown for 7 magnitude intervals. As shown in the figure, the mean values of ϵ_e are close to zero for all magnitude intervals; suggesting that the first stage of the regression, whose main predictor variable was magnitude, yields an equation that is unbiased with respect to the magnitudes of the earthquakes included in the regression.

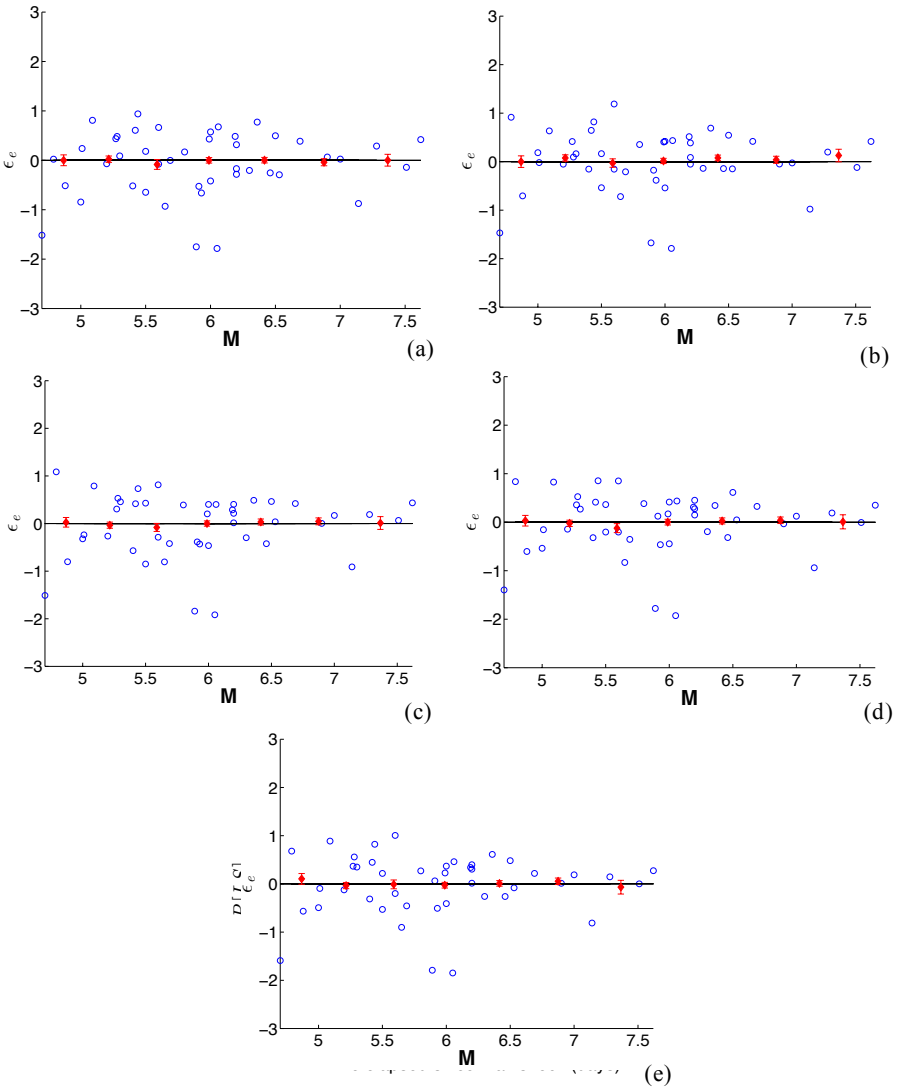

Figure 4.2: Earthquake-to-earthquake residuals (ϵ_e) versus earthquake magnitude for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

One of the main predictors in the second stage of the regression, which was performed on the predictors that would vary from one recorded time series to another, was the distance

from the buildings' location to the rupture plane of the fault. Figure 4.3 shows how the residuals in this step of regression (ϵ_r) are distributed with respect to the distances of all the earthquakes involved in the second step of the regression. The ϵ_r values are again shown for each of the five case study buildings. Error bars, which show the mean and standard deviation of ϵ_r for 11 distance intervals, are also plotted on each curve. For very short distances ($R_{jb} < 10km$), the lack of sufficient datapoints in the regression steps results in ϵ_r values whose average deviate from zero whereas for the larger distances the averages of ϵ_r values are close to zero. Moreover, the lack of sufficient datapoints has a notable impact of the variance of the error term ϵ_r. As can be seen in Figure 4.3, for $R_{jb} < 10km$ the variance of ϵ_r is much larger than the variance of the ϵ_r term at higher diatances. However, for larger distances, which have more datapoints in the regression stage, the variance of ϵ_r does not change significantly with distance as shown in Figure 4.3.

Figure 4.4 shows the ratio of the observed building seismic responses (measured by applying the set of ground motions to the numerical models of the buildings) to the seismic responses predicted by Equation 3.1. These plots serve as a metric for examining the accuracy of the prediction equations developed for each of the five studied buildings' seismic response. In each plot, both the ratios are grouped into four sets based on the magnitude of the earthquake under which the seismic response is measured. As we can see in Figures 4.5 and 4.6, magnitude is one the most important characteristics of an earthquake that affects a building's seismic response. As such, we would be able to identify any systematic bias in the predictions for a certain magnitude range by grouping the observations based on their magnitudes. The response ratios are plotted against distance following the same logic as the one behind grouping observations based on their magnitudes and since distance is also one of the main variables that affects a building's seismic response. Error bars are also plotted for 12 distance intervals. As can be seen, for most of the observations the average of the ratios of the actual to predicted response variable is within 0.8 to 1.5 with the ratios randomly scattered on both sides of $y = 1$ line. This indicates that the prediction equations developed for the five studies buildings have an acceptable degree of accuracy and do not show a systematic bias with respect to magnitude or distance when predicting the seismic

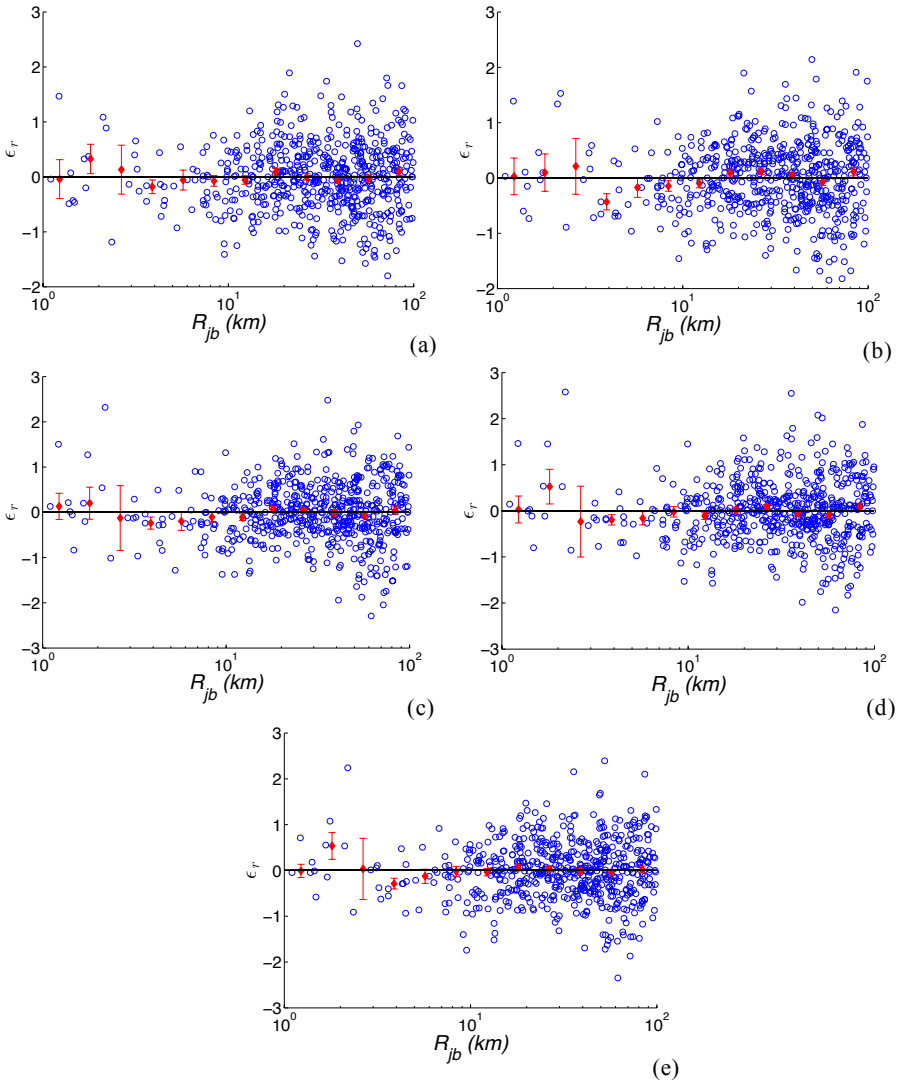

Figure 4.3: Record-to-record residuals (ϵ_r) versus earthquake distance for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

response of the case study buildings.

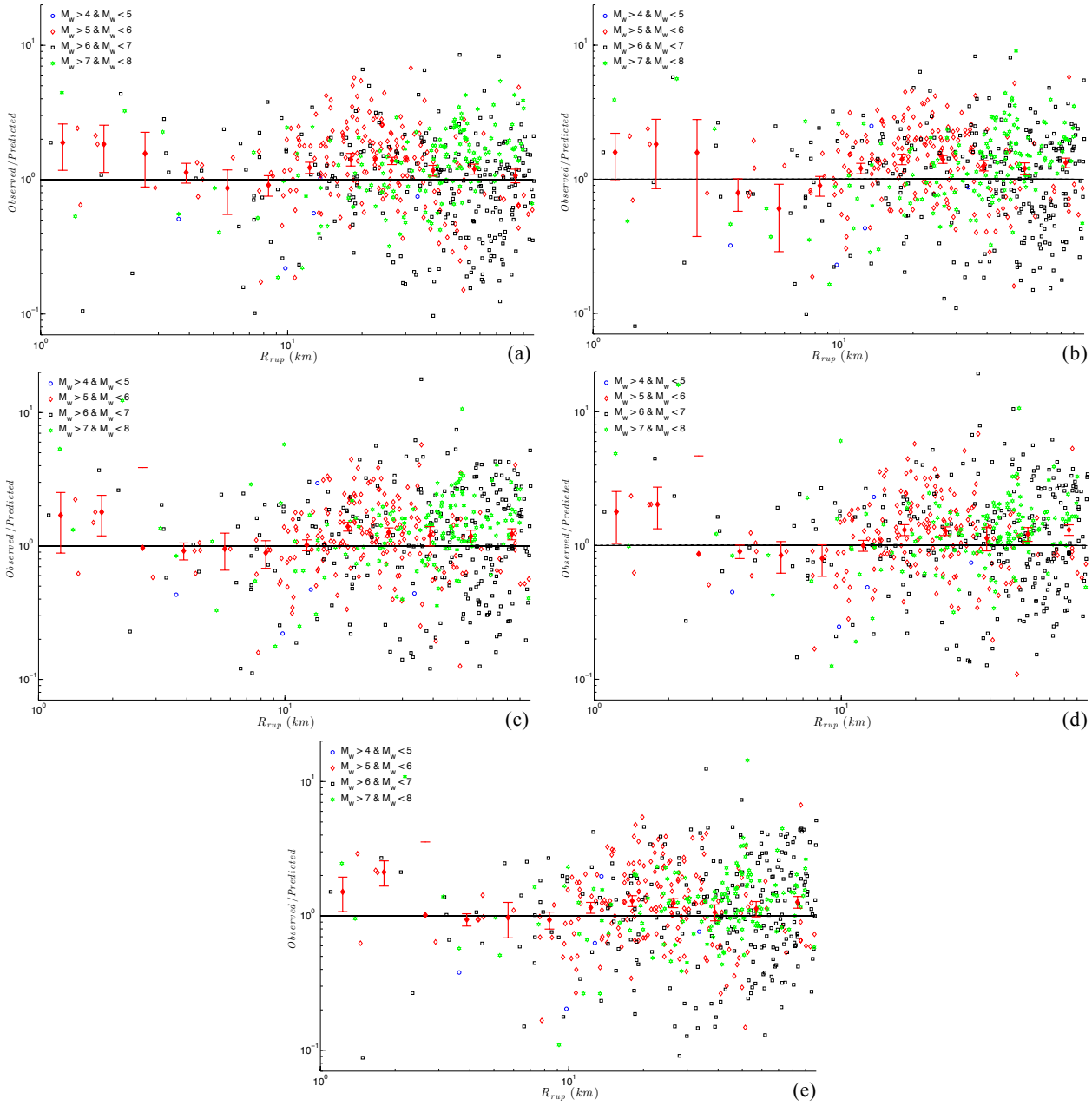

Figure 4.4: Ratio of observed to recorded building seismic response for four different magnitude bins and for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

As another metric for accuracy of the predictions of the buildings' seismic response, we plotted the maximum relative story displacement values (SDR) versus the distance from the

buildings' site to the rupture surface of the faults in Figure 4.5. It is expected that as the distance increases, the SDR values decrease as the waves generated by the earthquake would get more extensively attenuated as the distance between the building and the fault increases. Moreover, we would expect the seismic demand in the form of SDR values increase as the magnitude of the earthquake increases. Both trends can be seen in Figure 4.5 for all the five studied buildings. At distances bewlo 10 km, there is not a significant decline in SDR as the distance increases. This is an expected outcome since at distances very close to the fault large pulse-like velocity periods, which mostly depend on the magnitude of the earthquake rather than its distance, are the main factor in determining the building's seismic response. However, as distance exceeds 10 km, there is a clear decline in SDR as distance grows larger which, as explained above, is an expected trend. For each of the three magnitude values, the increase in the magnitude of the earthquake has a very notable impact on the building's seismic response; under a magnitude 7.5 event, the SDR value could be up to 10 times larger compared to a magnitude 5.5 event. The gap between the SDR values under the three magnitude values remain almost intact throughout the entire distance range of Figure 4.5.

Finally, Figure 4.6 shows how the relatonship between the predicted SDR values, distance and V_{s30}. Recall that V_{s30} is a measure for the type of the soil that the building is located on. As V_{s30} increases, the soil underneath the building becomes harder, i.e., the soil would get closer to rock. For very soft soils with low V_{s30} values, the soil can amplify the earthquake waves such that the final shaking that the building experiences after it is filtered by the soil profile that the building is constructed on could be much larger than the original shaking amplitude. As such, it is not surprising that, in Figure 4.6, there is an inverse relationship between SDR values and V_{s30} such that SDR values could be up to 5 times larger when $V_{s30} = 180\frac{m}{s}$ compared to when $V_{s30} = 760\frac{m}{s}$.

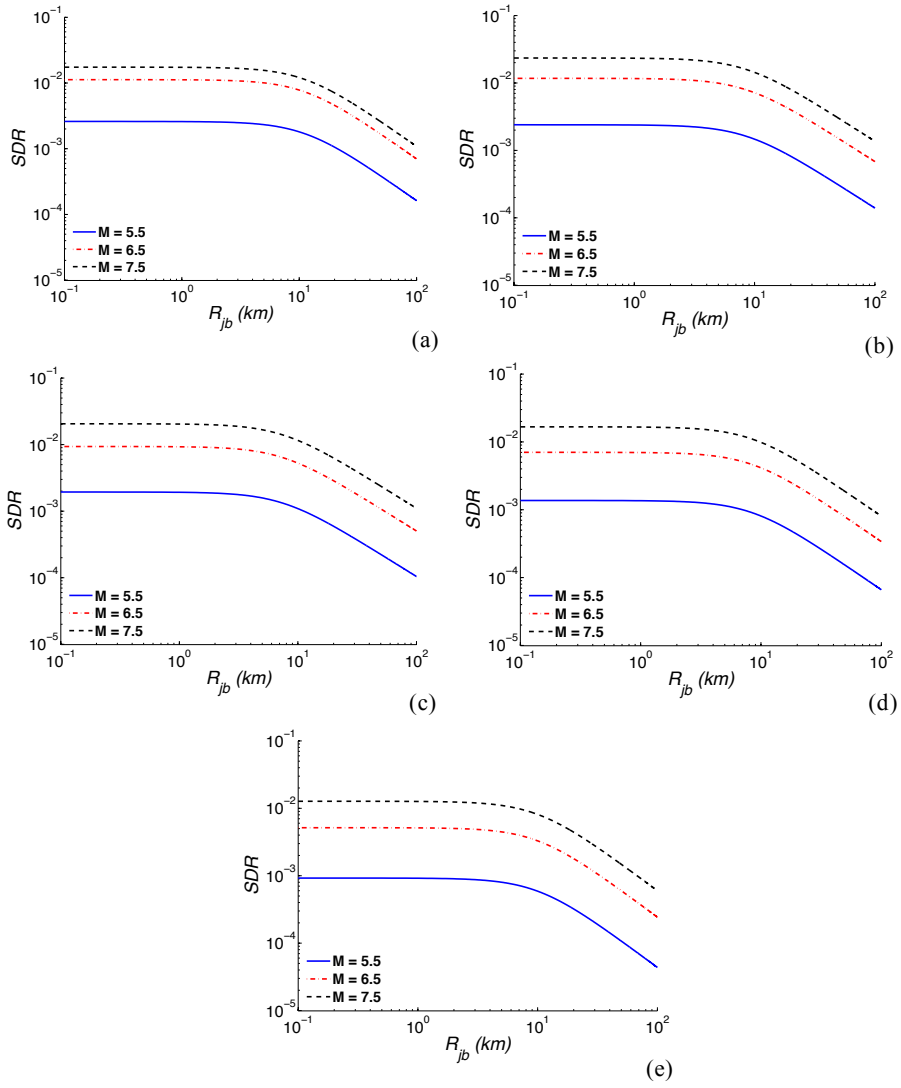

Figure 4.5: Trends between maximum relative story displacement and source-to-site distance for 3 magnitude bins and for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

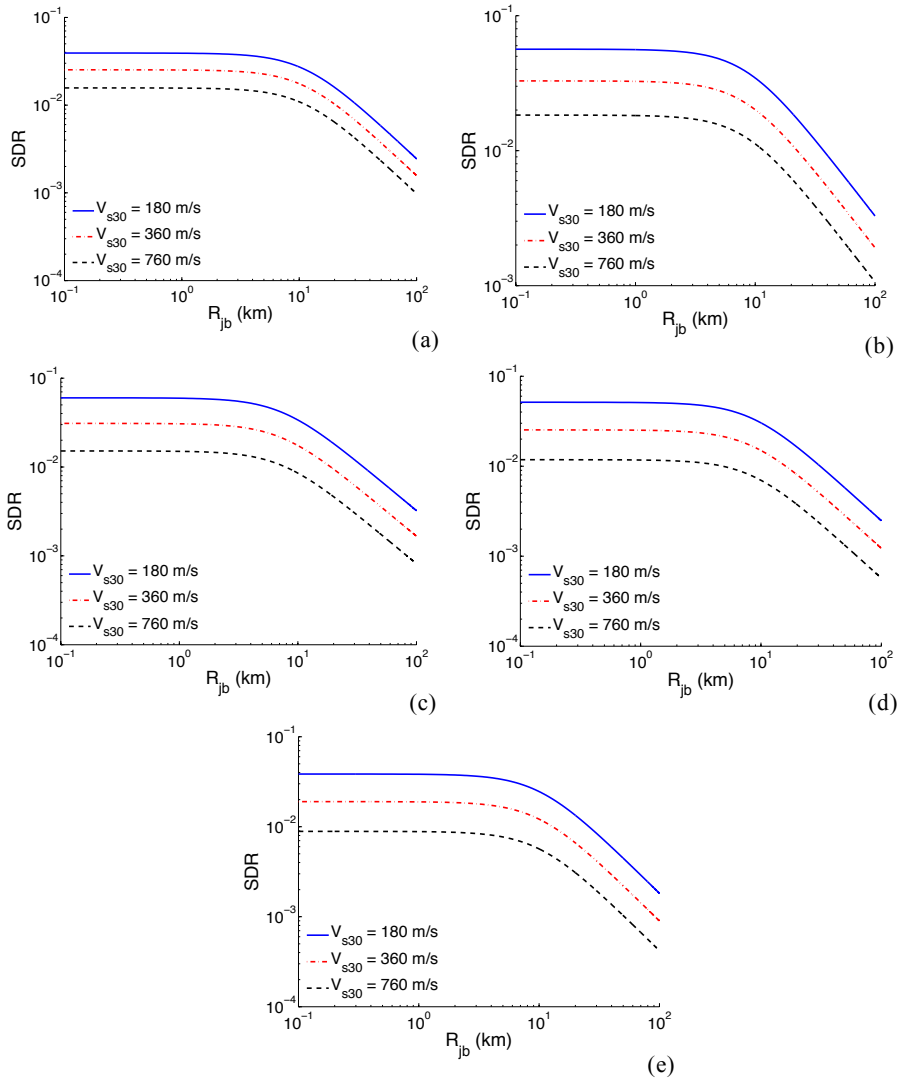

Figure 4.6: Trends between maximum relative story displacement and source-to-site distance for 3 different soil types and for a) 2-story, b) 4-story, c) 8-story, d) 12-story and e) 20-story buildings

CHAPTER 5

Conclusion

A key step in having a community that is resilient to earthquakes is the ability to accurately predict the seismic response of the buildings that the population of the community reside in. While the structural engineering community has witnessed a substantial increase in the accuracy of the numerical models employed in predicting building seismic response, the complexity of such models and the computational effort associated with simulating the building seismic response are usually the two main prohibitive factors in accurately predicting the building seismic response. The main focus of this study was placed on proposing a relationship that can be used in estimating the seismic response of reinforced concrete structures. This equation would establish a relationship between the key dynamic features of the building and the characteristics of earthquakes that have the most significant impact on the building's seismic response. Such a relationship can be used in obtaining an estimate of the extend of the demand that a building is likely to experience when it is subjected to an earthquake with a certain magnitude that has occured within a certain distance of the building.

We adopted a two-step maximum likelihood regression method when developing our prediction equation. The first stage of the regression incorporates the characteristics of earthquakes that vary from one seismic event to another; namely magnitude and the fault type. The second stage of the regression would provide a picture of how the parameters that could vary within a single seismic event (distance from the fault's rupture plane to individual buildings and the type of the soil the building is located on) would impact the building's seismic response response.

To ensure the unbiasedness of the proposed prediction equation with respect to the

predictors involved in the regression steps, we examined the residuals from both stages of the regression and no systematic bias in the prediction equations developed for each of the five case study buildings was found. To examine the accuracy of the predictions made using the proposed equations, we compared the impact that changes in the key parameters of the equation would have on the buildings' seismic response to what we were expecting from physics-based models and we found that the predicted response follows the expected trends.

REFERENCES

[1] S Atzori, C Tolomei, A Antonioli, JP Merryman Boncori, S Bannister, E Trasatti, P Pasquali, and S Salvi. The 2010-2011 canterbury, new zealand, seismic sequence: multiple source analysis from insar data and modeling. *Journal of Geophysical Research: Solid Earth*, 117(B8), 2012.

[2] M. Kazama and T. Noda. Damage statistics (summary of the 2011 off the pacific coast of tohoku earthquake damage). *Soils and Foundations*, 52(5):780–792, 2012.

[3] *A framework methodology for performance-based earthquake engineering*, 2004.

[4] Jack W. Baker and C. Allin Cornell. *Vector-valued ground motion intensity measures for probabilistic seismic demand analysis*. Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, 2006.

[5] Beno Gutenberg and Charles F. Richter. Frequency of earthquakes in california. *Bulletin of the Seismological Society of America*, 34(4):185–188, 1944.

[6] M Raghunandan, AB Liel, and N. Luco. Aftershock collapse vulnerability assessment of reinforced concrete frame structures. *Earthquake Engineering & Structural Dynamics*, 44:419–439, 2015.

[7] GL Yeo and C. Allin Cornell. *Stochastic characterization and decision bases under time-dependent aftershock risk in performance-based earthquake engineering*. Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, 2005.

[8] TL Holzer. *Implications for earthquake risk reduction in the United States from the Kocaeli, Turkey, earthquake of August 17, 1999*. US Geological Survey, Information Services, 2000.

[9] Curt B Haselton. *Assessing Seismic Collapse Safety of Modern Reinforced Concrete Moment Frame Buildings*. PhD thesis, Stanford University, 2007.

[10] American Concrete Institue (ACI). *ACI 318M-02: Building code requirements for structural concrete and commentary*. American Concrete Institue, Farmington Hills, MI, 2002.

[11] American Society of Civil Engineers (ASCE). *ASCE/SEI 7-05: Minimum design loads for buildings and other structures*. American Society of Civil Engineers, Reston, VA, 2005.

[12] Silvia Mazzoni, Frank McKenna, and Gregory L. Fenves. *OpenSees command language manual*. Pacific Earthquake Engineering Research Center, University of California, Berkeley, 2006.

[13] L.F. Ibarra, R.A. Medina, and H. Krawinkler. Hysteretic models that incorporate strength and stiffness deterioration. *Earthquake engineering & structural dynamics*, 34(12):1489–1511, 2005.

[14] T.D. Ancheta, R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, and T. Kishida. Nga-west2 database. *Earthquake Spectra*, 30(3):989–1005, 2014.

[15] Katsuichiro Goda. Record selection for aftershock incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics*, 44(7):1157–1162, 2015.

[16] J Ruiz-Garca. Mainshock-aftershock ground motion features and their influence in building's seismic response. *Journal of Earthquake Engineering*, 16(5):719–737, 2012.

[17] Norman Alan Abrahamson, Joseph Silva Walter, and Ronnie Kamai. *Update of the AS08 Ground-Motion Prediction equations based on the NGA-west2 data set.* Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, 2013.

[18] BrianS-J. Chiou and Robert R. Youngs. An nga model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, 24(1):173–215, 2008.

[19] Kathryn E. Wooddell and Norman A. Abrahamson. Classification of main shocks and aftershocks in the nga-west2 database. *Earthquake Spectra*, 30(3):1257–1267, 2014.

[20] Kenneth W. Campbell and Yousef Bozorgnia. Nga-west2 ground motion model for the average horizontal components of pga, pgv, and 5 *Earthquake Spectra*, 30(3):1087–1115, 2014.

[21] David M. Boore, Stewart Jonathan P., Seyhan Emel, , and Gail M. Atkinson. Nga-west2 equations for predicting pga, pgv, and 5shallow crustal earthquakes. *Earthquake Spectra*, 30(3):1057–1085, 2014.

[22] Brian S-J. Chiou and Robert R. Youngs. Update of the chiou and youngs nga model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra*, 30(3):1117–1153, 2014.

[23] Jonathan Hancock, Julian J. Bommer, and Peter J. Stafford. Numbers of scaled and matched accelerograms required for inelastic dynamic analyses. *Earthquake Engineering & Structural Dynamics*, 37(14):1585–1607, 2008.

[24] Yoshimitsu Fukushima and Teiji Tanaka. A new attenuation relation for peak horizontal acceleration of strong earthquake ground motion in japan. *Bulletin of the Seismological Society of America*, 80(4):757–783, 1990.

[25] David R. Brillinger and Haiganoush K. Preisler. An exploratory analysis of the joyner-boore attenuation data. *Bulletin of the Seismological Society of America*, 74(4):1441–1450, 1984.

[26] William B. Joyner and David M. Boore. Methods for regression analysis of strong-motion data. *Bulletin of the Seismological Society of America*, 83(2):469–487, 1993.