Title
Performance-Based Analytics-Driven Seismic Retrofit of Woodframe Buildings

Author
Yi, Zhengxiang

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Performance-Based Analytics-Driven Seismic Retrofit of Woodframe Buildings

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy in Civil Engineering

by

Zhengxiang Yi

2020

ABSTRACT OF THE DISSERTATION

Performance-Based Analytics-Driven Seismic Retrofit of Woodframe Buildings

by

Zhengxiang Yi

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2020

Professor Henry Burton, Chair

Woodframe construction is commonly used for single and multifamily residential buildings in the United States. In many parts of California, multifamily woodframe residential buildings are constructed with open first stories, which have much less strength and stiffness compared to the ones above. In older single-family residences, the "crawl space" is constructed with unbraced and unbolted cripple walls. Both these conditions lead to a soft-story response during seismic loading, resulting significant damage, economic losses and even collapse. This type of vulnerability is often addressed through seismic retrofits, which can be mandated by local jurisdictions (e.g., the Los Angeles Soft-Story Ordinance) or incentivized by state or local entities (e.g., the California Earthquake Authority Brace and Bolt Program). A key challenge in implementing these retrofit programs (mandated or incentivized) is quantifying the improvements in performance at the individual and portfolio scale and creating design procedures that maximize the overall benefit. This research integrates nonlinear structural modeling, performance-based assessments and

advanced statistical and machine learning techniques to quantify the benefit of soft-story woodframe building retrofit and develop optimal design solutions that maximize regional performance.

The considered construction types include single-family houses with unbraced cripple walls developed as part of the recently completed Pacific Earthquake Engineering Research Institute (PEER) and California Earthquake Authority (CEA) project and multi-family residences with soft, weak and open front wall lines (SWOF). An end-to-end computational platform is developed to automate the construction and analysis of archetype numerical models in OpenSees and conduct seismic evaluations based on the PEER performance-based earthquake engineering framework. The performance of existing and retrofitted buildings is assessed in terms of collapse safety and direct (due to earthquake damage) economic losses.

The effect of retrofit and various structural characteristics is illuminated for the single-family cripple wall houses. 2^k full factorial experiment design combined with hypothesis testing is used to identify the most influential structural properties. Two story buildings performed worse than their one-story counterparts and pre-1945 buildings performed better than pre-1955 construction. Building performance is found to be positively correlated with cripple wall heights and cripple wall retrofits provided significant overall improvements. Surrogate models are developed as a compact statistical link between key structural characteristics and seismic performance. Several machine learning algorithms are investigated for predicting the building median collapse intensity and expected annual loss using the cripple wall height, seismic weight, damping ratio and material properties as features. The XGBoost algorithm provides the most accurate prediction and on average, limits the prediction error to less than 10%. Using the well-developed machine learning

models, additional sensitivity analyses are conducted and the effect of model uncertainty on collapse safety and expected annual losses is quantified using Monte Carlo simulation.

For the SWOF buildings, a multi-scale cost-benefit analysis of the Los Angeles Soft-Story Ordinance Retrofit is performed. Individual buildings take an average of four to five years for the reduced earthquake losses to exceed the one-time retrofit cost. At the portfolio-scale, the average cost-benefit ratio is found to be 0.32 for the hypothetical M 7.1 Puente Hills scenario earthquake. A stochastic event-set cost-benefit assessment is also performed, where all events (approximately 8,000) that are significant to the region are considered. From this assessment, it is determined that the probability of achieving a desirable cost-benefit ratio (value between 0.0 and 1.0) within a 50-year period is approximately 0.9.

Lastly, a retrofit design optimization framework is proposed with the goal of maximizing performance-based benefits at the regional scale. The methodology relies a machine learning-based surrogate model to predict seismic performances of retrofitted buildings given the design parameters. Then, a stochastic optimization algorithm is implemented to find the retrofit designs that maximize the improvement in seismic performance for the entire portfolio under a set of pre-defined constraints. The algorithmic retrofit leads to collapse losses that are comparable to the Los Angeles Ordinance guidelines while using only 60% of the resources. The performance-oriented framework is shown to address the inefficiency of conventional strength-based retrofit policies.

The dissertation of Zhengxiang Yi is approved.

Yousef Bozorgnia

Gregory G. Deierlein

Jian Zhang

Henry Burton, Committee Chair

University of California, Los Angeles

2020

Table of Contents

List of Figures

List of Tables

Acknowledgement

The research presented in the thesis is supported by Pacific Earthquake Engineering Research Center (PEER) – The California Earthquake Authority (CEA) Project "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-frame Buildings" and National Science Foundation Division of Civil, Mechanical and Manufacturing Innovation research grant number 1538747. Any opinions, findings, and conclusions or recommendations expressed in this dissertation are those of the authors and do not necessarily reflect the views of the sponsor.

I would like to express my sincere gratitude to my Ph.D. advisor, Dr. Henry V. Burton, for his passion of investigating new possibilities, continuous support and encouragement. I would also like to thank Dr. Yousef Bozorgnia, Dr. Gregory G. Deierlein and Dr. Jian Zhang for their dedicated guidance.

Besides, I really appreciate all my supportive and considerate friends and their delicious dishes. Hope we all can live a skinny and healthy life in the future. Lastly, I want to deliver my special thanks to my parents for their trust, support and unconditional love. I cannot imagine I could graduate with a doctoral degree without my family.

VITA

2012 – 2016	B.Sc. in Civil Engineering, Tongji University, Shanghai, China
2016 – 2017	M.Sc. in Civil Engineering, University of California, Los Angeles
2017 – 2020	Ph.D. Candidate in Civil Engineering, University of California, Los Angeles

CHAPTER 1. INTRODUCTION

1.1. Motivation and Background

Buildings constructed with adjacent stories having large differences in strength and stiffness can lead to the formation of a single-story mechanism during earthquake shaking. This type of behavior is undesirable because inelastic deformation is concentrated in a few non-ductile components of a single story instead of being distributed along the height of the building, which can lead to sidesway collapse. In California, soft- and weak-story wood frame construction is prevalent in both single- and multi-family residential buildings. In multi-story apartment buildings, some or all of the first story is often used for parking or commercial space leading to less walls and partitions compared to the upper stories. Single-family residences with unbraced and unbolted cripple walls (surrounding the crawl space) are also susceptible to soft/weak-story failure.

The vulnerability of soft and weak story wood frame buildings has been highlighted in several historical events, including as far back as the 1971 San Fernando earthquake (FEMA, 2012). Since then, wood frame building collapses attributed to the formation of soft-story mechanisms have occurred during the 1989 Lomo Prieta (Harris and Egan 1992) and 1994 Northridge earthquakes, where soft-story damage to apartment buildings resulted in tens of thousands of unoccupiable residential units (Holmes and Sommers 1996). Besides the human casualties and loss of housing, the concentration of collapsed and irreparably damaged soft-story buildings within a community can slow the overall pace of post-earthquake recovery (Comerio, 2006).

Policy actions are often implemented to mitigate the effects soft-story vulnerabilities. The cities of Los Angeles and San Francisco have enacted Ordinances to mandate the retrofit of soft-weak and open-front (SWOF) multi-family residential buildings. For single-family residences, the earthquake brace and bolt program (https://www.earthquakebracebolt.com/) provides up to $3,000

to help pay the cost of cripple wall retrofits. Guidelines such as the FEMA P807 (FEMA, 2012a) and FEMA P50 (FEMA 2012b) documents have sought to develop systematic procedures for seismic retrofit and performance evaluation of soft/weak-story structures. A major challenge with developing such guidelines is being able to adequately capture the variations in structural configurations and material properties of the affected structures. Furthermore, multiple sources of uncertainties (i.e. structure dimensions, construction quality) must be addressed when developing retrofit techniques than can be generalized for a given portfolio. Lastly, the tools, methods and guidelines that are currently available for designing retrofits to provide adequate strength and inelastic deformation capacity. However, within the framework of performance-based seismic design (PBSD) (Moehle and Deierlein 2004), little or no consideration is given to earthquake-induced economic losses.

To address the challenges associated with developing portfolio-based seismic retrofit solutions, a performance-based analytical-driven (PBAD) methodology is proposed. The key elements of PBAD method include an (i) end-to-end computational platform that links retrofit design, nonlinear response simulation and economic loss assessment, (ii) surrogate models for robust uncertainty quantification and performance exploration in high-dimensional variable space and (iii) strategies for achieving optimal retrofits based on portfolio-scale performance objectives.

1.2. Research Significance

The main goal of the proposed research is to develop the tools, models and methods for implementing optimal portfolio-based retrofit strategies for wood frame structures. The detailed objectives are listed as follows:

- Develop an “end-to-end” computational platform that links seismic retrofit, nonlinear response analyses and earthquake-induced losses of wood frame structures.

- Illuminate the relationship between the key structural parameters and the seismic response and earthquake-induced losses in wood frame structures. The goal here is to quantitively describe how different building features (e.g. cripple wall height, roof weight, and structure material parameters) influence building performance (e.g. stricture collapse capacity, engineering demand parameters, and expected losses).
- Introduce surrogate models to quantify the uncertainty in the seismic performance and earthquake-induced losses in wood-frame structures. The surrogate models will enable robust uncertainty propagation and performance exploration in high-dimensional variable spaces while minimizing the number of explicit (or mechanics-based) (e.g. nonlinear response history, loss assessment) simulations.
- Formulate a framework for realizing optimal portfolio-based retrofit schemes for wood frame structures.

A schematic representation of the models, tools and methods developed as part of the proposed scope of work is shown in Figure 1.1. The entire body of work can be roughly divided into five modules. The "modeling module" is where the end-to-end tool for constructing nonlinear models for structural analysis and performance-based assessment is developed. The input building variants and output seismic performances are used to develop a woodframe building performance database. The "analysis module" uses the database for multi-scale seismic performance quantification, sensitivity analysis, uncertainty quantification and cost-benefit analysis. Machine learning (ML) surrogate models are developed using the structural responses and performance assessment results contained in the database. These ML-based surrogate models can either utilizes in the analysis module or optimization module, where the latter is used to find the most desirable retrofit scheme based on a predefined portfolio-scale performance objective. The effectiveness and efficiency of

this retrofit scheme is assessed within the "evaluation module". Finally, a retrofit policy is proposed based on the results of the optimization and evaluation modules. Multi-scale seismic performance assessment and the development and application of analytics-driven methods are the core contribution of this dissertation.

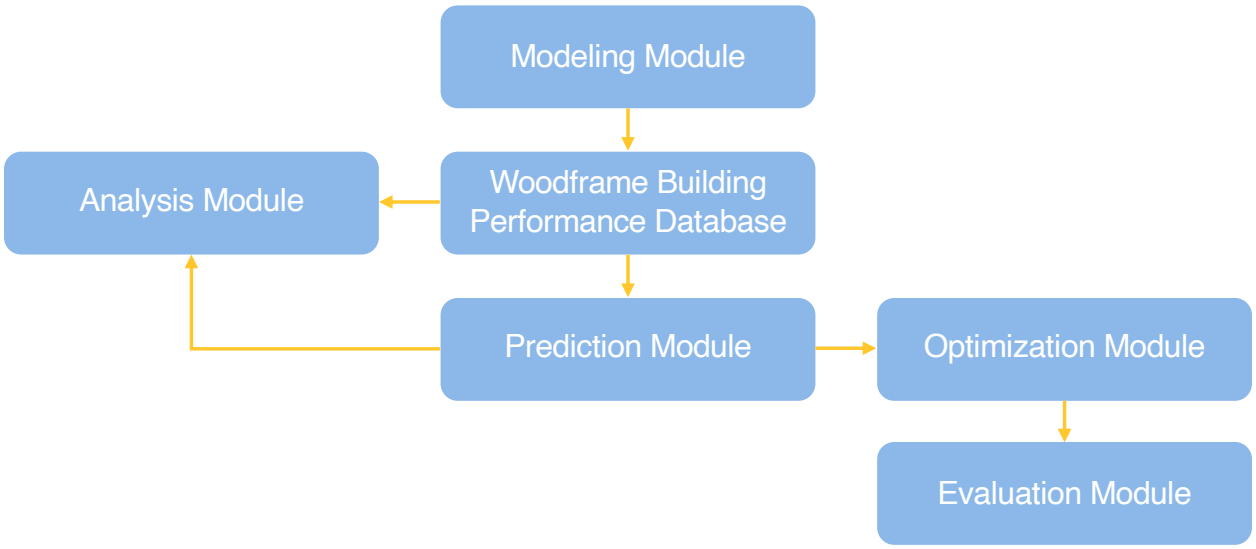


Figure 1.1 Schematic representation of the relationships among the models, methods and tools developed as part of the proposed scope of work

1.3. Organization and Outline

This thesis consists of seven chapters and the content of each chapter is elaborated below:

Chapter 1 discusses the background, motivation, and objective of the research.

Chapter 2 starts by providing a brief summary of prior studies on woodframe building numerical analysis and performance assessment. The end-to-end Python-based tool to generate mechanics-based models in OpenSEES (McKenna et al. 2003), perform nonlinear numerical analysis and conduct post-processing and FEMA-P58 (2012) based loss assessment is then introduced. Details of the modeling and analysis procedures for single-family unbraced and unbolted cripple multi-family soft-story woodframe buildings are presented. Some sample analyses results are presented for the purpose of demonstrating the modeling and analysis procedures.

Chapter 3 provides an overview of ML algorithms and their applications in solving structural engineering problems. Multiple ML models are constructed to predict building collapse safety and expected annual losses (EAL). A formal machine learning model derivation process for structural engineering regression problems is presented alongside the model development.

Chapter 4 discusses the nonlinear structural response simulation, performance-based assessment, sensitivity analyses and uncertainty propagation for collapse risk and loss estimations of single family woodframe buildings. A preliminary sensitivity analysis study is first introduced using the full factorial experimental design for isolating the effects of different parameters on the seismic performances of single-family cripple wall buildings. Then, the ML models developed in Chapter 3 are adopted for a more comprehensive sensitivity analysis. Lastly, the effect of model uncertainty on collapse risk and expected annual losses is assessed.

Chapter 5 investigates various issues related to multi-family woodframe residences with soft, weak and open-front wall lines (referred to as SWOF buildings in the remainder of the dissertation). The effect of alternative retrofit schemes and key structural parameters on seismic performance are evaluated. A multi-scale cost-benefit analysis of SWOF building seismic retrofit is performed.

Chapter 6 discusses a comprehensive framework for achieving optimal portfolio-based retrofit schemes for wood frame structures. The framework produces optimal retrofit designs that achieve the most desirable regional performance objective. The framework is implemented on the SWOF building inventory that is under the purview of the Los Angeles Soft-Story Ordinance.

Chapter 7 summarizes the findings from the various studies and discusses ideas for future work.

CHAPTER 2. WOODFRAME BUILDING MODEL DEVELOPMENT AND PERFORMANCE ASSESSMENT

2.1. Literature Review

This section provides a detailed review of the development of woodframe building finite element modeling and performance assessment.

2.1.1. Woodframe Building Modeling

Three-dimensional models of woodframe buildings are developed as an aggregation of individual components. The major components considered in woodframe structures include the walls, floors and roof. These components are connected by intercomponent fasteners and connectors (i.e. nails and metal plates) (Kasal et al. 1994). The physical properties of wood, gaps between components and connections, connector properties govern the nonlinear behavior of woodframe structures. Among all components, the shear walls serve as the primary lateral force resisting system (Folz and Filiatrault 2001) and therefore have a significant impact on the performance of the entire structure.

There is a large body of research on modeling woodframe shear walls, which comprise the framing, sheathing and sheathing-to-framing connectors. Lateral forces are resisted by sheathing materials and transferred to framing members by the connectors. The design of woodframe structures have been informed by experiments involving monotonic or reversed-cyclic testing (Stewart 1987; Dolan and Madsen 1992; Yamaguchi 1998). Through these experiments, researchers have reached a consensus on characterizing the hysteretic behavior of wood shear walls. First, the force-deformation behavior of the material combination is nonlinear under monotonic loading and experiences strength and stiffness degradation under cyclic loading (Folz and Filiatrault 2002). The global nonlinear response of the shear wall is similar to that of the individual sheathing-to-

framing connector used in the wall (Dolan and Madsen 1992; Folz and Filiatrault 2001). Also, the hysteretic behavior of wood shear walls is superimposable when different materials are used on the interior and exterior (Folz and Filiatrault 2004; Pang et al. 2007; Van de Lindt et al. 2007). Multiple hysteretic models had been proposed to capture the load-deformation relationship for shear walls and sheathing-to-framing connectors. In the 1980s, piecewise linear models for different stages of the loading-unloading curve were meticulously fitted to experimental data (Polensek and Laursen 1984; Stewart 1987; Dolan and Madsen 1992). Then, the Bouc-Wen-Baber-Noori (BWBN) model, which is a single element pinching model generally used for modeling steel and concrete structure under random vibration load, was introduced to timber structures by Foliente (1995). The Consortium of Universities for Research in Earthquake Engineering (CUREE) – Caltech Woodframe project was initiated in 1998 with the goal of reducing woodframe construction earthquake induced losses. Comprehensive shake table tests were conducted to investigate the damage characteristics and seismic behavior of wall finish materials, connections, anchorages, single shear walls, diaphragms, foundations, cripple walls and full-scale houses. A computing software, cyclic analysis of wood shear walls (CASHEW), specialized in nonlinear seismic analysis of woodframe buildings was developed as part of that project. Folz and Filiatrault (2001) proposed using a 10-parameter SDOF hysteretic model for capturing force-deformation responses of wood shear walls with nailed sheathing-to-framing connector. The modeled cyclic response is shown in Figure 2.1. The material monotonic force-deformation response is defined based on the initial stiffness, yielding strength, two stiffness reduction factors, displacement at peak strength and residual displacement. An additional four parameters were introduced to capture responses under cyclic loading. The proposed model was validated using a large body of test data (Folz and Filiatrault 2002, 2004; Gatto and Uang 2002).

The seismic analysis of woodframe structures (SAWS) hysteretic model was incorporated in CASHEW. The SAWS model had been broadly used in light frame wood structure seismic response analysis. Christovasilis et al. (2009) used SAWS models for the index buildings developed in CUREE-Caltech project and performed incremental dynamic analysis (IDA) to assess building collapse safety. Pei and Van de Lindt (2011) delivered a 'blind-prediction' analysis between the numerical simulated seismic responses of a six-story light frame wood building modeled using SAWS with the testing results. Good agreements between the experimental and numerical responses were obtained for multiple engineering demand parameters (EDP). Bahmani et al. (2016) created numerical models using SAWS hysteretic springs for a full scale four-story soft-story building and evaluated its seismic performance. Even today, the SAWS model is still one of the most favorable hysteretic models for wood shear wall dynamic response modeling. Based on the SAWS model, Pang et al. 2007 proposed the evolutionary parameter hysteretic model (EPHM), which adopts exponential functions for the backbone curve instead of the linear functions used in the SAWS model. The NEESWood project which began in 2005 sought to investigate the dynamic behavior of woodframe constructions under seismic loading and develop a performance-based seismic design philosophy for mid-rise woodframe buildings. Full-scale shake table tests and numerical analyses were conducted for a benchmark two-story residential building. The Seismic Analysis Package for Woodframe Structures (SAPWood) was developed as part of the NEESWood project (Pei and van de Lindt 2010a). The software is capable of conducting nonlinear analysis, incremental dynamic analysis (IDA), and loss assessment for bi-axial structure models, simplified MDOF models, and triaxial shear and bending models. Both the SAWS and EPHM models are included in the SAPWood software. Furthermore, the Nail Pattern (NP) module of SAPWood allows users to model structures at the fastener level. van de Lindt et al. (2010)

presented a numerical analysis for a full-scale woodframe building to verify the SAPWood model with testing data.

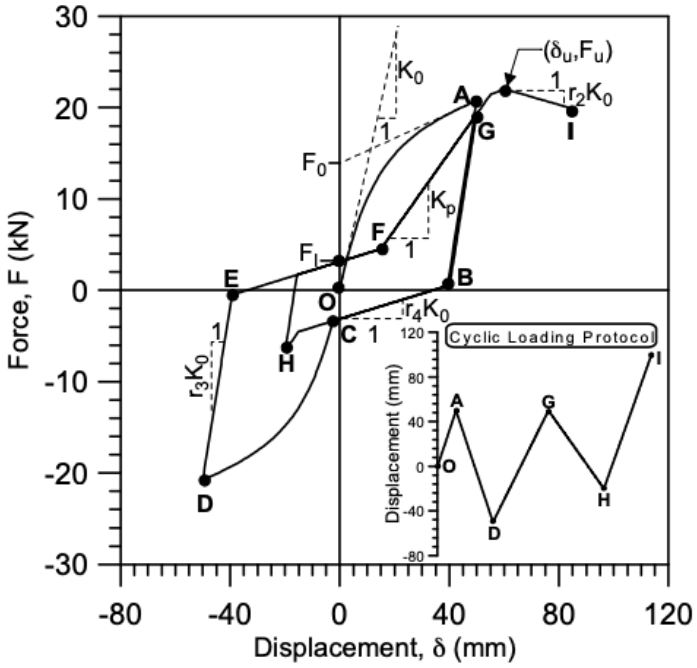


Figure 2.1 SAWS hysteretic model (figure adapted from (Folz and Filiatrault 2002))

The floors and roof are the other components considered in woodframe building structural modeling. Diaphragms are mainly used to transfer lateral loads to the lateral force resisting system. Similar to wood shear walls, wood diaphragm properties are influenced by the sheathing type and orientation, fastener type and spacing, and diaphragm dimensions (Falk and Itani 1989). Amana (1967) first introduced the nail modules to account for fastener stiffness. In the 1980s, Itani and Cheung (Cheung and Itani 1983; Itani and Cheung 1983, 1984) developed finite element models for predicting static and dynamic responses of diaphragms considering the distribution and stiffness of sheathing-to-framing fasteners. Those models have shown great agreement with test data. Further, investigations into diaphragm flexibility were made in the 1990s and 2000s. Through performing 330 numerical analyses on rigid and flexible diaphragms in one-story timber structures,

Dolce et al. (1992) proposed that if the vertical stiffness is uniformly distributed, the diaphragm flexibility will not significantly impact the nonlinear responses. Folz and Filiatrault (2001) performed 14 experiments on different diaphragms and structure configurations to study the in-plane rigidity of the diaphragm. Owing to the fact that nails provide diaphragm with enough strength and the vertical wood shear wall can be regarded as additional chord members, small rotation and rigid diaphragm assumptions generally hold when modeling the entire structure. However, the results from the NEESWood project suggested that the rigid diaphragm assumption may not be valid under some circumstances. Pang et al. (2012) also argued that these assumptions will lead to inaccurate boundary conditions and would be unable to predict incipient collapse properties of wood buildings. They proposed a new 3D frame element modeling method incorporating corotational formulation and large displacement theory to address the problems that arise from overly simplistic assumptions. In the analysis, SAPWood permits the consideration of connector stiffnesses to model semi-rigid or flexible diaphragms.

2.1.2. Woodframe Building Performance Assessment

The current state-of-art design philosophy advocates for the use of performance targets. Specifically, in the performance-based earthquake engineering (PBEE) framework, structures are supposed to be designed with enough strength to resist earthquake load and meeting target performance specifications. A brief introduction into the evolution of PBEE and its applications to woodframe buildings is presented in this section.

The origin of PBEE can be traced back to displacement-based design. In the 1990s, the direct-displacement design method was first developed by Priestley (Priestley 1993, 1999, 2000) and applied to concrete structures. Then, the Structural Engineering Association of California (SEAOC 1996) and NEHRP Guidelines (Building Seismic Safety Council 1997) proposed the concept of

seismic design performance levels measured by drift limits. Different from traditional force-based design, these methodologies advocated for using drift as the basis for measuring performance. FEMA 273 (1997) and ATC-40 (1996) represent the first generation of PBEE assessment and design procedures. Within the 1st generation of the PBEE framework, buildings subjected to earthquake lateral loading experience nonlinear responses and results in different levels of damages. Quantitative measurements of component-level damage can be established on the basis of dynamic responses. Building performance targets such as immediate occupancy, life safety and collapse prevention were introduced. The performance of the entire structure was based on the worst-case-scenario for component-level damage. However, in this era, the engineering demands were mostly based on simplified analysis procedures which lacked calibration. No consistent and reliable component data was available to define the mapping from demands to damage. To address these issues, the Pacific Earthquake Engineering Research Center (PEER) initiated a research program to develop a more systematic and robust methodology for PBEE. In the 2nd generation of PBEE, probabilistic analysis procedures were developed. The overall methodology included hazard assessment, structural analysis, damage evaluation and loss estimation. The final outcome of the framework is a set of decision variables (DV), which represent the probability aggregation of engineering demand parameters (EDP), damage variables (DM) and intensity measures (IM) as illustrated by Equation (2.1) (Cornell and Krawinkler 2000; Krawinkler and Miranda 2004). The details of the methodology are provided in FEMA P-58. The discussion and analysis of woodframe building performances for the entire thesis will be based on FEMA P-58.

$$v(DV) = \int\int\int G(DV|DM)dG(DM|EDP)dG(EDP|IM)d\lambda(IM) \quad (2.1)$$

Prior studies on performance-based assessment of woodframe buildings is presented in this section. Filiatrault and Folz (2002) summarized the detailed procedure for conducting displacement-based

design and applied numerical analysis to verify the process. Rosowsky and Ellingwood (2002) provided an overview of performance-based design and its application to residential woodframe buildings. Performance requirements, limit state definition, system reliability and fragility modeling were examined in detail. Van de Lindt (2005) first proposed a reliability-based approach to calibrating a seismic damage model and quantify wood-frame building performance. Five component level criteria, including nailing behavior and sheathing out-of-plane movement, were used to define limit states. Fastener spacing was selected as the primary design variable, which was then regressed against the Park-Ang damage. The probability of failure calibrated to a lognormal distribution. Kim and Rosowsky (2005) applied fragility analysis to light-frame wood shear walls using CASHEW and developed the technical basis for the seismic design. The application of a new generation of PBSD to low- and mid-rise light-frame wood buildings was investigated in the NEESWood project. At this stage, building performance on the basis of economic losses was explored. Crowley and Bommer (2006) developed loss exceedance curves in Italy using the probabilistic seismic hazard analysis (PSHA) method and Monte Carlo simulation. Pei and van de Lindt (2010) proposed a probabilistic framework to estimate long-term earthquake related economic loss for wood frame buildings. A vulnerability-based approach was introduced to quantify model uncertainties, and Bayesian inferencing was used to improve the fragility prediction accuracy. Four levels of component damage were defined, including non-detectable, repairable, borderline repairable and not repairable/replace. Li and Ellingwood (2009) developed a multi-hazard performance evaluation and mitigation framework for woodframe residential building construction. Building hazard risk was quantified based on the probabilities of being in different damage states, which is a function of demand and capacity. The performance of two typical one-story woodframe residential buildings with minimum and strengthened hazard-

resistant systems were evaluated under wind and seismic loading to compare the effectiveness of different hazard mitigation strategies. van de Lindt et al. (2013) summarized the outcomes of the NEESWood project and proposed a new performance-based design procedure. The design procedure is an iterative process where direct displacement design is performed, the design detailing and code requirements are evaluated for a given building plan, target performance and hazard level. The procedure was validated on a 6-story woodframe building used in the NEESWood capstone full-scale testing.

2.2. Python Based End-to-End Tool for Nonlinear Structural Analysis and Performance-Based Assessment of Woodframe Structures

An efficient end-to-end Python based tool is developed to construct and analyse 3D woodframe building numerical models in OpenSees and conduct performance-based assessments. The tool is capable of constructing OpenSees 3D models given the necessary structural parameter inputs, performing nonlinear static and dynamic analysis and evaluating seismic performances. The detailed modeling considerations and structural analysis procedures are introduced in this chapter. Modeling is discussed for single- and multi-family houses. The baseline single-family houses developed as part of the PEER-CEA project (Welch and Deierlein 2020) and multi-family SWOF buildings under the purview of the Los Angeles ordinance are used to demonstrate the modeling strategies. Two sample analysis results are presented to show the capability of the platform.

An end-to-end workflow was programmed in Python to fulfill the large demand of woodframe building nonlinear modeling, analyses, and performance assessment. Figure 2.2 demonstrates the entire process. Overall, the tool takes in information about the structure configuration, material properties, analysis parameters and ground motions as input to generate 3D OpenSees (McKenna et al. 2003) models, and perform nonlinear analysis and performance-based assessment to obtain

EDPs, collapse performance and losses. Three different sets of OpenSees models are generated by the tool for a given set of inputs, which are used for eigen, nonlinear static and nonlinear dynamic analysis. The analysis and post-processor can conduct the desired analyses and data cleaning on a local machine or high performance clusters, such as the UCLA Hoffman2 and DesignSafe (Rathje et al. 2017). The analysis and post-processor are capable of parallel computing to speed up the nonlinear analysis. EDPs, collapse safety and losses can be extracted from the raw results as the final outcome of the end-to-end simulation. Furthermore, the input-output pairs constitute a woodframe building performance database. The database can be used for sensitivity analysis, uncertainty quantification and the development and verification of analytics-driven models. The entire platform was designed to be generic for any types of woodframe structures.

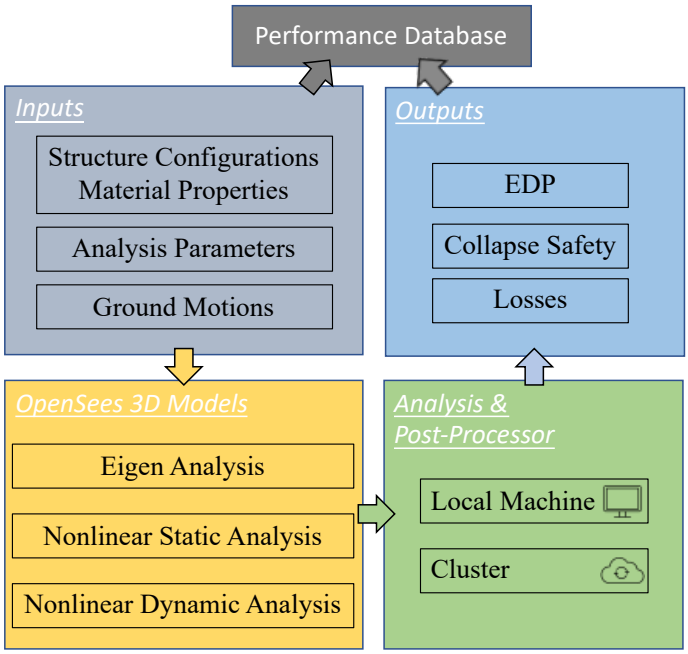


Figure 2.2 Workflow of the Python-based end-to-end woodframe building modeling and performance assessment tool

Figure 2.3 provides a schematic representation of a 3D OpenSees model of a woodframe building. It shows the general composition of a model. The considered components include the diaphragms/roof, shear walls (interior, exterior and cripple wall), leaning columns, and moment

frames (only applied in retrofitted cases). The wall elements are idealized as hysteretic springs representing the location and geometry of the wall. The material properties are either calibrated from experimental results or numerical nonlinear analysis under cyclic loadings. Rigid diaphragms are assumed at each story level. The detailed modeling process and analysis settings are demonstrated using prototype single- and multi-family houses.

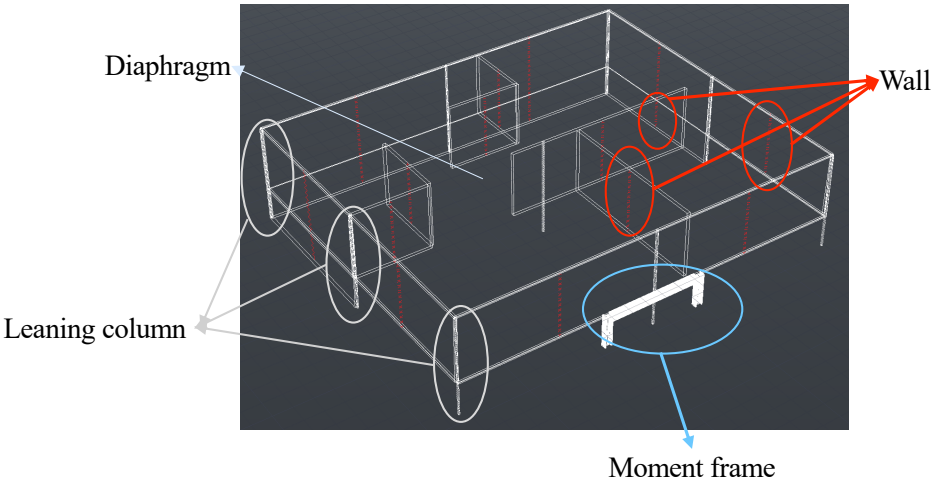


Figure 2.3 Schematic figure of 3D OpenSees model of a woodframe building

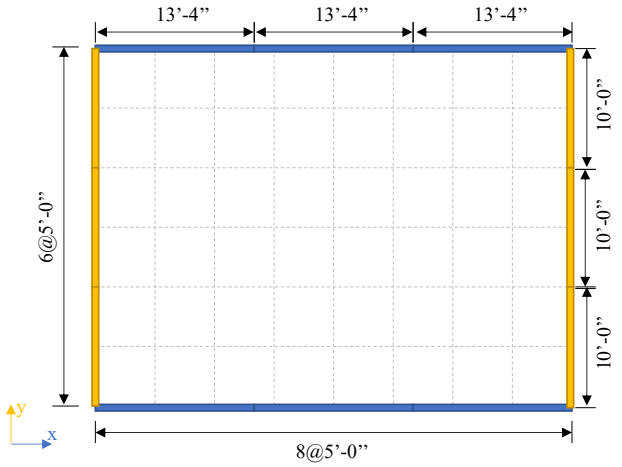
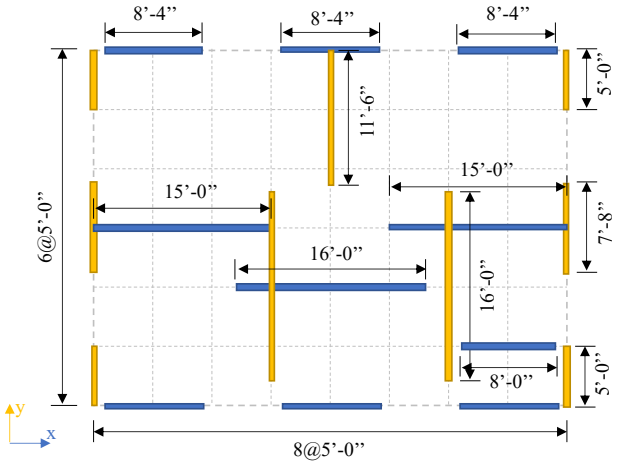
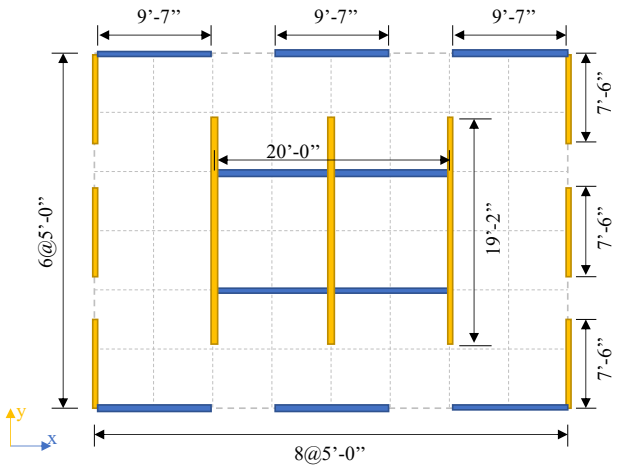
2.3. General Description of the Considered Buildings

2.3.1. Single-Family Dwellings

The single-family house archetypes developed as part of the PEER-CEA Project are used herein (Welch and Deierlein, 2020). The layouts and wall dimensions of each floor are listed in Table 2.1. The overall building dimension is of 40ft × 30ft. For each wall presented in the floor plan, two parallel connected Pinching4 hysteretic springs (discussed later) located at the centroid of the wall are used to capture the location, dimension, and mechanical properties of the wall. The typical superstructure story height is 9ft and different heights are considered for the cripple wall level. Wood structural panels are introduced to retrofit the unbraced and unbolted cripple walls, which, in their original condition, represents a soft-story vulnerability. For the retrofitted buildings, the

wood structural panels are placed at the two ends of each line of cripple walls. The lengths of the structural panels vary based on the FEMA P-1100 specification (FEMA 2018). The details of the retrofit designs are presented in section 1.2.1 of the PEER-CEA project technical report (Welch and Deierlein 2020).

Table 2.1 Floor plan and wall dimensions for the single-family woodframe houses

Story Level	Floor Plan
Cripple Wall	
1st Story	
2nd Story	

The archetypes developed as part of the PEER-CEA project include buildings constructed pre-1945, 1945 to 1955, and 1956 to 1970 era. The main distinction between the different construction eras is the interior panel sheathing material. Listed in Table 2.2 are the exterior and interior wall material combinations for different construction era considered in the current study. Other building variants under consideration include the cripple wall height, seismic weight (discussed later), and number of stories. Cripple walls ranging in height from 2ft to 6ft are considered. For each material combination, 1- and 2-story buildings, with and without cripple wall (fixed base) are considered.

Table 2.2 Wood panel sheathing material combinations for different construction eras

Cripple Wall	Superstructure Exterior Wall	Superstructure Interior Wall	Construction Era
Stucco	Stucco	Gypsum	1945 – 1955, 1956 – 1970
Horizontal wood siding	Horizontal wood siding	Gypsum	1945 – 1955, 1956 – 1970
Stucco	Stucco	Plaster on wood lath	Pre-1945, 1945 – 1955
Horizontal wood siding	Horizontal wood siding	Plaster on wood lath	Pre-1945, 1945 – 1955

2.3.2. Multi-Family Dwelling

The multi-family archetypes are based on the SWOF buildings under the purview of the Los Angeles Soft Story Ordinance. A survey of the SWOF inventory was conducted and used as the basis for developing the archetypes. A dataset containing the addresses of approximately 12,000 of the SWOF buildings impacted by the LA Ordinance Retrofit Project was obtained from the Los Angeles Times (http://graphics.latimes.com/soft-story-apartments-needing-retrofit/). A survey of approximately 25% of the buildings on the Los Angeles Times list was then conducted using Google Street View. Figure 2.4 shows a map of the approximately 12,000 buildings from the Los Angeles Times list with the approximately 3,000 surveyed buildings highlighted. About 72% of

the surveyed buildings had two stories and about 23% had three stories. The remaining 5% included 4-, 5- and 6-story buildings.

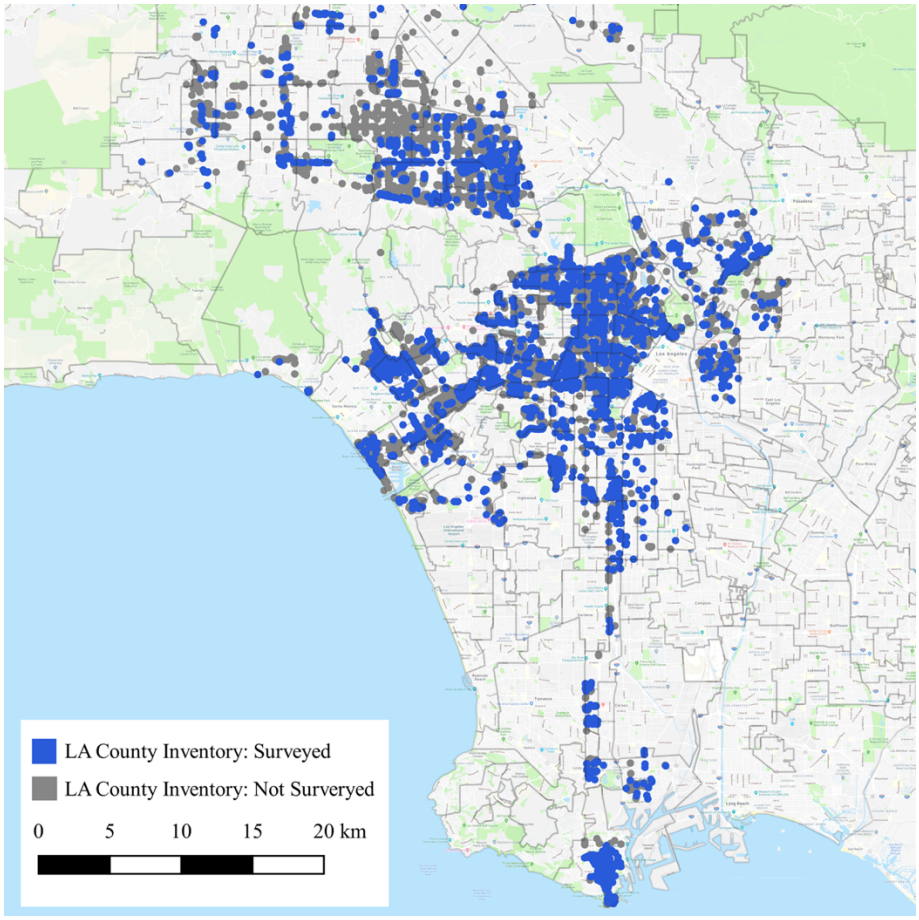


Figure 2.4 Map showing inventory of SWOF woodframe buildings in the City of Los Angeles with surveyed buildings highlighted

The wall layout in the first story of SWOF woodframe buildings has strong implications to their collapse performance. From the results of the survey, four typical first-story wall layouts were identified. Figure 2.5 shows pictures of real SWOF buildings with the four typical first-story wall layouts, which are identified as L1 through L4. All four SWOF layouts have partially open first stories that include parking and an enclosed area that consists of living, laundry or storage space. The SWOF layouts identified as L1 and L2 have one wall line that is completely open and two that are partially open. For L1, the completely open wall line is in the “short” direction whereas the

"long" direction of L2 has the completely open wall line. L3 has a single completely open wall line that serves as the entrance to a parking garage that has walls on three sides. The parking area in L4 is located at the corner of the building where there are two partially open wall lines. Schematic isometric views of the four typical wall layouts are shown in Figure 2.6. The first-story wall layout of approximately one-third of the surveyed buildings (the ones that were clearly notable from Google Street View) was documented. Among those buildings, approximately 17%, 2%, 61% and 20% had layouts L1, L2, L3 and L4, respectively.

Figure 2.5 Photos of typical Los Angeles SWOF woodframe building configurations: (a) L1, (b) L2, (c) L3 and (d) L4

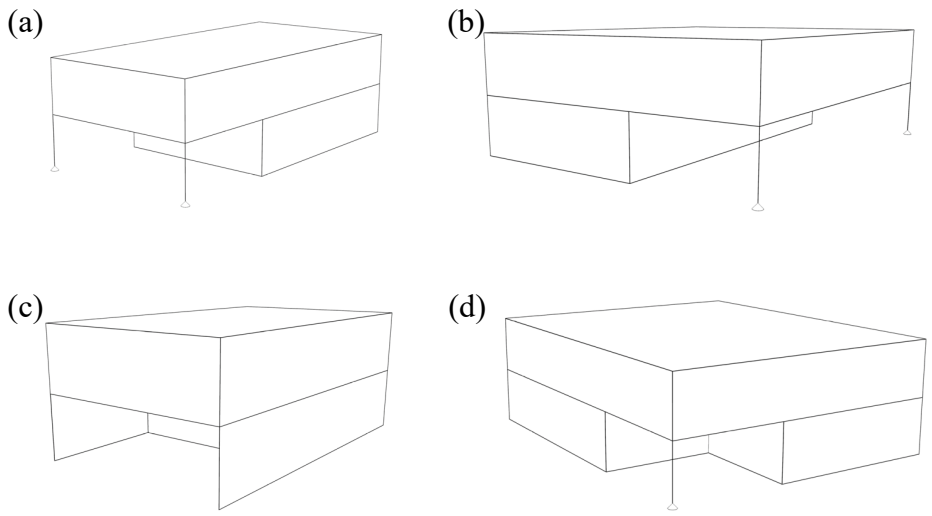


Figure 2.6 Schematic isometric views of typical SWOF woodframe building configurations identified from survey: (a) L1, (b) L2, (c) L3 and (d) L4

A total of thirty-two archetype buildings are developed and used to represent the inventory of SWOF woodframe buildings in the City of Los Angeles. The properties of each archetype are summarized in Table 2.3. In addition to the four first-story wall layouts described earlier, variations in the number of stories, plan aspect ratio and type of interior panels are considered. As noted earlier, 95% of the surveyed buildings had either two or three stories. Therefore, only 2-story and 3-story variants are included in the set of archetypes. Figure 2.7 shows the wall layout of the four typical 1st story plan configurations and the typical plan layout of the upper floors is shown in Figure 2.8. The floor layouts (1st and upper) are configured with one, two- and three-bedroom apartments, such that the wall densities (total wall length normalized by the floor area), which range from $0.12 ft/ft^2$ in the first story to $0.20 ft/ft^2$ in the upper stories, are comparable to real buildings. The story heights for all archetypes is taken to be 9'-3".

For each layout, the shorter plan dimension is denoted as "B" and the longer dimension is labeled "L". Two variations in plan dimensions, which are based on the ratio L/B, are included for each 1st floor wall layout. The variant with higher L/B is denoted as having a "large aspect ratio" the

other is described as having a “small aspect ratio”. The value of B and L for each building case is summarized in Table 2.3. For each layout, detailed plan views (8 in total) showing all wall dimensions are provided in Appendix A.

The exterior walls are taken to be constructed with stucco on the outside and either gypsum wallboard (GWB) or horizontal wood siding (HWS) on the inside. The interior partitions consist of either GWB or HWS on both sides. Including the GWB/HWS interior panel variant enables an evaluation of the effect of upper story strength and ductility on the shift in collapse from the 1st to the 2nd story after retrofit. None of the walls in the existing buildings have wood structural panels.

Table 2.3 Summary of SWOF woodframe building archetypes

Building ID	No. of Stories	Plan Dimensions	Interior Panel Type	Exterior Panel Type
L1-2S-60X30-GWB	2	L = 60'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L1-3S-60X30-GWB	3	L = 60'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L1-2S-100X30-GWB	2	L = 100'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L1-3S-100X30-GWB	3	L = 100'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L1-2S-60X30-HWS	2	L = 60'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L1-3S-60X30-HWS	3	L = 60'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L1-2S-100X30-HWS	2	L = 100'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L1-3S-100X30-HWS	3	L = 100'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L2-2S-60X50-GWB	2	L = 60'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L2-3S-60X50-GWB	3	L = 60'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L2-2S-100X50-GWB	2	L = 100'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L2-3S-100X50-GWB	3	L = 100'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L2-2S-60X50-HWS	2	L = 60'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L2-3S-60X50-HWS	3	L = 60'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L2-2S-100X50-HWS	2	L = 100'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L2-3S-100X50-HWS	3	L = 100'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L3-2S-50X30-GWB	2	L = 50'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L3-3S-50X30-GWB	3	L = 50'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L3-2S-80X30-GWB	2	L = 80'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L3-3S-80X30-GWB	3	L = 80'-0", B = 30'-0"	2 Layers GWB	Stucco + GWB
L3-2S-50X30-HWS	2	L = 50'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L3-3S-50X30-HWS	3	L = 50'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L3-2S-80X30-HWS	2	L = 80'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L3-3S-80X30-HWS	3	L = 80'-0", B = 30'-0"	2 Layers HWS	Stucco + HWS
L4-2S-60X50-GWB	2	L = 60'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L4-3S-60X50-GWB	3	L = 60'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L4-2S-100X50-GWB	2	L = 100'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L4-3S-100X50-GWB	3	L = 100'-0", B = 50'-0"	2 Layers GWB	Stucco + GWB
L4-2S-60X50-HWS	2	L = 60'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L4-3S-60X50-HWS	3	L = 60'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L4-2S-100X50-HWS	2	L = 100'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS
L4-3S-100X50-HWS	3	L = 100'-0", B = 50'-0"	2 Layers HWS	Stucco + HWS

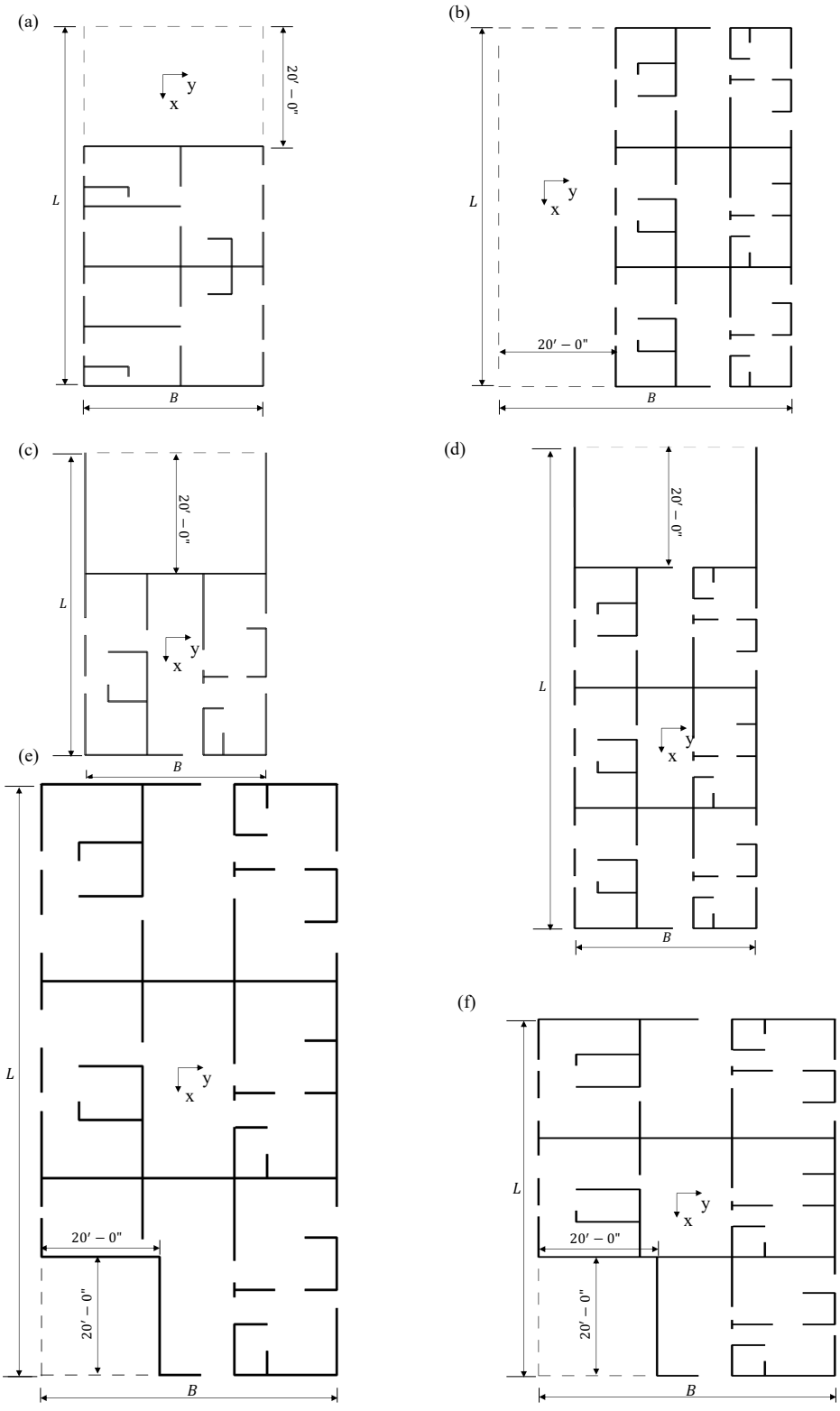


Figure 2.7 First floor plan showing wall layouts: (a) L1, (b) L2, (c) L3 (L = 50’-0’’), (d) L3 (L = 80’-0’’), (e) L4 (L = 60’-0’’) and (f) L4 (L = 100’-0’’)

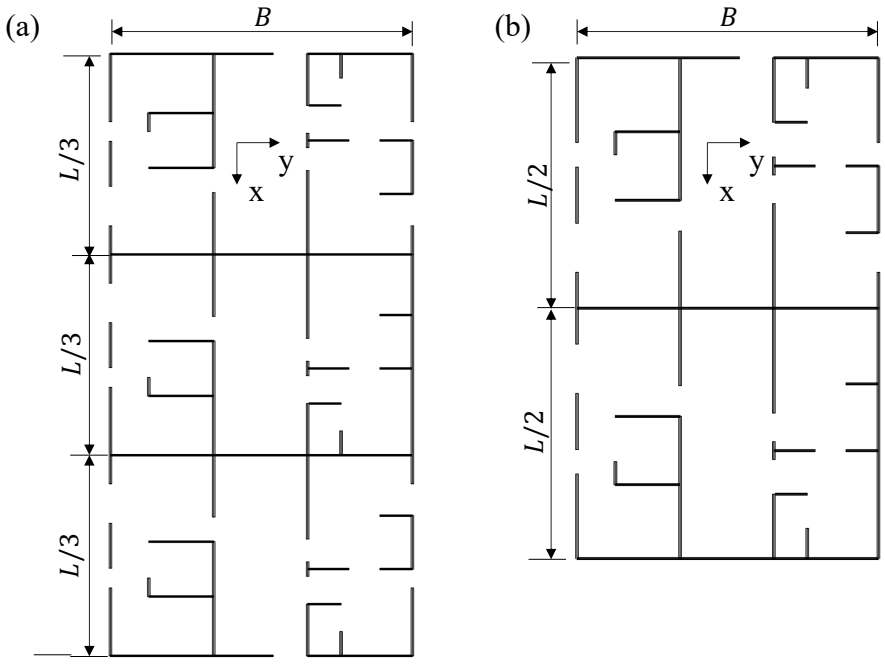


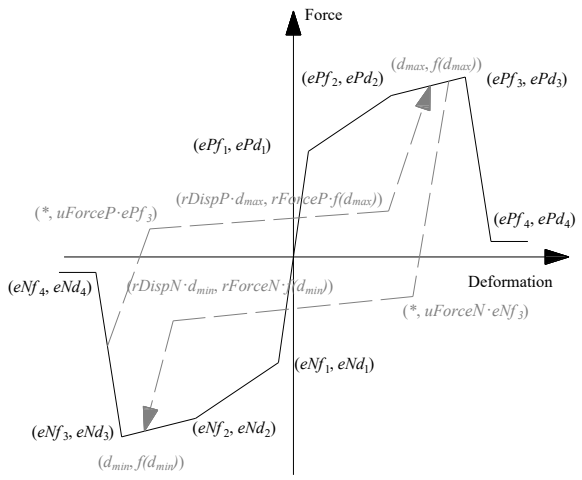
Figure 2.8 Plan configuration in upper floors: (a) L1, L2, L3 (L = 80′-0″) and L4 and (b) L3 (L = 50′-0″)

2.4. Woodframe Shear Wall Material Models

2.4.1. Introduction to Pinching4 Hysteretic Model

Lowes et al. (2003) developed a two-dimensional beam-column joints nonlinear hysteretic model, which is known as Pinching 4 model. Compared to other models, Pinching4 can update the strength envelopes and unloading/reloading paths during the analysis. This property makes the Pinching4 model capable of accurately capturing sheathing connection behavior. Thus, the Pinching4 material is used to model the nonlinear response of the panels. Figure 2.9 shows the hysteretic response and associated parameter definitions of Pinching4 model in OpenSees. The response envelope (positive and negative) is multilinear, which includes degrading and constant-residual-strength branches. Force and deformation parameters at each point where there is a change in stiffness (a total of sixteen parameters) are used to define the envelopes. Each unload-reload path is defined by six parameters, which includes the load-deformation point at which unloading occurs

and four other parameters defined by some fraction of the force and deformation at the unloading point. Fourteen parameters are used to define the hysteretic damage rules.



Parameter	Description
ePf_i/eNf_i	Force points on the positive/negative response curve
ePd_i/eNd_i	Deformation on the positive/negative response curve
$rDispP/rDispN$	Ratio of the deformation at which reloading occurs to the maximum/minimum historic deformation demand
$rForceP/rForceN$	Ratio of the force at which reloading occurs to the maximum/minimum historic deformation demand
$uForceP/uForceN$	Ratio of strength developed upon unloading from negative load to the maximum/minimum strength developed under monotonic loading
$gK1, gK2, gK3, gK4, gKLim$	Controlling cyclic degradation model for unloading stiffness degradation
$gD1, gD2, gD3, gD4, gDLim$	Controlling cyclic degradation model for reloading stiffness degradation
$gF1, gF2, gF3, gF4, gFLim$	Controlling cyclic degradation model for strength degradation
gE	Define maximum energy dissipation under cyclic loading.

Figure 2.9 A schematic representation of Pinching4 model under monotonic loading

2.4.2. Single-Family Dwelling

Loading protocols (Zareian and Lanning 2020) and a series of component experiments (Cobeen et al. 2020; Schiller et al. 2020a, b, c) was designed and conducted as part of the PEER CEA project to develop the required sheathing material mechanical properties to support numerical analyses of the single-family houses. Two parallelly connected Pinching4 hysteretic springs are used for the calibration. The two-spring assumption is capable of fitting a large variety of experimental data. The details of the calibration are presented in the PEER CEA technical background report (Welch and Deierlein 2020). During the calibration, ‘spring 1’ was used to capture the material cyclic behaviors under small displacement, and ‘spring 2’ was used to capture the strength degradation segments. Best estimates of the material properties, which are based on experimental data, knowledge of behavior, detailing of actual buildings, and observations from historical earthquakes, were adopted in the single-family building models (Welch and Deierlein 2020). The best estimates of the Pinching4 backbone and cyclic parameters for the materials used in the analyses are listed in Table 2.4 and

Table 2.5, respectively. The given backbone parameters are normalized by panel height and length.

Table 2.4 Single-family house material Pinching4 backbone parameters

Location	Material Name		d1 (%)	d2 (%)	d3 (%)	d4 (%)	f1 (lbs/ft)	f2 (lbs/ft)	f3 (lbs/ft)	f4 (lbs/ft)
Superstructure Exterior Wall	Horizontal Wood Siding	Spring 1	0.24	1.16	4	15	32.8	78.75	57	12.1
		Spring 2	0.24	1.16	4	15	8.2	26.25	133	108.9
	Horizontal Siding + Plaster on Wood Lath	Spring 1	0.08	0.3	1.1	2.2	121.6	303	157.5	20
		Spring 2	0.08	0.3	1.1	2.2	30.4	101	367.5	180
	Stucco + Gypsum Wallboard Best Estimate	Spring 1	0.08	0.72	1.5	5.4	205.6	548.25	240	24
		Spring 2	0.08	0.72	1.5	5.4	51.4	182.75	560	216
	Stucco + Plaster on Wood Lath	Spring 1	0.08	0.36	1.2	3.7	285.6	621.75	315	31.5
		Spring 2	0.08	0.36	1.2	3.7	71.4	207.25	735	283.5
Superstructure Interior Wall	Gypsum Wallboard Best Estimate	Spring 1	0.12	0.36	0.8	5.65	84	138.75	63	6.3
		Spring 2	0.12	0.36	0.8	5.65	21	46.25	147	56.7
	Plaster on Wood Lath	Spring 1	0.08	0.28	1.2	2.9	92	214.5	133.5	12.8
		Spring 2	0.08	0.28	1.2	2.9	23	71.5	311.5	115.2
Cripple Wall	2ft Stucco Best Estimate	Spring 1	0.1	0.77	1.6	6.7	178.4	373.5	161.7	18.9
		Spring 2	0.1	0.77	1.6	6.7	44.6	124.5	377.3	170.1
	6ft Stucco	Spring 1	0.08	0.32	1	3.17	201.6	369	196.8	28.4
		Spring 2	0.08	0.32	1	3.17	50.4	123	459.2	255.6
	2ft Retrofit Stucco Best Estimate	Spring 1	0.1	1.4	5.5	12.5	279.2	1260	580.2	58
		Spring 2	0.1	1.4	5.5	12.5	69.8	420	1353.8	522
	6ft Retrofit Stucco	Spring 1	0.08	0.65	2.9	9	308	1137.75	561	56.1
		Spring 2	0.08	0.65	2.9	9	77	379.25	1309	504.9
	Horizontal Wood Siding Best Estimate	Spring 1	0.1	0.75	2	5	49.6	157.5	106.2	10
		Spring 2	0.1	0.75	2	5	12.4	52.5	247.8	90
	2ft Retrofit Horizontal Wood Siding Best Estimate	Spring 1	0.1	2.06	6.5	12.1	283.2	1164.75	565.8	45.3
		Spring 2	0.1	2.06	6.5	12.1	70.8	388.25	1320.2	407.7
	6ft Retrofit Horizontal Wood Siding Best Estimate	Spring 1	0.17	1.16	3.5	8.4	369.6	1006.5	546.3	39.6
		Spring 2	0.17	1.16	3.5	8.4	92.4	335.5	1274.7	356.4

Table 2.5 Single-family house material Pinching4 cyclic parameters

Location	Material Name		rDisp	rForce	uForce	gK1	gKlim	gD1	gDlim
Superstructure Exterior Wall	Horizontal Wood Siding	Spring 1	0.18	0.35	-0.08	-0.1	-1	0.14	0.5
		Spring 2	0.5	0.12	-0.05	0	0	0.09	0.2
	Horizontal Siding + Plaster on Wood Lath	Spring 1	0.06	0.31	0.14	-0.07	-0.5	0.14	0.3
		Spring 2	0.28	0.18	-0.11	-0.05	-0.2	0.11	0.3
	Stucco + Gypsum Wallboard Best Estimate	Spring 1	0.06	0.26	-0.2	0	0	0.13	2
		Spring 2	0.06	0.17	-0.23	0.3	2	0.13	2
	Stucco + Plaster on Wood Lath	Spring 1	0.06	0.26	-0.2	0	0	0.13	2
		Spring 2	0.06	0.17	-0.23	0.3	2	0.13	2
Superstructure Interior Wall	Gypsum Wallboard Best Estimate	Spring 1	0.15	0.22	-0.21	-0.3	-1	0.1	2
		Spring 2	0.4	0.12	-0.19	0.2	2	0.12	2
	Plaster on Wood Lath	Spring 1	0.06	0.31	-0.1	-0.07	-0.5	0.14	0.3
		Spring 2	0.28	0.18	-0.11	-0.05	-0.2	0.11	0.3
Cripple Wall	2ft Stucco Best Estimate	Spring 1	0.05	0.2	-0.1	0	0	0.3	3
		Spring 2	0.25	0.16	-0.1	0	0	0.2	2
	6ft Stucco	Spring 1	0.04	0.31	-0.05	-0.12	-0.75	0.3	3
		Spring 2	0.25	0.16	-0.25	-0.09	-0.3	0.2	2
	2ft Retrofit Stucco Best Estimate	Spring 1	0.05	0.37	-0.05	0	0	0.15	2
		Spring 2	0.55	0.13	-0.07	0	0	0.1	0.5
	6ft Retrofit Stucco	Spring 1	0.09	0.31	-0.05	-0.1	-0.3	0.21	2
		Spring 2	0.49	0.16	-0.15	0	0	0.1	1
	Horizontal Wood Siding Best Estimate	Spring 1	0.18	0.37	-0.1	-0.1	-1	0.14	0.3
		Spring 2	0.4	0.34	-0.12	0	0	0.09	0.2
	2ft Retrofit Horizontal Wood Siding Best Estimate	Spring 1	0.03	0.4	0.02	0	0	0.14	0.3
		Spring 2	0.55	0.14	-0.09	0	0	0.05	1
	6ft Retrofit Horizontal Wood Siding Best Estimate	Spring 1	0.06	0.38	0.02	-0.1	-1	0.2	0.3
		Spring 2	0.55	0.2	-0.15	-0.1	-0.5	0.1	1

2.4.3. Multi-Family Dwelling

The development of Pinching4 parameters for the multi-family dwelling modeling is slightly different from that of single-family houses. To obtain the Pinching4 parameters for each panel used in the retrofitted and un-retrofitted multi-family building, calibrations on numerical data was performed. More specifically, for the GWB, HWS and stucco panel material, the calibration was conducted using the hysteretic response of the 10-parameter SAWS material as the benchmark. For the WSP used in the retrofit, the hysteretic response obtained from the CASHEW program (Folz and Filiatrault 2001, 2002), which includes the nailing sizes and spacings as input. The SAWS/CASHEW-based hysteretic response of an 8'-0" long, 9'-3" tall panel is developed from the OpenSees/CASHEW program using the CUREE-Caltech loading protocol (Krawinkler et al. 2001). The Pinching4 material parameters were then selected such that the hysteretic response under the CUREE-Caltech loading protocol is comparable with the SAWS/CASHEW response. A comparison between the hysteric response of panels with all six studied materials is shown in Figure 2.10, where it is observed that the calibrated Pinching4 material adequately captures the SAWS hysteric response. The Pinching4 model parameters for all panel types are summarized in Table 2.6.

Table 2.6 Pinching4 model parameters for multi-family dwellings[1]

Panel Type	Force (lb/ft)				Drift (%)				Cyclic		
	f1	f2	f3	f4[2]	d1	d2	d3	d4	rDisp	rForce	uForce
GWB	48	155	199	0	0.06	0.44	0.98	4.86	0.4	0.18	0.04
HWS	18	59	159	0	0.11	0.87	7.30	8.32	0.02	0.25	0.05
Stucco	140	424	693	0	0.04	0.21	0.97	7.75	0.3	0.22	0.04
WSP 10d @ 4 inches O.C.	277	752	845	0	0.05	0.38	0.96	6.43	0.45	0.16	0.03
WSP 10d @ 2 inches O.C.	526	1406	1521	0	0.06	0.44	0.98	6.17	0.45	0.15	0.06
WSP 8d @ 4 inches O.C.	280	754	829	0	0.06	0.46	1.09	6.44	0.3	0.1	0.02

[1]The Pinching4 parameters not given in the table were set to 0
[2]Residual strength not considered

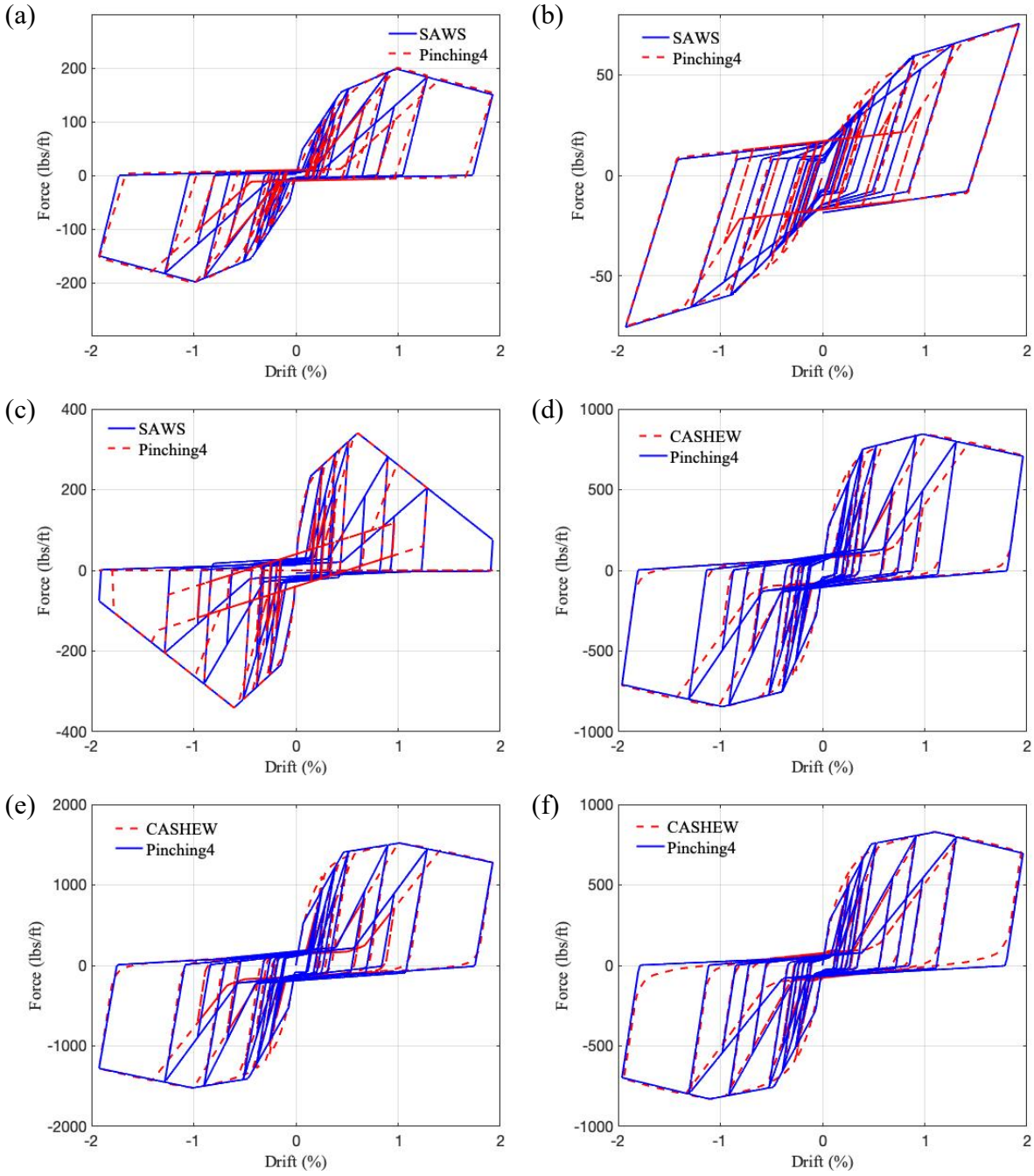


Figure 2.10 Comparison between the hysteric response of the SAWS and Pinching4 materials for (a) GWB, (b) HWS, (c) Stucco, (d) WSP, 10d @ 4'' O.C. panels, (e) WSP, 10d @ 2" O.C. panels, and (f) WSP, 8d @ 4'' O.C. panels

2.5. Component Modeling

Similar modeling strategies are applied to the single- and multi-family houses. The wood panels are idealized using a two-node link element with horizontal springs that captures the force–

deformation behavior of the panel. The two nodes are located at the top and bottom of each panel at mid-length. It should be noted that this type of element is unable to capture the overturning response of the panels. Therefore, the effect of the flexibility of the connection between the panel and the foundation (e.g. via hold-downs) on the overall response, is not incorporated. For the SWOF buildings, ordinary moment frames (OMFs) were used as the retrofit elements and the beams and columns are modeled using elastic elements with concentrated plastic hinges, which incorporate the Modified Ibarra–Krawinkler deterioration model (Ibarra et al. 2005). The model parameters for the hinges are obtained from the empirical equations developed by Lignos and Krawinkler (2012).

Nine leaning columns (one in each corner, one at mid-length of each exterior wall line and one at the centroid) are used to capture the spatial distribution of masses and P–Δ effects. Figure 2.11 shows the position of the leaning columns. Each leaning column is modeled using an elastic-beam-column element in OpenSees. The element has an artificially large cross-section area, torsional moment of inertia and second moment of area to ensure the leaning columns can capture P-Δ effects without providing additional stiffness.

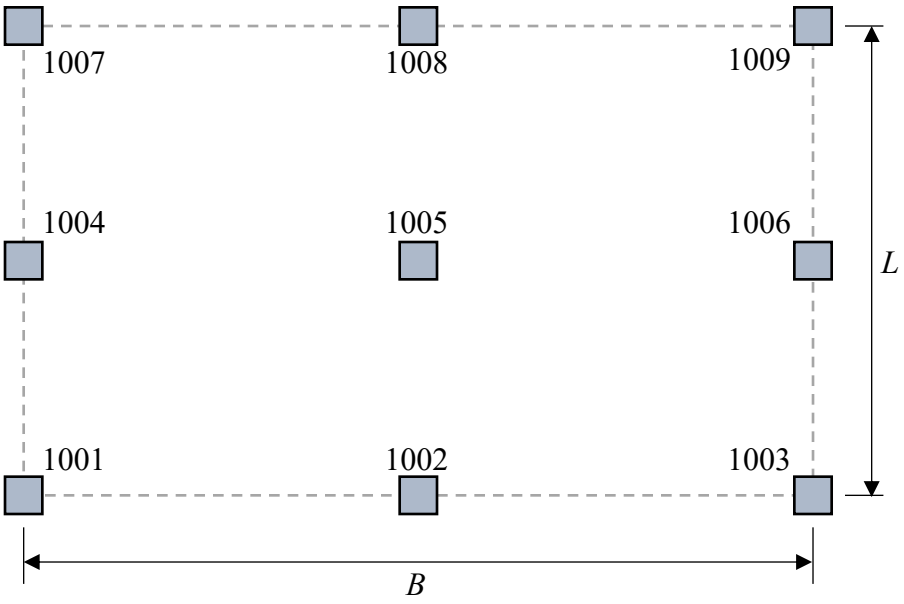


Figure 2.11 Leaning column locations

From prior numerical studies on woodframe building modeling, rigid diaphragm assumptions were generally used though it has inefficiencies under some scenarios. Considering all single- and multi-family houses involved in the study are of rectangular shape with an aspect ratio (long to short dimension ratio) less than 2, the rigid diaphragm constraint is deemed a reasonable assumption. This assumption is one of the limitations of the current study, and more comprehensive analyses can be carried out in the future by modeling semi-rigid or flexible diaphragms.

2.6. Gravity Loads and Masses

2.6.1. Single-Family Dwelling

As introduced in the preceding section, nine leaning columns are used to capture P–Δ effects. For single-family houses, the total gravity loads are computed based on contributions from the floors, panels (interior and exterior wall), and roof. The numbers vary by sheathing material, construction era and diaphragm type. Then, the resulting loads are assigned to each leaning column based on tributary diaphragm areas and wall lengths. Masses are directly calculated based on the assigned

gravity load. The weights in pounds per square foot for different diaphragms, roofs, exterior, interior and cripple walls are listed in Tables 2.7 through 2.11. The information in these tables are based on the PEER CEA technical background report (Welch and Deierlein 2020).

Table 2.7 Summary of floor load weight take-offs for floor diaphragms (Welch and Deierlein 2020)

Load Type	Description	Weight (psf)
1st Floor	First floor diaphragm, used for all variants, no finish on underside	12
2nd Floor, pre-1945, Light	Second floor diaphragm, lath and plaster on ceiling below, used when shingle or composition roof assumed (light flooring)	17.5
2nd Floor, pre-1945, Heavy	Second floor diaphragm, lath and plaster on ceiling below, used when concrete tile roof assumed (hardwood flooring)	21
2nd Floor, post-1955, Light	Second floor diaphragm, gypsum wallboard on ceiling below, used when shingle or composition roof assumed (light flooring)	12
2nd Floor, post-1955, Heavy	Second floor diaphragm, gypsum wallboard on ceiling below, used when concrete tile roof assumed (hardwood flooring)	15.5

Table 2.8 Summary of roof weight take-offs (Welch and Deierlein 2020)

Roof Type	Description	Weight (psf)	Weight 6:12[1] (psf)
pre-1945, Light	Asphalt or composition shingle roofing, lath and plaster on ceiling below	18.5	20.7
pre-1945, Heavy	Concrete tile roofing, lath and plater on ceiling below	26	29
post-1955, Light	Asphalt or composition shingle roofing, gypsum wallboard on ceiling below	13	14.5
post-1955, Heavy	Concrete tile roofing, gypsum wallboard on ceiling below	20.5	22.9

[1]Roof flat load considering horizontal projection of an 6:12 roof pitch

Table 2.9 Summary of exterior wall load weight take-offs for superstructure walls (Welch and Deierlein 2020)

Exterior Wall Material Type	Description	Weight (psf)
Stucco, pre-1945	Stucco exterior, lath and plaster interior	23
Stucco, post-1955	Stucco exterior, gypsum wallboard interior	17
Wood Siding, pre-1945	Horizontal wood siding exterior, lath and plaster interior	15
Wood Siding, post-1955	Horizontal wood siding exterior, gypsum wallboard interior	7

Table 2.10 Summary of interior partition wall load weight take-offs for superstructure walls (Welch and Deierlein 2020)

Interior Wall Material Type	Description	Weight (psf)
Plaster on Wood Lath, pre-1945	Two sides of finish material	18
Gypsum Wallboard, post-1955	Two sides of finish material	7

Table 2.11 Summary of cripple wall load weight take-offs for cripple walls (Welch and Deierlein 2020)

Cripple Wall Material Type	Description	Weight (psf)
Stucco	Stucco exterior, includes weight of wood sheathing	14
Wood Siding	Horizontal wood siding or plywood siding exterior	4 (6)[1]

[1]Add 2psf to include 1'' nominal wood sheathing behind siding for pre-1945 wood siding cases

2.6.2. Multi-Family Dwelling

For the multi-family houses, the seismic weight of each building is computed using $35\ psf$ as the typical floor dead load, $25\ psf$ for roof dead loads, $10\ psf$ for the weight of the interior partitions and $15\ psf$ for the exterior wall weight per square foot of wall. Details of the seismic weight each

archetype are summarized in Table 2.12. The calculated gravity loads are distributed to each of the leaning column based on the tributary area.

Table 2.12 Multi-family baseline archetype seismic weight

Building ID	Seismic Weight (kips)
L1-2S-60X30	185
L1-3S-60X30	306
L1-2S-100X30	284
L1-3S-100X30	467
L2-2S-60X50	287
L2-3S-60X50	469
L2-2S-100X50	437
L2-3S-100X50	709
L3-2S-50X30	151
L3-3S-50X30	248
L3-2S-80X30	242
L3-3S-80X30	394
L4-2S-60X50	292
L4-3S-60X50	475
L4-2S-100X50	444
L4-3S-100X50	716

2.7. Damping

A brief review on the damping model development and treatment in woodframe building models is provided in this section. Table 2.13 presents a comparison among commonly adopted classical damping models in structure dynamic analysis. Classical damping refers to the case, where $\Phi^T C \Phi$ is a diagonal matrix (C is the damping matrix and Φ is the modal shape). The assumption can be made where the damping mechanisms throughout the structure are similar. In woodframe building, damping effects primarily originate from inherent friction in materials, components and fasteners. Within the scope of current study, no soil-structure-interaction effects are incorporated in modeling. Caughey (Caughey and O'Kelly 1965) and Rayleigh damping are common choices for classical damping. Rayleigh damping takes a simplified form of Caughey damping, and it avoids the matrix

inversion calculation. Therefore, it is the most popular damping model in practice. Since the stiffness matrix is involved in the construction of the damping matrix, Rayleigh damping models could take two forms, which are initial and tangent stiffness proportional Rayleigh damping. For the former, the initial stiffness matrix of the structure is adopted and results in a relatively simple form of the damping matrix. For such damping model, ignoring stiffness degradation in inelastic analysis would overestimate the actual damping, eventually creating artificially high damping forces, which can lead to higher collapse resistance (Symans et al. 2008; Hardyniec and Charney 2015). Also, as the stiffness changes, the modal periods changes in the analysis so that damping ratios at each mode will not correspond to the original damping ratio curve (Symans et al. 2008; Jehel et al. 2014). As for the latter, the tangent stiffness with an adaptively updated stiffness matrix is used in the damping matrix calculation. One tangent stiffness proportional Rayleigh damping is calculated, and the coefficients are fixed at the beginning of analysis. The other is updated with the periods and the coefficients are recomputed at each time step. The fixed coefficients method can create damping ratios that lie exactly lay the original curve (Symans et al. 2008; Jehel et al. 2014). However, as stiffness degrades, the modal damping ratios corresponding to periods are less than the initial period, will have gradually larger differences between the original damping ratio in elastic stage. Using the updated coefficients method can re-anchor modes' damping ratios at each time step and solve this issue. Nevertheless, performing eigen analysis at each step significantly increases the computational expense. Similar to the initial stiffness proportional Rayleigh damping, tangent stiffness proportional damping still potentially produces high artificial damping forces. More unrealistic, stiffness degradation might lead to negative damping forces. Modal damping could resolve the issues of spurious damping forces in Rayleigh damping model. Only the mass

matrix is involved in its formulation, and massless degree-of-freedoms would not contribute to the system damping.

Table 2.13 Comparison among different classical damping models

Damping Model	Damping Matrix	Pros	Cons
Caughey damping	$C = M\sum_{i=0}^{N-1} a_i\,[M^{-1}K]^i$	Damping matrix is diagonalizable, appropriate for classical damping assumption	Matrix inversion is involved in implementation, computationally expensive
Initial Stiffness Proportional Rayleigh Damping	$C = a_0 M + a_1 K_0$	Simple form, easy to solve	Overestimate actual damping forces, modal damping ratios change over the analysis
Tangent Stiffness Proportional Rayleigh Damping with Fixed Coefficients	$C = a_0 M + a_1 K_t$	Damping ratios lay on the original curve, update stiffness matrix real-time	Lack of physical explanation, risk of negative damping force
Tangent Stiffness Proportional Rayleigh Damping with Updated Coefficients	$C = a_{0,t} M + a_{1,t} K_t$	Update stiffness matrix and coefficients real-time, damping ratios maintain as set	Computationally expensive, lack of physical explanation
Modal Damping	$C = M\left[\sum_{i=1}^{N} \frac{2\xi_i \omega_i \phi_i \phi_i^T}{\phi_i M \phi_i^T}\right] M$	Will not generate 'spacious' damping force in the analysis	

Table 2.14 shows the damping models and ratios used prior woodframe building numerical studies. For woodframe structure modeling, a linear damping ratio is usually set within a range of 0.01 to 0.05 (Chui and Smith 1989; Foliente 1995). In the CUREE project, viscous damping was assigned based on initial stiffness proportional damping (Folz and Filiatrault 2001). The damping characteristics were derived from shaking table testing using steady state input. 4.2% critical damping ratio was calculated under small amplitude excitation from the seismic test, while under large amplitude excitation reduced critical damping ratio should be used because shear walls' hysteretic behaviors can account for most of structure damping (Folz and Filiatrault 2001). In other types of woodframe building numerical analysis and performance assumptions, 1% to 5% damping ratios were commonly adopted.

Table 2.14 Damping models and ratios in prior woodframe building studies

Study	Damping Model	Damping Ratio
Foliente 1995	SDOF constant	1% to 5%
Dinehart and Shenton III 2000	Constant damping	5%
Folz and Filiatrault 2001	SDOF constant	1%
Folz and Filiatrault 2004	Rayleigh damping	1%
Pang et al. 2007	SDOF constant	1%
Ellingwood et al. 2008	NA	1%
Park and van de Lindt 2009	Constant damping	1%
Yin and Li 2010	NA	1%
Pozza et al. 2014	Rayleigh damping	2%
Ghehnavieh 2017	Initial stiffness proportional Rayleigh damping	5%
Roohi et al. 2019	Initial stiffness proportional Rayleigh damping	2%

In the current study, a 2.5% tangent stiffness proportional Rayleigh damping is assumed for single-family buildings (Welch and Deierlein 2020) and 1% initial stiffness proportional Rayleigh damping is used for the multi-family dwellings. The damping is anchored to the 1st and 3rd mode. The mass proportional damping is assigned to the leaning column nodes, where the masses are concentrated. The stiffness proportional damping is distributed to any components providing lateral resistances, including wood panels and moment frames.

2.8. Nonlinear Static Analysis

Nonlinear static (pushover) analyses are performed on the numerical models to evaluate the strength and overall drift capacity of the buildings. Given that most of the woodframe buildings are of one- or two-story, the pushover analyses are conducted using the equivalent lateral force (ELF) specified in ASCE 7-16, Section 12.8-3 (ASCE 2016). The displacements are assigned to the central leaning column (labelled 1005 in Figure 2.11). The story drift ratios (SDR) and roof drift ratios are recorded at three leaning columns on the centerline perpendicular to the loading direction. Base shears are recorded at the base nodes of the components, including wood panels, leaning columns and moment frames (if present).

2.9. Nonlinear Dynamic Analysis

2.9.1. Ground Motion Selection

2.9.1.1. Single-Family Dwelling

The performance assessment for single-family houses is conducted through multiple stripe analysis (MSA). As part of the PEER-CEA Project, four sites representing a range of seismicity were selected for the numerical analyses. For each site, probabilistic seismic hazard analysis was conducted for 10 return periods ranging from 15 to 2500 years (Mazzoni et al. 2020). The ground motions selected for the San Francisco $V_{s,30} = 270m/s$ site is used in the single-family house MSA. Figure 2.12 shows the mean spectral accelerations of the selected ground motions for 16 hazard levels with return periods ranging from 15 years to 4000 years. For the first 10 hazard levels (15- to 2500-year return period), ground motions were selected using the conditional spectra with a target conditioning period of 0.25s. Compared to the uniform hazard spectra selected ground motions, the mean conditional spectra are of lower values at periods away from the conditional period. As for the remaining six hazard levels, the ground motions selected for the 2500-year return

period hazard level were used with amplification factors of 1.2, 1.5, 1.8, 2.0, 2.2 and 2.5. These six hazard levels were used to assess the seismic performance of very stiff buildings (retrofitted or rigid-base buildings).

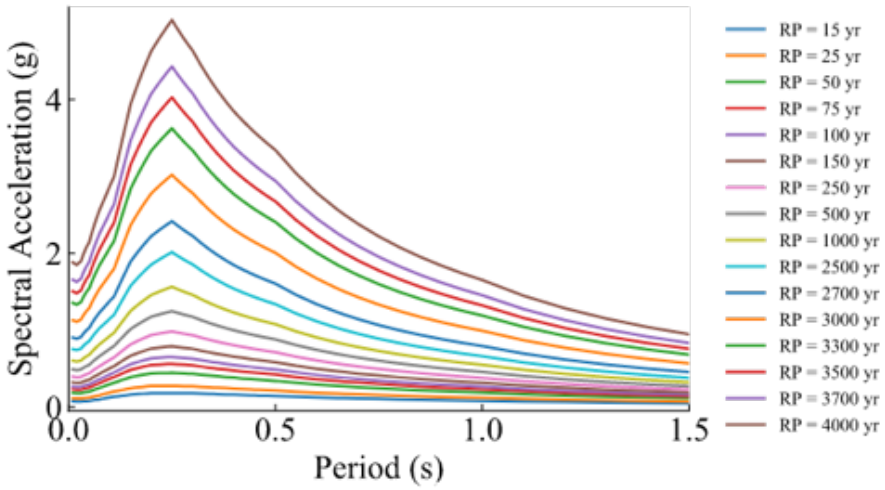


Figure 2.12 Mean conditional spectra (conditioned on a period of $T^* = 0.2s$) of San Francisco $V_{s,30} = 270m/s$ site for 16 hazard levels (Mazzoni et al. 2020)

2.9.1.2. Multi-Family Dwelling

The performances of the multi-family dwellings are assessed using incremental dynamic analyses (IDAs). Nonlinear response history analyses are conducted using the 22 pairs of far-field ground motions specified in FEMA P695 (FEMA, 2009). For each of the pre-defined intensity levels, a uniform constant scaling factor is applied to all 22 pairs of ground motions. Then, the intensity measure of the corresponding intensity level is the mean spectra value of the 22 scaled ground motion pairs given the specified period. In the implementation, 15 intensity levels were determined using scaling factors ranging from 0.2 to 3.0 with 0.2 increment. Additionally, in the multi-family building IDA, the directions of the record-pairs are switched such that 44 analysis cases are conducted at each intensity.

2.9.2. Analysis Settings

Collapse is defined as the condition where dynamic instability occurs, or the maximum story drift ratio exceeds 20% and 10% for single- and multi-family houses, respectively. At each intensity level, the maximum story drift ratio in all stories and the peak acceleration at all floor levels in both directions are recorded. These EDPs are needed to compute the non-collapse losses in the FEMA P-58 methodology. The maximum residual drift over all stories are also recorded and used to consider demolition losses for the multi-family dwellings. The median collapse intensity is calculated by minimizing the squared difference or maximizing the likelihood function to the empirical data. For model uncertainty consideration in single family houses, a FEMA P-58 default value of 0.35 is adopted (FEMA 2012). The final dispersion is given by the squared root sum of squares (SRSS) of record-to-record and model uncertainty. As for the multi-family buildings, a single dispersion value of 0.6 is used for all existing and retrofitted archetypes. This value includes both record-to-record variation and model parameter uncertainty (FEMA 2009). The spectral shape factor is computed for the individual archetypes based on their fundamental period and the period-based ductility obtained from nonlinear static analyses (FEMA 2009).

2.10. FEMA P-58 Loss Assessment

The FEMA P-58 methodology, which integrates seismic hazard assessment, nonlinear structural response simulation, damage evaluation and quantification of decision-metrics, is used to obtain the economic losses for the existing and retrofitted archetypes. The mean annual frequency of exceedance of specific ground motion intensity levels (or hazard curves) is obtained from probabilistic seismic hazard analysis applied to the site of interest. As described in Section 2.10, EDPs are generated from NRHAs. Component level fragility curves are used to link these EDPs to physical damage and statistical loss functions are used to describe the relationship between

physical damage and repair or replacement costs. Within the context of the FEMA P-58 methodology, the cost of collapse, demolition and component-level repairs are probabilistically combined using an expected value calculation.

The detailed analytical formulation of the loss estimation methodology can be found in several publications (Porter et al. 2001; Miranda et al. 2004; Ramirez et al. 2012). As part of the end-to-end analysis platform, a FEMA P-58 based loss assessment module is developed to perform the intensity-based loss assessments. The module uses a Monte Carlo simulation-based procedure that is illustrated in Figure 2.13. For each intensity level, the NRHA results are used to compute the multivariate lognormal distribution parameters for the EDPs (SDRs, PFAs and RDRs). The probability of non-collapse is also computed from the collapse fragility curve. A set of randomly sampled $EDPs$ is generated from the multivariate lognormal distribution. For a given random sample, if collapse occurs (i.e. maximum drift exceeds 10%) or the residual drift limit for demolition is exceeded, replacement cost is assumed (sometimes an additional 25% is for the cost of debris removal). Otherwise, the EDPs are used to sample realizations of the component-level damage from the fragility curves. The component-level repair or replacement costs, which are conditioned on the damage state, are then sampled from the loss functions. The non-collapse building repair cost is taken as the sum of the repair costs for the individual components over all directions and stories. The losses conditioned on a single intensity is taken as the mean loss over all Monte Carlo realizations.

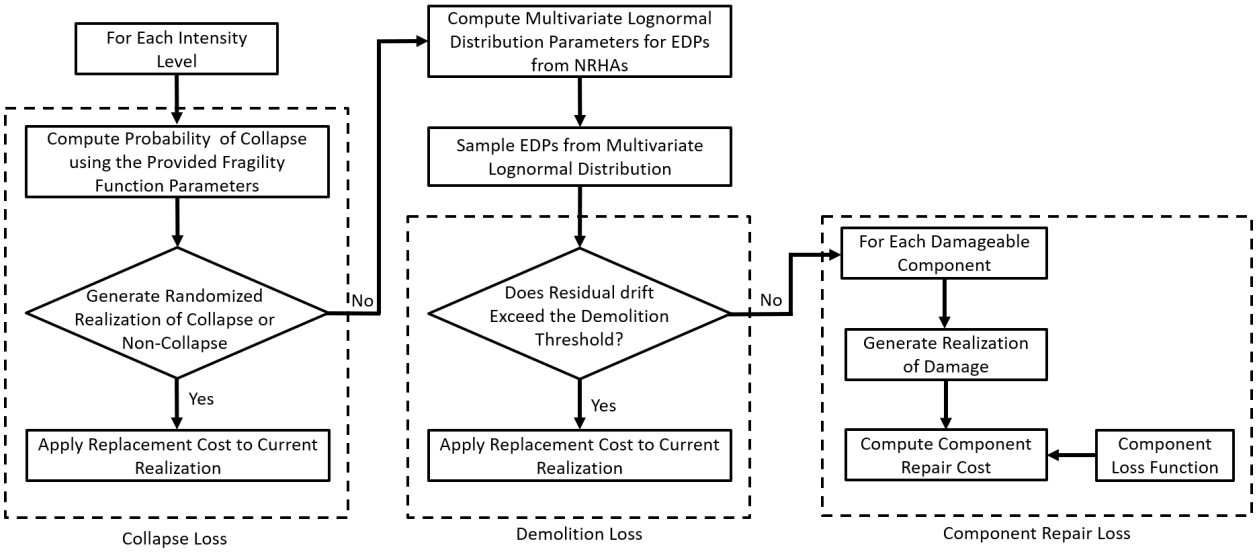


Figure 2.13 FEMA P58 earthquake-induced building loss assessment procedure

The considered damageable components for the single- and multi-family houses are shown in Table 2.15 and Table 2.16, respectively. For single-family houses, a series of damage fragilities and cost functions were adjusted on the basis of FEMA P-58 component database (Welch and Deierlein 2020). The modification was based on critical review of existing component database and additional information from PEER-CEA Earthquake Damage Workshop (Vail et al. 2020). Table 2.15 provides the assembly type and associated quantities for the components. The specific loss functions to use in the assessment are selected to match building materials. The fragility function and repair cost parameters are listed in the PEER CEA technical background report (Welch and Deierlein 2020). As for the multi-family houses, the component damage information from the FEMA P-58 database was adopted. Since the baseline archetypes vary by structure layouts, the dimensions and material types, only component names, types and applied cases are listed here. The associated quantities are summarized in Table B1 of Appendix B. A building replacement cost of $200 per square foot is assumed for both single- and multi-family dwellings (FEMA 2012; Welch and Deierlein 2020). To support the loss assessment, a SQL component database containing fragility and loss functions is established.

Table 2.15 Components and quantities considered in the single-family building loss assessment (Welch and Deierlein 2020)

Assembly Type	Unit	First-Story or Cripple Wall		Second Story	
		QX-1	QY-1	QX-1	QY-1
Exterior Walls	100SF	7.20	5.40	7.20	5.40
Interior Partitions/ Interior Wall Material	100LF	0.94	0.735	0.80	0.875
Ceramic Tile	100LF	0.13	0.12	0.13	0.12
Wallpaper Finish	100LF	0.25	0.25	0.25	0.25
2-foot tall cripple walls	100SF	1.60	1.20	NA	NA
6-foot tall cripple walls	100SF	4.80	3.60	NA	NA

Table 2.16 Components considered in multi-family building loss assessment

Fragility Name	Component Type	Applied Case
Light framed wood lateral walls	Structural	Existing and Retrofitted
Steel Column Base Plates	Structural	Retrofitted
Welded Steel Moment Connection	Structural	Retrofitted
Prefabricated steel stair no seismic joint	Non-structural	Existing and Retrofitted
Potable Water Piping	Non-structural	Existing and Retrofitted
Heating Water Piping	Non-structural	Existing and Retrofitted
Heating Water Pipe Bracing	Non-structural	Existing and Retrofitted
Sanitary Waste Piping	Non-structural	Existing and Retrofitted
HVAC Ducting	Non-structural	Existing and Retrofitted
Fire Sprinkler Water Piping	Non-structural	Existing and Retrofitted
Fire Sprinkler Drop	Non-structural	Existing and Retrofitted
Low Voltage Switchgear	Non-structural	Existing and Retrofitted
Wall Partition, Wood Stud	Non-structural	Existing and Retrofitted
Clay tile roof	Non-structural	Existing and Retrofitted

2.11. Sample Analysis Results

2.11.1. Single-Family Dwelling

The nonlinear static, MSA, and loss assessment results for a light weight 1-story 2ft stucco cripple wall building with stucco exterior wall and gypsum interior wall (1L-S2-G2-2C-S2-EX) are presented in this section.

Figure 2.14 shows the pushover analysis results for the baseline building. The curves are expressed as the relationship between the peak strength normalized by seismic weight and roof drift. Through comparing Figure 2.14 (a) with (b) and (c), the overall behavior of the single-family building with cripple wall can be assessed. The overall pushover curve is the superimposition of the curves of the individual panels. The first sign of damage occurs in the cripple wall at a roof drift approximately 0.2% in both of the two directions. The cripple wall level continues to undergo damage and normalized (by seismic weight) peaks strengths of 65% and 48% are achieved in the X- and Y-direction, respectively. The majority of story drift demands are concentrated at cripple wall level. For the X-(Y-) direction, at peak strength level, the drift at the cripple wall level is approximately 2% (2%) while the superstructure only has 0.3% (0.2%) story drift. The curve starts to descend due to strength loss in the cripple wall. The ductility in two directions, defined as the roof drift corresponding to 80% of peak strength in the descending segment, are 0.73% and 0.72% in X- and Y-direction, respectively. At this point, the cripple wall level has already generated approximately 3% story drift demand, while superstructure only has around 0.2% story drift demand.

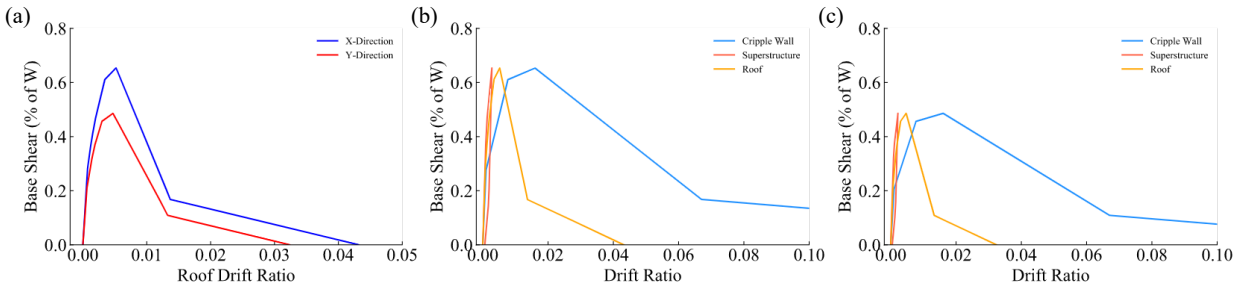


Figure 2.14 Pushover analysis results for 1L-S2-G2-2C-S2-EX: (a) base shear vs. roof drift in two perpendicular directions, (b) base shear vs. story drift in X-direction and (c) base shear vs. story drift in Y-direction

Figure 2.15 shows the median story drift demands and mean non-collapse deformed shape at each MSA level. From Figure 2.15 (a), it is observed that, for all hazard levels, the superstructure maintains lower than 0.2% median story drift ratios. However, for the cripple wall level, median drift demands explode at around1.2g. More than 50% of cases collapse due to the cripple wall story drift exceeding 20%. Due to the differences in the strength and stiffness between the two levels, all collapse cases occur at cripple wall. For non-collapse cases, displacement demands concentrate at cripple wall level as expected.

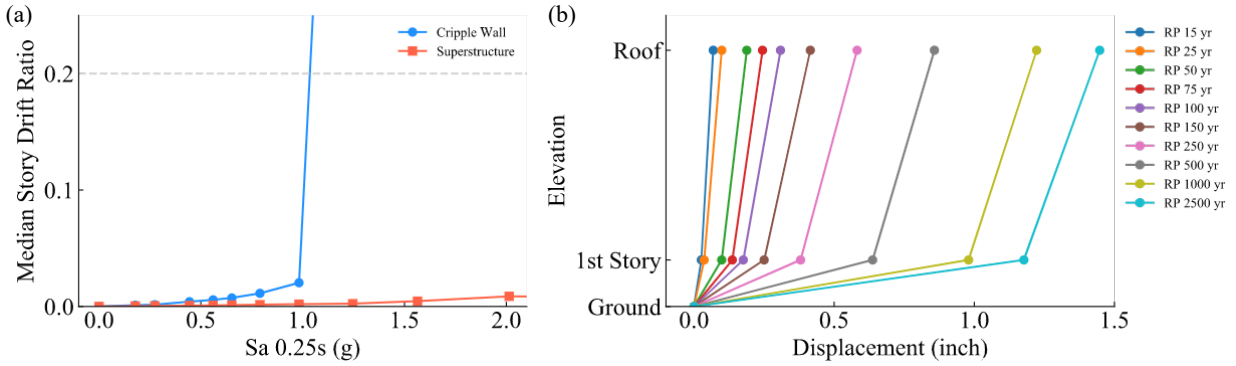


Figure 2.15 (a) Median story drift ratio at cripple wall and superstructure for each hazard level, (b) mean non-collapse structure deformed shape for each hazard level

Figure 2.16 shows the empirical probability of collapse and fitted fragility curves. The maximum likelihood method fitted fragility has a median of 1.21g and ground motion record-to-record uncertainty 0.29. At the maximum considered earthquake (MCE) level, the probability of collapse

without considering model uncertainty is 0.92. If the FEMA P-58 recommended 0.35 is adopted for model uncertainty, the probability of collapse at MCE level decreases to 0.81, which is an approximately 10% reduction compared to without model uncertainty cases.

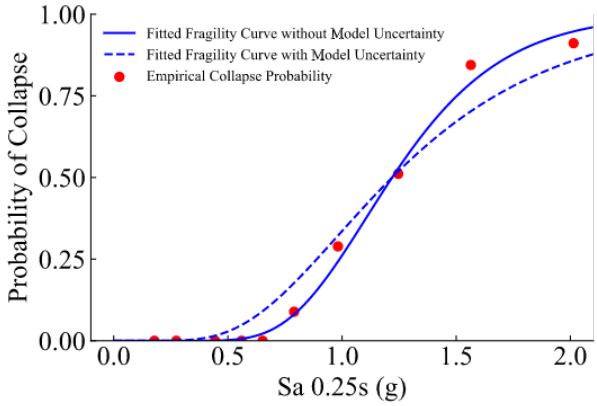


Figure 2.16 Fragility curve for 1L-S2-G2-2C-S2-EX

Figure 2.17 presents the results from the FEMA P-58 based loss assessment for the 1L-S2-G2-2C-S2-EX building. Due to the weak cripple wall present in the building, collapse loss governs in all intensity levels. The component loss curve has an increasing segment before 0.8g, and experiences a gradual reduction as collapse loss dominates the building performance. Since only story drift sensitive components are considered in the single-family house loss assessment, the component loss curve is consistent with the structure deformed shape shown in Figure 2.15 (b). The expected annual loss (EAL) is computed to be 0.46% of the building value. As expected, the majority of the contribution comes from collapse risk, which takes 83% of the total EAL. Component loss takes the rest 17% of EAL.

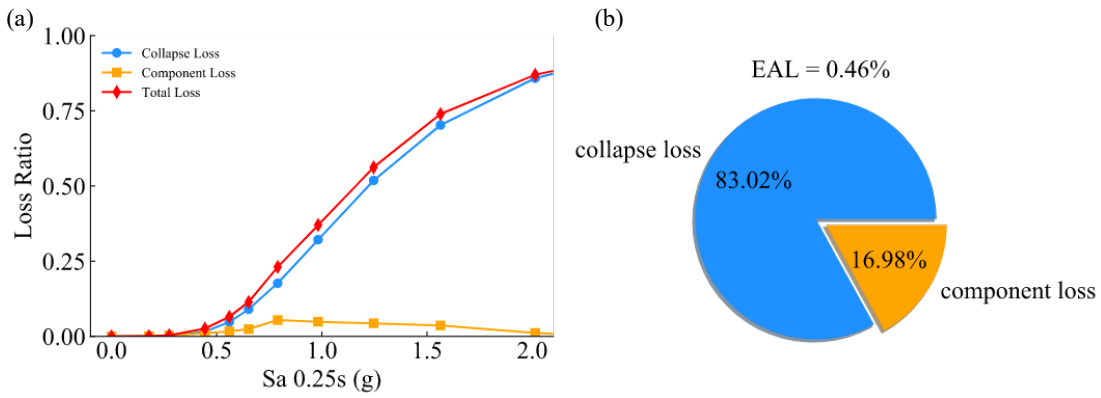


Figure 2.17 Loss Assessment results for 1L-S2-G2-2C-S2-EX: (a) intensity-based loss curve and (b) expected annual loss

2.11.2. Multi-Family Dwelling

Following the same analysis procedure, the nonlinear analysis and loss assessment results for the L1-2S-60x30-GWB archetype are presented as an illustration for the multi-family residential buildings. Figure 2.18 shows the pushover analysis results for the L1-2S-60x30-GWB archetype. Similar to the single-family residence, the soft story issue can be observed in the SWOF buildings as shown by the pushover analysis. For L1-2S-60x30-GWB, the X- and Y-direction have peak strengths of 33% and 35% of the seismic weight, respectively. The pushover curves are governed by drift demands in the 1st story. In other words, most of story drift demands concentrate at the SWOF level. In the X-direction, at peak strength point, the 1st story has a drift of approximately 1% while the 2nd story only has a 0.2% drift demand. Since no residual strengths are considered in multi-family house materials, the pushover curves directly degrade to zero after peak strength. The pushover results will be used for computing the period-based ductility to account for spectra shape variations in the collapse performance assessment. The period-based ductility μ_T is defined as the ratio between the effective yield roof drift δ_{eff} and ultimate roof drift δ_u. In practice, δ_u is the roof drift corresponding to 20% strength loss, which are 0.28% and 0.44% in the two directions in

this case. δ_{eff} is the roof drift when building achieves its peak strength with initial stiffness, which are 0.12% and 0.20% for X- and Y-direction, respectively.

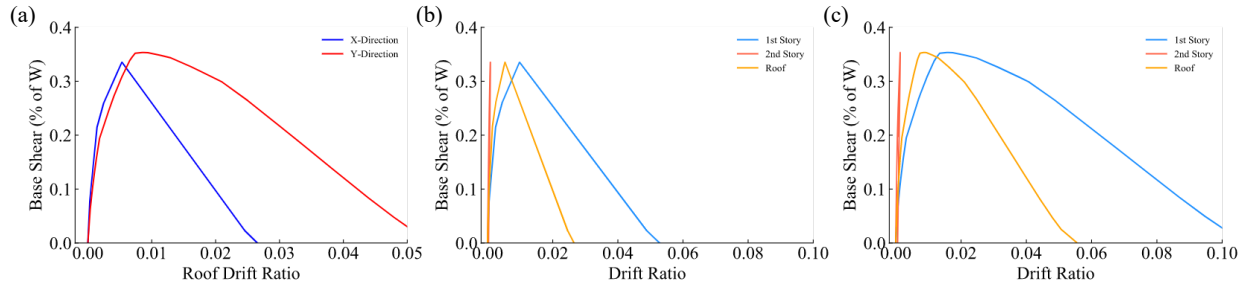


Figure 2.18 Pushover analysis results for L1-2S-60x30-GWB: (a) base shear vs. roof drift in two perpendicular directions, (b) base shear vs. story drift in X-direction and (c) base shear vs. story drift in Y-direction

Figure 2.19 presents the median story drift demands at the two stories and the mean non-collapse deformed shape. The soft-story mechanism similar to what was observed in the single-family residence with cripple walls is observed. Due to the presence of the missing wall line, almost all displacement occurs at the 1st story under dynamic loading. For all computed 15 intensity levels, collapse happens to the 1st story with the upper story remaining undamaged.

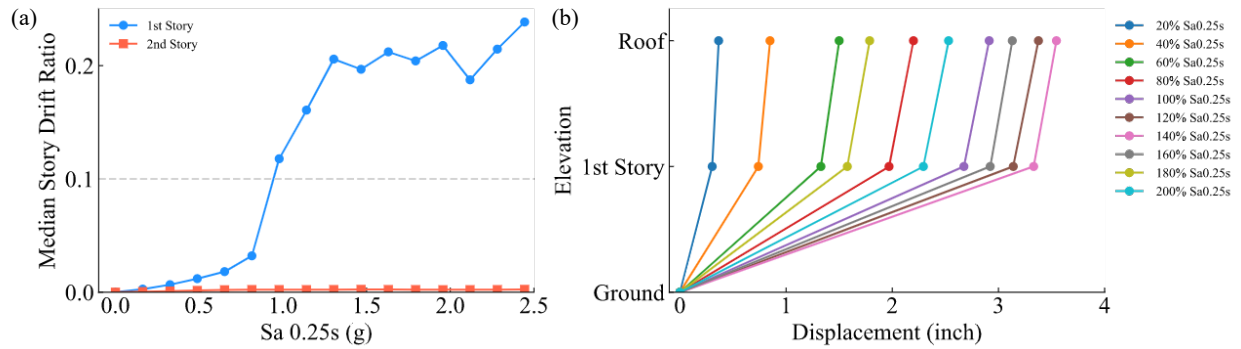


Figure 2.19 (a) Median story drift ratio at 1st and 2nd story for each intensity level, (b) mean non-collapse structure deformed shape for each intensity level

Figure 2.20 shows the fragility curve fitted from the IDA empirical results. Based on FEMA P-695 specification, the collapse margin ratio is required to be adjusted to account for spectral shape. A spectral shape factor (SSF) is obtained by interpolating the period and period-based ductility in

Table 7-1a of FEMA P-695 (FEMA 2009). Since no instructions were provided for 3D buildings, the period-based ductility is computed as the average of the two directions. Based on the previous calculation, μ_T is 2.26 for the studied building, and the resulting SSF is 1.065. The solid blue curve in Figure 2.20 shows the adjusted fragility curve, where a deviation can be observed from the empirical collapse probabilities. Similar to the single-family residence example, when 0.35 is used for model uncertainty, the probability of collapse is lower than the one without model uncertainty.

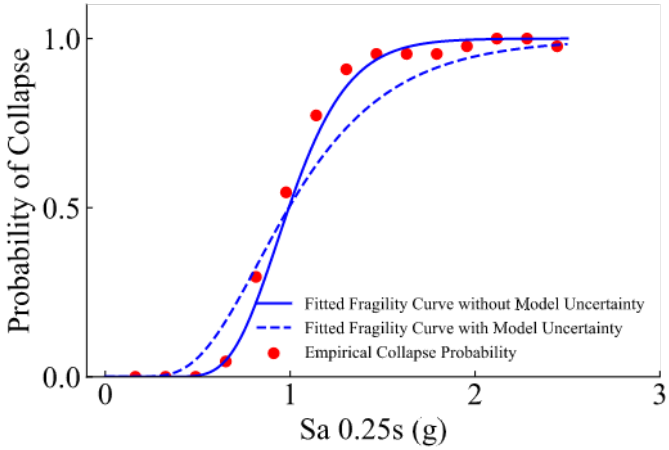


Figure 2.20 Fragility curve for the L1-2S-60x30-GWB archetype

Figure 2.21 presents the loss assessment results for the L1-2S-60x30-GWB archetype. Demolition is based on 1% residual drift ratio and 25% of the building value is assumed for debris removal in the case of demolition. For intensity levels below 0.6g, the demolition loss increases with the intensity measure, and it dominates the total loss. Because most of the story drift demands occur at the 1st story, where the quantity of components is less than the second story, component loss is of lower in magnitude than the collapse and demolition losses. As the intensity increases, the building collapse risk increases and governs the total loss. The analyses with large residual drifts turn into collapse cases and leads to reduced demolition losses. Since component loss calculation is conditioned on non-collapse and non-demolition Monte Carlo realizations, the component loss degrades to almost zero in this stage. As for the aggregated EAL, it is computed to be

approximately 3% of the building value. The demolition loss is the predominant category, which accounts for 61% of the EAL. Collapse and component losses account for 30% and 9%, respectively. Though collapse loss dominates under high shaking intensities, the associated exceedance rates are low. Thus, the EAL contribution from collapse risk is less significant than demolition.

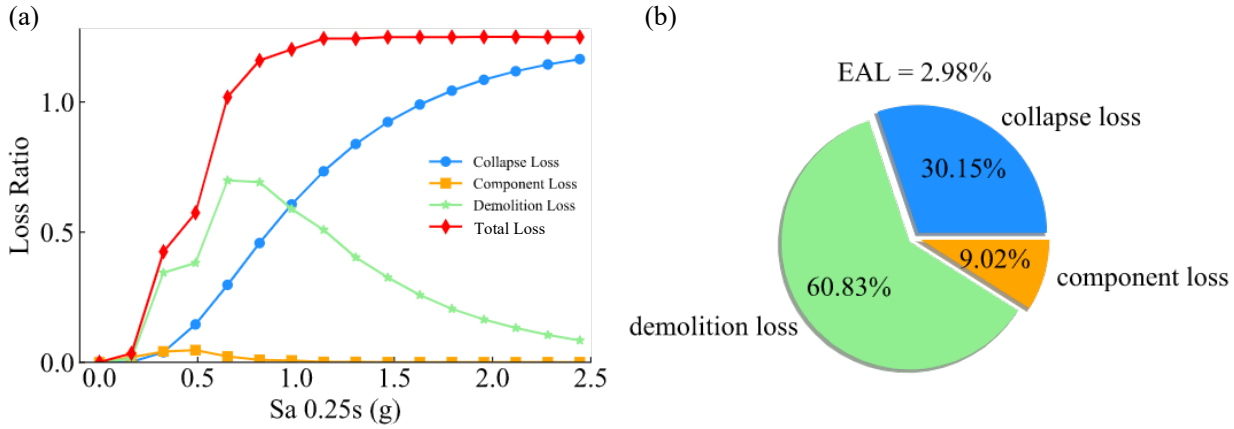


Figure 2.21 Loss Assessment results for L1-2S-60x30-GWB: (a) intensity-based loss curve and (b) expected annual loss

2.12. Summary

To address the need for assess the performance of large numbers of buildings, a Python based end-to-end platform for nonlinear structural analysis and performance-based assessments is developed. This chapter introduces the details involved in the numerical modeling and analysis procedure. The baseline archetypes developed in the PEER-CEA Project are used to demonstrate the analysis and performance assessment features for single-family dwellings. An archetype representing a SWOF building under the purview of the Los Angeles soft-story ordinance is considered for multi-family dwellings.

Two sample analysis results for single- and multi-family dwellings are provided. From the analysis, the soft story vulnerability is observed. The presence of a soft story (cripple wall in single-family

and SWOF in multi-family residence) results in significant differences in strength and stiffness discrepancies along the building height. As a result, large displacement demands concentrate at the soft story under both static and dynamic loading. These observations motivate the investigation of the effect of different retrofit strategies on the seismic performance of single and multi-family woodframe residences.

Use the end-to-end platform, performance evaluation and loss assessment for large number of archetypes can be realized, which can be incorporated in sensitivity analyses, uncertainty quantification, regional seismic performance evaluation and the development of analytics-driven models. In the following chapters, all numerical models, analyses and performance assessments are based on the procedure described here.

CHAPTER 3. DEVELOPMENT AND VERIFICATION OF ANALYTICS-DRIVEN MODEL

3.1. Introduction

3.1.1. Background on Machine Learning Application in Structural Engineering

With the rapid evolution of computing hardware performance and better-developed open source libraries, machine learning (ML) has been widely adopted to solve different types of engineering problems. Within structural/earthquake engineering, there has been several state-of-the-art reviews of ML applications. Xie et al. (2020) summarized the state-of-art in applying ML methods to earthquake engineering problems. The studies reviewed in the paper were categorized based on the ML methods, topic area, data resources and scale of analysis. Popular algorithms including Artificial Neural Network (ANN), Support Vector Machine (SVM), Response Surface Model, Logistic Regression and Hybrid methods were discussed. The author also discussed a broad variety of prior ML studies on seismic hazard analysis, system identification and damage detection, seismic fragility assessment and structural control for earthquake mitigation. Sun et al. (2020) conducted a comprehensive review of ML applications for structural design and performance assessment (SDPA). The details of the formal model development procedure was presented along with related studies. Similarly, the ML-SDPA publications were characterized based on the following four topics: 1. Building response and performance assessment prediction, 2. Model development using empirical data from physical experiments, 3. Multi-media information retrieval, and 4. ML models developed using structural health monitoring and field reconnaissance data. In addition to summarizing the prior work, the authors also discussed the challenges that need to be addressed to enable practical applications of ML to build SDPA problems.

3.1.2. Objective and Scope

As introduced in Chapter 1, the potential vulnerability of soft-story buildings is generally resolved by policy actions, which are implemented on a regional scale. From this perspective, when modeling woodframe buildings, there will be large numbers of design variants. These design variants will explicitly or inexplicitly impact building seismic performances and retrofit designs. The ability to quantify these underlying impacts would contribute to a better and more comprehensive understanding of the effect of the proposed retrofit on seismic performance. Based on this understanding, specific target performance outcomes can be specified. However, on a regional scale, this computational expense is likely to be significant even with the end-to-end platform described in Chapter 2. Therefore, this chapter investigates the use of ML-based surrogate models to reduce the computational expense of regional impact assessments. These surrogate models will be 1. used to evaluate the sensitivity of seismic performance to different structural characteristics, 2. enable efficient and effective quantification of model uncertainties, and 3. embedded within a regionally-targeted seismic retrofit optimization framework.

First, an overview of different ML models is provided with a focus on regression. The adopted models include linear regression, response surface method (RSM), LASSO regression, ridge regression, regression tree and ensemble trees. Then, a set of ML models are developed to predict the median collapse intensities and expected annual losses (EAL) for a set of building cases for different features that are generated using Latin Hypercube Sampling. The ML model development procedure is introduced alongside with the application. A baseline single family dwelling with stucco as the exterior wall material, gypsum wallboard as interior wall material, and stucco cripple wall is used to illustrate the model development. Lastly, the model performances are discussed and compared. The investigation here aims to present a general ML model development procedure,

their application and the interpretation of the results. More specifically, the ML models developed in this chapter will be used for sensitivity analysis and uncertainty quantification in Chapter 4 and regional performance-based retrofit design optimization in Chapter 6.

3.2. Overview of Machine Learning Methods

Machine learning (ML) refers to the mathematical models or algorithms that can 'learn' the inherent nature of a set of sample data (so-called training data) without being explicit programmed and perform decision making (Koza et al. 1996). An overview of the different categories of ML and a few classical algorithms are provided in **Error! Reference source not found.**. Generally speaking, ML algorithms can be characterized as supervised learning, unsupervised learning and reinforcement learning models depending on the objectives. For supervised learning, the task is to make predictions given input data/features. The training set is labeled, which means there exists responses/labels for each input datapoint. The supervised learning model parameters are derived through minimizing the differences between the predictions and actual responses. Based on the tasks, supervised learning can be further characterized as regression and classification. As for unsupervised learning, the training data is not labelled, and the algorithm tries to capture the inherent data structures and patterns. Dimension reduction and clustering are the two main categories of unsupervised learning models. Reinforcement learning mainly aims to make an agent adaptively learn the environment, take actions and maximize the cumulative rewards from environment feedbacks. A well-known reinforcement learning application is AlphaGo, which performed surprisingly well in board games. The main objective of this Chapter is to develop machine learning models to predict building seismic responses and seismic performances given structural characteristics. The current study will only focus on regression models.

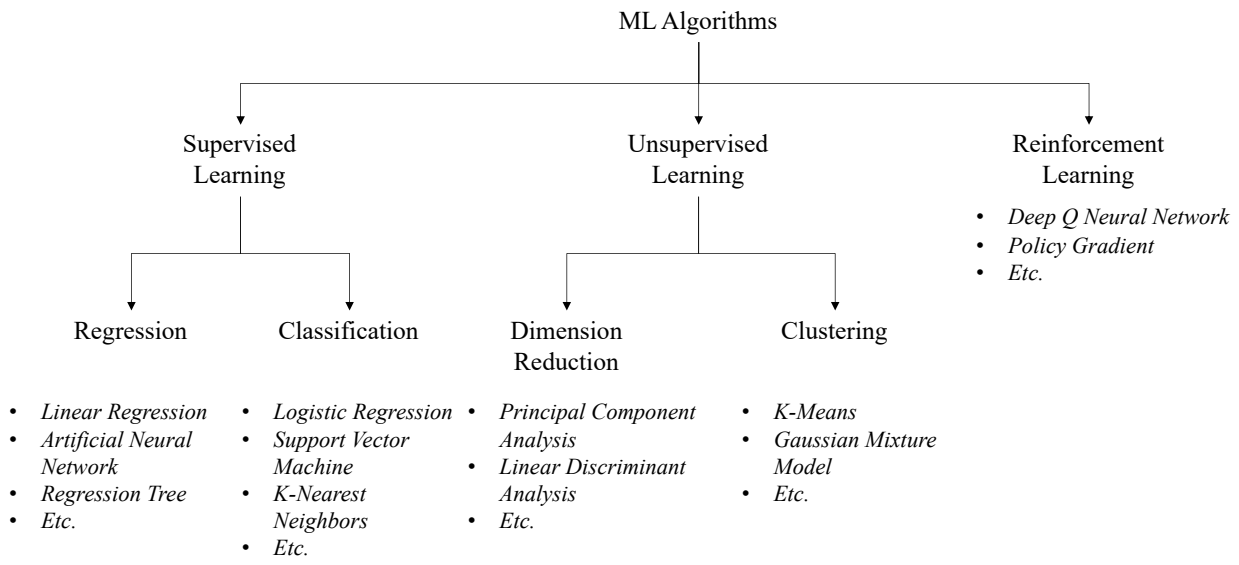


Figure 3.1 Brief overview of machine learning algorithm categories and examples

3.2.1. Introduction of General Regression Algorithms

Listed in Table 3.1 are some classical supervised learning regression models that will be introduced in this section. The selected algorithms are effective and efficient in handling regression tasks. These algorithms can be characterized as linear regression models and tree-based models. The mathematical background provided later would bring more insights into the differences between the two model classes.

Table 3.1 Summary of some common regression models

Algorithm	Original Work	Model Structure
Linear Regression	Galton 1894; Pearson 1896	$y = X\beta + \epsilon$
Response Surface Method	Box and Wilson 1951	$y = X\beta + \epsilon$
Lasso Regression	Tibshirani 1996	$y = X\beta + \epsilon$
Ridge Regression	Hoerl and Kennard, 1970	$y = X\beta + \epsilon$
Classification and Regression Tree	Breiman et al. 1984	$y = \sum_i w_i + \epsilon$ (Non-parametric)
Random Forest	Breiman 2001	$y = \sum_i w_i f_i(x) + \epsilon$ (Non-parametric)
Extreme Gradient Boosting	Chen and Guestrin, 2016	$y = \sum_i f_i(x) + \epsilon$ (Non-parametric)

Before presenting the detailed derivation of each model, some commonly used terminology are specified. As noted earlier, regression tasks seek to make predictions for continuous variables given the input vectors. x_i with $p \times 1$ dimension, denoting the i^{th} input vector with p features. The matrix form $X_{n\times p}$ is the row stacking of n examples, where $X_{n\times p} = (x_1, x_2, \dots, x_n)^T$. The corresponding i^{th} ground truth response is denoted as y_i, and the matrix form is $y = (y_1, y_2, \dots, y_n)^T$. A regression model $f(.;\theta)$ is essentially a mapping (parametrized by θ) from the input vector to the response. The model outcome or prediction can be written as $f(x_i;\theta)$ ($f(x_i)$ for simplicity) or $\widehat{y_i}$. The input matrix is split into training (X_{train}, y_{train}) and testing sets (X_{test}, y_{test}), with the goal of finding the best model parameters and examining the model performance, respectively. Model parameters θ are obtained by minimizing the differences between the ground truth response and model output on the training data. This process is called model training. Quantitative measurements of the aforementioned differences are defined as loss functions (denoted as $L(f(X), y)$). In summary, the training process seeks to acquire model

parameters that minimize the loss function $L(f(X), y)$ to on the training set. The model performance on the testing set informs the model generalizability on 'unseen' data.

An important characteristic that distinguishes ML methods from other statistical analysis techniques is generalizability. The ML model is trained on a small set of known data, and is expected to perform predictions on unseen data. This desire has spurred the broad topic of error analysis in statistical learning theory. A brief introduction to error analysis, overfitting and bias-variance trade-off is presented here. The discussion aims to provide a general sense on the sources of issues that may arise in the model development rather than delivering a comprehensive mathematical derivation. The generalizability of a ML model is measured by generalized error/out-of-sample error, which is defined as the prediction accuracy on unseen data. The theoretical formulation of generalized error is given in Equation (3.1).

$$I\big(f(x;\theta)\big) = \int_{X\times y} L(f(X), y)\rho(X, y)dXdy \tag{3.1}$$

Where $\rho(X, y)$ is the joint probability distribution of X and y. Since the predictions are performed on unknown data, the ground truth distribution $\rho(X, y)$ is usually unknown. Instead, the empirical/training error shown in Equation (3.2) is accessed on the training set with finite a number of data.

$$I_S\big(f(x;\theta)\big) = \frac{1}{n}\sum_{i=1}^{n} L(f(x_i), y_i) \tag{3.2}$$

Therefore, to make reliable predictions on unseen data, the differences between the empirical error and generalized error are needed to be minimized. Practically, the minimization is achieved by bounding the differences in probability as presented in Equation (3.3).

$$P = P\big(I\big(f(x;\theta)\big) - I_S\big(f(x;\theta)\big) \le \epsilon_n\big) \ge 1 - \delta(\epsilon_n) \tag{3.3}$$

Where ϵ_n is the error bound and δ is a function of ϵ_n. As shown in previous research (Abu-Mostafa 1989; Dudley et al. 1991; Alon et al. 1997), the above expression converges with a rate for some types of models trained on finite number of data. The aforementioned references provide a solid foundation supporting the use of a finite number of training data to develop a broadly applicable predictive model. It is also the motivation behind splitting the dataset for the usage of training and testing.

Recall **Error! Reference source not found.**, all listed ML models take a general form of $y = f(x) + \epsilon$, where $f(x)$ is the model systematic structure and ϵ is random noise. It turns out that for any given form of $f(x)$, the mean squared error (MSE) can be decomposed into the linear combination of the model bias and variance as illustrated in Equations (3.4) through (3.6) (Trevor et al. 2009; James et al. 2013).

$$E\left[\left(y - \hat{f}(x;D)\right)^2\right] = \left(Bias_D\left(\hat{f}(x;D)\right)\right)^2 + Var_D\left(\hat{f}(x;D)\right) + \sigma^2 \tag{3.4}$$

$$Bias\left(\hat{f}(x;D)\right) = E\left(\hat{f}(x;D)\right) - f(x) \tag{3.5}$$

$$Var\left(\hat{f}(x;D)\right) = E\left(\left(E\left(\hat{f}(x;D)\right) - \hat{f}(x;D)\right)^2\right) \tag{3.6}$$

Where D is a dataset containing a finite number of examples, $\hat{f}(x;D)$ refers to the prediction on the dataset D, and σ^2 is the irreducible error. Intuitively, bias refers to the prediction error on the unseen data, and the variance represents the model performance on the training data. Given a constant unavoidable error σ^2, the model MSE inevitably makes a trade-off between model generalizability and training prediction accuracy. Consequently, in the process of minimizing training/empirical error, the overfitting issue may arise. Overfitting refers to the case where the model is too flexible or complex to absorb some spurious patterns within the training set and leads to high generalized error. In this case, the model performs perfectly on the training set while

sacrificing the accuracy on the testing data as a result of the variance-bias trade-off. This issue is affected by the model selection and complexity, the size of the dataset, the number of features and other factors. In general, overfitting happens when training a more complex model on less abundant data. This issue can be avoided by using cross-validation and regularization in the models. Overfitting and the bias-variance trade-off is further discussed later in the chapter in the context of the surrogate model development.

3.2.1.1. Linear Regression

Linear regression uses a linear combination of features to predict responses, and its matrix form as shown in Equation (3.7).

$$y = \hat{y} + \epsilon = X\beta + \epsilon \tag{3.7}$$

Where X is feature matrix with $n \times p$ dimension, including an intercept term. β is a $p \times 1$ coefficient vector. β describes the average change in the response $\hat{y}$ caused by a unit change in the feature vector, which can be interpreted as feature sensitivity or importance. ϵ is the random error, which is assumed to independently and identically normal $N(0, \sigma^2)$ with a constant variance σ^2. Based on the additivity of expectation, the observation y also follows normal distribution $N(X\beta, \sigma^2)$. The main objective is to minimize the differences between prediction $\hat{y}$ and observation y to obtain the coefficient vector β. Both minimizing residual sum of squares (Equation (3.8)) and maximum likelihood (Equation (3.9)) can give the estimator. The corresponding estimators are denoted as ordinary least square (OLS) estimator and maximum likelihood estimator (MLE).

$$\hat{\beta}_{OLS} = \arg_\beta \min \sum_{i=1}^{n} (y_i - x_i^T \beta)^2 = \arg_\beta \min (y - X\beta)^T (y - X\beta) \tag{3.8}$$

$$\hat{\beta}_{MLE} = \arg_{\beta} \max \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - x_i^T \beta)^2}{2\sigma^2}\right) \tag{3.9}$$

The estimation of β derived from the two expressions take the same form as shown in Equation (3.10).

$$\hat{\beta} = \hat{\beta}_{OLS} = \hat{\beta}_{MLE} = (X^T X)^{-1} X^T y \tag{3.10}$$

The estimator takes a simple form and has a desirable unbiased property. The expectation of the coefficient $\hat{\beta}$ is the same as the ground truth value β as proven in Equation (3.11).

$$E(\hat{\beta}) = E((X^T X)^{-1} X^T y) = E((X^T X)^{-1} X^T X \beta) = E(\beta) \tag{3.11}$$

The simple form and desirable property make linear regression suitable for the first model to be explored as part of supervised learning. Additionally, linear regression has good interpretability and is widely used in statistical analysis. The variance of $\hat{\beta}$ can be computed using Equation (3.12).

$$Var(\hat{\beta}) = Var((X^T X)^{-1} X^T y) = (X^T X)^{-1} X^T Var(y)((X^T X)^{-1} X^T)^T = \sigma^2 (X^T X)^{-1} \tag{3.12}$$

Based on the normality assumption for ϵ $\hat{\beta}$ also follow a normal distribution $N(\beta, \sigma^2 (X^T X)^{-1})$. With the distribution information, hypothesis testing can be applied to examine the statistical significance of one or several features' impacts on the responses. Typical testing statistics include t-Test (Student 1908), F-Test/ANOVA (Welch 1951), likelihood ratio test (Wilks 1938) and Wald Test (Wald 1939). Other than hypothesis testing, a confidence interval can be obtained from the distribution.

Another important use of the estimator is outlier detection. Several statistical measures can be used to evaluate a data point's impact on the model and leverage. The leverage of the i^{th} data point h_{ii} can be calculated using Equation (3.13).

$$h_{ii} = H_{ii} = (X(X^T X)^{-1} X^T)_{ii} \tag{3.13}$$

Where leverage is the i^{th} diagonal element of the matrix H. The variance for the i^{th} error term e_i is given by Equation (3.14).

$$Var(e_i) = Var(y - x_i^T\beta) = (1 - h_{ii})\sigma^2 \tag{3.14}$$

For a close to 1 leverage, the variance of the error is close to zero such that the i^{th} data point 'pulls' the regression line towards itself. Intuitively, the i^{th} data point will have a greater impact on the slope of the regression line. h_{ii} is minimized when $x_i = \bar{x}$, and vise versa. Therefore, for a large h_{ii}, the data point could potentially be an outlier. A general rule of thumb to identify outliers is $h_{ii} > \frac{2p}{n}$, which is equivalent to two times the average leverage. This calculation will be adopted later to examine whether the dataset contains outliers.

Recall Equation (3.10), X^TX is needed to be full rank to calculate the matrix inversion. Therefore, the full rank requirement limits the model capability on correlated features, the so-called collinearity issue. When perfect or strong collinearity occurs, where some features are highly correlated or linear combinations of other features, X^TX is not a full rank matrix. For such cases, though the inversion can be computed by pseudo inversion ($(X^TX)^+$), the estimator is sensitive to newly added data. Also, the coefficients would be unstable such that the estimation of feature sensitivity is not reliable. This is one of the primary motivations for introducing other ML algorithms to assess building performance sensitivity.

3.2.1.2. Response Surface Method (RSM)

RSM, which can also be regarded as second order polynomial regression, is a special case of linear regression. In RSM, when constructing the feature matrix X, the independent features are their interactions are considered. The scalar form of RMS is presented in Equation (3.15).

$$\hat{y}_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j} + \sum_{j,k} \beta_{jk} x_{i,j} x_{i,k} \tag{3.15}$$

where β_0 is the intercept term, β_j is the coefficient of the j^{th} independent feature, and β_{jk} is the coefficient of the $j^{th}$and k^{th} features' interaction effect. RSM can provide a statistical basis for capturing how the response variable changes when multiple features change at the same time. Though higher than two order effects can be introduced to form higher order polynomial regression, RSM is robust enough to handle the regression task covered in this chapter.

As introduced earlier, complex models can easily overfit a given dataset. To reduce the risk of overfitting in linear regression, penalizations are added to loss functions. Based on different types of regularizations (l_1 and l_2), the LASSO and Ridge are two commonly used penalized linear regression models. The elastic net (Zou and Hastie 2005), which uses a weighted average of l_1 and l_2 penalization to address the shortcoming of Lasso Regression, is also sometimes used. However, elastic net is not considered here.

3.2.1.3. LASSO Regression

Introducing l_1 penalization on coefficients to the loss function as presented in Equation (3.16) gives Lasso Regression.

$$L(f(X), y) = (y - X\beta)^T(y - X\beta) + \lambda\left|\left|\beta\right|\right|_{l_1} = (y - X\beta)^T(y - X\beta) + \lambda|\beta| \qquad (3.16)$$

Where λ is a hyperparameter that affects the level of penalization. However, the Lasso Regression loss function is not differentiable near $\beta = 0$, which forms a non-smooth optimization problem. By applying the subgradient method, the solution shown in Equation (3.17) can be derived for the case where all features are independent ($X^TX = I$) (Donoho and Johnstone 1995; Tibshirani 1996).

$$\hat{\beta}_k^{LASSO} = sign\left(\hat{\beta}_k^{OLS}\right)\left(\left|\hat{\beta}_k^{OLS}\right| - \frac{\lambda}{2}\right)_+ = sign\left(\hat{\beta}_k^{OLS}\right)\max\left(0, \left|\hat{\beta}_k^{OLS}\right| - \frac{\lambda}{2}\right) \qquad (3.17)$$

Where $\hat{\beta}_k^{LASSO}$ is the LASSO coefficient for the k^{th} feature. As for more general cases where the features are correlated, $\hat{\beta}_{LASSO}$ can be solved through numerical methods such as coordinate

descent and proximal gradient method. One desirable property of the LASSO estimator is that the coefficients of the features with low impacts on the responses are shrunk to zero. Thus, Lasso Regression can be used for feature selection.

3.2.1.4. Ridge Regression

Substituting the l_1 with l_2 penalization in Equation (3.16) gives the loss function for Ridge Regression as given by Equation (3.18).

$$L(f(X), y) = (y - X\beta)^T(y - X\beta) + \lambda\left|\left|\beta\right|\right|_{l_2} = (y - X\beta)^T(y - X\beta) + \lambda\beta^2 \tag{3.18}$$

Much simpler than the Lasso estimator, the ridge estimator can be obtained by directly taking the partial derivative with respect to β and setting the derivative to be zero. The ridge estimator takes the form shown in Equation (3.19).

$$\hat{\beta}_{ridge} = \left(X^T X + \lambda I_p\right)^{-1} X^T y \tag{3.19}$$

Where I_p is a p dimensional identity matrix. Compared to the OLS estimator, ridge regression estimator is shrunken by a constant value of $1 + \lambda$. Recall Equation (3.7), the model formed by the prediction $X\beta$ and random noise ϵ. Intuitively, such penalization prevents the model from leanrning too much unavoidable error (ϵ_{train}) from the training set.

Other than the preceding formular, ridge regression can take various forms. For example, mapping the inner product $X^T X$ to Hilbert Space can lead to kernel ridge regression. Considering the scope of this Chapter and some initial results from pilot studies, ridge regression transformations are not considered.

3.2.1.5. Classification and Regression Tree (CART)

The classification and regression tree (CART) model refers to a non-parametric function class which recursively partitions the input space and defines a model for each of the resulting sub-spaces. The general procedure is to recursively: (1) perform prediction on the split sub-spaces, (2)

find the best split of the input space, until no further partition can be performed or meet some specified stopping rules. The process is an analogy of a family tree. The earlier partitions can be regarded as the ancestors of the later partitioned features and points. The formal process is described in the pseudo code in Table 3.2. The pseudo code uses group average for regression predictions. Alternatively, linear regression or other basic functions can be fitted locally to obtain a more robust CART model.

Table 3.2 Pseudo code for CART generator (Algorithm 16.1 in Murphy 2012)

Input: training set $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$*, node node, tree depth* d
function TreeGenerator$(D, node, d)$*:*
node.prediction $=$ *mean*$(y_i : i \in D)$ *or label distribution*$(y_i : i \in D)$*;*
$(j^*, v^*, D_{left}, D_{right}) =$ *split*(D)*;*
if *StopSplitting*$(d, cost, D_{left}, D_{right})$*,* ***then***
return node
else
node.left=TreeGenerator$(D_{left}, node, d+1)$
node.right=TreeGenerator$(D_{right}, node, d+1)$
return node

One critical point in the CART development process is to find the best 'cut-off' feature and value to split the dataset D into D_{left} and D_{right}. Generally, binary partition is performed to avoid non-binary trees, where a small number of data may fall into a subtree and cause overfitting. For each partition manipulation, there exists an associated cost in the resulting subspaces. The value and feature to spit the dataset are determined through greedily minimizing the cost as illustrated in Equation (3.20) (Murphy 2012).

$$(j^*, v^*) = \arg \min_{j \in \{1,\dots,p\}} \min_{v \in V_j} cost\left(\{x_i, y_i : x_{ij} \le v\}\right) + cost\left(\{x_i, y_i : x_{ij} > v\}\right) \tag{3.20}$$

Where j^* and v^* are the optimal feature and value of the feature to split the dataset. p is the total number of features and V_j are all possible values of the j^{th} feature. For regression tasks, a common choice for the cost function is the squared sample deviation shown in Equation (3.21) (Murphy 2012).

$$cost(D) = \sum_{i \in D} (y_i - \bar{y})^2 \tag{3.21}$$

Where y_i is the response in dataset D and $\bar{y}$ is the sample mean. The StopSplitting function determines whether a specified stop rule is satisfied to stop growing a tree. Some commonly used stopping criteria include the cost relative to a target threshold, the tree depth relative to a defined maximum depth and the number of data points in the resulting subspaces relative to a predefined threshold. The stopping rules can help to prevent generating a deep tree and overfitting.

The divide-and-conquer nature gives CART some superior properties compared to simple linear regression. Recursively splitting the data improves model interpretability. Each step in CART generation is analogous to decision making. Different policies/models are determined for each of the local regions. The selection of features as the basis of the data partitioning informs feature importance. In general, the responses are more sensitive to the features that are used early in the data splitting process. Quantitatively, features' impacts can be computed by the cost reduction/information gain after applying split based on the feature under consideration. Additionally, CART can easily handle missing values in the dataset by assigning higher weights to the features with more complete data.

Though as introduced above, CART is interpretable, informative, and not sensitive to missing data, it also easily overfits, especially for trees with complicated structure. More importantly, its divide-and-conquer nature makes CART more appropriate to handle categorical variables than continuous variables, which is the main focus of the thesis. Ensemble methods provide a good supplement to

CART to address such problems. Ensemble models makes decisions through aggregating the outcomes from a series of so-called weak/base learners, herein CART. Large numbers of weak learners can be developed either parallelly or sequentially, which can then contribute to the final prediction. Some commonhhly used aggregation methods including bagging, boosting, and stacking. Random Forest and Extreme Gradient Boosting (XGBoost) have shown outstanding performances on the similar datasets generated in some pilot studies compared to other types of machine learning algorithms (e.g. Kernel Ridge Regression, Gradient Boosting and Ada Boosting). Additionally, they are representative of bagging and boosting models. Therefore, they are introduced here in details and implemented for building seismic performance predictive models.

3.2.1.6. Random Forest

As suggested by the name 'Forest', Random Forest is formed by combining a large number of CART models. Its fundamental structure is presented in Figure 3.2. Based on the Hoeffding inequality (Hoeffding 1994), the error rate of aggregated independent CART models declines exponentially with the number of models. Though the development of CART cannot be fully independent (given a same training set), introducing fluctuations when growing each tree can bring enough discrepancies in each base learner. As demonstrated by Figure 3.2, n sub-training sets are first formed by bootstrap sampling from the original training set. Then, each data point has a $\frac{1}{N}$ (N is the size of the original training set) probability of appearing in a sub-training set. Besides the input data, the features can also be sampled to determine whether they will be used when growing the CART. Then, n CART models are derived on the basis of each sub-training set. Lastly, the outcomes of each CART model are aggregated to predict responses. The process can be parallelized.

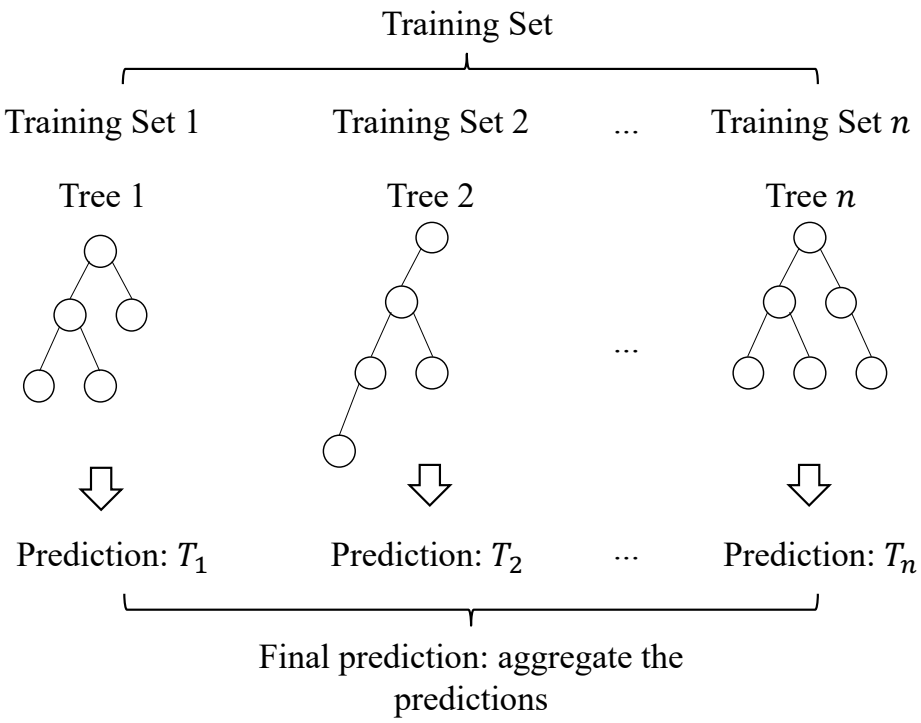


Figure 3.2 Structure of the Random Forest model

Different strategies can be adopted to bag single CART predictions. Common options include simple averaging, weighted averaging, majority voting and weighted voting. For the regression task this Chapter, simple averaging calculated using Equation (3.22) is applied.

$$f(x) = \frac{1}{n}\sum_{i=1}^{n} T_i \tag{3.22}$$

Where, n is the total number of tree models and T_i is the prediction of the i^{th} tree. From the perspective of bias-variance trade-off, the above expression explains how Random Forest reduces prediction variance. For approximately independent CART predictions, the variance of the aggregated predictions can be estimated using Equation (3.23). On average, the variance of the final prediction is lower than that of a single CART model. Therefore, Random Forest mainly focuses on reducing variance and improves generalizability.

$$Var\big(f(x)\big) = Var\left(\frac{1}{n}\sum_{i=1}^{n} T_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(T_i) \tag{3.23}$$

3.2.1.7. Extreme Gradient Boosting (XGBoost)

XGBoost combines the accuracy of ensemble learning, the high efficiency and flexibility of the gradient boosting model, the suitability for parallel computing and a low risk of overfitting. Different from Random Forest, XGBoost uses a sequenced rather than parallel set of tree-based models' aggregation to predict the responses. Its fundamental structure is presented Figure 3.3. A decision tree model performs a piecewise division of the input domain generating predictions locally. In each step of XGBoost, a sub regression decision tree is developed on the residuals from previous trees, which adaptively boosts the performance of the new tree.

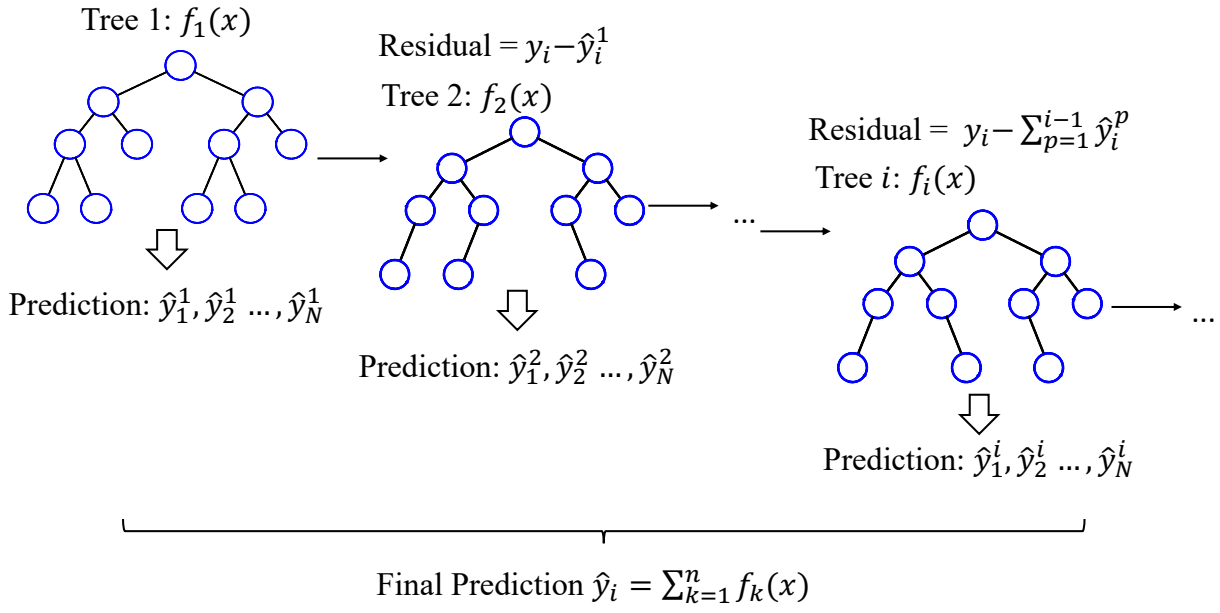


Figure 3.3 Structure of the XGBoost model

The XGBoost model predicts responses by summing over the predictions from individual tree-based models as shown in Equation (3.24).

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{3.24}$$

where f_k is the k^{th} tree model computed by minimizing the objective function. x_i is the feature vector of data point i. $\hat{y}_i$ is the prediction/response for the i^{th} data point.

In practice, when a large number of individual trees are involved in XGBoost, the complexity of the model increases and can lead to overfitting, where high accuracy is achieved on the training set but with low generalizability. To balance the model performance and risk of overfitting, the objective function of the XGBoost model shown in Equation (3.25) is used.

$$J = L(\theta) + \Omega(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{3.25}$$

where $L(\theta)$ computes the differences between the real response y_i and prediction $\hat{y}_i$ over n training samples to measure the model's goodness of fit, $\Omega(\theta)$ is the regularization function that penalizes deep and complex models and θ represents the model parameters. Common examples of loss functions used for regression include the mean squared error (MSE) and median absolute error (MAE). The regularization function sums the penalties for function f_i over a total of K models. Specifically, for the XGBoost model, the penalty placed on model complexity is defined in Equation (3.26).

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{3.26}$$

where γ and λ are the regularization magnitude, which are tuned as part of the model development process. T and w_j are the number of leaves and the score of leaf j in the tree-based model f, respectively. The model complexity and penalty increase with the number of leaves and the scores. XGBoost accumulates new trees by reducing the residual from the previous step. At step t, the prediction can be computed using Equation (3.27).

$$\hat{y}_i^t = \sum_{k=1}^{t} f_k(x_i) \tag{3.27}$$

To develop the model for step $t+1$, the residual from step t is minimized and the introduction of a new model f_{t+1} is penalized. The objective function for f_{t+1} is given by Equation (3.28).

$$J_{t+1} = L_{t+1}(\theta) + \Omega_{t+1}(\theta) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^t + f_{t+1}(x_i)\right) + \sum_{k=1}^{t} \Omega(f_k) + \Omega(f_{t+1}) \tag{3.28}$$

Considering the MSE loss function, the penalty term from the previous t steps is fixed when developing the model for step $t+1$ and the previous regularization term can be waived when minimizing the objective function. Equation (3.28) can be further simplified to Equation (3.29).

$$J_{t+1} = \sum_{i=1}^{n} \left(y_i - (\hat{y}_i^t + f_{t+1}(x_i))\right)^2 + \Omega(f_{t+1}) = \sum_{i=1}^{n} 2(\hat{y}_i^t - y_i) f_{t+1}(x_i) + f_{t+1}(x_i)^2 + \Omega(f_{t+1}) \tag{3.29}$$

Equation (3.29) can be estimated using the Taylor series expansion keeping the second order term. The analytical solution of f_{t+1} is computed by minimizing Equation (3.30).

$$J_{t+1} = \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^t) + g_i f_{t+1}(x_i) + \frac{1}{2} h_i f_{t+1}^2(x_i) \right] + \Omega(f_{t+1}) \tag{3.30}$$

where g_i and h_i are the first and second order partial derivative of the loss function with respect to the prediction for training i at step t, respectively. By removing the constant loss from the previous step, the finalized objective function for $f_{t+1}$is achieved through Equation (3.31).

$$J_{t+1} = \sum_{i=1}^{n} \left[g_i f_{t+1}(x_i) + \frac{1}{2} h_i f_{t+1}^2(x_i) \right] + \Omega(f_{t+1}) \tag{3.31}$$

In the development of XGBoost, the penalization term $\Omega(f_{t+1})$ is defined as Equation (3.32) (Chen 2014).

$$\Omega(f_{t+1}) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2 \tag{3.32}$$

Where, γ is the penalization coefficient for the number of leaves (T) in the tree, and w_j is the score of the j^{th} leaf. The aggregation of the leaves' socres give final prediction ($f_{t+1} = \sum_{j=1}^{T} w_j$). l_2 regularization is applied to penalize the leaf giving high scores. Thus, Equation (3.31) can be organized into Equation (3.33) (Chen 2014).

$$J_{t+1} = \sum_{i=1}^{n}\left[g_i w_{q(x_i)} + \frac{1}{2}h_i w_{q(x_i)}^2\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 = \sum_{j=1}^{T}\left[\left(\sum_{i\in I_j} g_i\right)w_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)w_j^2\right] + \gamma T \quad (3.33)$$

Where, $q(x_i)$ refers to an indicator function $i \in I_j$, which yields to 1 if the i^{th} data point is assigned to the j^{th} leaf and 0 otherwise. The items in the square brackets in the equation above is a quadratic function parametrized by w_j. Thus, for a given tree structure, optimal w_j can be derived by minimizing the quadratic function. An optimal objective function value $-\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j+\lambda} + \gamma T$ is achieved at $w_j = -\frac{G_j}{H_j+\lambda}$, where $G_j = \sum_{i\in I_j} g_i$ and $H_j = \sum_{i\in I_j} h_i$. The derivation further helps to quantify the feature importance. Recall the development of tree requires to select the best feature and value to split the dataset minimizing the cost. Each split manipulation would generate a unique indicator function $i \in I_j$ and resulting in different objective function values. The feature and value maximize the cost reduction after splitting are selected as a node the generate the following tree. Thus, the cost reduction can be used to quantify the feature importance. Mathematically, the cost reduction or total can be calculated by differences between objective function values of the resulting left and right trees and the original tree using Equation (3.34) (Chen 2014).

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right] - \gamma \quad (3.34)$$

3.3. Development and Verification of Performance Prediction Model

3.3.1. Literature Review of Regression Model Application in Structural Engineering

Conventionally, nonlinear response history simulations are performed to assess building seismic performance under the framework of performance-based earthquake engineering (PBEE). Large numbers of structure parameter combinations, site conditions and uncertainties may create unaffordable computational burden in the assessment, especially for portfolio level analysis. ML methods are found to be good supplements to mechanics-based models by statistically linking critical analysis parameters to structural responses and seismic performance. Some prior investigations using ML algorithms to construct regression predictive models for structural

response and seismic performances are introduced here. The summary focuses on the studies adopting the algorithms discussed in section 3.2 for regression tasks. For each study, the applied algorithm, structure type, model predictors and predictions, performance metrics and model performances are listed, which help to inform the parameter selection and model assembling in the current study. These studies cover a broad range of construction types, which demonstrates ML methods' capability on structural engineering problems. Among them, Response Surface method is the most popular model for structure response and fragility prediction. The common response variables of interest are fragility parameters (median and dispersion) and engineering demand parameters (EDPs), which could serve as intermediate responses for loss associated performances. Though these works focus on different construction types, the predictors can generally be categorized as structure geometric parameters, material properties, component configuration and hazard exposure parameters. Commonly used model performance metrics for structural responses prediction include root mean squared error (RMSE), root mean absolute error (RMAE), coefficient of determination (R^2), and mean absolute relative difference (MARD). Besides developing promising ML models for regression tasks, these works also applied ML models in sensitivity analysis (Seo et al. 2012; Sichani et al. 2018; Xie and DesRoches, 2019), uncertainty analysis (Xie and DesRoches, 2019), performance-based optimization (Moradi and Burton, 2018) and regional/multi-hazard performance assessment.

Table 3.3 Summary of prior work on the development of regression ML models in structural response and performance prediction

Study	ML Algorithm	Structure Type	Predictors	Prediction	Performance Measure	Model Performance*
Seo et al. 2012	Response Surface model	Steel Moment Frame	Earthquake direction, steel yield strength, damping ratio, bottom floor height, year built, eccentricity, steel Young's modulus	Mean & standard deviation of peak inter story drift ratio	NA	NA
Kameshwar and Padgett 2014	Response surface model	Multi-span simply supported concrete girder bridge	Concrete nominal strength, reinforcement yield strength, span length, column height & diameter, width of deck, number of spans, longitudinal & transverse reinforcement ratio,	Column curvature	RMSE, RMAE, R^2	0.93(R^2), 0.30 (RMSE), 1.02 (RMAE)
Burton et al. 2017	OLS, LASSO regression, Ridge regression	RC Infilled Frames	Ground motion intensity measures, structural response parameters and physical damage indicators	Aftershock building fragility parameters	MARD	0.16
Sichani et al. 2018	Two-layer stepwise regression	Concrete dry cask structures	Structural, geometric, material properties, cask accelerations	Cask horizontal displacement and rocking	R^2	0.94
Moradi and Burton, 2018	Response surface model	Controlled rocking steel braced frames	Yield strength, initial stiffness, strain hardening ratio of the fuse, initial force and modulus of elasticity of the post-tensioning stands	Peak roof drift, residual roof drift, peak floor acceleration	R^2	1.0
Xie and DesRoches 2019	OLS, stepwise regression, LASSO regression	Highway bridge	Soil structure interaction parameters, ground motion parameters	Bridge component demand parameters and fragility estimates	R^2	0.826
Mangalathu and Jeon 2019	Random Forest	Bridge system	Super structure, interior bent, deep foundation, exterior bent, bearing, gap and other parameters	Seismic fragility parameters	R^2, MSE	0.82(R^2), 0.04 (MSE)

*Best model performance among all investigated algorithms are listed in the table

3.3.2. Objective

As introduced in Chapter 1, retrofit policies have been widely enacted in the west coast region to address woodframe building vulnerability issues. These policies focus on improving existing woodframe building seismic performance and strengthening the community-level resilience to seismic hazard. Within the framework of PBEE, structural seismic performance is derived through the combinations of Engineering Demand Parameters (EDPs), fragility functions and hazard information. On each stage of analysis, structural characteristics (e.g. structure configurations and mechanical properties) are expected to have significant impacts on the outcomes. Aggregating such impacts, evaluating performance metrics, convolving feature variations, studying feature sensitivities and quantifying model uncertainties would contribute to insightful understandings of individual building and regional scale seismic performances. Eventually, this type of knowledges would benefit the entire community including decision makers, insurers, site engineers and homeowners. Based on prior work, using predictive models to supplement mechanics-based models within the PBEE framework can inexplicitly integrate and simplify the analysis process while maintaining accuracy.

Several ML models are developed following for performance-based assessment of single-family woodframe buildings with cripple walls using the end-to-end simulation tool. A standard regression ML model development process that involves dataset development, model training, model evaluation and selection. Linear regression, RSM, LASSO regression, ridge regression, Random Forest and XGBoost are compared in terms of predictive accuracy and generalizability. Then, a deep dive into the model applications is provided. The primary focus here is the adoption of the selected models for sensitivity analysis and quantification of uncertainty. Additionally, the entire process is universally applicable for similar structural engineering problems.

3.3.3. Model Development

This section aims to develop different ML models to predict building level seismic performance given structural characteristics. Figure 3.4 shows a flow chart of the ML model development procedure. Overall, the entire process comprises of dataset development, model training, performance evaluation and model selection. At first, the dataset is generated to feed algorithms with training and testing datasets. The development process is supposed to be with a clear objective in mind. Based on the scope of the task, the considered variables and associated ranges should be proposed. Then, a set of parameters, which serve as input features, will be sampled from target distributions and ranges with appropriate sampling techniques. These parameters are further passed to the end-to-end tool to construct OpenSees models. Model seismic performance are obtained by conducting nonlinear response history analysis and performance assessment. The performance measures of interest will be the ML model output responses. Input feature and output response pairs form a complete dataset, which will be split into training and testing set for model training and performance measurement purposes. Focusing on the model training process, the training set is further split into training and validation sets to find the best set of model parameters and fine-tuning, respectively, where the latter seeks to acquire the optimal model hyperparameters or pre-specified model settings that are not learnt from data. Cross-validation should be adopted to reduce the risk of overfitting. Finalized models are obtained by combining the outcome of model training and fine-tuning. Model performance is then evaluated on the testing set and will be used for the model selection criterion. The detailed process is presented in the following discussion on a cripple wall single-family woodframe dwelling.

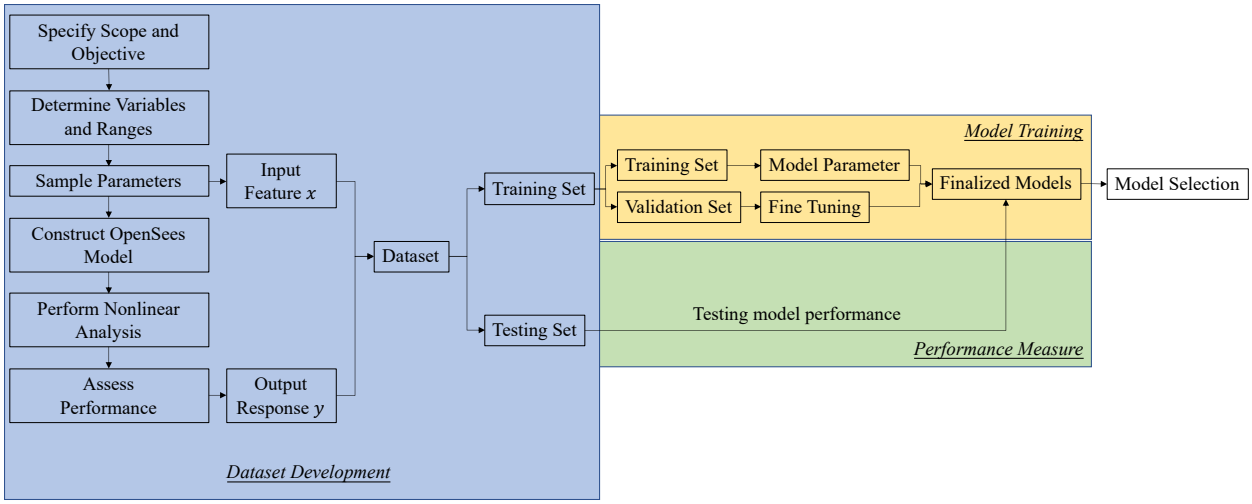


Figure 3.4 Overview ML model development workflow

3.3.3.1. Dataset Development

As introduced in section 3.3.2, eventually the developed predictive model will be used for feature sensitivity analysis and uncertainty quantification. Thus, the generated dataset should satisfy some prerequisites to serve these purposes. A couple of guiding principles are used to design the dataset. To evaluate sensitivity, the features that might significantly contribute to building seismic performance must be considered. The random variables may or may not be independent from each other.

To quantify model uncertainty, the considered sources must be explicitly used as the ML model input. The sampled variable distributions are not necessarily the same as the uncertainty distribution. However, the proposed variable ranges should at least cover all possible values for uncertainty consideration.

Following the above principles, one baseline building with a fixed configuration, and sheathing material combination is selected. If different archetypes and materials types are introduced, either multiple predictive models or one predictive model with higher-level features (e.g. nonlinear static analysis results) are required to achieve desirable model performance. In such cases, the

uncertainties are accumulated through inter-dependent variables, which results in over-counted model uncertainties. Also, different archetypes usually have different structural and non-structural components to compute losses, which cannot be easily handled by ML models. Therefore, the post-1955 stucco exterior gypsum interior cripple wall single family house is selected for demonstration.

A detailed feature generating procedure is discussed here. Table 3.4 lists the assumed ranges and distributions used for generating the random samples. The listed features are both performance sensitive-features and uncertainty sources. The critical features proposed in the ATC 110 project (Blaney et al. 2018) including the number of stories, cripple wall height, seismic weight and sheathing anchorage combination are incorporated. For single family woodframe buildings, cripple wall height usually ranges from 2 to 6ft. To quantify uncertainties on both sides of 2 and 6 ft, 1.5 to 6.5 ft, a uniform distribution is assumed. The variations of seismic weight are introduced via uniformly distributed seismic weight adjustment factor. Seismic weights are first calculated using sampled cripple wall height and self-weight of the material, then a sampled adjustment factor is applied to intentionally decrease or increase the weight. The treatment of material properties is not trivial. Figure 3.5 provides a schematic figure showing the variations in the material force and displacement parameters. Recall that the Pinching4 model uses four force-displacement pairs to define the backbone curves. An adjustment parameter is sampled from truncated a normal distribution and applied to the best estimation of four force/displacement parameters at the same time. The superstructure and cripple wall materials are considered separately. For the superstructure, the exterior and interior wall materials are treated with a same factor. The relative contributions from the two parallel springs and Pinching4 hysteretic parameters remain unchanged. In summary, a total of four random numbers are used here to vary the superstructure and cripple

wall Pinching4 force and displacement parameters. Such simplification has the following advantages.

100% correlation assumption for the force/displacement parameters reduces the number of random variables in the model. Also, it avoids possible numerical issues when different factors are applied to each of the variable (e.g. generating a yield strength that is higher than the peak strength). Considering the force and displacement parameters separately helps identify the relative importance of strength and ductility. The impact can be interpreted as the expected change in building seismic performance caused by a unit change in the superstructure/cripple wall strength/ductility. Woodframe building seismic performance is highly dependent on vertical distribution of strength and stiffness. Separating the superstructure from the cripple wall material helps to model the relative available strength and stiffness differences between them.

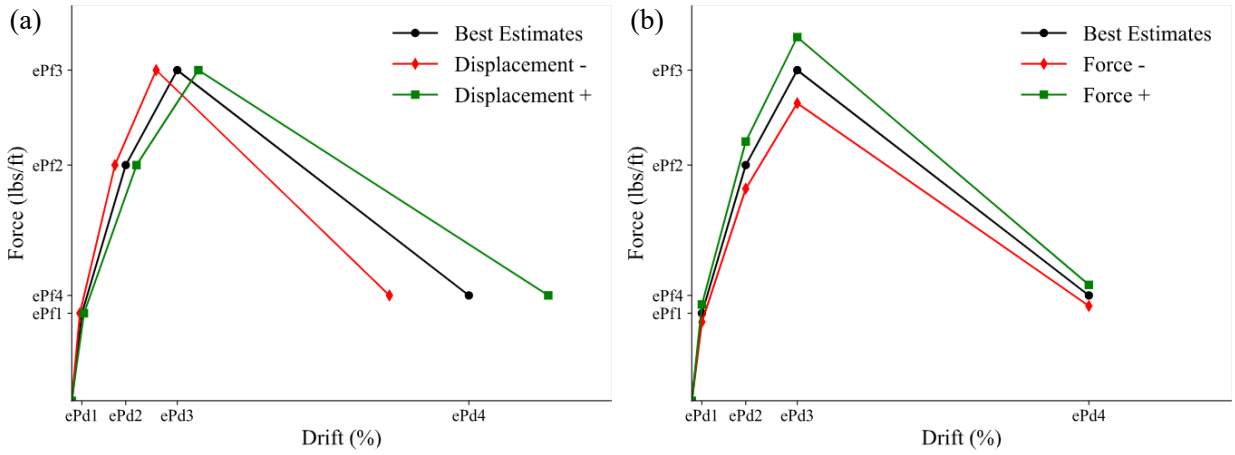


Figure 3.5 Schematic representation of the treatment of Pinching4 force and displacement parameters

Besides the parameters mentioned in ATC 110 project, the damping ratio and retrofit panel length are considered as additional building variants. Uniformly distributed damping ranging from 1% to 5%, and the length of the retrofit panel is assumed to be uniformly distributed between10 to 15ft. The amount of retrofit panel is supposed to be determined following FEMA P-1100. Introducing

structure panel length helps to simplify the development of retrofit design, meaning for each of the building variants combinations structural panel don't have to be designed specifically.

Table 3.4 Considered structural features range and distribution

Variable	Lower Bound	Upper Bound	Distribution
Seismic Weight Adjustment Factor	80% (light)	120% (heavy)	Uniform
Damping Ratio	1%	5%	Uniform
Force Amplification (Super Structure)	0.2	NA	$N(1,0.5)$
Drift Shift (Super Structure)	0.2	NA	$N(1,0.5)$
Force Amplification (Cripple Wall)	0.2	NA	$N(1,0.5)$
Drift Shift (Cripple Wall)	0.2	NA	$N(1,0.5)$
Cripple Wall Height	1.5ft	6.5ft	$N(1,0.5)$
Retrofit Length*	10ft	15ft	Uniform

*Retrofit length is only sampled for retrofitted cripple wall buildings, it refers to the length of a single retrofit panel.

For all the random variables mentioned above, 200 Latin Hypercube samples (LHS) are generated independently for 1- and 2-story, existing- and retrofitted-cripple wall buildings (a total of 800 random models). OpenSees models with the sampled parameter sets are then created using the end-to-end modeling tool. For each model, nonlinear static analysis was conducted to obtain structural properties (peak strength, ductility, and etc.). Multiple stripe analysis (MSA) was performed using ground motions selected based on a conditional spectrum for the San Francisco VS30 = 270m/s site. The EDPs and collapse safety were extracted from the MSA and used for FEMA P-58 analysis. The FEMA P-58 loss assessments were performed following the procedure described in Chapter 2. The quantities of components remain the same as shown in Table 2.15 except for the cripple wall elements, which are adjusted based on the sampled cripple wall heights. The performance metrics of interest are the median collapse intensity and expected annual losses (EAL). A few concerns for selecting the above two metrics as response variables are provided here.

Most woodframe building seismic retrofit policies aim to reduce collapse risk. Using ML models to predict median collapse intensities helps to inform the collapse performance under different scenarios at the building and regional level.

Seismic induced loss is one of the critical outcomes of FEMA P-58 performance assessment framework. However, it is usually implicitly considered in building retrofit policies. Additionally, prior work on ML predictions in structural engineering has not placed a major focus on nloss-based response variables.

3.3.3.2. Model Training

Table 3.5 provides basic information about the features used in the ML models. The features under consideration take different data type and ranges. Though this situation can be addressed by the tree-based algorithms (Random Forest and XGBoost), linear regression models are sensitive to input data scales and may lead to inappropriate interpretation of feature sensitivities. To address this issue, feature engineering is conducted here to unify the feature scales. Common choices for data normalization algorithm including min-max and standard scaler for continuous variables. Equation (3.32) and (3.33) demonstrate how min-max and standard scaler transfer the data, respectively.

$$X_{min-max,i} = \frac{X_i - \min(X)}{\max(X) - \min(X)} \times (\max(X) - \min(X)) + \min(X) \tag{3.32}$$

$$X_{standard,i} = \frac{X_i - \text{mean}(X)}{std(X)} \tag{3.33}$$

Where X_i is the i^{th} data point and X is the entire dataset. As shown above, min-max scaler transfers the data into a unit size by its relative magnitude in the dataset, and standard scaler standardizes the data into a standard normal distribution. Intuitively, the min-max and standard scaler suit the uniform- and normal-like distributions, respectively. Therefore, min-max scalers are

applied for the random variables sampled from the uniform distribution and standard scalers are applied to the truncated normal distributed variables. The number of stories and retrofit indicator are the only two categorical variables involved. Either continuous or categorical data type can be assigned to the number of stories, and retrofit indicator can be simply transferred to Boolean type.

Table 3.5 ML model features

Features	Unit	Type	Sampled Distribution
Number of Stories	NA	Categorical	NA
Retrofit Indicator	NA	Categorical	NA
Seismic Weight	kips	Continuous	Uniform
Damping Ratio	NA	Continuous	Uniform
Cripple Wall Height	ft	Continuous	Uniform
Force Amplification (Super Structure)	NA	Continuous	Truncated Normal
Drift Shift (Super Structure)	NA	Continuous	Truncated Normal
Force Amplification (Cripple Wall)	NA	Continuous	Truncated Normal
Drift Shift (Cripple Wall)	NA	Continuous	Truncated Normal
Retrofit Length	ft	Continuous	Uniform

A total of 800 feature-response pairs are further split into 70% for training the ML models and 30% for testing. To maintain similar sample sizes for existing- and retrofitted-cripple wall buildings with different number of stories, the split manipulation is performed separately within each building group. The same training and testing sets are used across all six ML methods such that their model performance can be suitably compared.

As discussed earlier, cross-validation is implemented in the model training stage to reduce the risk of overfitting and determine the optimal set of model hyperparameters. This process can be waived for linear regression since it does not contain any hyperparameters and model penalization. K-fold cross-validation is one of the commonly adopted cross-validation techniques (Stone 1974), and K-fold with random hyperparameter search is used. K-fold cross-validation divides the training set into K equal-sized groups. For each iteration, a set of model hyperparameters is independently

sampled from a specified set of ranges and used to formulate the model. The model parameters are then trained using the $K-1$ data subsets and the validation score is computed using the remaining data. Taking the average validation score over the K validation cases as the final cross-validation score, the hyperparameter set with the highest score is selected. The negative median absolute error, as defined in Equation (3.34), is used to compute the cross-validation scores.

$$s = -\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.34}$$

3.3.4. Model Performance Evaluation

After obtaining the hyperparameter set with the best validation score, the testing dataset is fed into the model to check the performance on an "unseen" dataset. The $MARD$ and R^2 computed using Equations (3.35) and (3.36) are selected to evaluate the model performance. $MARD$ can provide a quantified measure on how the model over- or under-estimates the actual value on average. R^2 represents the proportion of response variance 'explained' by the model.

$$MARD = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3.35}$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \tag{3.36}$$

3.3.4.1. Median Collapse Intensity Predictive Model

Six models using the previously introduced algorithms are constructed to predict the median collapse intensities. Table 3.6 reports the training and testing $MARD$ and R^2 scores for each the model. As expected, ordinary linear regression has the poorest performance since it takes the simplest model structure. No substantial difference between ordinary linear regression and penalized linear regression (LASSO and ridge regression) is observed. Their testing $MARD$ and R^2 are approximately the same. This observation can be interpreted as the current features under

consideration are not complex enough to cause overfitting. Introducing penalization cannot boost the model predictive capability. Ensemble learning algorithms significantly enhances the model prediction accuracy relative to the linear regression algorithms. No overfitting issues are observed in all listed algorithms.

Table 3.6 Median collapse intensity ML model $MARD$ and R^2 scores

Model	$MARD$		R^2	
	Training Set	**Testing Set**	**Training Set**	**Testing Set**
Linear Regression	37.6%	37.5%	0.740	0.739
RSM	25.5%	22.2%	0.887	0.908
LASSO Regression	37.3%	37.3%	0.741	0.738
Ridge Regression	37.4%	37.4%	0.740	0.738
Random Forest	14.4%	14.7%	0.940	0.927
XGBoost	6.54%	6.95%	0.963	0.968

Figure 3.6 presents a box plot of the relative difference between the median collapse intensity predictions of the six ML models on the testing dataset. A box plot gives a simplified distribution representation of a target quantity. In the box plot, the green bar shows the position of median value. The box in the middle covers the 25% to 75% quantile of the population, which is denoted as interquartile range (IQR). The black bars on the top and bottom refer to the 75% quantile + 1.5IQR and 25% quantile – 1.5IQR values of the population, respectively. The range defined by the two black bars approximately covers 99.7% of the data in any normal distribution. From Figure 3.6, all models have close to zero relative errors, while the distributions of individual data points vary significantly. For linear, LASSO and ridge regression, the median collapse intensities can be overestimated by approximately 90% and underestimated by 75%. The error distribution is long-tailed and slightly left-skewed, which illustrates that the model is biased towards under-estimating

the actual median collapse intensities. Also, the relative errors of these three algorithms are highly divergent. As for RSM, which accounts for higher order effects, the maximum and minimum relative error reduces to around 50% level. By comparing the testing R^2 scores of linear regression and RSM in Table 3.6, the proportion of the variance accounted by the model increases from 0.74 to 0.91. The comparison shows that second order feature interaction effects have unignorable impacts on the median collapse intensities. The relative error distributions of Random Forest and XGBoost are remarkably narrower than the linear models'. This is especially true for the XGBoost model, where the relative errors of more than 80% of the testing predictions are less than 10%.

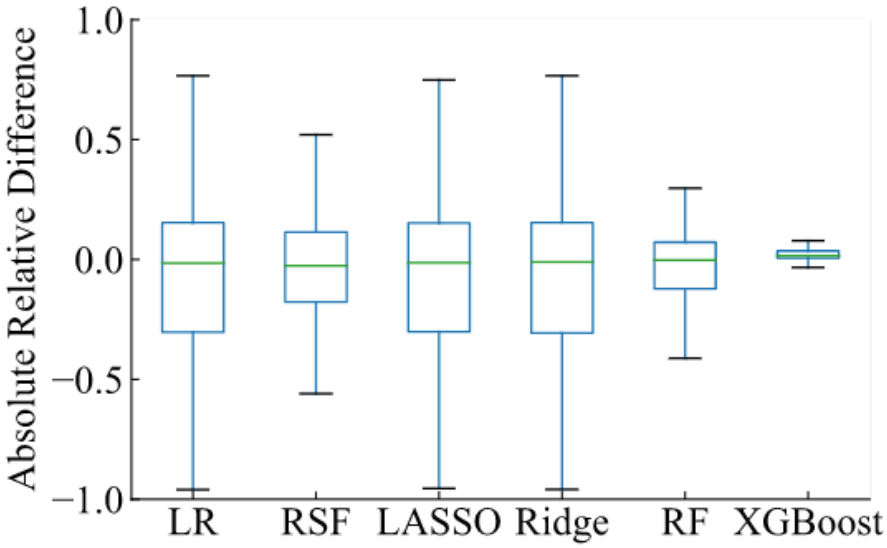


Figure 3.6 Box plot of relative difference for median collapse intensities

Though the linear regression model has the lowest accuracy, the results useful for checking the model assumptions and for data outlier detection.

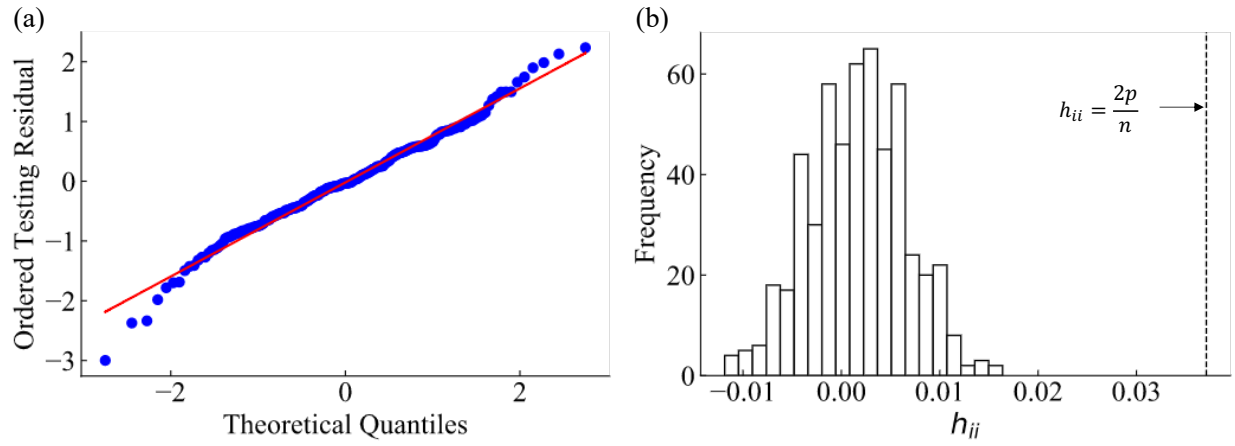


Figure 3.7 shows the quantile-quantile (Q-Q) plot and histogram of the diagonal elements of the hat matrix. As introduced in section 3.2, the linear model has a fundamental normality assumption, where the unavoidable error term ϵ_i is independent and identically follows a normal distribution $N(0, \sigma^2)$ with a constant variance σ^2. The Q-Q plot helps to inform whether the prediction residuals follow normal distribution. In the Q-Q plot, all residual values are sorted in ascending order, and the corresponding percentiles are plotted against the theoretical quantile of a normal distribution. By visually inspecting whether the plotted data distribute on the 45-degree straight line (strict normal distribution), the normality assumption can be examined. In Figure 3.7(a), most of the residuals lay around the theoretical red line, except for the two tails. The residuals are distributed closely to the normal distribution, but with slightly heavier tails at both ends. The normality assumption holds in this case. Figure 3.7(b) shows the histogram of h_{ii}, where h_{ii} is the i^{th} diagonal element of hat matrix from Equation (3.13). Greater than $\frac{2p}{n}$ h_{ii} indicates that the i^{th} data point 'drags' the regression line towards to itself, which can be identified as an outlier. By comparing $\frac{2p}{n}$ with the histogram, there is no h_{ii} that exceeds the $\frac{2p}{n}$ limit. It can also be concluded

that, the generated dataset is consistent with the process discussed in section 3.3.3.1 satisfies the requirements for deriving confident ML models.

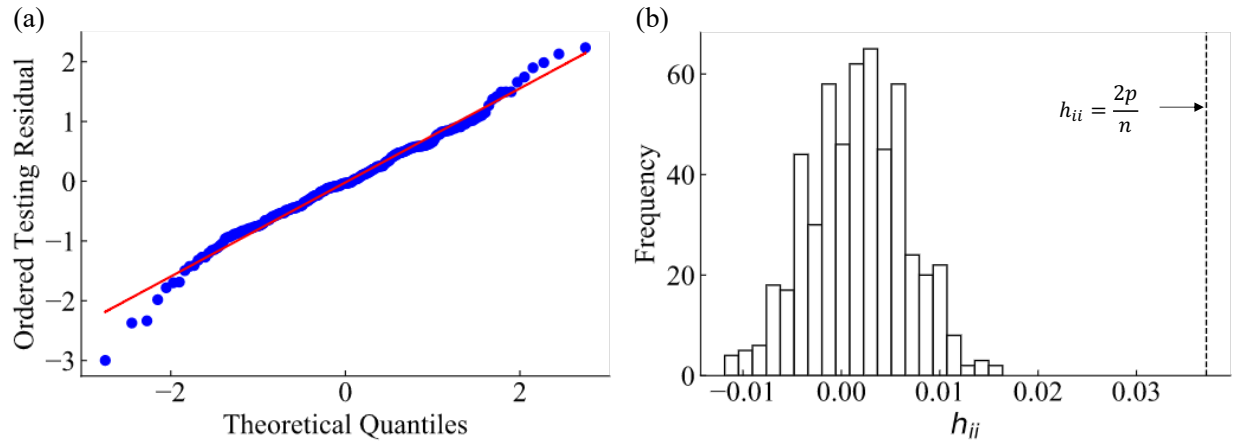


Figure 3.7 Linear regression diagnostics: (a) residual Q-Q plot and (b) h_{ii} histogram

Some additional model performance diagnostics are presented in this section. Figure 3.8 shows the predicted and ground truth median collapse intensities of the testing dataset. Closer to 45-degree straight line data points suggest more accurate predictions. The plots disaggregate the predictions into 1- and 2-story, existing- and retrofitted-buildings to examine how models perform under different scenarios. Similar patterns can be observed for linear, LASSO and ridge regression (Figure 3.8 (a), (c) and (d)). For the buildings with lower than 4g median collapse intensities, though predictions lay on both sides of the straight line, the errors are highly divergent. Model performances are biased towards specific building types. They tend to overestimate (above the straight line) most of 1-story existing and retrofitted and 2-story retrofitted buildings, especially for the buildings with lower than 2g median collapse intensities. For greater than 4g cases, the models constantly under-estimate ground truth values. Most of these cases are 1-story retrofitted cripple wall buildings. Higher level features, e.g. building peak strength and ductility, can be introduced to boost model performance. As observed in Figure 3.8 (b), RSM stabilizes the predictions within a band. The under-estimation issue for strong buildings is resolved, which

indicates that the second order feature interactions have higher impacts on stronger buildings than weaker buildings. A similar over-predicting issue can be found for buildings with greater than 5g median collapse intensities in the random forest model (Figure 3.8 (e)). Also, the random forest predictions for the existing building median collapse intensities are better than those of retrofitted buildings. Lastly, the XGBoost model delivers the most robust median collapse intensity predictive capability among all models. Most of the data points distribute closely to the 45-degree line. The model is unbiased across all building types, except for a few cases with greater than 6g median collapse intensities.

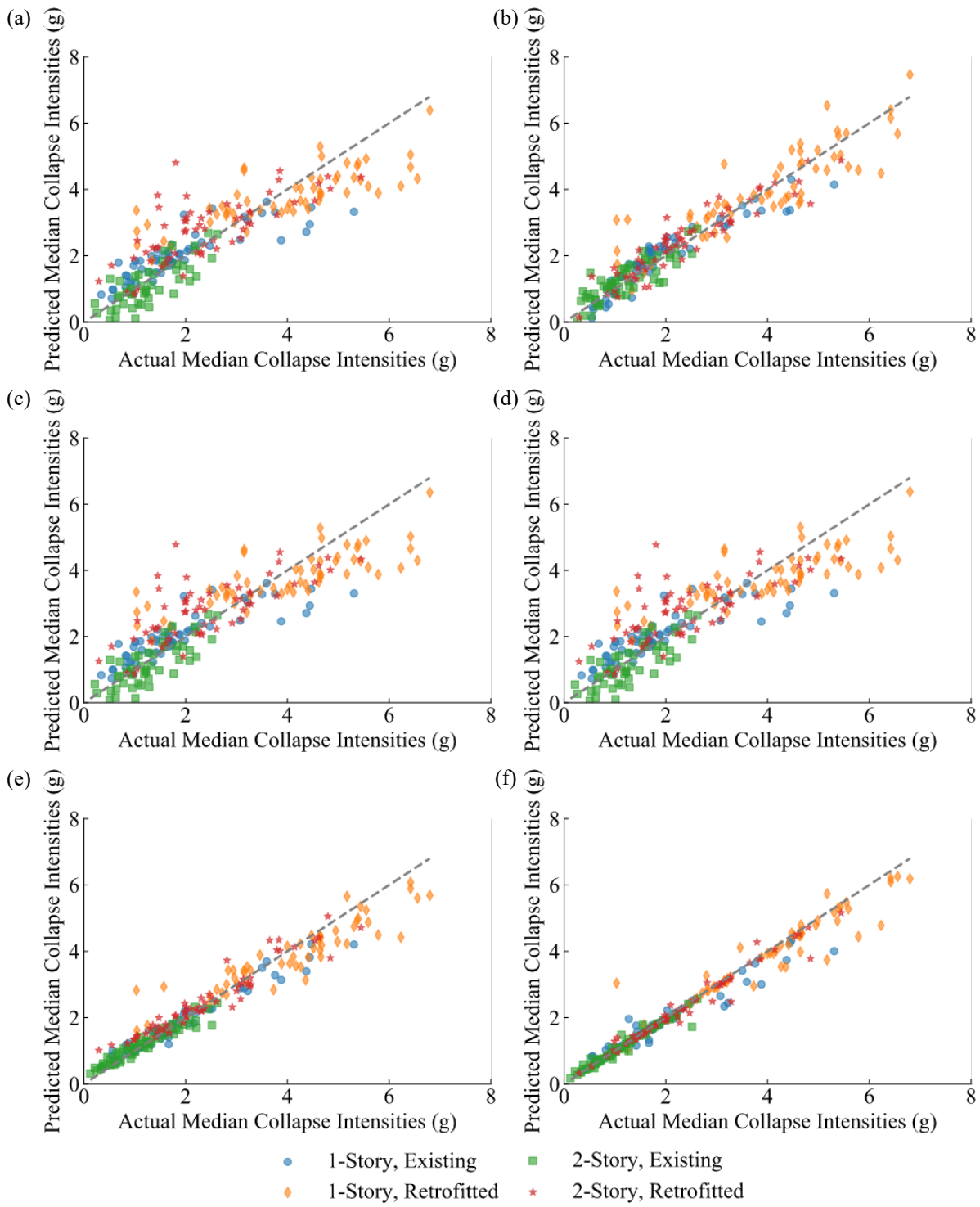


Figure 3.8 Testing set median collapse intensity prediction vs. ground truth value: (a) linear regression, (b) RSM, (c) LASSO regression, (d) ridge regression, (e) random forest and (f) XGBoost

3.3.4.2. EAL Prediction Model

Similarly, six regression models are established for the EAL predictions. Table 3.7 summarizes the $MARD$ and R^2 scores for each of the predictive models. Since the magnitudes of EAL usually follow an order of 10^{-4} to 10^{-3}, the response herein is transferred to logarithm scale. The $MARD$ and R^2 are calculated using the exponential of the model outcomes and original EALs. Compared to the median collapse intensity prediction models, the EAL models do not perform as well. Linear and ridge regression models have around 0.1 R^2 values on the testing set, which indicates the linear combinations of the features under consideration cannot explain building EAL well. Though LASSO has a slightly higher testing R^2 score of 0.320, its $MARD$ implies it over- or under-estimates EAL by approximately 75% on average. RSM has a significantly superior performance than other linear models. The second order feature interaction effects help to reduce the testing $MARD$ from 63.3% in the linear regression to 22.4%. The two ensemble tree models random forest and XGBoost show more robust capabilities in the EAL prediction as expected. XGBoost has the best performance on $MARD$ and R^2. Besides, its testing and training scores are close, which indicates that overfitting is not a concern.

Table 3.7 EAL ML model $MARD$ and R^2 scores

Model	***MARD***		***R^2***	
	Training Set	**Testing Set**	**Training Set**	**Testing Set**
Linear Regression	60.4%	63.3%	0.231	0.063
RSM	22.6%	22.4%	0.842	0.447
LASSO Regression	75.8%	75.7%	0.468	0.320
Ridge Regression	60.7%	63.7%	0.276	0.101
Random Forest	18.4%	17.8%	0.904	0.890
XGBoost	14.3%	13.2%	0.955	0.964

The box plot of the relative errors on the testing set shown in Figure 3.9 provides additional information for model selection. The relative errors of the linear, Lasso and ridge regression are

left-skewed. For the worst case, they exaggeratedly over-predict the EAL by more than 150%. Considering the EAL computation for single family cripple wall buildings, engineering demand parameters (EDP) and median collapse intensities contribute to component losses and collapse losses, respectively. The poor performances of the linear models indicate that the EDPs are less explainable than the median collapse intensities using the linear combination of the features under consideration. RSM significantly reduces the relative error ranges of other linear models, and random forest has a slightly narrower error distribution than RSM. XGBoost model gains one point in the relative error distribution compared with other models. It successfully limits the relative error on testing set within $\pm 20\%$ range.

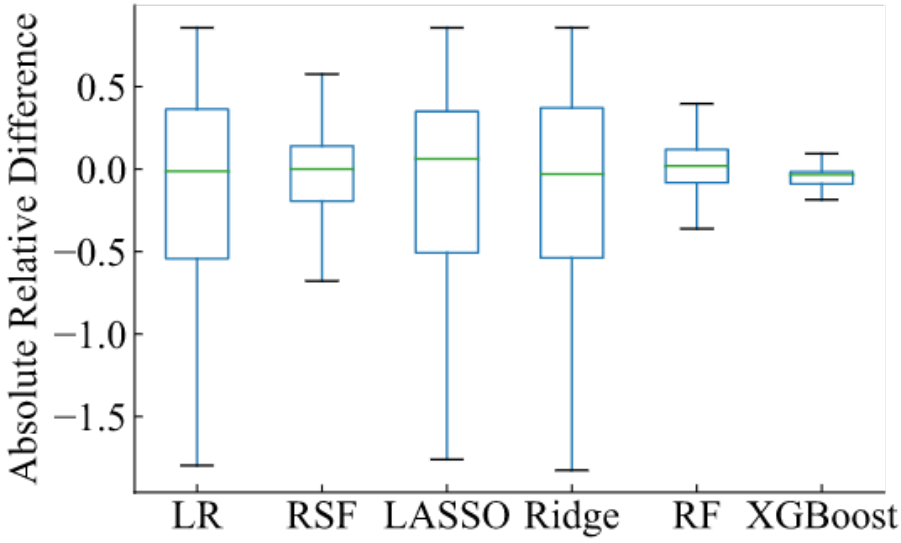


Figure 3.9 Box plot of absolute relative difference for EAL

The EAL predictions against ground truth EAL values of the six models are shown in Figure 3.10. Overall, for all six models, the deviations from ground truth values are found to increase with increased EAL values. By comparing Figure 3.10 with Figure 3.8, though linear, LASSO and ridge regression have acceptable performance on the weak buildings' median collapse intensity predictions, the EAL performance is not nearly as good. The observation further indicates that the

linear combination of the features cannot explain the component loss well. Besides, the linear models are biased towards the 1-story retrofitted buildings, for which they provide lower than ground truth predictions. As for the 1-story existing and 2-story retrofitted buildings, these three algorithms give an approximate constant prediction for the cases where the EAL is greater than 1%. Compared to these linear models, RSM reduces the differences between the predictions and ground truth values, also it eliminates the biases in 1-story retrofitted buildings. However, the predictions are getting worse with increased EAL. As for random forest, the model provides better predictions overall, but it is biased on the testing set for buildings with greater than 1% EAL. It consistently under-estimates the EALs by approximately 50%. Lastly, the XGBoost model shows the most robust predictions for EALs. There are no biases that can be observed for specific building types. Additionally, it provides the best prediction for the largest EAL case (the greater than 4% case) among all models.

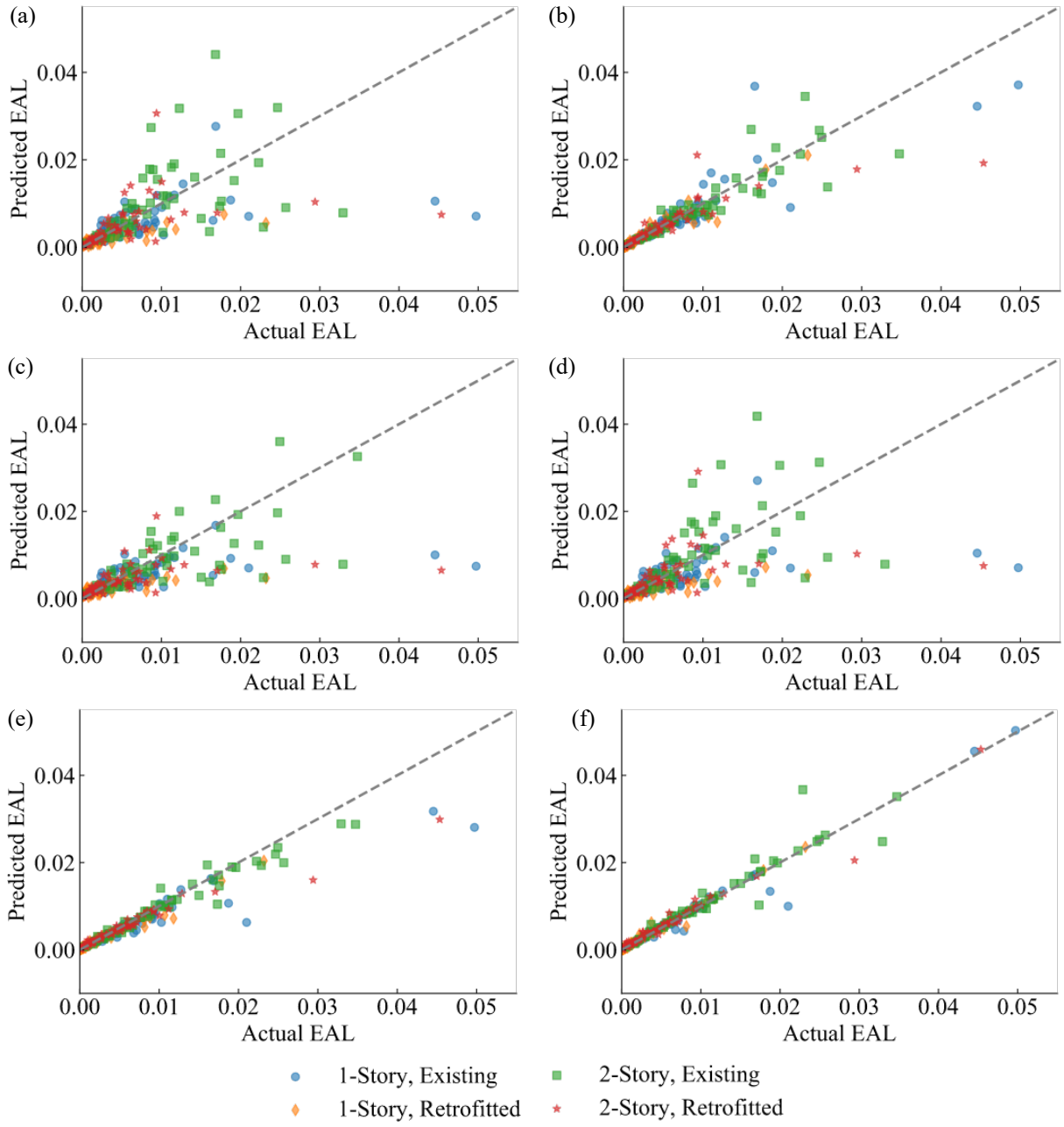


Figure 3.10 The testing set EAL prediction vs. ground truth value: (a) linear regression, (b) RSM, (c) LASSO regression, (d) ridge regression, (e) random forest and (f) XGBoost

3.4. Summary

To support sensitivity analysis and model uncertainty quantification involving large numbers of numerical analyses, ML based methods are investigated in this chapter to create a statistical link between building features and seismic responses. Several classical regression models including

linear regression, RSM, LASSO regression, ridge regression, Random Forest and XGBoost are investigated to predict the median collapse intensity and EAL. A stucco cripple wall single-family house with stucco as the exterior superstructure wall material, gypsum wallboard as interior superstructure wall material, is used to demonstrate the model development.

For the median collapse intensity model, all investigated ML algorithms provide some predictive capability. XGBoost, which has 7% MARD and 0.97 R^2 score on the testing set, performs the best among all models. Random Forest is the next best model. RSM provides more accurate predictions than the ordinary linear regression model, indicating that the second order feature interaction creates unignorable impacts on the median collapse intensities. Though linear regression gives the least preferable performance, its underlying statistical model is useful for hypothesis testing. The analysis of the residual distribution and outlier detection illustrates the LHS generated dataset is representative and no outliers are present in the dataset. As for the EAL model, all linear regression-based models are not capable of giving stable predictions. Biased and diverged predictions are observed for these models under higher magnitudes of EAL. Still, XGBoost model performs the best, with only 13.2% MARD and 0.96 R^2 score on the testing set. Boosted tree-based models have revealed more powerful prediction capability given the limited number of inputs compared to the linear regression models.

The model development and verification procedure presented in this section is generally applicable to other structural engineering regression tasks. A stable and reliable ML model serves as a solid foundation for further investigations in terms of sensitivity analysis, uncertainty quantification and optimization. In Chapter 4, the developed ordinary linear regression and XGBoost models are adopted to analyze the impacts of the building parameters on collapse risk and losses. In Chapter

6, a more comprehensive seismic retrofit optimization framework is established on the basis of the ML surrogate models.

CHAPTER 4. APPLICATIONOF THE ANALYTICS DRIVEN MODEL

4.1. Background and Introduction

4.1.1. Literature Review

Prior research on sensitivity analyses on the woodframe building seismic performance and earthquake induced losses are introduced in this paragraph. Yin and Li (2010) conducted a comprehensive sensitivity analysis on the effect of woodframe building material properties impacts on building collapse risk. The SAWS material hysteretic parameters were treated as random variables, and the collapse performance of a one-story single-family house as assessed using Latin Hypercube sampled hysteretic parameters. The material initial stiffness, peak strength, intercept strength and material degradation were found to have the largest effect on building collapse performance. Jayamon et al. (2019) conducted a similar sensitivity analysis for the SAWS hysteretic parameters. The implications to building collapse safety of P-delta effect and damping ratios were also examined. A total of 126 unique woodframe building cases were developed on the basis of the FEMA P-695 example. The intercept strength, displacement at peak strength and post-peak stiffness were found to significantly impact the building collapse margin ratio. P-delta effects reduced the collapse margin ratios by about 10%, and 5% damping ratio buildings had on average 13% higher collapse resistance than the 1% damping cases. Pei and van de Lindt (2010) performed a sensitivity analysis for seismic-induced losses. A two-story residential woodframe building was modeled in SAPWood, and three common nailing patterns with six levels of construction quality were used to vary the structural panel properties. 20 ground motions were selected for an event-based loss estimation, and long-term loss assessments were also performed for three sites assuming 5-, 30- and 75-year ownership. For the event-based assessment, the expected losses were not significantly affected by the structural properties for spectral acceleration levels below 0.3g and

greater than 2.6g. The effects were more observable under intermediate level shaking intensities. The high seismicity region was observed to have greater than collapse loss median long-term loss, while the low seismicity regions had essentially zero loss. The reduction in long-term loss obtained by altering the nailing pattern was limited, but the construction quality had a significant influence. Within the performance-based seismic design (PBSD) framework, the performance is assessed using a probabilistic-based procedure while considering collapse safety, demolition, the non-collapse/non-demolition EDPs and seismic hazard. Uncertainties rise at all stage of performance assessment and eventually affect the loss estimation. The sources of uncertainties can be placed into three categories: resistance/strength, load/demand, and the analysis method. Resistance or strength uncertainties originates from variations in the material properties and construction quality. The uncertainty associated with variations in the construction quality cannot be avoided but may be reduced through strict construction supervision. The associated uncertainty is very difficult to consider within the PBSD framework and is not address here. The uncertainty in material properties influence the loss assessment through the EDPs generated from nonlinear response history analysis. Pei and van de Lindt (2010) noted that the initial stiffness and ultimate strength will have a significant impact on the demands generated from nonlinear response history analysis (NRHA). Pei and van de Lindt (2010) performed a loss sensitivity analysis on a two-story residential woodframe building incorporating material properties, low construction qualities and the seismic hazard level. Yin and Li (2010) performed an uncertainty quantification study on the SAWS model parameters using Monte Carlo simulation and Latin Hypercube sampling. They sampled model parameters from a truncated normal distribution and quantified the impact on the collapse capacity of a single wood shear wall. Gokkaya (2015) performed a comprehensive uncertainty quantification for concrete structures. Monte Carlo simulation with Latin Hypercube

sampling was implemented to investigate the performance sensitivity to the concrete frame backbone curve parameters. Neural Networks and the Response Surface method were explored to propagate uncertainties. Seismic hazard uncertainty is another generally considered source of uncertainty in the performance-based assessments. Seismic hazard uncertainty affects the loss assessment through NRHAs and probabilistic seismic hazard analysis (PSHA). In the structural nonlinear response history analysis, ground motion record-to-record uncertainty are known to have a significant impact on the response demands. Incremental dynamic analysis (Vamvatsikos and Cornell 2002) was proposed to address the record-to-record uncertainty. A set of ground motions are selected and scaled to match the spectral acceleration at the fundamental period of the structure, and NRHA is performed using this set of ground motions. The record-to-record uncertainty can be illustrated using a series of IDA curves describing the engineering demand parameters (EDP) at each intensity level from different ground motion records.

4.1.2. Objective and Organization

Sensitivity analyses and uncertainty quantification are performed for the single-family woodframe residences with cripple walls. The number of stories, seismic weight, construction era and retrofit, are considered. The 2^k factorial experimental design technique is applied when designing required numerical analyses. Two-way ANOVA is presented to statistically quantify the features' and feature interactions' impacts on building collapse risk and expected annual losses (EAL). Then, the ordinary least square and XGBoost models developed in Chapter 3 are used for more comprehensive sensitivity study. Though the 2^k experiment is able to provide insights into the parameter influences, the relative importance of the different features cannot be accessed. Machine learning (ML) models could deliver such information from the perspective of features' contributions to the predictions. Lastly, the effect of model uncertainty building collapse

performance and EAL are quantified using the XGBoost model. The latter two analyses demonstrate the capability of ML methods beyond performing predictions.

4.2. Statistical Analysis Sensitivity Study

4.2.1. 2^k Full Factorial Experiment Design

The 2^k full factorial refers to the experimental design with k different factors and each factor is at two levels (denoted as + and -). A total of 2^k analyses/experiments covers all possible factor level combinations. Each factor level appears the same number of times (2^{k-1}) in the entire run, and each possible factor level pair also appears the same number of times (2^{k-2}). Thus, 2^k full factorial experiment is a balanced and orthogonal experiment design method. The 2^k full factorial experiment can be used for evaluating the effect of individual factors and factor interactions on the responses. The main effect (ME) of variant A on the response y could be quantified simply using Equation (4.1).

$$ME(A) = \bar{y}(A+) - \bar{y}(A-) \tag{4.1}$$

Where $\bar{y}(A+)$ is the average of the response with feature A set to its higher level and so on. Intuitively, it measures the difference between the average response with factor A set to its higher and lower level. The higher the main effect, the strong the impact of the feature. Similarly, the conditional main effect of factor B given factor A at the upper level is defined as shown in Equation (4.2).

$$ME(B|A+) = \bar{y}(B+|A+) - \bar{y}(B-|A+) \tag{4.2}$$

With the conditional main effect, the second order factor interaction effects can be derived as shown in Equation (4.3)

$$INI(A,B) = \frac{1}{2}\{ME(B|A+) - ME(B|A-)\} \tag{4.3}$$

Through the procedure above, the average influence of the features and feature interactions can be quantified. Statistical significance of the factors can be accessed through two-way analysis of variance (ANOVA). ANOVA is a frequentist methodology that is used to test whether the impact of a feature is statistically significant. ANOVA decomposes variance into the regression error and estimation residuals as shown in Table 4.1. The default null hypothesis H_0 is that the coefficient of the i^{th} factor β_i is zero. The p-value is the probability of obtaining an observation that is at least as extreme as violating null hypothesis given null hypothesis is true. A lower than pre-specified significance level (α) p-value indicates that the factor is statistically significant to the responses. To be noted, the magnitude of the p-value cannot infer the relative importance of the factors. Two-way ANOVA refers to the ANOVA test for both categorical and continuous variables.

Table 4.1 ANOVA table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratios
Model	$SSR = \sum_{i=1}^{n} \left(\hat{Y}_i - \bar{Y}_i\right)^2$	$p-1$	$MSR = \frac{SSR}{p-1}$	$F = \frac{MSR}{MSE} \sim F(p-1, n-p)$
Residual	$SSE = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$	$n-p$	$MSE = \frac{SSE}{n-p}$	
Total	$SST = SSR + SSE$	$n-1$		

4.2.2. Building Variants

The 2^k full factorial experiment is used to design the required numerical analyses. For each factor or property, two levels are specified. The details of the considered factors and their variations is discussed in this section. Four baseline archetypes developed as part of the PEER-CEA Project are used here for the sensitivity study. The naming conventions are listed in Table 4.2. For each building, one- and two-story cases with un-retrofitted and retrofitted 2ft and 6ft cripple wall are modeled. Since only ground motions for San Francisco $V_{s,30} = 270m/s$ site were adopted for the multiple stripe analysis (MSA), the retrofit design based on high seismicity ($S_{DS} = 1.2g$) is applied.

Heavy (concrete tile roof) and light (shingle or composition roof) seismic weights are considered. The details of the considered loading is provided in Table 2.7 and Table 2.8. The construction era is implicitly considered through the interior wall material. For pre-1945 construction, plaster on wood lath is adopted for the interior wall. For post-1955 cases, gypsum wallboard is used. From 1945 to 1955 is a transition period, where both materials were used for the interior wall. Therefore, the effect of construction era can be assessed by comparing the performance of gypsum wallboard and plaster on wood lath buildings. The variations in material properties are not considered herein, and the best estimates for all materials are used in modeling. A total of 64 cases were modeled and analyzed. For each model, pushover analyses, MSAs and loss assessments are performed as introduced in Chapter 2 to obtain building nonlinear static response properties, the median collapse intensities (MCI) and the expected annual losses (EAL). Additional information such as the modal periods, peak strength, drift capacities, ductility, and engineering demand parameters (EDPs) are also recorded.

Table 4.2 The 2^k full factorial experiment building variants index (table adapted from section 7.2 of Welch and Deierlein 2020)

Baseline	Description	Variation
$nRW - C1 - LP - mC - HS2 - EXRET$	Horizontal wood siding exterior with plaster on wood lath interior (pre-1945 era), best estimate horizontal wood siding as cripple wall material	n: number of stories, RW: roof weight, H for heavy weight, L for light weight m: cripple wall height, $2C$ for 2ft cripple wall, $6C$ for 6ft cripple wall $EXRET$: cripple wall design, EX for un-retrofitted cripple wall, $SDS12$ for $S_{DS} = 1.2g$ retrofitted cripple wall
$nRW - S2 - G2 - mC - S2 - EXRET$	Stucco exterior with gypsum wallboard interior (post-1955 era), best estimate stucco as cripple wall material	
$nRW - SLP2 - LP - mC - S2 - EXRET$	Stucco exterior with plaster on wood lath interior (pre-1945 era), best estimate stucco as cripple wall material	
$nRW - W2 - G2 - mC - HS2 - EXRET$	Horizontal wood siding exterior with gypsum wallboard interior (post-1955 era), best estimate horizontal wood siding as cripple wall material	

4.2.3. Results Analysis

The MSA and loss results together with detailed statistical analyses are presented and discussed in this section. Figures Figure 4.1 through Figure 4.4 present the median collapse intensities (MCI) and EAL for all analyzed 64 cases. The cubic representation of the 2^3 experiments can demonstrate the effect of a single factor. In this representation, each vertex represents a factor level combination. For example, the case on the lower left vertex refers to 1 story, 2 ft cripple wall, light weight roof and un-retrofitted case, where all variants are set to their lower levels. The two cases connected by one edge differ from each other by a single factor level. By comparing the cases on the left surface with those on the right surface, the number of stories is found to negatively contribute to collapse risk and EAL (i.e. 2-story buildings have a higher collapse risk than 1-story cases). Considering the C1-LP un-retrofitted buildings (Figure 4.1 (a)), the two-story buildings on average have 30% lower MCI and 60% higher EAL. The impact is consistent across all four existing and retrofitted archetypes. The effect of the number of stories on the overall performance is complex. 2-story buildings have a larger seismic weight and longer period than their 1-story counterparts and the P-Δ effect is more significant. Seismic weight is observed to have a similar but less significant effect as the number of stories when the cases on the bottom surface are compared with those on the top. The cripple wall height is observed to have an overall positive influence on the building seismic performance (i.e. buildings with tall cripple walls perform better). Based on experimental testing data (Zareian and Lanning 2020), 6ft tall cripple wall was found to be about 20% stronger and have almost twice displacement capacity than 2ft tall cripple wall. Additionally, since the material properties were normalized by panel height and length, cripple wall level displacement capacity increases proportionally to panel height. These factors contribute to a stiffer cripple wall. Under such scenarios, the weak-story mechanism is reduced and the overall

collapse resistance improves. Interestingly, the impact of cripple wall height is not consistent between un-retrofitted and retrofitted buildings. For multiple retrofitted buildings, the cases with taller cripple walls have lower or the same MCI (e.g. 2L-S2-G2-2C-S2-SDS12 and 2L-S2-G2-6C-S2-SDS12 in Figure 4.2 (b)). One possible explanation is that the increased cripple wall strength obtained from the retrofit and taller cripple wall reaches a critical point such that the adjacent story is weaker, which shifts the collapse mechanism to upper story. This effect has been highlighted and discussed in multiple prior soft-story building studies (Buckalew et al. 2015; Yi et al. 2020). The retrofit impacts can be assessed by comparing the results in (a) and (b) in Figures Figure 4.1Figure 4.4. With strengthened structural wood panels, significant improvement in MCI and reduction in EAL is observed for all four archetypes.

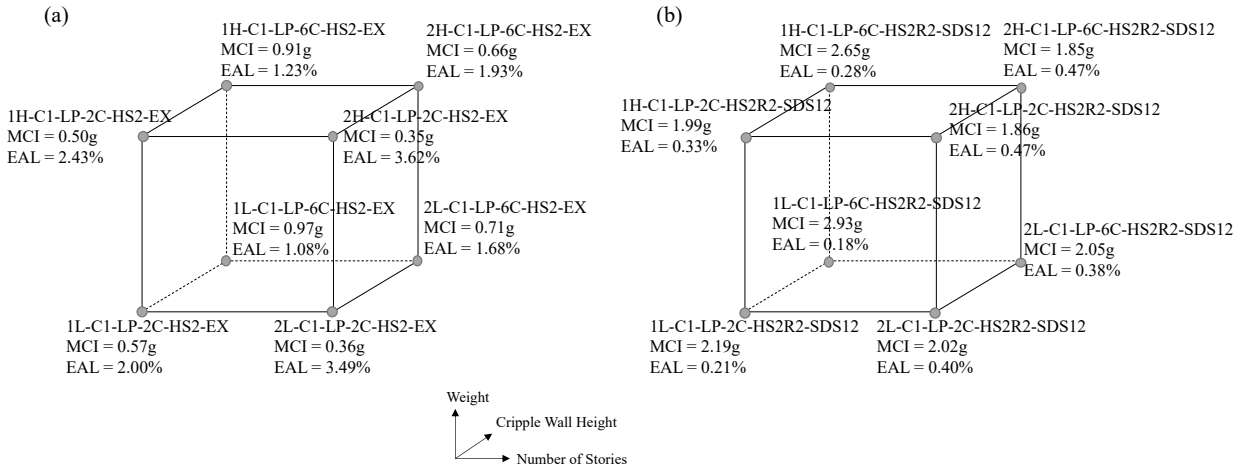


Figure 4.1 MCI and EAL for horizontal wood siding with plaster on wood lath buildings: (a) un-retrofitted cripple wall cases, (b) retrofitted cripple wall cases

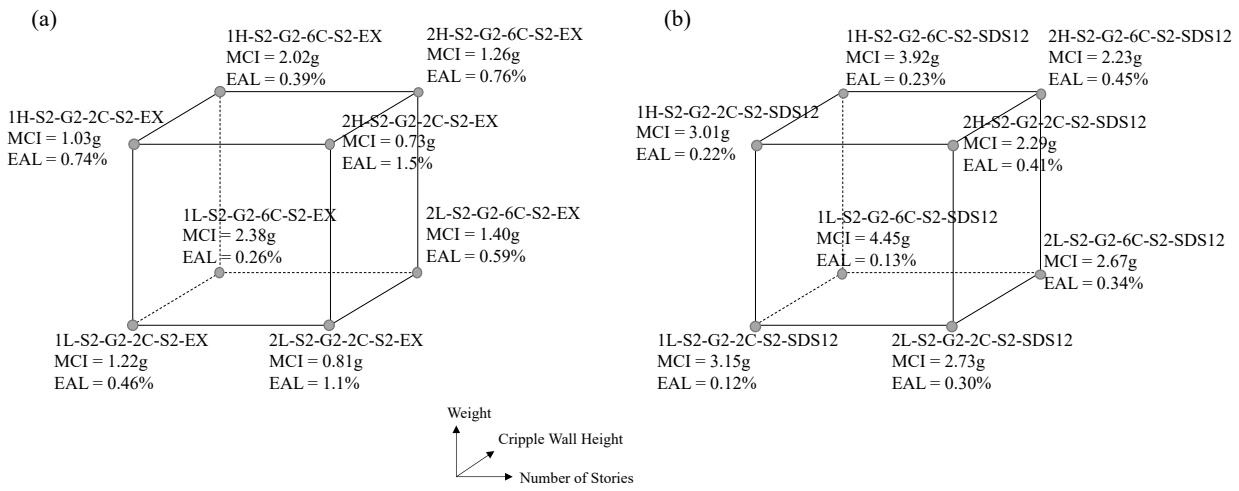


Figure 4.2 MCI and EAL for stucco with gypsum wallboard buildings: (a) un-retrofitted cripple wall cases, (b) retrofitted cripple wall cases

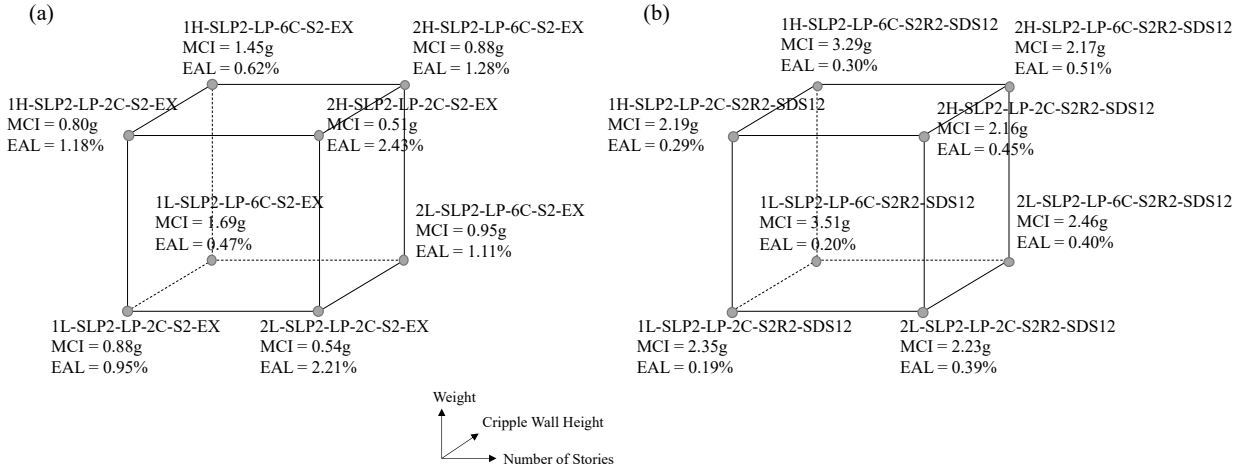


Figure 4.3 MCI and EAL for stucco with plaster on wood lath buildings: (a) un-retrofitted cripple wall cases, (b) retrofitted cripple wall cases

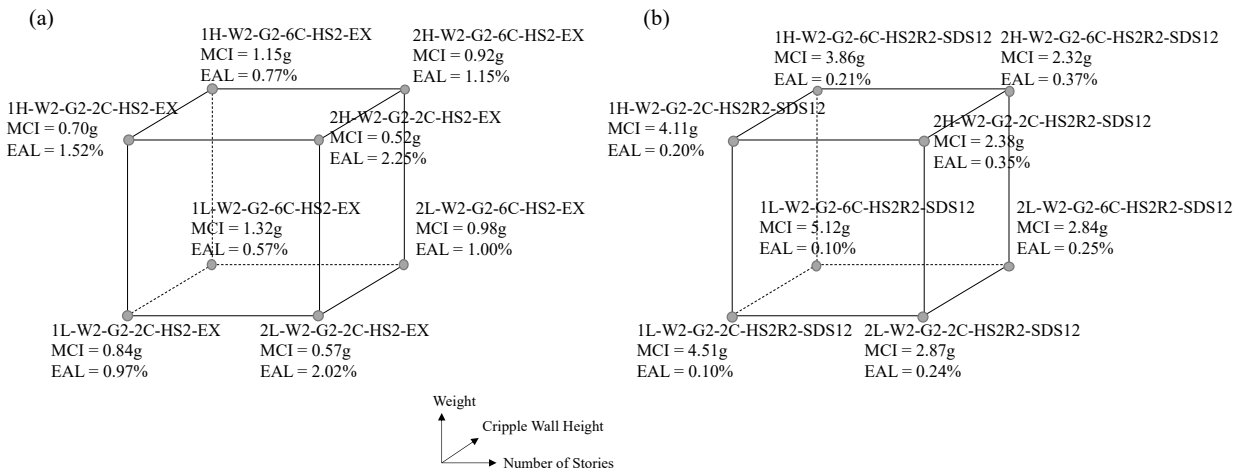


Figure 4.4 MCI and EAL for horizontal wood siding with gypsum wallboard buildings: (a) un-retrofitted cripple wall cases, (b) retrofitted cripple wall cases

Although the preceding figures enable comparison of the buildings' seismic performances, the results are not always easily interpretable. Also, the effect of the construction year is difficult to isolate. The box plots of the MCI and EAL in Figure 4.5 and Figure 4.6 provides a clearer illustration of the main effect of the factors of interest. For the buildings where horizontal wood siding or stucco is used as the exterior wall panel, 32 cases are aggregated to compare the average seismic performance. The relative main effect of different factors can also be compared. MCI has an overall negative correlation with EAL. Similar to the findings in the cubic representation of the 2^3 experiment, the number of stories and seismic weight are negatively correlated with MCI, while the cripple wall height and retrofit are positively correlated. Unsurprisingly, the cripple wall retrofit is found to provide the most significant improvement in the building seismic performances followed by the cripple wall height. The construction era is also negatively correlated with seismic performance. Post-1945 buildings with gypsum wallboard as the interior panels have on average higher collapse resistance and lower EAL than pre-1955 construction with plaster on wood lath on the interior. Although collapse is expected to occur at the cripple wall level, higher available strength in the superstructure still plays an important role in the collapse resistance. The difference

between construction era are more significant for horizontal wood siding buildings than stucco buildings.

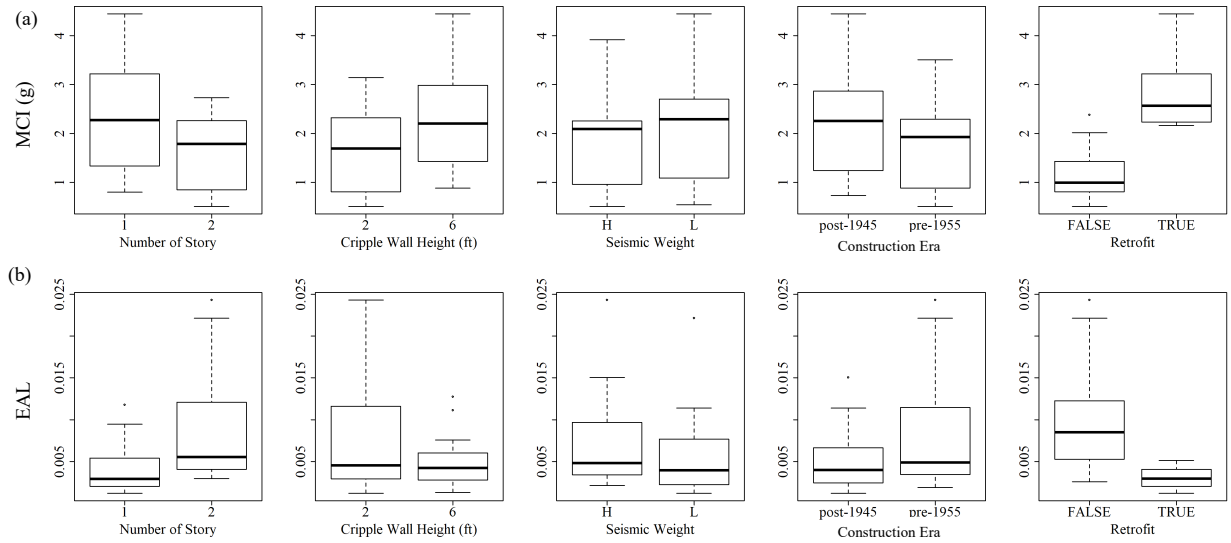


Figure 4.5 Box plot of responses for stucco as exterior buildings: (a) MCI, (b) EAL

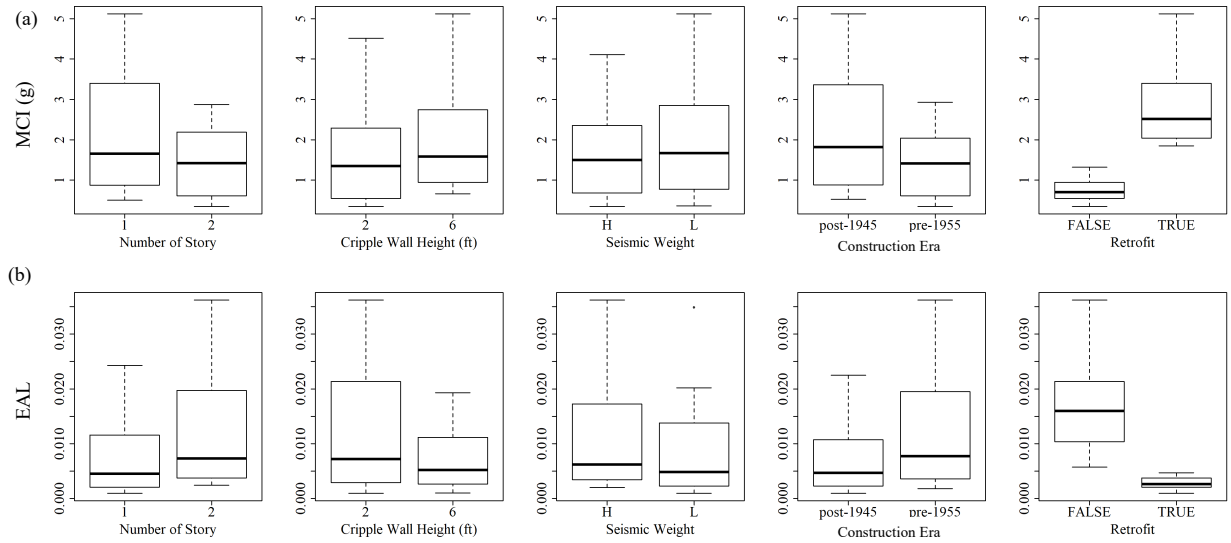


Figure 4.6 Box plot of responses for horizontal wood siding as exterior buildings: (a) MCI, (b) EAL

Figures Figure 4.7 through Figure 4.10 show feature interaction plots for MCI and EAL for stucco and horizontal wood siding buildings, respectively. The '*A*-against-*B*' interaction plot

demonstrates the change in the response caused by the joint effect of A and B. If the two lines in the interaction plot are parallel, the effect of A is synergistic with changes in factor B. If the slopes of the two lines are of opposite sign (plot is antagonistic), the effect of factor A is reversed when B is set to different levels. Overall, almost all feature interactions are synergistic. The effect of the number of stories on MCI and EAL is unform when the seismic weight is changed for both stucco and horizontal wood siding buildings. Thus, their interaction is not significant on MCI and EAL. Similar observations can be observed for the interaction between seismic weight and cripple wall height and construction era. The effect of the number of stories is not uniform for 2ft and 6ft cripple wall buildings. Increased number of stories leads to higher reduction in the MCI for 6ft cripple wall buildings compared to the 2ft case. It worth noting that cripple wall height and retrofit have antagonistic effects on EAL (Figure 4.8 and Figure 4.10). More specifically, taller cripple walls lead to reduced EAL on un-retrofitted buildings, while for retrofitted buildings, the EAL remains almost the same or increases slightly. The interaction plots provide insight into the biased MCI and EAL predictions by the linear regression model developed in section 3.3.4, since this second order interaction effect was not included.

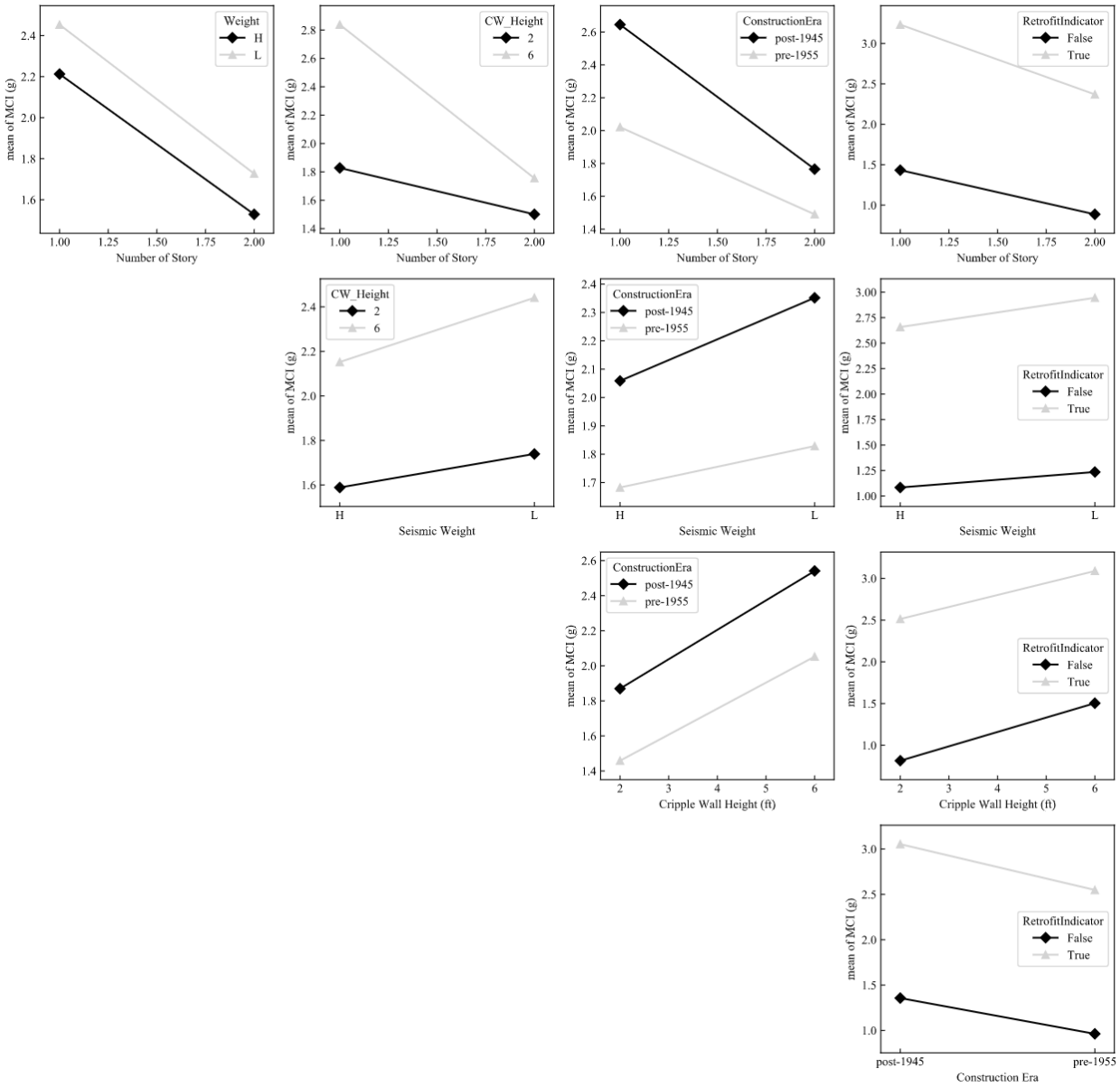


Figure 4.7 MCI factor interaction plot for stucco exterior buildings

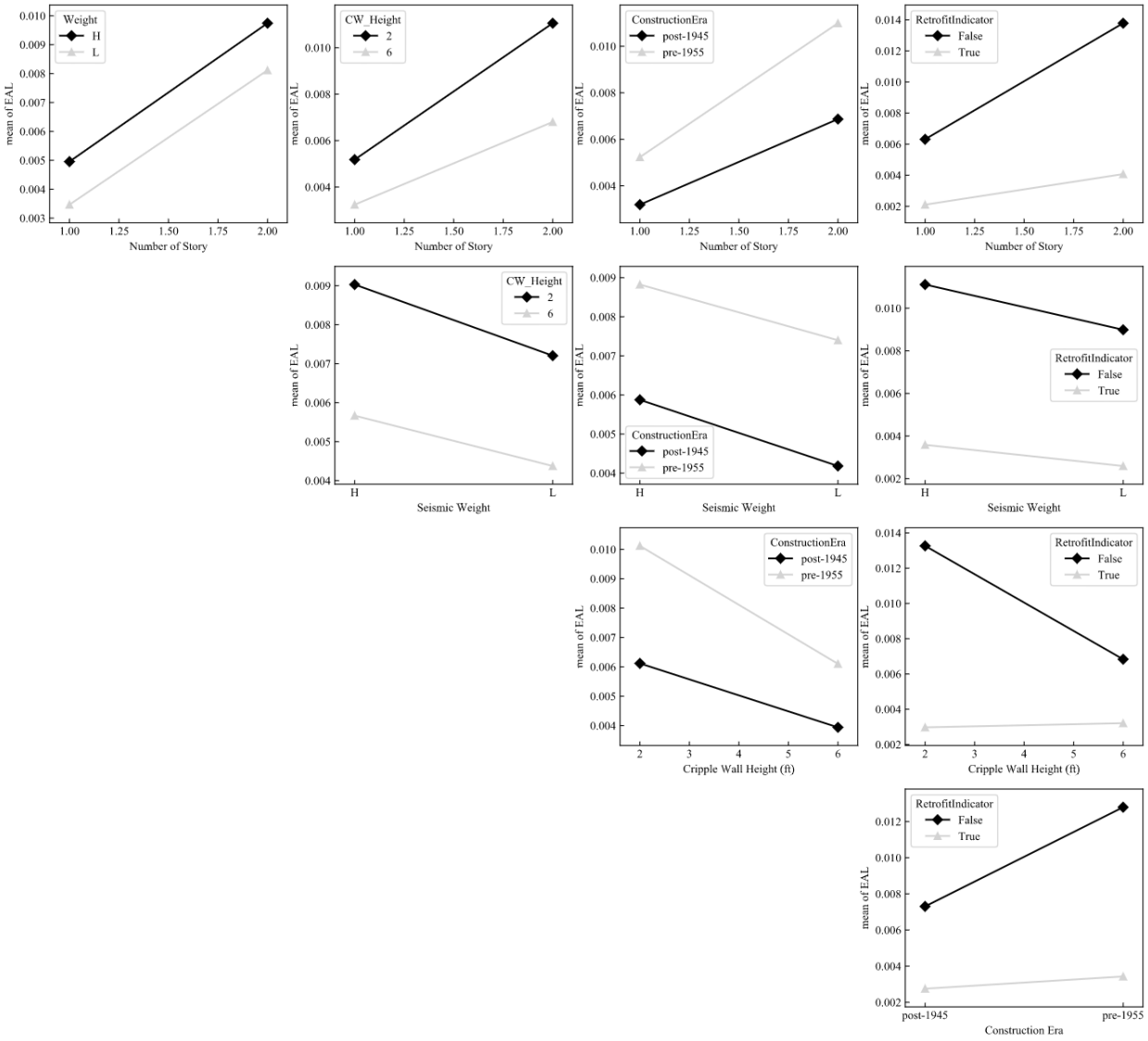


Figure 4.8 EAL factor interaction plot for stucco exterior buildings

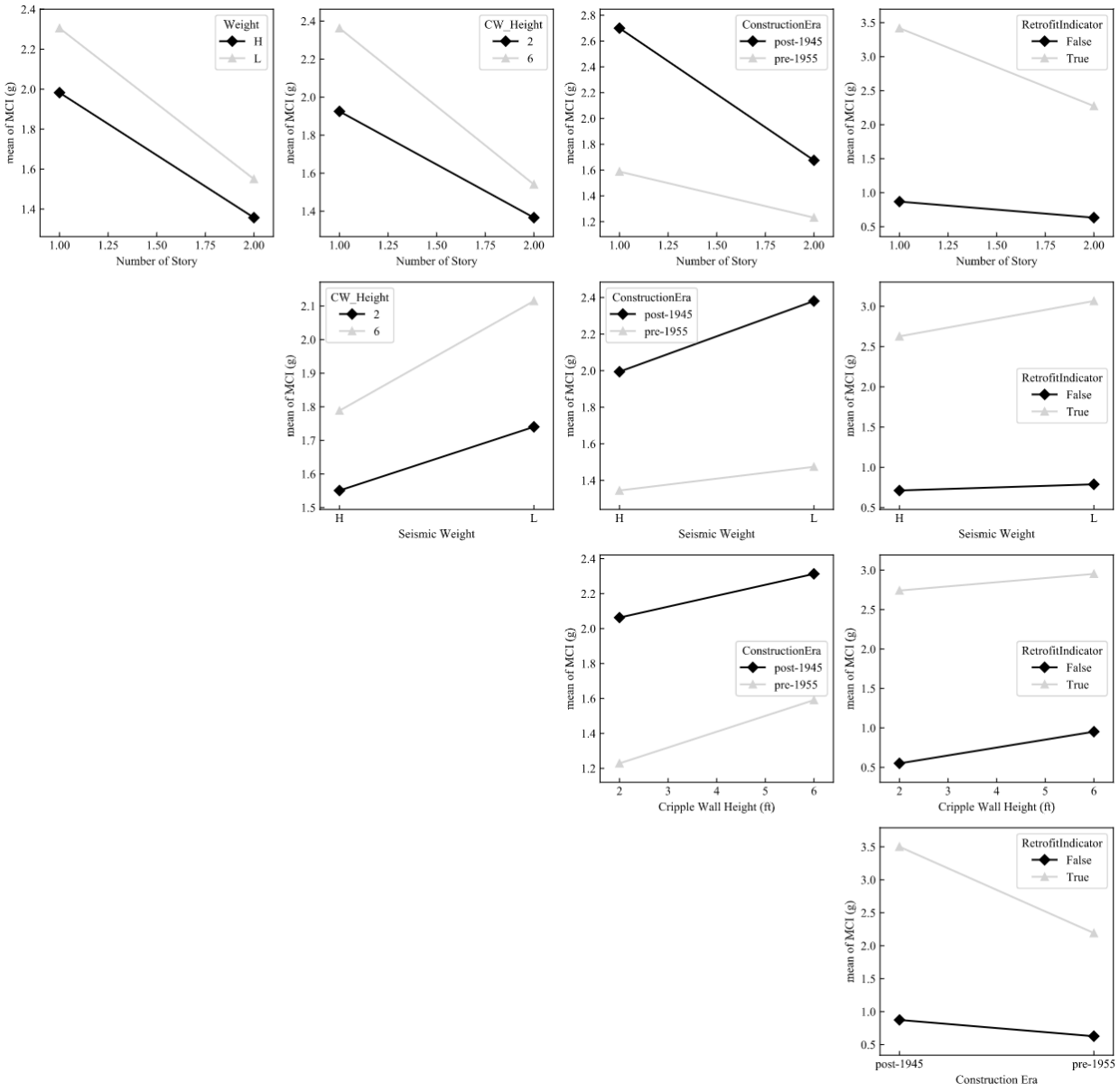


Figure 4.9 MCI factor interaction plot for horizontal wood siding exterior buildings

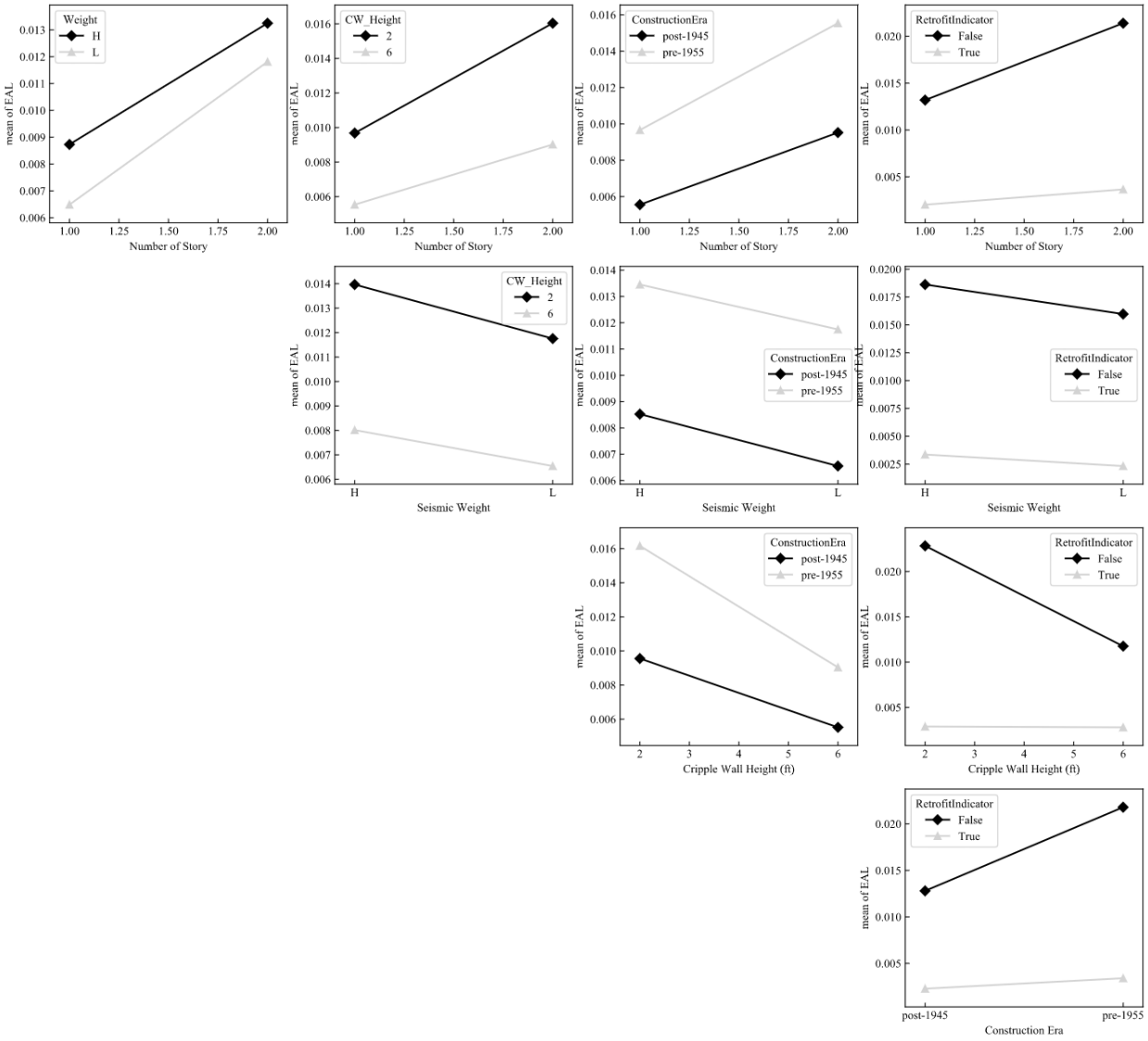


Figure 4.10 EAL factor interaction plot for horizontal wood siding exterior buildings

Most of the preceding discussion was based on visual inspections of the response plots. Quantitative measurements are required to determine whether the features of interest are statistically significant to the responses. The two-way ANOVA is used for this purpose. The p-values are calculated for main and interaction effects on MCI and EAL for the stucco and horizontal wood siding buildings (shown in Tables Table 4.3 through Table 4.6). For a 95% of confidence level, all individual factors are statistically significant to both MCI and EAL. The p-values are consistent with the observations from the box plots. For the MCI, the only significant

interaction effects are between the number of stories and cripple wall height and construction era and retrofit. On the other hand, three additional interaction effects are statistically significant for EAL. The differences in the feature sensitivities explain the linear regression models poorer performance on EAL relative to MCI. Also, it provides additional motivation for the response surface method where the compound effects of different features are incorporated.

Table 4.3 P-values for individual and interacting factors for stucco exterior buildings' MCI

	Number of Stories	**Cripple Wall Height**	**Seismic Weight**	**Construction Era**	**Retrofit**
Number of Stories	2.74E-09*	1.05E-05*	7.26E-01	1.03E-02*	1.78E-02*
Cripple Wall Height		1.30E-08*	2.69E-01	5.29E-01	3.55E-01
Seismic Weight			2.10E-03*	2.38E-01	2.78E-01
Construction Era				1.27E-06*	3.71E-01
Retrofit					7.18E-15*

*Statistically significant for 95% confidence level

Table 4.4 P-values for individual and interacting factors for stucco exterior buildings' EAL

	Number of Stories	**Cripple Wall Height**	**Seismic Weight**	**Construction Era**	**Retrofit**
Number of Stories	5.12E-08*	3.24E-02*	8.91E-01	5.12E-02*	4.20E-05*
Cripple Wall Height		1.11E-05*	5.97E-01	8.10E-02*	4.62E-06*
Seismic Weight			6.10E-03*	7.90E-01	2.70E-01
Construction Era				1.17E-05*	1.69E-04*
Retrofit					1.93E-10*

*Statistically significant for 95% confidence level

Table 4.5 P-values for individual and interacting factors for horizontal wood siding exterior buildings' MCI

	Number of Stories	Cripple Wall Height	Seismic Weight	Construction Era	Retrofit
Number of Stories	3.19E-06*	2.03E-01	5.21E-01	3.93E-03*	3.02E-04*
Cripple Wall Height		7.05E-03*	4.99E-01	5.79E-01	3.51E-01
Seismic Weight			1.93E-02*	2.14E-01	8.63E-02*
Construction Era				7.07E-07*	6.39E-05*
Retrofit					4.06E-13*

*Statistically significant for 95% confidence level

Table 4.6 P-values for individual and interacting factors for horizontal wood siding exterior buildings' EAL

	Number of Stories	Cripple Wall Height	Seismic Weight	Construction Era	Retrofit
Number of Stories	3.91E-07*	2.89E-02*	5.10E-01	1.30E-01	4.93E-05*
Cripple Wall Height		7.27E-08*	5.45E-01	2.02E-02*	9.06E-08*
Seismic Weight			7.28E-03*	8.26E-01	1.95E-01
Construction Era				2.70E-07*	6.47E-06*
Retrofit					5.20E-14*

*Statistically significant for 95% confidence level

4.3. Machine Learning Model-Based Sensitivity Analysis

A preliminary sensitivity analysis for the median collapse intensity and EAL was presented in the preceding section where the main and interaction effects were evaluated for a full factorial experiment. Although this approach provides a simple and interpretable analysis, there are a number of drawbacks. The required number of numerical analysis cases associated with the full

factorial design increases exponentially with the number of factors of interest. In the prior analyses, only 6 variants were examined, and other dynamic response parameters such as the damping ratio and material properties were fixed. To consider additional features, a much larger number of numerical analyses must be incorporated. The stated statistical analysis procedure was based on two-level experiment design. Whether the conclusion is universally applicable is questionable. More importantly, the ANOVA is established based on an underlaying linear regression model. However, such models (ordinary least square and RSM) showed predictive performance for EAL so the conclusions might be misleading. Using the ML models developed in section 3.3, the building seismic performance sensitivity to the input features can be conveniently assessed. Since the ML models were established on large number of randomly sampled parameters and numerical simulations, more generalizability is expected. Also, tree-based models showed superior predictive performance relative to linear regression. Therefore, the conclusion is expected to be more reliable. In this section, the linear regression and XGBoost models are selected evaluate the effect if different features on the building seismic performance. Although it has been shown linear regression has the poorest predictive capability among the derived models (especially for EAL), the regression coefficients are still a useful measure of feature importance. Linear regression has the scaler form shown by Equation (4.4).

$$\hat{y}_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{i,j} \tag{4.4}$$

Where, β_0 is the intercept term, β_j is the regression coefficient of the j^{th} feature, and $x_{i,j}$ is the value of the j^{th} feature of the i^{th} data point. The regression coefficient β_j can be interpreted as the expected change in the prediction $\hat{y}_i$ caused by a unit variation in $x_{i,j}$. Since all features have been

unified during the development of the linear regression model, the coefficients can be directly compared to measure relative importance.

XGBoost explains the importance of individual features based on their contribution to the prediction. Recall that in Equation (3.34), the subtrees are developed on the basis of ancestor nodes by maximizing the gain. More specifically, a binary split of the dataset would create costs, which are measured by within sub-dataset deviations for regression tasks. The feature and value generating the lowest costs (highest reduction in the cost) are selected to execute the split. Conceptually, the features that are selected to split the tree in the earlier stages of the model development and generates lower cost or higher gain have the greatest impact on the response. The detailed mathematical derivation was provided in section 3.2.1.3.

4.3.1. Median Collapse Intensity Sensitivity Analysis

Listed in Table 4.7 are the linear regression model coefficients and XGBoost gain scores used for evaluating the sensitivity of the median collapse intensity to the input features. Overall, the relative feature impacts derived from linear regression and XGBoost models are similar. For the linear regression coefficients, the seismic weight is the most impactful feature and retrofit length and retrofit indicator have the next highest influence. More specifically, the linear regression model shows that the building retrofit could increase the median collapse intensity, on average, by 0.8g. Also, for every 5 feet increase in the length of the retrofit wood panel (a total of 40 feet in the entire building), the median collapse intensity increases by 1.3g. As expected, the cripple wall height and strength also contribute significantly to the building collapse safety. The cripple wall height is positively correlated since the material properties are defined based on drift. Taller cripple walls provide more displacement capacity and stiffness compared to shorter ones. These two features together with the superstructure strength affects the strength distribution along the

building height and determines the position of the weak story. The number of stories and seismic weight decrease the collapse safety at the same time since they are synchronized. Considering the same building and weight distribution, an additional one story approximately reduces the median collapse intensity by 1g. The two displacement parameters are the least impactful features. The coefficients are consistent with the observations from the box plots in section 4.2.3.

As for the XGBoost gain scores, retrofit length has the highest score while the retrofit indicator does not contribute to the prediction. The reason is that existing buildings have zero retrofit length and zero retrofit indicator, and there are no differences when splitting the dataset on either of the two features. The retrofit length is more effective in reducing the cost, so the retrofit indicator was not adopted in the tree growth. For the same reason, the seismic weight and number of stories have very high and low gain scores, respectively. The seismic weight and cripple wall strength parameters are the 2nd and 3rd predominant features. Other than the number of stories, the two material displacement parameters are the least impactful features, which is consistent with the linear regression coefficients. Although the XGBoost gain score captures the relative importance of features from the perspective of contribution to the predictions, the feature impacts cannot be directly quantified.

Table 4.7 Linear regression and XGBoost median collapse intensity model feature sensitivity

Features	Linear Regression Coefficient	XGBoost Gain Score
Seismic Weight	-2.48	1740
Retrofit Length	1.30	2838
Retrofit Indicator	0.80	0
Cripple Wall Height	0.68	91
Cripple Wall Force	0.63	1438
Damping Ratio	0.57	78
Super Structure Force	0.27	832

Number of Stories	-0.16	18
Cripple Wall Displacement	0.08	71
Super Structure Displacement	0.06	22

Additionally, a sensitivity analysis using variable control method and XGBoost model is presented here. Figure 4.11 shows how XGBoost predicted median collapse intensity changes when shifting different building variants. The horizontal axis in the figure refers to a reduction/amplification factor applied to the feature, and the vertical axis shows the associated EAL values predicted by the XGBoost model. 11 ratios starting from 0.75 to 1.25 with 0.05 increment are adopted. The point with applied ratio equal to 1 refers to the baseline building with 2ft high cripple wall, standard seismic weight (67 kips) and best estimation of the material properties. One- and two-story un-retrofitted cases share very similar sensitivity patterns. Same as the previous discussion, building seismic weight and cripple wall strength are the two most predominant features. Seismic weight negatively and cripple wall strength positively influence the collapse resistance. Increasing seismic weight by 20% (compared to baseline case) approximately reduce the median collapse intensity by 40%. The two-story case benefits more than one-story case from increased cripple wall strength. 20% stronger cripple wall level nearly improves collapse resistance by 40%. Other factors don't reveal significant contribution to median collapse intensities. Situation is more complicated on the retrofitted buildings. For the one-story retrofitted case, structural panel length is the most influential factor, while only increased structural panel length shows a difference. As for the two-story retrofitted building, superstructure strength rather than retrofit structural panel length is the most impactful factor. From comparing these two cases, the two-story retrofitted building reaches the retrofit threshold, where upper level is relatively weaker than the retrofitted cripple wall level. Therefore, increasing structural panel length cannot further benefit collapse resistance while

strengthening superstructure could. The 'one-side' impacts of weaker cripple wall could also provide additional clue to the explanation.

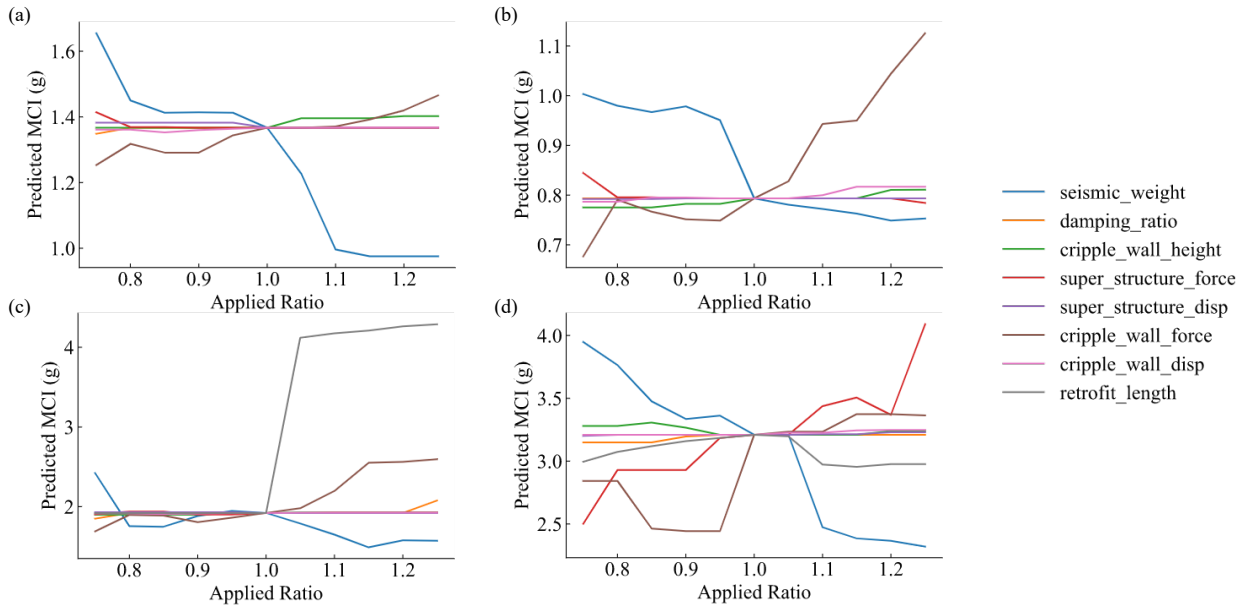


Figure 4.11 XGBoost feature impacts on MCI: (a) 1-story existing building, (b) 2-story existing building, (c) 1-story retrofitted building, and (d) 2-story retrofitted building

4.3.2. EAL Sensitivity Analysis

A slightly different sensitivity analysis is conducted for EAL. Since the linear regression model delivers poor predictions for EAL, the coefficients cannot be used for quantifying the relative feature importance. As for the XGBoost gain scores, although the relative impacts can be compared, the scores are difficult to interpret. Also, the gain score cannot reflect the whether the feature is positively or negatively correlated with the performance. Therefore, the variable control method is adopted here on the basis of XGBoost model to examine the sensitivity of EAL.

Figure 4.12 presents the variations in the EAL when changing an individual feature while fixing the rest. The analysis settings remain the same as the preceding section. The results are shown for the 1- and 2-story, existing- and retrofitted-buildings, respectively. Aside from a few exceptions, the two existing-/retrofitted-buildings share similar feature sensitivities. For the existing buildings,

the seismic weight, superstructure displacement and cripple wall strength parameters are the three predominant features. Compared with the baseline building, increasing the seismic weight by 25% almost double the EAL and vice versa. The effect of seismic weight on EAL is complex. As discussed earlier, the seismic weight is negatively correlated with the median collapse intensity, resulting in higher collapse losses under heavier loads. As for non-collapse losses, the seismic weight determines building natural period and mode shape, which further impacts the damping assignment and ground motion spectral acceleration. These factors contribute to the magnitudes and vertical distributions of EDPs. The impact of seismic weight on EAL is approximately uniform over the investigated range. The superstructure displacement and cripple wall strength parameters are the other two most influential features. Increasing the superstructure displacement parameters leads to reduced stiffness and higher displacement capacity. Thus, higher story drift ratios are expected in the superstructure. The effect of the cripple wall force parameter is similar, However, the effect of the two features are not the same on the one-story un-retrofitted case. A stiffer superstructure and weaker cripple wall minimally affect the EAL. In the baseline building, the presence of the cripple wall creates a weak story mechanism. Collapse and the highest drift demands occur at cripple wall level. Reducing these two parameters exacerbate the stiffness and strength differences. The cripple wall level is expected to suffer slightly higher component damage and the collapse risk is reduced, but the changes in EAL are not significant. More variations can be observed in retrofitted cripple wall buildings when the parameters are changed. As with the existing buildings, the seismic weight, superstructure force and displacement parameters and cripple wall strength parameters are significant. However, the superstructure displacement parameters contribute in both directions. The superstructure strength parameters dominate the EAL of the retrofitted buildings. These observations are related to the relative strength and stiffness of

the superstructure and cripple wall levels. The retrofit wood panel can balance the strength and stiffness differences between the cripple wall and superstructure, and the weak story mechanism is suppressed. In such cases, the superstructure strength, stiffness and ductility determine the EDP magnitude, vertical distribution and collapse safety. Given that most of the damageable components are in the superstructure, the retrofitted buildings are more sensitive to the superstructure material properties. Thus, the effect of a stronger superstructure is more observable, and the reduced superstructure displacement has an impact on the retrofitted building EAL. For the rest of the parameters, the EAL fluctuations are comparable with the model prediction error. Consequently, they can be categorized as EAL non-sensitive features.

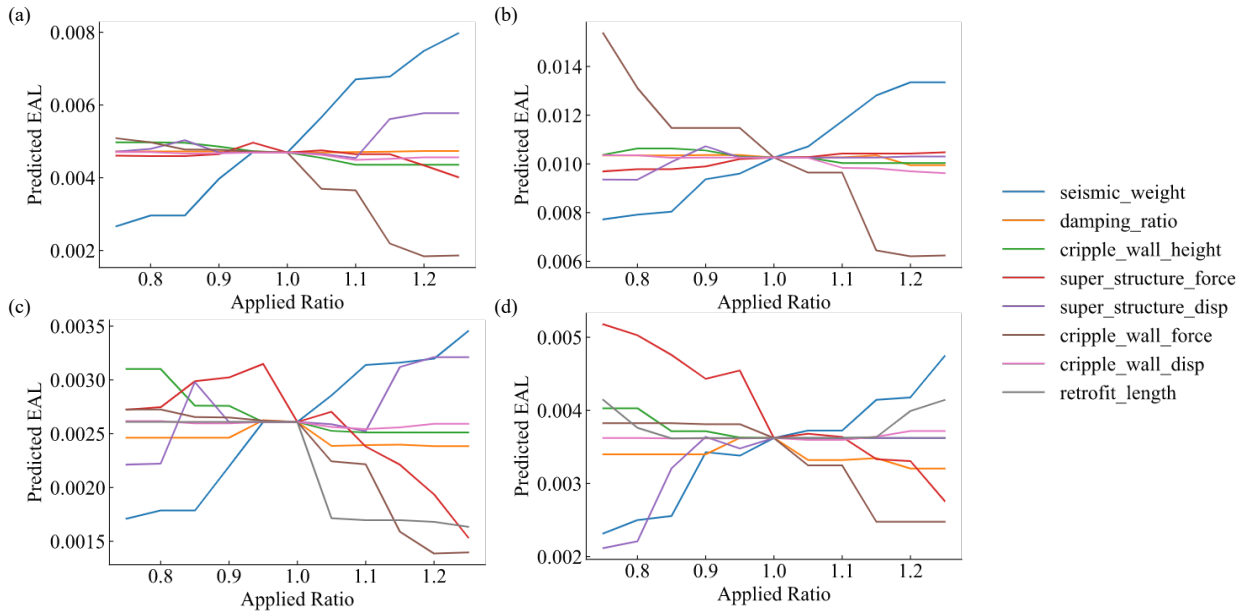


Figure 4.12 XGBoost feature impacts on EAL: (a) 1-story existing building, (b) 2-story existing building, (c) 1-story retrofitted building, and (d) 2-story retrofitted building

4.4. Machine Learning Model Uncertainty Quantification

Within the scope of current study, the ground motions for assessing building collapse performance are selected based on the conditional spectrum therefore effect of spectra shape is explicitly considered. Figure 4.13 provides a flow chart showing the procedure used to numerically perform

uncertainty quantification in PBSD. To begin with, a set of random parameters that are known to affect the building performance is specified. For each random parameter, the range and distribution of interest are needed. Then, a set of random samples are generated from these distributions using an appropriate technique (e.g. Monte Carlo, Latin Hypercube Sampling). A set of finite element models are constructed using the sampled parameters. The target response variables of interest (e.g. collapse performance, earthquake induced losses, EDPs) are obtained through formal numerical analysis and performance assessment procedure. Lastly, the response variables distributed are analyzed to obtain the propagated uncertainties. In this procedure, a large computational burden arises during the numerical analyses. Besides, the number of required samples are usually unknown to achieve a stable outcome. ML models can be used to complement the numerical analyses. For any given set of sampled parameters, the responses can be estimated using the ML model without numerical evaluation. Additionally, the number of random samples can be arbitrarily large until stable results are obtained with the associated computational expense remaining unchanged. This section demonstrates the process of using previously-developed XGBoost models to quantify the effect of model uncertainties on collapse safety.

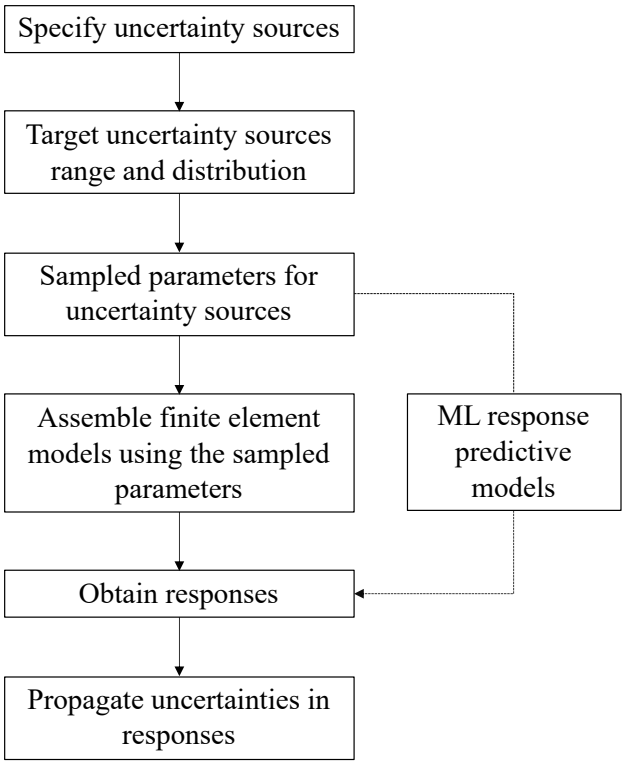


Figure 4.13 Flow chart showing uncertainty propagation methodology

4.4.1. Effect of Model Uncertainty on Median Collapse Intensity

Based on the sensitivity analysis, the features used to construct the ML models will be introduced as the sources of model uncertainty. These features are statistically significant to the wood frame building collapse performance, and are also commonly considered in building performance uncertainty analysis (Porter et al. 2002; Liel et al. 2009; Yin and Li 2010; Frangopol and Liub 2011; Jiang and Ye 2020). Table 4.8 lists the statistical distribution information of the uncertainty sources.

Table 4.8 Uncertainty sources statistical characteristics

Random Variable	Distribution	Mean	CV	Reference
Seismic Weight	Normal	67 kips	0.1	Ellingwood 1980; Haselton 2006
Cripple Wall Height	Normal	2 ft	0.2	Ellingwood 1980; Haselton 2006
Damping Ratio	Normal	0.025	0.4	Porter et al. 2002

Cripple Wall Displacement Parameters	Lognormal	1.0	0.3	Jiang and Ye 2020
Cripple Wall Strength Parameters	Lognormal	1.0	0.3	
Superstructure Wall Displacement Parameters	Lognormal	1.0	0.3	
Superstructure Wall Strength Parameters	Lognormal	1.0	0.3	

To estimate the probability of collapse given the intensity measures, other than building median collapse intensity, the dispersion of the lognormal distribution is required. Thus, an additional XGBoost model is developed to predict the dispersion. The detailed model derivation follows the same procedure described in section 3.2. The final XGBoost model has a testing $MARD$ of 15% and R^2 score of 0.81.

With the specified uncertainty distributions, the median collapse intensity and dispersion prediction models, the Monte Carlo (MC) method is executed. A total of 1000 independent random samples are generated, and the corresponding collapse medians and dispersions are estimated using the XGBoost models. Then, a fragility curve with model uncertainty is estimated through aggregating all 1000 fragility curves. Specifically, at each intensity level, the expected empirical number of collapse cases with model uncertainty is taken as the average collapse probability of the 1000 cases. A new lognormal distribution is fitted using the aggregated collapse probability at each intensity level. The resulting dispersion includes both model and ground motion record-to-record uncertainty. The empirically estimated β_M is taken as the square root of the difference between the square of the total and record-to-record dispersion, which is shown in Equation (4.5).

$$\beta_M = \sqrt{\beta_T^2 - \beta_{RTR}^2} \tag{4.5}$$

where β_{RTR} is the record-to-record dispersion and β_T is the total dispersion. Using this procedure, the model uncertainty associated with the collapse fragility is propagated for the 2-ft high cripple wall stucco exterior gypsum wallboard interior building. Table 4.9 lists the median collapse intensity with and without β_M. One major difference between the FEMA P-58 and ML-based approach to evaluate the effect of model uncertainty is the median collapse intensity. Within FEMA-P58 framework, model uncertainty is directly added to the ground motion dispersion using SRSS without affecting the median collapse intensity. However, as shown by the ML models, the fragility curve with model uncertainty has a different median collapse intensity. Except the 2-story retrofitted cases, the rest median collapse intensities with model uncertainty have on average 15% higher collapse resistance. As for the 2-story retrofitted case, with considering model uncertainty, the MCI decreases by 10%. One possible explanation could be the features' impacts on the building median collapse intensity and dispersion are not uniform and consistent across different cases. For example, the cripple wall force strength parameters were shown to only affect the building collapse risk when it increases in the 1-story buildings. Table 4.10 presents the estimated model uncertainty for the 1- and 2-story, existing- and retrofitted-cases. FEMA P-58 recommends that the total uncertainty is calculated using SRSS with β_{RTR} assuming a β_M of 0.35. Overall, with the introduction of model uncertainty, the total dispersion increases by approximately 20% to 50% depending on the building type. Table 4.9 shows that, the measured β_M is lower than the 0.35 recommended by FEMA P58 for the 1-story existing and 2-story retrofitted buildings. However, the opposite is true for the 2-story existing and 1-story retrofitted buildings

Table 4.9 Measured median collapse variations due to model uncertainty obtained from XGBoost model

Building Type	Median Collapse Intensity without β_M	Median Collapse Intensity with β_M
1-story Existing	1.22	1.45
2-story Existing	0.81	0.97
1-story Retrofitted	3.15	3.39
2-story Retrofitted	2.73	2.46

Table 4.10 Measured model uncertainty obtained from XGBoost model

Building Type	β_{RTR}	Estimated β_T	Estimated β_M
1-story Existing	0.310	0.457	0.335
2-story Existing	0.398	0.568	0.405
1-story Retrofitted	0.303	0.443	0.434
2-story Retrofitted	0.517	0.609	0.322

The fragility curve without model uncertainty, with the FEMA P-58 recommended model uncertainty and with the model uncertainty derived from XGBoost models are provided in Figure 4.14. The light grey fragility curves are from the individual MC samples. The estimated fragility curve with β_M (blue curve) is developed by averaging the probability of collapse at each intensity level of all MC samples. Note that the model uncertainty also affects the median collapse intensity value. The red dash curve was derived through applying β_M to the fragility curve without model uncertainty (orange curve). The green curve, which uses 0.35 as β_M, is provided as reference. By considering additional model uncertainty, fragility curves have flattened. The probabilities of

collapse beyond median collapse intensity decrease and vice versa. The detailed probabilities of collapse under the design basis earthquake (DBE) (1.2g) and maximum considered earthquake (MCE) (1.8g) are listed in Table 4.11. At MCE level, for the 1- and 2-story existing buildings. Incorporating model uncertainty reduces the probability of collapse by 7% to 10%. However, for the retrofitted buildings, the probability of collapse at the DBE increases. The magnitude of the change in the probability of collapse is determined by the intensity level relative to the median collapse intensity.

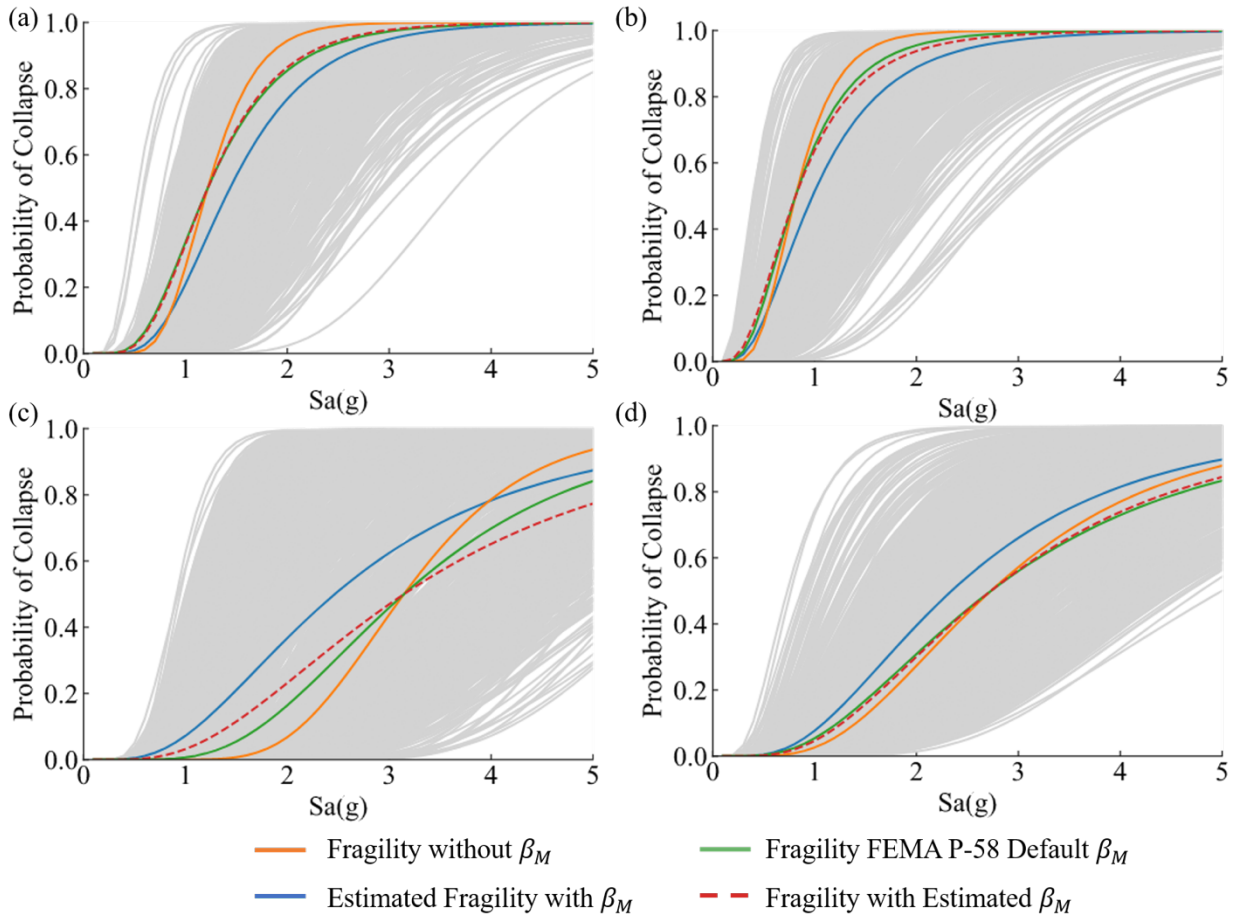


Figure 4.14. Collapse fragility with propagated model uncertainty using the XGBoost model: (a) 1-story existing building, (b) 2-story existing building, (c) 1-story retrofitted building, and (d) 2-story retrofitted building

Table 4.11 Probabilities of collapse under the MCE and DBE with different uncertainty considerations

Building Type	MCE				DBE			
	Fragility Without β_M	**Fragility with 0.35 β_M**	**Fragility with Estimated β_M**	**Estimated Fragility with β_M**	**Fragility Without β_M**	**Fragility with 0.35 β_M**	**Fragility with Estimated β_M**	**Estimated Fragility with β_M**
1-story Existing	0.90	0.80	0.80	0.68	0.48	0.49	0.49	0.34
2-story Existing	0.98	0.93	0.92	0.86	0.84	0.77	0.75	0.65
1-story Retrofitted	0.03	0.11	0.10	0.08	0.00	0.02	0.01	0.01
2-story Retrofitted	0.21	0.25	0.25	0.30	0.06	0.09	0.09	0.12

4.4.2. EAL Uncertainty Quantification

A similar approach is used to quantify the effect of model uncertainty on EAL. Recalling the FEMA P-58 loss assessment procedure, the ground motion uncertainty is incorporated in the resulting loss curve and EAL through the fragility dispersion and EDP covariance matrix. Model uncertainty can be combined with the dispersion of the fragility curve and covariance matrix using SRSS. With the help of the ML model, this step can be bypassed and the variations in EAL caused by model uncertainty can be quantified. To quantify the effect of model uncertainty on EAL, a ML model is constructed on EAL without model uncertainty incorporated in the loss assessment is required. Following the same procedure introduced in section 3.2, an XGBoost model is derived for EAL, where no model uncertainty is considered. MC simulation is again performed, where 1000 independent random samples are generated to estimate the associated EALs using the developed XGBoost model. The simulation was performed on the same 2ft high cripple wall stucco exterior and gypsum wallboard interior building.

The MC simulation results are presented in terms of a histogram for 1- and 2-story, existing and retrofitted buildings as shown in Figure 4.15. Overall, incorporating model parameter uncertainty into EAL, the empirical distributions are left-skewed with long right tails. A beta distribution was used to fit the empirical probability density functions. The detailed statistics of the EAL with model uncertainty, the fitted beta distributions, and EAL without model uncertainty are provided in Table 4.12. The derived coefficients of variation (COV) across the MC samples range from 0.49 to 0.54 with an average of 0.51. COV values are stable across different archetypes. The average EALs from the MC samples are observed to be higher than the EALs with 0.35 β_M for the existing buildings, but lower for the retrofitted cases. The EALs without incorporating model uncertainty are slightly lower in magnitude than those with model uncertainty. For the FEMA P-58 approach,

introducing model uncertainty increases the probability of collapse under the intensities lower than the median collapse intensity. These intensities have higher rates of occurrence, therefore, the expected annual collapse loss increases. As for the EALs with ML-based model uncertainty, the effect of model uncertainty varies based on the archetype.

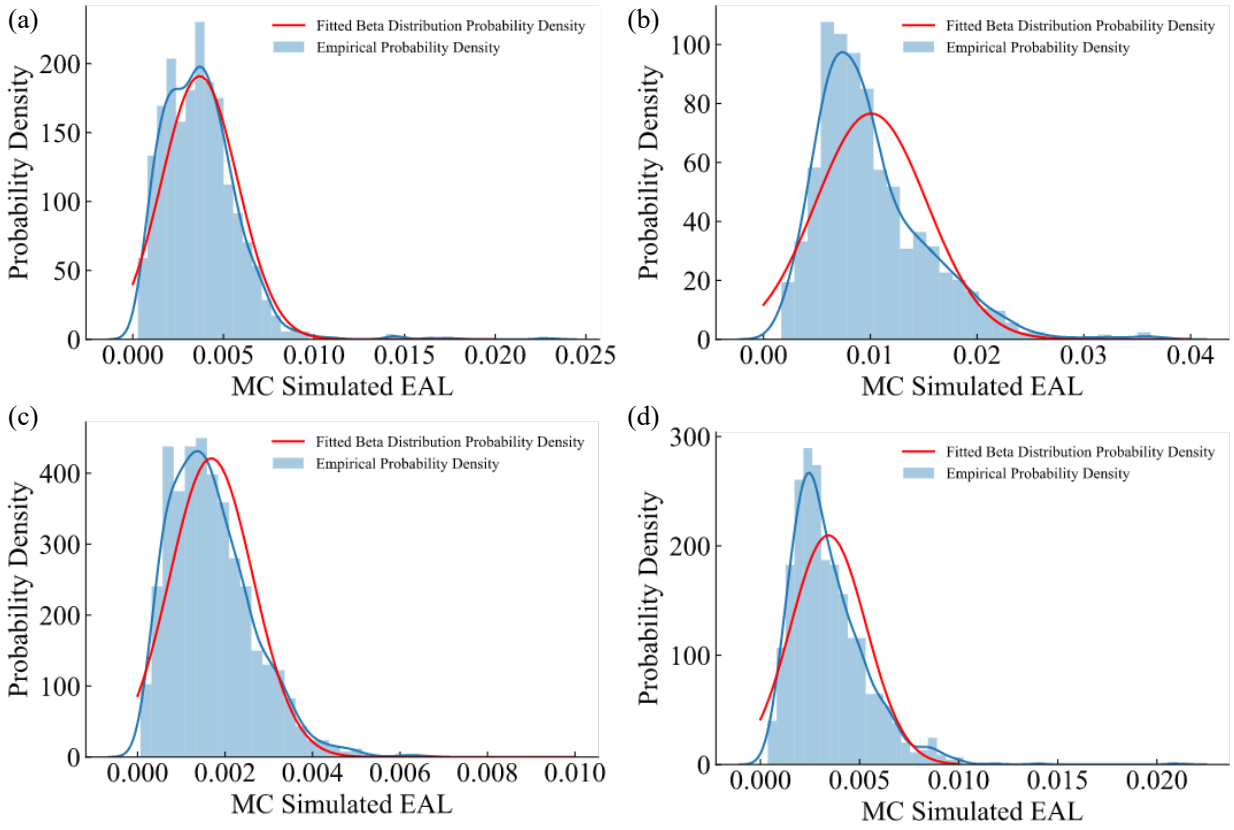


Figure 4.15 EAL with propagated model uncertainty using XGBoost model: (a) 1-story existing building, (b) 2-story existing building, (c) 1-story retrofitted building, and (d) 2-story retrofitted building

Table 4.12 Statistics of EAL with different β_M consideration

Building Type	Mean EAL with Estimated β_M	Coefficient of Variation	Mean EAL of Fitted Beta Distribution	Coefficient of Variation of Fitted Beta Distribution	EAL without β_M	EAL with 0.35 β_M
1-story Existing	0.34%	0.51	0.34%	0.51	0.34%	0.46%
2-story Existing	1.00%	0.50	1.00%	0.50	0.98%	1.1%
1-story Retrofitted	0.18%	0.49	0.18%	0.49	0.09%	0.12%
2-story Retrofitted	0.32%	0.54	0.32%	0.54	0.25%	0.30%

4.4.3. Convergence Analysis

As discussed earlier, the ML model enables computation of the intensity-conditioned collapse probability without additional NRHAs. Thus, a simple convergence check is presented using the 1-story existing to determine the total required computational expense without the ML model. In Figure 4.16, model uncertainty in the 1-story existing single-family house is calculated using 15 different numbers of MC samples ranging from 10 to 1000. As shown by the trend, the fluctuation of effect of model uncertainty decreases with the number of samples. After approximately 600 samples, the quantified effect of the model uncertainty remains relatively stable. This suggests that at least 600 randomly generated samples are required to achieve a stable estimation of model uncertainty for either the ML model-based method or conventional numerical method. Recall a total of 800 models and NRHA are involved in the development of ML model, and this includes the 1- and 2-story, existing and retrofitted buildings. To achieve the estimation accuracy, 2400 numerical models are needed. The number increases with the number of random parameters. A

60% computational cost saving can be expected by using ML models. More importantly, the ML models can be used for other tasks, e.g. sensitivity analysis and optimization.

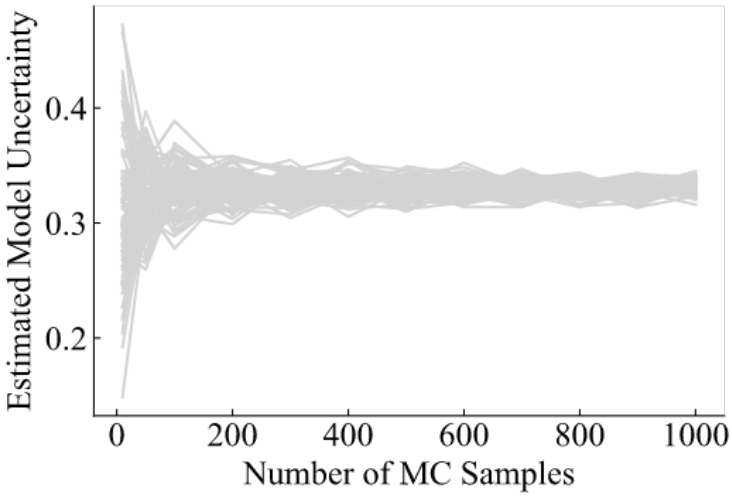


Figure 4.16 Estimated model uncertainty using different number of MC samples

4.5. Summary

In this chapter, a sensitivity study using 2^k experiment design method was first presented to illuminate the building variants' impact on seismic performances. In this preliminary analysis, factors' main effect and interaction effect were evaluated. For single-family cripple wall buildings, the collapse resistance and EAL contribution from building seismic weight, number of stories, cripple wall height, construction era and retrofit benefit were examined. Among the five variants, seismic weight and number of stories were found to negatively and positively affect MCI and EAL, respectively. The opposite is true for the rest three factors. Retrofitting contributes most significantly to the improved collapse safety and reduced earthquake losses. Additionally, ANOVA tables suggest that aforementioned feature interactions have unignorable contribution to building seismic performances, which explains the poor predictive capability of ordinary linear regression models in Chapter 3.

Not limited to predictive capabilities, the ML model can also be adopted for feature sensitivity analysis and model uncertainty quantification. A more comprehensive sensitivity analysis involving more building factors was presented for building collapse safety and EAL using linear regression and XGBoost models. The critical takeaways are summarized here:

(1) Seismic weight is found to be the most predominant feature for median collapse intensity in both linear regression and XGBoost models. The length of the panel used in the retrofit, cripple wall height and cripple wall strength parameters also impact the collapse safety by changing the difference between the cripple wall and superstructure strength. The seismic weight and number of stories are the only two features that are negatively correlated with building the collapse performance.

(2) The superstructure and cripple wall level displacement parameters are the least influential features for the median collapse intensity.

(3) The EAL sensitivities to different features are different for existing- and retrofitted-buildings. The is explained by the fact that the relative strength and stiffness between the cripple wall level and superstructure have significant influence on the engineering demand parameter distributions.

(4) The seismic weight is the most influential feature, and its impact on EAL is approximately linear.

(5) For the existing buildings, the superstructure material displacement parameter and cripple wall level material strength parameters are the other two critical features. The he impact is especially prominent in weaker superstructures and stronger cripple wall. Other than these two features, the superstructure level material strength parameters dominate for the retrofitted buildings. All these features influence the EAL through the mechanism discussed in bullet (3).

Motivated by the conclusions from the sensitivity analysis, the effect of model uncertainty on collapse performance and EAL is quantified. In this regard, the ML model provides with an efficient alternative to conventional numerical analyses. First, the statistical distributions of the random parameters are specified. Then, Monte Carlo simulation is implemented to generate sample sets of these random parameters. The building collapse performance is then estimated using the sampled features. Finally, the collapse fragility curve with model uncertainty was derived by aggregating the fragility curves for the randomly sampled buildings. The model uncertainty dispersions for 1- and 2-story, existing- and retrofitted- buildings are 0.335, 0.405, 0.434 and 0.322, respectively. These values have varying levels of differences with the FEMA P-58 recommended 0.35. With the introduction of model uncertainty, the probability of collapse decreases for intensity levels higher than the median collapse intensity, and vice versa. A similar approach was used to propagate uncertainty in the EAL. The average coefficient of variation in the EAL with model uncertainty is 0.51. The value is found to be fairly stable across different building heights and for the existing and retrofitted cases. Lastly, convergence check is used to assess the number of random samples required to achieve stable uncertainty quantification. As shown from the results of multiple rounds of MC simulations, at least 600 samples are needed to have reliable estimates. Additionally, the results demonstrate that the ML model reduces the computational expanse by at least 60% compared to conventional numerical analyses.

In this chapter, the benefit of ML models for sensitivity analysis, uncertainty propagation and quantification were demonstrated. this chapter is focused on a baseline single family woodframe house with cripple walls and fixed material type. Obviously, the formal model derivation procedure is generally applicable for other construction types (e.g. multi-family woodframe

buildings, steel moment frames, concrete buildings) and responses (e.g. EDPs, resiliency). In Chapter 6, ML models are also used for retrofit design optimization.

CHAPTER 5. EVALUATING THE IMPLICATIONS OF LOS ANGELES ORDINANCE OF SOFT, WEAK, OPEN-FRONT WALL LINE BUILDINGS

5.1. Introduction

5.1.1. Background

Woodframe construction is commonly used for residential buildings in the United States. For multi-unit residential woodframe buildings in the Los Angeles metropolitan area, it is common for the first story to be used for parking or commercial spaces. To increase the accessibility to these types of spaces, a lower wall density is used relative to the upper stories. This can lead to substantial differences in the stiffness and strength of adjacent stories and the formation of a single-story mechanism during earthquake shaking. Numerous complete or partial woodframe building collapses have been attributed to soft-story damage in prior seismic events including the 1971 San Fernando, California (FEMA 2012b), 1989 Loma Prieta, California (Harris and Egan 1992) and 1994 Northridge, California earthquakes (Holmes and Sommers 1996).

Los Angeles is one of several cities in California that has established policies to address the seismic risk to soft-story woodframe buildings. The City of San Francisco enacted the first mandatory soft-story retrofit program for woodframe multi-family buildings in 2013. In southern California, the Cities of Santa Monica and West Hollywood established soft-story retrofit programs in 2017. A similar program was instituted in the City of Oakland in 2019. The Los Angeles Soft-Story Ordinance, which was signed into law in 2015 and later amended, was part of the city's Resilience by Design initiative (https://www.lamayor.org/resilience-design-building-stronger-los-angeles). The ordinance mandates the retrofit of multi-family (more than three units) woodframe buildings with permit applications for new construction submitted prior to January 1, 1978 and at least one soft, weak and open front (SWOF) wall line. A SWOF wall line is defined as having at least one

of the following characteristics: (a) stiffness that is less than 70% of the wall immediately above (i.e. soft), (b) strength that is less than 80% of the wall immediately above (i.e. weak) or (c) there is a diaphragm above with a cantilever that is more than 25% of the distance between lines of lateral resistance from which the cantilever extends (i.e. open). Approximately 12,060 SWOF buildings have been identified in the City of Los Angeles (Xia and Schleuss, 2016).

With the objective of reducing the collapse risk of SWOF buildings, the Ordinance permits four alternative retrofit procedures. The Los Angeles Department of Building and Safety (LADBS) has established its own Structural Design Guidelines (LADBS 2015). Alternatively, the retrofit procedures described in the 2012 IEBC (International Existing Building Code) Appendix Chapter A4 (IEBC 2012), ASCE 41-13 (Seismic Evaluation and Retrofit of Existing Buildings) (ASCE 2017) and FEMA P807 (Seismic Evaluation and Retrofit of Multi-Unit Woodframe Buildings with Weak First Stories) (FEMA 2012b), can be adopted.

5.1.2. Previous Research on Seismic Performance of Soft-Story Woodframe Buildings and Cost-Benefit Analysis

This section provides a brief overview of prior research on assessing the seismic performance of woodframe buildings. The review is primarily focused on seismic evaluation of soft-story buildings, however, several other more general studies on numerical modeling and performance assessment of woodframe buildings are also discussed. As part of the Network for Earthquake Engineering Simulation Soft Story (NEES-Soft) project (van de Lindt et al. 2012), two full scale test programs were conducted to investigate the seismic performance of retrofitted and unretrofitted soft-story buildings. Bahmani et al. (2016) tested a full-scale 4-story building to collapse. The test specimen was designed to represent typical soft-story woodframe building construction in the San Francisco Marina and Mission Districts from the 1950's. The upper stories

of the building had two 2-bedroom apartment units and the first story had an open layout for parking and laundry and storage rooms. The maximum story drift at incipient collapse was 19.3% and the maximum residual drift in the penultimate test was 16.4%. Collapse occurred in the 1st story with the upper stories exhibiting near rigid body displacement. Hybrid testing to collapse of a 3-story soft-story building was also included in the series of NEES-Soft experiments (Jennings et al. 2015). The primary goal of this experiment was to evaluate the shift in collapse to the upper stories caused by over-strengthening the first story. Collapse occurred in the second story, which was constructed using archaic materials, at an intensity corresponding to approximately 125-150% of the maximum considered earthquake (MCE) and a story drift of 8%. Several numerical studies on the seismic performance of the NEES-Soft test specimen were also conducted. One, by Bahmani et al. (2016), used incremental dynamic analyses (IDAs) to evaluate the performance of the as-built and retrofitted 3-story specimen. One of the earliest numerical studies on collapse performance assessment of woodframe buildings was conducted by Christovaslis et al. (2009). Construction quality and ground motion directionality were found to have a significant impact on the seismic collapse fragility of two multi-story woodframe buildings. Li et al (2010) evaluated the effect of construction practices and site-specific hazard on the collapse risk of a 1-story woodframe building. Yin and Li (2010) quantified the effect of aleatory and epistemic uncertainty on the collapse risk of a 1-story woodframe building. As part of an independent review of the FEMA P-807 guidelines, Maison et al. (Maison et al. 2014) developed a numerical model and performed collapse simulation on a four-story soft-story building. Sutley and van de Lindt (2016) chronicled the evolution of seismic design requirements of woodframe buildings from 1959 to the current state-of-the-art. In the same study, the performance of 37 archetypes was evaluated, including as-built and retrofitted (using FEMA P-807) variants of a 4-story soft-story woodframe

building. The retrofit provided life-safety performance at the design basis earthquake (DBE) level and reduced the probability of collapse at the MCE. Park and van de Lindt (2015) used a genetic algorithm to find the optimal FEMA P807-based retrofit solution for multistory residential woodframe buildings with soft stories.

A brief summary of prior studies on cost-benefit analysis of building and bridge seismic retrofit is presented in this section. As part of the CUREE-Caltech Woodframe Project, Porter et al. (Porter et al. 2006) examined the cost-effectiveness of seismic enhancements to four hypothetical woodframe dwellings. Stochastic structural models incorporating uncertainties in mass, damping ratio and force-deformation properties were developed for four types of woodframe buildings that differed based on the quality of construction and whether vulnerable buildings are retrofitted or replaced with modern code-based designs. Earthquake induced losses were quantified using the Assembly Based Vulnerability method. The benefit in terms of US dollars and benefit-cost ratios were computed for each design across different California zip codes considering a 30-year planning period. Retrofit and redesign were found to be cost-effective in half of the investigated locations in California. The median savings ranged from $1,000 to $10,000 over 30 years. Padgett et al. (Padgett et al. 2010) incorporated cost-benefit analyses as part of a broader project on bridge seismic retrofit. Expected life-cycle costs were quantified by integrating probabilistic damage and loss with seismic hazard for four types of multi-span concrete and steel bridges. Five different retrofit measures were considered. With the goal of maximizing cost-benefit ratio, the ideal seismic retrofits were selected for each studied location. The authors concluded that the best retrofit option depends on the seismic hazard and/or location. Liel and Deierlein (Liel and Deierlein 2013) performed cost-benefit evaluations for a set of pre-1970 and modern reinforced concrete frame buildings with assumed locations in high seismicity areas. Performance-based economic loss

assessment was conducted and the cost-benefit implications of retrofitting or replacing (with new modern designs) the pre-1970 buildings was evaluated based on the reduction in expected annual losses (EAL) considering the repair/replacement and fatality costs of earthquake damage. The cost-benefit ratio over a 50-year period and the break-even retrofit cost were used to assess the replacement and retrofit cost-effectiveness, respectively. Replacement of the vulnerable buildings significantly improved life safety and reduced repair cost. Retrofitting was found to be favorable when the cost is less than 50% of the replacement cost. As part of a comprehensive cost-benefit analysis, the National Institute of Building Science (NIBS) showed that federally funded earthquake mitigation projects can save $3 for every $1 spent when both direct and indirect losses are considered (NIBS 2017). Cardone et al. (Cardone et al. 2019) performed a cost-benefit analysis considering different retrofit strategies for reinforced concrete frame residential buildings in Italy. The FEMA P-58 loss assessment methodology was implemented. Net present values of the retrofit investment were computed over a 50-year period, and the breakeven time for retrofit interventions was calculated and used to evaluate the effectiveness of different strategies. Replacing masonry infills and partitions was found to be preferable in terms of cost efficiency while strength-based retrofit techniques had longer break-even times and lower reduction in losses.

5.1.3. Scope and Organization

Given the vulnerability of SWOF buildings in earthquake events and widely implemented retrofitting policy, in Chapter 4, the main objective is investigating how retrofit impacts seismic performance at individual building and regional scale. A comprehensive assessment focusing on Los Angeles Ordinance retrofit is presented here. The discussion mainly aims to address (1) horizontally comparing different retrofit alternatives' strength and collapse safety enhancements; (2) vertically studying Basic Ordinance retrofit benefits in earthquake induced losses and cost-

effectiveness. The set of existing multi-family woodframe building archetypes (32 unique archetypes) developed in Chapter 2 is adopted in the study. Firstly, four alternative retrofit procedures permitted by the Ordinance Guideline, including the requirements documented in the LADBS Structural Design Guidelines (LADBS, 2015), 2012 IEBC (International Existing Building Code) Appendix Chapter A4 (henceforth referred to as IEBC A4) (IEBC, 2012), ASCE 41-13 (Seismic Evaluation and Retrofit of Existing Buildings) (ASCE, 2013) and FEMA P807(Seismic Evaluation and Retrofit of Multi-Unit Woodframe Buildings with Weak First Stories) (FEMA, 2012) are introduced and implemented. A performance-based assessment of the set of existing and retrofitted SWOF woodframe building archetypes to compare the incremental increase in collapse safety that results from applying the four alternative retrofit procedures is then presented. Lastly, a FEMA P-58 based loss assessment and cost-benefit study seeking to assess the cost-benefit implications of the Los Angeles Ordinance at the individual building and regional scales is developed. In the cost-benefit analysis, only the LADBS Structural Design Guideline retrofit procedure is considered.

5.2. Summary of Retrofit Alternatives

LADBS developed a set of structural guidelines that provide the basic engineering requirements for retrofitting SWOF buildings in accordance with the Ordinance. In lieu of the requirements provided in the Structural Design Guidelines (LADBS, 2015), the following alternative retrofit methods that enhance the performance of the entire first story and are at least equivalent to the Ordinance requirements are also permitted: ASCE 41-13, FEMA P-807 and IEBC A4. In other words, if either of these three alternative guidelines are used, the entire weak first story must be analyzed and designed. The engineering requirements of the four alternative retrofit methods are described in this section with the key differences summarized in Figure 5.1.

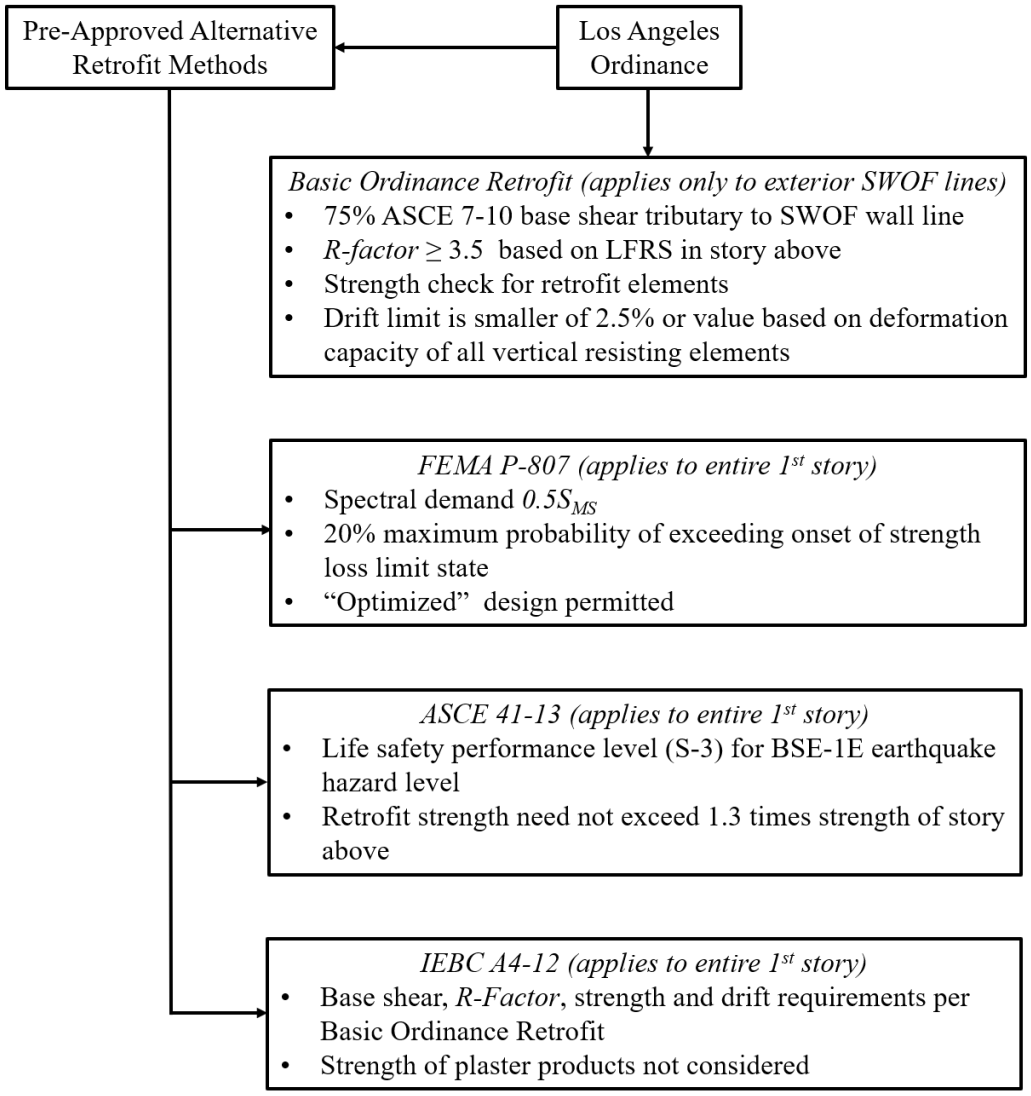


Figure 5.1 Summary of engineering requirements for alternative retrofit procedures

5.2.1. Basic Ordinance (BO) Retrofit

The engineering requirements for retrofitting SWOF buildings in accordance with the Ordinance are outlined in the Structural Design Guidelines prepared by LADBS. The forces used to retrofit the *SWOF wall lines* are based on 75% of the design base shear obtained from the ASCE 7-16 standard. The design forces can be obtained from a two-dimensional analysis of the SWOF wall line using the equivalent lateral force procedure in Section 12.8 of ASCE 7-16. The tributary base shear for the wall line is determined based on a flexible diaphragm assumption. Steel moment

frames, wood structural panels (*WSP*) and cantilevered columns are permitted to be used as strengthening elements. In accordance with Table 12.2-1 of ASCE 7-16, ordinary moment frames (*OMFs*) are generally not permitted in seismic categories D, E and F and intermediate moment frames are not permitted in seismic design categories D and E. However, in cases where (a) the building height does not exceed 35 feet, (b) the roof and floor dead loads do not exceed 35 psf and (c) the wall dead loads do not exceed 20 psf, steel *OMFs* can be used in seismic design categories D and E and intermediate moment frames can be used in seismic design category F. Concrete and masonry walls and steel braced frames cannot be used. The R-Factor used to design the retrofit elements must be less than or equal to that of the existing lateral force resisting elements in the story above but does not need to be less than 3.5. The story drift limit is based on the smaller of the allowable deformation compatible with all vertical load resisting elements (generally assumed to be 2%) and 2.5% (SEAOSC, 2017).

5.2.2. ASCE 41-13

ASCE 41-13 is a comprehensive standard that provides procedures and performance criteria for conducting seismic evaluation or retrofit of all building types and sizes. The Ordinance requires SWOF buildings retrofitted using ASCE 41-13 to achieve the Life Safety performance level in the BSE-1E event (20% in 50-year hazard level). However, the strength of the retrofitted story should not exceed 1.3 times the strength of the story above (LADBS, 2015). Four analysis procedures are included in the ASCE 41-13 guidelines: linear static, nonlinear static, linear dynamic and nonlinear dynamic. The linear static procedure (*LSP*) is consistent with the analysis approach permitted by the Ordinance. Two component-level strength checks are required for the LSP. First, the product of the expected strength of the deformation-controlled element (Q_{CE}), the component capacity modification factor (m), which accounts for ductility and knowledge factor (κ), cannot exceed the

demand (Q_{UD}). For buildings with structural irregularities such as soft-stories, the ratio Q_{UD}/Q_{CE} cannot exceed 3.0. There are no drift requirements in the ASCE-41 procedures.

5.2.3. IEBC A4

The retrofit procedures in IEBC A4 were developed specifically for SWOF woodframe buildings. The analysis procedures, design base shear, strength reduction factor, element strength and drift requirement are the same as the Basic Ordinance Retrofit. However, IEBC A4 does not permit vertical elements that are not sheathed structural panels (WSPs) to be considered as providing structural resistance.

5.2.4. FEMA P-507

The FEMA P-807 retrofit guidelines incorporate a performance-based approach and relies on the statistical evaluation of hundreds of surrogate models that were analyzed using nonlinear response history analyses. The guideline considers the consequence associated with providing "too much" strength in the first story which could lead to excessive damage in the upper stories. The performance of the existing and retrofitted buildings is defined in a probabilistic manner. According to the LADBS Structural Design Guidelines, the spectral acceleration corresponding to $0.5S_{MS}$ per ASCE 7-10 must be used for a FEMA P807 retrofit, except for buildings located in site class E, where the value of F_a must be taken as 1.3 (LADBS, 2015). The target performance objective is based on achieving a 20% maximum probability of exceeding (POE) drift demands corresponding to the onset of strength loss (OSL) in the seismic force-resisting woodframe elements.

5.3. Retrofit Designs for SWOF Buildings

In this section, each of the 32 SWOF archetypes is retrofitted using the four procedures described in the previous section for a total of 128 retrofitted archetypes. A histogram showing the

distribution of S_{MS} values for the sites of the surveyed buildings, which ranges from 0.16g to 2.99g, is shown in Figure 5.2. All retrofits are developed based on the median S_{MS} (2.2g) for the surveyed sites. The corresponding $S_{M1} = 1.2g$. Risk Category II, importance factor $I = 1.0$ and soil site class *D* is assumed for all archetypes. The seismic weight of each building is computed following the same way introduced in Chapter 2. Details of the seismic weight and empirical period (ASCE 7-16 Equation 12.8-7) of each archetype are summarized in Table 5.1.

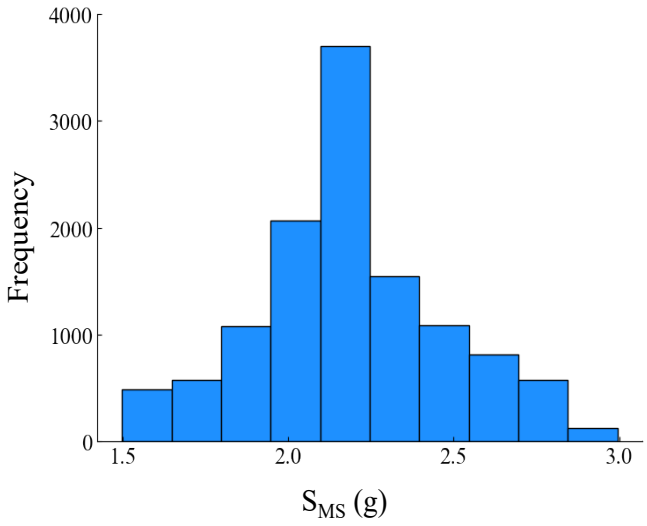


Figure 5.2 Histogram showing distribution of S_{MS} values at sites of surveyed SWOF buildings

Table 5.1 Summary of seismic weights and ASCE 7-16 estimated period for the existing woodframe building archetypes

Building ID	Seismic Weight (kips)	Code Period, C_uT_a (s)
L1-2S-60X30	185	0.25
L1-3S-60X30	306	0.25
L1-2S-100X30	284	0.34
L1-3S-100X30	467	0.34
L2-2S-60X50	287	0.25
L2-3S-60X50	469	0.25
L2-2S-100X50	437	0.34
L2-3S-100X50	709	0.34
L3-2S-50X30	151	0.25
L3-3S-50X30	248	0.25
L3-2S-80X30	242	0.34
L3-3S-80X30	394	0.34
L4-2S-60X50	292	0.25
L4-3S-60X50	475	0.25
L4-2S-100X50	444	0.34
L4-3S-100X50	716	0.34

The locations and types of retrofit elements used for each of the four procedures is shown in Figure 5.3. Since only a single floor plan is shown for each unique layout, the elements shown encompass all associated archetypes and retrofit methods. The details of each retrofit element (frame beam and column wide flange section and WSP length and nailing) including the applicable archetype and retrofit method are summarized in Appendix C of the thesis. A 15'-0" one-bay steel OMF is used in the open wall lines for all four retrofit methods and archetypes. However, as noted earlier, the Basic Ordinance procedure only requires the SWOF wall lines to be retrofitted. Therefore, the steel OMF placed on the partially or completely open wall line is the only retrofit element used for this method. The force demands used to design these elements are computed using a seismic response modification coefficient and deflection amplification factors of $R = 3.5$ and $C_D = 3.0$, respectively, which are consistent with stucco, GWB and HWS serving as the lateral force resisting

elements in the upper stories (LADBS, 2015). Since the strength of the existing panels cannot be considered for the IEBC A4 retrofits, WSPs are used together with the OMFs, where the demands are computed using the same *R* and C_D factors as the Basic Ordinance Method. As shown in Figure 5.3, the WSPs are used to replace portions of the existing stucco walls.

The OMF-WSP combination is also used as needed for the ASCE-41 retrofit procedure, where the site parameters corresponding to the BSE-1E event (20% in 50-year earthquake) are $S_{MS} = 2.2g$ and $S_{M1} = 1.2g$. The site coefficients F_a and F_v (from ASCE 41-13 Tables 2-3 and 2-4) are the same as the ones obtained from ASCE 7-16. As discussed earlier, the LSP is used for the ASCE 41-13 retrofit. To compute the pseudo seismic force (ASCE 41-13 Equation 7-21), the product of C_1 (modification factor used to relate expected inelastic displacements to elastic displacements) and C_2 (modification factor to account for hysteresis shape) is taken as $C_1C_2 = 1.1$ based on Table 7-3 of ASCE 41-13. The effective mass factor to account for higher modal mass participation effects (Table 7-4 of ASCE 41-13) $C_m = 1.0$. The component-demand modification factor (m-factor) corresponding to the life safety limit state is 3.6, 4.7, 2.3 and 3.0 for stucco, GWB, HWS and WSP (Table 7-4 of ASCE 41-13), respectively and 6.0 for the OMF beams and columns (Table 9-4 of ASCE 41-13). The expected panel strengths (Q_{CE}) are based on Table 12-1 of ASCE 41-13. The knowledge factor is taken as $\kappa = 0.75$ because the default strength and stiffness values are used. Note that all ASCE 41-13 retrofits are controlled by the LADBS Structural Guideline provision, which states that the strength of the retrofitted 1st story need not exceed 1.3 times the strength of the story above.

The FEMA P-807 retrofits are developed using the target performance objective of a 20% maximum POE drift demands corresponding to the OSL for a spectral acceleration corresponding to $0.5S_{MS} = 1.1g$, in accordance with the LADBS Structural Guidelines. First, the performance of

the existing structure is evaluated. The plan layout and seismic weight of the existing building are imported into the FEMA P-807 computer-aided tool including the location of every interior and exterior panel and the short period spectral acceleration demand corresponding to $0.5S_{MS}$. The nonlinear load-deflection curves for the stucco, GWB and HWS are based on Table 4-1 of FEMA P-807. The performance of the structure is evaluated and retrofit is required for the cases/directions where the spectral capacity of the existing first story corresponding to 20% POE at the OSL limit state is less than $0.5S_{MS}$. For the cases where the required strength of the retrofit elements exceeds the upper limit (to avoid excessive damage in the 2nd story), an "optimized" solution is adopted, whereby the between 90% of the required strength and 110% of the upper limit strength is provided. Like ASCE 41-13 and IEBC A4, a combination of steel OMFs on the open wall lines and WSPs replacing portions of the existing panels, are used for the FEMA P-807 retrofits.

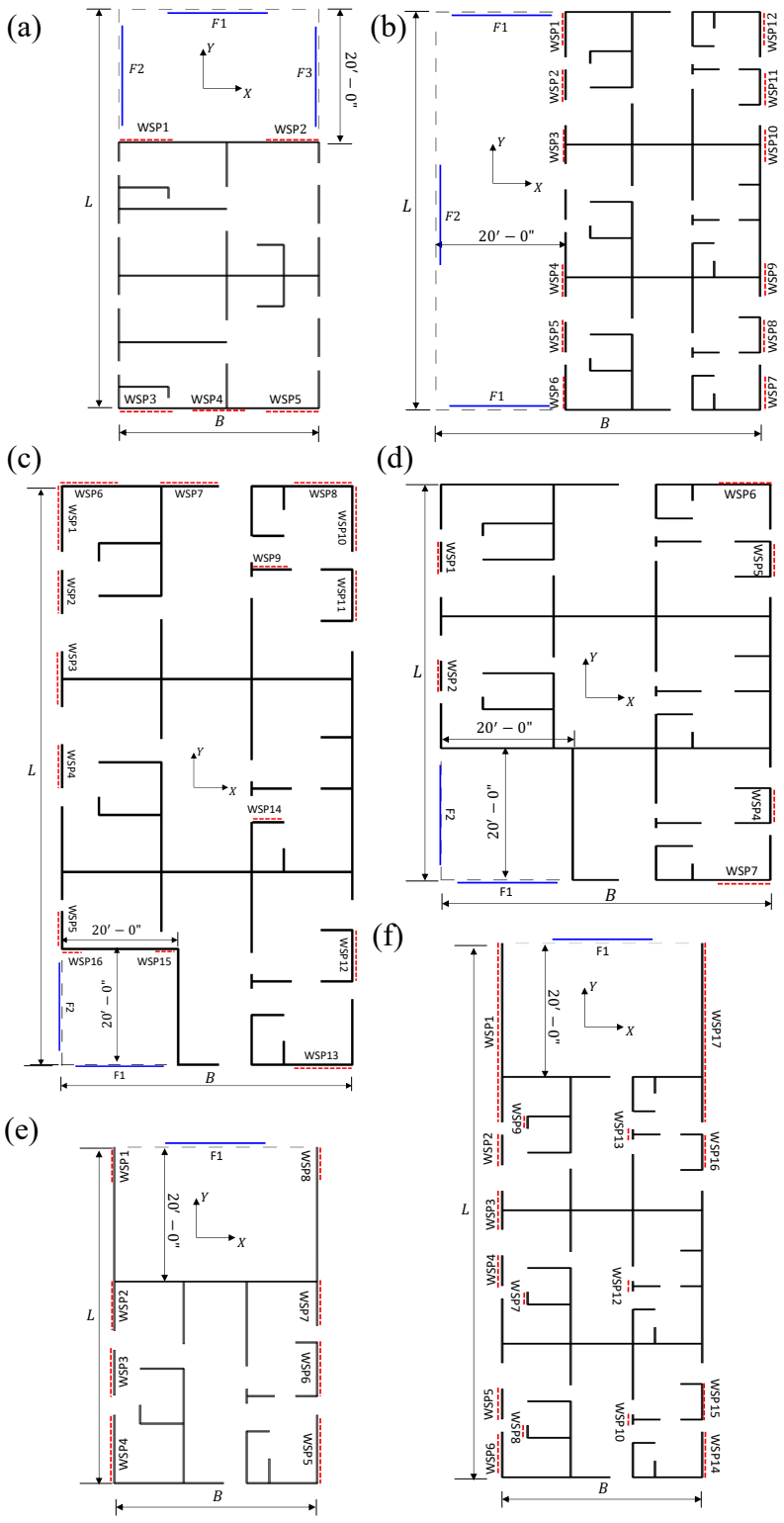


Figure 5.3 First floor plan showing wall layouts and locations of retrofit elements including moment frames and wood structural panels (WSPs): (a) L1, (b) L2, (c) L3 (L = 50’-0”), (d) L3 (L = 80’-0”), (e) L4 (L = 60’-0”) and (f) L4 (L = 100’-0”)

5.4. Performance Evaluation of Existing and Retrofitted Buildings

To evaluate the aforementioned four retrofit alternatives' strength and collapse safety improvement, three-dimensional numerical model is created following the procedure described in Chapter 2 in OpenSees for each of the existing building and retrofitted buildings (a total of 156 buildings with 32 existing and 128 retrofitted buildings). Details of results are presented and discussed in this section.

5.4.1. Nonlinear Static Analysis

Nonlinear static (pushover) analyses are performed on the numerical models to investigate the effect of retrofit on the strength and overall drift capacity of the SWOF buildings. The overall base shear versus drift for the Y-Direction (longer dimension direction) of buildings L1-S2- 60X30-GWB and L1-S2-60X30-HWS is shown in Figure 5.4. The pushover responses for the existing SWOF buildings are shown in Figure 5.4 (a), (b) including the roof, 1st and 2nd story drift where, most of the drift is concentrated in the first story with very low demand in the upper story. Consistent with earlier observations regarding the panel force–deformation envelope, the strength of the GWB building is slightly greater than the HWS building. However, the response of the HWS building is the more ductile of the two.

Figure 5.4 (c), (d) show the pushover response in terms of roof drift for the existing and retro-fitted GWB and HWS structures, respectively. As expected, all four retrofit cases increase both the strength, V_{max}, and drift capacity, θ_u, of the existing building, where the latter is defined as the roof drift demand corresponding to 20% strength loss. The IEBC A4 retrofit is the only one that completely ignores the existing panels and therefore adds the most strength. As discussed later, this is an observation that holds true for all buildings. For the GWB building case shown in Figure 5.4 (c), the ASCE 41-13 and FEMA P807 retrofit designs are the same and thus, the pushover

curves overlap exactly. The θ_u, for that same building ranges from 0.95% in the existing SWOF to 2.4% in the IEBC A4 retrofitted case. For the HWS building, θ_u is 1.3% and 4.2% in the existing SWOF and IEBC retrofitted cases, respectively.

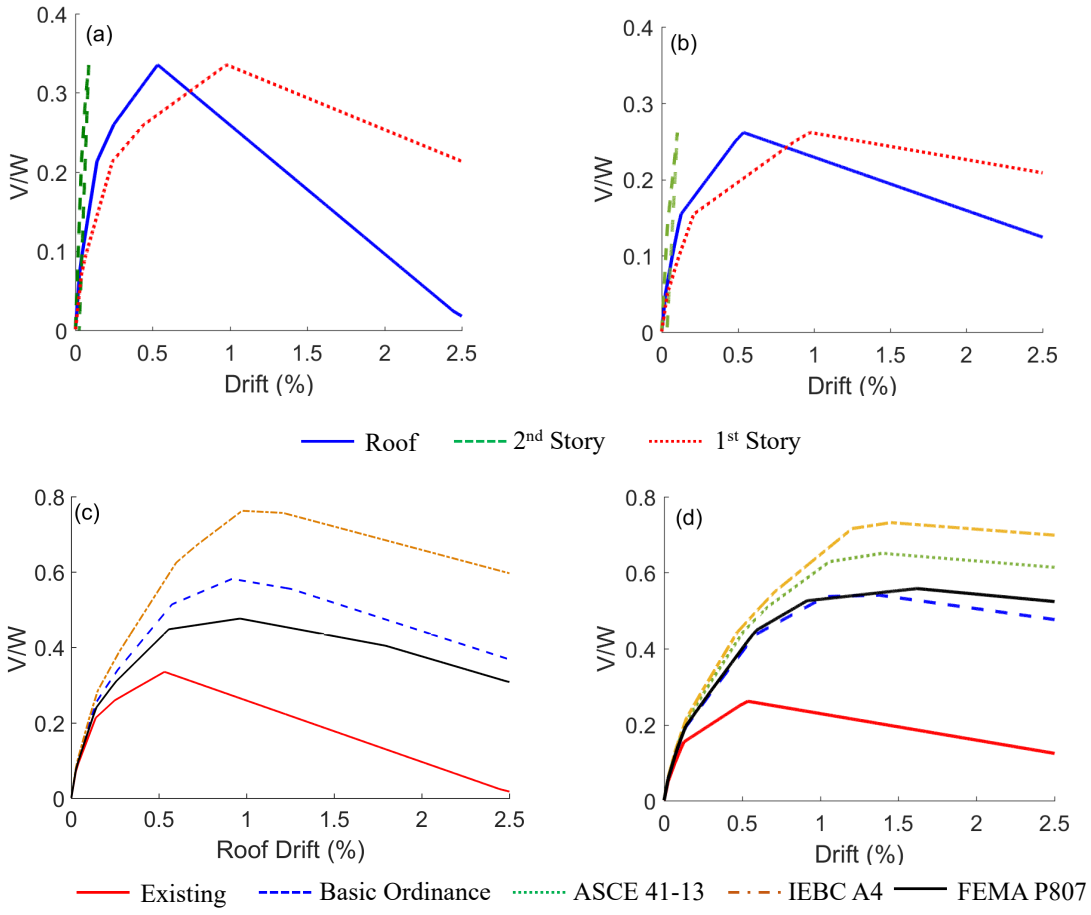


Figure 5.4 Y-Direction normalized base shear (V/W) versus drift for existing SWOF (a) L1-S2-60X30-GWB and (b) L1-S2-60X30-HWS buildings and the four retrofitted (c) L1-S2-60X30-GWB and (d) L1-S2-60X30-HWS buildings

Figure 5.5 shows the ratio between the maximum base shear of the existing and retro- fitted cases ($V_{max,retrof.}/V_{max,exist.}$) for all buildings. The goal is to inform the relationship between an increase in pushover strength and the improvement in collapse safety of the retrofitted structures. All four building properties including the type of interior wall panel, number of stories, plan layout and aspect ratio, are identified in Figure 5.5, where the following conclusions can be drawn:

- The Y-Direction of the buildings with layout L1 derive the greatest benefit in terms of increase in pushover strength. The same is true about the X-Direction of buildings with layout L2.
- On average, $V_{max,retrof.}/V_{max,exist.}$ is 20% higher in buildings with HWS interior panels compared to the cases where the interior panels are GWB.
- The size of the building in plan has only a small effect on $V_{max,retrof.}/V_{max,exist.}$, which on average, is only approximately 4% higher for the large aspect ratio cases.
- For the IEBC A4 retrofit, $V_{max,retrof.}/V_{max,exist.}$ for the 2-story buildings is higher (31%) than the 3-story cases. However, for the other three retrofit cases the average $V_{max,retrof.}/V_{max,exist.}$ is only approximately 3% higher for the 2-story buildings.
- The overall increase in pushover strength is highest for the IEBC A4 retrofit (an average of 84%) and comparable for the other three retrofit methods where the average $V_{max,retrof.}/V_{max,exist.}$ is 35%, 33% and 36% for the Basic Ordinance, FEMA P807 and ASCE 41-13 retrofits, respectively.

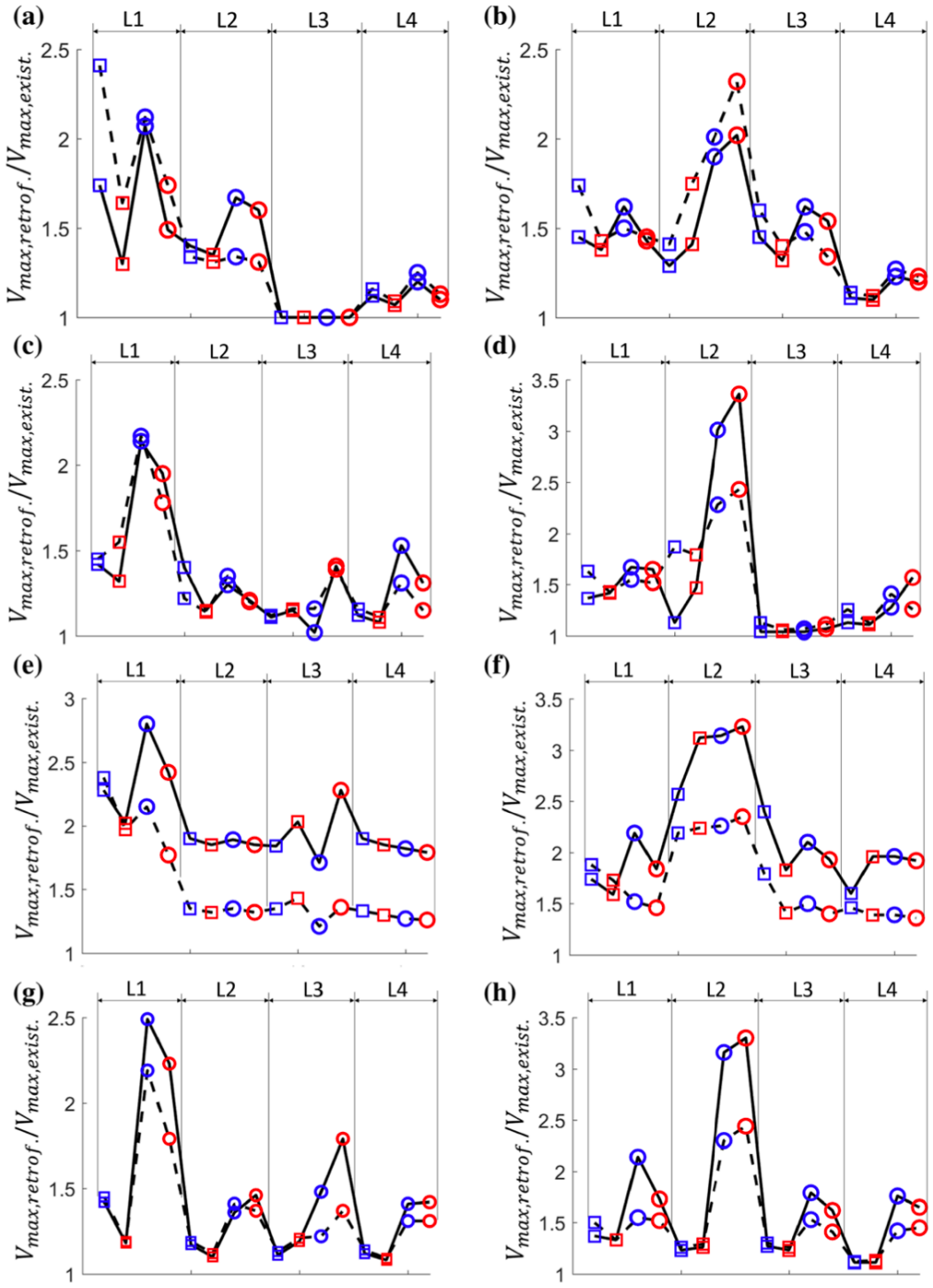


Figure 5.5 Ratio of maximum base shear from pushover response between existing and retrofitted SWOF building ($V_{max,retrof}/V_{max,exist}$): (a) basic ordinance, Y-Direction, (b) basic ordinance, X-Direction, (c) FEMA P807, Y-Direction, (d) FEMA P807, X-Direction, (e) IEBC

A4, Y-Direction, (f) IEBC A4, X-Direction, (g) ASCE 41-13, Y-Direction and (h) ASCE 41-13, X-Direction

Figure 5.6 shows a scatter plot with the ratio between the ultimate roof drift capacity for the existing and retrofitted cases ($\theta_{u,retrof.}/\theta_{u,exist.}$) versus $V_{max,retrof.}/V_{max,exist.}$. Overall, a strong trend is observed between the two metrics. However, Figure 5.6 also reveals that, while the trend is generally very strong for the Basic Ordinance, FEMA P807 and ASCE 41-13 retrofit methods, it is moderate for the IEBC retrofit. A detailed examination of the $\theta_{u,retrof.}/\theta_{u,exist.}$ dataset provided further insight into this and other issues related to the effect of each retrofit method on drift capacity.

For some building properties (plan layout, plan aspect ratio and type of interior panel), the effect on $\theta_{u,retrof.}/\theta_{u,exist.}$ and $V_{max,retrof.}/V_{max,exist.}$ are very similar. However, there are also some key differences. For instance, unlike $V_{max,retrof.}/V_{max,exist.}$, the mean value of $\theta_{u,retrof.}/\theta_{u,exist.}$ for all IEBC A4 retrofits (48%) is comparable to the other methods (between 44 and 60%). Also, when the effect of retrofit on θ_u is disaggregated based on the number of stories, IEBC A4 provides the least benefit for the 3-story cases where the mean $\theta_{u,retrof.}/\theta_{u,exist.}$ is only 14%. This is significantly smaller than the other three retrofit methods where the mean $\theta_{u,retrof.}/\theta_{u,exist.}$ is between 41 and 54%. Whereas, for the 2-story building cases, the mean $\theta_{u,retrof.}/\theta_{u,exist.}$ is highest for the IEBC A4 retrofits (approximately 81%) compared to the other three (49% to 70%). This series of observations points to a key finding: the high levels of strength that the IEBC retrofit adds to the 1st story (because the existing panels are ignored), on average, also provide a benefit in terms of roof drift capacity. However, this benefit is significantly eroded as the number of stories is increased where the drift demands begin to concentrate in the 2nd story. The implications of this finding to collapse safety is discussed in the next section.

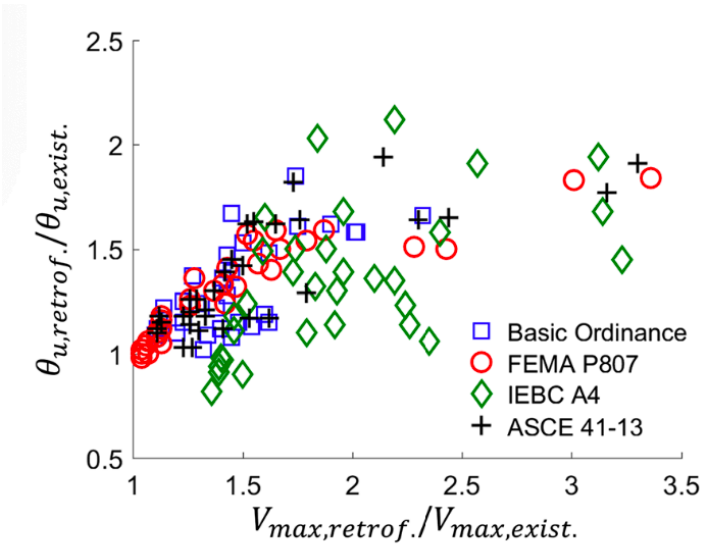


Figure 5.6 Plot showing relationship between increase in strength and drift capacity provided by alternative retrofit methods

5.4.2. Collapse Performance Assessment

The collapse safety of the existing and retrofitted building cases is assessed using incremental dynamic analyses (IDAs), where each ground motion pair is scaled until the collapse point is reached. The collapse analysis is performed using the far-field record set of 22 component pairs of the ground motions specified in the FEMA P695 (FEMA, 2009) guidelines using bi-directional loading. Since the primary focus of the current study is the relative change in median collapse capacity due to the alternative retrofit methods, adjustments for spectral shape and modeling uncertainty are not considered.

Collapse fragility curves for the existing and retrofitted SWOF buildings L1-2S-60X30- GWB and L1-3S-60X30-GWB are shown in Figure 5.7. The only difference between these two building cases is the number of stories. The median collapse intensity for the existing 2- (Figure 5.7a) and 3-story (Figure 5.7b) building cases is 1.08 g and 0.77 g, respectively. It is not surprising that the existing 3-story case has a lower median collapse capacity com- pared to its 2-story counterpart

since the two buildings have the same first story wall layout (where most of the drift demands are concentrated) but the former has a higher seismic weight. For the 2-story case, the IEBC A4 retrofit, which increases the median collapse capacity by a factor of approximately 1.9, provides the greatest benefit to collapse safety. The increase in the median collapse capacity is approximately 40% for the FEMA P807 and ASCE 41-13 retrofits and 50% for the Basic Ordinance retrofit. Relative to the 2-story, the IEBC A4 retrofit results in a smaller increase in the median collapse capacity (approximately 53%) for the 3-story case. The increase in the median collapse capacity is 60% for Basic Ordinance and FEMA P807 and 45% for ASCE 41-13. This finding is consistent with earlier results where, even though the IEBC A4 retrofit provides the greatest increase in the pushover strength for the 3-story case, the overall impact on drift capacity is minimal and less than the other three retrofit methods.

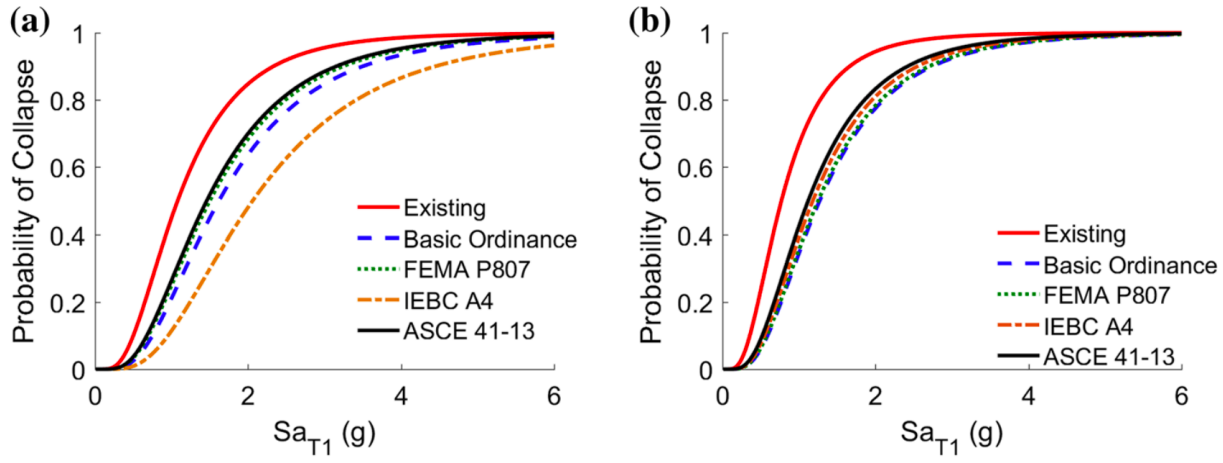


Figure 5.7 Collapse fragility curves for existing and retrofitted SWOF buildings: (a) L1-2S-60X30-GWB and (b) L1-3S-60X30-GWB

Figure 5.8 shows the ratio of the median collapse capacity between the existing and retrofitted buildings $\left(\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}\right)$ for the complete set of archetypes. Like Figure 5.5, each plot

corresponds to a different retrofit method and is annotated to highlight the effect of the various building properties. The almost identical mean $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ values for the high and low aspect ratio cases points to the fact that the size of the building in plan did not play a major role in the effect of retrofit collapse safety. Consistent with the findings related to pushover strength, the overall benefit of retrofit to collapse safety is highest for layouts L1 and L2, where the average $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ is 1.63 and 1.74, respectively. $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ is approximately 20% higher for the buildings with HWS interior panels compared to GWB. This small difference reflects the few cases where the 1st story retrofit results in 2nd story collapse (no collapses occurred in the 3rd story) and the HWS buildings benefit from the higher ductility of the interior panels. Overall, retrofit increases the collapse safety of the 2- and 3-story building cases with the same mean $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ (50%). While the IEBC retrofit provides the highest overall improvement in collapse safety, benefits vary significantly between the 2- and 3-story cases. For the 2-story cases, the improvement in collapse safety is significantly higher than the other retrofit methods with a mean $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ of 78% compared to 39% - 44% for the other three methods. On the other hand, compared to the other three methods, IEBC A4 provides approximately the same benefit to the 3-story building cases with a mean $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$of 50%. This result further highlights the importance of considering the building height or number of stories when selecting an appropriate retrofit method.

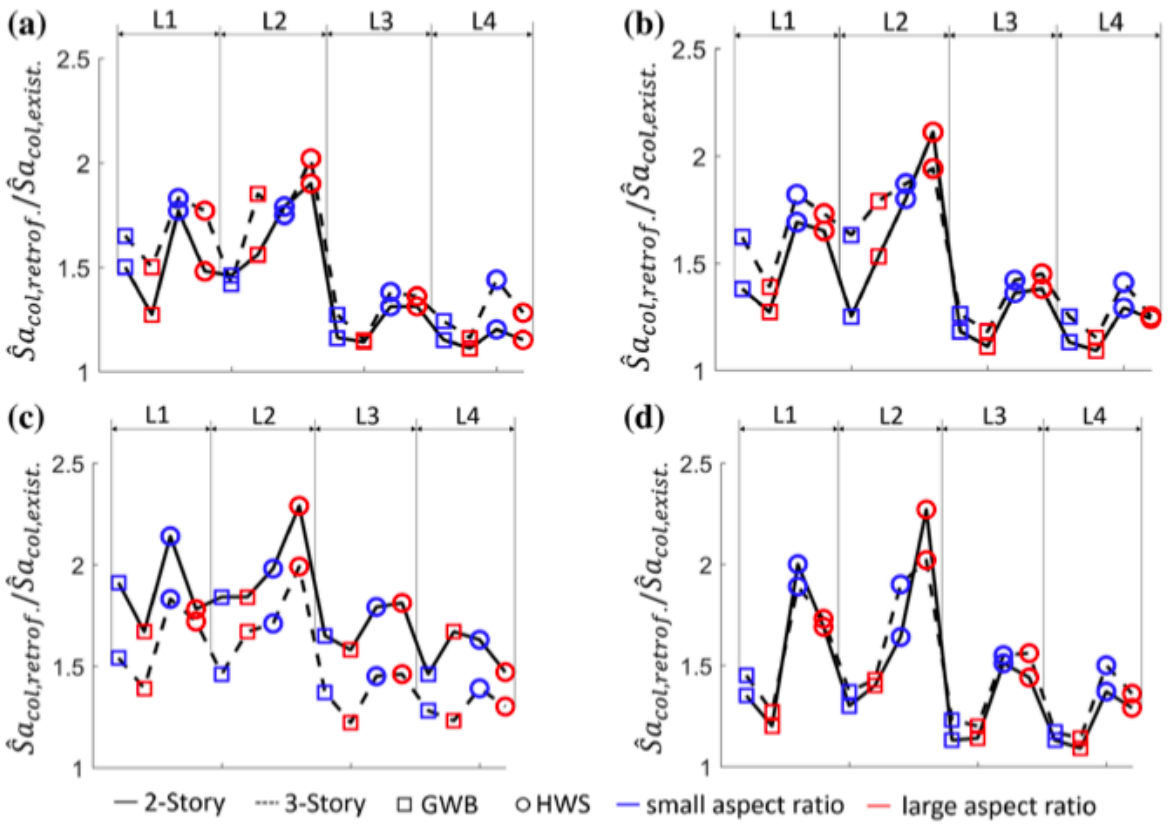


Figure 5.8 Ratio of median collapse capacity between existing and retrofitted SWOF buildings: (a) basic ordinance (b) FEMA P807, (c) IEBC A4 and (d) ASCE 41-13

Figure 5.9 shows a plot of $V_{max,retrof.}/V_{max,exist.}$ versus $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$ where it can be observed that the increase in pushover strength has an overall strong correlation with the improvement in collapse safety. However, consistent with earlier observations, the benefit of added strength varies across the different retrofit methods. The higher degree of "scatter" in the IEBC A4 datapoints reflects the differences in the how much the 2- and 3-story SWOF buildings benefit from retrofit. More importantly, the plot provides critical insights on the development of machine learning building performance prediction models, and it further motivates the implementation of prediction module discussed in Chapter 6.

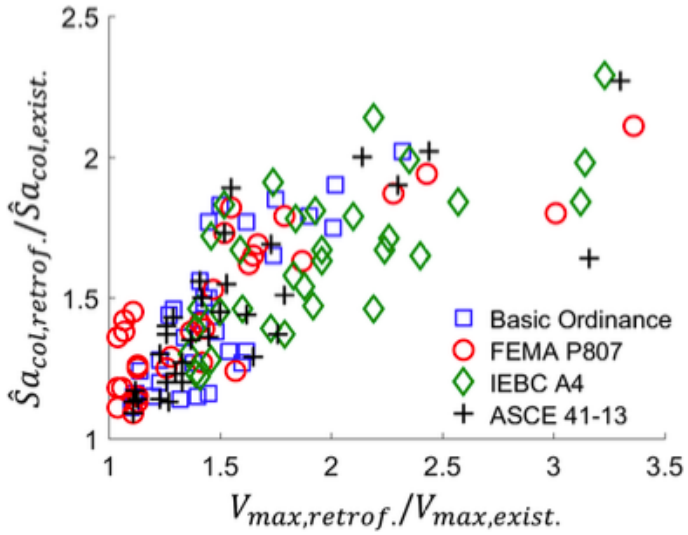


Figure 5.9 Effect of $V_{max,retrof.}/V_{max,exist.}$ on $\hat{S}a_{col,retrof.}/\hat{S}a_{col,exist.}$

5.5. Cost-Benefit Analysis for Los Angeles Ordinance Retrofit

In the previous section, the strength and collapse safety improvements were compared across different retrofit alternatives permitted by Ordinance Retrofit Guideline. In the past 5 years since the implementation of mandatory retrofit plan, more than 43% inventories have been in a compliance status of fully retrofitted (level 3 – complete construction). Therefore, from the perspective of engineering practice, the evaluation of earthquake induced losses is informative for house owners, decision makers, insurers and etc. In this section, a multi-scale cost-benefit study is conducted to assess the cost-benefit implications of the Los Angeles Ordinance at the individual building and regional scales. The set of existing archetypes developed in Chapter 2 is adopted for the current study, however, only the LADBS Structural Design Guideline retrofit procedure is considered. For each existing archetype, retrofits are developed based on two site seismicity groups. At the individual building scale, comparative loss assessments based on the FEMA P-58 methodology are performed on the existing and retrofitted archetypes. The cost-effectiveness of the individual building retrofits is assessed based on intensity-based and expected annual losses.

At the regional scale, the cost-benefit of the ordinance-mandated retrofits for the approximately 12,000 affected buildings are first assessed for a Mw 7.1 Puente Hills event (Chen et al. 2016). Lastly, a first-of-its-kind stochastic-event-set-based cost-benefit analysis is performed using 8,313 events from the Uniform California Earthquake Rupture Forecast Version 2 (UCERF2) model (Field et al. 2009).

The retrofits for the current study are based on the procedure described in the LADBS Structural Design Guidelines (LADBS, 2015). Recall Figure 5.2, to consider the variation in design forces based on building location, the existing archetypes are retrofitted using two S_{MS} values and the corresponding seismic design parameters. More specifically, the maximum of the 0-50th and 50th-100th percentile values (2.2g and 2.96g) are used for the retrofit designs. The corresponding S_{M1} values are 2.2 g and 3.0 g, respectively. All retrofit designs are based on site class D, Risk Category II and an importance factor $I = 1.0$ in accordance with the LADBS guidelines (LADBS 2015). All other retrofit design settings remain the same as discussed in section 4.5. For all archetypes, the retrofit element consists of a 15'-0" one-bay steel *OMF* that is placed at the center of the SWOF wall line (recall Figure 5.3). The member sizes used for all retrofits (a total of 64 corresponding to 2 retrofit cases for each of the 32 archetypes) are summarized in Appendix D of the dissertation.

5.5.1. Nonlinear Structure Response Simulation

Three-dimensional numerical models were created for all existing and retrofitted buildings (a total of 96 models). The modeling procedure follows the details introduced in Chapter 2. To access the earthquake induced losses, FEMA P-58 methodology is used as the basis for the cost-benefit analysis of the ordinance-mandated SWOF retrofits. The process requires probabilistic quantification of engineering demand parameters ($EDPs$) and the demolition and collapse limit state. These are obtained by performing bi-directional nonlinear response history analyses using

the 22 pairs of far-field ground motions specified in FEMA P695 (FEMA 2009). For each structural model, incremental dynamic analyses (IDAs) are performed using increments corresponding to 10% of the first mode spectral acceleration level at the maximum considered earthquake (MCE) ($Sa_{T_1,MCE}$). A single scale factor is applied such that the median spectra for the ground motion set matches the target intensity. Additionally, the directions of the record-pairs are switched such that 44 analysis cases are conducted at each intensity. Collapse is defined as the condition where dynamic instability occurs (zero or near-zero slope of the IDA plot), or the maximum story drift ratio exceeds 10% (FEMA 2000). At each intensity level, the maximum story drift ratio in all stories and the peak acceleration at all floor levels in both directions are recorded. These EDPs are needed to compute the non-collapse losses in the FEMA P-58 methodology. The maximum residual drift over all stories are also recorded and used to consider demolition losses. The median collapse intensity is calculated by assuming a lognormal distribution and minimizing the squared loss in the empirical data. A single dispersion value of 0.6 is used for all existing and retrofitted archetypes, which includes both record-to-record variation and model parameter uncertainty (FEMA 2009). The spectral shape factor is computed for the individual archetypes based on their fundamental period and the period-based ductility obtained from nonlinear static analyses (FEMA 2009).

Figure 5.10 shows the ratio between the median of the maximum non-collapse and non-demolition story drift ratio (SDR) (considering the 1st story and longer dimension direction) for the retrofitted (SDR_{retrof}) and existing (SDR_{exist}) archetypes. These SDR_{retrof}/SDR_{exist} values provide some initial insight into the effect of the SWOF retrofit on losses associated with repairing drift-sensitive non-structural components. An intensity level corresponding to 25% of the design basis earthquake (DBE) is used to ensure that the analysis results are not dominated by either collapse or demolition

responses. Figure 5.10a and Figure 5.10b show results for the retrofit cases based on the median and maximum S_{MS} (considering all sites in the study region), respectively. Most SDR_{retrof}/SDR_{exist} values shown in Figure 5.10 are less than 1.0, which means that, on average, the SWOF retrofits are able to reduce the peak non-collapse, non-demolition story drift ratios. The average reduction is 11.2% for the median S_{MS} retrofits and 5.6% for the maximum S_{MS} retrofits. Figure 5.10 also shows that, compared to those with 2 stories, SDR_{retrof}/SDR_{exist} is consistently higher for the 3-story archetypes. It is also observed that the archetypes with the L2 and L1 1st story wall layout derive slightly greater benefit from the SWOF retrofit (in terms of non-collapse, non-demolition story drift ratios). Note that there are several cases with greater than 1 ratio, which suggests that after being retrofitted, the building has greater non-collapse non-demolition SDR. In these situations, the residual drift threshold for demolition is exceeded for several ground motions (which are not excluded) in the existing building but not in the retrofitted case, which can lead to higher non-demolition SDRs in the latter.

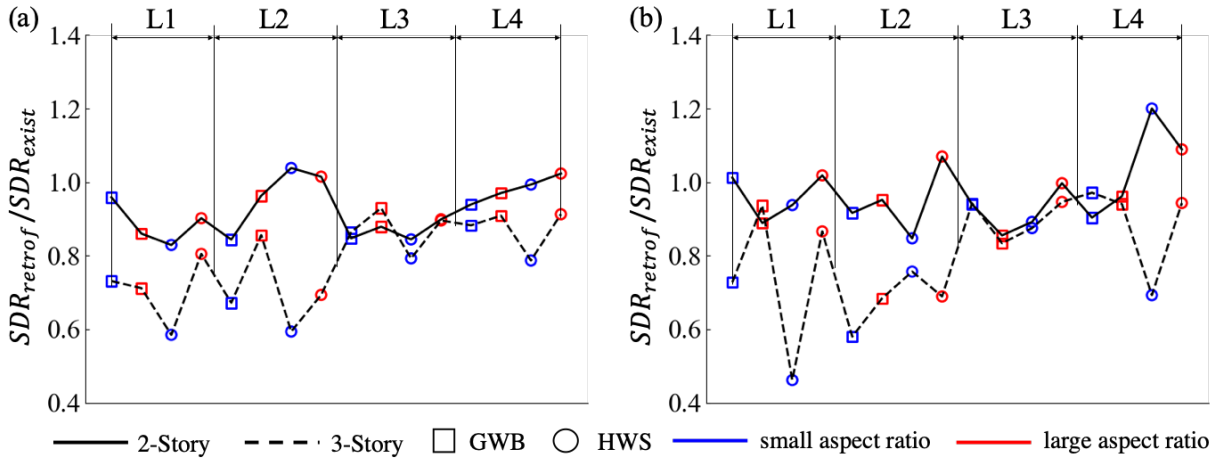


Figure 5.10. Ratio of the maximum story drift ratio between the retrofitted and existing SWOF buildings: (a) median S_{MS} and (b) maximum S_{MS} retrofit

Figure 5.11 compares the ratio between the peak non-collapse, non-demolition story drift ratio in the 2nd and 1st stories (SDR_{2nd}/SDR_{1st}) of the retrofitted and existing buildings i.e. $(SDR_{2nd}/SDR_{1st})_{retrof}/(SDR_{2nd}/SDR_{1st})_{exist}$. The goal here is to provide some insight into how the SWOF retrofits affect the distribution of peak story drifts (and the associated component-level repair costs) along the height. A high $(SDR_{2nd}/SDR_{1st})_{retrof}/(SDR_{2nd}/SDR_{1st})_{exist}$ value means that the retrofit reduces the concentration of drifts in the first story. This also means that, relative to the existing case, the 2nd story drift demands could be higher in the retrofitted case. For both the median and maximum S_{MS} retrofit cases, the $(SDR_{2nd}/SDR_{1st})_{retrof}/(SDR_{2nd}/SDR_{1st})_{exist}$ values are highest in the L1 and L2 archetypes. In other words, these two archetypes experience the highest level of drift "redistribution" from the 1st to the 2nd story. As discussed later in the paper, this observation has implications to where damage to drift-sensitive non-structural components is concentrated.

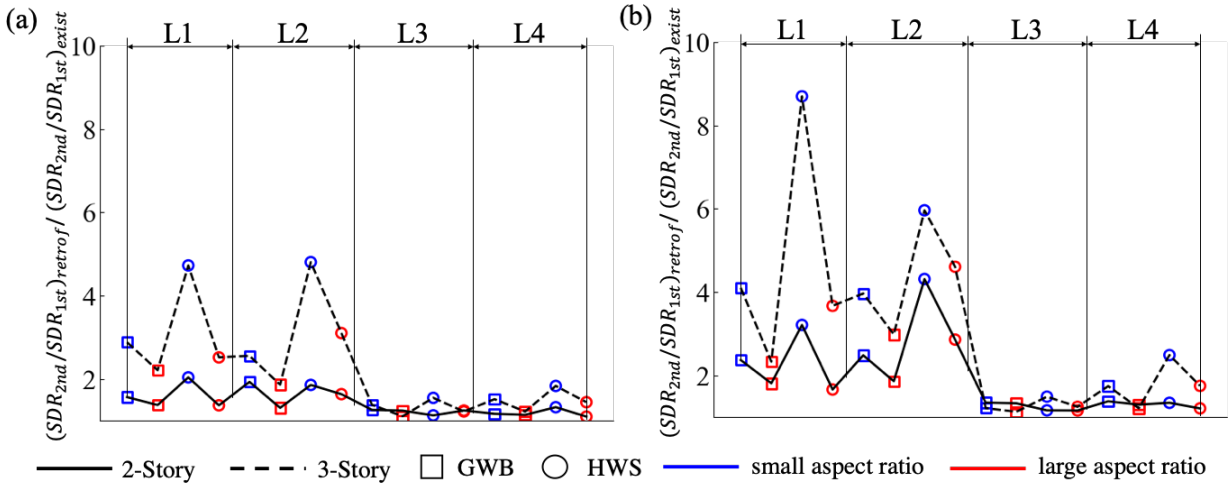


Figure 5.11. Ratio of 2nd to 1st story maximum drift ratio between the retrofitted and existing SWOF buildings: (a) median S_{MS} based retrofit and (b) maximum S_{MS} based retrofit

The ratio between the median non-collapse, non-demolition peak floor acceleration (PFA) for the retrofitted (PFA_{retrof}) and existing (PFA_{exist})archetypes is shown in Figure 5.12. $PFA_{retrof}/$

PFA_{exist} is observed to be greater than 1.0 for almost all cases, indicating that the SWOF retrofit has an overall stiffening effect on the structure, which increases the peak floor accelerations. The impact is highest for the 3-story archetype with the L1 and L2 1st story wall layout. It is not surprising that PFA_{retrof}/PFA_{exist} is 50% higher in the maximum S_{MS} retrofit cases (compared to the median S_{MS} retrofit cases) for the L1 and L2 archetypes because the designs produce stronger and stiffer moment frames. These results provide some initial insight into the effect of the SWOF retrofit on damage and losses associated with acceleration-sensitive components.

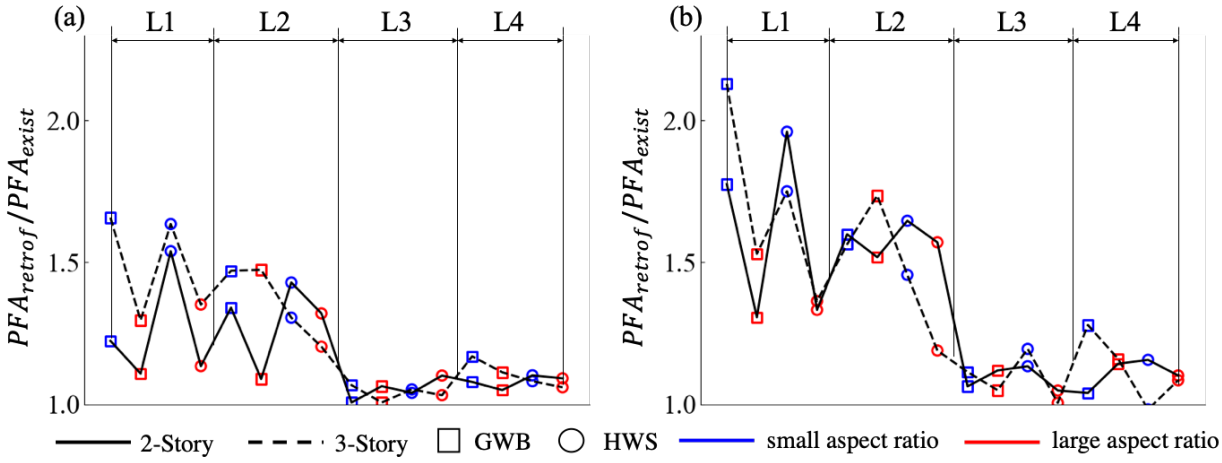


Figure 5.12. Ratio of peak floor acceleration between retrofitted and existing SWOF buildings: (a) median S_{MS} and (b) maximum S_{MS} retrofit

The effect of the SWOF retrofits on peak residual story drifts (RDR) serves as a preliminary indication of the implication to demolition losses. To this end, Figure 5.13 shows the ratio of non-collapse $RDRs$ between the retrofitted (RDR_{retrof}) and existing (RDR_{exist}) buildings. The analyses used to generate these results are also performed at the 25% DBE hazard level. The RDR_{retrof}/RDR_{exist} values for all cases is less than 1.0, indicating that the SWOF retrofit reduces the non-collapse residual drift demands. Unlike SDR and PFA, the number of stories and design intensity (median versus maximum S_{MS}) have less of an influence on how the SWOF retrofit

affects residual drift demands. However, consistent with the other two EDPs, the archetypes with the L1 and L2 1st story wall layout benefit the most from the retrofit in terms of reduced residual drifts.

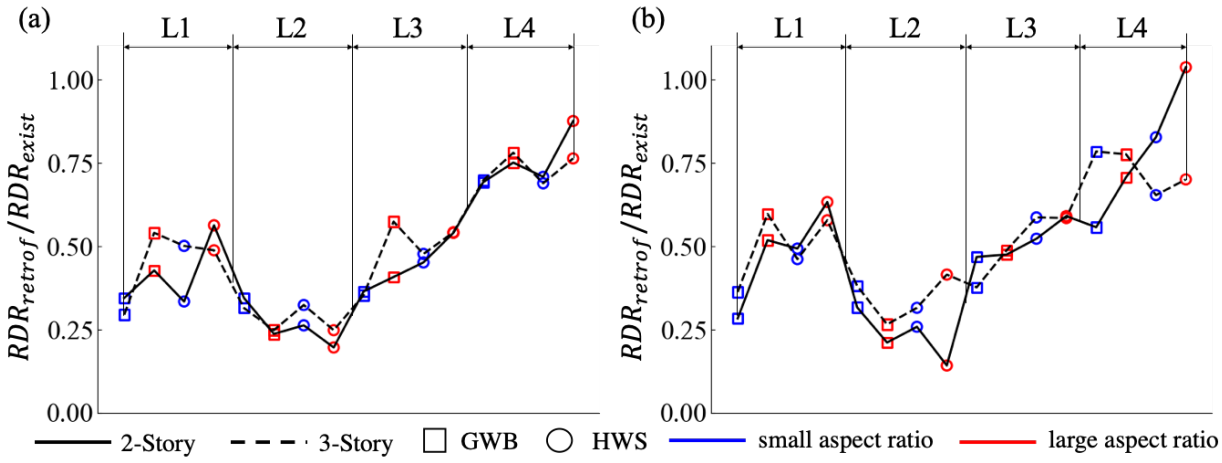


Figure 5.13. Ratio of maximum residual drift between retrofitted and existing SWOF buildings: (a) median S_{MS} and (b) maximum S_{MS} retrofit

The effect of the SWOF retrofit on collapse performance is summarized in Figure 5.14, which shows the collapse margin ratio (CMR) between the retrofitted and existing (CMR_{retrof}/CMR_{exist}) archetypes. The CMR is defined as the ratio between median collapse intensity and $Sa_{T1,MCE}$. This metric is chosen because it informs the relative improvement given different seismic design parameters. It's interesting that the average CMR_{retrof}/CMR_{exist} value for retrofits based the median (47%) and maximum (51%) S_{MS} are comparable. In other words, the additional strength and stiffness provided in the maximum S_{MS} retrofit does not provide significant additional benefit in terms of collapse safety. The reason is that, at some threshold of strength added to the 1st story, the collapse mechanism moves to the 2nd story and providing additional 1st strength beyond this upper limit does not improve the collapse safety. In fact, 6 out of the 32 archetypes retrofitted based on the maximum S_{MS} have slightly lower CMR_{retrof}/CMR_{exist}

values than the median S_{MS} retrofitted cases. Similar to the EDPs, the archetypes with the L1 and L2 1st story wall layouts benefit the most from the SWOF retrofit. Both have average CMR_{retrof}/CMR_{exist} values of approximately 1.7 compared to 1.3 and 1.2 for the L3 and L4 retrofits, respectively. Compared to the 2-story buildings, CMR_{retrof}/CMR_{exist} is 18% and 10% higher in the 3-story archetypes for the median and maximum S_{MS} retrofits, respectively. Because of the higher ductility, CMR_{retrof}/CMR_{exist} is approximately 20% higher for the archetypes with HWS panels compared to those with GWB. The aspect ratio has a minimal effect of how much the SWOF retrofit affects collapse safety.

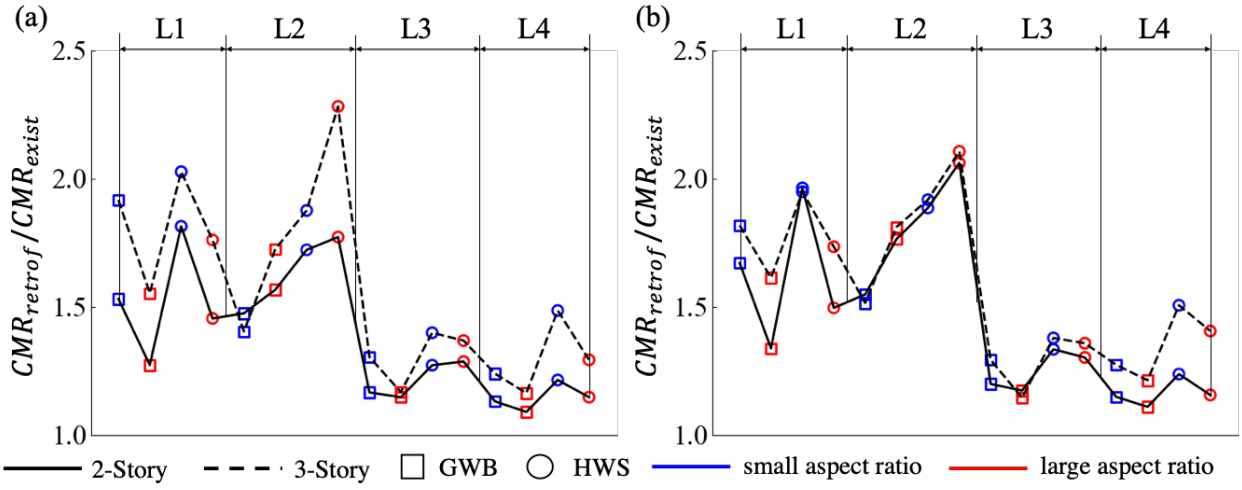


Figure 5.14. Ratio of collapse margin ratio between retrofitted and existing SWOF buildings: (a) median S_{MS} and (b) maximum S_{MS} retrofit

5.5.2. FEMA P-58 Loss Assessment

The FEMA P-58 methodology, which integrates seismic hazard assessment, nonlinear structural response simulation, damage evaluation and quantification of decision-metrics, is used to obtain the economic losses for the existing and retrofitted archetypes. The mean annual frequency of exceedance of specific ground motion intensity levels (or hazard curves) is obtained from probabilistic seismic hazard analysis applied to the site of interest. As described in Section 4.7.1,

$EDPs$ are generated from NRHAs. Component level fragility curves are used to link these $EDPs$ to physical damage and statistical loss functions are used to describe the relationship between physical damage and repair or replacement costs. Within the context of the FEMA P-58 methodology, the cost of collapse, demolition and component-level repairs are probabilistically combined using an expected value calculation.

The process outlined in the Chapter 2 produces a building-level loss function, which describes the expected losses conditioned on the ground motion intensity level. Figure 5.15 and Figure 5.16 present the loss function for the L1-2S-60x30-GWB and L4-2S-100x50-GWB archetypes, which experience the most and least reduction in expected annual losses after being retrofitted, respectively. The losses are disaggregated based on the contributions from collapse, demolition and component repairs. Using a $S_a$value of 0.5g (34% and 25% DBE for median and maximum S_{MS} retrofits, respectively) as the intensity-based point of comparison, the reduction in total losses for the L1-2S-60x30-GWB archetype is approximately 70% for both the median and maximum S_{MS} retrofits. At the same intensity, the losses in the L4-2S-100x50-GWB archetype are reduced by 5% and 17% for median and maximum S_{MS} retrofits, respectively. For both buildings, the two retrofit cases (based on the median and maximum S_{MS}) result in an overall reduction in the collapse and demolition losses and an increase in the losses due to component repairs. Figure 5.15 shows that the collapse loss at $S_a = 0.5g$ is reduced by 73% and 82% for the median and maximum S_{MS} retrofits, respectively. At the same intensity level, the demolition losses decrease by 89% and 86%, while the component-level repair costs increase by 53% and 57%, respectively. The reduction in the collapse and demolition losses is consistent with the results shown in Figure 5.15 and Figure 5.16. Two factors explain the increase in component-level repair costs after retrofit. First, there is an increase in damage to acceleration sensitive components, which is also reflected in the PFA

comparison shown in Figure 5.12. Second, the first story retrofit increases the drift demands in the upper stories, where there are more damageable components (relative to the more open first story). Consequently, at lower intensity levels where component repair dominates the total losses, the retrofitted buildings experience a higher total loss than the existing buildings.

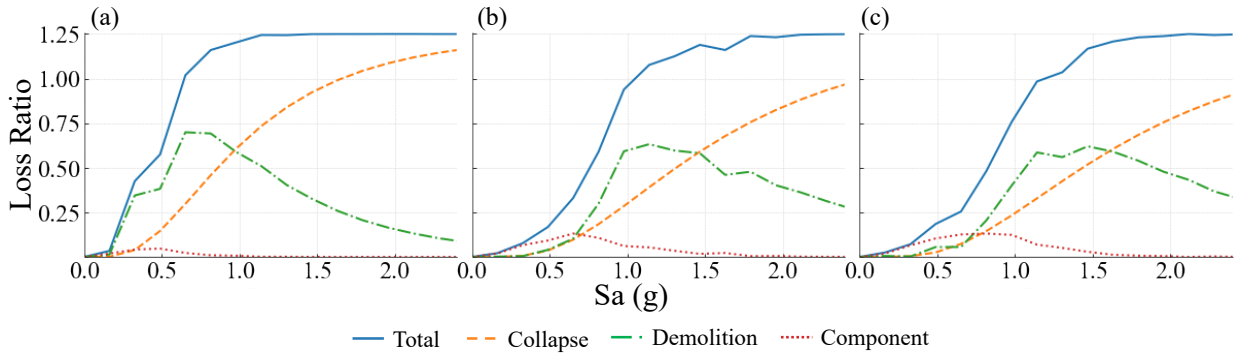


Figure 5.15. L1-2S-60x30-GWB intensity-based loss curve: (a) existing building, (b) median S_{MS} and (c) maximum S_{MS} retrofit

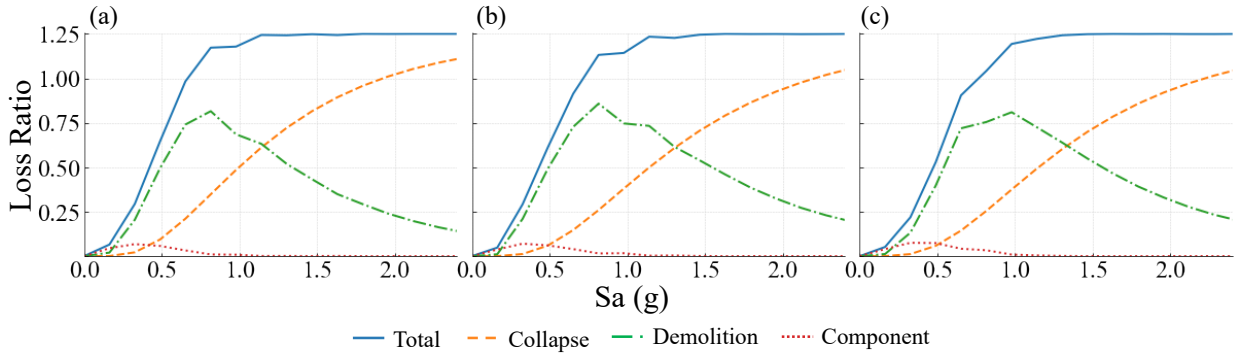


Figure 5.16. L4-2S-100x50-GWB intensity-based loss curve: (a) existing building, (b) median S_{MS} and (c) maximum S_{MS} retrofit

Expected annual losses (EAL) are computed by numerically integrating the loss functions with site-specific hazard curves. To highlight the building-specific EAL comparisons, hazard curves are obtained from the United States Geological Survey (USGS) (USGS 2019) for two sites corresponding to the median (34.0453, -118.456) and maximum S_{MS} (34.3237, -118.449) retrofits. The total and disaggregated EAL for the L4-2S-100x50-GWB and L1-2S-60x30-GWB archetypes

are shown in Figure 5.17 and Figure 5.18, respectively. The SWOF retrofit has a small but measurable effect on the EAL for L4-2S-100x50-GWB, with reductions of 18% and 6.5% for the median and maximum S_{MS} retrofits, respectively. Whereas the total EAL is reduced by approximately 60% for the two retrofit cases of the L1-2S-60x30-GWB archetype. Most of this reduction is attributed to the demolition losses, which decreases by almost one half after retrofit (both cases). Consistent with earlier observations, the $EALs$ attributed to component repairs are higher for the retrofitted cases.

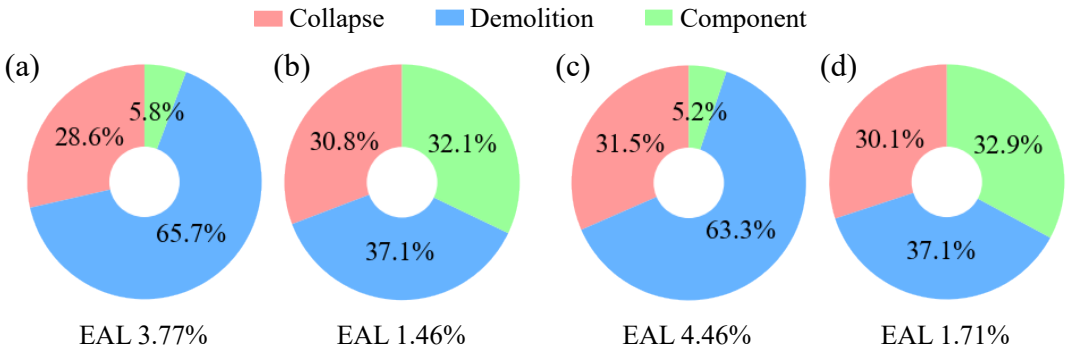


Figure 5.17. Disaggregated EAL for L1-2S-60x30-GWB archetype: (a) median S_{MS} existing, (b) median S_{MS} retrofit, (c) maximum S_{MS} existing and (d) maximum S_{MS} retrofit

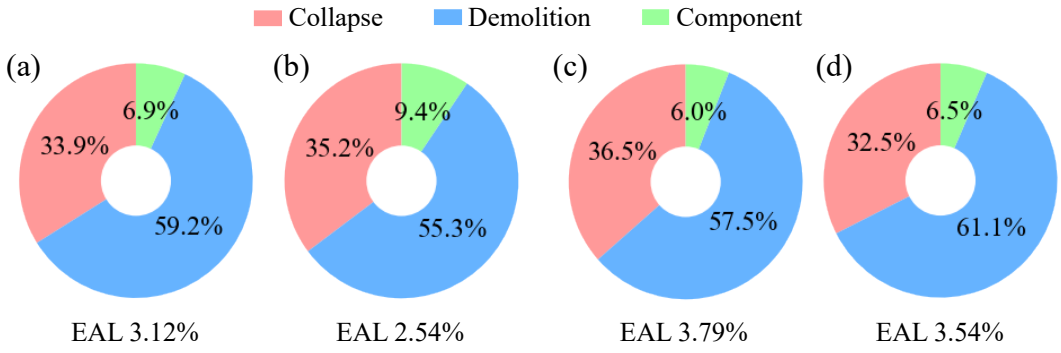


Figure 5.18. Disaggregated EAL for L4-2S-100x50-GWB archetype: (a) median S_{MS} existing, (b) median S_{MS} retrofit, (c) maximum S_{MS} existing and (d) maximum S_{MS} retrofit

5.5.3. Cost-Benefit Analysis

5.5.3.1. Introduction and Assumptions

The cost-benefit analysis of the SWOF retrofit is presented in this section. While the main focus is on the regional scale, building-specific results for the best and worst case (in terms of cost-benefit) are also presented. The cost of the retrofit is determined based on information provided by general contractors (Soft Story Retrofit Pros 2019). Generally, this cost primarily depends on the number and type of retrofit elements, which, in this study, comprise of ordinary steel moment frames and the associated strip footing foundation. The average retrofit cost per unit ranges from approximately $4,000 to $6,250 and is proportional to total number of units (Soft Story Retrofit Pros 2019). In other words, the per unit retrofit cost is generally lower for buildings with more units. To consider inherent uncertainties, the per unit retrofit cost is assumed to take on a uniform distribution that is centered at the average per unit cost (conditioned on the number of units) with a range that is equal to the average cost plus/minus $400. The number of units associated with each archetype is determined by the plan layout specified in Appendix A.

Two types of earthquake induced losses are considered: direct losses due to repair or replacement of the damaged building and the cost of fatalities. Fatality costs are determined by assigning a dollar value to human life. Previous studies have adopted values ranging from $2 to $5 million per human life (Mitrani-Reiser 2007). The current study initially adopts a lower-bound value of $2 million for each human life and the sensitivity of the cost-benefit ratio to this assumption is examined. The fatality assessment is performed using the SP3 tool, which implements the FEMA-P58 methodology. For both the building-specific and regional-scale assessments, the cost-benefit ratio is computed as the retrofit cost divided by the reduction in losses due to the retrofit. The only difference is that, for the latter, the retrofit cost and loss-reduction for the individual buildings are

aggregated for the entire portfolio before computing the regional cost-benefit ratio. The value of the cost-benefit ratio can be interpreted as follows: (i) a negative cost-benefit ratio means that the loss associated with the retrofitted building is greater than that of the existing one i.e. the denominator is negative (least desirable), (ii) a positive cost-benefit ratio that is less than 1.0 means that the cost of the retrofit is less than the reduced losses (most desirable), (ii) a positive cost-benefit ratio that is greater than 1.0 means that the cost of the retrofit is higher than the reduced losses. Besides the cost-benefit ratio, the effectiveness of the retrofit can be assessed based on the break-event time, which is obtained by solving for T in Equation (5.1) (Padgett et al. 2010)

$$\sum_{t=0}^{T} C_{exist,t} = C_{retrof} + \sum_{t=0}^{T} C_{retrof,t} \tag{5.1}$$

Where $\sum_{t=0}^{T} C_{exist,t}$ and $\sum_{t=0}^{T} C_{retrof,t}$ are the accumulated annual costs for the existing and retrofitted buildings, respectively, over a period T and C_{retrof} is the cost of the retrofit. The break-even time evaluates the duration needed to recoup the cost of the investment with lower values being more desirable. Unlike the cost-benefit ratio, which considers event-based or ground motion intensity-based losses, the break-even time considers expected annual losses (i.e. the probabilistic combination of all shaking intensities). However, like the event-based or intensity-based losses, it considers the cost of earthquake-induced damage and fatalities. Another more tangible interpretation of the break-even time can be obtained if there is a reduction in the cost of earthquake insurance if the building is retrofitted. In that case, the break-even time would consider the time it takes for annual premium discount plus EAL reduction to equal the upfront cost of the retrofit. A 3% annual discount rate is considered when computing the break-even time (Lee and Ellingwood 2015).

5.5.3.2. Building-Specific Cost-Benefit Analysis

First, a building-specific cost-benefit assessment is performed for the L4-2S-100x50-GWB and L1-2S-60x30-GWB archetypes and the median and maximum S_{MS} retrofit designs. Considering the losses due to earthquake damage and fatalities, the cost-benefit ratios corresponding to the DBE and MCE events are shown in Figure 5.19. Among the cases that are shown, only the median S_{MS} retrofit for the L1-2S-60x30-GWB evaluated at the DBE level has a cost-benefit ratio that is less than 1.0. A general observation is that the cost-effectiveness of the retrofit decreases as the ground motion intensity increases from the DBE to the MCE hazard level where collapse and demolition losses dominate for both the retrofit and existing cases.

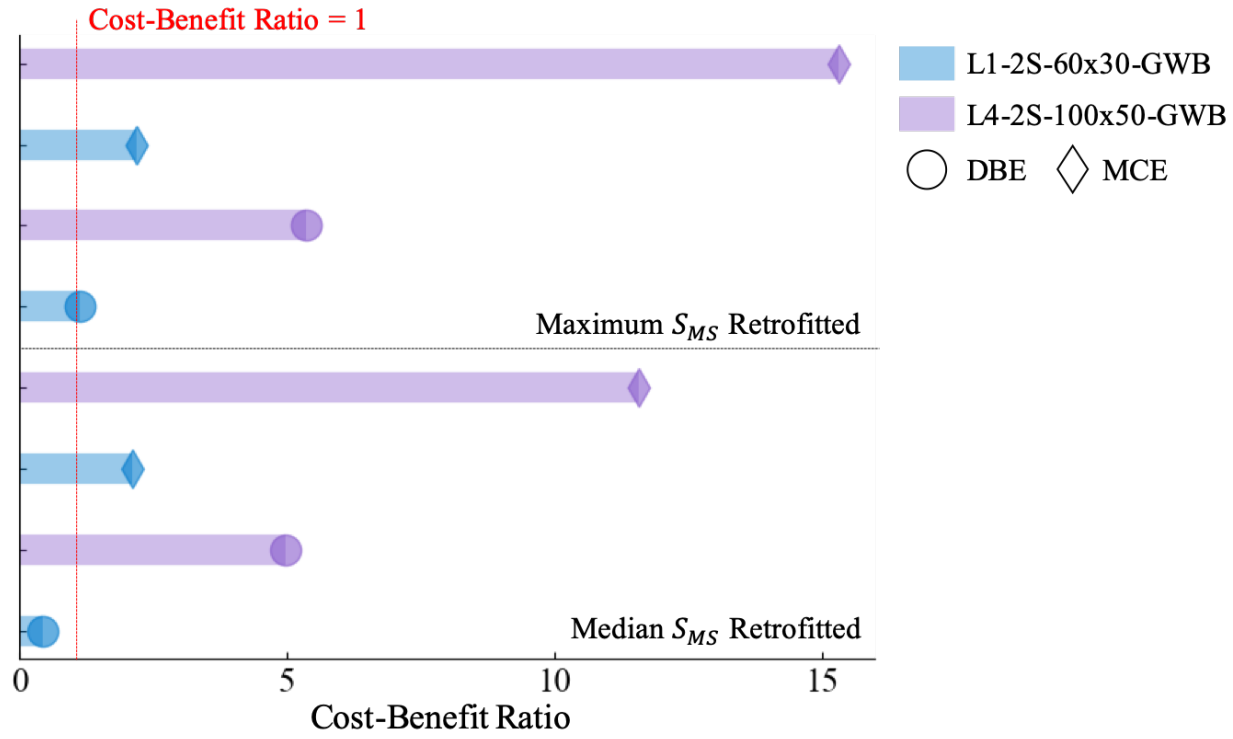


Figure 5.19. Cost-benefit ratios for the L1-2S-60x30-GWB and L4-2S-100x50-GWB archetypes at the DBE and MCE shaking intensities

Using Equation (5.1), the break-even time is computed for the L4-2S-100x50-GWB and L1-2S-60x30-GWB archetypes and the median and maximum S_{MS} retrofit designs. The cost

considerations include the expected annual earthquake losses (building damage and fatalities), the retrofit investment and the annual cost of earthquake insurance premiums. In other words, the break-even analysis is performed with and without considering earthquake insurance. The expected annual losses are obtained directly from SP3. Estimating earthquake insurance premiums requires a detailed assessment that considers the building age, configuration, contents, seismic vulnerabilities and site hazard. However, in the current study, upper and lower bound values of $800 and $5,000, respectively are used for the annual premium cost (Andrew Blankstein and Monica Alba 2014). The California Earthquake Authority offers insurance premium discounts ranging from 10% to 25% for retrofitted multi-family homes (California Earhtquake Authority 2019). In the current study, a 10% reduction in the earthquake insurance premium is assumed for the retrofitted buildings. For the cases where earthquake insurance is considered, the premium is included in the annual cost.

Figure 5.20 through Figure 5.23 show the yearly accumulated costs in US dollars for L1-2S-60x30-GWB and L4-2S-100x50-GWB. The considered cases include combinations of the existing and retrofitted (median and maximum S_{MS}) buildings, with and without considering earthquake insurance for annual premiums of $800 and $5,000. The accumulated costs are disaggregated based on the earthquake losses and insurance premiums (where considered). Figure 5.20 and Figure 5.22 show that, for the L1-2S-60x30-GWB archetype, the break-even time is between one and two years with and without earthquake insurance. In contrast, Figure 5.21 and Figure 5.23 show that the accumulated cost of the retrofitted L4-2S-100x50-GWB archetype (including the upfront retrofit cost) is higher than the existing one during the first ten years. Also, the break-even time for the maximum S_{MS} design is approximately 40 years for both insured and uninsured buildings. It is noteworthy that, at the $5000 level and considering a 40-year period, the

accumulated annual insurance premium is still less than 20% of the total accumulated cost. On average, the break-even time for the Ordinance retrofit considering all archetypes is between 4 to 5 years. Also, consistent with earlier results, the L1 and L2 archetypes have the shortest break-even times compared to L3 and L4.

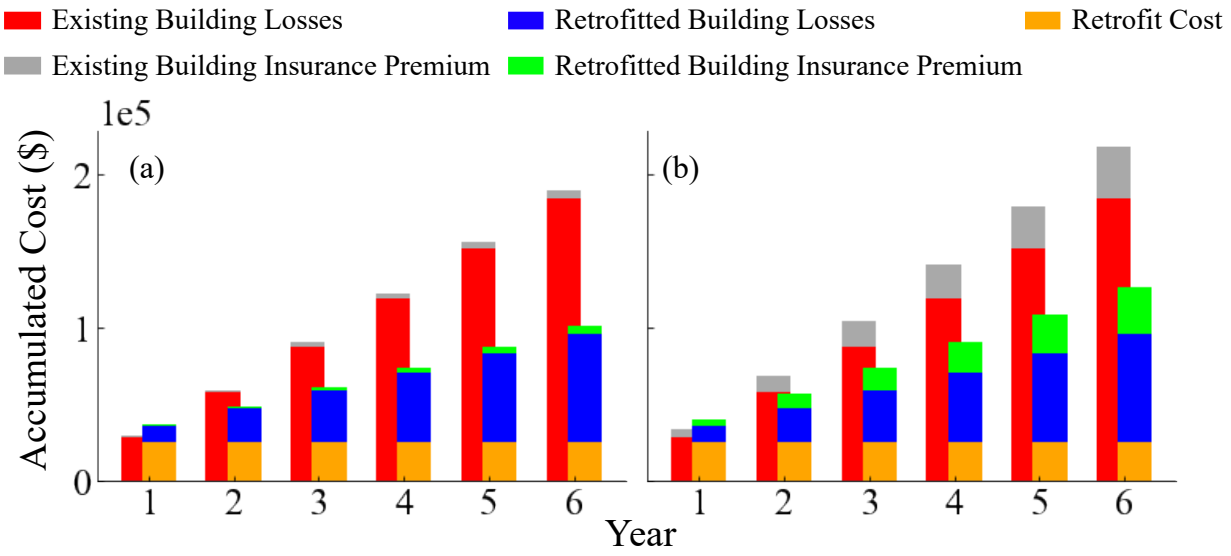


Figure 5.20. Accumulated cost versus time for the L1-2S-60x30-GWB archetype and median S_{MS} retrofit: (a) with 800$ earthquake insurance premium and (b) with 5,000$ earthquake insurance premium

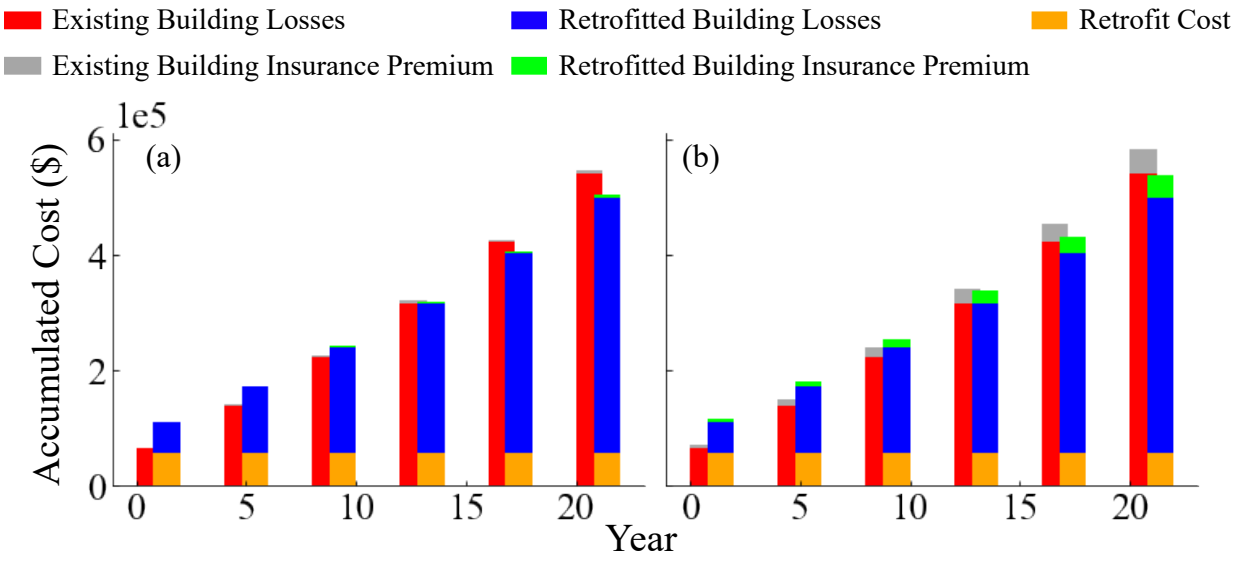


Figure 5.21. Accumulated cost versus time for the L4-2S-100x50-GWB archetype and median S_{MS} retrofit: (a) with 800$ earthquake insurance premium and (b) with 5,000$ earthquake insurance premium

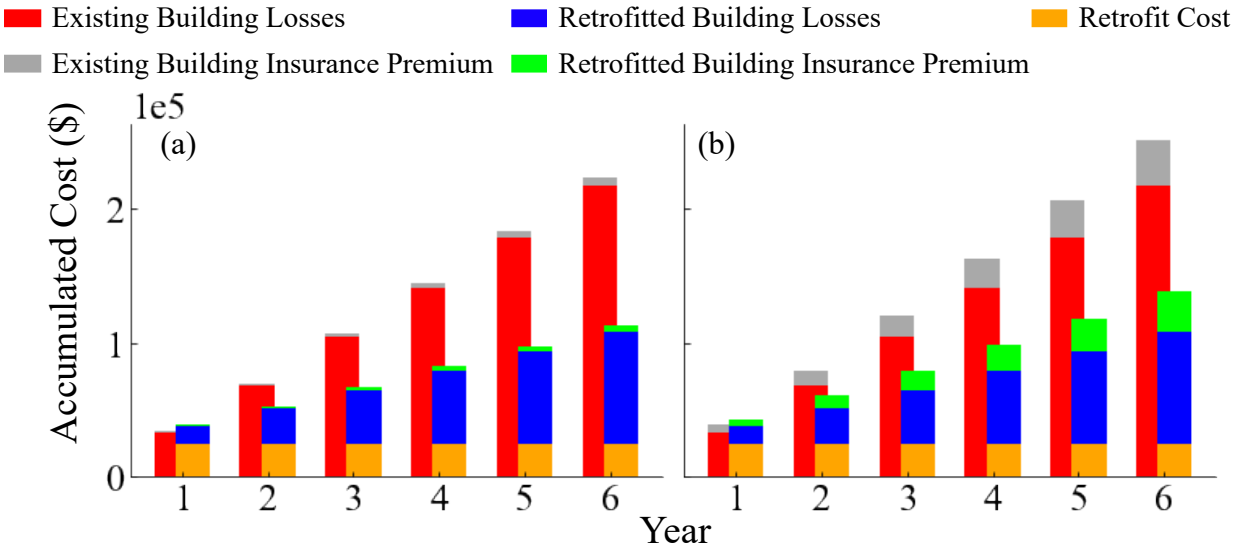


Figure 5.22. Accumulated cost versus time for the L1-2S-60x30-GWB archetype and maximum S_{MS} retrofit: (a) with 800$ earthquake insurance premium and (b) with 5,000$ earthquake insurance premium

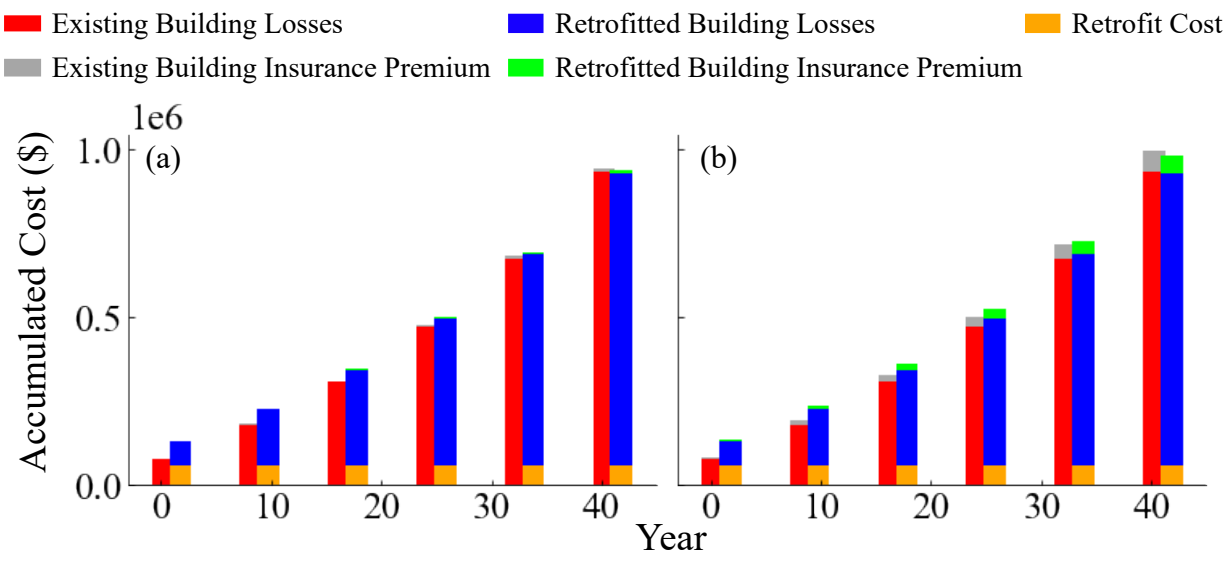


Figure 5.23. Accumulated cost versus time for the L4-2S-100x50-GWB archetype and maximum S_{MS} retrofit: (a) with 800$ earthquake insurance premium and (b) with 5,000$ earthquake insurance premium

5.5.3.3. Regional-Scale Single-Scenario Assessment

The cost-benefit analysis is performed at the regional scale considering the approximately 12,000 SWOF buildings in the City of Los Angeles. The M7.1 Puente Hills event is used for this assessment (Chen et al. 2011). Figure 5.24 shows the spatial distribution of median spectral acceleration at a period of 0.2 seconds ($Sa_{0.2s}$) for this scenario obtained from the Scenario ShakeMap Calculator application in *OpenSHA* (Field et al. 2003). The latitude and longitude of the earthquake epicenter, the boundary of the study region and several earthquake rupture parameters (e.g. rupture type, fault surface, magnitude) are the *OpenSHA* inputs. The median $Sa_{0.2s}$ map and residuals (inter-event and intra-event) are based on the Boore and Atkinson ground motion prediction equation (Boore and Atkinson 2008). The Jayaram and Baker model (Jayaram and Baker 2009) is used to generate the spatially correlated shaking intensities. As the epicenter is located in East Los Angeles, the magnitude of the median simulated shaking intensities decays from the east to west.

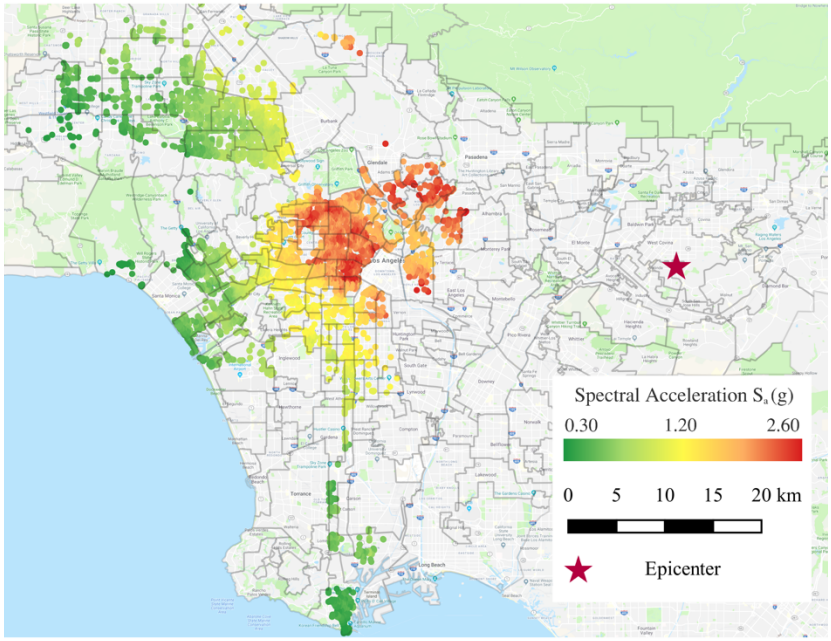


Figure 5.24. Spatial distribution of median $Sa_{0.2s}$ values for the Mw 7.1 Puente Hills Scenario

Given the shaking intensity at each site, the earthquake-induced losses (building damage and fatalities) are obtained from the building specific loss functions described earlier (e.g. Figure 5.15 and Figure 5.16). Fifty shake maps are generated for the considered event and the cost-benefit ratio for each building is computed as the retrofit cost divided by the reduced losses over all events. In addition to the spatial distribution of building-specific cost-benefit ratios, the value for the entire portfolio is computed as the total retrofit cost divided by the total reduced earthquake losses. The mean portfolio-based cost-benefit ratio is 0.32, which implies that, under the M7.1 Puente Hills event, on average, the aggregated cost of the SWOF retrofit is only 32% of the reduced building damage and fatality losses. The spatial distribution and a histogram of the building-specific cost-benefit ratios are presented in Figure 5.25. The histogram shows that 94% of the buildings in the portfolio have cost-benefit ratios between 0.0 and 1.0. Note that the shaking intensities from the Puente Hills event are generally lower than the DBE intensity for the two S_{MS} categories. This explains the much better cost-benefit ratio obtained for the regional scenario-based assessment

compared to the building-specific assessments (Figure 5.19). The cases with greater than 1.0 cost-benefit ratios are all associated with the L4-2S-100x50-GWB, L4-2S-100x50-HWS and L4-2S-60x50-GWB archetypes. Among these, only one is the L4-2S-60x50-GWB archetype, and 41% and 59% are associated with the L4-2S-100x50-GWB and L4-2S-100x50-HWS archetypes, respectively. It is also important to note that, considering all L4-2S-100x50-GWB and L4-2S-100x50-HWS archetypes, 36% and 53.4%, respectively, have greater than 1.0 cost-benefit ratios. These three cases correspond to the archetypes with the least improvement in collapse safety and reduction in expected annual losses after being retrofitted. Figure 5.25 (a) shows the spatial distribution of cost-benefit ratios under the scenario event. Comparing Figure 5.25 (a) and Figure 5.24, the distribution of cost-benefit ratios does not follow that of the shaking intensities. In the high shaking intensity areas of East Los Angeles, many buildings have cost-benefit ratios between 0.0 and 1.0. Also, greater than 1.0 cost-benefit ratios can be found in West and South Los Angeles.

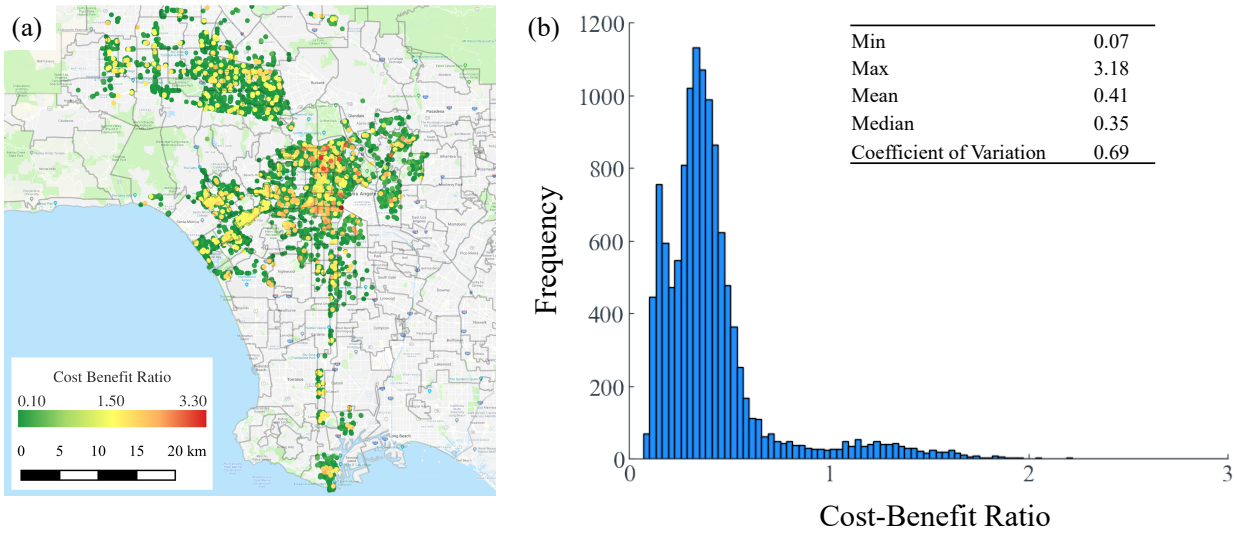


Figure 5.25. Results for Mw 7.1 Puente Hills scenario: (a) spatial distribution and (b) histogram of building-specific cost-benefit ratio

Recall that the archetypes are randomly placed at the various sites of the building portfolio with probabilities corresponding to their percentage of the inventory (obtained from the survey of

approximately 25% of the portfolio). To evaluate the sensitivity of the portfolio cost-benefit ratio to this randomized placement, a Monte Carlo simulation-based assessment is performed. For each realization, a randomized archetype distribution is generated (maintaining the same probability of each archetype) and the portfolio-cost-benefit ratio is computed. Using 100 realizations, a coefficient of variation of less than 1% is obtained for the portfolio cost-benefit ratio, which indicates that the randomized placement has a minimal effect. The sensitivity of the portfolio cost-benefit ratio to the assumed loss of $2 million per fatality is also examined. The portfolio cost-benefit ratio corresponding to fatality costs ranging from $2 to $5 million is 0.33 to 0.30, and the relationship is shown to be approximately to be linear. The retrofit cost is expected to have a direct influence on the cost-benefit ratio. Therefore, the sensitivity of the cost-benefit ratio to the average per unit retrofit cost is investigated. As noted earlier, the per unit retrofit cost is assigned a uniform distribution that is centered at $4,000 to $6,250 (depending on the number of units). Also, the range corresponds to the center-value plus and minus $400. The cost-benefit ratios are recomputed after applying factors of 0.5 and 2.0 to the center values while maintaining the $400 on both sides of the mean. For these two factors, the corresponding mean regional cost-benefit ratios range from 0.16 to 0.65, respectively and the relationship is approximately linear. This shows that the cost-benefit ratio is very sensitive to the retrofit cost. However, the mean cost-benefit ratio for the portfolio is still less than 1.0 even when the default average per unit retrofit cost is doubled.

5.5.3.4. Regional-Scale Stochastic Event Assessment

In the previous section, the cost-benefit analysis was performed for a single scenario event. To consider a range of possible events and their associated rates of occurrence, a stochastic event set-based cost-benefit analysis is performed at the regional scale. To this end, the *OpenSHA* Intensity Measure (IM) Event Set Calculator application (Field et al. 2003), the UCERF2 source model

(Field et al. 2009, p. 2) and the Boore and Atkinson (Boore and Atkinson 2008) ground motion prediction equation, are used to consider all the seismic events that are significant to the study region. The catalogue of events includes more than 8,000 rupture scenarios generated by the IM Event Set Calculator application of *OpenSHA*. Stochastic ground motion maps can be used to calculate the exceedance rate for a quantity of interest (Miller 2014), including the ground motion intensity hazard maps or other metrics more closely related to building seismic performance. The rate of exceedance for a target quantity or performance measure (X) is computed using Equations (5.2) and (5.3).

$$\lambda_{X \geq x} = \sum_{i=1}^{N_E} \lambda_i \times P(X \geq x|i) \tag{5.2}$$

$$P(X \geq x|i) = E\left[I[X \geq x|i]\right] = \sum_{X} I[X \geq x|i] p_{X|i}(x|i) \tag{5.3}$$

The summation in Equation (5.2) is over all earthquake scenarios in the catalogue of events, λ_i is the rate of exceedance of event i and $P(X \geq x|i)$ is the probability of exceeding the quantity of interest X given the occurrence of event i. $P(X \geq x|i)$ is calculated using Equation (5.3) where $I[X > x|i]$ is an indicator function that yields 1.0 for $X \geq x$ and zero otherwise. $p_{X|i}(x|i)$ represents the probability mass function of the quantity of interest X given the occurrence of event i. The expectation is taken over all ground motion maps generated for event i.

Figure 5.26 shows the annual rate of exceedance for cost-benefit ratios. The grey lines are cost-benefit ratios for individual buildings at different locations and the black line is for the portfolio. When calculating the indicator function value in Equation (5.3) for the portfolio cost-benefit ratio, the total retrofit cost and reduced losses for the entire inventory is adopted. The annual rate of exceedance for the portfolio cost-benefit ratio between 0 and 1 is 0.048 (computed as the difference

between the rates corresponding to the two values). It is observed that lower cost-benefit ratios have higher annual rates of exceedance or lower return periods.

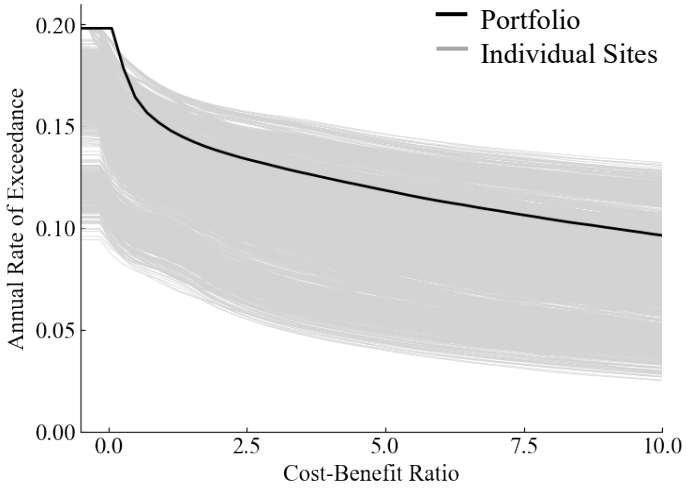


Figure 5.26. Annual rate of exceedance for cost-benefit ratios

A more interpretable result from the stochastic event set assessment can be obtained by examining the probability of observing desirable portfolio cost-benefit ratios (between 0 and 1.0) over a specified duration. To do so, the event corresponding to achieving a portfolio cost-benefit ratio between 0 and 1.0 is assumed to follow a Poisson Distribution. Therefore, the probability that this even occurs at least once in a period of t can be computed using Equation (5.4)

$$Prob(T < t) = 1 - e^{-\lambda t} \tag{5.4}$$

For 10-, 20- and 50-year periods, the probability that the portfolio cost-benefit ratio is between 0 and 1.0 for the Los Angeles Ordinance retrofit is 0.38, 0.62 and 0.91, respectively. As expected, the probability of a desirable cost-benefit ratio increases with the length of the considered period. In other words, the effectiveness of the Ordinance retrofit at the portfolio scale increases with the post-retrofit service life of the building stock.

5.6. Summary and Key-Takeaways

Given the vulnerability of woodframe buildings (soft, weak, open-front wall lines in multifamily houses or single-family houses with cripple wall) observed in the past earthquakes, policies are enacted often to reduce the earthquake risk. The primary focus of this Chapter is the performance improvement, including strength/drift capacity, collapse safety, earthquake induced losses and retrofit cost-effectiveness, brought by retrofitting. Started in 2015, Los Angeles Ordinance that mandates the retrofit of woodframe residential buildings with soft, weak and open-front (SWOF) wall lines was enacted by the City of Los Angeles to reduce their collapse risk. The study aims to quantify retrofitted building seismic performances from aforementioned perspectives. The Los Angeles Department of Building and Safety (LADBS) developed a prescriptive method for retrofit- ting soft-story buildings in accordance with the Ordinance (herein referred to as the Basic Ordinance method), which only addresses the SWOF wall lines. Besides, the LADBS guidelines also permit the use of three additional alternative full-story retrofit procedures based on Appendix A4 of the 2012 International Existing Building Code (IEBC A4) and the ASCE 41-13 and FEMA P807 documents.

Firstly, a comparative assessment of the improvement in collapse safety provided by the four alternative retrofit methods was conducted. Incremental Dynamic Analyses (IDAs) to collapse using bi-directional ground motions was performed on three-dimensional structural models of a set of archetype buildings that capture variations in the salient building properties of Los Angeles SWOF buildings. The archetypes were developed based on a survey of real SWOF buildings under the purview of the Ordinance, which has been discussed in detail in Chapter 2. The effect of the (a) number of stories, (b) first (soft) story wall layout, (c) type of interior wall sheathing and (d)

building size in plan, on the improvement in collapse safety provided by the four retrofit methods was also assessed.

The important observations from the comparative assessment are documented here. Overall, the retrofit strength added to the first story was found to increase both the drift capacity (defined by the roof drift ratio corresponding to a 20% loss of post-peak strength) and collapse safety. However, the benefit obtained from increased first story strength varied significantly between 2- and 3-story archetypes. For instance, the IEBC A4 retrofit is the only one among the four methods requiring the lateral resistance provided by walls without wood structural panels (all existing walls) to be ignored and therefore results in the greatest increase in first story strength of the SWOF buildings. As a result, the greatest benefit to the 2-story SWOF buildings, both in terms of drift capacity and enhanced collapse safety, was achieved by the IEBC A4 retrofit. To put this into further perspective, on aver- age, the IEBC retrofit provided a 78% increase in the median collapse capacity compared to 39–44% for the other three retrofit methods. Relative to the 2-story cases, the 3-story SWOF buildings derived much less benefit from the IEBC A4 retrofit where the average increase in the median collapse capacity was 50%. The value of the same metric ranged from 48 to 51% for the other three methods. While it was not the main focus of the study, it is worth noting that, for both the existing and retrofitted cases, the 2-story buildings always had higher collapse capacities than their 3-story counterparts.

The number of SWOF wall lines in a building had a significant effect on the overall improvement in collapse safety provided by the retrofits. The average increase in the median collapse capacity was found to be 68% and 36% for buildings with three and one open (partially or completely) wall line(s) respectively. When the increase in first story strength caused the collapse mechanism to shift to the second story (no collapses were observed in the third story), the retrofitted archetypes

with the more ductile interior panels performed better. As a result, the buildings with horizontal wood sheathing (HWS) interior panels benefitted more from the SWOF retrofits compared to those with gypsum wall board (GWB). More specifically, the average increase in median collapse capacity provided by the four retrofit methods was 38% and 64% for the GWB and HWS archetypes, respectively. The size of the building in plan did not have a significant effect on the improvement in collapse safety provided by the retrofits.

Later in this Chapter, a comprehensive loss assessment was conducted within the framework of FEMA P-58 for the existing and retrofitted archetypes. Only Basic Ordinance retrofit is considered in this study. Using the LADBS design guideline, two sets of ordinary moment frame (*OMF*) retrofits were implemented for each SWOF archetype based on two site seismicity categories based on the median and maximum S_{MS} values for all sites associated with the inventory. Several key observations were made when examining the intensity-based loss curves for each archetype. Under high shaking intensities (e.g. greater than the design basis earthquake level), both the existing and retrofitted archetypes are dominated by near complete (100%) losses so there is very little difference between the two. Whereas, for very low shaking intensities, component-level repair costs dominate and was often higher in the retrofitted cases. Under intermediate level of shaking intensities, the retrofit provides significant reduction in collapse and demolition losses, which lead to high reductions in the total loss. Expected annual losses (EAL) were computed by integrating the intensity-based loss curve over site-specific hazard curves. The archetypes that benefitted the most and least from the retrofit had 60% and 6% EAL reductions, respectively. After disaggregation, it was determined that the reduction in the demolition losses played the most significant role in lowering the total EAL. Consistent with the intensity-based losses curve, the retrofit caused a slight overall increase in the EAL associated with component repairs.

In the last part of the discussion, the cost-effectiveness of the retrofit is assessed at the individual building (intensity- and time-based) and portfolio scale (single scenario and stochastic event-set) based on the reduction in the building damage and fatality related losses. More specifically, the cost-benefit ratio is computed as the ratio between the retrofit cost and the reduction in earthquake-induced losses where positive values between 0.0 and 1.0 are most desirable. For the building-specific intensity-based assessment, the retrofit is generally not cost effective for higher intensities (e.g. higher than design basis earthquake level). The reason is that, as noted earlier, at these intensities, both the existing and retrofitted are dominated by complete (100% of replacement cost) losses. The concept of break-even time was used to evaluate the duration required for accumulated annual reduction in earthquake-induced losses to cover the initial retrofit costs. Across all archetypes, it takes an average of 4 to 5 years to recoup the initial investment of the retrofit. The range is from approximately two years for the archetype that benefitted the most from the retrofit (L1 first story layout) to as much as forty years for the one that derives the least benefit. Retrofitted buildings that are insured and receive a reduction in premiums require a shorter time to recoup the initial cost of the retrofit.

Some of the key findings from the cost-benefit analysis are summarized as following. A regional (or portfolio) scale cost-benefit analysis was first performed by only considering the hypothetical Mw 7.1 Puente Hills event. For this single scenario, the average portfolio cost-benefit ratio was 0.32. This number implies that the reduction in earthquake-induced losses that result from the Ordinance retrofit is approximately three times the total retrofit cost. A stochastic-event-set cost benefit analysis was also performed at the regional scale, where all events that are significant for the region and their associated rates of occurrence are considered. The annual rate of occurrence for the event that the portfolio cost-benefit ratio is between 0 and 1 is 0.048. More insightfully, for

10-, 20- and 50-year periods, the probability that the Ordinance retrofit benefit is greater than the cost is 0.38, 0.62 and 0.91, respectively. Sensitivity analyses were performed to evaluate the implications of the assumed values (based on the best available information) for several key variables. It was determined that the portfolio-scale cost-benefit ratio is not very sensitive to the assumed per person fatality cost. Variations in the assumed retrofit cost (on a per unit basis) significantly affected the portfolio cost-benefit ratio. However, it is important to note that, even when the baseline retrofit cost was doubled, the average cost-benefit ratio for the Puente Hills scenario was still less than 1.0.

CHAPTER 6. DEVELOPMENT AND IMPLEMENTATION OF AN EFFECTIVE AND EFFICIENT METHODOLOGY FOR PORTFOLIO-SCALE SEISMIC RETROFIT USING MACHINE LEARNING AND STOCHASTIC OPTIMIZATION

6.1. Introduction

6.1.1. Building Seismic Retrofit Policies

Numerous past earthquakes have highlighted the seismic risk to older existing buildings both in terms of human lives and the financial well-being of communities. Across different construction types (concrete, steel, masonry and wood), there are buildings with known seismic vulnerabilities. One mechanism that has been used to mitigate the risk to buildings with known seismic deficiencies is to enact policies that mandate their retrofit. Several California cities including Los Angeles, Santa Monica, West Hollywood, San Francisco and Oakland, have enacted ordinances that mandate the retrofit of soft-weak and open-front (SWOF) wall line multi-family residential buildings. The earthquake brace and bolt program (https://www.earthquakebracebolt.com/) provides up to $3,000 to offset the cost of retrofitting single-family woodframe buildings with unbraced and/or unbolted cripple walls. In 2014, the city of Portland launched a mandatory retrofit plan for unreinforced masonry structures (https://www.portlandoregon.gov/bds/article/588418). Guidelines such as the ones provided in FEMA P807 (FEMA 2012a) and FEMA P50 (FEMA 2012b) have sought to develop systematic seismic retrofit and performance evaluation procedures for soft/weak-story structures. ASCE 41 (ASCE 2017) and IEBC (IEBC 2012) are often recommended as permissible alternative (to the local ordinance approach) guidelines for retrofitting a range of structure types and vulnerabilities including soft/weak and non-ductile buildings. The general design philosophy used in many retrofit methods (e.g. local ordinance procedures or national standards) is based on achieving some building-specific performance target.

Retrofit components are designed to supply the building with the necessary strength and ductility to achieve the desired performance objective. Most procedures utilize a strength check at the component level, and seismic performance factors (e.g. R factors in IEBC and m factors in ASCE 41) are used to account for nonlinear behavior. Additionally, inelastic drift limits are either implicitly or explicitly stipulated as a design constraint. To limit the financial burden that will be placed on building owners, there is often a need to develop ordinance-mandated retrofit solutions that are both efficient and effective. Because of differences in the structural configuration and hazard exposure, achieving both effectiveness (in terms of target performance) and efficiency (in terms of cost) in portfolio-targeted retrofits is very challenging.

Several prior studies have examined the performance implications of seismic retrofits to SWOF woodframe buildings. Buckalew et al. (2015) compared the performance of three- and four-story San Francisco buildings retrofitted using the FEMA P807, ASCE 41 and IEBC procedures. Three retrofit designs were proposed based on each specification, and pushover and incremental dynamic analysis (IDA) were conducted to quantify the strength and performance improvement. The IEBC and ASCE 41 designs provided the building with significantly higher strength than the FEMA P807 design. From the results of the IDA, it was observed that adding too much strength in the soft/weak first story could lead to collapse in the upper story. Maison et al. (2014) performed an independent review of the FEMA P807 guidelines using a four-story soft-story building. The authors concluded that the P807 procedure can provide efficient retrofit design solutions but raised questions about the accuracy of building performance prediction module. Burton et al. (2018) performed a comparative assessment of the collapse performance of SWOF woodframe buildings designed using alternative procedures, including the LA Ordinance, FEMA P807, IEBC, and ASCE 41. Detailed finite element models were created for 32 representative archetypes, and

nonlinear static and IDA analyses were conducted to examine and compare the collapse safety improvements provided by different retrofit alternatives. It was concluded, on average, the IEBC provides the greatest increase (78%) in the median collapse capacity. The number of SWOF wall lines was found to have the most significant impact on the retrofit benefits. The same authors conducted a follow-up study to evaluate the regional scale cost effectiveness of the LA Ordinance retrofit (Yi et al. 2020). Cost-benefit ratios were computed at the individual building and regional scales. For the latter, both a single event and stochastic earthquake catalogue was considered. Under a hypothetical M7.1 event, the Ordinance retrofit achieved a cost-benefit ratio of 0.32. Based on the results of the stochastic event assessment, the probability that the Ordinance retrofit produces an effective cost-benefit ratio (between 0 and 1) within a 50-year period was computed to be 0.9.

6.1.2. Seismic Retrofit Optimization

There have been a small number of studies on the performance-based optimization of seismic design and retrofit of new and existing buildings, respectively. Pei and van de Lindt (2009) developed a loss-based design optimization framework for light frame wood structures. Conventional force-based design procedures were embedded in an iterative binary search until there is a match between the resulting and target loss curves. The procedure was used to optimize the design of a single-family house to match two loss-based performance targets. Dong et al. (2014) proposed a bridge network retrofit optimization framework to mitigate economic impacts. A bridge network in Orange County was selected to illustrate the framework. The main goal was to find the best retrofit timeline for the bridges in the network that achieves the maximum reduction in the expected economic loss while minimizing the upfront cost of the retrofit. A genetic algorithm was implemented to obtain the optimal retrofit plan. Jennings (2015) developed a multi-objective

seismic retrofit optimization framework for a portfolio of woodframe buildings. The goal was to optimize for community resilience as measured by several socioeconomic performance indices. The objective function was defined based on a weighted combination of the retrofit cost, earthquake-induced economic loss, the number of fatalities and the recovery time. A set of 37 archetypes buildings was developed based on 5 plan configurations, 5 design codes and 2 retrofit methods. A genetic algorithm was used to determine the optimal allocation of funding to maximize the resilience of the neighborhood. Note that while the Dong et al. and Jennings studies focused on regional scale seismic performance, neither addressed the optimization of the final structural design.

6.1.3. Objective and Scope

This chapter proposes a framework for determining the optimal building seismic retrofit designs to achieve the most desirable performance at the portfolio scale. The overall methodology comprises distinct but connected modules for performance prediction, optimization and evaluation of the final solution. The prediction module utilizes machine learning-based surrogate models to provide a compact but accurate link between the building seismic retrofit parameters and building-specific performance outcomes. The surrogate models are embedded in an optimization routine that finds the retrofit parameters that give the best overall performance for the portfolio. The effectiveness and efficiency of the optimized retrofit solution is then evaluated by benchmarking its performance against conventional ordinance-mandated procedures. The framework is demonstrated by applying it to the portfolio of SWOF woodframe buildings that are under the purview of the Los Angeles ordinance. The remainder of the chapter is organized as follows: Section 2 provides and overview of the framework including a discussion of the three modules. Section 3 applies the framework to the inventory of SWOF buildings in Los Angeles. The paper

concludes with a brief summary of the study, a discussion of the key findings and limitations and ideas for future work.

6.2. Overview of Proposed Framework

A schematic overview of the methodology for developing retrofit procedures that achieve optimal portfolio-scale seismic performance is presented in Figure 6.1. The framework is comprising of prediction, optimization and evaluation modules. The algorithm takes in a set of preliminary retrofit design parameters as input, which are linked to performance outcomes via the prediction module. The optimization algorithm then iteratively searches for the set of retrofit design parameters that achieve the most desirable value of the objective function. The retrofit design that is proposed by the optimization module is then evaluated for one or more earthquake scenarios. The outcome of the framework is the best retrofit design recommendation based on structural characteristics of the target inventory, the spatial variation of seismic hazard and any predefined constraints.

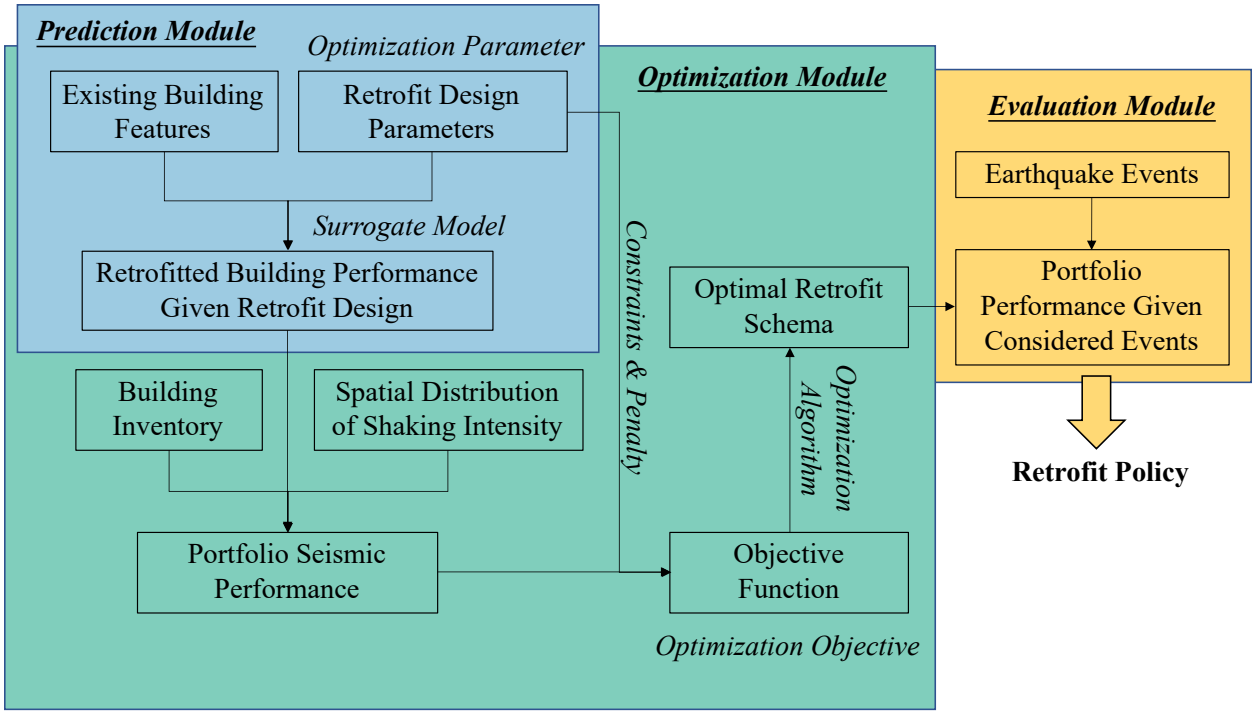


Figure 6.1 Schematic overview of the portfolio-scale seismic retrofit optimization framework

The prediction module links the performance outcome of individual retrofitted buildings to the design parameters without the explicit need for nonlinear response history analyses (NRHAs). Considering the differences in structural parameters, the spatial variation in seismic hazard, and the iterative search that is embedded in the algorithm, performing NRHAs in each stage of distributed system optimization is intractable. The introduction of a surrogate model in the prediction module significantly reduces the number of NRHAs that need to be performed to achieve an optimal solution. A performance prediction model provides a directly link between the retrofit design and the performance improvement relative to the existing building, which makes the iterative search for an optimal solution more feasible. To develop such a model, variables related to the building configuration and structural and dynamic properties, are initially considered as potential input parameters (or features). Then, an appropriate model selection procedure can be used to reduce the number of predictors. To maintain the interpretability of the surrogate model,

parameters that are familiar to designers should be used even if their contribution to predict performance of the model is limited. The chosen model response variable (or output parameter) must be consistent with the target portfolio performance metric. This facilitates computation of the regional performance metric through aggregating the individual building performances.

The optimization module, which is the core of the framework, uses an appropriate algorithm to search for the best retrofit strategy under a given spatial distribution of seismic hazard, building configuration and structural properties and the specified design constraints. A key decision in this stage of the framework is selecting the seismic hazard definition that is used as the basis for the optimal retrofit search. One approach is to use the spatial distribution of shaking intensities corresponding to a single representative scenario as the basis for finding the optimal retrofit solution. However, this event-specific approach would not be general enough for a retrofit policy that is intended to minimize the impacts associated with future possible events. Alternatively, a stochastic catalogue approach could be used where all possible events and their associated ground motion maps can be used in the optimization search. Sometimes, for large urban centers such as Los Angeles where the number of impactful events in the stochastic catalogue is on the order of tens of thousands, the computational burden could be onerous, even with the use of the surrogate models described in the previous section. To reflect a diverse and more realistic set of spatial shaking intensity distributions, a subset of representative events from the entire stochastic event is used within the optimization framework.

Besides the hazard definition, another important decision is the choice of the portfolio performance metric. One or some combination (weighted by level of importance) of metrics related to building damage, economic loss, fatalities, recovery or even environmental impacts, can be used. The selection of this (or these) performance metric(s) should be consistent with the objective of the

retrofit design. The objective function, which is used as the basis of the optimization, is then created by introducing the appropriate constraints (applied to the retrofit design parameters) and penalty terms (applied to the objective function). Note that the optimality of the final design is closely associated with how the objective function is formulated. Generally, constraints are created to place practical limits on the search space of the retrofit design. To ensure the efficiency of the retrofit scheme, a penalty function is introduced, which represents the "cost" (not necessarily in dollars) of the retrofit. An appropriate algorithm has to be chosen to serve as the search engine that finds the optimal retrofit design in the specified target space. Common choices for this purpose include gradient based (Newton Raphson Algorithm, Quasi-Newton Algorithm and etc.) and stochastic (Genetic Algorithm, Bayesian Optimization and etc.) optimization algorithms. In the context of portfolio-scale retrofit, the feasibility of a given optimization algorithm is influenced by the surrogate model used in the prediction module and the formulation of the objective function.

The evaluation module takes the optimal retrofit solution that is generated by the optimization module and assesses the portfolio performance under one or multiple events. Regional performance metrics (same or different from the one adopted in the optimization module) are recomputed using the proposed retrofit design for the scenario(s) of interest. Ultimately, this module is used to benchmark the performance of the portfolio for the optimal retrofit solution relative to the existing inventory and/or more conventional (e.g. code-based) procedures.

6.3. Application of the Methodology to the Seismic Retrofit of Soft, Weak, and Open-Front Wall Line Buildings in the City of Los Angeles

6.3.1. Background on Existing Ordinance and Its Inefficiencies

For multi-unit residential woodframe buildings in the Los Angeles metropolitan area, it is common for the first story to be used for parking or commercial spaces such that a lower wall density is

used relative to the upper stories. This can lead to substantial differences in the stiffness and strength of adjacent stories and the formation of a single-story mechanism during earthquake shaking. Numerous complete or partial woodframe building collapses have been attributed to soft-story damage in prior seismic events including the 1971 San Fernando, California (FEMA, 2012a), 1989 Loma Prieta, California (Harris and Egan 1992) and 1994 Northridge, California (Holmes and Sommers 1996) earthquakes. Several cities in California have established policies to address the seismic risk to soft-story woodframe buildings. In Los Angeles, an ordinance was passed to mandate the retrofit of the approximately 12,060 SWOF wall line buildings in the city (Xia and Schleuss, 2016).

Prior numerical analysis studies have highlighted some inefficiencies in woodframe soft story retrofits. For example, providing too much strength in the first story increases the level of damage to the upper stories (Burton et al. 2016; Buckalew et al. 2015). Additionally, Yi et al. (2020) determined that, for some configurations, a 40% increase in the strength of the lateral system provided by the retrofit elements, only achieves a 6% average increase in the median collapse capacity. In summary, at some threshold of strength difference between the 1st and 2nd story, the collapse mechanism moves to the 2nd story. Strengthening the 1st story beyond that limit significantly reduces the incremental benefit or efficiency of the retrofit.

6.3.2. Prediction Module

The primary goal of prediction module is to create a surrogate model that links important structural properties to the building-specific seismic outcomes, such that the computational expense associated with parameter variations and iterative search can be minimized. The two main stages that comprise this module are the dataset preparation and prediction model development, the details of which are summarized in Figure 6.2. In the data preparation stage, the building

configuration and structural model parameters are first sampled to create a diverse dataset. Nonlinear structural models are then constructed in the Open System of Earthquake Engineering Simulation (OpenSees) (McKenna et al. 2003) on the basis of the sampled parameters. Modal, nonlinear static and dynamic analyses are then performed. Some structural properties together with additional building information can be used as the features in the surrogate model to predict responses. In practice, the selection of features is an iterative process. One approach is to start with all available parameters and reduce the number using a feature-selection technique. The results from nonlinear analyses will be further used to assess the building seismic performances, which are the response variables that are to be estimated by the surrogate model. The generated data is divided into training and testing sets. The former is used to determine the best prediction model parameters by minimizing the differences between the prediction and actual values. Once an appropriate model is derived, parameter selection can be executed to reduce the number of features. However, this step is optional and will not be discussed in detail in this paper. Lastly, the testing set is used to reveal the model performance on an "unseen" dataset. If the model performance is acceptable under the testing set, the surrogate model can be finalized. The implementation of prediction module for the retrofitted SWOF buildings is discussed in the next section.

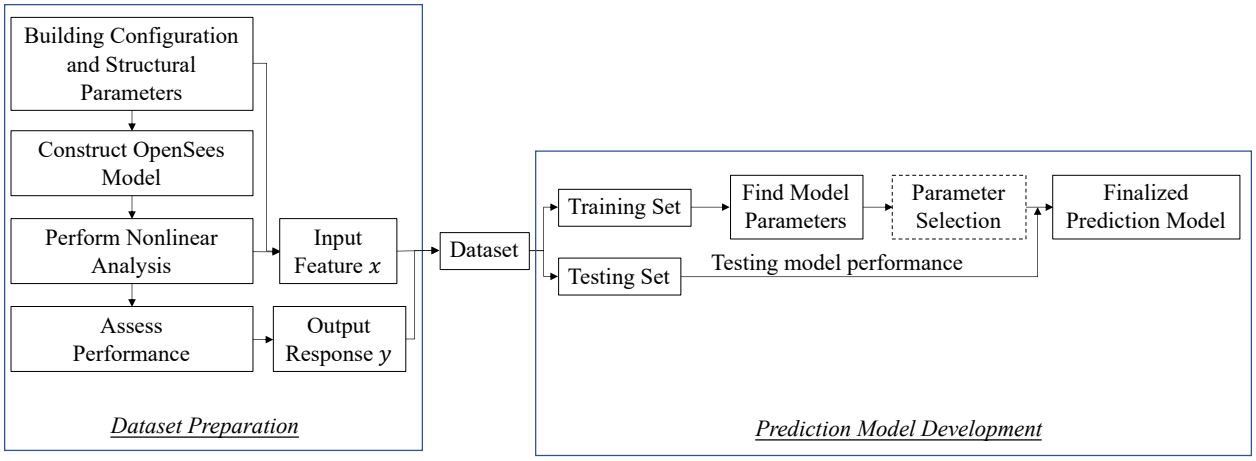


Figure 6.2. Prediction module workflow

6.3.2.1. Woodframe Building Modeling and Performance Assessment

To develop a surrogate model, a large dataset needs to be generated and used for training and testing. Considering the number of buildings in the entire portfolio, the 32 representative existing SWOF archetypes described in Chapter 2 are used. Ordinance moment frames (*OMF*) are designed following the LADBS guideline (LADBS 2015) to retrofit the existing buildings. Eight designs are established for each archetype (a total of 256 retrofitted cases) to capture the variations in the seismic design parameters. The design details are provided in Appendix E. The performance outcomes of these retrofitted buildings will be used to construct the surrogate model and for the portfolio-scale seismic evaluation. For each existing and retrofitted building, a three-dimensional structural model is constructed using the end-to-end modeling tool. Incremental dynamic analyses were conducted following the procedure introduced in Chapter 2. The median collapse intensities were obtained, and a single dispersion value of 0.6 is used for all existing and retrofitted archetypes. This value includes both record-to-record variation and model parameter uncertainty.

6.3.2.2. XGBoost Model Development and Performance Evaluation

Several machine learning algorithms are investigated as candidates for the surrogate model used to estimate the performance improvement provided by the retrofit relative to the existing building given some features. The Extreme Gradient Boosting (XGBoost) model (Chen and Guestrin, 2016), which is an ensemble learning technique, is found to be the most robust for accomplishing this task. XGBoost combines the accuracy of ensemble learning, the high efficiency and flexibility of the gradient boosting model, the suitability for parallel computing and a low risk of overfitting. The superior performance of XGBoost on the datasets generated in the current study has been determined through a preliminary comparative assessment with other types of machine learning algorithms (e.g. Kernel Ridge Regression, Gradient Boosting and Random Forest).

Several machine learning algorithms are investigated as candidates for the surrogate model used to estimate the performance improvement provided by the retrofit relative to the existing building given some features. The Extreme Gradient Boosting (XGBoost) model (Chen and Guestrin, 2016), which is an ensemble learning technique, is found to be the most robust for accomplishing this task. XGBoost combines the accuracy of ensemble learning, the high efficiency and flexibility of the gradient boosting model, the suitability for parallel computing and a low risk of overfitting. The superior performance of XGBoost on the datasets generated in the current study has been determined through a preliminary comparative assessment with other types of machine learning algorithms (e.g. Kernel Ridge Regression, Gradient Boosting and Random Forest).

The 256 retrofitted archetypes (8 designs based on 8 intensity bins for each of the 32 archetypes), their features, and the associated $\frac{CMR_{retrof}}{CMR_{exist}}$ values, are split into 70% for training the XGBoost model and 30% for testing. Additionally, 32 edge cases without retrofit $\left(\frac{V_{retrof}}{V_{exist}} = 1; \frac{CMR_{retrof}}{CMR_{exist}} =\right.$

$1\big)$ are added to the training dataset to improve the model performance near the boundary condition. The final training and testing set include 212 and 76 cases, respectively i.e. 88% for training and 12% for testing. During the training process, cross-validation with random parameter search is implemented to reduce the risk of overfitting while finding the optimal XGBoost parameters. K-fold cross-validation divides the training set into K equal-sized groups. For each iteration, a set of model hyperparameters is independently sampled from a target range and used to formulate the model. The model parameters are then trained using the $K-1$ data subsets and the validation score is computed using the remaining data. Taking the average validation score over the K validation cases as the final cross-validation score, the hyperparameter set with the highest score is selected. The Huber loss function (Huber 1992), as defined in Equation (6.1), is used to compute the cross-validation scores because it is less sensitive to outliers than MSE and also has a median-unbiased property similar to MAE.

$$l_\delta(y_i, \hat{y}_i) = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2, & \text{for } |y_i - \hat{y}_i| \le \delta \\ \delta\left(|y_i - \hat{y}_i| - \frac{1}{2}\delta\right), & \text{otherwise} \end{cases} \tag{6.1}$$

After obtaining the hyperparameter set with the best validation score, the testing dataset is fed into the model to check the performance on an "unseen" dataset. Figure 6.3 shows the predicted versus the actual $\frac{CMR_{retrof}}{CMR_{exist}}$ for both the training and testing datasets. The mean absolute relative difference ($MARD$) computed using Equation (6.2) is selected to evaluate the model performance.

$$MARD = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{6.2}$$

The $MARD$ values for the training and testing sets are 1.50% and 1.44%, respectively. This implies that the median absolute error is less than 2% of the actual values for both the training and testing

sets. Figure 6.3 (a) shows that the data points are evenly distributed around a 45-degree angled straight line, which indicates low model bias. The residuals (difference between predicted and actual value) for the testing set in Figure 6.3 (b) have normal-like distribution that is centered at zero which also indicates low model bias. Also, more than 99% residuals are within mean plus and minus one standard deviation range demonstrating the presence of very few outliers.

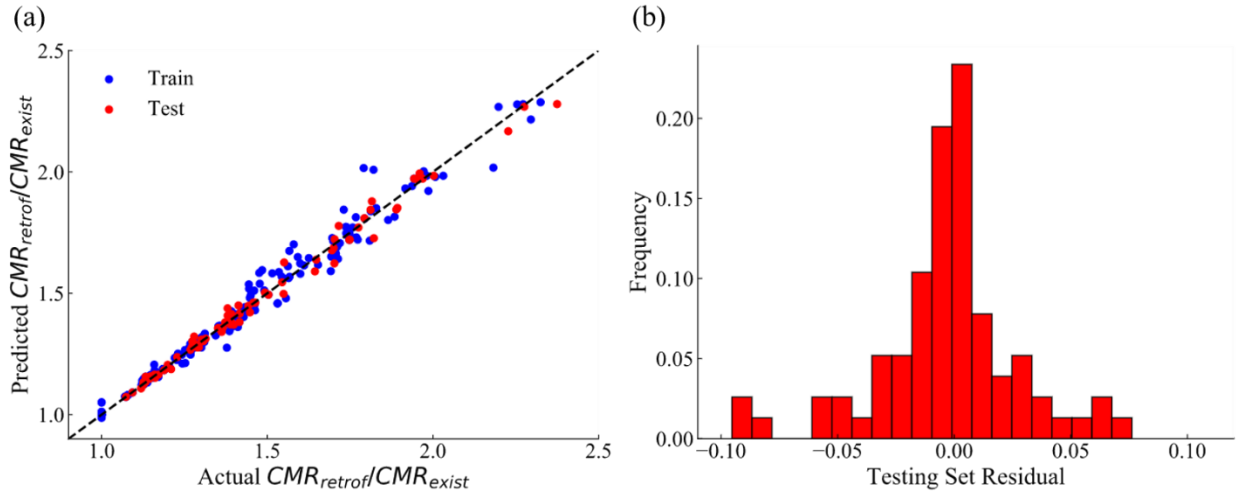


Figure 6.3. XGBoost model diagnostics: (a) predicted versus actual $\frac{CMR_{retrof}}{CMR_{exist}}$ and (b) histogram of testing set residuals

6.3.3. Optimization Module

Overall, the optimization module takes in a preliminary set of retrofit designs for the portfolio as input and iteratively adjusts them to maximize the objective function value under the specified constraints. The workflow implemented in the SWOF building retrofit optimization for the city of Los Angeles is shown in Figure 6.4. The parameter to be optimized is the additional strength provided by the retrofit normalized by the peak strength of the existing building. Facilitated by the prediction module, the median collapse intensities of the SWOF retrofitted building conditioned on the percentage of added strength can be rapidly assessed. The collapse losses for the portfolio is used as the performance metric because the ordinance is specifically targeted at reducing

collapse risk and does not focus on other metrics such as repair costs or recovery. The optimization is performed on a representative subset of the stochastic event catalogue for Los Angeles that includes 20 scenarios. An inversion function is then developed and used to penalize excessively strong retrofits. The penalty function introduces a cost-related measure into the objective function to improve the retrofit efficiency. The objective function is taken as the sum of the penalty function and portfolio collapse loss. Constraints are also introduced to define upper and lower bound retrofit designs. After setting up the objective function, Bayesian optimization is implemented to find the retrofit scheme that maximizes the objective function. The details of each step are outlined in the following subsections.

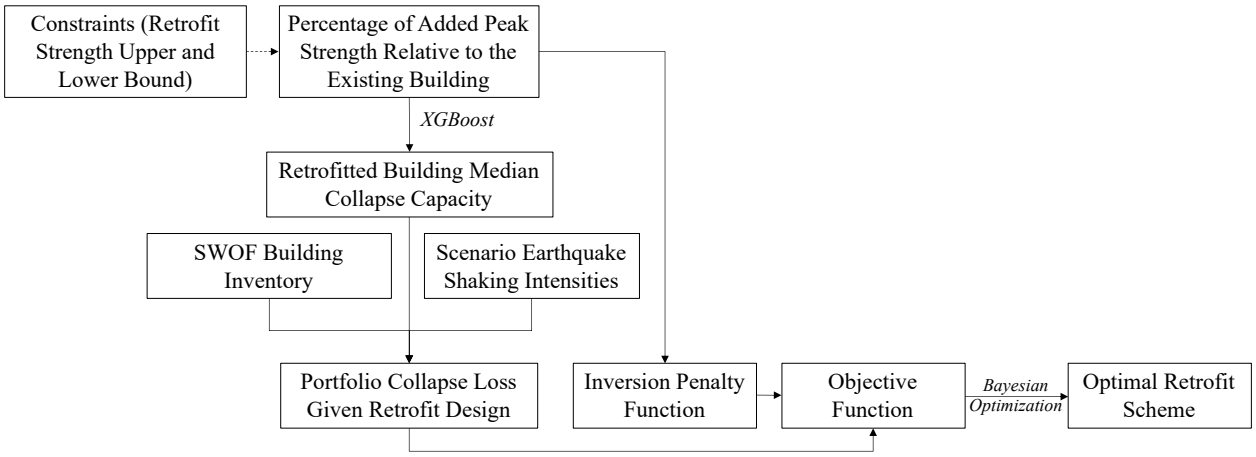


Figure 6.4. Optimization module workflow

6.3.3.1. Objective Function

6.3.3.1.1. Retrofit Design (Optimization Parameter)

As observed from the nonlinear analysis results and consistent with prior studies, the improved collapse margin ratio is highly correlated with the increase in peak strength provided by the retrofit. To better capture this phenomenon, the retrofit strategy is quantified as the percentage of peak

strength added to the existing building. Listed in Table 6.1 are the peak strength (in the percentage of seismic weight) for each existing archetype in the two orthogonal directions.

Table 6.1 Normalized (by seismic weight) peak strength of the existing archetypes

Archetype ID	Peak Strength X-Direction (% of Seismic Weight)	Peak Strength Z-Direction (% of Seismic Weight)
L1-S2-60x30-GWB	34%	35%
L1-S3-60x30-GWB	20%	21%
L1-S2-100x30-GWB	44%	25%
L1-S3-100x30-GWB	26%	15%
L1-S2-60x30-HWS	26%	25%
L1-S3-60x30-HWS	15%	15%
L1-S2-100x30-HWS	34%	17%
L1-S3-100x30-HWS	20%	10%
L2-S2-60x50-GWB	38%	33%
L2-S3-60x50-GWB	22%	20%
L2-S2-100x50-GWB	42%	21%
L2-S3-100x50-GWB	25%	12%
L2-S2-60x50-HWS	27%	19%
L2-S3-60x50-HWS	16%	11%
L2-S2-100x50-HWS	30%	12%
L2-S3-100x50-HWS	18%	7%
L3-S2-50x30-GWB	60%	34%
L3-S3-50x30-GWB	36%	20%
L3-S2-80x30-GWB	60%	36%
L3-S3-80x30-GWB	36%	21%
L3-S2-50x30-HWS	48%	28%
L3-S3-50x30-HWS	29%	17%
L3-S2-80x30-HWS	45%	22%
L3-S3-80x30-HWS	27%	13%
L4-S2-60x50-GWB	37%	48%
L4-S3-60x50-GWB	23%	29%
L4-S2-100x50-GWB	42%	33%
L4-S3-100x50-GWB	26%	20%
L4-S2-60x50-HWS	28%	29%
L4-S3-60x50-HWS	17%	18%
L4-S2-100x50-HWS	30%	19%
L4-S3-100x50-HWS	18%	12%

6.3.3.1.2. Constraints

Constraints are used to establish upper and lower bound strengths for the retrofits. As discussed earlier, the optimization parameter is parametrized as the additional percentage of peak strength provided by the retrofit. Also, at some threshold where the first story is stronger than the second story, the collapse mechanism is pushed to the second story and there is no benefit provided by additional strength. ASCE 41-17 (ASCE 2017) addresses this concern by specifying that the strength of first story should not exceed 30% of the adjacent story. This 30% is used as the upper bound for the additional peak strength added to the existing building.

It is also important to consider that the improvement provided by the retrofit varies across different shaking intensities. Under extremely high shaking intensities, the existing and retrofitted building have equally high probabilities of collapse and associated losses. This is considered by creating eight shaking intensity bins, and the same percentage of added strength is used to retrofit all sites falling into the same intensity bin. To be consistent with Ordinance retrofit, eight intensity bins ranging from 0.7g to 1.5g with 0.1g increments are created based on the spectral acceleration associated with 75% DBE (Reference LADBS Guideline). Figure 6.5 shows the spatial distribution and a histogram of the shaking intensities corresponding to 75% DBE. In summary, 8 parameters are defined to be optimized, which represent the percentage of additional peak strengths for the buildings within each of the intensity bins.

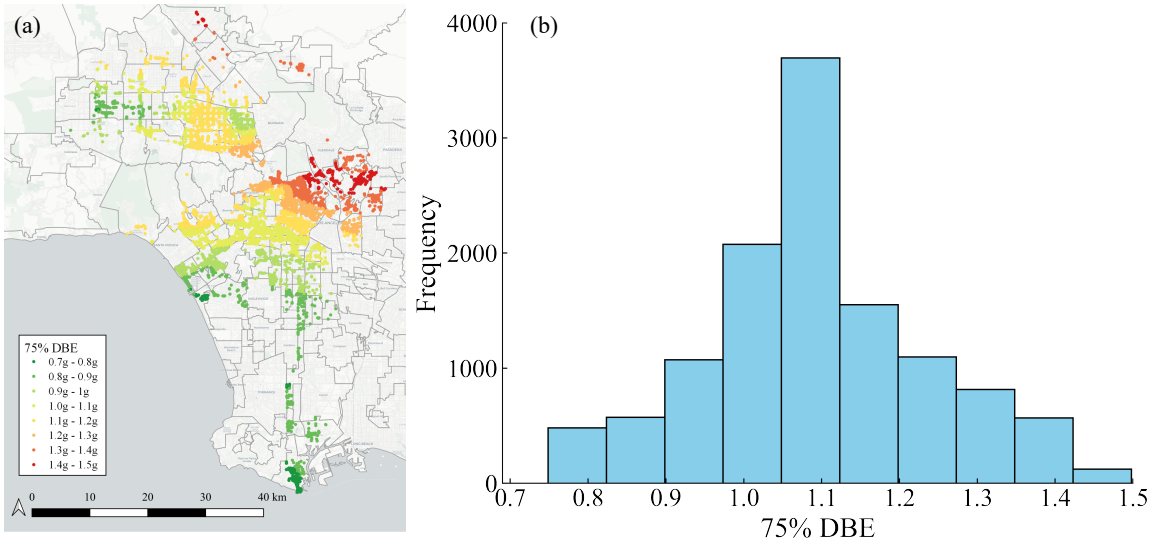


Figure 6.5. Shaking intensities corresponding to 75% DBE: (a) spatial distribution, (b) histogram

6.3.3.1.3. Portfolio Performance

The objective of the optimization is to develop a SWOF retrofit strategy that efficiently achieves the best performance for the entire building portfolio under the specified constraints. The optimization is performed for 20 representative scenarios and the total collapse loss in United States Dollars (USD) is used as the performance metric. The procedure for selecting the representative events and performance metric are discussed in this section.

The *OpenSHA'* IM Event Set Calculator application (Field et al. 2003) is used to generate ground motion maps for each of the SWOF sites located in the city of Los Angeles. At each site, all faults that have a significant contribution to the seismic hazard in the region are incorporated by using the Uniform California Earthquake Rupture Forecast 2 (UCERF2) source model (Field et al. 2009). For a given magnitude-distance pair, the Boore and Atkinson (2008) ground motion model (GMM) is used to calculate the median intensity measure (IM) along with the corresponding inter-event and intra-event standard deviations at each site. A catalogue of 8,313 UCERF2 events is generated by the *OpenSHA* Event Set Calculator. The Jayaram and Baker (2009) model is used to generate

100 shake maps with spatially correlated shaking intensities. A representative subset of 20 scenarios are selected using the K-Means algorithm, which first groups all events into clusters based on a measure of the distance between the median spectral acceleration at a period of 0.2 seconds ($Sa_{0.2s}$) at each site (Jayaram and Baker 2010). The event closest to the cluster centroid is selected as the representative event for that cluster. The return period, magnitude and source information for the 20 scenarios are listed in Table 6.2 below.

Table 6.2. Basic information for the 20 representative scenarios

Event ID	Return Period	Magnitude	Source/Fault
1	476744	7.95	Elsinore;W+GI+T+J+CM
2	62026	7.25	San Jacinto;SJV+A
3	1132780	7.25	Elsinore
4	801413	7.25	S. San Andreas
5	138001	6.65	Anacapa-Dume, alt 2
6	12427	6.55	Hollywood
7	105551	7.55	Hosgri
8	371731	6.65	Newport-Inglewood, alt 1
9	1602211	7.25	Newport Inglewood Connected alt 1
10	1635059	6.65	Newport Inglewood Connected alt 2
11	768713	7.45	Newport Inglewood Connected alt 2
12	707708	6.85	Palos Verdes Connected
13	129575	6.65	Puente Hills
14	130612	6.75	San Gabriel
15	68992	6.65	Santa Monica, alt 2
16	269044	6.65	Santa Monica Connected alt 1
17	346173	6.65	Santa Monica Connected alt 2
18	247448	7.05	Santa Monica Connected alt 2
19	125181	6.95	Sierra Madre Connected
20	27148	6.85	White Wolf

The portfolio performance is calculated by averaging the regional performance over all shake maps and events. The total financial losses due to building collapse is computed using the FEMA P-58 methodology. The expected collapse loss (L_{total}) is computed using Equation (6.3).

$$L_{total} = \frac{1}{N_{events} N_{maps}} \sum_{i=1}^{N_{events}} \sum_{j=1}^{N_{maps}} \sum_{k=1}^{N_{sites}} \mathrm{P}\left(Collapse \middle| IM_{ijk}\right) L_k \tag{6.3}$$

where, N_{events} is the number of considered events, N_{maps} is the number of shake maps for each event, N_{sites} is the number of considered building sites, $\mathrm{P}\left(Collapse \middle| IM_{ijk}\right)$ is the probability of collapse at site k under the shaking intensity, IM_{ijk}, corresponding to the jth shake map and the ith scenario, and L_k is the collapse loss corresponding to the building at site k. The collapse loss is taken as the sum of the building value and the fatality loss. The fatality loss for a single building is a function of its use purpose, dimension, the time of day when the earthquake occurs and the assigned financial losses for human life. The $L_k's$ are obtained from the Seismic Performance Prediction Program (SP3) (Haselton Baker Risk Group, 2019) and assumed to be the same for each archetype. The probability of building collapse conditioned on the shaking intensity is computed using Equation (6.4).

$$\Pr\left(collapse \middle| IM_{i_j,k}\right) = \Phi\left(\frac{ln\left(IM_{ijk}\right) - \ln(CM_k)}{\sigma}\right) \tag{6.4}$$

where, Φ is the standard normal cumulative distribution function (CDF), CM_k is the median collapse intensity at location k, which can be predicted using the XGBoost model given the building features and the increase in peak strength and σ is the lognormal standard deviation, which is set to 0.6 to account for record-to-record dispersion and model parameter uncertainties (FEMA, 2012).

6.3.3.1.4. Penalty Function

With only the above constraints, the algorithm would simply retrofit all buildings to achieve the maximum reduction in collapse losses i.e. 30% peak strength added since the portfolio performance metric selected in this case is a monotonically increasing function. However, this will

lead to an inefficient retrofit scheme from the perspective of benefit-cost ratio. To address this issue, a penalty term that is a function of the retrofit strategy, herein the percentage of peak strength, is added as part of the objective function. The penalty term can be interpreted as the cost of the retrofit such that unnecessarily high strength retrofits would be suppressed. The second order inversion function shown in Equation (6.5) is adopted.

$$P = \sum_{i=1}^{N_{bins}} \frac{w}{(u_i - x_i + \varepsilon)^2} \tag{6.5}$$

where u_i is the upper limit for the optimization target x_i and w is the penalty weight, N_{bins} is the number of shaking intensity bins. The magnitude of the penalty weight directly influences the retrofit design efficiency. The results from a sensitivity analysis (discussed later) showed that using $w = 300{,}000$ results in a penalty weight that provides high efficiency while maximizing the reduction in collapse losses. To avoid numerical conversion problems, a small number (ε) is added to the denominator.

Combining the collapse loss, the specified constraints and the penalty function, the final objective of the optimization problem is to find the strength x_k added to the existing building that maximizes Equation (6.6).

$$J = -\frac{1}{N_{events}N_{shakemaps}} \sum_{i=1}^{N_{events}} \sum_{j=1}^{N_{shakemaps}} \sum_{k=1}^{N_{sites}} \sum_{p=1}^{N_{bin}} \Phi\left(\frac{ln(IM_{ijk}) - \ln(XGBoost(x_k))}{\sigma}\right) L_k I\left(Sa_{75\%DBE_k} \in p\right) - P(x_k) \tag{6.6}$$

where, $I\left(Sa_{75\%DBE_k} \in p\right)$ is an indicator function that yields 1 if the spectral acceleration corresponding to 75% DBE at the kth site $Sa_{75\%DBE_k}$ falls in the pth intensity bin and 0 otherwise. A conceptual interpretation of the objective function is that the goal is to find the retrofit design that achieves the highest reduction in portfolio collapse losses while adding the least amount of retrofit strength.

6.3.3.2. Stochastic Optimization Algorithms

6.3.3.2.1. Bayesian Optimization

Bayesian optimization is a generic algorithm for finding the maximum or minimum value of a function $f(x)$ over the input domain X. The algorithm recursively creates a probabilistic surrogate model to link input features to function responses and recommends the next subspace to explore based on the chances that the optimal value is in that subspace. This probabilistic model is the critical element that distinguishes Bayesian optimization from other algorithms. Using the posterior information, the algorithm gives the probability distribution over the function values and quantifies the uncertainties in the function space. This property balances exploration and exploitation trade-off (Auer 2002) and avoids local optima. Also, in each step, the algorithm recommends a small number of points to evaluate the target function, which limits the total number of times the objective function is evaluated. This is especially advantageous to cases where evaluating the target function is expensive. Additionally, Bayesian optimization does not require the target function $f(x)$ to take on a specific form, which makes it suitable for non-parametric and non-convex optimization problems. In this study, the objective function formulated in Equation (5.6) is solved by the XGBoost model and is therefore non-parametric and non-differentiable, which makes Bayesian optimization a suitable algorithm.

Figure 6.6 summarizes the main steps in the Bayesian optimization algorithm. A preliminary evaluation of the target function is first conducted to construct the prior function. The prior function is a probabilistic model linking input features to function responses. The posterior distribution is then updated by the prior function and further used to create an acquisition function, which estimates the objective function on the unevaluated domain. In the next step, the algorithm then evaluates the target function at the point that maximizes the acquisition function. This

evaluation point is used as the input to calculate the objective function value, and the input-output pair updates the prior function that is used in the next step. The detailed derivation of the prior function, posterior distribution and acquisition function are discussed in the following subsections.

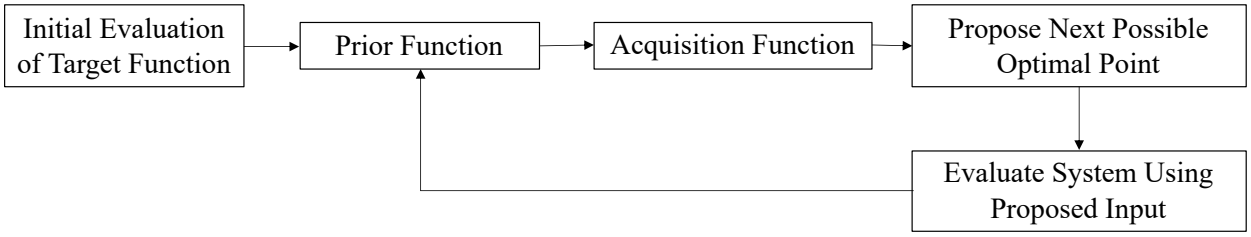


Figure 6.6. Bayesian Optimization algorithm flow chart

- Prior Function

The prior represents the probability distribution of function values in the input space given a finite set of observations. The target function values are drawn from the prior function. The Gaussian process (Rasmussen 2003) $\left(f \sim GP(\mu, \Sigma)\right)$ is the most natural and robust choice for the prior function because its conditional and marginal distribution follow the same distribution. It is assumed that the function values at any finite n pre-specified points follow an n-dimensional multivariate normal distribution as shown in Equation (6.7).

$$f(x_1), \dots, f(x_n) \sim N\left(\left(\mu(x_1), \dots, \mu(x_n)\right), \Sigma(x_1, \dots, x_n)\right) \tag{6.7}$$

where $f(x)$ is the target unknown function, $\mu(x)$ and $\Sigma(x_1, \dots, x_n)$ are the mean and covariance functions, respectively. The mean function could be either parametric or non-parametric and the covariance function reflects the deviations from the mean. One common choice for the covariance function for a pair (x, x') is the squared exponential shown in Equation (6.8). The intuition behind the exponential covariance function is that adjacent points would share close function values.

$$\Sigma(x, x') = \tau^2 \exp\left(-\frac{|x - x'|^2}{\sigma^2}\right) \tag{6.8}$$

where τ and σ are unknown hyperparameters that control the magnitude and smoothness, respectively. The choice of hyperparameters generally follows the principal of maximum a posteriori probability (MAP), the details which are provided in Gelman et al. (2014).

The prior function is constructed on the basis of a set of known values. Given the n points in the input space where function values have been revealed as vector X, the mean, covariance and probability density function can be estimated by Equations (6.9) through (6.11).

$$E\big(f(X)\big) = \mu(X) \tag{6.9}$$

$$E\left[\big(f(X) - \mu(X)\big)^T\big(f(X) - \mu(X)\big)\right] = \Sigma(X, X) \tag{6.10}$$

$$p(f(X)|X, \mu, \Sigma) = (2\pi)^{-\frac{n}{2}}|\Sigma(X, X)|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\big(f(X) - \mu(X)\big)^T\Sigma(X, X)^{-1}\big(f(X) - \mu(X)\big)\right) \tag{6.11}$$

Since the above model is estimated through finite observations of the target function, the observations may or may not incorporate noise as illustrated in Equations (6.12) and (6.13), respectively. In the case study, each observation is acquired by plugging the input vector into the objective function, therefore, a noise free model is a reasonable choice.

$$y = f(x) + \epsilon \tag{6.12}$$

$$y = f(x) \tag{6.13}$$

Based on the noise free model, the joint probability distribution of the function value at the observed points and any unseen point x^* in the input domain can be derived using Equation (6.14).

$$p\left(\begin{pmatrix} f(X) \\ f(x^*) \end{pmatrix}\right) = N\left(\begin{pmatrix} \mu(X) \\ \mu(x^*) \end{pmatrix}, \begin{pmatrix} \Sigma(X, X) & \Sigma(X, x^*) \\ \Sigma(x^*, X) & \Sigma(x^*, x^*) \end{pmatrix}\right) \tag{6.14}$$

By partitioning the joint distribution, the probability of the function value at the unexamined points $f(x^*)$ conditioned on available information $f(x)$ is computed using Equation (6.15).

$$p\big(f(x^*)\big|f(X)\big) = \frac{p\big(f(X), f(x^*)\big)}{p\big(f(X)\big)} = N\left(\mu_{f(x^*)|f(X)}(x^*), \Sigma_{f(x^*)|f(X)}(x^*, X)\right) \tag{6.15}$$

From Equation (6.15), the function values at new points still follow a multivariate normal distribution, where the mean and covariance function for new data are computed using Equations (6.16) and (6.17).

$$\mu_{f(x^*)|f(X)}(x^*) = \mu(x^*) + \Sigma(x^*, X)\Sigma(X, X)^{-1}\big(f(X) - \mu(X)\big) \tag{6.16}$$

$$\Sigma_{f(x^*)|f(X)}(x^*, X) = \Sigma(x^*, x^*) - \Sigma(x^*, X)\Sigma(X, X)^{-1}\Sigma(X, x^*) \tag{6.17}$$

Equation (6.16) can either be used as the posterior distribution of the Gaussian Process or a predictive distribution for unseen points. The posterior distribution is passed into the acquisition function to query the next point used to evaluate the objective function.

- Acquisition Function

The acquisition function is used to recommend the next step in the optimization search based on the posterior distribution. The queried point for the next evaluation x_{next} is the ones that maximizes the acquisition function (Equation (6.18)).

$$x_{next} = \arg\max_x acq(x) \tag{6.18}$$

where $acq(x)$ is the acquisition function parametrized by the input feature vector x. Multiple forms can be used for the acquisition function. The following discussion will concentrate on the acquisition function options incorporated in the Bayes-Opt python library (Brochu et al. 2010; Snoek et al. 2012). The expected improvement acquisition function is selected for Bayesian optimization since it balances exploration and exploitation trade-off and does not involve hyperparameter tuning.

Expected Improvement Acquisition Function

The expected improvement acquisition function defined in Equation (6.19) represents the increase in the target function value at any point compared to the current known maximum function value.

$$acq_{EI}(x^*) = E(\max(0, f(x^*) - f_n)) \tag{6.19}$$

where x^* is the new evaluation point, $f(x^*)$ is the corresponding observation, f_n is the maximum of all previous evaluations. A closed form solution to Equation (6.19) can be derived as Equation (6.20) using integration by parts and the Gaussian Process posterior distribution from Equation (6.15) (Frazier 2018).

$$acq_{EI}(x^*) = \max(0, \mu(x^*) - f_n) + \sigma(x^*)\phi\left(\frac{\mu(x^*) - f_n}{\sigma(x^*)}\right) - |\mu(x^*) - f_n|\Phi\left(\frac{\mu(x^*) - f_n}{\sigma(x^*)}\right) \tag{6.20}$$

where $\mu(x^*)$ and $\sigma(x^*) = \sqrt{\Sigma(x^*, x^*)}$ are the posterior mean and standard deviation from Equation (6.16) and (6.17), $\sigma(x^*)$. $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density and cumulative distribution functions of the standard normal distribution, respectively.

Probability of Improvement

The probability of improvement computed in Equation (6.21) represents the probability that the target function value is higher than the current known maximum function value, which can be calculated from the posterior distribution. This form of the acquisition function tends to sample from locations that result in the highest probability of improvement rather than the highest improvement.

$$acq_{PI}(x^*) = \Pr(f(x^*) \geq f_n) = \Phi\left(\frac{\mu(x^*) - f_n}{\sigma(x^*)}\right) \tag{6.21}$$

Gaussian Process Upper Confidence Bound

The upper confidence bound is the upper bound of the function value with a specified level of confidence. Given the posterior distribution of the Gaussian Process, the upper confidence bound of the function value is computed using Equation (6.22).

$$acq_{GP-UCB}(x^*) = \mu(x^*) + \zeta\sigma(x^*) \tag{6.22}$$

where ζ is the hyperparameter governing the exploration and exploitation trade-off. When a large ζ is selected, the algorithm tends to query points with high uncertainties (exploration) and the algorithm recommends points with high rewards (exploitation) when ζ is small.

6.3.3.2.2. Genetic Algorithm

Genetic algorithm (GA) is a heuristic search technique that mimics the process of natural selection on the fittest individuals to produce offspring. It has been widely applied in the aforementioned regional scale structure optimization problems (Dong et al. 2014; Jennings 2015). Each iteration in the GA search algorithm follows the flow chart shown in Figure 6.7. First, a set of individuals forms an initial population. Each individual is a solution to the problem. An individual contains Genes, which represent the values of different parameters. For each individual in the initial population, the fitness is calculated to evaluate its performance. Then, a selection procedure is conducted to find the fittest individuals to pass their genes to the next generation. The selected 'parents' form a mating pool, where each pair can generate their offspring. The procedure of generating offspring involves crossover and mutation. By conducting crossover, new offspring is generated combining half Genes from both sides of parents. After a new offspring formed, some of their genes can be subjected to a mutation with a specified random probability to reflect the randomness. All offspring forms the new initial population for the next round of search.

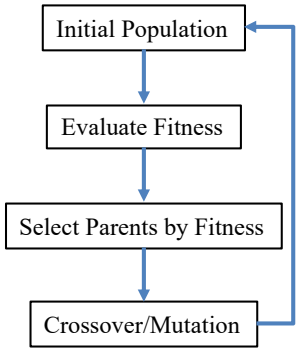


Figure 6.7 Genetic algorithm flow chart

Similar to Bayesian Optimization, GA is generally applicable for non-parametric and non-differential objective functions. Its interpretability and generalizability make it widely applied in structural engineering field. However, due to its crossover nature, where offspring inherit properties from ancestor populations, the convergence rate and associated computational cost are highly sensitive to the choices of initial population selection, crossover strategy and mutation rate. Back to the optimization problem this chapter focusing on, large number of retrofit strengths is generated at beginning as initial population. Each of strength combinations are evaluated to obtain portfolio seismic performance as fitness. When large number of earthquake scenarios are involved in the regional performance assessment, significantly heavier computational burden (than Bayesian Optimization) would rise in fitness evaluation process in GA. Though GA can provide similar optimization results, the computational expanse makes it less feasible for the objective function constructed in section 6.3.3.1 than Bayesian Optimization.

6.3.3.3. Optimization Results

For the surveyed buildings (25% of the inventory), the archetype assignment is consistent with the observed configuration. For the buildings not included in the survey, a randomized archetype assignment is adopted based on the proportions determined from the survey. One hundred randomized archetype assignment realizations are generated, and the optimization algorithm is implemented for each one. Figure 6.8 shows the median and median $\pm$ one standard deviation (based on the 100 simulations) of the optimal ratio (as determined by the Bayesian optimization algorithm) between the strengths of the retrofitted and existing buildings conditioned on the intensity bins. The median values are also listed in Table 6.3. Recall that the optimal percentage of peak strengths are based on an intensity map corresponding to 75% of the DBE. It is observed from Figure 6.8 that below 1.1g, the strength added by the retrofit has a strong positive correlation

with the intensity level. Whereas, at values higher than 1.1g, the added percentage of peak strength becomes negatively correlated with the ground motion intensity. Since almost two thirds of the buildings fall within the 1.0g to 1.2g intensity (see histogram in Figure 6.5), the algorithm tends to provide those buildings with higher strength to acquire more rewards in the objective function. The regional performance under the 20 scenarios following presented in the evaluation module.

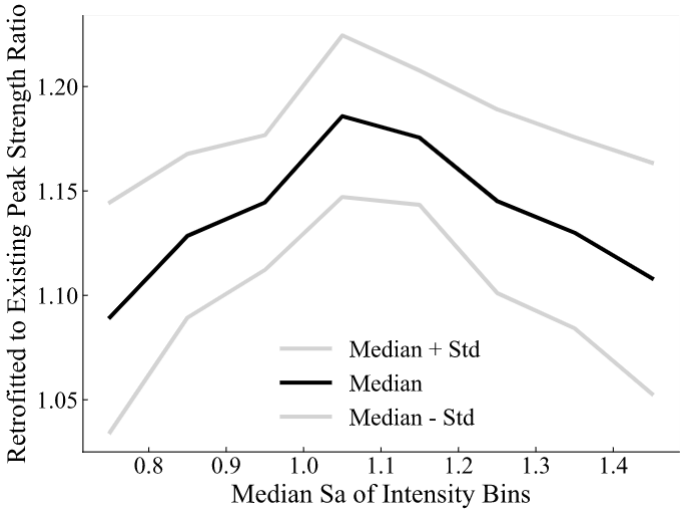


Figure 6.8. Percentage of peak strength added by the optimal retrofit scheme as a function of the intensity bin

Table 6.3. Median ratio of peak strength between the optimally retrofitted and existing buildings for each intensity bin

75% of DBE	**Median** $\frac{V_{retrof}}{V_{exist}}$
0.7g ~ 0.8g	1.09
0.8g ~ 0.9g	1.13
0.9g ~ 1.0g	1.14
1.0g ~ 1.1g	1.19
1.1g ~ 1.2g	1.18
1.2g ~ 1.3g	1.15
1.3g ~ 1.4g	1.13
> 1.4g	1.11

The number of iterations needed for the Bayesian optimization routine provides insight into the computational efficiency that is achieved through the introduction of the surrogate model. The construction of XGBoost the model required performing IDAs on 300 OpenSees models. The XGBoost surrogate model enables the evaluation of regional seismic performance within the optimization algorithm without additional IDAs. If the optimization algorithm is executed using mechanics-based-based models, IDAs would need to be performed on each unique retrofitted archetype during each iteration. For 100 Bayesian optimization iterations (Figure 6.6 shows the steps involved in each iteration), a minimum of 3200 IDAs are required. Moreover, the number would proportionally increase with the number of iterations. The net result is that the optimal portfolio scale retrofit is achieved using 10% of computational expense compared to if the surrogate is not used.

6.3.3.4. Sensitivity Analysis

As noted earlier, the algorithm achieves a high level of efficiency because it balances the gain from reducing collapse losses with the cost or penalty from adding lateral strength. Obviously, the chosen penalty function and penalty weight have significant impacts on the optimal solution. Although the penalty is supposed to reflect the cost of the retrofit, it is difficult to assign an appropriate value to each retrofit design during the iterative search. A penalty weight of 300,000 is adopted in the optimization based on the outcome of a sensitivity analysis. More specifically, the sensitivity of the optimal retrofit scheme to the chosen penalty weight is evaluated by implementing the algorithm using 10 values of the latter that range from 0 to 800,000. Figure 6.9 shows how the reduction in collapse loss (red curve) (as a percentage of the losses for the existing inventory) and added strength (blue curve) (as a percentage of the existing building strength) provided by the algorithmic retrofit are affected by the penalty weight. It is observed that, as the

penalization increases, the reduction in losses achieved by the algorithm as well as the total added strength decrease. In the areas where the slope of the red line is high, the rate of reduction in the effectiveness of the retrofit with increasing penalty is higher. However, beyond some threshold (a penalty weight of approximately 400,000), the reduction in effectiveness with increasing penalty becomes negligible. For the blue line, a high slope indicates that the gain in efficiency with increasing penalty is high. The most desirable penalty weight is one that is in the region where the rate of loss of effectiveness is balanced by the rate of efficiency gain. A penalty of 300,000 is chosen because it corresponds to the point beyond which an increase in the penalty provides a minimal decrease in strength (blue curve) at the cost of a substantive reduction in effectiveness (red curve).

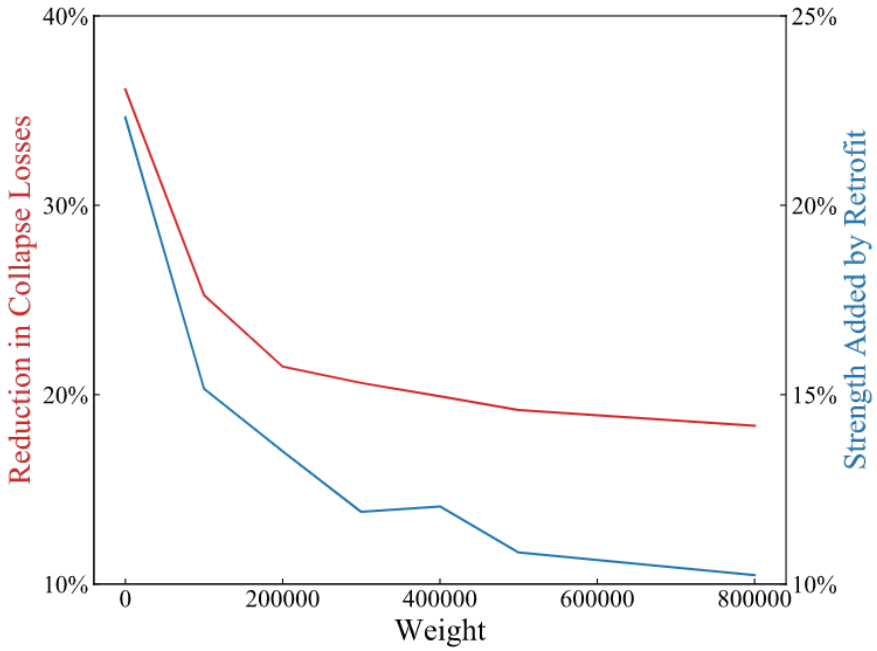


Figure 6.9. Effect of the penalty weight on the reduction in collapse losses and added strength, both normalized by values associated with the existing inventory

6.3.4. Evaluation Module

In this section, the algorithmic retrofit scheme is first evaluated by comparing the regional collapse losses and cost (added lateral strength) for the 20 scenarios used for optimization. Then, the

algorithm retrofit scheme is similarly evaluated using events outside the ones used as the basis of the optimization. Three events are considered for this purpose, which vary based on the severity of the shaking intensities and associated collapse losses. The first is the hypothetical **M** 7.1 Puente Hills scenario which has been used in several other studies (e.g. Field et al. 2005; Rowshandel et al. 2006). The Puente Hills scenario is generally considered to be a "rare" event that would cause severe earthquake shaking in the city of Los Angeles. The two additional **M** 6.4 and **M** 6.8 hypothetical events are selected from the stochastic event set to represent moderate and low levels of earthquake induced shaking and losses. Due to the further distance between the target region and the earthquake sources, the relative impacts of these two events are significantly lower than the Puente Hill event. The goal here is to evaluate the effectiveness and efficiency of the algorithmic retrofit for "unseen" events over a range of impact levels.

6.3.4.1. Evaluation for 20 Representative Scenarios

First, the retrofit scheme is evaluated with respect to the scenarios used in the optimization. The inventory is retrofitted based on the median percentage of peak strength values shown in Table 6.3. The mean regional collapse loss ratios (as a percentage of the inventory replacement cost) for the 20 scenarios are listed in Table 6.4. The existing inventory loss ratio ranges from 0.01% to 45%, which demonstrates the diversity in terms of impact level for the selected subset of scenarios. The mean loss ratio for the existing inventory is 18.5% compared to 13.3% for the Ordinance-retrofitted inventory, which represents a 29% reduction. When averaged over all scenarios, the algorithmic retrofit reduces the collapse loss of the existing inventory by 21% and the corresponding collapse losses are only 11% higher than that of the Ordinance-retrofitted inventory. To compare the efficiency of the algorithmic and Ordinance retrofit, the total strength added to the inventory is used as a metric of the retrofit cost. When the inventory is retrofitted based on the

optimization algorithm, the total added strength is approximately 60% of what Ordinance provides. In other words, the algorithm uses 60% of the strength provided and results in a regional collapse loss that is 11% higher than the Ordinance.

Table 6.4. Mean loss ratios for the existing and retrofitted inventories subjected to the 20 scenarios

Event ID	Existing Inventory Mean Loss Ratio	Algorithm Retrofitted Inventory Mean Loss Ratio	Ordinance Retrofitted Inventory Mean Loss Ratio
1	19.6%	14.8%	12.9%
2	0.8%	0.5%	0.4%
3	2.6%	1.7%	1.4%
4	5.6%	3.8%	3.2%
5	19.4%	14.7%	12.9%
6	44.9%	37.4%	33.9%
7	0.017%	0.008%	0.006%
8	34.3%	27.7%	24.8%
9	7.9%	5.7%	4.8%
10	23.0%	17.8%	15.7%
11	22.8%	17.6%	15.4%
12	6.3%	4.5%	3.8%
13	35.9%	29.4%	26.4%
14	16.6%	12.7%	11.1%
15	44.8%	37.3%	33.8%
16	6.2%	4.3%	3.7%
17	11.5%	8.4%	7.2%
18	40.5%	33.5%	30.5%
19	26.6%	21.1%	18.7%
20	0.51%	0.32%	0.26%

Figure 6.10 shows the spatial distribution of the mean normalized (by the replacement cost) collapse losses for the existing, ordinance-retrofitted and algorithm-retrofitted inventories. For the buildings located in downtown and west Los Angeles where the mean loss ratios exceed 20% (yellowish and reddish points), both the algorithm and Ordinance retrofit reduce the losses to below 17%. The Ordinance retrofit has observably better performance for buildings located in the

downtown area. Figure 6.10 also provides insight into the spatial pattern of the collapse risk for the stochastic event catalogue. For the existing inventory, downtown, eastern, and western Los Angeles suffered more collapse loss compared to northern and southern Los Angeles. Therefore, the retrofit work for these regions should be prioritized to maximize the collapse risk reduction.

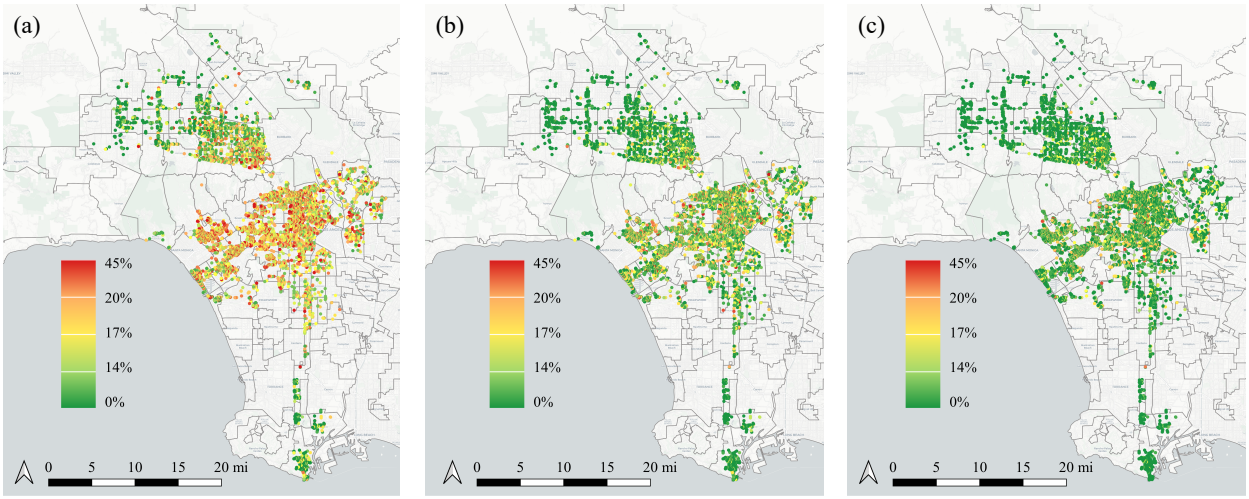


Figure 6.10. Spatial distribution of normalized (by building replacement cost) collapse loss: (a) existing, (b) ordinance-retrofitted and (c) algorithm retrofitted inventories

6.3.4.2. Evaluation of Three Single Events

Other than the events adopted in optimization module, the algorithmic retrofit scheme is further evaluated using 'unseen' events. Three events (denoted as the Puente Hill, **M** 6.8 and **M** 6.4 events) representing a range of shaking intensities and collapse losses are selected. The basic information for the three events are summarized in Table 6.5, and Figure 6.11 shows the associated spatial distribution of median spectral accelerations at 0.2s ($Sa_{0.2}$) for 50 shake maps. For the three scenarios, the median $Sa_{0.2}$ ranges from 0.2g to 1.8g, 0.12 to 0.83g and 0.03g to 0.12g for the Puente Hill, **M** 6.8 and **M** 6.4 events, respectively. In addition to differences in the severity of ground shaking, there are also spatial variations where the highest intensities are located for each

event. For the Puente Hill, **M** 6.8 and **M** 6.4 events, the highest shaking intensities are in the east, west and south, and north Los Angeles, respectively.

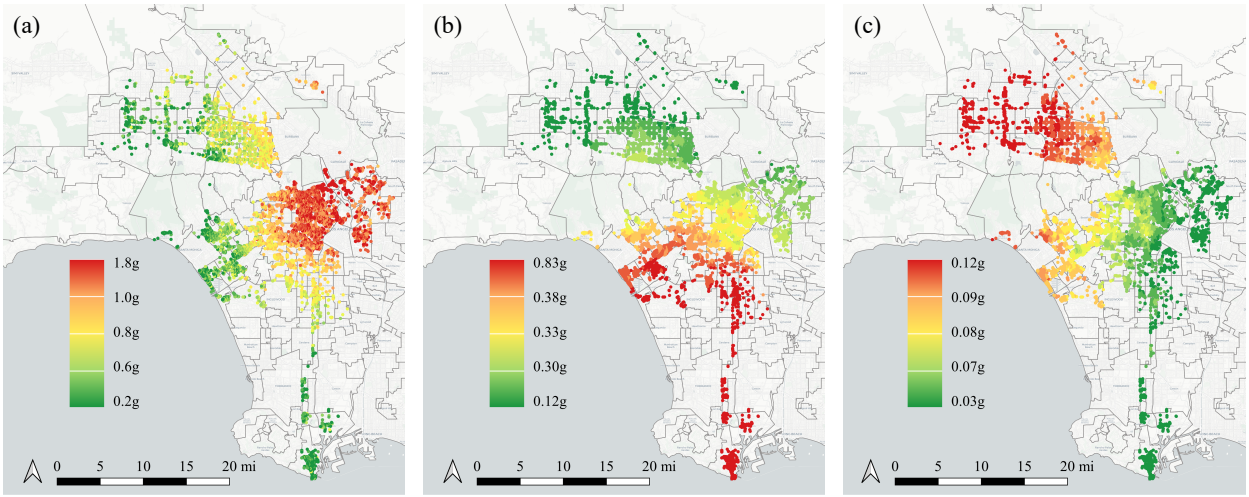


Figure 6.11. Spatial distribution of $Sa_{0.2}$ for the: (a) Puente Hill Scenario, (b) **M** 6.8 event, (c) **M** 6.4 event

Table 6.5. Basic information for the three considered scenarios

Event ID	Magnitude	Source
Puente Hill	7.1	Puente Hill
M 6.8	6.8	Palos Verdes
M 6.4	6.4	North Channel

Some key results are summarized in Table 6.6. The normalized collapse losses for the existing inventory averaged over all shake maps is 50.5%, 13.6% and 0.15%, for the Puente Hill, **M** 6.8 and **M** 6.4 events, respectively. It is observed that the benefit derived from the retrofit (algorithmic and Ordinance) decreases with the severity of the impact. For example, the Ordinance retrofit reduces the collapse loss for the existing inventory by 16.4%, 37.5% and 61.5% for the Puente Hill, **M** 6.8 and **M** 6.4 events, respectively. For the same three events, the losses associated with the algorithmic retrofit are 6%, 21% and 33% higher than the Ordinance retrofit, respectively. This result indicates that the effectiveness of the algorithm relative to the Ordinance increases with the

severity of the event. Recall that the algorithm only uses 60% of the strength required by the Ordinance retrofits.

Table 6.6. Evaluation results for the three "unseen" events

	Puente Hill Event	**M 6.8 Event**	**M 6.4 Event**
Existing	50.5%	13.6%	0.15%
Algorithm Retrofitted	46.4%	10.4%	0.08%
Ordinance Retrofitted	43.7%	8.6%	0.06%

The spatial distribution of the mean collapse loss ratios associated with the Puente Hills event considering all shake maps are shown in Figure 6.12. Similar to $Sa_{0.2}$, the collapse losses for the existing and retrofitted buildings decay from east to west. Note that the actual loss ratios sometimes exceed 100% of building replacement cost because fatality costs are incorporated. Due to the extremely high $Sa_{0.2}$ in East Los Angeles, even some of the retrofitted buildings suffer complete losses, which leads to lower efficiency of Ordinance retrofit. This also results in relatively comparable regional performance of the algorithmic and ordinance retrofitted inventory.

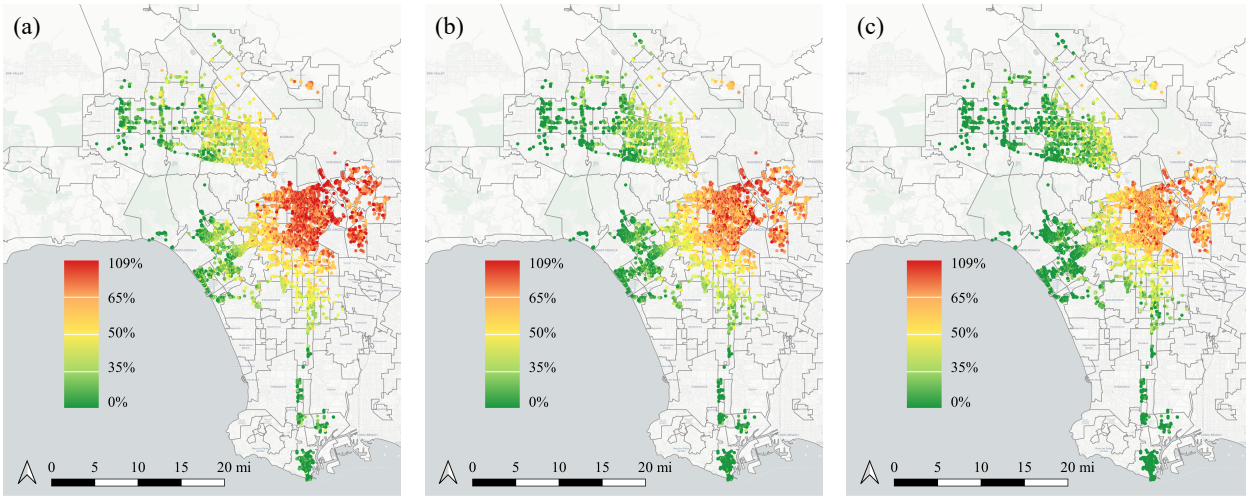


Figure 6.12. Spatial distribution of building collapse loss under the Puente Hills scenario: (a) existing, (b) algorithm retrofitted and (c) ordinance retrofitted buildings

Figure 6.13 presents the spatial distribution of mean loss ratios under the M6.8 event. The normalized collapse loss for the existing inventory of 13.6% is reduced to 10.6% and 8.4% by the algorithmic and ordinance retrofit, respectively. Unlike the Puente Hill event, none of the buildings in the existing inventory suffer complete losses. Additionally, only approximately 1% of the buildings in the existing inventory have collapse losses that exceed 50% of the building value (including fatalities). The percentage of buildings with losses greater than 25% of the building value is reduced from 7.6% for the existing inventory to 3.9% (algorithmic retrofit) and 3.1% (ordinance retrofit). As for the existing buildings with greater than 10% loss ratios, the algorithmic and ordinance retrofit reduces the proportion from 60% to 38.6% and 26.5%, respectively. Based on these different threshold comparisons, it is observed that the ordinance retrofit provides greater benefit than the for intermediate loss ratios (from 10% to 20%). Comparing the loss distribution in Figure 6.13 (a), (b) and (c), the most significant reduction in collapse loss is in the downtown and west Los Angeles areas for both retrofit designs, where most of existing buildings have loss ratios ranging from 9% to 18%. Additionally, the ordinance retrofit shows noticeably superior performance compared to the algorithmic retrofit scheme in these areas.

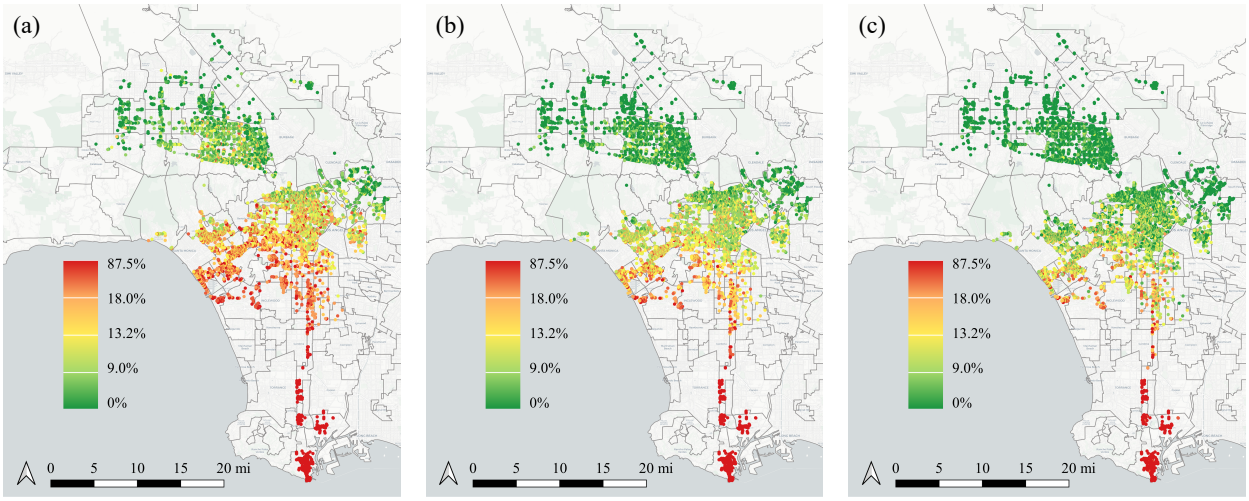


Figure 6.13. Spatial distribution of building collapse loss under M 6.8 event: (a) existing buildings, (b) algorithm retrofitted buildings, (c) ordinance retrofit schema retrofitted buildings

Figure 6.14 presents the same performance metric for the **M** 6.4 event. Major losses in existing buildings are concentrated around northern Los Angeles. Due to low shaking intensities, the maximum collapse loss ratio for the existing buildings is only 2.6% and is to 1.4% and 0.9% by the algorithmic and ordinance retrofit schemes, respectively. It is also observed that the reduced loss is positively correlated to existing building losses. In regions with higher existing building losses (north Los Angeles), the differences between the ordinance and algorithmic retrofit are more noticeable. On the other hand, in the area where the existing building collapse losses are lower (relative to the other areas for the same event), the effects of retrofitting are subtle (downtown and

east Los Angeles). As a result, no major differences in the benefit derived from the two retrofit approaches is observed.

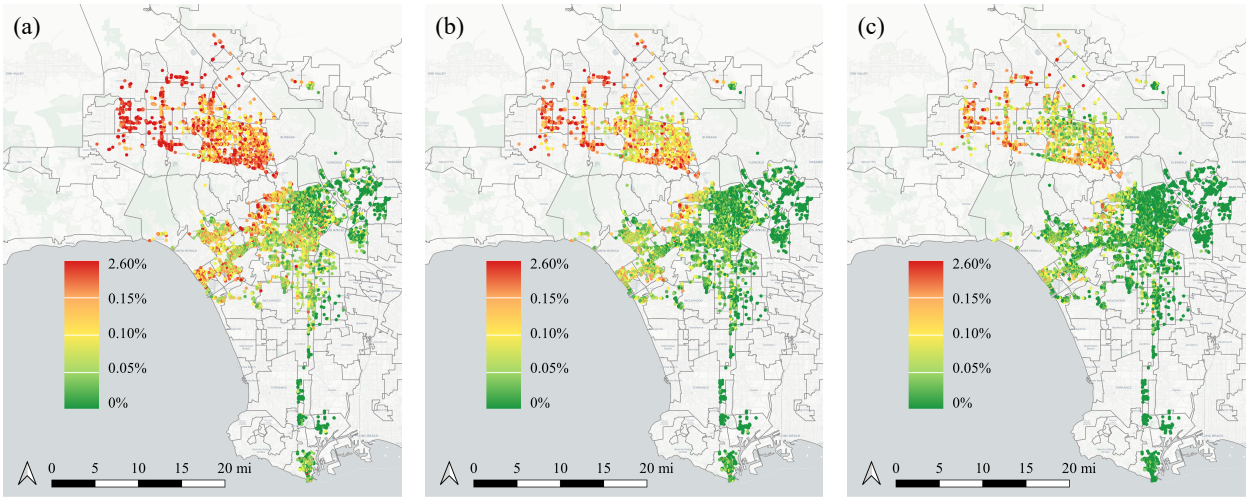


Figure 6.14. Spatial distribution of building collapse loss under M 6.4 event: (a) existing (b) algorithm retrofitted and (c) ordinance retrofitted buildings

Besides the associated efficiency, the results from the algorithmic retrofit can be used to inform the priority of retrofit implementation. For example, buildings with moderate to severe site seismicity can be prioritized since they contribute the most to the regional seismic collapse risk. Whereas buildings corresponding to the lowest hazard level can be given the least priority since the associated risk is relatively low. Ultimately, the adopted retrofit scheme should be based on a more comprehensive evaluation that the one presented in this study.

6.4. Development of A Retrofit Policy Based on Results from Optimization

Practically speaking, the retrofit scheme generated by the optimization algorithm cannot be used as a retrofit policy in its current form. More specifically, the level of granularity in the algorithm-optimized retrofit scheme is such that it cannot be directly used as a retrofit policy. However, the results from the optimization can be used to develop a set of rules that are general enough to serve

as a policy. The following guidelines are presented as a roadmap for generating a retrofit policy based on the optimization results:

1. Identify the structures that warrant retrofitting: Examples of such criteria include (i) buildings with first story wall lines that are more than 20% less than that of the above story should be retrofitted or (ii) buildings where the strength and/or stiffness of the 2nd story is more than 20% that of the first. Note that example (i) is used in the Los Angeles Ordinance retrofit.
2. The next step is to determine the design force that will be used to the retrofit buildings. First, the peak lateral strength in the direction under consideration should be determined using a pushover analysis. Then, the intensity level corresponding to 75% of the DBE hazard level $\left(Sa_{T1,75\%DBE}\right)$ is computed using ASCE 7-16 (ASCE 2016) and the associated site seismicity parameters. The design base shear is taken as 20% of the peak strength if $Sa_{T1,75\%DBE} > 1.0g$ (based on the results summarized in Table 6.3). For the remaining cases, 14% of the peak strength should be used to compute the design force (also based on results in Table 6.3).
3. Select and design the lateral force resisting system to provide the required additional strength.
4. The designed retrofit alternative should have at least the same drift capacity at peak strength as the existing system. This recommendation is based on the results from this and other studies (Buckalew et al. 2015; Burton et al. 2019) which have shown that the collapse safety is less sensitive to the ductility of the retrofit components than to strength.

6.5. Conclusion and Future Work

A framework for optimizing seismic retrofits based on regional seismic performance metrics is developed in this paper. The overall procedure comprises a prediction module, an optimization module and an evaluation module. The prediction module uses surrogate models to estimate retrofitted building performance conditioned on key retrofit design variables. The surrogate model

is shown to significantly reduce the large computational expense associated with assessing regional seismic performance while considering variations in the structural characteristics of the inventory, the spatial variation of seismic hazard and the iterative search used in the optimization algorithm. The optimization module is used to find the retrofit design that maximizes the objective function value, which is defined based on regional scale seismic performance. Reasonable constraints on the input parameters are used to confine the search space. Additionally, a penalty function is added to the portfolio performance metric to suppress unnecessarily high retrofit strengths. The evaluation module benchmarks the regional performance of the algorithmic retrofit scheme against the existing inventory and another retrofit scheme that uses more conventional approaches. A case study that is based on the Los Angeles Ordinance that mandates the retrofit of soft, weak and open front (SWOF) wall line buildings is used to demonstrate the capabilities of the framework.

As part of the prediction module, a machine learning model based on the XGBoost algorithm is used to rapidly estimate the collapse margin ratio of individual retrofitted buildings using their structural characteristics and the strength added by the retrofit (relative to existing buildings) as features or input variables. The performance of the XGBoost model was evaluated by computing the median absolute relative error, which was 1.4%. the results from residual diagnostics also revealed low bias and high predictive accuracy.

Bayesian optimization was used to find the most desirable retrofit solution for the approximately 12,000 SWOF buildings in Los Angeles. A set of 20 scenarios deemed representative for the region is used as the basis of the optimization and the collapse losses for the inventory is used as the performance metric. In terms of the retrofit design, the percentage of the existing building peak strength added by the retrofit was the parameter to be optimized. To account for the spatial

variation in seismic hazard, eight intensity are established based on 75% of the shaking intensity corresponding to the design basis earthquake. The objective function was formed by coupling the collapse losses with an inversion penalty. Based on the results of a sensitivity study, a penalty weight of 300,000 is used to balance the effectiveness (reduction in collapse losses) and efficiency (avoided unnecessarily added strength) of the algorithm retrofit. The XGBoost model was shown to reduce the computational expense (relative to if only mechanics-based assessments were performed) by more than 90%.

For the 20 events used as the basis of the optimization, the average regional collapse loss ratio (with respect to inventory replacement cost) for the existing inventory is reduced from 18.5% to 14.7% and 13.3% by the algorithm and Ordinance retrofit, respectively. The algorithm retrofit was shown to be highly efficient, utilizing only 60% of the added lateral strength relative to the Ordinance. The algorithm retrofit as also evaluated for three "unseen" (i.e. not considered in the optimization routine): the **M** 7.1 Puente Hill event and **M** 6.8, and **M** 6.4 events. These events were chosen to evaluate the effectiveness and efficiency of the algorithm retrofit using events with varying severity in terms of collapse losses. The collapse loss ratios for the **M** 7.1, **M** 6.8, and **M** 6.4 events were 50.2%, 13.6% and 0.15%, respectively. For the same three scenarios, algorithm retrofit reduces the losses for the existing inventory by 11.2%, 24%, and 44.6%, respectively, which suggests that its effectiveness increases with the severity of the event. Relative to the Ordinance, the much lower cost (as quantified by the added strength) and comparable performance of the algorithm serve as compelling evidence of its overall ability to achieve effective and efficient retrofits.

This study has demonstrated the great potential of using advanced optimization and machine learning methods for balancing the effectiveness and efficiency of portfolio retrofit designs.

However, there are areas where additional work is needed to solidify the usefulness as the proposed framework. The collapse performance of the archetype buildings was assessed using the FEMA P695 ground motions (FEMA 2009). This was then used as the basis for regional seismic assessments where there are significant variations in the site characteristics. Issues such as near-source ground motion and local site effects should be considered when selecting ground motions for similar future studies. K-means clustering was the primary means by which the set of 20 scenarios (intended to represent the stochastic catalogue for the region) used as the basis of the optimization was selected. Future studies should explore the use of more advanced techniques such as active learning for selecting such an event subset. Alternatively, if the computational power permits, the entire stochastic catalogue can be used as the basis of the optimization to increase its generalizability. This approach would also enable the use of risk-based performance metrics in the objective function. Lastly, because the SWOF retrofits are targeted towards life safety, the regional collapse losses were used as the performance metric. However, the proposed framework can also be used to establish design parameters for new buildings that maximize regional performance where other more resilience-related metrics (e.g. total economic losses, recovery trajectory) can be considered.

CHAPTER 7. CONCLUSION

7.1. Summary

The presence of crawlspace in single-family houses and soft, weak, open, front wall in multi-family houses forms single story mechanism and results in severe damage risk under earthquakes. Policy actions are usually taken to address such issues. The thesis focuses on the implications of woodframe building retrofit and proposing a quantitative way to optimize performance-oriented retrofit policy.

To fulfill the objective, reliable numerical analyses are required to access building mechanical properties, seismic performances and earthquake induced losses. In Chapter 2, an end-to-end Python based woodframe building OpenSees modeling and analyzing workflow was developed. The tool is capable of generate numerical models in OpenSEES, perform nonlinear numerical analysis, conduct post-processing and FEMA-P58 based loss assessment given structural configurations, material properties and specified analysis settings. Detailed modeling procedure at OpenSees component level for single-family cripple wall houses developed in PEER-CEA Project and Los Angeles soft, weak, open-front wall line (SWOF) multi-family houses were discussed. The sample pushover and nonlinear dynamic response analyses identify the vulnerability of soft story buildings.

With the help of the modeling tool, larger number of numerical analyses can be realized with affordable computational expanse. Chapter 3 investigated the potential opportunities of developing machine learning (ML) models using the woodframe building performance database generated by the tool. Detailed mathematical derivation of the algorithms and formal ML model development procedure involving dataset development, model development and verification were introduced. Six ML algorithms, including ordinary linear regression, response surface method (RSM), LASSO

regression, ridge regression, Random Forest and XGBoost, were used for establishing predictive models for a PEER-CEA Project single family cripple wall house's median collapse intensities (MCI) and expected annual losses (EAL). Derived XGBoost models were found to be capable of delivering the most accurate and stable predictions on both MCI and EAL.

Chapter 4 started with a preliminary sensitivity analyses, where a full factorial experiment design was adopted to detect the impacts of single-family house number of stories, seismic weights, cripple wall heights, construction era and retrofit. Using the analysis of variance (ANOVA) method, all studied five variants were found to be statistically significant on both building MCI and EAL. Then, a more comprehensive sensitivity study was presented with the ML involved. Ordinary linear regression and XGBoost model developed in Chapter 3 were used to quantitatively identify the relative contributions to building seismic performance of number of stories, seismic weights, material properties, damping ratios and retrofitting panel lengths. Lastly, the XGBoost models developed on the same dataset were used for measuring model uncertainties on building MCI and EAL. The introduction of ML method provides promising estimation on the uncertainties while maintaining affordable level of computational burden.

Chapter 5 concentrates on the implications of Los Angeles Ordinance retrofit of SWOF buildings. The benefits of four different retrofit alternatives permit by Ordinance Guideline (LADBS 2015) on building strength, ductility and collapse risk were quantified and compared. Specifically, for Guideline based retrofit, a multi-scale cost-benefit analysis was presented. Two sets of retrofit designs were proposed capturing the site seismicity variation following the Guideline. Cost-benefit ratios and break-even time were calculated for each of the archetype to assess retrofit design cost-effectiveness under different assumptions (with and without earthquake insurance premium). Innovatively, cost-effectiveness was further evaluated for the entire city of Los Angeles under a

single earthquake event and a stochastic event catalogue. Such analyses facilitate a deeper understanding on retrofit benefit and its long-term implications.

In Chapter 6, an exhaustive framework was proposed for performing portfolio seismic performance oriented retrofit policy optimization for wood frame structures. The framework comprises of prediction module, optimization module and evaluation module. The prediction module constructs ML models to rapidly access retrofitted building seismic performance given retrofit design. The outcome of prediction module helps to form the objective function in the optimization module with building exposure, shaking intensity spatial distribution and appropriate constraints applied on retrofit designs. Then, powerful optimization engine solves the stochastic optimization problem and propose an optimal retrofit schema. Lastly, the derived retrofit schema is passed into evaluation module to obtain regional seismic performance outcomes. An optimization was performed on LA SWOF building retrofit, and the proposed retrofit schema was compared with ordinance retrofit. The outcome was found to be able to address the inefficiencies in ordinance retrofit.

7.2. Conclusion

The presence of soft story, crawlspace in single-family dwellings and open wall line in multi-family dwellings, forms single-story mechanism. From the pushover and response history analyses, the majority of story drift demands concentrate at the weak story. Under dynamic loading, collapse occurs at soft story, and the adjacent story almost suffer no structural damage. These observations result from the relative stiffness and available strength between the soft story and upper level. Retrofitting, either through structural wood panels or moment frames, addresses the imbalanced vertical distribution of strength and stiffness issues. Building strength capacity, ductility capacity and collapse safety are significantly improved, and associated earthquake losses reduce. Such

benefits are limited by an upper bound. For 'soft-story-only' retrofit, increasing additional strength added at the soft story, the strength differences between the soft story and adjacent level are eliminated until some threshold. It results in a shift of soft story from original level to its adjacent level. Beyond the threshold, providing extra strength cannot further improve building seismic performance. The impacts of retrofit also differ by shaking intensities. Under extremely intensive ground shaking, even retrofitted building experience a total loss. As for subtle shaking intensities, un-retrofitted building is expected to have close to zero losses. In both cases, the reductions in earthquake induced losses are limited, which further result in inefficient retrofit cost-benefit ratios. Woodframe building seismic performances under portfolio scope are highly diverged. Building number of stories and seismic weight are found to negatively contribute to collapse safety and losses. Construction era, damping ratio, and retrofitting work through a positive way. The impacts of these features can be summarized into two categories: strength and dynamic behavior. Changes in the building available strength and strength vertical distribution significantly variates building capacity. Number of stories, seismic weight, construction era and material properties could lead to such differentiations. Dynamic behavior including the period, inherent damping and hysteretic energy dissipation. Number of stories, seismic weight and material properties either explicitly or inexplicitly contribute to this aspect. Though such impacts can be assessed theoretically, no analytical form could be derived to quantitatively describe the influences. ML models are found to be able to handle this situation. A well-developed ML model can consistently deliver accurate predictions on building seismic performances. Such models could be utilized to quantify the uncertainties, measure features' importance and even exploit high-dimensional feature space for an optimal solution. More importantly, the investigation with ML model could control the number

of nonlinear analysis to an affordable level. It provides alternatives to the earthquake engineering problems with exhausting amount of analyses involved.

7.3. Limitations and Future Work

This study has demonstrated the great potential of using advanced machine learning methods for prediction, sensitivity analyses and uncertainty quantification. However, there are areas where additional work is needed to solidify the usefulness as the proposed methodologies. In the modeling part, rigid diaphragm assumption was adopted. The impact on seismic performances of semi-rigid or flexible diaphragm should be investigated. The ground motion uncertainty was not well addressed in the thesis. The analyses on single family houses only used conditional spectra selected ground motions for San Francisco $V_{s,30} = 270\ m/s$ site. The variations in building seismic performances caused by site effect were not incorporated. The collapse performance of the multi-family buildings was assessed using the FEMA P695 ground motions. Issues such as near-source ground motion and local site effects should be considered when selecting ground motions for similar future studies. In the retrofit design optimization, K-means clustering was the primary means by which the set of 20 scenarios (intended to represent the stochastic catalogue for the region) used as the basis of the optimization was selected. Future studies should explore the use of more advanced techniques such as active learning for selecting such an event subset. Alternatively, if the computational power permits, the entire stochastic catalogue can be used as the basis of the optimization to increase its generalizability. This approach would also enable the use of risk-based performance metrics in the objective function. Lastly, because the SWOF retrofits are targeted towards life safety, the majority discussion on regional scale seismic performance focuses on collapse safety and loss. Regional performances with other more resilience-related metrics (e.g. total economic losses, recovery trajectory) can be considered.

7.4. Future

One of the most attractive findings from the thesis is the utilization of ML method in solving structural engineering problems. Right now, with the help of high-performance clusters and powerful personal computers, we are standing at a point where any desired analyses with different fidelities and numbers of variants are doable. Built on the top of mechanics-based-based models, ML allows researchers to step further into a more chaotic domain to model the real-world. A few personal thoughts on the possibilities of earthquake engineering are presented here to intrigue other researchers. Figure 7.1 presents a conceptual workflow for an evolutional performance-based design facilitated by design database, traditional ML predictive model, cloud computing and deep learning algorithms. Under the traditional performance-based design catalogue, the design procedure is iterative. Nonlinear analysis is required to determine whether the performance objective is achieved. In this process, the numerical analyses performed prior to the final design are 'wasted'. Within the evolutional design scope, all analyses histories could be gathered and form into a performance database, where design and associated performance outcome pairs are stored. The database could be used to construct predictive model using conventional ML algorithms. Similarly, the evolutional framework starts with a desired performance. More abundant metrics not limited to seismic performance could serve for performance objective, such like resiliency index and socioeconomic impact. A direct search on the performance database could be performed to check whether any analyses histories deliver desired performance objective. Alternatively, the prediction model could be used for adaptively proposing possible designs in the unknown domain achieving performance objective. Once a proposal design is specified, cloud computing and deep learning are able to boost the nonlinear analyses and performance assessment.

The outcome of ground truth performance is then compared with objective to guide the next round of iteration. The analyses generated meanwhile could enrich the database and prior knowledge.

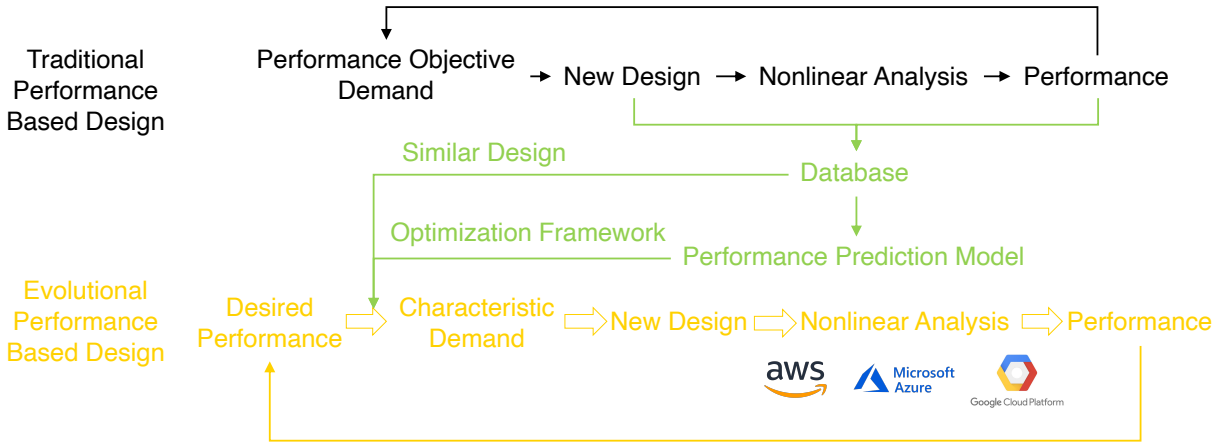
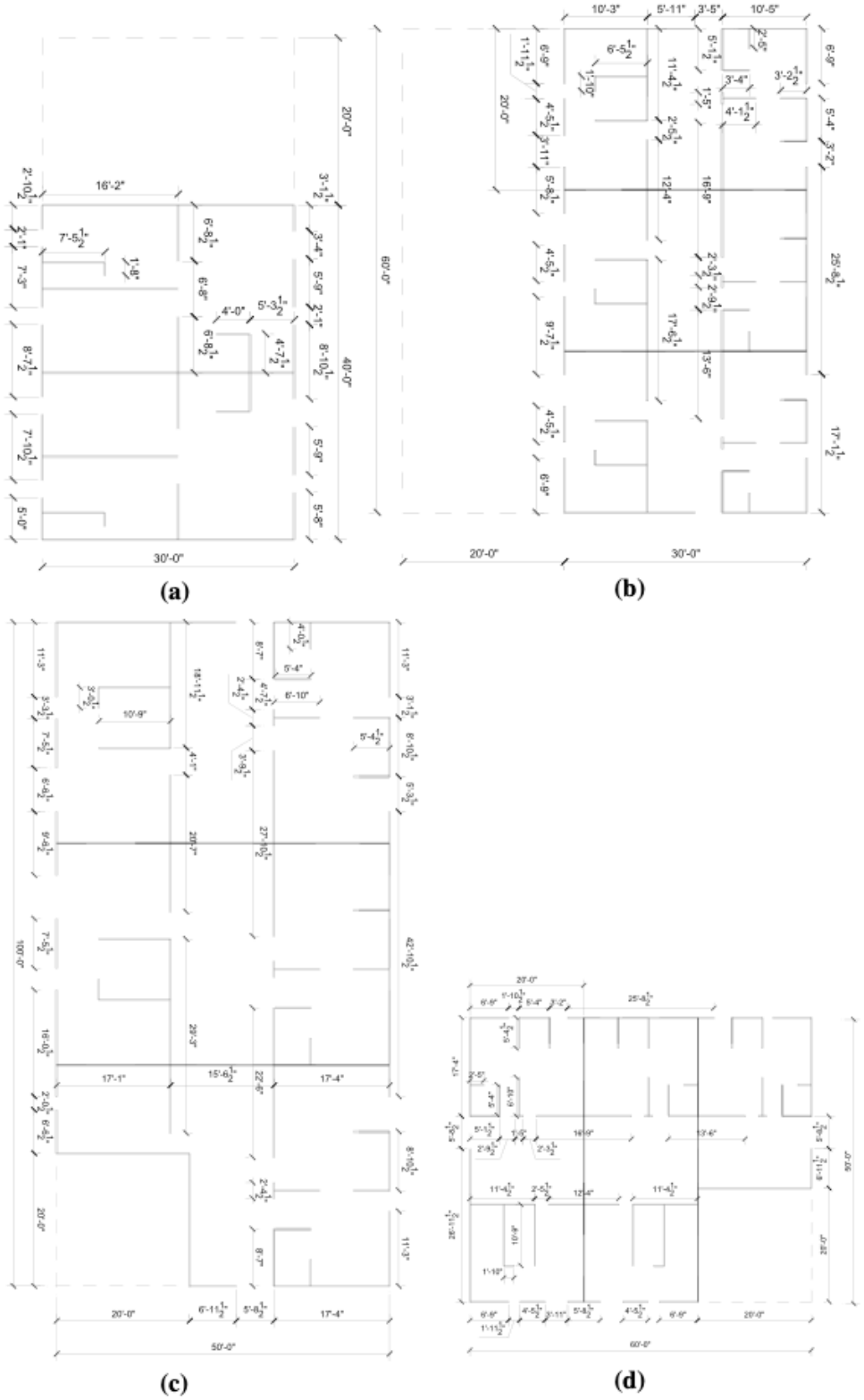
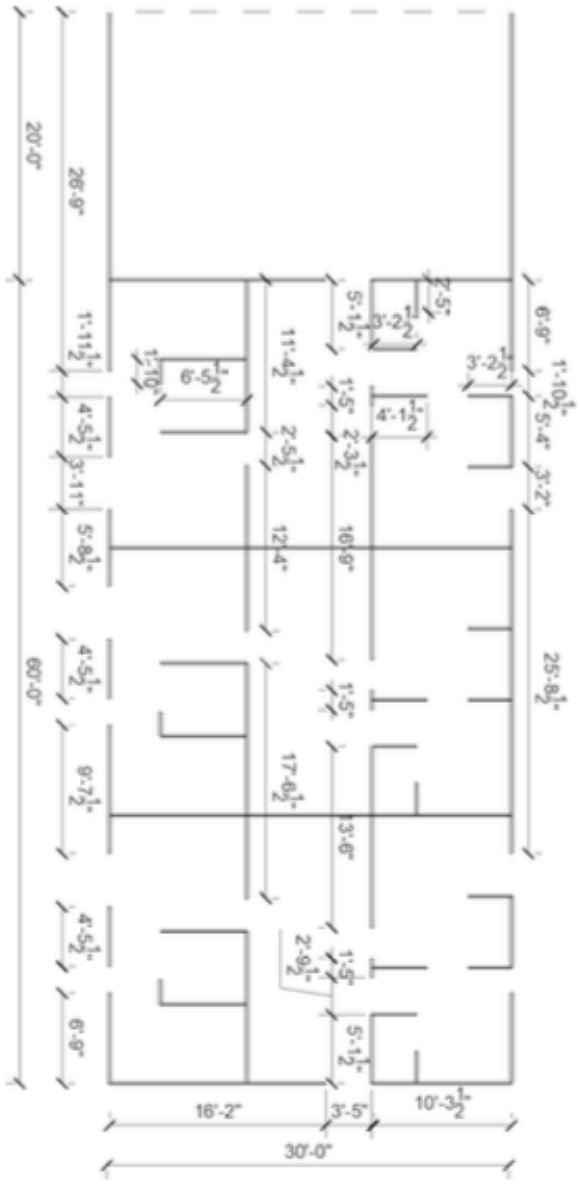


Figure 7.1 A conceptual idea on evolutional performance-based design

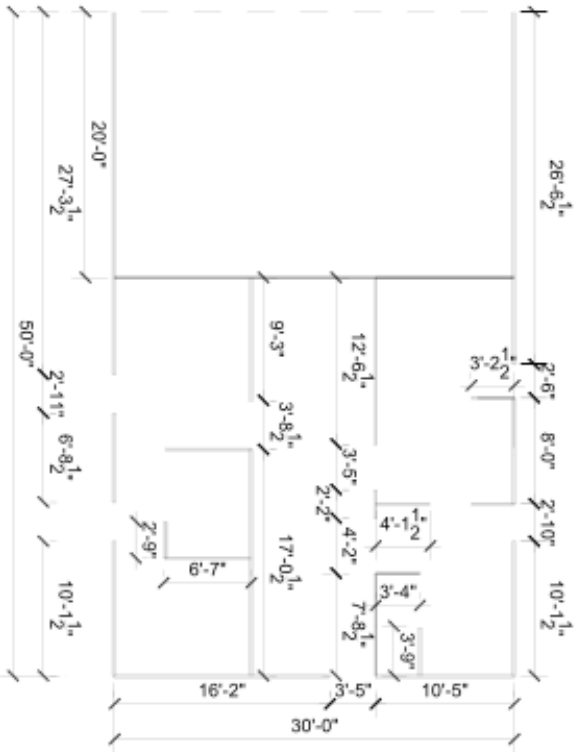
Another thinking the author wants to deliver here is ML is not supposed to be abused. There is a well-known aphorism saying, “all models are wrong, but some are useful”. ML models are constructed on the basis of data. They are expected to explain the inherent structure and relationship of the data rather than the nature. ML model is more used as a supplement to mechanics-based-based model to tackle the problems without analytical solutions in the thesis. It has its own limitation. Therefore, every details of ML model development including data wrangling, experiment design, feature engineering, model training, fine tuning, parameter selection, model selection and interpretation are supposed to be well treated. Additionally, engineering researchers are supposed to chase more fundamental and interpretable representations of ML model to fill in the gap between engineering practice and researches.

Appendix A. Floor Plan Dimensions of SWOF Building Archetypes




20'-0"
26'-9"
60'-0"
1'-11½"
4'-5½"
3'-11"
5'-8½"
4'-5½"
9'-7½"
4'-5½"
6'-9"
1'-10"
6'-5½"
11'-4½"
2'-5½"
12'-4"
5'-1½"
3'-2½"
2'-5"
1'-5"
2'-3½"
4'-1½"
16'-9"
17'-6½"
2'-9½"
6'-9"
1'-10½"
5'-4"
3'-2"
25'-8½"
13'-6"
5'-1½"
10'-3½"
16'-2"
3'-5"
30'-0"

(e)


20'-0"
27'-3½"
50'-0"
2'-11"
6'-8½"
10'-1½"
26'-6½"
9'-3"
3'-8½"
2'-9"
6'-7"
12'-6½"
3'-5"
2'-2"
4'-2"
17'-0½"
4'-1½"
3'-4"
7'-8½"
3'-9"
2'-6"
8'-0"
2'-10"
10'-1½"
16'-2"
3'-5"
10'-5"
30'-0"

(f)

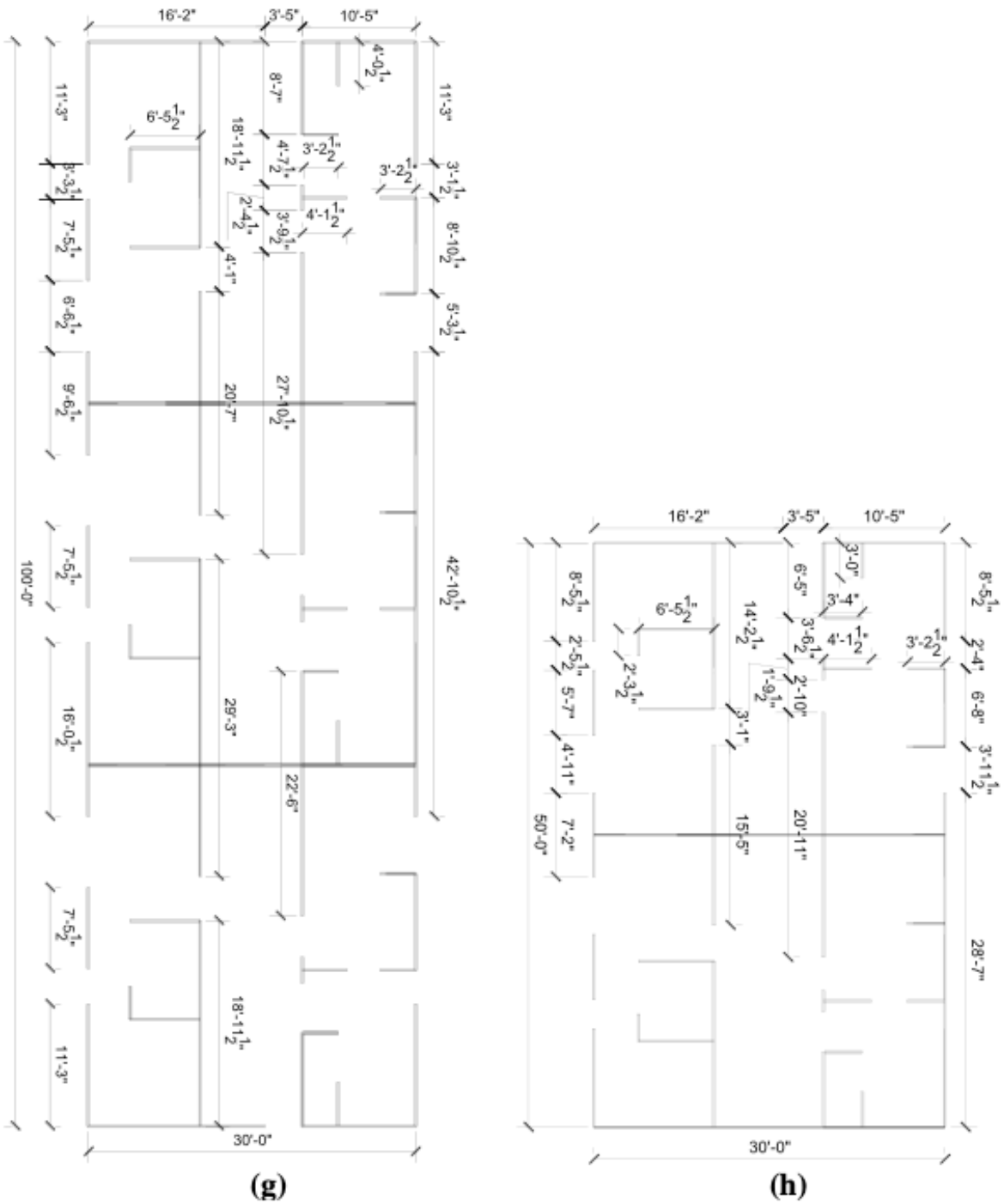


Fig. A.1 SWOF building floor plans and dimensions

Appendix B. Component Information for SWOF Building Archetypes

Table B.1 Component quantities of SWOF building archetypes (for entire building)

Building Name	Structural Components		
	Light framed wood lateral walls	Steel Column Base Plates (only retrofitted case applied)	Welded Steel Moment Connection (only retrofitted case applied)
L1-2S-60X30-GWB/HWS	25.9	6	3
L1-3S-60X30-GWB/HWS	40.17	6	3
L1-2S-100X30-GWB/HWS	38.62	6	3
L1-3S-100X30-GWB/HWS	58.63	6	3
L2-2S-60X50-GWB/HWS	32.8	6	3
L2-3S-60X50-GWB/HWS	50.83	6	3
L2-2S-100X50-GWB/HWS	44.94	6	3
L2-3S-100X50-GWB/HWS	69.09	6	3
L3-2S-50X30-GWB/HWS	27.03	2	1
L3-3S-50X30-GWB/HWS	40.17	2	1
L3-2S-80X30-GWB/HWS	32.64	2	1
L3-3S-80X30-GWB/HWS	47.09	2	1
L4-2S-60X50-GWB/HWS	36.44	4	2
L4-3S-60X30-GWB/HWS	54.59	4	2
L4-2S-100X50-GWB/HWS	49.4	4	2
L4-3S-100X50-GWB/HWS	73.85	4	2

Table B.1 cont. Component quantities of SWOF building archetypes (for entire building)

Building Name	Non-Structural Components					
	Prefabricated Steel Stair No Seismic Joint	Potable Water Piping	Heating Water Piping	Heating Water Pipe Bracing	Sanitary Waste Piping	HVAC Ducting
L1-2S-60X30-GWB/HWS	4	1.306	0.036	0.018	0.872	0.18
L1-3S-60X30-GWB/HWS	6	1.959	0.054	0.027	1.308	0.27
L1-2S-100X30-GWB/HWS	4	2.178	0.06	0.03	1.452	0.3
L1-3S-100X30-GWB/HWS	6	3.267	0.09	0.045	2.178	0.45
L2-2S-60X50-GWB/HWS	4	2.178	0.06	0.03	1.452	0.3
L2-3S-60X50-GWB/HWS	6	3.267	0.09	0.045	2.178	0.45
L2-2S-100X50-GWB/HWS	4	3.63	0.1	0.05	2.42	0.5
L2-3S-100X50-GWB/HWS	6	5.445	0.15	0.075	3.63	0.75
L3-2S-50X30-GWB/HWS	4	1.09	0.032	0.016	0.728	0.15

L3-3S-50X30-GWB/HWS	6	1.635	0.048	0.024	1.092	0.225
L3-2S-80X30-GWB/HWS	4	2.904	0.08	0.04	1.936	0.4
L3-3S-80X30-GWB/HWS	6	2.613	0.072	0.036	1.74	0.36
L4-2S-60X50-GWB/HWS	4	2.178	0.06	0.03	1.452	0.3
L4-3S-60X30-GWB/HWS	6	3.267	0.09	0.045	2.178	0.45
L4-2S-100X50-GWB/HWS	4	3.63	0.1	0.05	2.42	0.5
L4-3S-100X50-GWB/HWS	6	5.445	0.15	0.075	3.63	0.75

Table B.1 cont. Component quantities of SWOF building archetypes (for entire building)

Building Name	Non-Structural Components				
	Fire Sprinkler Water Piping	Fire Sprinkler Drop	Low Voltage Switchgear	Wall Partition, Wood Stud	Clay tile roof
L1-2S-60X30-GWB/HWS	0.792	0.432	2	5.32	18
L1-3S-60X30-GWB/HWS	1.188	0.648	3	8.87	18
L1-2S-100X30-GWB/HWS	1.32	0.72	2	13.9	30
L1-3S-100X30-GWB/HWS	1.98	1.08	3	11.57	30
L2-2S-60X50-GWB/HWS	1.32	0.72	2	8.5	30
L2-3S-60X50-GWB/HWS	1.98	1.08	3	13.42	30
L2-2S-100X50-GWB/HWS	2.2	1.2	2	10.58	50
L2-3S-100X50-GWB/HWS	3.3	1.8	3	16.54	50
L3-2S-50X30-GWB/HWS	0.66	0.36	2	3.96	15
L3-3S-50X30-GWB/HWS	0.99	0.54	3	6.74	15
L3-2S-80X30-GWB/HWS	1.76	0.96	2	5.14	24
L3-3S-80X30-GWB/HWS	1.584	0.864	3	7.91	24
L4-2S-60X50-GWB/HWS	1.32	0.72	2	8.3	30
L4-3S-60X30-GWB/HWS	1.98	1.08	3	13.5	30
L4-2S-100X50-GWB/HWS	2.2	1.2	2	10.58	50
L4-3S-100X50-GWB/HWS	3.3	1.8	3	17.06	50

Table B.2 Component information of SWOF building archetypes

Component Name	Fragility Unit	Demand Parameter	Direction
Light framed wood lateral walls	100 SF	Story Drift Ratio	Directional
Steel Column Base Plates	1 EA	Story Drift Ratio	Directional
Welded Steel Moment Connection	1 EA	Story Drift Ratio	Directional
Prefabricated steel stair no seismic joint	1 EA	Story Drift Ratio	Directional
Potable Water Piping	1000 LF	Peak Floor Acceleration	Non-directional
Heating Water Piping	1000 LF	Peak Floor Acceleration	Non-directional
Heating Water Pipe Bracing	1000 LF	Peak Floor Acceleration	Non-directional
Sanitary Waste Piping	1000 LF	Peak Floor Acceleration	Non-directional
HVAC Ducting	1000 LF	Peak Floor Acceleration	Non-directional
Fire Sprinkler Water Piping	1000 LF	Peak Floor Acceleration	Non-directional
Fire Sprinkler Drop	100 EA	Peak Floor Acceleration	Non-directional
Low Voltage Switchgear	1 EA	Peak Floor Acceleration	Non-directional
Wall Partition, Wood Stud	100 LF	Story Drift Ratio	Directional
Clay tile roof	100 SF	Peak Floor Acceleration	Non-directional

Appendix C. Details of SWOF Building Retrofit Elements for Four Alternatives

Table C.1 Retrofit elements information for SWOF buildings

Building ID	Basic Ordinance						FEMA P807					
	F1		F2		F3		F1		F2		F3	
	Beam	Column	Beam	Column	Beam	Column	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15	8X13	8X13	8X13	8X13	8X13	8X13
L1-3S-60X30-GWB	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22	10X12	10X12	8X13	8X13	8X13	8X13
L1-2S-100X30-GWB	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15	10X12	10X12	10X12	10X12	10X12	10X12
L1-3S-100X30-GWB	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22	10X12	10X12	10X17	10X17	10X17	10X17
L1-2S-60X30-HWS	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15	10X12	10X12	10X12	10X12	10X12	10X12
L1-3S-60X30-HWS	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22	10X19	10X19	10X19	10X19	10X19	10X19
L1-2S-100X30-HWS	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15	10X22	10X22	10X19	10X19	10X19	10X19
L1-3S-100X30-HWS	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22	10X22	10X22	10X26	10X26	10X26	10X26
L2-2S-60X50-GWB	W10x15	W10x26	W8x10	W8x15	W10x15	W10x26	8X13	8X13	8X13	8X13	Adequate	Adequate
L2-3S-60X50-GWB	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39	10X17	10X17	10X17	10X17	8X13	8X13
L2-2S-100X50-GWB	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39	10X12	10X12	10X12	10X12	10X12	10X12
L2-3S-100X50-GWB	W10x33	W10x54	W8x13	W10x22	W10x33	W10x54	10X17	10X17	10X17	10X17	8X13	8X13
L2-2S-60X50-HWS	W10x15	W10x26	W8x10	W8x15	W10x15	W10x26	10X19	10X19	10X19	10X19	8X13	8X13
L2-3S-60X50-HWS	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39	10X30	10X30	10X30	10X30	8X13	8X13
L2-2S-100X50-HWS	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39	10X22	10X22	10X22	10X22	10X12	10X12
L2-3S-100X50-HWS	W10x33	W10x54	W8x13	W10x22	W10x33	W10x54	10X26	10X26	10X26	10X26	8X13	8X13

L3-2S-50X30-GWB	W8x10	W10x17					8X13	8X13				
L3-3S-50X30-GWB	W8x13	W10x22					8X13	8X13				
L3-2S-80X30-GWB	W8x10	W10x17					8X10	8X10				
L3-3S-80X30-GWB	W8x13	W10x22			NA		8X10	8X10			NA	
L3-2S-50X30-HWS	W8x10	W10x17					10X19	10X19				
L3-3S-60X30-HWS	W8x13	W10x22					10X22	10X22				
L3-2S-80X30-HWS	W8x10	W10x17					10X22	10X22				
L3-3S-80X30-HWS	W8x13	W10x22					8X10	8X10				
L4-2S-60X50-GWB	W8x10	W8x15	W8x10	W8x15			8X13	8X13	8X10	8X10		
L4-3S-60X30-GWB	W8x10	W10x15	W8x10	W10x15			10X19	10X19	8X13	8X13		
L4-2S-100X50-GWB	W8x10	W8x15	W8x10	W8x15			8X10	8X10	8X10	8X10		
L4-3S-100X50-GWB	W8x10	W10x15	W8x10	W10x15			8X10	8X10	8X13	8X13		
L4-2S-60X50-HWS	W8x10	W8x15	W8x10	W8x15	NA		8X13	8X13	10X19	10X19	NA	
L4-3S-60X50-HWS	W8x10	W10x15	W8x10	W10x15			10X30	10X30	10X19	10X19		
L4-2S-100X50-HWS	W8x10	W8x15	W8x10	W8x15			10X19	10X19	10X19	10X19		
L4-3S-100X50-HWS	W8x10	W10x15	W8x10	W10x15			10X12	10X12	8X13	8X13		

Table C.1 cont. Retrofit elements information for SWOF buildings

Building ID	IEBC A4						ASCE 41-13					
	F1		F2		F3		F1		F2		F3	
	Beam	Column	Beam	Column	Beam	Column	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W10x12	W10x49	W8x13	W10x22	W8x13	W10x22	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L1-3S-60X30-GWB	W10x22	W10x39	W10x12	W10x49	W10x12	W10x49	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L1-2S-100X30-GWB	W10x33	W10x54	W10x12	W10x39	W10x12	W10x39	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L1-3S-100X30-GWB	W10x33	W10x54	W10x33	W10x54	W10x33	W10x54	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L1-2S-60X30-HWS	W10x12	W10x49	W8x13	W10x22	W8x13	W10x22	W10x12	W10x30	W10x12	W10x17	W10x12	W10x17
L1-3S-60X30-HWS	W10x22	W10x39	W10x12	W10x49	W10x12	W10x49	W10x12	W10x30	W10x12	W10x17	W10x12	W10x17
L1-2S-100X30-HWS	W10x33	W10x54	W10x12	W10x39	W10x12	W10x39	W10x12	W10x30	W10x12	W10x33	W10x12	W10x33
L1-3S-100X30-HWS	W10x33	W10x54	W10x33	W10x54	W10x33	W10x54	W10x12	W10x30	W10x12	W10x33	W10x12	W10x33
L2-2S-60X50-GWB	W10x12	W10x45	W10x33	W10x54	W10x12	W10x45	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L2-3S-60X50-GWB	W10x33	W10x54	W10x33	W12x58	W10x33	W10x54	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L2-2S-100X50-GWB	W10x33	W10x54	W10x33	W12x58	W10x33	W10x54	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L2-3S-100X50-GWB	W12x35	W14x61	W10x33	W12x58	W12x35	W14x61	W8x10	W8x10	W8x10	W8x10	W8x10	W8x10
L2-2S-60X50-HWS	W10x12	W10x45	W10x33	W10x54	W10x12	W10x45	W10x12	W10x12	W10x12	W10x12	W10x12	W10x12
L2-3S-60X50-HWS	W10x33	W10x54	W10x33	W12x58	W10x33	W10x54	W10x12	W10x12	W10x12	W10x12	W10x12	W10x12
L2-2S-100X50-HWS	W10x33	W10x54	W10x33	W12x58	W10x33	W10x54	W10x12	W10x30	W10x12	W10x30	W10x12	W10x30
L2-3S-100X50-HWS	W12x35	W14x61	W10x33	W12x58	W12x35	W14x61	W10x12	W10x30	W10x12	W10x30	W10x12	W10x30
L3-2S-50X30-GWB	W10x12	W10x49					W8x10	W8x10				
L3-3S-50X30-GWB	W10x33	W10x54					W8x10	W8x10				
L3-2S-80X30-GWB	W10x33	W12x58					W8x10	W8x10				
L3-3S-80X30-GWB	W12x35	W14x61		NA			W8x10	W8x10		NA		
L3-2S-50X30-HWS	W10x12	W10x49					W10x12	W10x17				
L3-3S-60X30-HWS	W10x33	W10x54					W10x12	W10x17				
L3-2S-80X30-HWS	W10x33	W12x58					W10x12	W10x12				

L3-3S-80X30-HWS	W12x35	W14x61				W10x12	W10x12			
L4-2S-60X50-GWB	W12x35	W14x61	W12x35	W14x61		W8x10	W8x10	W8x10	W8x10	
L4-3S-60X30-GWB	W12x35	W14x61	W12x35	W14x61		W8x10	W8x10	W8x10	W8x10	
L4-2S-100X50-GWB	W10x33	W10x54	W10x33	W10x54		W8x10	W8x10	W8x10	W8x10	
L4-3S-100X50-GWB	W12x35	W14x61	W12x35	W14x61	NA	W8x10	W8x10	W8x10	W8x10	NA
L4-2S-60X50-HWS	W12x35	W14x61	W12x35	W14x61		W10x12	W10x39	W10x12	W10x17	
L4-3S-60X50-HWS	W12x35	W14x61	W12x35	W14x61		W10x12	W10x39	W10x12	W10x17	
L4-2S-100X50-HWS	W10x33	W10x54	W10x33	W10x54		W10x12	W10x33	W10x12	W10x30	
L4-3S-100X50-HWS	W12x35	W14x61	W12x35	W14x61		W10x12	W10x33	W10x12	W10x30	

Appendix D. Location and Sizes of Two Ordinance Guideline Based SWOF Building Retrofit Elements

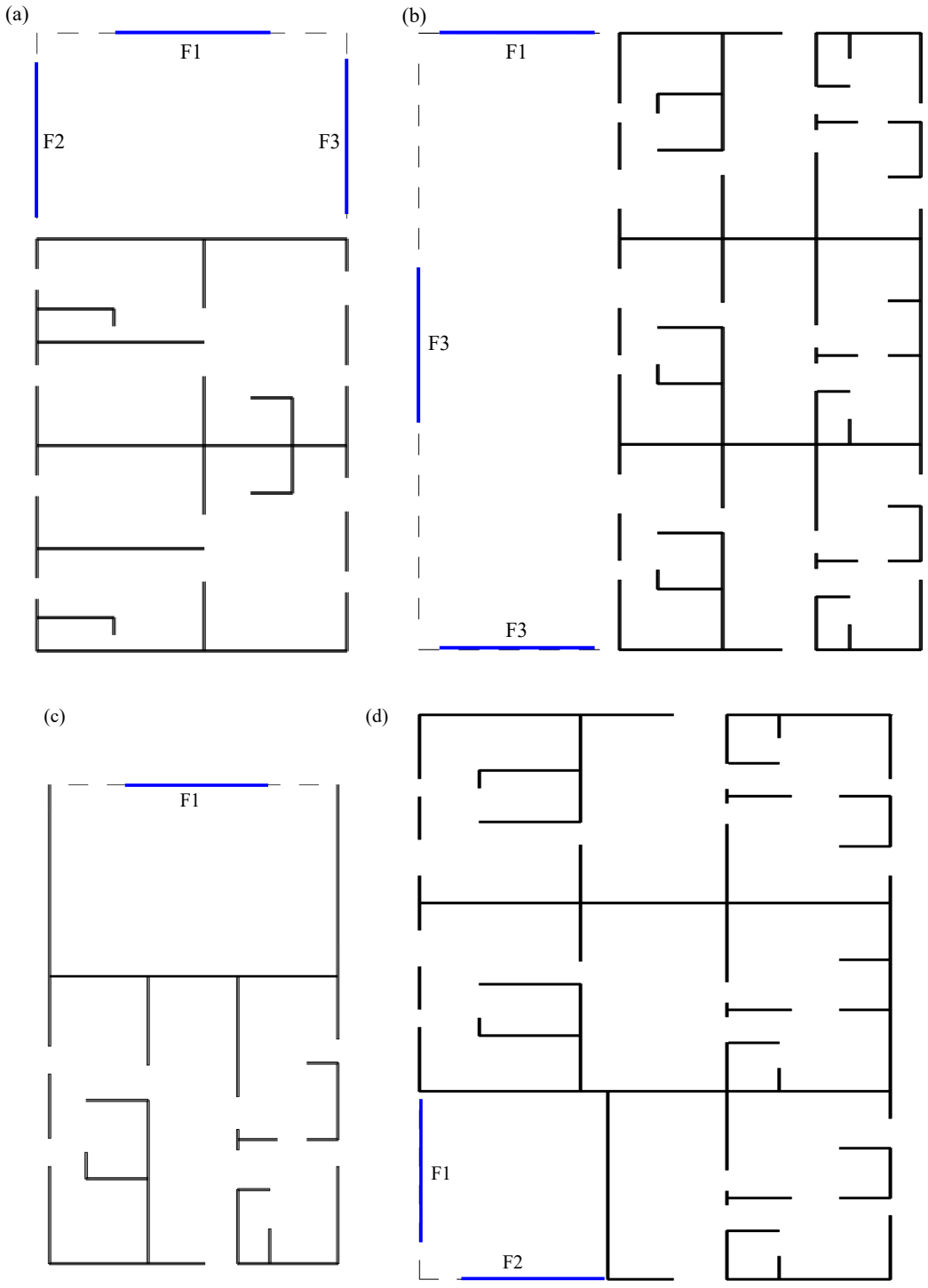


Figure D.1 Location of moment frames used in retrofit: (a) L1, (b) L2, (c) L3 and (d) L4

Table D.1 Frame sizes for median S_{MS} retrofit

Building ID	Median SMS Retrofit					
	F1		F2		F3	
	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15
L1-3S-60X30-GWB	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22
L1-2S-100X30-GWB	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15
L1-3S-100X30-GWB	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22
L1-2S-60X30-HWS	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15
L1-3S-60X30-HWS	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22
L1-2S-100X30-HWS	W8x10	W10x15	W8x10	W10x15	W8x10	W10x15
L1-3S-100X30-HWS	W8x13	W10x22	W10x15	W10x22	W10x15	W10x22
L2-2S-60X50-GWB	W10x15	W10x26	W8x10	W8x15	W10x15	W10x26
L2-3S-60X50-GWB	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39
L2-2S-100X50-GWB	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39
L2-3S-100X50-GWB	W10x33	W10x54	W8x13	W10x22	W10x33	W10x54
L2-2S-60X50-HWS	W10x15	W10x26	W8x10	W8x15	W10x15	W10x26
L2-3S-60X50-HWS	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39
L2-2S-100X50-HWS	W10x22	W10x39	W8x10	W10x15	W10x22	W10x39
L2-3S-100X50-HWS	W10x33	W10x54	W8x13	W10x22	W10x33	W10x54
L3-2S-50X30-GWB	W8x10	W10x17	NA			
L3-3S-50X30-GWB	W8x13	W10x22				
L3-2S-80X30-GWB	W8x10	W10x17				
L3-3S-80X30-GWB	W8x13	W10x22				
L3-2S-50X30-HWS	W8x10	W10x17				
L3-3S-60X30-HWS	W8x13	W10x22				
L3-2S-80X30-HWS	W8x10	W10x17				
L3-3S-80X30-HWS	W8x13	W10x22				
L4-2S-60X50-GWB	W8x10	W8x15	W8x10	W8x15	NA	
L4-3S-60X30-GWB	W8x10	W10x15	W8x10	W10x15		
L4-2S-100X50-GWB	W8x10	W8x15	W8x10	W8x15		
L4-3S-100X50-GWB	W8x10	W10x15	W8x10	W10x15		
L4-2S-60X50-HWS	W8x10	W8x15	W8x10	W8x15		
L4-3S-60X50-HWS	W8x10	W10x15	W8x10	W10x15		
L4-2S-100X50-HWS	W8x10	W8x15	W8x10	W8x15		
L4-3S-100X50-HWS	W8x10	W10x15	W8x10	W10x15		

Table D.2 Frame sizes for maximum S_{MS} retrofit

Building ID	Maximum S_{MS} Retrofit					
	F1		F2		F3	
	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W10x15	W8x18	W10x15	W10x17	W10x15	W10x17
L1-3S-60X30-GWB	W10x15	W10x26	W10x17	W10x30	W10x17	W10x30
L1-2S-100X30-GWB	W10x15	W8x18	W10x15	W10x17	W10x15	W10x17
L1-3S-100X30-GWB	W10x15	W10x26	W10x17	W10x30	W10x17	W10x30
L1-2S-60X30-HWS	W10x15	W8x18	W10x15	W10x17	W10x15	W10x17
L1-3S-60X30-HWS	W10x15	W10x26	W10x17	W10x30	W10x17	W10x30
L1-2S-100X30-HWS	W10x15	W8x18	W10x15	W10x17	W10x15	W10x17
L1-3S-100X30-HWS	W10x15	W10x26	W10x17	W10x30	W10x17	W10x30
L2-2S-60X50-GWB	W10x15	W10x45	W8x10	W10x15	W10x15	W10x45
L2-3S-60X50-GWB	W10x33	W10x54	W10x15	W10x17	W10x33	W10x54
L2-2S-100X50-GWB	W10x33	W10x54	W10x15	W10x17	W10x33	W10x54
L2-3S-100X50-GWB	W10x45	W10x68	W10x15	W10x26	W10x45	W10x68
L2-2S-60X50-HWS	W10x15	W10x45	W8x10	W10x15	W10x15	W10x45
L2-3S-60X50-HWS	W10x33	W10x54	W10x15	W10x17	W10x33	W10x54
L2-2S-100X50-HWS	W10x33	W10x54	W10x15	W10x17	W10x33	W10x54
L2-3S-100X50-HWS	W10x45	W10x68	W10x15	W10x26	W10x45	W10x68
L3-2S-50X30-GWB	W8x13	W10x22	NA			
L3-3S-50X30-GWB	W10x22	W10x26				
L3-2S-80X30-GWB	W8x13	W10x22				
L3-3S-80X30-GWB	W10x22	W10x26				
L3-2S-50X30-HWS	W8X13	W10x22				
L3-3S-60X30-HWS	W10x22	W10x26				
L3-2S-80X30-HWS	W8X13	W10x22				
L3-3S-80X30-HWS	W10x22	W10x26				
L4-2S-60X50-GWB	W8x10	W10x15	W8x10	W10x15	NA	
L4-3S-60X30-GWB	W10x15	W10x17	W10x15	W10x17		
L4-2S-100X50-GWB	W8x10	W10x15	W8x10	W10x15		
L4-3S-100X50-GWB	W10x15	W10x17	W10x15	W10x17		
L4-2S-60X50-HWS	W8x10	W10x15	W8x10	W10x15		
L4-3S-60X50-HWS	W10x15	W10x17	W10x15	W10x17		
L4-2S-100X50-HWS	W8x10	W10x15	W8x10	W10x15		
L4-3S-100X50-HWS	W10x15	W10x17	W10x15	W10x17		

Appendix E. Location and Sizes of Eight Ordinance Guideline Based SWOF Building Retrofit Elements

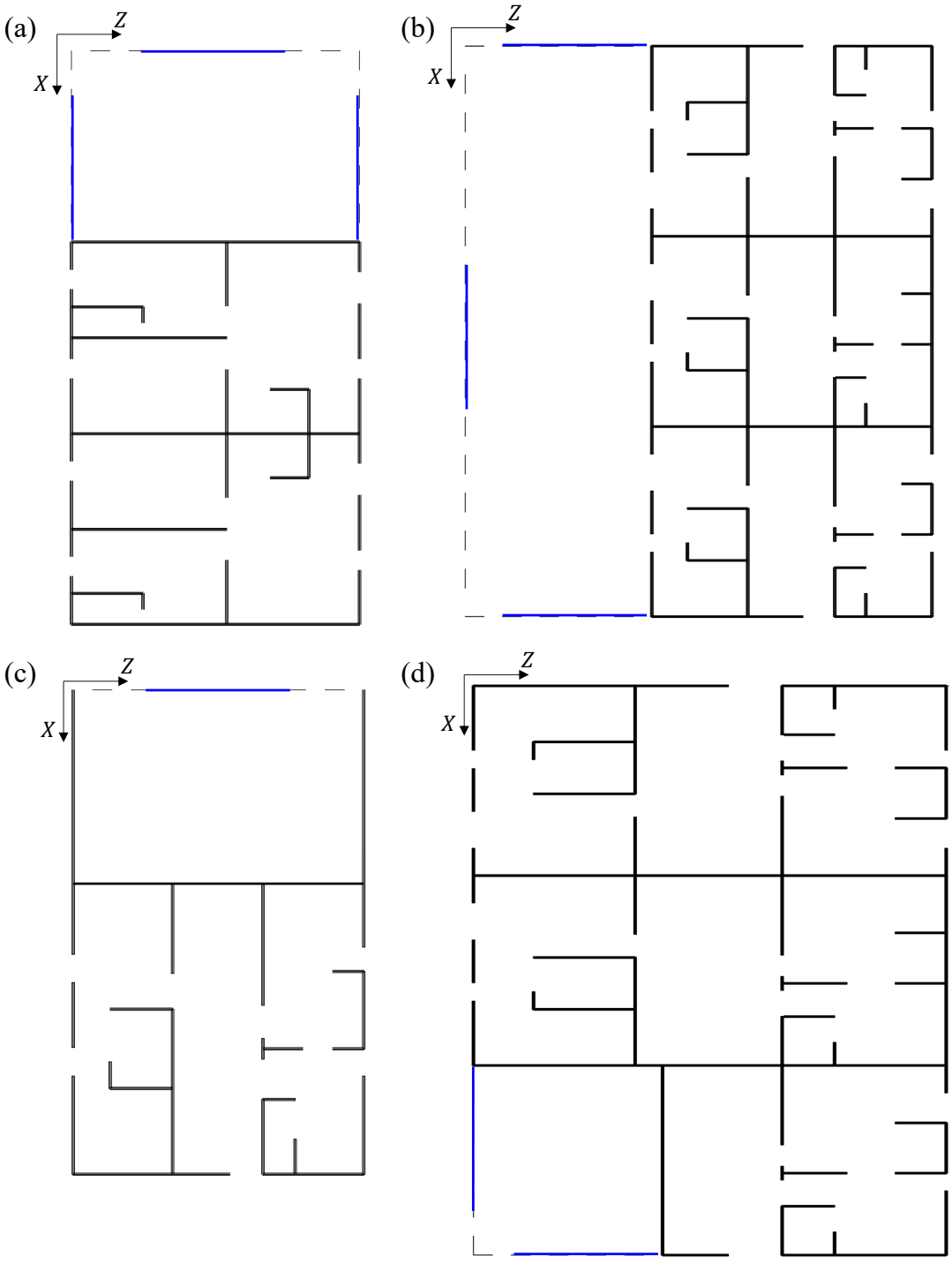


Figure E.1 Location of moment frames used in retrofit: (a) L1, (b) L2, (c) L3 and (d) L4

Table E.1 Ordinance Retrofit Ordinary Moment Frame in the X Direction[1]

Building Index ID	S_{MS}							
	1.6g		1.8g		2g		2.2g	
	Beam	Column[2]	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L1-2S-60X30-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L1-3S-60X30-GWB	W8x13	W10x17	W8x15	W10x19	W8x15	W10x22	W10x12	W12x19
L1-3S-60X30-HWS	W8x13	W10x17	W8x15	W10x19	W8x15	W10x22	W10x12	W12x19
L1-2S-100X30-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L1-2S-100X30-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L1-3S-100X30-GWB	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L1-3S-100X30-HWS	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L2-2S-60X50-GWB	W8x18	W10x19	W8x18	W10x22	W10x15	W12x19	W10x12	W12x22
L2-2S-60X50-HWS	W8x18	W10x19	W8x18	W10x22	W10x15	W12x19	W10x12	W12x22
L2-3S-60X50-GWB	W10x12	W12x26	W10x15	W12x26	W10x12	W12x35	W10x12	W12x35
L2-3S-60X50-HWS	W10x12	W12x26	W10x15	W12x26	W10x12	W12x35	W10x12	W12x35
L2-2S-100X50-GWB	W10x12	W12x26	W10x15	W12x26	W10x12	W12x35	W10x17	W12x35
L2-2S-100X50-HWS	W10x12	W12x26	W10x15	W12x26	W10x12	W12x35	W10x17	W12x35
L2-3S-100X50-GWB	W10x17	W12x35	W10x19	W12x40	W10x19	W12x45	W10x22	W12x50
L2-3S-100X50-HWS	W10x17	W12x35	W10x19	W12x40	W10x19	W12x45	W10x22	W12x50
L3-2S-50X30-GWB								
L3-2S-50X30-HWS								
L3-3S-50X30-GWB								
L3-3S-50X30-HWS								
L3-2S-80X30-GWB								
L3-2S-80X30-HWS								
L3-3S-80X30-GWB								
L3-3S-80X30-HWS								
L4-2S-60X50-GWB	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13	W8x10	W8x15
L4-2S-60X50-HWS	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13	W8x10	W8x15
L4-3S-60X50-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-3S-60X50-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-2S-100X50-GWB	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13
L4-2S-100X50-HWS	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13
L4-3S-100X50-GWB	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-3S-100X50-HWS	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18

[1]The table summarizes the OMF component sizes in X direction specified in Figure A1. If more than 1 OMFs presented in X direction, they are of the same size.

[2]Each OMF has 2 columns that of the same size.

Table E1 cont. Ordinance Retrofit Ordinary Moment Frame in the X Direction

Building Index ID	SMS							
	2.4g		2.6g		2.8g		3g	
	Beam	Column	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L1-2S-60X30-HWS	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L1-3S-60X30-GWB	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26	W10x12	W12x26
L1-3S-60X30-HWS	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26	W10x12	W12x26
L1-2S-100X30-GWB	W8x10	W10x17	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L1-2S-100X30-HWS	W8x10	W10x17	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L1-3S-100X30-GWB	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L1-3S-100X30-HWS	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L2-2S-60X50-GWB	W10x12	W12x26	W10x15	W12x26	W10x15	W12x26	W10x12	W12x35
L2-2S-60X50-HWS	W10x12	W12x26	W10x15	W12x26	W10x15	W12x26	W10x12	W12x35
L2-3S-60X50-GWB	W10x17	W12x35	W10x19	W12x40	W10x19	W12x40	W10x19	W12x45
L2-3S-60X50-HWS	W10x17	W12x35	W10x19	W12x40	W10x19	W12x40	W10x19	W12x45
L2-2S-100X50-GWB	W10x19	W12x40	W10x19	W12x40	W10x19	W12x45	W10x22	W12x50
L2-2S-100X50-HWS	W10x19	W12x40	W10x19	W12x40	W10x19	W12x45	W10x22	W12x50
L2-3S-100X50-GWB	W10x26	W12x50	W10x30	W12x50	W10x33	W12x53	W10x39	W12x58
L2-3S-100X50-HWS	W10x26	W12x50	W10x30	W12x50	W10x33	W12x53	W10x39	W12x58
L3-2S-50X30-GWB								
L3-2S-50X30-HWS								
L3-3S-50X30-GWB								
L3-3S-50X30-HWS								
L3-2S-80X30-GWB								
L3-2S-80X30-HWS								
L3-3S-80X30-GWB								
L3-3S-80X30-HWS								
L4-2S-60X50-GWB	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-2S-60X50-HWS	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-3S-60X50-GWB	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L4-3S-60X50-HWS	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L4-2S-100X50-GWB	W8x10	W8x15	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-2S-100X50-HWS	W8x10	W8x15	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-3S-100X50-GWB	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L4-3S-100X50-HWS	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19

Table E.2 Ordinance Retrofit Ordinary Moment Frame in the Z Direction[3]

Building Index ID	SMS							
	1.6 (g)		1.8 (g)		2 (g)		2.2 (g)	
	Beam	Column	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L1-2S-60X30-HWS	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L1-3S-60X30-GWB	W8x10	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L1-3S-60X30-HWS	W8x10	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L1-2S-100X30-GWB	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L1-2S-100X30-HWS	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L1-3S-100X30-GWB	W8x10	W10x15	W8x10	W10x17	W8x13	W10x19	W8x15	W10x19
L1-3S-100X30-HWS	W8x10	W10x15	W8x10	W10x17	W8x13	W10x19	W8x15	W10x19
L2-2S-60X50-GWB	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x15
L2-2S-60X50-HWS	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x15
L2-3S-60X50-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L2-3S-60X50-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L2-2S-100X50-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L2-2S-100X50-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L2-3S-100X50-GWB	W8x10	W10x17	W8x13	W10x17	W8x15	W10x19	W8x18	W10x19
L2-3S-100X50-HWS	W8x10	W10x17	W8x13	W10x17	W8x15	W10x19	W8x18	W10x19
L3-2S-50X30-GWB	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17
L3-2S-50X30-HWS	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17
L3-3S-50X30-GWB	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L3-3S-50X30-HWS	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L3-2S-80X30-GWB	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17
L3-2S-80X30-HWS	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17
L3-3S-80X30-GWB	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L3-3S-80X30-HWS	W8x13	W10x17	W8x13	W10x19	W8x18	W10x19	W8x15	W10x22
L4-2S-60X50-GWB	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13	W8x10	W8x15
L4-2S-60X50-HWS	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13	W8x10	W8x15
L4-3S-60X50-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-3S-60X50-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-2S-100X50-GWB	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13
L4-2S-100X50-HWS	W8x10	W8x10	W8x10	W8x10	W8x10	W8x13	W8x10	W8x13
L4-3S-100X50-GWB	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-3S-100X50-HWS	W8x10	W8x13	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18

[3]The table summarizes the OMF component sizes in Z direction specified in Figure A1. If more than 1 OMFs presented in Z direction, they are of the same size.

Table E2 cont. Ordinance Retrofit Ordinary Moment Frame in the Z Direction

Building Index ID	SMS							
	2.4 (g)		2.6 (g)		2.8 (g)		3 (g)	
	Beam	Column	Beam	Column	Beam	Column	Beam	Column
L1-2S-60X30-GWB	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L1-2S-60X30-HWS	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L1-3S-60X30-GWB	W8x18	W10x22	W10x12	W12x19	W10x12	W12x22	W10x12	W12x22
L1-3S-60X30-HWS	W8x18	W10x22	W10x12	W12x19	W10x12	W12x22	W10x12	W12x22
L1-2S-100X30-GWB	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17
L1-2S-100X30-HWS	W8x10	W8x18	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17
L1-3S-100X30-GWB	W8x18	W10x19	W8x18	W10x22	W10x12	W12x19	W10x15	W12x19
L1-3S-100X30-HWS	W8x18	W10x19	W8x18	W10x22	W10x12	W12x19	W10x15	W12x19
L2-2S-60X50-GWB	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18
L2-2S-60X50-HWS	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18
L2-3S-60X50-GWB	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L2-3S-60X50-HWS	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L2-2S-100X50-GWB	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L2-2S-100X50-HWS	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L2-3S-100X50-GWB	W8x18	W10x22	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22
L2-3S-100X50-HWS	W8x18	W10x22	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22
L3-2S-50X30-GWB	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L3-2S-50X30-HWS	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L3-3S-50X30-GWB	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L3-3S-50X30-HWS	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L3-2S-80X30-GWB	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L3-2S-80X30-HWS	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19	W8x18	W10x19
L3-3S-80X30-GWB	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L3-3S-80X30-HWS	W10x12	W12x19	W10x15	W12x19	W10x12	W12x22	W10x12	W12x26
L4-2S-60X50-GWB	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-2S-60X50-HWS	W8x10	W8x18	W8x10	W8x18	W8x10	W8x18	W8x10	W10x15
L4-3S-60X50-GWB	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L4-3S-60X50-HWS	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19	W8x15	W10x19
L4-2S-100X50-GWB	W8x10	W8x15	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-2S-100X50-HWS	W8x10	W8x15	W8x10	W8x15	W8x10	W8x18	W8x10	W8x18
L4-3S-100X50-GWB	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19
L4-3S-100X50-HWS	W8x10	W10x15	W8x10	W10x17	W8x13	W10x17	W8x13	W10x19

References

Abu-Mostafa YS (1989) The Vapnik-Chervonenkis dimension: Information versus complexity in learning. Neural Computation 1:312–317

Alon N, Ben-David S, Cesa-Bianchi N, Haussler D (1997) Scale-sensitive dimensions, uniform convergence, and learnability. Journal of the ACM (JACM) 44:615–631

Amana E (1967) Theoretical and experimental studies on nailed and glued plywood stressed-skin components, Part I. Theoretical study. J of the Institute of Wood Sci 4:43–69

Andrew Blankstein, Monica Alba (2014) Why do so few California homeowners have earthquake insurance? NBC News

ASCE (2016) Minimum design loads and associated criteria for buildings and other structures. American society of civil engineers

ASCE (2017) ASCE 41-17: Seismic Evaluation and Retrofit Rehabilitation of Existing Buildings. Proceedings of the SEAOC

ASCE (2013) ASCE 41-13: Seismic evaluation and retrofit rehabilitation of existing buildings. Proceedings of the SEAOC

Auer P (2002) Using confidence bounds for exploitation-exploration trade-offs. Journal of Machine Learning Research 3:397–422

Bahmani P, van de Lindt JW, Gershfeld M, et al (2016) Experimental seismic behavior of a full-scale four-story soft-story wood-frame building with retrofits. I: Building design, retrofit methodology, and numerical validation. Journal of Structural Engineering 142:E4014003

Blaney C, Cobeen C, Filiatrault A, et al (2018) Vulnerability Based Seismic Assessment and Retrofit of One-and Two-Family Dwellings (ATC-110 Project)

Boore DM, Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s. Earthquake Spectra 24:99–138

Box GE, Wilson KB (1951) On the experimental attainment of optimum conditions. Journal of the royal statistical society: Series b (Methodological) 13:1–38

Breiman L (2001) Random forests. Machine learning 45:5–32

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. CRC press

Brochu E, Cora VM, De Freitas N (2010) A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:10122599

Buckalew J, McDonald B, McCormick D, et al (2015) Example Case Studies of Soft-Story Retrofits Using the San Francisco Ordinance. In: Improving the Seismic Performance of Existing Buildings and Other Structures 2015. pp 548–559

Building Seismic Safety Council (1997) NEHRP guidelines for the seismic rehabilitation of buildings

Burton H, Rad AR, Yi Z, et al (2019) Seismic collapse performance of Los Angeles soft, weak, and open-front wall line woodframe structures retrofitted using different procedures. Bulletin of Earthquake Engineering 17:2059–2091

Burton HV, Sreekumar S, Sharma M, Sun H (2017) Estimating aftershock collapse vulnerability using mainshock intensity, structural response and physical damage indicators. Structural safety 68:85–96

California Earhtquake Authority (2019) California Earthquake Authority Premium Discounts. https://www.earthquakeauthority.com/California-Earthquake-Insurance-Policies/Earthquake-Insurance-Policy-Premium-Discounts

Cardone D, Gesualdi G, Perrone G (2019) Cost-benefit analysis of alternative retrofit strategies for RC frame buildings. Journal of Earthquake Engineering 23:208–241

Caughey T, O'Kelly ME (1965) Classical normal modes in damped linear dynamic systems

Chen R, Branum DM, Wills CJ (2011) Hazus scenario and annualized earthquake loss estimation for California. California Geological Survey, California Department of Conservation

Chen R, Jaiswal K, Bausch D, et al (2016) Annualized earthquake loss estimates for California and their sensitivity to site amplification. Seismological Research Letters 87:1363–1372

Chen T (2014) Introduction to boosted trees. University of Washington Computer Science 22:115

Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp 785–794

Cheung CK, Itani RY (1983) Analysis of sheathed wood-stud walls. In: Electronic computation. ASCE, pp 683–696

Christovasilis IP, Filiatrault A, Constantinou MC, Wanitkorkul A (2009) Incremental dynamic analysis of woodframe buildings. Earthquake engineering & structural dynamics 38:477–496

Chui Y, Smith I (1989) Quantifying damping in structural timber components

Cobeen K, Mahdavifar V, Hutchinson T, et al (2020) Large-Component Seismic Testing for Existing and Retrofitted Single-Family Wood-Frame Dwellings, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single

Family Wood-Frame Buildings" Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Cornell C, Krawinkler H (2000) Progress and challenges in seismic performance assessment. PEER Center News, Spring 2000

Council (ATC) AT (1996) Methodology for evaluation and upgrade of reinforced concrete buildings. California Seismic Safety Commission Sacramento, CA

Council (US) BSS, Council AT (1997) NEHRP guidelines for the seismic rehabilitation of buildings. Federal Emergency Management Agency

Crowley H, Bommer JJ (2006) Modelling seismic hazard in earthquake loss models with spatially distributed exposure. Bulletin of Earthquake Engineering 4:249–273

Dinehart DW, Shenton III HW (2000) Model for dynamic analysis of wood frame shear walls. Journal of engineering mechanics 126:899–908

Dolan J, Madsen B (1992) Monotonic and cyclic nail connection tests. Canadian Journal of Civil Engineering 19:97–104

Dolce M, Lorusso V, Masi A (1992) Inelastic seismic response of building structures with flexible diaphragm. In: Proceedings of the Tenth World Conference on Earthquake Engineering, Madrid. pp 3967–3972

Dong Y, Frangopol DM, Saydam D (2014) Pre-earthquake multi-objective probabilistic retrofit optimization of bridge networks based on sustainability. Journal of Bridge Engineering 19:04014018

Donoho DL, Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. Journal of the american statistical association 90:1200–1224

Dudley RM, Giné E, Zinn J (1991) Uniform and universal Glivenko-Cantelli classes. Journal of Theoretical Probability 4:485–510

Ellingwood B (1980) Development of a probability based load criterion for American National Standard A58: Building code requirements for minimum design loads in buildings and other structures. US Department of Commerce, National Bureau of Standards

Ellingwood BR, Rosowsky DV, Pang W (2008) Performance of light-frame wood residential construction subjected to earthquakes in regions of moderate seismicity. Journal of structural engineering 134:1353–1363

Falk RH, Itani RY (1989) Finite element modeling of wood diaphragms. Journal of Structural Engineering 115:543–559

FEMA (2012a) Seismic Performance Assessment of Buildings: Volume 1—Methodology (P-58-1)

FEMA (2018) Vulnerability-Based Seismic Assessment and Retrofit of One- and Two-Family Dwellings, Volume 1 – Prestandard. Applied Technology Council for the California Earthquake Authority (Sacramento, CA) and the Federal Emergency Management Agency (Washington, DC)

FEMA (2009) Quantification of Building Seismic Performance Factors

FEMA (2012b) Seismic evaluation and retrofit of multi-unit wood-frame buildings with weak first stories

FEMA (2000) Recommended seismic design criteria for new steel moment-frame buildings. Fema Washington, DC.

FEMA (2012c) Simplified Seismic Assessment of Detached, Single-Family, Wood-Frame Dwellings

Field EH, Dawson T, Felzer K, et al (2009) Uniform California earthquake rupture forecast, version 2 (UCERF 2). Bulletin of the Seismological Society of America 99:2053–2107

Field EH, Jordan TH, Cornell CA (2003) OpenSHA: A developing community-modeling environment for seismic hazard analysis. Seismological Research Letters 74:406–419

Field EH, Seligson HA, Gupta N, et al (2005) Loss estimates for a Puente Hills blind-thrust earthquake in Los Angeles, California. Earthquake Spectra 21:329–338

Filiatrault A, Folz B (2002) Performance-based seismic design of wood framed buildings. Journal of Structural Engineering 128:39–47

Foliente GC (1995) Hysteresis modeling of wood joints and structural systems. Journal of Structural Engineering 121:1013–1022

Folz B, Filiatrault A (2001) Cyclic analysis of wood shear walls. Journal of Structural Engineering 127:433–441

Folz B, Filiatrault A (2002) A computer program for seismic analysis of woodframe structures. Consortium of Universities for Research in Earthquake Engineering

Folz B, Filiatrault A (2004) Seismic analysis of woodframe structures. II: Model implementation and verification. Journal of Structural Engineering 130:1361–1370

Frangopol D, Liub M (2011) Structure and infrastructure engineering: Maintenance, management, life-cycle design and performance. Struct Infrastruct Eng 7:389–413

Frank McKenna, Gregory L Fenves, Michael H Scott, others (2003) the open system for earthquake engineering simulation. University of California, Berkeley, CA

Frazier PI (2018) A tutorial on bayesian optimization. arXiv preprint arXiv:180702811

Galton F (1894) Natural inheritance. Macmillan and Company

Gatto K, Uang C-M (2002) Cyclic response of woodframe shearwalls: Loading protocol and rate of loading effects. Consortium of Universities for Research in Earthquake Engineering

Gelman A, Carlin JB, Stern HS, Rubin DB (2014) Bayesian data analysis (Vol. 2). Boca Raton, FL: Chapman

Ghehnavieh EZ (2017) Seismic Analysis of Light-Frame Wood Building with a Soft-Story Deficiency

Gokkaya BU (2015) Seismic reliability assessment of structures incorporating modeling uncertainty and implications for seismic collapse safety. PhD Thesis, Stanford University

Hardyniec A, Charney F (2015) An investigation into the effects of damping and nonlinear geometry models in earthquake engineering analysis. Earthquake Engineering & Structural Dynamics 44:2695–2715

Harris SK, Egan JA (1992) Effects of ground conditions on the damage to four-story corner apartment buildings. The Loma Prieta, California, Earthquake of October 17, 1989–Marina District F181–F194

Haselton Baker Risk Group (2019) Seismic Performance Prediction Program

Haselton CB (2006) Assessing seismic collapse safety of modern reinforced concrete moment frame buildings. PhD Thesis, Stanford University

Henry V. Burton, Aryan Rezaei Rad, Jonathan Buckalew (2016) A Comparative Assessment of the Collapse Performance of Soft, Weak or Open-Front Wall Woodframe Buildings Retrofitted using Alternative Procedures. Maui, HI

Hoeffding W (1994) Probability inequalities for sums of bounded random variables. In: The Collected Works of Wassily Hoeffding. Springer, pp 409–426

Hoerl AE, Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12:55–67

Holmes W, Sommers P (1996) Northridge earthquake of January 17, 1994. Reconnaissance Report, Vol. 2. Supplement C to. Earthquake Spectra 12:1–278

Huber PJ (1992) Robust estimation of a location parameter. In: Breakthroughs in statistics. Springer, pp 492–518

Ibarra LF, Medina RA, Krawinkler H (2005) Hysteretic models that incorporate strength and stiffness deterioration. Earthquake engineering & structural dynamics 34:1489–1511

IEBC (2012) 2012 International Existing Building Code

Itani R, Cheung C (1983) Dynamic response of wood diaphragms in low-rise wood-framed buildings. Final Report, Part I, NSF Grant No CEE 8114530:

Itani RY, Cheung CK (1984) Nonlinear analysis of sheathed wood diaphragms. Journal of Structural Engineering 110:2137–2147

James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning. Springer

Jayamon JR, Line P, Charney FA (2019) Sensitivity of Wood-Frame Shear Wall Collapse Performance to Variations in Hysteretic Model Parameters. Journal of Structural Engineering 145:04018236

Jayaram N, Baker JW (2009) Correlation model for spatially distributed ground-motion intensities. Earthquake Engineering & Structural Dynamics 38:1687–1708

Jayaram N, Baker JW (2010) Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment. Earthquake Engineering & Structural Dynamics 39:1109–1131

Jehel P, Léger P, Ibrahimbegovic A (2014) Initial versus tangent stiffness-based Rayleigh damping in inelastic time history seismic analyses. Earthquake Engineering & Structural Dynamics 43:467–484

Jennings E, van de Lindt JW, Ziaei E, et al (2015) Full-scale experimental verification of soft-story-only retrofits of wood-frame buildings using hybrid testing. Journal of Earthquake Engineering 19:410–430

Jennings EN (2015) Multi-objective community-level sesimic retrofit optimization combining social vulnerability with an engineering framework for community resiliency, A. PhD Thesis, Colorado State University. Libraries

Jiang L, Ye J (2020) Quantifying the effects of various uncertainties on seismic risk assessment of CFS structures. Bulletin of earthquake engineering 18:241–272

Kameshwar S, Padgett JE (2014) Multi-hazard risk assessment of highway bridges subjected to earthquake and hurricane hazards. Engineering Structures 78:154–166

Kasal B, Leichti RJ, Itani RY (1994) Nonlinear finite-element model of complete light-frame wood structures. Journal of Structural Engineering 120:100–119

Kim JH, Rosowsky DV (2005) Fragility analysis for performance-based seismic design of engineered wood shearwalls. Journal of structural engineering 131:1764–1773

Koza JR, Bennett FH, Andre D, Keane MA (1996) Automated design of both the topology and sizing of analog electrical circuits using genetic programming. In: Artificial Intelligence in Design'96. Springer, pp 151–170

Krawinkler H, Miranda E (2004) Performance-based earthquake engineering. Earthquake engineering: from engineering seismology to performance-based engineering 9:1–9

LADBS (2015) Mandatory wood frame soft-story retrofit program: structural design guidelines. Los Angeles

Lee JY, Ellingwood BR (2015) Ethical discounting for civil infrastructure decisions extending over multiple generations. Structural Safety 57:43–52

Li Y, Ellingwood BR (2009) Framework for multihazard risk assessment and mitigation for wood-frame residential construction. Journal of structural engineering 135:159–168

Li Y, Yin Y, Ellingwood BR, Bulleit WM (2010) Uniform hazard versus uniform risk bases for performance-based earthquake engineering of light-frame wood construction. Earthquake Engineering & Structural Dynamics 39:1199–1217

Liel AB, Deierlein GG (2013) Cost-benefit evaluation of seismic risk mitigation alternatives for older concrete frame buildings. Earthquake Spectra 29:1391–1411

Liel AB, Haselton CB, Deierlein GG, Baker JW (2009) Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Structural Safety 31:197–211

Lignos DG, Krawinkler H (2012) Development and utilization of structural component databases for performance-based earthquake engineering. Journal of Structural Engineering 139:1382–1394

Lowes LN, Mitra N, Altoontash A (2003) A beam-column joint model for simulating the earthquake response of reinforced concrete frames

Maison B, McDonald B, McCormick D, et al (2014) Commentary on FEMA P-807 for retrofit of wood-frame soft-story buildings. Earthquake Spectra 30:1359–1380

Mangalathu S, Jeon J-S (2019) Stripe-based fragility analysis of multispan concrete bridge classes using machine learning techniques. Earthquake Engineering & Structural Dynamics 48:1238–1255

Mazzoni S, Gregor N, Al Atik L, et al (2020) Probabilistic Seismic Hazard Analysis and Selecting and Scaling Ground-Motion Records, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings" Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Miller MK (2014) Seismic risk assessment of complex transportation networks. PhD Thesis, Stanford University

Miranda E, Aslani H, Taghavi S (2004) Assessment of seismic performance in terms of economic losses. In: Proceedings, International Workshop on Performance-Based Seismic Design: Concepts and Implementation. Pacific Earthquake Engineering Research (PEER) Center, University of …, pp 149–160

Mitrani-Reiser J (2007) An ounce of prevention: probabilistic loss estimation for performance-based earthquake engineering. PhD Thesis, California Institute of Technology

Moehle J, Deierlein GG (2004) A framework methodology for performance-based earthquake engineering. In: 13th world conference on earthquake engineering

Moradi S, Burton HV (2018) Response surface analysis and optimization of controlled rocking steel braced frames. Bulletin of Earthquake Engineering 16:4861–4892

Murphy KP (2012) Machine learning: a probabilistic perspective. MIT press

NIBS (2017) Natural hazard mitigation saves 2017 interim report: summary of findings

Padgett JE, Dennemann K, Ghosh J (2010) Risk-based seismic life-cycle cost–benefit (LCC-B) analysis for bridge retrofit assessment. Structural Safety 32:165–173

Pang W, Rosowsky D, Pei S, Van de Lindt J (2007) Evolutionary parameter hysteretic model for wood shear walls. Journal of structural engineering 133:1118–1129

Pang W, Ziaei E, Filiatrault A (2012) A 3D model for collapse analysis of soft-story light-frame wood buildings. World 15:19

Park S, van de Lindt JW (2009) Formulation of seismic fragilities for a wood-frame building based on visually determined damage indexes. Journal of performance of constructed facilities 23:346–352

Pearson K (1896) VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. Philosophical Transactions of the Royal Society of London Series A, containing papers of a mathematical or physical character 253–318

Pei S, Van de Lindt J (2011) Seismic numerical modeling of a six-story light-frame wood building: Comparison with experiments. Journal of Earthquake Engineering 15:924–941

Pei S, van de Lindt J (2010a) User's manual for SAPWood for Windows: Seismic analysis package for woodframe structures. NEEShub (nees org)

Pei S, van de Lindt J (2010b) Influence of structural properties and hazard level on seismic loss estimation for light-frame wood structures. Engineering structures 32:2183–2191

Pei S, van de Lindt J (2009) Systematic seismic design for manageable loss in wood-framed buildings. Earthquake spectra 25:851–868

Polensek A, Laursen HI (1984) Seismic Behavior of Bending Components and Intercomponent Connections of Light Frame Wood Buildings. Oregon State University

Porter KA, Beck JL, Shaikhutdinov RV (2002) Investigation of sensitivity of building loss estimates to major uncertain variables for the Van Nuys testbed. PEER Report 2002/03

Porter KA, Kiremidjian AS, LeGrue JS (2001) Assembly-based vulnerability of buildings and its use in performance evaluation. Earthquake spectra 17:291–312

Porter KA, Scawthorn CR, Beck JL (2006) Cost-effectiveness of stronger woodframe buildings. Earthquake Spectra 22:239–266

Pozza L, Scotta R, Trutalli D, et al (2014) Experimental and numerical analyses of new massive wooden shear-wall systems. Buildings 4:355–374

Priestley M (1999) Displacement-based approaches to rational limit states design of new structures. In: Proc. 11th Eur. Conf. Earthquake Eng.: Invited Lectures. pp 317–335

Priestley M (2000) Performance based seismic design. Bulletin of the New Zealand society for earthquake engineering 33:325–346

Priestley MN (1993) Myths and fallacies in earthquake engineering-conflicts between design and reality. Bulletin of the New Zealand National Society for Earthquake Engineering 26:329–341

Ramirez C, Liel A, Mitrani-Reiser J, et al (2012) Expected earthquake damage and repair costs in reinforced concrete frame buildings. Earthquake Engineering & Structural Dynamics 41:1455–1475

Rasmussen CE (2003) Gaussian processes in machine learning. In: Summer School on Machine Learning. Springer, pp 63–71

Rathje EM, Dawson C, Padgett JE, et al (2017) DesignSafe: new cyberinfrastructure for natural hazards engineering. Natural Hazards Review 18:06017001

Roohi M, Hernandez EM, Rosowsky D (2019) Nonlinear seismic response reconstruction and performance assessment of instrumented wood-frame buildings—Validation using NEESWood Capstone full-scale tests. Structural Control and Health Monitoring 26:e2373

Rosowsky DV, Ellingwood BR (2002) Performance-based engineering of wood frame housing: Fragility analysis methodology. Journal of Structural Engineering 128:32–38

Rowshandel B, Reichle M, Wills C, et al (2006) Estimation of future earthquake losses in California. California Geological Survey, Menlo Park, California

Schiller B, Hutchinson T, Cobeen K (2020a) Cripple Wall Small-Component Test Program: Dry Specimens, A Report for the “Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings” Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Schiller B, Hutchinson T, Cobeen K (2020b) Cripple Wall Small-Component Test Program: Wet Specimens I, A Report for the “Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings” Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Schiller B, Hutchinson T, Cobeen K (2020c) Cripple Wall Small-Component Test Program: Wet Specimens II, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings" Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

SEAOSC (2017) SEAOSC design guide: City of Los Angeles soft, weak and open-front wall line building ordiance. Structural Engineers Association of Southern California, Los Angeles

Seo J, Dueñas-Osorio L, Craig JI, Goodno BJ (2012) Metamodel-based regional vulnerability estimate of irregular steel moment-frame structures subjected to earthquake events. Engineering Structures 45:585–597

Sichani ME, Padgett JE, Bisadi V (2018) Probabilistic seismic analysis of concrete dry cask structures. Structural Safety 73:87–98

Snoek J, Larochelle H, Adams RP (2012) Practical bayesian optimization of machine learning algorithms. In: Advances in neural information processing systems. pp 2951–2959

Soft Story Retrofit Pros (2019) Soft Story Retrofit Pricing. : http://www.softstoryretrofitpros.com/pricing

Stewart W (1987) The seismic design of plywood sheathed shear walls.

Stone M (1974) Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B (Methodological) 36:111–133

Student (1908) The probable error of a mean. Biometrika 1–25

Sun H, Burton HV, Huang H (2020) Machine Learning Applications for Building Structural Design and Performance Assessment: State-of-the-Art Review. Journal of Building Engineering 101816

Sutley EJ, van de Lindt JW (2016) Evolution of predicted seismic performance for wood-frame buildings. Journal of Architectural Engineering 22:B4016004

Symans M, Charney F, Whittaker A, et al (2008) Energy dissipation systems for seismic applications: current practice and recent developments. Journal of structural engineering 134:3–21

Tibshirani R (1996) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58:267–288

Trevor H, Robert T, Jerome F (2009) The elements of statistical learning. Springer Science+ Business Media, LLC

USGS (2019) USGS Unified Hazard Tool

Vail K, Lizundia B, Welch DP, Reis E (2020) Earthquake Damage Workshop, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings" Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley

Vamvatsikos D, Cornell CA (2002) Incremental dynamic analysis. Earthquake engineering & structural dynamics 31:491–514

Van de Lindt J (2005) Damage-based seismic reliability concept for woodframe structures. Journal of Structural Engineering 131:668–675

Van de Lindt JW, Liu H, Pei S (2007) Performance of a woodframe structure during full-scale shake-table tests: Drift, damage, and effect of partition wall. Journal of performance of constructed facilities 21:35–43

van de Lindt JW, Pei S, Liu H, Filiatrault A (2010) Three-dimensional seismic response of a full-scale light-frame wood building: Numerical study. Journal of structural engineering 136:56–65

van de Lindt JW, Rosowsky DV, Pang W, Pei S (2013) Performance-based seismic design of midrise woodframe buildings. Journal of Structural Engineering 139:1294–1302

van de Lindt JW, Symans MD, Pang W, et al (2012) Seismic risk reduction for soft-story woodframe buildings: the nees-soft project. World 15:19

Wald A (1939) Contributions to the theory of statistical estimation and testing hypotheses. The Annals of Mathematical Statistics 10:299–326

Welch BL (1951) On the comparison of several mean values: an alternative approach. Biometrika 38:330–336

Welch DP, Deierlein GG (2020) Technical Background Report for Structural Analysis and Performance Assessment, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings" Project, Report PEER 2020/XX. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. The annals of mathematical statistics 9:60–62

Xia R, Schleuss J (2016) LA releases addresses of 13,500 apartments and condos likely to need earthquake retrofitting. Los Angeles Times

Xie Y, DesRoches R (2019) Sensitivity of seismic demands and fragility estimates of a typical California highway bridge to uncertainties in its soil-structure interaction modeling. Engineering Structures 189:605–617

Xie Y, Ebad Sichani M, Padgett JE, DesRoches R (2020) The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. Earthquake Spectra 8755293020919419

Yamaguchi N (1998) Dynamic performance of wooden bearing walls by shaking table test. Proceedings of the 5th WCTE, Montreux, Switzerland, 1998 2:26–33

Yi Z, Burton HV, Shokrabadi M, Issa O (2020) Multi-scale cost-benefit analysis of the Los Angeles Soft-Story Ordinance. Engineering Structures 214:110652

Yin Y-J, Li Y (2010) Seismic collapse risk of light-frame wood construction considering aleatoric and epistemic uncertainties. Structural Safety 32:250–261

Zareian F, Lanning J (2020) Development of Testing Protocol for Cripple Wall Components, A Report for the "Quantifying the Performance of Retrofit of Cripple Walls and Sill Anchorage in Single Family Wood-Frame Buildings" Project. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology) 67:301–320